DEPARTMENT OF PLANT MOLECULAR BIOLOGY
UNIVERSITY OF DELHI SOUTH CAMPUS
NEW DELHI-110021

January 12, 2024

The strength of M.Sc. students in our Department in the last five years is as follows:

M.Sc. Batch	No. of Students
2017-2019	10
2018-2020	8
2019-2021	10
2020-2022	13
2021-2023	14

Professor Sanjay Kapoor
Head of the Department
Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

Characterization of drought responsive *OsEATB* transcription factor in rice

Thesis Submitted to the University of Delhi

in the Partial fulfilment for the Degree of

Master of Science

in

Plant Molecular Biology and Biotechnology

2019



Professor Paramjit Khurana
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110 021 India

ANKIT

Department of Plant Molecular Biology

University of Delhi South Campus

New Delhi 110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Promoter Analysis and Sequence Variability of Begomoviruses Infecting Okra**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in the planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: May 21, 2019

Atreyee Sardar

Signature of the Candidate

Atreyee Sardar

Indranil Dasgupta

Signature of the Supervisor

Prof. Indranil Dasgupta

Paramjit Khurana

Signature of the HOD

Prof. Paramjit Khurana

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Functional characterization of stress inducible transcription factor gene, *OsHOX6*, in rice**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi, South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in the planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Himanshi

Date: 21st May, 2019

Signature of the Candidate

(Himanshi Gautam)

Signature of the Supervisor

(Prof. Akhilesh K. Tyagi)

Paramjit Khurana

Signature of the HOD

(Prof. Paramjit Khurana)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation thesis entitled, **"Cloning of *Arabidopsis LHY* and *TOC1* genes and characterization of their mutants in response to phytohormones and sugar"** during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in the planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures originality of the work as well as presentation.

Date: May 21, 2019

Monika

Monika Shrivastava

(Signature of the Candidate)

Jitendra Khurana

Prof. Jitendra P. Khurana

(Signature of the Supervisor)

Paramjit Khurana

Prof. Paramjit Khurana

(Signature of the HOD)

Benito Juarez Road, New Delhi.110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, **"*At*HSP101 plays role in seed germination under heavy metal stress in Arabidopsis thaliana"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology.The work presented here demonstrates that the candidate has been trained in the planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 21-05-2019

Signature of the Candidate

NIDHI YADAV

Signature of the Supervisor

PROF. ANIL GROVER

Signature of the HOD

PROF. PARAMJIT KHURANA

Benito Juarez Road, New Delhi.110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, **"Identification of targets of rice PP42 and *Arabidopsis* DSP4 using yeast two- hybrid and cloning of *OsPP42* into pGADT7 vector"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in the planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 20 May, 2019

Signature of the Candidate

(Sarvesh Jonwal)

Signature of the Supervisor

(Prof. Girdhar K. Pandey)

Signature of the HOD

(Prof. Paramjit Khurana)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Comparative Analysis of Different Promoters under various Abiotic Stress**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in the planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: May 21, 2019

Signature of the Candidate

SHIVAM CHAUDHARY

Signature of the Supervisor

Prof. Paramjit Khurana

Signature of the HOD

Prof. Paramjit Khurana

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Preparation of dual reporter construct expressing fluorescent proteins under DR5 and TCS*n* promoters for simultaneous visualization of auxins and cytokinins levels in BY-2 cell line**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in the planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: May 21, 2019

Signature of the Candidate

Shruti Saini

Signature of the Supervisor

Prof. Sanjay Kapoor

Signature of the HOD

Prof. Paramjit Khurana

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus
DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Analysis of the promoter controlling the expression of the coat protein gene of Sri Lankan cassava mosaic virus**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Monica Biswas

Date: 6th June, 2020

Signature of the Candidate

(Monica Biswas)

Indranil Dasgupta

Signature of the Supervisor

(Prof. Indranil Dasgupta)

Signature of the HOD

(Prof. Anil Grover)

Benito Juarez Road, New Delhi. 110021, **India.** Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus
DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, **"Kinship of APX2 and HSP101 in birthing and working together is important in protection of *Arabidopsis thaliana* seedlings from heat stress"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in the planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 06-06-2020

Haneet

Signature of the Candidate

(HANEET PAPREJA)

Signature of the Supervisor

(PROF. ANIL GROVER)

Signature of the HOD

(PROF. ANIL GROVER)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Characterization of *OsPLC3* in *Arabidopsis thaliana* and rice**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 06.06.2020

Signature of the Candidate

(AASTHA)

Signature of the Supervisor

(PROF. GIRDHAR K. PANDEY)

Signature of the HOD

(PROF. ANIL GROVER)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**An insight into the Class C of Heat Stress Transcription Factor (HSF) protein family in plants**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Ritesh.

Date: June 6, 2020

Signature of the Candidate
(RITESH)

Surekha

Signature of the Supervisor
(Dr. Surekha Katiyar Agarwal)

Signature of the HOD
(Prof. Anil Grover)

University of Delhi South Campus
DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Overexpression of rice stress associated protein gene *OsiSAP15* in *Arabidopsis thaliana* confers tolerance to drought stress**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus by **Anurag Panchal** and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Anurag

Date: 6th June 2020

(Anurag Panchal)

(Prof. Akhilesh K. Tyagi)

(Prof. Anil Grover)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax: 24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**A COMPENDIUM OF THE TRANSCRIPTION FACTORS ASSOCIATED WITH PLANT SENESCENCE AND AN INSIGHT INTO THE GENETIC VARIABILITY IN PROMOTER AND PROTEIN SEQUENCE OF HSFC1 IN NATURAL ACCESSIONS OF *ARABIDOPSIS THALIANA***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Apurva Singh

Date: 28-06-21

Signature of the Candidate

(APURVA SINGH)

Surekha

Signature of the Supervisor

(Prof. SUREKHA KATIYAR-AGARWAL)

Indranil Dasgupta

Signature of the HOD

(Prof. INDRANIL DASGUPTA)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Functional characterization of Rice Phospholipase C (*OsPLC3*)**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 28/06/2021

Signature of the Candidate

(Atul Bahukhandi)

Signature of the Supervisor

(Prof. Girdhar K. Pandey)

For

Signature of the HOD

(Prof. Anil Grover)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Heat Sensing Mechanism in Lower Organisms**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: June 28, 2021

Signature of the Candidate

(**Gopika Unni P J**)

Signature of the Supervisor

(**Prof. Paramjit Khurana**)

Signature of the HOD

for (**Prof. Anil Grover**)

अध्यक्ष, पादप आणविक जीव विज्ञान विभाग
Head, Department of Plant Molecular Biology
दिल्ली विश्वविद्यालय दक्षिणी परिसर
University of Delhi South Campus
नई दिल्ली-११००२१ / New Delhi-110021

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Toward elucidating the function of a plant-specific transcription factor, OsTCP17, in rice**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus by **K. Thanduanlung** and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates thatthe candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 28th June 2021

K. Thanduanlung

Signature of the Candidate

(K. THANDUANLUNG)

Signature of the Supervisor

(PROF. AKHILESH K. TYAGI)

Indranil Dasgupta

Signature of the HOD

(PROF. INDRANIL DASGUPTA)

Benito Juarez Road, New Delhi. 110021, India. Tel. 24111208, Fax: 24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Comparative *in-silico* analysis of PIFs in *Arabidopsis* and Rice and *in-silico* based functional characterization of *TaPIF4***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Karishma Sahu

Date: 28 JUNE 2021

Signature of the Candidate

(**KARISHMA SAHU**)

Paramjit Khurana

Signature of the Supervisor

(**Prof. Paramjit Khurana**)

Indrid Dasgupta

Signature of the HOD

for (**Prof Anil Grover**)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Identification of miRNA biogenesis genes in Rice and *in-silico* analyses of their regulation under drought conditions**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Nidhi

Date: 28 June 2021

(Signature of the Candidate)

NIDHI MAURYA

Indrani Dasgupta

Signature of the Supervisor

For (PROF. SAURABH RAGHUVANSHI)

Indrani Dasgupta

Signature of the HOD

(PROF. INDRANIL DASGUPTA)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Study of interaction between SlMBD8 & SlIAA17 and characterization of their role in tomato fruit ripening by VIGS**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Nitika Gupta

Date: 28.06.2021

Signature of the Candidate
(NITIKA GUPTA)

Arun K Sharma

Signature of the Supervisor
(Prof. Arun Kumar Sharma)

Indu Dasgupta

for Signature of the HOD
(Prof. Anil Grover)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Expression analysis of heat stress regulated genes at diverse developmental stages by *in silico* tools and role of Hsp101 in metal stress response by wet laboratory experiments in *Arabidopsis thaliana***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 28th June 2021

Signature of the Candidate

(SHITIJ GUPTA)

Signature of the Supervisor

(PROF. ANIL GROVER)

Signature of the HOD

for

(PROF. ANIL GROVER

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Regulation of Translation in Plants: Its Impact on Developmental Transitions**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: June 28, 2021

Signature of the Candidate

(Sritama Das)

Signature of the Supervisor

(Prof. Sanjay Kapoor)

Signature of the HOD

for

(Prof. Anil Grover)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**A comparative study of differentially expressed mRNAs in *Oryza sativa* and *Nicotiana benthamiana* upon viral infection**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 15.06.2021

Signature of the Candidate

(Aashmeen Kaur)

Signature of the Supervisor

(Prof. Indranil Dasgupta)

Signature of the HOD

for (Prof. Anil Grover)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Determination of the Role of Epigenetic Regulation in MADS -Box Genes during Reproductive Development in Rice**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 20-05-2022.

Signature of the Supervisor
(Prof. Sanjay Kapoor)

Abhishek Kumar

Signature of the Candidate
Abhishek Kumar

Signature of the HOD
(Prof. Anil Grover)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**RNAi and genome editing techniques for generating plant virus resistance and testing resistance against Tomato Leaf Curl New Delhi Virus in *Nicotiana benthamiana* by a CRISPR construct**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 20-05-22

Signature of the Candidate

(Aditi Soni)

Signature of the Supervisor

(Prof. Indranil Das Gupta)

Signature of the HOD

(Prof. Anil Grover)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Developing Strategies to Predict Calmodulin Binding Sites in Rice Mads-Box Family of Proteins**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarising research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as the presentation.

Date: 20-05-2022

Signature of the Candidate
Ankur

Signature of the Supervisor
(Prof. Sanjay Kapoor)

Signature of the HOD
(Prof. Anil Grover)

Benito Juarez Road, New Delhi 110021, INDIA. Tel: +91 11 24111208, Fax: 24110669, Web: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, **"Transactivation of the CP promoter of Sri Lankan Cassava Mosaic Virus with AC2"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as the presentation.

Date: 20/5/22

Signature of the Candidate

HARDIK

Signature of the Supervisor

Prof. INDRANIL DASGUPTA

Signature of the Head of Department

Prof. ANIL GROVER

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research **work embodied in this M.Sc. dissertation entitled, "Virus Induced Gene Silencing in Crops"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 20.05.2022

Kirti

Signature of the Candidate

(Kirti Sharma)

Indranil Dasgupta

Signature of the Supervisor

(Prof. Indranil Dasgupta)

Signature of the HOD

(Prof. Anil Grover)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M. Sc. dissertation entitled **"Unravelling roles of TET3 and TET13 in abiotic stress response in *Arabidopsis thaliana*"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 20.05.22

Kirti

Signature of the Candidate
(Kirtishree Yadav)

Surekha

Signature of the Supervisor
(Prof. Surekha Katiyar Agarwal)

Signature of the HOD
(Prof. Anil Grover)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, **"Analysis of the expression of *Arabidopsis thaliana* HSP101 gene in heat and fungal stresses"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as the presentation.

Date:
May 20, 2022

Signature of the Candidate

MANSI PANT

Signature of the Supervisor

Prof. ANIL GROVER

Signature of the Head of Department

Prof. ANIL GROVER

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Studies on the role of Hsp101 in heat and fungal stresses in *Arabidopsis thaliana***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 20th May 2022

Signature of the Candidate
(NEHA RAWAT)

Signature of the Supervisor
(PROF. ANIL GROVER)

Signature of the HOD
(PROF. ANIL GROVER)

Professor Anil Grover
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "STUDY OF THE EFFECT OF VIGS OF *VIM1* GENE ON PHENOTYPE OF TOMATO PLANTS" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 19.05.2022

Nishita

Signature of the candidate

(Nishita)

Signature of the Supervisor

(Prof. Arun Kumar Sharma)

Signature of the HOD

(Prof. Anil Grover)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Functional characterization of calcium and potassium transporters under abiotic stress conditions in plants**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 20th May, 2022

Niyaz Ahmed

Signature of the Candidate

(NIYAZ AHMED)

Signature of the Supervisor

(Prof. Girdhar K. Pandey)

Signature of the HOD

(Prof. Anil Grover)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Characterization of the function of novel miRNA M00107 which targets the *ERF72* mRNA *in silico* during tomato fruit ripening using Virus-Induced Gene Silencing technology**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution, and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as the presentation.

Date: 20.05.2022

Signature of the Candidate
(SHWETA)

Signature of the Supervisor
(Prof. Arun Kumar Sharma)

Signature of the HOD
(Prof. Anil Grover)

Professor Anil Grover
Head
Department of Plant Molecular Bology
University of Delhi South Campus
New Delhi-110021

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, ***"In silico* expression analysis of Arabidopsis clade B type 2C protein phosphatases and identification of interacting proteins of AtPP2C5 by yeast-two-hybrid method"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 20/05/2022

Signature of the Candidate

(SONIA)

Signature of the Supervisor

(Prof. Girdhar K. Pandey)

Signature of the HOD

(Prof. Anil Grover)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M. Sc. dissertation entitled "**Towards functional characterization of TruB and Pus10 Pseudouridine Synthase family in *Arabidopsis thaliana***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 20.05.22

Vaishali Gupta

Signature of the Candidate

(Vaishali Gupta)

Surekha

Signature of the Supervisor

(Prof. Surekha Katiyar Agarwal)

Signature of the HOD

(Prof. Anil Grover)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Study on the roles of Ethylene Responsive Factors(ERFs) of tomato against salt stress.** " was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: May 24,2023

Signature of the Candidate

(Aditya)

Signature of the Supervisor

(Prof. Arun K.Sharma)

Signature of the HOD

(Prof. Sanjay Kapoor)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, **"*In silico* expression analysis of rice calcium signaling associated kinases, *OsCIPK9* & *OsCIPK23* and identification of interacting partners of OsCIPK23 by yeast-two-hybrid library screening"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 24.05.2023

Signature of the candidate

(ANKIT KUMAR)

Signature of the Supervisor

(Prof. Girdhar K. Pandey)

Signature of the HOD

(Prof. Sanjay Kapoor)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Role of COP9 signalosome in meiosis and plant reproduction during heat stress**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date:

Signature of the Candidate

(ASHWANI KUMAR)

Signature of the Supervisor

(DR. AMIT KUMAR SINGH)

Signature of the HOD

(PROF. SANJAY KAPOOR)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

Benito Juarez Road, New Delhi.110021, India. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**RNAi and genome editing techniques for generating plant virus resistance and testing resistance against Tomato Leaf Curl New Delhi Virus in *Nicotiana benthamiana* by a CRISPR construct**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 20-05-22

Signature of the Candidate

(Aditi Soni)

Signature of the Supervisor

(Prof. Indranil Das Gupta)

Signature of the HOD

(Prof. Anil Grover)

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Exploring the Impact of Single Amino Acid Substitutions on BIG SIZE and NINJA Protein Interaction in Legumes**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 24.5.23

Bhawana
Signature of the Candidate

(BHAWANA)

Kamal Kumar
24/05/23
Signature of the Supervisor

(Dr. KAMAL KUMAR)

Signature of the HOD

(Prof. SANJAY KAPOOR)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus
DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Vitamin C supplementation serves as an antidote to the heat stress damages in Arabidopsis thaliana**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Deepshikha Sharma

Date: 24-05-2023

Signature of the Candidate

(DEEPSHIKHA SHARMA)

Signature of the Supervisor

(PROF. ANIL GROVER)

Signature of the HOD

(Prof. SANJAY KAPOOR)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, **"INVESTIGATION OF ACTIVE CHROMATIN MARKS FUNCTION IN MEIOSIS AND PLANT REPRODUCTION DURING HEAT STRESS"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 24/05/2023

K.V.V.K. Arjun Chowdary

Signature of the Candidate

(KOTA VENKATA SAI KRISHNA ARJUN CHOWDARY)

Signature of the Supervisor

(DR. AMIT KUMAR SINGH)

Signature of the HOD

(PROF. SANJAY KAPOOR)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

Benito Juarez Road, New Delhi.110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M. Sc. dissertation entitled, "**Unravelling the Role of PUS9 (Pseudouridine Synthase 9) in Regulation of Development and Abiotic Stress Responses in *Arabidopsis thaliana***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution, and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 22.05.2023

Signature of the Candidate

(Milinda Lahiri)

Surekha

Signature of the Supervisor

Prof. Surekha Katiyar-Agarwal

Signature of the HOD

Prof. Sanjay Kapoor

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

University of Delhi South Campus
DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M. Sc. dissertation entitled, "**Unraveling the Role of Tetraspanins (TET2 and TET3) in Modulating Abiotic Stress Responses in *Arabidopsis thaliana***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date : 24/5/23

Muskaan
Signature of the Candidate

(Muskaan Johnson)

Surekha
Signature of the Supervisor

(Dr. Surekha Katiyar-Agarwal)

Signature of the HOD

(Dr. Sanjay Kapoor)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in



University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Development of Molecular Markers for Mapping Race-Specific Fusarium Wilt Resistance Loci in Chickpea**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 24.5.23

Nikita Gupta
Signature of the Candidate

(Nikita Gupta)

Kamal Kumar
24/05/23
Signature of the Supervisor

(Dr. Kamal Kumar)

Signature of the HOD

(Prof. Sanjay Kapoor)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M. Sc. dissertation entitled, **"Study on the roles of Ethylene Responsive Factors (ERFs) of tomato in tolerance against oxidative stress"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 24/05/2023

Signature of the Candidate

(Pooja Solanki)

Signature of the Supervisor

(Prof. Arun Kumar Sharma)

Signature of the HOD

(Prof. Sanjay Kapoor)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, **"ESTABLISHMENT OF CYTOKININ AND AUXIN SENSORS, TCSv2 AND DR5 IN ARABIDOPSIS FOR SPATIO-TEMPORAL ANALYSIS OF THESE HORMONES IN PLANTA"**was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 24th May 2023

Signature of the Supervisor
(Prof. Sanjay Kapoor)

Rinki

Signature of the Candidate
RINKI

Signature of the HOD
(Prof. Sanjay Kapoor)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**ESTABLISHMENT AND CHARACTERISATION OF FLUORESCENCE-BASED REPORTER SYSTEMS FOR SPATIO-TEMPORAL ANALYSIS OF CYTOKININ ACCUMULATION AND CELL DIVISION IN ARABIDOPSIS**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 24th May, 2023

Signature of the Candidate

(Sheena Shah)

Signature of the Supervisor

(Prof. Sanjay Kapoor)

Signature of the HOD

(Prof. Sanjay Kapoor)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110607, Website: www.dpmb.ac.in

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Glutathione, the major ROS scavenger, supplementation serves as an antidote against heat stress damages in *Arabidopsis thaliana***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

23-05-2023

Date:

Smrity Jha

Signature of the Candidate
(Smrity Jha)

Signature of the Supervisor
(Prof. Anil Grover)

Signature of the HOD
(Prof. Sanjay Kapoor)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, **"Comparative *in silico* investigation of putative interactors of *AtCIPK9* and their interaction analysis in Arabidopsis"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Sneha Pathak

Date: 24.05.2023

Signature of the candidate

(SNEHA PATHAK)

Signature of the Supervisor

(Prof. Girdhar K. Pandey)

Signature of the HOD

(Prof. Sanjay Kapoor)

Professor Sanjay Kapoor
Head
Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021

Overexpression of rice stress associated protein gene *OsiSAP15* in *Arabidopsis thaliana* confers tolerance to drought stress

Thesis Submitted to the University of Delhi
in the Partial fulfilment for the Degree of

Master of Science
in
Plant Molecular Biology and Biotechnology

2020

ANURAG PANCHAL

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi 110021, India

Overexpression of rice stress associated protein gene *OsiSAP15* in *Arabidopsis thaliana* confers tolerance to drought stress

Thesis Submitted to the University of Delhi
in the Partial fulfilment for the Degree of

Master of Science
in
Plant Molecular Biology and Biotechnology

2020

ANURAG PANCHAL

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi 110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Overexpression of rice stress associated protein gene *OsiSAP15* in *Arabidopsis thaliana* confers tolerance to drought stress**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus by **Anurag Panchal** and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Anurag

Date: 6th June 2020

(**Anurag Panchal**)

(**Prof. Akhilesh K. Tyagi**)

(**Prof. Anil Grover**)

Benito Juarez Road, New Delhi. 110021, India. Tel. 24111208, Fax: 24110669, Website: www.dpmb.ac.in

Preface

Constantly changing environment and anthropogenic activities in world have created adverse stressed conditions for plants which are ultimately being reflected in crop yield and productivity in recent years. Being sessile, plants encounter different biotic and abiotic stresses which affect their growth and development. Stress is likely to enhance the severity of problem to be faced by plants in near future. In the next 50 years, drought stress alone is projected to limit the crop productivity in more than half of the arable land (Cattivelli et al., 2008). Drought is a multidimensional response in plants to high salinity, dehydration stress, heat stress, oxidative stress which ultimately results in different morpho-physiological changes in plant.

To cope up with the continuously changing environment, several genes are induced to protect the plant for proper growth and development. One such newly emerging class of stress inducible genes called stress associated proteins (SAPs) are induced in response to multiple abiotic stresses and have been proved to be ideal and potential candidates that can impart tolerance to several stresses. These multiple stress inducible class of zinc finger protein (ZFP) coding genes upon their overexpression in plants are able to confer tolerance to different stresses.

Stress associated proteins (SAPs) are the zinc finger proteins containing A20 or AN1 domains or both A20 and AN1 domains at their N and C terminal, respectively. Their animal homologs are known to regulate immunity (Beyaert et al., 2000). Mukhopadhyay et al. (2004) characterized first rice SAP gene (*OsiSAP1*) which conferred tolerance to abiotic stresses like dehydration and cold stress at seedling stage. Subsequently, Vij and Tyagi (2006) reported 18 *OsiSAP* genes in rice and 14 *AtSAP* genes in *Arabidopsis thaliana,* respectively. In last decade, some unique roles have also been accredited to them like they can sense the redox status of cell during stressed conditions and encompass E3 ubiquitin ligase activity.

In this study with rice AN1 containing protein gene *OsiSAP15*, attempts have been made to dig out its role in providing drought stress tolerance upon its overexpression in model system *Arabidopsis thaliana*. Second part is to clone the gene in pET28a (+) vector and transform host *E.coli* strain DH5α. In future, isolated protein will be used for checking its molecular activities and interactions during stress.

Acknowledgements

It is indeed my pleasure to express my sincere thanks to my esteemed teachers, my seniors, friends who have contributed to my academics and supported me in one way or the other during this indelible journey of masters.

First and foremost, I would like to express my deep and sincere regards, profound sense of gratitude and indebtedness to my teacher and supervisor **Prof. Akhilesh K. Tyagi** *for his time, advice, constant motivation and excellent guidance that have gone a long way in completion of my dissertation. His enthusiasm, dedication and commitment towards work are the qualities, which I would always like to emulate.*

I am highly obliged to all the faculty members and teachers **Prof. Anil Grover, Prof. J.P. Khurana, Prof. Paramjit Khurana, Prof. Indranil Dasgupta, Prof. Arun K. Sharma, Prof. Girdhar K. Pandey, Prof. Sanjay Kapoor, Dr. Saurabh Raghuvanshi, and Dr. Surekha K. Agrawal** *for helping me in gaining a clear understanding of this molecular biology coursework. I would also like to thank* **Mrs. Neelam** *and* **Mrs. Renu** *for their guidance and help throughout.*

I truly want to thank my mentors cum friends **Mr. Vishal Vashisht** *and* **Dr. Poonam Mehra** *for providing me the constant support, commendable guidance and steadfast encouragement to achieve optimal level of my performance.*

I would like to thank **Dr. Priyanka deweshwar**, **Ms. Ankita Prusty**, **Mr. Shivam Sharma** *for their immense support and guidance and for always taking interest in progress of my work and helping me out whenever possible.*

I also appreciate day to day lab organisation and helping in whatever way possible by **Mr. Dharmendra**, **Mr. Ahmed**, **Mr. Raju**.

I want to thank all of my batch-mates **Ritesh, Aastha, Haneet, Vartika, Frankie, Monica and Mahima** *and my seniors* **Swati, Shivam and Himanshi** *for their useful discussions and timely motivation.*

Last but most important I am deeply thankful to my beloved parents and my sister who have constantly motivated and encouraged me to do my job with best of my way possible. Cordial thanks to the almighty god for continuously being with me in this entire journey.
I want to dedicate this work to my grandparents.

Abbreviations

1. ˚C - Degree Celsius
2. bp - Base pair
3. cDNA - Complementary DNA
4. C-terminus - Carboxy terminus
5. DNA - Deoxyribonucleic acid
6. g - Gram
7. hr. - Hour
8. L - Litre
9. LB - Luria Bertani
10. MCT - Micro-centrifuge tube
11. ml - Millilitre
12. mM - Millimolar
13. mRNA - Messenger Ribonucleic acid
14. nm - Nanomolar
15. N – terminus - Amino terminus
16. OD - Optical Density
17. OE - Over-expression
18. PB-1 - Pusa Basmati 1
19. PCR - Polymerase Chain Reaction
20. *RD29A* - Responsive to dehydration 29A
21. RNA - Ribonucleic acid
22. ROS - Reactive oxygen species
23. rpm - Revolutions per minute
24. RT - Room temperature
25. RT PCR - Real Time PCR
26. SAP - Stress Associated Protein
27. WDS - Water-deficit stress
28. WT - Wild-type
29. ZFP - Zinc Finger Protein
30. µg - Microgram
31. µM - Micromolar

Table of Contents

1. **Chapter 1. Review of literature……7-31**
2. **Chapter 2. Materials and Methods……..32-39**
3. **Chapter 3. Results and Discussion……...40-46**
4. **Chapter 4. Conclusions and Future work……47-48**
5. **Chapter 5. Bibliography……49-53**
6. **Appendix...……..54**

Chapter 1

Review of literature

SAPs: An emerging class of stress regulators in plants

1.1 INTRODUCTION

The rate of world population growth has exceeded the rate of growth in food grain production. It is predicted that the world population will exceed 8 billion people by 2025 and to meet these global food demands, the production of food grains needs to be increased up to 50% more by the year 2025 (Peng et al., 1999). The major hurdles in attaining a good yield in crop plants for increasing population are the biotic and abiotic stresses that they encounter during their growth and development. Among them abiotic stresses like drought, high light intensities, extreme temperature and exposure to heavy metals play pivotal role in less productivity and yield. Abiotic stress is the primary cause of crop loss worldwide, reducing average yields for most major crops by more than 50% (Boyer et al., 1982; Bray et al., 2000). This burden of environmental stress is likely to increase in upcoming time due to increased global warming and industrialization of cultivable land for agriculture.

In current scenario, to meet the demands of growing population in changing climate, development of improved crop varieties is a major challenge. Classical breeding programmes and genetic modifications are the two powerful methods to overcome this problem. To save our time and labour of genetic introgression in plant breeding, molecular modification of plants to overcome stress is thought to be an intelligent approach.

In accordance with that, *SAP* is an emerging gene family in crop plants which has the potential to combat the abiotic stresses at different developmental stages. *SAP* (Stress Associated Protein) is a newly identified class of zinc-finger proteins (ZFPs) in plants that confers tolerance to multiple abiotic and biotic stresses (Giri et al., 2011; Tyagi et al., 2014). This gene family is found to be conserved across a wide class of plant species. Characteristic feature of SAPs is the presence of unique A20 or AN1 zinc

finger domains or both A20 and AN1 domains at N- or C-terminal of their protein structure. This unique SAP family is prevalent in many organisms including plants, animals, protists, and fungi. These proteins in animal kingdom are known for their role in immune responses. A20 domain is characterized by the presence of multiple Cys2Cys2 finger motifs which was first identified in TNFα-inducible protein of the human endothelial cells (Dixit et al., 1990; Opipari et al., 1990) and AN1 Zinc finger domains are characterized by presence of Cys2His2 finger motifs and discovered as proteins encoded by the *Xenopus laevis* Animal Hemisphere1 (AN1) maternal RNA (Linnen et al., 1993). Some SAPs also contain additional Cys2His2 RING motifs at the C-terminus.

Till date many SAPs have been discovered and characterized in several plant species like rice, thale cress, cotton, banana, poplar, wheat, etc. for their respective roles in stress tolerance. Almost all of them at any growth stage provide tolerance to one or more abiotic and biotic stresses. First plant *SAP* to be discovered was *OsiSAP1* from *Oryza sativa* subspecies *indica* (Mukhopadhyay et al., 2004) and its overexpression in tobacco provided tolerance to salt, cold and dehydration. Subsequently, other plant species were searched for role of SAPs in mitigating stress and Ben Saad et al. (2012) reported that expression of *AlSAP* gene in rice could help tolerate a range of abiotic stresses like drought, cold and salt. Similar results were found in *Medicago, Aleuropus, Populus, Gossypium, Cucumis* and many more.

In a detailed study of this unique class of genes by Vij and Tyagi (2006), this gene family was reported to have a number of *SAPs* in rice (18) and *Arabidopsis* (14) by genome wide analysis of derived A20/AN1 ZFPs. Different abiotic stress related cis-regulatory elements present in the promoter region also provided the clues for the possibility of regulation of *SAPs* by different abiotic stresses.

Structural analysis showed the presence of A20/AN1 domains in plant species in various patterns providing unique identity for their categorization. Earlier studies showed *SAPs* as intron less genes but later many of these were found to have introns. Transient expression/ sub-cellular localization of these proteins revealed that most of them are nuclear and cytoplasmic or membrane associated.

Uniqueness to this gene family was its early stress inducible nature. Stress exposure led the transcript accumulation in response to stress within very less time and a decrease in expression levels was seen thereafter indicating the adaptive property of the gene in response to continuous stress (Mukhopadhyay et al., 2004). Upon overexpression of

SAPs in plants, many other stress responsive genes were also found to be induced like *DREBs* which are known for their role in drought stress.

With a variety of reports on SAPs, they have been accredited with different key roles in biotic and abiotic stress tolerance like plant hormone homeostasis and redox sensing (Ströher et al., 2009). NO is an important and well known signalling molecule that is also produced in higher amounts in response to stress upon the overexpression of SAP in plants (Charrier et al., 2012). Although poly-ubiquitination and their interaction with RLCK have been also been reported (Giri et al., 2011; Kang et al. 2011), their mode of action requires more work.

Futuristic approaches including mechanism of action for these novel regulators of stress inducible genes may open new pathways to capture their role in stress tolerance and exploit this property in enhancement of crop plants growing in areas with stressed environment and to dig out unexplored roles in plant defence.

1.2 DISCOVERY OF SAPs IN PLANT KINGDOM

Although A20/AN1 zinc finger domains were found to be present in several eukaryotic proteins and were well characterized much before in animals for their roles in providing immune responses, their role in plant kingdom was unraveled for the very first time when these zinc finger protein genes from rice were overexpressed in tobacco. As a result, Mukhopadhyay et al. (2004) found increased tolerance to cold, dehydration, and salt stress in transgenic tobacco. Overexpression could also result in faster recovery from stress associated injuries such as chlorosis and cell death. This gene was found to be induced at high levels on exposure to different abiotic stress and was named as *OsiSAP1* (*Oryza sativa* subspecies *indica* Stress Associated Protein 1) and all these similar kind of genes with A20/AN1 zinc finger domain were categorized in SAP gene family.

Using protein homology search tools with first identified A20/AN1 ZF protein in rice, 14 such proteins in *Arabidopsis* (see Table 1.1), and 18 in *Oryza sativa* (see Table 1.2) were reported by Vij and Tyagi (2006). Subsequently, Jin et al. (2007) identified 18 *SAP* genes in *Populus trichocarpa*, 11 in *Zea mays* and 2 *SAP* genes in *Chlamydomonas reinhardtii* by phylogenetic analysis. In another study by Kanneganti and Gupta (2008), *OsiSAP8* was found to be a cytoplasmic protein and interaction within its A20 and AN1 domains was also shown. Overexpression of *OsiSAP8* in transgenic rice (homologous) and tobacco (heterologous) provided fair tolerance against high salt, drought, and cold stress (Kanneganti and Gupta, 2008).

Table 1.1. Distribution of *AtSAP* gene family encoding A20/AN1 zinc-finger proteins in the *Arabidopsis* genome.

S.No.	Zinc-finger	Gene_ID	Accession	Chr	Coordinates (5′ to 3′)	Nucleotide length (bp)	Predicted protein (aa)	No. of introns
1	A20-AN1	*AtSAP1*	AT1G12440	1	4,241,590–4,242,970	507	169	1
2	A20-AN1	*AtSAP2*	AT1G51200	1	18,988,102–18,990,236	522	174	2
3	A20-AN1	*AtSAP3*	AT2G27580	2	11,783,488–11,784,539	492	164	1
4	A20-AN1	*AtSAP4*	AT2G36320	2	15,235,867–15,237,575	486	162	0
5	A20-AN1	*AtSAP5*	AT3G12630	3	4,012,615–4,013,590	483	161	0
6	A20-AN1	*AtSAP6*	AT3G52800	3	19,580,289–19,581,660	513	171	1
7	A20-AN1	*AtSAP7*	AT4G12040	4	7,214,699–7,215,983	528	176	1
8	A20-AN1	*AtSAP8*	AT4G14225	4	8,198,772–8,199,197	378	126	1
9	A20-AN1	*AtSAP9*	AT4G22820	4	11,987,720–11,989,012	531	177	1
10	A20-AN1	*AtSAP10*	AT4G25380	4	12,975,946–12,976,338	393	131	0
11	AN1-AN1-C2H2-C2H2	*AtSAP11*	AT2G41835	2	17,464,864–17,466,632	840	280	1
12	AN1-AN1	*AtSAP12*	AT3G28210	3	10,521,758–10,522,769	561	187	1
13	AN1-AN1-C2H2	*AtSAP13*	AT3G57480	3	21,288,917–21,290,331	750	250	1
14	AN1	*AtSAP14*	AT5G48205	5	19,564,163–19,565,013	576	192	3

(Taken from Vij and Tyagi, 2006)

Table 1.2. Distribution of *OsiSAP* gene family encoding A20/AN1 zinc-finger proteins in the rice genome.

S.No	Zinc-finger	Gene_ID	Locus ID_TIGR	KOME	Chr	Coordinates (5' to 3')	Nucleotide length (bp)	Predicted protein (aa)	No. of introns
1	A20-AN1	*OsSAP1*	LOC_Os09g31200.1	AK106128	9	18,759,705–18,760,622	495	165	0
2	A20-AN1	*OsSAP2*	LOC_Os01g52030.1	AK111476	1	30,231,637–30,230,953	447	149	0
3	A20-AN1	*OsSAP3*	LOC_Os01g56040.1		1	32,593,903–32,596,306	1068	356	1
4	A20-AN1	*OsSAP4*	LOC_Os02g10200.1	AK067403	2	5,347,486–5,345,611	522	174	1
5	A20-AN1	*OsSAP5*	LOC_Os02g32840.1	AK063708	2	19,482,320–19,483,065	465	155	0
6	A20-AN1	*OsSAP6*	LOC_Os03g57890.1	AK102503	3	32,915,847–32,913,214	483	161	1
7	A20-AN1	*OsSAP7*	LOC_Os03g57900.1	AK108205	3	32,921,903–32,921,106	510	170	0
8	A20-AN1	*OsSAP8*	LOC_Os06g41010.1	AK069720	6	24,490,982–24,493,135	516	172	2
9	A20-AN1	*OsSAP9*	LOC_Os07g07350.1	AK121813	7	3,676,709–3,674,265	486	162	1
10	A20-AN1	*OsSAP10*	LOC_Os07g07400.1		7	3,697,428–3,696,970	459	153	0
11	A20-AN1	*OsSAP11*	LOC_Os08g39450.1	AK101969	8	24,821,094–24,822,077	513	171	0
12	A20-A20-AN1	*OsSAP12*	LOC_Os08g33880.1		8	21,072,646–21,073,320	675	225	0
13	AN1	*OsSAP13*	LOC_Os01g51990.1		1	30,207,127–30,205,282	576	192	1
14	AN1	*OsSAP14*	LOC_Os03g57920.1	AK108384	3	32,927,052–32,925,978	714	238	0
15	AN1	*OsSAP15*	LOC_Os05g23470.1	AK072411	5	1,3345,932–13,343,520	525	175	0
16	AN1-AN1-C2H2-C2H2	*OsSAP16*	LOC_Os07g38240.1		7	22,929,069–2,2931,587	873	291	1
17	AN1-AN1	*OsSAP17*	LOC_Os09g21710.1		9	13,154,831–13,153,949	567	189	0
18	A20	*OsSAP18*	LOC_Os07g07370.1		7	3,683,966–3,683,751	216	72	0

(Taken from Vij and Tyagi, 2006)

Following these studies, in 2009 a cDNA clone from salt treated roots was isolated from *Aleuropus littoralis* plant which could encode a polypeptide of 159 amino acids (Ben Saad et al., 2010). InterProScan at NCBI and conserved domain search could predict the presence of two zinc-finger domains in the gene and led to categorization of this gene in SAP gene family.

Till now, many plant SAPs have been characterized by genome wide analysis and single gene approach to explore their roles in stress tolerance reportedly in *Medicago* by Charrier et al. (2012), *Malus domestica* by Dong et al. (2018), *Poplar* by Lloret et al. (2017), cucumber by Lai et al. (2020), and many more. And with these studies, key roles of SAPs in abiotic and biotic stress tolerance with providing immunity to plants have been revealed.

1.3 BASIC MOLECULAR STRUCTURE OF SAPs

SAP genes have a general feature of containing either A20 or AN1 or both A20 and AN1 zinc finger domains at the N and C terminal of their protein structure (see Figure 1.1). "Zinc fingers" are basically sequence patterns encoded in a variety of genes for nucleic acid binding purposes and they make tetrahedral zinc coordination complexes with loop configurations (Miller et al., 1985). Presence of zinc fingers helps SAP proteins to fold and provide stability to overall protein conformation.



Figure 1.1. Basic domain organisation of *SAP* gene coded protein with an example of both A20 and AN1 organised at N and C terminals.

❖ A20 type ZFP domain

A20 protein gene is a member of large family of zinc finger protein coding genes (Opipari et al., 1990). Novel A20 zinc-finger domain was first identified as a part of TNFα-inducible protein in endothelial cells of human umbilical vein (Dixit et al., 1990). Opipari et al. (1990) were successful in isolating A20 cDNA clone for the first time. TNF (Tumor Necrosis Factor) was likely to induce the expression of many primary response genes. Among them a novel cDNA clone was obtained that was transient and rapidly inducible (within 15 min) on TNF treatment. This newly identified protein was novel in a way that it was not identical to any known gene product as far as sequence similarity was concerned. A20 protein in itself contained repeated elements of seven new zinc finger motifs. Alignment results of repetitive sequences showed the presence of conserved cysteine residues in a pattern similar to the Cys_2/Cys_2 zinc finger motifs (see Figure 1.2 and 1.3) in the form of Cys-X_4-Cys-X_{11}-Cys-X_2-Cys (6 times) ,Cys-X_2-Cys-X_{11}-Cys-X_2-Cys (1 time), present continuously. Finger loop domain of 11 amino acids defined this class as a new family of protein coding genes (Opipari et al., 1990). In humans, A20 zinc finger protein acts as a negative regulator of tumor necrosis factor (TNF)-induced signalling pathways leading to apoptosis, stress response and inflammation (Lademann et al., 2001).

A20 protein domain in human, also acts as a cellular inhibitor of Nuclear Factor-kB dependent gene expression and apoptosis (Beyaert et al., 2000). In humans, A20 gene encodes a 790 amino acids long protein with molecular size of 90 kDa. This kind of inhibition of apoptosis was correlated with inhibition of phospholipase A2 activation, reduced collapse of mitochondrial membrane potential, reduced ROS (Reactive Oxygen Species) production. Such inhibitory role has potential to prove itself to be a potent therapeutic treatment of inflammatory diseases in future.

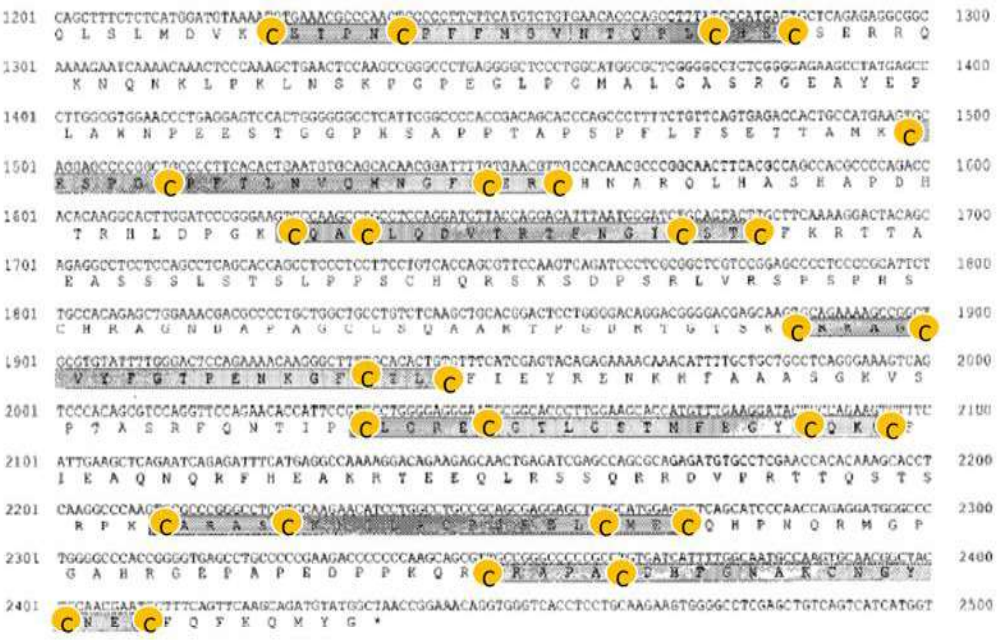


Figure 1.2. Nucleotide and predicted protein sequence of A20 cDNA (Cysteine residues at N and C-terminal of the repeated motifs are highlighted in yellow colour).

(Adapted from Opipari et al. (1990))

❖ AN1 type ZFP domain

Originally, this domain was first identified in a protein localized in animal hemisphere 1 at vegetal axis of *Xenopus laevis eggs* and its early embryos. Three kb long maternal mRNA of AN1 protein codes for two isoforms of protein named An1a and AN1b. Each protein contains N-terminal domain that has 50 per cent identity to *Xenopus* ubiquitin (Dworkin et al., 1984). Homology results also indicated the presence of a Cys/His-rich consensus sequence with zinc finger in all three proteins at both termini. In AN1 protein, two pairs of cysteine residues or alternatively cysteine2-histidine2 residues were found (Linnen et al., 1993) (see Figure 1.3).

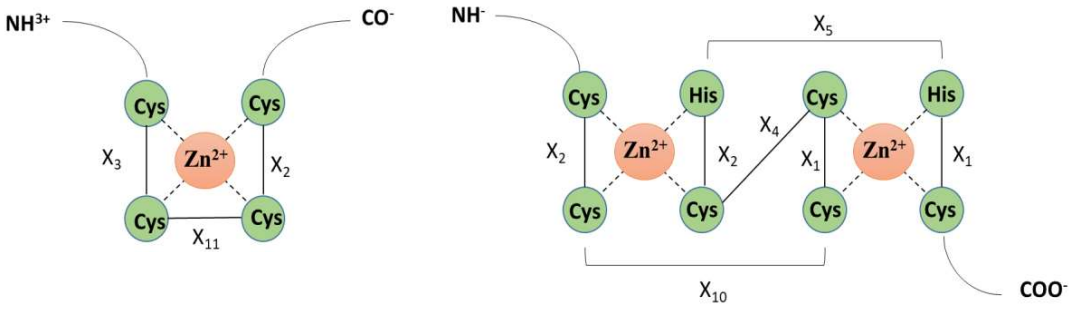


Figure 1.3. Primary sequence organization of A20 and AN1 domains. (Xn: the number of amino acid residues between zinc ligands).

A20/AN1 zinc finger genes widely express in plants and their encoded proteins can also be divided into two different types, Type I and Type II. Type I genes contain the traditional pattern like CX(2)CX(9,12)CX(1,2)CX(4)CX(2)HX(5)HXC and Type II contain the expanded pattern like CX(4)CX(9,12)CX(1,2)CX(4)CX(2)HX(5)HXC of the ZFP domain (Jin et al., 2007). Another characteristic difference between both the types is the absence of introns with A20 domain in Type I and presence of introns in Type II. Basic organisation of A20 and AN1 domains is represented in figure 1.3. Jin et al. (2007) suggested that type I may have a role in ubiquitin pathway and Type II may have a role in transcriptional regulation.

1.4 DIFFERENT TYPES OF STRESS ASSOCIATED PROTEINS (SAPs) IN DIFFERENT PLANT SPECIES

On basis of domain organization and different combinations of A20/AN1 domains, seven different types of SAPs are known till date (see Table 1.3 for examples) and are listed below:

1. A20
2. A20-AN1
3. A20-A20-AN1
4. AN1
5. AN1-AN1
6. AN1-AN1- C2H2
7. AN1-AN1- C2H2-C2H2

The plant species in which these special domains are found are given below in details:

1. **A20**: This type of domain organization with only A20 at N-terminal of protein can be found in species like *Oryza sativa* (1) as OsSAP18 and *Brassica napus* (7) etc.
2. **A20-AN1**: The most prevalent type of arrangement of A20/AN1 domains at N and C terminal of protein is in SAPs of all plants species searched including 11 in rice, 10 in *Arabidopsis thaliana*, 14 in *Populus tricocarpa,* and 32 in *Brassica napus.*
3. **A20-A20-AN1**: Very rare kind of arrangement which includes 2 successive A20 domains positioned at N terminal and 1 A20 domain at C terminal, respectively. It is found in *Oryza sativa* as OsSAP12.
4. **AN1:** In this type, only AN1 domain at N terminal is present and is found in *Oryza sativa* (3), *Gossypium hirsutum* (3), and *Brassica napus* (9).
5. **AN1-AN1:** This kind of arrangement contains 2 AN1 domains successively at N terminus and is present in *Oryza sativa* (1) and *G.hirsutum* (4).
6. **AN1-AN1-C2H2:** Again a very rare kind of arrangement present in Arabidopsis as AtSAP13.
7. **AN1-AN1- C2H2-C2H2:** Unique kind of arrangement found in *Oryza sativa* as OsSAP16 and in *Arabidopsis* as AtSAP11.

Table 1.3. Number and domain organization of various SAPs of some major crop plants.

S. no.	Organism	No. of SAP members	A20	A20-AN1	A20-A20-AN1	AN1	AN1-AN1	AN1-AN1-C2H2	AN1-AN1-C2H2-C2H2	References
1	*Oryza sativa*	18	1	11	1	3	1	0	1	Vij and Tyagi., 2006
2	*Arabidopsis thaliana*	14	0	10	0	1	1	1	1	Vij and Tyagi., 2006
3	*Zea mays*	11	0	8	0	1	2	0	0	Jin et al., 2007
4	*Populus trichocarpa*	19	1	14	0	2	1	0	1	Li et al., 2019, Jin et al., 2007
5	*Solanum lycopersicum*	13	0	9	0	1	2	0	1	Solanke et al., 2009
6	*Camellia sinensis*	11	0	10	0	1	0	0	0	Paul & Kumar., 2015
7	*Gossypium hirsutum*	37	0	28	0	3	4	0	2	Gao et al., 2016
8	*Populus euphratica*	18	0	15	0	0	2	0	1	Jia et al., 2016
10	*Medicago truncatula*	17	0	10	0	4	1	0	2	Zhou et al., 2017
12	*Malus domestica*	30	3	23	0	0	3	0	1	Dong et al., 2018
14	*Artemisia annua*	16	4	9	0	3	0	0	0	Wang et al., 2019
15	*Vitis vinifera*	15	0	12	0	3	0	0	0	Ding et al., 2019
16	*Brassica napus*	57	7	32	0	9	5	4	0	He et al., 2019
18	*Glycine max*	27	1	18	0	2	3	3	0	Zhang et al.,2019
19	*Cucumis sativus*	12	0	9	0	1	1	0	1	Lai et al.,2020

1.5 SUB-CELLULAR LOCALIZATION OF *SAP* GENES

Transient expression or sub-cellular localization assays basically help in determining the compartmentalization of the proteins concerned. Following are some selected examples which explain the localization of SAPs in plant cells:

- In silico analysis of *OsiSAP8* gene product did not find any kind of NLS (Nuclear Localization Signal) in its protein sequence. Transient expression study was done for *OsiSAP8* by Kanneganti and Gupta (2008) by fusing coding region of *OsiSAP8* with GFP in epidermal cells of *Nicotiana benthamiana* leaves. As a result, SAP8-GFP fusion protein was found to be localized to cell cytoplasm only, indicating clearly that SAP8 was a cytoplasmic protein (Kanneganti and Gupta, 2008) (see Figure 1.4).

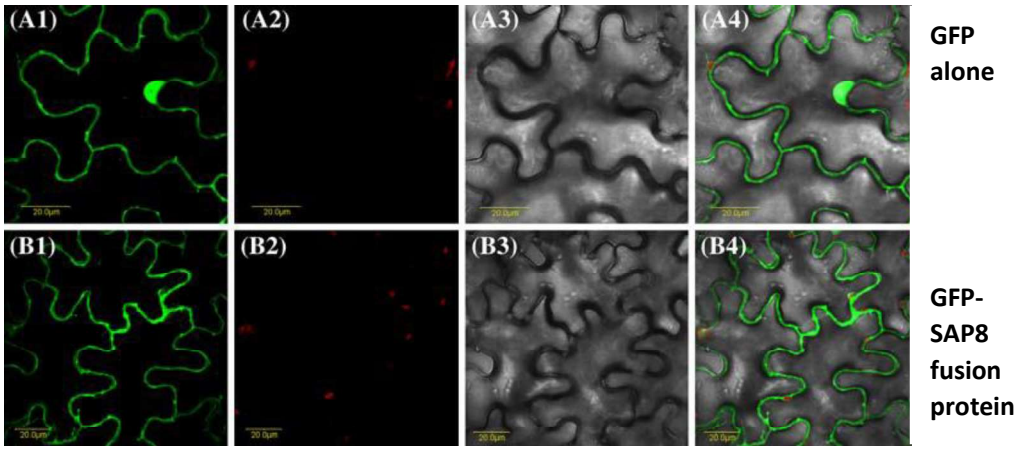


Figure 1.4. Sub-cellular localization of GFP alone and GFP-SAP8 fusion protein during transient expression in *N. benthamiana* leaves. Picture 1-4 represent GFP fluorescence (green channel), red channel (autofluorescence of plant tissue), bright field channel and channel overlays of green, red and bright field channel, respectively. Free GFP was expressed both in the cytoplasm and nuclei (A1–A4) whereas GFP-SAP8 fusion protein was localized only in the cytoplasm (B1–B4).

(Adapted from Kanneganti and Gupta (2008))

- In another experiment, nuclear localization of *AtSAP10* gene of *A. thaliana* was predicted using the WoLF PSORT protein localization program (Dixit and Dhankher, 2011). For confirmation of results, leaf protoplasts from *Arabidopsis* plants with *AtSAP10-eGFP* chimeric fusion construct were studied

for sub-cellular localization experiment of *AtSAP10* gene product. When observed under different planes, typical nuclear as well as cytoplasmic localization could be clearly seen. Further validation was done in root tissue samples which confirmed the presence of nuclear and cytoplasmic localization on whole root sections with *AtSAP10-eGFP* fusion protein.

- Another experiment of sub-cellular localization proved OsiSAP1 to be nuclear localized and OsiSAP11 to be cytoplasmic localized in addition to the nucleus and plasma membrane (Giri et al., 2011).
- A single bipartite NLS was predicted in *MusaSAP1* protein in another study by Sreedharan et al. (2012) using cNLS mapper. Nuclear and cytoplasmic localization of *MusaSAP1* protein confirmed the *in silico* analysis when fusion constructs were bombarded on onion peel cells.
- When SAP gene from *Arabidopsis thaliana AtSAP13* was checked for its transient expression, *in silico* analysis results perfectly matched the expression assay done in vitro. Both cytoplasmic and nuclear localization of the AtSAP13 protein was predicted using WoLF PSORT which was even confirmed by sub-cellular localization of *AtSAP13* gene product clearly demarcating that protein was cytoplasmic and nuclear localized (Dixit et al., 2018).

In conclusion, most of the SAPs do have a nuclear and/or cytoplasmic localization which basically helps in understanding the nature of action of SAPs for downstream functioning by protein-protein interactions. Nuclear localization may show a role of SAP proteins acting as transcription factors and performing further molecular function by DNA protein interaction during stress.

1.6 EARLY STRESS INDUCIBILITY OF *SAP* GENE FAMILY MEMBERS

Healthy growth and development of a plant is controlled by proper functioning and regulation of some crucial genes at various time points in its life cycle. This proper functioning can be affected by various environmental factors which can include several different types of abiotic stresses. These stresses may ultimately lead to either non-

functioning of these genes or induction of some genes for the production of unique transcripts in response to combat these stresses for proper physiology and metabolism.

Stress induced expression of *OsiSAP1* was seen by Mukhopadhyay et al. (2004) on treatment of rice seedlings with various abiotic stresses like cold, heat, ABA, heavy metals, etc. Transcript levels were found to be continuously increasing for 3 hours and remained elevated till 12 long hours when seedlings were treated with cold stress. Different kind of *OsiSAP1* induction kinetics could be observed with different type of abiotic stresses. As the mRNA levels of *OsiSAP1* were high after stress, it could also be noticed that, in addition to stress inducibility, the gene was early inducible in nature (see Figure 1.5). Early stress inducibility of *OsiSAP1* may have a role in protection of plant from stress and its gene product might be required to do so.

Kanneganti and Gupta (2008) showed that *SAP8* transcript levels were increased within 2 hours of cold stress in seedlings, remained same till next 6 hours, and then started declining. Salt stress of seedlings led the mRNA levels to be elevated till 12 hrs and drought stress till 24 hrs. This all proved *SAP8* to be a multiple stress inducible gene. Solanke et al. (2009) characterized all 13 *SAP* genes of tomato to check the involvement of these genes in 9 different kinds of environmental stress. Among them, *SlSAP6* showed ~83-fold higher transcript levels in salt treated tomato seedlings as compared to unstressed samples. Approximately 24 fold increase in expression levels of *SlSAP11* was observed when seedlings were given wounding stress. Almost all tomato *SAP* genes were induced by salt, desiccation and ABA treatment of seedlings.

Time-dependent expression profiling of *Arabidopsis AtSAP10* gene revealed an increase in its transcript levels after 12 and 24 hours when treated with heavy metals like AsIII, AsV, Cd, and Zn. An increase in *AtSAP10* mRNA levels within 30 minutes and continuous rise till next 6 hours was seen on exposure of seedlings to heat and cold stress (Dixit and Dhankher, 2011). Exposure of *A. littoralis* plants to PEG, heat and cold stress led the transcript levels remain elevated till 6 days and then they start declining (Ben Saad et al., 2010). PEG and mannitol if present in media generally mimic osmotic stress. 20% PEG6000 solution was used to check the expression level of transcripts of *TaSAP5* gene (Zhang et al., 2017) and levels of homologous *TaSAP5-A, TaSAP5-B*, and *TaSAP5-D* were found to be increasing till 8 hrs. These results suggested that *TaSAP5* can be induced by osmotic stress.

Differential induction patterns can also be seen in many cases like *ZFP177/OsSAP9* from rice which is upregulated in response to heat, cold, and drought stress but is downregulated by treating rice seedlings with salt stress (Huang et al., 2008). Jia et al. (2016) showed that all desert poplar *PeuSAP*s are induced by drought, salt, and heat stresses with different induction patterns, respectively. For example, induction within 1 hour of PEG treatment could be seen in *PeuSAP2, PeuSAP6,* and *PeuSAP18* mRNA levels with a dramatic decline after 6 and 12 hours of stress. Contrastingly, downregulation of *PeuSAP3, PeuSAP5, PeuSAP7/9, PeuSAP13*, and *PeuSAP15* for 1 hour was prominent before induction at 6 and 12 hours. All cucumber SAPs, *CsSAPs* are upregulated during drought stress within 6 hours of treatment and then transcript levels started declining till next 24 hrs *CsSAP5* and *CsSAP9* were upregulated and further peaked at 24 hrs of salt stress treatment (Lai et al., 2020).

In conclusion, early induction/expression of *SAP*s results in rapid transcript accumulation in response to stress within hours and this transcript accumulation can vary for various stresses following different induction kinetics. These studies also helped in elucidating potential role of *SAP* genes in stress responses.

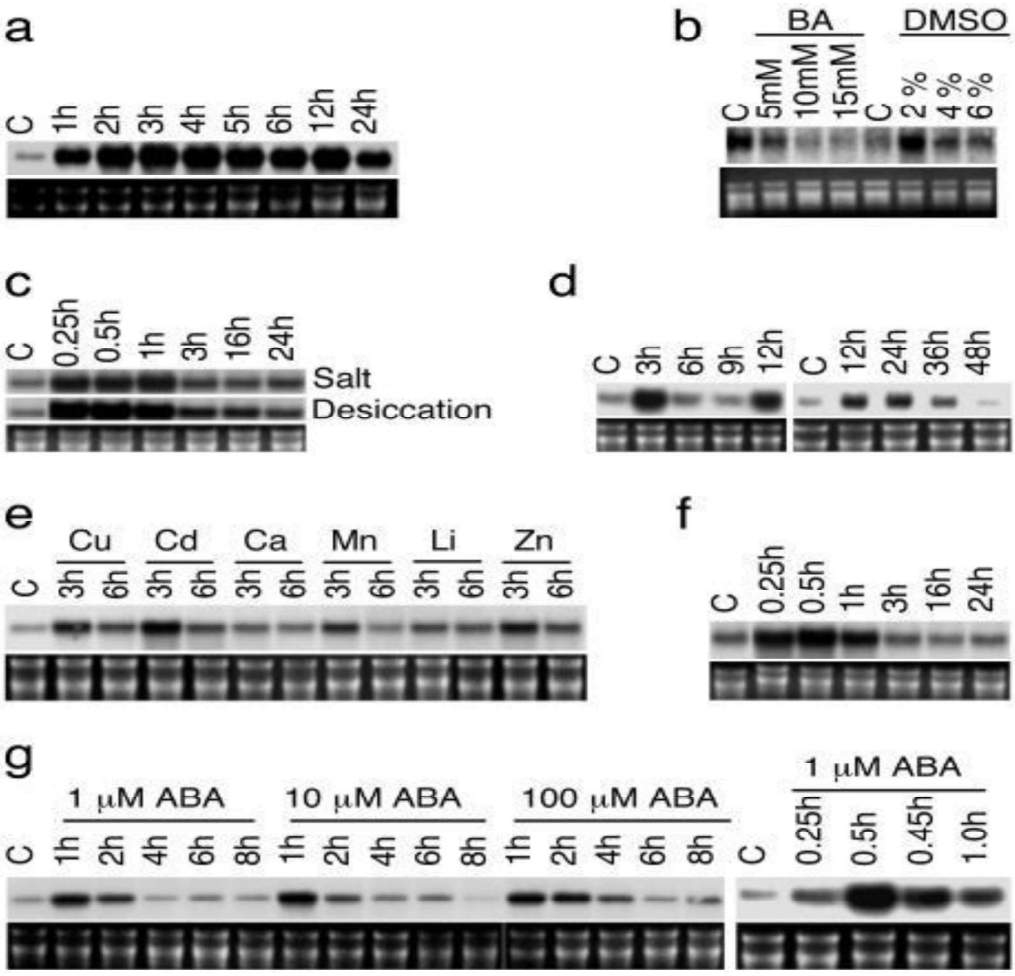


Figure 1.5. Expression pattern of *OSISAP1* after different stresses to rice seedlings.

(Taken from Mukhopadhyay et al. (2004))

1.7 TISSUE AND DEVELOPMENTAL STAGE SPECIFICITY OF *SAP* GENE EXPRESSION

Expression of SAP genes in most of the research work done was found to be tissue specific in nature. *CsSAP09* from *Crocus sativus* was found to be highly expressed in stigma of flower during anthesis stage as compared to other floral parts (Malik and Ashraf, 2017). Histochemical GUS analysis of *AtSAP10* gene promoter showed prominent GUS staining in floral parts like petals, stamen and in roots whereas no GUS expression was seen in epicotyl, stems, and leaves of *AtSAP10p-GUS* transgenic lines at any developmental stages (Dixit and Dhankher, 2011).

Tissue-specific expression analysis of *AtSAP13* gene resulted in GUS expression in shoots significantly in vascular tissue, hydathodes and apical meristem tissue. Staining was also seen in developing seeds in siliques (Dixit et al., 2018). Almost all of **(***Gossypium hirsutum SAP) GhSAP* genes do express in either petal or stamen. These results may indicate that this gene may be associated with developmental regulation of *Gossypium* flowers. In contrast, only a few *SAP* genes (*GhSAP1A/D*, *GhSAP14A/D*, *GhSAP12D*, and *GhSAP18A*) exhibit high expression in ovule and flowers (Gao et al., 2016).

1.8 *CIS* ACTING REGULATORY ELEMENTS OF SAP GENES

Cis regulatory elements (CREs) play a vital role in gene regulatory network which controls the overall growth and development of plants. They themselves do not code for anything but control the expression of neighbouring genes by binding to the transcription factors. *In silico* analysis of SAP genes also revealed such CREs in their vicinity which are thought to control the expression of *SAP* genes. Ultimately, CREs help *SAP* genes to upregulate or downregulate in response to stress.

Several abiotic stress responsive elements from rice *SAP* family members like ABA responsive element (ABRE) in ABA stress, coordinate regulatory element for antioxidant defence (CORE) in antioxidant stress, dehydration-responsive element (DRE/CRT) in dehydration stress, low temperature responsive element (LTRE) in low temperature stress, and a heat-shock element (HSE) in heat stress have been reported by Vij and Tyagi (2006).

In silico analysis of *AtSAP10* gene with 1 kb upstream region predicted the presence of CCAAT box which was thought to be involved in increasing promoter activity of *AtSAP10* gene with help of HSE (Dixit and Dhankher, 2011). Similar analysis by Kang et al. (2011) with *AtSAP5* detected various *cis*-acting elements like ABA response element (ABRE), drought response element (DRE), low temperature response elements (LTRE), and MYB/MYC elements, which are well known for their roles in transcriptional regulation pathways.

CREs in promoter of *GhSAPs* (*Gosssypium*) were reported by Gao et al. (2016) where they found a variety of *cis*-elements related to plant stress, pathogen defence and hormone response. HSE (heat shock element) was found in 28 *SAP* promoters, which are involved in heat stress response. 20 *SAP* promoters were found to be containing ABA-responsive element (ABRE). An MBS motif, that is involved in drought response was found in 26 *SAP* promoters explaining the role of cotton *SAP* genes in various abiotic responses. Gibberellin-responsive elements (GARE), wound response elements (WUN-motif), cold-related motifs (LTR) and pathogen response elements (Box-W1 and TC-rich repeats) indicated the possible role of SAPs in hormone signalling and biotic stress with overall plant defence in stressed conditions.

A recent study with cucumber *SAP*s, to find out *cis* acting elements in the promoter region of *CsSAP*s was done by Lai et al. (2020). With the help of PlantCARE database, they were successful in finding out seven elements for stress and nine types of *cis* elements for hormone response, respectively. Out of total 12 *CsSAP* genes, ARE element was present in 10 genes with a frequency of at least one. Different responsive elements for various hormones including those for ABA (ABRE), ethylene (ERE), salicylic acid (TCA-element), auxin (AuxRR-core and TGA-element), MeJA (CGTCA-motif) and gibberellin (P-box, GARE-motif and TATC-box) were also detected.

However, mere presence of any *cis* acting element does not support its involvement in the upregulation of gene expression in abiotic stress response, similarly absence of CREs does not mean no upregulation of stress responsive gene expression. But, all these findings hinted the potential role of *SAP* genes in stress and hormonal responsive signalling.

1.9 CHARACTERIZED AND PUTATIVE ROLES OF *SAP* GENES

From discovery of *SAPs* in plants till now, diverse roles of them have been reported in various species in providing tolerance from one or the other stress. Majority of them help in mitigating abiotic stress. Some *SAPs* do have roles in hormone homeostasis, heavy metal detoxification, redox sensing, increasing yield, pathogen attack tolerance, etc. (Giri et al., 2013). Some of the major characterized roles of SAPs are as follows:

1.9.1 SAPs are involved in abiotic stress tolerance in plants

First characterized rice *OsSAP1* showed increased tolerance to various abiotic stresses like drought, cold, mechanical wounding and submergence etc. (Mukhopadhyay et al., 2004). *MusaSAP1* overexpressing plants showed higher tolerance and less damage by drought and salt stress (Sreedharan et al., 2012). Role of *Arabidopsis SAP* genes in heavy metal tolerance was shown by (Dixit and Dhankher, 2011). *AtSAP10* overexpressors were found to be more tolerant to heavy metals like Ni, Zn, Mn in salt forms. Even at toxic concentrations of these metals, these overexpressor plants were having approximately three times higher shoot biomass and longer roots as compared to WT control plants. Similarly, overexpression of *AtSAP13* in *Arabidopsis* showed severe tolerance to As, Cd, and Zn toxicity (Dixit et al., 2018). The germination rate of *LmSAP*-overexpressing plants was significantly higher in presence of heavy metals as compared to NT lines (Saad et al., 2018). Increased tolerance could be clearly observed in *LmSAP* transgenic lines which was even supported by higher shoot biomass and well developed root system in presence of heavy metal stress of Zn, Cu, Cd and Mn as compared to NT lines of *Lobularia maritima. TaSAP17-D* when overexpressed in *Arabidopsis* improved tolerance to salt stress (Xu et al., 2018).

Medicago SAP when overexpressed in tobacco revealed a better germination rate as well as fresh weight as compared to WT control line when given osmotic stress and were more salt stress tolerant than WT control tobacco plants (Charrier et al., 2012). *MdSAP15* overexpression in *Arabidopsis* showed enhanced tolerance and higher survival rate when they were analysed on grounds of water deficit stress i.e. *MdSAP15* could enhance the degree of drought tolerance in *Arabidopsis* (Dong et al., 2018). *ZFP177* overexpression in tobacco showed an enhanced heat and cold stress tolerance at seedling stage (Huang et al., 2008). Another study with *OsiSAP8* overexpression in rice and tobacco showed lower chlorophyll loss due to salt stress than WT plants (Kanneganti and Gupta, 2008). In drought stress, *PpSAP1* overexpression in *Prunus persica* helps in higher water retention as compared to WT plants echoing the role of *PpSAP1* in drought stress tolerance (Lloret et al., 2017).

With a string of positive roles, differential expression of *SAP*s can also be seen in some cases to increase sensitivity. OE of *ZmAN13* showed an enhanced

tolerance to cold stress but plants proved to be sensitive to salt and drought stress (Xuan et al., 2011). By observing above results, it is clear that SAPs have a potential to be exploited to mitigate abiotic stress in plants. Particularly, their overexpression with stress inducible promoters can prove them as ideal candidates for conferring abiotic stress tolerance in plants.

1.9.2 Plant SAPs confer biotic stress tolerance

SAPs are also involved in mitigating various biotic stresses in plants like resistance to fungi, bacteria, etc. *SAP* genes respond to various biotic stimuli, and some examples demonstrate that overexpression of SAPs can impart sensitivity or tolerance to biotic stimulus. *OsSAP1* overexpressing transgenic line showed delayed disease symptoms and restricted lesions on tobacco leaves as compared to WT plants (Tyagi et al., 2014). Transient expression of *SlSAP* gene in *Nicotiana benthamiana* could resist the attack of fungus *Botrytis cinerea* to a much higher extent with smaller lesions as compared to control plants. Silencing of *SlSAP4* and *SlSAP10* attenuated the immunity of tomato (*Solanum lycopersicum*) plants from necrotrophic fungal attack by *B. cinerea* (Liu et al., 2019). Overexpressing *AtSAP9* gene in *Arabidopsis* plants increased susceptibility of plants to non-host pathogen *Pseudomonas syringae* pv. *Phaseolicola* which clearly reflected the possible role of *AtSAP9* gene in disease resistance (Kang et al., 2017).

1.9.3 Hormone homeostasis is also regulated by SAPs

Plant hormone biosynthesis and degradation is controlled by various set of regulatory pathways which in turn are responsible for the major phenotypic characters of these sessile organisms. Regulation of pathways can be controlled by feed-back or feed-forward mechanisms based on requirements for active forms of hormones for predestined growth and development of plants. A20/AN1 ZFP gene may also have a role in homeostasis of plant hormones was found by Liu et al. (2011). One such A20/AN1 ZFP gene from rice, *OsDOG* (*Oryza sativa dwarf rice with overexpression of gibberellin-induced gene*), when overexpressed in rice showed a dwarf phenotype as compared to WT plants. OE also led to reduction of

GA1, the major bioactive GA in rice to approximately half the levels as WT had. These results could show the possible role of *OsDOG* in GA homeostasis. Another example is when OE of *OsiSAP7* in *Arabidopsis* showed insensitivity to ABA at seed germination stage and hindered the stomatal closure even after external application of ABA (Sharma et al., 2015). This may have occurred due to impaired ABA signalling at seed germination stage which is necessary to break the dormancy of seeds for germination.

1.10 DIFFERENT OPERATIONAL MECHANISMS OF *SAP*s

Despite the fact that SAPs containing A20/AN1 domains play pivotal roles in stress tolerance in transgenic plants and with continuously increasing knowledge of differential expression of these genes with their protein structure being categorized for its variety in organization, still the way they work is not fully unrevealed. However, with some cues from earlier work and newly emerging roles, SAPs have been attributed with some roles of being involved in E3 ubiquitin ligase activity, acting as redox sensor, and regulation of gene expression by protein-protein interactions. Some of these roles are mentioned below in details:

1.10.1 SAPs show E3 ubiquitin ligase and protein degradation activity for stress responsive signalling

A20/AN1 proteins may have E3 ubiquitin ligase activity was first suggested by Kanneganti and Gupta (2008). AtSAP5 showed auto-ubiquitination in vitro by E3 ubiquitin ligase activity (Kang et al., 2011). When ubiquitin, yeast E1, and human E2 were present in the system, AtSAP5-GST was able to make poly-ubiquitin chains within 15 minutes. However, when tested without these components, these chains were not being formed proving that AtSAP5 had ubiquitin ligase activity**.** In another study by Choi et al. (2012) proved that AtSAP5 does not bind with mono-ubiquitin and can recognise linkage specific poly-ubiquitin chains. OsiSAP7 also showed *in vitro* E3 ubiquitin ligase activity as reported by Sharma et al. (2015).

Under stressed conditions, plants modify their proteome through ubiquitin 26S proteasome pathway to tolerate stress. In wheat, TaSAP5 was able to make poly-ubiquitin chains by ligase activity and could also degrade the DRIP proteins via 26S proteasome degradation pathway (Zhang et al., 2017). Kang et al. (2017) showed the interaction AtSAP9 with Rad23d through yeast two-hybrid assays. Rad23d is a member of RADIATION SENSITIVE23 family which are known to act as shuttle factor during the delivery of putative protein moieties to 26S proteasome. As per BiFC analysis, in vivo interaction of both the proteins was restricted to nucleus by formation of a complex with Rad23d isoforms. This study suggested the possible binding of ubiquitinated proteins to Rad23d and finally their degradation by 26S proteasome. Above results indicate that SAPs with recognition and binding of ubiquitin moieties may play cellular roles in regulating cytoplasmic proteins in stressed conditions by participating in stress responsive signalling pathways.

1.10.2 Redox sensing: a unique feature of SAPs

To tolerate the stressed conditions in environment, plants may activate some tolerance mechanisms to protect its proper physiological functioning. These signalling mechanisms may cause oxidation or reduction of biomolecules making the system change redox status in a way which can result in change in conformation of such molecules. The most redox reactive amino acid is cysteine which in ZFPs coordinates the zinc ions and is a characteristic feature of domain organization of SAPs. *AtSAP12* may undergo redox dependent change in conformation of its molecular structure. It was seen that under oxidized conditions, zinc ions were released from ZFP structure and can regain the conformation again when the redox potential changes. This clearly indicates that changing redox conditions of environment can lead the cysteine of ZF structure residues to undergo redox reactions and finally result in reversible conformational changes of SAPs (Ströher et al., 2009).

This feature of SAPs acting as redox sensors may provide information to neighbouring proteins that have binding affinities to them. This may further activate or inhibit the binding of interacting proteins under stressed conditions for their downstream regulatory functions.

1.10.3 SAPs work by modulating gene expression

OE of SAPs in plants results in differential gene expression patterns of stress related genes which can mitigate stress or may show sensitive phenotype for that particular stress. OE of *AlSAP* in rice enhanced transcript accumulation of eight stress responsive genes in cold and drought stress conditions. These included variety of genes like *OsDREB1A, OsDREB1B* (DRE-binding proteins), *OsP5Cs* (pyrroline-5-carboxylatesynthetase), *OsHKT1* (a Na+/K+ transporter), *OsSOS1* (Na+/H+ antiporter), *OsProT* (a proline transporter), *OsSodA* and *OsAPX1 (genes* encoding manganese superoxide dismutase and ascorbate peroxidase) (Ben Saad et al., 2012). These genes were having two to three times higher transcript levels as compared to WT control plants suggesting that SAPs can induce the expression of other regulatory proteins for stress adaptation.

In many cases it has been seen, that when a plant encounters stress the amount of some specific amino acids increases with an increase in sugar metabolism genes as well. Zahur et al. (2012) identified the binding of cotton SAP *GaZnF* to a MYB-box element present in the promoter of *GUSP1* a cotton stress responsive gene. However, direct binding of SAPs to DNA structure or acting as a TF and driving the transcription has not been reported till now but the change in transcriptome due to SAP OE can definitely be called as a major mechanism of action of these novel proteins.

1.10.4 SAPs work by protein-protein interactions

Bio-molecular interactions in the cell are responsible for physiological and metabolic functions and ultimately proper growth and development of plant. Rice A20/AN1 proteins OsSAP1/OsSAP11 were found to have an interaction with a stress responsive receptor like cytoplasmic kinase OsRLCK253 through A20 domain in *Arabidopsis* (Giri et al., 2011). SAP1/11 could interact with themselves and each other as well. FRET results reflected the strong interaction of OsSAP11 and OsSAP1 with OsRLCK253 at nuclear as well as plasma membrane.

These results suggested the possible mode of action of SAP by protein-protein interactions by receiving the signal at membrane and transduction through cytoplasmic kinase which help in conferring biotic and abiotic stress tolerance to plants.

TASAP5 was also found to be interacting *in vivo* with TaDRIP, AtDRIP1 and AtDRIP2 as well. OE of *TaSAP5* in *Arabidopsis* enhanced the degradation of DRIP proteins during drought stress and increased the levels of DREB2A proteins which have a role in drought tolerance (Zhang et al., 2017).

Above studies indicate that after perceiving any stress, plant SAPs interact with other proteins through formation of homodimers or oligomers to drive the downstream regulation in mitigating stress tolerance.

1.11 CONCLUSIONS AND FUTURE PROSPECTS

A20/AN1 containing ZFPs confer biotic and abiotic stress tolerance in plants. This tolerance can be achieved by signal perception and further downstream regulation of gene expression achieved by reversible conformational changes of these proteins. In animal counterparts, A20 and AN1 domains have their unique roles in providing immunity to the organism and act as negative regulators of tumour. In plants they can provide immunity against pathogens, but its mechanism of action is not fully understood. Ubiquitination and protein dimerization also need further clarity in their action. Nuclear and cytoplasmic localization of SAPs is a prompt for their mechanism of action and their possible role of acting as TFs which after binding to *cis* regulatory elements may drive the gene expression in defence against stress. However, a lot of further analysis is required to unravel the possible roles of these proteins and their mechanism by which they work. In future, this property of these proteins can be fairly exploited in protecting crop plants from abiotic stresses which leads to heavy losses in crop yield per year.

OE of these proteins with stress inducible promoters can mitigate the limitations of crop plants to be grown in regions with high abiotic stresses like drought, cold, heavy metal, heat, etc. Immunity provided by SAPs to plants like tomato against bacterial and fungal pathogens should be explored more to propagate this property in plants which are more prone to these biotic stresses. Molecular approaches to overexpress these proteins with stress inducible promoters in crop plant species like rice, wheat and corn can be a good approach to avoid any trade-off with yield of crop in field. This would be a new shift of crop improvement from traditional breeding programmes and transformation moving to just orchestrating the resident genome in mitigating biotic and abiotic stress for higher and quality yield production of crop plants.

Chapter 2

Materials and Methods

2.1 *OsiSAP15*

OsiSAP15 is a 525 bp intron less gene encoding a protein containing AN1 domain at C terminal. This gene is present on rice chromosome 5 and codes for ~175 amino acid long protein (see Appendix 1 for CDS). Overexpression lines of *OsiSAP15* in *Arabidopsis* under stress inducible promoter *rd29A* were obtained from the Ph.D. work of Sharma (2015). To know the putative role of the gene in drought tolerance, T5 seeds of *Arabidopsis* with 3:1 segregation ratio were used.

2.2 Phenotypic analysis of *OsiSAP15* overexpression plants in water deficit stress

2.2.1 Plant materials and growth conditions

Seeds of wild-type (WT) and transgenic *Arabidopsis thaliana* ecotype Columbia were surface-sterilized with 70% ethanol (v/v) for 2 minutes and rinsed twice with sterile MQ water. Then, these seeds were treated with 4% sodium hypochlorite solution (v/v) for 1 minute and rinsed with sterile MQ water for 1 minute in sterile laminar air flow for 4-5 times. Seeds were inoculated onto square shaped Petri plates containing half-strength Murashige and Skoog (MS) medium with vitamins (Murashige and Skoog, 1962), 1% sucrose and 0.4% Phytagel. WT and transgenic seeds were inoculated on the same plate in sections.

Before transfer to growth conditions, seeds were stratified at 4°C for 48 hours. Then plates with inoculated seeds were transferred to the *Arabidopsis* growth room with cycling growth condition of 16 hours light and 8 hours dark at 22°C and 18°C, respectively.

2.2.2 Drought stress treatment at seedling stage

For phenotypic assay in Petri plates, wild-type and transgenic seeds were grown on MS Phytagel medium supplemented with 100, 200 and 300 mM mannitol and

sealed Petri plates were incubated vertically in stand in optimum *Arabidopsis* growth conditions. Growth of plants was monitored for 7 to 14 days with respect to root and shoot development. Seedlings were photographed on 15th day and primary root length was recorded with ImageJ software.

2.2.3 Water-deficit stress response phenotype analysis at mature plant stage

To mimic the water-deficit stress, one-week-old WT and transgenic *Arabidopsis* seedlings were transferred to pots containing 1:1 soilrite and cocopeat. Plastic tray containing pots was covered with saran plastic wrap to make a humidity chamber. After 3 days, plastic wrap was removed and seedlings were allowed to grow in pots for 2 weeks. After 3 weeks, pot grown plants were kept devoid of water for 15 days. Unstressed and stressed plants were photographed after stress period. Meanwhile, control plants were regularly watered. Recovery of 1 week was given after stress to allow resurrection of plants. For phenotypic analysis, plants were photographed (bolts cut) before and after stress as well as after recovery.

2.3 Restriction based cloning of *OsiSAP15* gene in pET-28a(+) expression vector

2.3.1 Primer designing

Locus ID (LOC_Os05g23470.1) was searched to retrieve complete CDS sequence of *OsiSAP15* from RGAP (Rice Genome Annotation Project) database (http://rice.plantbiology.msu.edu/). The primers for cloning in pET-28a (+) were designed using Gene Runner software (version 3.05). The unique restriction sites (*EcoRI* and *SalI*) present in multiple cloning site (MCS) of vector, which are absent in gene to be amplified, were added to 5'- end of forward primer and 3'-end of reverse primer, respectively.

2.3.2 PCR amplification of *OsiSAP15* gene from rice (PB-1) cDNA

For full length gene amplification, Q5® High-Fidelity DNA Polymerase (©NEB, USA) was used and annealing temperature was calculated using NEB Tm

calculator (https://tmcalculator.neb.com) software. Approximately, 10-50 ng of template cDNA and gene-specific primers were used. The PCR reaction was carried out in Veriti 96-Well Thermal Cycler (Applied Biosystems, USA) with following cycle (see Table 2.1 and Figure 2.1).

Table 2.1. PCR reaction mixture for amplification of Os*SAP15* gene using reverse and forward primer.

COMPONENTS	VOLUME (in µL)
Q5 Buffer	5
10 mM dNTPs	0.5
10 µm forward primer	2.5
10 µm reverse primer	2.5
cDNA template	1
5X GC enhancer	5
Q5 DNA polymerase	0.25
MQ	8.25
Total	**25**

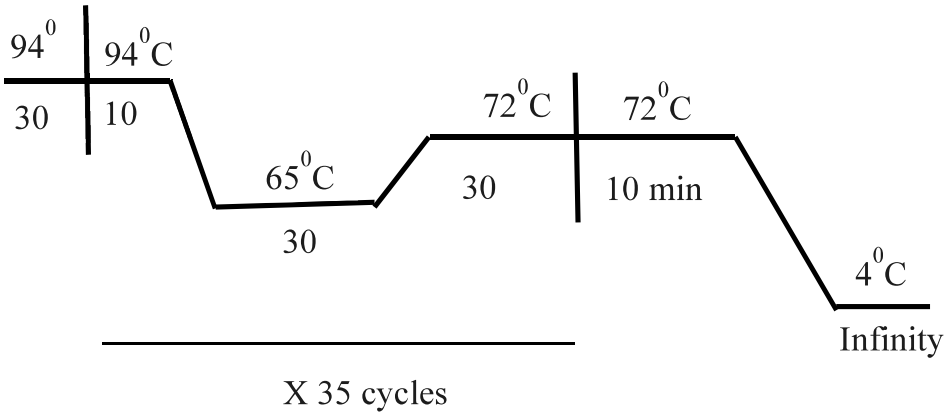


Figure 2.1. PCR cycle used for the amplification of *OsSAP15* gene.

The amplified product was analysed on 1% agarose gel with 1 kb DNA ladder.

2.3.3 Gel extraction of PCR-amplified product

PCR-amplified DNA was purified from agarose gel using GeneJET® PCR Purification Kit (©Thermo Fisher Scientific, USA) following manufacturer's

protocol. By visualizing gel under UV-transilluminator, specific band was excised with help of a fresh scalpel and weighed in a 2 ml MCT. 1:1 volume of Binding Buffer was added to completely immerse the gel in buffer and placed in dry bath at 50°C.

After melting the gel completely, one volume of isopropanol was added, mixed and content was passed through GeneJET® purification column by spinning at 13,000 rpm for 1 min. Flow through was discarded. 700 µl of wash buffer (with ethanol) was added and centrifuged at 13000 rpm for 1 min. Empty column was again centrifuged for 1 min. to completely remove residual wash buffer. GeneJET® purification column was transferred to a clean 1.5 ml MCT. 50 µl of sterile MQ was added to the center of the column membrane and centrifuged for 1 min. Column was discarded and purified DNA was stored at -20° C.

2.3.4 Restriction digestion of insert DNA and vector

For restriction digestion based cloning, purified PCR product of *OsiSAP15* gene and pET-28a (+) vector DNA were digested with their restriction sites specific enzymes i.e. *EcoRI* and *SalI*, respectively, for linearization. Approximately 1 µg of DNA was mixed with 1X restriction digestion buffer and digested with 1 U of Fast Digest® (Thermo Fisher Scientific, USA) restriction enzymes in 50 µl reaction mix.

The restriction enzyme mix was incubated at optimum 37°C temperature for 10 min for Fast Digest restriction enzymes (see Table 2.2).

Table 2.2. Restriction digestion reaction mixture of pET-28a (+) vector and *OsiSAP15* insert with *EcoRI* and *SalI*.

Components	**Vector volume (in µl)**	**Insert volume (in µl)**
10X Fast digest buffer	5	5
Sterile MQ	19	26
EcoRI	2	1
SalI	2	1
Plasmid (DNA template)	22	17
Total	**50**	**50**

2.3.5 Ligation of insert DNA and vector

The digested and purified insert DNA and vector were ligated using T4 DNA ligase (New England Biolabs, USA) in 3:1 molar ratio (see Table 2.3) with help of following formula:

$$\text{Insert (ng)} = \frac{\text{Ratio of insert x vector (ng) x insert size (kb)}}{\text{Ratio of vector x vector size (kb)}}$$

Table 2.3. Components of ligation reaction mixture containing vector and insert.

Components	**Volume (in µl)**
Sterile MQ water	7.5
10X Ligase Buffer	2
Insert	1.5
Vector	8
T4 DNA Ligase	1 U
Total volume	**20 µl**

The reaction components were gently mixed and centrifuged followed by incubation at 16ºC, overnight.

2.3.6 Preparation of chemically competent *E. coli* (DH5α strain) cells

A single colony of DH5α strain of *E. coli* from streaked plate was inoculated in 5 ml LB broth and incubated at 37ºC while shaking at 200 rpm overnight. Approximately 2 ml of this full grown culture was used as starter culture for 200 ml LB broth and kept at 37ºC on shaking at 200 rpm till OD_{600} reached 0.4-0.5. Any further growth was stopped by keeping culture at 4ºC for 45 min and cells were pelleted by centrifugation at 4,000 rpm for 10 min at 4ºC in SS-34 tubes. Pellet was re-suspended by adding 15-20 ml of pre-chilled 0.1 M $CaCl_2$ in each

tube. Tubes were incubated on ice for 20-30 min and were centrifuged at 4,000 rpm for 10 min at 4ºC. Again pellet was re-suspended in 4 ml of pre-chilled 0.1 M $CaCl_2$ and incubated overnight on ice. Following day, 4 ml of sterile 50% glycerol was added to the cell suspension. 100 µl cell suspension was aliquoted into 1.5 ml MCTs and snap frozen in liquid N_2 prior to storage at -80ºC.

2.3.7 Transformation in *E. coli*

Frozen competent *E.coli* DH5α strain cells were thawed on ice for 10 minutes. Insert DNA and vector containing ligation mix was added to competent cells prepared as described in section 2.2.6 and kept on ice for 30 min. After 30 min of incubation, heat shock was given to competent cells containing ligation mix for 90 sec at 42ºC in water bath followed by incubation on ice for 5 min. After heat shock treatment, 1 ml of LB broth was added to the tube containing transformation mix. The transformed *E. coli* cells were kept for revival in incubator shaker at 37ºC for 90 min on 200 rpm. After incubation, transformed cells were pelleted by centrifugation at 6,000 rpm for 5 min. LB broth was decanted in laminar air hood and pellet was re-suspended into 200 µl of fresh LB broth followed by plating onto 30 ml LB agar plate containing 30 µl kanamycin antibiotic. Plate was incubated at 37°C overnight.

2.3.8 Plasmid isolation from transformed cells by alkaline lysis method

About 5 ml LB broth was taken in a culture vial and a single positive colony was picked up and inoculated in LB broth with kanamycin antibiotic (50 mg/ml) and kept at 37˚C in incubator shaker at 200 rpm overnight. Next day, the culture was centrifuged at 6000 rpm for 5 minutes to pellet down the cells. Then, 250 µl of alkaline lysis solution I (25 mM Tris–Cl pH 8.0, 10 mM EDTA and 50 mM glucose) was added to the pelleted cells. After proper mixing, lysis solution II (0.2 N NaOH, 1% SDS) was prepared freshly and 250 µl of it was added into the tube and incubated for 5 min. at room temperature. 300 µl chilled alkaline lysis solution III (3M potassium acetate pH 4.8) was added to the suspension. 250 µl of phenol: chloroform: isoamylalcohol (25:24:1) were added for removing protein contamination. Contents in the tube were mixed by gentle shaking. Centrifugation

was done at 13,000 rpm for 15 minutes and supernatant was carefully transferred to the fresh 1.5 ml MCT and 500 µl of isopropanol was added. Centrifugation was again done at 13,000 rpm for 10 minutes. Supernatant was discarded and pellet was washed with 500 µl of 70% ethanol and again centrifugation was repeated. Supernatant was discarded and pellet were air-dried to remove the traces of ethanol. Pellet was dissolved in nuclease free MQ with 2 µl RNase and incubated for 15-30 min at 37˚C for proper dissolution before storage at 4°C.

2.3.9 Confirmation of cloning by restriction digestion

To confirm the cloning of *OsiSAP15* gene in PET-28a (+) expression vector, isolated plasmid DNA was restriction digested with *EcoRI* and *SalI* enzymes, respectively. One µg of plasmid DNA was digested with 1U of enzyme in 1X restriction digestion buffer at 37ºC for 5 min. After digestion, electrophoresis of digestion mix was on done 1% agarose gel with undigested vector as control. The Components of reaction mixture are shown in Table 2.4 below:

Table 2.4. Components of restriction digestion reaction of vector for confirmation of cloning of insert into the vector.

Components	**Volume (in µl)**
10X fast digest buffer	2
Sterile MQ	13
EcoRI	1
SalI	1
Plasmid (DNA template)	3
Total	**20**

Chapter 3

Results and Discussion

3.1 RESULTS

3.1.1 Phenotypic analysis of *OsiSAP15* overexpression plants in water deficit stress

3.1.1.1 *rd29A:OsiSAP15* overexpressing seedlings are tolerant to water deficit stress

To understand the *in planta* function of *OsiSAP15,* coding region of *OsiSAP15* gene was overexpressed under the influence of stress inducible *rd29A* promoter. Three independent T5 homozygous transgenic *Arabidopsis* lines with heterologous expression, namely, *OsiSAP15-18*, *OsiSAP15-22* and *OsiSAP15-23* were selected for further analysis as reported by Sharma (2015) in her Ph.D. thesis work.

When these overexpression lines were analysed for their efficiency to tolerate water deficit stress at germination stage, Sharma (2015) could observe strong tolerance to mannitol concentrations (100, 200 and 300 mM) at **germination stage**. Transgenic lines at all three stress conditions had a higher germination efficiency as compared to WT seeds.

To mimic drought stress at **seedling stage**, when seedlings were grown on ½ strength MS media without mannitol, no significant phenotypic difference between the growth of transgenic lines and WT was observed. Contrastingly, all three transgenic lines showed tolerance to mannitol stress as compared to WT when grown in Petri plates with mannitol for comparative analysis (see Figure 3.1 and 3.2). At 100 mM mannitol concentration, transgenic plants had well developed roots and longer roots with many lateral branching, whereas WT seedlings had smaller roots. Similarly at 200 mM mannitol concentration, transgenic lines outperformed WT as far as root length was concerned. 300 mM mannitol concentration created most stressed conditions for the growth of seedlings as compared to the previous two concentrations. At 300 mM concentration of mannitol, transgenic lines had higher root length as compared to WT seedlings.

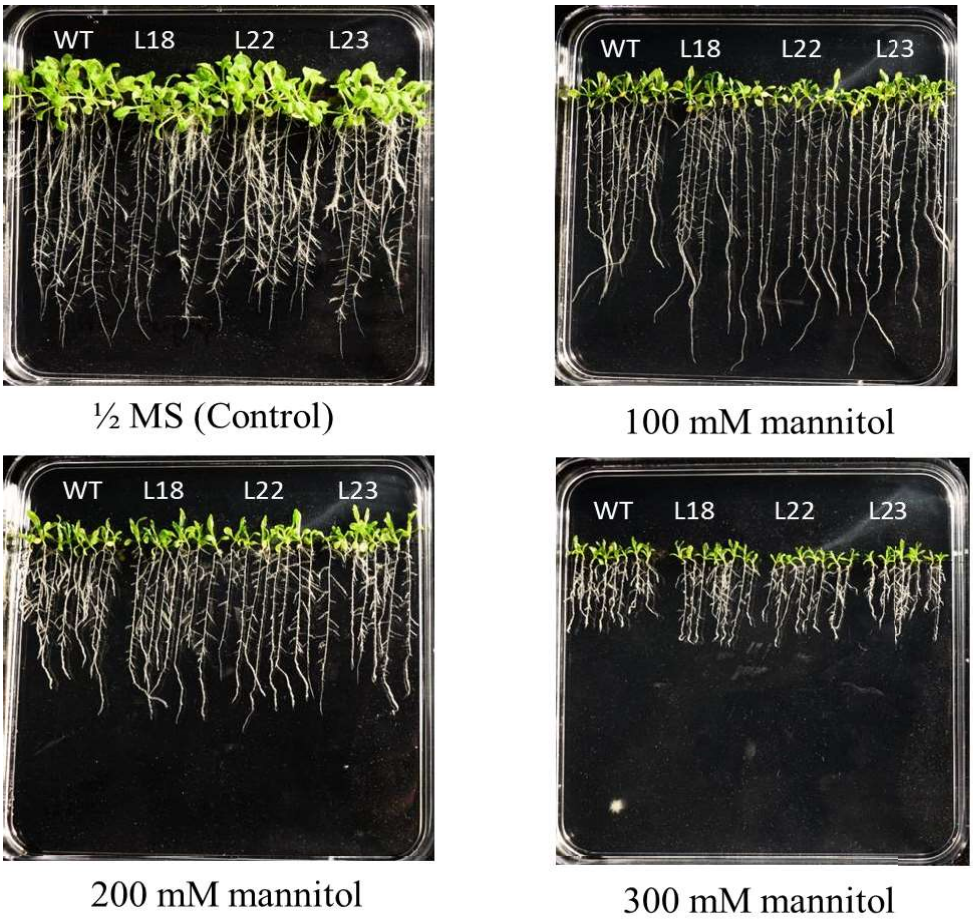


Figure 3.1. Drought stress tolerance phenotype of *Arabidopsis OsiSAP15* overexpression lines in 14-day-old *Arabidopsis* seedlings.

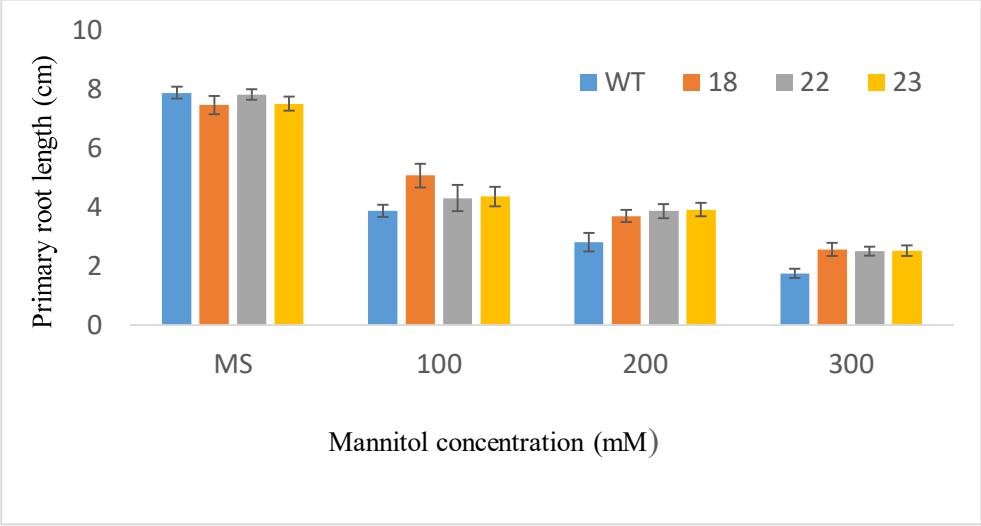


Figure 3. 2. Graphical representation showing the effect of different mannitol concentrations on primary root length of *OsiSAP15* overexpressing *Arabidopsis* seedlings.

3.1.1.2 Overexpression of *rd29A:OsiSAP15* in *Arabidopsis* confers tolerance to water deficit stress (WDS) at mature plant stage

To analyse overexpression lines in order to impart tolerance to WDS, 3-week-old pot grown plants were kept devoid of water for 15 days (~2 weeks). After 6-7 days of WDS, leafy tissue started drying because of loss of moisture from soil inside pots. First sign of dryness was observed in rosette leaves that travelled up to cauline leaves of *Arabidopsis* plant. Two weeks of WDS led to complete water loss from soil, as a result, leaves started turning pale yellow from green and ultimately complete brown in colour.

After 3 weeks of growth and 2 weeks of WDS, re-watering of pots was done and plants were regularly observed for their recovery. After 2 days of watering, revival of leafy portion in some plants of transgenic lines was observed. One-week of recovery period was able to revive most of the transgenic plants with arrival of new bolts. All of the 9 plants of transgenic line-23 survived. Four out of 9 plants of transgenic line-22 revived. Three out of 9 plants of transgenic line-18 revived. None of the WT *Arabidopsis* plants was able to revive. Overall, *OsiSAP15* overexpression plants had significantly higher survival rate as compared to WT as far as revival efficiency was concerned (see Figure 3.3).

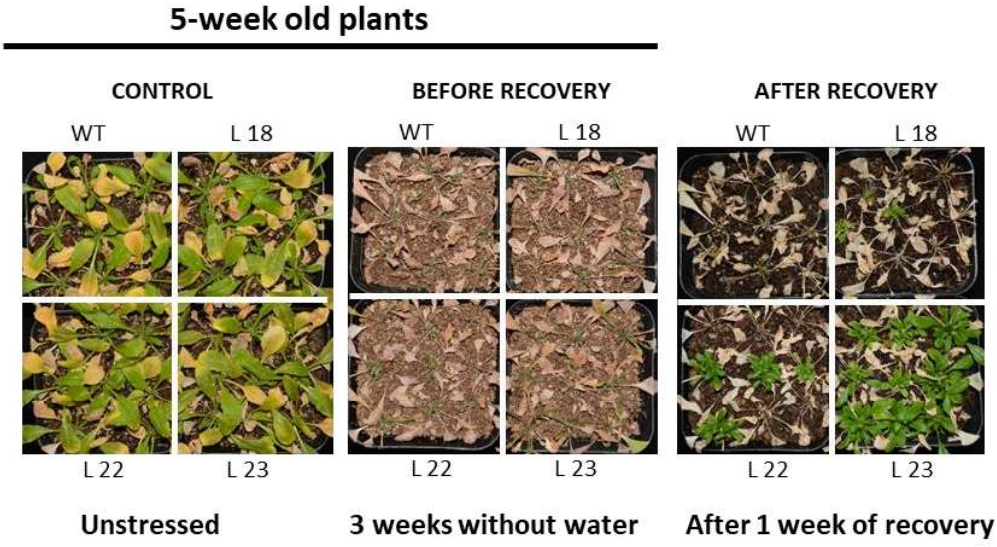


Figure 3.3. Phenotypic analysis of water-deficit stress response of *rd29A:OsiSAP15* overexpressing *Arabidopsis* lines (L18, L22, L23) in pots. Three-week-old light grown WT and transgenic *Arabidopsis* plants were kept devoid of water for 15 days and then re-watering (1 week of recovery) was done. Plants in set 1 and 2 were photographed just before recovery (5-week) and plants in set 3 were photographed after recovery period (i.e. after 6 weeks).

3.1.2 Cloning of *OsiSAP15* gene in pET-28a(+) expression vector

PCR amplification of *OsiSAP15* gene was done from rice PB-1 cDNA using 5'-TATTGAATTCATGGCGCAGGAGAGCT-3' as forward primer and 5'-ATTAGTCGACTCAGATCTTAATGATCTTGGCAGC-3' as the reverse primer, respectively, with help of Q5® High-Fidelity DNA Polymerase (©NEB, USA). When the amplified product was loaded on 1% agarose gel and electrophoresis was carried out, an intense band could be seen in every lane with loaded product. Band between 500-600 bp was seen when compared to the reference ladder, which was the approximate size of our gene i.e. 525 bp (see Figure 3.4).

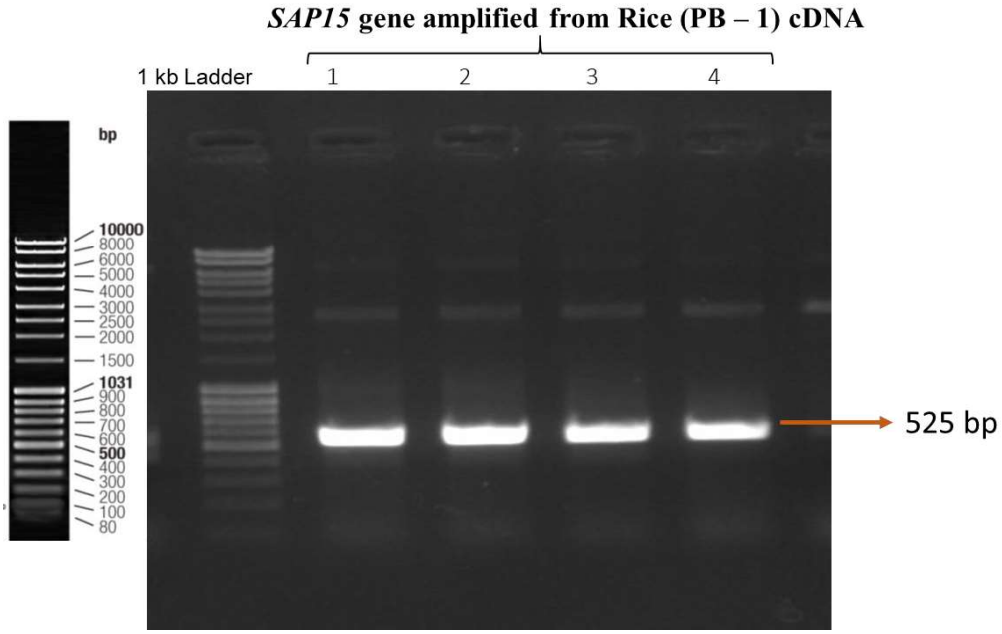


Figure 3.4. PCR amplification of *OsiSAP15* gene from rice cDNA using forward and reverse primers. Products were loaded in 4 different lanes and run on 1% agarose gel.

Bands from all four lanes were pooled, eluted and purified with help of GeneJET® PCR Purification Kit (©Thermo Fisher Scientific, USA). After digestion of vector and insert (linearization), ligation reaction resulted in final cloning of the gene in MCS of the respective vector. *E.coli* DH5α cells were transformed with ligated vector by heat shock method and cells were plated on

kanamycin selection plate. Next day, plasmid was isolated from colonies and restriction digested for confirmation of cloning into pET-28a (+) vector.

Restriction digested product was visualized on 1% agarose gel after electrophoresis. A faint band between 500-600 bp was seen which was approximately equal to the size of our gene (525 bp), another intense band near reference ladder 5000 bp was observed, which was the approximate size of our digested vector (see Figure 3.5). No such band at position of 500 bp was observed in lane with undigested vector. Two different forms of plasmid DNA were prominent in undigested lane.

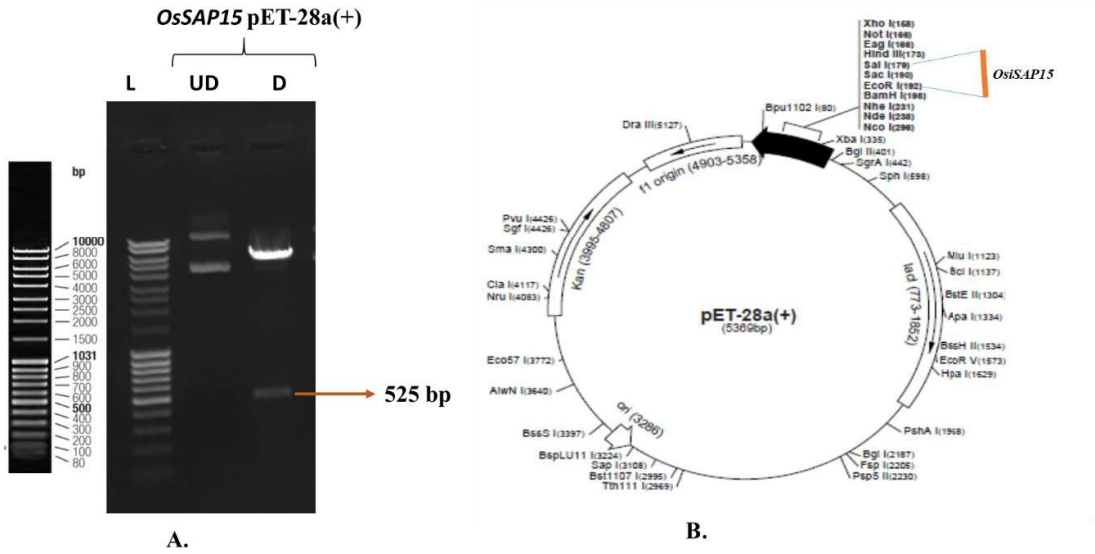


Figure 3.5. Restriction digestion of *OsSAP15* pET28-a (+) with *EcoRI* and *SalI* for confirmation of cloning. **A**. Digested product was run on 1% agarose gel. L represents 1 kb ladder, UD = Undigested product and D = Digested product. **B**. pET 28-a (+) vector map showing *SAP15* gene cloned between restriction sites.

3.2 DISCUSSION

Overexpression lines of *rd29A:OsiSAP15* were raised by Sharma (2015) for characterization of the respective gene. In first part of our study, we checked the *rd29A:OsiSAP15* overexpression lines for their role in mitigating drought/WDS tolerance in *Arabidopsis thaliana* seedlings and mature plant stage. A significant tolerance in transgenic lines at **germination stage** was already observed earlier. However, same transgenic lines when analysed for ABA stress tolerance, showed a stress sensitive phenotype (Sharma, 2015) hinting a putative role of *OsiSAP15* in relation to ABA sensitivity. Role of this gene in ABA dependent gene expression under dehydration stress conditions would be a good futuristic approach (Yamaguchi-Shinozaki and Shinozaki, 2006).

In our analysis at **seedling stage**, all three transgenic lines (L18, L22 and L23) had higher primary root length as compared to WT *Arabidopsis* Columbia seedlings, under water-deficit stress given by mannitol, indicating the potential of *OsiSAP15* in providing tolerance in WDS conditions. However, no significant difference in the primary root architecture was seen.

At **mature plant stage**, most transgenic *Arabidopsis* plants were able to revive from WDS when re-watering was done but even a single plant from WT lot was not able to revive. This observation can help us to understand the role of *OsiSAP15* in revival of plants in WDS conditions. Efficiency to regain the ceased physiology and metabolism in stressed conditions was higher in overexpressing plants under osmotic stress inducible promoter *rd29A* (Yamaguchi-Shinoaki and Shinozaki, 1993). In conclusion to this study, overexpression of *OsiSAP15* in its heterologous host *Arabidopsis thaliana* was able to confer drought/WDS tolerance to a significant extent.

Two unique activities had been previously reported in rice SAP by Sharma et al. (2015) and in *Arabidopsis* SAP by Ströher et al. (2009), respectively. Hence, our future goal is to transform *E.coli* BL21 strain with the *OsiSAP15* cloned product for its high protein expression ability to check the E3 ubiquitin ligase activity and ability to act as a redox sensor in *OsiSAP15*. Therefore, in second part of study, focus was on the cloning of *OsiSAP15* gene in expression vector pET28-a (+). That was largely successful as evident from restriction digestion of isolated recombinant plasmids from transformed *E.coli* strain DH5α that gave expected positive results, but further validation using sequencing technique is required.

Chapter 4

Conclusions and Future work

Multiple stress inducible protein gene, *SAP,* from its discovery in rice till date, has been characterized by various research groups around the globe. In last few years, this family of A20/AN1 ZFP coding genes has emerged to be a potential candidate in regulating various biotic and abiotic stresses. From crop plants like rice and wheat to plants like *Arabidopsis, Aleuropus* and *Lobularia,* their overexpression has conferred tolerance to different environmental stresses.

Although, their mode of action and their signalling cascades are not yet fully unravelled, but in recent years, some unique roles have been accredited to them like their ability to act as redox sensor and their E3 ubiquitin ligase activity.

From 18 known rice stress associated proteins with different protein domain organization, OsiSAP15 has AN1 domain at N terminal of protein structure. In the study, *OsiSAP15* was found to have a role in drought stress tolerance at seedling and mature plant stage upon its heterologous overexpression in *Arabidopsis thaliana* with stress inducible *rd29A* promoter. Under water deficit conditions, physiology of the plant changes leading to abnormal phenotypes and biochemical changes like production of higher amount of osmolytes like proline for osmotic balance, ROS production and cell death. These changes remain to be confirmed by biochemical assays.

Second part of our study was to clone *OsiSAP15* gene in pET28-a (+) expression vector which was confirmed by restriction digestion of the isolated recombinant plasmid DNA from *E.coli* strain DH5α. Ultimate aim was to transform *E.coli* strain BL21 with that to check the E3 ubiquitin ligase activity and redox sensing activity of OsiSAP15 which will be helpful in understanding its working mechanism.

In conclusion, *SAPs* are the potential candidates to protect the crop yield from the adverse nature of environmental stresses. Overexpression of *SAPs* in many plants under stress inducible promoter has helped variety of plant species to tolerate different abiotic stresses. The final futuristic approach is to utilize this potential of *SAPs* from growth rooms to the fields after validations and exploit this unique class of stress inducible genes in agriculture in regions with stressed conditions.

Chapter 5

Bibliography

1. **Ben Saad, R., Fabre, D., Mieulet, D., Meynard, D., Dingkuhn, M., Al-Doss, A., Guiderdoni, E., and Hassairi, A**. (2012). Expression of the *Aeluropus littoralis AlSAP* gene in rice confers broad tolerance to abiotic stresses through maintenance of photosynthesis. Plant. Cell Environ. **35**: 626–643.
2. **Ben Saad, R., Zouari, N., Ben Ramdhan, W., Azaza, J., Meynard, D., Guiderdoni, E., and Hassairi, A.** (2010). Improved drought and salt stress tolerance in transgenic tobacco overexpressing a novel A20/AN1 zinc-finger "AlSAP" gene isolated from the halophyte grass *Aeluropus littoralis*. Plant Mol. Biol. **72**: 171–190.
3. **Beyaert, R., Heyninck, K., and Huffel, S. Van.** (2000). A20 and A20-binding proteins as cellular inhibitors of nuclear factor-kB-dependent gene expression and apoptosis. Biochem Pharmacol. **60**: 1143–1151.
4. **Boyer, J.S.** (1982). Plant productivity and environment. Science **218**: 443–448.
5. **Bray, E.A., Bailey-Serres J., Weretilnyk, E.** (2000) Responses to abiotic stresses in Biochemistry and Molecular Biology of Plants 1st ed. 1158–1249. (American Society of Plant Physiologists, Rockville, MD).
6. **Cattivelli, L., Rizza, F., Badeck, F.W., Mazzucotelli, E., Mastrangelo, A.M., Francia, E., Marè, C., Tondelli, A., and Stanca, A.M.** (2008). Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. F. Crop. Res. **105**: 1–14.
7. **Charrier, A., Planchet, E., Cerveau, D., Gimeno-Gilles, C., Verdu, I., Limami, A.M., and Lelièvre, E.** (2012). Overexpression of a *Medicago truncatula* stress-associated protein gene (*MtSAP1*) leads to nitric oxide accumulation and confers osmotic and salt stress tolerance in transgenic tobacco. Planta **236**: 567–577.
8. **Choi, H., Han, S., Shin, D., and Lee, S.** (2012). Polyubiquitin recognition by AtSAP5, an A20-type zinc finger containing protein from *Arabidopsis thaliana*. Biochem. Biophys. Res. Commun. **419**: 436–440.
9. **Dansana, P.K., Kothari, K.S., Vij, S., and Tyagi, A.K.** (2014). *OsiSAP1* overexpression improves water-deficit stress tolerance in transgenic rice by affecting expression of endogenous stress-related genes. Plant Cell Rep. **33**: 1425–1440.
10. **Ding, L., Gu, B., Li, P., Shu, X., and Zhang, J.** (2019). Genome-wide identification and expression analysis of *SAP* family in Grape. Scientia Agricultura Sinica **52**: 2500-2514.
11. **Dixit, A.R. and Dhankher, O.P.** (2011). A novel stress-associated protein "AtSAP10" from *Arabidopsis thaliana* confers tolerance to Nickel, Manganese, Zinc, and high temperature stress. PLoS One **6**: e20921.
12. **Dixit, A., Tomar, P., Vaine, E., Abdullah, H., Hazen, S., and Dhankher, O.P.** (2018). A stress-associated protein, AtSAP13, from *Arabidopsis thaliana* provides tolerance to multiple abiotic stresses. Plant. Cell Environ. **41**: 1171–1185.
13. **Dixit, V.M., Green, S., Sarma. V., Holzman, LB., et al.** (1990). Tumor necrosis factor-alpha induction of novel gene products in human endothelial cells including a macrophage-specific chemotaxin. J Biol Chem. **265**: 2973–2978.
14. **Dong, Q., Duan, D., Zhao, S., Xu, B., Luo, J., Wang, Q., Huang, D., Liu, C., Li, C., Gong, X., Mao, K., and Ma, F.** (2018). Genome-wide analysis and cloning of the apple stress-associated protein gene family reveals *MdSAP15*, which confers tolerance to drought and osmotic stresses in transgenic Arabidopsis. Int. J. Mol. Sci. **19**: 2478.
15. **Gao, W., Long, L., Tian, X., Jin, J., Liu, H., Zhang, H., Xu, F., and Song, C.** (2016). Genome-wide identification and expression analysis of stress-associated proteins (SAPs) containing A20/AN1 zinc finger in cotton. Mol. Genet. Genomics **291**: 2199–2213.
16. **Giri, J., Dansana, P.K., Kothari, K.S., Sharma, G., Vij, S., and Tyagi, A.K.** (2013). SAPs as novel regulators of abiotic stress response in plants. BioEssays. **35**: 639–648.
17. **Giri, J., Vij, S., Dansana, P.K., and Tyagi, A.K.** (2011). Rice A20/AN1 zinc-finger containing

stress-associated proteins (SAP1/11) and a receptor-like cytoplasmic kinase (OsRLCK253) interact via A20 zinc-finger and confer abiotic stress tolerance in transgenic Arabidopsis plants. New Phytol. **191**: 721–732.

18. **He, X., Xie, S., Xie, P., Yao, M., Liu, W., Qin, L., Liu, Z., Zheng, M., Liu, H., Guan, M., and Hua, W.** (2019). Genome-wide identification of stress-associated proteins (SAP) with A20/AN1 zinc finger domains associated with abiotic stresses responses in *Brassica napus*. Environmental and Experimental Botany. **165**: 108–119.
19. **Huang, J., Wang, M.M., Jiang, Y., Bao, Y.M., Huang, X., Sun, H., Xu, D.Q., Lan, H.X., and Zhang, H.S.** (2008). Expression analysis of rice A20/AN1-type zinc finger genes and characterization of ZFP177 that contributes to temperature stress tolerance. Gene. **420**: 135–144.
20. **Jia, H., Li, J., Zhang, J., Ren, Y., Hu, J., and Lu, M.** (2016). Genome-wide survey and expression analysis of the stress-associated protein gene family in desert poplar, *Populus euphratica*. Tree Genet. Genomes. **12**: 78
21. **Jin, Y., Wang, M., Fu, J., Xuan, N., Zhu, Y., Lian, Y., Jia, Z., Zheng, J., and Wang, G.** (2007). Phylogenetic and expression analysis of ZnF-AN1 genes in plants. Genomics. **90**: 265–275.
22. **Kang, M., Fokar, M., Abdelmageed, H., and Allen, R.D.** (2011). Arabidopsis SAP5 functions as a positive regulator of stress responses and exhibits E3 ubiquitin ligase activity. Plant Mol. Biol. **75**: 451–466.
23. **Kang, M., Lee, S., Abdelmageed, H., Reichert, A., Lee, H.K., Fokar, M., Mysore, K.S., and Allen, R.D.** (2017). Arabidopsis stress associated protein 9 mediates biotic and abiotic stress responsive ABA signaling via the proteasome pathway. Plant, Cell Environ. **40**: 702–716.
24. **Kanneganti, V. and Gupta, A.K.** (2008). Overexpression of OsiSAP8, a member of stress associated protein (SAP) gene family of rice confers tolerance to salt, drought and cold stress in transgenic tobacco and rice. Plant Mol. Biol. **66**: 445–462.
25. **Lademann, U., Kallunki, T., and Jäättelä, M.** (2001). A20 zinc finger protein inhibits TNF-induced apoptosis and stress response early in the signaling cascades and independently of binding to TRAF2 or 14-3-3 proteins. Cell Death & Differentiation **8**: 265–272.
26. **Lai, W., Zhou, Y., Pan, R., Liao, L., He, J., Liu, H., Yang, Y., and Liu, S.** (2020). Identification and expression analysis of stress-associated proteins (SAPS) containing A20/AN1 zinc finger in cucumber. Plants. **9**: 1–16.
27. **Linnen, J.M., Bailey, C.P., and Weeks, D.L.** (1993). Two related localized mRNAs from *Xenopus laevis* encode ubiquitin-like fusion proteins. Gene. **128**: 181–188.
28. **Liu, S., Yuan, X., Wang, Y., Wang, H., Wang, J., Shen, Z., Gao, Y., Cai, J., Li, D., and Song, F.** (2019). Tomato stress-associated protein 4 contributes positively to immunity against necrotrophic fungus *Botrytis cinerea*. Mol. Plant-Microbe Interact. **32**: 566–582.
29. **Liu, Y., Xu, Y., Xiao, J., Ma, Q., Li, D., Xue, Z., and Chong, K.** (2011). OsDOG, a gibberellin-induced A20/AN1 zinc-finger protein, negatively regulates gibberellin-mediated cell elongation in rice. J. Plant Physiol. **168**: 1098–1105.
30. **Lloret, A., Conejero, A., Leida, C., Petri, C., Gil-Muñoz, F., Burgos, L., Badenes, M.L., and Ríos, G.** (2017). Dual regulation of water retention and cell growth by a stress-associated protein (SAP) gene in *Prunus*. Sci. Rep. **7**: 332.
31. **Malik, A.H. and Ashraf, N.** (2017). Transcriptome wide identification, phylogenetic analysis, and expression profiling of zinc-finger transcription factors from *Crocus sativus* L. Mol. Genet. Genomics. **292**: 619–633.
32. **Mukhopadhyay, A., Vij, S., and Tyagi, A.K.** (2004). Overexpression of a zinc-finger protein gene from rice confers tolerance to cold, dehydration, and salt stress in transgenic tobacco. Proc Natl Acad Sci USA. **101**: 6309–6314.
33. **Murashige, T. and Skoog, F.** (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiologia Plantarum. **15**: 473–497.
34. **Opipari, A.W. Jr., Boguski, MS., Dixit, V.M.,** (1990). The A20 cDNA induced by tumor necrosis factor alpha encodes a novel type of zinc finger protein. J Biol Chem. **265**: 14705–14708.
35. **Paul, A. and Kumar, S.** (2015). An A20/AN1-zinc-finger domain containing protein gene in tea is differentially expressed during winter dormancy and in response to abiotic stress and plant growth regulators. Plant Gene. **1**: 1–7.

36. **Peng, S., Cassman, K.G., Virmani, S.S., Sheehy, J., and Khush, G.S.** (1999). Yield potential trends of tropical rice since the release of IR8 and the challenge of increasing rice yield potential. Crop Sci. **39**: 1552–1559.
37. **Saad, R. Ben, Hsouna, A. Ben, Saibi, W., Hamed, K. Ben, Brini, F., and Ghneim-Herrera, T.** (2018). A stress-associated protein, LmSAP, from the halophyte *Lobularia maritima* provides tolerance to heavy metals in tobacco through increased ROS scavenging and metal detoxification processes. J. Plant Physiol. **231**: 234–243.
38. **Sharma, G.** (2015). Functional characterization of stress associated protein (SAP) gene family in rice. Ph.D. thesis. NIPGR, New Delhi.
39. **Sharma, G., Giri, J., and Tyagi, A.K.** (2015). Rice OsiSAP7 negatively regulates ABA stress signalling and imparts sensitivity to water-deficit stress in Arabidopsis. Plant Sci. **237**: 80-92.
40. **Solanke, A.U., Sharma, M.K., Tyagi, A.K., and Sharma, A.K.** (2009). Characterization and phylogenetic analysis of environmental stress-responsive SAP gene family encoding A20/AN1 zinc finger proteins in tomato. Mol. Genet. Genomics **282**: 153–164.
41. **Sreedharan, S., Shekhawat, U.K.S., and Ganapathi, T.R.** (2012). *MusaSAP1*, a A20/AN1 zinc finger gene from banana functions as a positive regulator in different stress responses. Plant Mol. Biol. **80**: 503–517.
42. **Ströher, E., Wang, X.J., Roloff, N., Klein, P., Husemann, A., and Dietz, K.J.** (2009). Redox-dependent regulation of the stress-induced zinc-finger protein SAP12 in *Arabidopsis thaliana*. Mol. Plant **2**: 357–367.
43. **Tyagi, H., Jha, S., Sharma, M., Giri, J., and Tyagi, A.K.** (2014). Rice SAPs are responsive to multiple biotic stresses and overexpression of OsSAP1, an A20/AN1 zinc-finger protein, enhances the basal resistance against pathogen infection in tobacco. Plant Sci. **225**: 68–76.
44. **Vij, S. and Tyagi, A.K.** (2006). Genome-wide analysis of the stress associated protein (SAP) gene family containing A20/AN1 zinc-finger(s) in rice and their phylogenetic relationship with *Arabidopsis*. Mol. Genet. Genomics **276**: 565–575.
45. **Wang, Y., Fu, X., Xie, L., Qin, W., Li, L., Sun, X., Xing, S., and Tang, K.** (2019). Stress associated protein 1 regulates the development of glandular trichomes in *Artemisia annua*. Plant Cell, Tissue and Organ Culture (PCTOC). **139**: 249–259.
46. **Xuan, N., Jin, Y., Zhang, H., Xie, Y., Liu, Y., and Wang, G.** (2011). A putative maize zinc-finger protein gene, *ZmAN13*, participates in abiotic stress response. Plant Cell. Tissue Organ Cult. **107**: 101–112.
47. **Xu, Q. F., Mao, X. G.,, Wang, Y. X., Wang, J. Y., Xi, Y. J, and Jing, R. L.** (2018). A wheat gene TaSAP17-D encoding an AN1/AN1 zinc finger protein improves salt stress tolerance in transgenic *Arabidopsis*. J. Integr. Agric. **17**: 507–516.
48. **Yamaguchi-Shinozaki, K. and Shinozaki, K.** (1993). Characterization of the expression of a desiccation-responsive *rd29* gene of *Arabidopsis thaliana* and analysis of its promoter in transgenic plants. Mol. Gen. Genet. **236**: 331–340.
49. **Yamaguchi-Shinozaki, K. and Shinozaki, K.** (2006). Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses.The Annu. Rev. Plant Biol. **57**:781–803.
50. **Zahur, M., Maqbool, A., Irfan, M., Jamal, A., Shahid, N., Aftab, B., and Husnain, T.** (2012). Identification and characterization of a novel gene encoding myb-box binding zinc finger protein in *Gossypium arboreum.* Biol. Plant. **56**: 641–647.
51. **Zhang, N., Yin, Y., Liu, X., Tong, S., Xing, J., Zhang, Y., Pudake, R. N., Izquierdo, E. M., Peng, H., Xin, M., Hu, Z., Ni, Z., Sun, Q., and Yao, Y.** (2017). The E3 ligase TaSAP5 alters drought stress responses by promoting the degradation of DRIP proteins. Plant Physiol. **175**: 1878–1892.
52. **Zhang, X. Z., Zheng, W. J., Cao, X. Y., Cui, X. Y., Zhao, S. P., Yu, T. F., Chen, J., Zhou, Y. B., Chen, M., Chai, S. C., Xu, Z. S., and Ma, Y. Z.** (2019) Genomic analysis of stress associated proteins in soybean and the role of *GmSAP16* in abiotic stress responses in *Arabidopsis* and soybean. Front. Plant Sci. **10**: 1453.
53. **Zhang, Y., Lan, H., Shao, Q., Wang, R., Chen, H., Tang, H., Zhang, H., and Huang, J.** (2016). An A20/AN1-type zinc finger protein modulates gibberellins and abscisic acid contents and increases sensitivity to abiotic stress in rice (*Oryza sativa*). J. Exp. Bot. **67**: 315–326.

54. Zhou, Y., Zeng, L., Chen, R., Wang, Y., and Song, J. (2018). Genome-wide identification and characterization of stress-associated protein (SAP) gene family encoding A20/AN1 zinc-finger proteins in *Medicago truncatula*. Archives of Biological Sciences. **70**: 87–98.

Appendix

1. CDS of *OsiSAP15* gene

>*OsSAP15*

ATGGCGCAGGAGAGCTGTGATCTCAACAAGGACGAGGCTGAGATCTTGAAGCCATCATCA
TCATCCTCACCTTCTCCTTCCCCAACCACAGCCTCCCCATCGCCACCAACAGCACAAATG
ACAGAACCACCGCCTCCGCAATCGACACCACCAACACCACCGGCAGCAGCAGCAGCGGCA
TCGGCAGCAGCAGCTCCTCAATTCTCAGCCAAGAACTGCGAAGGTATTCTCATAGAGGTT
TCCAAGAAGAGGAAACTAGCCGAAGCCACCGCCACCGATGCTAATGCCGTGGTCGTGGCG
GCGGTAGCCGAGCCATTATCTCCGGTGCTGTTTGTTAACCGGTGCAACGTGTGCCGCAAG
AGGGTCGGTTTGACCGGATTCCGGTGCCGCTGCGGGGAGCTCTTCTGTCCTCGCCACCGG
CATTCCGAGACCCATGAGTGCTCCTTCGATTATAAAACCGCGGGCAGGGAGGAGATTGCC
CGCGCGAATCCTGTGATCAGGGCTGCCAAGATCATTAAGATCTGA

Analysis of the promoter controlling the expression of the coat protein gene of *Sri Lankan cassava mosaic virus*

Thesis Submitted to the University of Delhi
in the Partial fulfilment for the Degree of
Master of Science
in
Plant Molecular Biology and Biotechnology
2020

MONICA BISWAS

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi 110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Analysis of the promoter controlling the expression of the coat protein gene of Sri Lankan cassava mosaic virus**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Monica Biswas

Date: 6th June, 2020

Signature of the Candidate

(Monica Biswas)

Indranil Dasgupta

Signature of the Supervisor

(Prof. Indranil Dasgupta)

Signature of the HOD

(Prof. Anil Grover)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

Preface

Cassava (*Manihot esculenta* crantz), a staple food crop belonging to the family *Euphorbiaceae* is an important source of carbohydrate. Cassava mosaic disease (CMD) is an important viral disease which is caused by whitefly borne (*Bamisia tabaci)* viruses of the genus *Begomovirus (*family *Geminivirdae*) reported from Africa and Asia. Geminiviruses are single-stranded (ss) DNA containing virus that constitutes the largest virus group, infecting economically important crops. The virus associated with CMD in India and Sri Lanka is *Indian cassava mosaic virus (ICMV)* and *Sri Lankan cassava mosaic virus (SLCMV)* respectively.

A significant advance in plant biotechnology to improve crop productivity and sustainable agriculture, is contributed to the availability of novel promoters for precise spatial and temporal control of transgene expression. Viral promoters have been used to express transgenes in plants, including that of geminiviruses. Understanding the promoter of SLCMV CP-promoter is the major aim of this thesis. The aim of this thesis was to further analyze the positive element, IR1P (2725 to 2620) of SLCMV CP-promoter (whether it enhances the strength of other heterologous promoter) by constructing a series of deletions of SLCMV CP-promoter. The aim was to study the expression of all the constructs to be analyzed by transient expression study in *Nicotiana benthamiana* by studying Gus activity and histochemical staining.

ACKNOWLEDGEMENT

I would like to thank God Almighty for giving me the strength and opportunity to undertake this journey and complete it satisfactorily. There is a universal truth we all have to face, whether we want to or not, everything eventually ends. It has been an amazing two year with lots of memories in the department. At the end of these two memorable years in the department and in the laboratory, I feel privileged to be part of this department.

Foremost, I would like to express my sincere gratitude to my teacher and supervisor **Prof. Indranil Dasgupta** for his patience, constant motivation, immense knowledge and excellent guidance that have gone a long way in the completion of my dissertation. He works with enthusiasm, a well-informed dedicated researcher. His enthusiasm, dedication and commitment and patience towards work are the qualities, which inspires me.

I am obliged to all the faculty members **Prof. Jitendra P. Khurana, Prof. Akhilesh K. Tyagi, Prof. Anil Grover, Prof. Paramjit Khurana, Prof Arun Sharma, Prof Sanjay Kapoor, Dr. Saurabh Raghuvansi, and Dr Surekha K. Aggarwal** for helping me gain a clear understanding of molecular biology course work.

I would like to express my special thanks of gratitude to Miss Fauzia Zarreen who always showed immense support, guidance throughout my dissertation period. I thank all my wonderful seniors in the laboratory, for their wonderful support and help. I would like to thank Mr. Rohit kumar, Miss Neelam Jagram for always taking interest in progress of my work and being always there to help me and answer my queries. I acknowledge Dr. Gaurav kumar, Dr. Zainul Abdeen Khan, Miss Kanika Gupta, Miss Kamlesh Kumari for their advice and support. I would like to thank Mr. Trilok, Mr. Brij Kishore for all their help during day to day lab organisation.

I would like to thank my batchmates, my friends in the department for providing with a wonderful and memorable two years experience in this department.

I deeply thank my family, my parents for their unconditional love, trust and endless patience. I am grateful to my best friends and friends . I would like to thank them for all their timely encouragement and unfailing emotional support despite the long distance between us. It was their love that raised me up again when I got weary.

List of commonly used Abbreviation

ACMV	*African cassava mosaic virus*
ATP	Adenosine triphosphate
BCTV	*Beet curly top virus*
BGMV	*Bean golden mosaic virus*
bp	Base pairs
CaMV 35S	*Cauliflower mosaic virus* 35S promoter
CaMV	*Cauliflower mosaic virus*
CMD	Cassava mosaic disease
CMV	*Cassava mosaic virus*
CP	Coat protein
CR	Common region
c-sense	Complementary sense
DNA	Deoxyribonucleic acid
dNTP	Deoxy nucleoside triphosphate
ds	Double stranded
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium bromide
FAOSTAT	Food and Agriculture Organization Corporate Statistical Database
g	Gram
GUS	ß-glucornidase enzyme
hr	Hour
ICMV	*Indian cassava mosaic virus*
IR	Intergenic region
kb	Kilo base pair
LB	Luria Bertani
LIR	Long intergenic region
MCS	Multiple cloning site
MCT	Microcentrifuge tube
min	Minutes
mM	Millimolar
MP	Movement protein
ng	Nanogram
nM	Nanometer
nmoles	Nanomoles
NSP	Nuclear shuttle protein

nt	Nucleotide
O/N	Overnight
OD	Optical density
ORF	Open reading frame
PCR	Polymerase chain reaction
Rep	Replication associated protein
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolution per minute
RT	Room temperature
s	Second
SDS	Sodium dodecyl sulphate
SLCMV	*Sri Lankan cassava mosaic virus*
ss	Single stranded
TBE	Tris borate EDTA
TE	Tris EDTA
TGMV	*Tomato golden mosaic virus*
TrAP	Transcriptional activator protein
Tris	Tris (hydroxymethyl) aminomethane
UV	Ultraviolet
v/v	Volume/volume
v-sense	Virion-sense
µg	Microgram
µm	micrometre
µM	micromolar

CONTENTS

1 General introduction 10-12

2. REVIEW OF LITERATURE

2.1 Introduction 13

2.2 Taxonomy 13-14

2.3 Begomovirus 15

2.4 Genome organization 15-16

2.5 Begomovirus promoter element and its regulation 16-18

2.6 Replication of geminivirus 18-19

2.7 Transcriptional regulation 20

2.8 Geminivirus life-cycle 20-21

2.9 Sri Lankan cassava mosaic virus 21

2.10 Sri Lankan cassava mosaic virus promoter element 22-23

3 MATERIALS AND METHODS

3.1 Strategy used to study the positive element IR1P in heterologous promoter 25-26

3.2 Isolation of plasmid (From *E. coli* DH5-α strain) 26

3.3 Gel isolation of desired DNA fragment 26

3.4 PCR amplification 27

3.5 Poly A tailing for blunt end PCR products 27-28

3.6 Ligation with vector with T-overhangs 28

3.7 Preparation of *E. coli* chemical competent cells 28-29

3.8 Transformation using chemical competent cells by heat shock 29

3.9 Colony PCR for Confirmation of positive colonies 29-30

3.10 Confirmation of positive clones by Restriction digestion 30

4 Results and Discussion

4.1 Result 32-37

4.2 Discussion 38

5 Conclusions and future work 40

5. Appendices and References 42-44

Chapter-1 & 2

Introduction

&

Literature review

1 GENERAL INTRODUCTION

Agriculture, which is the main source of food supply also plays a crucial role in the economy of developing countries like India. In India Gross value added by agriculture along with forestry is estimated at US$265.51 billion in FY19(PE) (https://www.ibef.org/industry/agriculture-india.aspx). However, different abiotic and biotic factors affect crop yield drastically, among which viruses causes some of the most damaging and economically important diseases of crop plants. Geminiviruses are single-stranded (ss) DNA containing virus that constitutes the largest virus group, infecting economically important crops. Some important diseases of plants caused by geminiviruses are given in figure 1.A.



Figure 1.A: Symptoms of geminivirus infection (a) *Chickpea chlorotic dwarf virus* (chickpea). (b) *Beet curly top virus* (tomato). (c) *Grapevine red blotch virus* (grapevine). (d) *Tomato yellow leaf curl virus* (tomato). (e) *Tomato mottle leaf curl virus* (tomato). (f) *Bean golden mosaic virus* (common bean). (g) *Bhendi yellow vein mosaic virus* (okra). (h) *Maize streak virus* (maize). (I*) Wheat dwarf virus* (wheat) (Photo source: Annual review of phytopathology, 2018)

Cassava (*Manihot esculenta* crantz) is a woody shrub, belongs to the family- *Euphorbiaceae.* Cassava is an important source of carbohydrates in the tropics after cereals, that makes it the most important root and tuber crop. Numbers of industrial products are prepared from cassava starch having wide applications in textile industries as sizing agent, in pharmaceutical industries, making adhesive, paper industry etc. Cassava was introduced into India by the Portuguese when they landed in the Malabar region, presently part of Kerala during the 17th century, from Brazil. India ranked among top twenty of the world's highest cassava production, producing nearly five million tons of cassava in the year 2018 (FAOSTAT, 2019).

Cassava mosaic disease

Cassava is vulnerable to a broad range of diseases. Cassava mosaic disease (CMD) is the severe and widespread, limiting its yield. CMD produces a variety of foliar symptoms that includes mosaic, molting/shedding of leaves, misshapen and twisted leaflets and an overall reduction in size of leaves, plants and produces few or no tubers depending on the severity of the disease (Figure 1.B). Historically, the disease was first reported by Warburg (1894) from Usambaras Mountains range in northeast Tanzania. The first sequence of Cassava mosaic virus (CMV) was published by Stanley and Gay, 1983. Probably it was called cassava latent virus and subsequently, called African cassava mosaic virus (ACMV). The virus associated with CMD in India and Sri Lanka is *Indian cassava mosaic virus* (ICMV; Malathi and Sreenivasan, 1983, Saunders et al, 2002) and *Sri Lankan cassava mosaic virus* (SLCMV: Dutt et al.,2005) respectively. Whitefly (*Bemisia tabaci*) acts as a carrier of the virus.



Figure 1.B: Picture comparing a healthy cassava plant to plants affected by Cassava mosaic disease (CMD). (A) A healthy cassava tuber. (B) (C) (D) compares the levels of CMD affected, from mild symptoms to moderate to severe symptoms respectively (photo source; Alabi et al.,2011).



Figure 1.C: A cassava plant with severe mosaic symptoms on leaves (puckering and malformation)

A significant advance in plant biotechnology to improve crop productivity and sustainable agriculture, is contributed to the availability of novel promoters for precise spatial and temporal control of transgene expression. The promoters which can drive constitutive/inducible/ tissue-specific expression of heterologous genes in plants are extremely limited. Viral promoters have been used to express transgenes in plants, including that of geminiviruses. Sri Lankan cassava mosaic virus. (SLCMV) belongs to the family *Geminivirdae* containing ssDNA comprising two component DNA-A and DNA-B. SLCMV DNA is transcribed bi-directionally having its promoter in the intergenic region. The analysis of sequence of the promoter of SLCMV DNA-A (PhD thesis Basanta Kumar Borah, 2009) revealed a complex regulation of promoter that depends upon the direction of transcription. The virion sense promoter expressed strength almost equal to the full length CaMV 35S promoter, but complementary sense strand promoter had only 60% strength of the CaMV 35S promoter. Deletion analysis of the promoter region revealed that the promoter had three elements, a eukaryotic positive element, a prokaryotic positive element and a eukaryotic and prokaryotic negative element.

This thesis deals with the analysis of the positive element (IR1P) of SLCMV promoter by transient expression study in *Nicotiana benthamiana*. Based on earlier work (PhD thesis, Basanta Kumar Borah, 2009) a positive element in the SLCMV CP-promoter was found. The aim of this thesis was to further analyze this element (whether it enhances the strength of other heterologous promoter) by studying its effect on heterologous promoter by constructing a series of deletions of SLCMV CP-promoter. The expression of all the constructs to be analyzed by transient expression study in *Nicotiana benthamiana* by studying Gus activity and histochemical staining.

2. REVIEW OF LITERATURE

2.1 Introduction

Geminiviruses constitutes one of the largest plant infecting viruses with >440 recognized species distributed throughout all land ecosystem. These viruses infect a broad variety of wild plants and agricultural crops causing significant shrinkage of staple food supplies and large economic losses worldwide. Diseases associated with this family Geminivirdae include: *Chickpea chlorotic dwarf virus* (chickpea), *Tomato yellow leaf curl virus* (tomato), *Bean golden mosaic virus* (common bean), etc. The name comes from their characteristic germinate (twinned) incomplete icosahedral particles (Figure 2.1) It is the largest known single stranded DNA viruses.

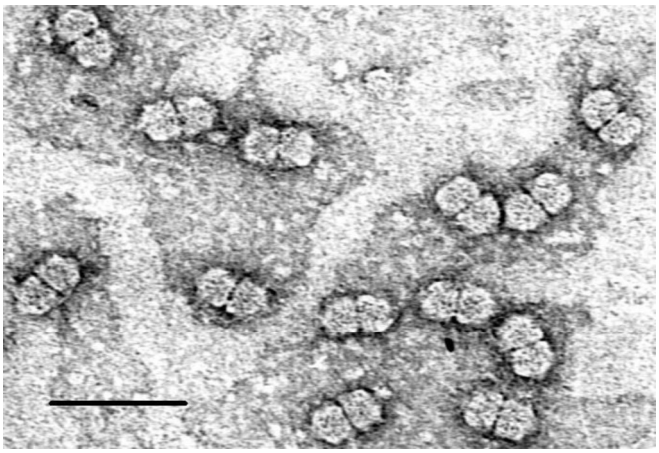


Figure 2.1: An EM of a typical germinate structure of geminivirus particle(Purified *Maize streak virus* (MSV) particle stained with uranyl acetate. Picture taken by Kassie Kasdorf, at Plant Protection Research Institute, Pretoria, for the University of Cape Town)

2.2 Taxonomy

Geminiviruses are transmitted by sap-sucking insects in the order Hemiptera, and the transmission process is highly specific and dependent on the viral coat protein (CP). Nine genera are currently recognized in the family *Geminivirdae* and the classification criteria is based on the type of insect vector (Figure 2.2A), type of plant they infect and the virus genome organization. Table (2.2.B) shows the general description of different taxons classified under *Geminivirdae* family.





Figure 2.2.A: Picture of common vector of geminiviruses, whitefly (*Bamisia tabaci*) (at left), leafhopper (right).

Genera	**Type species**	**Genome organization**	**Vector**	**Host range**
Becurtovirus	*Beet curly top Iran virus, spinach curly top Arizona virus*	monopartite	leafhopper	Dicots,
Begomovirus	*Bean golden mosaic virus*	Monopartite or bipartite	Whitefly	Dicots
Capulavirus	*Alfalfa leaf curl virus*	Monopartite	Aphids	Dicots
Curtovirus	*Beet curly top virus*	Monopartite	leafhopper	Dicots
Mastrevirus	*Maize streak virus*	Monopartite	Leaf hopper *(cicadellidas*)	Monocots
Eragovirus	*Eragrostic curvula streak virus*	Monopartite	Treehopper, leafhopper	Dicots
Grablovirus	*Grapevine red blotch virus*	Monopartite	Treehopper	Grapevine
Topocuvirus	*Tomato pseudo leaf curl virus*	Monopartite	Leafhopper (*Micrutalis malleifera*)	Dicots
Turncurtovirus	*Turnip curly top virus*	monopartite	leafhopper	Dicots,

Table 2.2.B: Table showing taxonomic classification of family *Geminivirdae.*

2.3 Begomovirus

The genus *Begomovirus* (BGVs) is the largest group within the family *Geminiviridae*, comprising nearly 88% of all geminivirus species. BGVs solely infect dicots, and are transmitted by *Bemisia tabaci*, a complex of cryptic and morphologically indistinguishable whitefly species (Figure 2.2A). Begomoviruses are responsible for numerous diseases of economically important crops, such as cassava, cotton, bean, pepper and tomato (Brown, 1992). In begomoviruses, the tendency for recombination and acquisition of extra DNA components has resulted in emergence of new viruses that infect new hosts and cause new diseases (Varma and Malathi, 2003; Chakraborty *et al*., 2003). Local evolution has resulted in multiple begomovirus species that infect a given crop or weed species and cause similar symptoms; e.g., there are more than 90 species that infect tomato. Thus, the precise identification of the begomovirus involved is important for the selection of management measures (Rojas, 2018).

2.4 Genome organization

The genome of bipartite begomoviruses consists of two ssDNA components, DNA-A and DNA-B, each approximately 2.8kb in size. Genome of monopartite begomovirus contains a single circular DNA molecule of approximately 2.8kb in size. The International Committee on the Taxonomy of Viruses (ICTV) has recommended a classification and scheme of nomenclature, based on complete nucleotide sequences of the genome (DNA-A and DNA-B) and geographical distribution of begomovirus. Begomovirus originating in the New World (NW: the Americas) have a bipartite (genome consisting of two circular DNA molecules, DNA-A and DNA-B) genome organization whereas those from the Old World (OW: Eurasia, Africa, and Oceania) have either bipartite or monopartite(genome consisting of a single circular DNA molecule) genome (Brown et al., 2002). Begomoviruses from the OW possess an additional ORF (AV2), not found in NW begomoviruses. DNA-A (2645 nt) component of the genome has six ORFs, two in the viral-sense strand (clockwise) AV1 or CP encodes for coat protein and AV2 (pre-coat protein) role in virus movement (Padidam et al., 1996). Four genes in the complementary-sense strand (anti-clockwise) (AC1 or Rep; AC2 or Trap; AC3 or Ren; and AC4 encoding proteins involved in a variety of essential functions for the virus infective cycle: replication, interference of the plant cell cycle, temporal regulation of viral gene expression, and suppression of host antiviral responses. The DNA-B component which is present only in bipartite begomovirus encodes two proteins, BV1 or NSP (nuclear shuttle protein), and BC1 or MP (movement protein), which participate in the intra- and intercellular movement of viral DNA respectively. The DNA-A and DNA-B molecules of bipartite BGVs are very different in sequence, with the exception of a shared DNA segment ranging from 140 to 200-nt in length, which contains the virus replication origin. A nonanucleotide sequence (TAATATT/AC), called the “common region” (CR) is common between the two components. This nonanucleotide sequence is fully conserved in all geminiviruses, and essential for the replication of the viral DNA. Figure 2.4 shows the arrangement of genes in DNA-A and DNA-B components of bipartite begomovirus.

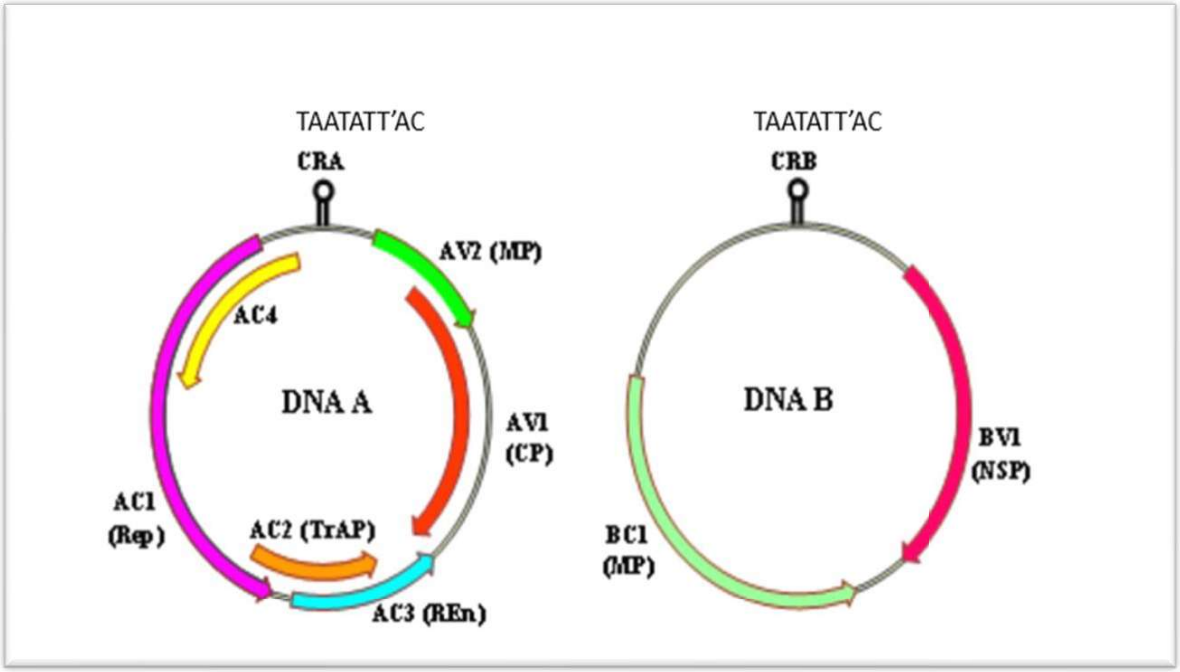


Figure 2.4: Genome organization of DNA A and DNA B components of cassava-infecting bipartite begomoviruses showing various ORFs (genes) in virion sense and complementary sense: DNA-A and DNA-B. CP, coat protein; MP, movement protein; Rep, replication-associated protein; TrAP, transcriptional activator protein; REn, replication enhancer protein (Alabi and Kumar, 2011).

2.5 Begomovirus promoter element and its regulation

In the bipartite begomovirus, DNA-A component encodes all the protein functions necessary for virus replication while the DNA-B component provides movement functions required for systemic spread as well as cell-to-cell spread. Geminiviral promoters are generally located within the intergenic region, although promoters have also been detected within the genes. The promoter and regulatory elements show a very complex regulation that is mediated by several host as well as viral proteins and it exhibit a strictly regulated temporal gene expression during infection, early and late stage infection as a result of a structural differences within the promoter architecture. The early genes represented by complementary-sense transcription essential for viral DNA replication at the early stage of infection, whereas the late genes represented by the viral-sense promoter involved in systemic spread of pathogenicity and are inherently weak one, requiring trans-activation by the complementary-sense viral gene product i.e. transcriptional activator protein (TrAP/AC2/C2/AL2) encoded by the early gene TrAP (Haley et al. 1992; Shivaprasad et al. 2005).

The virus encoded gene products which are essential for geminivirus/begomovirus are described below.

ORF AV1 (Coat Protein)

The ORF AV1 encodes for coat protein (CP) (Kallender et al., 1988). The CP gene plays the important role in encapsidation of the viral DNA and is implicated in viral movement within the plant as well as in whitefly transmission (Wartig et al., 1997; Hanley-Bowdoin *et al*., 1999; Harrison and Robinson, 1999; Sharma and Ikegami, 2009).

ORF AV2 (Pre-coat protein)

It is speculated that for monopartite begomoviruses found only in the Old World, ORF AV2 is complementing the function of virus movement rendering DNA-B redundant (Harrison and Robinson, 1999).

ORF AC1 (Replication initiator protein, Rep)

AC1 that encodes for Replication initiator protein is a multifunctional, oligomeric protein, which possesses site-specific DNA binding to the reiterated motifs at the intergenic region (IR) and initiates DNA replication (Hanley Bowdoin et al., 1999). AC1 functions as a negative-feedback regulator of transcription from its own promoter by binding to iterative sequences (iterons) between the TATA box and the AC1 transcription start site.

ORF AC2 (Transcription activator protein)

Transcription activator protein (TrAP) induces virion sense promoters in DNA-A (AV1) as well as in DNA-B (BC1 and BV1). (Hanley-Bowdoin et al., 1999; Rajeswaran, 2007). It has also been reported as a suppressor of RNA silencing in bipartite begomoviruses (Voinnet et al., 1999; Wang et al., 2003) as well as monopartite begomoviruses (Dong et al., 2003). AC2 of the bipartite begomovirus Mungbean yellow mosaic virus has been shown to suppress the RNAi mediated anti-viral defence response in plants. It was observed that mutations in the AC2 NLS or zinc-finger, or deletions of its activator domain abolished the function related to silencing suppression. (Trinks et al., 2005). TrAP determines pathogenicity by interference with silencing suppressors in the miRNA pathway by TrAP in plants (Rajeswaran et al., 2007).

ORF AC3 (Replication enhancer protein)

Rep/REn interaction is important for viral DNA replication (Gutierrez, 2002).

ORF AC4

ORF AC4 is present within Rep protein but in a different frame thus it encodes for a different protein. C4/AC4 found to induce hyperplasmia and tumorigenic growth in infected plant tissue by controlling cell cycle progression, targeting the cyclin CDK-complexes, preventing programmed cell deaths (Fondong, 2009). C4/AC4 proteins have also been shown to be RNA silencing suppressors, by interacting with the receptor-like kinase at the plasma membrane and regulating the cell-to-cell spread of the RNA silencing signal (Carluccio et al., 2018).

ORFs BV1: Nuclear shuttle protein (NSP) and BC1: Movement protein (MP)

For bipartite begomoviruses, ORF BV1 functions as NSP i.e. in transporting the virus DNA between nucleus and cytoplasm, while ORF BC1 (MP) functions in cell-to-cell movement of

the viral DNA (Gafni and Epel, 2002; Hussain et al., 2005). In figure 2.5, a schematic representation of geminivirus intracellular movement is shown.

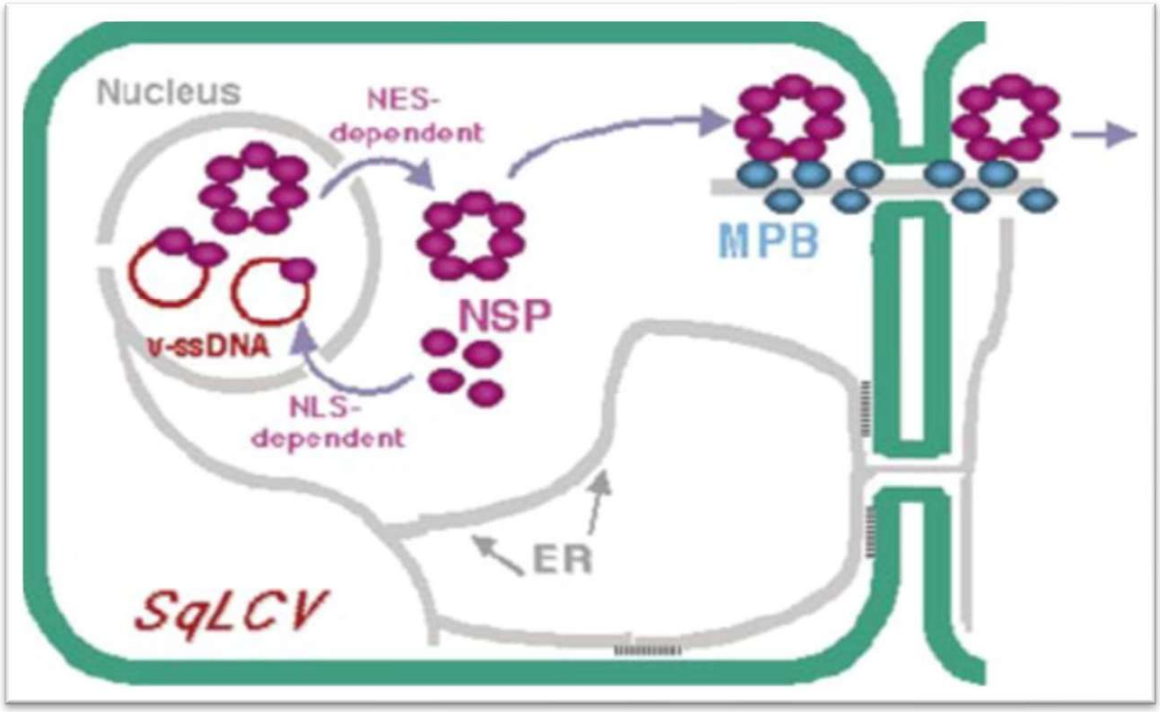


Figure 2.5: Schematic representation showing the process of geminiviral intracellular movement and the different geminiviral proteins that are known to have role in viral movement. (Lazarowitz and Beachy, 1999).

2.6 REPLICATION OF GEMINIVIRUS

The geminivirus replication cycle relies entirely on DNA intermediates and occurs within the nucleus of the infected cell through two basic stages:

Conversion of ssDNA to dsDNA intermediates.

Rolling circle replication (RCR).

Geminiviruses utilize the host cell machinery for their replication. Rep is necessary for geminivirus replication. For initiation of replication, ss DNA of the geminivirus first converted into dsDNA (Kammann et al., 1991; Saunders et al., 1992). This step is carried out entirely by host cellular enzymes.

The next stage, the rolling circle phase, requires the concerted action of viral Rep protein and other viral proteins with cellular factors and leads to the production of new dsDNA and ssDNA viral forms (Stenger et al., 1991; Stanley, 1995). The ds replicative form serves as a template for transcription as well as ssDNA synthesis. The ssDNA thus produced may enter in the replicative cycle or gets transported from cell-to-cell with the help of viral movement protein or gets encapsidated by viral coat protein (Gutierrez, 2002).

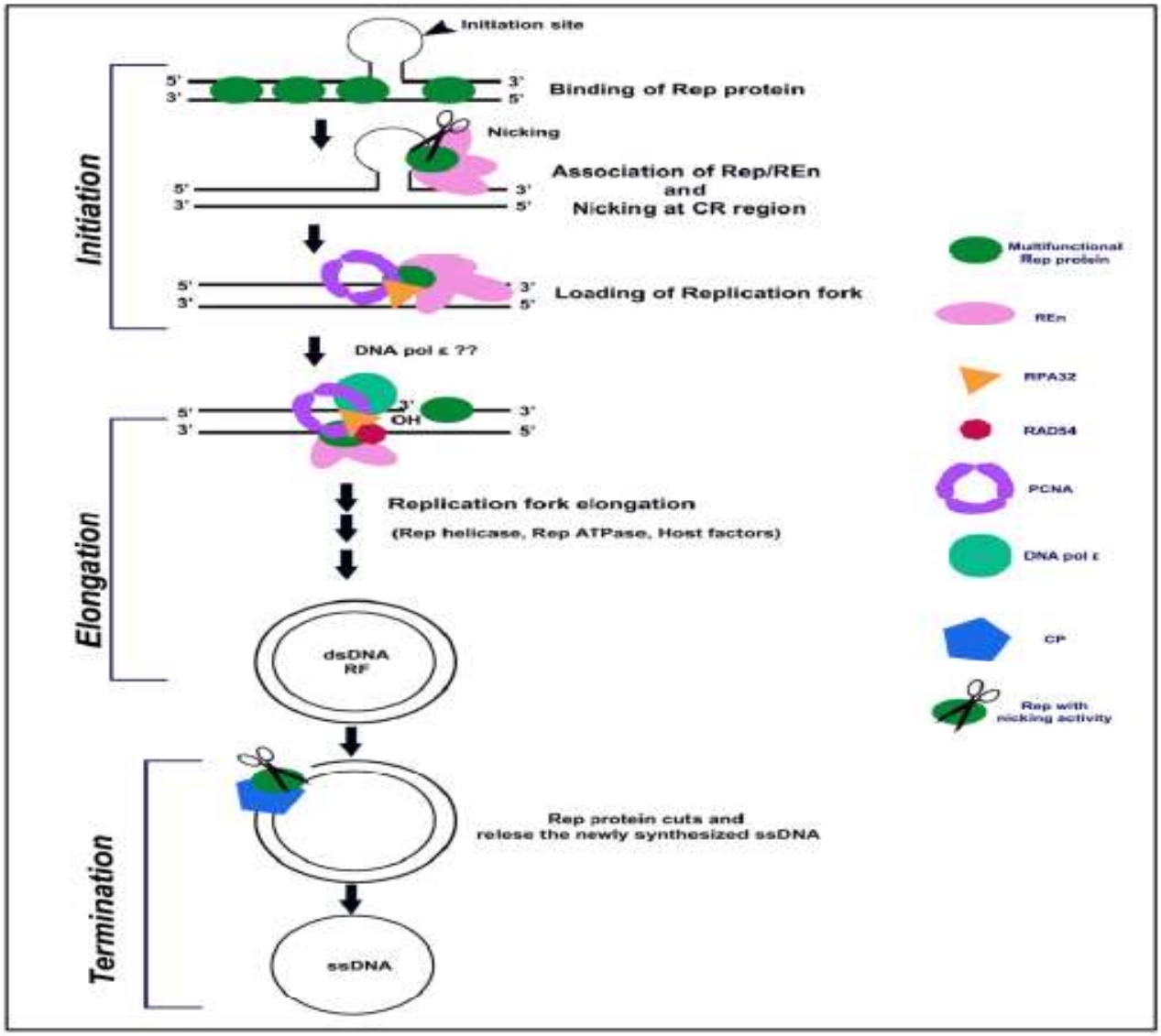


Figure 3: Schematic representation of geminiviral replication. Initiation: The geminivirus replication cycle relies mainly on DNA intermediates and occurs within the nucleus of the infected cell Replication initiation, Rep protein binds at CR region of DNA-A followed by nicking of DNA at initiation site. A Replication fork complex is formed with helicase and DNA synthesis activities. Final step termination, where Rep protein cuts and releases newly

synthesized ssDNA to generate many copies of viral ssDNAs (Molecular Biology of Geminivirus DNA Replication, Pradhan et al., 2017).

2.7 Transcriptional regulation

Geminivirus Rep is known to auto-regulate its own expression and viral gene expression. Rep is necessary for replication, it negatively regulates its own transcription (Hayley *et al*., 1992). In the analysis of bipartite begomovirus *African cassava mosaic virus*, it was found that both AC1 and AC4 promoters are negatively regulated by AC1 (Hong and Stanley, 1995). In another bipartite begomovirus *Tomato golden mosaic virus* (TGMV), AC1 promoter is negatively regulated both by its own product as well as by AC4 (Groning et al., 1994, Sunter et al., 1993). The Rep is known to bind between TATA-box and the transcription start site, probably hindering its own transcription by interfering with RNA polymerase (Kaerlas et al., 2000). The C-terminus domain of begomovirus Rep has a trans-activation domain, responsible for its transcriptional regulation (Gutierrez, 2002) i.e. responsible for Rep auto-regulation.

Sequence analysis of geminiviral promoter identified control elements such as conserved late elements (CLE) motif (consensus: GTGGTCCC; Arguello-Astorga et al., 1994) in AC2 of the bipartite begomovirus TGMV and many other begomoviruses. Study reported that CLE is essential for trans-activation by AC2 (Arguello-Astorga et al., 1994). Mutation in CLE motifs resulted in loss of AC2 responsiveness, which lead to conclusion of CLEs involvement in trans-activation (Ruiz-Medrano et al., 1999).

The LIR contained cis-acting DNA elements that are involved in begomovirus replication and transcriptional regulation in both strands i.e. complementary and viral-sense strand (Heyraud et al., 1995). In the bipartite begomovirus TGMV TATA-box and G-box motif mutational analysis impaired viral DNA replication which points towards the importance of these as important regulatory element that function as replication efficiency elements. (Eagles and Hanley-Bowdoin, 1997).

2.8 Geminivirus life-cycle

Whitefly carrying geminate particles, upon un-coating, release viral ssDNA to the infected plant cells cytoplasm and then the viral DNA is imported to the nucleus for amplification of viral DNA. The ssDNA subsequently gets converted to (dsDNA) with the use of host cell machinery. The viral DNA in association with the nuclear proteins can transcribe viral mRNAs at the nucleus and the latter ones are exported to cytoplasm for translation to different viral proteins. The viral proteins that are required in the nucleus to carry out replication, transcription or movement of viral DNA enter to the nucleus through nuclear pores. With the help of movement protein, viral DNA can enter adjacent cells through the plasmodesmatal channels and thus virus spread occurs (Figure 2.8).

Tissue specificity of the geminivirus infection is genetically determined. Experimental evidence suggesting that geminivirus tissue tropism has a viral genetic component (Rushing et al., 1987). Most geminiviruses are seen to be restricted to the vascular tissues. Some

geminiviruses in the genus begomovirus exhibit phloem limitations and are restricted to cells of the vascular system, whereas others can invade mesophyll tissue.

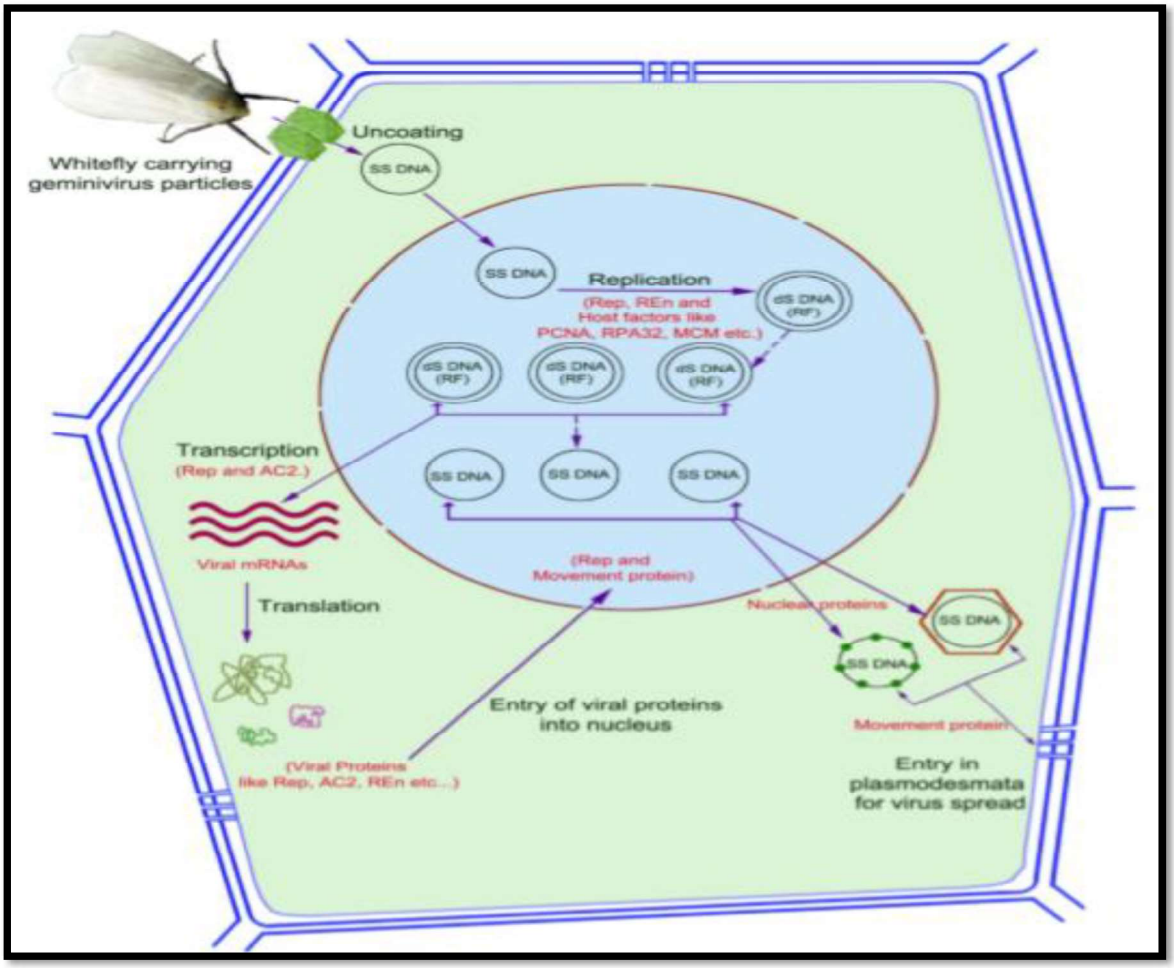


Figure 2.8: Representation of the geminivirus life-cycle. (Molecular Biology of Geminivirus DNA Replication, Pradhan et al., 2017).

2.9 Sri Lankan cassava mosaic virus

SLCMV is more closely related to ICMV (DNA-A, 84%; DNA-B, 94% nucleotide identity) than African cassava mosaic virus (ACMV) (DNA-A, 74%; DNA-B, 47% nucleotide

identity). SLCMV DNA-A has biological characteristics of a monopartite begomovirus, and the virus probably evolved by acquisition of a DNA-B component from ICMV (Saunders et al., 2002).

2.10 Sri Lankan cassava mosaic virus promoter element

In the bipartite SLCMV, bidirectional promoters have been identified in the intergenic regions (IGR) of DNA-A and DNA-B that controls bi-directional transcription. In the quest to investigate SLCMV promoter, work done by Basanta Kumar Borah, 2009 in his PhD thesis. Correlating all the data obtained from the deletion analysis of five forward and five reverse deletions and the transcript mapping experiment, a putative model of the CP-promoter region of SLCMV was put forwarded (Figure 2.10). The putative promoter region spanning from nucleotide position 2528 to 354 containing the intergenic region, as well as 73 nucleotides from the 5' end of Rep coding region (AC1) and 57 nucleotides from the 5' end of the coat protein (AV1) coding region was selected. The study confirmed that the intergenic region (IR) having varied regulation, depending upon the direction of transcription. The complementary-sense promoters have only 60% strength of the CaMV 35S promoter whereas the virion-sense promoter has strength almost equal to the full length CaMV 35S promoter.

The expression of the deletions was investigated using Gus reporter gene in different systems-

The prokaryotic system in *Agrobacterium.*

The eukaryotic system in BY2 protoplast.

The putative SLCMV-CP promoter region is divided into three regulatory elements as in the given table below.

Regulatory element	**Position w.r.t. SLCMV-[Adi] (AJ579307).**
Eukaryotic positive element.	2620 to 2725
Prokaryotic positive element.	2725 to 65
Eukaryotic and Prokaryotic negative element.	66 to 255

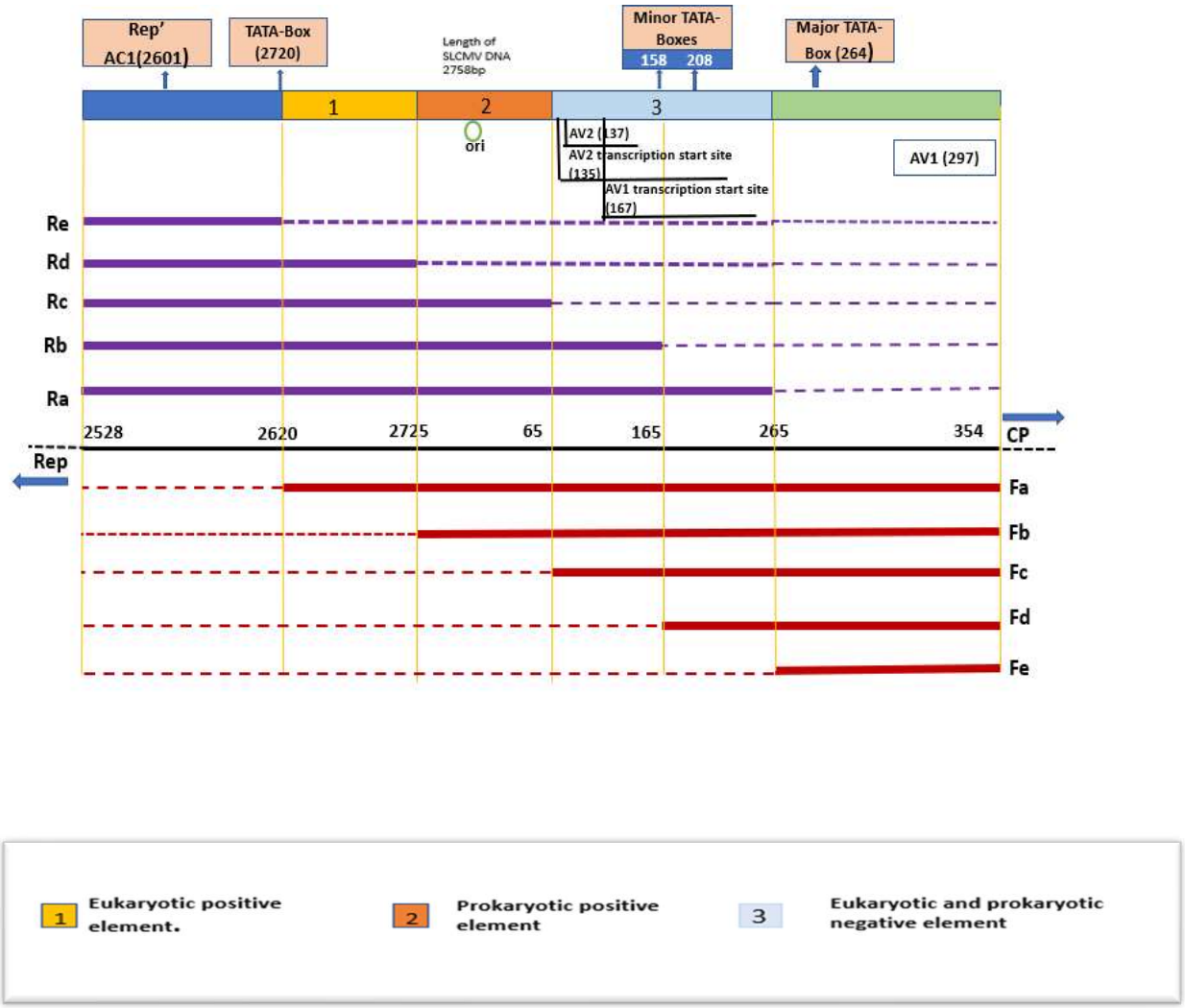


Figure 2.10: A putative model of the SLCMV CP-promoter region showing various regulatory region. (PhD Thesis, Basanta Kumar Borah, 2009).

Chapter-3

Materials and methods

3 MATERIALS AND METHODS

3.1 Strategy used to study the positive element IR1P in heterologous promoter

Viral clones: The positive element (IR1P) of the putative SLCMV-[Adi] (AJ579307) CP-promoter was already cloned earlier in the laboratory. The deletion constructs used Fc-TA, Fe-TA, a full length IR1-pUC19 GUS and CaMV 35S-pUC19 GUS (already constructed earlier in the laboratory).

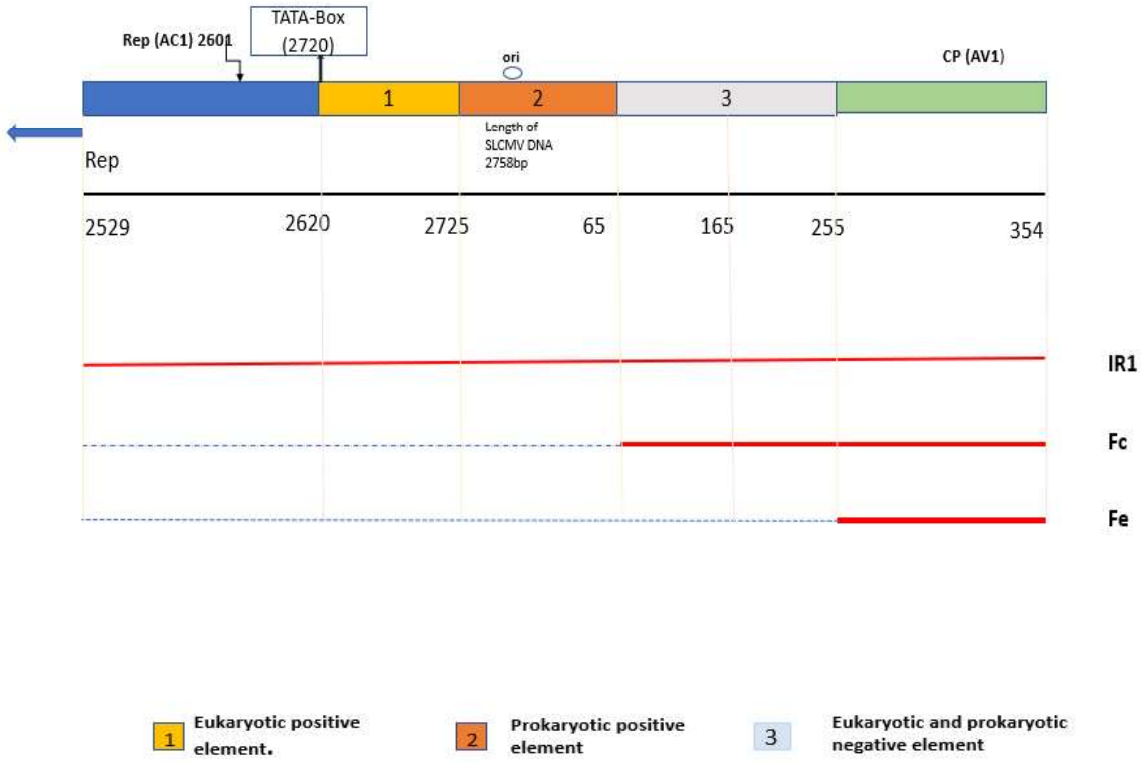


Figure 3.1: A representation of the SLCMV CP-promoter region with the positions of the deletions Fc, Fe and full length IR1 containing the positive element IR1P in itself

In the first step of cloning, the positive element IR1P was fist amplified using Phusion™ High fidelity DNA polymerase enzyme using common forward primer (IR1P-F with restriction site HindIII) and respective reverse primers (IR1P/R/Fc, IR1P/R/Fe, IR1P/R/35S, IR1P/R/IR1). Here, an overlap of the construct was created with IR1P so as to amplify it in the next round of cloning. In the second step, joining the positive element (IR1P) containing overlap of the deletion constructs with upstream of the deletion fragments of Fc (56 to 354), Fe (247 to 354), full length IR1(2529 to 354) containing the positive element in itself and one control using CaMV 35S. The PCR product was ethanol washed. It was then A-tailed followed by a second ethanol wash and finally ligated to T/A vector. The ligated

DNA was transformed by Heat shock protocol using chemical competent *E. coli* strain DH5-α. The colonies were screened for positive clones by colony PCR protocol. The putative positive clones were further confirmed by plasmid isolation and restriction digestion using restriction enzyme HindIII and BamHI (NEB enzyme). The positive clones were then confirmed by sequencing using universal M13 forward primer.

After confirming the clone, the insert was digested out of the T/A clones by restriction digestion with the enzymes HindIII and BamHI. The vector backbone (pBI121) was similarly digested. The desired DNA was eluted using gel extraction kit. The insert was then ligated to the vector backbone using the set up in table 3.4. The ligated DNA was then transformed by heat shock using chemical competent cells *E. coli* DH5-α strain. The colonies were screened for the positive clones which was then supposed to be further tested for confirmation using restriction digestion.

3.2 Isolation of plasmid (From *E. coli* DH5-α strain)

Plasmid isolation was done using alkaline lysis method.

5 ml of *E. coli* DH5-α strain Culture was grown overnight in incubator shaker maintained at 37ºC at 200 rpm. The cells were harvested by centrifugation at 13,000 rpm, 3 minutes in a 2 ml MCT. The pellet was re-suspended in 250 µl resuspension solution and vortexed till the pellet was dissolved. 250 µl of freshly prepared Lysis II was added and the contents were mixed by gentle inversion and kept at room temperature (RT) for 2-3 minutes. 300 µl of ice-cold Lysis III was then added and mixed gently by inversion and kept on ice for 15 min. It was then centrifuged at 13,000 rpm for 10 min at 4ºC. Supernatant was collected in fresh MCT and 610 µl of isopropanol was added to it and was kept on ice for 10-20 min for precipitation. The MCT was spin at 13,000 rpm for 10 min at 4°C. The supernatant was discarded and 300µl volumes of ethanol added to it. DNA was pelleted by centrifugation at 13,000 rpm, at 4°C for 5 minutes followed by air drying the pellet at RT for about 30 min. The isolated plasmid was dissolved in 20 µl of MQ.

The isolated plasmid was checked on 1% agarose gel made in 1XTBE buffer with a molecular weight marker.

3.3 Gel isolation of desired DNA fragment

Using a sharp scalpel, excised the band by cutting the gel surrounding the band very carefully with minimize size of the gel slice to just contain the DNA band. The agarose band with the DNA was dissolved in 1:1 volume of binding buffer (GeneJET gel extraction kit). It was then incubated at 50°C for 10 minutes (until the gel slice was completely dissolved) followed by loading into the column provided in the kit and centrifuged. The column was again centrifuged after adding wash buffer, and given an empty spin. 15 µl elution buffer was added to the extracted DNA. This was again checked on agarose gel for checking the final DNA concentration purified.

3.4 PCR amplification

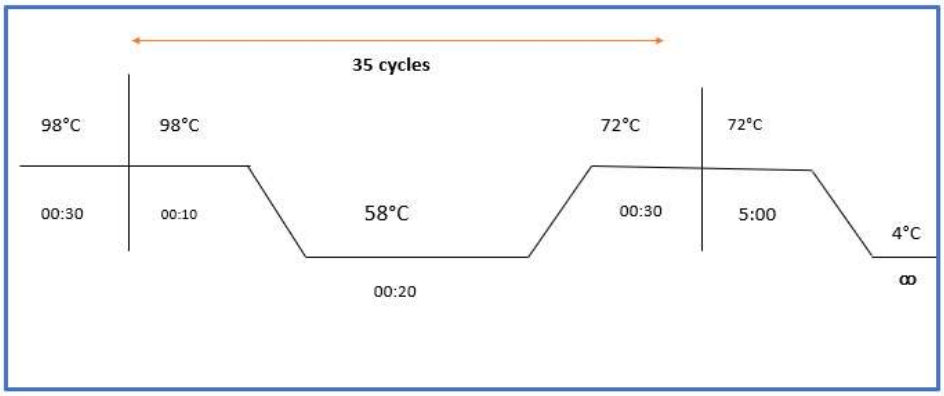
Amplification of DNA was carried out using Phusion™ High-Fidelity DNA polymerase using the primers designed. (The details of the primers used are given in appendices). The annealing temperature was optimized according to the melting temperature of the primer pair. The PCR amplicons were checked on 1% to 1.5% agarose gel (different gel concentration for varying amplicons sizes).

The PCR conditions and composition of the reaction are given below.

Table 3.2: Composition of reaction mixture used for PCR amplification

PARTICULARS	volume(µl) for 1X reaction
Buffer (HF 5X)	1.5µl
50mM Mgcl2	0.3µl
10mM dNTP	0.4µl
Forward primer (10µM)	0.5µl
Reverse primer (10µM)	0.5µl
Template	5.0µl
Phusion DNA polymerase	0.2µl
MQ	6.6µl
Total	15µl

The thermal cycle for PCR amplification is given below



3.5 Poly A tailing for blunt end PCR products

Adding of A-tail at the end of blunt ends of the PCR products generated by Phusion™ High fidelity DNA polymerase. The PCR product was first purified by ethanol washing. This step is important so as to remove any residual salt surrounding the DNA. The reaction was set-up according to table 3.3. The

reaction mix was incubated at 72°C for 30 minutes, followed by snap cooling on ice and stored at -20°C.

Table 3.3: Composition of reaction mixture for Poly-A tailing

POLY A tailing for IR1P	
PARTICULARS	volume (µl)
PCR Product	10µl
10X Taq buffer	5µl
10mM Datp	2.0µl
Taq polymerase (NEB)	2.0µl
MQ	31µl

3.6 Ligation with vector with T-overhangs

The poly-A tailed PCR product was ligated with T- overhang containing vector pTZ57R/T. The reaction protocol is given in table 3.4. The ligation mix was incubated overnight at 16°C.

Table 3.4: Composition of reaction mixture for Ligation

LIGATION of IR1P-TA	
PARTICULARS	volume (µl)
pTZ57R/T	1.2µl
5X Ligation	4.0µl
PCR Product	10µl
T4 DNA Ligase	1.0µl
MQ	3.8µl
Total	20µl

3.7 Preparation of *E. coli* chemical competent cells

Single fresh colony of bacterial strain DH5-α was streaked one day before on LB agar media plate containing Nalidixic acid (stock 30mg/ml) and incubated overnight at 37°C. A single colony of the bacterial strain E.coli DH5-α was inoculated in 5ml Luria Bertani (LB) broth with 5 µl of Nalidixic acid and was kept overnight in incubator shaker at 37°C at 200rpm. One percent of overnight culture was inoculated in 100 ml LB broth and incubated at 37°C at 200 rpm till the O.D. reaches 0.4 to 0.5 at 595nm. The culture was kept on ice for 10 minutes and pelleted by centrifugation at 5,000 rpm for 10

minutes at 4°C in Sorvall SS34 rotor. Supernatant was discarded and the pellet was resuspended gently in 10 ml pre-chilled 100mM $CaCl_2$. After complete resuspension of pellet, the centrifugation was carried out at 6000 rpm for 10 minutes at 4°C and the supernatant was discarded. This step was repeated thrice. The pellet obtained after three washes was dissolved in final 2 ml ice-cold 100 mM $CaCl_2$. The pellet was resuspended in 13% of glycerol (prepared in 100 mM $CaCl_2$) for storing in -80ºC. The cells were then aliquoted (100 µl) in 80 sterile 1.5 ml MCTs tubes and were frozen in liquid nitrogen. The chemically competent cells were then freshly used or stored at -80°C for further transformations protocols. The method was based on Hanahan method for preparation of E. coli chemical competent cells.

3.8 Transformation using chemical competent cells by heat shock

The competent cell which were stored in -80°C were thawed on ice for 5-10 minutes. The competent cell (*E. coli DH5-α*) was mixed with 10 µl ligated sample and kept on ice for 30 minutes followed by a cycle of heat shock at 42°C (heat shock for 1 min/quick thaw on ice for 1 min/heat shock for 30 sec) and the cells were then placed back on ice for 10 min. After this 1 ml of LB media was added to the cell suspension & was incubated at 37ºC at 200 rpm for 1 hour. The cells were subsequently pelleted at 5000 rpm for 5 minutes, re-suspended in 100 µl of LB media and plated on LB-agar plate with appropriate antibiotics (Carbenicilin and Nalidixic acid) and kept in incubator at 37ºC overnight.

3.9 Colony PCR for Confirmation of positive colonies

Screening of positive bacterial colonies was done by performing colony PCR. The observed colonies on plate were screened using colony PCR with gene specific primers. Single colony were picked with help of toothpick and patched on LB-agar plate and the very same colonies were dislodged in PCR tubes for screening positive clones. The reaction mixture is given in table 3.5.

Table 3.5: Composition of reaction mixture for colony PCR

PARTICULARS	volume(µl) for 1X reaction
10X buffer	1.5µl
50mM Mgcl2	0.3µl
10mM dNTP	0.4µl
Forward primer (10µM)	0.5µl
Reverse primer (10µM)	0.5µl
Template	5.0µl
Taq DNA polymerase	0.2µl
MQ	6.6µl
Total	15µl

3.10 Confirmation of positive clones by Restriction digestion

The bacterial clones which screened positive in colony PCR were further confirmed by restriction digestion. The restriction digestion was done using restriction enzymes (NEB) HindIII and BamHI. The reaction set-up components are given in table 3.6. Prepared reaction was kept at 37º C for two hours for digestion and then checked on agarose gel (1-1.5% concentration for varying amplicon sizes) for results.

Table 3.6: Composition of reaction mixture for Restriction digestion

PARTICULARS	Volume (µl)
MQ	5.2
Buffer	2
HindIII	0.4
BamHI	0.4
DNA	10
Total	20

Chapter-4
Results and Discussion

4 Results and Discussion

4.1 Results

To study the positive element of the putative CP-promoter region (IR1) of SLCMV promoter by deletion analysis.

The cloning of IR1P with four different constructs was first performed by amplifying the fragments by PCR with the respective primer pairs. Figure 4.1 showing the amplification of the positive element of putative CP-promoter region IR1P (107bp). In the first step, an overlap of deleted construct was created with IR1P fragment (~20bp), so as to amplify it in the next round (Figure 4.2). The PCR products were A-tailed and ligated to T/A vector.

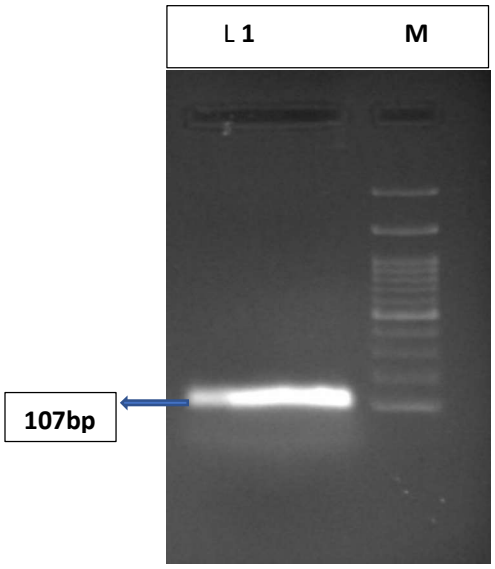


Figure 4.1: PCR amplification of positive element of SLCMV promoter IR1P (lane1). M1: 100bp ladder.

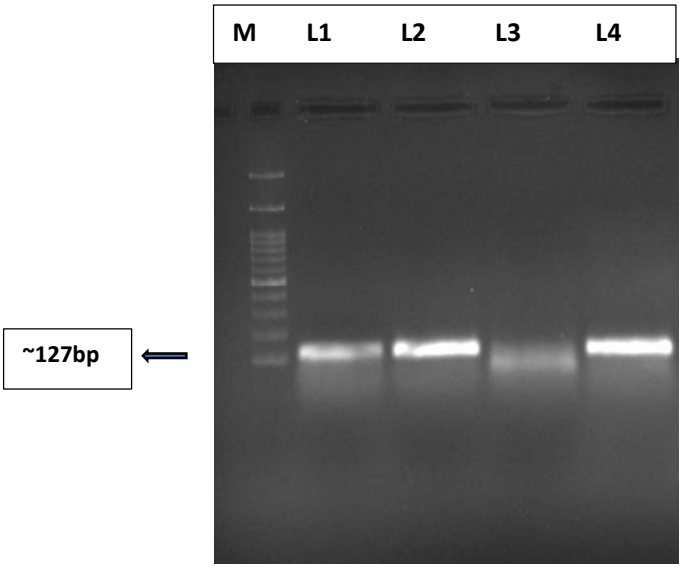


Figure 4.2: PCR amplification of IR1P amplified with ~20bp of each deletion construct primers. M-100bp DNA ladder, Lane 1- IR1P/Fc-TA, lane 2-IR1P/Fe-TA, lane 3-IR1P/35S-TA, lane 4- IR1P/IR1-TA.

IR1P with the four different constructs overlap is represented as IR1P/Fc-TA, IR1P/Fe-TA, IR1P/IR1-TA, IR1P/35S-TA.

The positive T/A clones were further confirmed by sequencing using universal M13 forward primer. Alignment analysis following sequencing showed that these were the positive clones. Next the strategy was to join the positive element (IR1P) containing overlap of the deletion constructs with upstream of the deletion fragments of Fc (56 to 354), Fe (247 to 354), full length IR1 containing the positive element in itself and one control using CaMV 35S (Figure: 4.4 and 4.5). The restriction digestion pattern for the deletion with restriction enzyme BamHI an HindIII is shown in figure 4.6 and 4.7. The binary vector PBI121 was digested using BamHI and HindIII to release the vector backbone to prepare it for ligation. The eluted fragments of IR1P Fc, IR1P Fe, IR1P 35S were cloned into this binary vector pBI121 and confirmed for the presence of insert by colony PCR protocol (Figure 4.9, 4.10, 4.11).

(confirmation using restriction digestion for the three final clones in binary vector is not done yet, also for IR1P IR1 cloning I could not proceed till binary clone)

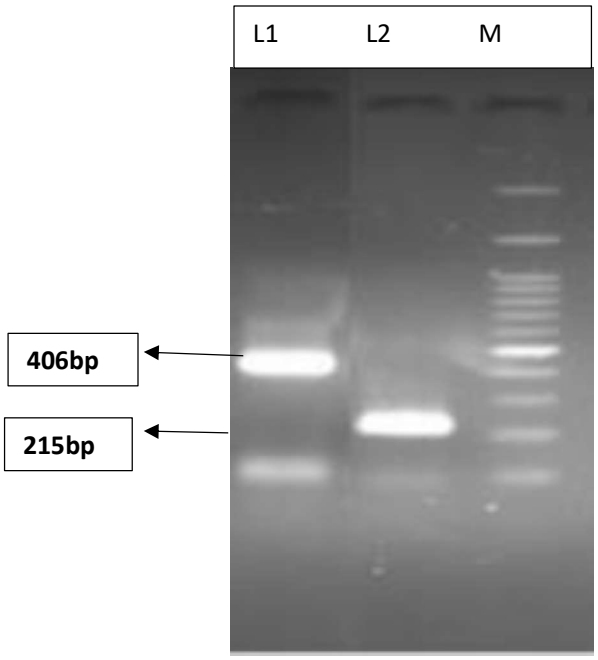


Figure 4.4: PCR amplification of IR1PFc-TA (L1) and IR1PFe-TA (L2). M: size marker 100bp ladder.

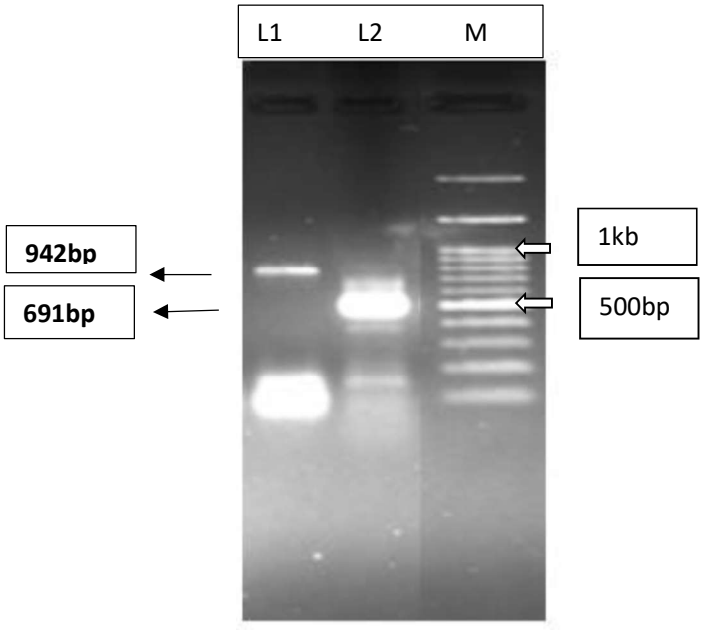


Figure 4.5: PCR amplification of IR1P35S-TA (L1) and IR1PIR1-TA (L2). M: size marker 100bp ladder.

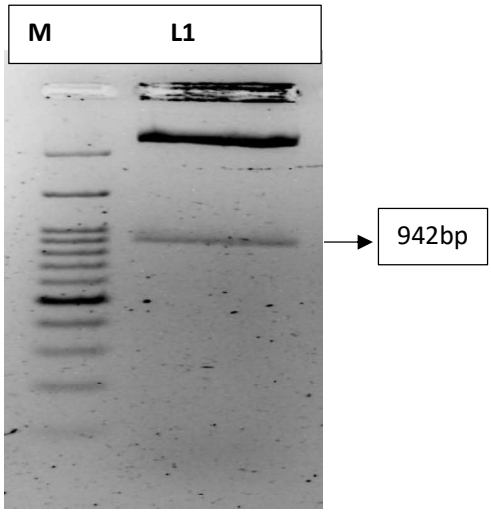


Figure 4.6: Restriction digestion of IR1P35-TA by BamHI and HindIII to get release (lane 2).

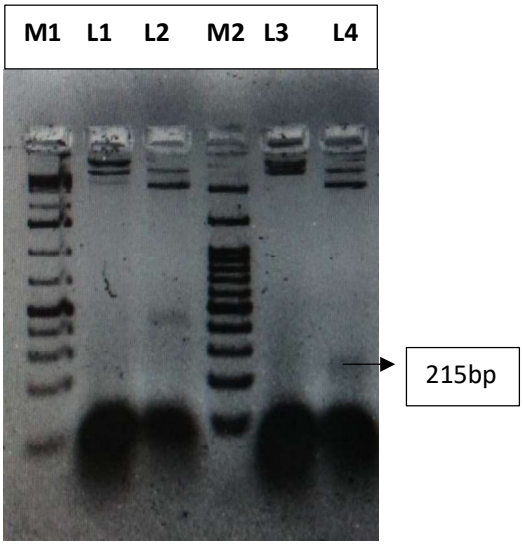


Figure 4.7: Restriction digestion by BamHI and HindIII to get release. L1- Undigested IR1PFc-TA. L2- Digested IR1PFc-TA L3: Undigested IR1PFe-TA L4-Digested IR1PFe-TA

The binary vector PBI121 was digested using BamHI and HindIII to release the vector backbone to prepare it for ligation (Figure 4.8). The eluted fragments of IR1P Fc, IR1P Fe, IR1P 35S were cloned into this binary vector pBI121 and confirmed for the presence of insert by colony PCR protocol (Figure 4.9, 4.10, 4.11).

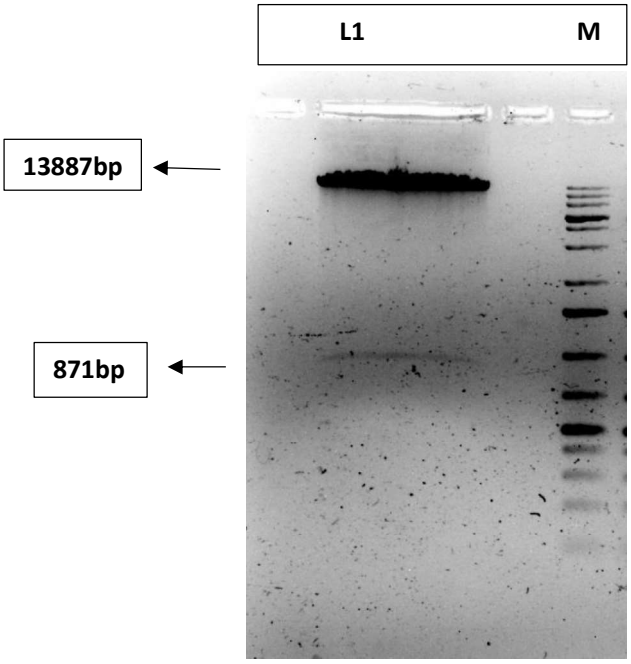


Figure 4.8: Restriction digestion pattern for Vector PBI121 using BamHI and HindIII, to get release and vector backbone. M- size marker DNA ladder of 1kb plus.

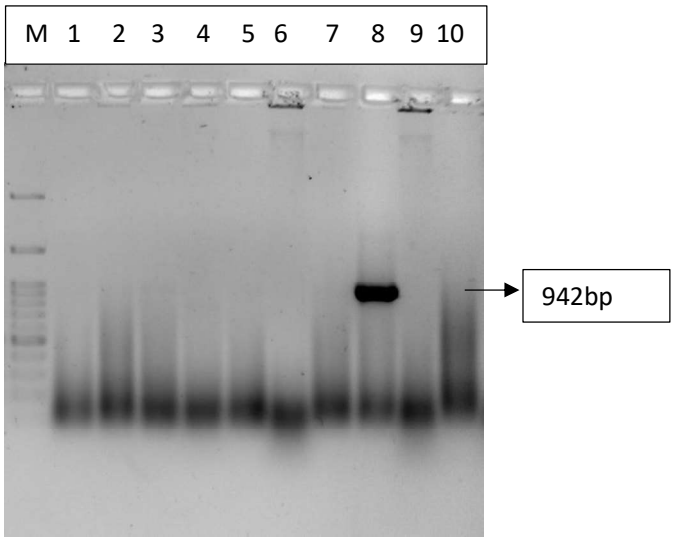


Figure 4.9: Amplicons showing confirmation of positive clone (lane 8) of IR1P35S-PBI121 by colony PCR.

In figure 4.10, the correct size of IR1PFe is 215bp. Here, lane 8 shows the correct size and no non-specific band. Hence, it could be a positive clone, which would be further confirmed by digestion.

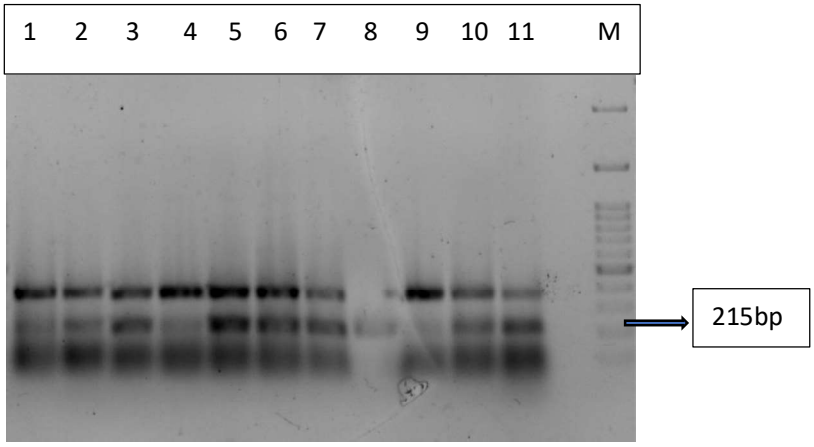


Figure 4.10: amplicons showing confirmation of positive clone **(lane 8)** of IR1PFe-PBI121 by colony PCR. M: size marker of 100bp DNA ladder

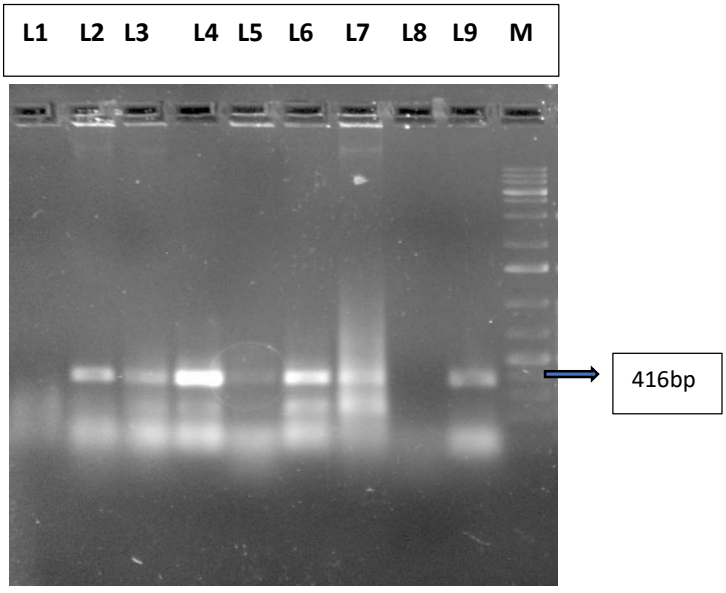
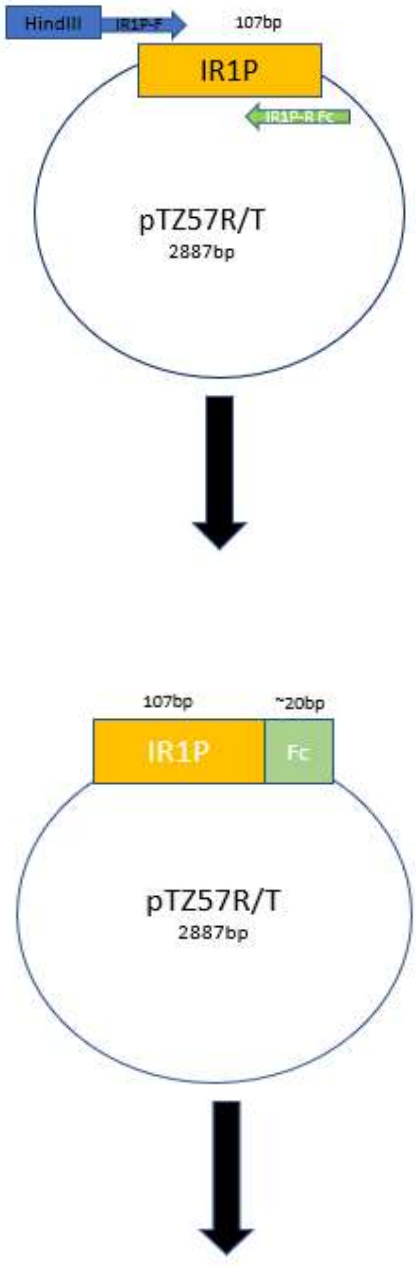


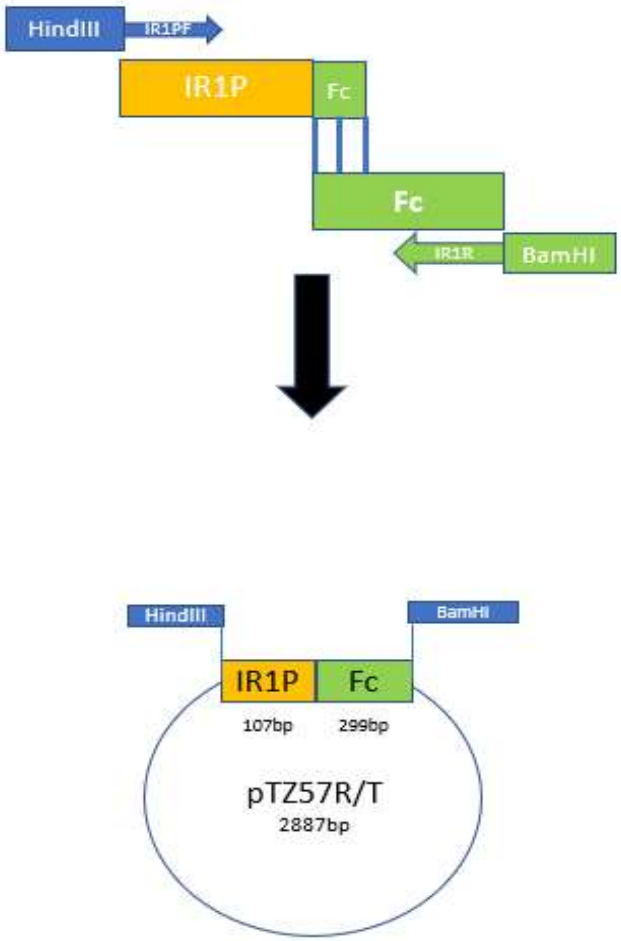
Figure 4.11: Amplicons showing confirmation of positive clones of IR1PFc-PBI121 by colony PCR. M: size marker of 100bp DNA ladder.

Here, in figure 4.11,all clones are positive, except lanes 1, 5 and 8 in which a very fainted smear like bands is visible. It might be possible that the correct ligation/insertion of insert DNA into the bacteria (E.coli DH5-α) was not successful/untransformed or it might be due to loading error.

Figure 4.12: A schematic representation of the cloning strategies which includes the different stages of TA cloning.



1st step of TA cloning: An overlap of deleted construct was created with IR1P fragment (~20bp).



2nd step of TA cloning: Joining the positive element (IR1P) containing overlap of the deletion constructs with upstream of the deletion fragments with desired restriction site added.

In the final/3rd step in binary clone: The insert IR1P Fc is digested out followed by elution and ligated in the final binary clone, pBI121.

The same cloning strategy was followed for all the constructs.

DISCUSSION

4.2 Deletion analysis of the positive element of IR1 Promoter

The putative CP-promoter region has already been analyzed in the laboratory (Ph.D thesis, Basanta Kumar Borah, 2009) by five forward and five reverse deletions which indicated the presence of 107 bp (nt position 2620 to 2725) positive element in the putative SLCMV CP-Promoter region. The positive element of the putative CP-promoter region was analysed further by construction of two forward deletions Fc (nt 56 to 354) Fe (nt 247 to 354) , a full length IR1 (2529 to 354) where the positive element (IR1P) is already present and one control using CaMV 35S. The aim of the study was to outline the positive element IR1P (nt 2620 to 2725) by placing the IR1P upstream of Fc, Fe, and cheeck how the expression/activity is influnced by transient expression study using GUS assay. With these clonings, i have obtained some of the constructs, which can be used for further work in the future.

Chapter-5

Conclusion and future work

5.1 CONCLUSIONS

This dissertation is aimed at further delineating of the 107bp (nt 2620 to 2725) positive element IR1P of the putative CP-Promoter region (IR1) by studying its effect on heterologous promoter by constructing a series of forward deletions Fc, Fe, full length IR1 having the positive element (IR1P) already and a positive control of full length CaMV 35S promoter. The aim was to study the expression of all the four constructs by studying GUS activity and histochemical staining of the stem sections. Three clones have been cloned in binary vector i.e. the final stages before infiltration, whereas others are in different stages of TA cloning.

5.2 FUTURE WORK

In the future work, after cloning the positive elements upstream of the different deletion, I would have done the GUS assay and check the expression of Gus gene driven by the promoter construct. It is expected some increase in the expression of Gus as compared to the control because of the positive element which is supposed to enhance the expression is attached to the promoter.

Chapter-6

Appendices and References

Appendices

Primer	Primers Sequences	Length of the primer. (bp)
IR1P-F	***AAGCTT***GGAGACACCCCTGAGCAA	17
IR1P-R	ATTCGAATTTAAAATTTGACTGTACCT	27
IR1P/R/IR1	CCTAAGTGTTCTCTCACCAAAGAATTCGAATTTAAAATTTGACTGTACCT	50
IR1P/R/Fc	TATACAGGCTTTGAGGGAGCGTTATTCGAATTTAAAATTTGACTGTACCT	50
IR1P/R/Fe	AGAGATCAAATCTCTGATTAAATATTCGAATTTAAAATTTGACTGTACCT	50
IR1P/R/35S	TCTGAAATTGAAAAGGCTAATCTATTCGAATTTAAAATTTGACTGTACCT	50
IR1R	**G*GGATCC***GGCGACGAACCTTC	
35S/F/IR1P Pink overlaps CaMV35S promoter	AGGTACAGTCAAATTTTAAATTCGAATAGATTAGCCTTTTCAATTTCAGA	50
35S/R/BamHI	***GGATCC***CGTGTTCTCTCCAAATGAAATG	22

References

Borah, B., Zarreen, F., Baruah, G., & Dasgupta, I. (2016). Insights into the control of geminiviral promoters. *Virology, 495*, 101-111. doi:10.1016/j.virol.2016.04.033

Dutt, N., Briddon, R. W., & Dasgupta, I. (2005). Identification of a second begomovirus, Sri Lankan cassava mosaic virus, causing cassava mosaic disease in India. *Archives of Virology, 150*(10), 2101-2108. doi:10.1007/s00705-005-0579-9

Ashraf, M., Shahid, A., Rao, A., Bajwa, K., & Husnain, T. (2014). Functional Characterization of a Bidirectional Plant Promoter from Cotton Leaf Curl Burewala Virus Using an Agrobacterium-Mediated Transient Assay. *Viruses, 6*(1), 223-242. doi:10.3390/v6010223

Cantú-Iris, M., Pastor-Palacios, G., Mauricio-Castillo, J. A., Bañuelos-Hernández, B., Avalos-Calleros, J. A., Juárez-Reyes, A., . . . Argüello-Astorga, G. R. (2019). Analysis of a new begomovirus unveils a composite element conserved in the CP gene promoters of several Geminiviridae genera: Clues to comprehend the complex regulation of late genes. *Plos One, 14*(1). doi:10.1371/journal.pone.0210485

Eini, O., Behjatnia, S. A., Dogra, S., Dry, I. B., Randles, J. W., & Rezaian, M. A. (2009). Identification of sequence elements regulating promoter activity and replication of a monopartite begomovirus-associated DNA satellite. *Journal of General Virology, 90*(1), 253-260. doi:10.1099/vir.0.002980-0

Fondong, V. N. (2019). The Ever-Expanding Role of C4/AC4 in Geminivirus Infection: Punching above Its Weight? *Molecular Plant, 12*(2), 145-147. doi:10.1016/j.molp.2018.12.006

Guerrero, J., & Sunter, G. (2019). Complementary-sense gene regulation in beet curly top virus-SpCT. *Archives of Virology, 164*(11), 2823-2828. doi:10.1007/s00705-019-04385-9

Rajeswaran, R., Sunitha, S., Shivaprasad, P. V., Pooggin, M. M., Hohn, T., & Veluthambi, K. (2007). The Mungbean Yellow Mosaic Begomovirus Transcriptional Activator Protein Transactivates the Viral Promoter-Driven Transgene and Causes Toxicity in Transgenic Tobacco Plants. *Molecular Plant-Microbe Interactions, 20*(12), 1545-1554. doi:10.1094/mpmi-20-12-1545

Shivaprasad, P. V., Akbergenov, R., Trinks, D., Rajeswaran, R., Veluthambi, K., Hohn, T., & Pooggin, M. M. (2005). Promoters, Transcripts, and Regulatory Proteins of Mungbean Yellow Mosaic Geminivirus. *Journal of Virology, 79*(13), 8149-8163. doi:10.1128/jvi.79.13.8149-8163.2005

Fondong, V. N. (2013). Geminivirus protein structure and function. *Molecular Plant Pathology, 14*(6), 635-649. doi:10.1111/mpp.12032

Tu, J., & Sunter, G. (2007). A conserved binding site within the Tomato golden mosaic virus AL-1629 promoter is necessary for expression of viral genes important for pathogenesis. *Virology, 367*(1), 117-125. doi:10.1016/j.virol.2007.04.018

Brown, J. K. (1995). First Report of an Epidemic in Tomato Caused by Two Whitefly-Transmitted Geminiviruses in Puerto Rico. *Plant Disease, 79*, 1250. doi:10.1094/pd-79-1250a

Hanley-Bowdoin, L., Settlage, S. B., Orozco, B. M., Nagar, S., & Robertson, D. (1999). Geminiviruses: Models for Plant DNA Replication, Transcription, and Cell Cycle Regulation. *Critical Reviews in Plant Sciences, 18*(1), 71-106. doi:10.1080/07352689991309162

Lazarowitz, S. G., & Beachy, R. N. (1999). Viral Movement Proteins as Probes for Intracellular and Intercellular Trafficking in Plants. *The Plant Cell, 11*(4), 535. doi:10.2307/3870882

Saunders, K., Lucy, A., & Stanley, J. (1992). RNA-primed complementary-sense DNA synthesis of the geminivirus African cassava mosaic virus. *Nucleic Acids Research, 20*(23), 6311-6315. doi:10.1093/nar/20.23.6311

Kong, L. (2000). A geminivirus replication protein interacts with the retinoblastoma protein through a novel domain to determine symptoms and tissue specificity of infection in plants. *The EMBO Journal, 19*(13), 3485-3495. doi:10.1093/emboj/19.13.3485

Gutierrez, C., Ramirez-Parra, E., Castellano, M. M., Sanz-Burgos, A. P., Luque, A., & Missich, R. (2004). Geminivirus DNA replication and cell cycle interactions. *Veterinary Microbiology, 98*(2), 111-119. doi:10.1016/j.vetmic.2003.10.012

Gutierrez, C. (2002). Strategies for geminivirus DNA replication and cell cycle interference. *Physiological and Molecular Plant Pathology, 60*(5), 219-230. doi:10.1006/pmpp.2002.0401

Kong, L., & Hanley-Bowdoin, L. (2002). A Geminivirus Replication Protein Interacts with a Protein Kinase and a Motor Protein That Display Different Expression Patterns during Plant Development and Infection. *The Plant Cell, 14*(8), 1817-1832. doi:10.1105/tpc.003681

Wezel, R. V., Dong, X., Blake, P., Stanley, J., & Hong, Y. (2002). Differential roles of geminivirus Rep and AC4 (C4) in the induction of necrosis in Nicotiana benthamiana. *Molecular Plant Pathology, 3*(6), 461-471. doi:10.1046/j.1364-3703.2002.00141.x

Kinship of APX2 and HSP101 in birthing and working together is important in protection of *Arabidopsis thaliana* seedlings from heat stress

Thesis Submitted to the University of Delhi

in the Partial fulfilment for the Degree of

Master of Science

in

Plant Molecular Biology and Biotechnology

2020

Haneet Papreja

Department of Plant Molecular Biology

University of Delhi South Campus

New Delhi -110021, India

Kinship of APX2 and HSP101 in birthing and working together is important in protection of *Arabidopsis thaliana* seedlings from heat stress

Thesis Submitted to the University of Delhi
in the Partial fulfilment for the Degree of

Master of Science

in

Plant Molecular Biology and Biotechnology

2020



Haneet Papreja

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi-110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, **"Kinship of APX2 and HSP101 in birthing and working together is important in protection of *Arabidopsis thaliana* seedlings from heat stress"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in the planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 06-06-2020

Signature of the Candidate

(HANEET PAPREJA)

Signature of the Supervisor

(PROF. ANIL GROVER)

Signature of the HOD

(PROF. ANIL GROVER)

PREFACE

Average global combined temperature of land and ocean surface has increased by 0.85°C between 1880 and 2012 (IPCC, https://www.ipcc.ch/). The concentrations of atmospheric carbon dioxide (CO_2) has been rising consistently, from 315 ppm (1959), 385 ppm (2008), 392.6 ppm (2014) and 403.95 ppm (2017). Current projections show that by the end of this century, the concentrations of CO_2 may reach 500–1000 ppm. The global mean temperature has increased due to increase in concentration of greenhouse gases (GHG) as they trap heat radiated from the earth. The mean global temperature would rise higher by 1°C and 3°C above the present temperature by the years 2025 and 2100 respectively, ultimately causing an acute global warming situation.

The increasing threat of climate change is already having a substantial impact on agricultural production worldwide as heat waves cause significant yield losses with great risks for future global food security. This higher temperature may alter distribution and growing season of crops by affecting the threshold temperatures for life processes.

With a 1°C increase in global temperature, global wheat production is projected to decrease between 4.1% and 6.4%.

High temperature stress or heat stress may be defined as the period in which ambient air temperature remains high for a sufficient period to harm the plant functions and/or its development in an irrecoverable manner.

There is an urgent need to genetically modify the crops to withstand heat stress. The discovery of candidate genes is a first step in this venture.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the help of many people in so many ways. It was also the product of a large measure of serendipity and fortuitous encounters with people that has changed the course of my academic career. I would take this opportunity to express my gratitude towards them.

First and foremost, I am highly grateful to my supervisor ***Prof. Anil Grover*** *for his guidance, support, kind advice and caring nature throughout the course of this dissertation. I am grateful for his supervision, time and ideas regarding this thesis and giving me complete freedom to think and work. I am highly grateful for the patience that he showed in teaching me even the basic concepts. The joy and enthusiasm he has for research is motivational for me and I would like to sincerely thank him for making every difficult task achievable and supporting me in every possible way.*

I would like to thank all my teachers, ***Prof. A.K. Tyagi****,* ***Prof. J.P. Khurana****,* ***Prof. P. Khurana****,* ***Prof. Indranil Dasgupta****,* ***Prof. Sanjay Kapoor****,* ***Prof. Arun K. Sharma****,* ***Prof. Girdhar K Pandey****,* ***Assoc. Prof. Saurabh Raghuvanshi*** *and* ***Asstt. Prof. Surekha Katiyar-Agarwal*** *for sharing their vast scientific experience and help me develop a scientific temperament.*

I wish to express my sincere thanks to ***Rinchuila Shimphrui*** *for her support, guidance and valuable time. I thank her for support and guidance throughout my stay in the lab. Without her help, accomplishing many experiments would not have been possible.*

I thank ***Dr. Neelam Sarkar*** *for teaching me the nitty-gritty of the protocols. She has been my fruit companion.* ***Dr. Chhavi Agarwal*** *for being cooperative during my lab tenure. I also wish to thank* ***Ritesh****,* ***Gayatri****,* ***Lisha****,* ***Garima****,* ***Richa*** *and* ***Manoj*** *for being exceptionally helpful.*

I would like to thank ***Vijyesh****,* ***Remya*** *and* ***Anjaly*** *for providing a wonderful environment in the lab to work in.* ***Brij Kishor*** *and* ***Vakil*** *also needs special thanks for keeping things in order in the lab and especially providing miscellaneous items on time. I once again would like to thank all of my lab members for arranging parties very frequently.*

I also thank my friends ***Mahima****,* ***Vartika****,* ***Monica****,* ***Anurag****,* ***Aastha****,* ***Ritesh*** *and* ***Frankie*** *for their supportive nature and all good times we had together.*

Lastly, I thank ***Almighty*** *and* ***my Family*** *for their constant encouragement, love and support without which this assignment would not have been possible.*

LIST OF ABBREVIATIONS

Abbreviation	Definition
APS	Ammonium per sulphate
bp	Base pair
BSA	Bovine serum albumin
˚C	Degree centigrade
cDNA	Complementary DNA
CBB	Coomassie Brilliant Blue
Da	Dalton
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
dNTP	Dinucleotide triphosphate
d	Day
EtBr	Ethidium bromide
EDTA	Ethylene diamine tetra acetic acid
g	Gram
h	Hour
HS	Heat stress
HSP	Heat shock protein
sHSP	Small heat shock protein
HSF	Heat shock factor
l	Liter
MQ	Milli-Q
MCTs	Microcentrifuge tube
µg	Microgram
mg	Miligram
µl	Microliter

ml	**Milliliter**
µM	**Micromolar**
mM	**Millimolar**
M	**Molar**
mA	**Milliampere**
min	**Minutes**
MOPS	**Morpholinopropranesulphonic acid**
MS medium	**Murashige-Skoog medium**
nm	**Nanometer**
NBT	**Nitroblue tetrazolium**
O.D.	**Optical density**
PAGE	**Polyacrylamide gel electrophoresis**
rpm	**Revolution Per Minute**
RT	**Room Temperature**
RNA	**Ribonucleic acid**
s	**seconds**
sqPCR	**Semi-quantitative PCR**
TAE	**Tris-acetate EDTA**
TEMED	**Tetra methyl ethylene diamine**
TRI	**TRIzol reagent**
V	**Voltage**
v/v	**Volume by volume**
w/v	**Weight by volume**

CONTENTS

i.	Preface	2
ii.	Acknowledgements	3
iii.	List of Abbreviations	4-5
1.	Introduction	7-9
2.	Review of literature	10-27
3.	Materials and Methods	28-35
4.	Results	36-46
5.	Discussion	47-50
6.	Conclusion and Future Work	51-52
7.	Literature cited	53-57
8.	Appendix	58-60

Introduction

INTRODUCTION

Growth and development of the plants is greatly affected in terms of morphological, biochemical and physiological changes by high temperature stress (Grover et al., 2013). Heat stress (HS) differentially affects the stability of various proteins, membranes, RNA species, cytoskeleton structures and alters the efficiency of enzymatic reactions in the cell, causing a state of metabolic imbalance.

The extent of damage from heat stress is determined by the environmental factors like intensity, duration, and rate of change in temperature and plant factors like stage of development and growing conditions before exposure to high temperature.

For survival in stressful conditions, plants continuously struggle to modify their metabolic process in many ways in response to HS, specifically by generating key solutes that leads to establish proteins and osmotic adjustment and re-establish the redox balance of cell and homeostasis by modifying the antioxidant system. A plant in defence from HS causes modifications at the molecular level in the expression of genes. In HS conditions, change in biochemical and physiological activities by gene expression alter gradually, resulting in the development of thermotolerance.

Heat shock proteins (HSPs) are produced in plants as a common phenomenon upon encountering stress. The proteins are largely chaperone by nature and manage misfolded and aggregated proteins during periods of stress and post-stress recovery. Cytosolic HSP101 is important for thermotolerance in plants. It is the only chaperone that apart from refolding misfolded proteins, also helps in reactivation of aggregated proteins. Thus, it is an important candidate for maintaining cellular protein homeostasis.

Oxidative damage is usually a subsequent stage of most of the abiotic stresses in plants. Exposure of plants to heat stress initially causes oxidative damage by the formation of reactive oxygen species (ROS) generally hydrogen-peroxide (H_2O_2), hydroxyl-radical ($OH^{\bullet}$), superoxide-radical ($O_2^{\bullet-}$), and singlet-oxygen (O_2). These ROS pose serious threat to the cell functioning by damaging lipids and proteins mainly produced in the chloroplast however, reaction of oxygen with the components of electron transport chain in mitochondria also results in the generation of ROS (Møller, 2001). The mechanisms

involved in the scavenging of the ROS can be either enzymatic or non-enzymatic (Apel and Hirt, 2004).
The damaging effects of HS can be reduced by inducing thermotolerance in plants using different transgenic and genetic engineering techniques.

Ascorbate peroxidase (APX) belongs to the family of heme containing peroxidases that catalyse the H_2O_2-dependent oxidation of a wide range of organic molecules. Ascorbate peroxidase activity has been established in higher plants, eukaryotic algae and certain cyanobacteria. The earliest event in ascorbate peroxidase evolution resulted in the formation of separate groups of APX-cytosol and chloroplast ascorbate peroxidases. There are multigenic families of APX in higher plants for e.g., *Arabidopsis* has 9 *APX* genes (*AtAPX1-AtAPX6, sAPX, tAPX, mAPX*). The accumulation of *APX1* transcripts occurs in response to high light stress. It is speculated that the expression of *APX1* is indirectly controlled by HsfA4a via another transcription factor ZAT12. Since APX1 is a heat-labile protein and its function might be compensated by APX2 under heat stress. Studies show that HsfA2 is responsible for the adequate expression of APX2.

HSP101 protein performs protein disaggregase activity. This protein has important role to play in proteostasis. This protein is strongly implicated in acquisition of heat stress tolerance.

This dissertation work aimed at understanding the relationship between HSP101 and APX2 transcript and protein accumulation in relation to development of heat tolerance. The major objectives of this study were as follows:

1. Establishment of heat stress phenotyping of HSP101 over-expressing *Arabidopsis* lines which are *a priori* generated in the laboratory
2. Analysis of *HSP101* transcript expression in relation to heat tolerance phenotype
3. Analysis of *APX* gene family in *Arabidopsis* and analysis of heat inducibility of *APX2* transcript
4. Unveiling relationship, if any, between *HSP101* expression and *APX2* expression in temporal and spatial sense

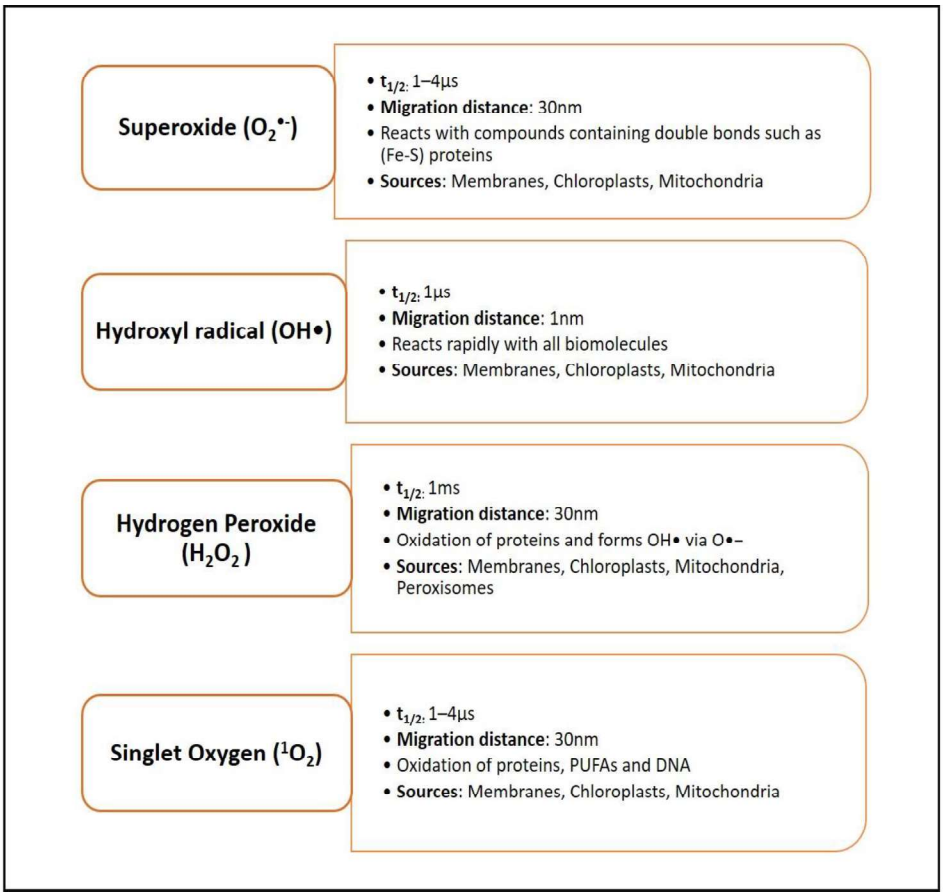
Review of Literature

REVIEW OF LITERATURE

Higher plants are sessile, they are often exposed to wide range of environmental stresses, such as high light (HL), heat shock (HS), drought, chilling and salinity that frequently imposes limitations on their growth and productivity. It is generally known that the imposition of environmental stress results in an increase in the concentration of ROS and that much of the damage caused to plants by exposure to stress involves oxidative damage at the cellular level (Shigeoka et al., 2002). Therefore, plant responds to a change in cellular redox status under several types of abiotic stress by activating a complex variety of tolerance mechanisms and integrating it with the expression of thousands of genes encoding proteins involved in several biological processes (Nishizawa-Yokoi et al., 2010).

ROS (e.g. $O_2^{\bullet-}$, H_2O_2, $OH^{\bullet}$) are partially reduced or activated forms of atmospheric oxygen (O_2). They are considered to be ineluctable by-products of aerobic metabolism that have accompanied life on Earth ever since the appearance of oxygen-evolving photosynthetic organisms about 2.2–2.7 billion years ago (Mittler et al., 2011).The evolution of higher plants has thus occurred in the presence of ROS and they have acquired dedicated pathways to protect themselves from ROS toxicity and to use ROS as signaling molecules as well (Foyer and Noctor, 2013; Choudhury et al., 2017). If the levels of ROS are kept unchecked, their concentration increases in cells and cause oxidative damage to membranes, proteins, RNA and DNA molecules, and even lead to the oxidative destruction of the cell in a process termed as oxidative stress (Mittler, 2002).

Under normal growth conditions, the production of ROS in cells is low (240 μM s^{-1} O_2^- and a steady-state level of 0.5 μM H_2O_2 in chloroplasts), however the production of ROS enhances (240–720 μM s^{-1} O_2^- and a steady-state level of 5–15 μM H_2O_2) as a consequence of disruption of the cellular homeostasis of cells under stress conditions (Polle, 2001)



$t_{1/2}$ (half life): time required for ROS to reduce to half its initial value.

Migration distance: distance the oxidant travels before being consumed.

Fig 1: Different members of ROS family and their attributes.

However, this process is alleviated by the presence of a large number of ROS detoxifying proteins [e.g. ascorbate peroxidase (APX), superoxide dismutase (SOD), catalase (CAT), glutathione peroxidase (GPX) and peroxiredoxin (PRX)] as well as by antioxidants such as ascorbic acid and glutathione (GSH) in almost all subcellular compartments in the cell (Mittler et al., 2004a). The dynamic process of ROS detoxification in plant cells is also supported by diverse metabolic adaptations that decrease ROS generation, by maintaining the level of free transient metals such as Fe^{2+} under control, to prevent the formation of the highly toxic hydroxyl radical ($OH^{\bullet}$) via the Fenton reaction (Halliwell and Gutteridge, 2015). On the contrary, plants actively produce ROS that are used in the process of signal transduction. They are mainly produced at the apoplast by NADPH oxidases (termed respiratory burst oxidase homologs; RBOHs) and some oxidases and peroxidases and at the chloroplast, mitochondria, peroxisome and other cellular compartments via different pathways (Suzuki et al., 2011; Gilroy et al., 2016; Choudhury et al., 2017)

There are two major sources of ROS during abiotic stress, they include-

- **Metabolic ROS**: They are produced as a consequence of disruptions in metabolic activity and could directly change the redox status of rate-limiting enzymes and influence metabolic fluxes in the cell, thereby altering different metabolic reactions in order to counteract the effects of stress (Miller et al., 2010). Additionally, it could affect transcription and/or translation by changing the function of key regulatory proteins via ROS-derived redox modifications (Foyer and Noctor, 2013)
- **Signaling ROS**: The purpose of production of these ROS is to function as signaling molecules as a part of the abiotic stress–response signal transduction network and result as a response to stress perception by stress sensors (e.g. cyclic nucleotide-gated channels activated by heat stress (Mittler et al., 2012). It is facilitated by calcium- and/or phosphorylation-derived activation of NADPH oxidases (RBOH) at the plasma membrane (Suzuki et al., 2011); they are also believed to directly alter the redox state of regulatory proteins and modify transcription and translation leading to the activation of an acclimation response that would diminish the effects of stress on metabolism and decrease the level of metabolically produced ROS (Choudhury et al., 2017)

Even though metabolic and signaling ROS could be produced at different subcellular compartments (e.g. metabolic ROS in the chloroplast and signaling ROS at the apoplast), but they can affect the level of each other and even move between the compartments.

The frequent process of ROS production (metabolically or for signaling purposes) and ROS scavenging occurs at all cellular compartments of the cells and is controlled by the ROS gene network. In *Arabidopsis thaliana* approximately 152 genes are involved in managing the levels of ROS (Mittler et al., 2004a).

The different ROS levels in the different subcellular compartments results in generation of a particular ROS signature. This signature can change depending on the type of cell, its developmental stage or magnitude of stress. Different ROS signatures are formed in plant cells depending on different abiotic stresses and/or different combinations of abiotic stresses (stress combination) and deciphering these signatures by different ROS sensors can produce a stress-specific signal that will adjust the acclimation response to the type of stress/combination affecting the plant (Choudhury et al., 2017).

MAJOR ROS PRODUCING SITES DURING ABIOTIC STRESS

The major ROS-producing sites during abiotic stress are the mitochondria, chloroplast, peroxisome and apoplast (Gilroy et al., 2016; Choudhury et al., 2017).

Mitochondrial ROS accumulation during abiotic stress is facilitated by electron leakage from complex I and III to produce O_2^-, which can be converted to H_2O_2 by MnSOD (Choudhury et al., 2017). This process can be alleviated by alternative oxidase (AOX), type II NAD(P)H dehydrogenase and uncoupling proteins in the inner mitochondrial membrane.

Chloroplastic ROS: Abiotic stresses that constrain CO_2 availability due to stomatal closure increase the production of ROS such as O_2^- and $^1O_2$ in chloroplasts and is relieved by a range of ROS-scavenging enzymes and pathways such as Fe- and Cu/Zn-SODs and the Asada–Foyer– Halliwell pathway, as well as high concentrations of antioxidants such as ascorbic acid and GSH (Das and Roychoudhury, 2014).

Generation of ROS in peroxisomes during abiotic stress is mainly the consequence of increased photorespiration leading to the production of H_2O_2 by glycolate oxidase (Foyer and Noctor, 2009; Gill and Tuteja, 2010).

EFFECTS OF ACCUMULATION OF ROS

1. Membrane lipid peroxidation

At the cellular level, the major effect of ROS production and accumulation is lipid peroxidation of both cell and organelle membranes. The two main targets of the ROS in membrane phospholipids are the double bond between C-atoms and the ester linkage between glycerol and fatty acids.

The membrane polyunsaturated fatty acids (PUFA) [linoleic acid (18:2) and linolenic acid (18:3)] are the main targets of $O_2^{\bullet-}$ and $OH^{\bullet}$ and elevate the formation of the polar lipid hydroperoxide complex mixtures, resulting in enhanced membrane fluidity, cytosolic solutes efflux and membrane associated protein functional loss (Møller et al., 2007). Aldehydes such as malondialdehyde (MDA) and 4-hydroxy-2-nonenal are created as a result of PUFA peroxidation which attack amino acid side chains in proteins and also cause fragmentation of DNA (Gill and Tuteja, 2010).

2. Protein oxidation

The important ROS-mediated post-translational modifications are sulfonylation, carbonylation, glutathionylation and S-nitrosylation.

Sulfonylation is the oxidation of sulfhydryl groups, one of the major mechanisms that regulate the activity of many enzymes and transcription factors (TFs) in plants. H_2O_2 generating sulfenic acid (R-SOH) mainly induces the oxidation that can cause disulfide (S–S) bond formation between cysteine residues and resulting in conformational changes that modify protein/enzyme activity. The first step in the ROS-dependent redox signalling pathway is the oxidation of cysteine residues to sulfenic acid. This process can be reverted mainly by thioredoxins (Trxs), peroxiredoxins (PRXs) and glutathione (GSH) system, which act in response to stress and regulate redox homeostasis (Khedia et al., 2019).

S-Glutathionylation is the further alteration of the side-chains of proteins containing sulfenic acid through covalent bond formation with low molecular weight thiols (e.g. GSH) and can control redox-driven signal transduction cascades and metabolic pathways (Khedia et al., 2019).

Carbonylation is yet another form of protein oxidation which occurs at arginine, histidine, lysine, proline and threonine residues and considered to be irreversible (Shacter, 2000).

During stress, S-nitrosylation, the covalent binding of nitric oxide (NO) to thiol groups of Cys results in the modification of some proteins, modulating the signalling cascades. S-nitrosylation is also reported to affect the DNA binding activity of some transcription factors and render RBOH inactive (Choudhury et al., 2017).

3. DNA

The mitochondrial and chloroplastic DNA both bear the brunt of the ROS attack due to the absence of protective histones unlike nuclear DNA and also due to the close proximity to ROS production machinery. The damage to DNA as a result of ROS occurs at various levels which include oxidation of the deoxyribose sugar residue, alteration of the nucleotide base, abstraction of a nucleotide, breaks in either DNA strand, and cross-linking of the DNA and protein.

$OH^{\bullet}$ disrupts the deoxyribose sugar backbone by extracting hydrogen atom and also reacts with double bonds of the purine and pyrimidine bases (Das and Roychoudhury, 2014). The ROS abstracts the C-4H atom of the deoxyribose sugar and forms a deoxyribose radical which reacts further to cause single strand breaks in the DNA (Evans et al., 2004).

The base oxidation results in the production of 8-hydroxyquinine and other less common products like hydroxyl methyl urea, dehydro-2'-deoxyguanosine, thymine glycol, and thymine and adenine ring opened. $OH^{\bullet}$ also reacts with either DNA or associated proteins creating DNA-protein cross-links which are not easily reparable and may be lethal to the plant cell, if not repaired in time before commencement of critical cellular processes like replication or transcription (Das and Roychoudhury, 2014).

INTERACTION OF ROS WITH PLANT HORMONES

The accumulation of ROS during abiotic stress has an impact on the level and function of different plant hormones, such as abscisic acid (ABA), auxin, brassinosteroids (BRs), gibberellins (GAs) and nitric oxide. Interactions between ROS and auxin cause auxin imbalance and result in altered growth. Stress-induced ROS generation can modify auxin gradients in the plant and also diminish auxin-mediated signaling (Xia et al., 2015).

Though the genetic basis of a crosstalk between brassinosteriods (BRs) and ABA remains to be explained, the idea of a positive feedback mechanism in which BRs induce the transient accumulation of H_2O_2, through the activation of RBOH and hence the induction of ABA biosynthesis has been proposed (Zhou et al., 2014).

Of late ROS were shown to facilitate rapid systemic signaling in plants in response to abiotic stress in a process that is linked to calcium signaling and maybe even electric waves (Miller et al., 2009; Mittler et al., 2011; Gilroy et al., 2016). In order for a plant to attain maximal fitness in the field, the response of all of its organs and leaves to abiotic stress must be coordinated. This coordination is believed to be achieved by an auto-propagating wave of ROS production (the ROS wave) that is initiated in a group of cells that first senses the stress and spreads it to the entire plant at a rate of up to 8.4 cm min^{-1} (Miller et al., 2009).

Recent studies have indicated that the ROS wave is facilitated by RBOHD in *Arabidopsis* and that it is coordinated with a systemic calcium wave (Gilroy et al., 2016). Additionally, the ROS wave is required for the generation of some electric signals during abiotic stress (Suzuki et al., 2013) and also for induction of a systemic acclimation response to light or heat stress (Suzuki et al., 2013; Choudhury et al., 2017)

ROS SCAVENGING ENZYMES

Since abiotic stresses generate abundant ROS which have deleterious effects, its detoxification is of prime importance to protect the integrity of the cell. To maintain a redox homeostasis, antioxidant defence systems are constantly activated in the plant (Table 1; modified version from the publication of Das and Roychoudhury (2014).

These antioxidant systems remove the excess ROS which is not required for basic plant processes and stabilize the internal biochemical state of the cell during various abiotic stresses, leading to acclimatization as well as tolerance (Bowler and Fluhr, 2000).

ROS scavenging pathway also have a role in maintaining a low steady baseline of ROS on which different signals can be registered (Mittler et al., 2004b).

Table 1: List of all the enzymatic and non- enzymatic antioxidants along with their functions.

Enzymatic antioxidants	Reaction catalysed	Non-Enzymatic antioxidants	Functions
Superoxide dismutase (SOD)	$O_2^{\bullet-} + O_2^{\bullet-} + 2H+ \rightarrow 2\ H_2O_2 + O_2$	Ascorbic Acid (AA)	Reacts with H_2O_2, OH•. Preservation of activities of metal binding enzymes
Catalase (CAT)	$2H_2O_2 \rightarrow O_2 + 2H_2O$	Reduced Glutathione (GSH)	scavenges H_2O_2, $_1O^2$, OH•, and $O_2^{\bullet-}$
Ascorbate peroxidase (APX)	$H_2O_2 + AA \rightarrow 2H_2O + DHA$	α -Tocopherol	Protects the membrane by reacting with O_2 halts the chain propagation step of lipid peroxidation cycle
Monodehydroascorbate reductase (MDHAR)	$2MDHA + NADH \rightarrow 2AA + NAD$	Carotenoids	Protects photosynthetic machinery by quenching excess energy
Dehydroascorbate reductase (DHAR)	$DHA + 2GSH \rightarrow AA + GSSG$	Flavonoids	scavengers of H_2O_2 and $^1O_2$ and $OH^{\bullet}$
Glutathione reductase (GR)	$GSSG + NADPH \rightarrow 2GSH + NADP^+$	Proline	scavenger of $OH^{\bullet}$ and $^1O_2$ and prevent damages due to lipid peroxidation
Guaiacol peroxidase (GPX)	$H_2O_2 + DHA \rightarrow GSSG + 2H_2O$		

Superoxide dismutase (SOD) is ubiquitous in all aerobic organisms. It forms the first line of defence against ROS-induced damages under environmental stresses. On the basis of metal ion it binds, SODs are classified into three isozymes : Mn-SOD (localized in mitochondria), Fe-SOD (localized in chloroplasts), and Cu/Zn-SOD (localized in cytosol, peroxisomes, and chloroplasts) (Mittler, 2002). Abiotic stress conditions cause upregulation of their expression (Das and Roychoudhury, 2014).

Catalase (CAT) has high affinity for H_2O_2, but lesser specificity for organic peroxides (R-O-O-R). It is distinctive from other antioxidant enzymes as it does not require a reducing equivalent. Three CAT genes have been reported in Angiosperms: CAT1 (localized in peroxisomes and cytosol), CAT2 (localized in peroxisomes and cytosol) and CAT3 (localized in the mitochondria). Stressful conditions demand for greater energy generation and expenditure of the cell which is realised by increased catabolism that leads to generation of H_2O_2. CAT eliminates the H_2O_2 in an energy efficient manner (Das and Roychoudhury, 2014). The role of CAT is mainly focused in neutralizing the excess of ROS during stressful conditions, whereas APX is more involved in the fine modulation of ROS for signalling (Lin et al., 2010).

Glutathione reductase catalyzes the reduction of glutathione disulfide (GSSG) to reduced glutathione (GSH), which plays an important role in maintaining the GSH level and redox state and safeguarding the cells against ROS. The non-enzymatic antioxidants include ascorbic acid, GSH (reduced glutathione), flavonoids, α-tocopherol, phenolics, carotenoids and amino acid cum osmolyte proline. In addition to protecting different components of the cell from damage, they also play a critical role in plant growth and development by regulating cellular processes like mitosis, cell elongation, senescence and cell death (Das and Roychoudhury, 2014).

Ascorbate peroxidase is a class of plant peroxidases. Plant peroxidases have been classified into three classes on the basis of the amino acid sequences (Welinder, 1992): (i) Class I: It includes the intracellular peroxidases of prokaryotic origin. APX, yeast cytochrome *c* peroxidase (CCP) and cyanobacterium CPX belong to this class (Shigeoka et al., 2002; Welinder, 1992), (ii) Class II: It contains the fungal peroxidase enzymes including manganese peroxidase and (iii) Class III: It includes the classical secretory peroxidases e.g., horseradish peroxidase.

ASCORBATE PEROXIDASE

Ascorbate peroxidase activity has been established in higher plants, eukaryotic algae and certain cyanobacteria. They are the intracellular enzymes which are encoded in the nucleus.

The initial detection of APX was in isolated intact chloroplasts and algae (Groden and Beck, 1979; Shigeoka et al., 1980). It contributes maximally to hydrogen peroxide detoxification and is therefore reported to be an efficient regulator of ROS.

APX belongs to the family of heme containing peroxidases that catalyse the H_2O_2-dependent oxidation of a wide range of organic molecules. It differs from other peroxidases due its dependency on ascorbate as the source of reducing power and becomes unstable in its absence (Shigeoka et al., 2002).

It participates in the elimination of H_2O_2 as a part of the ascorbate–glutathione or Asada–Halliwell–Foyer pathway (Foyer and Halliwell, 1976). APX uses AsA as its specific electron donor to reduce H_2O_2 to water with the concomitant generation of monodehydroascorbate (MDAsA), a univalent oxidant of AsA. MDAsA is spontaneously disproportionated to AsA and dehydroascorbate (DAsA). MDAsA is also directly reduced to AsA by the action of NAD(P)H-dependent MDAsA reductase. DAsA reductase utilizes glutathione (GSH) to reduce DAsA and thereby regenerate AsA. The oxidized GSH is then regenerated by GSH reductase, utilizing reducing equivalents from NAD(P)H. Thus, APX in combination with the effective AsA–GSH cycle functions to inhibit the accumulation of harmful levels of H_2O_2 in photosynthetic organisms (Asada, 1992; Shigeoka et al., 2002).

The earliest event in the evolution of ascorbate peroxidase resulted in the formation of separate groups of APX-cytosol and chloroplast ascorbate peroxidases. The divergence of chloroplast branch into stromal and thylakoid-bound types has occurred recently only via gene duplication in *Arabidopsis* or alternative splicing in pumpkin and spinach. On the contrary, the origin of microsomal APX in all species might have occurred from the APX ancestor or the cytosolic APX1 branch via gene duplication. On the basis of the analysis performed with APX sequences from other plants, cytosolic, chloroplastic (i.e., stromal and thylakoid APX) and microsomal APXs have originally generated from the ancient *APX* gene (Shigeoka et al., 2002).

There are multigenic families of APX in higher plants for e.g., *Arabidopsis* has 9 *APX* genes (*AtAPX1-AtAPX6, sAPX, tAPX, mAPX*) and *Oryza sativa* has 8 isozymes (OsAPX1-OsAPX8), having two APX each in cytosol, chloroplast, mitochondria and peroxisome whereas tomato has 7 genes encoding different APX isozymes (Pandey et al., 2017).

On the basis of amino acid composition, five isoforms of APX have been classified in plants viz cytosolic (cAPX: APX 1 and 2), mitochondrial (mitAPX), chloroplastic (chlAPX: stromal-APX and thylakoidal-APX), and peroxisomal/glyoxysomal (mAPX).They originated as an outcome of alternative splicing, which contributes to the differential regulation of expression of various isoforms (Caverzan et al., 2012; Pandey et al., 2017).

The different isozymes of APX, which are classified on the basis of their sub-cellular localization, show different structural and kinetic properties including the presence of specific conserved domains and signal peptides. The isoforms vary in their properties like molecular weight, optimal pH, stability, catalytic rate, and substrate affinity (Pandey et al., 2017).

The occurrence of multiple molecular forms of APX within the organelles and cells indicates the significant role played by them in developmental processes and stress tolerance (Shigeoka et al., 2002).

As APXs are heme-dependent oxidoreductases, iron is crucial for the catalytic activity and iron deficiency decreases its activity. Similar effects are noted when the concentration of ascorbate is reduced, wherein the activity and stability of the enzymes is negatively affected (Pandey et al., 2017).

APX is unable to scavenge lipid hydroperoxides. APX being a heme-containing enzyme whose prosthetic group is protoporphyrin, is inhibited by cyanide and azide (Shigeoka et al., 1980) and also by thiol- modifying reagents such as *p*-chloromercuribenzoate and suicide inhibitors such as hydroxylamine, *p*-amino-phenol, and hydroxyurea (Pandey et al., 2017).

The level of expression of *APX* corresponds directly to the duration and intensity of the stress imposed and the multiplicity of stresses.

Arabidopsis thaliana possess nine APX isoenzymes with different localizations: APX1, APX2 and APX6 are soluble cytosolic isoenzymes; APX3, APX4 and APX5 are microsomial isoenzymes, and sAPX and tAPX are localized in the chloroplast. The molecular details of the nine APX forms are shown in Table 2 (modified version from the publication of Mittler et al. (2004b).

Table 2: Gene and protein annotation for the ascorbate peroxidase (APX) gene family in *Arabidopsis*.

Protein name	Gene Identifier	Localisation	Molecular Weight(Da)	Isolelectric point
APX1	At1g07890.1	Cytosol	27561	5.72
APX2	At3g09640.1	Cytosol	27415	6.01
APX3	At4g35000.1	Microsomes	31571	6.47
APX4	At4g09010.1	Lumen	37934	8.59
APX5	At4g35970.1	Microsomes	30895	8.8
APX6	At4g32320.1	Chloroplast	18214	5.92
APX7	At1g33660.1	Mitochondria	11721	10.15
Stromal APX	At4g08390.2	Stroma/Mitochondria	40407	8.31
Thylakoid APX	At1g77490.1	Thylakoid	46092	6.81

APX STRUCTURE AND SUBSTRATE BINDING

The first APX crystal structure to be resolved was of pea cytosolic APX (Patterson et al., 1995). The domain structure of the monomer unit consists of Domain I (N- and C-termini, helices A–D) and Domain II (helices F–J). The domains are connected via the E helix. The J helix, which begins in Domain II, extends into Domain I where the C-terminus ends within 4 Å of the N-terminus. APX has a truncated C-terminus. The access channel connects the molecular surface to the distal heme pocket. APX has Arg172 which is one of the key amino acids that interacts with a heme propionate and is involved in ascorbate binding critical for ascorbate peroxidase activity. The phenolic substrates possibly bind near the exposed heme edge.

APX forms a homodimer. The dimer interface of the APX homodimer is mediated primarily by electrostatic interactions among Asp, Glu, Arg and Lys side chains. Although the monomer is fully active but the dimer contributes substantially to stability (Mandelman et al., 1998). There also is a major structural difference between the cytosolic and chloroplastic APXs. The chloroplastic APX has an extra loop, residues 176–193, near the heme propionates (Wada et al.,2003). In addition, the chloroplastic enzyme has no direct salt bridges or H-bonds between the heme propionates and amino side chains

found in the cytosolic enzyme. The cytosolic enzyme is substantially more stable than the chloroplastic enzyme in the absence of ascorbate.

FUNCTIONAL ANALYSIS OF APX

APX activity increases in response to a number of stresses including drought, high light, high temperature, chilling, iron, and salt stress.

APX1 gene is expressed constitutively in roots, leaves, stems and many other plant tissues. It has a relatively high level of expression in the absence of stress in contrast to other APXs and its expression is enhanced in response to sulfur dioxide, ozone, excessive light, and also by heat stress (Mittler, 2002; Shigeoka et al., 2002). Induction by heat stress is attributed to heat shock element (HSE) sequences that are the binding site for HSFs. The expression of *APX1* is also controlled post-transcriptionally during pathogen attack and recovery from drought stress (Mittler and Zilinskas, 1994). *A. thaliana* plants deficient in APX1 experienced growth suppression, stomatal closure, higher sensitivity to oxidative stress, accumulation of increased levels of H_2O_2 and oxidised proteins under light stress. (Pnueli et al., 2003).

APX2 is not expressed under controlled growth conditions. Its expression is related to a systemic signaling by hydrogen peroxide, and it is involved in the signaling of a variety of environmental stress responses including hypersensitive response, systemic acquired resistance, tolerance to chilling, and cross-tolerance to a range of biotic and abiotic stresses (Panchuk et al., 2002). The heat inducibility of *APX2* mRNA shows much stronger induction than that of *APX1* mRNA upon heat stress (Panchuk et al., 2002).

The heat tolerance of *Arabidopsis* seeds may be because of the presence of cytosolic APX6, an APX isoenzyme replacing APX1 in developing and germinating seeds (Chen et al., 2014). The seeds lacking APX6 accumulate high levels of ROS, display increased oxidative damage, and show reduced germination under control conditions. These effects escalate during HS. In addition, ripening *APX6*-deficient seeds exposed to HS display reduced germination vigour (Chen et al., 2014).

Genes of microsomal *APX3* and *APX4* show increased mRNA levels after short-term heat stress but no induction is observed upon exposure to stress for long duration.

The expression of *APX5* occurs at only extremely low levels of temperature. Genes for chloroplastic *tAPX, sAPX* mRNAs show only little variation with growth temperatures and no clear response to heat stress.

ROS AND HEAT STRESS

Heat stress can be described as an increase in the temperature above a threshold level for a period of time, sufficient to cause irreparable damage to plant growth and development. Sudden increase in temperature to higher levels may result in cell death within a few minutes as an outcome of extensive protein denaturation and aggregation and loss of membrane integrity (Driedonks et al., 2015).

Abiotic stress factors such as heat disrupt the equilibrium between the process of ROS generation and scavenging (Foyer and Noctor, 2005). There are enough evidence that when plants are exposed to heat, ROS production rapidly becomes extreme (Volkov et al., 2006; Bhattacharjee, 2013; Hasanuzzaman et al., 2013; Das and Roychoudhury, 2014).

One of the best analysed anti-stress mechanisms is the production of heat shock proteins (HSPs) upon exposure to high temperatures (Wang et al., 2004). HSPs act as molecular chaperones and prevent deleterious protein conformations and remove non-native aggregates, which are formed during stress (Vierling, 1991; Morimoto, 1998). The heat shock transcription factors (HSFs) (Driedonks et al., 2015) regulate the expression of HSPs and other heat-responsive genes. HSFs associate to a palindromic binding motif (5′-nAGAAnnTTCTn-3′) called the heat shock element (HSE) (Driedonks et al., 2015; Scharf et al., 2012) present in the promoter (upstream of the TATA box) region of the heat-responsive genes. Activation of HSFs upon stress occurs in a multistep process which involves homotrimer formation and acquisition of transcriptional competence for target gene induction (Liu et al., 2013).

Although they vary significantly in size and sequence, the basic structure of HSFs is conserved. HSFs have a modular structure comparable to many other transcription factors, with a highly conserved N-terminal DNA binding domain (DBD), which is required for recognition of the heat stress promoter elements. Sequencing of the *Arabidopsis* genome (The Arabidopsis Genome Initiative, 2000) revealed the presence of a unique complexity of the plant HSF family with 21 members: 15 type A, 5 type B, and

1 type C, differentiated by their domain architectures. Only class A HSFs are thought to be involved in the transcriptional activation of HS-inducible genes due to the presence of an activation domain at the C-terminus (Czarnecka-Verner et al., 2000; Davletova et al., 2005).

In *Arabidopsis* so far only HSFs A1a, A1b, A4a and B1 are functionally characterized in some detail. The functional analysis of the C-terminal domains (CTD) of the 21 HSFs described the CTD of HsfA2 as a potent activation domain with two AHA motifs which interact with the transcriptional machinery and a nuclear export signal (NES) being responsible for nucleocytoplasmic shuttling (Schramm et al., 2006).

In non-stressed conditions, the HSFs are localised in the cytoplasm of most eukaryotes in an inactive monomeric form due to their association with Hsp70, Hsp90 and potentially other proteins (Morimoto, 1998; Schöffl et al., 1998). According to the chaperone titration model, heat causes higher load of denatured proteins, which pull HSPs away from the HSF complexes through competition in order to act as molecular chaperones. This then results in the release of HSFs, which form trimers and move to the nucleus to activate expression of HSP and other heat-responsive genes (Driedonks et al., 2015).

Toxic amounts of heat-induced ROS also stimulate protein denaturation. ROS increases dissociation of the HSP–HSF complex, as described by the titration model (Schöffl et al., 1998). H_2O_2 might directly alter HSFs and induce their trimerization (Miller and Mittler 2006). There is an evidence that the transcription factors may be sensitive to H_2O_2 via "single-Cys" or "two-Cys" redox sensory mechanisms (Mittler et al., 2011). These cysteine residues typically respond to oxidative stress. HsfA1a contains one Cys residue located at the N-terminal portion of the trimerization domain (Driedonks et al., 2015).

It may be hypothesized that the heat shock proteins positively affect thermotolerance by protecting the conformation and activity ROS scavenging protein, resulting in a lower ROS concentration (Kong et al.,2014a).

JUB1 (ANAC042), a member of the NAC family of transcription factors, positively influences juvenility, tolerance to H_2O_2 treatment and survival of heat-stressed seedlings. JUB1 enhances general thermotolerance in the absence of priming and also acts during priming to further diminish the detrimental effect of high temperature. Thus, it resembles HsfA2 which also play an critical role in thermomemory in *Arabidopsis* (Shahnejat-bushehri et al., 2012).

APX2 EXPRESSION IN HEAT STRESS

Exposure of plants to high light stress results in the accumulation of *APX1* transcripts. It is hypothesised that the expression of *APX1* is indirectly regulated by HsfA4a via another transcription factor ZAT12. APX1 is a heat-labile protein so its function might be compensated by APX2 under heat stress. HsfA2 might be responsible for the sufficient expression of *APX2* (Schramm et al., 2006).

Among the 21 members of the *Arabidopsis* Hsf family, HsfA1a and HsfA1b initiate heat stress response by inducing downstream genes, including HsfA2, which then acts as a positive feedback regulator to sustain the expression of heat-induced genes during recovery or prolonged HS. In tomato, HsfA1a/b is required for the induction of HsfA2 by HS (Mishra et al., 2002), but in *Arabidopsis,* induction of HsfA2 is independent of HsfA1a/b (Busch et al., 2005). HsfA2 is the strongest expressed HSF under heat stress as shown by microarray analysis and is induced in almost all the tissues tested including leaves, stem, flowers, roots and green siliques. It accumulates to levels comparable to the abundant HSPs. HsfA2 is associated with a network of interacting proteins like HsfA1a, Hsp17-CII and Hsp17-CI influencing its solubility, intracellular localization and activator function (Schramm et al., 2006). HsfA2 plays a role in extending the duration of acquired thermotolerance and not in the induction of heat stress response (Charng et al., 2007).

The accumulation of denatured proteins in the cytosol cause induction of HsfA2 expression (Sugio et al., 2009).

In order to understand the function of HsfA2 in heat stress and identify its interacting proteins, (Schramm et al., 2006) created HsfA2 knockout mutant. Microarray was done to evaluate the changes in the transcriptome of the HsfA2 knockout (*HsfA2 ko*) plants compared to the wild-type plants under heat stress. This group identified subset of genes whose expression was reduced in *HsfA2 ko* plants upon heat stress. The strongly affected core cluster consist of Hsp26.5-P(r), Hsp25.3-P, Hsp22.0-ER, Hsp18.1-CI, Hsp70b, Hsp101-3 and DnaJ-C5. Comparison at the transcription level between wild-type and *AtHsfA2 ko* plants revealed that transcripts of *APX2* were present in the wild-type plants but absent in heat shock induced leaves of the knock out background (Schramm et al., 2006).

APX2 displayed a stronger induction by heat stress (Panchuk et al., 2002). AtHsfA2 has also been found to act as an *APX2* activator. Transient reporter assay was performed to analyse the transcription activation potential of HsfA2 suggesting the presence of

necessary target sequences for HsfA2 binding which initiate transcription by recruiting endogenous transcriptional machinery (Schramm et al., 2006).

GUS reporter construct assay with different fragments of *APX* promoter was used to map the binding site of HsfA2 in the promoter. Deletion of whole promoter upstream of the TATA box lead to complete loss of activity.

Results indicate that TATA proximal HSE dimer is essential for a transcriptional activation by HsfA2 and an adjacent downstream HSE cluster had some added positive impact (Schramm et al., 2006).

The reports on involvement of *APX1* and *APX2* on heat stress response of *Arabidopsis thaliana* are summarised in Table 3.

Table 3: Reports on involvement of *APX1* and *APX2* genes in heat stress in *Arabidopsis*.

S.No	Experimental set-up	Heat stress	Comments	Reference
1.	7 weeks old *Arabidopsis thaliana* WT (Col-24) and Hsf3 overexpressing transgenic plants (Hsf3-TP) used.	Moderate and severe HS at 37°C and 44°C (1,2 or 4h) followed by cultivation at 28°C for 3d. Incubation at 28°C and 34°C for 3d (long term)	*APX1* and *APX2* show stronger HS inducibility as compared to other APX forms. Induction of *APX2* more than *APX1*. In Hsf3-TP, APX activity enhanced after HS and appearance of a new isoform APX^S. In Hsf3-TP, APX^S activity was present at normal and severe HS whereas in WT it was induced at moderate HS. APX^S and APX2 are identical. Involvement of HSF in *APX2* induction indicates link between oxidative and heat stress signalling.	Panchuk et al., (2002)
2.	25 days old *A. thaliana* Col-0, *apx1, apx2* (Knockout mutants), *apx1/apx2* (Double mutant)	Mutants were subjected to 42°C/ 45°C for 2h for 28 days and control plants were grown under control conditions for 21-28 days.	*APX1* disruption resulted in enhanced ROS accumulation. It had no significant role in HSR of reproductive tissue. *APX2* disruption led to reduced HS tolerance at seedling stage and enhanced stress tolerance at reproductive stage. Signals generated were stage specific. Double mutant had no significant advantage over WT and single mutant.	Greco et al., (2012)

3.	7 days old *A. thaliana* seedlings	Seedlings incubated for different length of light and dark period before HS (45°C for 45min)	Temperature responses and light signalling are closely associated. Stronger induction of *APX2* in plants exposed to light prior to heat stress as compared to those in dark because of HsfA1 induction by light. Phytochrome B mediated light priming is important for development of thermotolerance and ROS detoxification. Light enhances the ROS scavenging capacity of *APX2*.	Han and Park, (2019)
4.	7-day-old seedlings of transgenic lines (35S:CtHsfA2b -a, b and c) and WT were used.	Plants were treated at 45°C for 4h for and 7d recovery.	CtHsf2b is HS-inducible. Its upregulation results in induction of subset of HS genes. Ectopic expression in *A.thaliana* improved the heat tolerance by upregulating *APX2* for ROS scavenging and *HSP101* for protein homeostasis. Transgenic plants have low level of H_2O_2 accumulation than WT indicating enhanced antioxidant capacity.	Wang et al., (2016)

APX2 is also controlled by the reduced/oxidized (redox) state of the plastoquinone (PQ) pool in *Arabidopsis*. The reduced PQ pool that is generated by either excess light (EL) plays a significant role in the induction of *APX2* expression. EL-driven *APX2* expression requires H_2O_2 and functional photosynthetic electron transport. On the basis of these observations, it is possible that *APX2* expression is regulated by both light and high temperature in a synergistic manner (Jung et al., 2013).

HsfA2 seems to play a principal role in the expression of ROS scavenging enzyme and thus establish an important link between heat shock and oxidative stress responses. The specificity of HsfA2 is due to its rapid transcriptional induction under heat stress, its accumulation as a protein and its regulatory functions by an amplified transcriptional induction.

Materials and Methods

MATERIALS AND METHODS

2.1. EXPERIMENTAL MATERIAL

Experimental material and chemicals used in the present work were obtained from diverse sources as shown in Table 2.1

Table 2.1: Sources of various experimental materials and chemicals used.

Material	Source
Arabidopsis seeds (Col-V, *GF 30-7*, *GF 10-5 mutant*)	Professor A. Grover's laboratory stocks, DPMB, UDSC
Oligo-nucleotides	Sigma-Aldrich, USA
Other fine chemicals and enzymes	Diverse sources including Qiagen (Germany), Sigma-Aldrich (Germany), Roche (US), Merck (Germany), Invitrogen (US) and Fermentas (US).

2.2. GROWTH CONDITIONS

2.2.1. Seed sterilization and plating

Arabidopsis seeds were first surface sterilised with 70% ethanol for 30s followed by treatment with 2% (v/v) sodium hypochlorite and a drop of Tween-20 for 5min. The seeds were then washed with autoclaved water 4-5 times until no frothing was observed and finally suspended in 0.1% (w/v) agar. The seeds were plated on MS medium (½MS, 1% sucrose and 0.8% agar, pH 5.8-5.85). The plates were placed in cold room (4°C) for 48h for stratification and then transferred into the culture room maintained at 22±1°C with 16h light/8h dark conditions and light intensity of 100-125µmol$m^{-2}s^{-1}$.

2.3. HIGH TEMPERATURE STRESS

2.3.1. For seedling expression analysis

Five-day old intact seedlings obtained were stressed on 0.5X MS media plates at 42°C for 1h in water bath after sealing the plates using petri-seal. After the treatment, the seedlings were harvested at different time points and flash chilled in liquid nitrogen and were stored in -80°C.

2.3.2. Phenotypic assays

Five-day old intact seedlings obtained were stressed on 0.5X MS media plates at 42°C for 1h in water bath after sealing the plates using petri-seal. The plates were then transferred to the growth room for recovery under control conditions.

2.4. ISOLATION OF TOTAL RNA AND QUALITY ASSESSMENT

Total RNA was isolated using TRI Reagent (Sigma-Aldrich) as per manufacturer's instructions as described below.

1. 50-100mg tissue was ground in liquid nitrogen. 1ml TRI reagent was added to the pestle and the samples were allowed to thaw at RT.
2. The samples were then transferred to labelled MCTs. 200µl chloroform was added to each sample and vortexed for 15s. The contents were allowed to stand for 5min at RT. Resulting mixture was centrifuged at 12,000×g for 15min at 2-8°C.
3. The upper phase obtained after centrifugation was carefully transferred to a fresh tube avoiding any mixing. Organic phase was stored to be used later for protein isolation.
4. 0.5ml chilled propan-2-ol was added to each MCT. Samples were allowed to stand for 5-10min at RT and centrifuged at 12,000×g for 10min at 2-8°C to precipitate RNA. The RNA appeared as a pellet on the bottom of the tube.
5. Supernatant was removed and RNA pellet was washed by adding 500µl of 75% ethanol. The pellet was dislodged and the MCTs were centrifuged at 12,000×g for 5min at 2-8°C.
6. RNA pellet was air dried for 5-10min. It was dissolved in 20µl DEPC treated double distilled water.
7. The RNA was quantified and used for further analysis or stored at -80°C.
8. RNA quantification was done using nanodrop. The O.D. was recorded at 260nm and 280nm. $A_{260/280}$ of all RNA samples were analysed.
9. Qualitative analysis was performed by formaldehyde- agarose gel electrophoresis in 1X MOPS buffer [200mM Morpholino-propanesulphonic acid, 50mM Sodium acetate and 10mM EDTA (pH 7.0)]

2.5. cDNA SYNTHESIS

First strand cDNA was synthesized by reverse transcription using 1µg of total RNA isolated from the *Arabidopsis* seedlings in 10µl of reaction volume using Verso cDNA synthesis Kit (Thermo Scientific, USA) as per manufacturer's instructions. Actin specific primers were used to amplify and check quality of cDNA synthesized.

2.6. POLYMERASE CHAIN REACTION

PCR was performed using i-Taq DNA polymerase enzyme (NEB, USA) and gene specific primers in Verity thermo-cycler (Applied Biosystem, USA) for gene amplification. Sequences of the primers used for the amplification in this study are shown in Table A1 (shown in Appendix). PCR products were analysed by agarose gel electrophoresis. PCR mix was prepared as shown in Table A6.

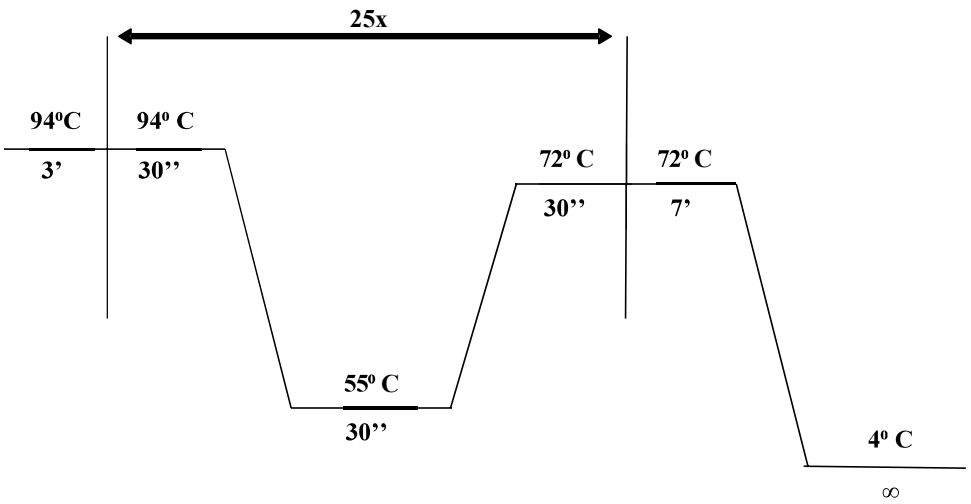


Fig. 2.1. PCR cycle for DNA amplification.

2.7 Q-PCR analysis

cDNA samples were used for Q-PCR analysis with 200nM of each primer mixed with Brilliant III Ultra-Fast SYBR Green (Agilent Technologies) Q-PCR master mix as per manufacturer's instructions. The reaction was carried out in 96-well optical reaction plate (Applied Biosystem, USA), using Mx3000P Q-PCR System (USA and Canada). To normalize the variance among samples, actin was used as an endogenous control. Three

technical replicates were used for each Q-PCR analysis. For Q-PCR, primers used are shown in Table A2.

2.8. GEL ELECTROPHORESIS OF NUCLEIC ACIDS

2.8.1. DNA

DNA (PCR product) samples were resolved by electrophoresis on 1-1.5% agarose gels made in 1X TAE buffer (Tris-acetate-EDTA). Ethidium bromide was added at a final concentration of 0.5µg/ml. DNA was mixed with 6X sample buffer [50% glycerol, 30% (v/v) xylene cyanol and 0.05% (v/v) bromophenol blue] prior to loading.

2.8.2. RNA

RNA samples were initially heat denatured at 70°C for 10min in 4 volume of sample buffer [25% (v/v) formaldehyde, 60% (v/v) formamide, 12.5% (v/v) 10X morpholino-proprane sulphonic acid (MOPS) with 250µgml^{-1} ethidium bromide (EtBr)]. The samples were snap chilled in ice for 2min. The samples were resolved by electrophoresis under denaturing conditions. Denaturing gel was casted by dissolving 1% agarose in 1X MOPS buffer [200mM MOPS, 50mM sodium acetate and 10mM EDTA (pH7.0)]. The gel solution was boiled to melt agarose, cooled to 55 to 60°C, followed by addition of formaldehyde (as a denaturant) to a final concentration of 5% (v/v). The gel apparatus was washed with 10% SDS and rinsed with sterile double distilled water prior to electrophoresis of RNA to make sure RNAase free environment.

2.9. PROTEIN EXTRACTION

1. Soluble leaf protein was extracted by pulverising the frozen tissue in liquid nitrogen using pestle and mortar and resuspending in extraction buffer [100mM Sodium phosphate buffer (pH 7), 5mM Ascorbate, 1mM EDTA]
2. The suspension was centrifuged at 15000×g for 15min at 4°C.
3. The supernatant was collected and stored at -80°C.

2.10. PROTEIN QUANTIFICATION

2.10.1. BSA standard curve

1. Standard curve for quantification of proteins was made using different concentrations of BSA (2µg, 4µg, 6µg, 8µg, 10µg, 12µg and 14µg).
2. The absorbance of each protein sample was measured at 595nm wavelength and standard curve was prepared.

Absorbance was plotted against BSA concentrations and the linear equation thus obtained was used for further protein quantification

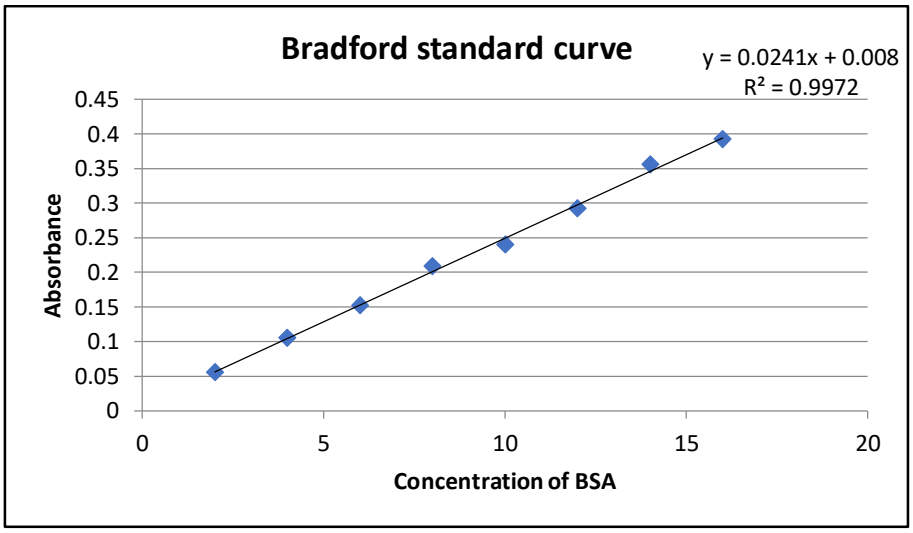


Fig 2.2. Graph showing BSA standard curve

2.10.2. Bradford Assay

The samples were quantified as follows:

1. 5µl of protein sample was mixed in 95µl of 0.15M NaCl. 1ml Bradford reagent [10% (w/v) CBB G-250, 5% (v/v) ethanol, 10% (v/v) o-phosphoric acid, RO] was added in dark and mixed thoroughly followed by 10min incubation in dark.

2. Absorbance was measured in spectrophotometer (Eppendorf) at 595nm.

2.11. NATIVE PAGE FOR APX

1. Gel casting apparatus was setup and checked for leakage. 10% resolving gel was prepared according to Table A3 and was poured in between gel plate sandwich.

2. The surface of the gel was covered by propan-2-ol to allow polymerization, and the apparatus was placed at 37°C for 10min.

3. Once the resolving gel was set, 3.9% stacking gel was prepared according to Table A4 and was added onto it after removing traces of isopropanol.

4. The wells were immediately flushed with electrode buffer which was prepared according to Table A5.

5. The gel was pre-run for 30min to allow the ascorbate present in the electrode buffer to enter the gel, prior to the application of the samples.

6. Equal concentration of protein samples were mixed with a pinch of bromophenol blue and glycerol and then loaded onto the gel. The gel was allowed to run at 30mA at 4°C until the ladder is properly resolved.

2.11.1. DETECTION OF APX ISOENZYMES IN GEL

The following steps were performed at room temperature (20-25°C)

1. Subsequent to the electrophoretic preparation, the gel was equilibrated with 50mM sodium phosphate buffer (pH 7) and 2mM ascorbate for a total of 30min; the equilibration buffer was changed every 10min.
2. The gel was then incubated with 50mM sodium phosphate buffer (pH 7) containing 4mM ascorbate and 2mM H_2O_2 for 20min. H_2O_2 was added to this solution immediately prior to incubation with the gel.
3. The gel was subsequently washed with sodium phosphate buffer (50mM pH 7) for 1min and submerged in a solution of 50mM sodium phosphate buffer (pH 7.8), 28mM TEMED and 2.45mM NBT with gentle agitation.
4. The reaction was allowed to continue for 3-5min and stopped by a brief wash with deionised water.

2.12. *IN-SILICO* ANALYSIS

2.12.1. Generation of Phylogenetic tree

1. The protein sequence of *Arabidopsis thaliana* APX2 was obtained from Phytozome database.
2. BLASTp was performed to find the most similar protein sequences in the organisms specified.
3. Protein sequence alignment was performed using the global alignment tool (with free end gaps) of the Geneious software.
4. Phylogenetic tree was generated using the neighbor joining method. Jukes-Cantor genetic distance model was used for the same.

2.12.2. Detection of HSEs and other elements in the promoter of *APX2*

1. The genomic sequence 1500bp upstream of *APX2* gene was obtained from Phytozome. HSEs in the *APX2* promoter was detected manually.
2. The other elements in the promoter were detected using PlantCare database. (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)

Results

RESULTS

3.1. Phenotypic analysis of Col-V and *GF* line under HS

Five-day old *A. thaliana* Col-V (Col-0 transformed with empty vector) and *GF* line seedlings were subjected to HS at 42°C for 1h. Plants were allowed to recover for 10d at 22°C. *GF* plants (HSP101 over-expressing transgenic line; Professor Anil Grover's laboratory) recovered from HS and survived to a much higher extent than Col-V seedlings most of which die.

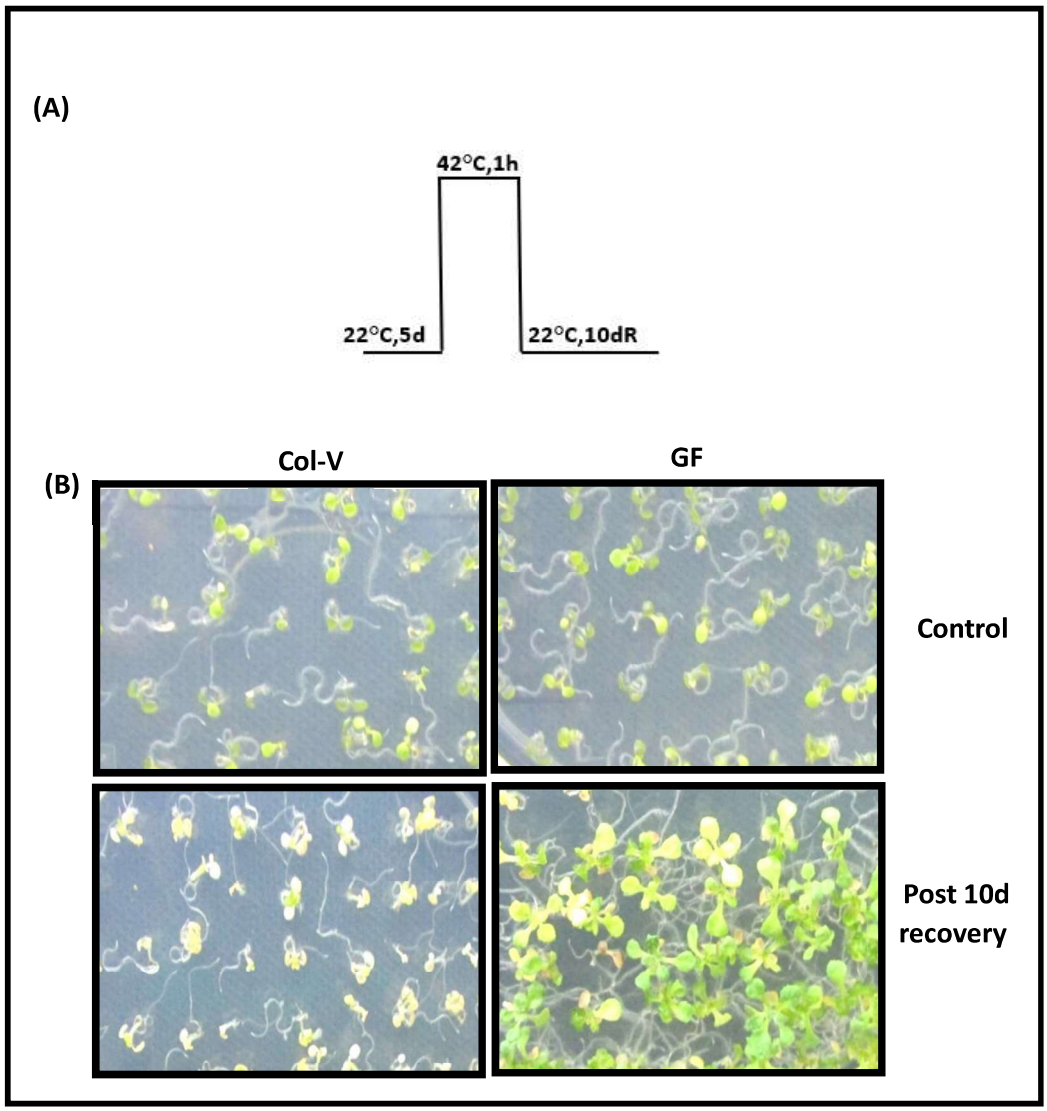


Fig 3.1. Heat stress response of Col-V and *GF* plants. **(A)** Stress regime followed in the experiment. **(B)** Basal thermotolerance assay of Col-V and *GF* seedlings.

3.2. Expression of *AtHSP101* in Col-V, *GF 30-7* and *GF 10-5* plants in response to HS

The expression of *AtHSP101* was analysed in five-day old *A. thaliana* Col-V, *GF 30-7* and *GF 10-5* (HSP101 over-expressing transgenic line; Professor Anil Grover's laboratory) seedlings in response to HS at 42°C for 1h followed by 2d recovery at 22°C. Samples were harvested at three different time points namely control (C), heat stress (HS) and post 2d recovery (2dR) respectively. RNA was extracted using TRIzol method followed by cDNA preparation. Atactin_F and Atactin_R primers were used to equalize the cDNA concentration. sqPCR was carried out to analyse *AtHSP101* expression in Col-V, *GF 30-7* and *GF 10-5* using AtHSP_101F and AtHSP_101R primers. Col-V, *GF 30-7* and *GF 10-5* seedlings showed up-regulation of *AtHSP101* transcript upon encountering HS.

3.3. Expression of *APX2* in Col-V, *GF 30-7* and *GF 10-5* plants in response to HS

Five-day old Col-V, *GF 30-7* and *GF 10-5* seedlings were subjected to HS at 42°C for 1h followed by 2dR at 22°C. Samples were harvested at three different time points namely control (C), heat stress (HS) and post 2d recovery (2dR) respectively. RNA was extracted using TRIzol method followed by cDNA preparation. Atactin_F and Atactin_R primers were used to equalize the cDNA concentration. sqPCR was carried out to analyse *AtAPX2* expression in Col-V, *GF 30-7* and *GF 10-5* using AtAPX2_F and AtAPX2_R primers. Col-V, *GF 30-7* and *GF 10-5* showed up-regulation of *APX2* transcript at 2dR and no transcript was detected under control and heat stress conditions. The expression was maximum in Col-V followed by *GF 10-5* and *GF 30-7* at post 2dR timepoint.

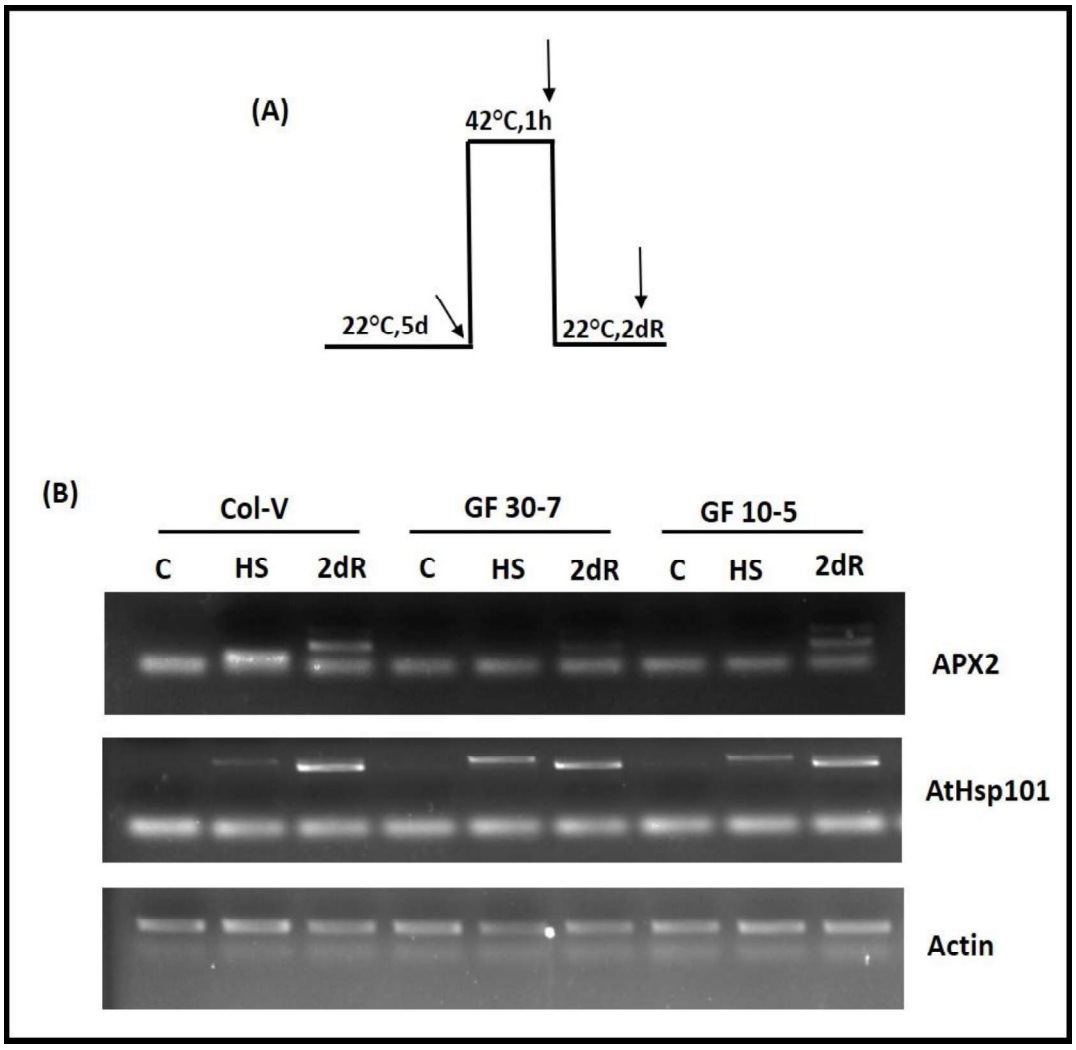


Fig 3.2. (A) Stress regime followed in the analysis. Arrows depict time of sample harvest for transcript analysis. **(B)** sqPCR showing *AtHSP101* and *APX2* expression level.

3.4. Expression of *AtHSP101* and *AtAPX2* in Col-V, *GF 30-7* and *GF 10-5* seedlings in response to HS

Five-day old Col-V, *GF 30-7* and *GF 10-5* seedlings were subjected to HS at 42°C for 1h followed by 2 day recovery at 22°C. Samples were harvested at three different time points namely control (C), heat stress (HS) and post 2d recovery (2dR) respectively. RNA was extracted using TRIzol method followed by cDNA preparation. The transcript level of *AtHSP101* and *APX2* were analysed by Q-PCR. ΔΔCt method was used to calculate fold changes.

Q-PCR results showed transient *AtHSP101* expression at HS and Rec (2dR) timepoint in all the three lines. (Fig.3.3B,C,D). The fold change of the expression of *HSP101* was highest for *GF 30-7* at HS and at Rec (2dR), the fold change was comparable for both overexpressing lines (Fig 3.3E). The expression increased significantly at 2dR in all the three lines.

Q-PCR analysis of *APX2* expression revealed that the expression is considerably low at C and HS timepoints and a significant fold change is observed at 2dR timepoint in all the lines. The maximum fold change was observed in *GF 30-7* followed by *GF 10-5*.

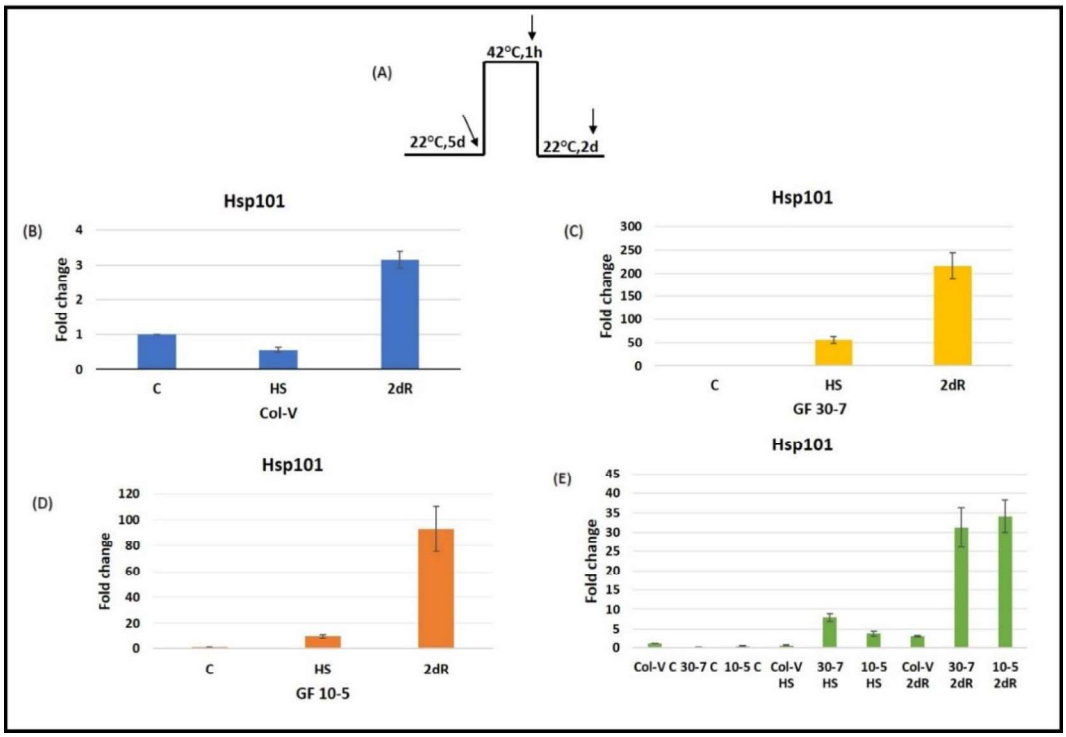


Fig. 3.3. (A) Stress regime followed in the analysis. Arrows depict time of sample harvest for transcript and analysis. Q-PCR showing *AtHSP101* expression level at various time points in Col-V **(B),** *GF 30-7* **(C)** and *GF 10-5* **(D)** seedlings respectively. **(E)** Comparative analysis of Q-PCR showing *AtHSP101* expression level at various time points in different lines.

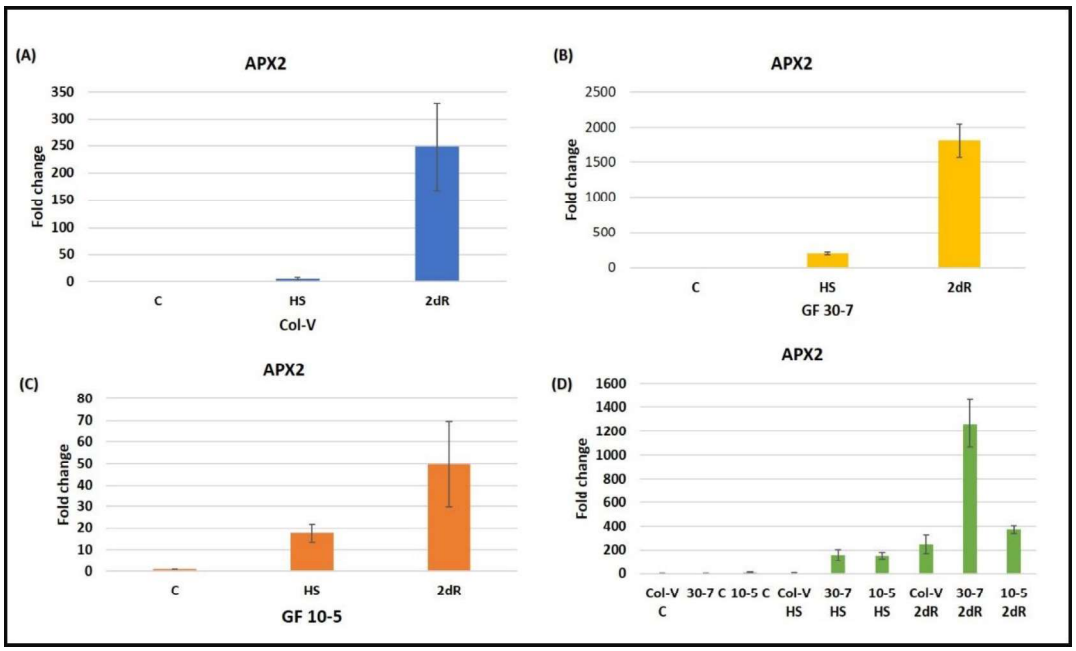


Fig 3.4. Q-PCR showing *AtAPX2* expression level at various time points in Col-V **(A),** *GF 30-7* **(B)** and *GF 10-5* **(C)** seedlings respectively. **(D)** Comparative analysis of Q-PCR showing *AtHSP101* expression level at various time points in different lines.

3.5. Native gel analysis for detection of APX isoenzymes

Five-day old Col-V, *GF 30-7* and *GF 10-5* seedlings were subjected to stress regime as shown in Fig. 3.5A. Soluble protein was isolated post HS and recovery time points. 40µg protein was loaded and Native-PAGE gel was run. After electrophoretic preparation, the gel was equilibrated with 50mM sodium phosphate buffer (pH 7) and 2mM ascorbate for a total of 30min. The gel was next incubated in 50mM sodium phosphate buffer (pH 7) containing 4mM ascorbate and 2mM H_20_2 for 20min. The gel was subsequently washed with 50mM sodium phosphate buffer (pH 7) for 1min.The gel was then submerged in a solution of 50mM sodium phosphate buffer (pH 7.8), 28mM TEMED and 2.45mM NBT with gentle agitation. The APX activity is observed as achromatic band on purple background. The isoenzymes that detected were APX1 and APX2. The isoenzyme activity of APX declined at elevated temperatures.

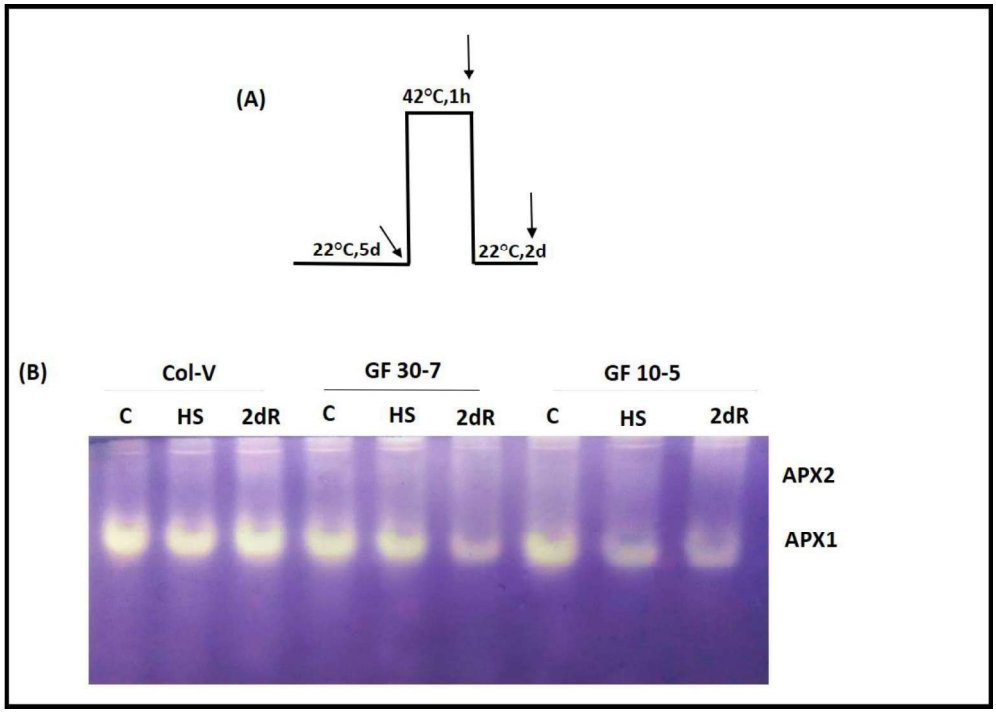


Fig 3.5. (A) Stress regime followed for the analysis. **(B)** Native-PAGE gel showing APX isoenzyme profiling.

3.6. Construction of phylogenetic tree to determine relatedness of *A. thaliana* APX2 to other plant species

We identified APX2 protein in over 44 plant species by performing a BLAST analysis using the *Arabidopsis thaliana* APX2 homologues on the plant sequence database ‘‘Phytozome’’. We recognized that the APX2 of various dicot species are clubbed together. Likewise, it was noted that APX2 of various monocot species come together in the phylogenetic analysis, indicating that APX2 origin in these species are from distinct from each other.

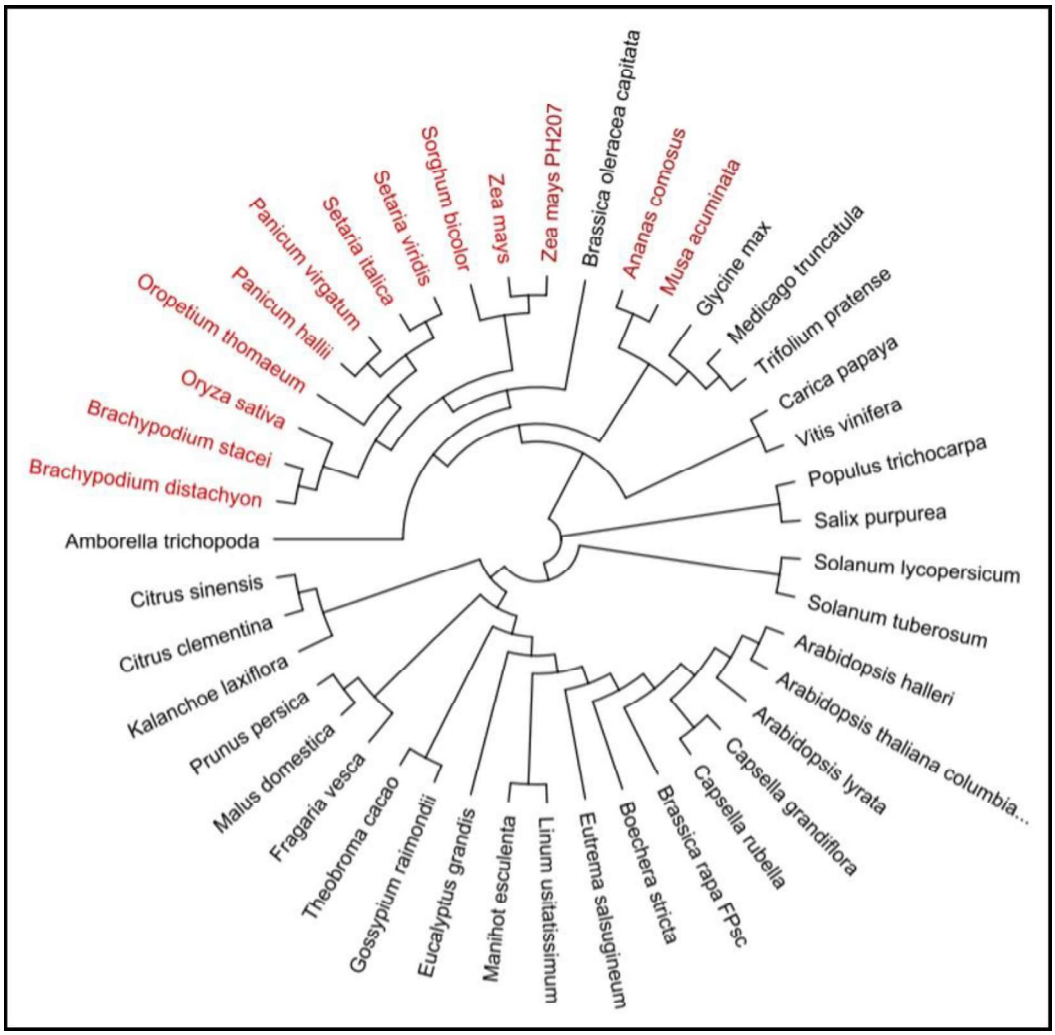


Fig 3.6. The phylogenetic relatedness of APX2 from various plant species. The species indicated in black belong to clade eudicots and those in red belong to clade monocots.

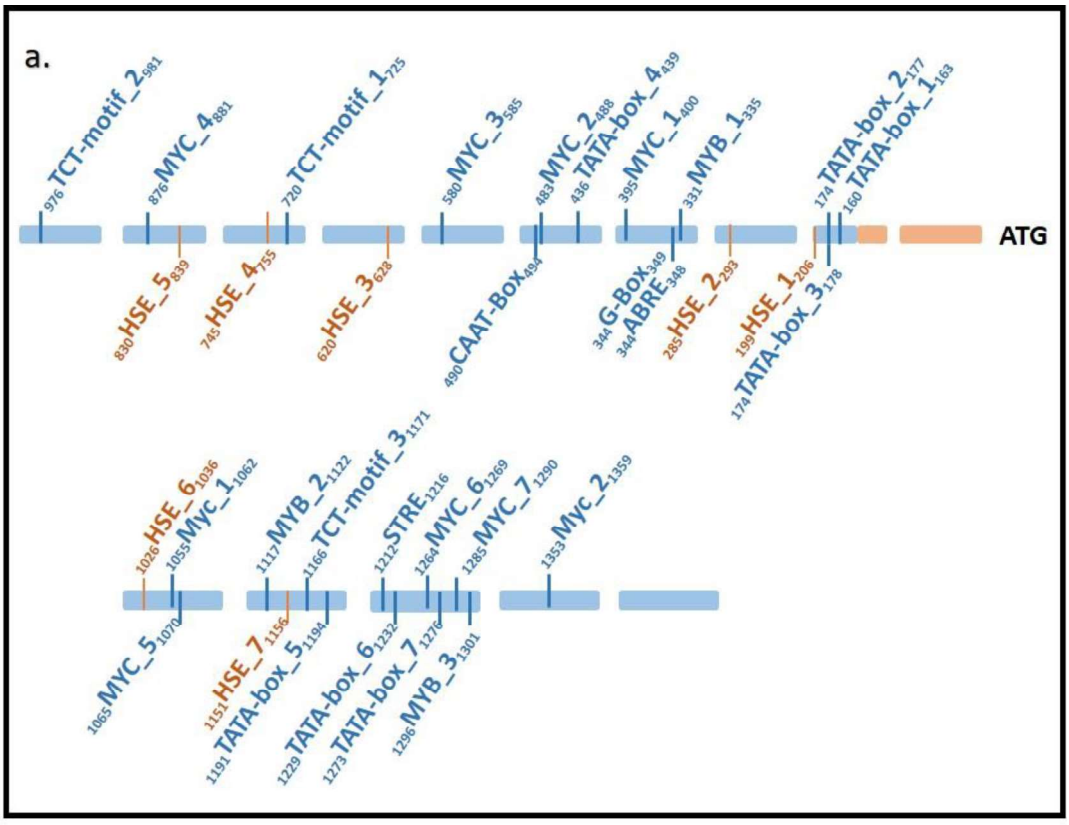
3.7. Detection of HSEs and other elements in the promoter of *APX2*

On manual detection of 1500bp sequence upstream of *APX2* gene of *Arabidopsis*, 7 heat shock elements (HSEs) were found (Fig. 3.7). The sequence of these HSEs were **GAA**AATTG**TTC**, **TTC**T**GAA**, **GAA**CACT**TTC**, **GAA**CAAGT**TTC**, **GAA**GCT**TTC**, **TTC**TG**GAA** and **GAA**GC**TTC.** The HSFs bind to these HSEs and enhance transcriptional activity of *APX2* under stress conditions.

Other elements detected in the promoter of *APX2* using PlantCare database include - TATA-Box (core promoter element around -30 of transcription start), ABRE (involved in the response to abscisic acid), G-box (cis-acting regulatory element involved in light responsiveness), MYB, Myc, STRE, TCT-motif (part of a light responsive element), MYC and CAAT-box (element present in promoter and enhancer regions).

The elements in the promoter of *AtHSP101* include HSEs, ABRE, G-box, STRE, MYC, CAAT-box, AE-box and GATA-motif.

Comparison was made between *Arabidopsis APX2* and *HSP101* gene promoter (1133bp upstream of gene). The results are presented in Table 3.1.



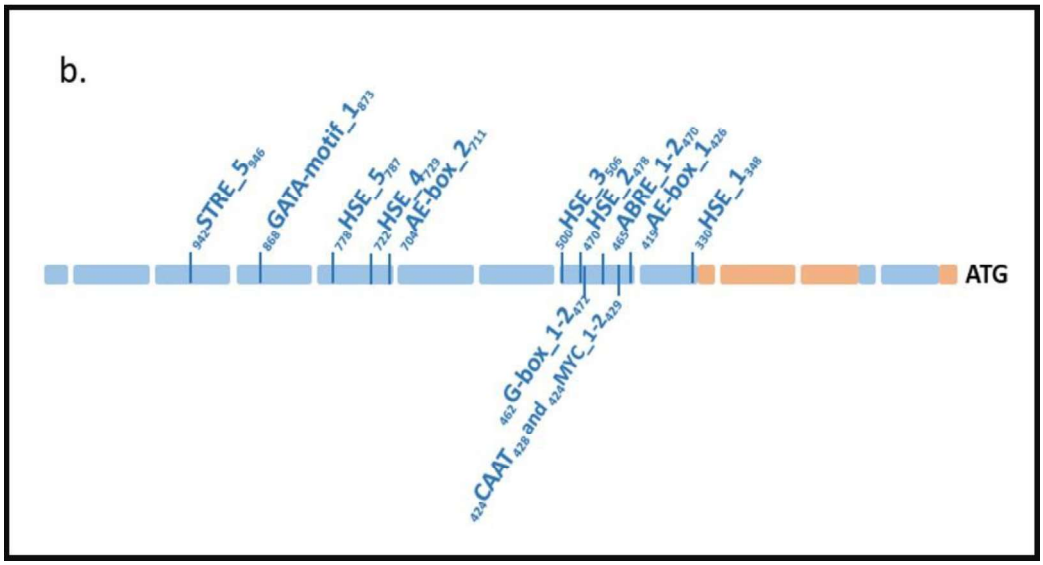


Fig 3.7. a. Linear representation of the native promoter region of *Arabidopsis* 1500 upstream to *APX2* gene. All elements are marked taking A of ATG as +1. HSEs were found manually and other elements were found using PlantCare database and only elements known present in *Arabidopsis* were used to show in each promoter sequence. Each block represent 100bp length where blue blocks represent upstream promoter region and orange blocks represent 5'UTR sequence. Sequences for the *Arabidopsis* HSEs **(GAAAATTGTTC, TTCTGAA, GAACACTTTC, GAACAAGTTTC, GAAGCTTTC, TTCTGGAA, GAAGCTTC)**.

b. *Arabidopsis* 1133bp upstream to *HSP101* gene.

Name of the element	*AtHSP101* promoter	*AtAPX2* promoter
COMMON ELEMENTS		
Heat Shock elements (HSEs)	1. 330-348bp 2. 470-478bp 3. 500-506bp 4. 722-729bp 5. 778-787bp	1. 199-206bp 2. 285-295bp 3. 620-628bp 4. 745-755bp 5. 830-839bp 6. 1026-1156bp
ABRE	465-470bp	344-348bp
G-Box	462-472bp	344-348bp
MYC	424-429bp	1. 395-400bp 2. 483-488bp 3. 580-585bp 4. 1065-1070bp
CAAT-Box	424-428bp	490-494bp
DIFFERENT ELEMENTS		
GATA-motif	868-873bp	
AE-box	1. 419-873bp 2. 704-711bp	
STRE	942-946bp	
MYB		1. 330-335bp 2. 1117-1122bp
TCT-motif		1. 720-725bp 2. 976-981bp
Myc		1055-1062bp

Table 3.1. Comparison of the elements present in the promoter of *AtHSP101* and *AtAPX2* (1133bp upstream of the gene in each case) on the basis of results obtained from PlantCare database.

Discussion

DISCUSSION

ROS such as H_2O_2, $OH^{\bullet}$, $O_2^{\bullet-}$ are molecular species that have undergone electron addition(s) and thus are reduced form of oxygen. They are unavoidable byproducts of aerobic metabolism. They are generated when plant is exposed to environmental stresses such as heat stress, cold, drought etc. At low concentration, they act as key signaling molecules for activation of stress response and defense pathway whereas at high concentration, they cause non-controlled oxidation of a variety of cellular structures including membrane lipids, DNA and proteins resulting in their damage and disruption of metabolism. In order to maintain cellular redox balance, plants possess a range of enzymatic (CAT, SOD, APX, GR) and non-enzymatic (Ascorbic acid, proline, Reduced glutathione) antioxidant mechanisms.

Ascorbate peroxidase catalyses the H_2O_2 dependent oxidation of a large number of molecules. In *A. thaliana*, 9 APX isozymes- APX1, APX2, APX3, APX4, APX5, APX6, mAPX, sAPX and tAPX are present.

The objective of the present study was to analyze the expression of *APX2* in *A. thaliana* Col-V and HSP101 overexpression lines (*GF 30-7* and *GF 10-5*) in response to heat stress and in recovery period post heat stress. Col-V, *GF 30-7* and *GF 10-5* showed induction of *APX2* transcript at 2dR timepoint. Q-PCR analysis revealed maximum fold change in the transcript of *APX2* in *GF 30-7* at 2dR timepoint.

Expression of *AtHSP101* under HS was also studied in the three lines. The transcript was observed at HS and 2dR timepoint. The transcript level was maximum in *GF 30-7* followed by *GF 10-5* and then Col-V as shown by Q-PCR analysis.

Phenotypic analyses revealed that the survival rate of overexpression lines was more as compared to the WT under stress conditions.

The APX enzyme activity was detected using Native-PAGE. This assay is based on the ability of ascorbate to reduce nitroblue tetrazolium (NBT) in the presence of N,N,N',N'-tetramethylethylenediamine (TEMED) to formazan, a poorly soluble, highly coloured derivative. In the presence of H_2O_2, APX is capable of preventing the accumulation of formazan due to the rapid H_2O_2 dependent oxidation of ascorbate, and thus APX activity is observed as an achromatic band. The enzyme activity was comparable in all the samples. Previous studies provide enough evidence that APX enzyme activity reach a maximum at 37°C and then declines further.

A phylogenetic tree was generated to determine the relatedness of *A. thaliana* APX2 to other plant species. From the analysis of 44 species, APX2 of dicots was clubbed together and that of monocots were grouped together suggesting that the origin of the protein in these species is separate from each other. Monocot APX2 was closer to *Brachypodium distachyon* showing a possible relatedness of their ancestry. Three *Arabidopsis* species taken in this analysis were clubbed together. While *Brassica rapa* was closer to *Arabidopsis* in this analysis, it was surprising to see that B. *oleracea* was quite far off in its positioning.

Analysis of 1500bp promoter upstream of *AtAPX2* gene was done. Various elements such as HSEs, ABRE, G-box, TCT-motif, MYB, MYC, Myc and STRE were detected which respond to different stresses. Comparison with *AtHSP101* promoter revealed the presence of several common elements in both. From the presence of multiple HSEs both in *APX2* and *HSP101*, the basis of their heat inducibility can be accounted for.

We found that both *HSP101* and *APX2* display similar trends. Both the genes are induced by HS. In both cases, high transcript amount is observed at recovery timepoint than during heat stress. This implies that both the genes are important for post stress recovery processes.

HSP101 is a molecular chaperone involved in disaggregation of the misfolded and aggregated proteins formed under HS. Previous studies show that this protein is more essential in recovery period than stress period. It is interesting to find that APX2 also display similar expression pattern.

It is therefore possible that *HSP101* and *APX2* transcripts are protected during heat stress. During stress, translational repression occurs and most cellular RNAs and proteins are sequestered into stress granules (SGs). When stress is over, SGs disassemble which allow the reactivation of mRNA translation and protein activity.

The study shows that both *GF 30-7* and *GF 10-5* display superior phenotypes, contain more amount of *APX2* than WT under stress suggesting the role of APX2 in stress protection. Likewise, both *GF 30-7* and *GF 10-5* contain more amount of *HSP101* transcript than WT under stress suggesting the role of HSP101 in stress protection. It thus comes out that these two proteins work together in protection of seedlings under heat stress. Previous studies have shown that both *APX2* and *HSP101* genes are regulated in the same regulon involving HsfA1s and HsfA2 (Schramm et al. 2006). From these facts, we conclude that there is a kinship in birthing and working of APX2 and HSP101 towards heat protection. We present a model below which highlights this viewpoint (Fig. 4.1).

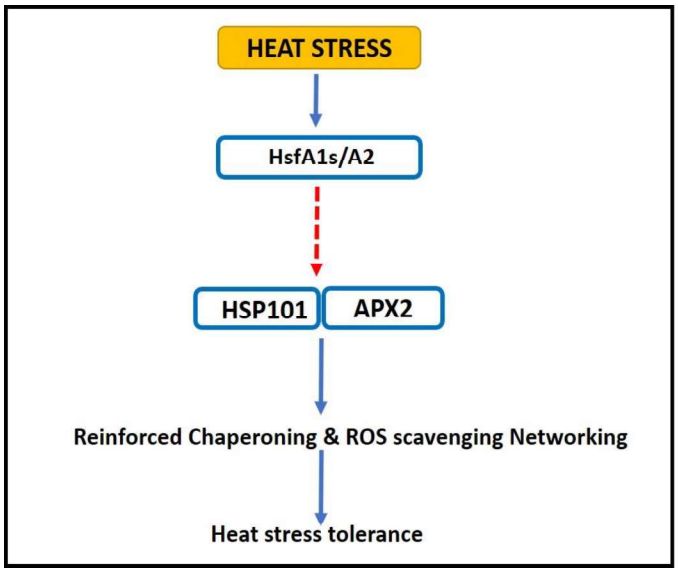


Fig 4.1: Model showing how HSP101 and APX2 work together in mitigating heat stress injuries to *Arabidopsis* plants. Past studies show that both *HSP101* and *APX2* are the target genes of HsfA2. HSP101 is important in removal of protein aggregates and hence it is a chaperone. APX2 removes ROS which are generated during stress conditions. We show in this model that both these genes come from the same regulon and work together in making cells survive under heat stress.

Finally, we submit that further replications and experimentation are needed to validate the results obtained. It will be interesting to provide further evidence in support of the kinship in the expression of *HSP101* and *APX2* under stress conditions to counter heat stress effects.

Conclusion & Future Work

CONCLUSION AND FUTURE WORK

Climate change predictions signify a gradual increase in global mean temperature and an increase in the frequency of heatwaves, which will have an adverse impact on crop productivity. Several mechanisms are adopted according to the intensity and the duration of heat stress, as well as the capacity of the different species or genotypes to overcome stress, an alteration in redox homeostasis is one of those events. The mechanisms that act to enhance redox systems allow crops to better survive the heat stress. Based on work carried out in this dissertation, we propose that the interplay of HSP101 and APX2 confers protection to plants under stress conditions and is important for post stress recovery processes. We find that the HSP101 overexpression lines displayed superior stress tolerance phenotype, contained high levels of APX2 in contrast to WT, indicating role of both proteins in stress tolerance. We propose that HSP101 and APX2 play important roles in stress tolerance, HSP101 through its disaggregase activity and APX2 through its ROS scavenging capacity. It can be inferenced that more *APX2* is transcribed when more *HSP101* is expressed during stress and recovery. The transcriptional induction of the stress responsive genes represents the first line of defense against imbalances in cellular homeostasis. Identification of regulators involved in abiotic stresses is an essential first step toward understanding the molecular mechanisms. The production of transgenic plants with enhanced resistance to a stress can be achieved on this basis. The work presented in this dissertation can be extended to study the response of these lines when exposed to different stress combinations and the expression levels of various other ROS scavenging enzymes as well as chaperone proteins.

Literature cited

LITERATURE CITED

1. Apel, K. and Hirt, H. (2004). REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction. Annu. Rev. Plant Biol. 55: 373–399.
2. Asada, K. (1992). Ascorbate peroxidase – a hydrogen peroxide-scavenging enzyme in plants. Physiol. Plant. 85: 235–241.
3. Bhattacharjee, S. (2013). Heat and chilling induced disruption of redox homeostasis and its regulation by hydrogen peroxide in germinating rice seeds (Oryza sativa L., Cultivar Ratna). Physiol. Mol. Biol. Plants 19: 199–207.
4. Bowler, C. and Fluhr, R. (2000). The role of calcium and activated oxygens as signals for controlling cross-tolerance. Trends Plant Sci. 5: 241–246.
5. Busch, W., Wunderlich, M., and Schöffl, F. (2005). Identification of novel heat shock factor-dependent genes and biochemical pathways in Arabidopsis thaliana. Plant J. 41: 1–14.
6. Caverzan, A., Passaia, G., Rosa, S.B., Ribeiro, C.W., Lazzarotto, F., and Margis-Pinheiro, M. (2012). Plant responses to stresses: Role of ascorbate peroxidase in the antioxidant protection. Genet. Mol. Biol. 35: 1011–1019.
7. Charng, Y.Y., Liu, H.C., Liu, N.Y., Chi, W.T., Wang, C.N., Chang, S.H., and Wang, T.T. (2007). A heat-inducible transcription factor, HsfA2, is required for extension of acquired thermotolerance in Arabidopsis. Plant Physiol. 143: 251–262.
8. Chen, C., Letnik, I., Hacham, Y., Dobrev, P., Ben-Daniel, B.H., Vanková, R., Amir, R., and Miller, G. (2014). ASCORBATE PEROXIDASE6 protects arabidopsis desiccating and germinating seeds from stress and mediates cross talk between reactive oxygen species, Abscisic acid, And auxin. Plant Physiol.166: 370-383.
9. Choudhury, F.K., Rivero, R.M., Blumwald, E., and Mittler, R. (2017). Reactive oxygen species, abiotic stress and stress combination. Plant J. 90: 856–867.
10. Czarnecka-Verner, E., Yuan, C.X., Scharf, K.D., Englich, G., and Gurley, W.B. (2000). Plants contain a novel multi-member class of heat shock factors without transcriptional activator potential. Plant Mol. Biol. 43: 459–471.
11. Das, K. and Roychoudhury, A. (2014). Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. Front. Environ. Sci. 2, 53.
12. Davletova, S., Rizhsky, L., Liang, H., Shengqiang, Z., Oliver, D.J., Coutu, J., Shulaev, V., Schlauch, K., and Mittler, R. (2005). Cytosolic ascorbate peroxidase 1 is a central component of the reactive oxygen gene network of Arabidopsis. Plant Cell 17: 268–281.
13. Driedonks, N., Xu, J., Peters, J.L., Park, S., and Rieu, I. (2015). Multi-level interactions between heat shock factors, heat shock proteins, and the redox system regulate acclimation to heat. Front. Plant Sci. 6: 1–9.
14. Evans, M.D., Dizdaroglu, M., and Cooke, M.S. (2004). Oxidative DNA damage and disease: Induction, repair and significance. Mutat. Res. - Rev. Mutat. Res. 567: 1–61.
15. Foyer, C.H. and Halliwell, B. (1976). The presence of glutathione and glutathione reductase in chloroplasts: A proposed role in ascorbic acid metabolism. Planta 133: 21–25.
16. Foyer, C.H. and Noctor, G. (2005). Redox homeostasis and antioxidant signaling: A metabolic interface

between stress perception and physiological responses. Plant Cell 17: 1866–1875.

17. Foyer, C.H. and Noctor, G. (2009). Redox regulation in photosynthetic organisms: Signaling, acclimation, and practical implications. Antioxidants Redox Signal. 11: 861–905.
18. Foyer, C.H. and Noctor, G. (2013). Redox signaling in plants. Antioxidants Redox Signal. 18: 2087–2090.
19. Gill, S.S. and Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol. Biochem. 48: 909–930.
20. Gilroy, S., Białasek, M., Suzuki, N., Górecka, M., Devireddy, A.R., Karpiński, S., and Mittler, R. (2016). ROS, calcium, and electric signals: Key mediators of rapid systemic signaling in plants. Plant Physiol. 171: 1606–1615.
21. Greco, M., Chiappetta, A., Bruno, L., Bitonti, M.B., Miller, G., Sejima, H., Harper, J., and Mittler, R. (2012). Enhanced seed production under prolonged heat stress In Posidonia oceanica cadmium induces changes in DNA conditions in Arabidopsis plants deficient in methylation and patterning cytosolic ascorbate peroxidase 2. 64: 253–263.
22. Groden, D. and Beck, E. (1979). H2O2 destruction by ascorbate-dependent systems from chloroplasts. BBA - Bioenerg. 546: 426–435.
23. Grover, A., Mittal, D., Negi, M., and Lavania, D. (2013). Generating high temperature tolerant transgenic plants: Achievements and challenges. Plant Sci. 205–206: 38–47.
24. Halliwell, B. and Gutteridge, J.M.C. (2015). Free Radicals in Biology and Medicine.
25. Han, S. and Park, Y. (2019). Light Primes the Thermally Induced Detoxification of Reactive Oxygen Species During Development of Thermotolerance in Arabidopsis. 1: 230–241.
26. Hasanuzzaman, M., Nahar, K., Alam, M.M., Roychowdhury, R., and Fujita, M. (2013). Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. Int. J. Mol. Sci. 14: 9643–9684.
27. Jung, H.S., Crisp, P.A., Estavillo, G.M., Cole, B., Hong, F., Mockler, T.C., Pogson, B.J., and Chory, J. (2013). Subset of heat-shock transcription factors required for the early response of Arabidopsis to excess light. Proc. Natl. Acad. Sci. U. S. A. 110: 14474–14479.
28. Khedia, J., Agarwal, P., and Agarwal, P.K. (2019). Deciphering hydrogen peroxide-induced signalling towards stress tolerance in plants. 3 Biotech 9.
29. Lin, K.H., Huang, H.C., and Lin, C.Y. (2010). Cloning, expression and physiological analysis of broccoli catalase gene and Chinese cabbage ascorbate peroxidase gene under heat stress. Plant Cell Rep. 29: 575–593.
30. Liu, Y., Zhang, C., Chen, J., Guo, L., Li, X., Li, W., Yu, Z., Deng, J., Zhang, P., Zhang, K., and Zhang, L. (2013). Arabidopsis heat shock factor HsfA1a directly senses heat stress, pH changes, and hydrogen peroxide via the engagement of redox state. Plant Physiol. Biochem. 64: 92–98.
31. Mandelman, D., Jamal, J., and Poulos, T.L. (1998). Identification of two electron-transfer sites in ascorbate peroxidase using chemical modification, enzyme kinetics, and crystallography. Biochemistry. 37: 17610-17617.
32. Miller, G., Schlauch, K., Tam, R., Cortes, D., Torres, M.A., Shulaev, V., Dangl, J.L., and Mittler, R. (2009). The plant NADPH oxidase RBOHD mediates rapid systemic signaling in response to diverse stimuli. Sci. Signal. 2.
33. Miller, G., Suzuki, N., Ciftci-Yilmaz, S., and Mittler, R. (2010). Reactive oxygen species homeostasis

and signalling during drought and salinity stresses. Plant, Cell Environ. 33: 453–467.

34. Mishra, S.K., Tripp, J., Winkelhaus, S., Tschiersch, B., Theres, K., Nover, L., and Scharf, K.D. (2002). In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of thermotolerance in tomato. Genes Dev. 16: 1555–1567.
35. Mittler, R. (2002). Oxidative stress, antioxidants and stress tolerance. Trends Plant Sci. 7: 405–410.
36. Mittler, R., Finka, A., and Goloubinoff, P. (2012). How do plants feel the heat? Trends Biochem. Sci. 37: 118–125.
37. Mittler, R., Vanderauwera, S., Gollery, M., and Van Breusegem, F. (2004a). Reactive oxygen gene network of plants. Trends Plant Sci. 9: 490–498.
38. Mittler, R., Vanderauwera, S., Gollery, M., and Van Breusegem, F. (2004b). Reactive oxygen gene network of plants. Trends Plant Sci. 9: 490–498.
39. Mittler, R., Vanderauwera, S., Suzuki, N., Miller, G., Tognetti, V.B., Vandepoele, K., Gollery, M., Shulaev, V., and Van Breusegem, F. (2011). ROS signaling: The new wave? Trends Plant Sci. 16: 300–309.
40. Mittler, R. and Zilinskas, B.A. (1994). Regulation of pea cytosolic ascorbate peroxidase and other antioxidant enzymes during the progression of drought stress and following recovery from drought. Plant J. 5: 397–405.
41. Møller, I.M. (2001). P LANT M ITOCHONDRIA AND O XIDATIVE S TRESS : Electron Transport, NADPH Turnover, and Metabolism of Reactive Oxygen Species . Annu. Rev. Plant Physiol. Plant Mol. Biol. 52: 561–591.
42. Møller, I.M., Jensen, P.E., and Hansson, A. (2007). Oxidative Modifications to Cellular Components in Plants. Annu. Rev. Plant Biol. 58: 459–481.
43. Morimoto, R.I. (1998). Regulation of the heat shock transcriptional response: Cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. Genes Dev. 12: 3788–3796.
44. Nishizawa-Yokoi, A., Tainaka, H., Yoshida, E., Tamoi, M., Yabuta, Y., and Shigeoka, S. (2010). The 26S proteasome function and Hsp90 activity involved in the regulation of HsfA2 expression in response to oxidative stress. Plant Cell Physiol. 51: 486–496.
45. Panchuk, I.I., Volkov, R.A., and Schöffl, F. (2002). Heat stress- and heat shock transcription factor-dependent expression and activity of ascorbate peroxidase in Arabidopsis. Plant Physiol. 129: 838–853.
46. Pandey, S., Fartyal, D., Agarwal, A., Shukla, T., James, D., Kaul, T., Negi, Y.K., Arora, S., and Reddy, M.K. (2017). Abiotic stress tolerance in plants: Myriad roles of ascorbate peroxidase. Front. Plant Sci. 8: 1–13.
47. Patterson, W.R., Poulos, T.L., and Goodin, D.B. (1995). Identification of a Porphyrin π Cation Radical in Ascorbate Peroxidase Compound I. Biochemistry 34: 4342-4345.
48. Pnueli, L., Liang, H., Rozenberg, M., and Mittler, R. (2003). Growth suppression, altered stomatal responses, and augmented induction of heat shock proteins in cytosolic ascorbate peroxidase (Apx1)-deficient Arabidopsis plants. Plant J. 34: 187–203.
49. Polle, A. (2001). Dissecting the superoxide dismutase-ascorbate-glutathione-pathway in chloroplasts by metabolic modeling. Computer simulations as a step towards flux analysis. Plant Physiol. 126: 445–462.
50. Scharf, K.D., Berberich, T., Ebersberger, I., and Nover, L. (2012). The plant heat stress transcription factor (Hsf) family: Structure, function and evolution. Biochim. Biophys. Acta - Gene Regul. Mech. 1819: 104–119.

51. Schöffl, F., Prändl, R., and Reindl, A. (1998). Regulation of the heat-shock response. Plant Physiol. 117: 1135–1141.
52. Schramm, F., Ganguli, A., Kiehlmann, E., Englich, G., Walch, D., and Von Koskull-Döring, P. (2006). The heat stress transcription factor HsfA2 serves as a regulatory amplifier of a subset of genes in the heat stress response in Arabidopsis. Plant Mol. Biol. 60: 759–772.
53. Shacter, E. (2000). Protein oxidative damage. Methods Enzymol. 319: 428–436.
54. Shahnejat-bushehri, S., Mueller-roeber, B., and Balazadeh, S. (2012). Arabidopsis NAC transcription factor JUNGBRUNNEN1 affects thermomemory- associated genes and enhances heat stress tolerance in primed and unprimed conditions. 7: 1518–1521.
55. Shigeoka, S., Ishikawa, T., Tamoi, M., Miyagawa, Y., Takeda, T., Yabuta, Y., and Yoshimura, K. (2002). Regulation and function of ascorbate peroxidase isoenzymes. In Journal of Experimental Botany, pp. 1305–1319.
56. Shigeoka, S., Nakano, Y., and Kitaoka, S. (1980). Purification and some properties of l-ascorbic acid-specific peroxidase in Euglena gracilis z. Arch. Biochem. Biophys. 201: 121–127.
57. Sugio, A., Dreos, R., Aparicio, F., and Maule, A.J. (2009). The cytosolic protein response as a subcomponent of the wider heat shock response in arabidopsis. Plant Cell 21: 642–654.
58. Suzuki, N., Miller, G., Morales, J., Shulaev, V., Torres, M.A., and Mittler, R. (2011). Respiratory burst oxidases: The engines of ROS signaling. Curr. Opin. Plant Biol. 14: 691–699.
59. Suzuki, N., Miller, G., Salazar, C., Mondal, H.A., Shulaev, E., Cortes, D.F., Shuman, J.L., Luo, X., Shah, J., Schlauch, K., Shulaev, V., and Mittler, R. (2013). Temporal-spatial interaction between reactive oxygen species and abscisic acid regulates rapid systemic acclimation in plants. Plant Cell 25: 3553–3569.
60. Vierling, E. (1991). The roles of heat shock proteins in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 42: 579–620.
61. Volkov, R.A., Panchuk, I.I., Mullineaux, P.M., and Schöffl, F. (2006). Heat stress-induced H2O2 is required for effective expression of heat shock genes in Arabidopsis. Plant Mol. Biol. 61: 733–746.
62. Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci.
63. Wang, X., Huang, W., Yang, Z., Liu, J., and Huang, B. (2016). Transcriptional regulation of heat shock proteins and ascorbate peroxidase by CtHsfA2b from African bermudagrass conferring heat tolerance in Arabidopsis. Nat. Publ. Gr.: 1–12.
64. Welinder, K.G. (1992). Superfamily of plant, fungal and bacterial peroxidases. Curr. Opin. Struct. Biol. 2: 388–393.
65. Xia, X.J., Zhou, Y.H., Shi, K., Zhou, J., Foyer, C.H., and Yu, J.Q. (2015). Interplay between reactive oxygen species and hormones in the control of plant development and stress tolerance. J. Exp. Bot. 66: 2839–2856.
66. Zhou, J., Wang, J., Li, X., Xia, X.J., Zhou, Y.H., Shi, K., Chen, Z., and Yu, J.Q. (2014). H2O2 mediates the crosstalk of brassinosteroid and abscisic acid in tomato responses to heat and oxidative stresses. J. Exp. Bot. 66: 2839-2856.

Appendix

APPENDICES

Table A1: List of primers used for PCR amplification.

S. No.	Oligo Name	Sequence (5'-3')
1.	Atactin_F	TATCGCTGACCGTATGAGCAAAG
2.	Atactin_R	TGGACCTGCCTCATCATACTCG
3.	AtHSP101_F	CTCCTTGCAGGGCTAACTG
4.	AtHSP101_R	TTAATCCTCGATCATTTCC
5.	AtAPX2_F	TCTGGTGTCCATCTCCTACCTCC
6.	AtAPX2_R	CAGCAAGCAAGTGCTCTTATTTCA

Table A2: List of primers used for gene expression analysis by Q-PCR.

S. No.	Oligo Name	Sequence (5'-3')
1.	Atactin_F	TATCGCTGACCGTATGAGCAAAG
2.	Atactin_R	TGGACCTGCCTCATCATACTCG
3.	qAtClpB-C_F	CTGTCTCTCGCCTCTTGG
4.	qAtClpB-C_R	AGTTGTCCACCTTCCTCGTG
5.	qAtAPX2_F	TCTGGTGTCCATCTCCTACCTCC
6.	qAtAPX2_R	CAGCAAGCAAGTGCTCTTATTTCA

Table A3: Composition of 10% resolving gel mixture.

Component	Volume taken
Distilled water (RO)	4.97ml
3M Tris (pH-8.8)	1.25ml
Acrylamide (30%)	3.33ml
APS (1.5%)	0.450ml
TEMED	7µl

Table A4: Composition of 3.9% stacking gel mixture.

Components	Volume taken
Distilled water (RO)	2.675ml
0.5M Tris (pH-6.8)	1.25ml
Acrylamide (30%)	0.625ml
APS (1.5%)	0.450ml
TEMED	7µl

Table A5: Composition of electrode buffer.

Components	Amount
Glycine	7.2g
Tris	2.6g
Ascorbate	0.176g
Distilled water (RO)	Volume make up (500 mL)

Table A6: PCR mixture used for amplification of genes.

Component	10µl Reaction	Final Concentration
MQ water	5.85µl	
10X i-Taq buffer	1µl	1X
2.5mM dNTPs	1µl	250µM
10µM Forward Primer	0.5µl	0.5µM
10µM Reverse Primer	0.5µl	0.5µM
Template DNA	1µl	100ng
i-Taq DNA Polymerase	0.2µl	1 Unit

Functional Characterization of Rice Phospholipase C (*OsPLC3*)

Thesis Submitted to the University of Delhi
in the Partial fulfilment for the Degree of
Master of Science
in
Plant Molecular Biology and Biotechnology

2021

Atul Bahukhandi

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi 110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Functional characterization of Rice Phospholipase C (*OsPLC3)***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 28/06/2021

Signature of the Candidate

(ATUL BAHUKHANDI)

Signature of the Supervisor

(PROF. GIRDHAR K. PANDEY))

Signature of the HOD

(PROF. ANIL GROVER)

Preface

Rice is one of the most important food grains and provides more than 20% calories in the human diet. It is the staple food for half of the world's population. Like most cultivated crop species, it is highly sensitive to extreme environmental conditions and displays reduced productivity under stress. Various abiotic stresses affects rice almost at every stage of development from germination to grain filling stage. In response to these adverse conditions, plant evolves highly complex signaling pathways which help them to mitigate these adverse environmental conditions. Lipid mediated signaling is one major pathway occurring in response to various environmental stresses, such as temperature change, salinity, drought and pathogen attack. Several lipids and lipid derived signaling molecules are shown to be involved in abiotic stress responses (Munnik et al., 2014). Understanding the molecular mechanism and functional characterization of signaling molecules will help us to engineer better stress tolerant varieties with higher productivity and yields.

In this dissertation, I attempted to functionally characterize the rice Phospholipase C (*OsPLC3*). Genome wide analysis of rice genome shows the rice *OsPLC3* gene is upregulated in drought, salt and cold stress (Singh et al., 2013). To understand the role of *OsPLC3* function overexpression transgenic lines were generated. For finding the interacting partner of *OsPLC3,* yeast two-hybrid library screening was performed.

This dissertation consists of four chapter. First chapter provide the detailed review of available literature related about the Phospholipase C. Second chapter includes the detail of materials and method employed for conducting the experiments. The third chapter provides results obtained and discussion for observed results. The final chapter includes the conclusion of work carried out and its future aspects.

Acknowledgement

It's my privilege to have joined the laboratory of Prof. Girdhar Pandey for doing my M.Sc. Dissertation. I express my heartfelt regards and indebtedness towards him and, am thankful to him for his trust, invaluable thoughts and academic guidance. I would also like to acknowledge all my teachers of my department Prof. A.K Tyagi, Prof. J.P. Khurana, Prof. Paramjit Khurana, Prof. Anil Grover, Prof. Indranil Dasgupta, Prof. Arun Sharma, Prof. Saurabh Raghuvanshi and Prof. Surekha K. Agarwal for, their valuable suggestion for improving my study and their constant encouragement throughout the study.

I would like, to thank Dr. Israr Ahmed. It was a great privilege to work with him. He displayed immense patience in making me understand the tactics of planning and execution of experiments and to present the research works as clearly as possible. I would also like to thank him for his friendship, empathy and encouragement. I've learned a lot from him, without his help, I could not have finished my M.Sc. dissertation successfully.

I would like to thanks Ms. Kanwaljeet and Ms. Swati for guidance and suggestions to perform crucial experiments. Their help is always appreciated.

I would like to thanks all the lab members Dr. Pooja, Dr. Ajit, Mr. Lokesh Ms. Soma, Ms. Sonam, Ms. Bharkha, Ms. Diksha, for their valuable suggestion and for supporting me throughout my dissertation.

A special thanks to Ms. Ankita and Mr. Shivam for allowing me to use the rice growth room.

I am grateful to my M.Sc. friends Aashmeen, Apurva, Gopika, K. Thanduanlung, Karishma, Nitika, Shitij and Sritama for their cheerful company and zestful enthusiasm. Joyous moments spent with my M.Sc. friends will always be cherished.

List of commonly used Abbreviations

Col-0	Columbia 0
Ca^{2+}	Calcium ion
ER	Endoplasmic Reticulum
GFP	Green Fluorescent Protein
GST	Glutathione S-Transferase
hr.	Hour
min	Minutes
µM	Micromolar
mM	Millimolar
µg	Microgram
ml	Milliliter
NPC	Non-Specific Phospholipases
PLC	Phospholipase C
35S *CaMV*	Cauliflower Mosaic Virus 35S Promoter
ROS	Reactive Oxygen Species
Y2H	Yeast Two-Hybrid
PtdIns4P	Phosphatidylinositol 4-phosphate
NPC	Non-specific phospholipase C
PLC	Phospholipase C
DAG	Diacylglycerol
IP3	Inositol 1,4,5-trisphosphate

Table of contents

Preface **i**

Acknowledgement **ii**

List of commonly used Abbreviations **iii**

Chapter 1 Review of Literature **1-13**

1.0 Introduction 1

1.1 Phospholipases 1

1.2 Phospholipase C in plants 1

1.3 Phospholipase C classification and domain structure 2

1.4 Subcellular localization of PLCs 4

1.5 PLC signaling in plants 5

1.6 Functional role of PLC in plants 6

1.7 Phospholipase C in rice 12

Chapter 2 Materials and methods **14-24**

2.1 Sequence retrieval and *in-silico* analysis of rice PLCs 14

2.2 Generation of overexpression lines of *OsPLC3* in rice 14

2.3 Yeast Two-Hybrid library screening to find interacting partners of OsPLC3 18

2.4 Preparation of GFP-O*sPLC3* construct for subcellular localization 21

2.5 Floral dip method for *Arabidopsis thaliana* to generate OsPLC3 overexpression lines 22

2.6 Cloning of *OsPLC3* in *pGEX4T-1* to generate for protein expression construct 23

Chapter 3 Results and discussion **25-38**

3.1 Sequence retrieval and homology modelling of rice PLCs 25

3.2 *In-silico* prediction of subcellular localization of rice PLCs 29

3.3 *In-silico* analysis for predicting interacting partners of OsPLC3 30

3.4 Generation of OsPLC3 overexpression lines in rice 31

3.5 Yeast two-hybrid library screening to identify interacting partner of OsPLC3 34

3.6 Subcellular localization of OsPLC3 36

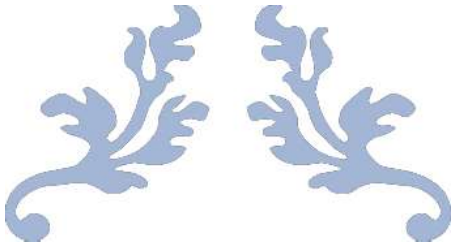
3.7 Floral dip method for *Arabidopsis thaliana* to generate *OsPLC3* overexpression lines 37

3.8 Cloning of *OsPLC3* in *pGEX4T-1* to generate for protein expression construct 37

Chapter 4 Conclusions and Future work 39

Bibliography 41

Appendix 45



CHAPTER 1

REVIEW OF LITERATURE



1.0 Introduction

Phospholipids are major building blocks of membranes. Several environmental signals have been identified to trigger the hydrolysis of membrane phospholipid which, results in the production of different classes of lipid and lipid-derived signaling molecules such as diacylglycerol, diacylglycerol pyrophosphate, phosphatic acid phosphoinositides, inositol polyphosphates, lysophospholipids and free fatty acid. Phospholipase is the main class of lipid hydrolyzing enzyme that catalyzes the preliminary step in lipid hydrolysis and generation of lipid signaling molecules.

1.1 Phospholipases

Phospholipases have been broadly classified as phospholipase A (PLA), phospholipase C (PLC) and phospholipase D (PLD) based on the action of the respective enzyme at different sites on a glycerophospholipid molecule. Phospholipase A hydrolyzes acyl group from sn-1 and sn-2 position of glycerophospholipid molecule to generate free fatty acids and lysophospholipids. Phospholipase C hydrolyzes the entire phosphate group hydrolyzing the glycophosphate bond producing the diacylglycerol (DAG) and phosphorylated head group. Phospholipase D cleaves the phosphodiester bond between the phospholipids to produce phosphatic acid (PA) and a free phosphate head group. These phospholipases play critical role in plant like stress-response signaling, nutrient deficiency, plant growth and development.

1.2 Phospholipase C in plants

Phospholipase C constitutes an important group of lipids hydrolyzing enzymes in the plant and animal kingdom. Phosphatidylinositol-specific PLCs (PI-PLCs) are thoroughly investigated and they specifically act upon phosphatidylinositides (PIP_2) at the plasma membrane and results in the production of second messengers such as diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP_3). IP_3 diffuses in the cytoplasm where it binds to the ligand-gated IP_3 receptors in the intracellular membranes, leading to the release of Ca^{2+} ion from intracellular reservoirs. In

animal system, DAG stays attached to the membrane and recruits the Protein kinase C class protein via its conserved C2 domain, which further relays the signal to the downstream signaling molecules (Rupwate et al., 2012). But the role of PLCs in plant remains enigmatic because of lack of IP_3 receptor and the homologues of Protein kinase C (PKC). Also, the level of PIP_2 in the plant is low as compared to animal membranes (Leeuwen et al.,2007; Munnik et al., 2009; Vermeer et al., 2009).It has been hypothesized that the phosphorylated protein of DAG like phosphatic acid (PA), diacylglycerol pyrophosphate (DGPP) and inositolhexaphosphate (IP_6), phosphorylated product of IP_3 may play a role in downstream signaling (Munnik., 2009). Phospholipase C constitute two major groups based on specificity towards the lipid's substrate. Phosphatidylinositol phospholipase C hydrolysis the phosphatidylinositides (PIP_2) at the membrane and produce downstream signaling molecules like diacylglycerol and inositol 1,4,5-trisphosphate (IP_3) whereas other group shows preferential activity toward the phosphatidylcholine (PC) but also hydrolyzes phosphatidylethanolamine (PE) also called non-specific phospholipases (NPC). Phospholipase C has been reported in abiotic and biotic stress response signaling, development and nutrient remodeling (Hong et al., 2016).

1.3 Phospholipase C classification and domain structure

Based on the domain structure and organization, eukaryotic PI-PLCs are classified into six sub-families, i.e., β, ϒ, δ, ε, η (Munnik et al., 2014). Mammalian cells comprise a total of 13 members divided into six isoforms, whereas plants exhibit only one i.e., PLCζ. Plant PI-PLC shows structural similarity to mammalian PLCζ because this class lacks Pleckstrin homology (PH) domain, which is a characteristic feature of other isoforms. *Arabidopsis* comprises of nine PI-PLCs and six NPCs (Nakamura et al., 2005), rice has four PI-PLCs and five NPCs (Singh et al., 2013). PI-PLCs are reported in many other plant species such as mung bean (Kim et al., 2004), *Physcomitrella* (Mikami et al., 2004), lily (Pan et al., 2005), pea (Liu et al., 2006), maize (Wang et al., 2008), *Brassica napus* (Georges et al., 2009), oat (Huang et al., 2009), tomato (Vossen et al., 2010), wheat (Wang et al., 2020), soybean (Wang et al., 2015), and chickpea (Sagar et al., 2020).Plant PI-PLCs has catalytic X and Y domains that form a TIM barrel structure essential for

phosphodiesterase activity, followed by Ca^{2+}/lipid binding C2 domain at C-terminus and EF hand-like motif at N-terminus but was poorly conserved (figure 1). Plant PI-PLCs are grouped into two classes, based on the Ca^{2+} requirement and intracellular localization (cytosolic or membrane-bound) (Wang et al., 2001, Dowd et al., 2006). The first-class is present in the cytosol and shows preferential activity towards the phosphatidylinositol (PtdIns) and required millimolar Ca^{2+} concentrations. The second-class is membrane-bound and exhibits preferential activity towards the phosphatidylinositol 4-phosphate (PtdIns4P) and $PtdInsP_2$ and require micromolar Ca^{2+} concentrations.

NPCs harbors single Phosphoesterase domain. The phosphoesterase domain of plant NPC is highly conserved (figure 2), containing invariable ENRSFDxxxG, TxPNR, DExxGxxDHV and GxRVPxxxxxP motifs predicated critical for PC-PLC activity. The 50-100 amino acids at C-terminus are highly variable, with different length and sequence conservation between various NPCs that may contribute to the diverse function of NPCs by affecting their subcellular localization and interaction with diverse proteins. NPC4 and NPC5 show the highest homology with 88% similarity in overall amino acid sequences. Though NPC4 is 17 amino acids longer than NPC5 at C-terminus. The N-terminus of NPC1, NPC2 and NPC6 have found to contain a putative signal peptide-containing more than ten amino acids. The N-terminus of NPC3, NPC4 and NPC5 doesn't contain putative signal peptide but have longer C-terminus than NPC1, NPC2 and NPC6.

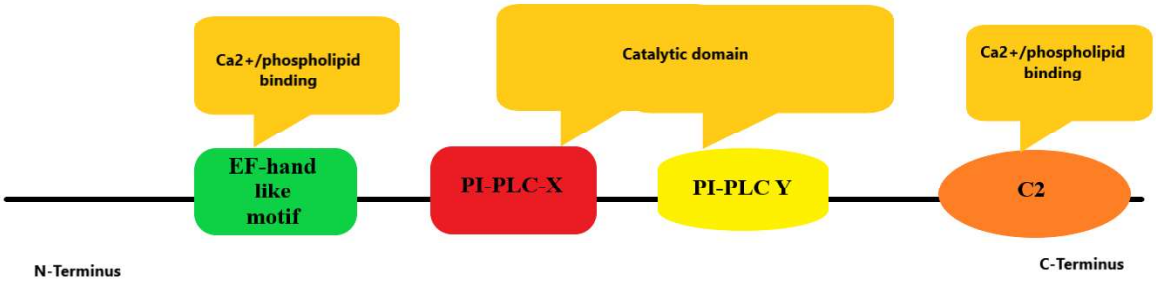


Figure 1: Schematic representation of different domain of PI-PLCs. Plant PLCs consist of N-terminal EF-hand motif domain which was poorly conserved, catalytic PI-PLC-X and PI-PLC-Y domain for phosphoesterase activity and C-terminus C2 lipid binding domain.



Figure 2: Schematic representation of domain of plant NPCs. Plants harbors single Phosphoesterase domain.

1.4 Subcellular localization of PLCs

Plant PI-PLCs are classified into two main groups, one is predominantly localized in the plasma membrane and the other in the cytoplasm. Arabidopsis AtPLC2, AtPLC9 are predominantly localized to the plasma membrane (Zheng et al., 2012; Chen et al., 2019). Rice OsPLCs and soybean PLCs and two PI-PLC from oat seedling are localized in both membrane and cytoplasm (Singh et al., 2013; Shi et al., 1995). Recently, Li et al., (2017) showed the rice OsPLC3, which is localized in the cytosol in normal condition, targeted to the membrane under stress conditions.

In *Arabidopsis thaliana,* subcellular localization of NPCs is highly diverse because of the presence of N-terminal signal peptide. N-terminal signal peptide present in NPC1, NPC2, and NPC6 while missing in others. Based on the published experimental data, NPC1 is localized in secretory pathway compartments like Golgi apparatus, and ER (Krčková et al., 2015). NPC4 is localized in the plasma membrane (Nakamura et al., 2005), NPC5 is mainly present in cytoplasm, but also in ER (Gaude et al., 2008). The localization of NPC2 and NPC6 was shown to be localized in plastids in mesophyll cell (Ngo et al., 2018). In roots, NPC2 protein localized in Golgi and a small fraction is present in ER (Krčková et al., 2018). Similarly, NPC6 protein is localized in the chloroplast and in the microsomal membrane when transiently expressed in tobacco mesophyll cells (Cai et al., 2020). This evidence indicates that NPC2 and NPC6 are localized in diverse subcellular localization in different tissues.

1.5 PLC signaling in plants

PLC signaling pathway well-characterized in the animal system. In the animal systems, diacylglycerol (DAG) produces by PLC activity remains bound to the membrane and activates the Protein Kinase C (PKC). The other product of PLC activity, IP_3 diffuses in the cytosol, and binds to its receptors i.e., ligand-gated Ca^{2+} channels in intercellular membrane to release calcium ions from intercellular reservoirs (Munnik et al., 2009). However, this model is debatable in the plant system because the levels of $PI(4,5)P_2$ is lower in plants membranes than animal's membranes also, the plant PLCs lacks the PH domain might make it difficult to sense these low levels of $PI(4,5)P_2$ (Munnik et al., 2009). In plants, the IP_3 level raised rapidly, followed by an increase in cytosolic Ca^{2+} levels in response to various stimuli such as salt, heat, cold stress and drought (Hong et al., 2016). However, the IP_3 receptor has not been identified in plants. Recently it was proposed that the IP_3 might be converted in inositol hexaphosphate (IP_6) by polyinositolphosphate kinases and IP_6 may lead to the release of Ca^{2+} from intercellular reservoirs (Chen et al., 2011). This hypothesis further supported by the identification of the IP_6 receptor in plants (Tan et al., 2007). Similarly, the plants lack equivalents of Protein Kinase C, which is a major target of DAG in the animal system (Munnik et al., 2009). The DAG is produced by both PI-PLCs and NPCs. In plants, DAG is converted into phosphatidic acid (PA) ,diacylglycerol pyrophosphate (DGPP) by DAG kinases, which can function as downstream signaling molecules. Furthermore, seven DAG kinase reported in the *Arabidopsis* genome and PA receptors identified in the *Arabidopsis* genome, wheat (*Triticum aestivum*; TaDGK; Zhang et al., 2008), rice (*Oryza sativa*; OsDGK; Ge et al., 2012). An overview of PLC signaling, given in figure 3. However, the exact role of DAG, phosphatic acid and IP_6 in plants remains open to scientific studies.

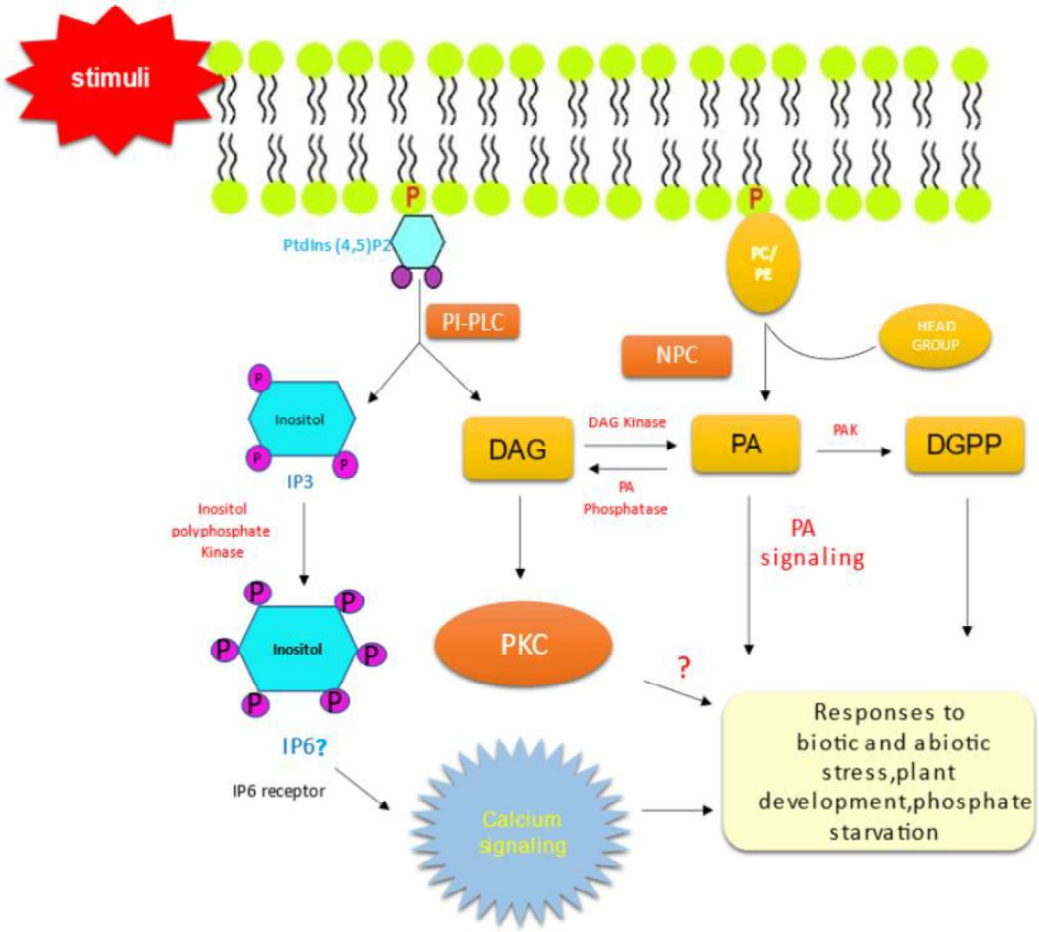


Figure 3: An overview of PLC signaling pathway in plants. Upon sensing the external stimuli PI-PLCs hydrolyze the PI(4,5)P_2 and generate IP_3 and DAG. IP_3 may be converted into IP_6 by inositol polyphosphate kinase, which might bind to IP_6 receptor and release Ca^{2+} from intercellular reservoirs. NPCs act on PE/PC/PS to produce DAG and corresponding head group. The DAG produce by both PI-PLC and NPC might be converted to PA and DGPP by DAG kinase, which may activate downstream signaling molecules, leading to responses to biotic and abiotic stress, development, phosphate starvation.

1.6 Functional role of PLC in plants

1.6.1 Role in Abiotic stress

Several studies reveal that Phospholipase C mediated signaling triggers when plants encounter abiotic stress. *Arabidopsis* PI-PLC, AtPLC9 was shown to be associated with conferring thermotolerance (Zhang et al.,2012). Genetic investigation of *AtPLC9* mutant exhibited sensitivity towards the heat stress. The complementation of *Atplc9* mutant with *AtPLC*9 resulted in the rescue of the thermosensitive phenotype. In

addition, *AtPLC9* overexpressing transgenic plants showed enhanced thermotolerance. Overexpression lines were shown to increase the Ca^{2+} signal in response to heat stress, indicating the role of AtPLC9 in raising Ca^{2+} levels in heat stress. Also, the expression of small HS protein HSP18.2 and HSP25.3 enhanced in *AtPLC9* overexpressing lines (Zhang et al.,2012). The related role of AtPLC3 is established in conferring heat tolerance in *Arabidopsis. AtPLC3* mutants shows sensitivity to heat stress, complementation of mutant with *AtPLC3* shown to rescue the mutant thermosensitive phenotype. Overexpression lines exhibited higher resistance to heat stress compared to WT plants. *AtPLC3* mutants exhibited reduction in heat stress-induced calcium signals. Also, the *HSP* gene, *HSP18.2, HSP25.5, HSP70-1* and *HSP83* were up-regulated in *AtPLC3* overexpressing lines after heat stress. These experiments indicate that AtPLCs are involved in heat stress response by modulating intracellular calcium levels. Liu et al., (2020) showed heterologous expression of *Arabidopsis PLC9* (*AtPLC9*) in rice results in enhanced thermotolerance in transgenic lines. Recently, Krckova et al., (2015) provided evidence of the involvement of NPC1 in response to basal thermotolerance. Loss-of-function mutants for *NPC1* exhibited a reduced basal thermotolerance, whereas the overexpression of *NPC1* results in enhanced resistance to heat stress compared to wild types when exposed to a higher temperature. However, the molecular mechanism of thermotolerance is yet to be identified.

High salinity is a major abiotic stress that severely affects growth, development and yield in plants. Salt stress stimulates the rapid accumulation of phosphatidylinositol 4,5-bisphosphate ($PI(4,5)P_2$), IP_3, diacylglycerol (DAG), phosphatidic acid (PA), diacylglycerol pyrophosphate (DGPP), and phosphatidylcholine (PC) in plant cells (Munnik et al., 2009). Phosphoinositide-specific phospholipase C (PI-PLC) is a class of enzymes that hydrolyze PIP_2 into two second messengers, IP_3 and DAG, and participate in various signal transduction pathways in eukaryotic cells. Calcium spike is a hallmark of salt stress perception in plants (Munnik et al., 2009). Recently, OsPLC1 was reported to confer stress-induced salt tolerance in rice (Li et al., 2017). Knock-out mutant of *OsPLC1* leads to higher accumulation of Na^+ ions, whereas its overexpression (OsPLC*1-OE*) lines accumulate less Na^+ compared to wild type in the shoot. Loss-of-function lines showed reduced chlorophyll content and seedling growth and are highly sensitive to salt stress.

However, the overexpression and knock-out doesn't alter Na^+ in roots. This result implicit OsPLC1 may involve in the transport of Na^+ to shoots. Subcellular localization studies show OsPLC1 is localized in the cytosol, but in response to salt stress, OsPLC1 recruited from cytosol to plasma membrane where it hydrolyzes PtdIns4P. Ion chromatography analysis reveals that OsPLC1 prefers PtdIns4P to PIP_2.OsPLC3 mediated Ca^{2+} signaling prevents accumulation of Na^+ in leaf blades (Li et al., 2017). Similarly, AtPLC4 is involved in negatively regulating salt stress response in *Arabidopsis* seedling (Xia et al., 2017). Mutant seedling shows hyposensitivity to salt stress compared to wild type seedling. Expression of *AtPLC4* complements *atplc4* knock out mutant phenotype. Transgenic lines overexpressing *AtPLC4* have a salt hypersensitive phenotype. Using aequorin as calcium indicator in mutant and overexpression lines, it was demonstrated that AtPLC4 positively regulates stress-induced Ca^{2+} increase (Xia et al., 2017). The transcript levels of salt stress responsive genes *RD29B, MYB15 and ZAT10* inversely regulated in *AtPLC4* mutant and *PLC4-OE* seedlings.

Arabidopsis NPC4 regulates salt and drought stress responses (Deng et al., 2018). *AtNPC4* knock-out mutant show, reduced levels of DAG. Phenotypic analysis demonstrated that *AtNPC4* knock-out mutants with a reduction of sensitivity to ABA in seed germination, root elongation and stomatal movements and exhibited sensitivity to hyperosmotic stress (Deng et al., 2018). Overexpression of At*NPC4* leads to enhanced salt tolerance and higher ABA sensitivity. Similarly, NPC5 shown to involve in the development of lateral root under mild salt stress (Peters et al., 2014). *NPC5* mutant develop fewer lateral roots under mild salt stress. Complementation with full length *NPC5* rescue the mutant phenotype. *NPC5*-OE lines have increased the number of root length (Peters et al., 2014). The level of DAG was higher in *NPC5*-OE, while lower in *npc5* mutant. Exogenous supplementation of DAG rescues the *npc5* mutant root phenotype (Peters et al., 2014).

Recently, many evidences showed the possible role of NPC4 in Aluminum toxicity (Pejchar et al. 2015). Using microsomal fraction from BY-2, tobacco cells incubated with fluorescent PC, shows reduced fluorescent DAG when treated with $AlCl_3$, indicating PC hydrolysis affected by high aluminum concentration (Pejchar et al., 2015; Pejchar et al., 2010). GUS reporter assay showed that expression of NPC4

decrease in root. Knockout mutant of *NCP4* exhibit higher sensitivity to $AlCl_3$ stress than wild type (Pejchar et al., 2015).

1.6.2 Plant phospholipase C and development

Role of PI-PLC in the pollen tube development is well known. PLC2 and PLC3 were reported to be involved in pollen tubes growth in *Petunia inflata* and *Nicotiana tabacum,* respectively (Dowd et al., 2006; Helling et al., 2006). PLC2 can promote early flowering and maturation in *Brassica napus* (Georges et al., 2009). Recently, Bose et al., (2021) provide evidence for the function of NPC2 and NPC6 in pollen tube development. NPC2 and NPC6 are highly expressed in developing pollen tube. Overexpression of NPC2 and NPC6 enhanced the pollen germination and pollen tube length and suppression lead to the reduction of pollen germination and pollen tube length. Gain- and loss-of-function studies indicate the NPC2 and NPC6 involved in lipid remodeling by hydrolyzing the phospholipids to produce triacylglycerol (TAG), which is crucial for pollen development. Intriguingly, NPC members in *Arabidopsis* have been involved in hormone signaling, which regulates diverse aspects of plant growth and development. Wimalasekera et al., (2010) show the NPC members are differentially expressed after auxin, cytokinin and brassinolide treatments. NPC3, NPC4 promoter *GUS* seedling showed enhanced expression of GUS in root tips, cotyledons, leaf margins and floral organs in response to α-Naphthaleneacetic acid (1-NAA) treatment (Wimalasekera et al.,2010). Moreover, *NPC3* and *NPC4* knock-out mutants showed shorter primary root and lesser lateral root density in response to low brassinolide treatment, but higher in increased lateral root density in higher (0.05–1.0 mM) exogenous BL treatment. Further, the expression of BR induced genes *TCH4* and *LRX2* genes was also decreased, which regulates the cell expansion. These results indicate the potential role of NPC3 and NPC4 in root development in response to brassinolide treatment.

Recently, PLC2 has been shown to be involved in auxin-modulated root development (Cheng et al., 2019). PLC2 changes the polar distribution of PIN2 efflux carrier, thus regulates, auxin response and root development. The PLC2 mutants exhibit defective root phenotype, including short primary growth, impaired root hair growth and impaired root gravitropism.

A recent study on *Arabidopsis* NPCs shows the NPC2 and NPC6 act redundantly in gametophyte development (Ngo et al., 2018). Double mutants *npc2npc6* exhibits lethal phenotype due to the defect in the development of male and female gametophyte (Ngo et al., 2018). Transmission electron microscopy revealed the development of many pollens in anthers of *npc2-1/–* *npc6-2/+* or *npc2-1/+* *npc6-2/–* was arrested in mitosis I stage. Knockdown double mutant of NPC2 and NPC6 showed that root phenotype has significantly retarded, with an increase in lateral root density. Interesting, exogenous supplementation of phosphocholine rescued the short root phenotype. The expression of *PMT1* phosphobase N-methyltransferase 1 is higher in knock-down lines, which encodes the enzyme that produces phosphocholine from phosphoethanolamine and the reduced levels of *PMT1* are reduced to wild type by exogenous application of phosphocholine (Pcho). These results indicate that NPC2 and NPC6 interact with the PMT pathway to produce Phosphocholine (Pcho). Thus, NPC2 and NPC6 redundantly play tissue-specific roles in flowers and roots.

1.6.3 Phospholipase C in nutrient starvation

Phosphorus is an indispensable element for plant growth, development, and reproduction. Phosphorus is required for many enzyme's activity, also the major component of nucleic acids and membrane lipids. Under the Pi starving plant, NPC4 and NPC5 are shown to induce and assumed to promote mobilization of phosphate from the membrane phospholipids by cleaving the phosphate head group from phosphatidylcholine and other membrane phospholipids (Tjellstrom et al., 2010). The removal of the phosphate group doesn't affect the integrity of the membrane and functionality of the membrane because the DAG yield through NPC activity acts as a precursor of other non-phosphorous membrane lipids like galactolipid and sulfolipid. Nakamura et al., (2005) conducted a comprehensive analysis of *the Arabidopsis* NPC family under phosphate limiting conditions and found NPC4 was highly upregulated under low-phosphate (0-0.01mM Pi) condition for 6-10 days among all the NPC members. In addition, knock-out mutants displayed decreased PC- phospholipase activity and higher DAG accumulation. However, loss of NPC4 activity leads to a decrease in phosphatidylcholine (PC) in roots and leaves only at an early stage during the phosphate deficiency, implicating the PC may not be the principal substrate of NPC

during phosphate deprivation conditions. Recently, NPC4 was reported to hydrolyze Glycosylinositolphosphoceramide (GIPC), a major phosphosphingolipid present in the membrane rafts (Yang et al.,2021). In Arabidopsis, *NPC4* mutant inhibits the GIPC decrease induced by phosphate deficiency. Yang et al., (2021) showed in phosphate limiting conditions, NPC4 is S-acylated and targeted to plasma membrane which hydrolyses the Glycosylinositolphosphoceramide. S-acylation is a post-transcriptional and reversible protein modification by which a fatty acid primarily palmitate is attached to the cysteine residue via thioester bond, resulting in the change in the stability, activity and localization of protein. Protein S-acylation reported function in various cellular function such as hormone regulation, stress response, disease resistance (Hemsley et al., 2009; Li et al., 2017; Turnbull et al., 2017). Subcellular localization study with full-length NPC4 fused with GFP and C- terminally truncated NPC4$^{\Delta 17}$ fused with GFP in *Arabidopsis* protoplast and tobacco leaves show the plasma membrane localization of full-length NPC4-GFP whereas the C-terminus truncated protein NPC4$^{\Delta 17}$-GFP was cytoplasm localized (Yang et al.,2021). Protein fractionation studies showed the presence of NPC4 in the microsomal fraction, whereas the NPC4$^{\Delta 17}$ in the soluble fraction implicating 17 amino acids at the C-terminal harbors the amino acid residue crucial for membrane association. C-terminal analysis reveals amino acid cysteine at position 533 identified as potential acylation site. Mutation of Cys-533 to Ala-533 results in loss of acylation and association of NPC4 to the membrane. Complementation of knock-out mutant with NPC4^{C533A} didn't rescue the mutant phenotype. This implicit non-acylated NPC4 fails to hydrolyze GIPC in phosphorus deprivation. Also, NPC4^{C533A} also fails to rescue the primary root growth phenotype under stress. Similarly, NPC4 of *Brassica napus BnaC01*.NPC4 was S-acylated at Cys-531 residue. Mutation of Cys-531 to Ala-531 results in loss of membrane association.

1.6.4 Plant phospholipase and biotic stress

The role of plant PLCs are implicated not only in abiotic stresses but also in biotic stress, including plant immunity (Abd-El-Haliem et al., 2017). PLC are involved in mediating MAMP trigger immunity, effector trigger immunity, phytoalexin, and ROS generation along with HR in response to the invasion of biotic agents like bacterium, fungi (Abd-El-Haliem et al., 2017). Silencing of *SlPLC6* results in HR response and increased colonization of cf-4 containing tomato plants by Avr4 carrying *Cladosporium fulvum.Also,Avr4* induced HR is reduced in *SlPLC4* silenced *cf-4* carrying tomato plant

(Vossen et al. 2010). Heterologous expression of tomato *SILPC4* in *Nicotiana bethamina* also results in higher Avr4/Cf-*4* induced HR. Studies with PI-PLC inhibitor U73122 results in phytoalexin and ROS production together with HR response in Cf-4/Avr4 elicited tobacco cells (Wang et al., 2013). Xylanase treatment of tomato suspension cell results in rapid induction of NO and HR that leads to an increase in expression of *SIPLC5* (Gonororazky et al., 2014). Gonororazky et al., (2014) showed *SIPLC2* also required for HR. Silencing of *SIPLC2* shows decrease in the expression of a defense- related gene such as, pathogen- related (*PR*) genes *and hrs203J.*

A recent study on *Arabidopsis* PLC2 suggested their role in MAMP trigger immunity (MTI). *PLC2* silenced lines are more susceptible to sprayed treatment of *P. syringae* pv tomato *DC3000 hrcC*- strain lacking functional TTSS which are important for the MTI suppression but not to wild type *P. syringae* pv tomato *DC3000*. Bacterial infiltration with *Pst DC3000 hrcC*- in *PLC2* silenced yielded no significant difference in susceptibility between silenced and non-silenced plants.

These indicate the role of PLC2 in MTI. In response to flg2, *PLC2* silenced plant shows wild type MAPK activation and *PHI, WRKY33 and FRK1* immune marker gene expression but shows reduced ROS dependent responses i.e., stomatal closure and callose deposition. *PLC2* silenced plants show reduction of ROS upon the treatment of flg22, indicating the role of PLC2 in full activation of ROS and ROS dependent MTI (D'Ambrosio et al., 2017). Also, PLC2 is shown to interact with NADPH oxidase and act upstream of it.

1.7 Phospholipase C in rice

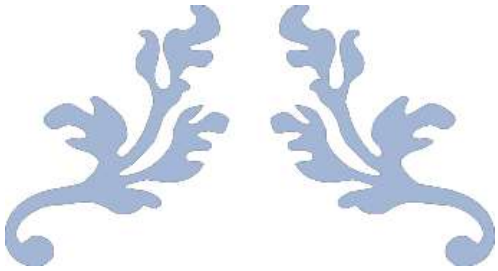
Genome-wide exploration results in the identification of nine phospholipase C in the rice genome (Singh et al., 2013). Phospholipase C in rice can be divided into two major groups; PI-PLCs and NPCs based upon their substrate specificity. A total of four PI-PLCs and six NPCs have been identified.

Domain and motif analysis of PI-PLC showed the presence of catalytic PI-PLC-X and PI-PLC-Y domain at N-terminus and C2 lipid/Ca^{2+} binding domain at C terminus. The EF-hand like motif found only in OsPLC1 and OsPLC4 (Singh et al., 2013). Structural analysis of rice PLC genes revealed all PI-PLCs contain 6-8 introns. Likewise, all the NPC members contain a highly conserved phosphoesterase domain. However, the NPCs contain 0-3 introns.

Analysis of rice PI-PLCs and NPCs in different abiotic stress conditions such as, cold, drought and salt stress, revealed that PI-PLCs and NPCs expressed differentially under these conditions. The rice PI- PLC *OsPLC3* is highly upregulated during cold, drought and salt stress, indicating this gene might have a significant role in signaling during abiotic stress (Singh et al., 2013). Functional characterization of *OsPLC3* helps us to understand the underlying signaling pathway occurring during stress conditions. Moreover, the functional knowledge of this gene will enable us to engineer better rice varieties which can tolerate adverse conditions without losing yield.

Major objectives

- Generation of overexpression of *OsPLC3* in rice.
- Expression of *OsPLC3* in *E. coli* to study its enzyme activity.
- Yeast two-hybrid library screening using OsPLC3 as bait to find out its interacting partners.
- To find the subcellular localization of OsPLC3.



CHAPTER 2

MATERIALS AND METHODS



2.0 Sequence retrieval and *in-silico* analysis

Nucleotide and protein sequences of genes used in this study were retrieved from the rice genome annotation project database (http://rice.plantbiology.msu.edu/) using locus search. The *in-silico* subcellular localizations were predicted using various prediction tools such as WoLF PSORT (https://www.genscript.com/wolf-psort.html), ChloroP (http://www.cbs.dtu.dk/services/TargetP/),
TargetP-2.0 (http://www.cbs.dtu.dk/services/TargetP/).
The structure of rice phospholipases was predicted using protein modelling tool Swiss-Model (https://swissmodel.expasy.org/). Also, the interacting partners of OsPLC3 were predicted using STRING software.

2.1 Generation of overexpression lines of *OsPLC3* in rice

2.1.1 Preparation of *Agrobacterium tumefaciens* competent cells

A single bacterial colony of *A. tumefaciens* (*EHA105*) was inoculated in 5 mL of Luria Bertani (LB) media with rifampicin (50 µg/mL) and incubated at 28ºC on a rotary shaker at 200 rpm for overnight. Secondary culture was grown by inoculating 1 mL of primary culture in 100 mL of fresh LB media and incubated until the O.D reached 0.5. The secondary culture was centrifuged at 4000 rpm for 5 minutes at 4ºC. The supernatant was discarded and the pellet was resuspended in 20 mL of 0.1M $CaCl_2$ and incubated in ice for 20 mins. After incubation, the bacterial suspension was again centrifuged under similar conditions and the pellet was resuspended in 5 mL of 0.1M CaCl2, containing 15% glycerol. Resuspended cells were aliquoted (200 µl each) in 1.5 ml microcentrifuge tubes (MCTs) and frozen in liquid nitrogen. The competent cells were stored at -80 ºC.

2.1.2 Transformation of *Agrobacterium* cells with plasmid DNA

A competent cell was thawed by incubating on ice for 5 min. 5 µL of plasmid DNA was added to the competent cell and incubated in ice for 5 min. The competent cells were frozen in liquid N_2 for 2 min. Then heat shock was given at 37ºC for 5 min., followed by incubation on ice for 5 min. Then 800 µL of LB media was added for recovery of bacterial cells from heat shock and incubated at 28ºC, 200 rpm for 5 hrs. After 5 hrs., the bacterial cells were pelleted down at 4000 rpm for 2 min. The 800 µL of supernatant was discarded and the pellet was resuspended in 200 µL LB and plated

on desired selection media (rifampicin (50 µg/mL), spectinomycin (50 µg/mL)). The plates were covered with aluminum foil and incubated at 28ºC for 2 days.

2.1.3 Plasmid isolation by alkaline lysis method

For plasmid isolation, a single bacterial colony was inoculated in 5 mL of LB with rifampicin (50 µg/mL) and spectinomycin (50 µg/mL) as selective antibiotic and incubated overnight on a rotary shaker. The 2 mL of overnight culture was transferred to MCT, followed by centrifugation at 13000 rpm for 2 min. The supernatant was discarded and to the pellet, 200 µL of Lysis I solution was added, followed by vortexing until the pellet was completely dissolved. Then 200 µL freshly prepared lysis II solution was added and mixed gently by inverting the MCT for 2 min. After this, 300 µL of lysis III solution was added and mixed gently by inverting the MCT, followed by incubation on ice for 10 min. After 10 min, they were centrifuged at 12000 rpm for 10 min and the supernatant obtained transferred to fresh MCT. An equal volume of isopropanol was added to the supernatant to precipitate down plasmid DNA and centrifuged at 12000 rpm to obtain the pellet. Pellet were washed with 70% ethanol by centrifuging at 8000 rpm for 5 min. The supernatant was discarded and the pellet was air-dried at room temperature. Finally, the pellet was dissolved in 10 µL of sterile MQ water.

2.1.4 Analysis of constructs

The isolated plasmid was confirmed by PCR. The construct was validated by using *CaMV 35S* forward and *OsPLC3* gene-specific reverse primers. The description of primers used for amplification and cycling conditions are given in table 1 and figure 4). The positive clone was selected for preparation of glycerol stock and further used for transformation of rice calli.

Table 1: Primer sequences for *OsPLC3* amplification

Primers	Sequence
CaMV 35S forward	CAAGACCCTTCCTCTATATAA
OsPLC3 TOPO reverse	GAGGAAGTCGAAGCGCATGA

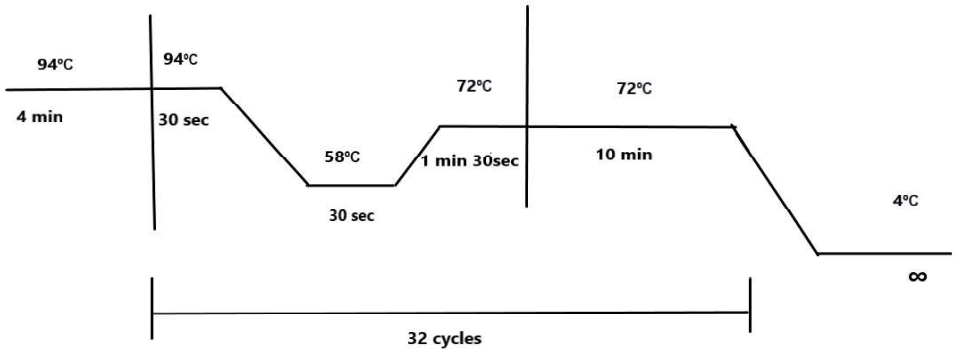


Figure 4: CYCLING CONDITIONS FOR PCR

2.1.5 Rice tissue culture

Pusa basmati-1 (PB-1) was used for rice transformation. PB-1 seeds were dehusked and sterilized using 70% ethanol, then 0.1% $HgCl_2$ with few drops of teepol for 20 min. with shaking at 200 rpm, followed by 6-7 washes with autoclaved MQ. The seeds were soaked in MQ water and kept in the dark for overnight. Next day, the seeds were blotted, dried on sterile tissue paper and then transferred to callus induction media (Table 2) using sterile forceps. The plates are wrapped with aluminum foil and incubated at 28ºC for 21 days in the rice growth chamber.

Table 2: Composition of Callus Induction Media

S.NO.	Components	Concentration
1	NB media	24 g/L
2	Proline	2.8 g/L
3	Glutamine	500 mg/L
4	Casein hydrolysate	200 mg/L
5	2,4-D	2 mg/L
6	Sucrose	3%
7	Phytagel	0.4%
		pH-5.85

2.1.6 Subculturing rice calli

The 21 days old rice calli were transferred to fresh callus induction media for 4 days under similar conditions. Parallelly, two days before co-cultivation primary culture was set up by inoculating the glycerol stock of *Agrobacterium* harboring the desired construct in 5 mL YEP media with desired antibiotic selection and incubated at 28ºC, 200 rpm for overnight. The secondary culture was set up by inoculating 1 mL primary culture into 100 mL of fresh YEP media with desired antibiotic selection and incubated at 28ºC, 200 rpm until O.D reached 0.4-0.5.

2.1.7 Co-cultivation

The secondary culture was centrifuged at 6,000 rpm for 20 min. at 28ºC. The pellet was resuspended in callus induction media containing 200 µL Acetosyringone (Agro inoculation media) in the sterile conical flask. The subcultured calli were infected with Agro inoculation media at 150 rpm for 15 min. at 28ºC. After 15 min. the Agro-inoculation media was discarded and calli were blotted dry on sterile tissue paper. The calli were transferred on co-cultivation media (callus induction media supplemented with 200 µL Acetosyringone), wrapped with aluminum foil and incubated at 28ºC for 2 days in the rice growth chamber.

2.1.8 Selection

After 2 days the calli were checked for growth of *Agrobacterium* cells around the calli. The calli were washed 6-7 times with autoclave MQ. Finally, calli were blotted on sterile tissue paper and transferred to selection media (callus induction media supplemented with 50 µg/mL hygromycin, and 300 mg/L cefotaxime) and incubated at 28ºC in 16h/8h light/dark cycle for 14 days. Calli were subjected to two rounds of selection by repeated subculturing on selection media for every 14 days interval.

2.1.9 Regeneration

The healthy white calli were transferred to selection regeneration media (Regeneration media (Table 3), supplemented with hygromycin (50µg/L), and cefatoxime (300mg/L) for 14 days at 28ºC in 16h/8h light/dark cycle. The calli were subjected to successive

rounds of selection regeneration by repeated subculturing on fresh regeneration medium after every 14 days interval, until the shoots emerged from the calli.

Table 3: Composition of Regeneration Media

S.No.	Components	Concentration
1	NB media	24 g/L
2	Casein Hydrolysate	2 g/L
3	Sorbitol	30 g/L
4	Kinetin	2 mg/L
5	NAA	0.02 mg/L
6	Sucrose	3%
7	Phytagel	0.4%
		pH-5.85

2.2 Yeast two-hybrid library screening to identify interacting partners of OsPLC3

2.2.1 Yeast transformation protocol

The OsPLC3 gene was previously cloned in pGBKT7 vector (Gandhi., 2020). The construct was then mobilized into yeast strain Y2H GOLD. A single colony (Y2H GOLD) was inoculated in 5 mL YPD media and incubated at 30ºC for 2 days. The secondary culture was set up by inoculating 1 mL of primary culture in 100 mL of fresh YPD media and grown at 30ºC until the O.D reached 0.4. The secondary culture was centrifuged at 5000 rpm for 5 min. at 20ºC. Then the supernatant was discarded and the pellet was resuspended in 25 mL sterile MQ followed by centrifugation at 5000 rpm for 5 min. 20ºC. The above step was repeated with a similar condition. The pellet was resuspended in 1 mL sterile MQ and transferred to 1.5 mL MCT. The MCT was centrifuged at 14000 rpm for 30 sec. at RT and the pellet was resuspended in 1 mL of sterile MQ. The resuspended cells were aliquoted as 100 µl in 1.5 ml MCTS. 100µl of yeast cell

suspension aliquoted in 1.5 mL MCT and centrifuged at 14000 rpm for 30 sec. and 360µl of transformation mix (Table 4) added to the yeast cell suspension. The pellet was resuspended by vortexing. The cells were subjected to heat shock at 42ºC for 40 min. then centrifuged at 13000 rpm for 30 sec. The supernatant was discarded and the pellet was resuspended in 1 mL sterile MQ and plated on SD/-T dropout media, then incubated at 30ºC for 2-3 days.

Table 4: Composition of transformation mix

S.NO	Transformation mix components	Volume(µL)
1	PEG 3350 (50% (w/v))	240
2	LiAc 1.0 M	36
3	Single-stranded carrier DNA (2.0 mg ml^{-1})	50
4	Plasmid DNA plus sterile water	34
5	Total volume	360

2.2.2 Yeast plasmid isolation

The single colony was inoculated in 5 mL of YPD media and incubated at 30ºC for 2 days. The 250 µL of primary culture was transferred to 1.5 mL MCT. 250 µL of freshly prepared lysis solution (100mM NaCl,10mM Tris-Cl pH 8.0,1mM EDTA,0.1% SDS) was added to the 250 µL primary culture and mixed properly. The mixture was extracted using 500 µL of Phenol: Chloroform: Isoamyl alcohol (25:24:1). The aqueous phase was transferred to fresh MCT. Equal volume of 100% ethanol was added and pelleted down at 13000 rpm for 10 min. Supernatant was discarded and pellet was washed with 70% ethanol and pellet was air-dried. The pellet was resuspended in 10 µL of sterile MQ.

2.2.3 Analysis of constructs

The isolated plasmid was verified by PCR. The construct was validated by gene-specific primer. The description of primers used for amplification of ORF and cycling conditions are given in figure 5 and table 5.

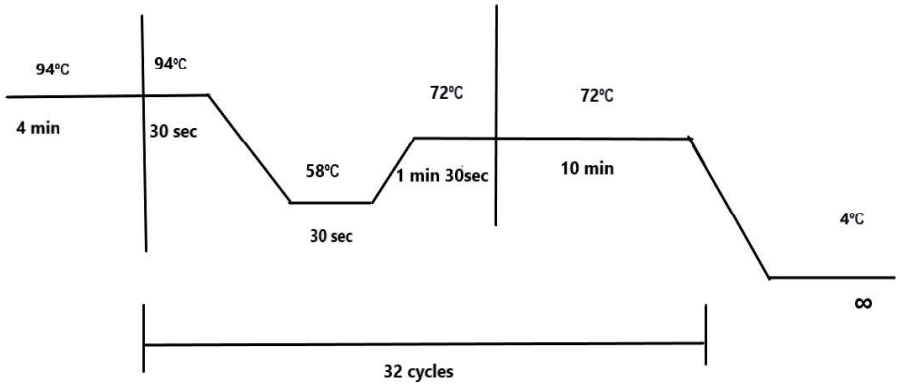


Figure 5: CYCLING CONDITION FOR PCR

Table 5: Primer sequences for OsPLC3 amplification

Primer	Sequence
OsPLC3 TOPO forward	CACCATGGGGAGCTACGCGTACAA
*OsPLC*3 TOPO reverse	GAGGAAGTCGAAGCGCATGA

2.2.4 Mating and screening of library

A single 2-3 mm colony from freshly streaked plate was inoculated in 50 mL of SD/-Trp dropout media to prepare bait culture and incubated at 30ºC until O. D reached to 0.8. The culture was centrifuged for 5 min. at 1000g. Supernatant was discarded and pellet was resuspended in 4-5 mL of SD/-Trp liquid media.1 mL of library strain (Y187), thawed on ice and mixed with 4-5 mL of bait culture in 2 L sterile flask. 45 mL of 2X YPDA (with 50 g/ml kanamycin) was added to the flask. The library was rinsed twice with 1 mL of 2X YPDA and added to the flask. The flask was incubated at 30ºC, 50 rpm for 20 hrs. to allow mating. After 20hrs., a drop of culture was observed under the microscope for the presence of Zygotes. The cells were centrifuged for 10 min at 1000g. The empty flask was rinsed twice with 50 mL of 0.5X YPDA (with 50 g/ml kanamycin) this was added to resuspend the pellet and again centrifuged for 10 min. at 1000g. Supernatant

was discarded and pellet was resuspended in 10 mL of 0.5X YPDA (with 50 g/ml kanamycin). The mated culture was plated in SD/-LT agar plates with 100-200 µL each. The plates were incubated at 30ºC for 2-3 days. After 4-5 days the colonies were patched on high stringency SD/-HLT agar plates.

2.3 Preparation of *GFP-OsPLC3* construct for subcellular localization

The *OsPLC3* gene was previously cloned in entry vector *pENTR-D/TOPO* (Invitrogen) by Gateway® technology (Gandhi., 2020). The *OsPLC3/TOPO* was digested with *Nru*I to linearize plasmid and then subsequently mobilized to gateway compatible binary destination vector *pSITE2CA* (Figure 6) by LR clonase reaction. In *pSITE2CA* vector, the expression of cloned gene is controlled by 2X*CaMV 35S* promoter. The construct will be confirmed by PCR and sequencing. The LR Clonase reaction composition is given in table 6.

Table 6: Components for LR Clonase reaction

S.No.	Components	Volume (µL)
1	*OsPLC3/TOPO NruI* digest	1.0
2	*pSITE2CA* vector	1.0
3	LR Clonase enzyme	0.5
4	TE buffer	2.5
	Total volume	5.0

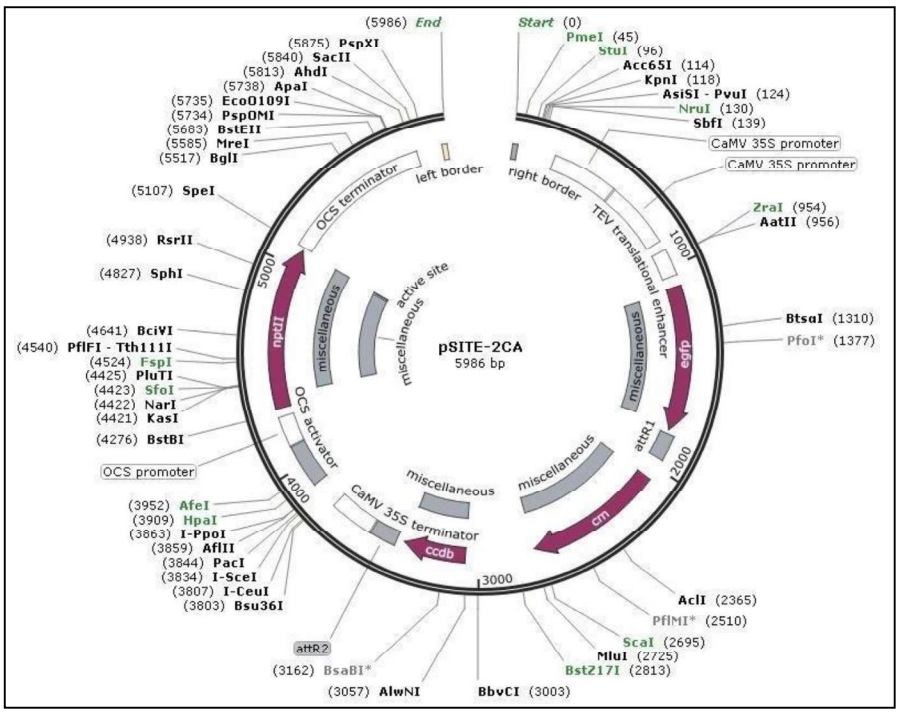


Figure 6: Vector map for pSITE-2CA

2.4 Floral dip method for *Arabidopsis thaliana* to generate *OsPLC3* overexpression lines

For the floral dip *Col-0*, Arabidopsis plants were grown until they start bolting. Primary bolts were cut off to induce the development of secondary bolts. Two days before the floral dip, the primary culture of *Agrobacterium* harboring the *OsPLC3-pGWB520* construct was initiated by inoculating glycerol stock of *Agrobacterium* in 5 mL of YEB and incubated overnight at 28ºC, 200 rpm. A day before the floral dip, 1 mL of primary culture was inoculated into 200 mL of fresh YEB media, incubated at 28ºC 200 rpm overnight (Secondary culture). The next day, the secondary culture was centrifuged at 4000 g for 15 min. at 28 in GSA rotor in the ThermoFisher Lynx 6000 Sorvell centrifuge. The supernatant was discarded and the pellet was resuspended in 500 mL of 5% sucrose solution. Before floral dip 180 µL of silwet L-77 was added and mixed thoroughly. The pot was inverted such that all the buds get fully immersed in the bacterial suspension for 2 min. The beaker rotated gently slowly. The pots, kept back in the tray and covered by a humidity chamber placed in the dark for 16-18 hours. The was plant transferred to light condition, supplied water until they started setting seeds.

2.5 Cloning of *OsPLC3* in *pGEX4T-1* to generate for protein expression construct

OsPLC3 gene was amplified from *OsPLC3-TOPO* construct using gene-specific primers having *NotI* and *Sal*I restriction sites in forward and reverse primers, respectively. The *OsPLC3* PCR product was digested with *Not*I and *Sal*I. Also, the *pGEX4T-1* vector (Figure 7) was digested with *Not*I and *Sal*I. The digestion reaction was prepared by adding the following components, as given in table 7 and table 8.

Table 7: Composition for pGEX4T-1 restriction digestion

S.No.	Component	Volume (µL)
1	*pGEX4T-1 vector*	2.0
2	Cut smart buffer(10X)	2.0
3	*Not*I	0.5
4	*Sal*I	0.5
5	MQ	15.0
Total volume		20.0

Table 8: Composition for PLC3 PCR product restriction digestion

S.No.	Component	Volume (µL)
1	*PLC3* PCR product	10.0
2	Cut smart buffer(10X)	2.0
3	*Not*I	0.5
4	*Sal*I	0.5
5	MQ	5.0
Total volume		20.0

The restriction digested samples were quantified on gel and eluted using gel extraction kit. The ligation reaction set up for *OsPLC3* and *pGEX4T-1* restriction digested products, as given in Table 9. The ligation mixture incubated at 22^{o}C overnight and next day the ligation product was transformed into DH5α cells.

Table 9: Composition for *pGEX4T-1* and *OsPLC3* ligation reaction

S.No.	Component	Volume (µL)
1	*OsPLC3* PCR product	7.0
2	*pGEX4T-1*	10.0
3	T4 DNA Ligase	1.0
4	Ligase buffer (10X)	2.0
Total volume		20.0

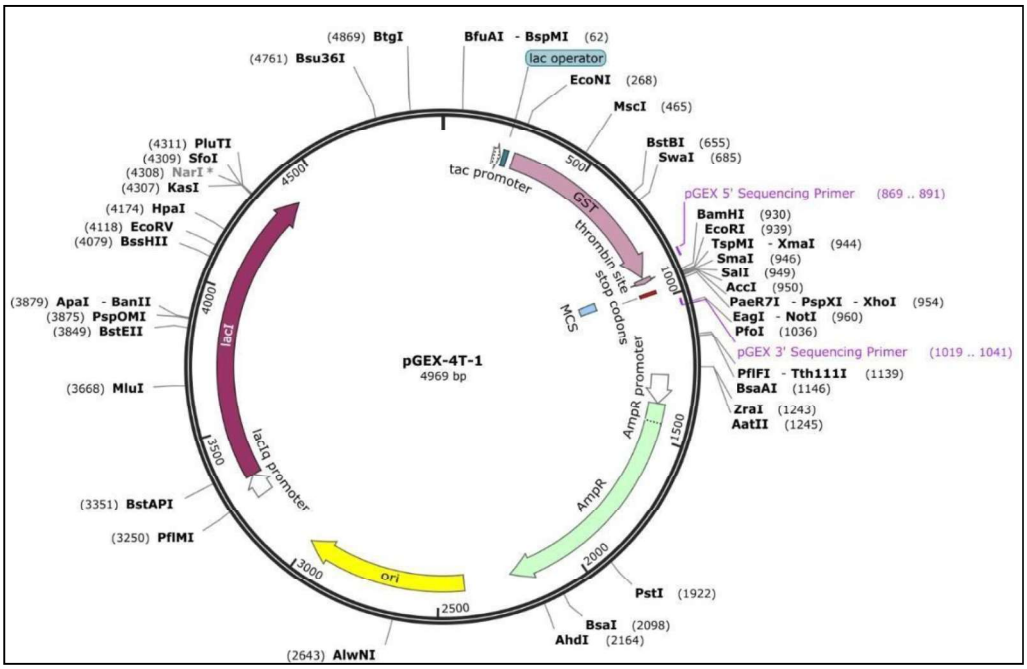


Figure 7: Vector map for pGEX4T-1

CHAPTER 3
RESULT AND DISCUSSION

3. Results and discussions

Genome-wide analysis by Singh at el., (2013) led to identify four PI-PLCs and six NPCs in rice genome. On the basis of expression studies of *OsPLCs* at transcript level, it was found that *OsPLC3* gets induced under salt, cold and drought conditions (Singh et al. 2013). The aim of this study is to functionally characterize this rice phospholipase C (*OsPLC3*).

3.1 Sequence retrieval and homology modelling of rice phospholipase C

3.1.1 Homology modelling of rice phospholipases

The protein sequences were retrieved from the rice genome browser. Swiss-Model programme was used to generate the 3-dimensional structure of rice phospholipase C. The structure of *Sepia officinalis* PLC21 (3qr0.1.A), phospholipase C beta 2 (2zkm.1A), human phospholipase C-beta3 delta (94qj.1B,) *Loligo pealei* PLC21 (3qr1.1A) was used to build the structure of OsPLC1, OsPLC2, OsPLC3 and OsPLC4, respectively. The structure of *Francisella tularensis* Acid Phosphatase A (AcpA) bound to orthovanadate (2d1g.1.A) was used to build structural model for the OsNPCs. Models displayed sequence similarity 17.13-31.24%, coverage 0.67- 0.96%. The GMQE and QMEAN values indicates the quality of structure. For OsPLCs structures GMQE values lie between 0.40-0.63 and QMEAN value lie between -3.53 to -7.42 as given in table 10.

Table 10: Homology modelling of rice phospholipase C

Model	Template	Sequence identity	Oligomeric state	Resolution	Sequence similarity	Coverage	GMQE	QMEAN
OsPLC1	3qr0.1. A	26.72	Monomer	2.00 Å	0.33	0.88	0.59	-3.53
OsPLC2	2zkm.1A	29.16	Monomer	1.62	0.34	0.96	0.63	-4.09
OsPLC3	94qj.1B	31.24	Monomer	3.00	0.35	0.84	0.56	-3.90
OsPLC4	3qr1.1A	25.09	Monomer	3.20	0.33	0.89	0.58	-4.00
OsNPC1	2d1g.1A	19.34	Monomer	1.75	0.30	0.66	0.43	-5.66
OsNPC2	2d1q.1A	19.28	Monomer	1.75	0.30	0.70	0.44	-5.68
OsNPC3	2d1q.1A	17.13	Monomer	1.75	0.30	0.67	0.43	-5.76
OsNPC4	2d1q.1A	18.21	Monomer	1.75	0.30	0.68	0.42	-7.42
OsNPC5	2d1q.1A	19.23	Monomer	1.75	0.30	0.67	0.40	-5.24

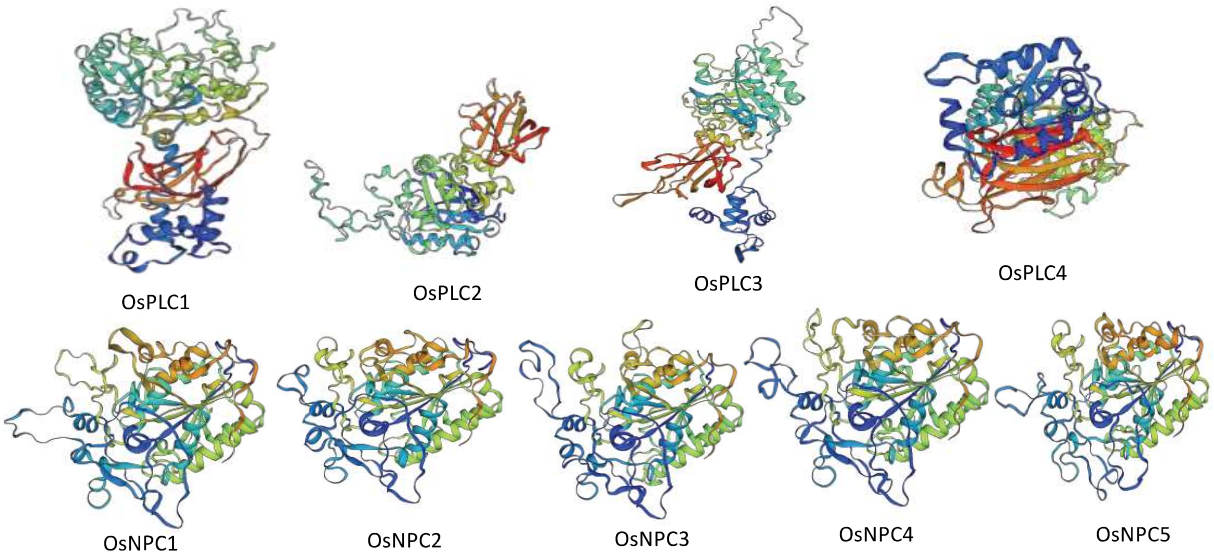


Figure 8: **Predicted 3-dimensional models for rice phospholipase C**. Three-dimensional structures were predicted using SWISS-MODEL. Structure of *S. officinalis* PLC21, phospholipase C beta 2, human phospholipase C-beta3 delta, *Loligo pealei* PLC21 are used to build structure of OsPLC1, OsPLC2 OsPLC3 OsPLC4, respectively. Structure of *Francisella tularensis* Acid Phosphatase A (AcpA) bound to orthovanadate (2d1g.1.A) was used to build a structural model for the OsNPCs.

3.1.2 Structure validation by Ramachandran plot

Ramachandran plot showed the distribution of φ and ψ angle in four PI-PLCs and five NPCs protein models. The OsPLC1 showed 89.10% amino acid residues are in the favored region and 10.90% in the disallowed region, OsPLC2 showed 90.44% amino acid residues are in the allowed region. OsPLC3 displayed 90.27% amino acid residues are in the favored region. OsPLC4 showed 89.44% amino acid residues are in the favored region. NPC1 displayed 87.94% amino acid residues are in the favored region. NPC2 showed 85.71% amino acid residues are in the favored region. NPC3 showed 86.17% amino acid residues are in the favored region. NPC4 displayed 85.07% amino acid residues are in the favored region. NPC5 showed 85.07% amino acid residues are in the favored region.

Ramachandran plots of PLCs proteins confirmed that model structures are following dihedral angles of Ramachandran plot occupied chiefly favorable positions (figure 9).

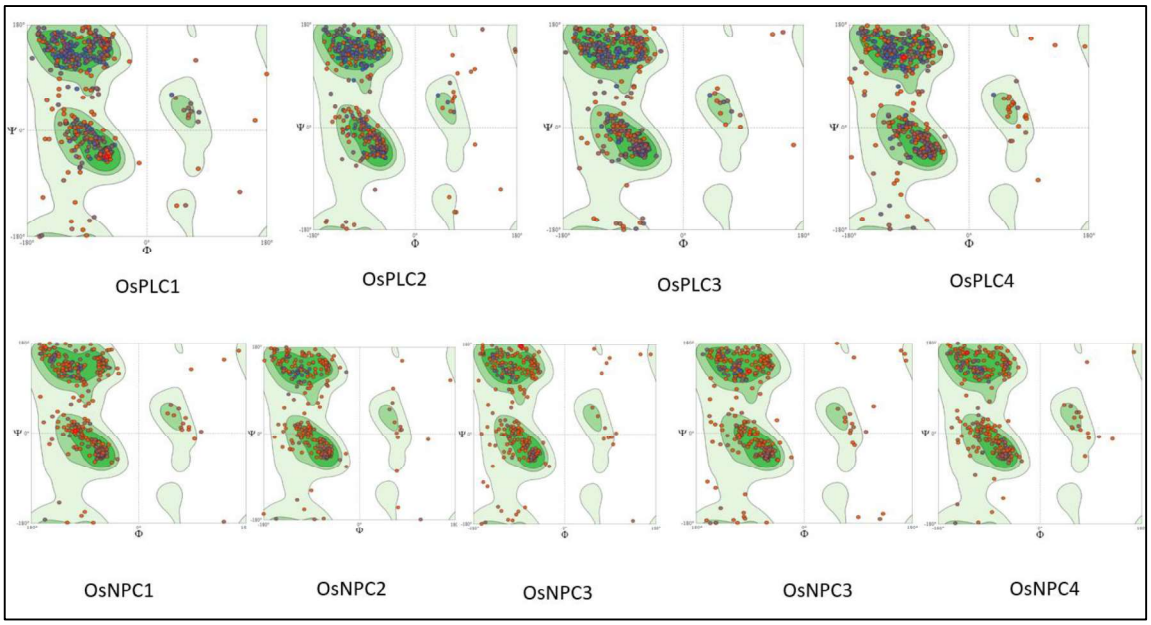


Figure 9: Ramachandran plot statistics of rice phospholipase C

3.1.3 Ligand and ligand interacting residues in rice PLCs

Swiss model also predicts calcium ion is present as a ligand in OsPLC1, OsPLC2, OsPLC3 and OsPLC4 *(*Table 11). The OsPLC1 ligand predicted to interact with amino acid residues ASN130, OsPLC2 ligand predicted to interact with amino acid residues ASN 15, ASP46, GLU 44, GLU 94. OsPLC3 predicted ligand-protein interaction sites are ASN 128, GLU 207, ASP 157, GLU 207. OsPLC4 predicted ligand-protein interaction sites are ASN 133, GLU 321, ASP 164, GLU 212 (figure 10). In the predicted NPC model, no biologically functional ligand predicted to be present.

Table 11: Ligand binding site prediction of rice phospholipase C

Model	Template	Ligand	Ligand- protein interacting residues
OsPLC1	3qr0.1. A	Calcium ion	N.130, E.159, D.161, E.209
OsPLC2	2zkm.1A	Calcium ion	N.15, E.44, D.46, E.94
OsPLC3	4qj.1B	Calcium ion	N.128, E.157, D.159, E.207
OsPLC4	3qr1.1A	Calcium ion	N.133, D.164, E.212, E.321

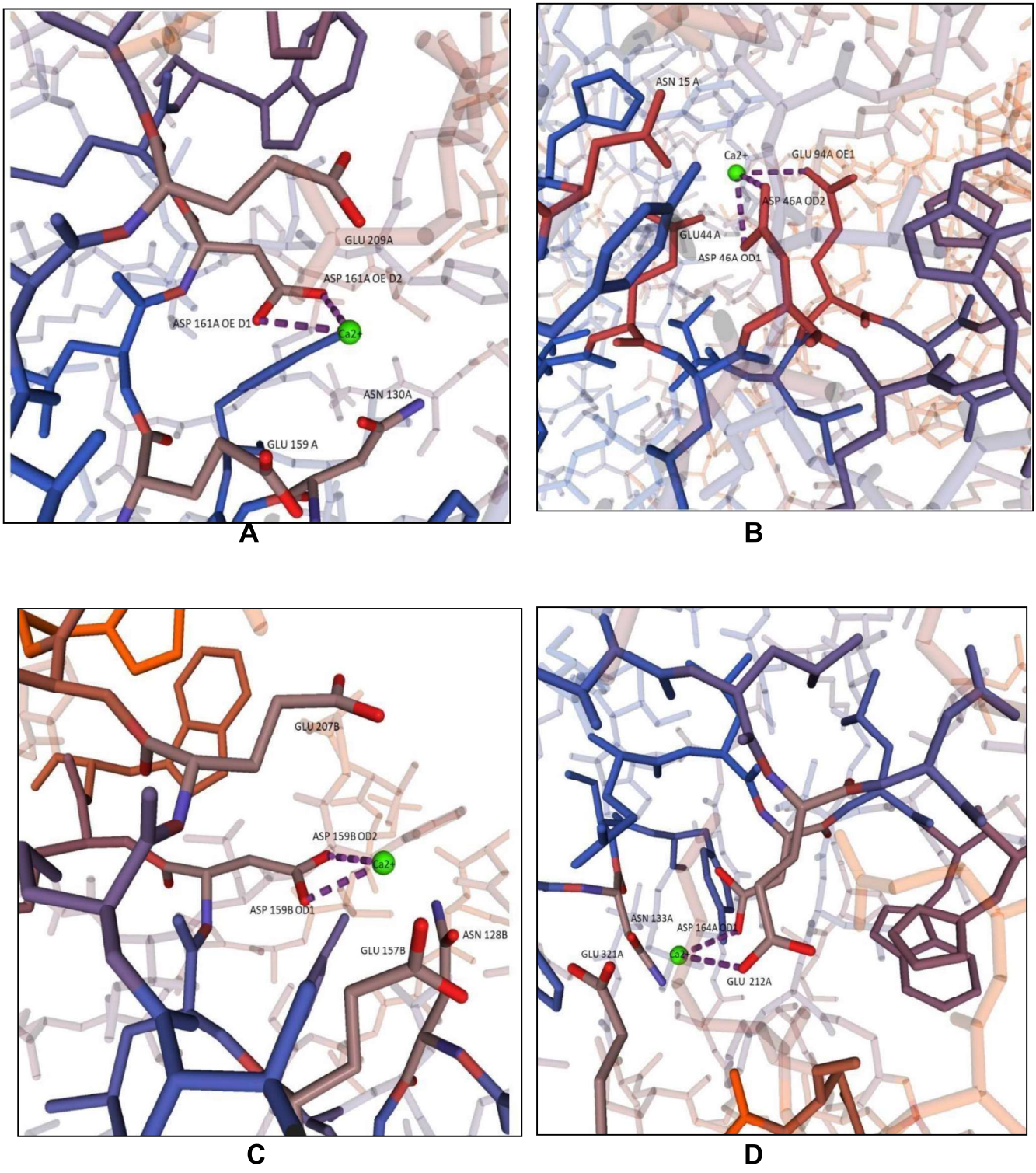


Figure 10: Predicted ligand protein interaction site in rice phospholipase C.A) OsPLC1 B) OsPLC2, C) OsPLC3 D) OsPLC4.

3.2 *In-silico* prediction of subcellular localization of rice PLCs

The intracellular localization of a protein is crucial evidence for understanding its function. The subcellular localization of OsPLCs and OsNPCs was predicted using WoLF PSORT, ChloroP and TargetP-2.0 (table 12). ChloroP and TargetP-2.0 predicts OsPLC3 doesn't harbor transit peptide for chloroplast and mitochondria localization (table 12). Also, signal peptide for secretory pathway localization was also absent. These results indicate that OsPLC3 might localize in membrane or cytosol (table 12). But the WoLF PSORT predicts both chloroplast and mitochondrial localization, only chloroplast or mitochondrial localization for OsPLC3. Also, it predicts cytoplasmic and mitochondrial localization, only cytoplasmic subcellular localization. Also, experimental evidences showed OsPLC1 and OsPLC4 were localized to cytosol and nucleus (Singh et al., 2013, Liu et al., 2017). The subcellular localization of NPCs was extensively studied in *Arabidopsis*. NPC1 is localized in secretory pathway compartments like Golgi apparatus, and ER (Krčková et al., 2015). NPC4 is localized in the plasma membrane (Nakamura et al., 2005), NPC5 is mainly present in cytoplasm, but also in ER (Gaude et al., 2008). The localization of NPC2 and NPC6 was shown to localized in plastids in mesophyll cell (Ngo et al., 2018). In roots, NPC2 localized in Golgi and a small fraction is present in ER (Krčková et al., 2018). Similarly, NPC6 is localized in the chloroplast and in the microsomal membrane when transiently expressed in tobacco mesophyll cells (Cai et al., 2020). This evidence indicates that NPC2 and NPC6 are localized in diverse subcellular localization in different tissues. However, not much is known about the rice NPCs subcellular localization.

Table 12: *In-silico* prediction for subcellular localization of rice PI-PLCs and NPCs.

Protein	Locus	WoLF PSORT	Target-2.0	ChloroP
OsPLC1	LOC_Os07g49330.1	chlo:6, mito:3.5, cyto:3, cyto_mito:2.5	Other	None
OsPLC2	LOC_Os03g18010.1	nucl: 4, cyto: 4, chlo: 2, vacu: 1, E.R.: 1, pero: 1	Other	None
OsPLC3	LOC_Os12g37560.1	chlo_mito: 5.33333, mito: 5, chlo:4.5, cyto_mito: 3.33333, cyto: 3	Other	None
OsPLC4	LOC_Os05g03610.1	mito: 7, chlo_mito: 6.83333, chlo: 5.5, cyto_mito: 4.33333	Other	None
OsNPC1	LOC_Os03g61130.1	chlo: 7, nucl: 2, cyto: 2, mito: 2	SP	None
OsNPC2	LOC_Os01g72520.1	nucl: 4, chlo: 3, mito: 3, E.R.: 2, vacu: 1	SP	None
OsNPC3	LOC_Os11g38050.1	cyto: 10, chlo: 2, nucl: 1	Other	None
OsNPC4	LOC_Os03g63580.1	cyto: 10, plas: 2, chlo: 1	Other	None
OsNPC5	LOC_Os01g01190.1	chlo:11, mito: 2	SP	None

3.3 *In-silico* analysis for predicting interacting partners of OsPLC3

In-silico analysis using STRING software predicts IPK2, guanine nucleotide-binding protein alpha-1 subunit; guanine nucleotide-binding proteins (G proteins), putative phosphatidylinositol 3-kinase, CDP-diacylglycerol-inositol 3-phosphatidyltransferase and putative kinases of PI-3/PI-4 kinase family as major interacting partners of OsPLC3 (figure 11 and table 13).

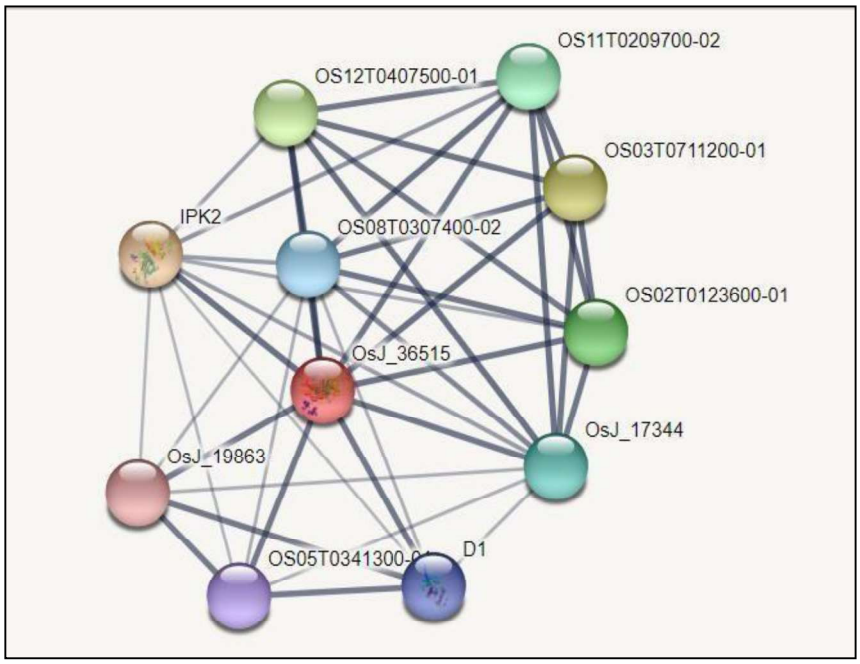


Figure 11: *In-silico* protein-protein interaction using STRING software. Network nodes represent proteins, edges represent protein-protein association (thickness of line indicate confidence of predicted interactions).

Predicted interaction of OsPLC3 with IPK2 and multiple putative kinases indicate that inositol 1,4,5-triphosphate generated by membrane phospholipid hydrolysis might be phosphorylated to produce Inositol hexaphosphate and IP_6 function as downstream signaling molecule as previously hypothesized (Chen et al., 2011). Also, G-protein shown to interact with PI-PLC for example, Gα1 can to bind to the C2 domain of PI-PLC in *Pisum Sativum* (Mishra et al., 2007). OsPLC3 also predicated to interact with G-protein (Guanine nucleotide-binding protein alpha-1 subunit, putative extra-large G-protein). These results suggest that OsPLC3 might be regulated by G-protein.

Table 13: Predicted interacting partners of OsPLC3.

Protein name	Confidence of prediction
Inositol polyphosphate multikinase IPK2	0.966
Os03g0711200 protein; Phosphatidylinositol 3-and 4-kinase family protein	0.962
Os12g0407500 protein	0.950
CDP-diacylglycerol-inositol 3-phosphatidyltransferase	0.948
Os11g0209700 protein	0.937
Os05g0180600 protein; Putative phosphatidylinositol 3-kinase	0.921
Os08g0307400 protein; Putative kinase; Belongs to the PI3/PI4-kinase	0.921
Guanine nucleotide-binding protein alpha-1 subunit; Guanine nucleotide binding proteins (G proteins)	0.919
Os06g0111400 protein; Putative extra-large G-protein	0.919

3.4 Generation of *OsPLC3* overexpression lines in rice

3.4.1 Mobilization of *OsPLC3-pGWR520* in *Agrobacterium*

The *OsPLC3* entry clone was mobilized to destination vector *pGWB520* vector (Gandhi., 2020), where it was under the control of *CaMV 35S* promoter. The vector backbone had *HygR* gene for hygromycin resistance and *SmR* gene for spectinomycin resistance. The construct was mobilized into the *Agrobacterium tumefaciens (EHA105*) strain. The transformed cells were selected on the LB agar medium containing spectinomycin (50 µg/L). The bacterial cells that grow on selection medium were further confirmed by PCR. First, plasmid was isolated from single bacterial colonies by using alkaline lysis method (Figure 12). The plasmids were screened for the presence of *OsPLC3* gene*,* this was done by using *CaMV 35S* forward primer and gene-specific reverse primer for amplification. The PCR products, resolved on 1% agarose gel and a band of size 1.5 kb band in colonies 1,2 and 5, was documented (figure 13). Positive clone was selected for the preparation of glycerol stock.

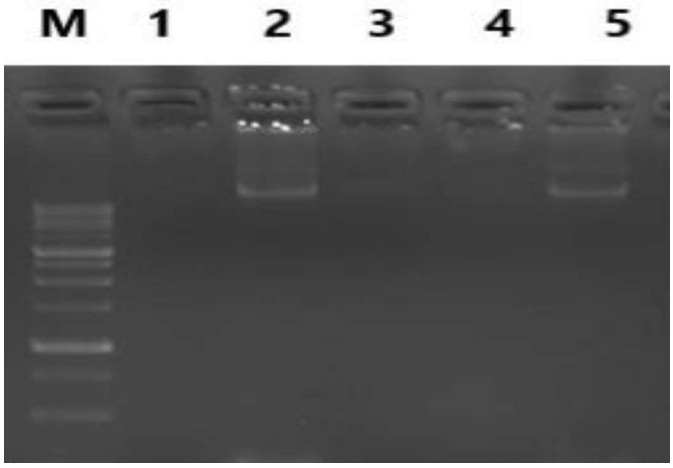


Figure 12: Plasmid isolation from transformed *Agrobacterium* colonies. The lanes 2 and 5 clearly shows presence of plasmid whereas in other lanes (1,3,4) plasmid was present in less concentrations.

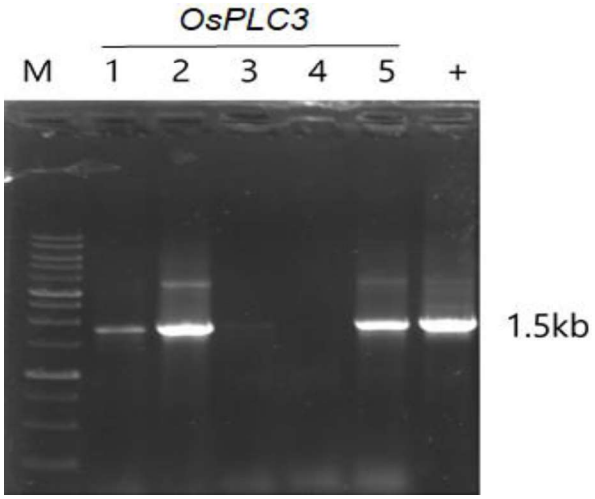


Figure 13: PCR for validation of *Agrobacterium* transformation. The transformation of *A. tumefaciens* (*EHA105*) was verified by amplification of *OsPLC3* gene. Lane M contain DNA ladder of 1 kb, lanes 1 to 5 represent different colonies. A band of 1.5 kb size clearly seen in transformed *Agrobacterium* colonies and absent in untransformed colonies. The Os*PLC3-pGWR520* plasmid was used as positive control.

3.4.2 Transformation of rice calli with *OsPLC3-pGWR520* overexpression construct to generate overexpression lines

The rice seeds (Pusa Basmati-1) Were surface sterilized and plated on callus induction media for 21 days (figure 14A). Twenty-one days old rice calli were transformed with *OsPLC3-pGWR520* construct (Figure 14B). The rice calli was infected with Agro inoculation media and was transferred on co-cultivation media for 2 days and after two days, transformed calli transferred to selection media (callus induction media supplemented with 50 µg/mL hygromycin, and 300 mg/L cefotaxime) for 14 days (Figure 14C). Every time, the black portion of calli, representing the untransformed part was discarded and only the white proliferating calli were carried forward. After two rounds of selection, the calli were transferred to regeneration medium (Figure 14D).

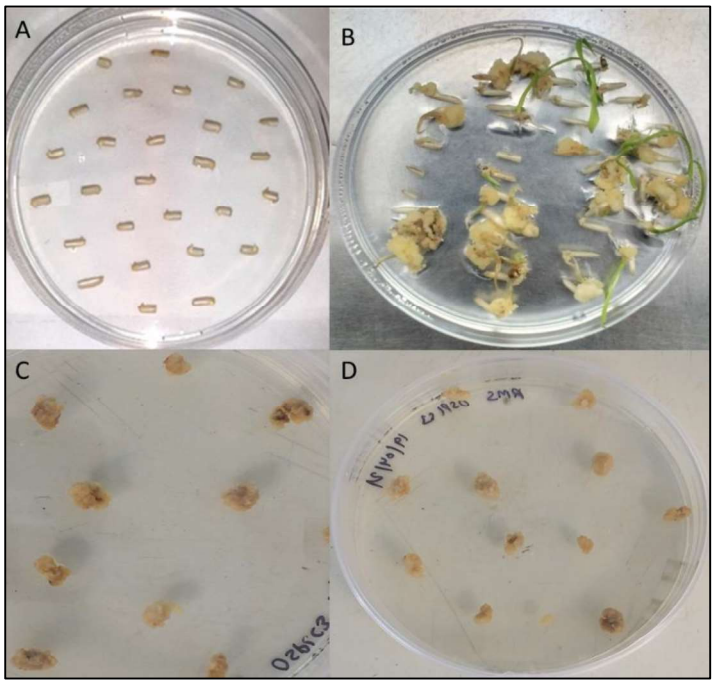


Figure 14: Generation of *OsPLC3* overexpression transgenic lines. A) Rice seed sterilization. B) Callus induction C) Selection of transformed calli D) Regeneration of transformed calli (7 Days).

3.5 Yeast two-hybrid assay to find interacting partners of OsPLC3

To identify the putative interacting partners of OsPLC3 yeast two-hybrid assay was performed. The bait OsPLC3 expressed in fusion with Gal4 DNA binding domain in the yeast expression vector *pGBKT7,* previously in lab (Gandhi., 2020). The vector expresses the *OsPLC3* gene under the control of T7 promoter and also harbors the *Kan*^{r} for kanamycin resistance. The vector was mobilized in the yeast strain *Y2H* Gold. The transformed yeast cells selected on SD/-T drop out media. Figure 15 shows the colonies were observed in yeast cell transformed with empty vector (*pGBKT7*) and vector harboring construct (*pGBKT7-PLC3*). In the negative control no colonies, were observed. The colonies growing on dropout media were confirmed with PCR. The plasmid was isolated from yeast colonies transformed with *pGBKT7-PLC3* vector (figure 16). Then, plasmid is validated for the presence of *OsPLC3* gene using gene-specific primers. The transformed yeast colonies show the 1.5 kb amplified PCR product but absent in non-transformed colonies. The positive was selected for preparation of glycerol stock, which was further used for mating and screening rice protein library.

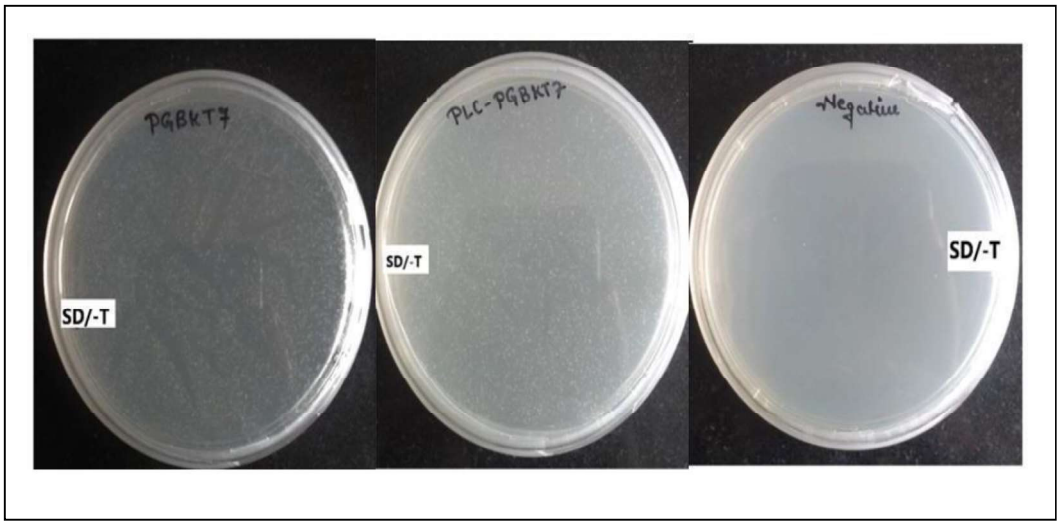


Figure 15: Transformation of *OsPLC3-pGBKT7* bait construct in Y2H GOLD. A) Positive control B) *OsPLC3-pGBKT7 C)* Negative control.

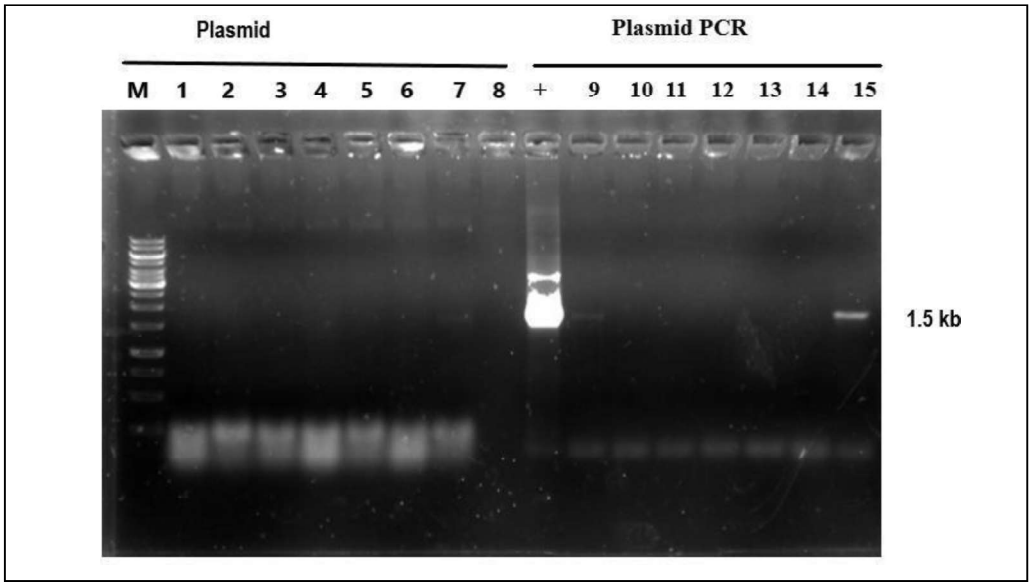


Figure 16: Plasmid isolation and PCR verification of yeast cell transformed with *pGBKT7-OsPLC3* construct. Lanes 1 to 8 represent plasmid isolated by different yeast colonies. The PCR verified by gene specific primer. Lane 9 to 15 represents different colonies. Band of 1.5 kb observed in transformed colonies which was absent in non-transformed colonies. The *OsPLC3-TOPO* plasmid was used as positive control. Lane M contains the 1kb DNA ladder.

For finding the putative, interacting partners the bait strain culture mated with rice library strain Y187. After mating for 20 hrs., the successful mating between bait and prey strain was confirmed by presence of diploids (figure 17). Mated culture was plated on SD/- LT media and incubated at 30ºC for 3-4days. After 4 days colonies were patched onhigher stringency SD/-HLT plates.

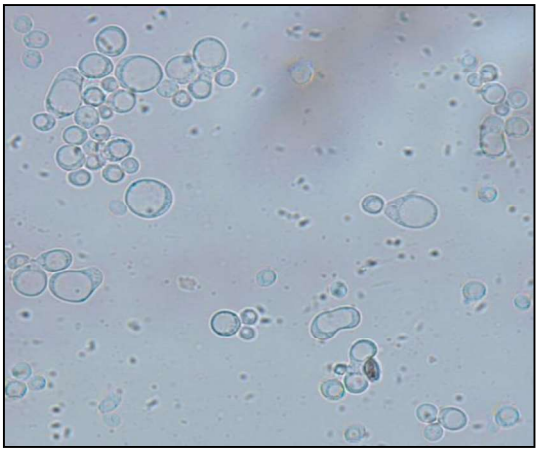


Figure 17: Yeasts mating and screening for diploids. The bait strain (*Y2H* GOLD mated with rice library *(Y187*). Three lobed structures represent the zygotes. The two lobes represent two haploid parents giving rise to diploid zygote. The mated culture visualized under Olympus light microscope at 40X magnification.

3.6 Subcellular localization of OsPLC3.

The *OsPLC3* gene cloned in *pENTR/D-TOPO* previously in the lab (Gandhi., 2020) was mobilized into destination vector *pSITE2CA* by LR Clonase reaction. The LR reaction mix was transformed into *DH5α*. The colonies grew on selection plates, further validated by PCR. PCR was performed using *pSITE2CA,* forward primer and gene-specific reverse primer. The results showed non-specific bands of 1kb and 750 bp (figure 18).) To further confirm, the cloned plasmid was isolated from transformed colonies and digested with *ApaI.* No release of insert was observed, indicating the absence of *OsPLC3* gene in the *pSITE2CA* plasmid (figure 19). The results indicate some problem in the LR Clonase or performing the PCR verification reactions.

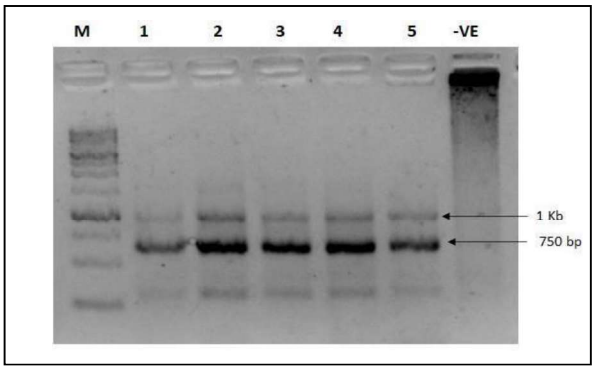


Figure 18: Cloning of *OsPLC3* in *pSITE2CA* GFP vectors. *OsPLC3* was mobilized to GFP destination vector *pSITE2CA* by LR Clonase enzyme through homologous recombination. Gel picture showing PCR amplification of non-specific bands indicating absence of desired gene.

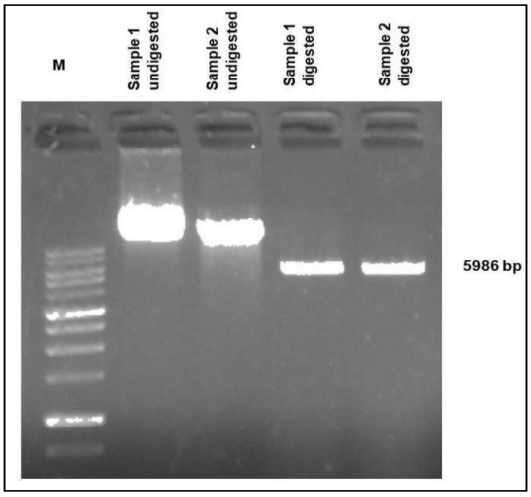


Figure 19: Cloning of *OsPLC3* in *pSITE2CA* GFP vectors. *OsPLC3* was mobilized to GFP destination vector *pSITE2CA* by LR Clonase enzyme through homologous recombination. The plasmids isolated from colonies that grew on selection plates. The lane 1 and 2 were loaded with undigested sample. Lane 3 and 4 loaded with sample digested with *Apa I*.

3.6 Floral dip method for *Arabidopsis thaliana* to generate OsPLC3 *overexpression lines*

To generate of *OsPLC3* overexpression lines in *Arabidopsis thaliana Col-0* plants were transformed with *OsPLC3-pGWR520* construct by floral dip. The transformed plant was kept in dark for 16-18 hrs. and then transferred to light condition and watered until seed setting. The harvesting of seeds to be done yet.

3.7 Cloning of *OsPLC3* in *pGEX4T-1* to generate for protein expression construct

Full length ORF of *OsPLC3* was cloned in fusion with GST tags of *pGEX4T-1.* The cloning was validated by PCR and band shift assay. PCR was done using *NotI* forward primer and *SalI* reverse primer. The gel image shows multiple, non-specific bands. (Figure 20). For further validation plasmid isolated from colonies were observed for band-shift. No band shift was observed (Figure 21) indicating absence of desired gene.

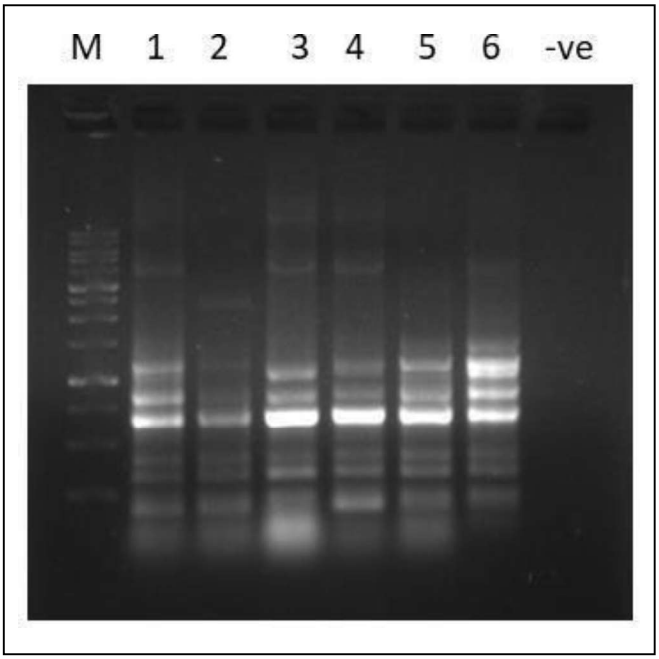


Figure 20: PCR validation of cloning of *OsPLC3* in *pGEX4T-1* vector.

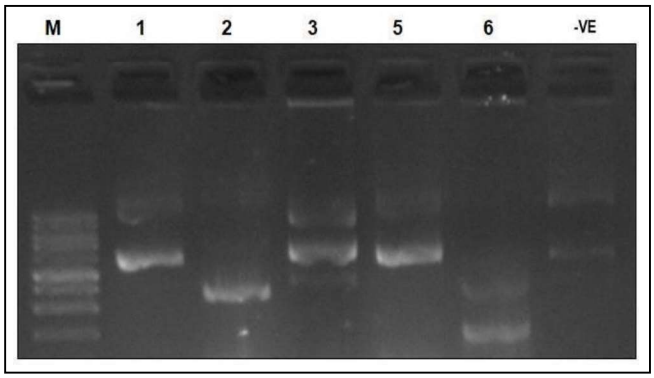
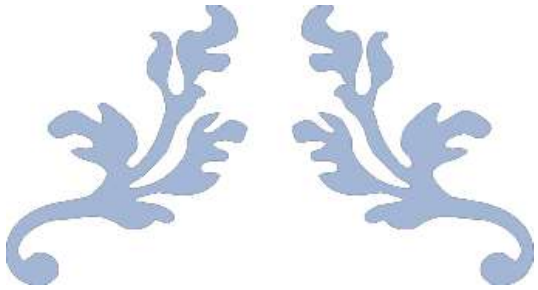
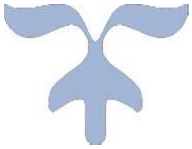


Figure 21: Confirmation of cloning of OsPLC3 in pGEX4T-1 by band shift assay. Gel picture showing isolated plasmid from colonies grew on selection. Lanes 1 to 6 indicates plasmid samples. No gel shift was observed in sample indicating absence of *OsPLC3* gene insert in the *pGEX4T-1* plasmid construct. *pGEX4T-1* used as a negative control and M represents 1kb DNA ladder.



CHAPTER 4

CONCLUSION AND FUTURE WORK



Conclusion and future work

Genome-wide expression profiling of rice PLC families under abiotic stresses such as salinity, cold, drought identified *OsPLC3* expressed differentially and significantly under these stress conditions (Singh et al., 2013). Here, I attempted to functionally characterize the *OsPLC3* gene to establish its role in lipid-mediated signaling under stress condition. To elucidate the function of *OsPLC3* in rice, *OsPLC3* over-expression plants, need to be generated in rice and work is in progress to achieve this objective. I have transformed the rice calli with *OsPLC3-pGWR520* construct and white transformed calli were shifted to regeneration medium for shoot regeneration. The transgenic plants generated by rice tissue culture will be confirmed by gene-specific PCR and RT- qPCR. Finally, phenotypic analysis of the transgenic plants will be done under salt, drought and cold stress conditions.

Spatial and temporal subcellular localization gives important clues about the function of the protein. *In-silico* studies with WoLF PSORT, TargetP-2.0 and ChloroP1.1 predicted the localization of the OsPLC3 in cytoplasm, mitochondria, cytosol. To experimentally validate its subcellular localization, I attempted to mobilize *OsPLC3* cloned in *pENTR/DTOPO (*Gandhi., 2020) into destination vector *pSITE2CA* but was not successful in cloning this gene in the *pSITE2CA* vector. This work is in progress to achieve this objective.

The three-dimensional protein structure was built using homology modelling, using the Swiss model to understand structural aspects of rice PLCs.The structure of *OsPLC1, OsPLC2, OsPLC3* and *OsPLC4* was built using structure of *Sepia officinalis* PLC21 (3qr0.1.A), phospholipase C beta 2 (2zkm.1A), human phospholipase C-beta3 delta (94qj.1B,), *Loligo pealei* PLC21 (3qr1.1A) as template respectively. The structure of *Francisella tularensis* Acid Phosphatase A (AcpA) bound to orthovanadate (2d1g.1.A) was used to build a structural model for the OsNPCs. Ramachandran plot analysis of OsPLCs shows the amino acids residues are present in favored region indicating the reliability of the models. *In-silico* studies with STRING software predicts IPK2, guanine nucleotide-binding protein alpha-1 subunit; guanine nucleotide-binding proteins (G proteins), putative phosphatidylinositol3-kinase, CDP-diacylglycerol-inositol 3-phosphatidyltransferase, putative kinases as major interacting partners of OsPLC3. To find the interacting partner of OsPLC3 *yeast* two-hybrid assay was performed. I mobilize the construct bait in Y2H gold and mated it with the rice library strain Y187. The

colonies, which grew on SD/-LT media will be patched on higher stringency SD/-HLT plates. Finally, the colonies will be selected by plating on higher stringency selection SD/-AHLT. Plasmids will be isolated from the colonies, that grow on the higher stringency conditions and will be sent for sequencing to identify the interacting protein. Also, I made several attempts to clone *OsPLC3* in the *pGEX4T-1* vector for protein expression but could not achieve this objective. This work will be undertaken in near future.

BIBLIOGRAPHY

Bibliography

Abd-El-Haliem AM, Vossen JH, van Zeijl A, Dezhsetan S, Testerink C, Seidl MF, Beck M, Strutt J, Robatzek S, Joosten M. (2016). Biochemical characterization of the tomato phosphatidylinositol-specific phospholipase C (PI-PLC) family and its role in plant immunity. Biochim Biophys Acta 1861:1365–13.

Bose D, Ngo AH, Van Nguyen C, Nakamura Y. (2021). Non-specific phospholipases C2 and 6 redundantly function in pollen tube growth via triacylglycerol production in Arabidopsis. Plant J
BY-2 cells and whole Arabidopsis seedlings. Plant J 52 1014-26.

De Wald DB, Torabinejad J, Jones CA, Shope JC, Cangelosi AR, Thompson JE, Prestwich GD, Hama H. (2001). Rapid accumulation of phosphatidylinositol 4,5-bisphosphate and inositol 1,4,5-trisphosphate correlates with calcium mobilization in salt-stressed Arabidopsis. Plant Physiology 126: 759–769.

Cai, G., Fan, C., Liu, S., Yang, Q., Liu, D., Wu, J. et al. (2020). Nonspecific phospholipase C6 increases seed oil production in oilseed Brassicaceae plants. New Phytologist, 226, 1055-1073.

Chen G, Snyder CL, Greer MS, Weselake RJ. (2011). Biology and biochemistry of plant phospholipases. Critical Reviews in Plant Sciences 30: 239–258.

Chen, X, Li, L, Xu, B, et al. (2019). Phosphatidylinositol-specific phospholipase C2 functions in auxin-modulated root development. *Plant Cell Environ*.; 42: 1441– 1457.

Deng X, Yuan S, Cao H et al. (2019). Phosphatidylinositol-hydrolyzing phospholipase C4 modulates rice response to salt and drought. Plant Cell Environ 42:536–548.

Dowd PE, Coursol S, Skirpan AL, Kao Th, Gilroy S. (2006). Petunia phospholipase C1 is involved in pollen tube growth. Plant Cell18: 1438–1453.

Emanuelsson O, Nielsen H, von Heijne G (2019). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites Protein Sci. 8(5):978-84.

Gandhi, A. (2020). Functional Characterization of rice phospholipase C. Dissertation, University of Delhi.

Gaude N, Nakamura Y, Scheible W-R, Ohta H, Dörmann P (2008). Phospholipase C5 (NPC5) is involved in galactolipid accumulation during phosphate limitation in leaves of Arabidopsis. Plant J 56:28–39.
Georges F, Das S, Ray H, Bock C, Nokhrina K, et al. (2009). Over-expression of Brassica napus phosphatidylinositol-phospholipase C2 in canola induces significant changes in gene expression and phytohormone distribution patterns, enhances drought tolerance and promotes early flowering and maturation. Plant Cell Environ 32: 1664–1681.

Gietz, R., Schiestl, R. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2, 31–34.

Gu, Y., Zhao, C., He, L. et al. *(2018).* Genome-wide identification and abiotic stress responses of *DGK* gene family in maize. *J. Plant Biochem. Biotechnol.* **27,** 156–166.

Helling D, Possart A, Cottier S, Klahre U, Kost B. (2006). Pollen tube tip growth depends on plasma membrane polarization mediated by tobacco PLC3 activity and endocytic membrane recycling. Plant Cell 18: 3519–3534.

Hong Y, Zhao J, Guo L et al. (2016). Plant phospholipases D and C and their diverse functions in stress responses. Prog Lipid Res 62:55–74.

Huang CH, Tate BF, Crain RC, Cote´ GG. (1995). Multiple phosphoinositide specific phospholipases C in oat roots. Characterization and partial purification. Plant J 8: 257–268.

Huang CH, Crain RC. (2009). Phosphoinositide-specific phospholipase C in oat roots: association with the actin cytoskeleton. Planta 230:925–33.

J.E. Vermeer, J.M. Thole, J. Goedhart, E. Nielsen, T. Munnik, T.W. Gadella, Jr. (2009). Imaging phosphatidylinositol 4-phosphate dynamics in living plant cells. Plant J 57,356-72.

Jones JD, Dangl JL. (2006). The plant immune system. Nature 444: 323–329.

Jose Juan Almagro Armenteros, Marco Salvatore, Ole Winther, Olof Emanuelsson, Gunnar von Heijne, Arne Elofsson, and Henrik Nielsen. (2019). Detecting Sequence Signals in Targeting Peptides Using Deep Learning Life Science Alliance 2 (5).

Kanehara K, Yu CY, Cho Y, Cheong WF, Torta F, Shui G, Wenk MR, Nakamura Y. (2015). Arabidopsis AtPLC2 is a primary phosphoinositide-specific phospholipase C in phosphoinositide metabolism and the endoplasmic reticulum stress response. PLoS Gene 11: e1005511.

Kim YJ, Kim JE, Lee JH, Lee MH, Jung HW, et al. (2004). The Vr-PLC3 gene encodes a putative plasma membrane-localized phosphoinositide-specific phospholipase C whose expression is induced by abiotic stress in mung bean (Vigna radiata L.). FEBS Lett 556: 127–136.

Kocourková D, Krčková Z, Pejchar P, Veselková Š, Valentová O, Wimalasekera R, Scherer GFE, Martinec J. (2011). The phosphatidylcholine hydrolyzing phospholipase C NPC4 plays a role in response of Arabidopsis roots to salt stress. J Exp Bot 62:3753– 3763.

Krčková Z, Brouzdová J, Daněk M, Kocourková D, Rainteau D, Ruel- land E, Valentová O, Pejchar P, Martinec J. (2015). Arabidopsis non-specific phospholipase C1: characterization and its involvement in response to heat stress. Front Plant Sci 6:928.

Li L, Wang F, Yan P et al. (2017). A phosphoinositide-specific phospholipase C pathway elicits stress-induced Ca2+ signals and confers salt tolerance to rice. New Phytol 214:1172–1187.

Liu HT, Liu YY, Pan QH, Yang HR, Zhan JC, Huang WD. (2006). Novel interrelationship between salicylic acid, abscisic acid, and PIP2-specific phospholipase C in heat acclimation-induced thermotolerance in pea leaves. J Exp Bot 57:3337–47.

Liu X, Ma D, Zhang Z et al. (2019). Plant lipid remodeling in response to abiotic stresses. Environ Exp Bot 165:174–184.

Mikami K, Repp A, Graebe-Abts E, Hartmann E. (2004). Isolation of cDNAs encoding typical and novel types of phosphoinositide-specific phospholipase C from the moss Physcomitrella patens. J Exp Bot 55: 1437–1439.

Misra S, Wu Y, Venkataraman G, Sopory SK, Tuteja N. (2007). Heterotrimeric G-protein complex and G-protein-coupled receptor from a legume (Pisum sativum): role in salinity and heat stress and cross-talk with phospholipase C. Plant J.

Munnik T, Testerink C. (2009). Plant phospholipid signaling: “In a nutshell”. Journal of Lipid Research 50 (suppl.): S260–S265.

Munnik T. (2014). PI-PLC: Phosphoinositide-phospholipase C in plant signaling. In: Wang X, ed. Phospholipases in plant signaling. Berlin, Germany: Springer, 27–54.

Munnik, T. and Nielsen, E. (2011). Green light for polyphosphoinositide signals in plants. Curr. Opin. Plant Biol. 14: 489–497.

Nakamura Y. (2013). Phosphate starvation and membrane lipid remodeling in seed plants. Prog Lipid Res 52:43–50.

Nakamura Y (2014). NPC: nonspecific phospholipase cs in plant functions. In: Wang X (ed) Phospholipases in plant signaling. Springer, Berlin, Heidelberg, pp 55–67.

Nakamura Y, Awai K, Masuda T, Yoshioka Y, Takamiya K, Ohta H. (2005). A novel phosphatidylcholine-hydrolyzing phospholipase C induced by phosphate starvation in Arabidopsis. J Biol Chem 280:7469–7476.

Ngo AH, Kanehara K, Nakamura Y. (2019). Non-specific phospholipases C, NPC2 and NPC6, are required for root growth in Arabidopsis. Plant J 100:825–835.

Ngo AH, Lin YC, Liu YC, Gutbrod K, Peisker H, Dörmann P, Nakamura Y. (2018). A pair of nonspecific phospholipases C, NPC2 and NPC6, are involved in gametophyte development and glycerolipid metabolism in Arabidopsis. New Phytol 219:163–175.

Pan YY, Wang X, Ma LG, Sun DY. (2005). Characterization of phosphatidylinositol-specific phospholipase C (PI–PLC) from Lilium daviddi pollen. Plant Cell Physiol 46:1657–1665.

Pejchar P, Martinec J. (2015). Aluminium ions alter the function of non-specific phospholipase C through the changes in plasma membrane physical properties. Plant Signal Behav 10: e1031938.

Pejchar P, Potocký M, Krčková Z, Brouzdová J, Daněk M, Martinec J. (2015). Non-specific phospholipase C4 mediates response to aluminium toxicity in Arabidopsis thaliana. Front Plant Sci 6:66.

Pejchar P, Potocký M, Novotná Z, Veselková Š, Kocourková D, Valentová O, Schwarzerová K, Martinec J. (2010). Aluminium ions inhibit the formation of diacylglycerol generated by phosphatidylcholine hydrolysing phospholipase C in tobacco cells. New Phytol 188:150–160.

Peters C, Kim SC, Devaiah S, Li M, Wang X. (2014). Non-specific phospholipase C5 and diacylglycerol promote lateral root development under mild salt stress in Arabidopsis. Plant Cell Environ 37:2002–2013.

Peters C, Li M, Narasimhan R, Roth M, Welti R, Wang X. (2010). Non- specific phospholipase C NPC4 promotes responses to abscisic acid and tolerance to hyperosmotic stress in Arabidopsis. Plant Cell 22:2642–2659.

Pokotylo I, Kolesnikov Y, Kravets V, Zachowski A, Ruelland E. (2014). Plant phosphoinositide-dependent phospholipases C: variations around a canonical theme. Biochimie 96:144–157.

Pokotylo I, Kravets V, Martinec J, Ruelland E. (2018). The phosphatidic acid paradox: Too many actions for one molecule class? Lessons from plants. Prog Lipid Res 71:43–53.

Pokotylo I, Pejchar P, Potocký M, Kocourková D, Krčková Z, Ruelland E, Kravets V, Martinec J. (2013). The plant non-specific phospholipase C gene family. Novel competitors in lipid signalling. Prog Lipid Res 52:62–79.

Rupwate, S.D., Rajasekharan, R. (2012). C2 domain is responsible for targeting rice phosphoinositide specific phospholipase C. *Plant Mol Biol* 78, 247–258.

Sagar S, Biswas DK, Singh A. (2020). Genomic and expression analysis indicate the involvement of phospholipase C family in abiotic stress signaling in chickpea (*Cicer arietinum*). Gene 753:144797.
salicylic acid, abscisic acid, and PIP2-specific phospholipase C in heat acclimation-induced thermotolerance in pea leaves. J Exp Bot 57:3337–47.

Shaoping Lu, Tarig Fadlalla, Shan Tang, Long Li, Usman Ali, Qing Li, Liang Guo. (2019). Genome-Wide Analysis of Phospholipase D Gene Family and Profiling of Phospholipids under Abiotic Stresses in *Brassica napus*, *Plant and Cell Physiology*, Volume 60, Issue 7, Pages 1556–1566.

Shi J, Gonzales RA, Bhattacharyya MK. (1995). Characterization of a plasma membrane-associated phosphoinositide-specific phospholipase C from soybean. Plant Journal 8: 381–390.

Singh A, Bhatnagar N, Pandey A, Pandey GK. (2015). Plant phospholipase C family: regulation and functional role in lipid signaling. Cell Calcium 58:139–146.

Singh A, Kanwar P, Pandey A, Tyagi AK, Sopory SK, Kapoor S, Pandey GK. (2013). Comprehensive genomic analysis and expression profiling of phospholipase C gene family during abiotic stresses and development in rice. PLoS ONE 8: e62494

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J., & von Mering, C. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic acids research*, *43*(Database issue), D447–D452.
Tan X, Calderon-Villalobos LIA, Sharon M, Zheng C, Robinson CV, Estelle M, Zheng N. (2007). Mechanism of auxin perception by the TIR1 ubiquitin ligase. Nature 446: 640–645.

Van Wijk R, Zhang Q, Zarza X et al. (2018). Role for Arabidopsis PLC7 in stomatal movement, seed mucilage attachment, and leaf serration. Front Plant Sci 9:1721.

Vossen JH, Abd-El-Haliem A, Fradin EF, van den Berg GC, Ekengren SK, Meijer HJ, Seifi A, Bai Y, ten Have A, Munnik T et al. (2010). Identification of tomato phosphatidylinositol-specific phospholipase-C (PI-PLC) family members and the role ofPLC4 and PLC6 in HR and disease resistance. Plant Journal 62: 224–239.

W. van Leeuwen, J.E. Vermeer, T.W. Gadella, Jr., T. Munnik. (2007). Visualization of phosphatidylinositol 4, 5-bisphosphate in the plasma membrane of suspension-cultured tobacco BY-2 cells and whole Arabidopsis seedlings. Plant J 52 1014-26.

Wang CR, Yang AF, Yue GD, Gao Q, Yin HY, et al. (2008). Enhanced expression of phospholipase C1 (ZmPLC1) improves drought tolerance in transgenic maize. Planta 227: 1127–1140.
Wang L, Zhu X, Liu J, Chu X, Jiao J, Liang Y (2013) Involvement of phospholipases C and D in the defense responses of riboflavin-treated tobacco cells. Protoplasma 250: 441–449

Wang F, Deng Y, Zhou Y, Dong J, Chen H, Dong Y, et al. (2015). Genome-Wide Analysis and Expression Profiling of the Phospholipase C Gene Family in Soybean (*Glycine max*). PLoS ONE 10(9): e0138467.

Wang, K., Li, Yl. & Chen, S. (2020). Genome-wide identification of phospholipase C related to chilling injury in peach fruit. *J. Plant Biochem. Biotechnol.*

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Research.

Wimalasekera R, Pejchar P, Holk A, Martinec J, Scherer GFE. (2010). Plant Phosphatidylcholine hydrolyzing phospholipases C NPC3 and NPC4 with roles in root development and brassinolide signaling in Arabidopsis thaliana. Mol Plant 3:610–625.

Wu L, Sadhukhan A, Kobayashi Y et al. (2019). Involvement of phosphatidylinositol metabolism in aluminium-induced malate secretion in Arabidopsis. J Exp Bot 70:3329–3342.

X. Wang, Plant Phospholipases. (2001). Annu Rev Plant Physiol. Plant Mol Biol 52 211-231.

Xia K, Wang B, Zhang J, Li Y, Yang H, Ren D. (2017). Arabidopsis phosphoinositide-specific phospholipase C 4 negatively regulates seedling salt tolerance. Plant Cell Environ 40:1317–1331

Yang B, Li M, Phillips A et al. (2021). Nonspecific phospholipase C4 hydrolyzes phosphosphingolipids and sustains plant root growth during phosphate deficiency. Plant Cell.

Yang, B., Zhang, K., Jin, X., Yan, J., Lu, S., Shen, Q., Guo, L., Hong, Y., Wang, X. and Guo, L. (2021). Acylation of non-specific phospholipase C4 determines its function in plant response to phosphate deficiency. Plant Journal.

Zhang B, Wang Y, Liu JY. (2018a). Genome-wide identification and characterization of *phospholipase C* gene family in cotton (*Gossypium* spp.). Sci China Life Sci 61:88–99

Zhang Q, van Wijk R, Shahbaz M, Roels W, Schooten BV, Vermeer JEM, Zarza X, Guardia A, Scuffi D, Garcia-Mata C, Laha D, Wiliams P, Willems LAJ, Ligterink W, Hoffmann-Benning S, Gillaspy G, Schaaf G, Haring MA, Laxalt AM, Munnik T. (2018b). Arabidopsis phospholipase C3 is involved in lateral root initiation and ABA responses in seed germination and stomatal closure. Plant Cell Physiol 59:469–486.

Zhang Q, van Wijk R, Zarza X, Shahbaz M, van Hooren M, Guardia A, Scuffi D, Garcia-Mata C, Van den Ende W, Hoffmann-Benning S, Haring MA, Laxalt AM, Munnik T. (2018c). Knock-down of Arabidopsis *PLC5* reduces primary root growth and secondary root formation while overexpression improves drought tolerance and causes stunted root hair growth. Plant Cell Physiol 59:2004– 2019.

Zheng SZ, Liu YL, Li B, Shang ZL, Zhou RG, Sun DY (2012). Phosphoinositide-specific phospholipase C9 is involved in the thermotolerance of Arabidopsis. Plant Journal 69: 689–700.

Zhu J-K. 2002. Salt and drought stress signal transduction in plants. Annual Review of Plant Biology 53: 247–27.

APPENDIX

Rice Phospholipase C Protein Sequences.

>LOC_Os07g49330.1 protein

MGTYKCCLIFKRRFRWNDAPPPDDVRALFANHSAGGGPHMAADGLRAYLQATGQDGDVDM
ERLVEQIRQLQGRGGRIPRVGRALPLLTVDDFHRFLFSHELNPPIRHGQGQVHHDMAAPL
SHYFIYTGHNSYLTGNQLSSDCSDLPIIRALQRGVRVIELDMWPNSSKDDISILHGRTLT
TPVSLLKCLLSIKQHAFEASPYPVIITLEDHLTPDLQDKAAKMVLEVFGDILYYPDKDHL
KEFPSPQDLKGRVLLSTKPPREYLQAKDGNAATIKEDAKAAATDDAAWGKEVPDIHSQIH
SATKHDQREDDDDTDEDEDDEEEEQKMQQHLAPQYKHLITIKAGKPKGTLLDALQSDPEK
VRRLSLSEQQLAKLADHHGTEIVRFTQRNLLRIYPKGTRVTSSNYNPFLGWVHGAQMVAF
NMQGYGRALWLMHGFYKANGGCGYVKKPDFLMQTDPEVFDPKKSLSPKKTLKVKVYMGDG
WRMDFTQTHFDQYSPPDFYARVGIAGVPADSVMKRTRAIEDNWVPVWEEDFTFKLTVPEI
ALLRVEVHEYDMSEKDDFGGQTVLPVSDLIPGIRAVALHDRKGIKLNNVKLLMRFEFE*

>LOC_Os03g18010.1 protein

MNAPLSHYFIYTGHNSYLTGNQLSSDCSDIPIIKALQIGVRVIELDMWPNSSKDDVDILH
GRTLTAPVSLIKCLKSIKEYAFVASPYPVIITLEDHLTSDLQAKVAKMVLEVFGDTLYYP
ESKHLQEFPSPEALRGRVILSTKPPKEYLESKGGTMKDRDIEPQFSKGQNEEAVWGTEVP
DIQDEMQTADKQHENDILYTQRDVEEDDEKKMCQHHPLEYKHLITIKAGKPKGAVVDALK
GDPDKVRRLSLSEQELAKVAAHHGRNIVSFTHKNLLRIYPKGTRFNSSNYNPFLGWVHGA
QMVAFNMQGYGRSLWLMHGFYKANGGCGYVKKPDFMMQTCPDGNVFDPKADLPVKKTLKV
KVYMGEGWQSDFKQTYFDTYSPPDFYAKVGIAGVPSDSVMQKTKAVEDSWVPVWEEEFVF
PLTVPEIALLRVEVHEYDVSEDDFGGQTALPVSELRPGIRTVPLFDHKGLKFKSVKLLMR
FEFV*

>LOC_Os12g37560.1 protein

MGSYAYKYCMCFTRKFRSPAADPPPDVRAAFLAAGGGDGGLRRFLAQAQGETPAEVDRIL
ALLSGGGGGGGGGGIAARLVGPRPGPAPSLDDFFGFLFNADLNPPIATQVHQDMSAPFSH
YYIYTGHNSYLTGNQLNSDSSDIPIIKALQRGVRVIELDMWPNSAKNNIDILHGGTLTAP
VQIIKCLKSIKEYAFCASPYPLVITLEDHLTPDLQAKVAEMLVKTFGNLLYIPSSDPINE
FPSPESLMKKIIISTKPPEEYKKFLKSKDNQNINGGLANLAEEGSLRRIDSNAEESDGKD
ELDDQDEDSSDEDDPKFQQETACEYRELITIHAGKPKGHLKDALKVDPDKVRRLSLSETQ
LAKATASHGADVIRFTQKNILRVYPKGTRINSSNYDPMNAWTHGAQMVAFNMQGHDKALR
LMQGFFRANGGCGYVKKPDFLLRTGPNGEVFDPNASMPVKKTLKVKVYMGDGWRMDFSKT
HFDTFSPPDFYTRVGIAGVRADCVMKKTRTIEDQWVPMWDEEFTFPLTVPELAVLRIEVH
EYDMSEKHDFGGQTCLPVSELKQGIRAVPLHDRRGTRYKSVRLLMRFDFL*

>LOC_Os05g03610.1 protein

MTTYRVCCFLRRFRAASNEPSEELGDVFQAYADGGGGVMGEEALRRFLREVQGEAAGGGD

DELEATAREVMAFAAEQRLLRKGGAAAAGGGLTVEGFHRWLCSDANAALDPQKRVYQDMG
LPLSHYFIYTGHNSYLTGNQLSSGCSEVPIVKALHDGVRVIELDLWPNAAKDAVEVLHGR
TLTSPVGLMKCLEAIREYAFVASPYPVILTLEDHLTPDLQSKVAKMIKETFGDMLYVSET
ENMAEFPSPDELKGKIIVSTKPPKEYLQTKNDADADEAGVWGEEITDDKVAATAMTTEEK
CAAAEEAVAAAAVDEEMQEAETDKKTQHGVDNEYRRLIAIPLTRRKHDMDQDLKVDPDMV
TRLSLGEKAYEKAIVTHGAHIIRFTQRKLLRIFPRSTRITSSNYNPLMGWRYGVQMVAAN
MQGHGRKLWLTQGMFRANGGCGYVKKPDILMNNDPDKLFDPTSKLPVKTRLKVTVYMGDG
WRFDFRKTHFDKCSPPDFYARVGIAGVEADTRMEQTKVKMDTWIPAWDHEFEFPLSVPEL
ALLRVEVHESDNHQKDDFGGQTCLPVWELRRGIRSVRLCDHRGEPLRSVKLLMRFDFT*

>LOC_Os03g61130.1 protein

MAGGGGRERRGGGRLLVGVLLLTLVVSGHCLESTHHRGLKRRRRKHEIHSPIKTVVVVVM
ENRSFDHILGWLSRTRPDIDGLNGTQSNRLNASDPSSPEIFVTDEAGYVDSDPGHGFEDI
REQIFGSADTSAVPAPMSGFAQNARGMGLGMPQNVMSGFKPESVPVYAALADEFAVFDRW
FASVPTSTQPNRLYVHSATSHGLTFNARKDLIHGFPQKTIFDSLEENGLSFGIYYQNIPA
TLFYQSLRRLKHLVKFHQYSLKFKLHAKWGKLPNYAVIEQRYFDCEMFPANDDHPSHDVA
RGQRFVKEVYETLRASPQWNETALIITYDEHGGFYDHVPTPVVGVPQPDGIVGPDPYYFK
FDRLGVRVPSFLISPWIEKRTVIHEPNGPQDSSQYEHSSIPATVKKLFNLHSNFLTKRDA
WAGTFENYFKIRKTPRTDCPEKLPEVTKSLQPFGPKEDSSLSEFQVELIQLASQLNGDHV
LNTYPDIGRTMTVGEANRYAEDAVARFLEAGRIALRAGANESALVTMRPALTSRASPSSD
LSSEL*

>LOC_Os01g72520.1 protein

MAVRRRRPGPVAAAVLLLLAVATQAAASPIKTVVVVVMENRSFDHMLGWMKRLNPEIDGV
TGGEWNPTNASDPSSGRVYFGEGAEYVDPDPGHSFQEIRQQIFGSDDASGPARMDGFVQQ
ARSLGDNMTAAVMNGFSPDSVAVYRELVGEFAVFDRWFASVPSSTQPNRLFVHSATSGGA
TSNNPELLAKGYPQRTIFDNVHDAGLSFGVYYQDVPAVLFYRNLRKLKYLTKFHPFHGAF
RDHAARGSLPNYAVVEQHYMDSKSHPANDDHPSHDVFQGQMLVKEVYETLRASPQWNQTL
MVVTYDEHGGFYDHVPTPVTGVPSPDGIVGPPPYNFAFDRLGVRVPAIVISPWINKGTVV
HGPNGSPTATSEYEHSSIPATVKKLFDLPQDFLTKRDAWAGTFESVVQGRTEPRTDCPEQ
LPMPMRIRLTEANEEAKLSEFQQELVQLASVLNGDHQLSSLQDTIRDRMNVREGIAYMRG
AVKRFFETGMSAKRMGVDDEQIVKMRPSLTTRTSPAIEQP*

>LOC_Os11g38050.1 protein

MAGKIKTVVVLVQENRSFDHMLGWMKSLNPEIDGVTGAEFNRATAGDPSSPAIHFGDGSG
YVDPDPGHSFQAIYEQVYGDAYTWGTTSPATKPGVPSPPMSGFAQEAEKERAGMSSTVMN
GFRPEKVPVYRELVREFAVCDRWFASVPTSTQPNRMFVHSATSHGLVSNDGKQLRAGLPQ
RTIFDALHDAGHSFGVYYQFPPSVLFYRNMRQLKYVGNFHPYDTAFKRDCKAGKLPNYVV
IEQRYFDLKLLPGNDDHPSHDVAHGQRLVKDVYEALRSSPQWHEILFVITYDEHGGFFDH

VPTPVAGVPSPDGIVSAAPVSFAFDRLGVRVPTLLVSPWIEPGTVVHDPASCGGAPEPTS
QFEHSSIPATVKRIFGLKEFLTRRDAWAGTFDTVLTRAAPREDCPATLPEPPRLRAAEAE
EHREVSEFQAELVQLGAALNGDHDGEGYDPEVFVRGMTVAGAAQYCRDAFDRFREECHRC
RDGGMDGSHVPMLQPASASASSSAPAADPPATAPAPATPSALSKLCGCFPCFNAS*
>LOC_Os03g63580.1 protein
MAAAAAGGKIKTVVVLVMENRSFDHMLGWMKSLNPEIDGVTGDEINHLDAADPTSRAIRF
GDGAEYVDPDPGHSMQAIYEQVYGTPFVDARATPITPPGVPSPPMAGFAQQAEKEKPGMA
DTVMNGFRPEAVPVYRELVRQFAVCDRWFASNPASTQPNRLFVHSATSHGLVSNDTKLLV
AGLPQRTIFDSLHDAGFSFGIYYQYPPSTLFYRSLRQLKYAGNFHPFDLAFRRHCAEGKL
PNYVVVEQRYFDLKMLPGNDDHPSHDVSEGQRFVKEVYEALRGGPQWEEALLVVTYDEHG
GFYDHVPTPVDVPSPDGIVSAAPFFFEFNRLGVRVPALFISPWIEPGTVVHRPSGPYPTS
EFEHSSIPATVKKLFNLKSFLTNRDAWAGTFDVVLTRDAPRTDCPATLPEPVKMRPATEA
AEQAALTEFQEELVQLGAVLNGDHADEDVYPRKLVEGMTVAEAASYCNAAFKAWMDECDR
CRKCGEDGSHIPTVVKPPPPPSTSSSGSSSFASKLLSCFACGRPNKN*
>LOC_Os01g01190.1 protein
MWPLSPHRRHLQQHGHSGAMGRRLLLLFLMLAQAPNSNGDSKIKNVVVLALENRSFDHML
GWMQRLLGLPIDGLTGAECNPAPGPGPADSLLHCVSPDADLVVPDDPAHAFEDVLEQLLG
FRPNDSTGAAASPSDMSGFVRSAVSVSALLTDAVMRGFTPSRLPAFSALASSFAVFDRWF
SSIPGPTQPNRLFLYSATSHGAVAHDKWNLLRGYPQRTIFDSLAADALDYRVYFKTIPTT
LFYRRLRTVANAARGTFRRYDAAFRDHARRGLLPALSVIEPRYFDLTGTPADDDHPAHDV
ANGQRLVKDVYEALRAGPQWNHTLLIITYDEHGGFYDHVPPPNVGVPSPDAIRGPLPFFF
RFDRLGVRVPTIMVSPWIRKGTVVGRPPGGPTPTSEYEHSSIPATIKKIFNLSSDFLTRR
DAWAGTFEHLFTDLDEPRTDCPETLPEIPPPSSSSSSTKKEDGGWLSDFQRELVQLAAFL
NGDYMLSSFAQEYESRMTMTVKQADAYVRRAVKSFLEASKRAKRLGANDSAIVTMRPSLT
TATTCCP*

Heat Sensing Mechanisms in Lower Organisms

Thesis Submitted to the University of Delhi

in the Partial fulfilment for the Degree of

Master of Science

in

Plant Molecular Biology and Biotechnology

2021



GOPIKA UNNI P J

Department of Plant Molecular Biology

University of Delhi South Campus

New Delhi 110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Heat Sensing Mechanism in Lower Organisms**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: June 28, 2021

Signature of the Candidate

(**Gopika Unni P J**)

Signature of the Supervisor

(**Prof. Paramjit Khurana**)

Signature of the HOD

(**Prof. Indranil Dasgupta**)

PREFACE

The thesis focuses on the various mechanisms adopted by lower organisms to sense temperature fluctuations in the environment and mount a suitable response. Temperature is one of the most important parameter that is monitored by free-living microorganisms and, any fluctuations in temperature can cause the microbes to alter their physiology in response to it. Each stressful situation can induce a typical stress response, which ultimately changes the pattern of gene expression so as to restore homeostasis in the cell. It is the presence of various biomolecules within a cell that makes it possible for the organism to sense such changes, thereby acting as thermosensors. Such responses thus ensures that the drastic fluctuations in the temperature are not lethal for the organism.

Structural lability in response to temperature changes is a broadly acting mechanism that underlies the thermosensing functions of a wide range of molecules. Among, the various biomolecules present, the membrane lipids, which is the major constituent of the plasma membrane, changes its lipid composition in response to temperature shifts, thereby acting as a thermosensor. Nucleic acids, both DNA and RNA, being made of thermally labile hydrogen bonds that joins complementary bases provides a satisfactory exposition as to why they behave as a thermosensor. As proteins are complex biomolecules that are comprised of thousands of interacting atoms, they too are highly susceptible to experiencing significant conformational changes in response to fluctuations in temperature. This review shall provide an insight on the various mechanisms by which biomolecules act as thermosensors.

ACKNOWLEDGEMENTS

First of all, praises and thanks to Almighty for his showers of blessings in these difficult times and nothing would not have been possible without his blessings. Next to God, I would like to express my deep and sincere gratitude to my research supervisor *Professor Paramjit Khurana* for her invaluable guidance and support, and encouraging me constantly with her optimistic views and ideas. It is indeed a great privilege to work under her guidance.

I am truly honored to have been able to pursue this M.Sc. degree under the excellent teaching and guidance of our eminent professors *Prof. Anil Grover, Prof. J.P. Khurana, Prof. Akhilesh K. Tyagi, Prof. Indranil Dasgupta, Prof. Arun Sharma, Prof. Sanjay Kapoor, Prof. Giridhar Pandey, Prof. Saurabh Raghuvanshi* and *Prof. Surekha Katiyar Agarwal.* The perception that I have developed towards molecular biology is the fruitful result of continuous efforts that our teachers made and for that I would be always thankful to them.

I am extremely thankful to my mentor Ms. Harsha Samtani, who guided me at each and every step of for this thesis. She was always there whenever I needed her help and was indeed a constant source of support. She reviewed my work regularly and gave me valuable suggestions.

This journey would not have been complete without my M.Sc. classmates, Aashmeen, Apurva, Athul, Thanduanlung, Karishma, Nidhi, Nitika, Shitij and Sritama with whom I spent most of my journey, and for helping and encouraging me in whatever ways they could. I would also like thank the Department and University of Delhi for providing me with ample learning opportunities.

Finally, I would like to thank my family for being always there for me, and without whose support, amidst this pandemic, this would not have been possible.

ABBREVIATIONS

- AT- Adenine Thymine
- CCW-Counter clockwise
- Csp- Cold shock protein
- *css* -encoding capsule biosynthesis
- CW-Clockwise
- DNA -Deoxyribonucleic Acid
- ECF- Extra cytoplasmic Factor
- fourU- Four Uridine
- H-NS - Heat-stable Nucleoid-Structuring
- HSP-Heat Shock Protein
- MCP- Methyl-accepting Chemotaxis Proteins
- mRNA- Messenger Ribonucleic Acid
- PAGE- Polyacrylamide Gel Electrophoresis
- RBS - Ribosome Binding Site
- RNA – Ribonucleic Acid
- RNAT- RNA Thermometer
- ROS-Reactive Oxygen Species
- ROSE – Repression Of heat Shock gene Expression
- SD – Shine-Dalgarno
- T3SS -Type III Secretion System
- T6SS – Type VI Secretion System
- UTR- Untranslated Region

Table of Contents

INTRODUCTION **1-3**

MEMBRANE LIPIDS **4-6**

Acyl-lipid desaturase 4

LpxD1 and LpxD2 5

DNA **6-10**

Supercoiling 6

Promoter curvature 7

Nucleoid-associated proteins 8

Copy number 9

RNA **10-18**

ROSE elements 11

fourU elements 13

Cyanobacterial thermometer 13

Css thermometer 14

Alternate sigma factor 15

Csp thermometer 16

cIII thermosesnor 17

Intra- and Inter-molecular RNA-RNA interactions 18

FRQ 18

PROTEINS **18-24**

Transcriptional regulators 19

Methyl-accepting chemotaxis proteins 21

Chaperones and Protease 21

MISCELLANEOUS **24-29**

TRP Channels 25

Kinases 25

Flagellar motor 27

Thermal Avoidance 27

Type VI Secretion System 28

CONCLUSION AND FUTURE PROSPECTS **30-31**

BIBLIOGRAPHY **32-35**

REVIEW OF LITERATURE
HEAT SENSING MECHANISM IN LOWER ORGANISMS

INTRODUCTION

Temperature is one of the most important parameters that affects the growth and development of an organism (Phadtare & Inouye, 2008). Any change in the environmental temperature can affect the homeostatic state of any organism (Digel et al., 2008). The ability of an organism to sense the environmental temperature is fundamental for its ability to regulate its body temperature (Kortmann & Narberhaus, 2012). The stability of various proteins, membranes, RNA species and cytoskeleton structures are affected and also alters the efficiency of enzymatic reactions inside the cell (Digel et al., 2008). Thus, in order to survive, organisms have to sense any change in the temperature so as to adapt their biological processes accordingly (Aguilar et al., 2001).Hence all organisms, from a single celled prokaryote to a multi cellular complex eukaryote, have an inherent property known as thermosensing (Sengupta & Garrity, 2013). Therefore, thermosensing refers to the ability of an organism to detect and codify both environmental and internal temperature changes and hence is indispensable for their very survival (García-Ávila & Islas, 2019).

For mounting temperature specific responses, various sensory mechanisms are present in the single cell itself which can translate the change in temperature into a suitable response (Digel et al., 2008). Usually, the sensory mechanisms contain a cascade of reactions that lead to the appropriate response. In most cases, there is a primary sensor which undergoes a conformational change in response to temperature that activates the cascade of reactions. This primary sensor is usually a biomolecule that translates the change in temperature into a suitable biological response (Sengupta & Garrity, 2013). Upon encountering a change, the sensory molecule triggers a series of signalling events , which ultimately results in activation of genes specific for this stimuli (Digel et al., 2008). So basically, these systems function as a thermostat that helps in performing sensory signal transductions that integrate and decode physical and chemical stimuli in the environment (Digel et al., 2008).

Amongst the various mechanisms which are temperature controlled, the most well-known ones are the expression of heat-shock genes and cold shock genes

(Narberhaus et al., 2006). Even though the initial launching mechanisms differ , the general response of all organisms from prokaryotes to plants and higher eukaryotes to sudden temperature shifts, be it heat or cold shock, are similar (Digel et al., 2008). Usually, the response after a temperature stress is the elite and rapid expression of the small group of proteins called the heat-shock and cold-shock proteins (Digel et al., 2008).

In the lower organisms, like the bacteria, a heat-shock response invokes a few heat shock proteins whose primary function is to deal with and alleviate the damaging effects caused by the temperature shift. Several of these proteins are involved in reconstituting and stabilising protein structures and removing misfolded proteins (Hazel, 1995). Primary sensing event is a bit more obscure during cold shock. There have been reports suggesting that lowering of temperature increases the rigidity of the membrane which further affects the membrane associated functions (Digel et al., 2008). However, lesser drastic temperature changes may not induce shock genes, instead can modulate the expression of certain other genes like the virulence genes in pathogenic fungi and bacteria (Digel et al., 2008). On the other hand, plants unlike the lower organisms experience daily and seasonal temperature fluctuation that vary in range, rate and duration. They hence respond to different temperature ranges with widely divergent physiological and developmental responses (Zhang et al., 2015). At the molecular level, temperature stress affects the protein stability, membrane fluidity, cytoskeletal organization, transport and enzymatic reactions, which leads to metabolic imbalances and pernicious accumulation of by-products such as reactive oxygen species (ROS) (Hayes et al., 2020).

Such mechanisms even though are widespread in nature, studies have been significantly done in the lower organisms like bacteria, fungi and viruses (Klinkert & Narberhaus, 2009). It has been found out that various biomolecules that function as major thermosensors are; Membrane lipids, DNA, RNA, and proteins. The plasma membrane which is composed of lipid bilayer are extremely sensitive to temperature. A sudden drop in the temperature reduces the membrane fluidity, which in turn inhibits normal membrane functioning (Vigh et al., 1998). Unless this is corrected rapidly, this can lead to physiological damage and ultimately, death. Hence, the membrane takes it upon itself to sense any temperature change and activates suitable genes. According to the central dogma, the genetic information is transferred from DNA via RNA to protein. Hence, anything that affects the DNA topology, ultimately impacts the gene

expression (Klinkert & Narberhaus, 2009). mRNAs apart from carrying the coding information for protein synthesis, can also act as regulators of expression of the encoded message. They detect changes in the cellular environment either directly or through interactions with proteins or small non-coding RNAs, thus, acting as thermosensors (Schumann, 2007). Most sensory pathways in organisms utilise the change in protein conformation as the primary perceptive event, which further activates signalling cascades as mentioned before (Digel et al., 2008). However, temperature sensors are not limited to these biomolecules (Murata et al., 2000).

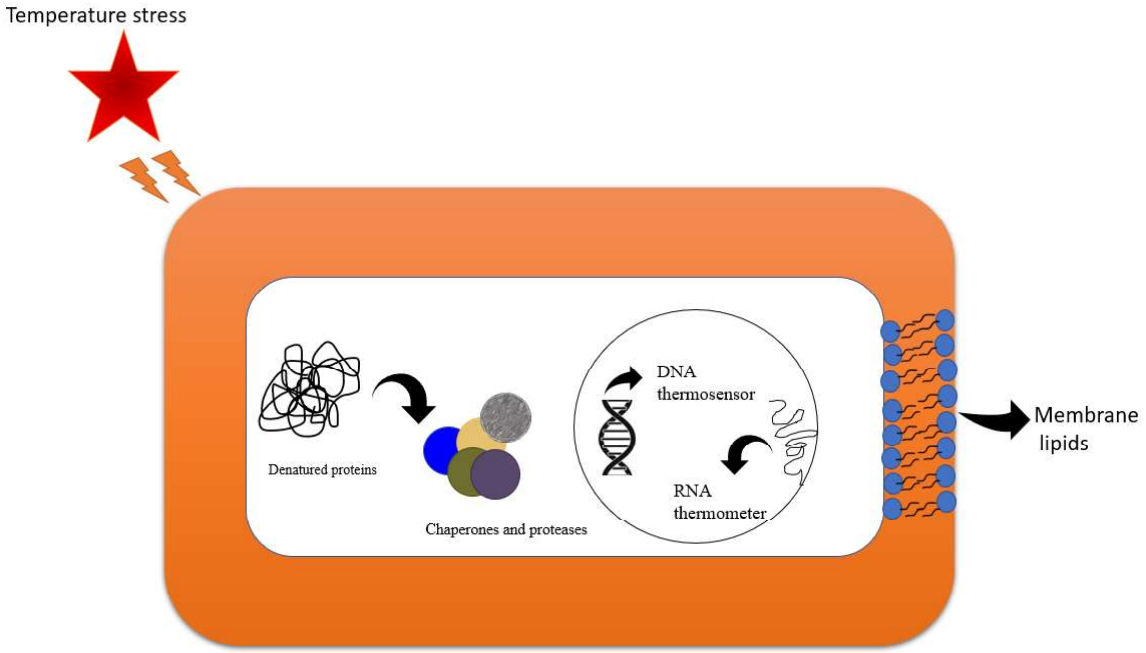


Figure 1. Major thermosensing mechanisms. The major biomolecules that acts as thermosensors in lower organisms are membrane lipids, DNA, RNA and Proteins.

1. Membrane lipids

Membranes provide the structural framework that separates the cells from their environment. In a wide range of microorganisms, it is the plasma membrane that first senses any change in the environmental temperature . Rather than just being simply the passive barriers that are liable to be damaged during environmental challenge, they are also involved in cellular responses and modulates intracellular signalling (Vigh et al., 1998). Depending on the shift in temperature, the membrane properties gets altered, which influences the expression of several genes that respond to such temperature changes, thereby acting as thermosensors.

1.1 Acyl -lipid desaturase

It has been observed that in variety of microorganisms, the level of unsaturation gets altered in response to temperature fluctuation (Digel et al., 2008). For example,, in the cyanobacterium *Synechocystis,* a drop in the temperature causes the organisms to increase the production of the enzyme acyl -lipid desaturases, which helps in increasing the level of cis-unsaturation of the membrane-bound lipid fatty acids (Digel et al., 2008). Thus, the membrane fluidity lost during the temperature drop is regained. This is due to the presence of four acyl-lipid desaturases, DesA, DesB, DesC, and DesD. Further, it has been reported that *desC* is expressed constitutively, whereas transcription of *desA, desB,* and *desD* is induced only when temperature decreases (Klinkert & Narberhaus, 2009) .

Another similar mechanism is observed in the case of the bacterium *Bacillus subtilis.* Here, the lipid modification is brought about by the two-component regulatory system that contains the two proteins; DesK and DesR (Aguilar et al., 2001). DesK is a multi-pass transmembrane histidine kinase that helps in maintaining the membrane fluidity at low temperatures (Sengupta & Garrity, 2013). DesK acts as a kinase at cold temperature (below 30°C) as it auto phosphorylates a conserved histidine residue within the kinase domain, and then transfers the phosphoryl group to aspartate in DesR; thereby remodelling the membrane fluidity. Such an event activates *des* gene that codes for the enzyme acyl lipid desaturase (Cybulski et al., 2010). Reports suggests that greater *des* transcription ensures more desaturase activity, i.e., greater amount of unsaturated fatty acids and hence, an increase in the membrane fluidity (Klinkert & Narberhaus, 2009).

Altering the lipid composition is not just limited to prokaryotes. Green algae, which is a eukaryote, too survives the fluctuating environmental conditions by this virtue. At lower temperature, the reduced membrane fluidity is counteracted by the action of the enzyme fatty acid desaturase (Sakamoto & Murata, 2002). Algal cells increase the content of membrane unsaturated fatty acids (UFAs) so as the lipids in the membrane are more loosely packed, thereby ensuring that solidification of the membrane doesn't occur at lower temperature (Lyon & Mock, 2014). Increase in the unsaturated fatty acid content in the membrane have also shown to stabilise the photosynthetic machinery under cold stress as well as salinity stress (Allakhverdiev et al., 2001). Conversely, raise in temperature increases the membrane fluidity due to the presence of membrane-stabilising proteins and by synthesizing saturated fatty acids (SFA) *den novo* which gets integrated into the membrane (Los et al., 2013). During the stress conditions it has been observed that cells show an increase in fatty acid synthesis. The key enzyme that controls fatty acid biosynthesis rate is ACCase and is present in almost all algae (Podkowinski and Tworak, 2011; Singh et al., 2016) . ACCase is comprised of four subunits aacA, accD, accC and accB. It was further seen in the algae *Chlorella sorokiniana* that there was an up-regulation of ACCase subunits along with the high ACCase activity and fatty acid synthesis (Barati et al., 2019).

1.2 LpxD1 and LpxD2

Lipopolysaccharide (LPS) is an endotoxin that is the major constituent of the outer membrane of Gram-negative bacteria (Raetz et al., 2007). Lipid A is the biologically active component of the LPS that can trigger an immune response in the host as it is the only component recognized by the immune system (Kusumoto et al., 2010). There are certain organisms, that can synthesize modified forms of lipid A in response to change in temperature (Wang & Quinn, 2010). Such a modification alters the bacterium's outer membrane's integrity, thereby protecting itself from being susceptible to the host's immune responses (Li et al., 2012). For example, the bacteria *Fransciella*, while entering a mammalian host remodels its membrane by altering lipid A (Phillips et al., 2004; Shaffer et al., 2007). The enzyme responsible for bringing about such a modification is acyltransferase LpxD1/2. LpxD1 enzyme adds longer acyl chains that are usually 18 carbons in length at the 2 and 2′ positions of lipid A, whereas LpxD2 adds shorter acyl chains(16 carbons) at the same positions (Shapiro & Cowen, 2012). It was reported by Li et al. (2012) that *lpxD1* gene was up-regulated at mammalian

temperature and *lpxD2* at environmental temperature. Moreover, they also found out the enzymatic activity of LpxD1 was higher at 37°C while LpxD2 was active at a much lower temperature i.e., 21°C. Thus, regulation of LpxD1/2 at both transcriptional as well as enzymatic level appears to be temperature dependent, and this helps the organism in remodelling its membrane fluidity in response to the change in the environment (Li et al., 2012).

2. DNA

DNA just like other biomolecules, also respond to change in temperature (Digel et al., 2008). It is the presence of the thermally labile hydrogen bonds which join complementary base pairs in nucleic acids that makes them sensitive to temperature (Sengupta & Garrity, 2013). Temperature affects the topology and integrity of DNA that in turn results in affecting gene expression (Shapiro & Cowen, 2012). DNA can act as a thermosensor in the following manner:

2.1 Supercoiling

DNA supercoiling is one of the many parameters affected by temperature. DNA topology influences transcription and, thus any change in the supercoiling pattern can result in a rapid and efficient cellular mechanism so as to alter the gene expression during stress conditions (Lopez-Garcia & Forterre, 2000; Pruss & Drlica, 1989) . It has been observed that plasmids from mesophiles and hyperthermophiles undergoes a reversible change in topology during temperature stress. In case of mesophilic bacteria like *Escherichia coli* and *Salmonella* it has been shown that the plasmid DNA is negatively supercoiled under unstressed conditions and heat stress induces a transient increase in positive supercoiling, leading to plasmid relaxation, while cold shock leads to a transient decrease in supercoiling (Kataoka et al., 1996; Mizushima et al., 1997) . But in the case of hyperthermophiles, like *Sulfolobus* and *Thermococcus*, DNA has been reported to be positively supercoiled, and higher temperature (temperature ≥ 80°C) increased positive supercoiling, while colder temperatures induced negative supercoiling(Shapiro & Cowen, 2012). Such a switch in conformation is brought about by enzymes like DNA gyrase and topoisomerase I, the nucleoid-binding protein HU and the molecular chaperone DnaK (Klinkert & Narberhaus, 2009).

Studies have shown that similar kind of mechanism exists in the cyanobacterium *Synechocystis* also, where low temperature increased the supercoiling of the genomic DNA region that contained the regulatory elements of the desB gene (Los, 2004). The desB gene codes for the omega-3 desaturase, a fatty acid desaturase that controls membrane fluidity by introducing double bonds into the acyl chains of lipid-bound fatty acids (Klinkert & Narberhaus, 2009).

Since, transcription efficiency is also very sensitive to changes in DNA topology, and thus DNA supercoiling is an indispensable factor in the thermoregulation of gene expression. Like in the case of *Sulfolobus*, at lower temperatures (≤ 75°C) polymerase was found to unable to overcome the thermodynamic barrier posed by an overwound DNA template due to the excessive negative supercoiling. However, the transcription was not affected at higher temperatures (Klinkert & Narberhaus, 2009).

2.2 Promoter curvature

Even the local DNA structures can affect the DNA topology. It has been shown that intrinsically curved DNA regions characterized by AT rich tracts are often located on the upstream of the promoter region (Mizushima et al., 1997). Moreover, these further affect the binding capability of RNA polymerase due to the topological changes caused by temperature fluctuations in these regions, ultimately affecting gene expression. In the bacterium *Clostridium perfringens*, *plc* gene encodes for the enzyme phospholipase C (Katayama et al., 1999). The RNA polymerase binds to the minor grooves of three A tracts present between -66 and -40 position of the transcription start site of the *plc* gene (Klinkert & Narberhaus, 2009). At low temperatures, increased bending of these A tracts enhanced the binding affinity for the polymerase, thereby facilitating open complex formation and hence, transcription (Katayama et al., 1999) .

Another example wherein promoter curvature controls gene expression was shown in enteropathogen *Shigella flexneri,* which are capable of penetrating replicating within human colonic epithelial cells (Labrec et al., 1964). A study conducted by Maurelli et al. (1984), established the dependence of temperature and bacterial invasion. The study showed that the bacteria grown at 37°C were virulent and was capable of invading tissue culture cells, but, at 30°C, they were avirulent and failed to invade both *in vitro* and *in vivo*. Such a loss in virulence was reversed by the subsequent growth of the bacteria at 37°C (Maurelli et al., 1984). This property of

virulence was due to *virF* gene that codes for the primary regulator of invasion functions in the pathogen and its transcription was found to be temperature dependent (Shapiro & Cowen, 2012). Interestingly, the *virF* promoter was characterized by the presence of a major DNA bend halfway between two H-NS (Heat-stable Nucleoid Structuring) binding sites (Prosseda et al., 2004). It was observed that at temperatures below 32°C, the *virF* promoter was in a closed conformation which kept the two H-NS binding sites sufficiently close to each other to favour a stable nucleoprotein bridge and hence, transcription of *virF* gene did not take place (Prosseda et al., 2004). But at temperatures greater than 32°C, the intrinsic bent melted abruptly, thereby forming a stable open complex for transcription and subsequently expressing the virF proteins, thereby causing virulence. (Shapiro & Cowen, 2012). So, the curvature of the promoter itself could act as a thermosensor.

2.3 Nucleoid- associated proteins

Nucleoid associated proteins can bring about temperature-induced conformational changes in DNA. Their binding not only affects its conformation, but also processes like DNA replication, transcription and recombination (Dame, 2005). A well-studied type of nucleoid-associated protein is the H-NS protein, which mostly acts as a negative modulator; in response to different environmental signals including temperature (Schumann, 2007). H-NS preferentially binds to AT-rich sequences and is itself is under the control of temperature. These proteins are not able to form higher-order oligomers and bind with DNA at temperature above 37°C (Ono et al., 2005). But at lower temperature, the ratio of H-NS proteins to DNA increases three-to-four-fold (Atlung & Ingmer, 1997). Temperature-modulated accessibility of promoter regions occupied by H-NS at low temperature plays a key role of virulence gene expression in many human pathogens (Schumann, 2012).

It has been reported that *Salmonella enterica,* a pathogen, has 77% of the temperature regulated genes that are under the control of the H-NS proteins directly or indirectly (Ono et al., 2005). An example of H-NS dependent thermoregulation is SPI-2, which is a horizontally acquired genomic island, which encodes for a type III secretion system (T3SS). It was observed that when bacteria were grown at a temperature less than 30°C, they were unable to express T3SS despite the presence of the inducible signals (Klinkert & Narberhaus, 2009). Moreover, Hha and H-NS were

the two nucleoid-like proteins which repressed the expression of virulence genes at non-permissive temperature by strongly binding with the promoter (Duong et al., 2007). But at higher temperature, the DNA binding capability of H-NS and Hha was affected. This opened up the promoter for the RNA polymerase to bind and thus the transcription occurred (Klinkert & Narberhaus, 2009).

Another case wherein nucleoid-associated proteins regulates gene expression is Pap pili, in uropathogenic *Escherichia coli,* which is encoded by the pyelonephritis-associated pili (*pap*) operon. They facilitate the attachment to uroepithelial cells and subsequent colonization of the host upper urinary tract (Göransson & Uhlin, 1984). Transcription of pap pili was shown to be temperature regulated (Göransson & Uhlin, 1984), wherein the expression was maximum at 37°C. At lower temperature (23°C and below), the level of transcripts was greatly reduced. The transcription was found to be controlled by two proteins H-NS and RimJ (Schumann, 2012). H-NS prevented transcription at the low temperature by binding within the pap regulatory region at 23°C but not at 37°C. RimJ is an N-terminal acetyltransferase of the ribosomal protein S5 and deletion of the *rimJ* gene lead to the loss of thermoregulation resulting in equivalent *papBA* transcript levels at both 37°C and 23°C (White-Ziegler & Low, 1992).The mechanism by which RimJ repressed *papBA* transcription is still not well understood.

2.4 Copy number

Some bacteria carry a plasmid DNA that has a constant plasmid copy number at one temperature and a much higher or totally uncontrolled copy number at a different temperature (Digel et al., 2008) For example, the plasmid pL088 was found in *Escherichia coli* (HB101), showed an excessive increase in the copy number at temperatures greater than 37°C (Lupski et al., 1986).

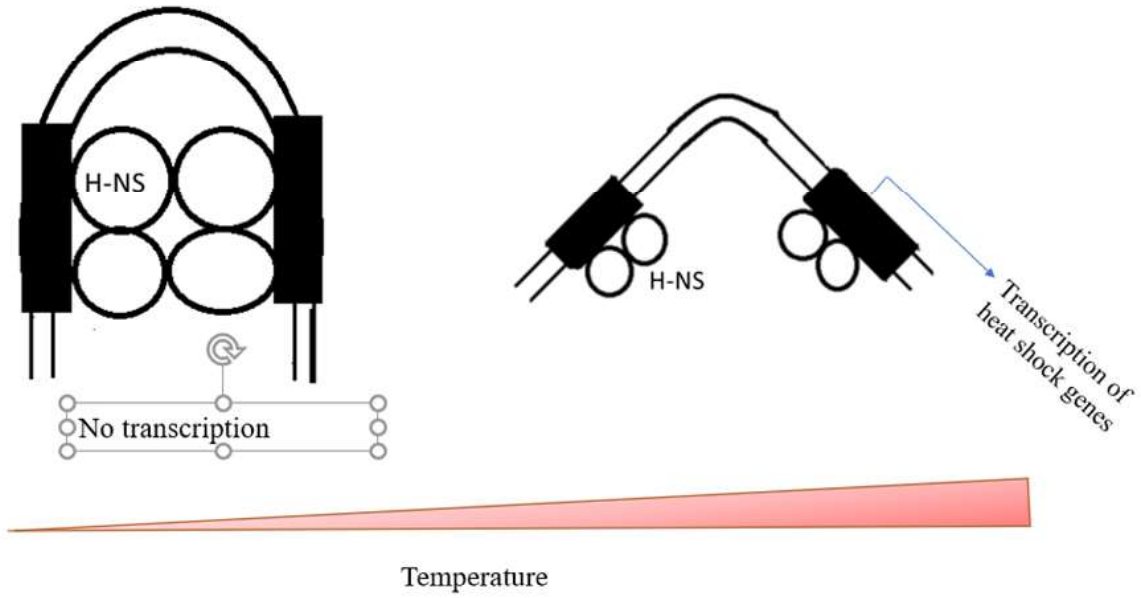


Figure 2. H-NS affects the DNA bending and access of RNA polymerase. H-NS strongly binds with the promoter at lower temperature. When temperature increases, H-NS proteins are not able to form higher-order oligomers, thereby freeing up the promoter for the RNA Polymerase to bind and transcription proceeds smoothly.

3. RNA

RNA based temperature-responding systems are another very common type of heat sensor found in nature (Digel et al., 2008). Their ability to respond to temperature caused the RNA sensor to be designated as RNA thermometer. RNA is well suited as a post-transcriptional intracellular thermosensor due to the natural tendency of the single stranded RNA molecules to change their secondary structure in response to shift in temperature. This in turn causes altered RNA stability or translation rates. Due to its temperature responsiveness, RNA thermosensors are often referred to as RNA thermometers (RNATs) (Narberhaus et al., 2006; Storz, 1999). All of the RNA thermometers that have been discovered thus far are translational control elements, and most of them are located in the 5' UTR of bacterial heat shock and virulence genes (Kortmann & Narberhaus, 2012). RNATs have been reported to work by forming pronounced secondary and tertiary structures, and also by forming intermolecular RNA

: RNA hybrids (Digel et al., 2008). Both of these processes greatly depend on the formation of complementary base pairing, and consequently one would anticipate these to be dependent on environmental temperature. RNA when compared to DNA, responds to temperature in a more immediate fashion, thus making them more efficient sensor (Kortmann & Narberhaus, 2012).

RNA thermometers provide an elegant mechanism for the microorganisms to adapt to different temperatures. They function independently of a response regulator and *de novo* transcription is not required (Kortmann & Narberhaus, 2012). Another major advantage of this system is that they are energetically economical. They are also easily activated and can switch from one form to the other depending on the environmental changes (Barnwal et al., 2016).

RNA thermometers are either heat-inducible or heat-repressible i.e., they activate or repress gene expression in response to high temperature (Narberhaus et al., 2006). In nature, most of the RNA thermosensors are heat-inducible (Chowdhury et al., 2006). They sequester the ribosome binding site (RBS) in a hair-pin loop structure at low temperature and when there is an upshift in temperature, the RBS gets exposed due to the destabilization of the hairpin structure . An example of heat-inducible thermosensor is the ROSE element (Narberhaus et al., 2006). Heat repressible thermosensor on the other hand has a variety of mechanisms and the expression of the alternative sigma factor RpoS is one such example (Schumann, 2012).

3.1 ROSE elements

ROSE elements (repression of heat shock gene expression), is the most common class of RNAT. This system was first discovered in *Bradyrhizobium japonicum* (Narberhaus et al., 1998) and has been later described in different *Rhizobium* species and *Agrobacterium tumefaciens* (Balsiger et al., 2004; Nocker et al., 2001). ROSE-type structures are composed of two, three or four individual hairpins and usually consists of 60 to more than100 nucleotides. All ROSE elements are located in the 5'-UTR of transcripts that code for small heat shock genes , acquire a complex structure comprising 2-4 stem loops where the 3'- proximal hairpin consists of the Shine-

Dalgarno sequence and sometimes even the AUG start codon (Schumann, 2012) . The 5´-most hairpins were found to be stable under heat shock conditions but the 3´-most hairpin that pairs with the Shine–Dalgarno sequence (SD sequence) was stable only at low temperatures (Chowdhury et al., 2003). Moreover, rise in temperature induced local melting of the hairpin structures, which exposed the SD sequence, thereby facilitating ribosome binding. The role of other hairpins could be to ensure proper folding of the labile hairpin during the process of transcription (Chowdhury et al., 2003; Nocker et al., 2001). It has been shown that the SD region remained sequestered at 30°C, but when temperature raised to 37°C, partial melting of the hairpin took place leading to low levels of transcription. At 42°C, full liberation of the SD and AUG start codon occurs due to the complete melting of the hairpin structure, facilitated mRNA-ribosome interaction (Chowdhury et al., 2006). Thus, due to its gradual response to temperature alterations *in vivo* and *in vitro,* the ROSE element acts like a typical zipper (Kortmann & Narberhaus, 2012).

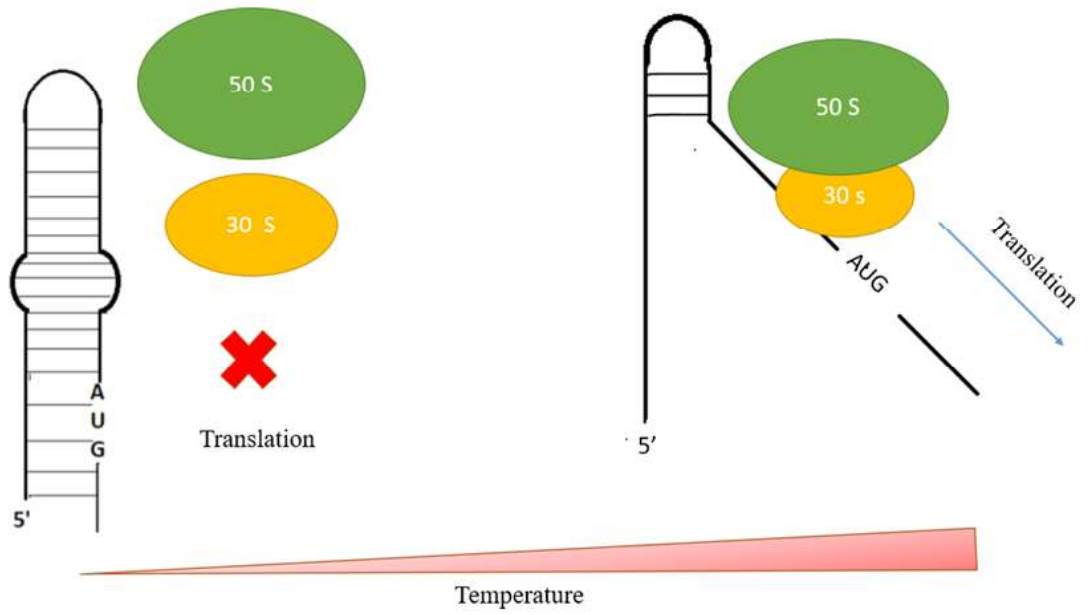


Figure 3. RNA Thermometer . The Shine Dalgarno sequence and the AUG start codon are base paired at low temperature. The melting of the structure at higher temperatures permits the binding of the ribosome (30S and 50S) allowing translation to proceed.

3.2 fourU elements

Another widespread RNA thermometer is the fourU element, and one of the widely studied fourU element was found to be present upstream of the small heat shock gene aggregation suppressing A (*agsA*) in *Salmonella enterica,* which controlled its translation initiation (Waldminghaus et al., 2007) . These RNATs are named so due to the presence of a stretch of four uridines that pairs with AGGA in the Shine-Dalgarno sequence (Kortmann & Narberhaus, 2012). Such fourU (also known as 4U) elements also existed in the 5´ UTRs of many heat shock and virulence genes as well (Waldminghaus et al., 2007). Secondary-structure prediction and experimental RNA probing revealed that the *agsA* were short thermosensors consisting of only 52 nucleotides folding in two hairpins (Schumann, 2007). Hairpin I was observed to be heat stable, whereas formation of the second hairpin, which sequesters the SD region through the fourU motif, was temperature controlled (Kortmann & Narberhaus, 2012). With the increase in temperature, hairpin II opened up and allowed the binding of the ribosomes to the SD sequence, thus, allowing transcription (Kortmann & Narberhaus, 2012). Ribosome binding studies showed that the formation of the ternary translation initiation complex occurred only at high (45 °C) and not at low (30°C) temperatures (Waldminghaus et al., 2007).

Similarly, in the case of *Yersinia pestis, the expression of* adhesin YadA was found to regulated by the virulence factor LcrF. It was found that *the UTR* of LcrF contained four consecutive uridines that are predicted to pair with SD sequence at ambient temperature of 26°C and inhibited transcription (Kortmann & Narberhaus, 2012). A rise in temperature (37°C) led to the opening of the stem loop, resulting in the expression of the adhesin proteins (Kortmann & Narberhaus, 2012). Here also the UTR acted in a zipper-like manner, similar to the *Salmonella agsA* thermometer (Waldminghaus et al., 2007).

3.3 Cyanobacterial thermometer

This is the simplest natural RNAT. It was first identified in the upstream region of the heat shock protein 17 gene (*hsp17*) in *Synechocystis* (Kortmann & Narberhaus, 2012). This RNAT is only 46 nucleotides long and consists a single stem–loop structure which can open and close in a reversible manner (Kortmann & Narberhaus, 2012). Hsp17 functions as a chaperone for misfolded proteins in the cytoplasm. But apart from

that, it also controls membrane integrity and is therefore crucial for photosynthetic activity, especially during stress conditions (Török et al., 2001).

3.4 Css thermometer

This sort of thermosensor is found in the gram negative, diplococcus, obligate human pathogenic bacteria *Neisseria meningitidis* (Barnwal et al., 2016). They usually colonize the nasopharynx region, which is the site of several thermal gradients (Yazdankhah & Caugant, 2004). They produce a capsule that has sialic acid as the major component, which triggers an immune response (Barnwal et al., 2016). It has also been observed that the bacteria has evolved in such a way that it is able to evade host responses. It expresses its immune evasion mechanism prominently in the nasopharynx region where the temperature range is 32°C or above (Loh et al., 2013). The biosynthesis of sialic acid is encoded by the capsule biosynthesis operon, *css* (Barnwal et al., 2016). The RNA thermometer, which mediates the thermal response in capsule production, was found to be present within 110 nucleotides of the 5' untranslated region of the *css* operon. It formed secondary structure at lower temperature, thereby sequestering the Ribosome-binding site, but an increase in temperature causes melting of the secondary structure, resulting in translation (Grosso-Becera et al., 2015). Now, unlike other thermometers, *css* thermometer precisely regulated a single biochemical pathway, rather than affecting global responses to temperature and responds to even very small changes in the nasopharynx (Barnwal et al., 2016). During narrow transition ranges, the thermometer was reported to equilibrate between 'closed' and 'open' states, wherein, closed configuration dominated at temperature below activation (Loh et al., 2013) .. It was observed that the ribosomes screened the population of open thermometers so as to recognize the RBS and hence, initiate translation. Even though the structure begins to melt from the upper helix, the RBS remained sequestered, which became partially accessible for translation only at 37°C when the lower helix too melted (Loh et al., 2013). During infection, the temperature rises to 42°C, destabilising the entire helix, liberating the RBS and ribosome is able to carry-on with the process of translation (Barnwal et al., 2016).

3.5 Alternate Sigma factor

The first RNA thermometer that acts via the melting mechanism was found in the *Escherichia coli rpoH* gene, which codes for the alternative sigma factor σ^{32} or RpoH (Morita et al., 1999). σ^{32} is a key regulator of the heat shock response . At low temperature, cells were found to contain very little σ^{32} (10-30 molecules at 30°C), but after a temperature upshift to 42°C, the amount of σ^{32} increased to about 15-fold. Unlike other RNATs, majority of the regulatory region resided in the coding region and not in the 5' UTR (Klinkert & Narberhaus, 2009). Open reading frame of *rpoH* consisted of two segments i.e., A & B that are about 220 nucleotides in length. It formed an extensive RNA structure at lower temperature, preventing the entry of ribosome to the SD sequence and melted at higher temperature, liberating the ribosome binding site for translation of the sigma factor that induces the heat shock response (Klinkert & Narberhaus, 2009).

Another alternative sigma factor RpoS (σ^S or σ^{38}) was found to regulate the virulence-associated major outer surface proteins OspC and OspA in the spirochete *Borrelia burgdorferi* which is the causative organism of Lyme disease (Klinkert & Narberhaus, 2009). Temperature had a direct effect on RpoS levels and it is $DsrA_{Bb}$, which is a small non-coding RNA, that is responsible for regulating temperature-induced increase in RpoS (Lybecker & Samuels, 2007). It was observed that $DsrA_{Bb}$ folded into a stable secondary structure at 23°C, which did not allow the base pairing with the rpoS mRNA. With the upshift in temperature, the secondary structure of $DsrA_{Bb}$ melted, and was able to interact with the anti-SD region of the rpoS mRNA. This led to the releases of the sequestered SD sequence and induction of translation under virulence conditions (37°C) (Klinkert & Narberhaus, 2009).

Alternative sigma factor also plays an important role during heat shock in the microorganisms *Lactobacillus casei and Lactobacillus rhamnosus,* that are widely used in the food fermentation industry and as probiotics. Their ability to survive various stresses like heat, cold, acidic and osmotic stress they are exposed to during processing is due to the presence of alternative sigma factors (Rossi et al., 2016). Some alternative factors of the σ^{70} family, namely σ^{24} and σ^{32}, which are the σ subunits of RNA polymerase that are expressed under particular conditions and recognizes specific promoter regions, were also reported to be upregulated in response to the heat shock response (Zhao et al., 2005). However, the heat shock response is not just limited to the upregulation of σ factors. When there is no stress, σ^{32} levels were found

to be low due to the action of an ATP dependent protease called as the FtsH protease which complexes with the chaperone DnaKJ (Rossi et al., 2016). Upon heat shock, level of FtsH increased in response to the level of misfolded proteins and thereby, the level of σ^{32} also increased (Zhao et al., 2005). Now, it has been reported that σ^{24} is encoded by *rpoE* gene found in the organisms like *Escherichia coli* and *Salmonella typhimurium* (Dartigalongue et al., 2001). This factor was found to be an extracytoplasmic (ECF) σ factor and the *rpoE gene* gets co-transcribed with a downstream gene that encoded for an anti- σ protein whose function was to keep the levels of HSR σ^{24} factor inactive in absence of stress (Zhao et al., 2005). During stress, membrane proteases degraded the anti- σ protein ,thereby activating the factor σ^{24} (Staron et al., 2009). On the other hand, nothing much is known about the role of σ^{24} in *Lactobacilli*.

3.6 Csp thermometer

In bacteria like *Escherichia coli ,* cold shock genes expression gets induced *de* novo massively during the lag phase following a cold shock from 37°C to 15°C (Klinkert & Narberhaus, 2009). Cold shock protein A (CspA) is an RNA chaperone that binds single-stranded RNAs to prevent the formation of any secondary structures during cold temperatures (Jiang et al., 1997). The cspA transcript was observed to be extremely unstable at 37°C, but was stabilized when the temperature reduced to 10°C (Yamanaka et al., 1999). cspA expression at lower temperature was enhanced due to the presence of an inbuilt RNAT, which was present 159 nucleotides upstream of the *cspA* gene. Unlike other RNATs mechanisms, the rearrangements of the *cspA* RNAT resulted not from the melting of a hairpin structure but from mutually exclusive alternative structures (Jiang et al., 1997). The 'cold shock structure' was translated in an efficient manner and less susceptible to degradation than the structure at 37°C. Here, the SD sequence was found to be exposed and the start codon got placed in an unstable helix. At 37°C, the translation initiation elements were buried within a double-stranded structure that hindered the translation efficiency (Kortmann & Narberhaus, 2012).

Facultative human pathogen *Vibrio cholerae* is usually found in aquatic reservoirs like estuaries, coastal waters and fresh water rivers and lakes, and temperature is the signal that helps them differentiate freshwater environment (12°C

to 30°C) from human host (37°C) . (Alam et al., 2006). And it was only in the human host that they are able to express the virulence factors as there is a temperature upshift (Townsley et al., 2016). It has been reported that only CspA and CspV, out of the four cold shock proteins (the other two being VCA0184 and VCA1142) encoded by *Vibrio cholerae* were induced upon shifting the bacterial cells from 37°C to lower temperature (Datta & Bhadra, 2003) . Unlike other cold shock genes, *cspV* consisted of a cold-inducible promoter that lacks a long 5' untranslated region in its mRNA, and their transcript levels in human shed *Vibrio cholerae* were high, which provides us an insight on how the bacteria survives once getting disseminated from human host (Townsley et al., 2016)..

3.7 cIII thermosensor

Bacteriophage λ belongs to the category of temperate phage. Within 10minutes of infection, the phage decides whether it enters the lytic phase or lysogenic phase. The cIII gene product is involved in the lysis–lysogeny decision, translation of which is temperature- dependent (Schumann, 2012). Hence the cIII gene acts as an RNAT. But unlike zipper-like RNATs, its translation is repressed with increasing temperature. Hence, this RNAT exists as two mutually exclusive structures which act in a switch-like fashion. It makes use of 85 nucleotides in the 5´ UTR along with 45 nucleotides in the coding region so as to form one of the two mutually exclusive RNA structures, known as A (translation OFF) and B (translation ON) (Altuvia et al., 1989). When temperature is below 37°C, the RNAT existed in the B form. This ensured the interaction between the anti-SD site and an anti-anti-SD site, making the SD site accessible for translation. The unpaired SD region facilitated the binding of the 30S ribosomal subunit and translation of the cIII mRNA, which in turn activated the lysogenic pathway. When the temperature upshifted to 45°C, it was demonstrated that the structural rearrangement occurred and the equilibrium shifted to structure A, an event that was stimulated by the processing of RNase III. Hence, the SD sequence of A got sequestered due to interaction with anti-SD site, which prevented the translation of cIII, thereby driving the phage into the lytic pathway (Kortmann & Narberhaus, 2012). Further, it appeared that Mg^{2+} plays a structural role in the cIII thermosensor, as at a low Mg^{2} concentration, equilibrium was shifted towards structure A, i.e., translation OFF and vice-versa (Altuvia et al., 1989).

3.8 Intra- and intermolecular RNA–RNA interactions

In *Listeria monocytogenes* ,the virulence regulator PrfA was found to be encoded by an mRNA that had a 127-nucleotide long 5´ leader sequence which formed a complex secondary structure with several internal loops (Johansson et al., 2002). At lower temperatures, in spite of the fact that it had exposed SD sequence and AUG start codon, the secondary structure was observed to be considerably stable and thus, inhibited translation (Kortmann & Narberhaus, 2012). But when the temperature increased to 37°C, this wild-type structure got destabilized and allowed the translation of PrfA mRNA (Johansson et al., 2002). PrfA has been shown to activate the expression of several virulence genes that encode adhesins, phagosome-escape factors and several immune-modulating factors that repress the host innate immune response (Kortmann & Narberhaus, 2012).

3.9 FRQ

In *Neurospora sp.,* the expression levels and ratio of isoforms of the circadian clock protein FREQUENCY (FRQ) are important for the temperature dependent circadian rhythms. The long (l) and short (s) isoform ratio was reported to be regulated by the thermosensive splicing of intron 6 of *frq* (Diernfellner et al., 2005). Accumulation of this protein hence occurred in a temperature dependent manner and was found to be controlled at the translational level. At low temperature translation initiation occurred at the non-consensus Kozak sequences upstream of the *frq* circadian gene. Reports suggested that an efficient initiation at these sequences decreased the abundance of FRQ protein at low temperatures and vice-versa (Sengupta & Garrity, 2013).

4. PROTEINS

A variety of environmental stresses can alter the protein conformation by damaging its structure that leads to protein aggregation, which can affect the ability of an organism to survive. So, organisms have regulatory networks which helps in refolding or degrading the damaged proteins (Klinkert & Narberhaus, 2009). Even though, these networks are expressed constitutively, they are induced strongly when there is protein damage (Klinkert & Narberhaus, 2009). Since, proteins are extremely sensitive to

changes in the environment, they are also very good thermosensors (Sengupta & Garrity, 2013).

There is a huge diversity in protein-based thermometers, because the tertiary and quaternary structures of proteins are very susceptible to temperature changes. Proteins are very sensitive to change in temperatures, and can easily shift from an active to inactive conformation. Temperature can also induce the disassembly of a dimer into a monomer (Schumann, 2012). Protein sensors include transcriptional regulators, chemosensory proteins, chaperones and proteases.

4.1 Transcription regulators

There are some bacterial heat shock and virulence genes which are under the direct control of temperature-responsive repressor proteins. In *Salmonella,* TlpA is an autoregulatory repressor protein that senses temperature (Digel et al., 2008). This repressor protein was found to be functional at low but not at high temperatures such as those encountered upon host infection. It was shown to have an alpha-helical protein that forms an elongated coiled-coil homodimer (Klinkert & Narberhaus, 2009). Studies have revealed that at lower temperatures, it adopted a coiled-coil homo-dimeric configuration that bound the promoter region of its own gene and repressed its own transcription (Sengupta & Garrity, 2013). When the temperature increased to 37°C (upon entry inside host), the dimers dissociated which led to a loss of its transcriptional repressor activity (Sengupta & Garrity, 2013).

RheA, a specific repressor of the small heat shock gene hsp18, is another example of auto-regulatory repressor found in the filamentous soil bacterium *Streptomyces albus*. It negatively regulated the expression of hsp18 coding for a small HSP (Klinkert & Narberhaus, 2009). Upon further analysis, it was observed that RheA complexed with the *hsp18* promoter at low temperature (30°C) . At 41°C this repression did not occur and transcription took place (Servant & Mazodier, 2001).

Another transcriptional regulator is the Phr protein present in *Pyrococcus furiosus* which was the first heat shock transcription factor in archaea to be documented (Liu et al., 2007). It is a dimeric protein that recognizes a specific palindromic sequence (Klinkert & Narberhaus, 2009). Studies have shown that at 95°C, it repressed the expression of heat shock genes by binding specifically to that

sequence, but, upon a heat shock (107°C), it was released from the promoters, thereby inducing the heat shock gene expression (Klinkert & Narberhaus, 2009).

Such a mechanism was also seen in dimorphic fungi like *Histoplasma capsulatum.* Autoregulatory transcriptional activator Ryp1 switched from the filamentous form to the pathogenic yeast form upon entering the host (37°C) (Klinkert & Narberhaus, 2009).

Another transcriptional repressor is the CtsR protein, which is a highly conserved heat-shock regulator that regulates the heat shock response in low GC containing Gram-positive bacteria (Elsholz et al., 2010). It is a helix-turn-helix (HTH) dimeric DNA-binding protein that recognizes a specific sequence within the promoter of the regulated genes and binds with it, thereby suppressing their expression (Derre et al., 2000). It was observed that in *Bacillus subtilis,* DNA binding capability of CtsR was dramatically reduced at heat-shock temperatures (50°C) and regained upon a downshift to 37°C. Thus, it can be concluded that CtsR senses the temperature shift as it binds to the DNA at the promoter region under normal conditions, thereby supressing transcription. However, it is unable to do so under heat stress, thus, making the promoter accessible for transcription of genes (Elsholz et al., 2010).

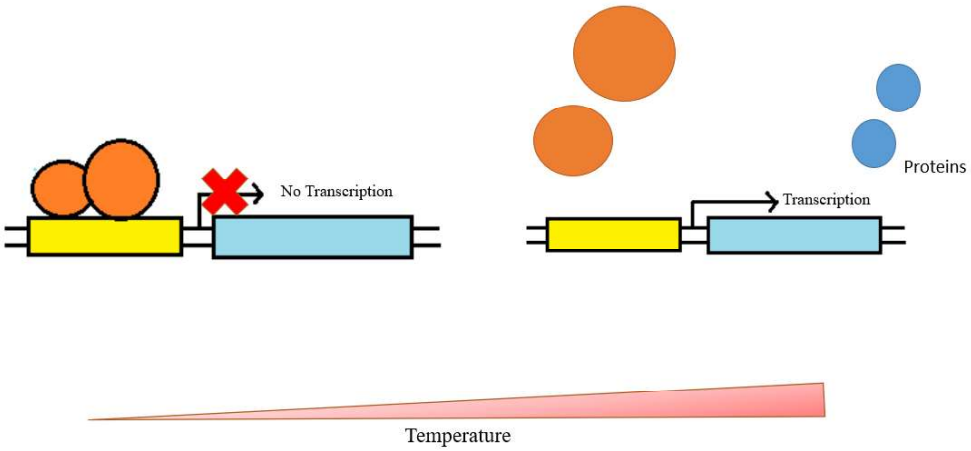


Figure 4. Protein Thermosensing via Transcriptional Repressors. Transcriptional repressors act as thermosensors in several lower organisms. The repressor proteins bind to the DNA associated with the heat shock response, thereby blocking its transcription. Upon a heat shock, the proteins loses its conformation and dissociates from the DNA, allowing transcription to occur.

4.2 Methyl-accepting chemotaxis proteins (MCPs)

Escherichia coli responds to small changes in temperature by altering its swimming behaviour so as to migrate in spatial temperature gradient (Sengupta & Garrity, 2013). Tsr and Tar are the two receptors that are involved in this chemotaxis mechanism. Tsr receptor sense glycine, while Tar senses aspartate (Mesibov & Adler, 1972). It was observed that upon deletion of *tsr*, the bacteria escaped from the heated region at all growth concentrations and deletion of *tar* resulted in the accumulation of the bacteria at the heated region, indicating that Tsr was responsible for attraction towards warmth and Tar towards cold (Salman & Libchaber, 2007). Tsr receptors which are abundant in the cells under normal growth conditions, attracted bacteria towards warmth, while methylation made the receptors thermo-insensitive (Klinkert & Narberhaus, 2009). Tsr which was methylated by glycine secreted by bacteria during its growth , now insensitive, caused Tar being the sole receptor that controlled the bacterial response to temperature changes (Park et al., 2003; Salman & Libchaber, 2007). It appeared that Tar acts as a warm- sensor as long as aspartate is absent and mediates attractant and repellent responses to temperature increase and decrease respectively (Salman & Libchaber, 2007). On the other hand, when aspartate is bound, Tar behaved as a cold-sensor (Klinkert & Narberhaus, 2009).

4.3 Chaperones and proteases

Several stresses, particularly heat stress disrupts the weak interactions among proteins and thereby affects the protein stability and conformation (Klinkert & Narberhaus, 2009). Heat stress usually causes aggregation and denaturation of proteins, that is extremely lethal to the cell. All these effects are counteracted by special group of proteins called as chaperones and proteases (Schumann, 2012).

The chaperone system in *Escherichia coli* consists of DnaK (HSP70) , its co-chaperone DnaJ and the nucleotide exchange factor GrpE (Klinkert & Narberhaus, 2009). DnaJ is the non-native protein sensor and GrpE is the thermosensor which monitors the cellular situation and induces the activity of DnaK system in response to heat shock (Siegenthaler & Christen, 2006). This system responds to change in temperature by the virtue of the conformational changes in the GrpE protein (Grimshaw et al., 2001). In studies, GrpE was observed to have a quaternary structure and it is the long N-terminal α-helix of the protein that behaved as a thermosensor (Harrison,

2003). This thermosensor is what that brought about a decrease in the nucleotide-exchange activity by GrpE during heat shock (Harrison, 2003). GrpE and DnaK were noticed to be complexed with each other in a stoichiometry of 2:1. The affinity for DnaK by GrpE was not affected by temperature above 40°C, but paired α-helices of GrpE underwent a helix-to-coil transition, thereby melting the helix, which inactivated its nucleotide exchange activity (Klinkert & Narberhaus, 2009). This inhibited the release of substrate from the chaperone DnaK at higher temperatures. Thus, this resulted in conversion of DnaK from an ATP consuming foldase into a holdase whose function is to tightly bind unfolded polypeptides and maintain them in a folding-competent state under conditions where protein folding is inefficient (Siegenthaler & Christen, 2006). At lower temperatures, GrpE reassociated, and enhanced the release of the respective proteins from DnaK (Klinkert & Narberhaus, 2009).The chloroplast homolog of GrpE in the unicellular algae *Chlamydomonas reinhardtii* was found to be CGE1 (Schroda et al., 2001). CGE1 existed in two isoforms; CGE1a and CGE1b, which is formed in response to alternative splicing due to change in temperature. The isoform 'a' existed predominantly at lower temperatures, while the other form was predominant at higher, elevated temperatures (Klinkert & Narberhaus, 2009).

Another heat shock protein, Hsc66 is the second HSP70 protein found in *Escherichia coli* (Sengupta & Garrity, 2013). Usually, Hsp70 proteins are associated with heat shock response, but Hsc66, which is an HSP70 homolog, was found to be induced following a cold shock (Lelivelt & Kawula, 1995). When the *Escherichia coli* cells were shifted from 37°C to 10°C, it was observed that the cellular growth and protein synthesis got inhibited (Digel et al., 2008). So as to compensate for this, a subset of proteins were induced which turned out be Hsc66 proteins. In order to confirm the role of Hsc66, the mutant *Escherichia coli* cells that lacked Hsc66 were exposed to cold shock treatment and expression pattern of other induced proteins were analysed. It was observed that the expression of those proteins were altered (Lelivelt & Kawula, 1995).

There are reports suggesting that on exposing the cells of the bacterium *Mycobacterium tuberculosis* to elevated temperature, a limited set of heat shock proteins were induced (Young & Garbe, 1991). The PAGE and subsequent autoradiography revealed that there were some distinct differences in the patterns of the proteins that were induced at different temperatures (Young & Garbe, 1991). Mild heat shock (37°C to 42°C) treatment resulted in the induction of 70-kDa and 65-kDa

proteins, while high temperature stress (48°C) resulted in the prominent induction of 90-kDa, 20-kDa and 15-kDa proteins (Young & Garbe, 1991). By performing two-dimensional gel electrophoresis, they observed that prominent heat-shock proteins induced at 70-kDa and 65-kDa matched the antigens which upon identification by sequence analysis was revealed to be the mycobacterial DnaK and GroEL homologs. Upon metabolic labelling with a ^{14}C-amino acid mixture, a 12-kDa protein was identified, further analysis of which revealed it to be the GroES, which resembled the DnaK protein with respect to temperature dependence. However, the remaining heat shock proteins did not correspond to any other antigen of *Mycobacterium tuberculosis* (Young & Garbe, 1991).

Just like in bacteria, several heat shock proteins are also induced in green algae so as to re-establish protein homeostasis (Barati et al., 2019). HSP70 and HSP90 have been known to have much more noticeable chaperone activity when compared to the other HSPs. Reports suggested that HSP90s stabilised and maintained the conformation of unstable proteins close to their native form (Richter & Buchner, 2001). HSP70 chaperones on the other hand, were mainly involved in protein translocation and protein folding. A chloroplast stromal HSP70B in *Chlamydomonas reinhardtii* and *Dunaliella* were reported to be involved in the molecular protection and repair of photosystem II (Barati et al., 2019).

Protease acting as a thermosensor has been elucidated only in the case of *Escherichia coli.* DegS, anchored in the inner membrane of the cells facing the periplasmic space has an N-terminal transmembrane domain ,followed by a protease domain and C-terminal PDZ domain (Alba et al., 2001). PDZ domains usually recognizes specific C-terminal polypeptides and is assumed to inhibit the protease domain by direct contact between both domains (Alba et al., 2001). Usually, heat shock results in over production of porins. So, studies have shown that PDZ domain in DegS recognized C-terminal peptides with Y-X-F motif, which is commonly found in many outer membrane porins (Schumann, 2007). The exposed C-terminal ends of the denatured porins caused the PDZ domain to release itself from the protease domain and interacted with the Y-X-F motif (Schumann, 2012). The newly free protease domain attacked the anti-sigma factor RseA which contained the periplasmic domain, transmembrane domain and a cytoplasmic domain. The cytoplasmic domain sequestered the alternative sigma factor σE. So, the free DegS protease, was able to cleave within the periplasmic domain of RseA, and remaining RseA got further

degraded (Walsh et al., 2003). These events caused the cytoplasmic domain of RseA to get destabilized, thereby releasing σE which in turn activated the transcription of genes in the σE regulon. Once the denatured porins are removed, PDZ domain again binds with the protease domain, thereby switching off the heat shock response(Schumann, 2012).

Apart from the previously mentioned ones, there is a thermosensor in *Escherichia coli* which has both chaperone and protease activity. DegP or HtrA is the most evolved temperature sensing protein in *Escherichia coli* and was found to have a remarkable chaperone-protease mechanism that switched between both the activities in a temperature-dependent manner (Clausen et al., 2002). They have two domains in their structure; proteolytic domain and the PDZ domain, with each domain forming a six-stranded β-barrel. The PDZ domains were found to be highly mobile swinging around to capture substrate proteins. This domain bound with the C-terminal of the target proteins (Schumann, 2012). It was observed that this protein had chaperone activity and had no proteolytic activity at normal temperature (20°C). When the temperature rose to 30°C, only proteolytic activity of the protein was prominent (Spiess et al., 1999). Such a mechanism has been also reported in the bacteria *Thermotoga maritima,* which consisted of the chaperone-protease Tm HtrA (Klinkert & Narberhaus, 2009). Here also, chaperone form predominated at lower temperature, while protease activity got induced at higher temperatures (Klinkert & Narberhaus, 2009).

The green algae *Chlamydomonas reinhardtii* harbor another class of proteins that aids in its response to temperature stress. Proteins that are irreversibly denatured or misfolded are acted upon by ubiquitins for degradation (Von Kampen et al., 1995). The ubiquitinated proteins are acted upon by proteasomes. Increase in ubiquitin level is normally associated with high levels of protein degradation and turn-over (Schroda et al., 2015). When *Chlamydomonas reinhardtii* was exposed to heat stress, an up-regulation in the levels of ubiquitin was observed (Von Kampen et al., 1995).

5. MISCELLANEOUS

Apart from the above-mentioned categories, diverse types of thermosensors are present in various living systems.

5.1 TRP channels

TRP channels refers to transient receptor potential channels, which form the key aspect our sensory stem, as it responds immediately to both physical and chemical stimuli and, thus, has multiple functions (Digel et al., 2008). The TRP superfamily consists of 30 mammalian members that are grouped into 6 major families: TRPV, TRPC, TRPP, TRPM, TRPN and mucolipins. Out of the various TRP members, only a few acts as thermo-TRPs. They are TRPV(1-4) which are heat activated, while TRPM8 and TRPA1 are activated by cold (Digel et al., 2008). Despite the great advancements made in this field, the molecular basis for regulation by temperature remains unknown because of the lack of the structural information (Schumann, 2007).

5.2 Kinases

Two-component histidine kinase system (TCS) in bacteria acts as a sensor due to its ability to respond to a variety of environmental stimuli (Klinkert & Narberhaus, 2009). TCS is comprised of membrane-bound sensor protein with an intracellular kinase domain that contains a histidine residue. When there is a change in environment, the histidine residue gets auto phosphorylated and then the phosphate group is then transferred to an aspartate on an intracellular response protein. This in turn regulates a number of cellular pathways, including transcription(Sengupta & Garrity, 2013) .

Drk1(dimorphism regulating kinase) is a conserved histidine kinase present in the dimorphic fungi *Histoplasma capsulatum* and *Blastomyces dermatitidis.* Drk1 has been shown to act as temperature sensor and has the ability to control morphogenesis as well as adapt to various stresses present inside the mammalian host (Shapiro & Cowen, 2012). Drk1 induced the yeast-like growth of the organism at higher temperature (37°C) (Klinkert & Narberhaus, 2009). This protein is also required for the expression of virulence genes and in establishing fungal pathogenicity inside their hosts. Thus, it can be inferred that there is direct relationship between fungal thermotolerance and virulence (Shapiro & Cowen, 2012).

A slightly different type of kinase system exists in *Pseudomonas syringae*, which is a phytopathogenic bacteria that synthesizes a phytotoxin, coronatine (COR), that helps it to invade without being detected by the host (Sengupta & Garrity, 2013). COR mimics the plant signalling and causes chlorosis, shrinkage of chloroplast and

synthesizes ethylene and proteinase inhibitors (Klinkert & Narberhaus, 2009). A specific strain of this bacteria PG4180, was found to synthesize COR in a temperature dependent manner such that highest amount of transcripts were produced at lower temperature of 18°C, while at 28°C, which is the optimal growth temperature for the pathogen, negligible amounts of COR was produced (Budde & Ullrich, 2000). This, occurred due to interaction between three proteins CorR, S and P. Response regulator CorR, was found to bind to the promoter region of the *cma* operon responsible for the biosynthesis of coronamic acid. Sensor kinase CorS transferred the phosphoryl group to CorR due to presence of a receiver aspartate (Klinkert & Narberhaus, 2009). Studies have further revealed that autophosphorylation of CorS was controlled by conformational changes of the H-box that contains histidine, which at 18°C was active, whereas, at 28°C it slipped into membrane and remain inaccessible for autophosphorylation (Klinkert & Narberhaus, 2009).

There have also been reports suggesting the presence of two-component system in other plant pathogens like *Agrobacterium*. *Agrobacterium tumefaciens* is a gram-negative bacteria that causes crown gall disease mostly in dicot plants. Virulence of this pathogen is also controlled by a two-component system that consist of the sensor kinase VirA and response regulator VirG (Sengupta & Garrity, 2013). VirA phosphorylates VirG which in turn activates other sets of *vir* genes. These set of *vir* gene aid in the excision and transfer of T-DNA from the Agrobacterium to the plant host (Klinkert & Narberhaus, 2009). Various experiments have revealed that the expression of virulence occurred only at temperatures below 32°C and temperatures above this led to its reversible inactivation and thus prevented VirG from being phosphorylated (Schumann, 2012) .

In *Synechocystis sp.,* the membrane fluidity changes were observed to be controlled by four acyl-lipid desaturases; DesA, DesB, DesC, and DesD (Murata & Los, 2006). It was reported that *desC* gene was expressed constitutively, while *desA*, *desB* and *desD* genes were transcribed only after a temperature drop. Two histidine protein kinases, i.e., Hik33 and Hik19, controlled the expression of desaturase (Klinkert & Narberhaus, 2009). Hik33(DspA) integrated into a network of histidine kinases and response regulators that perceived and transduced the cold shock as well as osmotic and oxidative stress conditions. Hik33 has been postulated to be a thermosensor that is embedded in the membrane (Murata & Los, 2006). When the temperature decreased and membrane fluidity reduced, Hik33 got auto phosphorylated and

transferred its phosphate group via Hik19 to its response regulator Rer1. Rer1 in turn regulated the transcription of *desB* and *desD* genes (Klinkert & Narberhaus, 2009).

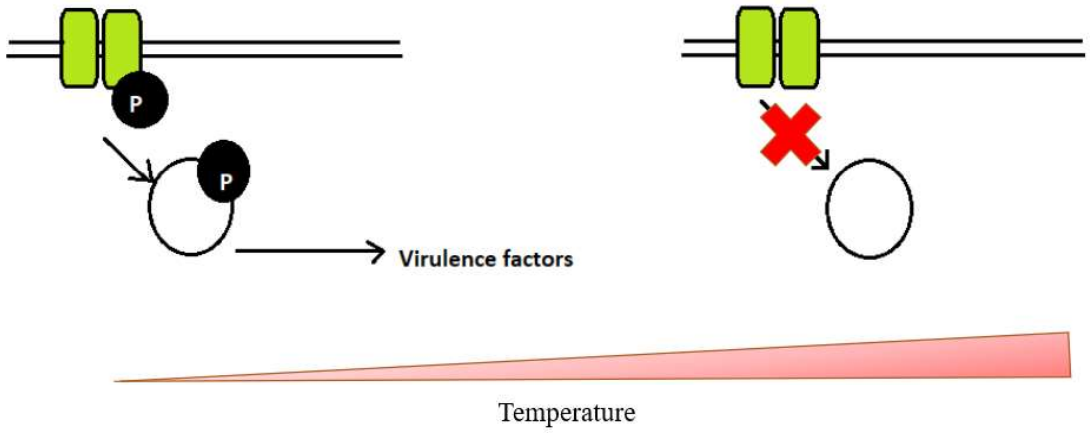


Figure 5. Histidine Kinase Thermosensor. The temperature-sensing kinase activates its downstream effector via autophosphorylation and effector phosphorylation, which allows the effector protein to activate different signalling pathways and virulence factors in response to temperature. Upon elevated temperature, changes in kinase protein conformation abolishes this signalling.

5.3 Flagellar motor

In *Escherichia coli,* flagellar rotary motor has been reported to switch between the clockwise (CW) and the counter clockwise (CCW) state in response to shift in temperature (Digel et al., 2008). So, it altered its motility to increase tumbling (CCW) while moving in an unfavourable direction. Similarly, tumbling rates were decreased and smooth swimming (CW) type of movement was adapted by the organism when moving in a favourable direction (Sengupta & Garrity, 2013). Thus, it is presumed that such a switch is due to the response regulator CheY-P (Digel et al., 2008).

5.4 Thermal avoidance

The phenomenon of movement of an organism towards or away from heat is called as thermotaxis. Such a phenomenon has been observed in various organisms like including the *Physarum polycephalum, Caenorhabditis elegans, Chlamydomonas,*

Escherichia coli, the protozoan *Paramecium etc* (Jennings, 1906). It was elucidated by Jennings in 1906 that *Paramecium* avoided regions of extreme temperature. When they encountered any hot surface, they swam long way back, thereby removing themselves as far as possible from it (Jennings, 1906). However, when the oral cilia encountered a weak stimuli, it either stopped, or progressed more slowly i.e., it altered its swimming pattern (Jennings, 1906). The reaction here is more delicate and precise when compared to the previous case. Various electrophysiological and genetic studies of *Paramecium* revealed that the stimuli can cause the depolarization of the surface membrane and the opening of a voltage-sensitive Ca^{2+} channel (Nakaoka & Oosawa, 2007). The resulting influx of Ca^{2+} raised the internal Ca^{2+} concentration above a critical value, which caused a temporary reversal in the direction of the effective stroke of the thousand cilia and a reversal in the swimming direction, followed by forward motion in a new direction. The net result of such avoiding reactions was hence, motion away from the stimulus (Hennessey & Nelson, 1979). In a temperature gradient, *Paramecium* was able to select the location having the temperature to which it had adapted and accumulated in that region which it deems more comfortable (Nakaoka & Oosawa, 2007). Thus, a thermo-sensory response is induced when there is a shift in temperature from its adaptation. During such an adaptation, the fatty acid composition of the membrane is also altered so as to keep the fluidity of the membrane constant (Nakaoka & Oosawa, 2007). Since a thermo-sensory response of *Paramecium* is induced when the temperature starts to change from the adapted temperature, it is believed that Prostaglandins (PGs), which are produced by the arachidonate cascade, may play a very vital role. Murata et al. (2000), did a study to understand the relationship between the thermo-sensory mechanism of *Paramecium* and various PGs, and only the prostaglandin PGI_2 and its analogs (iloprost and cicaprost) were found to cause a transient increase in the tumbling frequency. This meant that PGI_2 must be playing the role of a mediator thermo-sensory response (Murata et al., 2000).

5.5 Type VI secretion system (T6SS)

The type VI secretion system (T6SS) is responsible for mediating interspecies competition, virulence, and natural transformation in *Vibrio cholerae* (Townsley et al., 2016). It has been reported that it is composed of a base which spans the cell envelope, an inner tube that is composed of hemolysin-coregulated protein (Hcp) polymers, and an outer contractile sheath that is formed by VipA and VipB (Townsley

et al., 2016). Further, it was observed that temperature affected T6SS to a great extent as the message abundance, protein production and secretion, and activity of Hcp were highest at 25°C when compared to 37°C (Townsley et al., 2016).

CONCLUSION AND FUTURE PROSPECTS

Microorganisms respond to various environmental stresses by undergoing complex changes in the cell composition. All the cellular processes are monitored and regulated precisely. This helps the microorganisms to maintain the physiology of the cell, which helps them survive and grow under stressful conditions (Sengupta & Garrity, 2013). Temperature stress is one of the common stresses faced by any organism. Stress caused by high temperature itself can be of two types depending on the mode of induction of response; heat shock and high temperature response (Schumann, 2007). Heat shock response is a transient response and is caused by temperature increments. On the other hand, high temperature response is induced only at particular temperature, and proceeds until the organism is exposed to that temperature (Schumann, 2012).

A plethora of thermosensing mechanisms is present in various organisms. Their multifaceted response to a simple stimulus like temperature highlights how intricately the microorganisms integrate environmental signals and responds to it (Klinkert & Narberhaus, 2009). Plasma membrane is a very important organelle that separates the cellular environment from the outer one and happens to be the first responders of heat stress. Plasma membrane, being composed of lipid bilayer, changes its lipid composition with the aid of desaturases that helps them maintain their membrane fluidity at lower temperature (Sengupta & Garrity, 2013). Another class of thermosensors is the nucleic acid thermosensors which immediately respond to any temperature fluctuation in the environment. Any change in temperature is immediately sensed by the nucleic acids, which undergo changes so as to suit the stressful environment in which they are present (Sengupta & Garrity, 2013). The most common method by which DNA thermosensor responds is by changing its conformation upon a temperature stress. At times certain proteins known as the nucleoid associated proteins can also aid in change in conformation of the DNA. On the other hand, RNA exerts translational control with the help of certain special class of RNA called as RNA thermosensors, or RNATs. Usual mechanism by which RNATs act is by masking the ribosome binding site 5’ end at low temperatures by forming secondary structures. These structures melts at higher temperature for translation to proceed (Narberhaus et al., 2006). Apart from this, protein thermosensors also exists which can either affect

the activity of another gene or protein or itself undergoes a confirmation change in response to temperature changes. The heat shock response is caused by denatured proteins, and the primary aim of the heat shock response is to prevent the formation of protein aggregates by either protein refolding or degradation, a process taken care of by chaperones and proteases. In some organisms, transcriptional regulators which can activate or repress transcription of several temperature dependent genes depending on the temperature have also been identified.

While the molecular mechanisms by which certain bacterial proteins and nucleic acid thermosensors sense temperature have been elucidated, clear knowledge of exactly how lots of other thermosensors work is poorly understood and will be an important avenue for future investigations. The fact that primary sensors can't be distinguished from other temperature sensors which sense temperature indirectly pose a major challenge (Klinkert & Narberhaus, 2009). Moreover, the thermal response occurs within the physiological growth range and often induces conformational changes that are subtle and reversible which often act as a hurdle in the identification of the primary sensors.

In this era of climate change, the ability to sense and adapt to environmental temperature perturbations may be a life-or-death challenge for several species and may increase selective pressure on the evolution of biological thermometers. Hence, elucidating these thermosensors and discovering their mechanisms of thermosensing becomes important. The recent advancements in the field of RNA sequencing can be used to define secondary structures of all transcripts in bacteria to compare it at different temperatures (Mandin & Johansson, 2020). Narberhaus and co-workers have successfully used similar approach on RNA isolated from *Yersinia pseudotuberculosis* to identify and characterise more than a dozen new thermosensors (Righetti et al., 2016). In spite of the developments being made, thermosensing is still not understood entirety, and therefore will be an important avenue for future investigations. Thus, the diversity and complexity of molecular mechanisms involved in temperature-sensing underscores the importance of temperature as one of the most powerful forces in nature.

BIBLIOGRAPHY

Aguilar, P. S., Hernandez-Arriaga, A. M., Cybulski, L. E., Erazo, A. C., & de Mendoza, D. (2001). Molecular basis of thermosensing: a two-component signal transduction thermometer in *Bacillus subtilis*. The EMBO Journal **20**: 1681-1691.

Alam, M., Sultana, M., Nair, G. B., & Sack, R. B. (2006). Toxigenic *Vibrio cholerae* in the aquatic environment of Mathbaria, Bangladesh. Applied and Environmental Microbiology **72**: 2849-2855.

Alba, B. M., Zhong, H. J., Pelayo, J. C., & Gross, C. A. (2001). degS (hhoB) is an essential *Escherichia coli* gene whose indispensable function is to provide sigma (E) activity. Molecular microbiology **40**: 1323-1333.

Allakhverdiev, S. I., Kinoshita, M., Inaba, M., Suzuki, I., & Murata, N. (2001). Unsaturated fatty acids in membrane lipids protect the photosynthetic machinery against salt-induced damage in *Synechococcus*. Plant Physiology **125**: 1842-1853.

Altuvia, S., Kornitzer, D., Teff, D., & Oppenheim, A. B. (1989). Alternative mRNA structures of the cIII gene of bacteriophage λ determine the rate of its translation initiation. Journal of Molecular Biology **210**: 265-280.

Atlung, T., & Ingmer, H. (1997). H-NS: a modulator of environmentally regulated gene expression. Molecular microbiology **24**: 7-17.

Balsiger, S., Ragaz, C., Baron, C., & Narberhaus, F. (2004). Replicon-specific regulation of small heat shock genes in *Agrobacterium tumefaciens*. Journal of Bacteriology **186**: 6824-6829.

Barati, B., Gan, S. Y., Lim, P. E., Beardall, J., & Phang, S. M. (2019). Green algal molecular responses to temperature stress. Acta Physiologiae Plantarum **41**.

Barnwal, R. P., Loh, E., Godin, K. S., & Yip, J. (2016). Structure and mechanism of a molecular rheostat, an RNA thermometer that modulates immune evasion by *Neisseria meningitidis*. Nucleic Acids Research **44**: 9426-9437.

Budde, I. P., & Ullrich, M. S. (2000). Interactions of *Pseudomonas syringae* pv. glycinea with host and nonhost plants in relation to temperature and phytotoxin synthesis. Molecular plant-microbe interactions **13**: 951-961.

Chowdhury, S., Maris, C., Allain, F. H., & Narberhaus, F. (2006). Molecular basis for temperature sensing by an RNA thermometer. The EMBO Journal **25**: 2487-2497.

Chowdhury, S., Ragaz, C., Kreuger, E., & Narberhaus, F. (2003). Temperature-controlled structural alterations of an RNA Thermometer. Journal of Biological Chemistry **278**: 47915-47921.

Clausen, T., Southan, C., & Ehrmann, M. (2002). The HtrA family of proteases: implications for protein composition and cell fate. Molecular cell **10**: 443-455.

Cybulski, L. E., Martin, M., Mansilla, M. C., Fernandez, A., & de Mendoza, D. (2010). Membrane thickness cue for cold sensing in a bacterium. Current Biology **20**: 1539-1544.

Dame, R. T. (2005). The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. Molecular Microbiology **56**: 858-870.

Dartigalongue, C., Missiakas, D., & Raina, S. (2001). Characterization of the *Escherichia coli* σ^E Regulon. Journal of Biological Chemistry **276**: 20866-20875.

Datta, P. P., & Bhadra, R. K. (2003). Cold shock response and major cold shock proteins of *Vibrio cholerae*. Applied and Environmental Microbiology **69**: 6361-6369.

Derre, I., Rapoport, G., & Msadek, T. (2000). The CtsR regulator of stress response is active as a dimer and specifically degraded *in vivo* at 37°C. Molecular Microbiology **38**: 335-347.

Diernfellner, A. C. R., Schafmeier, T., Merrow, M. W., & Brunner, M. (2005). Molecular mechanism of temperature sensing by the circadian clock of *Neurospora crassa*. Genes & Development **19**: 1968-1973.

Digel, I., Kayser, P., & Artmann, G. M. (2008). Molecular processes in biological thermosensation. Journal of Biophysics **2008**: 1-9.

Duong, N., Osborne, S., Bustamante, V. H., Tomljenovic, A. M., Puente, J. L., & Coombes, B. K. (2007). Thermosensing coordinates a cis-regulatory module for transcriptional activation of the intracellular virulence system in *Salmonella enterica* serovar Typhimurium. Journal of Biological Chemistry **282**: 34077-34084.

Elsholz, A. K., Michalik, S., Zühlke, D., Hecker, M., & Gerth, U. (2010). CtsR, the Gram-positive master regulator of protein quality control, feels the heat. The EMBO Journal **29**: 3621-3629.

García-Ávila, M., & Islas, L. D. (2019). What is new about mild temperature sensing? A review of recent findings. Temperature **6**: 132-141.

Grimshaw, J. P., Jelesarov, I., Schönfeld, H. J., & Christen, P. (2001). Reversible thermal transition in GrpE, the nucleotide exchange factor of the DnaK heat-shock system. Journal of biological chemistry **276**: 6098-6104.

Grosso-Becera, M. V., Servín-González, L., & Soberón-Chávez, G. (2015). RNA structures are involved in the thermoregulation of bacterial virulence-associated traits. Trends in Microbiology **23**:

509-518.

Göransson, M., & Uhlin, B. E. (1984). Environmental temperature regulates transcription of a virulence pili operon in *E. coli*. The EMBO Journal **3**: 2885-2888.

Harrison, C. (2003). GrpE, a nucleotide exchange factor for DnaK. Cell Stress & Chaperones **8**: 218-224.

Hayes, S., Schachtschabel, J., Mishkind, M., Munnik, T., & Arisz, S. A. (2020). Hot topic: Thermosensing in plants. Plant,Cell & Environment 1-16.

Hazel, J. R. (1995). Thermal adaptation in biological membranes: is homeoviscous adaptation the explanation?. Annual Review of Physiology **57**: 19-42.

Hennessey, T., & Nelson, D. L. (1979). Thermosensory behaviour in *Paramecium tetraurelia*: a quantitative assay and some factors that influence thermal avoidance. Journal of General Microbiology **112**: 337-347.

Hoynes-O'Connor, A., Hinman, K., Kirchner, L., & Moon, T. S. (2015). *De novo* design of heat-repressible RNA thermosensors in *E. coli*. Nucleic Acids Research **43**: 6166-6179.

Jennings, H. S. (1906). Behavior of the lower organisms (The Columbia university press, The Macmillan company, agents: New York).

Jiang, W., Hou, Y., & Inouye, M. (1997). CspA, the major cold-shock protein of *Escherichia coli*, is an RNA chaperone. Journal of Biological Chemistry **272**: 196-202.

Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M., & Cossart, P. (2002). An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. Cell **110**: 551-561.

Kataoka, K., Mizushima, T., Ogata, Y., Miki, T., & Sekimizu, K. (1996). Heat shock-induced DNA relaxation *in vitro* by DNA gyrase of *Escherichia coli* in the presence of ATP. Journal of Biological Chemistry **271**: 24806-24810.

Katayama, S., Matsushita, O., Jung, C. M., Minami, J., & Okabe, A. (1999). Promoter upstream bent DNA activates the transcription of the *Clostridium perfringens* phospholipase C gene in a low temperature-dependent manner. The EMBO Journal **18**: 3442-3450.

Klinkert, B., & Narberhaus, F. (2009). Microbial thermosensors. Cellular and Molecular Life sciences **66**: 2661-2676.

Kortmann, J., & Narberhaus, F. (2012). Bacterial RNA thermometers: Molecular zippers and switches. Nature Reviews Microbiology **10**: 255-265.

Kusumoto, S., Fukase, K., & Shiba, T. (2010). Key structures of bacterial peptidoglycan and lipopolysaccharide triggering the innate immune system of higher animals: Chemical synthesis and functional studies. Proceedings of the Japan Academy, Series B **86**: 322-337.

Labrec, E. H., Schneider, H., Magnani, T. J., & Formal, S. B. (1964). Epithelial cells penetration as an essential step in the pathogenesis of bacillary dysentry. Journal of Bacteriology **88**: 1503-1518.

Lelivelt, M. J., & Kawula, T. H. (1995). Hsc66, an Hsp70 homolog in *Escherichia coli*, is induced by cold shock but not by heat shock. Journal of Bacteriology **177**: 4900-4907.

Li, Y., Powell, D. A., Shaffer, S. A., Rasko, D. A., Pelletier, M. R., Leszyk, J. D., Scott, A. J., Masoudi, A., Goodlett, D. R., Wang, X., Raetz, C. R., & Ernst, R. K. (2012). LPS remodeling is an evolved survival strategy for bacteria. Proceedings of the National Academy of Sciences **109**: 8716-8721.

Liu, W., Vierke, G., Wenke, A. K., Thomm, M., & Ladenstein, R. (2007). Crystal structure of the archaeal heat shock regulator from *Pyrococcus furiosus*: a molecular chimera representing eukaryal and bacterial features. Journal of Molecular Biology **369**: 474-488.

Loh, E., Kugelberg, E., Tracy, A., Zhang, Q., Gollan, B., Ewles, H., Chalmers, R., Pelicic, V., & Tang, C. M. (2013). Temperature triggers immune evasion by *Neisseria meningitidis*. Nature **502**: 237-240.

Lopez-Garcia, P., & Forterre, P. (2000). DNA topology and the thermal stress response, a tale from mesophiles and hyperthermophiles. BioEssays **22**: 738-746.

Los, D. A. (2004). The effect of low-temperature-induced DNA supercoiling on the expression of the desaturase genes in S*ynechocystis*. Cellular and Molecular Biology **50**: 605-612.

Los, D. A., Mironov, K. S., & Allakhverdiev, S. I. (2013). Regulatory role of membrane fluidity in gene expression and physiological functions. Photosynthesis Research **116**: 489-509.

Lupski, J. R., Projan, S. J., Ozaki, L. S., & Godson, G. N. (1986). A temperature-dependent pBR322 copy number mutant resulting from a Tn5 position effect. Proceedings of the National Academy of Sciences **83**: 7381-7385.

Lybecker, M. C., & Samuels, D. S. (2007). Temperature-induced regulation of RpoS by a small RNA in *Borrelia burgdorferi*. Molecular Microbiology **64**: 1075-1089.

Lyon, B. R., & Mock, T. (2014). Polar Microalage: New approaches towards understanding adaptations to an extreme and changing environment. Biology **3**: 56-80.

Mandin, P., & Johansson, J. (2020). Feeling the heat at the millennium: Thermosensors playing with fire. Molecular Microbiology **113**: 588-592.

Maurelli, A. T., Blackmon, B., & Curtiss, R., III. (1984). Temperature-dependent expression of virulence genes in *Shigella* species. Infection and Immunity **43:** 195-201.

Mesibov, R., & Adler, J. (1972). Chemotaxis toward amino acids in *Escherichia coli*. Journal of Bacteriology **112**: 315-326.

Mizushima, T., Kataoka, K., Ogata, Y., Inoue, R., & Sekimizu, K. (1997). Increase in negative supercoiling of plasmid DNA in *Escherichia coli* exposed to cold shock. Molecular Microbiology **23**: 381-386.

Morita, M. T., Tanaka, Y., Kodama, T. S., Kyogoku, Y., Yanagi, H., & Yura, T. (1999). Translational induction of heat shock transcription factor σ32: Evidence for a built-in RNA thermosensor. Genes & Development **13**: 655-665.

Murata, N., & Los, D. A. (2006). Histidine kinase Hik33 is an important participant in cold-signal transduction in *Cyanobacteria*. Physiologia Plantarum **126**:17-27.

Murata, T., Morii, H., Watanabe, Y., Matsumura, K., Watanabe, Y., & Nakaoka, Y. (2000). Prostaglandin I_2, a possible thermo-sensory mediator in *Paramecium*. Biochemical and Biophysical Research Communications **273**: 516-520.

Nakaoka, Y., & Oosawa, F. (1977). Temperature-sensitive behavior of *Paramecium caudatum*. The Journal of Protozoology **24**: 575-580.

Narberhaus, F., Käser, R., Nocker, A., & Hennecke, H. (1998). A novel DNA element that controls bacterial heat shock gene expression. Molecular Microbiology **28**: 315-323.

Narberhaus, F., Waldminghaus, T., & Chowdhury, S. (2006). RNA thermometers. FEMS Microbiology Reviews **30**: 3-16.

Nocker, A., Hausherr, T., Balsiger, S., Krstulovic, N. P., Hennecke, H., & Narberhaus, F. (2001). A mRNA-based thermosensor controls expression of rhizobial heat shock genes. Nucleic Acids Research **29**: 4800-4807.

Ono, S., Goldberg, M. D., Olsson, T., Esposito, D., Hinton, J. C., & Ladbury, J. E. (2005). H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. Biochemical Journal **391**: 203-213.

Park, S., Wolanin, P. M., Yuzbashyan, E. A., H, L., Darnton, N. C., Stock, J. B., Silberzan, P., & Austin, R. (2003). Influence of topology on bacterial social interaction. Proceedings of the National Academy of Sciences **100**: 13910-13915

Phadtare, S., & Inouye, M. (2008). The cold shock response. EcoSal Plus **3**.

Phillips, N. J., Schilling, B., McLendon, M. K., Apicella, M. A., & Gibson, B. W. (2004). Novel modification of Lipid A of *Francisella tularensis*. Infection and Immunity **72**: 5340-5348.

Podkowinski, J., & Tworak, A. (2011). Acetyl-coenzyme A carboxylase – an attractive enzyme for biotechnology. BioTechnologia **92**: 321-335.

Prosseda, G., Falconi, M., Giangrossi, M., Gualerzi, C. O., Micheli, G., & Colonna, B. (2004). The *virF* promoter in *Shigella*: more than just a curved DNA stretch. Molecular Microbiology **51**: 523-537.

Pruss, G. J., & Drlica, K. (1989). DNA supercoiling and prokaryotic transcription. Cell **56**: 521-523.

Raetz, C., R, Reynolds, C. M., Trent, M. S., & Bishop, R. E. (2007). Lipid A modification systems in gram-negative bacteria. Annual Review Of Biochemistry **76**: 295-329.

Richter, K., & Buchner, J. (2001). Hsp90: chaperoning signal transduction. Journal of Cellular Physiology **188**: 281-290.

Righetti, F., Nuss, A. M., Twittenhoff, C., Beele, S., K, U., Will, S., Bernhart, S. H., Stadler, P. F., Dersch, P., & Narberhaus, F. (2016). Temperature-responsive *in vitro* RNA structurome of *Yersinia pseudotuberculosis*. Proceedings of the National Academy of Sciences **113**: 7237-7242.

Rossi, F., Zotta, T., Iacumin, L., & Reale, A. (2016). Theoretical insight into the heat shock response (HSR) regulation in *Lactobacillus casei* and *L. rhamnosus*. Journal of Theoretical Biology **402**: 21-37.

Sakamoto, T., & Murata, N. (2002). Regulation of the desaturation of fatty acids and its role in tolerance to cold and salt stress. Current Opinion in Microbiology **5**: 208-210.

Salman, H., & Libchaber, A. (2007). A concentration-dependent switch in the bacterial response to temperature. Nature Cell Biology **9**: 1098-1100.

Schroda, M., Hemme, D., & Mühlhaus, T. (2015). The *Chlamydomonas* heat stress response. The Plant Journal, **82**: 466-480.

Schroda, M., Vallon, O., Whitelegge, J. P., Beck, C. F., & Wollman, F. A. (2001). The chloroplastic GrpE homolog of *Chlamydomonas*: two isoforms generated by differential splicing. Plant Cell **13**: 2823-2839.

Schumann, W. (2007). Thermosensors in eubacteria: role and evolution. Journal of Biosciences **32**: 549-557.

Schumann, W. (2012). Sensing in Nature 1st ed. (Springer: New York).

Sengupta, P., & Garrity, P. (2013). Sensing Temperature. Current Biology **23**: R304-R307.

Servant, P., & Mazodier, P. (2001). Negative regulation of the heat shock response in *Streptomyces*. Archives of Microbiology **176**: 237-242.

Shaffer, S. A., Harvey, M. D., Goodlett, D. R., & Ernst, R. K. (2007). Structural heterogeneity and environmentally regulated remodeling of *Francisella tularensis* subspecies novicida lipid A characterized by tandem mass spectrometry. Journal of the American Society for Mass Spectrometry **18**: 1080-1092.

Shapiro, R. S., & Cowen, L. E. (2012). Thermal control of microbial development and virulence: molecular mechanisms of microbial temperature sensing. mBio **3**.

Siegenthaler, R. K., & Christen, P. (2006). Tuning of DnaK chaperone action by non-native protein sensor DnaJ and thermosensor GrpE. Journal of Biological Chemistry **281**: 34448-34456.

Singh, P., Kumari, S., Guldhe, S., Misra, R., Rawat, I., & Bux, F. (2016). Trends and novel strategies for enhancing lipid accumulation and quality in microalgae. Renewable and Sustainable Energy Reviews **55**: 1-16.

Spiess, C., Beil, A., & Ehrmann, M. (1999). A Temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. Cell **97**: 339-347.

Staron, A., Sofia, H. J., Dietrich, S., Ulrich, L. E., Liesegang, H., & Mascher, T. (2009). The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) σ factor protein family. Molecular Microbiology **74**: 557-581.

Storz, G. (1999). An RNA thermometer. Genes & Development **13**: 633-636.

Townsley, L., Sison Mangus, M., Mehic, S., & Yildiz, F. (2016). Response of *Vibrio cholerae* to low-temperature shifts: cspv regulation of Type VI Secretion, biofilm formation, and association with zooplankton. Applied and Environmental Microbiology **82**: 4441-4452.

Török, Z., Goloubinoff, P., Horváth, I., Tsvetkova, N. M., Glatz, A., Balogh, G., Varvasovszki, V., Los, D. A., Vierling, E., Crowe, J. H., & Vigh, L. (2001). *Synechocystis* HSP17 is an amphitropic protein that stabilizes heat-stressed membranes and binds denatured proteins for subsequent chaperone-mediated refolding. Proceedings of the National Academy of Sciences **98**: 3098-3103.

Vigh, L., Maresca, B., & Harwood, J. L. (1998). Does the membrane's physicalstate control the expression ofheat shock and other genes?. Trends in Biochemical Sciences **23**: 369-374.

Von Kampen, J., Nieländer, U., & Wettern, M. (1995). Expression of ubiquitin genes in *Chlamydomonas reinhardtii* : involvement in stress response and cell cycle. Planta, **197**: 528-534.

Waldminghaus, T., Heidrich, N., Brantl, S., & Narberhaus, F. (2007). FourU: a novel type of RNA thermometer in *Salmonella*. Molecular Microbiology **65**: 413-424.

Walsh, N. P., Alba, B. M., Bose, B., C.A, G., & Sauer, T. R. (2003). OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. Cell **113**: 61-71.

Wang, X., & Quinn, P. J. (2010). Lipopolysaccharide: Biosynthetic pathway and structure modification. Progress in Lipid Research **49**: 97-107.

White-Ziegler, C. A., & Low, D. A. (1992). Thermoregulation of the pap operon: evidence for the involvement of RimJ, the N-terminal acetylase of ribosomal protein S5. Journal of Bacteriology **174**: 7003-7012.

Yamanaka, K., Mitta, M., & Inouye, M. (1999). Mutation analysis of the 5′ untranslated region of the cold shock *cspA* mRNA of *Escherichia coli*. Journal of Bacteriology **181**: 6284-6291.

Yazdankhah, S. P., & Caugant, D. A. (2004). *Neisseria meningitidis*: An overview of the carriage state. Journal of Medical Microbiology **53**: 821-832.

Young, D. B., & Garbe, T. R. (1991). Heat shock proteins and antigens of *Mycobacterium tuberculosis*. Infection and Immunity **59**: 3086-3093.

Zangrossi, S., Briani, F., Ghisotti, D., Regonesi, M. E., Tortora, P., & Dehò, G. (2000). Transcriptional and post-transcriptional control of polynucleotide phosphorylase during cold acclimation in *Escherichia coli*. Molecular Microbiology **36**: 1470-1480.

Zhang, N., Belsterling, B., Raszewski, J., & Tonsor, S. (2015). Natural populations of *Arabidopsis thaliana* differ in seedling responses to high temperature stress. AOB Plants **7**.

Zhao, K., Liu, M., & Burgess, R. R. (2005). The Global transcriptional response of *Escherichia coli* to Induced σ^{32} protein involves σ^{32} regulon activation followed by inactivation and degradation of σ^{32} *in vivo*. Journal of Biological Chemistry **280**: 17758-17768.

Toward elucidating the function of a plant-specific transcription factor, OsTCP17, in rice

Thesis Submitted to the University of Delhi
in the partial fulfillment for the degree of
Master of Science
in
Plant Molecular Biology and Biotechnology
2021

K. THANDUANLUNG

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi 110021, India

Toward elucidating the function of a plant-specific transcription factor, OsTCP17, in rice

Thesis Submitted to the University of Delhi
in the partial fulfillment for the degree of
Master of Science
in
Plant Molecular Biology and Biotechnology
2021

K. THANDUANLUNG

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi 110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Toward elucidating the function of a plant-specific transcription factor, OsTCP17, in rice**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus by **K. Thanduanlung** and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates thatthe candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 28th June 2021

Signature of the Candidate

(K. THANDUANLUNG)

Signature of the Supervisor

(PROF. AKHILESH K. TYAGI)

Signature of the HOD

(PROF. INDRANIL DASGUPTA)

Benito Juarez Road, New Delhi. 110021, India. Tel. 24111208, Fax: 24110669, Website: www.dpmb.ac.in

Preface

Rice is one of the most important cereal crops in the country and our economy depends immensely on rice as the majority of the population is a daily consumer of rice. Like many other crops, rice production is hugely affected by many environmental factors such as drought and salinity, flood, cold, pathogens (bacteria, fungi, and viruses), and pests. Besides, environmental factors like nutrients contents in the soil, rainfall, temperature, or duration of sunlight influence the extent of growth and development. However, the full potential of the plants' growth and development or the overall architecture of the plants is determined by its genetics components, and therefore, manipulation of its genetics components or factors that influences useful traits provides an efficient way of improving rice yields.

TEOSINTE BRANCHED1, **C**YCLOIDEA, and **P**ROLIFERATION OF NUCLEAR CELL ANTIGENS (PCF1/2), TCP, is a family of plant-specific transcription factors that plays an important role in the regulation of plant growth dynamics. Their potential to modulate several important traits by modulating a single endogenous factor makes TCPs a potential candidate for modulating plant architecture in a race towards developing sustainable rice varieties.

OsTCP17, a class I TCP, has overlapping expressing patterns with *OsTB1* (a known branching gene) i.e., specific to seedling and panicles stages. However, in contrast to *OsTB1*, *OsTCP17* shows the highest expression levels in panicle 5 stage, and its product is also predicted to have both chloroplast localization signal (transit peptide) and nuclear localization signal (NLS). Further, this gene seems to have a remarkable expression pattern vis-a-vis the panicle-specific gene SPL14/IPA1, which is known to regulate *OsTB1* and influence the panicle numbers. However, till date, no literature is available for functional aspects of OsTCP17. Hence, this work aims to elucidate the functional aspects of OsTCP17 in rice using certain molecular approaches available.

This work comprises a review of previous literature, in-silico analysis, as well as molecular cloning work including preparation of overexpression (OE) and CRISPR-cas9 knockdown constructs. Much work remains to be done on OsTCP17 spatiotemporal regulation, its interacting partners, predicted interacting factors binding to its cis-regulatory elements and its target genes, along with its function in rice.

Acknowledgements

The completion of this work could not have been possible without the participation and assistance of so many people whose names may not all be enumerated. Their contributions are sincerely appreciated and gratefully acknowledged.

First of all, I like to give thanks and express my sincere gratitude to my esteem teacher and supervisor Professor Akhilesh K. Tyagi for his guidance, availability, commitment and for priming me in every aspect of research work and at the same time allowing me the space much needed for me to finish my dissertation work.

I also thank my mentor Dr. Priyanka Deveshwar for her guidance and for giving me the space I greatly desired. I also acknowledge and express my gratitude to Mr. Shivam, Ms. Ankita and Mr. Anurag Panchal for their guidance, support and help in every possible way throughout my dissertation work.

I also express my sincere gratitude to all my teachers namely Prof. Anil Grover, Prof. J.P. Khurana, Prof. Paramjit Khurana, Prof. Indranil Dasgupta, Prof. Arun K. Sharma, Prof. Girdhar K. Pandey, Prof. Sanjay Kapoor, Prof. Saurabh Raghuvanshi, and Dr. Surekha K. Agarwal for helping me understand the concept of the molecular biology coursework, their constant support and guidance throughout the M.Sc. course.

I would also like to acknowledge the department as a whole and the University of Delhi for providing the opportunities and extending possible support to carry out our dissertation work despite the hardship faced by everyone in midst of the pandemic.

I also acknowledge the contribution of lab assistants Mr. Dharmendra and Mr. Ahmed in organising the lab, availability and help in every possible way.

I owe a deep sense of gratitude to my brother Guang who supported me morally and financially and my expression will not be complete if I failed to acknowledge his kind and sincere contribution for me.

I give thanks and express my gratitude to all my awesome friends for being an immense help to me, for their time and availability when I needed them and their encouragement that made me feel that I am not alone on this learning journey. Last but not the least; I would like to acknowledge our Almighty God, the author of Knowledge and Wisdom, for His countless love and help.

ABBREVIATIONS

1. µg - Microgram
2. µM – Micromolar
3. bp - Base pair
4. ˚C - Degree Celsius
5. cDNA - Complementary DNA
6. CHE- CCA1 HIKING EXPEDITION
7. CRISPR- Clustered Regularly Interspaced Short Palindromic Repeats
8. C-terminus - Carboxy terminus
9. DNA - Deoxyribonucleic acid
10. g - Gram
11. hrs. - Hours
12. l - Litre
13. LB - Luria Bertani
14. MCT - Micro-centrifuge tube
15. ml - Millilitre
16. mM - Millimolar
17. mRNA - Messenger Ribonucleic acid
18. N-terminus - Amino terminus
19. NEB - New England Biolabs, USA
20. nm - Nanomolar
21. OD - Optical Density
22. OE - Over-expression
23. PB-1 - Pusa Basmati 1
24. PCR - Polymerase Chain Reaction
25. RNA - Ribonucleic acid
26. rpm - Revolutions per minute
27. RT - Room temperature
28. RT PCR - Real Time PCR
29. sgRNA – Single guide RNA
30. SPL – Squamosa promoter-binding-like protein
31. TCP – Teosinte, Cycloidea, and proliferating cell antigens, (PCF)
32. TSS – Transcription start site
33. WT - Wild-type

TABLE OF CONTENTS

CERTIFICATE ..... **II**
PREFACE ..... **III**
ACKNOWLEDGEMENTS ..... **IV**
ABBREVIATIONS ..... **V**

CHAPTER 1. REVIEW OF LITERATURE **7-30**
1.1 INTRODUCTION 8
1.2 EXPRESSION PATTERN 9
1.3 ELUCIDATION OF TCP STRUCTURE 10
1.4 TCPs AND THE DEVELOPMENT OF RICE 11
1.4.1 Regulation of tillers 11
1.4.2 Floral morphology 14
1.5 ROLE OF TCP IN STRESS RESPONSE 17
1.5.1 Abiotic stress 17
1.5.2 Biotic stress 20
1.6 TCP AND DOMESTICATION OF RICE 21
1.7 COMPARATIVE FUNCTIONAL ASPECTS OF TCP GENES 23
1.7.1 Class I (TCP-P) 24
1.7.2 Class II (TCP-C) 24
1.8 CONCLUSION AND FUTURE PROSPECT 28

CHAPTER 2. MATERIALS AND METHODS **31-46**
2.1 IN SILICO WORK 32
2.2 BIOLOGICAL MATERIALS AND THEIR GROWTH CONDITIONS 33
2.3 PREPARATION OF SUBCELLULAR LOCALIZATION CONSTRUCT 34
2.4 *AGRO*-INFILTRATION 37
2.5 OVEREXPRESSION CONSTRUCT 38
2.6 CRISPR-CAS9 KNOCKDOWN CONSTRUCT 41
2.7 PREPARATION OF RICE TRANSGENICS 43
2.8 MEDIA COMPONENTS 45
2.9 PRIMERS 46

CHAPTER 3. RESULTS AND DISCUSSION **47-65**
3.1 IN SILICO ANALYSIS 48
3.2 OSTCP17 SUBCELLULAR LOCALIZATION 53
3.3 OVEREXPRESSION CONSTRUCT 58
3.4 KNOCKDOWN CONSTRUCT 61
3.5 DISCUSSION 64

CHAPTER 4. CONCLUSIONS AND FUTURE PROSPECTS **66-67**

BIBLIOGRAPHY **68-72**

CHAPTER 1

Review of Literature

1.1 INTRODUCTION

Rice is an important cereal crop as more than 50% of the global population consumes rice as their staple food (Khush, 2005; Jeon et al., 2011; Elert, 2014). Therefore, improving grain productivity or yield per unit area is of major interest for many plant breeders and the scientific community across the globe. Moreover, elite rice varieties planted in many areas and regions are highly sensitive to salt stress, common stress in modern times due to excessive irrigation and fertilization, drought stress, or submergence stress due to flooding in many low-lying areas and are susceptible to many pathogens and pests. Many efforts have been put forth in developing sustainable varieties for withstanding major biotic and abiotic stresses. One of them is by manipulation of endogenous transcription factors as transcription factors control intricate networks of regulation. This helps temper plant architecture and many other traits that play a major role in the development and can enhance crop yields. Hence, genes that control such novel characters become the prime interest of studies. One class of such factors that stands out is the plant-specific transcription factor TCP family.

TCP, a plant-specific transcription factor family, has been named after the first three characterized members, namely, **T**EOSINTE BRANCHED1/TB1 from maize, **C**YCLOIDEA/CYC from *Antirrhinum majus,* and **P**CF1/**P**CF2 from rice (Kosugi and Ohashi, 1997; Cubas et al., 1999; Navaud et al., 2007; Dhaka et al., 2017). Initially, this TF family was characterized to have a non-canonical basic helix-loop-helix (bHLH) domain, however, a recent report confirmed that the crystal structure of OsPCF6 shows TCP domain to have high structural similarity with RHH (Ribbon-helix-helix transcriptional repressor family) rather than bHLH (Sun et al., 2020a). Besides, additional R domain and bipartite nuclear localization signal (NLS) have also been reported in some subset of class II TCP proteins such as OsTCP10, OsTB1, 20, and 21 (Cubas et al., 1999; Martín-Trillo and Cubas, 2010). So far, 26 *TCP* genes have been reported in rice (Martín-Trillo and Cubas, 2010; Sharma et al., 2010). The length of TCP proteins in rice ranges from 141-480 amino acids with its TCP domain at the N-terminal region. Subcellular localization analysis indicates TCPs to be localized mainly in the nucleus with few exceptions that are predicted to have localization in other organelles (Sharma et al., 2010). Phylogenetically, two classes of TCP can be identified (Table 1). The role of TCPs in plant development and signaling network has been reviewed and extensively worked out in *Arabidopsis thaliana* (Navaud et al., 2007; Danisman, 2016; Nicolas and Cubas, 2016; Dhaka et al., 2017; Lan and Qin, 2020) but much remains to be done in rice. This review will mostly focus on rice TCPs.

Table 1: Types of TCPs and their binding sequences (Martín-Trillo and Cubas, 2010; Dhaka et al., 2017)

Class	Binding site	Associated with	Subclass
I (PCF class/ TCP-P)	GGNCCCAC	Activation	-
II (TCP-C)	GTGGNCCC	Repression	i) CINCINNATA (CIN) ubiquitous ii) CYC/TB1 restricted to angiosperms

1.2 EXPRESSION PATTERN

Spatiotemporal expression analysis done by Sharma et al. (2010) shows TCPs to have diverse expressions pattern in different tissues (Table 2).

Table 2: Expression pattern of all characterized rice TCPs [source: Sharma et al., (2010)]

S.no	TCP genes	Tissue
1	*OsTCP1, PCF2 & 3, OsTCP6,* and *OsTCP8*	Ubiquitously expressed in all the tissues
2	*OsTCP11* and *OsTCP17*	Preferential accumulation in panicles
3	*OsTCP7, 9, 19 & 21, PCF5, 6* and *PCF7*	Seedling, roots, SAM, mature leaf, panicles, early seed developmental stages as well as ovary, stigma, embryo and endosperm
4	*OsTCP8* and *PCF3*	Young vegetative and reproductive stages, roots and early panicles stages
5	*OsTB1* and *OsTCP17*	Specific to seedling and panicles
6	*OsTCP10, 22* and *PCF8*	Mature pollen
7	*OsTCP25*	Specific to endosperm

Assuming the relative abundance of the transcripts at a particular stage correlates with the function of the genes, TCPs seem to have a diverse range of functions. Besides, the significant numbers of intronless genes among the TCP family indicate their role that involves rapid regulation in response to an ever-changing environment (Sharma et al., 2010).

1.3 ELUCIDATION OF TCP STRUCTURE

TCP family members share a conserved ~60 amino acid residues domain called the TCP domain apart from other domains like the R domain present on some subset of class II TCP. Secondary structure prediction shows TCP to form non-canonical basic helix-loop-helix domain (bHLH) which facilitates protein-protein interaction and upon dimerization, it binds target DNA (sequence GGNCCC) (Kosugi and Ohashi, 1997, 2002; Cubas et al., 1999; Aggarwal et al., 2010; Martín-Trillo and Cubas, 2010; Danisman, 2016; Nicolas and Cubas, 2016). However, a recent report suggests that TCP be renamed as RHH family as OsPCF6, a member of rice TCP crystal structure resembles more of Ribbon helix-helix (RHH; DNA binding transcriptional repressor family) rather than that of bHLH protein. The structural comparison shows both OsPCF6 and CopG (RHH protein) shared similar 3D structures (although not that similar in amino acid sequence per se) with both using basic β-strand for interacting with the DNA. The structural superimposition fits well with the CopG (characterized RHH protein) while it does not fit well with the canonical bHLH protein MyoD. Also, the said basic region was reported to form three short β-strands (β1, β2, and β3) and two α helices forming HTH motif in OsPCF6. However, in CopG, instead of three interrupted β sheets it has a long antiparallel β-sheet (Sun et al., 2020a). Further, critical residues for TCP dimerization and DNA binding were characterized through site-directed mutagenesis and isothermal titration calorimetry (ITC) assay where a mutation in R16A, R17A, and R19A abolished the DNA binding capacity of TCP (Aggarwal et al., 2010; Sun et al., 2020a). Dimerization seems to be a necessity for DNA binding as the mutations in residues necessary for dimerization abolished the DNA binding capacity (Kosugi and Ohashi, 2002; Aggarwal et al., 2010). TCP can form either homodimer or heterodimer, e.g. PCF 1 and 2 (Kosugi and Ohashi, 1997), however, there seems to exist a preference for dimer formation of class I TCP with class I or class II with Class II depending upon the presence of glycine in position 11 (class I) or aspartate in position 15 for class II (Viola et al., 2012). Besides, depending upon its interacting partner, the preference of binding site may, however, alter; for example, TCP20 heterodimerized with either TCP18/21 binds to the promoter of *bHLH39* whereas homodimerized TCP20 does not (Andriankaja et al., 2014; Danisman, 2016).

1.4 TCPs AND THE DEVELOPMENT OF RICE

Development involves several factors such as positional information, boundaries, and regulatory switches working in concert to determine each cell fate that together forms the well-organized structure that is seen as a plant. Positional information can arise from simple signals moving away from the source, selective inhibition of certain mRNA by miRNAs, directed movement or asymmetric distribution of movement proteins, etc., Boundaries specification and determinate cell growth are necessary for proper development, e.g., the formation of abaxial and adaxial side of leaves which is determined by mutual antagonistic boundary genes action or floral transition (SAM to floral meristem). Therefore, the mechanism of action of the regulatory switches involved in intricate regulatory networks needs to be deciphered.

TCPs have an evolutionarily conserved role in diverse plant species (Danisman, 2016). Its prime members, PCF1 and PCF2, were believed to be the regulator of PCNA (proliferating cell nuclear antigen gene) by binding to its promoter and were the first reported to form homodimers as well as heterodimers in rice (Kosugi and Ohashi, 1997). Further studies on TCP genes affecting rice development are discussed below.

1.4.1 Regulation of tillers

Tillering in rice is controlled by multiple genes (multigenic trait) and a complex network of phytohormones and transcription factors such as auxin, brassinosteroids (BR), cytokinin (CK), strigolactone (SL), and D14, MADS, TCP, SPL14/IPA1, and OsCCA1, respectively (Figure 1).

Similar to maize TEOSINTE BRANCHED1/TB1, *OsTB1* negatively regulates lateral branching in rice, mainly the outgrowth, as the mutant still has the bud in its original position (Takeda et al., 2003). Overexpression of *OsTB1* reduces tiller number while allelic mutant *fc1* (loss of *OsTB1* function) increases tiller number (Takeda et al., 2003; Minakuchi et al., 2010; Guo et al., 2013).

FINE CULM (FC1), the allele of *OsTB1,* works downstream of the strigolactone-induced bud inhibition pathway and acts as an integrator in multiple pathways (Minakuchi et al., 2010). Also, OsTB1/FC1 integrates brassinosteroids pathway antagonistic to SL pathway (Umehara et al., 2008; Fang et al., 2020). However, both these phytohormones exert their function by stabilizing the OsBZR1-DLT-RLA1 module (Figure 2).

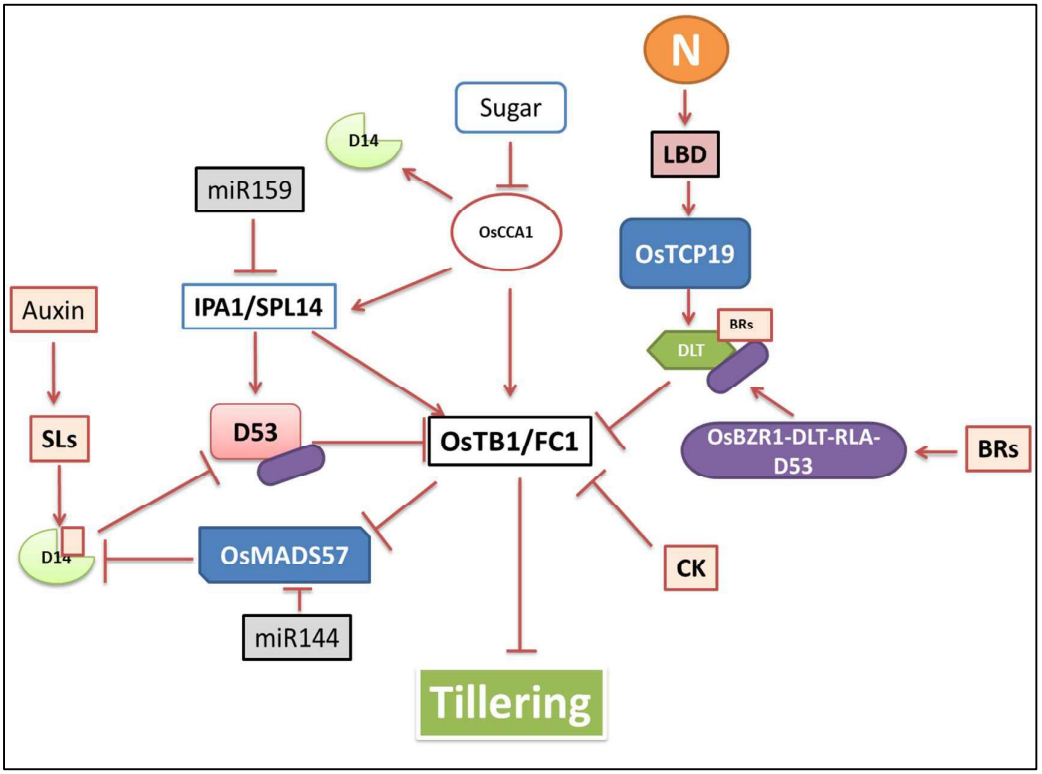


Figure 1: Regulatory network of tillering in rice.

TCP coordinates different regulatory factors so as to modulate and maintain plant architecture. OsTCP19 induces DLT which forms a complex with OsBZR1-RLA (binds to the promoter of *OsTB1*) and recruits D53 to inhibit *OsTB1* expression. This complex is stabilized by BRs and thus BRs induce tillering. OsCCA1 binds and induces the expression of *OsTB1, D14* and *SPL14* which in turn induce *D53* and *OsTB1*. D14, upon binding SLs, facilitates the degradation of OsBZR1 and D53 (D53 recruits TPL to the promoter and inhibit *OsTB1*). This allows *OsTB1* to express in the presence of SL, resulting in reduced tillering. Further, the SLs-induced OsTB1/FC1 sequesters OsMADS57 away from binding to *D14* cis-element thereby relieving its negative effect on *D14* expression and increases *D14* expression, creating a positive feedback loop to further facilitate SL perception. This shows OsTB1/FC1 as the main regulatory hub of tillering in rice.

Arrow indicates induction while cross bar represents inhibition and question mark indicates uncharacterized factor. BRs, brassinosteroids; CK, cytokinins; DLT, DWARF AND LOW-TILLERING; IPA1/SPL14, IDEAL PLANT ARCHITECTURE/SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE-14; LBD, LATERAL ORGAN BOUNDARY DOMAIN protein; N, nitrogen; OsBZR1, BRASSINAZOLE RESISTANT 1; SLs, strigolactones. (Adopted from *Takeda et al., 2003; Minakuchi et al., 2010; Guo et al., 2013; Lu et al., 2013; Song et al., 2017; Fang et al., 2020; Wang et al., 2020; Liu et al., 2021*).

D14 is a classified SL perception protein (Arite et al., 2009; Hu et al., 2010) and is induced by strigolactones (SL) and OsCCA1 while OsMADS57 inhibited its expression (Minakuchi et al., 2010; Guo et al., 2013; Wang et al., 2020). OsTB1 sequestered OsMADS57, making it unavailable to bind to its target *cis*-regulatory elements i.e., *D14* promoter. Since, binding of OsMADS57 inhibits the expression of *D14*, sequestering of OsMADS57 away from the *D14* promoter by OsTB1 increases the expression of *D14* and subsequently increases SL sensitivity, thereby inhibiting further tillering. Therefore, OsTB1/FC1 acts as a negative

regulator in SL induce bud inhibition pathway and OsMADS57 as a positive regulator (Guo et al., 2013).

Besides, *OsTB1* might also involve in floral organ initiation as indicated by a gradual increase in its transcript levels from SAM to early panicle development (Sharma et al., 2010). It is evident that tempering such genes has the potential of altering the plant architecture for improving grain yield. As a result, the TCP genes have become one of the prime targets of studies (Sharma et al., 2010) because secondary and tertiary tiller reduction may reduce the nutrient competition between primary tillers which is known to contribute to grain yield (Jeon et al., 2011). In addition, *OsTB1* might have an overlapping function with *OsTCP17* as it shares a similar pattern of expression (Sharma et al., 2010).

Use of nitrogen as fertilizers is a widely practiced system to improve grain yields. However, the molecular mechanism of its action remains obscure. Recent findings added new complexity that involves TCP in the regulation of tillers where TCP coordinates environmental factors (nitrogen content in the soil) with two phytohormones strigolactones and brassinosteroids to either enhance or reduce tillers outgrowth in rice (Figure 2).

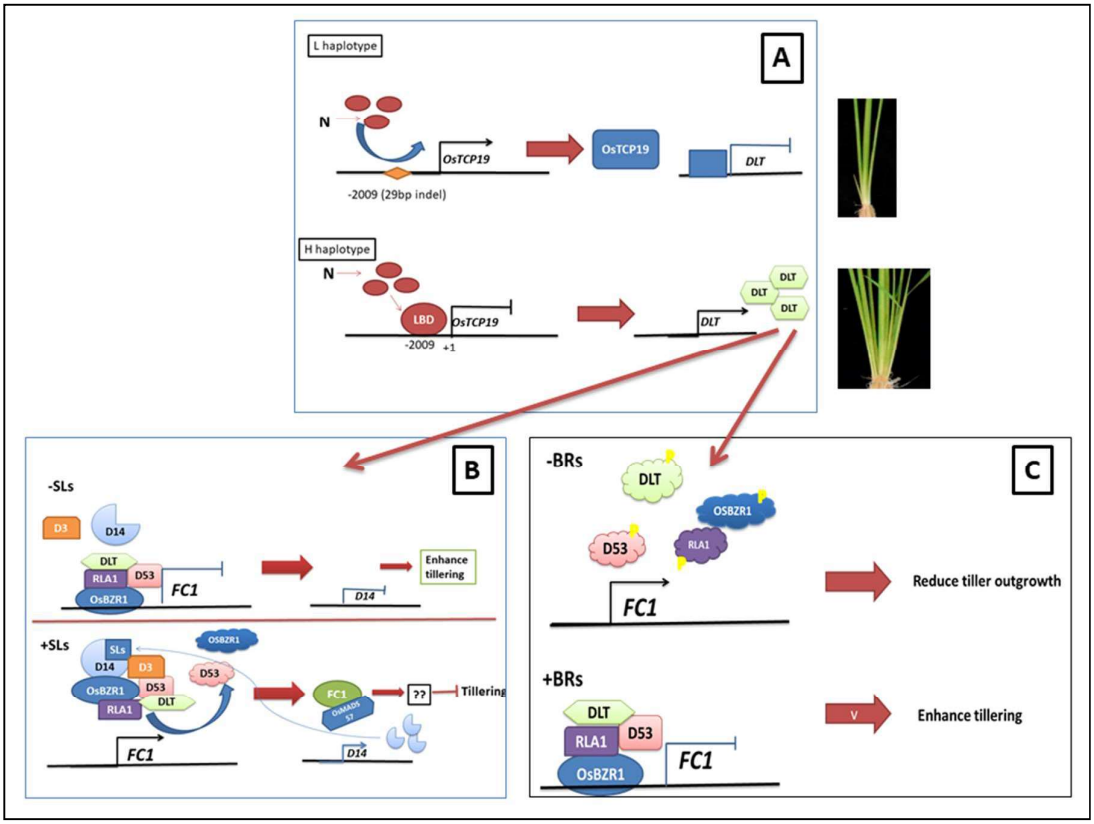


Figure 2: Working model of TCP regulated tillering response in rice.

A) Different landraces respond to nitrogen differently. This variation is brought about by the variation in the expression levels of *OsTCP19 as* a result of 29 bp indel in the promoter region (-2009 bp upstream of start site). Two haplotype group were classified; L haplotype having 29 bp indel associate with low tiller response to nitrogen and H-haplotype with high tiller response to nitrogen (most wild rice and Australian cultivar). Nitrogen induces LBD (LATERAL ORGAN BOUNDARY DOMAIN) protein and represses the expression of *OsTCP19* thereby allowing the expression of *DLT*. The 29-bp indel in L-haplotype *OsTCP19* promoter prevents binding of LBD while binding of LBD to *OsTCP19-H* promoter greatly reduces its expression resulting in the accumulation of DLT. DLT is required in both SLs and BRs signaling pathways to affect tillering.

B) In absence of SL, DLT interacts with OsBZR1 that binds directly to the *FC1/OsTB1* promoter and recruits D53 to inhibit the *FC1* expression. This enhances tillering. On the other hand, SLs induce degradation of D53 and OsBZR1 that results in accumulation of FC1. FC1 in turn sequesters OsMADS57 and prevent it from inhibiting *D14* expression. This increases SL sensitivity and reduces tiller outgrowth.

C) BRs inhibits the activity of OsGSK2 (GLYCOGEN SYNTHASE KINASE 2) to phosphorylate OsBZR1, DLT and RLA1. This stabilized the downstream component proteins (OsBZR1-DLT-RLA1) that in turn inhibits FC1 expression and subsequently enhances tiller outgrowth.

Arrow (black) indicates induction while cross bar represents inhibition and question mark indicates uncharacterized factor. BRs, brassinosteroids; D53, DWARF 53; DLT, DWARF AND LOW-TILLERING; *FC1, FINE CULM 1*; LBD, LATERAL ORGAN BOUNDARY DOMAIN protein, N, nitrogen; OsBZR1, BRASSINAZOLE RESISTANT 1; RLA1, REDUCED LEAF ANGLE 1; SLs, strigolactones. (Adopted from *Takeda et al., 2003; Minakuchi et al., 2010; Guo et al., 2013 Fang et al., 2020; Liu et al., 2021*).

1.4.2 Floral morphology

Lemma and palea are two protective organs that cover rice floret and ensure proper development of grain. Therefore, defect in any of these structures may have a deleterious effect on grain yield. Also, lemma and palea in some way enforce floral zygomorphy in rice. The first clue of floral zygomorphy regulation in rice comes from the identification of a mutant that has a defect in palea development, namely *rep1* mutant. *Rep1* mutant, however, shows no defect on any other organs and developed normally except a delay in the development of palea primordia supported by its asymmetric expression pattern indicating *REP1* mutation retards or suppresses palea development and subsequently the floral zygomorphy of rice (Yuan et al., 2009).

REP1 (RETARDED PALEA1) encodes OsTCP24, a plant-specific transcription factor known to enhance cell proliferation. REP1 physically interacts with DP1 (AT-hook protein) and regulates palea development and subsequently the floral zygomorphy by enhancing cell expansion and differentiation (Yuan et al., 2009; Yin et al., 2018).

OsPCF7 regulates a plethora of rice developmental processes such as growth, rice architecture or to be precise, the overall morphology of rice. OsPCF7 is a CIN clade of class II TCP of about 270 amino acids long and it is ubiquitously expressed in all the cell/tissues of rice (however, it shows a dramatical increase in mRNA levels in tillering stage and heading

stage in panicles). *OsPCF7* is characterized as a single dominant gene that controls rice panicle numbers, tillers numbers, height, and ultimately the grain yield by upregulating genes involved in basic developmental function such as boundary-specific genes, hormone regulation, and biosynthesis genes; *SHOOTMERISTEMLESS/STM, KNAT2, 6 (KNOX* genes) and *AS1* (ARP genes), *LOX2* (JA biosynthesis-related) and *IAA3* (auxin-related, – known to facilitate root growth) as shown in Figure 3.

As expected, overexpressing this single dominant gene resulted in higher shoot height, root numbers, and longer roots lengths, panicles numbers, and filled grain numbers per plant ultimately affecting the yield of the rice since the yield of the crops is determined by the numbers of tillers and panicle/plant (numbers of tillers and panicles influence grain number), grain number/plant and thousand-grain weight and numbers of filled grains (Li et al., 2020).

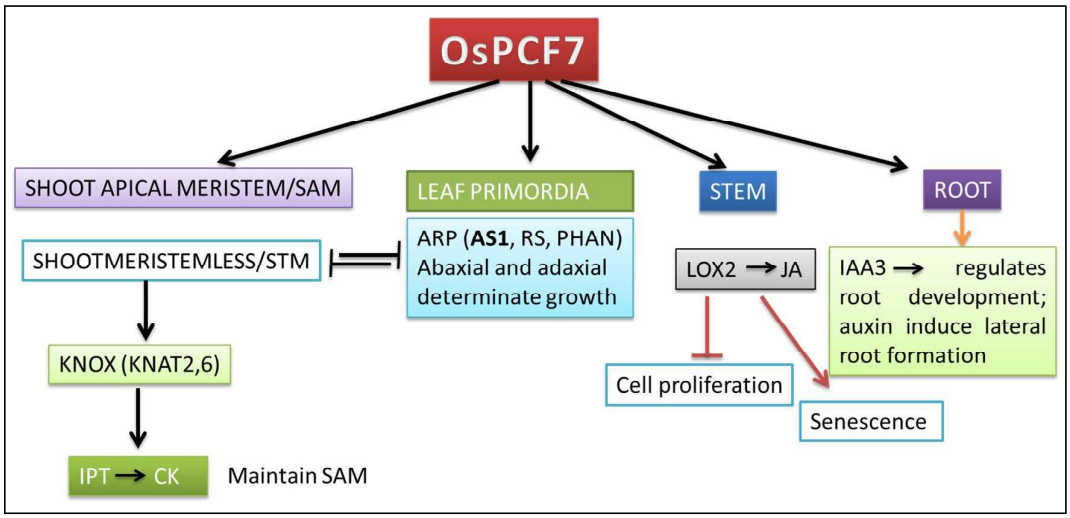


Figure 3: Diverse regulatory network of OsPCF7.

Induction of *STM* and *AS1* (leaf boundary specific genes) facilitates cell differentiation and at the same time upregulation of *LOX2* further promotes leaf maturity. Besides, *AS1* also influences inflorescence development. IAA is important for vascular differentiation and other auxin responses such as root system development, growth, hypocotyl elongation, leaf initiation and expansion, etc. Overall, OsTCP7 influences the proper development and architecture of the plant. Arrow indicates induction and cross bar indicates inhibition. AS1, ASSYMETRIC LEAVES1; CK, cytokinins; JA, jasmonic acid. *(Adopted from Li et al., 2020)*

OsTCP5, another CIN clade member of class II TCP, negatively regulates mesocotyl elongation in rice through the interaction of two novel phytohormones, cytokinins and strigolactone; CK inhibits *OsTCP5* expression leading to increase in cell numbers in mesocotyl whereas SL induces its expression and inhibits mesocotyl elongation (Hu et al., 2014; Figure 4).

Moreover, Ren et al. (2018) proposed a novel regulatory role of OsTCP5 in the acquisition of reproductive cells fate from somatic cells (transition from mitosis to meiotic division) via

SPOROCYTELESS regulatory pathway where OsSPL physically interacts with OsTCP5 and recruits OsTPL2 and OsTPL3 to inhibit the expression of OsTCP5 target (downstream) genes.

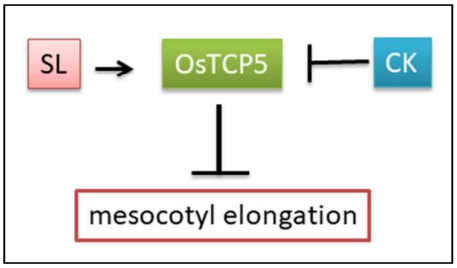


Figure 4: OsTCP5 negatively regulates mesocotyl elongation in rice. (Adopted from *Hu et al., 2014*). CK, cytokinin; SL, strigolactones.

Plants, unlike animals, acquire germline fate *de novo* during postembryonic development. The decision to acquire meiotic cell fate is a pivotal step in plant life, and therefore, involves a series of well-coordinated cell proliferation and differentiation. *MICROSPORELESS1 (MIL1), MEIOSIS ARRESTED AT LEPTOTENE2 (MEL2),* and *SPOROCYTELESS (OsSPL)* are some characterized genes vital in the meiotic entry, of which *SPL* works upstream of the former two. OsSPL contains two important domains, the SDTI domain, required for dimerization and interaction with OsTCP5 and EAR motif, required for interacting with TPL corepressors. Expression analysis shows an overlapping pattern of *OsSPL* with that of *OsTCP5* and *OsTPL3* in the early panicle stage. Moreover, the spatiotemporal expression analysis done by Sharma et al. (2010) shows a 26 folds downregulation of *OsTCP5* in panicle stage 1 as compared to SAM and also upon cold and salt treatment. This indicates that OsSPL links OsTCP5 with the corepressor TPL to repress those genes activated by OsTCP5. In other words, downregulation of *OsTCP5* is necessary to acquire meiotic fate, and that OsTCP5 target gene products might have an antagonistic relationship with those involved in reproductive development. As expected, *Osspl* mutation shows defects in microsporogenesis and megasporogenesis (Ren et al., 2018). If OsSPL acts as a repressor of OsTCP5 in the same pathway, overexpression of *OsTCP5* should have a similar effect but it remains to be tested in rice, and identifying the target downstream genes of OsTCP5 will greatly add to our current understanding.

1.5 ROLE OF TCP IN STRESS RESPONSE

Plants being sessile are constantly exposed to environmental cues. Our crops are experiencing stress (either biotic or abiotic) more than ever. Rice is no exception to such stress and due to excessive irrigation, drought, flood, fertilization, pesticides, etc., the arable land becomes more saline, depleted, or exhausted of available nutrients. Rapid changes in the environment are making it difficult for our crops to keep pace with the changes and they are more prone to pathogen attack, which indeed has been the major cause of crop loss each year. Plants' mechanism of stress tolerance can be thought in terms of signalling and perception, detoxification (metabolizing pathogen toxin, specific inhibition of pathogen genetic material, and downregulation of endogenous susceptibility genes), enhanced stress response genes expression, and evasion/avoidance. In short, tolerance is accomplished by a conglomerate of well-regulated or directed responses in proliferation and differentiation. Plants' response to stress could be of rapid (hypersensitive response) or long-term response.

1.5.1 Abiotic stress

Rice is very sensitive to salt stress, and salinity is often accompanied by drought. Salt stress response includes osmotic tolerance (reduced growth in root tip and shoot, closing stomata, etc.), tissue tolerance (Na^+ compartmentalization into vacuole), and ion exclusion (efflux of ion or compartmentalization). Stress tolerance is multigenic and involves a complex network of regulation which includes hormones, signaling, and perception, and regulatory switches which include mostly transcription factors that switch on several downstream genes upon stress. Under salt stress, *OsPCF2* along with other transcription factor genes like *OsNIN4, OsCCP5,* and *OsNIN-LIKE2* were found upregulated in rice. These positively (*OsPCF2 and OsNIN4*) or negatively (*OsCCP5 and OsNIN-LIKE2*) regulates the expression of *OsNHX1* (the most abundant Na^+, K^+/ H^+ ATPase exchanger localized in the tonoplast) by binding to its promoter, of which OsPCF2 binds the strongest. In addition, *OsPCF2* expression is strongest in root under salt stress indicating higher induction of *OsNHX1* in the root. This results in tolerance as OsNHX1 localized in the tonoplast pumps sodium ions into the vacuole with the exchange of protons, compartmentalizing sodium ions into the vacuole where a high concentration of ions is tolerated. This leads to the osmotic balance in the cells and also facilitates K+/PH homeostasis [Figure 5, (Almeida et al., 2017)].

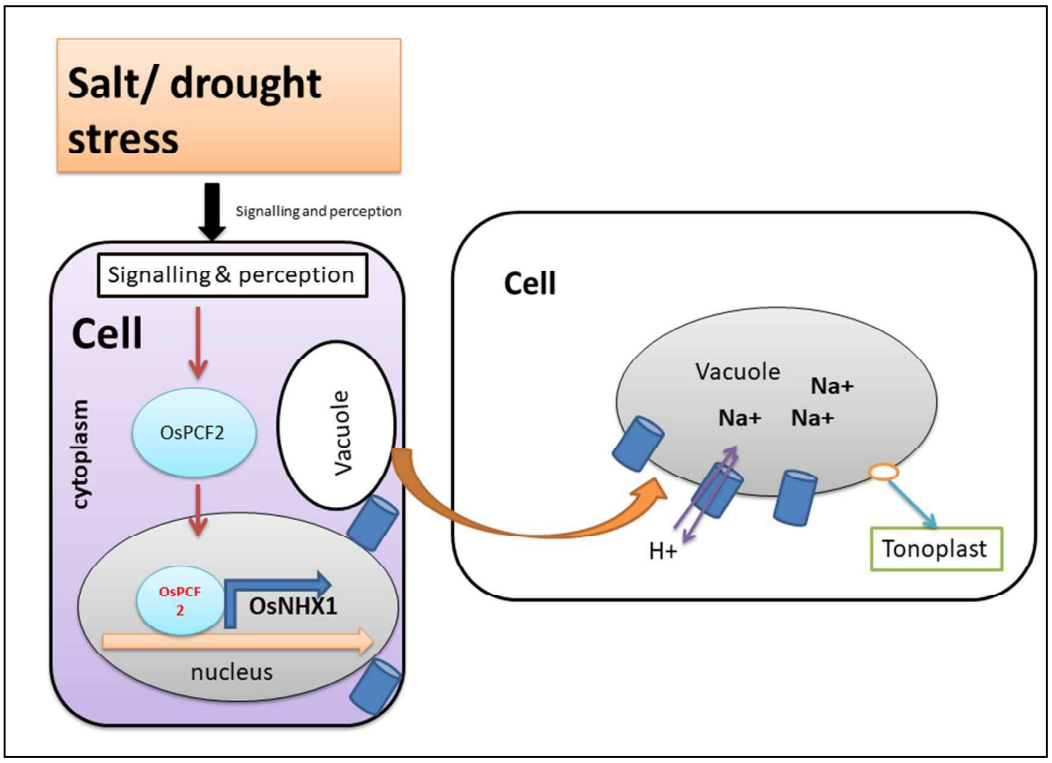


Figure 5: OsPCF2 induced OsNHX1 localized in the tonoplast of root cells and compartmentalized sodium ions (just one aspect of salt tolerance). Blue arrow indicates level of expression and red arrow represents the direction of signalling (Adopted from *Almeida et al., 2017*).

High salt concentration in the soil lowers the water potential making uptake of water by roots difficult because water uptake occurs mainly through osmotic diffusion i.e., from higher concentration to lower concentration. The higher solute level in the soil may even cause loss of water from the root as water tends to move towards lower water potential. Hence, higher induction in roots might facilitate the efficiency of water uptake by the root under stress and in turn influence the survival rate.

TCPs are also shown to be involved in coleoptile growth under oxygen-deprived conditions. Nine TCP genes were differentially expressed in coleoptile of which 8 show tip-specific expression. In addition, other differentially expressed genes like ribosomal proteins and proteins involved in folding show TCP target sites in their cis-regulatory elements indicating TCPs either directly or indirectly regulate these genes under oxygen stress conditions. In compliance with this, analysis shows downregulation of 47/64 differentially expressed r-protein and protein involved in protein folding in coleoptile tip with 52% showing TCP binding site on those tip-specific up-regulated genes (Narsai et al., 2015). The tip-specific transcript abundance of different genes under submergence conditions might be responsible for the rapid coleoptile elongation which is considered as an escape response.

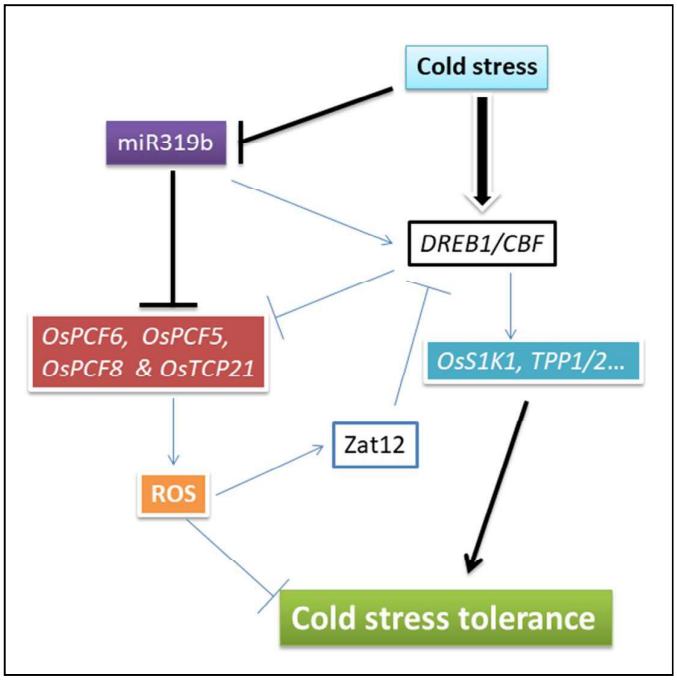


Figure 6: TCP negatively regulates cold tolerance.

Cold greatly reduces the production of miR319 and as a result OsPCF5, 6, 8 & OsTCP21 mRNA level increases. This reduces the expression of TCP target stress related genes rendering the plant sensitive. On the other hand, overexpression of miR319 reduces the level of OsTCPs and subsequently increases stress related genes expression. This results in tolerance. Arrow indicates positive regulation and cross bar indicates negative regulation. Small blue arrow represents possible pathways. DREB1A & 2A, DEHYDRATION-RESPOSIVE ELEMENT BINDING PROTEIN 1A & 2A; OsS1K1, H2O2 scavenging gene; TPP1/2, TREHALOSE-PHOSPHATE PHOSPHATASE 1/2. (Adopted from *Yang et al., 2013; Wang et al., 2014*).

Rice is a tropical crop and, therefore, it is sensitive to low temperatures (chilling 0-15°C and freezing <0°C). Low temperatures induce compaction of water molecules lowering the water potential outside the cell and as a result, water tends to move out and dehydrate the cells causing injury to the cell. *OsPCF5, OsPCF6, OsPCF8,* and *OsTCP21* negatively regulate cold tolerance in rice. *OsPCF5* and *8* expression levels were found to be highest in leaf and root, *OsPCF7* and *OsTCP21* in leaf and young panicles, and *OsPCF6* in young and mature panicles. These TCPs were the natural target of *miR319* and downregulation of these genes by overexpression of *miR319b* showed wider leaf (increase cell numbers), delayed development (entire life cycle), and enhanced cold tolerance after cold acclimation (Figure 6). Downregulated *OsPCF6* and *OsTCP21* lines show increased expression of cold stress-related genes such as *DREB1A, DREB2A, H2O2* scavenging gene, and *OsS1K1* indicating the tolerance may be a result of better ROS scavenging and higher proline accumulation (osmoprotectant) that protect the cell from dehydration due to cold (Yang et al., 2013; Wang et al., 2014).

OsTCP19, up-regulated by drought and cold stress, orchestrates hormonal signaling and aids in abiotic stress tolerance. OsTCP19 under stress condition induces the expression of *ABI4, ABI3, IAA3*, and down-regulates the expression of *LOX2*. Further, OsTCP19 physically interacts with UTL1 a trithorax protein and ABI4 which then activate *DGAT1, 2* and *pDAT1* resulting in the accumulation of lipid droplet. ABI3 mediates laterals roots formation, however, this auxin-mediated LR formation seems to be counteracted by the increase in accumulation of ABI4, IAA3, and downregulation of LOX2 by stress-induced *OsTCP19,* resulting in decreased LR formation. JA-induced accumulation of ROS initially helps stress response; however, uncontrolled accumulation of ROS will damage the cell. Therefore, negative control of JA biosynthesis gene *LOX2* by *OsTCP19* seems in favor of stress adaptation. Thus, *OsTCP19* negatively regulates auxin and JA, on the other hand, *OsTCP19* fine-tunes ABA regulation, thereby acting as a modulator in rice stress responses (Mukhopadhyay and Tyagi, 2015).

OsTB1 orchestrates cold response in rice indirectly by increasing the binding affinity of OsMADS57 to the promoter of *OsWRKY94S2,* enhancing its activation activity under low temperature. However, at normal temperature, *D14* out-competes *OsWRKY94S2* for OsMADS57 binding to its promoter. This results in the inactivation of *D14,* thereby promoting tillering (Guo et al., 2013). On the other hand, the binding affinity of OsMADS57 for *OsWRKY94S2* was higher at low temperature resulting in activation of *OsWRKY94S2* and subsequently enhanced stress response. At the same time, the release of MADS57 from *D14*, in turn, enhanced SL-dependent tillering inhibition (Chen et al., 2018).

In other words, this TCP trades-off plant architecture (i.e., reduce tiller number) for better stress response by enhancing the activation activity of OsMADS57 to promote stress-related gene expression under stressful conditions. Thus, OsTB1 modulates plant architecture and stress response via OsMADS57-D14 and OsMADS57-OsWRKY94S2 pathways.

1.5.2 Biotic stress

Oryza glumaepatula is resistant to leaf and neck blast while the cultivated rice varieties are susceptible. It was found that *pi64(t),* a dominant gene encoding 839 amino acid long two malectin and kinase domains, is responsible for the resistance in wild rice. The sequence analysis found SNPs in the promoter region of the susceptible allele that abolishes the cis-regulatory elements such as MYM element, W-box, TCP binding site, and BIHD10S, making cultivated rice susceptible. The wild rice contains all cis-regulatory elements and resistance indicating a possible role of TCP in the regulation of leaf and neck blast resistance in rice (Devi et al., 2020). Malectin domain is reported to be part of known plant receptor kinase

involved in defense mechanism so the hypothesis is that the malectin domain received the signal from the environment (blast infection) and transduced the signal through the kinase domain and up-regulated stress-related genes. Devi et al. (2020) found up-regulation of ATP binding protein, transcription factor, cell wall protein, and stress-related protein-encoding genes in the resistant rice, implying that TCP (and other associated factors) in response to infection up-regulates *pi64(t)* or precisely malectin and serine-threonine kinase which increases the perception of signal from the external environment (blast infection) forming a regulatory loop (Rajaraman et al., 2016; Bellande et al., 2017; Devi et al., 2020; Sun et al., 2020b).

Rice ragged stunt virus (RRSV), an RNA virus, causes stunted growth, excess tillering, and dark green leaves with a serrated edge or twisted tips in rice. RSV upon infection hijacks the plant regulatory system (miRNAs) by increasing the level of miR319 thereby decreasing the transcripts of its target gene *OsTCP21.*This reduces the level of endogenous jasmonic acid (JA), a phytohormone known to be involved in plant defense mechanism, and signaling related genes to facilitate virus propagation. As expected, overexpression of *OsTCP21* shows milder symptoms, and treatment of external JA reduces virus accumulation in rice. This indicates the role of *OsTCP21* in the rice defense mechanism against virus infection by regulating the jasmonate/JA and related genes/pathway (Zhang et al., 2016).

1.6 TCP AND DOMESTICATION OF RICE

Mutation in *TOESINTE BRANCHED1/TB1* is responsible for the phenotypic differences in the cultivated maize and its wild relative Teosinte meaning mutated *TB1* was selected in maize by humans and has been refined and cultivated since then. Similarly, OsTB1 negatively regulates lateral branching in rice (Takeda et al., 2003; Martín-Trillo and Cubas, 2010).

There are two important subspecies of *Oryza sativa; indica and japonica*. About 90% of the cultivated rice is planted in Asia, with India and China sharing the highest percentile. *Japonica* varieties are adapted to temperate and mountainous regions meaning cooler regions. They are high-yielding rice with narrow dark green leaves, average tillers number and smaller angle and medium plant height with compact plant architecture and sticky when cooked due to high gluten content. Whereas *indica* are adapted to the tropical region and have broad-narrow light green leaves, large tillers angle with medium height except for the semi-dwarf rice and are less sticky. These phenotypic differences were independently selected in the course of agricultural evolution. Plant architecture plays an important role in achieving desired crop yields and this is influenced by tillers numbers, plant height,

inflorescence morphology, and tillers angle; which have been the center of selection in rice simply because tiller angle determines either erect or slender phenotype (spread out as seen in wild rice) with erect posture facilitating compact/dense plantation without compromising the yield; a desirable feature of cultivated crops (Zhang et al., 2019).

TILLER INCLINE GROWTH 1 (TIG1) encoding a single dominant gene, *OsTCP15* CYC clade of type II TCP protein, controls tiller angle in wild rice, *O. sativa rufipogon*. TIG1/OsTCP15 directly regulates the expression of *EXPA3, EXPB5,* and *SAUR39* by binding to their transcription factor binding site (TBS) present in their promoter. Expansins are known to be involved in the proliferation of plant cells and, therefore, high expression of *TIG1* in the adaxial side increases the expression of expansin genes and maintains symmetric growth of the tillers. However, in cultivated *indica* cultivar, it was found that the SNPs in the promoter region of *TIG1/OsTCP15* decrease the mRNA level of *tig1* and its subsequent downstream genes *EXPA3, EXPB5,* and *SAUR39* on the adaxial side resulting in asymmetric growth that is faster growth on the abaxial side and slower growth on the adaxial side making the newly formed tiller to grow upward or erect rather than inclined tillers as observed in wild rice that harbor functional *TIG1* (Figure 7).

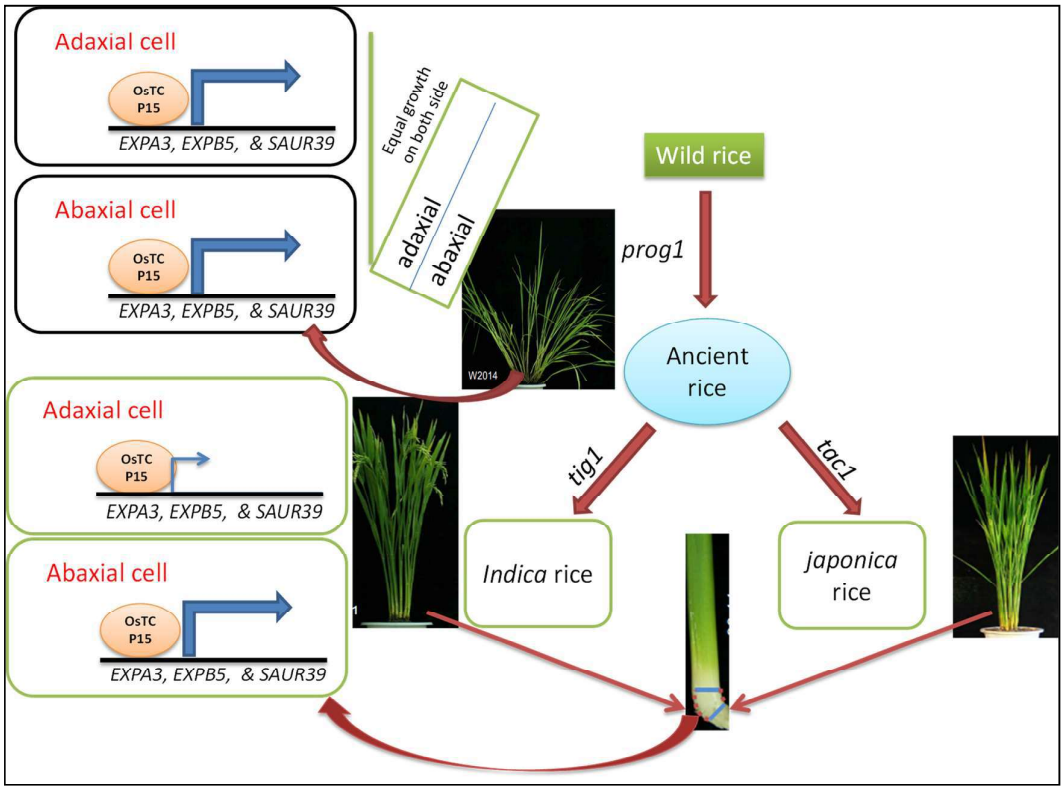


Figure 7: Asymmetric expression of *TIG1/OsTCP15* determines the tillers angle.

In *indica* rice cultivars, mutation in *tig1* resulted in unequal expression of *tig1* target genes. The higher expression of expansin in abaxial side enables the cells in the abaxial side to grow faster as compared to the adaxial side. This results in upward or erect growth of the newly formed tillers. *EXPA 3 & 5, EXPANSINS 3 & 5* and *SAUR39, SMALL AUXIN-UPREGULATED RNA*. The size of the blue arrow represents the level of expression. (Adopted from *Zhang et al., 2019*)

However, the *japonica* cultivar harbors a similar *TIG1* gene as that of the wild rice but SNP that causes splicing error in the *TAC1* gene has a similar effect leading to smaller tillers angle and compact architecture. Thus, *tac1* in *japonica* harboring normal *TIG1* while *indica* cultivar having normal *TAC1* but mutated *tig1* changed the plant architecture from inclined spread-out type to erect architecture making the two sub-species favourable for cultivation. The two corresponding mutated alleles, *tac1* in *japonica* and *tig1* in *indica* cultivars were independently selected by humans and domesticated (Yu et al., 2007; Zhang et al., 2019).

Modern days farming used fertilizers (nitrogen base) to obtained more grain yield. This led to the variations in the nitrogen used efficiency of domesticated rice and wild rice or those that are naturally grown in the region with low nitrogen soil. The genetic basis of these variations is the 29-bp indel in the *OsTCP19* promoter in the H-haplotype group (consists of *japonica* and *indica* varieties) that alters/changes the expression levels of *OsTCP19*. There seems to exist a selection pressure among these landraces planted in nitrogen-rich regions since overproduction of tillers results in low grain yield as a result of trade-off mechanism by the plant for its balanced growth. Moreover, breeder looks for its grain yield and therefore selected the low tiller response to nitrogen (TRN) varieties. As a result, domesticated rice almost lost this *OsTCP19-H* allele. However, this allele is retained in wild rice. Introgression of this lost allele back to the domesticated rice produces the same phenotype as TRN varieties (Liu et al., 2021). Thus, TCP provides an interesting aspect of geographical adaptation to soil nitrogen.

1.7 COMPARATIVE FUNCTIONAL ASPECTS OF TCP GENES

Transcription factors like TCP provide plasticity to the plant in terms of development as well as to adapt and respond to environmental cues. *TCP* genes act like the *first response team* (in our human context) for the plant and have become the central regulatory hub of plant development. As discussed earlier, *TCP* genes are divided into two main classes, class I and class II. Class II is further subdivided into CINCINNATA (CIN)-like TCPs and CYC/TB1-like TCPs, of which CYC subgroup is believed to be evolved from the CIN subgroup. Besides, the TCP family is evolutionarily ancient and is conserved in the plant kingdom; nevertheless, doubt persists whether TCP has a role in early aquatic plants. However, recent findings by Liu et al. (2019) clear this doubt that TCPs are specific to land-dwelling plants and that they evolved in land plants in course of evolution to facilitate dry land adaptation. In line with this, the numbers and diversity of TCP genes increase from lower to higher plants. Also, CYC/TB1 subgroups were found only in angiosperms meaning this subgroup evolved later, independently in eudicot and monocot. Members of TCP family are credited for their role in

regulating diverse aspects of the plants' life. In this section, we will discuss how each gene is functionally related or conserved in different plant species.

1.7.1 Class I (TCP-P)

Comparatively, much of the elucidated function of the TCPs has been on class II TCPs and much work remains to be done on class I TCPs. Similar functions of class I TCP in the regulation of cell proliferation were observed in both monocot (*PCF1/2* of rice) and dicot (*AtTCP20* of Arabidopsis) indicating functional conservation in these two diverse groups (Kosugi and Ohashi, 1997; Li et al., 2005). Besides, PCF2 and OsTCP19 of rice are implicated to have a regulatory role in stress response (Mukhopadhyay and Tyagi, 2015; Almeida et al., 2017), besides class II TCPs being the majority that is involved in stress responses. A recent review by Nicolas and Cubas (2016) summarised nicely the relationship of TCP in the hormonal regulation pathways. In addition, functional similarity between class I and class II TCP groups were also observed, for instance, *AtTCP7, 8, 9, 20, 22 & OsTCP19* (class I) and *OsPCF7* (class II) regulate KNOX genes to promote differentiation, and facilitate boundaries specification (Danisman et al., 2012; Aguilar-Martlnez and Sinha, 2013; Li et al., 2020) and also *CUPULIFORMIS (CUP)* of snapdragon interacts with TCP and establishes lateral organ boundaries in similar with that of Arabidopsis *CUC1,2 & 3* and petunia *NAM* (Weir et al., 2004). So far, class I TCPs have not been reported to be regulated by miRNAs as in the case of class II. Besides, one school of thought that says class II TCP antagonized Class I function seems to be shifting in that class II acts as a heterochronic gene that is spatiotemporally regulated or expressed to facilitate proper development rather than just retarding cell proliferation.

1.7.2 Class II (TCP-C)

CIN-like TCP

The founding member of the CIN-like TCP subgroup, *CINCINATA* gene was identified from *Antirhhinum majus,* in which mutation in *cin* produces abnormal leaves and petals. CIN promotes the expression of a cytokinin receptor homolog *HISTIDINE KINASE 4* (*AmHK4*) and *IAA3/SHY2 (INDOLE-3-ACETIC ACID INDUCIBLE 3/SHORT HYPOCOTYL 2)* in the transition zone, modulating the hormonal pathways to maintain the rate of proliferation and differentiation of the developing leaf (Das Gupta et al., 2014). A similar regulatory pathway is also observed in rice by *OsPCF7* (Li et al., 2020). However, the closest homolog in *Arabidopsis*, AtTCP4 together with BRM induces the expression of *ARR16* (a negative

regulator of cytokinin signaling) to promote determinate leaf growth (Efroni et al., 2013). This shows the involvement of TCP in epigenetic regulation. Also, *BrrTCP2* maintains leaf size in *Brassica rapa* (Du et al., 2017), and manipulation of *BrpTCP4* changes the phenotype of Chinese cabbage heads from round to cylindrical type (Figure 8, Mao et al., 2014).

A similar role of CIN TCPs was also observed in other plants like tomato (Ori et al., 2007) and cotton (Cao et al., 2020).

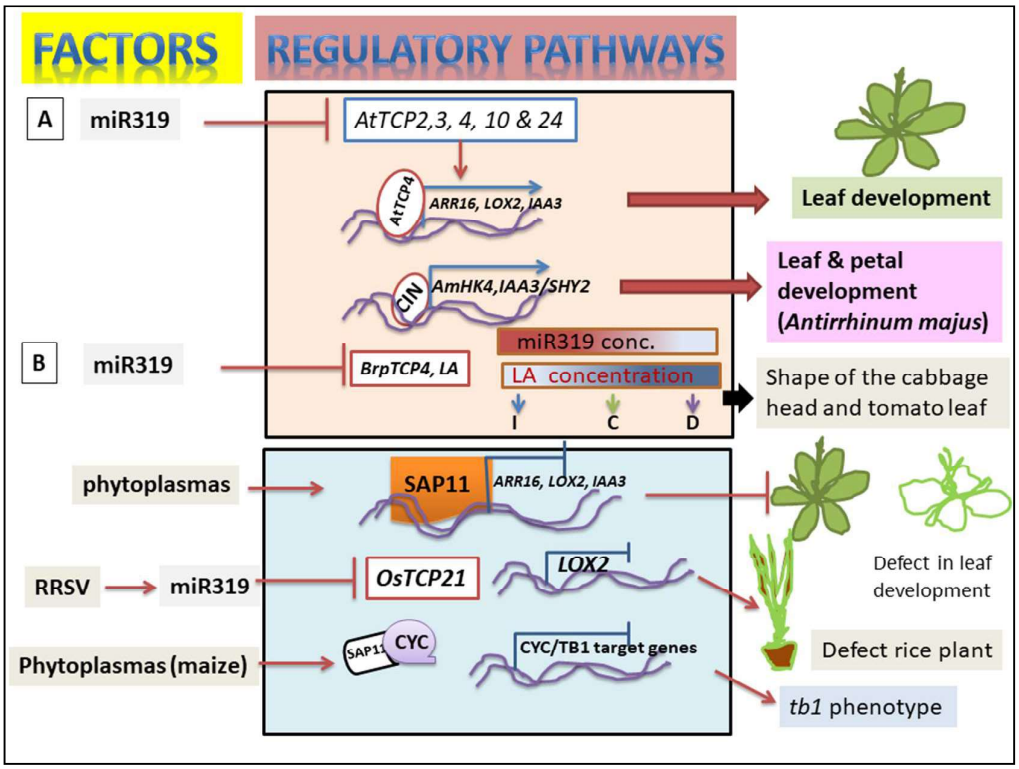


Figure 8: Overview of the regulation of leaf development by CIN-LIKE TCP.

A) Spatio-temporal regulation of TCP by miRNA319 is necessary for proper leaf development. Factors that influence TCP genes are given on the left column and the regulatory pathways involved are in the middle and the outcome/phenotype of the regulatory pathways are on the left column. The level of *BrrpTCP4* at particular place and time, regulated by miRNA319 define leaf shape in Chinese cabbage whereas high level of TCP4 at the margin is necessary for the round phenotype. Similarly, the level of LA determines initiation or differentiation in tomato.

B) Pathogen induced factors interfere with the host regulatory pathway and alter the developmental process causing disease; the pathogen SAP11 interacts with the TCP and destabilizes the protein, resulting in decreased expression level of CIN-LIKE TCP target genes. In maize however, SAP11 destabilized ZmCYC/TB1 rather than CIN-LIKE TCP causing *tb1* mutant like phenotype.

Blue arrow indicates gene induction while cross-bar indicates inhibition and protein associated with 26S indicates protein degradation through ubiquitin 26S proteasomal pathway. The genes in the box linking miR319 with cross bar arrow represent regulation at post transcriptional level. ARR16, ARABIDOPSIS RESPONSE REGULATOR 16; C (competence), D (differentiation), I (initiation); IAA3, INDOLE-3-ACETIC ACID INDUCIBLE 3; LA, LANCEOLATE; LOX2, LIPOXYGENASE 2; R:FR, red light: far-red light ratio; SAP11, SECRETED AY-WB PROTEIN 11. (Adopted from *Ori et al., 2007; Efroni et al., 2013; Das Gupta et al., 2014; Mao et al., 2014; Chang et al., 2018; Pecher et al., 2019; Lan and Qin, 2020; Li et al., 2020*)

In rice alone, there are nine genes belonging to CIN-TCP clade, *OsTCP10, 20, 21, PCF8 & 6,* and OsTCP5 & 18 and PCF5 & 7. Similarly, there are eight members in *Arabidopsis*, *TCP2, TCP3, TCP4, TCP10,* and *TCP24* (having miR319 regulatory site) and *TCP5, TCP13,* and *TCP17,* named as TCP5-like CIN-TCP (Lan and Qin, 2020). There exists an interesting functional similarity between OsTCP5 of rice and *Arabidopsis* AtTCP5/17. SL and CK modulate OsTCP5 which in turn controls mesocotyl elongation in dark (Hu et al., 2014) while AtTCP5/17 like subgroup controls thermomophogenesis in *Arabidopsis* where AtTCP 4 causes de-etiolation under light in contrast to PIF3 that causes etiolated growth under the dark condition.

Also, AtTCP5/17control hypocotyl elongation (Han et al., 2019; Dong et al., 2019). Besides, *OsTCP5* was downregulated by cold and salt stress (Sharma et al., 2010) whereas *AtTCP5/17* were induced by heat and stability of TCP5 increases under high temperature (Figure 9).

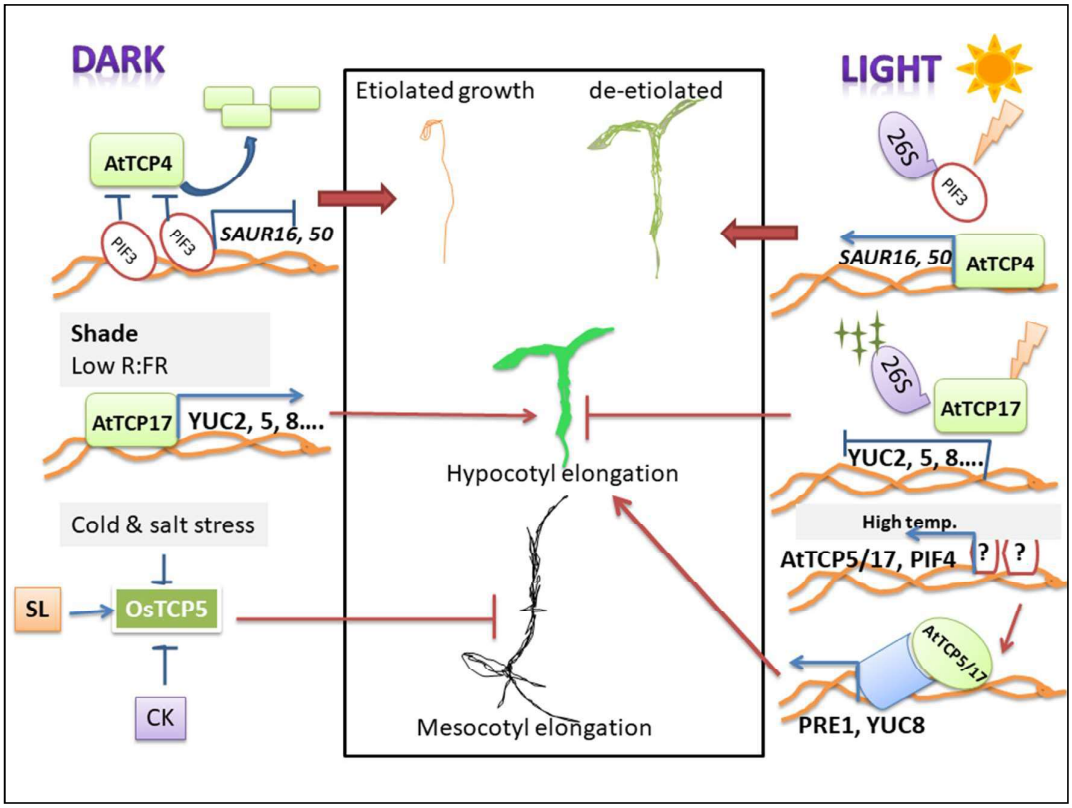


Figure 9: Overview of the regulatory mechanism of CIN-LIKE TCP in different environmental conditions.

The environmental factors are given on either side and the outcome of the regulation in the middle. The protein associated with 26S represents protein degradation through 26S proteasome degradation pathway, blue arrow represents induction and cross bar indicated inhibition. CK, cytokinins; PIFs, PHYTOCHROME-INTERACTION FACTORs; R: FR, red light: far-red light ratio; SAURs, SMALL AUXIN UPREGULATED RNAs; SL, strigolactone; YUCs, YUCCAs and question mark indicates unknown factors. (Adopted from *Sharma et al., 2010; Hu et al., 2014;Han et al., 2019; Dong et al., 2019; Lan and Qin, 2020*)

In addition, CIN-TCP shared similar regulatory functions in the reproductive development via interaction with the repressor protein SPOROCYTELESS and corepressor TOPLESS in both rice (monocot) and *Arabidopsis* (dicot) exemplified by the downregulation phenotype of *OsTCP5* (sterile), *AtTCP5, 13 & 17* (prolong mitotic division in *Arabidopsis* leaves) and overexpression of *Osspl*, miR319, and *AtTCP3 & 5* (Wei et al., 2015; Liu et al., 2017; Ren et al., 2018; Lan and Qin, 2020). In short, CIN-LIKE TCP determines timing, shape, and size, specific to each plant as exemplified by their mutant phenotypes where an alteration in CIN-LIKE TCP results in developmental defect associated with failure of timing between cell proliferation and expansion. In addition, CIN-like TCP provides tolerance to both biotic and abiotic stress. Pathogen derived protein often targets CIN-TCP and destabilizes the protein to enhance its proliferation as observed in rice and *Arabidopsis* (Chang et al., 2018), however, in maize, SAP11 homolog destabilized CYC/TB1 rather than CIN-TCP (Pecher et al., 2019).

For more detail, please see the review on the role of CIN-TCP genes by Sarvepalli and Nath (2018) and Lan and Qin (2020).

CYC/TB1

As mentioned above, this clade is the youngest and believed to be evolved specifically in angiosperm lineage. It may regulate the reproductive development of the newly diversified group. The analysis of data shows CYC/TB1 clade and MADS-box genes (*AP1, AP3, AG, and SEP*) sharing similar duplication events supported this line of thought (Howarth and Donoghue, 2006). MADS-box genes are organ identity genes of flowering plants commonly known as ABCDE genes. TCPs interaction with MADS-box containing protein, e.g., OsMADS57 with OsTB1/FC at the protein level, and overrepresentation of MAD-box binding site in TCPs upstream regulatory element (Sharma et al., 2010), suggest plausible co-evolution hypothesis to regulate each other's activity, adding more complexity in the regulatory pathways as plants struggle to become more and more independent of water in term of reproduction.

The first gene of this clade, *CYCLOIDEA* (*CYC*) was isolated by Luo et al. (1996) from snapdragon and is shown to regulate floral zygomorphy together with its close homolog *DICHOTOMA*. Since its discovery, many more genes have been identified and characterized; however, all seem to be directed toward the regulation of branching and floral zygomorphy (See review Wang et al., 2019).

TB1, first among the genes to be characterized as a branching gene and the mutation that alters its relative expression level, was responsible for maize domestication (Doebley et al., 1997). Similarly, mutation in *tig1/Ostcp15* is responsible for the domestication of *indica* cultivars of rice as shown in Figure 7 (Yu et al., 2007; Zhang et al., 2019). Later, a similar gene *OsTB1/FC1,* a homolog of maize *TB1,* was also found to regulate lateral branching and tillering in rice (Takeda et al., 2003; Minakuchi et al., 2010; Guo et al., 2013) and also in *Arabidopsis* (Finlayson, 2007). Besides, SNP in *TEN* (CYC/TB1, TCP transcription factor) produces branches instead of the tendril in cucumber indicating CYC-TCP gene *TEN* specifies organ identity (tendril identity) in cucumber (Wang et al., 2015).

In maize, *BRANCHED ANGLE DEFECTIVE 1 (BAD1)* regulates lateral branch angle by promoting cell proliferation in pulvinus (a grass-specific inflorescence structure between the main stem and lateral branches) and influences inflorescence architecture (Bai et al., 2012). On the other hand, *REP1/OsTCP24* influences palea development in rice (Yuan et al., 2009; Yin et al., 2018). Also, similar regulation exists in barley (*BD1/COM1)* in which, a mutation in *bdi1* causes indeterminate spikelets and extended branches. However, *rep1* mutant of rice did not show extended branch although the sizes of palea are affected (Levin and Boden, 2020; Poursarebani et al., 2020; Shang et al., 2020).

1.8 CONCLUSION AND FUTURE PROSPECT

Rice is an important cereal crop and has been used as a model plant to better understand the cereal family which has huge agricultural importance in the country as well as in other rice-consuming Asian countries and Africa. Since the full potential of a plant's growth is determined by its genetic component although environmental factors play a critical role in determining the extent of its growth and development, understanding the diverse endogenous regulatory factors and pathways involved in plant growth and development is highly valuable.

TEOSINTE BRANCHED1, **C**YCLOIDEA and **P**ROLIFERATION OF CELL ANTIGEN (PCF1/2), TCP, is an endogenous transcription factor class (specific to plant kingdom) that plays a critical role in regulating plant dynamic growth in response to external and internal cues, plant architecture, plant evolution in terms of reproductive development (duplicated alongside floral organ identity genes family) and domestication. Also, TCP is not limited just to cell proliferation but also modulates several regulatory pathways such as hormonal signaling, stress response, determining organ identity, meristem maintenance, and growth transition. Their characteristic features of being able to modulate several important traits by manipulating a single endogenous factor make TCP a potential candidate for modulating

plant architecture in a race toward developing sustainable variety circumventing the problem faced in the release of transgenic crop. However, the diverse and intricate regulatory networks of TCPs need to be deciphered and sadly much work remains to be done in this aspect especially in rice (Table 3).

Table 3: Expression and function of characterized TCP genes of rice

S.no.	Gene name	Expression	Function	Reference
1	*PCF1* and *2*	Specific to meristematic tissue	Meristem-specific expression of PCNA gene	(Kosugi and Ohashi, 1997)
2	*REP1/TCP24*	Palea	Palea development (floral zygomorphy)	(Yuan et al., 2009; Yin et al., 2018)
3	*OsTB1/FC1*	Floral initiation, SAM	Tiller numbers and Orchestrates cold response	(Martín-Trillo and Cubas, 2010; Sharma et al., 2010) (Chen et al., 2018)
4	*PCF5, PCF6, PCF7, PCF8* and *OsTCP21*	Leaves, roots, and panicles	Cold tolerance	(Yang et al., 2013; Wang et al., 2014)
5	*OsTCP5*	Coleoptile abundance in seedlings, mature leaf, SAM, early panicle stages, and embryos	Mesocotyl elongation, interact with CK and SL Involved in meiotic fate acquisition via OsSPL-TPL pathway	(Hu et al., 2014) (Ren et al., 2018)
6	*OsTCP19*	Higher expression in shoot and roots of resistant rice	Drought and cold stress, interacts with ABI4 and UTL1, fine-tunes the ABA pathway, and negatively regulates auxin and JA Adaptation to local soil type (soil nitrogen content)	(Mukhopadhyay and Tyagi, 2015) (Liu et al., 2021)
7	*OsTCP21*	leaves	Virus resistance	(Zhang et al., 2016)
8	*PCF2*	Roots and leaves	Salt stress	(Almeida et al., 2017)
9	*OsTCP15*	Tillering stage primarily in adaxial side	Controls tiller's angle	(Yu et al., 2007; Zhang et al., 2019)
10	*PCF7*	Ubiquitous, exceptionally high in tillering and heading stage	Regulates tiller number, panicle number, height, and grain development	(Li et al., 2020)
11	Novel TCP	SNPs in the targeted gene in the susceptible variety	Leaf and neck blast resistance	(Devi et al., 2020)

At the centre of every developmental process, there lie two fundamental processes i.e., cell proliferation and differentiation. Plants being sessile depend primarily on developmental plasticity for their survival. Developmental plasticity is an outcome of well-regulated proliferation and differentiation which are spatially and temporally separated and yet overlapping processes, brought about by tight regulation of dynamic regulatory factors such as TCP, at gene and protein level, exemplified by the CIN-LIKE TCP in the regulation of leaf development (Figure 8). This dynamic regulation determines the timing, shape, and size, specific to each species. Therefore, the mutant phenotype that we observed is often associated with defects in the timing of the transition between proliferation and differentiation (e.g., *Osspl/tcp5* or *tcp4* mutant where manipulation of this factor resulted in reproductive development failure or defect in leaf development).

The heterochronic nature required rapid upregulation and downregulation at specific time and space. However, the exact mechanism of spatio-temporal regulation of cell fate transition is not that clear for many rice TCPs. Analysis of upstream regulatory sequence and post-transcriptional regulatory factors such as small RNAs (whether other miRNAs apart from miR319 interact with TCP?) and at protein level (modification, interaction, and degradation) might give further insight into how TCPs regulate plant development spatio-temporally. Besides, TCPs target several downstream genes either to upregulate or downregulate them. TCPs belonging to class II may interact directly and inhibit activation or they may act indirectly in relay by activating kinase which further transduces the signal to inhibit certain pathways such that the effect is visible after the priming protein is gone. In such a case, we might not have a clear picture of how exactly this TCP exerts its function. For this, understanding the degradation pathway or stability of the protein i.e., how long it remains in an active form, or whether it is degraded rapidly or recycled after its action, will greatly broaden the understanding of its exact molecular action. For instance, OsTCP5 of class II TCP inhibits mesocotyl elongation in dark (here OsTCP5 inhibits cell division or proliferation). On the other hand, inhibition of OsTCP5 action by OsSPL through TPL promotes the acquisition of meiotic cell fate, and mutation of both *Osspl* and *Ostcp* (OE) has the same effect in *Arabidopsis* (indicating its involvement in the same pathway). This means, in early panicle stages, OsTCP5 seems to promote somatic cell division which is in contrast to what was seen in mesocotyl. Therefore, unravelling the target downstream genes of OsTCP5 and its degradation pathway will greatly help understand how exactly TCP facilitates cell fate transition. Besides, this particular TCP has more than one form, an outcome of alternative splicing therefore one splice variant may act differently from that of the other.

CHAPTER 2

Material and methods

2.1 IN SILICO WORK

Phylogenetic analysis

For analyzing the phylogenetic relationship, TCP proteins of five monocot crops (*Brachypodium distachyon, Oryza sativa, Sorghum bicolor, Triticum aestivum,and Zea mays*) were downloaded from Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/), Ensemble plant (http://plants.ensembl.org/) and Plant transcription database (http://planttfdb.gao-lab.org/). And, three legumes (*Cicer arietinum, Glycine max, Phaseolus vulgaris)* along with outgroup (*Arabidopsis thaliana*) TCP protein sequences were downloaded from the Plant transcription database (http://planttfdb.gao-lab.org/).

ClustalW software (https://www.genome.jp/tools-bin/clustalw) was used for multiple alignments and a neighbor-joining (N-J) phylogenetic tree was constructed using MEGA X software (Kumar et al., 2018) with 1000 bootstrap replicates.

Expression analysis

Expression profiles of all rice *TCPs* (Sharma et al., 2010) were examined for the selection of a gene for functional analysis. Selected *OsTCP17* gene expression profile was checked in the RGAP database (http://rice.plantbiology.msu.edu/) and microarray data (Deveshwar et al., 2011; Sharma et al., 2012) were used for comparative expression analysis.

OsTCP17 gene structure

The sequence and structure of *OsTCP17* were obtained from RGAP database (http://rice.plantbiology.msu.edu/index.shtml) and phytozome (https://phytozome.jgi.doe.gov/pz/portal.htm).

OsTCP17 protein structure

The full-length OsTCP17 protein sequence was downloaded from the RGAP database. Prosite (https://prosite.expasy.org/scanprosite/) and SMART (http://smart.embl-heidelberg.de/) and used for protein domain analysis. Transit peptide or nuclear localization signal were predicted using iPSORT Prediction (http://ipsort.hgc.jp/predict.cgi), ChloroP 1.1 Server (http://www.cbs.dtu.dk/services/ChloroP/).

2.2 BIOLOGICAL MATERIALS AND THEIR GROWTH CONDITIONS

Plant materials

Nicotiana benthamiana plants maintained at 28°C with 16 hrs light and 8 hrs dark conditions were used for the transient expression of *OsTCP17*.

15-20 days old rice (PB-1), calli were used for the transformation of rice with overexpression construct and CRISPR knockdown construct for *OsTCP17*.

Bacteria

E. coli stain DH5α was used for the cloning as well as for in vivo amplification of vector (pB4NU and pRGEB32). These were grown at optimum temperature (37°C) and, 200 rpm.

Agrobacterium tumefaciens strain GV3101 was used for the transient expression of *OsTCP17-YFP* and strain EHA105 was used for the stable transformation of rice calli.

The growth condition was maintained in an orbital shaker at optimum temperature (28°C) and 200 rpm.

2.3 PREPARATION OF SUBCELLULAR LOCALIZATION CONSTRUCT

Complete *OsTCP17* CDS was amplified from rice IR64 genomic DNA using Q5® High-Fidelity DNA Polymerase (©NEB, USA) and user design gene-specific primers (Table 4, Figure 10, Table 20).

Table 4: Reaction components for amplification of *OsTCP17* CDS from rice IR64 genomic DNA

COMPONENT	25 µl REACTION	FINAL CONCENTRATION
5X Q5 Reaction Buffer	5 µl	1X
10 mM dNTPs	0.5 µl	200 µM
10 µM Forward Primer	1.25 µl	0.5 µM
10 µM Reverse Primer	1.25 µl	0.5 µM
Template DNA	0.5 µl	< 1,000 ng
Q5 High-Fidelity DNA Polymerase	0.25 µl	0.02 U/µl
5X Q5 High GC Enhancer (optional)	(5 µl)	(1X)
MQ Water	11.25 µl	

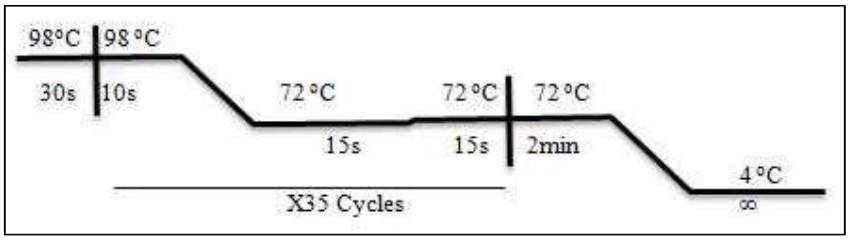


Figure 10: PCR cycle used for the amplification of *OsTCP17* gene.

The amplified product was cloned into the pENTR™/D-TOPO vector and further mobilized into destination vector, pSite-3CA (Figure 24) following manufacturer's protocol.

Restriction digestion

Double restriction digestion of pENTR-TCP17 with PstI (cut at 458 bp; insert) and NotI (cut at 669 bp; vector) was set up (Table 5) and incubated at 37°C for 30 mins and then checked on 1% agarose gel.

Table 5: Reaction components for restriction digestion

S.no	Components	Volume
1	10X FD buffer	2 µl
2	Plasmid	1 µl
3	PstI	0.5 µl
4	NotI	0.5 µl
5	MQ	16 µl

Sequencing

Two plasmids pENTR-TCP17 (11 and 14), selected from among those confirmed by digestion were further purified using kits (©Thermo Fisher Scientific, USA) and sent for sequencing to the CIF sequencing facilities, UDSC. Sequencing was done from both directions using gene-specific forward and reverse primers (Table 20).

Plasmid PCR (check for insert in pSite3CA-TCP17)

The following reaction components (Table 6) were prepared in the PCR tubes and set up the PCR reaction according to the following PCR cycle (Figure 11)

Table 6: Reaction components for plasmid PCR

S.no	Components	Volume
1	10X Taq PCR buffer	2.5 µl
2	dNTP	0.5 µl
3	FP	0.5 µl
4	RP	0.5 µl
5	Template	0.5µl
6	Taq Pol.	0.3
7	MQ	20.2
8	Total	25 µl

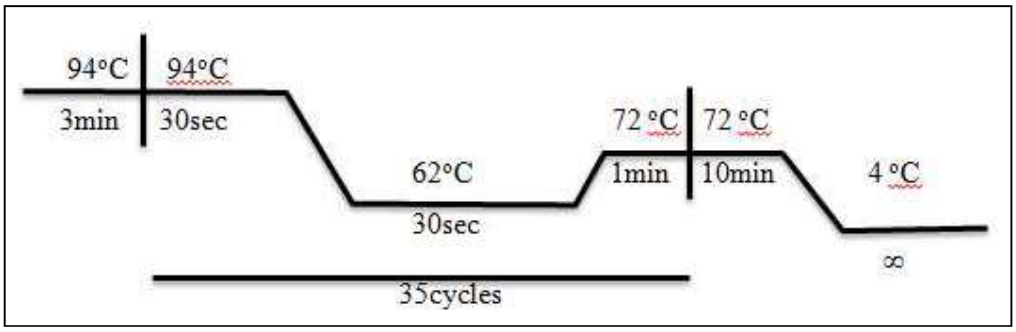


Figure 11: PCR cycle used for plasmid PCR to check the presence of insert. Tm was calculated using NEB Tm calculator tool.

Colony PCR

The transformed colonies were randomly picked from the plates and suspended in 10µl MQ pre-aliquoted in the PCR tubes. The PCR reaction was set up as shown below (Table 7, Figure 12) with DNA lysed from the suspended cells during PCR as a template. The PCR product was then checked by running in 1% agarose gel.

Table 7: Reaction components for colony PCR

S.no	Components	Volume (25µl)
1	10X Taq PCR buffer	2.5 µl
2	dNTP	0.5 µl
3	FP	0.5 µl
4	RP	0.5 µl
5	Template	10µl
6	Taq Pol.	0.3 µl
7	MQ	10.7 µl
8	Total	25 µl

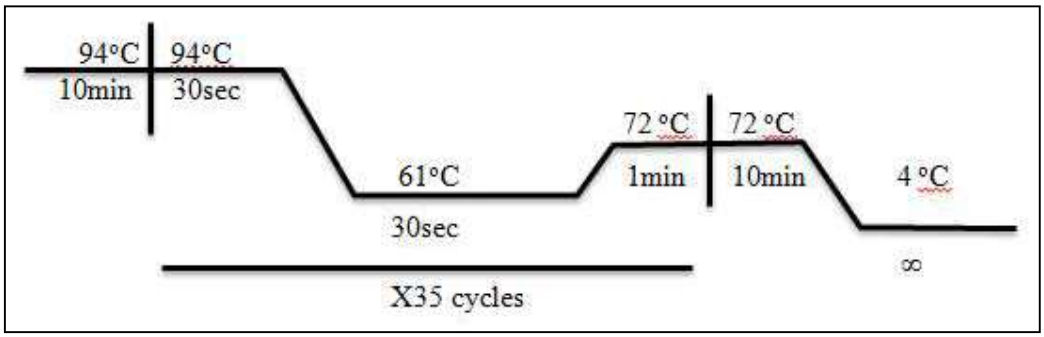


Figure 12: PCR cycles used to screen for the presence of insert (*OsTCP17*)

2.4 *AGRO*-INFILTRATION

The transformed *Agrobacterium* colonies, pSite3CA-TCP17 and P19 (helper strain) were grown in 10ml YEP medium (Table 18) containing appropriate antibiotics, and were mixed to make a final O.D of 0.5 (TCP17-YFP construct) and 0.3 for P19, respectively. The empty vector, pSite-3CA was used for control.

The cells mixture were activated by incubating in activation buffer (Table 8) for 2 hrs in dark following washing and infiltrated into the abaxial side of tobacco leaves with the help of 1ml syringe. The infiltrated plants were then incubated for 4days in dark and the infiltrated leaves were checked using a confocal laser scanning microscope (Leica TCS SP5 AOBS) for the subcellular localization of OsTCP17-YFP.

Table 8: Activation buffer components

S.no	Components	For 5ml	For 30ml
1	0.25M MES (pH 5.6)	200µl	1200µl
2	1M $MgCl_2$	50 µl	300 µl
3	100mM Acetosyringone	7.5 µl	45 µl
4	MQ	4742.5 µl	28.455ml

2.5 OVEREXPRESSION CONSTRUCT

Amplification of *OsTCP17* and pB4NU (vector)

Full-length *OsTCP17* was amplified from pENTR-TCP17 (Table 9 and Figure 13) using designed primers (Table 20) so that the amplified PCR product contains the desire restriction site (BamHI and KpnI). And vector, pB4NU was isolated from the transformed overnight DH5α culture using plasmid miniprep isolation kits (©Thermo Fisher Scientific, USA).

Table 9: PCR reaction components

COMPONENT	25 µl REACTION	FINAL CONCENTRATION
5X Q5 Reaction Buffer	5 µl	1X
10 mM dNTPs	0.5 µl	200 µM
10 µM Forward Primer	1.25 µl	0.5 µM
10 µM Reverse Primer	1.25 µl	0.5 µM
Template (pENTR-TCP17)	0.5 µl	< 1,000 ng
Q5 High-Fidelity DNA Polymerase	0.25 µl	0.02 U/µl
5X Q5 High GC Enhancer	5 µl	1X
MQ Water	10.25 µl	

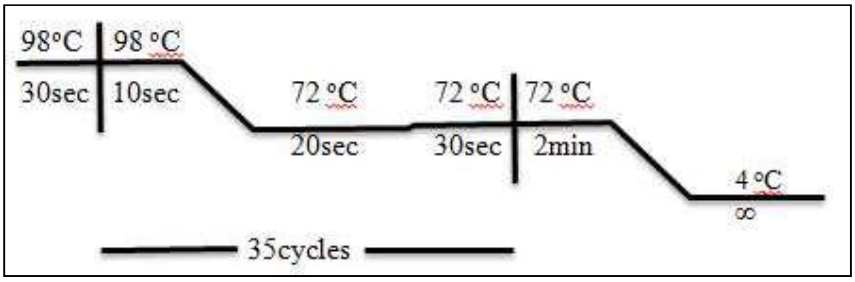


Figure 13: PCR cycles used for amplification of *OsTCP17* using specifically designed primers for cloning in pB4NU vector.

The purified PCR product (*OsTCP17*) and the vector pB4NU were then digested with BamHI and KpnI, respectively, to make a compatible end for ligation (Table 10) at 37°C, 10-30 min.

Table 10: Reaction components for restriction digestion to make compatible ends for ligation

S.NO	COMPONENTS	VOLUMES	
		pB4NU	PCR product
1	10X fast digest buffer	3 µl	3µl
2	BamHI	0.5 µl	0.5 µl
3	KpnI	0.5 µl	0.5 µl
4	DNA	5 µl	20µl
5	MQ	21 µl	6µl
6	Total	30 µl	30µl

Ligation of insert into linearized pB4NU vector using T4 ligase

The digested *OsTCP17* and linearized pB4NU vector were ligated using T4 DNA ligase (NEB) in 5:1 molar ratio (Table 11). Molar ratio was calculated using the formula:

$$\text{Insert (ng)} = \frac{\text{Ratio of insert x vector (ng) x insert size (kb)}}{\text{Ratio of vector x vector size (kb)}}$$

Table 11: Reaction components used for ligation of insert and vector pB4NU. The reaction components were gently mixed and centrifuged followed by incubated at 16°C overnight and transformed into *E. coli* DH5α.

Components	Volume (in µl)
Sterile MQ water	14.5
10X Ligase Buffer	2
Insert	0.5
Vector	1
T4 DNA Ligase	1
Total volume	20

Screening

The plasmids (pB4NU-TCP17) isolated from the transformed colonies were screened for the presence of insert using gene-specific primers (Table 20) followed by restriction digestion using the same enzymes used in the cloning step i.e., BamHI and KpnI (Table 12) for confirmation. Two selected plasmids, pB4NU-TCP17 (3 and 6) were sent for sequencing for final confirmation of the clone and its orientation.

Table 12: Components of restriction digestion reaction mixture for confirming the presence of insert in the vector. The reaction mixture was incubated at 37°C for 15-30 min.

Components	Volume (in µl)
10X fast digest buffer	2
Sterile MQ	16
BamHI	0.5
KpnI	0.5
Plasmid	1
Total	20

Colony PCR

Four transformed *Agrobacterium* colonies were randomly picked and used to perform colony PCR using vector-specific primers (Table 7 and Figure 12). The positive *Agrobacterium*-colonies were then used for *Agrobacterium*-mediated rice transformation.

2.6 CRISPR-CAS9 KNOCKDOWN CONSTRUCT

Primer designing

Single guide RNA specific to *OsTCP17* was designed using CRISPR GE database. A gRNA with less than 0.1 off-target score was selected from the detailed lists of gRNAs along with restriction site.

Golden gate assembly

The designed gRNA primers were diluted to 1 pM/µl of working concentration and annealed using PCR (Table 13, Table 14 and Figure 14).

Table 13: Reaction component for annealing of primer

S.no	Components	Volumes
1	F' primer	2 µl
2	R' primer	2 µl
3	Annealing buffer	1 µl
4	MQ	5 µl
5	**Total**	10 µl

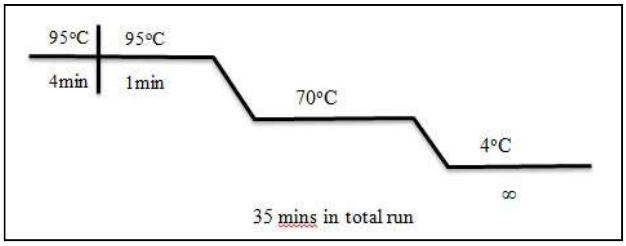


Figure 14: PCR cycle for gRNA annealing

Table 14: Components of annealing buffer

S.no	Components	Volume (1ml)	Working conc.
1	1M Tris (PH 8.0)	10µL	10nM
2	0.5M EDTA (PH 8)	2 µL	1mM
3	4M NaCl	12.5 µL	50mM
4	MQ	975.5 µL	

The annealed gRNA was assembled/inserted into the vector pRGEB32 (Xie et al., 2015) using golden gate assembly reaction components (Table 15 and Figure 15) that contained required enzymes for cutting and ligation followed by transforming chemically competent *E.coli* DH5α.

Table 15: Golden gate reaction components

S.no	Components	Volumes (10 µl)
1	10X GG buffer	1 µl
2	GG enzymes	0.5 µl
3	Vector (pRGEB32)	0.5 µl
4	Annealing template	2 µl
5	MQ	5.7 µl

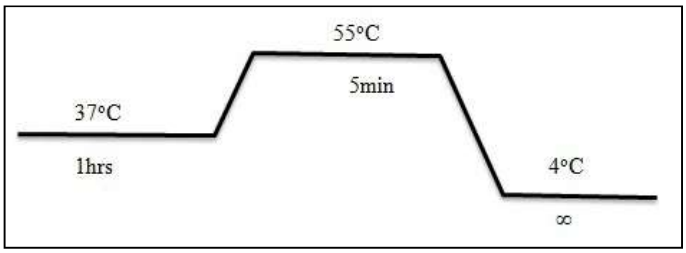


Figure 15: PCR set up for golden gate assembly (for insertion of *OsTCP17* gRNA into pRGEB32).

Sequencing

The transformed *E. coli* (DH5α) were checked for the presence of insert by sequencing using a vector-specific primer (Table 20) that annealed ~300 bp upstream and ~300 bp downstream to the insert site.

Agrobacterium-colony PCR

The selected plasmid confirmed by sequencing was then transformed into *Agrobacterium* (EHA105) by the freeze and thaw method. The transformed *Agrobacterium* colony was checked for the presence of vector pRGEB32 using a vector-specific primers.

This was then used to transform the 15-20 days old rice calli to generate a CRISPR-TCP17 line.

2.7 PREPARATION OF RICE TRANSGENICS

The *Agrobacterium* (EHA101) transformed with pB4NU-TCP17 and pGREB32-TCP17 were used for the transformation of PB1 rice callus. Customized transformation protocol from the lab and NIPGR (Mohanty et al., 1999) was followed for transforming the rice callus.

Step 1: Seed sterilization and inoculation

The de-husked half-cut endosperm of rice PB-1 seeds were taken in 250ml conical flask and sterilized using 70% ethanol solution and 1.75% sodium hypochlorite followed by washing with RO water 6-7 times and kept for drying on sterile Whatman paper. The dried seeds were then inoculated on callus induction medium (Table 16) and kept for 21 days at 27°C in dark condition.

Step 2: Callus subculture

With the help of sterile pair of forceps, the seed and coleoptile part were removed and calli were sub-cultured on fresh callus induction medium for another 5 days in dark at 27°C.

Step 3: Co-cultivation

Preparation of *Agrobacterium* culture

The Secondary culture of transformed *Agrobacterium* colonies from the overnight grown primary culture were allowed to grow till O.D reaches 0.3-0.4.

The culture was then transferred into 50 ml falcon tubes and centrifuged (pre-chilled) for 15 min, 6000 rpm at 4°C and suspended in 100 ml AAM medium (Table 17) containing 25mg/l acetosyringone.

Co-cultivation

The AAM medium containing the *Agrobacterium* cells was mixed with the calli and incubated for 10 min with shaking and calli were dried on sterile Whatman paper for 15-30 mins. The calli were then transferred to the co-cultivation plates (Table 16) and kept in dark for 2 days at 27°C.

Step 4: Selection of transformed calli

The co-cultivated calli were transferred to sterile conical flask and washed with sterile RO water 6-7 X at 100 rpm for 10 min each followed by another 6-7 times washing with RO water containing 250µg/ml amoxicillin to remove the *Agrobacterium* from the calli and kept it for drying. The calli were then transferred onto selection plates and kept for 15 days in dark.

White calli were sub-cultured on fresh selection medium (Table 16) and kept it for 7 days.

Step 5: Regeneration

Healthy white calli were transferred to regeneration medium, RM1 (Table 19), and kept in dark for 10 days at 27°C. Then, the calli were transferred on RM2 and kept in photoperiod 16 hrs light and 8 hrs dark condition for 7 days.

2.8 MEDIA COMPONENTS

Table 16: Media compositions for 1l

S.no	Components	Amount		
		Callus induction medium	Co-cultivation medium	Selection medium
1	MS with vitamin	4.4g	4.4g	4.4g
2	Sucrose	30g	30g	30g
3	Casein hydrolyses	300mg	300mg	300mg
4	L-Proline	600mg	600mg	600mg
5	2,4-D (1mg/ml)	2.5ml	2.5ml	2.5ml
6	BAP (1mg/ml)	0.2ml	0.2ml	0.2ml
7	Phytagel	3g	3g	3g
8	Acetosyringone	-	25mg	25mg
9	Hygromycin	-	-	1ml
10	Amoxicillin	-	-	0.3g
11	pH	5.8	5.2	5.8

Table 17: AAM medium (100 ml)

S.no	Components	Amount
1	MS with vitamin	0.44g
2	Sucrose	1.5g

Table 18: YEP medium (100 ml)

S.no	Components	Amount
1	Yeast Extract Powder	1g
2	Peptone	1g
3	Sodium Chloride	0.5g

Table 19: Regeneration medium (for 1l)

S.no	Components	Amounts			
		RM1	RM2	RM3	RM4
1	MS with vitamin	4.4g	4.4g	4.4g	4.4g
2	Sucrose	30g	30g	30g	30g
3	Casein hydrolyses	300mg	300mg	300mg	300mg
4	BAP (1mg/ml)	2.5ml	2.5ml	2.5ml	2.5ml
5	NAA (1mg/ml)	0.5ml	0.5ml	0.5ml	0.5ml
6	Agarose	10g	8g	8g	-
7	Hygromycin	1ml	600µl	600µl	600µl
8	Amoxicillin	0.3g	0.3g	0.3g	0.3g
9	Phytagel	-	-	-	3g
10	pH	5.8	5.8	5.8	5.8

2.9 PRIMERS

Table 20: Lists of primers used in the experiments (5'-3'). Sequences marked with red represent four additional bases added to the primers to make a compatible end with the vector and those marked with blue and yellow represent the recognition sites of BamHI and KpnI, respectively.

S.no	Primer name	Sequence	Vector/Plasmid	Experiment
1	TCP17 FP	CACCATGGCCTCCCGCGACGCCGC	pENTR™/D TOPO®	Topo cloning, PCR - colony/plasmid
2	TCP17 RP	CCGGGACTCCTCCTGTGCTG		
3	OsTCP17 FP	AAAGGATCCATGGCCTCCCGCGACG	pB4NU	Overexpression, colony & plasmid PCR
4	OsTCP17 RP	ACGGTACCCGGGACTCCTCC		
5	FP	GGCAGCAAGGTGAACGGGCGCGGC	pRGEB32	CRISPR gRNA
6	RP	AAACGCCGCGCCCGTTCACCTTGC		
7	T-F	TGATCTTGATATACTTGGATGAT	pB4NU-TCP17	Sequencing
8	T-R	ACGGTACCCGGGACTCCTCC		
9	TF	CACCATGGCCTCCCGCGACGCCGC	pSite-3CA-TCP17	Sequencing
10	TR	CCGGGACTCCTCCTGTGCTG		
11	Hy FP	TCTACACAGCCATCGGTCCAG	pB4NU-TCP17, pRGEB32-TCP17	Agro-colony PCR
12	Hy RP	GATGTAGGAGGGCGTGGATAT		
13	U3P-F	AGCCTTTCAGGACATGTATTG	pRGEB32-TCP17	Sequencing
14	U3P-R	CTTGACCCGAATTTGTGGAC		

CHAPTER 3

Results and discussion

3.1 IN SILICO ANALYSIS

Phylogenetic analysis of *OsTCP17*

The analysis identified three distinct clades, namely, PCF clade, CIN clade and TB1/CYC clade in both monocots and dicots (Figure 16 and 17). Analysis based on RGAP database identified *AtTCP21* as the closest gene of *OsTCP17* in *Arabidopsis*. Both *OsTCP17* and *AtTCP21* belong to PCF clade and appear to have one paralogous gene; *OsTCP7* and *AtTCP7*. Moreover, both monocots and dicots exhibit similar genes distributions with respect to *OsTCP17* and its homolog, *AtTCP21* (Figure 16 and 17). This indicates that *OsTCP17* and its related genes are highly conserved in both monocots and dicots.

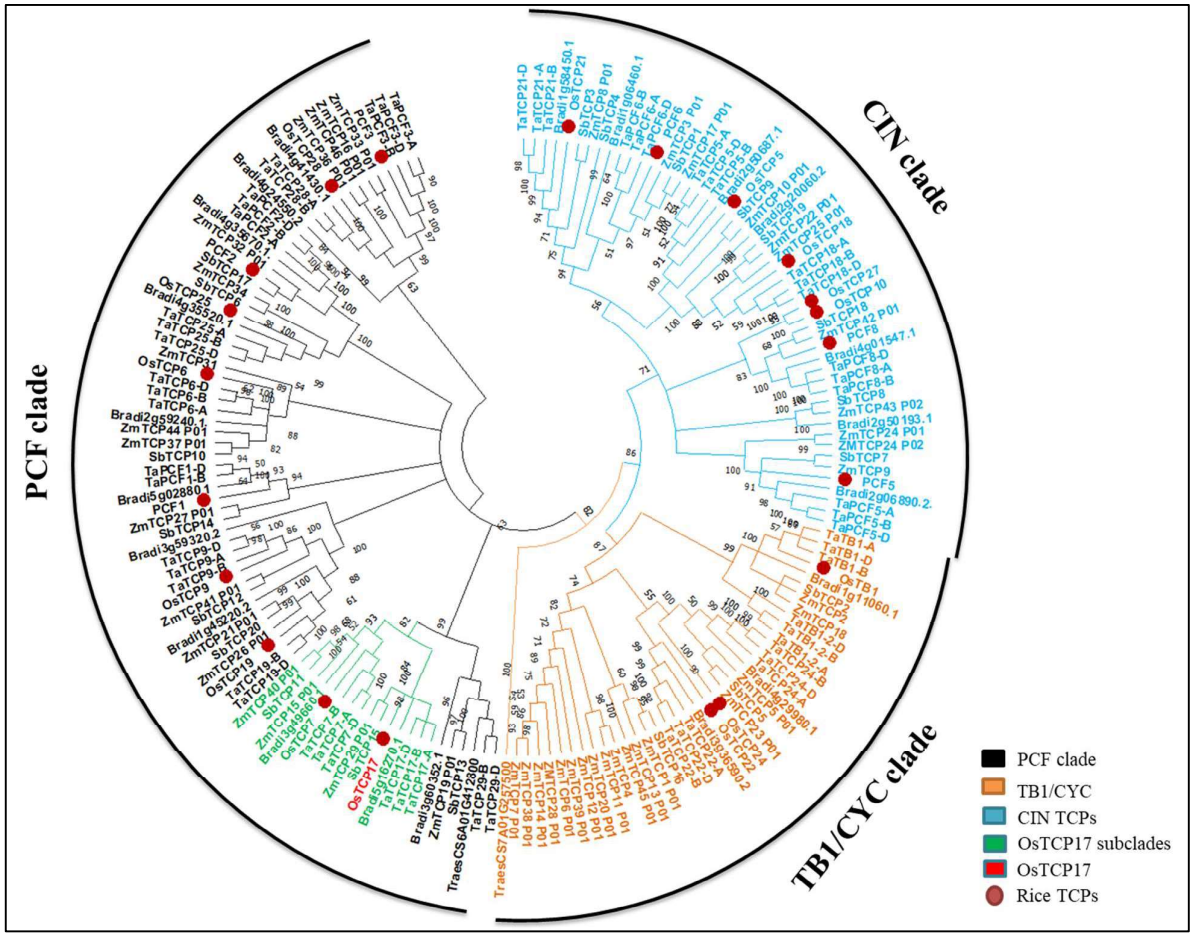


Figure 16: Evolutionary relationships of monocots TCPs (*Brachypodium distachyon, Oryza sativa, Sorghum bicolour, Triticum aestivum, and Zea mays*). The phylogenetic tree was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. This analysis involved 168 amino acid sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 1088 positions in the final dataset. Evolutionary analyses were conducted in MEGA X.

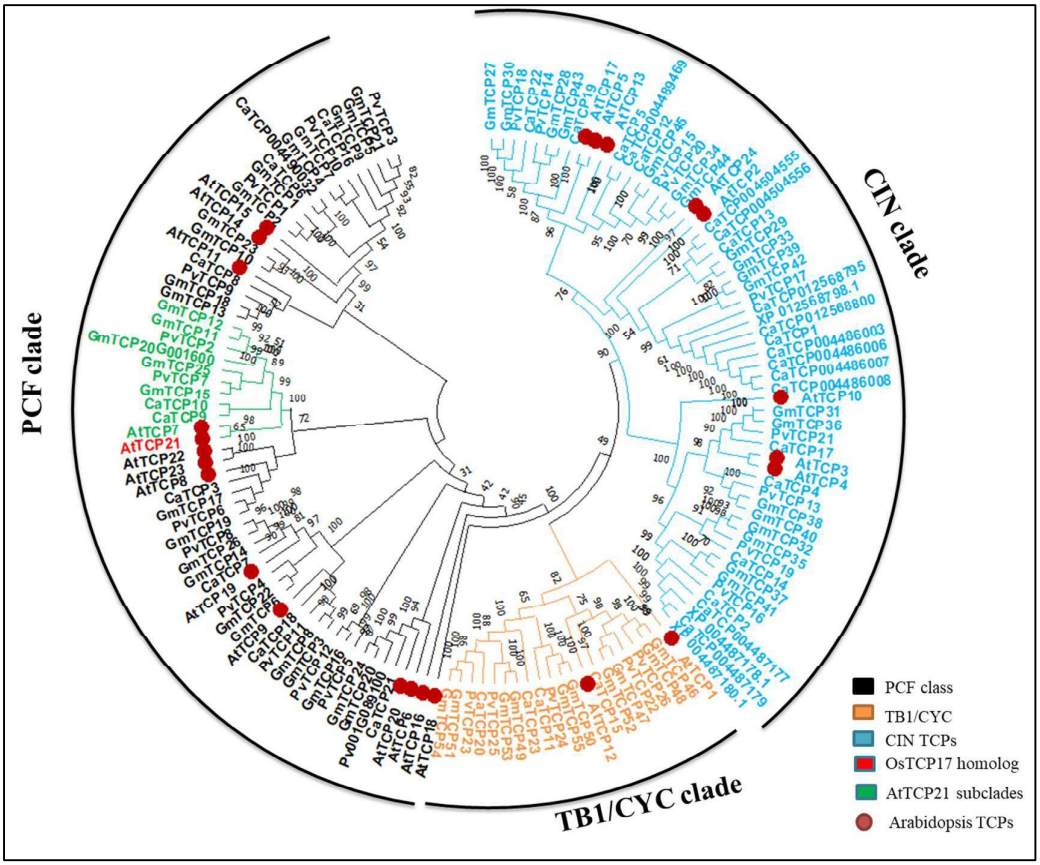


Figure 17: Evolutionary relationships of dicotsTCPs (*Arabidopsis thaliana* and three legumes, *Cicer arietinum, Glycine max, Phaseolus vulgaris*).
The phylogenetic tree was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. This analysis involved 145 amino acid sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 811 positions in the final dataset. Evolutionary analyses were conducted in MEGA X.

Classes and clades are classified base on the classification of Arabidopsis TCP.

Expression analysis of *OsTCP17*

The expression profile retrieved from microarray data shows *OsTCP17* to have a preferential expression in panicle with 500-600 folds up-regulation in stage 5 and significant down-regulation at stage 6 (Figure 18A). This suggests involvement of *OsTCP17* in the regulation of panicle development in rice that is spatio-temporally regulated. Besides, *OsTCP17* shares a similar expressions pattern with *OsTB1* (Figure 18C), a known branching gene in rice. *OsTB1* was reported to be regulated by OsSPL14 and these two were further regulated by OsCCA1. This prompted a look for its possible relationship between these genes in their expression

pattern. As expected, *OsTCP17* and *OsCCA1* show a significant overlapping expression pattern in panicle stage 5-6 (Figure 18B). This result suggests a possible regulatory role of OsCCA1 in regulating the spatiotemporal nature of *OsTCP17*. However, this remains to be tested.

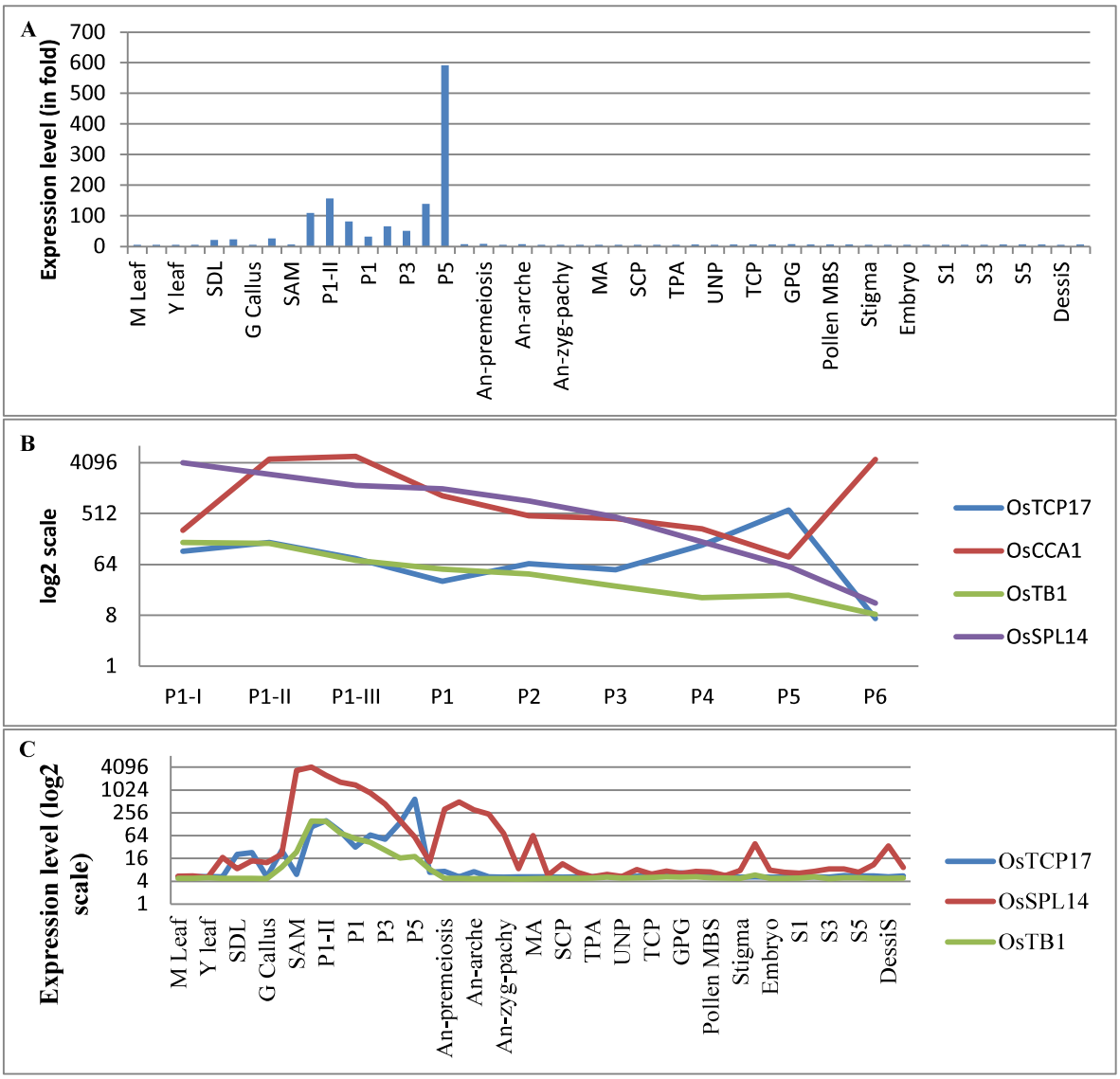


Figure 18: Expression analysis of *OsTCP17*. A) Expression profile of *OsTCP17* showing preferential expression at panicles stages. B) Relative expression pattern of *OsCCA1, OsTCP17, OsTB1* and *OsSPL14* in panicle stages; P1-P6 indicate the stages of panicles (i.e., P1- panicle stage 1) C) Relative expressions profile of *OsTCP17* with *OsTB1* and *OsSPL14*.

OsTCP17 gene structure

OsTCP17 is an intronless gene of about 621 bp long (CDS) with 51 bp long 5' UTR and 399 bp long 3' UTR (Figure 19).

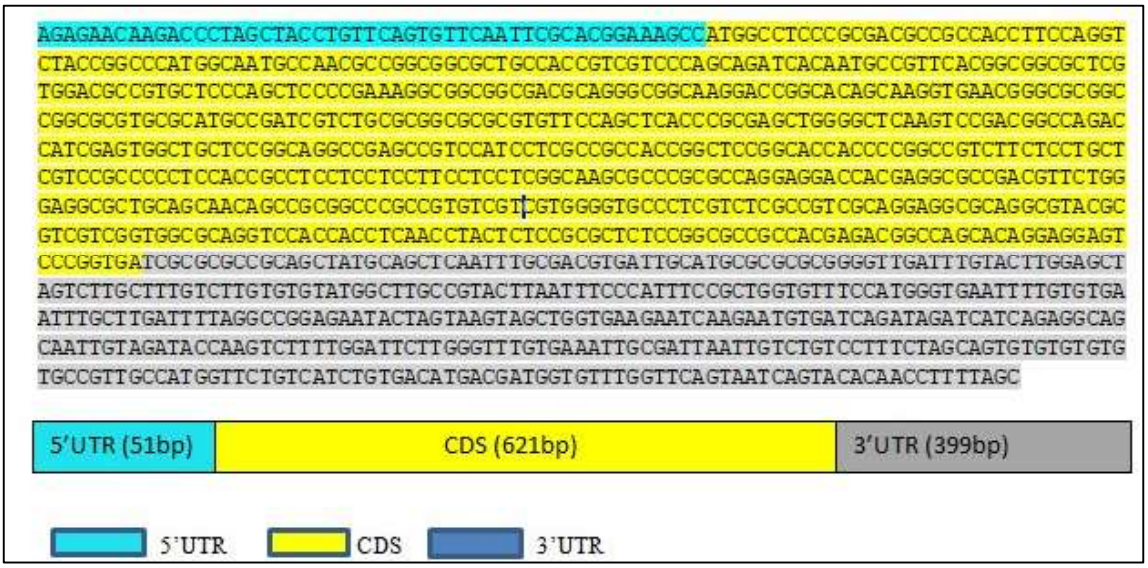


Figure 19: Detailed structure of *OsTCP17* gene.

OsTCP17 protein structure

OsTCP17 encodes 206 amino acids long protein that contains both chloroplasts localizing signal and nuclear localization signal beside conserved TCP domain toward N-terminal region (Figure 20).

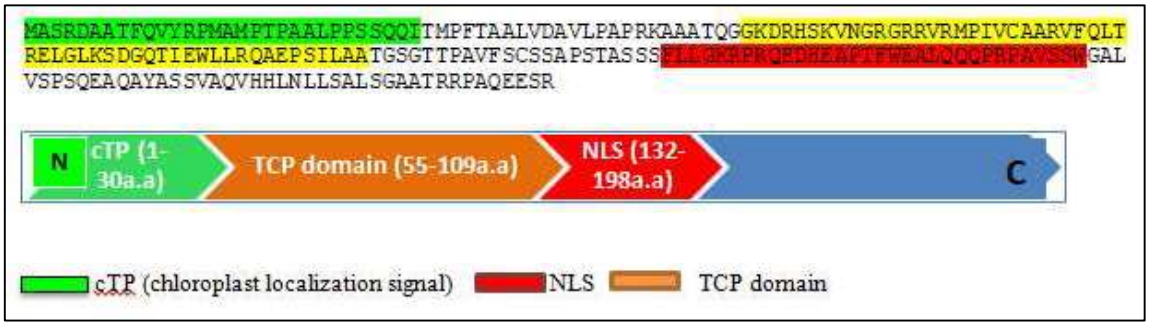


Figure 20: The sequence and the structure of OsTCP17. N: N-terminal C: C- terminal.

Promoter analysis of *OsTCP17*

The analysis of *OsTCP17* promoter shows the presence of binding sites for transcription factors like IPA1/SPL14, PCF1/2, OsCCA1 and other light induce elements (Figure 21) that are reported to regulate *OsTB1* and *AtTCP21*. Also, the binding site of OsBZR1, reported to bind *OsTB1* (involved in SL and BR response; see Figure 2) is found locating 37 bp downstream of TSS. This result corresponds to the expression analysis stated above.

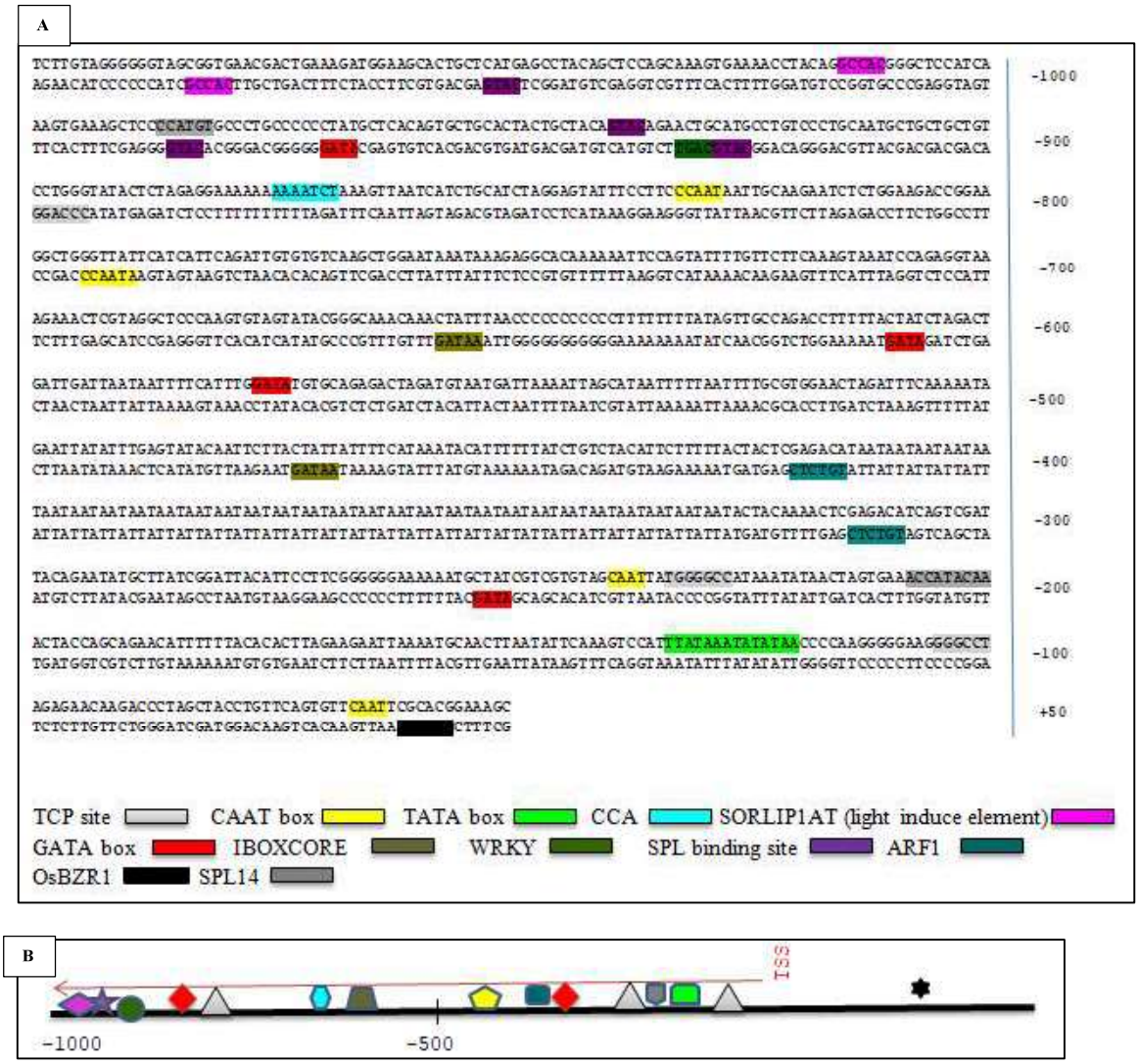


Figure 21: Promoter architecture of *OsTCP17*. A) Promoter sequence (1000bp upstream of TSS and 50bp downstream of TSS) showing the putative binding site of different transcription factor. B) Diagrammatic representation of *OsTC17* cis-regulatory elements.

The selected transcription factors represented here are those reported to regulate *OsTB1* and *AtTCP21*.

3.2 OSTCP17 SUBCELLULAR LOCALIZATION

Cloning of *OsTCP17* in pENTR™/D-TOPO® vector

The full-length *OsTCP17* CDS was amplified from IR64 rice genomic DNA using a user designed gene-specific primer set (Table 20) with added sequences to help the amplified product be compatible for Topo cloning procedure. The amplified PCR product was then checked in 1% agarose gel with 1 kb 6X MassRuler ladder as reference. A specific band was seen between 500 and 600 bp that correlates to desired gene amplicon size (Figure 22), indicating the successful amplification of *OsTCP17*.

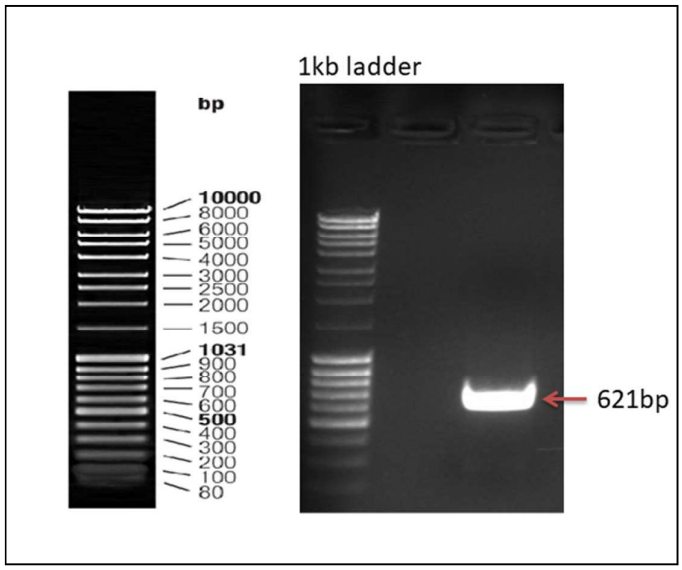


Figure 22: Agarose gel picture of the PCR amplified product, *OsTCP17* CDS from genomic DNA (IR64); Amplicon size ~621bp as indicated by red arrow. DNA size ladder is shown at left.

This was then purified using kits and directly cloned into pENTR™/ D-TOPO vector and transformed into chemically competent, *E. coli* DH5α. Transformed colonies (25) were obtained and plasmids from each colony were isolated to check the orientation of the cloned. The clone was named as pENTR-TCP17.

Restriction digestion to check the clone orientation and sequencing

Digestion of pENTR-TCP17 with Pst1 and Not1 will give two fragments; 2727 bp (vector) and 474 bp (insert) as fallout, provided the insert is in correct orientation. As expected, digested recombinant plasmids from nine colonies show the expected band (Figure 23) indicating successful cloning with correct orientation. This was further confirmed by sequencing.

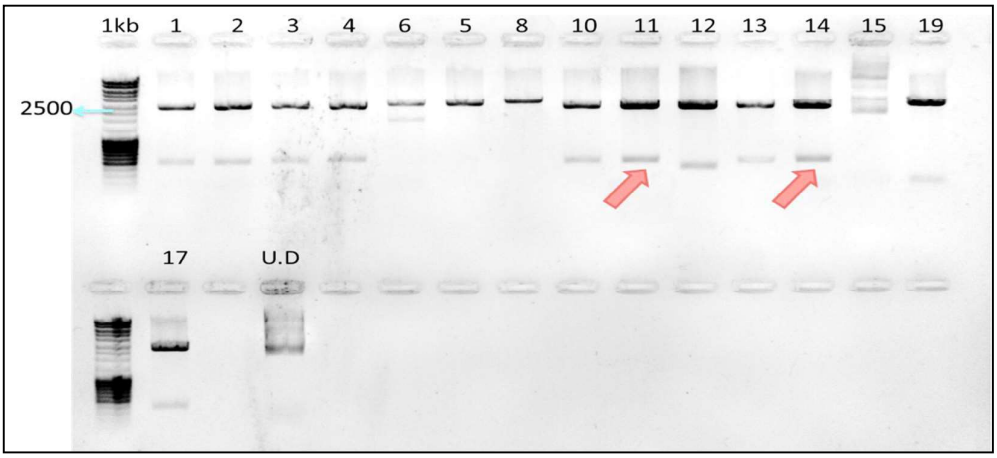


Figure 23: Agarose gel picture showing digested vector (~2727bp) and the fallout of expected size ~474 bp band indicated by red arrow. One kb DNA ladder is shown on left.

11 and 14 indicated by red arrow were purified and sent for sequencing to CIF sequencing facilities, UDSC. The number above each well indicates colony number patched in the plate for reference.

Cloning of *OsTCP17* in destination vector (pSite-3CA)

OsTCP17 fragment from pENTR-TCP17 was mobilized to the destination vectors, pSite-3CA, through LR recombination reaction (Figure 24).

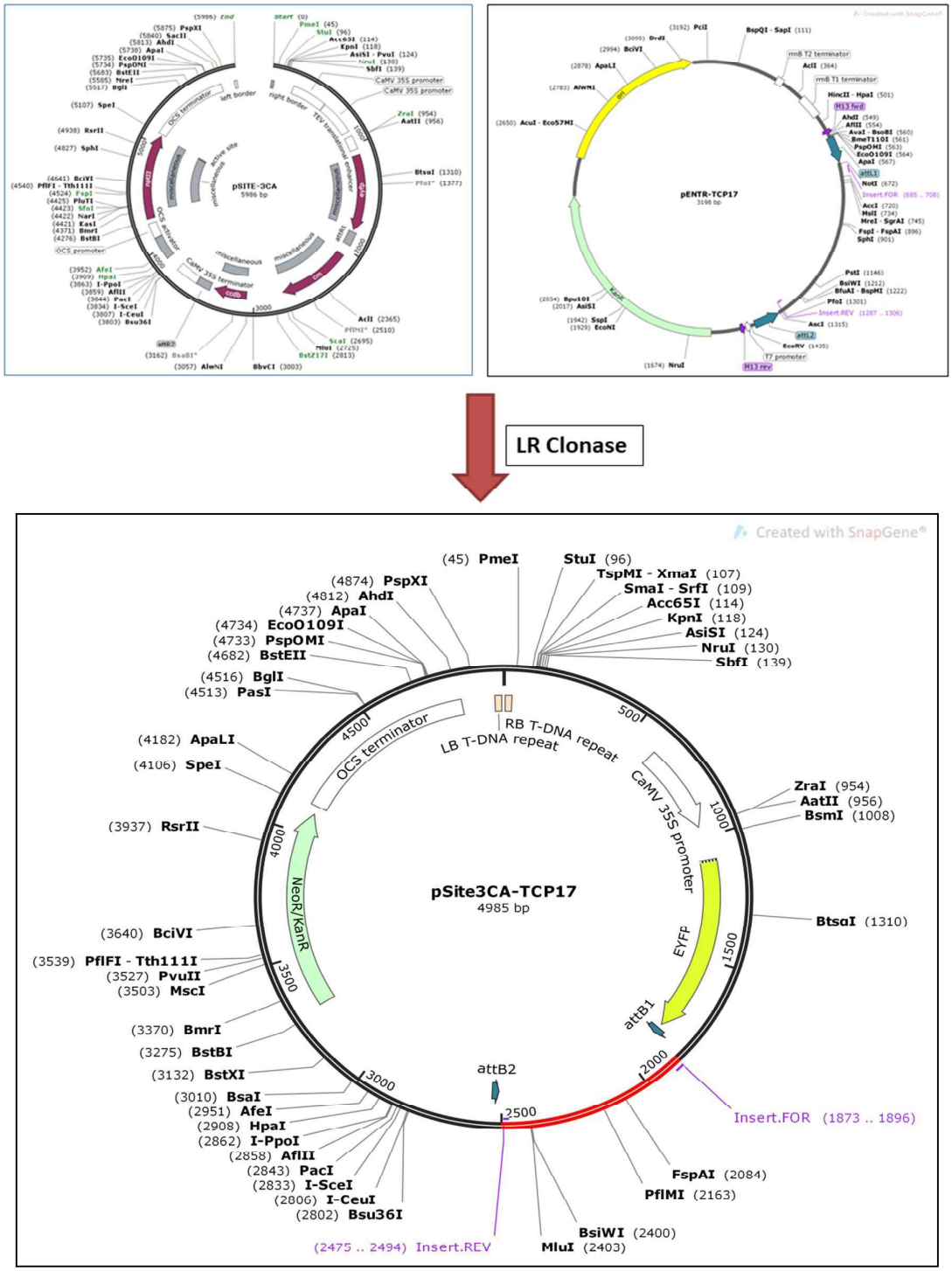


Figure 24: Diagrammatic representation of LR cloning. pENTR-TCP17 as an entry clone and Psite-3CA as a destination vector. The segment in red represents the insert (*OsTCP17*) flanked by two primer sites represented as Insert.FOR and Insert.REV in purple colour.

The plasmids (pSite3CA-TCP17) isolated from four transformed colonies were checked for the presence of insert using PCR amplification with gene-specific primer. An amplicon size of ~621 bp was observed in lane 1 and 2, indicating the presence of an insert (Figure 25).

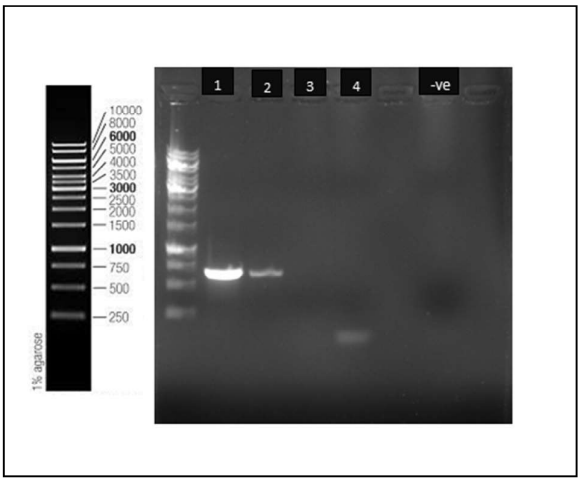


Figure 25: Agarose gel image of plasmid PCR showing gene-specific band between 700 and 500 bp in lane 1 and 2 in reference to 1 kb geneRuler ladder. –ve indicates negative control.

Lack of exact pSite-3CA (vector) sequence information hindered the orientation analysis by restriction digestion based screening. So, the positive clones, confirmed by PCR, were sequenced using a gene-specific primer. The sequencing results when BLAST analysed against genes in MSU RGAP Release 7-CDS sequence database show the best hit on desired gene *OsTCP17* with plus/plus matching. This confirms the successful mobilization of *OsTCP17* in pSite-3CA vector in correct orientation.

Agrobacterium-colony PCR to confirm the presence of pSite3CA-TCP17

Agrobacterium tumefaciens GV3101 was transformed with pSite3CA-TCP17. Four transformed *Agrobacterium* colonies were screened for the presence of pSite-3CA-TCP17 using a set of vector-specific primers (HyF and HyR, Table 20). The amplification of specific band around ~800 bp (Figure 26) confirms the successful transformation of pSite-3CA-TCP17 in GV3101.

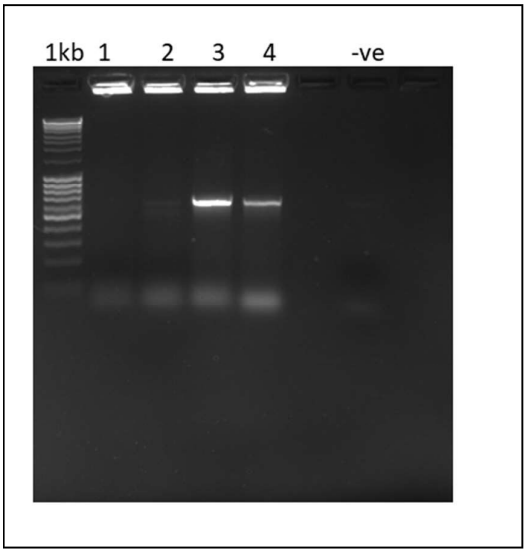


Figure 26: Colony PCR confirming the presence of vector. Lane 1, 1kb DNA ladder; lane 1-4 *Agrobacterium* colonies and lane 6 -negative control (water instead of template).

Subcellular localization

The translational fusion of OsTCP17 with YFP in pSite3CA-TCP17 expressed in vivo in *Nicotiana benthamiana* leaves under constitutive promoter *CaMV 35S*. The outcome indicates OsTCP17 to be localized in both nucleus and cytoplasm (Figure 27).

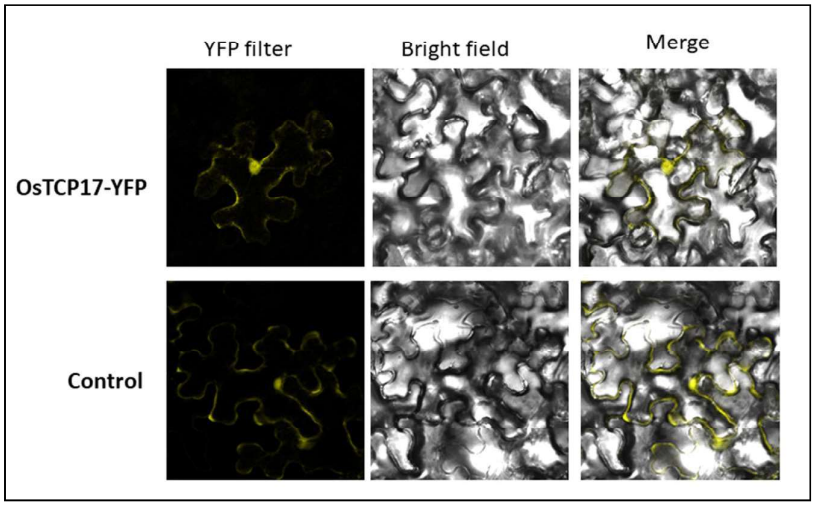


Figure 27: Subcellular localization of OsTCP17 fused with YFP in *Nicotiana benthamiana* leaves. The fluorescence signal of YFP indicates the localization of OsTCP17.

3.3 OVEREXPRESSION CONSTRUCT

OsTCP17 CDS was amplified (Figure 28) from pENTR-TCP17 using specifically designed gene-specific primer (Table 20) so that the amplified product contains desire restriction sites (BamHI and KpnI). The amplified product and the vector pB4NU were digested with BamHI and KpnI and ligated using T4 ligase to produce pB4NU-TCP17.

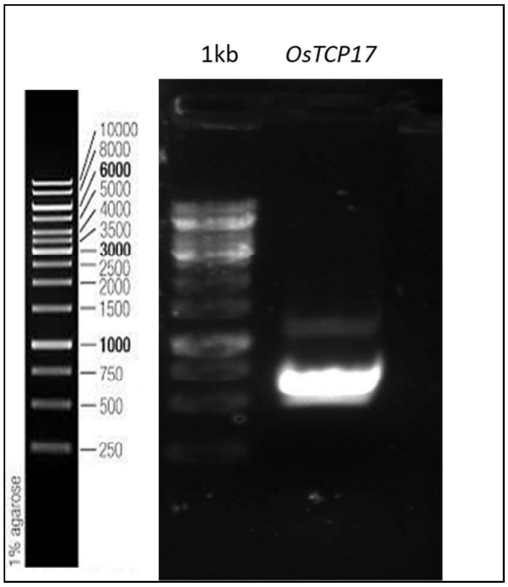


Figure 28: Amplification of full length CDS of *OsTCP17* using gene-specific primer (OsTCP17 FP & OsTCP17 RP). 1% agarose gel electrophoresis image showing high intensity band between 500bp and 750bp; which is an expected amplicon size of *OsTCP17* (621bp).

Plasmid PCR

Six plasmids (pB4NU-TCP17) isolated from the transformed colonies were checked for the presence of insert in correct orientation using a primers Ub1 (FP) and OsTCP17 RP (Table 20). Amplification of specific band at ~815 bp (Figure 29) indicates the successful cloning of *OsTCP17* in pB4NU vector with correct orientation.

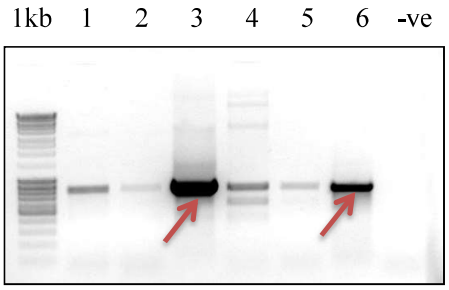


Figure 29: Plasmid PCR using gene-specific primer to check for the presence of insert. If the clone is in the correct orientation, then an amplicon size of ~815 bp will be observed.1% agarose gel shows gene-specific band around 800 bp in reference to 1kb MassRuler ladder. Red arrows indicate plasmids selected for sequencing.

To further validate the plasmid PCR result, the isolated plasmids used above were digested with the same restriction enzymes used for cloning (BamH1 and Kpn1). Since the insert was ligated to the vector pB4NU using BamH1 and Kpn1 overhang, digestion with the same enzymes will remove the insert and show as a fallout at ~621 bp. Out of 6, only two clones show the expected fallout band (Figure 30). These positive clones were further confirmed by sequencing.

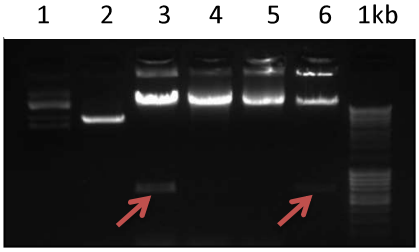


Figure 30: Restriction digestion of pB4NU-TCP17 using BamH1 and Kpn1. 1% agarose gel showing the expected fallout band ~621 bp indicated by red arrow. These two were then purified and sent for sequencing.

Agrobacterium-colony PCR to confirm transformation with pB4NU-TCP17 and rice transformation

The positive clone was then transformed into *Agrobacterium tumefaciens* strain EHA105 to be used for *Agrobacterium*-mediated rice transformation. Four transformed *Agrobacterium* colonies were randomly picked and checked for the presence of vector pB4NU using vector-specific primers; Hy FP and Hy RP (Table 20). Out of four, three colonies show a specific band between 800 and 900 bp which is an expected amplicon size (Figure 31) confirming the presence of vector in *Agrobacterium*.

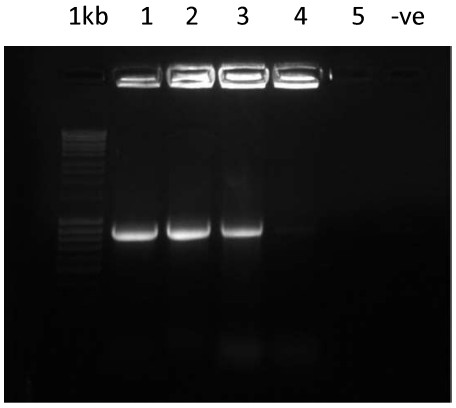


Figure 31: Screening for the transformed *Agrobacterium* colonies (for the presence of vector, pB4NU-TCP17). 1% agarose gel picture showing a vector-specific band between 800 and 900 bp in reference to 1 kb DNA ladder. Lanes 1-4 represent each colony picked; lane 5 was intentionally left blank to separate the negative control from test sample; lane 6, negative control (water instead of template).

All these screening results confirmed the successful cloning of the *OsTCP17* overexpression construct. *Agrobacterium* (EHA105) containing pB4NU-TCP17 was used to initiate co-cultivation with 15-20 days old rice (PB1) calli (Figure 32), which were also cultivated on selection medium after co-cultivation. These experiments need to be followed in future to obtain transgenic rice.

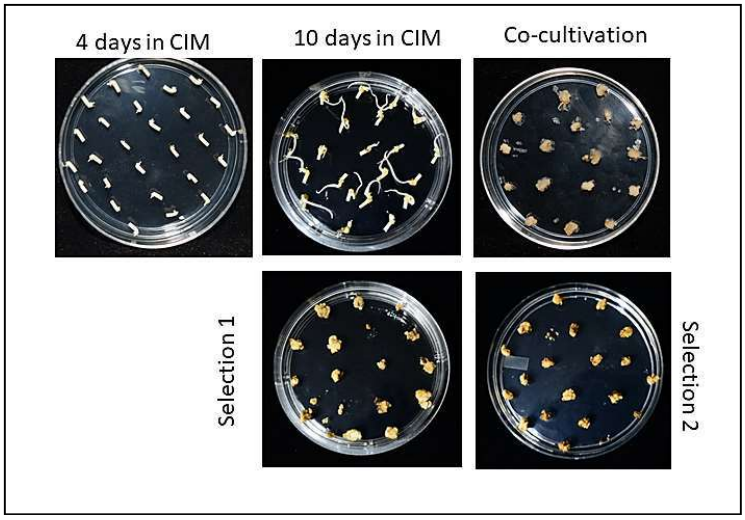


Figure 32: Different stages of rice transformation of 5-20 days old rice (PB1) calli using *Agrobacterium* carrying the overexpression construct pB4NU-TCP17. CIM - callus induction medium.

3.4 KNOCKDOWN CONSTRUCT

Designing of guide RNA

Guide RNA for targeting *OsTCP17* was designed using Target design software and the one with less than 0.1 off-target scores was selected (Figure 33A). Four extra bases were added (GGCA at 5' end and AAAC at 3'end) to the selected guide RNA sequence to make a compatible end with the vector (Figure 33B). The primers were then annealed in PCR and inserted into the vector pRGEB32 using Golden Gate assembly protocol (Table 15).

A

Off-target predictor:

Chrom	Position	Sequence	Off-score	Gene	Region
Chr11	25828082	GCGAGGGGAACGGCCGCCGC CGG	0.073		intergenic
Chr6	16991224	GCAAGGGGAAGGGGCGTGGC CGG	0.069		intergenic
Chr2	27386986	GCAAGGGGAGCTGGCGAGGC CGG	0.069	LOC_Os02g45150	five_prime_UTR
Chr2	25497584	GCAAGGTGAACGGGCGCGGG CGG	0.059	LOC_Os02g42380	CDS
Chr2	12004446	GCTAGGTGCAGCGGCGCGGC CGG	0.053		intergenic
Chr10	3541300	GCAAGGCGAACTGGCGAGGC CAG	0.032		intergenic
Chr12	964886	GCAAGGGGAACGGAGGCGGC TGG	0.022	LOC_Os12g02730	intron
Chr6	12642828	GCGGGGGGAACGGGGGCGGC CGG	0.011		intergenic
Chr2	24935052	GCGAGGAGATCGGGCGCGGC CGC	0.01	LOC_Os02g41580	CDS
Chr11	3760010	CCAAGGTGGAGGGGCGCGGG CGG	0.007	LOC_Os11g07460	CDS
Chr10	19214562	GCATCGTGAACGGGCGCGGC GGC	0.005	LOC_Os10g35950	CDS
Chr3	17598230	GCAAGGTCGACGGCCGCGGC CGC	0.004	LOC_Os03g30880	CDS
Chr2	35610452	GCAAGGTCGACGGCCGCGGC CGC	0.004	LOC_Os02g58180	CDS
Chr1	40477754	CCAAGGTGGACGGGCGGGGG CGG	0.002	LOC_Os01g69980	CDS
Chr2	27186048	GCACGGCGAACGAGCGCGCC CGT	0.002	LOC_Os02g44880	intron
Chr12	4097881	GCGAGGTTAACGGGCGCGGG GGA	0.002		intergenic
Chr8	24906196	GCAAGGTGGACGGCGGCGGC AGC	0	LOC_Os08g39390	CDS
Chr10	19928004	GCAAGGTGAACGAGCGCGCG CGC	0		intergenic

B

FP GGCAGCAAGGTGAACGGGCGCGGC

RP AAACGCCGCGCCCGTTCACCTTGC

Figure 33: Details of the selected guide RNA sequence obtained with the help of target design software. A) List of possible off-target site. B) Final primer with additional four bases mark with red colour.

Screening for cloning

The guide RNA is a very short nucleotide sequence so sequencing is the only option to check the successful insertion into the vector, pRGEB32. The sequencing result was then custom searched for the presence of a full guide RNA sequence with the help of Snapgene software that allowed identification and/or insertion of the primer (gRNA) as shown in Figure 34.

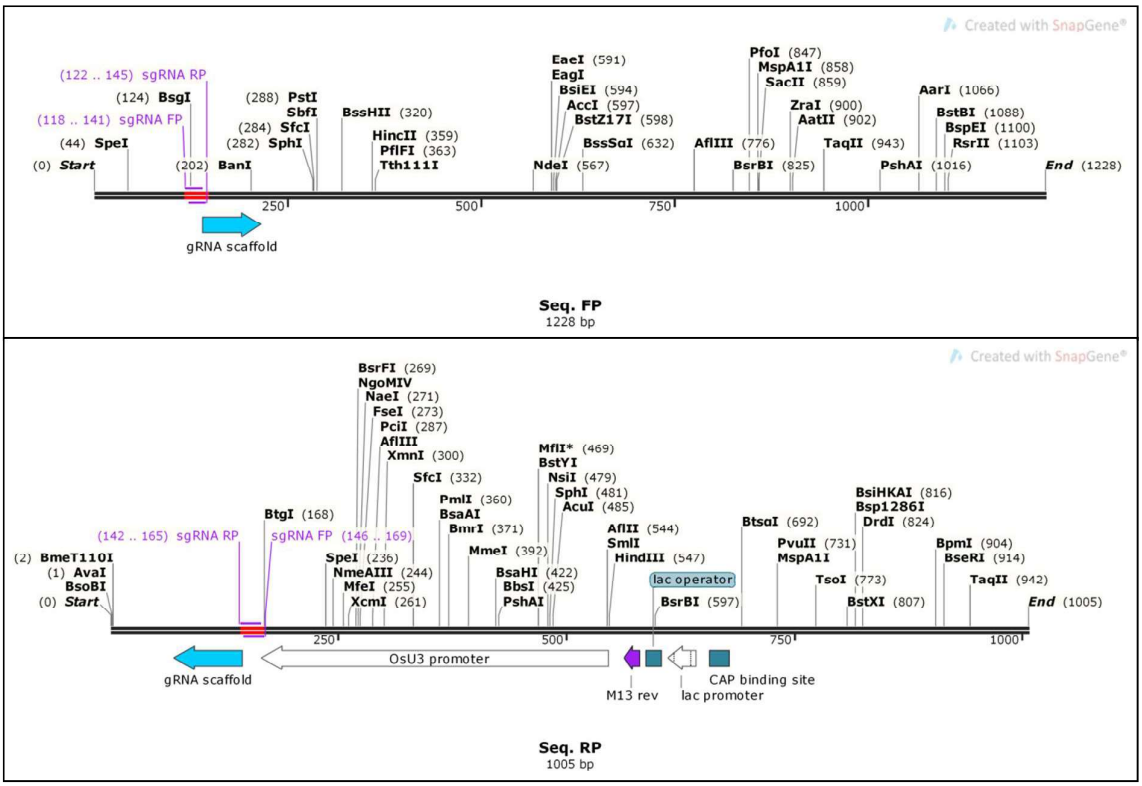


Figure 34: Sequencing result of pRGEB32-TCP17gRNA using U3P-F and U3P-R as sequencing primers. For clarity, the sequenced data was converted into linear restriction map using snapgene tool and checked for the presence of gRNA by annealing the designed primer (Figure 33B) to the converted linear sequence map. The two primers (sgRNA FP and RP) annealing perfectly (marked with purple color) indicating the presence of guide RNA (marked with red color).

This sequencing result confirms the successful insertion of gRNA into the vector pRGEB32. This construct was then transformed into *Agrobacterium* strain EHA105 to be used in the transformation of rice callus.

Agrobacterium-colony PCR to confirm transformation and raising rice transgenics

Four *Agrobacterium* colonies were picked randomly from the selection plates and screened for the presence of vector, pRGEB32-TCP17gRNA using a vector-specific primer (Hy FP and Hy RP, Table 20). Amplification of specific band between ~800-900 bp (Figure 35) confirmed

the successful transformation of *OsTCP17* CRISPR construct in *Agrobacterium* strain EHA105.

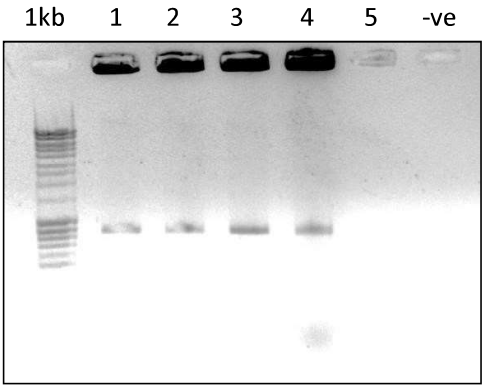


Figure 35: *Agrobacterium*-colony PCR to check the presence of vector, pRGEB32-TCP17gRNA. 1% agarose gel picture shows an expected amplicon size between ~900 bp-800 bp in all four colonies screened. Lanes 1-4, transformed *Agrobacterium* colonies; lane 5, intentionally left blank and –ve, negative control (water instead of template).

Fifteen to twenty days old dark-grown rice calli were co-cultivated with the transformed *Agrobacterium* EHA105 containing knockdown construct. The calli were kept in the selection medium for 7 days (selection 1) in dark and another 7days in the selection medium (selection 2) following which the healthy calli were transferred into regeneration medium 1 (Figure 36). This needs to be followed to get rice transgenics.

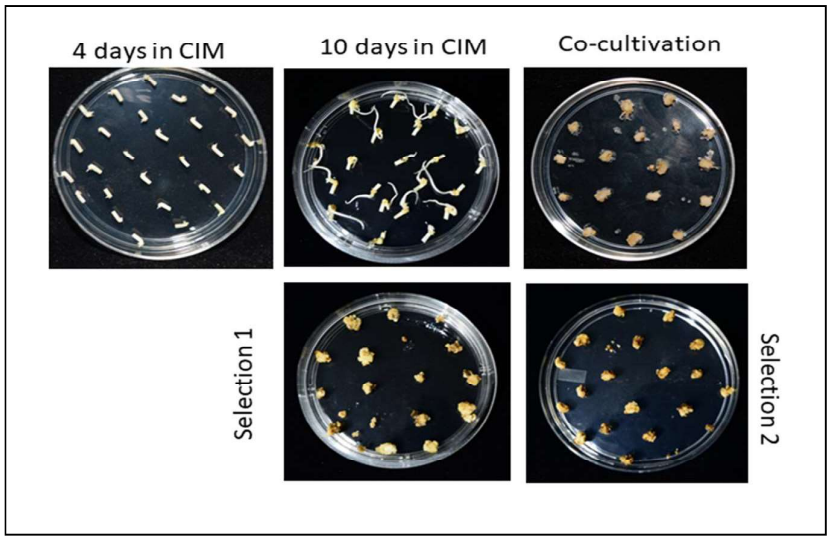


Figure 36: Different stages of rice calli (PB-1) transformed with CRISPR knockdown construct. CIM-callus induction medium.

3.5 DISCUSSION

OsTCP17 is an intron-less gene, encoding a protein of 206 amino acids long containing both transit peptide and nuclear localization signal (NLS) besides the conserve TCP domain located toward the N-terminus. This prediction suggests OsTCP17 to have localization in chloroplasts and nucleus. The localization result (Figure 27) confirmed its localization in the nucleus and cytoplasm but it is not clear where exactly in the cytoplasm. OsTCP17 might probably interact with some peripheral protein as some YFP fluorescence was also observed at the periphery. However, there is no conclusive result yet to confirm the localization in the chloroplast or nature of its interaction partner. For that, a construct with chloroplast and mitochondria markers has to be transformed in *Agrobacterium,* GV3101 for co-transformation.

The phylogenetic analysis of all TCPs from four cereal and three legumes identified three clades; PCF clade, CIN clade and TB1/CYC clade, respectively. *OsTCP17* and its related gene *AtTCP21* in *Arabidopsis* belong to PCF clade. *OsTCP17* has one paralogous gene (*OsTCP7*) and six homologs in wheat (hexaploid), three in maize and two each in *Brachypodium* and sorghum. Interestingly, similar genes distribution was observed in dicots as well, with *Glycine max* (tetraploid) sharing higher gene number compared to the others. The similar gene distributions with one duplicated gene (e.g., OsTCP17- OsTCP7 and AtTCP21-AtTCP7) in both monocots and dicots indicate that they are highly conserved. This maybe because they are very important to the survival or proper development of the plants and that stringent regulation or negative selection happens for its mutant or, the variant genes have now become highly diversified and are no more recognised as those belonging to the same subclades.

In order to understand its regulation, 1000 bp above TSS for both *AtTCP21* and *OsTCP17* were analysed. The analysis shows some similarity in their cis-regulatory sequence having a binding site for CCA1 and other light-responsive elements. AtTCP21 (CHE) is known to be involved in the regulation of circadian clock in *Arabidopsis*; AtTCP21(CHE) acts as a repressor of CCA1 and vice versa (Pruneda-Paz et al., 2009). Since OsCCA1 is homologous to *Arabidopsis* CCA1/LHY; binding with OsCCA1 to *OsTCP17* might have similar effects. This suggests a possible role of OsTCP17 in the clock regulation of rice. In addition, *OsTCP17* cis-regulatory sequence contains SPL binding site as well as PCF1 and PCF2 binding sites. IPA1 (IDEAL PLANT ARCHITECTURE 1) physically binds to TB1 (TGGGCC/T) to suppress rice tillering (Lu et al., 2013; Song et al., 2017). Moreover, OsCCA1 upregulates *IPA1* and *OsTB1* to regulate tillering (Wang et al., 2020). Also, *OsTCP17* has an inverse relationship with *OsCCA1* in panicle stages (Figure 18B). This

suggests a possibility of *OsTCP17* involving in the same regulatory pathways of OsCCA1, OsTB1, OsSPL14/IPA1, D53, and SL to regulate tillering and panicle development in rice.

OsTCP17 belongs to the PCF clade while OsTB1 belongs to class II TCP. Binding of OsCCA1 induces the expression of *OsTB1* and OsSPL14 (this reduces tillering), however, binding of OsCCA1 to *OsTCP17* might have the opposite effect on tillering and panicle development, because, PCF clades are known for its activation activity. If this is true, knockdown of *OsTCP17* will result in a significant decrease in tillers or panicles number, and overexpression of *OsTCP17* will either enhance tillering or panicles.

Also, if OsTCP17 has a similar role as that of its counterpart AtTCP21 (CHE), alteration of *OsTCP17* by knockdown or overexpression will have an immense effect on rice flowering time or defect in phase transitions.

To test this hypothesis, CRISPR knockdown construct as well as overexpression construct have been made and transformed in the rice. The subsequent raising/screening for transgenic lines and phenotypic analysis is yet to be done.

CHAPTER 4

Conclusions and future prospects

4 CONCLUSIONS AND FUTURE PROSPECTS

TCP is a family of plant-specific transcription factors that regulate and orchestrate diverse developmental aspects of the plant. OsTCP17 is a class I TCP (TCP-P) of 206 amino acids, localized in the nucleus. *OsTCP17* is preferentially expressed in panicle, reaching its peak at stage 5 and downregulated in stage 6. Also, *OsTCP17* shares an overlapping expression pattern with *OsTB1, OsSPL14,* and *OsCCA1* suggesting OsTCP17 to be involved in the same regulatory pathways to regulate tillering and panicle development.

Moreover, *OsTCP17* has OsCCA1 binding site in similar manner with its counterpart, *AtTCP21 (CHE)* in *Arabidopsis* where CHE inhibits CCA1 and vice versa. This suggests OsTCP17 might have a similar role in rice, as OsCCA1 or TOC1 (two clock components) can complement the functions of their counterpart mutants in *Arabidopsis* (Murakami et al., 2007). However, this remains to be tested.

Future work will involve screening and phenotyping of the CRISPR knockdown and overexpression lines. In addition, analysis of whether OsTCP17 and OsCCA1 have inverse relationship at protein level or whether OsTCP17 interacts with TOC1 in rice needs to be done. Also, identification of OsTCP17 target genes will add more insight into the functional aspect of OsTCP17.

Bibliography

BIBLIOGRAPHY

Aggarwal, P., Gupta, M.D., Joseph, A.P., Chatterjee, N., Sinivasan, N., and Nath, U. (2010). Identification of specific DNA binding residues in the TCP family of transcription factors in Arabidopsis. Plant Cell **22**: 1174–1189.

Aguilar-Martlnez, J.A. and Sinha, N. (2013). Analysis of the role of Arabidopsis class I TCP genes AtTCP7, AtTCP8, AtTCP22, and AtTCP23 in leaf development. Front. Plant Sci. **4**: 406.

Almeida, D.M., Gregorio, G.B., Oliveira, M.M., and Saibo, N.J.M. (2017). Five novel transcription factors as potential regulators of OsNHX1 gene expression in a salt tolerant rice genotype. Plant Mol. Biol. **93**: 61–77.

Andriankaja, M.E., Danisman, S., Mignolet-Spruyt, L.F., Claeys, H., Kochanke, H., Vermeersch, M., Milde, L.D., Bodt, S.D., Storme, V., Skirycz, A., Maurer, F., Bauer, P., Mühlenbock, P., Breusegem, F.V., Angenent, G.C., Immink, R. G. H., and Inzé, D. (2014). Transcriptional coordination between leaf cell differentiation and chloroplast development established by TCP20 and the subgroup Ib bHLH transcription factors. Plant Mol. Biol. **85**: 233–245.

Arite, T., Umehara, M., Ishikawa, S., Hanada, A., Maekawa, M., Yamaguchi, S., and Kyozuka, J. (2009). D14, a strigolactone-insensitive mutant of rice, shows an accelerated outgrowth of tillers. Plant Cell Physiol. **50**: 1416–1424.

Bai, F., Reinheimer, R., Durantini, D., Kellogg, E.A., and Schmidt, R.J. (2012). TCP transcription factor, BRANCH ANGLE DEFECTIVE 1 (BAD1), is required for normal tassel branch angle formation in maize. Proc. Natl. Acad. Sci. U. S. A. **109**: 12225–12230.

Bellande, K., Bono, J.J., Savelli, B., Jamet, E., and Canut, H. (2017). Plant lectins and lectin receptor-like kinases: How do they sense the outside? Int. J. Mol. Sci. **18**: 1164.

Cao, J.F., Zhao, B., Huang, C.C., Chen, Z.W., Zhao, T., Liu, H.R., Hu, G.J., Shangguan, X.X., Shan, C. M., Wang, L.J., Zhang, T.Z., Wendel, J.F., and Guan, X.Y. (2020). The miR319-targeted GhTCP4 promotes the transition from cell elongation to wall thickening in cotton fiber. Mol. Plant **13**: 1063–1077.

Chang, S.H., Tan, C.M., Wu, C.T., Lin, T.H., Jiang, S.Y., Liu, R.C., Tsai, M.C., Su, L.W., and Yang, J.Y. (2018). Alterations of plant architecture and phase transition by the phytoplasma virulence factor SAP11. J. Exp. Bot. **69**: 5389–5401.

Chen, L., Zhao, Y., Xu, S., Zhang, Z., Xu, Y., Zhang, J., and Chong, K. (2018). OsMADS57 together with OsTB1 coordinates transcription of its target OsWRKY94 and D14 to switch its organogenesis to defense for cold adaptation in rice. New Phytol. **218**: 219–231.

Cubas, P., Lauter, N., Doebley, J., and Coen, E. (1999). The TCP domain: A motif found in proteins regulating plant growth and development. Plant J. **18**: 215–222.

Danisman, S., van der Wal, F., Dhondt, S., Waites, R., de Folter, S., Bimbo, A., van Dijk, A.J., Muino, J.M., Cutri, L., Dornelas, M.C., Angenent, G.C., and Immink, R.G.H. (2012). Arabidopsis class I and class II TCP transcription factors regulate jasmonic acid metabolism and leaf development antagonistically. Plant Physiol. **159**: 1511–1523.

Danisman, S. (2016). TCP transcription factors at the interface between environmental challenges and the plant's growth responses. Front. Plant Sci. **7**: 1930.

Das Gupta, M., Aggarwal, P., and Nath, U. (2014). CINCINNATA in *Antirrhinum majus* directly modulates genes involved in cytokinin and auxin signaling. New Phytol. **204**: 901–912.

Deveshwar, P., Bovill, W.D., Sharma, R., Able, J.A., and Kapoor, S. (2011). Analysis of anther transcriptomes to identify genes contributing to meiosis and male gametophyte development in rice. BMC Plant Biol. **11**:78.

Devi, S.J.S.R., Singh, K., Umakanth, B., Vishalakshi, B., Rao, K.V.S., Suneel, B., Sharma, S.K., Kadambari, G.K.M., Prasad, M.S., Senguttvel, P., Syamaladevi, D.P., and Madhav, M.S. (2020). Identification and characterization of a large effect QTL from *Oryza glumaepatula* revealed Pi68(t) as putative candidate gene for rice blast resistance. Rice **13**: 17.

Dhaka, N., Bhardwaj, V., Sharma, M.K., and Sharma, R. (2017). Evolving tale of TCPs: New paradigms and old lacunae. Front. Plant Sci. **8**: 497.

Doebley, J., Stec, A., and Hubbard, L. (1997). The evolution of apical dominance in maize. Nature **386**: 485–488.

Dong, J., Sun, N., Yang, J., Deng, Z., Lan, J., Qin, G., He, H., Deng, X.W., Irish, V.F., Chen, H., and Wei, N. (2019). The transcription factors TCP4 and PIF3 antagonistically regulate organ-specific light induction of *SAUR* genes to modulate cotyledon opening during de-etiolation in Arabidopsis. Plant Cell **31**: 1155–1170.

Du, J., Hu, S., Yu, Q., Wang, C., Yang, Y., and Sun, H. (2017). Genome-wide identification and characterization of BrrTCP transcription factors in *Brassica rapa ssp. rapa*. Front. Plant Sci. **8**: 1588.

Efroni, I., Han, S.K., Kim, H.J., Wu, M.F., Steiner, E., Birnbaum, K.D., Hong, J.C., Eshed, Y., and Wagner, D. (2013). Regulation of leaf maturation by chromatin-mediated modulation of cytokinin responses. Dev. Cell **24**: 438–445.

Elert, E. (2014). A good grain. Nature **514**: S50–S51.

Fang, Z., Ji, Y., Hu, J., Guo, R., Sun, S., and Wang, X. (2020). Strigolactones and brassinosteroids antagonistically regulate the stability of the D53–OsBZR1 complex to determine FC1 expression in rice tillering. Mol. Plant **13**: 586–597.

Finlayson, S.A. (2007). Arabidopsis TEOSINTE BRANCHED1-LIKE 1 regulates axillary bud outgrowth and is homologous to monocot TEOSINTE BRANCHED1. Plant Cell Physiol. **48**: 667–677.

Guo, S., Xu, Y., Liu, H., Mao, Z., Zhang, C., Ma, Y., Zhang, Q., Meng, Z., and Chong, K. (2013). The interaction between OsMADS57 and OsTB1 modulates rice tillering via DWARF14. Nat. Commun. **4**: 1566.

Han, X., Yu, H., Yuan, R., Yang, Y., An, F., and Qin, G. (2019). Arabidopsis transcription factor TCP5 controls plant thermomorphogenesis by positively regulating PIF4 activity. iScience **15**: 611–622.

Howarth, D.G. and Donoghue, M.J. (2006). Phylogenetic analysis of the "ECE" (CYC/TB1) clade reveals duplications predating the core eudicots. Proc. Natl. Acad. Sci. U. S. A. **103**: 9101–9106.

Hu, Z., Yamauchi, T., Yang, J., Jikumaru, Y., Tsuchida-Mayama, T., Ichikawa, H., Takamure, I., Nagamura, Y., Tsutsumi, N., Yamaguchi, S., Kyozuka, J., and Nakazono, M. (2014). Strigolactone and cytokinin act antagonistically in regulating rice mesocotyl elongation in darkness. Plant Cell Physiol. **55**: 30–41.

Hu, Z., Yan, H., Yang, J., Yamaguchi, S., Maekawa, M., Takamure, I., Tsutsumi, N., Kyozuka, J., and Nakazono, M. (2010). Strigolactones negatively regulate mesocotyl elongation in rice during germination and growth in darkness. Plant Cell Physiol. **51**: 1136–1142.

Jeon, J.S., Jung, K.H., Kim, H.B., Suh, J.P., and Khush, G.S. (2011). Genetic and molecular insights into the enhancement of rice yield potential. J. Plant Biol. **54**: 1–9.

Khush, G.S. (2005). What it will take to feed 5.0 billion rice consumers in 2030. Plant Mol. Biol. **59**: 1–6.

Kosugi, S. and Ohashi, Y. (1997). PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell nuclear antigen gene. Plant Cell **9**: 1607–1619.

Kosugi, S. and Ohashi, Y. (2002). DNA binding and dimerization specificity and potential targets for the TCP protein family. Plant J. **30**: 337–348.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. **35**: 1547-1549.

Lan, J. and Qin, G. (2020). The regulation of CIN-like TCP transcription factors. Int. J. Mol. Sci. **21**: 4498.

Levin, K.A. and Boden, S.A. (2020). A new branch of understanding for barley inflorescence development. J. Exp. Bot. **71**: 6869–6871.

Li, C., Potuschak, T., Colón-Carmona, A., Gutiérrez, R.A., and Doerner, P. (2005). Arabidopsis TCP20 links regulation of growth and cell division control pathways. Proc. Natl. Acad. Sci. U. S. A. **102**: 12978–12983.

Li, W., Chen, G., Xiao, G., Zhu, S., Zhou, N., Zhu, P., Zhang, Q., and Hu, T. (2020). Overexpression of TCP transcription factor OsPCF7 improves agronomic trait in rice. Mol. Breed. **40**: 48.

Liu, J., Cheng, X., Liu, P., Li, D., Chen, T., Gu, X., and Sun, J. (2017). MicroRNA319-regulated TCPs interact with FBHs and PFT1 to activate CO transcription and control flowering time in Arabidopsis. PLoS Genet. **13**: e1006833.

Liu, Y., Wang,H., Jiang, Z., Wang, W., Xu, R., Wang, Q., Zhang, Z., Li, A., Liang, Y., Ou, S., Liu, X., Cao, S., Tong, H., Wang, Y., Zhou, F., Liao, H., Hu, B., and Chu, C. (2021). Genomic basis of geographical adaptation to soil nitrogen in rice. Nature **590**: 600-605.

Lu, Z., Yu, H., Xiong, G., Wang, J., Jiao, Y., Liu, G., Jing, Y., Meng, X., Hu, X., Qian, Q., Fu, X., Wang, Y., & Li, J. (2013). Genome-wide binding analysis of the transcription activator IDEAL PLANT ARCHITECTURE1 reveals a complex network regulating rice plant architecture. Plant Cell **25**: 3743–3759.

Luo, D., Carpenter, R., Vincent, C., Copsey, L., and Coen, E. (1996). Origin of floral asymmetry in *Antirrhinum*. Nature **383**: 794–799.

Mao, Y., Wu, F., Yu, X., Bai, J., Zhong, W., and He, Y. (2014). microRNA319a-targeted *Brassica rapa ssp. pekinensis* TCP genes modulate head shape in chinese cabbage by differential cell division arrest in leaf regions. Plant Physiol. **164**: 710–720.

Martín-Trillo, M. and Cubas, P. (2010). TCP genes: a family snapshot ten years later. Trends Plant Sci. **15**: 31–39.

Minakuchi, K., Kameoka, H., Yasuno, N., Umehara, M., Luo, L., Kobayashi, K., Hanada, A., Ueno, K., Asami, T., Yamaguchi, S., and Kyozuka, J. (2010). FINE CULM1 (FC1) works downstream

of strigolactones to inhibit the outgrowth of axillary buds in rice. Plant Cell Physiol. **51**: 1127–1135.

Mohanty, A., Sarma, N.P., and Tyagi, A.K. (1999). *Agrobacterium*-mediated high frequency transformation of an elite indica rice variety Pusa Basmati 1 and transmission of the transgenes to R2 progeny. Plant Sci. **147**: 127–137.

Mukhopadhyay, P. and Tyagi, A.K. (2015). OsTCP19 influences developmental and abiotic stress signaling by modulating ABI4-mediated pathways. Sci. Rep. **5**: 9998.

Murakami, M., Tago, Y., Yamashino, T., and Mizuno, T. (2007). Comparative overviews of clock-associated genes of *Arabidopsis thaliana* and *Oryza sativa*. Plant Cell Physiol. **48**: 110–121.

Narsai, R., Edwards, J.M., Roberts, T.H., Whelan, J., Joss, G.H., and Atwell, B.J. (2015). Mechanisms of growth and patterns of gene expression in oxygen-deprived rice coleoptiles. Plant J. **82**: 25–40.

Navaud, O., Dabos, P., Carnus, E., Tremousaygue, D., and Hervé, C. (2007). TCP transcription factors predate the emergence of land plants. J. Mol. Evol. **65**: 23–33.

Nicolas, M. and Cubas, P. (2016). TCP factors: New kids on the signaling block. Curr. Opin. Plant Biol. **33**: 33–41.

Ori, N., Cohen, A.R., Etzioni, A., Brand, A., Yanai, O., Shleizer, S., Menda, N., Amsellem, Z., Efroni, I., Pekker, I., Alvarez, J.P., Blum, E., Zamir, D., and Eshed, Y. (2007). Regulation of LANCEOLATE by miR319 is required for compound-leaf development in tomato. Nat. Genet. **39**: 787–791.

Pecher, P., Moro, G., Canale, M.C., Capdevielle, S., Singh, A., MacLean, A., Sugio, A., Kuo, C.H., Lopes, J.R.S., and Hogenhout, S.A. (2019). Phytoplasma SAP11 effector destabilization of TCP transcription factors differentially impact development and defence of Arabidopsis versus maize. PLoS Pathog**15**: e1008035.

Poursarebani, N., Trautewig, C., Melzer, M., Nussbaumer, T., Lundqvist, U., Rutten, T., Schmutzer, T., Brandt, R., Himmelbach, A., Altschmied, L., Koppolu, R., Youssef, H.M., Sibout, R., Dalmais, M., Bendahmane, A., Stein, N., Xin, Z., and Schnurbusch, T. (2020). COMPOSITUM 1 contributes to the architectural simplification of barley inflorescence via meristem identity signals. Nat. Commun. **11**: 5138.

Pruneda-Paz, J.L., Breton, G., Para, A., and Kay, S.A. (2009). A functional genomics approach reveals CHE as a novel component of the Arabidopsis circadian clock. Science. **323**: 1481–1485.

Rajaraman, J., Douchkov, D., Hensel, G., Stefanato, F.L., Gordon, A., Ereful, N., Caldararu, O.F., Petrescu, A.J., Kumlehn, J., Boyd, L.A., and Schweizer, P. (2016). An LRR/Malectin receptor-like kinase mediates resistance to non-adapted and adapted powdery mildew fungi in barley and wheat. Front. Plant Sci. **7**: 1836.

Ren, L., Tang, D., Zhao, T., Zhang, F., Liu, C., Xue, Z., Shi, W., Du, G., Shen, Y., Li, Y., and Cheng, Z. (2018). OsSPL regulates meiotic fate acquisition in rice. New Phytol. **218**: 789–803.

Sarvepalli, K. and Nath, U. (2018). CIN-TCP transcription factors: Transiting cell proliferation in plants. IUBMB Life **70**: 718–731.

Shang, Y., Yuan, L., Di, Z., Jia, Y., Zhang, Z., Li, S., Xing, L., Qi, Z., Wang, X., Zhu, J., Hua, W., Wu, X., Zhu, M., Li, G., and Li, C. (2020). A CYC/TB1-type TCP transcription factor controls spikelet meristem identity in barley. J. Exp. Bot. **71**: 7118–7131.

Sharma, R., Kapoor, M., Tyagi, A. K., and Kapoor, S. (2010). Comparative transcript profiling of TCP family genes provide insight into gene functions and diversification in rice and Arabidopsis. J. Plant Mol. Biol. Biotechnol. **1**: 24–38.

Sharma, R.,Agarwal, P., Ray, S., Deveswar, P., Sharma, P., Sharma, N., Nijhawan, A., Jain, M., Kumar, A.S., Singh, V.P., Khurana, J.P., Tyagi, A.K., and Kapoor, S. (2012). Expression dynamics of metabolic and regulatory components across stages of panicle and seed development in indica rice. Funct. Integr. Genomics **12**. 229-248.

Song, X., Lu, Z., Yu, H., Shao, G., Xiong, J., Meng, X., Jing, Y., Liu, G., Xiong, G., Duan, J., Yao, X. F., Liu, C.M., Li, H.,Wang, Y., Li, J. (2017). IPA1 functions as a downstream transcription factor repressed by D53 in strigolactone signaling in rice. Cell Res. **27**: 1128–1141.

Sun, L., Zou, X., Jiang, M., Wu, X., Chen, Y., Wang, Q., Wang, Q., Chen, L., and Wu, Y. (2020a). The crystal structure of the TCP domain of PCF6 in *Oryza sativa L.* reveals an RHH-like fold. FEBS Lett. **594**: 1296–1306.

Sun, Y., Qiao, Z., Muchero, W., and Chen, J.G. (2020b). Lectin receptor-like kinases: the sensor and mediator at the plant cell surface. Front. Plant Sci. **11**: 596301.

Takeda, T., Suwa, Y., Suzuki, M., Kitano, H., Ueguchi-Tanaka, M., Ashikari, M., Matsuoka, M., and Ueguchi, C. (2003). The OsTB1 gene negatively regulates lateral branching in rice. Plant J. **33**: 513–520.

Umehara, M., Hanada, A., Yoshida, S., Akiyama, K., Arite, T., Takeda-Kamiya, N., Magome, H.,

Kamiya, Y., Shirasu, K., Yoneyama, K., Kyozuka, J., and Yamaguchi, S. (2008). Inhibition of shoot branching by new terpenoid plant hormones. Nature **455**: 195–200.

Viola, I.L., Reinheimers, R., Ripoll, R., Uberti Manassero, N.G., and Gonzalez, D.H. (2012). Determinants of the DNA binding specificity of class I and class II TCP transcription factors. J. Biol. Chem. **287**: 347–356.

Wang, F., Han, T., Song, Q., Ye, W., Song, X., Chu, J., Li, J., and Jeffrey Chen, Z. (2020). The rice circadian clock regulates tiller growth and panicle development through strigolactone signaling and sugar sensing. Plant Cell **32**: 3124–3138.

Wang, M., Le Moigne, M.A., Bertheloot, J., Crespel, L., Perez-Garcia, M.D., Ogé, L., Demotes-Mainard, S., Hamama, L., Davière, J.M., and Sakr, S. (2019). BRANCHED1: A key hub of shoot branching. Front. Plant Sci. **10**: 76.

Wang, S., Yang, X., Xu, M., Lin, X., Lin, T., Qi, J., Shao, G., Tian, N., Yang, Q., Zhang, Z., and Huang, S. (2015). A rare SNP identified a TCP transcription factor essential for tendril development in cucumber. Mol. Plant **8**: 1795–1808.

Wang, S.T., Sun, X.L., Hoshino, Y., Yu, Y., Jia, B., Sun, Z.W., Sun, M.Z., Duan, X.B., and Zhu, Y.M. (2014). MicroRNA319 positively regulates cold tolerance by targeting OsPCF6 and OsTCP21 in rice (*Oryza sativa L.*). PLoS One **9**: e91357.

Wei, B., Zhang, J., Pang, C., Yu, H., Guo, D., Jiang, H., Ding, M., Chen, Z., Tao, Q., Gu, H., Qu, L.J., and Qin, G. (2015). The molecular mechanism of SPOROCYTELESS/NOZZLE in controlling *Arabidopsis* ovule development. Cell Res. **25**: 121–134.

Weir, I., Lu, J., Cook, H., Causier, B., Schwarz-Sommer, Z., and Davies, B. (2004). Cupuliformis establishes lateral organ boundaries in *Antirrhinum*. Development **131**: 915–922.

Xie, K., Minkenberg, B., and Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc. Natl. Acad. Sci. **112**: 3570–3575

Yang, C., Li, D., Mao, D., Liu, X., Ji, C., Li, X., Zhao, X., Cheng, Z., Chen, C., and Zhu, L. (2013). Overexpression of microRNA319 impacts leaf morphogenesis and leads to enhanced cold tolerance in rice (*Oryza sativa L.*). Plant, Cell Environ. **36**: 2207–2218.

Yin, D., Liu, X., Shi, Z., Li, D., and Zhu, L. (2018). An AT-hook protein DEPRESSED PALEA1 physically interacts with the TCP family transcription factor RETARDED PALEA1 in rice. Biochem. Biophys. Res. Commun. **495**: 487–492.

Yu, B., Lin, Z., Li, H., Li, X., Li, J., Wang, Y., Zhang, X., Zhu, Z., Zhai, W., Wang, X., Xie, D., and Sun, C. (2007). TAC1, a major quantitative trait locus controlling tiller angle in rice. Plant J. **52**: 891–898.

Yuan, Z., Gao, S., Xue, D.W., Luo, D., Li, L.T., Ding, S.Y., Yao, X., Wilson, Z.A., Qian, Q., and Zhang, D.B. (2009). Retarded Palea1 controls palea development and floral zygomorphy in rice. Plant Physiol. **149**: 235–244.

Zhang, C., Ding, Z., Wu, K., Yang, L., Li, Y., Yang, Z., Shi, S., Liu, X., Zhao, S., Yang, Z., Wang, Y., Zheng, L., Wei, J., Du, Z., Zhang, A., Miao, H., Li, Y., Wu, Z., and Wu, J. (2016). Suppression of jasmonic acid-mediated defense by viral-inducible microRNA319 facilitates virus infection in rice. Mol. Plant **9**: 1302–1314.

Zhang, W., Tan, L., Sun, H., Zhao, X., Liu, F., Cai, H., Fu, Y., Sun, X., Gu, P., Zhu, Z., and Sun, C. (2019). Natural variations at TIG1 encoding a TCP transcription factor contribute to plant architecture domestication in rice. Mol. Plant **12**: 1075–1089.

Regulation of Translation in Plants: Its Impact on Developmental Transitions

Thesis Submitted to the University of Delhi
in the Partial fulfilment for the Degree of
Master of Science
in
Plant Molecular Biology and Biotechnology
2021

Sritama Das

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi 110021, India

Regulation of Translation in Plants: Its Impact on Developmental Transitions

Thesis Submitted to the University of Delhi
in the Partial fulfilment for the Degree of
Master of Science
in
Plant Molecular Biology and Biotechnology
2021

Sritama Das

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi 110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Regulation of Translation in Plants: Its Impact on Developmental Transitions"** was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: June 28, 2021

Signature of the Candidate

(Sritama Das)

Signature of the Supervisor

(Prof. Sanjay Kapoor)

Signature of the HOD

(Prof. Anil Grover)

Preface

Protein synthesis is the defining aspect of all forms of life. It is the culminating step of central dogma and signifies the final regulatory checkpoint in the regulation of gene expression. Contrary to transcription, changes at the translational level are more dynamic and sought to establish more permanent changes in the cell fate. However, the information regarding various regulatory mechanisms that are instrumental in modulating protein translation levels is limited compared to its preceding stage (transcription).

This literary review presents a comprehensive analysis of our current understanding of translational regulation in the plant kingdom. The review opens up with a description of the general process of cytoplasmic mRNA translation along with the associated eukaryotic translation factors at each stage. It then proceeds towards a brief description of recently discovered additional protein factors and other key players; the ribosome, mRNA transcript features, the biocatalyst tRNA and the small RNAs. We will see how ribosomal proteins, covalent modifications in the transcript, it's UTRs and the presence of upstream reading frames (uORFs) in the 5'-leader sequence exert translational control. The impact of small non-coding RNAs on the translation process derived from the degradation of tRNAs known as tRNA-derived fragments (tRFs) has also been discussed. In an attempt to better understand the topic of plant gene regulation at the translational level some notable developmental events have been discussed, namely seed and pollen germination, diurnal regulation of global protein synthesis levels and establishment of root nodule symbiosis in legumes.

Acknowledgement

I would like to take this opportunity to express my gratitude to my mentor, Prof. Sanjay Kapoor for his vision, constant support and encouragement throughout the course of my dissertation. His meticulous nature has inspired me greatly. I am immensely grateful for his understanding and patience towards me.

I extend my gratitude to all the Professors of the Department of Plant Molecular Biology; Prof. Anil Grover, Prof. Jitendra P. Khurana, Prof. Akhilesh K. Tyagi, Prof. Paramjeet Khurana, Prof. Indranil Dasgupta, Prof. Arun Sharma, Prof. Girdhar K. Pandey, Prof. Saurabh Raghuvanshi and Prof. Surekha K. Aggarwal for their invaluable lessons which have taught me to strive hard and remain dedicated in every step I take.

I am thankful to Ms. Sanchi Bhimrajka without whose relentless help and guidance this thesis would not have been possible. I thank all the members of the department for providing us students with a holistic environment to learn and grow. I also thank my classmates for being a constant source of motivation in tough times. Lastly, I am grateful to God and my parents for their countless blessings and support in all endeavours of my life.

List of Abbreviations

ATP	Adenosine Triphosphate
BiFC	Bimolecular Fluorescence Complementation assay
CBP	Cap Binding Protein
DDM1	Decrease in DNA Methylation I
eIF	(eukaryotic) Initiation Factor
eEF	Elongation Factor
eRF	Release Factor
GTP	Guanosine Triphosphate
GDP	Guanosine Diphosphate
hnRNP	Heterogenous ribonucleoprotein
kDa	Kilo Dalton
lncRNA	Long non-coding RNA
mRNA	Messenger RNA
mORF	Main Open Reading Frame
miRNA	Micro RNA
min	Minutes
NAD	Nicotinamide Adenine Dinucleotide
NMD	Non-sense Mediated Decay
nt	Nucleotide
PABP	Poly (A) Binding Protein
PIC	Pre-Initiation Complex
uORF	Upstream Open Reading Frame
UTR	Untranslated Region
RNA	Ribonucleic Acid
rRNA	Ribosomal RNA
RP	Ribosomal Protein
RBP	RNA Binding Protein
sncRNA	Short non-coding RNA
tRNA	Transfer RNA
tRF	tRNA-derived Fragment
TE	Translation Efficiency
TIS	Translation Initiation Site
TOR	Target Of Rapamycin

Table of Contents

PREFACE..... i

ACKNOWLEGDEMENT..... ii

LIST OF ABBREVIATIONS..... iii

REVIEW OF LITERATURE..... 1

INTRODUCTION..... 1

1. THE GENERAL PROCESS OF mRNA TRANSLATION IN CYTOPLASM..... 2

1.1 PRELUDE TO INITIATION..... 2
1.2 INITIATION..... 3
1.3 ELONGATION..... 11
1.4 TERMINATION..... 13

2. TRANSLATIONAL CONTROL BY ACCESSORY TRANSLATION FACTORS..... 14

3. RIBOSOME IN REGULATION..... 19

4. IMPACT OF RNA ON TRANSLATIONAL REGULATION..... 20

4.1 mRNA STRUCTURAL FEATURES..... 20
4.2 TRANSFER RNAs..... 27
4.3 SMALL RNAs..... 28

5. IMPLICATIONS OF TRANSLATIONAL REGULATION ON PLANT GROWTH AND DEVELOPMENT 29

5.1 SEED GERMINATION..... 29
5.2 POLLEN TRANSLATOME..... 31
5.3 TRANSLATION AND CIRCADIAN CLOCK..... 31
5.4 ROOT NODULE SYMBIOSIS..... 33

6. REFERENCES 36

Review of Literature

Introduction

Protein synthesis is a fundamental process and constitutes a key step in continuing the flow of information from gene to protein to ultimately the phenotype. The translation process depends as much on the abundance and accessibility of the mRNA transcript, the protein-making templates, as it does on the activity of a diverse medley of proteins (translation factors), tRNA and the RNA-protein complex, ribosome, which make up the machinery that carries out the decoding task (Browning and Bailey-Serres, 2015; Merchante et al., 2017). Naturally, a process of such supreme importance that ultimately brings about the synthesis of the building blocks of life is ought to be placed under tight control by nature. Changes in gene expression levels are central to nearly all biological processes, be it cell differentiation, organ development, or the adaptation of an organism to the ever-changing environment. Compared to regulation at transcription checkpoints, the translational control of existing mRNAs allows for more dynamic changes in cellular concentrations of the encoded proteins. Hence, it is a greater determinant of the overall protein levels (Vogel and Marcotte, 2012; Sonenberg and Hinnebusch, 2009). The process of translation is divided into three main stages- initiation, elongation and termination, and the processes constituting these three phases ensure the precise start, timely progression and correct end of translation (Browning and Bailey-Serres, 2015; Raabe et al., 2019).

The molecular machinery for protein synthesis is profoundly similar between plants and other eukaryotes (yeast, mammalian systems) (Urquidi Camacho et al., 2020). However, the sessile lifestyle of plants lending them the capacity to respond to stresses *in situ* and the ability to perform biological activities unique to them, such as photosynthesis, nitrogen fixation, has led plants to evolve translational control mechanisms that are relevant to address their own needs (Browning and Bailey-Serres, 2015). Consequently, the last ten years have witnessed increasing interest in translational aspects. Supported by technical advancements, this has led to revelations of the remarkable ways in which plants utilise translational control mechanisms. This includes multiple situations, ranging from regulating protein synthesis rate in response to metabolites, defence against viruses, altering protein levels under environmental perturbations, and even in crucial developmental transitions such as nodulation,

flowering, seed, and pollen development. In addition, the resource partitioning that takes place over the diurnal cycle also incorporates control mechanisms, monitoring the dynamic proteome (Urquidi Camacho et al., 2020). Development of new-age *in silico* and molecular techniques such as RNA-seq, Ribo-seq, TRAP-seq, *in vitro* translation, coupled with refinement of existing molecular techniques such as polysome profiling, ribosome foot-printing, has significantly enriched our understanding of the intricacies of regulation that determine the fates of transcripts within the cytoplasm of plant cells and also sets the premise for exciting future research.

This review summarizes the current status of translation regulation in plants and highlights the instances of translation control mechanisms in developmental aspects.

1. The General Process of mRNA Translation in Cytoplasm

1.1 Prelude to Initiation

The export of an mRNA from the nucleus to the cytoplasm sets a dynamic chain of events in motion. Translation of mRNA on two-subunit ribosomes (the 70S in prokaryotes and 80S in eukaryotes) is a conserved process amongst all organisms, from the most ancient to the most recent of life forms. Upon export from the nucleus to the cytoplasm, the mRNA is bound at its 5'-terminus by a two-subunit nuclear cap-binding complex (CBC) composed of CAP BINDING PROTEIN20 (CBP20) and CBP80, and poly(A)- binding proteins (PABPs) bound to the 3'-poly(A) tail. The mechanisms of how the CBC is exchanged for a cytoplasmic cap-binding complex remain to be defined (Chantarachot and Bailey-Serres, 2018).

In the cytoplasm, the exported mRNAs undergo a pioneer round of translation as a quality-control mechanism. After the first round of translation and replacement of CBP20 by the eukaryotic translation initiation factor 4E (eIF4E), only mRNAs that pass the quality-control check can enter the active translation pool as templates. mRNAs with defects in the translation are subjected to degradation via Non-sense Mediated Decay (NMD), No-go Decay, Non-stop Decay pathways. While the NMD pathway has been well-characterised in plants, there are no reports of the no-go decay and the non-stop decay in plants. (Chantarachot and Bailey-Serres, 2018; Urquidi Camacho et al., 2020; Raxwal et al., 2020).

The quality-checked mRNAs now enter the first stage, i.e., initiation, wherein the maximum amount of regulation is exerted (Browning and Bailey-Serres, 2015; Urquidi Camacho et al., 2020). Subsequently, several translation initiation factors (eIFs) work together in a highly coordinated manner to ensure mRNA-ribosome association, scanning of the mRNA, start codon selection and its decoding by the initiator methionyl tRNA specialised for initiation (Met-tRNA$^{Met}_i$) (Merchante et al., 2017; Raabe et al., 2019).

1.2 Initiation

The initiation is the most well-studied and divergent phase of translation. Besides fundamental differences between bacteria and eukaryotes in start codon selection mechanisms (base-pairing in bacteria versus scanning in eukaryotes), translation initiation is significantly diverse amongst eukaryotes (Sonenberg and Hinnebusch, 2009; Urquidi Camacho et al., 2020). In bacteria, the translational control mainly involves moderating ease-of-access of the Shine-Dalgarno sequence. However, in eukaryotes, the start codon is detected by a scanning mechanism where the small (40S) ribosomal subunit, loaded with the initiator tRNA, Met-tRNA$^{Met}_i$ (coming from the ternary complex) in a pre-initiation complex (PIC), binds to the mRNA transcript near the 5'-end and starts scanning for an AUG (or rarely a near-cognate AUG) codon, upon successful recognition of which, initiation begins (Fig. 1) (Sonenberg and Hinnebusch, 2009; Browning and Bailey-Serres, 2015). Instances of translation initiation at non-AUG codons have been confirmed by Ribo-seq for some *Arabidopsis* mRNAs, such as AGAMOUS, FCA, and GDP-L-GALACTOSE PHOSPHORYLASE (GGP) (Merchante et al., 2017; Li and Liu, 2020).

All these processes of mRNA binding to the PIC and subsequent scanning are aided by eukaryotic initiation factors (eIFs) that recognise the mRNA's cap structure at the 5' end and the poly(A) tail at the 3' end.

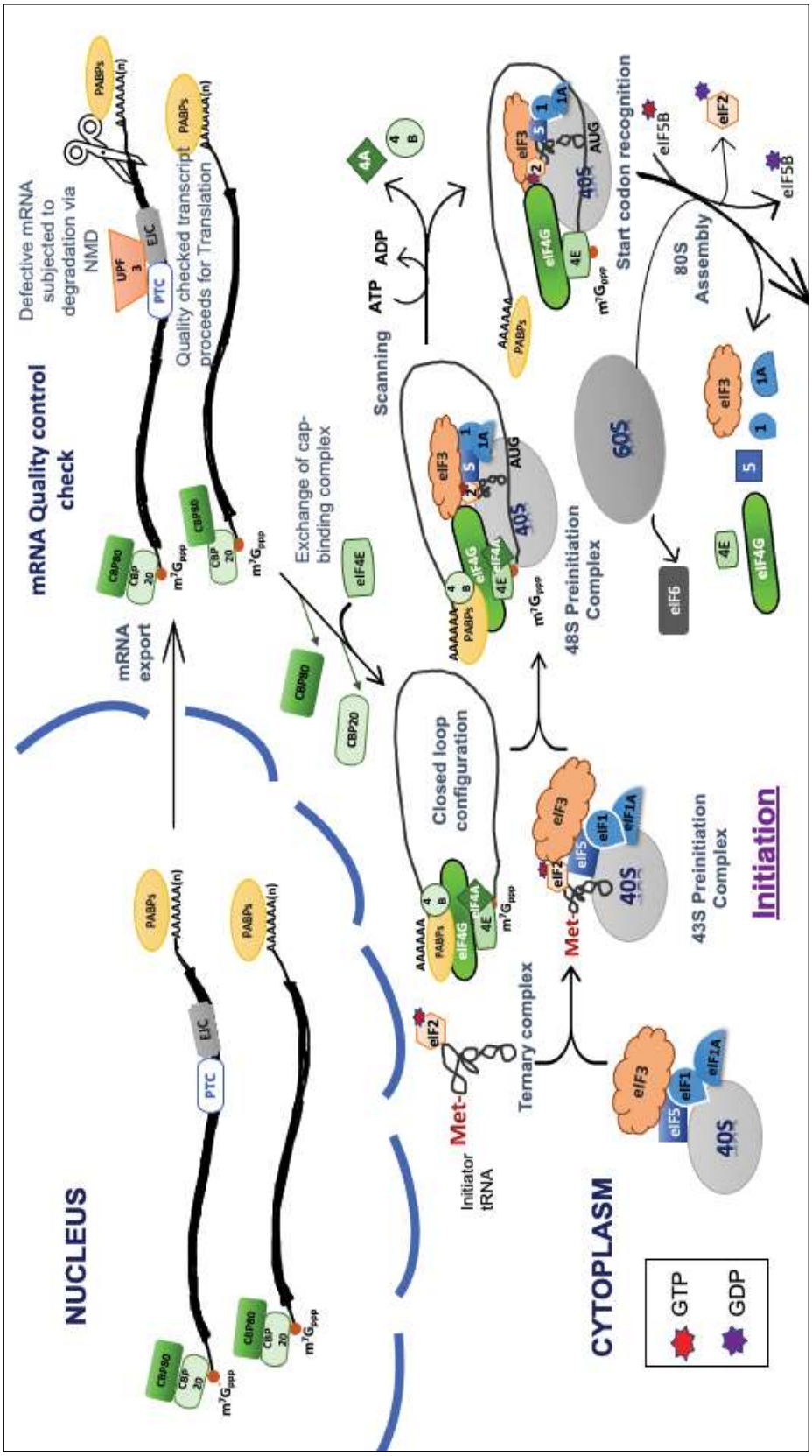


Fig. 1 General overview of cytoplasmic mRNA translation. mRNA transcript is exported from nucleus to cytoplasm with the nuclear CBP20/80 cap-binding complex bound to the 5'cap (m^7G_{ppp}) and poly(A)binding proteins (PABPs) bound to the 3'poly(A)tail. Following export, quality control check takes place; mRNAs with defects such as those containing a premature termination codon(PTC) upstream of an exon-junction complex(EJC), unusually long length etc. are subjected to degradation via Non-sense Mediated Decay(NMD) that requires certain proteins named UP Frameshift (UPF1,2,3) and other accessory components(not shown in figure).The quality-checked mRNA enters into active translation. Translation initiation begins with assembly of a 48S Preinitiation complex; formed by association of the 40S ribosomal subunit loaded with initiation factors (eIFs) 1,1A,3,5 along with eIF2-containing ternary complex (43S PIC). The linear mRNA is re-orientated in a closed-loop via association of PABPs with cap-binding complex (4F,4E+4G) with aid of scaffolding action of 4G. This mRNA is then loaded onto 40S subunit with eIF3'assistance forming the 48S PIC; aided by eIF4A helicase and its partner 4B. Then, scanning and start codon(AUG) recognition with help of eIF1,1A,5 takes place, eIF2 delivers the Met-tRNAi to P-site of 40S following GTP hydrolysis. Thereafter, all eIFs are released sequentially. eIF6 unbounds the 60S subunit which then joins the 40S subunit, catalyzed by GTPase eIF5B (contd. in next page).

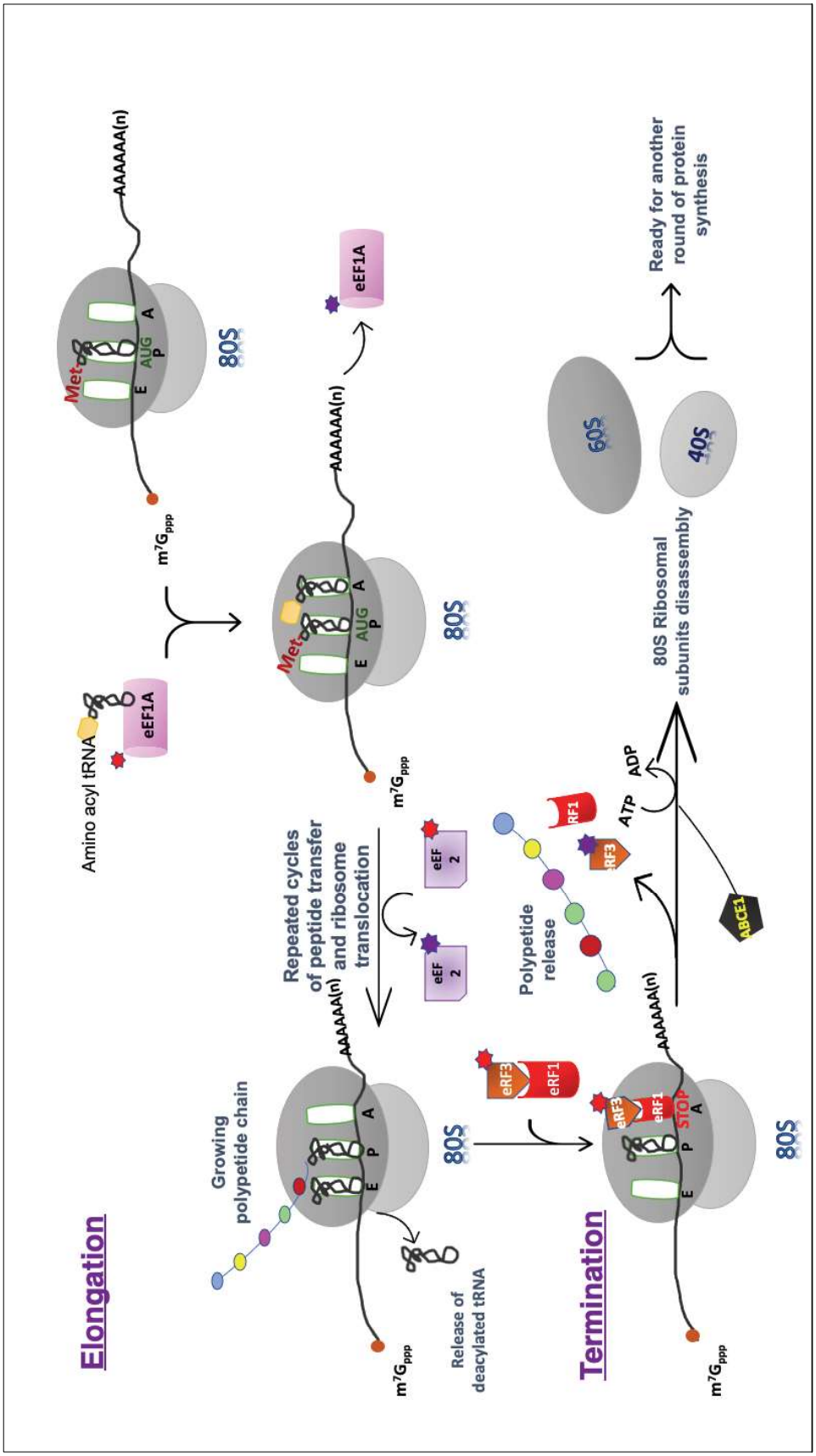


Fig. 2 General overview of cytoplasmic mRNA Translation. In the next stage, **Elongation,** elongation factor eEF1A delivers a charged amino acyl-tRNA to A-site of the 80S ribosme. A correct codon-anticodon match ensures peptide bond formation by peptidyl transferase activity of the ribosome and subsequent release of the eEF1A-GDP following GTP hydrolysis. Next, translocation follows aided by eEF2 which also utilizes GTP hydrolysis to shift the mRNA-tRNA complex to the next codon. The deacylated tRNA exits the ribosome through E-site. Decoding and simultaneous polypeptide formation continues in same manner until a stop codon is encountered, which triggers the last stage, i.e., **Termination.** A stop codon (UAA, UAG, UGA) signals recruitment of the release factor complex, eRF1-3-GTP in the A-site. Following GTP hydrolysis by eRF3, translation is terminated, and polypeptide is released from the P-site. The hydrolysis also opens binding surface for the recycling factor, ATP-binding cassette protein ABCE1 which presides over a final ATP hydrolysis event leading to 80S ribosomal subunits disassembly. The subunits can then be recycled for another round of translation.

Factors Regulating 43S PIC Binding to mRNA

The 43S pre-initiation Complex is formed by the joining of Ternary Complex (TC) with the 40S ribosome subunit; the process is facilitated by eIF1, eIF1A and eIF3 (Browning and Bailey-Serres, 2015; Merchante et al., 2017; Urquidi Camacho et al., 2020). The TC is composed of the GTP-binding protein eIF2, GTP and the initiator Met-tRNA$^{Met}_{i,}$ which can be distinguished from the Met-tRNA^{Met} used during elongation. The Initiator tRNA has evolved several strategies to avoid interaction with elongation factor eEF1A. One strategy that exists in plants and fungi is the modification of a certain nucleotide in the T-loop of Met-tRNA$^{Met}_{i}$ with a large O-ribosylphosphate moiety to prevent its interaction with eEF1A (Browning and Bailey-Serres, 2015).

a) eIF2 Group

This group of factors functions in formation of the ternary complex and the exchange of GDP for GTP from eIF2+GDP by eIF2B (guanine nucleotide exchange factor or GEF) (Urquidi Camacho et al., 2020). It is an eight member protein family. eIF2A and eIF2D are new members identified in animals. Although genes for eIF2A and eIF2D appear to be conserved in plants, but their role in plant initiation has not been elucidated (Browning and Bailey-Serres, 2015).

eIF2 is the principal member of this group and delivers Met-tRNA$^{Met}_{i}$ to 40S subunit; facilitated by its GTPase activity in presence of another initiation factor, eIF5. The eIF2 complex is composed of three subunits, eIF2α, eIF2β and eIF2γ. eIF2α and eIF2β interact with eIF2γ to form the core complex. It is known to contain the GTP nucleotide binding site (Browning and Bailey-Serres, 2015). It is worth mentioning here that one of the best-characterized mechanisms used by yeasts and especially mammals for regulating initiation of translation involves the phosphorylation of a single conserved serine (Ser51) residue on the eIF2α subunit. Evidences for this phosphorylation event has also been found in plants. Phosphorylation of the conserved Ser51 residue (by any kinase) in the eIF2α subunit inhibits eIF2B, the guanine nucleotide exchange factor for eIF2, dissociation from the eIF2-GDP complex and inhibits the formation of a new ternary complex. Hence, the gradual depletion and non-availability of ternary complex ultimately arrests translation initiation (Merchante et al., 2017; Sesma et al., 2017). In plants, the only kinase known to phosphorylate eIF2α is GENERAL CONTROL NONDEREPRESSIBLE2 (GCN2). The GCN2-dependent eIF2α

phosphorylation in plants regulates protein synthesis, even though the utility of this mechanism as for generally inhibition the process of translation is limited by the responses to 8-azaadenine, which is a purine synthesis inhibitor, and the amino acid synthesis inhibitor chlorsulfuron (a herbicide). Reports suggest a function for GCN2 and eIF2α phosphorylation in plant immunity, abiotic and nutritional stresses and in response to various phytohormones (jasmonate, ethylene and salicylic acid), although their actual role and underlying molecular signals remains open for future research (Merchante et al., 2017; Sesma et al., 2017; Urquidi Camacho et al., 2020). Phosphorylation of factors is a key mechanism that controls mRNA translation in eukaryotes and some more evidence for this has been discussed in a later section.

eIF2B is hypothesized to function as the guanine nucleotide exchange factor for eIF2, although its necessity in this capacity in plants has been questioned. No plant source depicts eIF2B-like activity yet, although genes with sequential similarity to mammalian eIF2B subunits are encoded by *Arabidopsis* (subunit ™), suggesting that although the proteins are expressed but probably the eIF2B recycling may not be necessary, and thus phosphorylation of eIF2α may not have as strong an inhibitory effect on translation in plants as in mammals or yeasts (Browning and Bailey-Serres, 2015; Sesma et al., 2017). Support for this claim stems from a study that reported that the binding of GDP to eIF2 is only about 10-fold higher than that of GTP in comparison to the ~100-fold difference for mammalian eIF2. There is still much to learn as to what extent eIF2 subunit phosphorylation regulates protein synthesis in plants (Browning and Bailey-Serres, 2015).

b) eIF3

The eIF3 works by binding to the 40S subunit and participates in the formation of the 43S pre-initiation complex (PIC). eIF3 basically functions as a bridge to accelerate binding of mRNA along with its associated factors, eIF4F, eIF4A, eIF4B and mRNA binding protein, Poly(A) binding protein (PABPs), with the PIC, thus orchestrating the 48S PIC (Scanning Complex) formation (Fig. 1) (Merchante et al., 2017).

eIF3 is by far the largest of all the eIFs and plays a pivotal role in the initiation phase as it interacts with numerous factors as well as the ribosome (Raabe et al., 2019). Hence, only few Arabidopsis eIF3 subunit mutants have been reported. In Arabidopsis, five of the subunits (a, e, f, h, k) for eIF3 are encoded by single genes and eight subunits (b, c, d, g, i, j, l, m) are encoded by two genes. According to two separate

studies, mutations in eIF3e or eIF3f cause male gametophytic lethality and genotypes that only express eIF3f in pollen indicate that the absence of eIF3f is also lethal for embryo. Also, transgenics with partial loss of functional alleles of the single eIF3h gene are although viable but display multiple developmental defects, including reduced male gamete transmission. So far, eIF3k appears to be non-essential under normal growth conditions (Browning and Bailey-Serres, 2015). eIF3 is also important in translation of endogenous plant mRNAs that possess a uORF (upstream Open Reading Frame); particularly the eIF3h subunit which is necessary for translation reinitiation of mRNAs with uORFs (Merchante et al., 2017; Raabe et al., 2019; Urquidi Camacho et al., 2020).

c) eIF4 group

The process of translation starts when the cytoplasmic cap-binding complex recognises the cap structure and the poly(A)-binding protein (PABP) binds to the mRNA poly(A) tail (Merchante et al., 2017).

The cap-binding complex, eIF4F is composed of the cap-binding protein eIF4E and the scaffolding protein eIF4G. Plants possess an abundant, alternate isoform of the cap-binding complex, named eIFiso4F, consisting of isoforms, eIFiso4E and eIFiso4G. These plant-specific isoforms are known to play key roles in the defence against viruses (Browning and Bailey-Serres, 2015; Urquidi Camacho et al., 2020). The viral 5'-linked protein (VPg) functions as a substitute for the canonical 5' cap in some plant viruses such as potyviruses and engage with the cap-binding complex for translation of its mRNA. eIF4iso4E have less affinity for the VPg thereby lending some passive anti-viral resistance by limiting translation initiation of viral mRNAs (Urquidi Camacho et al., 2020).

eIF4G brings the helicase eIF4A, eIF4E, PAPB and eIF3 together, circularizing the mRNA (Fig. 1). Circularisation of the mRNA serves as quality control to check for mRNA integrity, stimulates translation initiation, protects the mRNA ends from degradation and promotes 40S recycling (Chantarachot and Bailey-Serres, 2018; Urquidi Camacho et al., 2020). eIF4A is a DEAD-box motif containing RNA helicase, which in association with its co-factor eIF4B (RNA binding protein), unwinds secondary structures of the mRNA (ATP-dependent), hence allowing for the recruitment of the 43S pre-initiation complex (43S PIC). It further functions to remove the presence of any secondary structures and RNA binding proteins in the 5' UTR during the scanning of leader sequence by the PIC (Browning and Bailey-Serres, 2015; Merchante et al.,

2017; Urquidi Camacho et al., 2020; Tyagi et al., 2021). In addition to its involvement in translation initiation, eIF4A has also been shown to have regulatory roles in plant growth and stress response. From loss-of-function studies conducted in the grass family, *Oryza sativa* and *Brachypodium distachyon,* eIF4A was found to closely interact with the methyltransferase DRM2 (domains rearranged methyltransferase), indicating an implication in epigenetic regulation of plant development. A recent study on the model moss species, *Physcomitrella patens* identified eIF4A's interactions with chromatin remodeler protein DDM1 (Decrease in DNA Methylation 1) and a heterogenous ribonucleoprotein, LIF2L1 and between DDM1 and LIF2L1 through yeast two-hybrid assay and bimolecular fluorescence complementation assays. The *Arabidopsis* homolog of PpLIF2L1 regulates transcription of stress-induced genes. The salt-stress induced co-expression analysis suggests a close-knit regulatory network active in abiotic stress tolerance in *P. Patens* and a conservation of these coordinated stress-adaptive regulatory mechanisms across land plants (Tyagi et al., 2021).

Another member of this group, eIF4H is present in mammals; however, a counterpart for this factor has not been identified in yeast or plants (Browning and Bailey-Serres, 2015). A major form of regulation of mammalian translation is through 4E binding proteins (4EBPs) but the presence of these highly-conserved proteins, surprisingly remain unconfirmed in the plant kingdom (Sesma et al., 2017; Urquidi Camacho et al., 2020).

Factors Regulating Start Codon Selection

Once brought to the 5' end of the mRNA, the 48S scanning complex scans down the 5' leader of the mRNA until the Met-tRNA$^{Met}_{i}$ recognizes and base pairs with an AUG start codon and signals the beginning of protein synthesis (at the translation initiation site, TIS). The small protein factors eIF1, eIF1A and eIF5 assist in start codon recognition (Fig. 1). The nucleotide distribution around the TIS plays a crucial role in determining the translatability of a particular mRNA transcript (translation efficiency) (Browning and Bailey-Serres, 2015; Gupta et al., 2016). The scanning ribosome is able to recognise the AUG initiation codon better when it appears in a particular (favourable) "Kozak"-sequence context, which in eudicot plants is AANAUGGC, in monocotyledonous plants is GCNAUGGC and GCNAUGGC for plants in general, as

determined by protein abundance comparisons (Gupta et al., 2016; Merchante et al., 2017).

a) eIF1 group

eIF1 binds near the peptidyl (P)-site of the 40S subunit and prevents Met-tRNA$^{Met}_{i}$ bound to eIF2+GTP (the ternary complex) from being fully engaged in the peptidyl (P)-site of the 43S PIC until the initiation codon is identified. Similar to eIF1, eIF1A also participates in preventing full P-site engagement of the Met-tRNA$^{Met}_{i}$ until the correct initiation codon is identified (Browning and Bailey-Serres, 2015; Merchante et al., 2017). eIF1 and eIF1A thus play a key role in maintaining the fidelity of start codon selection mechanism.

b) eIF5 group

The principle member, eIF5 has GTPase-activating protein function whereas the other member, eIF5B promotes binding of the 60S subunit (subunit joining) (Browning and Bailey-Serres, 2015; Merchante et al., 2017; Urquidi Camacho et al., 2020). After start codon recognition, eIF1 is displaced which promotes hydrolysis of the GTP bound to ternary complex, a process catalysed by eIF2 and promoted by eIF5. eIF5 works as a GDP dissociation inhibitor for eIF2+GDP until eIF2B is able to complete the replacement of GDP with GTP (Browning and Bailey-Serres, 2015).

Factors Regulating Subunit Joining

a) eIF5B & eIF6

eIF5B GTPase is recruited next and mediates the joining of the 60S subunit to the 40S. The binding of eIF5B GTPase displaces the associated factors (eIF2, eIF3 and eIF5) from the 40S subunit which opens up surface for 60S subunit attachment. eIF5B is aided by another small protein factor eIF6 which interacts with the 60S subunit and prevents its premature association with the 40S subunit. Upon phosphorylation, eIF6 loses its affinity and detaches from 60S, hence enabling the formation of 80S ribosome (Fig. 1) (Merchante et al., 2017; Urquidi Camacho et al., 2020). eIF6 is also known to function in ribosome biogenesis and transport of ribosomal subunits from nucleus to the cytoplasm (Browning and Bailey-Serres, 2015; Urquidi Camacho et al., 2020).

The subunits assembly in turn promotes GTP hydrolysis and the release of eIF5B and eIF1A concludes initiation. At this stage, the ‘elongation-competent’ 80S ribosome is

fully assembled at the start codon site and all the eIFs have dissociated (Merchante et al., 2017).

1.3 Elongation

The second phase of translation, called elongation, constitutes the progression of ribosome catalysed polypeptide formation via mRNA decoding. At the end of initiation, the Methionyl-tRNA$^{Met}{}_{i}$ occupies the peptidyl (P) site of the assembled 80S ribosome and the aminoacyl (A) site of the ribosome is lying vacant; ready to recruit the second charged tRNA (aminoacyl-tRNA) with the assistance of corresponding elongation factors (eEFs) (Browning and Bailey-Serres, 2015; Merchante et al., 2017).

During the course of elongation, charged tRNA(s) is recruited into the empty aminoacyl (A)-site of the ribosome, followed by the peptidyltransferase activity of the large ribosomal RNA (25-26S rRNA) that catalyses the peptide bond formation. The mRNA–tRNA hybrid along with the ribosome keeps moving forward, one codon at a time (translocation), followed by subsequent ejection of the deacylated (uncharged) tRNA through the exit (E) site (Fig. 1) (Browning and Bailey-Serres, 2015; Merchante et al., 2017). It is worth mentioning that the speed (slowing down/acceleration) of elongation process in general, has no significant or lasting effects on the overall protein synthesis rate (translation efficiency) unless it blocks initiation, or elongation stalls completely. Pacing down elongation by say, two-fold will although increase the ribosome density by two-fold, but the actual rate at which protein synthesis will finish would eventually remain unaltered. It is however possible that if there is a slowing of elongation rate globally (across entire cell), it would certainly reduce the pool of free ribosomes, and in-turn may inhibit bulk initiation and bring about some sort of significant change in the overall translation efficiency (Urquidi Camacho et al., 2020).

Factors Involved in aa-tRNA Recruitment

a) eEF1A & eEF1B

The mode of action of this group of eEFs is quite similar to the activity of the eIF2 group in the process of translation initiation. eEF1A forms an elongation-specific ternary complex with GTP and an aa-tRNA; and delivers the aa-tRNA to the vacant A site (in the peptidyl transferase centre of the large subunit) (Fig. 2). Codon–anticodon recognition between the corresponding mRNA codon and the aminoacyl-tRNA results in the hydrolysis of GTP, followed by structural rearrangements of the tRNA, eEF1A and the ribosome. This leads to release of eEF1A-GDP which eventually is

regenerated by its guanine nucleotide exchange factor (GEF), eEF1B (Browning and Bailey-Serres, 2015; Merchante et al., 2017).

eEF1A has also been reported to have functions and interactions outside of its role in translation. This factor is known to interact with cytoskeleton (which may reflect an association of the translation process with cytoskeleton anchors), participate in processes such as nuclear export of tRNAs, targeting of misfolded damaged proteins to proteasome, proteolysis, apoptosis and viral propagation like the tombusvirus replication complex and the 3' tRNA-like structure of turnip yellow mosaic virus. Contrary to eEF1A, very little is known about eEF1B from plants; it has been shown to play a role in viral replication, like its partner and in cap-binding protein subunits (Browning and Bailey-Serres, 2015).

Factor Involved in Ribosome Translocation

a) eEF2

The ribosome moves forward by one codon with the help of another GTP-binding factor, eEF2 (Fig. 2). Post the peptide bond formation, the now uncharged tRNA must be moved from the P-site into the E-site, which would free the A-site for accepting the next incoming charged tRNA (next ternary complex). eEF2 (bound to GTP) binds to the ribosome and its GTP hydrolysis results in a three-nucleotides shift of the mRNA-tRNA hybrid, which coincides with the movement of the P-site deacylated tRNA to the E-site and ultimately its discharge from the translocating ribosome (Browning and Bailey-Serres, 2015; Merchante et al., 2017).

In mammals, eEF2 also functions as a substrate for phosphorylation by the Ca^{2+}/calmodulin-dependent eEF2 kinase (eEF2K), which reduces eEF2 association with the ribosome. This Ca^{2+}-regulated eEF2 phosphorylation which is thought to fine-tune protein synthesis at a local (cell type-specific) level is yet to be elucidated in plants *in vivo*. *In vitro* translation studies reported a reduction in eEF2 activity in presence of a mammalian Ca^{2+}/calmodulin-dependent eEF2 kinase. However, eEF2 has been hinted to have a role in cold stress acclimation (Browning and Bailey-Serres, 2015).

Interestingly, eEF5, which was earlier placed in the eIF5 group (previously known as *eIF5A*) is no longer considered to be an initiation factor, but an elongation factor involved in the efficient elongation of proteins containing proline (imino acid) or glycine residues. Studies on plant eEF5 have suggested a role in growth, development

(protoxylem development in *Arabidopsis* seedling roots) and stress responses (abiotic, pathogen, iron deficiency) (Browning and Bailey-Serres, 2015; Urquidi Camacho et al., 2020).

1.4 Termination

Once translocation places one of the three stop codons (UAA, UGA, or UAG) into the A-site of the ribosome, it marks the end of elongation phase. Termination is mediated by two release factors eRF1 and eRF3, that together with a GTP molecule form a termination-specific ternary complex i.e., eRF3+GTP bound to eRF1, which acts as a GTP dissociation inhibitor (Merchante et al., 2017; Urquidi Camacho et al., 2020). This complex is recruited when the elongating ribosome reaches a stop codon on the mRNA, preventing further entry of any elongation-specific ternary complexes (eRF3+GTP+eRF1). eRF1 is delivered to the A-site with the help of eRF3's GTPase activity which also leads to a simultaneous hydrolytic release of the nascent polypeptide (Fig. 2). The structure of these eRFs resembles a tRNA, which facilitates the N-terminal domain of the release factors to dock in the A-site and directly interact with the stop codon. The mechanism by which the eRFs recognise all the three codons is still lacking in molecular details (Browning and Bailey-Serres, 2015).

The polypeptide chain and tRNA are released from the ribosome and the 80S ribosome dissociates into its subunits which can now be recycled for following rounds of translation. The ATP-binding cassette protein ABCE1 (AtRLI in Arabidopsis) separates the 40S and 60S ribosomal subunits (Fig. 2) (Merchante et al., 2017).

Translation Reinitiation

In some translation events, after ribosome disassociation, the 40S subunit is retained to start translation of another ORF present in the same transcript. This is a frequent feature of uORF-containing transcripts. The 40S subunits reacquires lost initiation factors, 43S PIC reassembles, scanning recommences, and translation begins at a downstream ORF (Merchante et al., 2017; Urquidi Camacho et al., 2020).

The presence of eIF3h subunit in particular has been found to be necessary for reinitiating translation of uORF sequence containing genes including *LHY,CLV1,AS1,bZIP11*, as implicated by mutational studies. Mutants of *Arabidopsis* lacking *eIF3h* gene show growth arrest and a series of phenotypic defects like enlarged

shoot meristem, altered organ polarity, pin-formed inflorescences and are compromised in *ARF* mRNAs translation (Zhou et al., 2014; Merchante et al., 2017). Though the actual mechanistic role of eIF3h is unclear; it is thought to overcome the inhibitory effect of uORFs on the downstream ones (succeeding uORFs and/or of the main protein encoding ORF) by reinforcing translation initiation (Merchante et al., 2017). It has been seen that reinitiation is more likely to take place when the intercistronic region is lengthy and is after a short ORF (like an uORF) (Merchante et al.,2017; Urquidi Camacho et al.,2020). The ability of eIF3h to reinitiate after translation termination at uORFs is under phosphorylation control by S6K (AtS6K1), another TOR mediated (in presence of auxin which activates TOR via the GTPase ROP2) biochemical process (Schepetilnikov et al., 2013; Schepetilnikov and Ryabova, 2017, 2018).

2. Translational Control by Accessory Translation Factors

In the recent years, it has been unearthed that the core translation factors at times partner with a variety of additional proteins that influence their activity and in turn the efficiency of whole translation process (Urquidi Camacho et al., 2020).

eIF4E-Binding Proteins (4E-BPs)

4E-BPs repress translation initiation by competing with the 5' cap for binding to eIF4E. They are active during slow growth and are inhibited by TOR signalling (mTOR). This protein plays an abundant role in regulating the initiation phase in mammals; however, counterparts (homologs) for 4E-BPs are missing in plant genomes (Sesma et al., 2017; Urquidi Camacho et al., 2020).

Conserved Binding of eIF4E 1 (CBE1)

This group of factors also interact with cap-binding proteins and have been identified in plants as well. They play a similar role in suppressing initiation by interacting with eIF4E/iso4E. CBE1 action depends on TOR signalling pathway and its phosphorylation (activation) has been shown to be induced by sucrose and its activity rapidly reduced by TOR inhibitors. Furthermore, CBE1 attenuates (reduces) expression of some cell cycle genes (Urquidi Camacho et al., 2020).

CERES

CERES is a plant-specific eIF4E/iso4E-binding protein and promotes translation of light and carbohydrates related mRNAs, i.e., it is active during daytime. It can be found as a part of actively translating initiation complexes, such as eIF3. Strangely, eIF4G/iso4G is absent when CERES is present in the complex (Urquidi Camacho et al., 2020).

SOAR1

SOAR1 is also plant-specific and suppresses initiation by binding to eIFiso4G, thus obstructing eIF4iso4E and PABP. It is known to regulate translation of mRNAs that are involved in ABA signalling pathway (Urquidi Camacho et al., 2020).

Ribosome Inactivating Proteins (RIPs)

These proteins are encoded widely in plant genomes and possess anti-microbial and anti-viral functions. RIPs represent a class of glycosidases that depurinate viral RNAs and also 25S rRNA. Evidence of these potentially toxic proteins having roles in defence and biology has been discovered in the recent years (Urquidi Camacho et al., 2020).

RNA Binding Proteins (RBPs)

RBPs play crucial roles in RNA metabolism and are present in and associated with key biological processes including protein synthesis regulation. RBPs are characterized by RNA binding domains (RBDs) that interact with either single or double stranded RNAs. RBPs containing RNA Recognition Motif (RRM) are the most abundant in plant genomes (Mahalingam and Walling, 2020).

Of this vast protein family, Poly(A) binding proteins (PABPs) are so crucial that they are considered to be among the core translation factors, playing fundamental role in the closed-loop model formation. Also, the binding of PABPs to 3' end of mRNA transcripts has a direct positive correlation with the translation efficiency (Urquidi Camacho et al., 2020). Along the similar lines, binding of RBPs to a particular nucleotide stretch (170 nt downstream of the stop codon) in the Poly(A) of a group 6 Late embryogenesis abundant protein encoding transcript in the legume, *Phaseolus vulgaris* (*PvLEA6*-3'UTR) conferring tolerance in drought stress has been recently discovered (Battaglia et al., 2021). The 3' UTR enhances LEA6 protein synthesis levels

by supporting transcript loading into polysomes; aided by RBPs present in *P. vulgaris*. This recent finding is of relevance in widening our scope of understanding of the role of UTRs in translation as till date the 5' UTR is under the spotlight. The fact that proteins different to PABPs can also interact with the 3'UTR of some genes and mediate the translation process could now shift the focus to 3' UTRs as well and unravel the still elusive regulatory mechanisms (Battaglia et al., 2021).

PUFs are another well-characterized member that typically function to negatively modulate mRNA stability and translation. Moreover, functions of plant PUFs continue to expand owing to their heightened diversity in plants. *Arabidopsis* PUF mutants analyses have revealed involvement in development like stem cell self-renewal and maintenance (Group I APUMs), seed dormancy (Group II APUMs); and also in abiotic stress tolerance (APUM 5) (Joshna et al., 2020).

The functional roles of RBPs are ever-expanding and identifying their specific targets would be an exciting topic for future research.

Kinase Pathways

Numerous kinase pathways modulate translation via phosphorylation of core translation factors in response to environmental and developmental cues (Boex-Fontvieille et al., 2013; Choudhary et al., 2015; Dennis et al., 2009; Van Leene et al., 2019; Urquidi Camacho et al., 2020) (Table1). Recent advances indicate eIF4A phosphorylation by cyclin dependent protein kinase A (CDKA) during S, G2 and M phases represses its activity (Bush et al., 2016). Whereas, under dark and glucose starvation conditions, four paralogous MRF proteins (MA3-domain containing translation regulatory factors) phosphorylated by TOR-S6 Kinase pathway bind eIF4A and stimulate translation initiation (Lee et al., 2017). Also, the cap-binding complex offers sites for multiple phosphorylation events. Sucrose non-fermenting related kinase 1 (SnRK1) targets eIF4E and leads to translation suppression. This kinase also phosphorylates eIFiso4G on multiple sites during hypoxia stress (Bruns et al., 2019; Cho et al., 2019).

Table 1. Phosphorylation of Translation Initiation Factors (eIFs)

Initiation Factor	Kinase	Stimulus	Outcome	Organism	References
eIF2- α subunit	GCN2	Herbicide (chlorosulfuron), Stress and amino acid deprivation	Inhibition of ternary complex formation	*Arabidopsis thaliana, Oryza sativa,Nicotiana benthamiana*	Lageix et al., 2008; Zhang et al., 2008
eIF2- β subunit	TOR	Sucrose	Unknown	*Arabidopsis* cell culture	Van Leene et al., 2019
eIF2B-δ subnit	TOR	Sucrose	Unknown	*A. thaliana*, *Arabidopsis* cell culture	Boex-Fontvieille et al., 2013; Van Leene et al., 2019
eIF3-b subunit	Unknown	Darkness, High CO2	Unknown	*A. thaliana*	Boex-Fontvieille et al., 2013; Choudhary et al., 2015
eIF3-c subunit	CK2	–	Enhances interaction with eIF1,2,5	*in vitro*, *A. thaliana, Triticum spp.*	Boex-Fontvieille et al., 2013; Choudhary et al., 2015; Dennis et al., 2009
eIF3-h subunit	TOR-S6K1	Auxin	Translation reinitiation leading to increased translation of uORF-containing transcripts	*A. thaliana*	SchepetiInikov et al., 2013
eIF4A	CDKA	Darkness	Inhibition of eIF4A activity	*in vitro*	Boex-Fontvieille et al., 2013; Bush et al., 2016

	TOR-S6K1	Darkness and glucose starvation induced MRF accessory protein factors	Stimulate eIF4A activity	*in vitro*, *Nicotiana benthamiana*	Lee et al., 2017
eIF4B	CK2,TOR	Sucrose, High CO2	Unknown	*in vitro*, *Arabidopsis* cell culture	Boex-Fontvieille et al., 2013; Choudhary et al., 2015; Dennis et al., 2009; Van Leene et al., 2019
eIF4E	SnRK1	_	Translation repression	*in vitro*, *Nicotiana benthamiana*	Bruns et al., 2019
eIFiso4E	SnRK1		Translation repression	*in vitro*	Bruns et al., 2019
eIF4G	TOR	Darkness, Sucrose	Unknown	*Arabidopsis* cell culture	Boex-Fontvieille et al., 2013; Van Leene et al., 2019
eIFiso4G	SnRK1	Hypoxia stress	Stimulates core hypoxia gene translation (e.g. *ADH*), repress global translation	*Zea mays*	Cho et al., 2019
eIF5	CK2	_	Enhances interaction with eIF1,3c	*in vitro*, *Arabidopsis, Triticum spp.*	Boex-Fontvieille et al., 2013; Choudhary et al., 2015; Dennis et al., 2009
eIF6	TOR	Sucrose	Unknown	*Arabidopsis* cell culture	Van Leene et al., 2019

3. Ribosome in Regulation

In a cell, ribosome biogenesis is placed under tight regulation owing to the heavy energy investments that this process entails. The universal protein synthesis machinery is in fact a complex of ribonucleotides and proteins, ribosomal RNAs that impart catalytic prowess to the ribosome and ribosomal proteins (RPs) maintain the structural integrity of ribosome and also perform regulatory functions in translation process (Horiguchi et al., 2012; Browning and Bailey-Serres, 2015; Merchante et al., 2017). The plant ribosome comprises of four rRNAs (25-26S, 5.8S, and 5S in the 60S subunit and 18S in the 40S) accompanied by over 30 different RPs (81 in *Arabidopsis*) (Urquidi Camacho et al., 2020). Each *Arabidopsis* RP is encoded by two to seven paralogs and barring a few exceptions, all 81 of them have two paralogs at the least in active expression (Browning & Bailey-Serres, 2015).

Ribosomal Proteins

Like the core translation factors, RPs are also subjected to various post-translational modifications, particularly phosphorylation. Phosphorylation of several RPs and eIFs are stimulated by diurnal shifts in light and CO_2, like that of RPS6 by eS6K, RPS14 by uS11K, eIF4B. The eS6/RPS6 phosphorylation event has been studied quite extensively; phosphorylation of this RP on the 40S subunit via the TOR signalling pathway is involved in the selective translation of mRNAs involved in auxin, light, photo-assimilates, and photomorphogenetic signals and is repressed under heat stress and hypoxia (Enganti et al., 2018; Urquidi Camacho et al., 2020). However, the physiological significance of such phosphorylation dynamics remains to be resolved in most cases. Additionally, RP mutants in *Arabidopsis* share some abnormal developmental phenotypes such as reduced shoot and root growth and cell proliferation, increased ploidy in leaf cells, and a characteristic leaf defect called *pointed first leaf* or *denticulata* (Horiguchi et al., 2012). However, the vast range of observations on the effects of RP genes mutations, be it from embryo lethality to no-phenotype, some mutant phenotypes being completely unique to some paralogous pairs encoding the same RP exhibiting varied spatial and/or developmental expression domains propagates the possibility that the specific spatio-temporal expression patterns or the paralog composition of the ribosomes may have a potential role in translational control. The claims, however, await experimental validation (Horiguchi et al., 2012; Merchante et al., 2017).

4. Impact of RNA on Translational regulation

4.1. mRNA Structural Features

It is worth noting that not every mRNA is a substrate for translation; the transcripts that are acted upon by the translation apparatus contain certain intrinsic features that influences the translation process. These include covalent modifications and secondary structure elements at the two UTRs (5'cap and 3'poly(A) tail regions), and also the transcript's own sequence elements (uORFs). In general, transcripts with a shorter length and a low GC content are translated more efficiently than longer transcripts (Bai et al., 2017; Andrzejewska et al., 2020; Urquidi Camacho et al., 2020).

Covalent Modifications

The 5' cap and 3' poly(A) tail, both of these post-transcriptional additions are strong promoters of translation as they bring stability to the mRNA transcript and restrict its degradation (Andrzejewska et al., 2020). Other than capping and polyadenylation, evidences of methylation and 7-methylguanosine (m^7G) cap substitution has gathered interests in recent years (Urquidi Camacho et al., 2020).

Occurrence of a non-canonical NAD+ cap has been reported for about five percent of a gene's transcript from the *Arabidopsis* genome using a NAD sequence tagging method (NAD tagSeq) (Zhang et al., 2019; Urquidi Camacho et al., 2020). Overall, the NAD-capped mRNAs have almost the same nucleotide sequence as the canonical m^7G-capped mRNAs, suggesting that they could be also under active translation. Through gene ontology NAD-RNA producing genes were found to be linked to photosynthesis, cytokinin response and abiotic stress responses. Thus, it appears that the NAD-capped mRNAs may function as a signal to mediate cellular responses to intra-cellular energy status and environmental cues. However, the process by which NAD-mRNAs are processed and transported to cytosol for translation, is unknown presently (Zhang et al., 2019).

Methylation regulation at the N6 position of adenine N^6-methyladenosine (m^6A) is another notable example (Urquidi Camacho et al., 2020). The m^6A methylation mark has been identified across eukaryotic mRNAs including plant mRNAs. Although its presence and utility in gene regulation has been long known, its potential as a modulator of translation is a recent discovery. m^6A acts as a stabilizing/destabilizing

mark in the plant transcriptome depending upon the cell type and the developmental stage (Anderson et al., 2018; Reichel et al., 2019). Misregulation of the m^6A modification due to defects in the proteins involved in its metabolism result in flawed development of leaves and roots and altered timing of transition from vegetative to reproductive phase (Reichel et al., 2019; Shen et al., 2016). The m^6A sites are found predominantly near the stop codon and within the 3'UTR. Of all the different categories of proteins (writers, readers, erasers) partaking in this modification, YTH domain proteins, a family of 'reader proteins' play important roles in regulating translation. These proteins bind, interpret and bring about the downstream effects of m^6A methylation on various steps of RNA metabolism including translation (Reichel et al., 2019).

When located near stop codons or within 3'UTR, m^6A promotes canonical cap-dependent translation initiation via interaction of its reader protein YTHDF1 with eIF3 and other ribosome-associated protein factors. Whereas, when present in the 5'UTR, m^6A can direct non-canonical cap-independent initiation through internal recruitment of the 43S PIC by eIF3. Also, upon heat stress, relocation of ECT2 (of *Arabidopsis* ECT reader protein family) into stress granules indicate a role of m^6A in plant stress responses, similar to its role in heat shock mediated translational control in mammals identified previously (Reichel et al., 2019). Another study revealed its involvement in combating salt stress; m^6A boosts the stability of salt stress responsive transcripts by inhibiting internal (local) ribonucleolytic cleavage and hence promoting their translation (Anderson et al., 2018).

Secondary Structures

The secondary structures incorporated by an mRNA transcript can also strongly influence translation efficiency. Translationally relevant secondary structures like hairpin loops and pseudoknots located generally in the 5' leader sequence (UTR) impede translation of the main protein encoding ORF (mORF) in case the translating ribosome is unable to sort out and overcome the complex conformations (Merchante et al., 2017; Urquidi Camacho et al., 2020; Andrzejewska et al., 2020). Although instances of secondary structure mediated translation regulation are scarce in the plant

kingdom compared to prokaryotes and metazoans, few evidences of this mechanism occurring in plant systems have emerged over the years including that of Arabidopsis *PHYTOENE SYNTHASE (PSY)* gene and tobacco pollen gene *NPT303* (Merchante et al., 2017).

The translating ribosome can at times bypass the structural obstructions and resume translation downstream of an apparently large stretch of highly structured 5' UTR by a phenomenon known as ribosome shunting (Merchante et al., 2017; Urquidi Camacho et al., 2020). Instances of shunting are well-documented in plant RNA viruses such as CaMV (a pararetrovirus). The 40S ribosome is able to skip the structural motif by a non-linear scanning mechanism and the ribosome is able to assemble at a start codon downstream of the secondary structure (Urquidi Camacho et al., 2020). Ribosome shunting is quite uncommon outside viral transcripts and one such rare example of this phenomenon in plant cellular mRNAs comes from the green algae *Volvox carteri* wherein the repressive effects of stem loop containing uORFs in its regA mRNAs are overcome by shunting. This regA gene regulates (blocks) transcription of growth and reproduction related genes in somatic cells (Merchante et al., 2017).

Another interesting structure-based *cis* regulatory element is Internal ribosome entry site(s) (IRES) located within the 5'UTRs which can interact with various initiation factors and recruit the ribosome internally; thereby evading the need of canonical cap-dependent translation initiation. This can be particularly beneficial in continuation of the translation process under hazardous conditions. Majority cases of IRES comes from animal system and some from plant viruses; in endogenous plant mRNAs, reports of their occurrence is comparatively lower (Sonenberg and Hinnebusch, 2009; Merchante et al., 2017). However, some mRNAs involved in plant development and in stress-coping being translated via an IRES mechanism have been put forward. The best studied case of such a gene-specific translation regulation is that of the homeodomain transcription factor WUSCHEL (WUS) (Merchante et al., 2017). WUS protein maintains the stem-cell pool in the shoot apical meristem. In *Arabidopsis*, translation of WUS mRNAs is mediated by an IRES in its 5'UTR, which is further stimulated by binding of the AtLa1 protein (an RBP). Knockdown of AtLa1 gene expression leads to growth arrest due to translation repression. The La protein is nuclear-localized but under stress it is translocated to the cytoplasm where it enhances the IRES-dependent WUS mRNA translation (Cui et al., 2015). During heat stress too,

translation of heat shock proteins (HSPs) is most likely facilitated by IRES-like mechanisms. Upon heat stress, HSP101 translation remained uninterrupted in wheatgerm extracts absent in cap binding factors (eIF4E/eIFiso4E), implying a potential dependence of the transcripts on their 5'UTRs (Merchante et al., 2017). Plants possibly have devised such strategies to ensure translation of necessary genes like WUS, HSPs etc. is not switched off under unfavourable conditions which certainly pose a difficulty for cap-dependent translation to proceed (Cui et al., 2015; Merchante et al., 2017; Urquidi Camacho et al., 2020).

Upstream open reading frames

uORFs are sequence-based *cis*-regulatory elements present in the 5'UTRs of the same target mRNAs. uORF repress the translation initiation of the main ORF (mORF) which is downstream of the uORF by diverting the scanning ribosome from translating the mORF or by rendering the transcript unfit for translation by triggering degradation via non-sense mediated decay (stop codon of the uORF functions as a premature termination codon that signals NMD). Stalling of ribosome and/or failure to reinitiate translation at the mORF due to ribosome exit (dissociation) at an uORF stop codon, consequently lowers the translation efficiency of mRNAs harbouring uORFs in comparison to those without any uORFs (Merchante et al., 2017; Zhang et al., 2020; Urquidi Camacho et al., 2020). These short translationally active sequences have an almost universal presence in eukaryotic mRNAs; approximately 50% of mouse and human protein-encoding genes, and 24-30% of plants genes contain one or more functional uORFs (Zhang et al., 2020).

The extent of their negative impact on expression of the mORF depends upon the sequence context around uORF initiation codon (uAUG), uORF's own amino acid sequence, location with respect to the next ORF (the succeeding uORF/the mORF; whether overlapping/non-overlapping) and the number of uORFs present in the mRNA. However, the uORF length has not been shown to have any major effect on its inhibitory function; only except for uORFs longer than 16 amino acids. The further uORFs are from the 5'-cap or the closer the stop codons of uORFs are from the mORF start codon (mAUG), the more repressive the uORFs tend to be. Moreover, the uORFs that overlap with the mORF, i.e., the stop codons of which are downstream (either out-of-frame or in-frame) of the mAUG exert the strongest inhibitory effect. Majority of

identified uORFs belong to the non-overlapping category and have their stop codons separate (upstream) from the mAUG. When the uAUG is surrounded by a weak translation initiation context, i.e., an unfavourable Kozak sequence, it will be skipped by the scanning ribosome (leaky scanning) and the mORF will get translated. Similarly, uAUGs with a preferable sequence context will result in higher translatability of those uORFs, and subsequently, reduced/no mORF translation (Merchante et al., 2017; Zhang et al., 2020). Other than uAUG, near-cognate start codons like UUG, GUG, CUG, ACG can also act as uORF initiator, although less effective than the canonical AUG (Zhang et al., 2020), as identified by genome-wide ribosome profiling studies in yeast (Spealman et al., 2018), tomato (Li and Liu, 2020). Upstream ORFs (uORFs) of the bZIP transcription factors, GDP-L-galactose phosphorylase and adenosyl methionine decarboxylase, which are involved in the sugar/ polyamine/ ascorbate-mediated (vitaminC) translational suppression of the mORFs, have been identified to be translated with alternative AUG and ACG initiators. In addition, presence of a higher number of uORFs in a transcript is correlated with a more stronger translation inhibition of its mORF, as uORFs tend to work in a cumulative manner (Zhang et al., 2020).

Depending on the mode of regulation, uORFs are categorised as: Sequence-independent uORFs (SIuORFs) and Sequence-dependent uORFs, also known as conserved peptide uORFs (CPuORFs). In SIuORFs, their orientation on the 5'UTR affects their functionality. They maybe evolutionary conserved in their positions but not their sequence (Merchante et al., 2017). At present, most of identified uORFs are of this type and among them the minimum ORFs (initiation codon directly followed by a termination codon), designated as AUG-Stop uORFs are of significant importance from translational control viewpoint, including boron and phosphorus homeostasis maintenance (Zhang et al., 2020). Under high boron environment, two minimum ORFs (AUGUAA) present in the boron uptake channel AtNIP5;1 (nodulin 26-like intrinsic protein 5;1) cause ribosome stalling and subsequent translation repression of the mORF followed by mRNA degradation, thus avoiding B toxicity in plants. In boron-deficit, translation of the AtNIP5;1 mORF occurs through reinitiation due to scanning resumption by some of the 40S ribosomes that remain bound to the transcript post ribosome dissociation at the stop codon of minimum ORF. The 80S ribosome is able to reassemble and reinitiate translation at the mAUG. Thus, it could be suggested that high boron levels negatively impact reinitiation, although the underlying molecular mechanism remains to be elucidated (Tanaka et al., 2016). Moreover, boron uptake

channel BOR1 of *Arabidopsis* is also under B-dependent translation regulation by set of three uORFs with one of them an AUG-Stop uORF (Aibara et al., 2018; Tsednee et al., 2020). A similar B-dependent AUG-Stop mediated translational control has been detected in yeast recently and previously in animals, thus revealing the widely conserved nature of this unique regulation among eukaryotes (Tsednee et al., 2020).

The second and less abundant category of uORFs, CPuORFs brings about translation inhibition by virtue of their encoded peptide sequence (Merchante et al., 2017). The regulatory peptide synthesized when the uORF is functional, blocks the translocating ribosome and does not supports reinitiation. The CPuORFs contain certain rare codons that cause ribosomal arrest at the uORF stop codon and are an integral part of metabolite-mediated translational control with the cellular concentration of a particular metabolite controlling inhibitory activity of the uORF. Metabolites acting as effector molecules include polyamines (on the AdoMetDC mRNA), phosphocholine (XIPTOL1 mRNA), sucrose (bZIP11 mRNA), ascorbate (GGP mRNA) (Merchante et al., 2017; Urquidi Camacho et al., 2020; Zhang et al., 2020). Generally, the presence of the metabolite in its sufficient physiological levels activates the uORFs which then switches off the mORF translation (negative feedback) thereby limiting biosynthesis of the metabolite.

In summary upstream ORFs constitute an important section of translation regulators in plants encompassing several key aspects of a plant's life including metabolic pathways, nutrient uptake, morphogenesis, photomorphogenesis, growth, hormone signaling and defence (Urquidi Camacho et al., 2020; Dong et al., 2020; Kurihara et al., 2020; Zhang et al., 2020).

Table 2. Regulatory roles of uORFs

Controlled Process	Gene	uORF Type	Regulatory Role	Stimulus/ Signal	Organism	Reference
Plant Morpho-genesis	HOMEOBOX1 (HB1) AtHB1 ZmHB115	CPuORF (CPuORF33)	Maintenance of aerial tissue morphogenesis by translational repression of mORF and avoid aberrant phenotypes	Light	*A. thaliana* *Zea mays*	(Ribone et al., 2017)
Photomorph o-genesis	PHYTOCROME INTERACTING FACTOR (PIF3) AtPIF3	CPuORF	Phytochrome B-dependent alternative spliced variant contains a uORF which inhibits AtPIF3 mORF translation and promote seedling photomorphogenesis upon light exposure.	Light, Photo-receptor phyB	*A. thaliana*	(Dong et al., 2020; Kurihara et al., 2020)
Hormone Signalling and associated phenotypes Auxin	AUXIN RESPONSE FACTORS (At ARFs)	CPuORF	Translation of uORF(s)-containing ARFs is stimulated in presence of auxin.	TOR	*A. thaliana*	(Schepetilnikov et al., 2013; Rosado et al., 2012)
Brassinoster oid (BR)	BRASSINOSTER OID INSENSITIVE 1 (BRI1) AtBR1	CPuORF	Maintenance of steady-state levels of BR in vivo; uORF mutants show elongated hypocotyl length	-	*A. thaliana*	(Zhang et al., 2018)

Nutrient Absorption Phosphorus (Pi)	NITROGEN LIMITATION ADAPTATION (NLA) OsNLA1	-	Regulates Pi acquisition in roots; high Pi uptake leads to impaired pollen development and reduced grain production	Pi	*Oryza sativa* ssp. Japonica	(Yang et al., 2020)
Biotic stress	TL1 BINDING TRANSCRIPTION FACTOR (AtTBF1)	CPuORF (Phenyl-alanine-enriched)	Rapid reprogramming of gene expression under biotic stress (growth-to-defense transition)	Pathogen attack, Phenylalanine level in vivo	*A. thaliana*	(Pajerowska-Mukhtar et al., 2012)

4.2. Transfer RNAs

Apart from their primary role as biochemical reagents in protein biosynthesis, transfer RNAs appear to have multiple novel functions across different kingdoms. As per recent findings, their biogenesis and modifications are suggested to have a regulatory significance (Urquidi Camacho et al., 2020). In this regard, the production of small non-coding RNAs (ncRNAs) from degradation of tRNAs known as tRNA-derived fragments (tRFs) are of special interest. These fragments originate from either precursor or mature (most common) tRNA molecules via Dicer-dependent and RNase A/T2-dependent (dicer-independent) pathways. Mature tRNAs give rise to two main types of tRNAs based on the cleavage site, long tRFs (30-35 nt) from the anticodon loop, tRF-5A or tRF-3A and short tRFs (19-26 nt) generated from cleavage in the D (tRF-5D) or T loop (tRF-3T); named on the basis of which terminus whether 5' or 3' these RNA fragments are obtained from (Park and Kim, 2018; Urquidi Camacho et al., 2020; Lalande et al., 2020).

In general, cells subjected to stress conditions such as drought, salt, oxidative stresses and nutrient starvation (phosphorus) show elevated tRF levels, indicating a functional role in stress tolerance (Park and Kim, 2018). Their presence has been linked with suppression of the gene expression through post-transcriptional gene silencing (by interactions with AGOs) and translation inhibition (by interactions with ribosomes and eIFs) in diverse species from Archaebacteria, protozoa, yeast to human cells. In plant cells too tRFs have similar implications in repressing gene translation and combating

abiotic stress (Cognat et al., 2017; Park and Kim, 2018). Recently, their inhibitory function in plants has been validated by *in vitro* translation of GUS reporter gene using oligoribonucleotides mimicking the *Arabidopsis in vivo* tRF species (Lalande et al., 2020). Only a subset of the diverse tRF population was found to efficiently affect translation *in vitro*, besides tRF-5D (Ala-AGC) and tRF-5D (Asn-GUU) other synthetic tRFs could not cause strong inhibition of GFP expression. This observation suggests that the ability of tRFs to modulate protein synthesis is not correlated with their abundance in total tRF population; and the other subtypes associated with Argonaute immunoprecipitates (AGO1, AGO4) might have implications in the RNA silencing pathway (Park and Kim, 2018; Lalande et al., 2020). The tRFs are thus said to have two different modes of action in regulating gene expression. Further, the inhibition is both cap- and sequence-independent and rather based on the ability of the tRFs to bind with the actively elongating polyribosomes. This brings about an early arrest of translation by interfering with the binding of ribosomes to the transcript, as detected by enrichment of polyribosomal fraction with tRFs upon puromycin treatment (an antibiotic causing premature chain termination), however, the underlying mechanism of this binding process is poorly understood at present (Lalande et al., 2020). Further studies are needed to uncover the specific role of tRF subtypes in developmental transitions that can possibly incorporate an epigenetic regulatory layer; as supported by tRF level fluctuations in undifferentiated and differentiating rice calli (Cognat et al., 2017; Park and Kim, 2018).

4.3. Small RNAs

Small RNAs are important regulatory components of key biological processes including developmental and physiological events like abiotic/biotic stress responses in both plants and animals (Merchante et al., 2017). Small RNAs typically participate in various pathways to bring down the target gene expression. Among them, micro RNAs (miRNAs) function as guide RNA delivering the Argonaute-based RNA-Induced silencing complex (RISC) to the target mRNAs and induce their cleavage (Merchante et al., 2017; Urquidi Camacho et al., 2020). This cleavage can occur co-translationally, mediated by exoribonuclease XRN4 (in 5'-3' direction following the translocating ribosomes) or in special designated sites like the cytoplasm-localized ribonucleoprotein (RNP) complexes called P (processing) bodies and endoplasmic reticulum-bound polysomes (MBPs) (Chantarachot and Bailey-Serres, 2018; Afrin et

al., 2020). Interestingly, miRNAs and siRNAs (rarely) can alone also cause translational inhibition as demonstrated *in vitro*. How the decision is made as to whether the miRNA silences its target mRNA by translation repression or by degrading it, remains largely unknown, although some studies suggest it could depend on the cell type (Urquidi Camacho et al., 2020).

Micro RNAs have wide-ranging implications in numerous aspects of plant growth and development such as the miR397 family which mainly targets *LACCASE* genes (LAC) that are involved in lignin biosynthesis and plant development in conditions such as floral, seed and fruit organogenesis. By targeting a sister chromatid mucin during cell division miR397 also functions in somatic embryogenesis. Recent studies have suggested a role of miR397b in flowering regulation in *Arabidopsis* by targeting and cleaving *Casein Kinase II Subunit Beta3* (*CKB3 mRNAs*). Host of other genes linked to plant metabolism, growth and signal transduction may be translationally repressed by miR397 as revealed by bioinformatic predictions (Huang et al., 2020). Instances of other miRNAs functioning at translational level gene repression include miR172 (*APETTALA2 AP2* mRNA), miR156 (*SQUAMOSA BINDING PROTEIN-LIKE3, SPL3)* in flowering; miR398 (Cu/Zn *SUPEROXIDE DISMUTASE*, *CSD)* under copper deficiency; miR408 (*PLANTACYANIN)* in copper metabolism; and miR863 (*SERRATE)* and miR168 (*ARGONAUTE1)* in miRNA metabolism (Merchante et al., 2017).

5. Implications of Translational Regulation on Plant Growth and Development

5.1. Seed Germination

Seeds represent beginning of a new chapter in a plant's life. The germination episode is marked by transition from a quiescent state (dry seed) to a metabolically active state (young seedling) (Bai et al., 2020). It has been seen that the total RNA content does not differ much between dormant and non-dormant seeds, hinting at an additional translational level control of the germination process (Merchante et al., 2017). The dynamics of translational regulation during seed germination in *Arabidopsis* were explored by Bai et al. (2016, 2020) which revealed an elaborate stage-type specific control. The polysome profiles of total and polysome-associated mRNA over five consecutive stages of germination (0-72 hr after imbibition, HAI) were analyzed and

thousands of individual mRNAs whose polysome occupancy changed over the entire time course of germination were identified. Polysome occupancy is defined as the ratio between polysome-bound and total mRNA and can be taken as a measure of translation efficiency (Bai et al., 2017). However, this changeover in polysome occupancy was not uniform in each stage; with maximum mRNAs experiencing a change in their ribosome content (translational shift) coinciding with two distinct temporal phases, hydration (Stage I) (between 0-6 HAI) and root tip protrusion (following testa and endosperm rupture), i.e., germination (Stage III) (26-48 HAI). In contrast, Stage II (6-26 HAI) involving seed coat rupture and Stage IV (48-72 HAI) marked by seedling greening witnessed minor translational changes (Fig.2a) (Bai et al., 2017; Merchante et al., 2017; Urquidi Camacho et al., 2020).

The set of genes identified (upregulated/downregulated) in the two major physiological checkpoints (hydration HTS, germination GTS) were different and the mRNA transcript features that were found in relation to the translational shifts also differed significantly, suggesting that both these stages of regulation is being controlled by supplementary and independent mechanisms (Bai et al., 2017). For HTS, wounding response genes, chaperone genes etc. were upregulated and with respect to transcript features, a reduction in length, GC content, secondary structures, no. of uORFs and effective codon number favoured translation whereas, for GTS the CDS length, GC content and structural features affected translation and dominant gene groups identified include those involved in seed development, protein storage, hormonal response to auxin, ABA, lipid and sucrose metabolism, *LEA* genes, HSPs, cell wall and organelle related genes. Further, HTS upregulated transcripts also showed presence of adenosine enriched tract in the 5'UTR which may represent putative binding sites for translation initiation factors or some RBPs and result in translational enhancement (Bai et al., 2017; Urquidi Camacho et al., 2020).

A study was undertaken by the same group (Bai et al., 2020) focusing on translational aspects of dry seeds to gain some insights as to how seeds in a quiescent state remain translationally viable over several years. The seed-stored mRNAs were identified in association with monosome as compared to the non-dormant seeds where majority of the mRNAs are polysome-associated. The study suggested a potential role of RBPs in protection of the seed-stored RNAs and presents opportunities for future research.

5.2. Pollen Translatome

In angiosperms, the male gametogenesis takes place through coordinated activities of both gametophytic and sporophytic tissues and involves widespread dynamic changes in gene expression (Hafidh et al., 2016). An exceeding number of processes regulates pollen tube growth, guidance competence and reception by the target ovule. The diploid, sporophytic (pre-meiotic) phase is marked by an abundance in RBPs, molecular chaperones such as HSPs, cellulases and protein degrading enzymes. In comparison, the proteome of later haploid, gametophytic (post-meiotic) phase is abundant in late embryogenesis abundant (LEA) proteins, enzymes involved in glycolysis and fermentation processes (Fig. 2b) (Ischebeck et al., 2014; Hafidh et al., 2018).

The mature pollen grains (desiccated and shed) are also known to be enriched with an unusually rare and previously uncharacterized type of translationally-inactive monosomes called EDTA/puromycin-resistant particles (EPPs). EPPs contain singular ribosomes (monosomes) on mRNA transcripts along with initiation factors, elongation factor eEF1-A and ribosomal proteins (60S RPs). The stored transcripts are awaiting molecular signals to break free from their repressed state and reinstate their translation at the preferred time i.e., maturation phase of pollen germination following re-hydration (pollen tube growth from 4h-24h). The presence of these specialized structures has been reported in mature pollen for pollen-specific cell wall glycoprotein and pollen-specific LIM domain containing protein from tobacco (Hafidh et al., 2018; Urquidi Camacho et al., 2020). The mature pollen like seeds are in a metabolically quiescent state and upon re-hydration become activated followed by pollen tube emergence and growth. Pollen activation is accompanied by protein phosphorylation changes as identified by phosphoproteomics profiling done on three stages of tobacco pollen (mature, 5-min and 30-min activated pollen). Translation initiation was found to be modulated by protein phosphorylation; eIFs like eIF2, eIF3, eIF4A, eIF4B, eIF4G and eIF5 become dephosphorylated very early in pollen germination as a prelude to onset of translation of the EPP-stored mRNAs (Fíla et al., 2016).

5.3. Translation and Circadian Clock

One of the most perceivable environment perturbations experienced by plants is the day-to-day alterations between light and darkness (Pal et al., 2013; Merchante et al., 2017; Urquidi Camacho et al., 2020). Plants need to coordinate their metabolism and

growth with respect to the diurnal changes in carbon availability or else there will be severe carbon shortage at night-time and growth impairment. Protein synthesis is a major determinant of cellular growth and expends a considerable portion of cellular energy levels. It is therefore plausible that translation is put under a circadian control and diel regulation which enables the plants to respond to the dynamic environment in an energy-efficient manner and also cater to the metabolic status of photo-assimilates (Pal et al., 2013; Missra et al., 2015). Indeed, translation levels are found to be higher during the day compared to night based on the global shifts in ribosome loading into polysomes, (i.e., ribosomes in active translation) estimated between 60%-70% during day, 40%-45% during night and 20% following starch exhaustion during extended night period (Fig. 2c). No significant change has been observed in the ribosome abundance indicating protein translation is regulated by changes in the number of ribosomes in active translation pool over the diurnal cycle rather than their actual population. The global regulation of translation is also positively correlated with the cellular sucrose content as revealed by differences in polysome loading in wild type *Arabidopsis* rosette leaves and in starchless mutant *phosphoglutamase* (*pgm*) wherein the polysome loading decreased to less than 25% during night (Pal et al., 2013). In general, translation is supported by presence of carbohydrates and is true for non-vascular plants also, as it was found in a recent study on *P. patens*, wherein translation was upregulated globally in presence of glucose (Chamorro-Flores et al., 2020). While cytosolic polysome loading showed a correlation with sucrose content, plastidic polysome loading had a strong dependence on light and mitochondrial polysome loading remained high at all times, unaffected by periods of light and darkness. This suggests independent mechanisms are underway in regulation of cytosolic and organellar translation (Pal et al., 2013). mRNAs encoding for ribosomal proteins (RP) are preferably upregulated during night-time. Indeed, ribosome biogenesis constitutes a major portion of a cell's transcriptional and translational machinery and it's only natural to carry out the rRNAs and ribosomal proteins production during a time when demand for other energetically exhaustive processes such as photosynthesis, nitrogen fixation etc. are down. Certain post-translational modifications can also regulate the diurnal changes in translation like the differential phosphorylation of RPs between day and night (Missra et al., 2015; Merchante et al., 2017).

Circadian regulation of translation has been established by study of clock genes in *Arabidopsis.* Overexpression of the morning gene *CIRCADIAN CLOCK ASSOCIATED*

(CCA1) which suppress evening genes leads to disruption of the normal clock function. Such plants also exhibited differential patterns of polysome loading (6hr delay from noon to evening or from midnight to dawn) and subsequent phase shifts in peak translation state in contrast to the wild type seedlings. Another clock gene *LATE ELONGATED HYPOCOTYL (LHY),* counterpart of CCA1 is also translationally modulated at the initiation stage (Missra et al., 2015).

5.4. Root Nodule Symbiosis

Establishment of a nitrogen-fixing symbiotic relationship with the soil bacterium *Rhizobium* is a widespread feature of legume plants that helps them in overcoming nitrogen deficiency in soil. The atmospheric N_2-fixation takes place in specialized structures called root nodules which harbour the *Rhizobia*. Symbiosis involves activation of two gene sets- infection and nodule organogenesis (nodulation) which bring about extensive reprogramming of the root architecture and involves regulation of gene expression at multiple levels including cytoplasmic mRNA translation (Zanetti et al., 2020). The mRNA translational status during root nodule symbiosis as obtained by Translating Ribosome Affinity Purification followed by RNA-seq (TRAP-seq) analyses revealed differences in abundance of transcript subsets in cell-specific translatomes (root epidermal, cortical translatomes) in response to rhizobial infection (Fig. 3d). While the epidermal translatome showed accumulation of transcripts encoding for proteins with key functions in nodulation such as the transcription factors Ethylene Response Factor for Nodulation (ERN1), Nuclear Factor Y A-1 (NF-YA1), and the cell wall remodelling protein Nodule Pectate Lyase (NPL) at high levels in the polysomes; on the other hand the cortical translatome was specifically enriched in mRNAs encoding for the E3 ubiquitin ligase Seven in absentia (SINA) upon rhizobia inoculation. This spatial differences in transcript abundance in translatomes (selective mRNA translation) are not necessarily correlated with their overall cellular abundance, suggesting a translational level control of gene expression during establishment of the symbiotic association (Traubenik et al., 2020; Zanetti et al., 2020).

In addition to selective mRNA translation mechanisms, small RNAs (sRNAs) and long non-coding RNA (lncRNAs) contribute to modulate gene expression in response to exogenous and endogenous signals during symbiosis (Fig. 3d). Micro RNAs (miRNAs) function in auxin perception and response (e.g., miR160, miR167 targets *Auxin Response Factors*) during rhizobial infection and nodulation. Presence of mobile

miRNAs (miR2111 targets root-specific transcription factor Too Much Love, TML, in *Lotus japonicus*) that transport shoot originated signals to roots and limit further nodulation (autoregulation of nodulation mechanism) have also been discovered in recent years. The post-transcriptional functions of lncRNAs (>200nt) include protein and/or miRNA sequestration, control of protein sub-cellular localizations (RBPs export from nucleus to cytoplasm during nodulation, Early Nodulin 40 *ENOD40* lncRNA), association with polysomes resulting in either activation or repression of the specific transcripts and also act as precursors of trans-acting RNAs (tasiRNAs) which negatively impact nodulation (*TAS3* lncRNA and its isoform *ALT TAS3*). Most of our understandings regarding lncRNA roles in symbiosis has come from studies in *Medicago truncatula* and their involvement in regulating mRNA stability and translatabilty is an emerging field of research (Zanetti et al., 2020).

Another interesting aspect of root nodule symbiosis is the generation of *Rhizobia* bacteria transfer RNA derived sncRNAs, tRFs (Fig.3d). These bacterial origin tRFs then repress translation of their complementary host mRNAs, like in soyabean (*Glycine max*) wherein the tRFs block translation of mRNAs that negatively regulate rhizobial infection and nodule organogenesis. This presents a wonderful example of cross-kingdom RNA interference as well as an exciting research perspective involving genetic engineering of the tRFs to enhance this agronomically relevant N_2-fixing process (Zanetti et al., 2020).

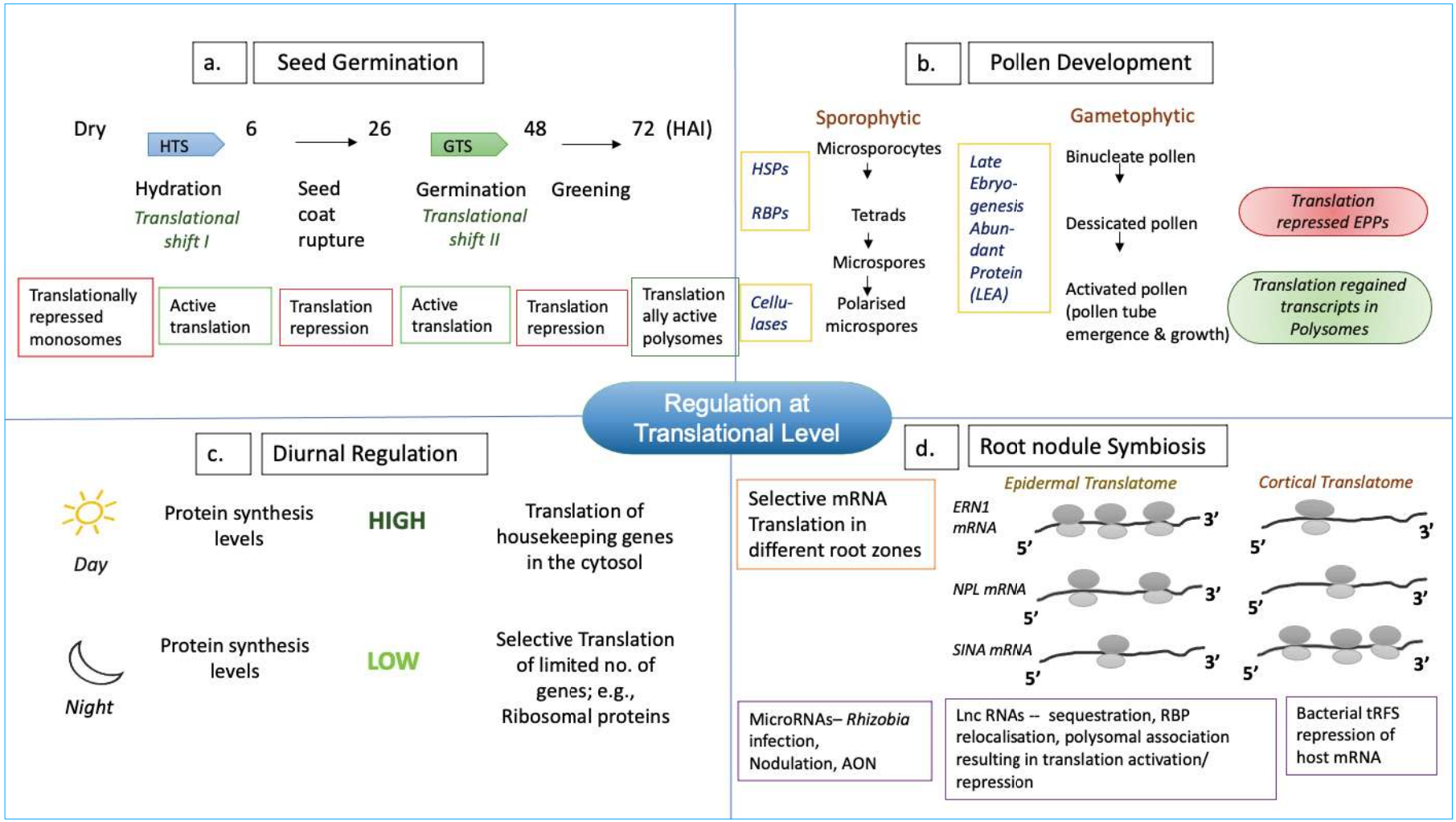


Fig. 3 Some notable instances of translational regulation in plant growth and development.
HAI, Hours after imbibition; HSPs, heat shock proteins; RBPs, RNA binding proteins; EPPs, EDTA/Puromycin-resistant particles; ERN1, Ethylene Response Factor for Nodulation; NPL,Nodule Pectate Lyase; SINA, Seven in absentia; Lnc RNA, Long non-coding RNA; tRF, tRNA-derived fragments; AON, Autoregulation of nodulation.

References

References

Afrin, T., Diwan, D., Sahawneh, K., and Pajerowska-Mukhtar, K. (2020). Multilevel regulation of endoplasmic reticulum stress responses in plants: where old roads and new paths meet. J. Exp. Bot. **71**: 1659–1667.

Aibara, I., Hirai, T., Kasai, K., Takano, J., Onouchi, H., Naito, S., Fujiwara, T., and Miwa, K. (2018). Boron-dependent translational suppression of the borate exporter BOR1 contributes to the avoidance of boron toxicity. Plant Physiol. **177**: 759–774.

Anderson, S.J., Kramer, M.C., Gosai, S.J., Yu, X., Vandivier, L.E., Nelson, A.D.L., Anderson, Z.D., Beilstein, M.A., Fray, R.G., Lyons, E., and Gregory, B.D. (2018). N6-Methyladenosine Inhibits Local Ribonucleolytic Cleavage to Stabilize mRNAs in Arabidopsis. Cell Rep. **25**: 1146-1157.e3.

Andrzejewska, A., Zawadzka, M., and Pachulska-Wieczorek, K. (2020). On the Way to Understanding the Interplay between the RNA Structure and Functions in Cells: A Genome-Wide Perspective. Int. J. Mol. Sci. **21**: 6770.

Bai, B., Van Der Horst, S., Cordewener, J.H.G., America, T.A.H.P., Hanson, J., and Bentsink, L. (2020). Seed-stored mRNAs that are specifically associated to monosomes are translationally regulated during germination. Plant Physiol. **182**: 378–392.

Bai, B., Peviani, A., van der Horst, S., Gamm, M., Snel, B., Bentsink, L., and Hanson, J. (2017). Extensive translational regulation during seed germination revealed by polysomal profiling. New Phytol. **214**: 233–244.

Battaglia, M.E., Martínez-Silva, A.V., Olvera-Carrillo, Y., Dinkova, T.D., and Covarrubias, A.A. (2021). Translational enhancement conferred by the 3' untranslated region of a transcript encoding a group 6 late embryogenesis abundant protein. Environ. Exp. Bot. **182**.

Boex-Fontvieille, E., Daventure, M., Jossier, M., Zivy, M., Hodges, M., and Tcherkez, G. (2013). Photosynthetic Control of Arabidopsis Leaf Cytoplasmic Translation Initiation by Protein Phosphorylation. PLoS One **8**.

Browning, K.S. and Bailey-Serres, J. (2015). Mechanism of Cytoplasmic mRNA Translation. Arab. B. **13**: e0176.

Bruns, A.N., Li, S., Mohannath, G., and Bisaro, D.M. (2019). Phosphorylation of Arabidopsis eIF4E and eIFiso4E by SnRK1 inhibits translation. FEBS J. **286**: 3778–3796.

Bush, M.S., Pierrat, O., Nibau, C., Mikitova, V., Zheng, T., Corke, F.M.K., Vlachonasios, K., Mayberry, L.K., Browning, K.S., and Doonan, J.H. (2016). EIF4A RNA helicase associates with cyclin-dependent protein kinase A in proliferating cells and is modulated by phosphorylation. Plant Physiol. **172**: 128–140.

Chamorro-Flores, A., Tiessen-Favier, A., Gregorio-Jorge, J., Villalobos-López, M.A., Guevara-García, Á.A., López-Meyer, M., and Arroyo-Becerra, A. (2020). High levels of glucose alter Physcomitrella patens metabolism and trigger a differential proteomic response. PLoS One **15**: e0242919.

Chantarachot, T. and Bailey-Serres, J. (2018). Polysomes, stress granules, and processing bodies: A dynamic triumvirate controlling cytoplasmic mRNA fate and function. Plant Physiol. **176**: 254–269.

Cho, H.Y., Lu, M.Y.J., and Shih, M.C. (2019). The SnRK1-eIFiso4G1 signaling relay regulates the translation of specific mRNAs in Arabidopsis under submergence. New Phytol. **222**: 366–381.

Choudhary, M.K., Nomura, Y., Wang, L., Nakagami, H., and Somers, D.E. (2015). Quantitative circadian phosphoproteomic analysis of Arabidopsis reveals extensive clock control of key components in physiological, metabolic, and signaling pathways. Mol. Cell. Proteomics **14**: 2243–2260.

Cognat, V., Morelle, G., Megel, C., Lalande, S., Molinier, J., Vincent, T., Small, I., Duchene, A.M., and Marechal-Drouard, L. (2017). The nuclear and organellar tRNA-derived RNA fragment population in Arabidopsis thaliana is highly dynamic. Nucleic Acids Res. **45**: 3460–3472.

Cui, Y., Rao, S., Chang, B., Wang, X., Zhang, K., Hou, X., Zhu, X., Wu, H., Tian, Z., Zhao, Z., Yang, C., and Huang, T. (2015). AtLa1 protein initiates IRES-dependent translation of WUSCHEL mRNA and regulates the stem cell homeostasis of Arabidopsis in response to environmental hazards. Plant, Cell Environ. **38**: 2098–2114.

Dennis, M.D., Person, M.D., and Browning, K.S. (2009). Phosphorylation of plant translation initiation factors by CK2 enhances the *in vitro* interaction of multifactor complex components. J. Biol. Chem. **284**: 20615–20628.

Dong, J., Chen, H., Deng, X.W., Irish, V.F., and Wei, N. (2020). Phytochrome B induces intron retention and translational inhibition of phytochrome-interacting factor3. Plant Physiol. **182**: 159–166.

Fíla, J., Radau, S., Matros, A., Hartmann, A., Scholz, U., Fecikova, J., Mock, H.P., Èapková, V., Zahedi, R.P., and Honys, D. (2016). Phosphoproteomics profiling of tobacco mature pollen and pollen activated *in vitro*. Mol. Cell. Proteomics **15**: 1338–1350.

Gupta, P., Rangan, L., Ramesh, T.V., and Gupta, M. (2016). Comparative analysis of contextual bias around the translation initiation sites in plant genomes. J. Theor. Biol. **404**: 303–311.

Hafidh, S., Fíla, J., and Honys, D. (2016). Male gametophyte development and function in angiosperms: a general concept. Plant Reprod. **29**: 31–51.

Hafidh, S., Potěšil, D., Müller, K., Fíla, J., Michailidis, C., Herrmannová, A., Feciková, J., Ischebeck, T., Valášek, L.S., Zdráhal, Z., and Honys, D. (2018). Dynamics of the pollen sequestrome defined by subcellular coupled omics. Plant Physiol. **178**: 258–282.

Horiguchi, G., Van Lijsebettens, M., Candela, H., Micol, J.L., and Tsukaya, H. (2012). Ribosomes and translation in plant developmental control. Plant Sci. **191–192**: 24–34.

Huang, S., Zhou, J., Gao, L., and Tang, Y. (2020). Plant miR397 and its functions. Funct. Plant Biol.: 361–370.

Ischebeck, T., Valledor, L., Lyon, D., Gingl, S., Nagler, M., Meijón, M., Egelhofer, V., and Weckwerth, W. (2014). Comprehensive cell-specific protein analysis in early and late pollen development from diploid microsporocytes to pollen tube growth. Mol. Cell. Proteomics **13**: 295–310.

Joshna, C.R., Saha, P., Atugala, D., Chua, G., and Muench, D.G. (2020). Plant PUF RNA-binding proteins: A wealth of diversity for post-transcriptional gene regulation. Plant Sci. **297**: 110505.

Kurihara, Y., Makita, Y., Shimohira, H., Fujita, T., Iwasaki, S., and Matsui, M. (2020). Translational landscape of protein-coding and non-protein-coding RNAs upon light exposure in Arabidopsis. Plant Cell Physiol. **61**: 536–545.

Lageix, S., Lanet, E., Pouch-Pélissier, M.N., Espagnol, M.C., Robaglia, C., Deragon, J.M., and Pélissier, T. (2008). Arabidopsis eIF2α kinase GCN2 is essential for growth in stress conditions and is activated by wounding. BMC Plant Biol. **8**: 1–9.

Lalande, S., Merret, R., Salinas-Giegé, T., and Drouard, L. (2020). Arabidopsis tRNA-derived fragments as potential modulators of translation. RNA Biol. **17**: 1137–1148.

Lee, D.H., Park, S.J., Ahn, C.S., and Pai, H.S. (2017). MRF family genes are involved in translation control, especially under energy-deficient conditions, and their expression and functions are modulated by the TOR signaling pathway. Plant Cell **29**: 2895–2920.

Van Leene, J. et al. (2019). Capturing the phosphorylation and protein interaction landscape of the plant TOR kinase. Nat. Plants **5**: 316–327.

Li, Y.-R. and Liu, M.-J. (2020). Prevalence of alternative AUG and non-AUG translation initiators and their regulatory effects across plants. Genome Res. **30**: 1418–1433.

Mahalingam, R. and Walling, J.G. (2020). Genomic survey of RNA recognition motif (RRM) containing RNA binding proteins from barley (Hordeum vulgare ssp. vulgare). Genomics **112**.

Merchante, C., Stepanova, A.N., and Alonso, J.M. (2017). Translation regulation in plants: an interesting past, an exciting present and a promising future. Plant J. **90**: 628–653.

Missra, A., Ernest, B., Lohoff, T., Jia, Q., Satterlee, J., Ke, K., and von Arnim, A.G. (2015). The circadian clock modulates global daily cycles of mrna ribosome loading. Plant Cell **27**: 2582–2599.

Pajerowska-Mukhtar, K.M., Wang, W., Tada, Y., Oka, N., Tucker, C.L., Fonseca, J.P., and Dong, X. (2012). The HSF-like transcription factor TBF1 is a major molecular switch for plant growth-to-defense transition. Curr. Biol. **22**: 103–112.

Pal, S.K. et al. (2013). Diurnal changes of polysome loading track sucrose content in the rosette of wild-type arabidopsis and the starchless pgm mutant. Plant Physiol. **162**: 1246–1265.

Park, E.J. and Kim, T.H. (2018). Fine-tuning of gene expression by trna-derived fragments during abiotic stress signal transduction. Int. J. Mol. Sci. **19**.

Raabe, K., Honys, D., and Michailidis, C. (2019). The role of eukaryotic initiation factor 3 in plant translation regulation. Plant Physiol. Biochem. **145**: 75–83.

Raxwal, V.K., Simpson, C.G., Gloggnitzer, J., Entinze, J.C., Guo, W., Zhang, R., Brown, J.W.S., and Riha, K. (2020). Nonsense-Mediated RNA Decay Factor UPF1 Is Critical for Posttranscriptional and Translational Gene Regulation in Arabidopsis. Plant Cell **32**: 2725–2741.

Reichel, M., Köster, T., and Staiger, D. (2019). Marking RNA: M6A writers, readers, and functions in Arabidopsis. J. Mol. Cell Biol. **11**: 899–910.

Ribone, P.A., Capella, M., Arce, A.L., and Chan, R.L. (2017). A uORF represses the transcription factor AtHB1 in aerial tissues to avoid a deleterious phenotype. Plant Physiol. **175**: 1238–1253.

Rosado, A., Li, R., Van De Ven, W., Hsu, E., and Raikhel, N. V. (2012). Arabidopsis ribosomal proteins control developmental programs through translational regulation of auxin response factors. Proc. Natl. Acad. Sci. U. S. A. **109**: 19537–19544.

Schepetilnikov, M., Dimitrova, M., Mancera-Martínez, E., Geldreich, A., Keller, M., and Ryabova, L.A. (2013). TOR and S6K1 promote translation reinitiation of uORF-containing mRNAs via phosphorylation of eIF3h. EMBO J. **32**: 1087–1102.

Schepetilnikov, M. and Ryabova, L.A. (2017). Auxin signaling in regulation of plant translation reinitiation. Front. Plant Sci. **8**: 1–15.

Schepetilnikov, M. and Ryabova, L.A. (2018). Recent discoveries on the role of tor (Target of rapamycin) signaling in translation in plants. Plant Physiol. **176**: 1095–1105.

Sesma, A., Castresana, C., and Castellano, M.M. (2017). Regulation of translation by TOR, eIF4E and eIF2α in plants: Current knowledge, challenges and future perspectives. Front. Plant Sci. **8**: 1–7.

Shen, L., Liang, Z., Gu, X., Chen, Y., Teo, Z.W.N., Hou, X., Cai, W.M., Dedon, P.C., Liu, L., and Yu, H. (2016). N6-Methyladenosine RNA Modification Regulates Shoot Stem Cell Fate in Arabidopsis. Dev. Cell **38**: 186–200.

Sonenberg, N. and Hinnebusch, A.G. (2009). Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Cell **136**: 731–745.

Spealman, P., Naik, A.W., May, G.E., Kuersten, S., Freeberg, L., Murphy, R.F., and McManus, J. (2018). Conserved non-AUG uORFs revealed by a novel regression analysis of ribosome profiling data. Genome Res. **28**: 214–222.

Tanaka, M., Sotta, N., Yamazumi, Y., Yamashita, Y., Miwa, K., Murota, K., Chiba, Y., Hirai, M.Y., Akiyama, T., Onouchi, H., Naito, S., and Fujiwara, T. (2016). The Minimum open reading frame, AUG-stop, induces boron-dependent ribosome stalling and mRNA degradation. Plant Cell **28**: 2830–2849.

Traubenik, S., Reynoso, M.A., Hobecker, K., Lancia, M., Hummel, M., Rosen, B., Town, C., Bailey-Serres, J., Blanco, F., and Zanetti, M.E. (2020). Reprogramming of root cells during nitrogen-fixing symbiosis involves dynamic polysome association of coding and noncoding RNAs. Plant Cell **32**: 352–373.

Tsednee, M., Tanaka, M., Kasai, K., and Fujiwara, T. (2020). Boron-dependent regulation of translation through AUGUAA sequence in yeast. Yeast **37**: 638–646.

Tyagi, V., Parihar, V., Singh, D., Kapoor, S., and Kapoor, M. (2021). The DEAD-box RNA helicase eIF4A1 interacts with the SWI2/SNF2-related chromatin remodelling ATPase DDM1 in the moss Physcomitrella. Biochim. Biophys. Acta - Proteins Proteomics **1869**.

Urquidi Camacho, R.A., Lokdarshi, A., and von Arnim, A.G. (2020). Translational gene regulation in plants: A green new deal. Wiley Interdiscip. Rev. RNA **11**: 1–40.

Vogel, C. and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat. Rev. Genet. **13**: 227–232.

Yang, S.Y., Lu, W.C., Ko, S.S., Sun, C.M., Hung, J.C., and Chiou, T.J. (2020). Upstream open reading frame and phosphate-regulated expression of rice OsNLA1 controls phosphate transport and reproduction. Plant Physiol. **182**: 393–407.

Zanetti, M.E., Blanco, F., Reynoso, M., and Crespi, M. (2020). To keep or not to keep: mRNA stability and translatability in root nodule symbiosis. Curr. Opin. Plant Biol. **56**.

Zhang, H., Si, X., Ji, X., Fan, R., Liu, J., Chen, K., Wang, D., and Gao, C. (2018). Genome editing of upstream open reading frames enables translational control in plants. Nat. Biotechnol. **36**: 894–900.

Zhang, H., Zhong, H., Zhang, S., Shao, X., Ni, M., Cai, Z., Chen, X., and Xia, Y. (2019). NAD tagSeq reveals that NAD+-capped RNAs are mostly produced from a large number of protein-coding genes in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. **116**: 12072–12077.

Zhang, T., Wu, A., Yue, Y., and Zhao, Y. (2020). uORFs: Important Cis-Regulatory Elements in Plants. Int. J. Mol. Sci. **21**: 6238.

Zhang, Y., Wang, Y., Kanyuka, K., Parry, M.A.J., Powers, S.J., and Halford, N.G. (2008). GCN2-dependent phosphorylation of eukaryotic translation initiation factor-2α in Arabidopsis. J. Exp. Bot. **59**: 3131–3141.

Zhou, F., Roy, B., Dunlap, J.R., Enganti, R., and von Arnim, A.G. (2014). Translational Control of Arabidopsis Meristem Stability and Organogenesis by the Eukaryotic Translation Factor eIF3h. PLoS One **9**: e95396.

Study of interaction between SlMBD8 & SlIAA17 and characterization of their role in tomato fruit ripening by VIGS

Thesis Submitted to the University of Delhi in the
Partial fulfilment for the Degree of

Master of Science
in
Plant Molecular Biology and Biotechnology

2021

NITIKA GUPTA

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi 110021, India

Study of interaction between SlMBD8 & SlIAA17 and characterization of their role in tomato fruit ripening by VIGS

Thesis Submitted to the University of Delhi in the
Partial fulfilment for the Degree of

Master of Science
in
Plant Molecular Biology and Biotechnology

2021



NITIKA GUPTA

Department of Plant Molecular Biology
University of Delhi South Campus
New Delhi 110021, India

DEDICATED

TO

THREE GEMS OF MY LIFE

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Study of interaction between SlMBD8 & SlIAA17 and characterization of their role in tomato fruit ripening by VIGS**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfillment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Nitika Gupta

Date: 28.06.2021

Signature of the Candidate
(NITIKA GUPTA)

Arun K Sharma

Signature of the Supervisor
(Prof. Arun Kumar Sharma)

Indrid Dasgupta

Signature of the HOD
(Prof. Anil Grover)

Benito Juarez Road, New Delhi. 110021, **India.** Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

PREFACE

Tomato is a fleshy fruit that delivers us immense quality of multi-nutrients such as poly-unsaturated fatty acids, carotenoids, vitamins, plant sterols, etc. Physiological and molecular mechanisms associated with the ripening of this fruit are being studied by scientists as well as plant breeders to enhance the nutrition quality and shelf-life of the ripened fruit. The science behind this is more fascinating now as there have been many discoveries which reveal that during the process of fruit ripening many genes are set into motion. Studies showed that the dynamics of such genes during ripening are significantly regulated by the mean of epigenetic modifications including histone modifications, chromatin remolding and DNA methylation.

Some proteins have methylated DNA binding domain which reads methylated cytosine marks present in the chromatin and transmits the code of gene activation or inactivation by recruiting some other proteins to the target. Information gathered from publicly available gene expression data and researchs in our laboratory indicated that some tomato methyl CpG binding domain proteins (SlMBDs) could be involved in the regulation of fruit ripening with hormonal prospective. In Arabidopsis, there is an interaction of MBD proteins with Aux/IAA proteins, therefore, identification of such interaction between their homologs in tomato can be a new insight for the role of the plant hormone auxin in regulation of the process of ripening through epigenetic regulation. It has already been reported that tomato AUX/IAA17 is involved in the regulation of size of the fruit during its development.

The present study comprises of characterization of protein-protein interaction between SlMBD8 and SlIAA17 and attempts to understand their role by employing VIGS in tomato fruits. As the expression of both the genes is maximum in small-sized immature fruits, their silencing via VIGS will give useful information. This may provide enough information to proceed further for the generation of stable transgenic plants of tomato to characterize the role of any such possible interaction between SlMBD8 and SlIAA17.

ACKNOWLEDGEMENT

"At times our own light goes out and is revitalize by a spark from another person. Each one of us has a cause to think with deep gratitude of those who have lighted the flame within us..."
-Albert Schweitzer

On top of everyone, I would like to thank the most gracious and merciful God for making me capable of doing this research and for all the blessings you have prepared for me.

It is a great pleasure and pride to acknowledge my deep gratitude towards my elite mentor ***Prof. Arun Kumar Sharma.*** *His presence in my day-to-day guidance and giving complete independence in my work ignites an enthusiasm for better understanding and validation of research. I am forever indebted to him for helping me throughout my dissertation.*

I would like to express my heartfelt gratitude and deep appreciation to all the faculty members of the department ***Prof. Anil Grover, Prof. J. P. Khurana, Prof. Paramjit Khurana, Prof. Akhilesh K. Tyagi, Prof. Indranil Das Gupta, Prof. Girdhar K. Pandey, Prof. Sanjay Kapoor, Prof. Saurabh Raghuvanshi and Prof. Surekha K. Aggarwal*** *for their encouragement, suggestions and ideas throughout my master's degree and being source of inspiration towards my work.*

I extend my sincere thanks to ***Vijendra Sir*** *for all the help and encouragement provided throughout. His unfailing support, untiring efforts and constant dedication have enabled me to carry out this work to the best of my capability. I also wish to extend my heartfelt thanks to* ***Priya Ma'am*** *for immense help and guidance to tackle all my problems with greatest fortitude. Thank you for always cheering me up and stood by me throughout all the time.*

I would extend my sincere thanks to Amrapali Ma'am, Sanskriti Ma'am, Utkarsh Sir, Nitin Bhaiya, Tiwari Ji and Raju Bhaiya. Most importantly I would like to express my gratitude towards three most important persons of my life: Mumma, Papa and Bhai. It gives me immense pleasure to thank all my batchmates for their all-time support. Special acknowledgment to Dr. Vishaw Vikas Sharma, Dr. Tanika Gupta, Bhabhi, Marvi, Aarushi and Nandita for my biggest strength in most difficult time.

LIST OF ABBREVIATION

SYMBOL	MEANING
µl	Microliter (s)
µg	Microgram
AFB	Auxin Signaling F Box
AD	Activation Domain
ARF	Auxin response factors
AUX/IAA	Auxin/Indole Acetic Acid
BD	Binding Domain
BiFC	Bimolecular Fluorescence complementation
bp	Base pair
CDS	Coding sequence
Co-IP	Co-Immunoprecipitation
cDNA	Complementary DNA
DNA	Deoxyribonucleic Acid
DDB	DNA binding protein
dNTP	Deoxyribonucleotide triphosphate
g	Gram
GS	Gene silencing
hrs.	Hours
IPTG	Isopropyl β-D-1 thiogalactopyranoside
kb	Kilobase
l	Liter
MCS	Multiple cloning sites
MCT	Micro centrifuge tube
MG	Mature green
mins	Minutes
mM	Millimolar
ml	Milliliter (s)
MQ	Milli q
ng	Nanogram
NLS	Nuclear localization signal
OD	Optical density
PCR	Polymerase chain reaction
PTGS	Post-transcriptional gene silencing
RNA	Ribonucleic acid
rpm	Revolutions per minute
SCF	SKP-Cullin-F Box
SKP	S-phase kinase-associated protein
TIR	Transport Inhibitor Response
TGS	Transcriptional gene silencing
TRV	Tobacco Rattle Virus
VIGS	Virus Induced Gene Silencing
Y2H	Yeast two-hybrid
YEM	Yeast Extract Mannitol

TABLE OF CONTENTS

Preface	**i**
Acknowledgments	**ii**
List of Abbreviations	**iii**
Review of Literature	**1-12**
Materials and Methodology	**13-29**
Results and Discussion	**29-53**
Conclusion and Future Work	**54-55**
References	**56-59**
Appendix	**60-61**

REVIEW OF LITERATURE

1. INTRODUCTION TO TOMATO:

History:

Tomato (*Solanum lycopersicum*) belongs to the Solanaceae family. It has been imported from the Andean region to Europe in about 16th century. The origin of its domestication was not cleared even after its introduction in Europe, and hence, two hypotheses came forward: Peruvian hypotheses and Mexican hypotheses. In 1886, DeCandolle used various approaches of sciences like botany, paleontology, history, archeology and philology to solve this hypothesis and he concluded that tomato has Peruvian origin of domestication (Bergougnoux, 2014). But today, it is pervasive all over the globe

Model Organism:

Tomato is considered as a model organism of the Solanaceae family. It is of major interest for both scientists as well as plant breeders because of its short life cycle, photoperiod insensitivity, controlled pollination and hybridization, self-fertility, homozygosity, availability of whole-genome sequence, regeneration of the whole plant from an explant and well-characterized mutants (Bergougnoux, 2014). Tomato is the second most important crop plant after potato. It is a superior source of many vital nutrients such as vitamin A and C, iron, phosphorus, lycopene and β-carotene. Surprisingly, this extremely important crop is a fruit and not a vegetable.

Tomato Ripening:

Ripening of fruit is a very complex process and certain changes are associated with ripening. For example, expression of definite genes and color of fruit changes during the process of ripening. Tomato ripening is characterized by the modifications in the softness or texture of a fruit. More alterations associated with ripening includes changes in the level of sugars, acids and volatile compounds. Parallelly, there is a transition of chloroplast to chromoplast, biosynthesis of carotenoids and hormonal regulation. At molecular level, there are certain changes in the DNA methylation level also.

2. CLIMACTERIC VS NON-CLIMACTERIC:

Climacteric fruits are associated with the increase in the biosynthesis of phytohormone “ethylene” upon onset of ripening. Initially, the term “climacteric” was used to highlight the burst of CO_2 production upon ripening but the actual cause of ripening in climacteric

fruits is ethylene (Pesaresi et al., 2014). Non-climacteric fruits do not manifest an increase in the biosynthesis of ethylene next to the initiation of ripening but exogenous application of ethylene expedites the process of ripening. Examples of climacteric fruits are apple, banana, tomato, avocado, peach and non-climacteric fruits: grapes, strawberry, citrus, etc.

3. HORMONAL REGULATION:

Ethylene plays a significant role in fruit ripening. Two ethylene biosynthesis systems are revealed in climacteric fruits. System 1 involves biosynthesis of basal level ethylene and is autoinhibitory whereas system 2 involves biosynthesis of ethylene after initiation of ripening and is autocatalytic (Quinet et al., 2019). Other phytohormones like abscisic acid, gibberellic acid, auxin, cytokinin and brassinosteroids are also known to play major role in fruit ripening. Abscisic acid and cytokinin escalate the biosynthesis of ethylene, hence, promote fruit ripening. Exogenous application of brassinosteroids also leads to the production of ethylene and therefore promotes fruit ripening whereas auxin and gibberellic acid have an antagonistic effect on fruit ripening.

4. DNA METHYLATION:

DNA methylation is a reversible epigenetic phenomenon in which there is addition of methyl group to 5^{th} carbon of cytosine, catalyzed by methyltransferases, which transfer a methyl group from S-adenosyl-L-methionine to the pyrimidine ring.

5' cytosine methylation can be symmetrical (CpG or CpNpG, where N is any nucleotide) as well as asymmetrical (CpHpH where H= A, T, C). It is a conserved process. The change in the interaction between proteins and DNA due to methylation leads to either increase or decrease in the rate of transcription.

DNA Methylation in Plants:

In plants, DNA methylation oversees many important traits like its role in biotic or abiotic stresses, ripening of tomato, heterosis etc. (Gallusci et al, 2016). Varying in their structure and function as well as having different preferences for cytosines in CpG or CpHpG, three families of DNA methyltransferases are present in plants (Finnegan and Kovac 2000):

- **MET1 family:** In Arabidopsis, DNA methyltransferase 1 (MET1) maintains methylation at 5'-CpG-3' sites. Though, primarily MET1 functions as maintenance

methyltransferases but it also plays a part in *de novo* methylation. A homolog of MET1 is also present in mammals.

- **CMT3 family:** In Arabidopsis, chromomethylases (CMT3) maintain methylation at 5'-CpHpG-3' sites. They are known to methylate heterochromatin DNA. This enzyme is unique to plants only.

- **Dnmt3 methyltransferases family:** Dnmt3 causes *de novo* methylation.

DNA Methylation in Fruit Ripening:

Evidence for the changes in DNA methylation during fruit development was initially given by Teyssier et al., 2008. In tomato, there is an overall decrease in the level of DNA methylation in specific tissues during fruit ripening. Treatment of immature tomato with azacytidine, an inhibitor of DNA methylation, causes premature ripening, revealing the relationship between DNA methylation level and fruit ripening (Tang et al., 2020). Before fruit ripening, the level of methylation and demethylation is maintained whereas during fruit ripening, this balance is disrupted and there is an increase in expression of genes causing hypomethylation and hence fruit ripening (Figure 1).



Figure 1: Shift in the balance between the level of DNA methylation and demethylation during fruit ripening

5. METHYLATED CpG BINDING DOMAIN (MBD) CONTAINING PROTEINS IN PLANTS:

Methylation at 5' cytosine plays a principal role in plant growth and development. It also plays a significant role in genome organization, chromatin structure, gene expression, pathogen infection, and transposon and transgene silencing. Methylated DNA binding proteins contain methylated CpG binding domain (MBD) that specifically recognizes and binds to 5' methylcytosine. But there are some proteins like, AtMBD4 and

AtMBD11, which can bind irrespective of 5' methylcytosine. Therefore, all MBDs cannot be considered as "methylation-specific" (Grafi et al., 2007).

These proteins are members of an evolutionarily conserved protein family present in eukaryotic lineage. MBDs are considered to function as an interpreter of signals of DNA methylation. 13 MBD genes are present in *Arabidopsis thaliana*, 18 in *Solanum lycopersicum*, 14 in maize, 17 in rice, 14 in Poplar (Grafi et al., 2007). MBD proteins help in systematizing the crosstalk between DNA methylation, chromatin organization and epigenome modifications to accomplish systematic transcription of a gene. There is recruitment of certain proteins on methylated DNA such as histone-modification proteins due to MBD proteins which causes alterations in gene expression. Mammalian MBD proteins are capable of binding to symmetrical methylated CpG sites only whereas plant MBDs can bind to both symmetrical CpG and CpNpG sites as well as asymmetrical CpHpH (H=C, A, T). But still, there is considerable homology between mammalian and plant MBDs.

Tomato MBDs:

In tomato, approximately 18 genes have been reported to encode for MBD proteins. With the already known sequences of MBD proteins of Arabidopsis and humans, tomato MBDs were also identified using manual BLAST search tool. The identified MBD proteins showed variation in their sizes with the smallest MBD i.e., SlMBD5 (made up of 75 amino acids) and largest MBD i.e., SlMBD11 *(*made up of 1886 amino acids) (Parida et al., 2018). Apart from the presence of MBD motif, certain other regions are also present in MBD proteins like Zinc finger domain, a coiled-coil region, internal repeats and nuclear localization signal (NLS).

Expression pattern studies revealed that the level of transcripts of *SlMBDs* varied in different stages of fruit development. For example, during fruit ripening, transcript levels of *SlMBD8* and *SlMBD15* gradually decrease whereas transcript levels of *SlMBD2*, *SlMBD3* and *SlMBD18* gradually increase in fruit tissue (Parida et al., 2018). In tomato, DNA methylation plays an inhibitory role in fruit ripening. Being a climacteric fruit, phytohormone ethylene is also engaged in tomato fruit ripening.

***SlMBD8*:** It is ubiquitously expressed in all tissues and has Zn finger domain along with MBD domain. Using quantitative RT-PCR, expression analysis of *SlMBD* genes was carried out in wild-type cultivar (Pusa Ruby), *rin* mutant (ripening inhibitor) and *Nr* (never ripe) mutant of tomato in three different stages of fruit ripening: mature green (MG), breaker (Br) and red-ripe (RR, here, Br+10). It was observed that the expression

level of *SlMBD8* was increased in *rin* mutant fruits but strongly decreased in *Nr* mutant fruits compared to wild-type (Parida et al., 2018). Also, the stage-specific expression level of *SlMBD8* was observed to be gradually decreased from MG to RR in wild-type fruits as well as in mutant fruits. According to Tomato Genome Consortium 2012, the maximum expression of *SlMBD8* can be seen in 1cm long tomato fruits. Many interacting partners are also known for MBDs in tomato, one of which is AUX/IAA protein.

6. AUX/IAA GENE FAMILY:

Auxin plays an indispensable role in many plant activities like adventitious root initiation, tropic responses, leaf formation, fruit formation, embryo development, apical dominance, vascular tissue formation, and in many cellular activities like cell division, cell differentiation and cell enlargement. Auxin/Indole acetic acid (AUX/IAA) genes usually encode for short-lived proteins and help in the auxin-transduction pathway. AUX/IAA proteins act as active repressor of auxin-dependent gene transcription (Audran-Delalande, 2012). Promoters of auxin-inducible genes contain Auxin Response Elements (AREs) bound with auxin response factors (ARFs). AUX/IAA proteins interact with ARFs in the absence of auxin and inhibit transcription of genes that are usually activated by ARFs. But when the level of auxin is high, these proteins interact with SKP-Cullin-F-Box TRANSPORT INHIBITOR RESPONSE 1/AUXIN SIGNALING F BOX ($SCF^{TIR1/AFB}$) receptors and are degraded via 26S proteasome pathway (Figure 2).

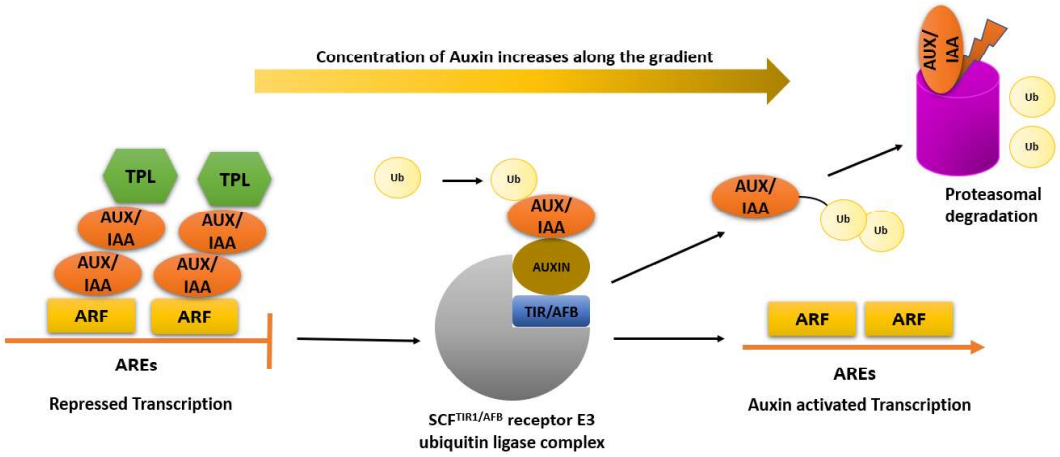


Figure 2: Auxin dependent proteasomal degradation of AUX/IAA proteins

These proteins comprise of extremely conserved domains, Domain I, Domain II, Domain III, Domain IV. Domain I is a repression domain in which ethylene response factor (ERF)-associated amphiphilic repression motif is present that recruits TOPLESS (TPL) co-repressor. Domain I is represented by "LxLxLx" motif. Domain II is essential for the interaction of AUX/IAA with TIR1/AFB via a conserved "GWPPV" motif present in it. This domain confers instability and regulates rapid turnover of AUX/IAA proteins. In between Domain I and II, there is a conserved lysine-arginine (KR) motif, which helps in regulating rapid turnover. Domain III contains ββα motif having role in homo or heterodimerization with AUX/IAAs and/or ARFs. Domain IV contains an acidic motif and three β-sheets and two α-helices that aid in dimerization. Also, there is presence of a "GDVP" motif between β4 and α2. (Wu et al., 2012; Luo et al., 2018) (Figure 3).

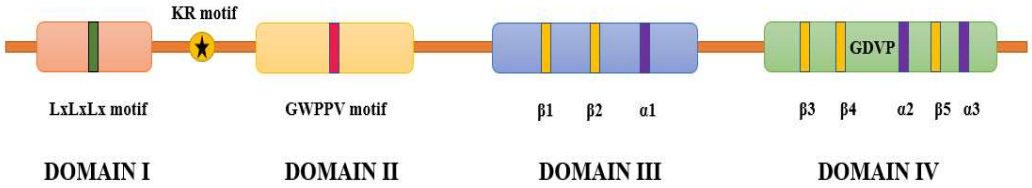


Figure 3: Structural illustration of AUX/IAA proteins

Auxin level varies among different plant tissue and TIR1/AFB-AIUX/IAA combinations have different binding affinities towards auxin. 29, 26, 27, 31 and 31 AUX/IAA gene family members have been identified in Arabidopsis (Luo et al., 2018), tomato, potato, rice and maize respectively (Wu et al., 2012).

Tomato AUX/IAA Gene Family:

In tomato, 26 AUX/IAA gene family members have been identified so far. Down-regulation of several *AUX/IAA* genes causes variation of auxin-related phenotypes such as apical dominance, tropic responses, leaf shape and leaf formation in tomato. These alterations are clearly visible, causing distinct phenotypes. The expression of 19 out of 26 *AUX/IAA* genes can be detected in all tissues of tomato whereas 7 *AUX/IAA* genes express specifically in some tissues only (Wu et al., 2012).

Out of 26, 16 SlIAA promoters contain conserved ethylene response motifs along with auxin-responsive elements. Hence, upon ethylene treatment, some auxin-responsive genes are upregulated while other genes are downregulated (Audran-Delalande, 2012). Great contrast can be seen in the sizes of AUX/IAA proteins from 147 amino acids in SlIAA33 to 349 amino acids in SlIAA9. Phylogenetic analysis

revealed that even proteins belonging to the same branching point shared low identity among themselves.

It has been experimentally proven that downregulation of *SlIAA3* causes phenotypic characteristic of reduced auxin and increased ethylene levels in mutant tomato plants, therefore it acts as a molecular connection between auxin and ethylene signaling-pathways (Kumar et al., 2015). Likewise, *SlIAA9* is involved with fruit development and leaf morphogenesis which are also associated with auxin signaling pathway and its downregulation causes formation of parthenocarpic fruits and irregular leaf shape (Wang et al., 2005). The expression pattern study of various *SlIAA* genes signifies that the proteins encoded may function both redundantly as well as specifically in a tissue.

SlIAA17: In tomato, IAAs shows organ-specific expression patterns like stem, root, leaf, flower bud or ovary. *SlIAA17* is highly expressed in stem as compared to other organs of the plant. Transcript level of *SlIAA17* continuously increased during early stages of fruit development (Wu et al., 2012) and is maximum in 2 cm long tomato fruit (Tomato Genome Consortium, 2012).

7. YEAST TWO-HYBRID (Y2H):

Yeast two-hybrid system is an effective system for examining protein-protein interactions *in-vivo*. It was first devised by Stanley Fields and Ok-Kyu Song (1989). It was known that transcription factors of eukaryotes are regimented into two functionally distinct domains namely activation domain and DNA binding domain. **DNA binding domain (BD)** recruit transcription factor on particular genomic DNA sequences and **activation domain (AD)** helps in the recruitment of transcription machinery that brings out transcription. Both domains are physically separable. To function properly, these domains need not to be present on the same protein. Gene can also be expressed if both the domains are present in the immediate vicinity (Paiano et al., 2019.

In yeast two-hybrid system, the cDNA which encodes for a protein of interest (say protein X) is usually cloned into a vector that permits the protein expression (known as "bait protein"), integrated with transcription factor BD. Another cDNA that encodes for other protein (say protein Y) is also cloned into a vector which further allows expression of a protein (known as "prey protein") integrated with transcription factor AD. When both expressed proteins interact in the nucleus of the yeast, it reconstitutes a functional

transcription factor in which AD and BD will come in close vicinity therefore, it will switch on the transcription of a reporter gene (Brückner, 2009; Paiano et al., 2019) (Figure 4).

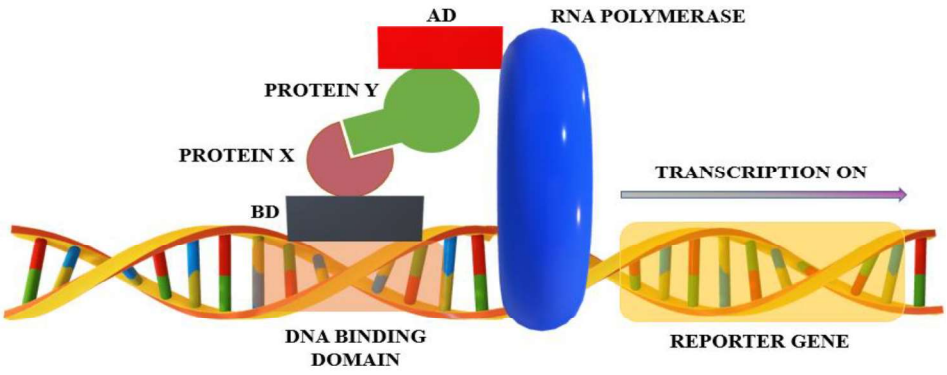


Figure 4: Principle of Yeast-Two-hybrid

It is already known that *SlIAA9* is involved in fruit development, leaf morphogenesis as well as fruit parthenocarpy. Y2H assay was used to identify interacting partners of *SlIAA9.* Therefore, it was found that many ARFs such as SlARF6A, SlARF8A, SlARF8B, SlARF24 interact directly with SlIAA9 in controlling tomato leaf forms (Wu, 2018).

Yeast two-hybrid assay was also demonstrated to check the physical interaction between SlMBD5 and UV damaged SlDDB1 proteins in tomato. DDB1 (DNA binding protein 1) protein complexed with DDB2 is involved in nucleotide excision repair mechanism and SlMBD5 is involved in multiple processes. SlMBD5 is a close homolog of AtMBD5. From Y2H analysis, it was confirmed that both proteins physically interact with each other and interaction is mediated by C-termini of SlMBD5 and β-propeller C (BPC) of SlDDB1 (Li, 2016)

8. GENE SILENCING IN PLANTS:

Gene silencing is one of the most powerful techniques to characterize the function of the target gene. Plants have evolved a mechanism of gene silencing due to introduction of sRNAs (small RNA) of 20-30 nucleotides in length into the plant and this leads to suppression of the expression of sequence complementary to small RNAs at transcriptional level, known as transcriptional gene silencing (TGS) or at post-transcriptional level, known as post-transcriptional gene silencing (PTGS).

Transcriptional gene silencing: Histone modifications or chromatin remodeling results in blocking the expression of the gene due to unavailability for transcriptional machinery causing transcriptional gene silencing. In TGS, there is an increase in DNA methylation level which further generates heterochromatic conditions (Souza, 2007).

Post-transcriptional gene silencing: Due to the introduction of small RNAs, either the transcript of a gene (mRNA) is completely degraded or there is translational repression, hence the gene product will not be formed. This technique was first reported in plants and is known as co-suppression (Pandey, 2015).

Gene silencing (GS) plays an essential role in regulating plant development as well as in response to biotic and abiotic stress. GS also plays a significant role in providing antiviral defense. Plants have evolved a mechanism to oppose viral infection by activating antiviral silencing as their defense mechanism. This involves PTGS as there is either cleavage or translational suppression of target RNAs.

VIGS is a PTGS technique that can also act at the transcriptional level by methylating either target DNA or histone proteins. This category involves TGS (Wang et al., 2019). TGS is further categorized into two categories: *cis* (silencing induced by transgene at one locus) and *trans* (silencing induced by transgene at another locus) depending upon the cause of induction of gene silencing.

9. VIRUS-INDUCED GENE SILENCING (VIGS) IN PLANTS:

The term "Virus Induced Gene Silencing" was originally used by van Kammen (1997). The technique involves the introduction of a recombinant vector into the plant to knock down the expression of endogenous gene to study its function. It involves cloning of an endogenous gene sequence and introduction of the cloned fragment in a viral vector, forming a recombinant vector. These recombinant vectors are then inoculated into the plants which further initiates PTGS. Therefore, plant sequence homologous to the introduced fragment gets silenced. VIGS has been successfully used to silence host genes in many plant species such as tomato, *Nicotiana*, pepper, potato, Arabidopsis, maize, rice, etc.

In monocots also, VIGS has divulged functions of many important plant genes which play major roles in multiple facets in plant growth and development, providing resistance against biotic or abiotic stresses, flowering and cell wall synthesis (Kant,

2019). In wheat, functions of many important genes such as *Col1*, *GBSS*, *DMC1* have been revealed using VIGS only (Bennypaul et al., 2012).

Multiple VIGS vectors are available that can be used for silencing more than one plant species. These are standard Ti plasmid-derived vectors used for *Agrobacterium tumefaciens* plant transformation. These vectors contain a cloned portion of the host genome. In the course of infection with the recombinant vector, the defense mechanism of a host is induced against the cloned part of the host gene, hence there is loss of function phenotype through which function of a gene can be identified.

For example, because of the broad range of hosts, TRV (Tobacco Rattle Virus) derived vector is a widely used system used in silencing dicotyledons plants. Other used VIGS vectors are given in the table below (Senthil-Kumar and Mysore, 2014).

Table 1: Commonly used VIGS vector and their target species

VIGS VECTOR	TARGET PLANT SPECIES
Alternanthera mosaic virus (AltMV)	*N. benthamiana* and *A. thaliana*
Bean pod mottle virus (BPMV)	Soybean and common bean
Cucumber mosaic virus (CMV)	*N. benthamiana* and soybean
Potato virus X (PVX)	*N. benthamiana* and potato
Beet curly top virus (BCTV)	Spinach and tomato
Cabbage leaf curl virus (CLCV)	*N. benthamiana* and *A. thaliana*
Tobacco curly shoot virus (TCSV)	Tobacco and tomato
Rice tungro bacilliform virus (RTBV)	Rice (Indica and Japonica)
Barley stripe mosaic virus (BSMV)	Wheat and barley
Tomato yellow leaf curl China virus (TYLCCNV)	Tobacco and tomato

Virus-induced gene silencing (VIGS) in tomato: Tomato leaf, stem, fruit, cotyledon as well as hypocotyl can be used for performing VIGS. TRV vector (having gene construct) can be either syringe-infiltrated or vacuum-infiltrated into the carpopodium of the detached fruit. It can also directly be infiltrated via syringe on attached tomato fruits (*on*-vine VIGS). There can be silencing of different genes simultaneously in different fruits present on the same plant. TRV vector cannot transfer from one fruit to another or to any other part of the plant. Therefore, any gene can be efficiently cloned into the TRV vector and infiltrated into the tomato fruit. The vector will start spreading systematically as well as start replicating within the same fruit, due to which there will be silencing of the cloned gene. Many ripening-related genes have been silenced using this method and hence their function is known now. These pieces of evidence clearly demonstrate that VIGS can be potentially used to study the function of tomato genes related to its development and ripening. For example, silencing of *LeEIN2* clearly resulted in the suppression of tomato fruit-ripening (Fu et al., 2005)

In the present study, *SlMBD8* and *SlIAA17* genes have been selected to study their effects on tomato fruit ripening. Further, to know whether there is any interaction between SlMBD8 and SlIAA17 proteins' just like their homologs in Arabidopsis, yeast two-hybrid approach was employed.

MATERIALS AND METHODOLOGY

1. BIOLOGICAL MATERIALS:

a) **Plant material:** Tomato (*Solanum lycopersicum*) fruit was used for plant transformation.

b) **Bacterial strain:** DH5α strain of *E. coli* was used for transformation. GV3101 strain of *Agrobacterium* was used for plant transformation.

GENERAL PROTOCOLS

1. ISOLATION OF PLASMID:

Plasmid isolation was performed using FavorGen™ Plasmid DNA Extraction Mini Kit. To start with, 4 ml of overnight grown bacterial culture was pelleted in a 2 ml MCT by repeated centrifugation at 13,000 rpm for 2 mins. The pellet cells were re-suspended in 200 µl of FAPD1 solution (RNase added) using a vortex mixer. Further, 200 µl of FAPD2 solution was added and mixed thoroughly by inverting not more than 6 times. It was then kept still for about 5 mins. After 5 mins, 300 µl of FAPD3 solution was added and mixed gently by inverting the tube 5-6 times. It was then kept on ice for about 10 mins. The solution was then centrifuged at 13,000 rpm for 5 mins at RT. With the help of the pipette, the supernatant was then transferred to the supplied FAPD column placed on to supplied collection tube and centrifugation was performed at 13,000 rpm for 1 min. Flow-through was discarded and a FAPD column was again placed onto the collection tube. Now, 400 µl of W1 buffer was added to a FAPD column and centrifuged at 13,000 rpm for 1 min. Again, the flow-through was discarded and a FAPD column was placed onto the collection tube. To the column, 700 µl of wash buffer was added and centrifuged at 13,000 rpm for 1 min. Flow-through was discarded and a FAPD column was again placed onto the collection tube. Another centrifugation was performed at 13,000 rpm for 3 mins to remove any residual wash buffer present in the column. The FAPD column was transferred to fresh 1.5 ml MCT and 50 µl of MQ water was added to the center of the membrane of a FAPD column and incubated for 5 mins. After 5 mins, it was centrifuged at 13,000 rpm for 2 mins to elute the plasmid DNA. FAPD column as well as collection tube were discarded and the eluted plasmid was quantified using NanoDrop and stored at -20^{o}C.

2. RESTRICTION DIGESTION:

Fast-Digest Green buffer compatible with restriction enzymes was used for digesting plasmid DNA after quantification. Restriction buffer was thawed on ice and stock concentration-10X was reduced to working concentration-1X. Reaction mixture having plasmid DNA, restriction buffer, restriction enzymes and MQ were incubated in a dry bath at 37^{o}C for about 120 mins.

3. GEL ELUTION OF DNA FRAGMENT:

Desired DNA band was cut from the agarose gel using a sterile blade, after the scale-up reaction. The excised fragment was taken into 1.5 ml of fresh MCT. Zymoclean™ Gel Recovery Kit was used for gel elution. Agarose dissolving buffer (ADB) was added in a 3:1 ratio of buffer and excised fragment. The tube was then incubated in a dry bath at 55°C for 10 mins, tapping in-between 2 times. Using a micropipette, the solution was transferred to a Zymo spin column placed on provided collection tube. It was then centrifuged at 13,000 rpm for 45 seconds at RT. Flow-through was discarded and a Zymo spin column was again placed onto the collection tube. Now, 200 µl of DNA Wash Buffer was added to the column and centrifuged at 13,000 rpm for 30 seconds. This step was repeated. It was drily centrifuged at 13,000 rpm for 2 mins. Zymo spin column was then placed onto a fresh 1.5 ml MCT. Further, 10 µl of MQ water was added into the center of the membrane of the column and incubated at RT for 5 mins. Final centrifugation was performed at 13,000 rpm for 2 mins. It was then run on an agarose gel to check the concentration of the eluted product.

4. POLYMERASE CHAIN REACTION:

Primer pairs were made for the genes of interest and ordered to Sigma-Aldrich. Amplification of DNA was performed using SimpliAmp™ Thermal Cycler. According to the gene, various parameters such as template quantity, sample tissue, presence or absence of $MgCl_2$, primers, extension time, annealing length were optimized. Depending upon the melting temperature of the primer pairs, annealing temperature as well as the number of amplification cycles were set up. Amplification reaction was performed using iTaq polymerase. The concentration of agarose gel was decided according to the size of the amplicon. After the amplification reaction, samples (mixed with 1 µl loading dye) were loaded into the wells of an agarose gel and agarose gel electrophoresis was performed. The size of the amplicon was verified with the help of a DNA ladder.

METHODOLOGY

1. SEQUENCE ANALYSIS OF GENES FOR VIGS AND DESIGNING OF PRIMERS:

CDS sequences of *SlMBD8* and *SlIAA17* were retrieved from NCBI database. Using the VIGS tool in Sol Genomics Network, 300 bp sequences specific to VIGS were selected. VIGS primers (Table 2) were then designed using Gene Runner software with *Xho*I and *Xba*I restriction sites in forward and reverse primers respectively, to perform restriction-based cloning.

Table 2: Primer sequences of *SlMBD8* and *SlIAA17* specific to VIGS region.

GENE	VIGS PRIMER SEQUENCE	SIZE (bp)
***SlMBD8* Forward Primer**	**5'-CTCGAG**ATGGCGAAAGATTCACCA-**3'**	24
SlMBD8 Reverse Primer	**5'-TCTAGA**CTTAGAGTAATCTCTCCTGAG-**3'**	27
***SlIAA17* Forward Primer**	**5'-CTCGAG**ATGAGTAGTAATAAGTTGGATTTT-**3'**	30
***SlIAA17* Reverse Primer**	**5'-TCTAGA**CATGCTCACCTTCACCAA-**3'**	24

2. PRIMER DESIGNING FOR YEAST TWO-HYBRID ASSAY:

CDS sequences obtained before were used for primer designing (Table 3) using Gene Runner Software. For Yeast two-hybrid, topo-based cloning was performed therefore, CACC sequence was added 5' upstream to the forward primer. It will enable directional cloning of the PCR product into the entry vector.

Table 3: Primer sequences of *SlMBD8* and *SlIAA17* specific to Y2H region.

GENE	Y2H PRIMER SEQUENCE	SIZE (bp)
***SlMBD8* Forward Primer**	**5'-CACC**ATGGCGAAAGATTCACCA-**3'**	22
***SlMBD8* Reverse Primer**	**5'-**TTTAGTGCTACTAGCTACACCC-**3'**	22
***SlIAA17* Forward Primer**	**5'-CACC**ATGAGTAGTAATAAGTTGGA-**3'**	24
***SlIAA17* Reverse Primer**	**5'-**TCCATTTCTGTTCTTGCA-**3'**	18

3. CLONING OF *SlMBD8* AND *SlIAA17* FOR VIGS:

PCR amplification of a specific region of *SlMBD8* and *SlIAA17* for VIGS:

Polymerase chain reaction was carried out to amplify regions of *SlMBD8* and *SlIAA17* specific to VIGS using SimpliAmp™ Thermal Cycler. Primers used for amplification had *Xho*I and *Xba*I restriction sites in forward and reverse primers respectively. Conditions such as dNTP concentration, cDNA concentration, annealing temperature, extension time, were optimized. Depending upon the melting temperature of the primer pairs, annealing temperature as well as the number of amplification cycles were set up. The master mix was prepared according to Table 4. After standardizing conditions (Figure 5), an amplification reaction was performed using iTaq polymerase. The amplified product was run on 1.5% (w/v) agarose gel.

Table 4: Master-mix setup for PCR amplification of VIGS specific region of *SlMBD8* and *SlIAA17*

REAGENTS	VOLUME REQUIRED FOR *SLMBD8*	VOLUME REQUIRED FOR *SLIAA17*
10X Buffer	2.5 µl	2.5 µl
Forward Primer	1 µl	1 µl
Reverse Primer	1 µl	1 µl
dNTPs	0.5 µl	0.5 µl
iTaq	0.2 µl	0.2 µl
cDNA	0.5 µl	1.5 µl
$MgCl_2$	-	0.5 µl
MQ	19.3 µl	17.8 µl
Total	25 µl	25 µl

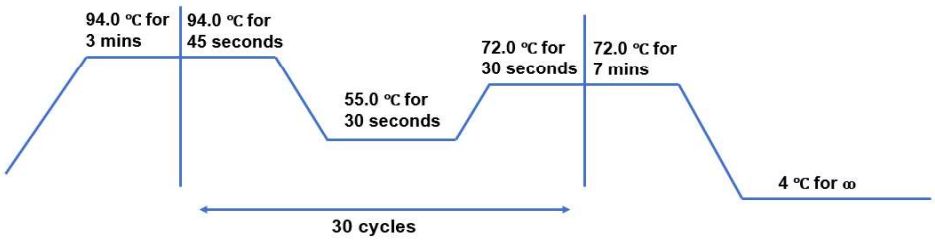


Figure 5: PCR conditions for PCR amplification of 300 base-pair specific region of *SlMBD8* and *SlIAA17* for VIGS

Cloning of DNA fragments into a pGEM®-T vector:

Primary cloning of *SlMBD8* and *SlIAA17* was done in pGEM®-T easy vector (Promega). Amplification of *SlMBD8* and *SlIAA17* was performed using iTaq DNA polymerase which produces poly-A tail at 3' end of the PCR product and pGEM®-T easy vector, which contains T overhangs. Therefore, both ends were compatible with each other. For ligation reaction, the molar ratio of vector to insert was 3:1. Table 5 shows the recipe for the ligation reaction mixture. The reaction mixture was kept at 16^{o}C overnight.

Table 5: Components of Ligation reaction mixture containing vector and insert:

COMPONENTS	AMOUNT REQUIRED
2X Buffer	10 µl
pGEM®-T easy vector	25 ng
Insert	75 ng
Ligase enzyme	1 µl
MQ	Volume makeup up to 20 µl

Next day, a ligation reaction mixture was used to transform *E. coli* Z-competent cells. The cells were plated on LB agar having IPTG (100mM), X-Gal (20 mg/ml) and 100 µg/ml carbenicillin (selection for pGEM®-T easy vector) and kept at 37^{o}C overnight. White and blue colonies were observed next day representing transformed vs non-transformed.

Screening of Transformed colonies:

Screening of transformed colonies from non-transformed ones was done by colony PCR as well as by restriction-digestion. Colony PCR was performed using forward as well as reverse primers, specific to the gene. DNA from colonies having gene of interest will amplify a band of 300 bp.

For restriction digestion, plasmids were isolated from an overnight grown culture of positive colonies of *SlMBD8* and *SlIAA17*. Screening for gene of interest was done by performing restriction-based digestion using *Xba*I and *Xho*I. The reaction mixture was

prepared using components as shown in Table 6. The mixture was incubated at 37°C for about 2 hrs and samples were run on 1.5% (w/v) agarose gel.

Table 6: Components of restriction digestion reaction of vector for confirmation of cloning of insert into the vector:

COMPONENTS	AMOUNT REQUIRED
10X Fast Digest Buffer	3 µl
Plasmid	400ng
***Xba*I**	0.5 µl
***Xho*I**	0.5 µl
MQ	Volume makeup up to 30 µl

Cloning of DNA fragments into Binary Vector-TRV2:

A scale-up reaction was set-up from which insert cloned into pGEM®-T easy vector was eluted. Both insert as well as vector were digested with *Xba*I and *Xho*I and assessed qualitatively through agarose gel electrophoresis. After that, the ligation reaction was set up in a 1:3 molar ratio of vector and insert using Fermentas Rapid Ligation Kit. The ligation reaction mixture was prepared on ice and kept at 22°C for 4 hrs. Recipe of the ligation reaction mixture made is given in Table 7.

Table 7: Components of ligation reaction mixture containing vector and insert:

COMPONENTS	VOLUME REQUIRED
5X Ligation Buffer	4 µl
Insert	According to the quality of insert
Vector	50 ng
T4 DNA Ligase	1 µl
MQ	Volume makeup up to 30 µl

Next day, a ligation reaction mixture was used to transform *E. coli* Z-competent cells and the cells were plated on LB agar having 50 µg/ml kanamycin (selection for TRV2) and incubated at 37°C overnight. Screening of transformed colonies from non-

transformed colonies was done by colony PCR as well as by restriction-digestion. Colony PCR was performed using forward and reverse primers of the genes. Colonies having gene of interest will amplify a band of 300 bp.

Plasmids were isolated from an overnight grown culture of positive colonies of *SlMBD8* and *SlIAA17*. Screening for gene of interest was done by performing restriction-based digestion using *Xba*I and *Xho*I. The reaction mixture was prepared using components as shown in Table 8. The mixture was incubated at 37°C for about 2 hrs and samples were run on 1.5% (w/v) agarose gel.

Table 8: Components of restriction digestion reaction of vector for confirmation of cloning of insert into the vector:

COMPONENTS	VOLUME REQUIRED
10X Fast Digest Buffer	3 µl
Plasmid	400ng
***Xba*I**	0.5 µl
***Xho*I**	0.5 µl
MQ	Volume makeup up to 30 µl

4. MOBILIZATION OF BINARY CONSTRUCTS INTO *AGROBACTERIUM*

Preparation of *Agrobacterium* competent cells:

To make *Agrobacterium* GV3101 competent cells, strain from glycerol stock was streaked on YEM agar plate having 50 µg/ml rifampicin (selection for GV3101) and incubated at 28°C. When enough growth of the colonies was observed on YEM agar, a single colony was picked for inoculation in 10 ml YEM medium (primary culture) having 50 µg/ml rifampicin and allowed to grow at 28°C with constant shaking at 200 rpm for 36-48 hrs. Further, 1 ml of primary culture was used for inoculating 100 ml of secondary culture (without any antibiotics). Again, it was grown at 28°C with 200 rpm constant shaking until $O.D._{600}$ reached 0.4-0.6. The culture was then kept on ice for 30 mins. *Agrobacterium* cells were then pelleted at 5000 rpm for 10 mins at 4°C in a pre-chilled rotor. Pellet was resuspended in 10 ml of 10 mM $CaCl_2$ solution. It was again kept on ice for 10 mins and centrifuged at 5000 rpm for 10 mins at 4°C. Cells were again resuspended in prechilled and autoclaved 10 mM of $MgCl_2$ solution and incubated on ice for 10 mins. Centrifugation was done again with the same conditions mentioned before. Finally, cells were dissolved in 1.5 ml of pre-chilled 10% glycerol and 100 µl of cells were aliquoted into fresh MCTs. These were frozen in liquid nitrogen and stored at -80°C.

Transformation of *Agrobacterium* with Binary Vector:

Freeze-thaw method was used for the transformation *of Agrobacterium* GV3101 with a binary vector. Competent cells were allowed to thaw in ice for about 10 mins. Following thawing, 500 ng of recombinant constructs were added into competent cells and were mixed by pipetting. Cells were then incubated on ice for 30 mins and freeze shocked in liquid nitrogen for 45 seconds. Immediately after freeze shock, cells were incubated at 37°C for 5 mins, then on ice for 2 mins. After 2 mins, 1 ml of LB broth was added into it and incubated at 28°C overnight with continuous shaking.

Next day, cells were pelleted at 4000 rpm for 15 mins and the supernatant was discarded just by inverting the MCT leaving behind 100 µl of solution for resuspending cells. Cells were plated on LB agar containing 50 µg/ml rifampicin and 50 µg/ml kanamycin. Plates were incubated in dark at 25°C for 2-3 days until clear colonies were observed.

Colony PCR of five random colonies was performed using forward as well as reverse primers of the gene. Positive clones were grown in 5 ml of YEM broth, having 50 µg/ml rifampicin and 50 µg/ml kanamycin, overnight at 28°C with continuous shaking at 200 rpm. Plasmid isolation was performed using FavorGen™ Plasmid DNA Extraction Mini Kit as mentioned before and the eluted plasmid was stored at -20°C.

TRV1 and TRV2 plasmids were back-transformed into DH5α cells and plated on LB agar plates having 50 µg/ml rifampicin and 50 µg/ml kanamycin and incubated at 37°C overnight. This was done to check for the confirmation of mobilization of plasmids into *Agrobacterium*. Again, plasmids were isolated from the back-transformed colonies and restriction digestion was performed to verify the size of the plasmids.

Transient expression of *SlMDB8* and *SlIAA17* in tomato fruits via *Agrobacterium*:

To perform transient expression in tomato fruits, LB-broth (primary culture) having 50 µg/ml rifampicin, 50 µg/ml kanamycin and 50 µg/ml gentamicin were inoculated with *Agrobacterium* positive clones of *SlMBD8* and *SlIAA17* and allowed to grow at 28°C overnight. From overnight grown primary culture, 1% secondary culture was inoculated without any antibiotics and was grown for 7-8 hrs. The O.D. at this point should not exceed 0.6. The culture was then centrifuged at 4000 rpm for 10 mins at 28°C. Pellet cells were resuspended in 10 ml of induction buffer (10 mM MES, 200 mM Acetosyringone, pH 5.5). Again, cells were kept at 28°C for 3 hrs at 50 rpm. Centrifugation was done at 4000 rpm for 10 mins at 28°C. Now, the pellet cells were dissolved in infiltration buffer (5 mM MES, pH 5.5, O.D. =1). This culture was further used for Agroinfiltration. Wild type, mature green (MG) tomatoes were surface sterilized and culture was injected through the carpopodium in tomato fruit using a 1 ml syringe. The needle was inserted 3-4 mm deep into the fruit. Fruits injected with an empty vector only (250 µl TRV1 and 250 µl TRV2) were used as a control for comparative fruit analysis with fruits that were injected with *SlMBD8*-TRV2 (250 µl TRV1 and 250 µl *SlMBD8*-TRV2) and *SlIAA17*-TRV2 (250 µl TRV1 and 250 µl *SlIAA17*-TRV2) modified vector. Fruits were kept at 25°C in the culture room and tissues from each fruit were harvested after Br+7 days. The same procedure was repeated with two more sets of tomato fruits.

5. CLONING OF *SlMBD8* AND *SlIAA17* FOR Y2H ASSAY:

PCR amplification of a specific region of *SlMBD8* and *SlIAA17* for Y2H:

Polymerase chain reaction was carried out to amplify the CDS region of *SlMBD8* and *SlIAA17* using SimpliAmp Thermal Cycler. Concentrations of reagents (dNTP, $MgCl_2$ and cDNA) along with PCR conditions (annealing temperature and extension time) were optimized. Depending upon the melting temperature of the primer pairs, annealing temperature as well as the number of amplification cycles were set up. Two master-mix (Table 9), one for each gene were prepared with Phusion DNA-Polymerase. PCR products after amplification were run on 1% (w/v) agarose gel. Conditions optimized for the amplification of both genes are shown in figure 6.

Table 9: Master-mix setup for PCR amplification of Y2H specific region of *SlMBD8* and *SlIAA17*

REAGENTS	VOLUME REQUIRED FOR *SLMBD8*	VOLUME REQUIRED FOR *SLIAA17*
5X Buffer	4 µl	4 µl
Forward Primer	1 µl	1 µl
Reverse Primer	1 µl	1 µl
dNTPs	0.5 µl	0.5 µl
Ph. DNA Poly.	0.2 µl	0.2 µl
cDNA	0.5 µl	1.5 µl
Mgcl2	-	0.5 µl
MQ	12.8 µl	11.3 µl
Total	20 µl	20 µl

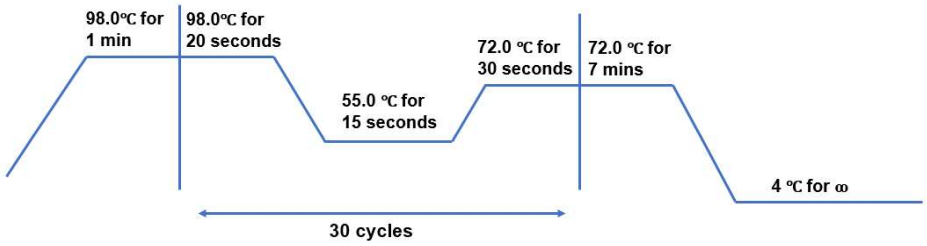


Figure 6: PCR amplification of a specific region of *SlMBD8* and *SlIAA17* for Y2H

Cloning of PCR-amplified *SlMBD8* and *SlIAA17* pENTR /D-TOPO Vector:

pENTR/D-TOPO vector which enables directional cloning was used as an entry vector for primary cloning of *SlMBD8* and *SlIAA17*. After that, the ligation reaction was set up in a 1:1 molar ratio of vector and insert. Following reagents (Table 10) were used for ligation reaction and the reaction mixture was kept at 22°C overnight.

Table 10: Components of Ligation reaction mixture containing vector and insert:

REAGENTS	VOLUME REQUIRED
Salt solution	1 µl
Template	200 ng
pENTR™ /D-TOPO® Vector	200 ng
MQ	2.5 µl
Total	6 µl

Next day, a ligation reaction mixture was used to transform *E. coli* Z-competent cells and plated on LB agar plates having 50 µg/ml kanamycin and incubated at 37°C overnight. Grown colonies on the selection medium were screened by colony PCR and restriction-digestion to confirm recombinants. Colony PCR was performed using M13 forward primer and gene reverse primer.

For restriction digestion plasmids were isolated from an overnight grown culture of positive colonies, confirmed by colony-PCR. For verification of orientation of the *SlMBD8* insert, plasmids were double digested with *Asc*I (site present in vector backbone) and *Pst*I (site present in gene of interest).

In the same manner, *SlIAA17* plasmid was also double digested with *Rru*I and *Asc*I (sites present in vector backbone only). Reaction mixtures were prepared using components as shown in Table 11. Mixtures were then incubated at 37°C for about 2 hrs and assayed on 1% agarose gel.

Table 11: Components of restriction digestion reaction of vector for confirmation of cloning of insert into the vector.

COMPONENTS	AMOUNT REQUIRED FOR *SlMBD8*	AMOUNT REQUIRED FOR *SlIAA17*
10X Fast Digest Buffer	3 µl	3 µl
Plasmid	400 ng	400 ng
***Pst*I**	0.5 µl	-
***Asc*I**	0.5 µl	0.5 µl
***Rru*I**	-	0.5 µl
MQ	Volume make up to 30 µl	Volume make up to 30 µl

Cloning of DNA fragments into pGBKT7 and pGADT7-Destination vectors:

SlMBD8-TOPO and *SlIAA17*-TOPO plasmids isolated above were used for LR recombination reaction with the destination vectors- pGBKT7 and pGADT7 respectively. pGBKT7 vector has kanamycin selection same as pENTR™ /D-TOPO® vector, therefore after LR cloning, we could not differentiate both on 50 µg/ml kanamycin agar-plates. For this reason, the kanamycin site was destroyed in the entry vector by digesting it with *Rru*I enzyme so that it becomes linearized, hence after transformation, only the destination vector can grow on 50 µg/ml kanamycin agar-plates. The reaction setup for *SlMBD8* linearization is given in Table 12.

Table 12: Components of linearization of SlMBD8 with *Rru*I restriction enzyme

COMPONENTS	AMOUNT REQUIRED
10X Fast Digest Buffer	3 µl
Plasmid	400ng
***Rru*I**	1 µl
MQ	Volume makeup up to 30 µl

Linearized *SlMBD8*-TOPO plasmid was eluted from an agarose gel and used for LR reaction with pGBKT7 vector and *SlIAA17*-TOPO plasmid was used for LR reaction

with pGADT7 vector. Table 13 shows the reagents used for LR reaction. The reaction mixture was set up on ice and molar ratio used for vector: insert was 1:1. It was kept at 25^{o}C overnight.

Table 13: Components for LR reaction of linearized *SlMBD8*-TOPO and *SlIAA17*-TOPO with pGBKT7 and pGADT7 vector respectively

COMPONENTS	**AMOUNT REQUIRED FOR *SlMBD8***	**AMOUNT REQUIRED FOR *SlIAA17***
pGBKT7 vector	200 ng	-
pGADT7 vector	-	200 ng
Linearized SlMBD8-TOPO	200 ng	-
SlIAA17-TOPO	-	200 ng
LR clonase	0.5 µl	0.5 µl
TE Buffer (10:1) pH 8.0	Volume makeup up to 10 µl	Volume makeup up to 10 µl

Next day, LR reaction mixture was used to transform *E. coli* Z-competent cells and *SlMBD8* mixture was plated on LB-agar having 50 µg/ml kanamycin (selection for pGBKT7 vector) and *SlIAA17* mixture was plated on LB-agar having 100 µg/ml carbenicillin (selection for pGADT7 vector) and incubated at 37^{o}C overnight.

Verification of colonies carrying recombinant plasmids was done by colony PCR and restriction-digestion. Colony PCR was performed using vector forward primer (BD vector for *SlMBD8* and AD vector for *SlIAA17*) and gene reverse primer. However, for restriction digestion, plasmids were isolated from an overnight grown culture of positive colonies of *SlMBD8-BD* and *SlIAA17-AD* and digested with *Xba*I and *Xag*I respectively. Table 14 shows the components used for digestion reaction.

Table 14: Components of restriction digestion reaction of vector for confirmation of cloning of insert into the vector.

COMPONENTS	**AMOUNT REQUIRED FOR *SlMBD8***	**AMOUNT REQUIRED FOR *SlIAA17***
10X Fast Digest Buffer	3 µl	3 µl
***SlMBD8*-BD plasmid**	400ng	-
***SlIAA17*-AD plasmid**	-	400ng
***Xba*I**	1 µl	-
***Xag*I**	-	1 µl
MQ	Volume makeup to 30 µl	Volume makeup to 30 µl

6. MORPHOLOGICAL ANALYSIS:

Phenotypic Analysis:

Tomato fruits injected with *SlMBD8*-TRV2 and *SlIAA17*-TRV2 constructs as well as control samples were kept at 25^{o}C in the culture room. Fruits were monitored visually for their softness as well as color and photographs were also taken day-to-day wise. Comparative outlines of all the batches are shown in Figures 16, 17 & 18.

Texture Analysis:

Texture profiles of *SlMBD8*-TRV2, *SlIAA17*-TRV2 as well as control (fruits injected with empty vectors) were analyzed using a device, known as Texture Analyzer. The height as well as force per unit weight were calibrated first using a known weight. For all the samples, force per unit distance was recorded and an average of reading was further evaluated and graphs were plotted using Microsoft Excel.

RESULTS AND DISCUSSION

1. EXPERIMENTAL PLAN:

Genes of Interest (GOI):

The present study aimed to characterize the role of *SlMBD8* and *SlIAA17* in tomato fruit development and fruit ripening. The idea that brought us to choose these genes in this study was the interaction between their homologs in Arabidopsis and their expression pattern in tomato plant tissues. Protein-protein interaction studies including Yeast two-hybrid and BiFC had been performed in our laboratory suggesting that homologs of both these proteins interact with each other in Arabidopsis. Expression profiling of both genes from RNA-Seq data available in the public domain shows their maximum level of transcripts in early fruit development stages. Co-expression and decreasing pattern in their expression from 1 cm-fruit to ripened stage directed us to check whether this pattern of expression is associated with the development and ripening of fruit.

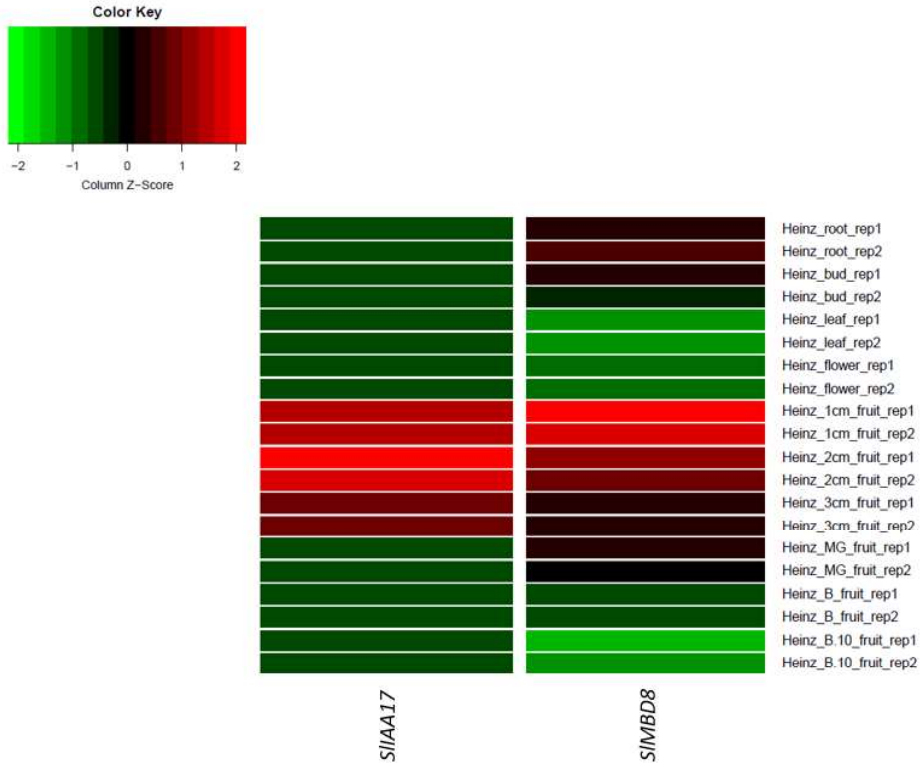


Figure 7: Using publicly available RNA seq data of ten tissues of "Heinz 1706" inbred tomato cultivar, transcript accumulation of *SlMBD8* (*Solyc06g084400*) and *SlIAA17* (*Solyc06g008590*) were analyzed and differential gene expression patterns were studied in different stages of fruit development [Heat map was generated using R-based software and color scale represents $\log_2$ (RPKM/FPKM)]

2. OBJECTIVES:

To find out the interaction between SlMBD8 and SlIAA17 protein and their role in fruit development and fruit ripening, the study was done under two main objectives as follows:

- To silence the expression of *SlMBD8* and *SlIAA17*
- To study the interaction between SlMBD8 and SlIAA17

3. OBJECTIVE 1: To silence the expression of *SlMBD8* and *SlIAA17*

Selection of Vector:

The tool picked for gene silencing in tomato fruits was VIGS and to begin with this objective there was a need for a vector to clone our GOIs: *SlMBD8* and *SlIAA17*. There are many vectors available for performing VIGS in tomato but TRV (Tobacco Rattle Virus) vector was the choice of the vector over other since it has comparably higher silencing efficiency and minimum viral symptoms. Vector design used was adapted in Senthil-Kumar and Mysore, 2014. Figure 8 illustrated the design of the vector:

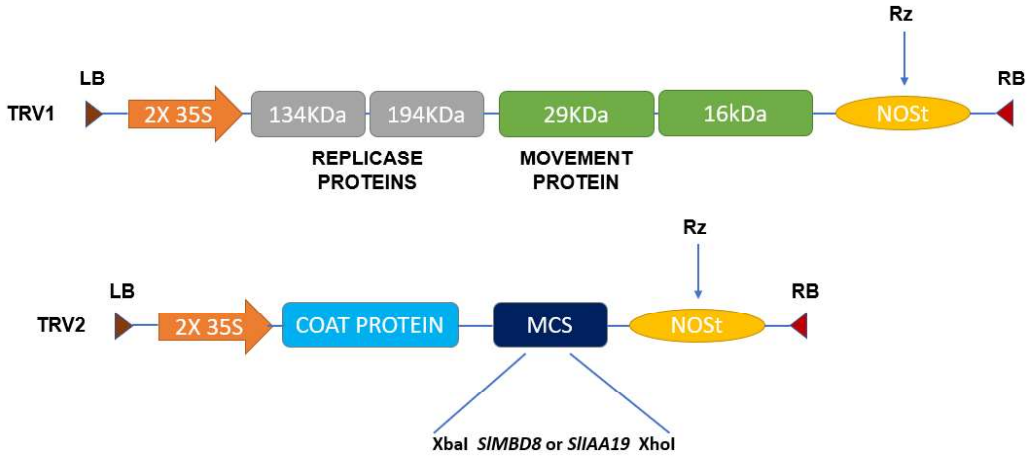


Figure 8: Design of TRV2-*SlMBD8* or TRV2-*SlIAA17* construct (Vector design as adapted from Senthil-Kumar and Mysore, 2014

PCR amplification of a specific region of *SlMBD8* and *SlIAA17* for VIGS:

cDNA used for PCR amplification was borrowed from the lab and reaction conditions were standardized. Primers designed having *Xba*I and *Xho*I restriction sites were used for amplifying VIGS specific region of 300 bp from the cDNA. After standardizing all reaction conditions, clear bands of 300 bp were obtained on an agarose gel (Figure 9), following which the elution was performed.

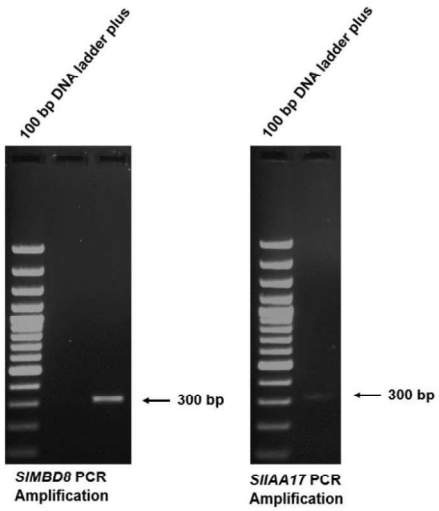


Figure 9: Gel image showing PCR amplification of VIGS specific region of *SlMBD8* and *SlIAA17*

To check the quantity of the eluted DNA, 1 µl of the eluted sample was run on an agarose gel (Figure 10). According to the concentration of the eluted DNA, further cloning was performed.

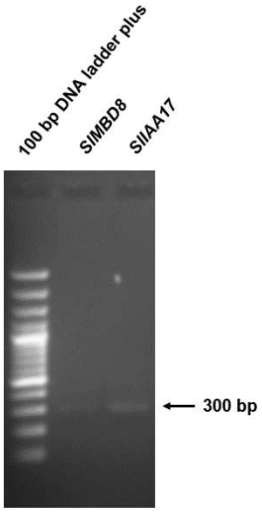


Figure 10: Gel image showing the quantity of the eluted fragments of *SlMBD8* and *SlIAA17*

Cloning of specific fragments of *SlMBD8* and *SlAA17* genes into a primary vector:

Primary cloning of *SlMBD8* and *SlIAA17* DNA fragment was done in a pGEM-T easy vector. Vector has 3' terminal thymidine present on both ends which prevents recircularization of the vector backbone. Since PCR product already had A-overhang (due to amplification via iTaq), it can easily be cloned into vector backbone having T-overhang. Primary cloning was done as it helps in increasing the efficiency of cloning of *SlMBD8* and *SlIAA17* DNA fragment into the final vector, TRV2 (9.2 kb). After ligating GOI in pGEM-T easy vector, it was transformed in DH5α bacterial cells and plated on LB-agar having 100 µg/ml carbenicillin, X-gal and IPTG for selection of recombinants.

The α-peptide coding region of β-galactosidase enzyme has T7 and SP6 RNA polymerase promoters which flank multiple cloning sites (MCS). Therefore, upon insertion of *SlMBD8* and *SlIAA17* DNA fragment into vector backbone, it causes insertional inactivation of α-peptide, hence can be observed directly via blue-white screening. If the gene of interest is inserted into the coding region, the functional β-galactosidase enzyme will not be formed and hence will give white-colored colonies. Therefore, white colonies present were expected to be recombinant colonies. Recombinant colonies were confirmed further via colony PCR as well as restriction digestion analysis.

Colony PCR to confirm recombinant colonies:

To certify efficient cloning, vigorous screening was done. Using gene and vector specific primers, colony PCR was performed, since, it is a rapid and high throughput method to check whether the desired construct is present. Each colony chosen was positive since 300 bp bands were released (Figure 11) in case of both *SlMBD8* and *SlIAA17* cloned vector. This confirmed that colonies were recombinant.

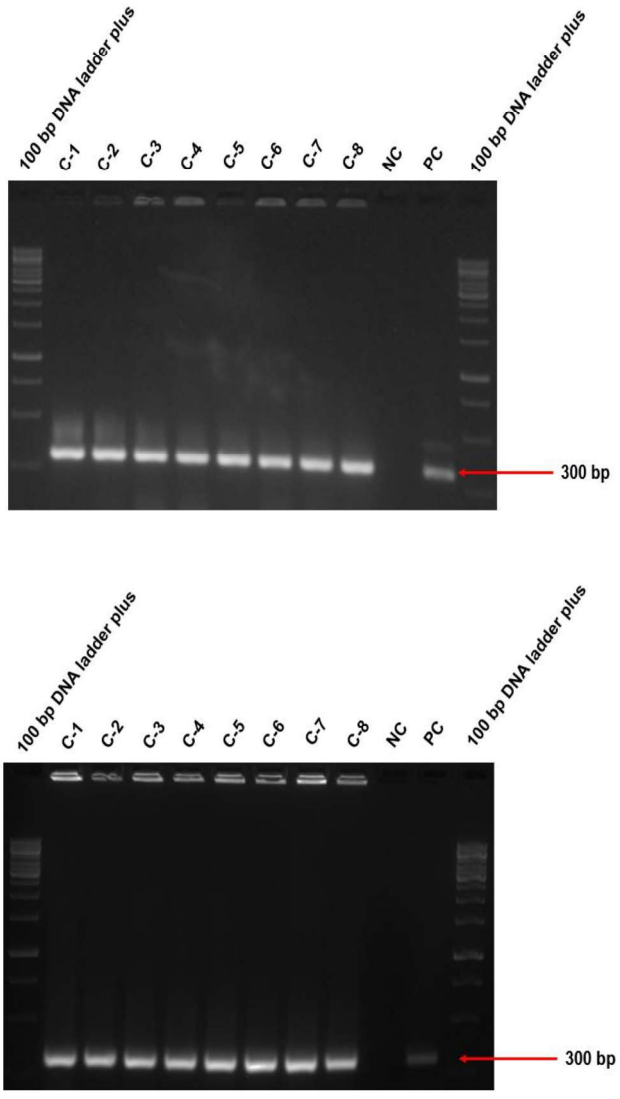


Figure 11: Gel image showing colony PCR of *SlMBD8*-pGEM-T and *SlIAA17*- pGEM-T easy vector respectively (PC: positive control, NC: negative control)

Plasmid isolation from positive colonies identified by colony PCR from preceding step:

To proceed further with downstream experiments, plasmids were required and hence isolated from the overnight grown bacterial culture at 37°C**.** The colonies used to inoculate 10 ml LB-broth having 50 µg/ml kanamycin were colony-1 for *SlMBD8* and colony-2 for *SlIAA17.* Isolated plasmids were sent to CIF for sequencing to check the orientation of the insert into the vector or to check for any abnormality.

Cloning of specific fragments of *SlMBD8* and *SlIAA17* gene into a binary vector:

To proceed with this step, the isolated plasmids were digested with restriction digestion enzymes *Xba*I and *Xho*I which gave a 300 bp long DNA fragment with flanking restriction sequence for *Xba*I and *Xho*I (Figure 12). These fragments were resolved on 2% agarose gel and the desired bands were excised from the gel. The DNA segment excised from the gel was eluted using a DNA gel recovery kit. TRV2 was also treated with the same restriction enzymes to produce complementary ends.

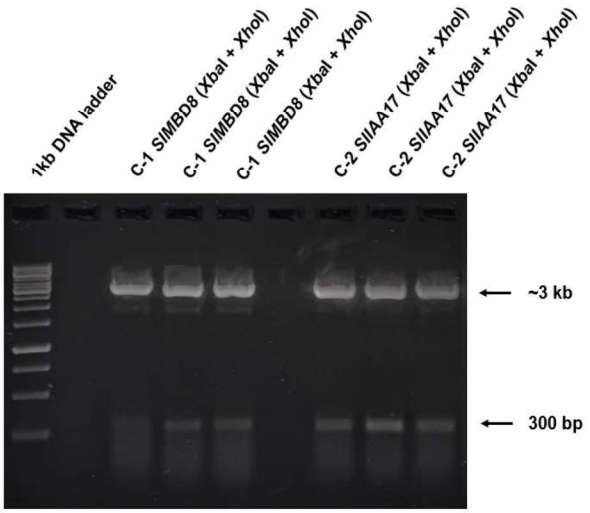


Figure 12: Gel image showing restriction digestion profiles of *SlMBD8* and *SlIAA17* cloned pGEM-T easy vector with enzymes *Xba*I and *Xho*I

Desired DNA fragments were ligated into TRV2 vector using Rapid DNA Ligation Kit and the ligation mixture was used to transform DH5α bacterial cells and the cells were plated on LB-agar containing 50 µg/ml kanamycin antibiotics as a selection for recombinants.

Colony PCR and restriction digestion to confirm recombinant colonies from preceding step:

Recombinant colonies were screen first by colony PCR and then via restriction-based screening of isolated plasmids. Colony PCR was done using gene-specific primers, to confirm recombinant colonies. In the CPCR of *SlMBD8*, each colony chosen was positive whereas, in *SlIAA17,* colony 2 and colony 3 were positive (Figure 13). From these positive colonies, plasmids were isolated. To further validate our transformants, restriction digestion of plasmids of *SlMBD8*-C1 and *SlIAA17*-C2 was performed by using *Xba*I and *Xho*I (Figure 14).

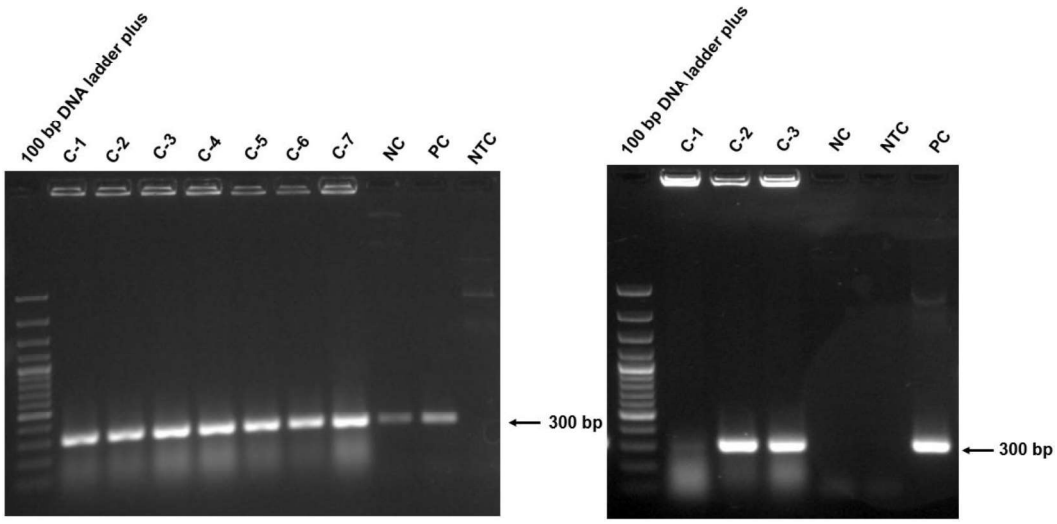


Figure 13: Gel image showing colony PCR of *SlMBD8*-TRV2 and *SlIAA17*-TRV2 vector (PC: positive control, NC: negative control, NTC: no template control)

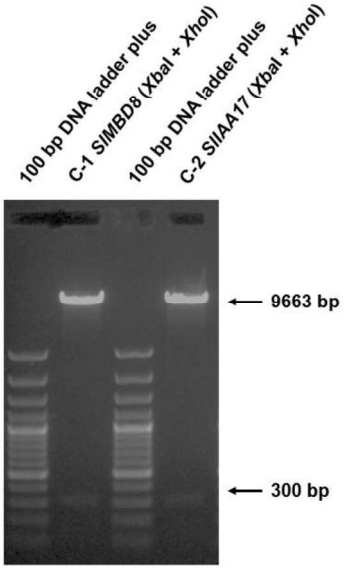


Figure 14: Gel image showing restriction digestion profiles of *SlMBD8* and *SlIAA17* cloned TRV2 vector with enzymes *Xba*I and *Xho*I

Mobilization of binary vectors containing desired DNA segments into *Agrobacterium* strain GV3101:

After successful cloning, *SlMBD8*-TRV2, *SlIAA17*-TRV2, TRV2 vector and TRV1 vector were mobilized into *Agrobacterium* strain GV3101. Colonies appeared on LB-agar supplemented with 50 µg/ml rifampicin and 50 µg/m kanamycin were screened for recombinant colonies by colony PCR. Colony PCR of five random colonies was performed using primers specific to genes. All colonies were positive and 300 bp bands were observed as shown in Figure 15.

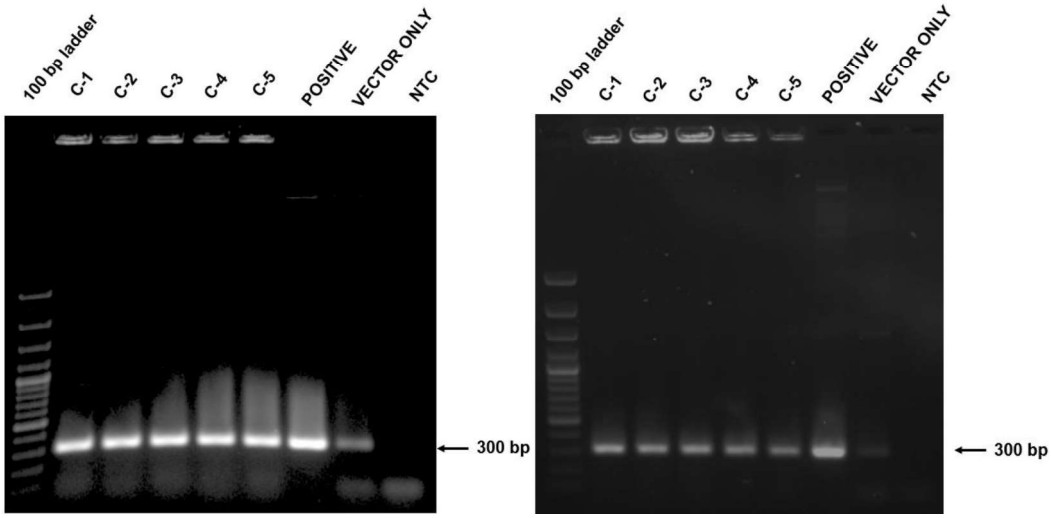


Figure 15: Gel image showing colony PCR of Agrobacterium transformed colonies of *SlMBD8* and *SlIAA17* (NTC: no template control)

4. KEY POINTS FOR MORPHOLOGICAL ANALYSIS OF TOMATO FRUITS:

There are predominately three factors influencing the process of fruit repining, namely, pigment, texture and flavor of the fruit.

The pigment primarily responsible for the deep-red color in red-ripe tomato is lycopene. In MG tomatoes, there is a greater amount of chlorophyll. At the onset of ripening, chlorophyll starts degrading and chloroplasts start converting to chromoplasts as well as there is biogenesis of carotenoids. This accumulation of lycopene provides the characteristic red color to ripe tomatoes.

Texture is one of the important characters to be considered while reviewing tomato fruit ripening. Texture reflects several attributes such as structure of cell wall, water status within the fruit, properties of cuticle and morphology of the fruit (Seymour et al., 2013). There is a modification of the cell wall of tomato during its ripening due to which it becomes soft and firm.

Flavor determining compounds are usually sugar, acids and volatile compounds. During the development of tomato, there is large accumulation of starch but as soon as ripening starts, starch is metabolized into sugar: fructose and glucose. Also, there is an increase in the content of organic acids, malic acid and citric acid, and flavor volatiles at the onset of ripening (Klee & Giovannoni, 2011).

5. VIGS ASSAY:

Based on expression patterns of our GOIs, the experiment was originally planned to be performed in 1 cm fruits. However, due to the COVID impact, the choice of fruit was shifted to fully developed mature green fruits. The assay was performed in mature green fruits of tomato to study the effect of silencing of *SlMBD8* and *SlIAA17* on fruit ripening. Tomato mature green fruits were injected with empty vector TRV2 (used as control) and TRV2 vector having gene-construct (*SlMBD8*-TRV2 and *SlIAA17*-TRV2). This was done in three different batches and analyzed comparatively. Photographs were taken every day until Br+7 days and on the same day, tissues were harvested for further transcript analysis and biochemical analyses like chlorophyll content, lycopene content and estimation of total soluble solids and stored at -80^{o}C. However, the COVID situation made us to put them on hold until the situation normalizes. Therefore, further research should be conducted to see how much percentage of transcripts were actually silenced or what biochemical changes were there in the fruit. From what had been done until now, ripening parameters such as texture/firmness as well as color of the fruits were taken to draw current results.

It was observed that on Br+7th day of fruit ripening, the lycopene content of *SlIAA17* silenced fruits in all three batches was less compared to control (empty vector). On Br+4th day of fruit ripening, clear orangish patches were observed on the *SlIAA17* silenced fruits giving a probable reason that lycopene formation might get hindered in that portion of the fruit. But in *SlMBD8* silenced fruits, lycopene content was almost same as that of control. Also, from the cut portion of the *SlIAA17* silenced fruits, it can be visualized that the pericarp tissue also had relatively less lycopene content than *SlMBD8* and control (Figure 16, 17 & 18). Therefore, *SlIAA17* might be involved in the regulation of lycopene accumulation, therefore its silencing affected the level of lycopene. However, this was not seen in the case of *SlMBD8* silenced fruits. Since we had not performed real-time analysis of the genes, further research is required for verification of results.

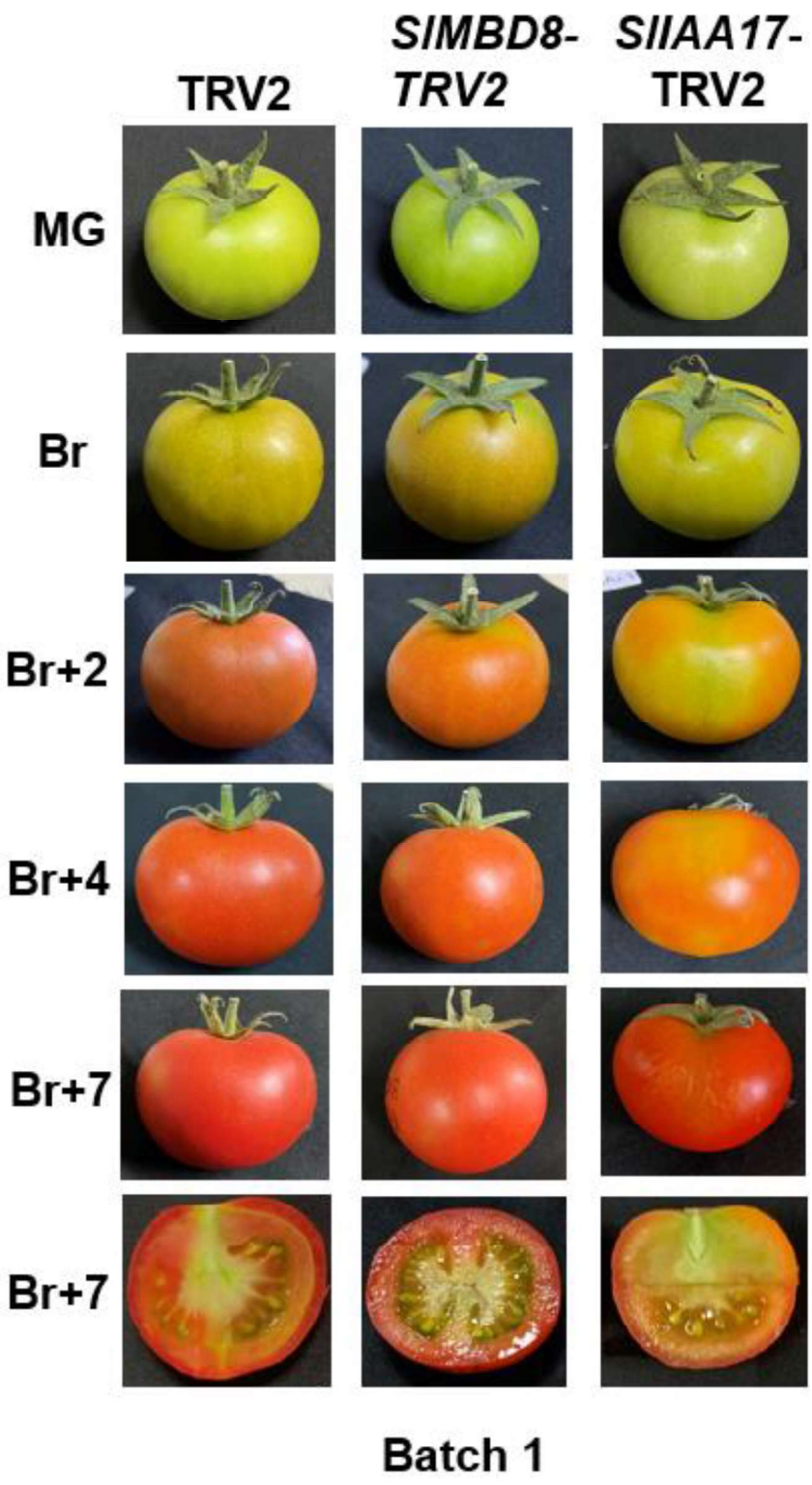


Figure 16: VIGS profile for Batch 1- *SlMBD8* and *SlIAA17* silenced mature green (MG) tomato fruits observed for Br+7 days

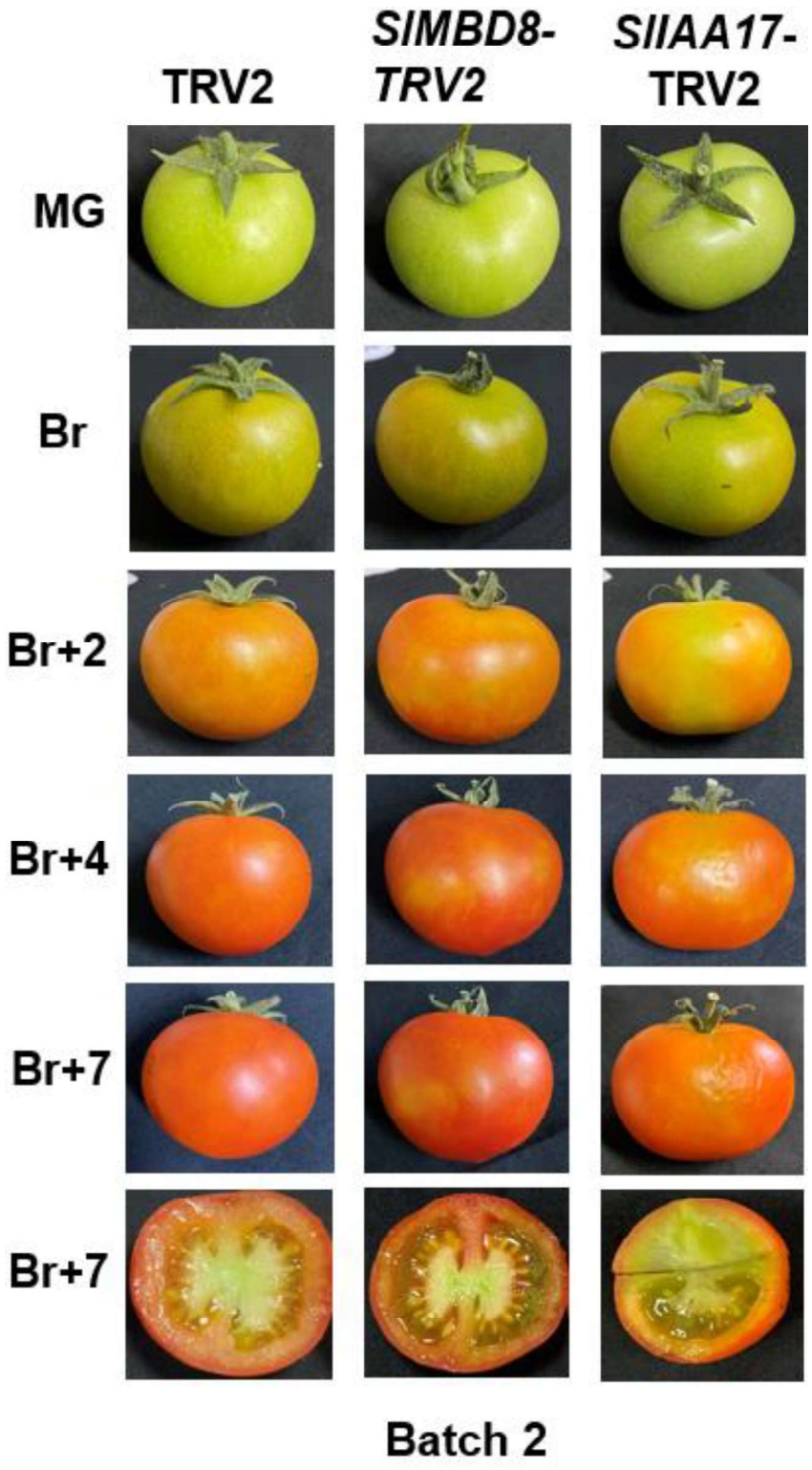


Figure 17: VIGS profile for Batch 2- *SlMBD8* and *SlIAA17* silenced mature green (MG) tomato fruits observed for Br+7 days

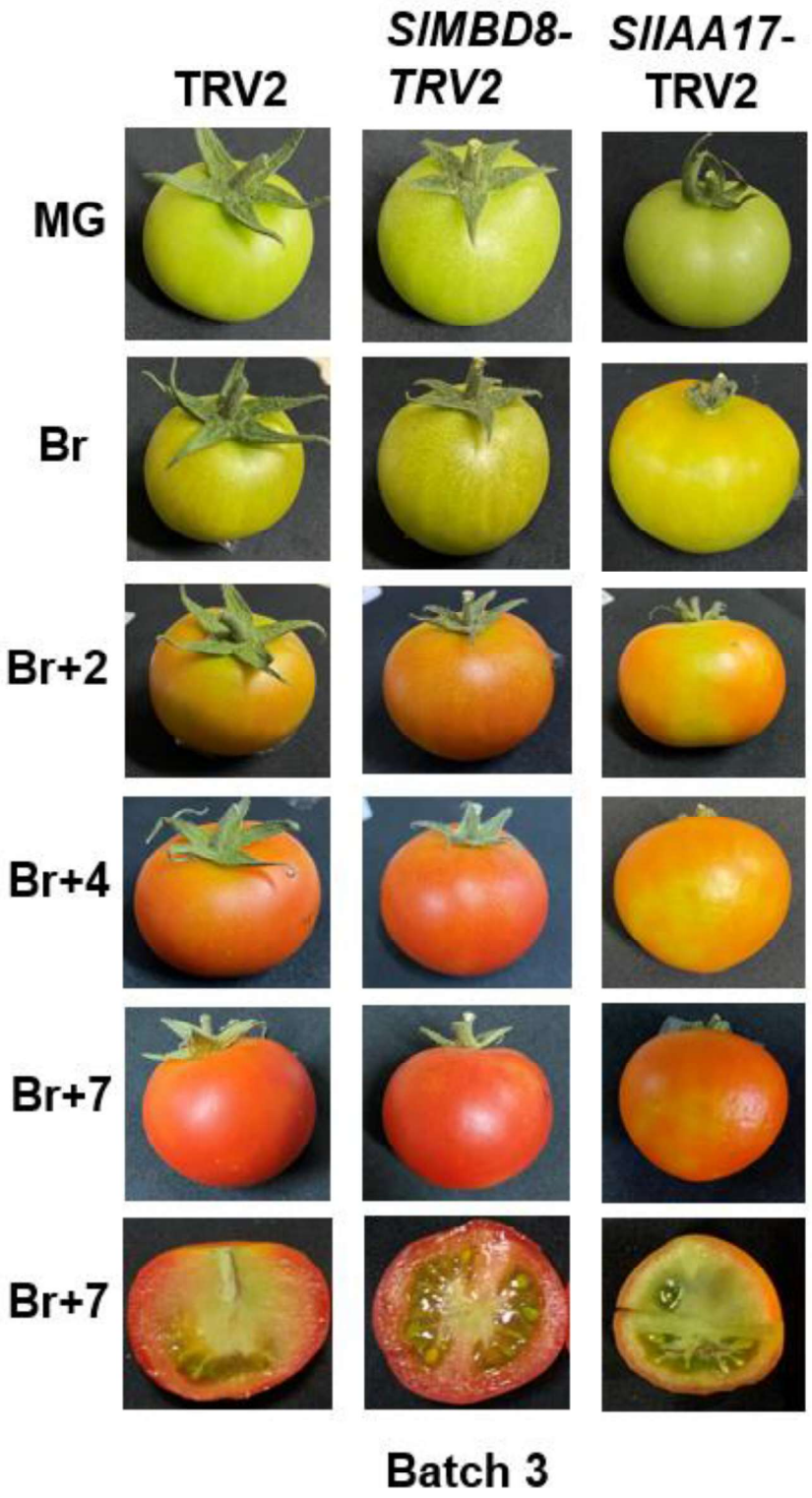


Figure 18: VIGS profile for Batch 3- *SlMBD8* and *SlIAA17* silenced mature green (MG) tomato fruits observed for Br+7 days

6. TEXTURE ANALYSIS:

A penetration test was used for the estimation of firmness by applying the force essential for the probe to penetrate inside the fruit to a specific distance. Firmness is primarily due to change in cell-wall composition. This test has an advantage as there is no such need for samples to be of similar sizes. However, the depth of the penetration differs according to the size as well as firmness of the fruit. Depending upon the firmness, the force required to puncture the fruit varies.

Fruits infiltrated with TRV2 (empty vector) as well as *SlIAA17*-TRV2 showed almost the same level of firmness. However, fruits infiltrated with *SlMBD8*-TRV2 were relatively firm than both control and *SlIAA17*. A probable reason could be that *SlIAA17* doesn't have any effect on cell wall metabolizing enzymes which is why the silencing of *SlIAA17* did not affect the firmness of the fruit. As it is already known that MBDs can act both as repressors as well as activators therefore, silencing of *SlMBD8*, in contrast, increases the firmness of the fruit considering that *SlMBD8* might be involved in regulating the level of cell wall metabolizing enzymes. The comparative analysis of the measured firmness of three fruits is shown in Figure 19.

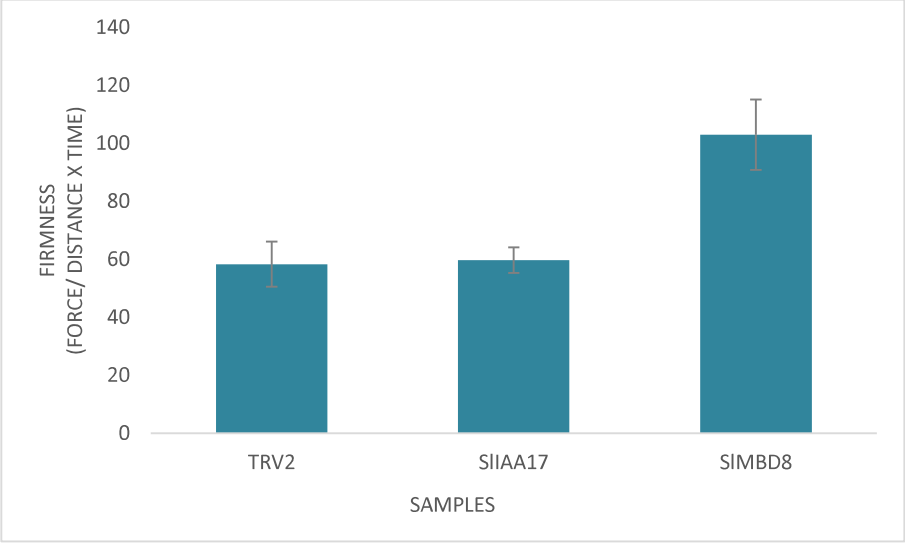


Figure 19: Comparative analysis of firmness analyzed via Texture Analyzer (Appendix-Table 1)

7. **OBJECTIVE 2:** To study the interaction between SlMBD8 and SlIAA17

Selection of Vector:

The yeast two-hybrid system was chosen to study protein-protein interaction between GOI and the binary vectors used for this study were pGBKT7-BD and pGADT7-AD. These both vectors are Gateway cloning vectors that utilize pENTR-D-TOPO as an entry vector. The Vector map of pGBKT7-BD and pGADT7-AD is shown below (created with Snap Gene).

- **pGBKT7-BD:** This vector is designed to use in the Y2H expression system. It expresses proteins that are fused between 1-147 amnio acids of Gal 4 DNA binding domain (DNA-BD). Due to the presence of an ADH1 promoter, fusion proteins are expressed at high rate. It also contains a T7 promoter. For its selection in *E. coli* and yeast, kanamycin and *TRP1* nutritional markers are present respectively (Figure 20).

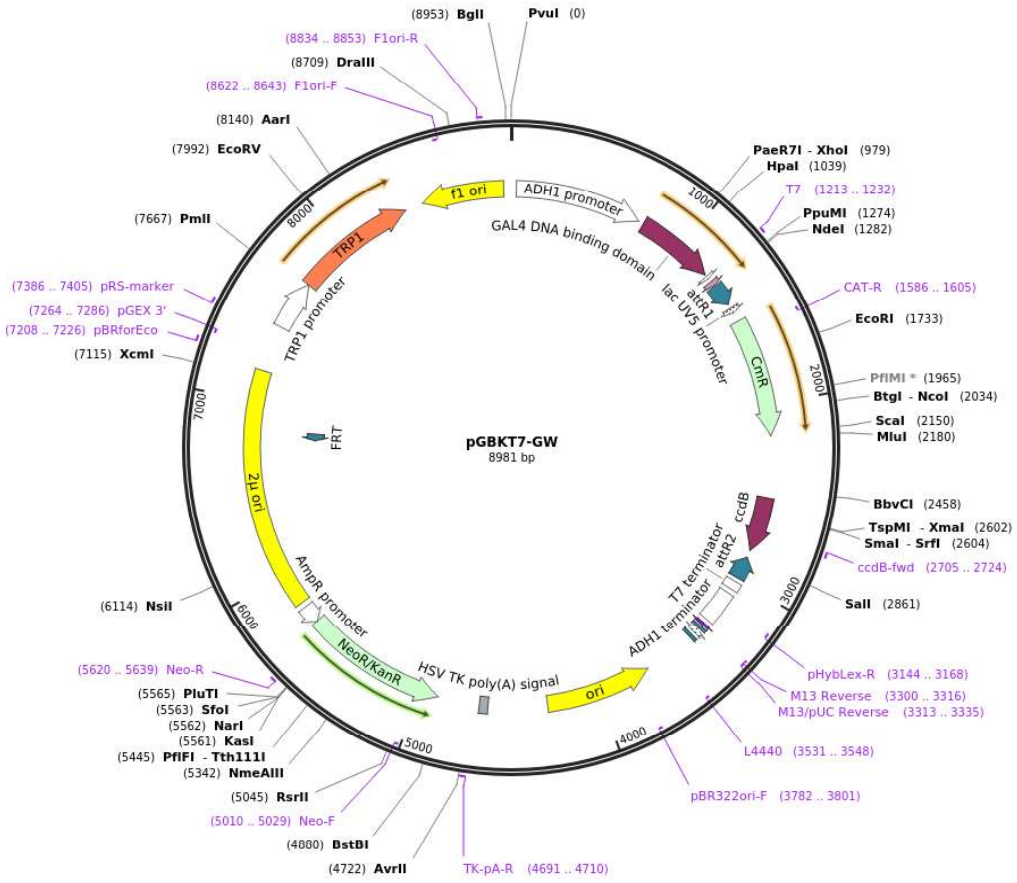


Figure 20: Vector map of pGBKT7 Gateway vector

- **pGADT7-AD:** This vector is also designed to use in the Y2H expression system. It expresses proteins that are fused between 768-881 amino acids of Gal 4 activation domain (AD). It also expresses protein at high level due to the presence of an ADH1 promoter. For its selection in *E. coli* and yeast, ampicillin and *LEU2* nutritional markers are present respectively (Figure 21).

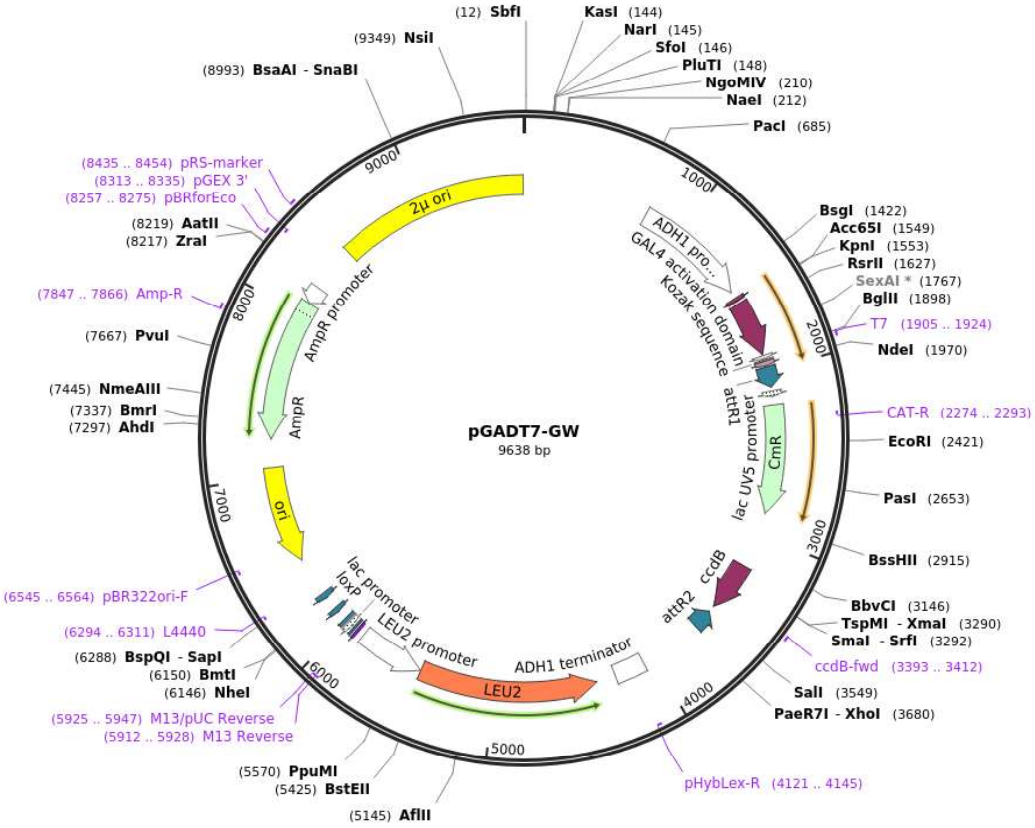


Figure 21: Vector map of pGADT7 Gateway vector

PCR amplification of a specific region of *SlMBD8* and *SlIAA17* for Y2H:

For PCR amplification, cDNA was borrowed from the lab and reaction conditions were standardized. While designing primers, CACC sequence was added at the 5' end of the forward primer to do TOPO-based cloning. After standardizing all reaction conditions, clear bands of 504 bp and 624bp were obtained on 1% (w/v) agarose gel (Figure 22).

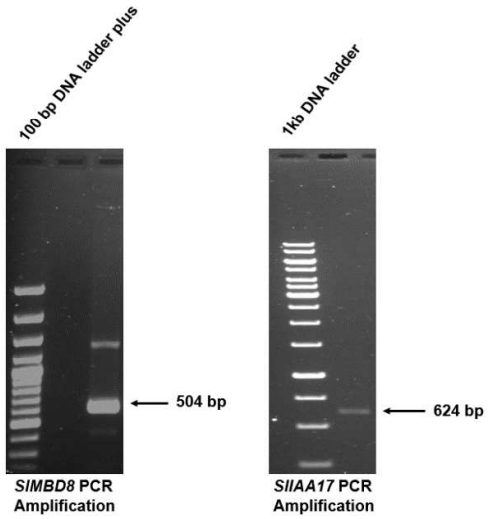


Figure 22: Gel image showing PCR amplification of Y2H specific region of *SlMBD8* and *SlIAA17*

To check the quantity of the eluted DNA, 1µl of the eluted sample was run on an agarose gel (Figure 23). According to the concentration of the eluted DNA, further cloning was performed.

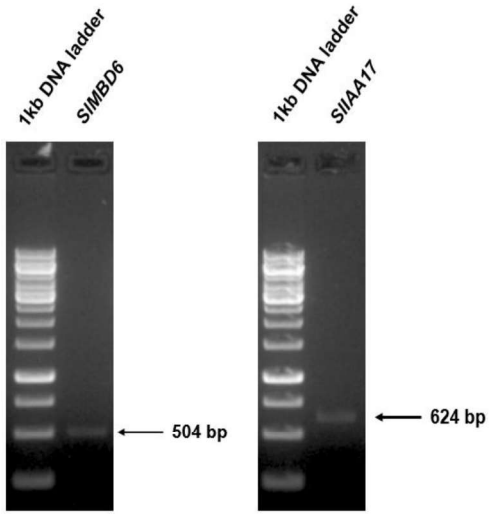


Figure 23: Gel image showing the quantity of the eluted fragments of *SlMBD8* and *SlIAA17*

Cloning of specific region of *SlMBD8* and *SlIAA17* into TOPO entry vector:

Primary cloning of *SlMBD8* and *SlIAA17* was done into pENTR™ TOPO® vector. This vector was used as it facilitates rapid as well as directional TOPO cloning of amplified product for entry into Gateway system. It is a linearized vector having GTGG overhang at one end and a blunt end at another end because of the presence of Topoisomerase I enzyme which has a dual function. This enzyme can act both as a restriction enzyme as well as a ligase. It can cleave Topo vector producing GTGG overhang at 3' end. At the 5' end of the PCR product, there was a CACC sequence present (added in the forward primer). GTGG overhangs invade into the CACC sequence of the PCR product and can rejoin with the help of Topoisomerase I.

Colony PCR to confirm recombinant colonies:

To certify efficient cloning, screening was done. Using specific primers for gene and vector, colony PCR was performed to check whether the desired construct is present. In case of *SlMBD8*, colonies 1, 3, 5, 6 and 8 were positive whereas, in case of *SlIAA17*, colonies 1, 3 and 6 were positive (Figure 24). This verified recombinant colonies.

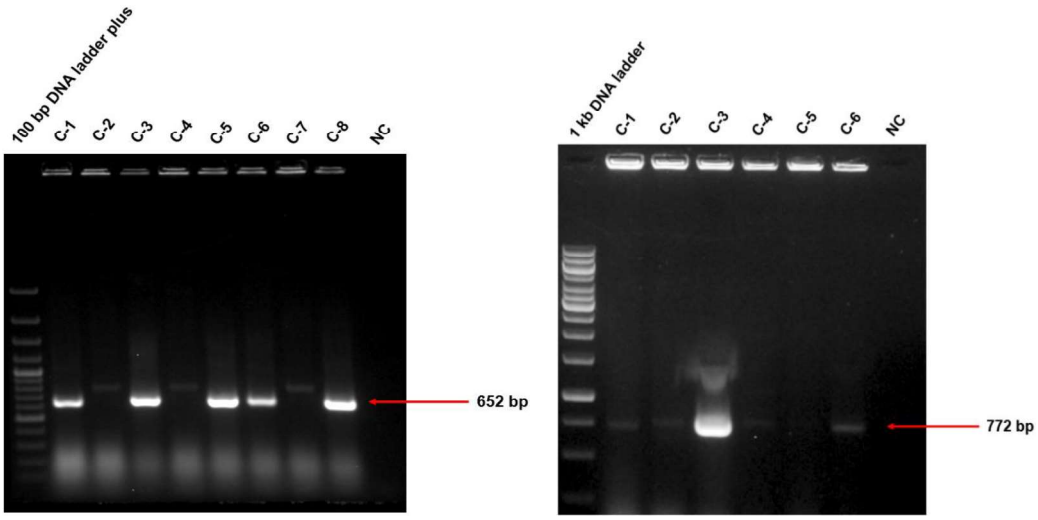


Figure 24: Gel image showing colony PCR of *SlMBD8*-TOPO and *SlIAA17*-TOPO vector

Plasmid isolation from positive colonies identified by colony PCR from preceding step:

To proceed further with downstream experiments, the plasmids were required and hence isolated from the overnight grown bacterial culture at 37°C. The colonies used to inoculate 10 ml LB-broth having 50 µg/ml kanamycin were colony-3 for *SlMBD8* and colony-1, 3 and 6 for *SlIAA17*. Isolated plasmids were sent to CIF for sequencing to check the orientation of the insert into the vector or to check for any abnormality. After digesting *SlMBD8*-TOPO with *Pst*I and *Asc*I restriction enzymes, two bands of sizes 312 bp and 2772 bp were released in case of colony-3 as expected. However, digesting *SlIAA17*-TOPO with *Rru*I and *Asc*I restriction enzymes, two bands of sizes 359 bp and 2845 bp were released as expected (Figure 25).

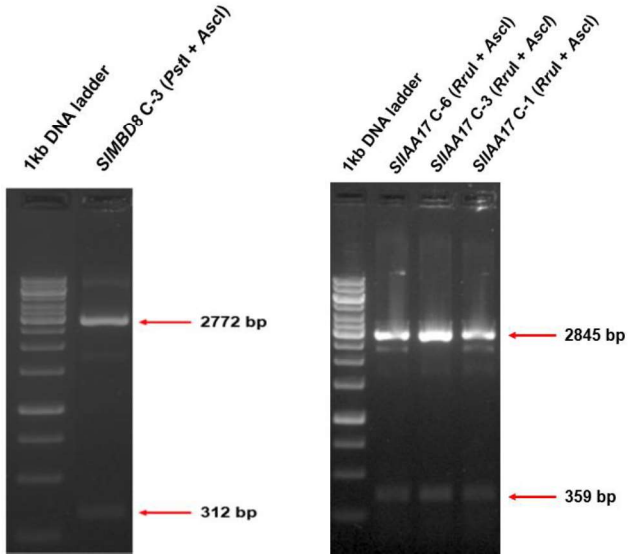


Figure 25: Gel image showing restriction digestion profiles of *SlMBD8* and *SlIAA17* cloned TOPO-vector with *Pst*I + *Asc*I and *Rru*I + *Asc*I enzymes respectively

Cloning of specific region of *SlMBD8* and *SlIAA17* into a binary vector:

SlMBD8-TOPO (linearized with *Rru*I) and *SlIAA17*-TOPO plasmids were used for LR recombination reaction with the destination vectors- pGBKT7-BD and pGADT7-AD respectively. Ligated products were used to transform DH5-alpha bacterial competent cells. Colonies appeared on LB-agar supplemented with 50 µg/ml kanamycin for *SlMBD8*-pGBKT7 and 100 µg/ml carbenicillin for *SlIAA17*-pGADT7 were screened for recombinant colonies.

Colony PCR to confirm recombinant colonies:

It was done using gene-specific primers for validation of efficient cloning at each step. All four colonies of *SlMBD8*-pGBKT7 and two colonies of *SlIAA17*-pGADT7 released a band of 504 and 624 bp respectively (Figure 26).

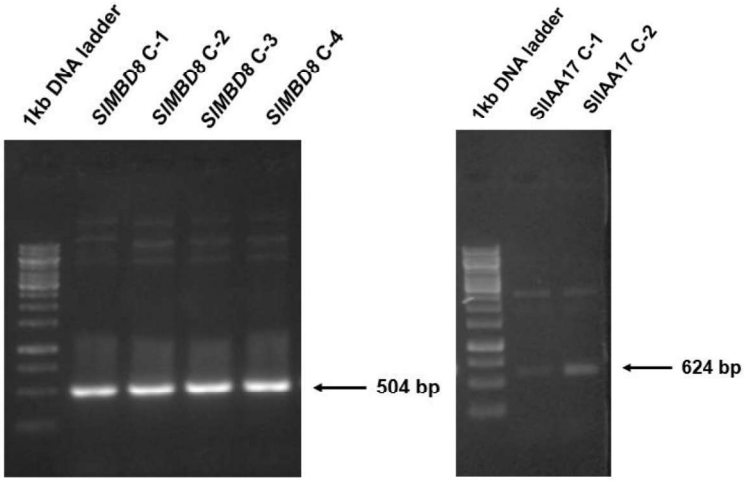


Figure 26: Gel image showing colony PCR profiles of *SlMBD8*-pGBKT7 cloned and *SlIAA17*-pGADT7 cloned vector

Plasmid isolation from positive colonies identified by colony PCR from preceding step:

To proceed further with downstream experiments, plasmids were required and hence isolated from the overnight grown bacterial culture at 37°C. Positive colonies of each were used to inoculate 10 ml LB-broth with 50 µg/ml kanamycin (*SlMBD8*-pGBKT7) and 100 µg/ml carbenicillin (*SlIAA17*-pGADT7). Isolated plasmids were sent to CIF for sequencing to check the orientation of the insert into the vector or to check for any abnormality.

Restriction digestion to confirm recombinant colonies from colony PCR:

Restriction digestion was performed using *Xba*I and *Xag*I restriction enzymes, sites present in pGBKT7 and pGADT7 vector respectively. pGBKT7 vector is of 8993 bp and has two *Xba*I restriction sites at 7803 and 6815 bp. Therefore, after digesting four random colonies of *SlMBD8*-pGBKT7, ~988 bp band was expected and released. Similarly, pGADT7 vector is of 9677 bp and has two *Xag*I restriction sites at 5233 and 4543 bp, which after digesting two random colonies of *SlIAA17*-pADGT7 gave a band of ~690 bp as expected (Figure 27).

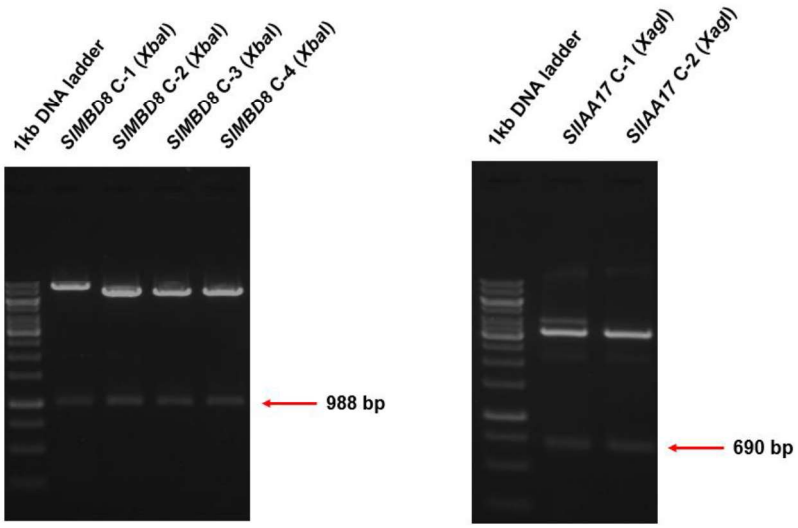


Figure 27: Gel image showing restriction digestion profiles of *SlMBD8*-pGBKT7 cloned and *SlIAA17*-pGADT7 cloned vector with *Xba*I and *Xag*I enzymes respectively

After cloning, we planned to study interaction between SlMBD8 and SlIAA17 but due to the pandemic situation, we are on hold until now.

8. DISCUSSION:

Ideally, the first objective of our study was to characterize protein-protein interaction between SlMBD8 & SlIAA17 and the second objective was to silence both genes in tomato fruits. If any interaction would have been observed in the Y2H assay, then to silence both genes together in tomato fruits to see effects in fruit ripening. We started our silencing experiment as well as the Y2H assay simultaneously. However, until the COVID situations have fallen worse, only cloning of *SlMBD8* and *SlIAA17* into their respective binary vector pGBKT7 and pGADT7 could be performed. The clue which led us to perform a protein-protein interaction study is the identification of the interaction between homologs of both the genes, interacting in *Arabidopsis*, assayed by Y2H and BiFC experiments. After this study, experiments that were supposed to be performed are BiFC and CoIP to validate the interaction *in-vitro* and *in-vivo* respectively.

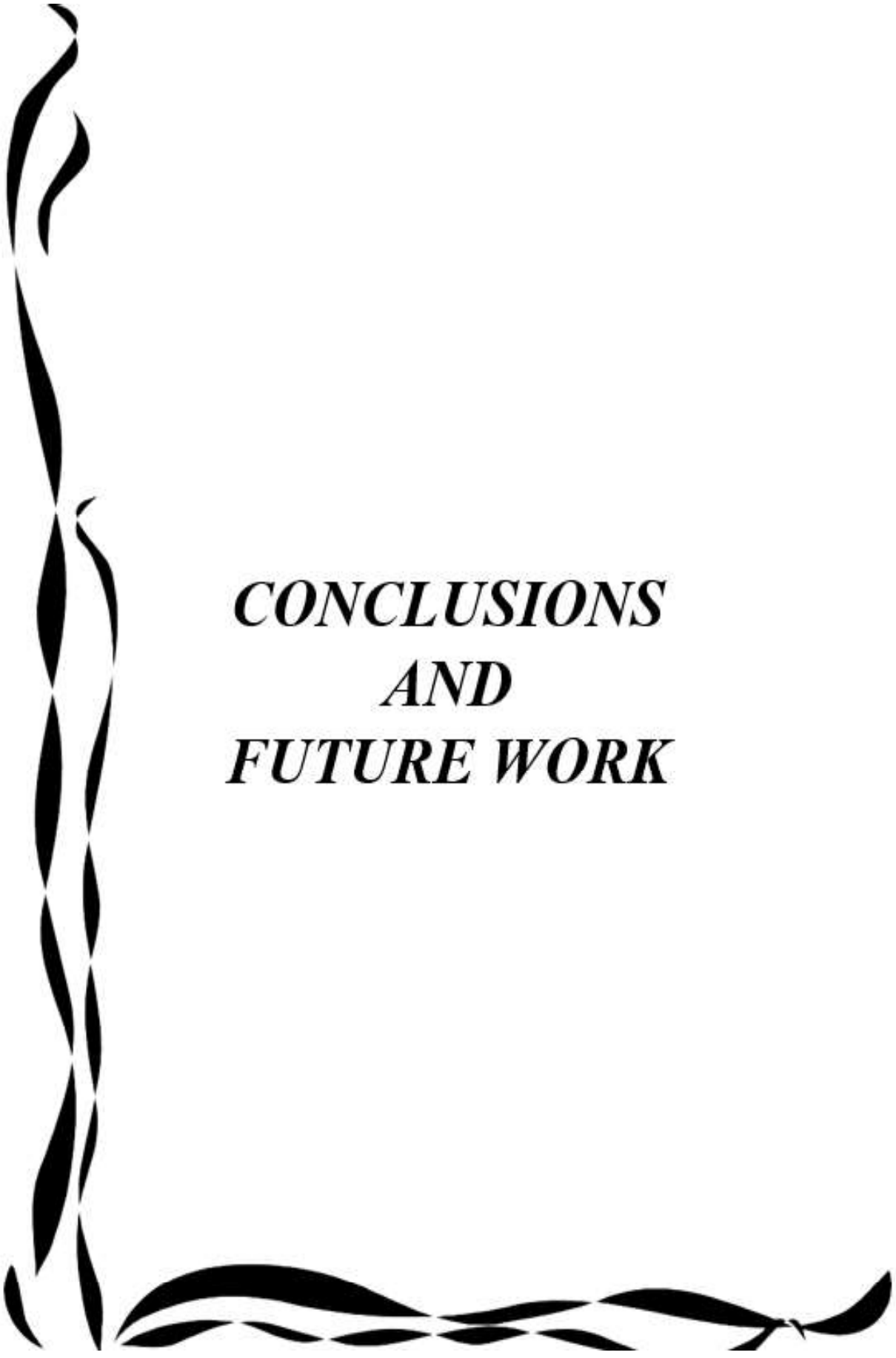
To gain primary insight into the involvement of GOI in fruit ripening, we tried to silence their expression transiently by injecting *Agrobacterium* carrying VIGS-based DNA segments specific to respective targets which were transcripts of *SlMBD8* and *SlIAA17*. As mentioned before, transcript level of both genes is maximum during the early stages of fruit development. However, the unavailability of such fruits led us to perform the assay in MG green fruits. So expected results were if genes are involved in fruit ripening, then either factor of ripening namely, firmness, pigment or flavor should be affected.

By silencing *SlMBD8*, fruits were relatively firm than TRV2 (empty vector) and *SlIAA17* silenced fruit. However, *SlIAA17* silenced fruit and control fruits showed the same level of firmness. It can be interpreted that *SlMBD8* may be involved in regulating the firmness of a fruit or it can potentially target genes which are involved in remodeling of cell wall during fruit maturation. However, further validation at the transcript level of such depolymerizing enzymes must be done through qPCR to rule out any off-target analysis.

Lycopene content of *SlIAA17*-silenced fruit is lower than control and *SlMBD8* silenced fruit. Some orangish patches were also seen in *SlIAA17* silenced fruit that might be due to inhibition of lycopene accumulation. It can be interpreted that *SlIAA17* might be involved in the regulation of lycopene accumulation, therefore silencing of the former affected the level of lycopene.

However, the study by Su et. al. in 2014 showed the role of *SlIAA17* in fruit enlargement by raising *SlIAA17* RNAi lines. They also mentioned that the ploidy level

was increased in *SlIAA17* down-regulated plants. We have collected tissue of the fruits showing phenotypes for isolation of RNA. Transcription studies are required to draw any conclusion from these observed phenotypes.



CONCLUSIONS AND FUTURE WORK

Ripening is one of the most coordinated and highly sophisticated known developmental process. This process is regulated by extremely organized network comprises of transcriptional regulators, epigenomic modifications, hormones and other regulatory elements at multiple levels (Chen et al., 2020). Hence it is essential to study some aspects of this network for the further improvement of its nutritional values for the betterment of mankind. A hormonally controlled molecular network underlying transcriptional and epigenetic regulation of fruit ripening and development is of great importance. Present study was carried out to find-out the position of the SlMBD8 and SlIAA17 in this molecular network that operates epigenetically under the influence of plant hormones, especially auxin.

In the present study, VIGS was used to transiently silence *SlMBD8* and *SlIAA17* to characterize their function in tomato fruit ripening. The potential of this technique (VIGS) to cause systemic silencing could enables us to extend the experiment on-vine for more precise and accurate results in future. For better results, stable transformation experiments for *SlMBD8* and *SlIAA17* overexpression as well as silencing should be performed. For example, phenotypes due to silencing of *SlMBD8* and *SlIAA17* confers to delay in fruit ripening and therefore, helps in improving the shelf life as well as nutritional quality of tomato.

Synthesis and accumulation of secondary metabolites related to flavor of fruit provide great benefits to human health. Functional studies of genes related to biosynthesis, accumulation and metabolism can be of great interest in improving the quality of the fruit.

The second objective of this study was to characterize protein-protein interaction using yeast two-hybrid assay which can helps us to draw an outline about the mechanism involved in the regulation of tomato fruit ripening. Further, BiFC and Co-IP can be the preferred experiments to validate the identified protein-protein interaction *in-vivo*.

REFERENCES

Assaf, Z. and Gideon, G. (2003). Characterization of Arabidopsis thaliana methyl-CpG-binding domain (MBD) proteins. Plant J. **34**: 565–572.

Audran-Delalande, C., Bassa, C., Mila, I., Regad, F., Zouine, M. and Bouzayen, M. (2012). Genome-wide identification, functional analysis and expression profiling of the Aux/IAA gene family in tomato. Plant and Cell Physiology, **53**: 659-672.

Bassa, C., Mila, I., Bouzayen, M. and Audran-Delalande, C., (2012). Phenotypes associated with down-regulation of Sl-IAA27 support functional diversity among Aux/IAA family members in tomato. Plant and Cell Physiology, **53**: 1583-1595.

Bennypaul, Harvinder S., Jasdeep S. Mutti, Sachin Rustgi, Neeraj Kumar, Patricia A. Okubara, and Kulvinder S. Gill. (2012). Virus-induced gene silencing (VIGS) of genes expressed in root, leaf, and meiotic tissues of wheat. Functional & integrative genomics, **12**: 143-156.

Bergougnoux, V. (2014). The history of tomato: from domestication to biopharming. Biotechnology advances, **32**: 170-189.

Brückner, A., Polge, C., Lentze, N., Auerbach, D., & Schlattner, U. (2009). Yeast two-hybrid, a powerful tool for systems biology. International journal of molecular sciences, **10**: 2763-2788.

Chen, T., Qin, G., & Tian, S. (2020). Regulatory network of fruit ripening: current understanding and future challenges. New Phytologist, **228**: 1219-1226.

Du, Q., Luu, P. L., Stirzaker, C., & Clark, S. J. (2015). Methyl-CpG-binding domain proteins: readers of the epigenome. Epigenomics, **7**: 1051-1073.

Finnegan, E. J., & Kovac, K. A. (2000). Plant DNA methyltransferases. Plant Gene Silencing, 69-81.

Fu, Da-Qi, Ben-Zhong Zhu, Hong-Liang Zhu, Wei-Bo Jiang, and Yun-Bo Luo. (2005). Virus-induced gene silencing in tomato fruit. The Plant Journal, **43**: 299-308.

Gallusci, P., Hodgman, C., Teyssier, E., & Seymour, G. B. (2016). DNA methylation and chromatin regulation during fleshy fruit development and ripening. Frontiers in plant science, **7**: 807.

Grafi, G., Zemach, A., & Pitto, L. (2007). Methyl-CpG-binding domain (MBD) proteins in plants. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression, **1769**: 287-294.

Kant, R., & Dasgupta, I. (2019). Gene silencing approaches through virus-based vectors: speeding up functional genomics in monocots. Plant molecular biology, **100**: 3-18.

Klee, H.J. and Giovannoni, J.J., (2011). Genetics and control of tomato fruit ripening and quality attributes. Annual review of genetics, **45**: 41-59.

Kumar, R., Agarwal, P., Pareek, A., Tyagi, A. K., & Sharma, A. K. (2015). Genomic survey, gene expression, and interaction analysis suggest diverse roles of ARF and Aux/IAA proteins in Solanaceae. Plant molecular biology reporter, **33**: 1552-1572.

Li, Yuxiang, Heng Deng, Min Miao, Huirong Li, Shengxiong Huang, Songhu Wang, and Yongsheng Liu (2016). Tomato MBD 5, a methyl CpG binding domain protein, physically interacting with UV-damaged DNA binding protein-1, functions in multiple processes. New Phytologist, **210**: 208-226.

Luo, J., Zhou, J. J., & Zhang, J. Z. (2018). Aux/IAA gene family in plants: molecular structure, regulation, and function. International journal of molecular sciences, **19**: 259.

Matzke, M.A. and Mosher, R.A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. Nature Reviews Genetics, **15**: 394-408.

Mauxion, J. P., Chevalier, C., & Gonzalez, N. (2021). Complex cellular and molecular events determining fruit size. Trends in Plant Science.

Nan, X., Ng, H.-H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature **393**: 386.

Paiano, A., Margiotta, A., De Luca, M., & Bucci, C. (2019). Yeast Two-Hybrid Assay to Identify Interacting Proteins. Current protocols in protein science, **95**: 70.

Pandey, Prachi, Muthappa Senthil-Kumar, and Kirankumar S. Mysore (2015). Advances in plant gene silencing methods. Plant gene silencing, 3-23.

Parida, A. P., Raghuvanshi, U., Pareek, A., Singh, V., Kumar, R., & Sharma, A. K. (2018). Genome-wide analysis of genes encoding MBD domain-containing proteins from tomato suggest their role in fruit development and abiotic stress responses. Molecular biology reports, **45**: 2653-2669.

Pesaresi, P., Mizzotti, C., Colombo, M., & Masiero, S. (2014). Genetic regulation and structural changes during tomato fruit development and ripening. Frontiers in Plant Science, **5**: 124.

Quinet, M., Angosto, T., Yuste-Lisbona, F. J., Blanchard-Gros, R., Bigot, S., Martinez, J. P., & Lutts, S. (2019). Tomato fruit development and metabolism. Frontiers in plant science, **10**: 1554.

Rogg, L.E., Lasswell, J. and Bartel, B., (2001). A gain-of-function mutation in IAA28 suppresses lateral root development. The Plant Cell, **13**: 465-480.

Senthil-Kumar, M. and Mysore, K.S., (2014). Tobacco rattle virus–based virus-induced gene silencing in Nicotiana benthamiana. Nature protocols, **9**: 1549-1562.

Seymour, G.B., Chapman, N.H., Chew, B.L. and Rose, J.K., (2013). Regulation of ripening and opportunities for control in tomato and other fruits. Plant Biotechnology Journal, **11**: 269-278.

Souza, A. J. D., Mendes, B. M. J., & Mourão Filho, F. D. A. A. (2007). Gene silencing: concepts, applications, and perspectives in woody plants. Scientia Agricola, **64**: 645-656.

Su, L., Bassa, C., Audran, C., Mila, I., Cheniclet, C., Chevalier, C., Bouzayen, M., Roustan, J.P. and Chervin, C., (2014). The auxin Sl-IAA17 transcriptional repressor controls fruit size via the regulation of endoreduplication-related cell expansion. Plant and Cell Physiology, **55**: 1969-1976.

Su, L. Y., Audran, C., Bouzayen, M., Roustan, J. P., & Chervin, C. (2015). The Aux/IAA, Sl-IAA17 regulates quality parameters over tomato fruit development. Plant signaling & behavior, **10**: 1001.

Tang, D., Gallusci, P., & Lang, Z. (2020). Fruit development and epigenetic modifications. New Phytologist, **228**: 839-844.

Wang, C., Wang, C., Zou, J., Yang, Y., Li, Z. and Zhu, S., (2019). Epigenetics in the plant–virus interaction. Plant cell reports, 1-8.

Wang, H., Jones, B., Li, Z., Frasse, P., Delalande, C., Regad, F., Chaabouni, S., Latche, A., Pech, J.C. and Bouzayen, M., (2005). The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. The Plant Cell, **17**: 2676-2692.

Wu, J., Peng, Z., Liu, S., He, Y., Cheng, L., Kong, F., Wang, J. and Lu, G., (2012). Genome-wide analysis of Aux/IAA gene family in Solanaceae species using tomato as a model. Molecular Genetics and Genomics, **287**: 295-311.

Wu, L., Tian, Z. and Zhang, J., (2018). Functional dissection of auxin response factors in regulating tomato leaf shape development. Frontiers in plant science, **9**: 957.

Zemach, A. and Grafi, G., (2007). Methyl-CpG-binding domain proteins in plants: interpreters of DNA methylation. Trends in plant science, **12**: 80-85.

APPENDIX

TABLE 1: Calculation for deducing firmness (force/distance X time)

SAMPLE	BATCH NUMBER	AVERAGE FORCE (F)	AVERAGE DISTANCE (D)	AVERAGE TIME (T)	AVERAGE FIRMNESS (F/D X T)	STANDARD DEVIATION
TRV2	BATCH-1	752.8	35.89	2.75	58.299	7.785
	BATCH-2	772.36	32.94	2.16		
	BATCH-3	889	32.51	2.42		
SlMBD8	BATCH-1	730.16	33.07	5.29	102.922	12.168
	BATCH-2	841.46	31.023	3.50		
	BATCH-3	796.13	27.19	3.30		
SlIAA17	BATCH-1	796.4	33.58	2.42	59.68	4.42
	BATCH-2	711.6	31.05	2.47		
	BATCH-3	902.3	29.94	2.15		

Comparative *in-silico* analysis of PIFs in *Arabidopsis* and Rice and *in-silico* based functional characterization of *TaPIF4*

Thesis Submitted to the University of Delhi

in the Partial fulfilment for the Degree of

Master of Science

in

Plant Molecular Biology and Biotechnology

2021

Karishma Sahu

Department of Plant Molecular Biology

University of Delhi South Campus

New Delhi 110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Comparative *in-silico* analysis of PIFs in *Arabidopsis* and Rice and *in-silico* based functional characterization of *TaPIF4***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 28 JUNE 2021

Signature of the Candidate

(**KARISHMA SAHU**)

Signature of the Supervisor

(**Prof. Paramjit Khurana)**

Signature of the HOD

(**Prof Anil Grover**)

Benito Juarez Road, New Delhi. 110021, **India**. Tel. 24111208, Fax:24110669, Website: www.dpmb.ac.in

Preface

The work presented in the thesis gives an overall idea about the various PIFs that are reported and their function. The presence of PIFs have been reported way back in 1998 with the discovery of PIF3 in *Arabidopsis*. Since then, numerous researches have been done and the all PIFs of *Arabidopsis* are well characterized. This report further discusses the various PIFs that have been discovered in dicots (*Arabidopsis*, tomato, apple, grapes), monocots (rice and maize), mosses (*Physcomitrella*) and liverworts (*Marchentia*). Further in the thesis we discuss briefly about the various functions of PIFs of *Arabidopsis* and rice.

A comparative analysis using phylogenetic tree, promoter analysis and gene expression in various tissue was performed for *Arabidopsis* and rice. This analysis gives an idea about how the genes get segregated in the course of evolution, yet sequence of genes gets conserved and show similarity in there functioning. Promoter analysis further gives clue about the regulation in biological process. As expected, a large number of light responsive elements were obtained from the promoter region of PIFs.

In-silico characterization of PIF4 in wheat is performed using public databases. Multiple sequence analysis was performed for bHLH domain and APB domain, which are characteristic domain of PIF4. The selected genes seem to show the presence of both the domain. These genes can be putative PIF4 genes. A further investigation using one-to-one protein interaction is needed to confirm the TaPIF4. Promoter analysis was carried out to see the presence of similar *cis*-acting regulatory elements of *Arabidopsis* and rice. The results showed presence of similar CAREs. This work can be the foundation of future work for confirming the *TaPIF*4 in wheat

Acknowledgments

With the completion of my thesis work for the fulfilment of my master degree I would like to acknowledge and convey my deepest and heartiest gratitude to everyone, who have been a constant support and guide throughout these two years.

First and foremost, I would like to express my sincere gratitude to my research supervisor *Prof PARAMJIT KHURANA*, FNA, FASc, FNASc, FNAAS, FTWAS, J.C. Bose National Fellow, Department of Plant Molecular Biology, University of Delhi, for taking me under her wing and providing me an opportunity to do my thesis under her guidance.

I would like to appreciate all the Professors of our department *Prof. Akhilesh Tyagi, Prof. J.P. Khurana, Prof. Anil Grover, Prof. Indranil Dasgupta, Prof. Arun Sharma, Prof. Sanjay Kapoor, Prof. Girdhar Pandey, Prof. Saurabh Raghuvanshi and Prof. Surekha Katiyar-Agarwal,* for imparting their knowledge and help us to sail through this process.

I would also like to sincerely appreciate my guide without whom the thesis would not have been possible *Mr. Aishwarye Sharma*. He has been a constant support and guiding light throughout the process of conception to the development of this report. Without his persistent help and suggestions this would not have been possible.

Last but not the least I would like to acknowledge all my classmates, who have been a constant support and encouragement in these trying times. I would also like to appreciate all the love and support I have received from my family.

Abbreviations

- ABA- Abscisic acid
- APA- Active phytochrome A-binding
- APB- Active phytochrome B-binding
- bHLH- basic helix-loop-helix
- BiFC- Bimolecular fluorescence complementation
- BLAST- Basic local alignment search tool
- CARE- *cis*- acting regulatory elements
- ChIP- Chromatin immuno-precipitation
- DNA- Deoxyribose nucleic acid
- EMSA- Electrophoresis mobility shift assay
- GA- Gibberellic acid
- LFR- Low light fluence response
- MeJA- Methyl jasmonate
- MEGA- Molecular evolutionary genetics analysis
- PCR- Polymerase chain reaction
- PHYA- Phytochrome A
- PHYB- Phytochrome B
- PIF- Phytochrome interacting factor
- QTL- Quantitative trait loci
- RNA- Ribose nucleic acid
- ROS- Reactive oxygen species
- SAS- Shade avoidance syndrome
- VLFR- very low light fluence response

Table of Contents

S No	CONTENT	PAGE NO
1.	**Review of Literature**	1-23
	Functions of PIFs	7-23
2.	**Material Methods**	24
3.	**Results and Discussion**	25-40
3.1	Phylogenetic analysis of PIFs in various organisms	25-28
3.2	Promoter analysis of PIFs in *Arabidopsis* and rice	29-31
3.3	Gene expression correlation analysis of PIFs in *Arabidopsis* and rice	32-33
3.4	Characterization of TaPIF4	33-40
3.4.1	Protein BLAST analysis of AtPIF4 in wheat database	33-34
3.4.2	Multiple sequence analysis of AtPIF4, OsPIL13 and TaPIFs	34-35
3.4.3	Phylogenetic analyis of TaPIFs	35-36
3.4.4	Promoter analysis of TaPIFs and putative TaPIF4	36-40
4.	**Conclusion and Future Prospects**	41
5.	**References**	42-46

1. Review of Literature

1.1 Introduction

A lot of work is being done to completely know the working and the fundamental developmental aspects of the plants. One would imagine with the amount of research that has gone into it, the scientific community would know a lot more but, in reality the vastness of research and knowledge has opened up only a minuscule of information. A lot more work needs to be done to completely understand the underlying mechanism. Phytochromes is amongst on the most extensively studied class of photoreceptor.

The discovery of phytochromes and their activity in light had opened up an arena to study about their function in light and dark. Phytochromes are present across the plant kingdom and widely conserved in the angiosperm, non-vascular plants, the fungi and are also found in bacteria's (Lee and Choi, 2017). The phytochromes are of various types, *Arabidopsis thaliana* has PHYA, PHYB, PHYC, PHYD, and PHYE, whereas in rice (*Oryza sativa*) has PHYA, PHYB, PHYC. The interaction of phytochromes with various proteins is the prime reason that is responsible for their activities, some amongst them are the basic-helix-loop-helix (bHLH) sub-family transcription factors known as the phytochrome interacting factors. The structure of the phytochrome interacting factors like any other protein has a N-terminal and a C-terminal. The N-terminal has a conserved motif for the binding of the PHYB (Pfr) called as the active phytochrome B-binding (APB) or the active phytochrome A-binding (APA) for the binding of the PHYA of the phytochrome to Pfr (Bae and Choi, 2008) which is an active form which is transitioned by the absorption of red light by the Pr. The PIFs have a very specific binding site that is found in the promoters of various transcription factors and they preferentially bind to a G-box (CACGTG) DNA sequence element, which is a subclass of an E-box element (CANNTG)(Moon et al., 2008).

The bHLH (basic helix-loop-helix) are transcription factors and they are found across the plant and animal kingdom. These have a DNA binding motif and possess dimerization capabilities. The bHLH has 60 amino acid and has two functional regions in them: the basic region and the HLH (helix-loop-helix) region. The N-terminal region is the basic region and contains the DNA binding motif in them. It has presence of15 amino acids, the HLH has two amphipathic α helices, which helps in the formation of the heterodimers and the homodimers. Most of the bHLH proteins

are seen to preferentially bind to either the E-Box (5'-CANNTG-3'), or the G-Boxes (5'-CACGTG-3') (Li et al., 2006). The basic region of the bHLH is very crucial as it has the ability for the DNA-binding and further functioning of the target gene.



Figure 1: The structural and functional domain of the PIFs

PIFs are seen to have a regulative role in the plant's developmental aspect and they are especially active in presence of the red light (Leivar and Monte, 2014). They have been reported to regulate a large range of functions in the plant ranging from seed germination, hypocotyl elongation, seedling photomorphogenesis, flowering, shade avoidance, leaf senescence, regulation of plant development in high temperature, stress response in plant. Therefore, PIFs have a very holistic function in plant development(Lymperopoulos et al., 2018).

Although phytochrome interacting factors are extensively studied in *Arabidopsis*, other findings have shown their presence in different plants including the lower plants. The first PIF that was discovered was the PIF3 that was reported in *Arabidopsis* in 1998 (Ni et al., 1998) and the studies following that has led to discovery of other PIFs in various plants. This has helped in deciphering their functions and also studying the evolutionary relationship amongst the plants.

Arabidopsis has eight PIF genes, PIF1, PIF3-8, and PIL 1(PIF-like1)/PIF2. All the PIFs in *Arabidopsis* has an APB motif, but PIF1 and PIF3 has an extra APA motif. The PIFs in *Arabidopsis* show a functional redundancy and has an overlap in the pattern of their expression, the difference in the functioning of the various PIF proteins is due to variation in protein stability in different light forms, and the varying regulators of the transcription (Lee and Choi, 2017).

In tomato (*Solanum lycopersicum*) around eight PILs are known in which the APB motifs are present and 3 PILs have APA motifs, which is believed to have an ancient origin according to the genome analysis. The PIFs in tomato are named as SlPIF1a, SlPIF1b, SlPIF3, SlPIF4, SlPIF7a, SlPIF8a and SlPIF8b (Rosado et al., 2016). The

transcriptional profiling done on the PIFs of tomato shows the functions are orthologous to that of the PIFs of *Arabidopsis* which are hypocotyl elongation, seedling de-etiolation, fruit ripening, carotenoid synthesis, maintaining the circadian rhythm of the plant (Llorente et al., 2016).

Apple (*Malus domestica*) is amongst the commercially important fruits and studying about the various aspect of the same has become a necessity Zheng et al., 2020 have reported that apple has 7 PIFs namely MdPIF1, MdPIF2, MdPIF3, MdPIF4, MdPIF5, MdPIF7, MdPIF8. These were deduced on the basis of conserved bHLH and APB domain in them. One interesting observation reported for the PIFs is that MdPIF1, 2, 3, 4 and 5 had a conserved APA domain very similar to that seen in AtPIF1 and 3. The promoter analysis for the MdPIFs revealed large number of ABA-responsive elements. MdPIF4 promoter had a large number of ABA-responsive elements. When MdPIF4 was overexpressed in the apple callus, it showed a reduced sensitivity towards the ABA. The ABA-responsive genes were downregulated in MdPIF4 (Zheng et al., 2020).

In grapes (*Vitis vinifera*), 4 PIFs were reported VvPIF1, VvPIF3, VvPIF4, VvPIF7 (Zhang et al., 2018). VvPIF3, 4 and 7 were seen having the conserved APB motif in them, whereas APA motif was seen in VvPIF1. An expression analysis done for evaluation of PIFs expression in all the tissues of the plant. PIF4 and PIF7 were seen to be highly expressive in the grape skin just before the period of maturation. And the PIFs expression was seen to be sensitive to the light and darkness.

The PILs of rice were also reported. Around six PILs are reported in rice plant namely OsPIL11 to OsPIL16. All these PIFs have an APB motif whereas only OsPIL15 has an APA motif in them. The various functions of the OsPILs that have been characterized are seen that OsPIL11 functions in the hypocotyl elongation and OsPIL13 functions in the internodal elongation of the stem of rice plant (Sakuraba et al., 2017). OsPIL15 shows its function in inhibiting the seedling growth in dark and also few studies carried out indicates its involvement in controlling the grain size (Ji et al., 2019). OsPIL16 was reported as positive regulator of cold stress (Nakamura et al., 2007). Maize is another monocot plant whose role of PIF is yet to be studied in detail and its functional classification to be made clearer and more prominent. It is seen that maize has seven PILs and has a conserved motif of APB. The PIF of maize are ZmPIF3.1, ZmPIF3.2, ZmPIF3.3, ZmPIF4.1, ZmPIF4.2, ZmPIF5.1, ZmPIF5.2, among these ZmPIF3.1 and ZmPIF3.2 has an APA motif in them. The over

expression of the ZmPIF3 has shown to confer drought and slat tolerance in rice(Wu et al., 2019; Lee and Choi, 2017).

The phytochromes that are seen in bryophytes like *Marchantia* are more similar to liverworts than that of the mosses, ferns and the seeded plants. MpPHY has a characteristic feature of PHYA and PHYB. The N-terminal also has PAS, GAF and PHY domain. They have no extension of the N-terminal, rather it translocates to the nucleus when both the red light and far-red light comes into play. The C-terminal of the MpPHY has PAS-PAS and a histidine kinase related domain (HKRD) domain. The PHY-PIF signaling module of the *Marchantia* is very same as seen in *Arabidopsis*. It has only one PIF named as MpPIF. It is seen that they have a conserved APA motif and do not possess the APB motif. Similarly, the four PIFs of *Physcomitrella* (PpPILs) also have a conserved APA motif and does not possess the APB motif. This conveys that the APB motif interaction seen in the seeded plant is a derived function within the course of evolution. MpPIF functions in inhibiting the germination of the gemma, along with the expression of the genes like POR and LHCB (Lee and Choi, 2017).

Table 1: The various PIFs found in the plant kingdom and their function

		Organisms	Accession No:	PIFs	Functions	References
Vascular Plants	Dicots	*Arabidopsis*	AT2G20180	PIF1	Seed germination (skotomorphic)/ Photoperiod development	(Moon et al., 2008; Paik et al., 2019; Gu et al., 2017; Lee and Choi, 2017)
				PIF2		(Lee and Choi, 2017)
			AT1G09530	PIF3	Photoperiod development/freeze tolerance	(Ni et al., 1998; Monte et al., 2004; Jiang et al., 2017; Lee and Choi, 2017)

			AT2G43010	PIF4	Thermosensory/Photoperiodic development	(Qiu, 2020; Franklin et al., 2011; Lee and Choi, 2017)
			AT3G59060	PIF5	Thermosensory/Photoperiodic development	(Khanna et al., 2007; Zhang et al., 2015; Lee and Choi, 2017)
			AT3G62090	PIF6	Seed dormancy	(Penfield et al., 2010; Lee and Choi, 2017)
			AT5G61270	PIF7	Thermosensory/ Shade avoidance	(Jiang et al., 2019; Leivar et al., 2020; Lee and Choi, 2017)
			AT4G00050	PIF8	Photoperiod development	(Oh et al., 2020; Lee and Choi, 2017)
		Tomato	Solyc09g063010	SlPIF1a	Fruit ripening	(L. B et al. 2016; R. D et al. 2016; N and G 2017)
			Solyc06g008030	SlPIF1b		(N and G 2017)
			Solyc01g102300	SlPIF3		(N and G 2017)
			Solyc07g043580	SlPIF4		(N and G 2017)
			Solyc03g115540	SlPIF7a		(N and G 2017)
			Solyc06g069600	SlPIF7b		(N and G 2017)
			Solyc01g090790	SlPIF8a		(N and G 2017)
			Solyc10g018510	SlPIF8b		(N and G 2017)
		Apple	MDP0000289642	MdPIF1		(Zheng et al., 2020)
			MDP0000205358	MdPIF2		(Zheng et al., 2020)
			MDP0000290263	MdPIF3		(Zheng et al., 2020)

			MDP00001984 04	MdPIF4	Reduce sensitivity of ABA	(Zheng et al., 2020)
			MDP00002546 50	MdPIF5		(Zheng et al., 2020)
			MDP00003192 48	MdPIF7		(Zheng et al., 2020)
			MDP00004395 40	MdPIF8		(Zheng et al., 2020)
		Grapes	LOC100264869	VvPIF1		(Zhang et al., 2018)
			LOC100247781	VvPIF3		(Zhang et al., 2018)
			LOC100262490	VvPIF4		(Zhang et al., 2018)
			LOC100262138	VvPIF7		(Zhang et al., 2018)
	Monocots	Rice	Os12g0610200	OsPIL11	Hypocotyl elongation	(Lee and Choi, 2017; Nakamura et al., 2007)
			Os03g0639300	OsPIL12		(Lee and Choi, 2017; Nakamura et al., 2007)
			Os03g0782500	OsPIL13	Internode elongation	(Lee and Choi, 2017; Nakamura et al., 2007)
			Os07g0143200	OsPIL14	Drought tolerance	(Lee and Choi, 2017; Nakamura et al., 2007)
			Os01g0286100	OsPIL15	Grain size	(Lee and Choi, 2017; Nakamura et al., 2007; Ji et al., 2019)
			Os05g0139100	OsPIL16	Cold tolerance	(Lee and Choi, 2017; Nakamura et al., 2007)

		Maize	GRMZM2G115960	ZmPIF3.1		(Lee and Choi, 2017; Wu et al., 2019)
			GRMZM2G387528	ZmPIF3.2		(Lee and Choi, 2017; Wu et al., 2019)
			GRMZM2G062541	ZmPIF3.3		(Lee and Choi, 2017; Wu et al., 2019)
			GRMZM2G065374	ZmPIF4.1		(Lee and Choi, 2017; Wu et al., 2019)
			GRMZM2G165042	ZmPIF5.1		(Lee and Choi, 2017; Wu et al., 2019)
			GRMZM2G016756	ZmPIF5.2		(Lee and Choi, 2017; Wu et al., 2019)
Non-Vascular Plants		Moss (Physcomitrella)		4 PpPILs		(Lee and Choi, 2017)
		Liverworts (Marchentia)		MpPIF	Inhibits gemma germination	(Lee and Choi, 2017)

2. Functions of phytochrome interacting factors

2.1 Light-mediated development

Light is an important factor when it comes to the development of the plant, from a seed to a plantlet. This occurs first in dark the process is called as skoto-morphogenesis and then once it has penetrated the soil, it changes to a light mediated development in them. This switch in the light helps in particular area of development for, example when the seed in under the ground in dark it tends to promote hypocotyl elongation. When the transition takes place for photomorphogenesis, development of the leaves and other lateral part takes place. The light mediated development can be either positive or negative (Qiu et al., 2015).

2.2 Seed germination

When a seed is sown in the ground, the process of germination takes place by the process of imbibition, followed by the rupturing of subsequent layers of the seeds. In *Arabidopsis,* it is known that the process of seed germination is initiated by PHYs. The *PHYA* of phytochrome promotes the seed germination in very low light fluence response (VLFR) and *PHYB* shows germination in low light fluence response (LFR)(Botto et al., 1996). Hence predominantly, it is seen that *PHYB* is important for seed germination. PIF1 acts as a repressor of seed germination. Studies show that *PHYB* destabilizes the PIF1and acts as a light mediated transcription factor and promotes the process of seed germination. PIF1 is responsible for the inhibition of the seed germination because it is very highly expressed in the imbibed seeds and represses the gibberellic acid pathway and activates the ABA pathway(Oh et al., 2006). One study had reported the interaction of Histone deacetylase 15(HDA15), PHYB and PIF. This study shows that HDA15 is a negative regulator for the *PHYB*-dependent seed germination and loss of function mutant of *hda15* showed the germination of seed. Protein interaction study done by yeast-two hybrid showed that HDA15 and PIF1 interacts with each other, suggesting their involvement in repressing the seed germination(Gu et al., 2017).

The seed that is sown in the ground can easily regulate the seed dormancy according to the seasonal variations faced by them. For the germination of seed, light plays an important role and the wavelength of light determines the process of germination. As seen above PIF1 plays a very important role in controlling the process of germination. Similarly, PIF6 is responsible for the seed dormancy. It was reported that an alternative spliced variety of PIF6 controls the primary seed dormancy. With the help of mutant analysis, authors confirmed that PIF6 plays a role in the process of seed germination in dormant seed. Using germination assay, PCR amplification and overexpressing the transgenics PIF6, it was confirmed that PIF6 expressed in seeds especially in dry seeds. The isoforms of PIF6, PIF6-α and PIF6-β were checked to see which one affected the primary seed dormancy and it was seen that the alternative spliced form PIF6-β is primarily responsible for the seed dormancy in *Arabidopsis* (Penfield et al., 2010).

Grain size in rice is an important agronomic character and this determines the yield potential of a rice. So, it is a matter of importance to know the genes that control the grain size and to work on possible way to improve the size and crop production. Over

the years various QTLs related to the grain size have been discovered and a number of transcription factors related to the grain size is known. Various interacting factors of rice like OsPL13 which is a QTL of GLW7 increases the grain size by increasing the cell size in the hull of the grain (Si et al., 2016) and another QTL GW8 that is OsPIL16 increases the grain size by active cell division or cell proliferation helps in increasing the grain size (Wang et al., 2012). There are various factors characterized for the improvement of the grain size. The function of OsPIL15 was characterized in the rice grain using the RNA sequencing method and then through various histochemical cross-section examination it was determined that OsPIL15 is responsible for the controlling of the grain size in rice by controlling the cell division in the grains. Using electrophoresis mobility shift assay (EMSA) the binding site of the OsPIL15 was determined to be N1-Box (CACGCG) and the various differentially expressed genes were then used to check the binding of the PIF. It was also seen that it binds to the OsPUP7 promoter and the yeast two hybrid analysis confirmed that it specifically binds to the promoter region of OsPUP7. OsPUP7 was seen to transport the isopentyle adenine and isopentyle adenosine, which plays a crucial role in cell proliferation, and mutation of OsPIL15 decreases the expression of OsPUP7 while the overexpression increases its level. These findings show that OsPIL15-OsPUP7 play an important role in regulating the grain size (Ji et al., 2019).

2.3 Hypocotyl elongation

PIF1 and PIF3 are actively involved in the skoto-morphorgenic development. During the hypocotyl elongation in dark, they get accumulated and when the PIFs are exposed to light their rapid degradation takes place to promote the protomorphogenetic development, which is initiated by the *PHYA/PHYB*. One of the components responsible for their degradation is the HEMERA (HMR), which works upstream of the PIFs to regulate them. Genetic evidences have proven that HMR binds to all the PIF and there is HMR-PIF interacts through the N-terminal of the HMR and the conserved motif of the PHYB in the PIFs. Along with this, the HMR complex has a nine-amino acid transcriptional activation domain (9aaTAD), which is seen to hinder the PIF targeted genes. All these in-vivo data provides a base where HMR acts as a transcription coactivator and regulate the PIFs and the PIF dependent gene and promotes the degradation of PIF1 and PIF3 (Qiu et al., 2015). In *Arabidopsis* CONSTITUTIVE PHOTOMORPHOGENESIS (COP) is a RING/WD40 repeat that express in dark, which encodes for E3 ubiquitin ligase. COP1 is seen to form a complex with SUPPRESOR OF PHYA (SPA1-4) in dark which targets genes

of photo-morphogenesis like LONG HYPOCOTYLE 5(HY5), LONG HYPOCOTYLE IN FAR RED LIGHT (HFR1), LONG AFTER FAR-RED LIGHT 1 (LAF1) (Kathare et al., 2020). It degrades the activity of target gene by ubiquitylation. PIF1 was seen to enhance the activity of the E3 ligase activity of COP, which degrades the HY5 and HFR1. This indicates that the PIF1 is responsible for the negatively regulating the photo-morphogenesis. This activity of PIF1 is seen to be in sync with other factors which in turns helps in repressing the photo-morphogenesis. During the dark to light transitions the COP1/SPA in association with CULLIN4 (CUL4) promotes light induced degradation of PIF1 (Zhu et al., 2015). To confirm the photomorphogenesis repression by the PIF1 in dark was checked by genetic analysis of the mutants of the *Arabidopsis.* Their protein analysis showed that the wild type seedling showed etiolated seedling whereas the *cop1*, *spa123* showed deetiolated seedling when given a dark treatment and then subsequently exposed to light (Xu et al., 2014). Recent study has shown the presence of PIF8 also known as UNFERTILIZED EMBRYO SAC 10 (UNE10). It is not studied very extensively and the data available is less. Various assays performed on PIF8 have shown that it has APB motif in them which allows in the partial binding of the PIF8 to PHYB and it does not have APA motif, which is the possible reason behind no to less interaction with PHYA and no degradation due to PHYA-induced in far red light as seen in previous literature. The data reported in paper shows that the PIF8 promotes hypocotyl elongation in presence of blue light and far-red light where as in presence of red light no such phenomenon is seen. This characteristic activity of the PIF8 is quite different from rest of the PIFs as they have shown hypocotyl elongation in red light (Oh et al., 2020).

MYB30 is one of the MYB proteins and are transcription factor that are highly conserved in vertebrates. MYB30 has two MYB domains in the N-terminal region and a variable C-terminal. MYB30 is seen to promote hypocotyl growth and is also seen to bind to both PhyA and PhyB as well as PIF4 and PIF5 (Yan et al., 2020). Yeast-two hybrid analysis showed that MYB30 strongly interacts with the Pfr form of PhyB and preferentially also interacted with the Pfr form of PhyA. To check the role of MYB30 in photomorphogenesis a mutant analysis study was done in red light, far red light and blue light and the results showed that except for the FR, short mutant phenotype was seen due to loss of the MYB30. MYB30 induced both the PIF4 and PIF5 when a prolonged exposure to red light was given. Both the PIFs had a MYB binding domain and it was seen that MYB30 bind to the promoter region of PIF4 and

PIF5 inhibiting them to bind to the active form of the phytochrome. Hence suggesting that the MYB30 regulates the photomorphogenesis by interacting with the active form of phytochrome(Bertoni, 2020).

Sucrose is a carbon source to the plants and it play an important role in the development of plant in all aspects including the germination of seed to the floral development to embryogenesis etc., and excess of the same can also lead to developmental arrest in the plants. In an experiment performed by Zhang et al., 2010 it was seen that when a seedling growing in sucrose free medium was transferred to a sucrose containing medium in presence of light. It was seen that the hypocotyl elongation was absent. However, when the same seedling was transferred to sucrose containing medium in absence of light the hypocotyl elongation was seen. Furthermore, it was seen that gibberellic acid was required for the hypocotyl elongation in the dark seedling that was transferred to the sucrose medium. The mechanism by which the GAs induces the hypocotyl growth in sucrose induced medium is not known. When exogenous GA was applied to the *pif*s, that did not help in reducing the effects caused by sucrose. This suggests that impairment is not due to endogenous GA, when endogenous GA was inhibited (Liu et al., 2011).

2.4 Chlorophyll biosynthesis

Along with the repression of the plant photomorphogenesis, it is also known that PIF1 negatively regulates the chloroplast development in dark and the seed germination in FR light. The transcriptional activity of PIF1 is seen when the plant is in dark while, when there is presence of light, the degradation of PIF1 takes places (Huq et al., 2004). Hence, it helps in inducing the production of the chlorophyll. To confirm the active involvement of the PIF1 in the chlorophyll biosynthesis, a gel shift assay was performed to see if it interacts with the genes involved in the chlorophyll biosynthesis pathway i.e., protochlorophyllide oxidoreductase (POR), ferrochelatase (FeChII), and heme oxygenase (HO3). The gel shift assay showed the binding of the PIF is direct to the G-Box of the PORC promoter. The FeChII also has a similar G-Box as that of the PORC, shows that it binds to the G-Box in a very sequence specific manner and confirm the involvement of PIF1 in chlorophyll biosynthesis (Moon et al., 2008).When a seedling is first exposed to light, the process of de-etiolation begins, but the earlier hypothesis that PIF3 functions exclusively to moderate the process of de-etiolation has been rejected by different experiment performed. Rather the experiment such as the microarray expression, RNA blot analysis done on the *pif3* mutant suggested that

the PIF3 functions as an intermediate and the very first light helps in mediating the chloroplast development in a seedling (Monte et al., 2004). *OsPIL13/OsPIL1* in rice is a homolog of *PIF4* and *PIF5* in *Arabidopsis* and its activity is seen in the developmental aspect of the plant. The overexpression of *OsPIL13* has shown its mantel in the shoot elongation process as well, but it has also shown an active involvement in the chlorophyll biosynthesis. In the study done to see the involvement of the OsPIL3 in the chlorophyll biosynthesis, *ospil13* mutant was exposed to natural growing condition of rice showed a pale green leaf. To further investigate about PIL in the chlorophyll biosynthesis a promoter binding assay along with transcription activation assay and it was seen that it directly binds to two chlorophyll biosynthesis gene *OsPORB* and *OsCAO*. The further investigation done shows that OsPIL13 also directly upregulated the genes like *OsGLK1* and *OsGLK2 (GOLDEN2- LIKE (GLK))* which involved in chlorophyll homeostasis (Sakuraba et al., 2017).

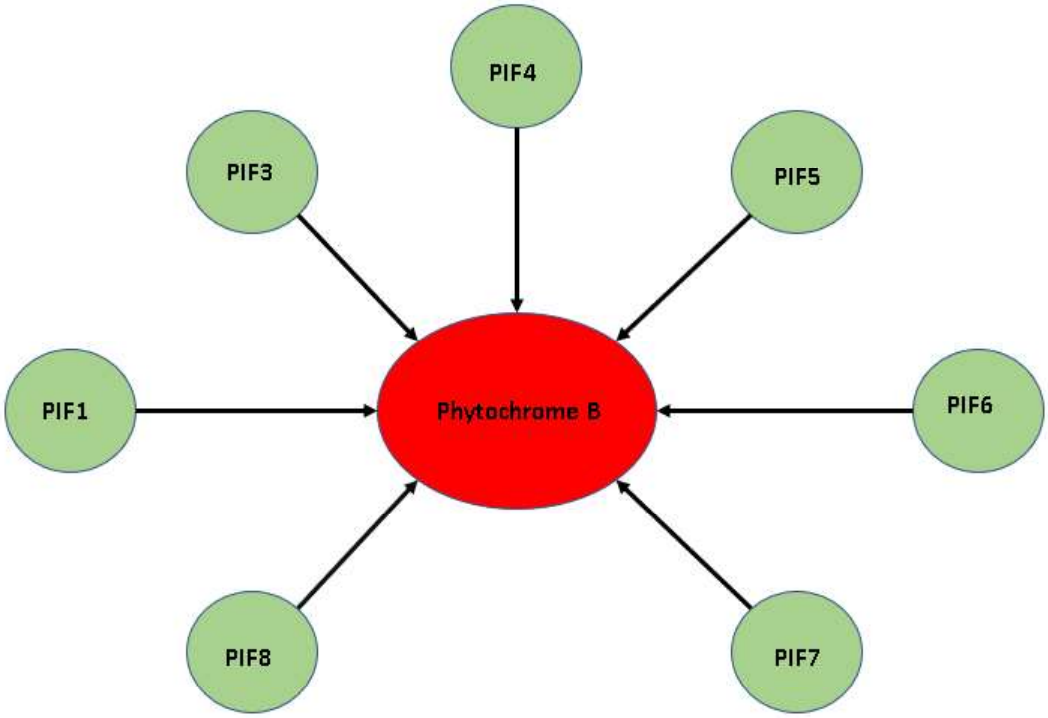


Figure 2: Phytochrome Interacting Factors (PIFs) interacts with the PHYB preferentially due to the presence of APB motifs in them, which allows them to interact with *PHYB*. APA motif is only present in PIF1 and PIF3.

2.5 Flowering

When a plant completes its vegetative phase, it enters the reproductive phase and the first sign of the plant entering the reproductive phase is flowering. It is seen that

when the temperature is increased, there is induction of flowering in *Arabidopsis.* When the temperature increases, the plants perceive it and it acts as a signal for the flowering, which in turn regulates many other pathways in the plant that also includes the photoperiodic cycle (Wigge, 2013). The photoperiodic cycle measures the day length and regulates the flowering time in the plants along with it the other two important gene in regulating the process of flowering in plant is *CONSTANS (CO)* and *FLOWERING LOCUS (FT)* (Suárez-López et al., 2001). During the process of flowering the CO is responsible to activate the *FT*. When the *FT* gets activated in a light dependent manner by the CO, it acts as a florigen, in apical meristem. When there is ample light, as that seen in long day plant the expression of *CO* increases, which in turn increases the expression of *FT*, which helps in promoting the flowering. It is also seen that when there is increase in the temperature, the PIF4 directly binds with the *FT* promoter region and expresses the FT protein. From earlier discussion, it is known that when there is warm temperature the expression of PIF4 is stabilized, which helps in the expression of the same. This indicates that the expression of FT and PIF4 is in a temperature dependent manner (Kumar et al., 2012). Along with the PIF4, other PIFs such as PIF 5 also helps in induction of flowering. Both PIFs works simultaneously to promote flowering in a temperature and light dependent manner. There is another protein that is the EVENING COMPLEX (EC), which is composed of three different proteins EARLY FLOWERING 3 (ELF3), LUX ARRHYTHMO (LUX), and EARLY FLOWERING 4 (ELF4). These were reported to repress the expression of the PIF4 and PIF 5 by binding to their respective promoters at dawn. The accumulation of both the PIFs is highest and at the peak at dawn. Hence simultaneous expression of both the PIFs further help in inducing flowering (Thines et al., 2014).

2.6 Leaf senescence

Leaf senescence is the last stage of developmental process of leaf. As leaf dies, senescence associated genes (SAGs) get upregulated. This happens when a plant is undergoing certain biological changes. For example, the process of flowering, this occurs when there are warmer temperature conditions. As the temperature increases the plants innate characteristic begins the process of flowering in plant and derive the source of nutrition from other part of the plant and gives it to the new developing part (Wigge, 2013). Leaves become the initial part of plant to sacrifice the energy source first. The process of leaf senescence is the last phase of development in the leaves where the degradation of chlorophyll takes place and the yellowing of leaves is seen.

When the temperature increases, level of *PIF 4* increases, along with it *ELF3 (EARLY FLOWERING 3)* decreases, which acts as a suppressor of leaf senescence(Box et al., 2015). Increase in the temperature increases the stability of the PIF4 and its expression and various other factors also help in the activation of the PIF4 in high temperature DE-ETIOLATED 1 (DET1) and CONSTITUTIVE PHOTOMORPHOGENIC 1/SUPPRESSOR OF PHYA-105 1 to 4 (COP1/SPAs), along with the substrate adaptor proteins in the CULLIN 4 (CUL4) E3 ligase complex activate *PIF4* mRNA expression and degrades the HY5 (ELONGATED HYPOCOTYLE 5) which is a repressor of *PIF 4*(Delker et al., 2014). Along with PIF4, it is seen that PIF3 and PIF5 also targets the downstream genes that are responsible for leaf senescence and activates them, *ETHYLENE INSENSITIVE 3 (EIN3), ABA INSENSITIVE 5 (ABI5), ORESARA 1 (ORE1), NON-YELLOWING 1 (NYE1), and GOLDEN2-LIKE 2 (GLK2)* (Sakuraba et al., 2014). All these downstream genes further activate the leaf senescence hormones like ABA, ethylene, brassinosteroid, jasmonic acid, salicylic acid or they promote it by repressing auxin, cytokinin, gibberellic acid. The involvement of PIF4 and PIF5 was confirmed by performing the leaf senescence assay and mRNA expression profiling and studying the *pif* mutants (Kim et al., 2020). PIF5 plays a crucial role in the process of leaf senescence by initiating the process of chlorophyll degradation in dark by directly binding to the G-box motif of the genes like ORE1, NYE, SGR. The study also proved that the ORE1 gets positively regulated in dark. Hence, PIF5 promotes the process of leaf senescence. The microarray data also provided the stance that *PIF5* was upregulated in dark (Zhang et al., 2015).

2.7 Shade avoidance

When a plant is growing under densely populated region of plants, there arises a condition called Shade Avoidance Syndrome (SAS), in which the plants have to compete for light. In plants, the light is received by various photoreceptors, for red light and far-red light the phytochromes are active, cryptochrome (CRY1-2), phototrophins (PHOT1-2) and ZTL-type receptor capture the blue light/UV-A light and one receptor of UV-B light is UVR8 (Pedmale et al., 2016). When the plant faces shade, it activates the shade avoidance response where the plants redirect its reserved resources for the plant's elongation and petiole development. When a plant is under shade for a longer time it speeds up its reproduction process and produces lower biomass and yield, which can be very costly for the crops (Casal, 2012). The shade is sensed by the plants by the decrease of the red light and far-red light in the

surrounding environment and PHYB is a major sensor of the shade as the red light and far-red light are received by the phytochrome as discussed above. When the ratio of red light to far-red light is more than 1, the Pr form of PHYB is converted to Pfr form and it moves to the nucleus and interacts with the PIFs and leads to it rapid degradation and phosphorylation and when the ratio of red light to far-red light is less than 1 the Pfr is converted to Pr form the PIFs gets accumulated in the cell and further binds to the hypocotyl elongation genes (Franklin and Whitelam, 2005). PIF4 and PIF5 accumulates in dark and when there is red light, they degrade and remain in subsequent amount when there is shade. This was proved by performing mutant analysis in shade mimicking conditions and in high red light and far-red light and the result showed elongated hypocotyl and petiole in low light. The degradation of these by phosphorylation requires the presence of the APB domain and this implicate that PIF4 and PIF5 are responsible for shade avoidance syndrome (Lorrain et al., 2008). The shade is sensed by the plant using the PHYB and the response shown by PHYB is not directly responsible for growth, but in paper by Li et al., 2012 has shown that, PIF7 is present in dephosphorylated form and gets accumulated in shaded condition, it interacts directly with the auxin biosynthesis genes and increases their gene expression. Hence, showing the role of PIF7 in hypocotyl elongation in presence of shade. To confirm this data, *pif7* was grown in shade and it was seen that the hypocotyl was short which showed that *PIF7* was involved in the shade avoidance response. Further studies were conducted using the *pif7* mutant to confirm its role in the auxin biosynthesis pathway and confirmed that PIF7 is either directly or indirectly involved in the shade induced auxin biosynthesis. AUXIN RESPONSE FACTOR 18 (ARF18) is a repressor of auxin biosynthesis. PIF mediated inhibition of ARF18 increases the auxin dependent hypocotyl elongation in shade. Both the PIFs and ARF18 suppress the (qua-quine starch) QQS proteins in the leaves, further resulting in the inhibition of growth of the leaves in shade (Jia et al., 2020).

Apart from the phytochromes the, cytochrome can also sense the decrease in the environmental light by the dip seen in the blue light. These changes are perceived by the CRY1 and CRY2 genes and these genes are known to further relay the various signals received to the PIF4 and PIF5 which seen earlier is responsible for the shade avoidance response (Ma et al., 2016). A bimolecular fluorescence complementation (BiFC) assay was performed to check interaction of the CRY1 and CRY2 to the PIF4 and PIF5. The results observed showed interaction strong fluorescence signal in the

nuclei. This indicates that the proteins have a strong interaction (Pedmale et al., 2016).

2.8 Thermo-sensory response

Climate change is a big concern of the scientists as the change in the climate is affecting the temperature of the earth and a slight fluctuation in the temperature affects the plant cell and the cell integrity, which in turn will affect the entire plant in whole be it in its growth and development or be it the internal functions and various pathways. Perception of heat in plants play a very crucial role not only in plants growth and development but also in plants ability to perceive them and overcome various types of stress related to them. Plants being sessile have to adapt to the various temperature range faced by them and slowly adapt to the same, many PIFs are seen to act towards this aspect and involved in the thermo-sensory growth and development of plants. PIF4 has a regulative activity on the thermo-sensory growth of the plant (MA et al., 2009). Multiple photoreceptors target PIF4 when the temperature increases and regulates its activity and the level of PIF. Amongst these are the photoreceptor PHYB, the ultraviolet-A (UV-A) receptor cryptochrome 1 (CRY 1), ultraviolet-B (UV-B) receptor UV RESISTANCE LOCUS 8 (UVR8) (Hayes et al., 2017; Ma et al., 2016). UV-A is a repressor of development of thermally responsive hypocotyl elongation with the help of CRY1. CRY1 protein inhibits the activity of PIF4 but does not affect the level of PIF4 when the temperature is high. UVR8 which is a UV-B receptor regulates the PIF4 transcription and its activity by COP1-dependent fashion (Hayes et al., 2017; Ma et al., 2016). The photoperiodic control and the temperature dependent growth of plant and the PIF4 expression all are co-dependant. It is seen that the hypocotyl elongation as seen in the plant in warm temperature is dependent on the warm daytime temperature for the long day plant (LD) and the warm night time temperature helps in the hypocotyl elongation in short day plant (SD) (Kumar et al., 2012). The EARLY FLOWERING 3 (ELF3) is seen to control the expression and transcription of the PIF4 in the short-day plants (SD) which is a part of the EVENING COMPLEX (EC) along with the ELF3 there is ELF4 and a Myb transcription factor LUX ARRYTHMO (LUX), these all bind to the promoter of PIF4 and represses its activity in early evening (Nusinow et al., 2011). The level of ELF3, ELF4and LUX decreases through the night thus accumulating the PIF4 and promoting the hypocotyl elongation at the end of night (Qiu, 2020). The circadian gating of the thermo-responsive function is only to the SD condition of the plant in spite of ELF3 regulating the expression and functioning of PIF4. A recent study for

LD has shown that GIGANTEA (GI) acts as a thermo-stabilizer of the DELLA protein of REPRESSOR OFga1-3 (RGA), which inhibits the PIF4 to target the auxin related gene hence stopping the hypocotyl elongation (Park et al., 2020).

HEMERA, DE-ETIOLATED 1 (DET1) and COP1, were reported to sense temperature by regulating PIF4. The HRM directly interacted with PIF4 and promoted the growth promotion in the hypocotyl by inducing the auxin related genes DET1 and COP1 (Qiu et al., 2019) work downstream of the PHYB and it transcriptionally and post-transcriptionally regulates the PIF4 and maintains its stability, whereas HRM is only prevalent in the accumulation of the PIF4 (Delker et al., 2014). Both the HRM and DET1/COP1 are essential for the thermo-sensory accumulation of PIF4. HRM is mainly responsible for the PIF4 accumulation at higher temperature but DET1/COP1 mediated stability is not restricted to warm temperature(Qiu, 2020). Apart from PIF4, PIF7 was also seen to regulate the thermo-morphogenesis. It interacted with the auxin biosynthesis genes which promotes the growth in plants (Fiorucci et al., 2020). It is also seen that PIF7 directly interacts with PIF4 and their combined interaction brings about their activity in elevated temperature. PIF7 has turned out as a novel factor apart from PIF4 to promote the early seedling growth in high temperature. In mutant analysis it was seen that similar to the mutant of *pif4*, the *pif7* mutants were unresponsive to the high temperature. PIF7 was seen to control the temperature dependent expression of the auxin related genes like YUC gene which helps in the hypocotyl elongation and cell elongation. The involvement of both the PIFs helped in enhancing the activity of the auxin gene regulation, this suggests that both the PIF forms a heterodimer (Fiorucci et al., 2020).

It has been mentioned that PIF3 acts negatively for photomorphogenesis but it plays a very important role in understanding the environmental cues that the plant receives and one amongst them is the perceivance of the colder temperature. Colder temperature affects the plants growth and development, and to survive the cold temperature cold accumulation takes place. This basically the triggers biochemical and physiological change in a plant. The C-REPEAT BINDING FACTOR (CBF) plays a pivotal role in the cold accumulation when the plant faces the colder temperature (Jiang et al., 2017). A study has shown that PIF3 is a negative regulator of plant freezing tolerance by directly supressing the CBF genes, where the protein stability is negatively regulated by F-Box proteins EBF1 and EBF2. These proteins degrade in presence of cold and enhances the stability of PIF3. PIF3 directly binds to the promoter of CBF gene and supresses its expression. The EBF1/2-PIF3 model

suggests the controlled expression of CBF pathway and controlling the leaky expression of the same and protecting the plants form adverse effect of the CBF gene expression such as excessive retardation of growth (Jiang et al., 2017).

2.9 Phytohormones regulation

Hormones are very crucial in regulating the various life function in all the organisms. In plants, it is seen that the PIFs play an important role in the regulating the various phytohormone signalling like auxin, gibberellin, abscisic acid, ethylene, jasmonate etc (Oh et al., 2009). They help in the integration of internal and external signals and optimize the plant development. The working of the GAs and its involvement in repressing the growth of the seedling in darkness was characterised by Alabadí et al., 2004 . And later, a group was able to articulate the role of DELLAs in repressing the functioning of PIF3 and PIF4 and hence playing a pivotal role in the process of photomorphogenesis repression(Li et al., 2016).

ABA plays a very important role in plant development and response to environmental cues, PIF1 was seen to promote the expression of ABA-biosynthetic genes ABA1, NCED6 and NCED9, and it represses the expression of the ABA catabolic gene CYP707A2, which in turns gives a high ABA levels (Oh et al., 2007). Many ABA related genes which are responsible for the ABA biosynthesis or ABA signalling are direct targets of the phytochrome interacting factor (Liang et al., 2020). *ABA1* is a target gene of PIF1 and PIF5, indicating the involvement of the PIFs in the ABA synthesis. ABA receptors like *PYL6* and *AB15* are a direct target of PIF3 and PIF5, similarly *PYL3* and *PYL12* were also targets of PIF5 (Liang et al., 2020). These evidences shows that the PIFs have a binding affinity to ABA responsive genes and promotes the seedling elongation after germination in plants. It is also seen that PIF1, PIF3, PIF4 and PIF5 can act with the ABA receptors PYL8 and PYL9 and they also directly bind to the G-box motif of the AB15. This indicates that PIFs play a very integral part in the transcriptional and post transcriptional regulating loop of the ABA levels (Liang et al., 2020).

When there is no sufficient light, it is known that the plant grows and that process is called etiolation and ethylene along with gibberellin acid and auxin plays an important role in hook formation that protects the hypocotyl when it penetrates the soil surface (Zhao et al., 2013). When an apical hook is formed there is accumulation of ethylene at the apical hook that is formed. When a seedling was overexpressed with *PIF5* it was seen that there is accumulation of ethylene in them than that of a wild-type

seedling(J et al., 2016). The *PIF5* overexpressed in the seedling plays a pivotal role in regulating the ACS (ACC SYNTHASE) level in the plants. PIF5 helps in maintain the apical hook formed and represses the cotyledon expansion which helps in protecting the primordia emerging from the ground. The PIF5 interacts with the photoactivated form of PHYB (Khanna et al., 2007). The interaction of PIF4 with the BZR1, which is a transcription factor of brassinosteriod, has shown to regulate the PRE family which is a helix loop helix (HLH) which is involved in the cell expansion process. The interaction of BZR1 and PIF4 is interdependent on each other and they show cell elongation in response to darkness, high temperature and BR (Oh et al., 2012).

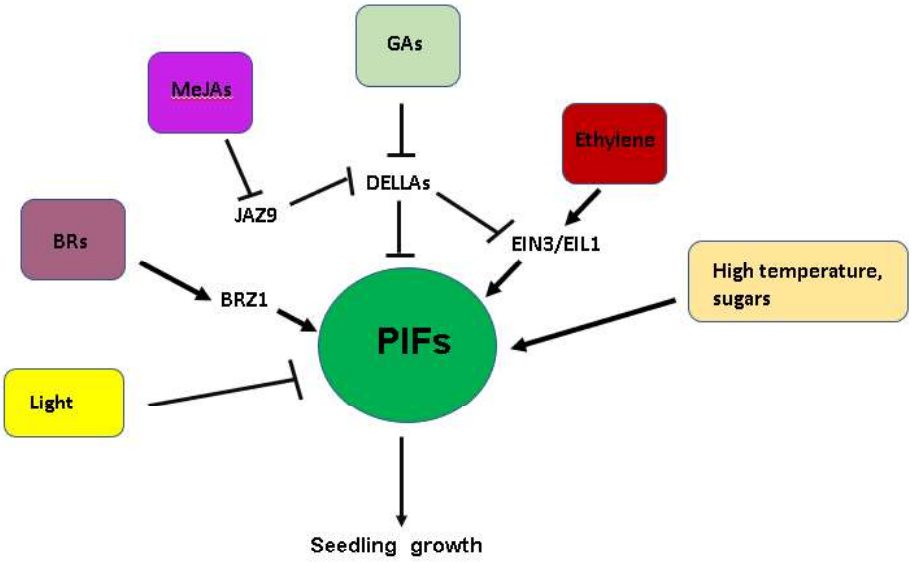


Figure 3: The following shows the interplay of light, temperature and hormone and their role in regulating the seedling photomorphogenesis.

2.10 Stress Response

When an organism faces stress, the very first instinct of an organism is to counter it and survive the stress. Stress faced by plants does not just affect one aspect of the plant's growth and development but it is a multifaceted program. It also triggers the change in the gene expression as well as the metabolism of the plant, hence becomes necessary for the plant to not just protect itself when facing stress but to be able to actively counter the effect of the same. In *Arabidopsis* HY5 and HYH (HY5 HYPOCOTYL) have a multifaceted function in the growth and development as well as in positively regulating de-etiolation of the plant (Oyama et al., 1997). It was seen that during the process of de-etiolation PIF1 and PIF3 prevent the excessive

production of reactive oxygen species and eventually prevent cell death (Chen et al., 2013). Loss of either PIF1 or PIF3 leads to the over-accumulation of the protochlorophyllide in the cells which are grown in dark, so when the seedlings of *Arabidopsis* were transferred from dark to light, due to ROS photo-bleaching occurred (Chen et al., 2013). Studies have shown that PIF1 and PIF3 bind to the G-box motifs and down regulates the activity of ROS-responsive gene in light, hence reducing photo-oxidative damage in plants during de-etiolation. PIF1/PIF3 and HY5/HYH binds to the same target, suggesting the co-regulation of the ROS-responsive genes and these act antagonistically to each other (Chen et al., 2013).

Another way through which the plant responds to protect itself from the abiotic and biotic stress, is by inhibiting its growth and development to conserve its resources and redirect these resources to help protect the plants in case it is needed. In *Arabidopsis* an AT-hook motif nuclear-localized (AHLs) transcription factors which are essentially abiotic stress response gene inhibits the rosette formation in response to the stress. The AHLs transcription factors have two domains in them: one is the plant and prokaryote conserved domain (PPC) and another DNA binding AT-hook motif. So, to confirm the role of AHLs, a *SUPRESSOR OF PHYTOCHROME B4-#3 (SOB3)* and other AHLs that restrict the petiole elongation by antagonising the growth promoting PIFs was studied (Favero et al., 2016). The mutant analysis study done on the mutants of *SOB3* to see the petiole growth, it showed that the mutant with high *SOB3* had a short petiole whereas the dominant negative mutant which had a AHL27 showed enhanced petiole growth. On further investigation, it was seen that AHLs represses the auxin signalling by inhibiting the transcription, as well as it inhibits the ethylene biosynthesis genes and the brassinosteriod genes whose functions have been discussed earlier. AHLs inhibits the petiole growth by reducing the PIFs binding to various transcriptional activation genes associated with growth as well as with the hormone signalling pathway (Favero et al., 2020; Monte, 2020).

For combating drought stress many transgenic plants have been produced but many of them show growth retardation in them. So, a study was carried out where the OsPIL13 of rice along with DREB1A (DEHYDRATION RESPONSIVE ELEMENT BINDING 1A) was over expressed in *Arabidopsis.* OsPIL13 is a homologous of PIF4. The double overexpression of these genes showed an improved drought tolerance along with improved hypocotyl elongation and flower induction. The transcription analysis done showed there was no hinderance when both the genes were

overexpressed together rather, they function in an independent manner by downregulating their respective genes like abiotic stress inducible genes for DREB1A and the cell elongation gene of the OsPIL13 (Kudo et al., 2017). Another transgenic for drought tolerance was made using the phytochrome interacting factor of *Zea mays* ZmPIF1. This was overexpressed in rice to see if it confers tolerance to drought and the results so obtained conferred the drought tolerance in rice. It was further seen that expression of ZmPIF1prevents water loss and actively participates in closing the stomata along with it the PIF increased the ABA sensitivity and it is well known that when a plant suffers from stress, ABA plays a significant role in avoiding the stress. The various drought resistance genes, that is the *OsRAB*16A, *OsRAB*16B, *OsRAB*16D, *OsPR*4b and OsPR4c, were up-regulated in ZmPIF1 transgenic rice (Gao et al., 2018). OsPIL14 a bHLH phytochrome interacting factor was seen to be binding with the *OsDREB*1B promoter which is a regulator of plant stress, this was confirmed by performing the Yeast-one screening. This interaction of OsPIL14 suggests, it has a role in the light mediated stress response as the OsPIL14 binds preferentially to the active form of PHYB. The OsDREB1B has shown to stress response to cold, drought, oxidative stress etc. This involvement of both the OsPIL14 and OsDREB1B shows a cross talk between the light and stress signalling (Cordeiro et al., 2016).

The salt stress affects the plants growth and development. When a seedling is exposed to salt at a very young stage, it affects their normal growth (Park et al., 2016). Gibberellic acid is one of the hormones that is involved in the salt stress response of a plant. When a plant receives a salt stress, the level of GA goes down. DELLA protein acts as a repressor of the GA synthesis and one of the DELLA protein that is involved is the SLENDER RICE1 (SLR1) (Itoh et al., 2002). OsPIL14 when overexpressed along with the DELLA protein SLR1, their interaction promoted the root development in dark and in presence of salt stress. When there is presence of salt, the OsPIL14 degrades but SLR1 accumulates. This OsPIL14-SLR1 interaction bring the facts that the light and GA signals works together to help the plant survive the salt stress (Mo et al., 2020). The expression pattern analysis done on the OsPIL16 of rice it was seen to be induced in the cold condition. He et al., 2016 showed that a mutant of PHYB provided tolerance to cold stress. When a protein interaction was done it was seen that OsPIL16 of rice interacted with the PHYB of the rice. The OsDREB1is a stress response gene as mentioned above, whose levels where very high in the mutant of *phyb.* The ChIP assay done on the OsDREB1 and

OsPIL16 showed that the rice PIL bind to the promoter region in the N-box of the stress response factor. All these data accumulated together propose that the deficiency of the PHYB can stimulate the production of OsDREB1via the OsPIL16 hence giving the plant a chance to survive the cold stress by maintain its cell membrane integrity and the less production of malondialdehyde, which confers that cold tolerance (He et al., 2016).

Table 2: Interacting partners of PIFs and the binding sites /promoter elements.

PIFs	Proteins	Elements the PIFs bind to	Reference
PIF1	COP1/SPA1, HY5, HECATE, MRI156S Gibberellic acid- GA1, RGA1 Abscisic acid- AB13, AB15 Chlorophyll biosynthesis- PORC, FeChll ROS responsive- APX2, ZAT10, SIB1, ERF4, NDB2 HEMERA	G-box element G-box element G-box element G-box element G-box element APB motif	(Moon et al., 2008; Chen et al., 2013; Xu et al., 2014; MA et al., 2009; Qiu et al., 2015)
PIF3	ROS responsive- APX2, ZAT10, SIB1, ERF4, NDB2 Freeze tolerance- CBF Abscisic acid- ABA1, PYL6 MRI156s HEMERA	G-box element G/E-box element G-box element G-box element APB motif	(Chen et al., 2013; Liang et al., 2020; Jiang et al., 2017; Qiu et al., 2015; Xie et al., 2017)
PIF4	Auxin biosynthesis - IAA19, IAA29, TAA1, CYP79B2, SAUR19, YUCA8 Brassinosteriod biosynthesis -BZR1, ARF6 ORE1, TCP5/13/17, TOC1, HRF1, ELF3, HY5, FCA, DELLAs, MRI156s, PRR MYB30 HEMERA, CRY1, CRY2	G/E-box element G-box element G-box element Promoter of PIF APB motif	(Qiu et al., 2015; Bertoni, 2020; Franklin et al., 2011; Sun et al., 2013; Qiu, 2020; Thines et al., 2014; Oh et al., 2012)

PIF5	ABA biosynthesis - ABA1, PYL6, PYL3, PYL12 COP1, MRI156s MYB30 CRY1, CRY2, HEMERA	G-box element G-box element Promoter of PIF APB motif	(Qiu et al., 2015; Bertoni, 2020; Liang et al., 2020)
PIF6	HEMERA	APB motif	(Qiu et al., 2015)
PIF7	Auxin biosynthesis genes- IAA19, IAA29, YUCA8 DREB1C, DREB1B ELF3, TOIC, PRR HEMERA	G- box element G- box element G- box element APB motif	(Qiu et al., 2015; Jiang et al., 2019; Fiorucci et al., 2020; Jia et al., 2020)
PIF8	COP1 HEMERA	G- box element APB motif	(Qiu et al., 2015; Oh et al., 2020)

3. Material and Methods

3.1 Phylogenetic analysis:

Using MEGA X software (Kumar et al., 2018) the phylogenetic tree for *Arabidopsis*, rice, maize, tomato, apple and grapes was performed. Unrooted phylogenetic tree was formed using the Neighbor joining method and with 1000 bootstrap value.

3.2 Identification of *Cis*-acting regulating elements:

Promoter analysis was done for *Arabidopsis* and rice by taking 2kb upstream region of various PIF genes from Phytozome (https://phytozome.net) database (Goodstein et al., 2012). The conserved cis-acting regulating elements were predicted using the PlantCARE database (Lescot et al., 2002).

3.3 Gene expression analysis:

Gene expression analysis for *Arabidopsis* and rice was done by using BAR (The Bio-Analytic Resource for Plant Biology) (http://bar.utoronto.ca/). The value of expression was obtained from the website and a graph was produced to study the same.

3.4 Characterization of *Arabidopsis* PIF4 in wheat

Using Plant ensemble database (https://plants.ensembl.org/index.html) Blast was performed against the protein sequence of PIF4 of *Arabidopsis* and sequence above the percentage identification of 80 was selected.

Multiple sequence analysis was performed using ClustalW. A phylogenetic tree was made using MEGA software. The 2kb upstream promoter sequence of the selected genes were obtained and a promoter analysis for the same was done using PlantCARE.

4. Results and Discussion

4.1 Phylogenetic analysis of PIFs in various organisms

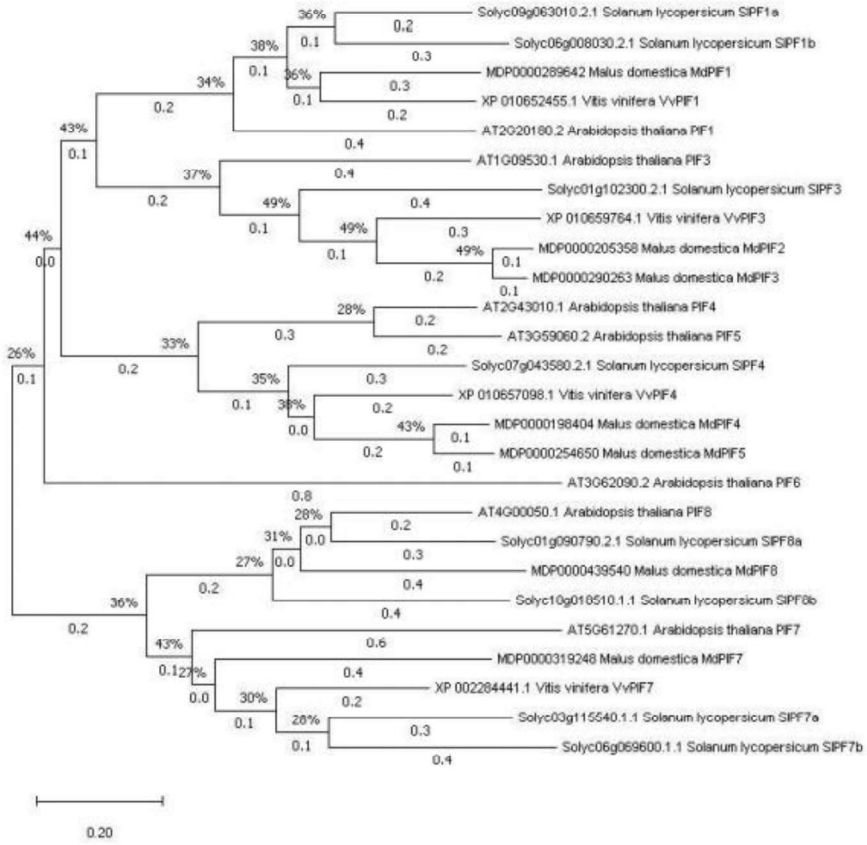


Figure 4: Phylogenetic tree representing the PIFs of dicots

The tree in Figure 4 is a phylogenetic tree of PIFs of dicot plants, which consists of *Arabidopsis thaliana, Solanum lycopersicum* (Tomato)*, Malus domestica* (Apple) and *Vits virifera* (Grapes). [The tree is formed using the MEGA software]. The process used to make the tree is the Neighbour joining tree, which is a distance-based clustering method where the evolutionary distance is defined by the distance between each of the sequence which is done by using the distance matrix to form a tree.

The horizontal lines in the above tree are called as the branch of the tree which gives the branch length which shows the genetic change (mutation) an organism has seen in comparison to the other over the time period of evolution. The percentage shown above the branches shows the bootstrapping value of the tree, when the tree is resampled, a subsequent tree is produced in each cycle the percentage gives an estimate to how many times the same branch is seen at the same place in the tree, hence allowing in gaining a confidence number.

In the above tree we can see that all the PIF 7 and PIF8 form one cluster and PIF6,5,4,3,2 and 1 form another cluster indicating that both the clusters evolved from the same ancestor and further divided according to the change in their sequence. In the tree clade with PIF1, we see that *Arabidopsis* and PIFs of *Solanum lycopersicum*, *Vits virifera* and *Malus domestica* have evolved from a common ancestor and divided further to form the leaf of the tree. The PIF1a and PIF1b of tomato is most closely related and evolve from the same ancestor, similarly we see that MdPIF1 and VvPIF1 is seen to evolved from the same ancestor and is more closely related to each other. When we move further down in the tree, we see that the PIF1 and PIF3 in the given tree have the most recent common ancestor and they have divided further to form another cluster where the PIF3s of the various organisms are seen to divide according to the change in their sequence. MdPIF2 and MdPIF3 are most closely related in the clade followed by VvPIF3 having a common ancestor to that of PIF of apple. PIFs in tomato are more closely related to that of the PIFs in grapes and then following is *Arabidopsis,* showing a closer relation to tomato. In the next clade of PIF4 and PIF5, of *Arabidopsis* are closely related, hence are seen in one clade. PIF4 of tomato, grapes and apple along with PIF5 have the most recent common ancestors to that of *Arabidopsis*. PIFs in apple i.e., MdPIF4 and MdPIF5 are more distantly related and VvPIF4 is more closely relate to that of PIF of apple. Followed by tomato which has a closer association to that of PIFs of grapes than that to the PIFs of *Arabidopsis*. PIF6 of *Arabidopsis* is present on a separate clade and has a common ancestor to that of PIF1,3,4 and 5. PIF8 and PIF7 have a more common ancestor and form a separate cluster altogether. PIF8 of *Arabidopsis* and tomato SPIF8a are more closely related in the clade. MdPIF8 has a common ancestor to that of PIF of *Arabidopsis* and tomato SPIF8a. SPIF8b of tomato has a common ancestor with that of MdPIF8. The entire clade of PIF8 has a common ancestor with PIF7 clade. In the PIF7 clade PIF of tomato SPIF7a and SPIF7b are more closely related to each other. Grapes have most recent ancestor with tomato, whereas apple has a common ancestor with that of grapes. *Arabidopsis* has most recent ancestor with the PIF of apple. One can see in this dicot phylogeny that each of the PIFs have evolved from a similar ancestor in different organisms, given us a base idea in their similarity in functions in different organisms.

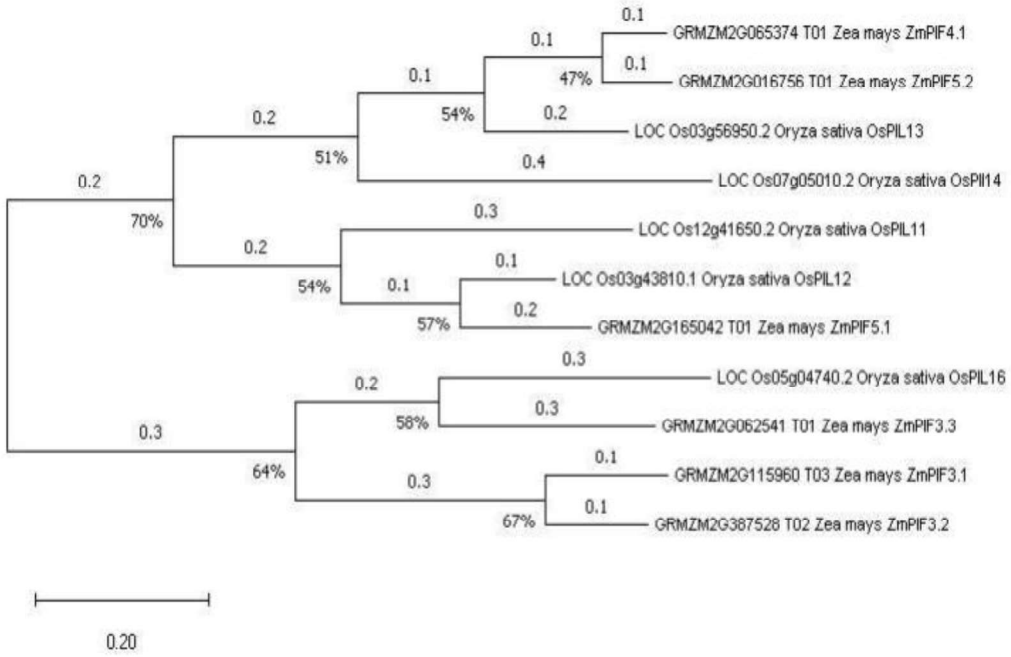


Figure 5: Phylogenetic tree representing the PIFs of monocots

Fig. 5 shows the phylogenetic tree of the monocots, which consists of rice and maize PIFs protein sequence. The PIFs of maize that is ZmPIF4.1 and ZmPIF5.2 are, most closely related in the cluster and have a more similarity with the rice PIF OsPIL13 than that of OsPIL14, though they may share some similarity with OsPIL14 because they are in the same clade. In the 2nd cluster we see that ZmPIF5.1 and OsPIL12 are most closely related to each other, followed by them sharing similarity with OsPIL11. In the final cluster OsPIL16 and ZmPIF3.3 are more closely related and ZmPIF3.1 and ZmPIF3.2 are more closely related, and these two clades together share a common ancestor. This tree indicates that the cluster 3 and cluster 1 and 2 share the same ancestor and they have evolved from there because of the genetic variation that occurred in them. Both the cluster 1 and cluster 2 share the same ancestor in the course of evolution.

Fig. 6 shows the phylogenetic tree showing the relationship of all the PIFs of both the monocots and dicots. The earlier phylogeny obtained for dicot and monocots remains the same, the only difference observed is that the PIFs of monocots place themselves in the tree as such that it helps in deciphering that to which PIF of dicot it is more similar to. The PIFs of rice and maize that is OsPIL13, OsPIL14, OsPIL11, OsPIL12 and that of the maize in the same clade are seen to share a common ancestor with that of PIF4 of *Arabidopsis*. Similarly, the OsPIL16 of rice and the

Maize PIFs of the same clade share a common ancestor with PIF3 of *Arabidopsis* this similarity of the PIFs also shows the similarity in functions as well, as seen earlier OsPIL16 in rice is responsible for cold tolerance and it is also seen that PIF3 of *Arabidopsis* is also involved in the cold tolerance.

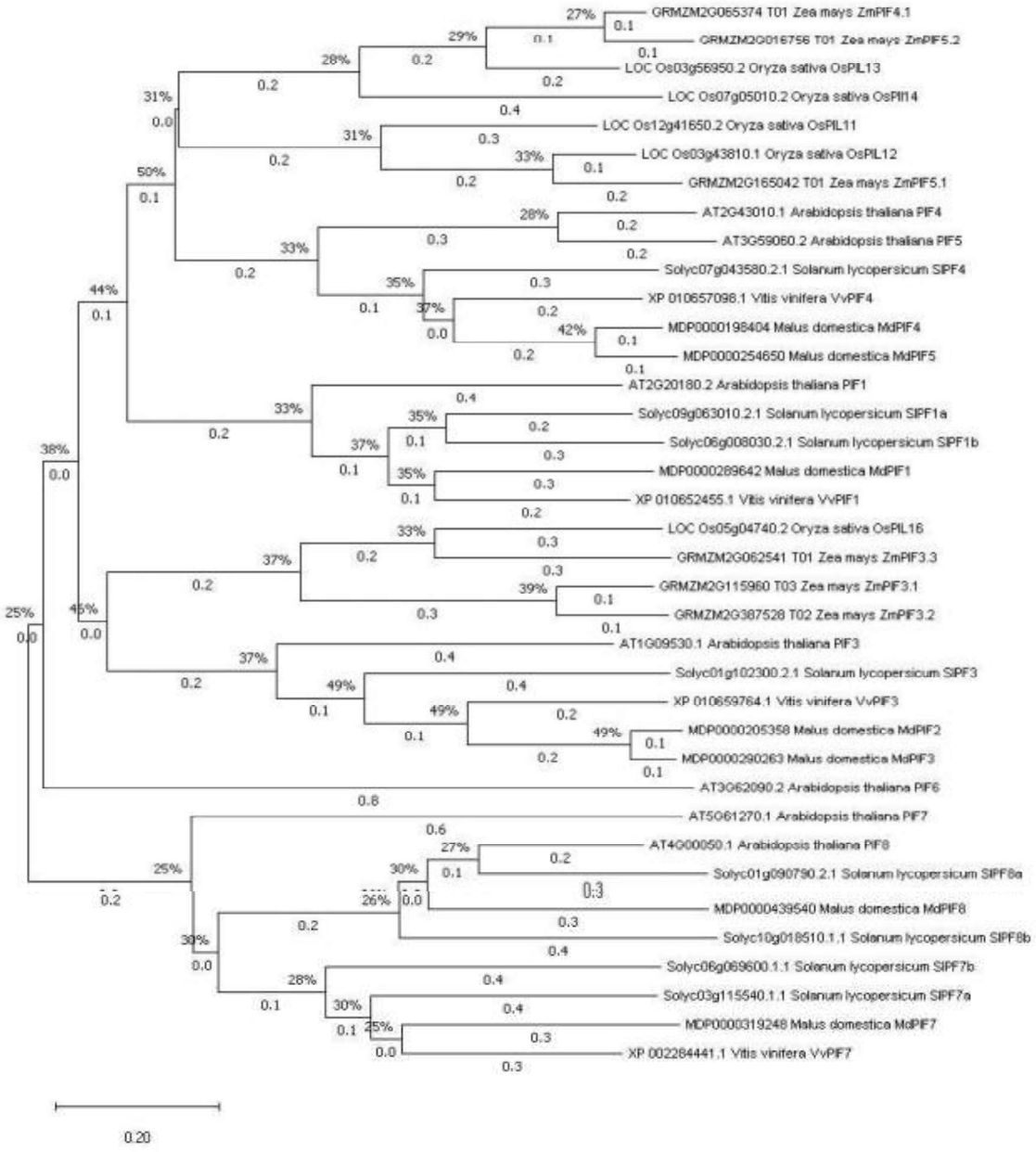


Figure 6: Phylogenetic tree representing the PIFs of monocots and dicots

4.2 Promoter analysis of PIFs in *Arabidopsis* and rice

The promoter analysis done on the PIFs of *Arabidopsis* and rice highlights the different type of regulatory elements that are present upstream of transcription initiation site. The analysis was done by taking 2kb upstream from the translational start point in the promoter region. The promoter elements so obtained can be divided into different groups based on their functions, such as light responsive elements, hormone regulation elements, stress responsive elements and elements involved in plant growth and development. Light responsive cis-acting regulating elements in the following analysis were found out to be highest in number 158 elements. These elements included G-Box (39), Box-4 (36), GT1- motif (16), AE- box (11), TCT- motif (10), ATCT- motif (8), GATA- motif (7), I-Box (7), TCCC- motif (6), MRE (5), ACE (3), AT1- motif (3), GA- motif (2), chs-CMA1a (2), chs-CMA2a (1), ATC- motif (1), LAMP element (1). The distribution of the light element is highest in PIF4 (18) of *Arabidopsis* and OsPIL12 (18) of rice, OsPIL13 ortholog of PIF4 only had (4) elements involved in light response. The presence of these elements in the promoter regions supports the role of the PIFs in the light mediated pathways that are seen in the plants and light plays a major role in the functioning of the PIFs. The 2 highest number of elements found are that of the elements involved in the hormone regulation of the plants. 110 elements for hormone regulation were seen in this analysis. These elements include the elements of the following hormones Gibberellic acid (TATC- Box (5), P- Box (2), GARE- motif (2)), Auxin (AuxRR-core (3), TGA-element (8)), Abscisic acid (ABRE (33)), Methyl-jasmonate (CGTCA-motif (20), TGACG-motif (34)), Salicylic acid (TCA-element (2)). Amongst the following the amount of element received for the MeJA is highest and is seen to be present in almost all of the PIFs except for PIF8. PIF5 of *Arabidopsis* is seen to be having elements for each of the hormones mentioned here. It is a well-known fact that the hormones have an interplay amongst themselves and the presence of these elements in the PIFs strongly suggests their role in the hormone regulation of the plants and the development that follows with them.

Stress responsive cis-regulating elements were also observed in this analysis. The number of these elements seen was 47 out of which ARE (anerobic responsive element) was found to be maximum (23), second to the ARE was the TC-rich repeats these are the cis-acting elements which are involved in the defense and stress responsiveness of the plants. Around 6 LTRs were observed which are involved in the low-temperature responsive action of the plants and 4 MYB binding elements that

regulates the drought response in the plants. In the following analysis few elements related to plant development were also found like circadian, A-Box, AACA_motif, GCN4_motif, MBSI, HD-Zip1, CAT-Box, RY-elements, and MSA-like. These elements are involved in various functions of development. The circadian elements are involved in the plants circadian-clock mediated development and these elements were seen in PIF3,8 and OsPIL13. The CAT-box is mostly involved in the meristematic expressions and can be one of the possible elements that is involved in the meristematic growth of the plant. RY-elements are the one to be involved in the seed specific regulation or expression of them, AACA_motif and the GCN4_motifs are the one involved in the endosperm specific expression. HD-Zip1is one element found in PIF7 which is responsible for the differentiation of the palisade and mesophyll cells of the plant which is responsible for photosynthesis. MSA-like is responsible for the regulating the cell cycle and this element is seen only in the rice OsPIL14.

Upon careful analysis it is concluded that PIFs are regulated by light responsive genes of the plant, which also supports the theory of PIFs major involvement in light mediated signaling pathways.

Table 3: Various elements present in promoter region of *Arabidopsis* and rice

Light responsive elements	No. of elements	Hormon regulated elements	No. of elements	Stress responsive elements	No.of elements
G-box	39	TATC-box	5	ARE	47
Box-4	36	P-Box	2	TC-rich repeats	23
GT1-motif	16	GARE-motif	2	LTRs	6
AE-box	11	AuxxRR-core	3	MBS	4
TCT-motif	10	TGA-element	8		
ATCT-motif	8	ABRE	33		
GATA-motif	7	CGTCA-motif	20		
I-Box	7	TGACG-motif	34		
TCCC-motif	6	TCA-element	2		
MRE	5				
ACE	3				
AT1-motif	3				
GA-motif	2				
chs-CMA1a	2				
chs-CMA2a	1				
ATC-motif	1				

LAMP element	1

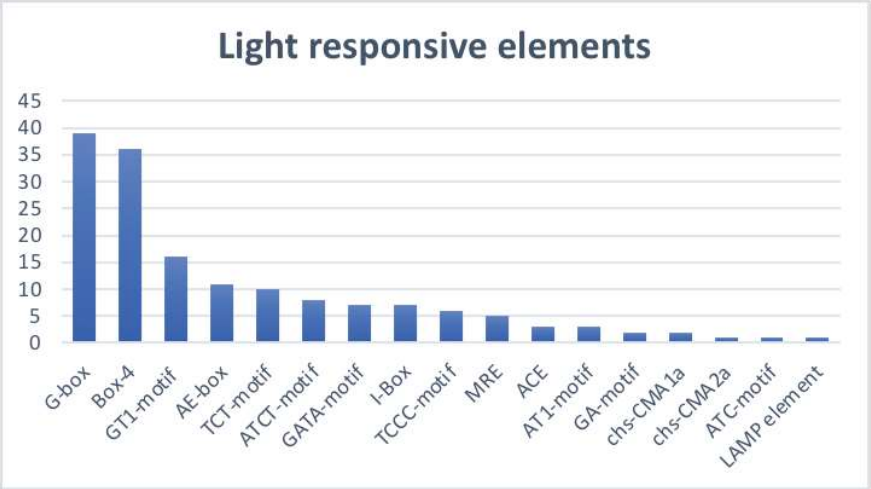


Figure 7: Representing the light responsive elements present in the promoter region of *Arabidopsis* and rice

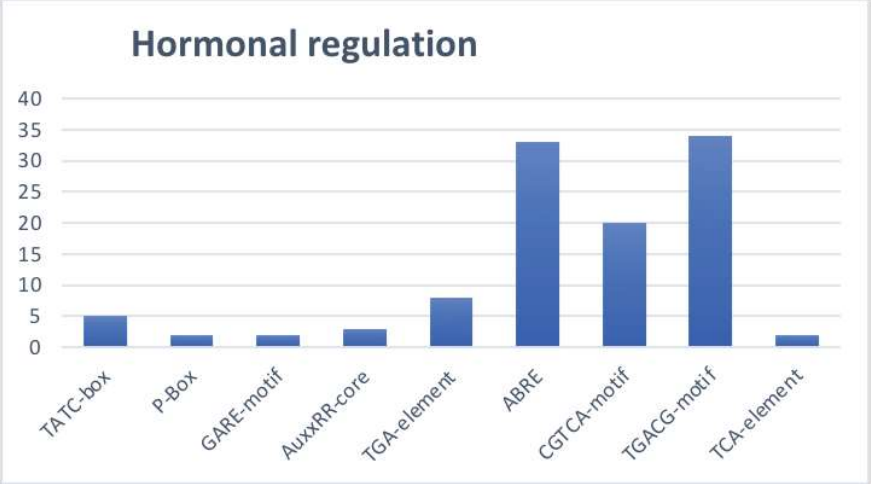


Figure 8: Representing the hormone regulating elements present in the promoter region of *Arabidopsis* and rice

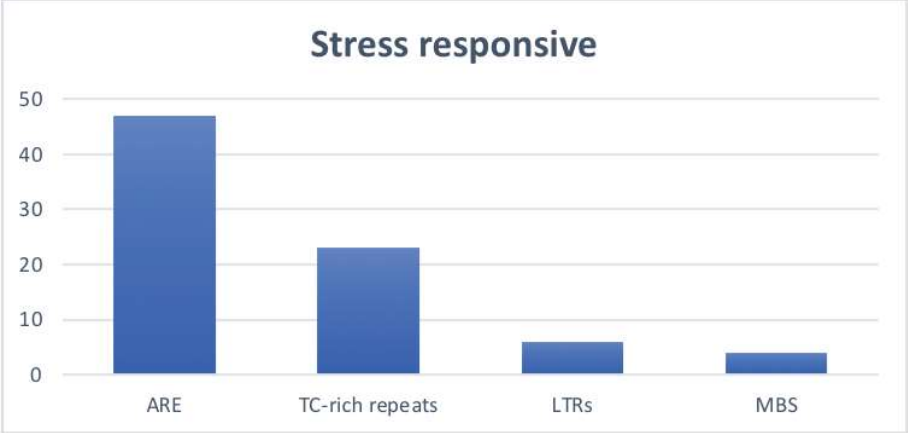


Figure 9: Representing the stress responsive elements present in the promoter region of *Arabidopsis* and rice

4.3 Gene expression correlation analysis of PIFs in *Arabidopsis* and rice

The expression pattern of the gene in plants gives an insight into their functioning. When a gene is expressed in a plant it shows its expression in different levels in various part of plants. To have a better understanding of working of the PIF a comparative analysis for the gene expression for *Arabidopsis* and rice was done and the expression pattern of the same were seen in the various tissues present in the plant. And it was seen that the expression level of *Arabidopsis* PIF4 and rice OsPIL13 which are orthologs are similar to each other. The expression pattern of both the PIFs for shoot apical meristem and root is of similar kind and seen to be the highest expressed gene in them compared to the carious other PIFs. The high level of expression of the PIFs in the shoot apical meristem confirms their involvement in the development of the shoot also in the process of formation of the shoot inflorescence and flowering.

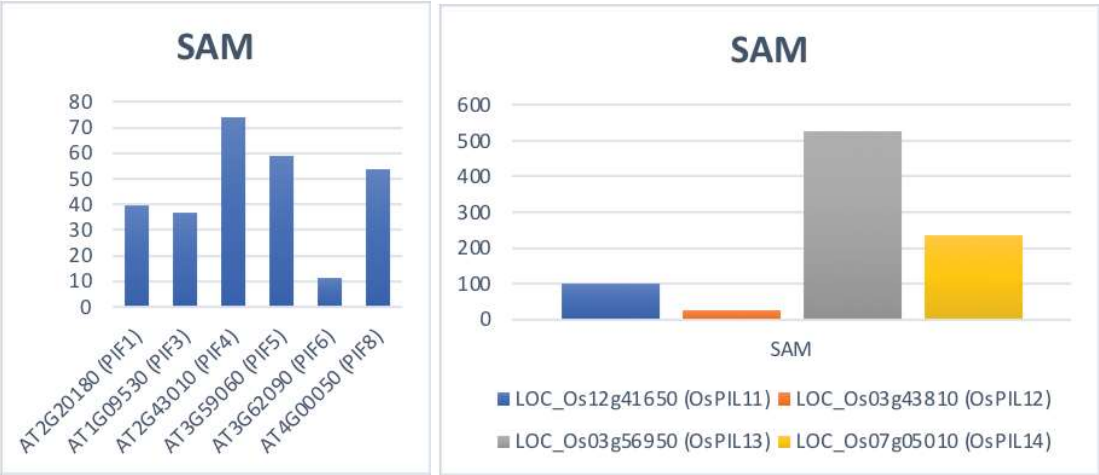


Figure 10: Representing the expression level of PIFs in *Arabidopsis* and rice in shoot apical meristem

A very similar kind of expression pattern is also seen for both the PIFs present in the roots of young seedlings in the plants. The expression level of OsPIL13 is at the lowest compared to other PIFs in the rice plant similarly the expression pattern of PIF4 and PIF5 in *Arabidopsis*, which are highly similar sequence (Sakuraba et al., 2017) supporting the reports of PIF4 and OsPIL13 being an ortholog.

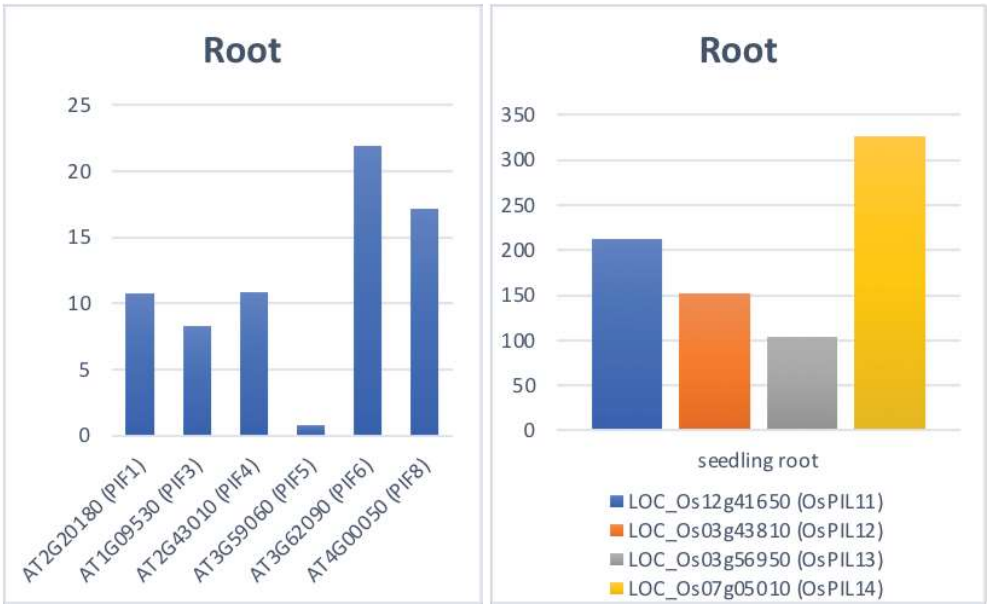


Figure 11: Representing the expression pattern of PIFs in *Arabidopsis* and Rice in the seedling roots.

4.4 Characterization of *TaPIF4*

4.4.1 Protein BLAST analysis of AtPIF4 in wheat database

In-silico characterization of *PIF4* of wheat (*Triticum aestivum*) was carried out to find out the putative orthologs. For the analysis, peptide sequence of AtPIF4 was taken and a BLASTP analysis was carried out, where E-value<1e-5 and sequence above percentage identification 80 was considered (Table 4). A total of 14 sequences were selected. Pfam was checked to find conserved domain and all the genes obtained conatained bHLH domain.

Table 4: Representing the result of BLASTP analysis

Gene name	Gene Hit	Genomic location	Orientation	% ID
TraesCS5A02G049600.1	TraesCS5A02G049600	5A:45370341-45370689	Reverse	87.7
TraesCS5B02G054800.2	TraesCS5B02G054800	5B:59926528-59926919	Reverse	86
TraesCS5B02G054800.1	TraesCS5B02G054800	5B:59926528-59926919	Reverse	86
TraesCS5D02G060300.1	TraesCS5D02G060300	5D:56951364-56951762	Reverse	86
TraesCS5B02G380200.1	TraesCS5B02G380200	5B:558121032-558121419	Reverse	84.2
TraesCS5D02G386500.1	TraesCS5D02G386500	5D:456341823-456342198	Reverse	84.2
TraesCS5A02G376500.1	TraesCS5A02G376500	5A:574197476-574197821	Reverse	84.2
TraesCS5D02G386500.2	TraesCS5D02G386500	5D:456341823-456342198	Reverse	84.2

TraesCS5B02G422000.2	TraesCS5B02G422000	5B:597219876-597220433	Forward	82.5
TraesCS5D02G428400.2	TraesCS5D02G428400	5D:486168064-486168615	Forward	82.5
TraesCS5A02G420200.2	TraesCS5A02G420200	5A:606894429-606894970	Forward	82.5
TraesCS5A02G420200.1	TraesCS5A02G420200	5A:606894429-606894970	Forward	82.5
TraesCS5D02G428400.1	TraesCS5D02G428400	5D:486168064-486168615	Forward	82.5
TraesCS5B02G422000.1	TraesCS5B02G422000	5B:597219876-597220433	Forward	82.5

4.4.2 Multiple sequence analysis of *AtPIF*4, *OsPIL*13 and *TaPIF*s

To further confirm if these obtained sequences are similar ortholog candidates of PIF4, multiple sequence alignment was performed using ClustalW. Along with *AtPIF*4, *OsPIL*13 and best matched sequence of *Horedum vulgare* and *Brachypoidum distachyon* obtained by a BLASTP analysis were used, as both *Horedum vulgare* and *Brachypodium distachyon* are closely related species of wheat. bHLH domain was checked in all the sequences and it was found in all the sequences (Fig. 12).

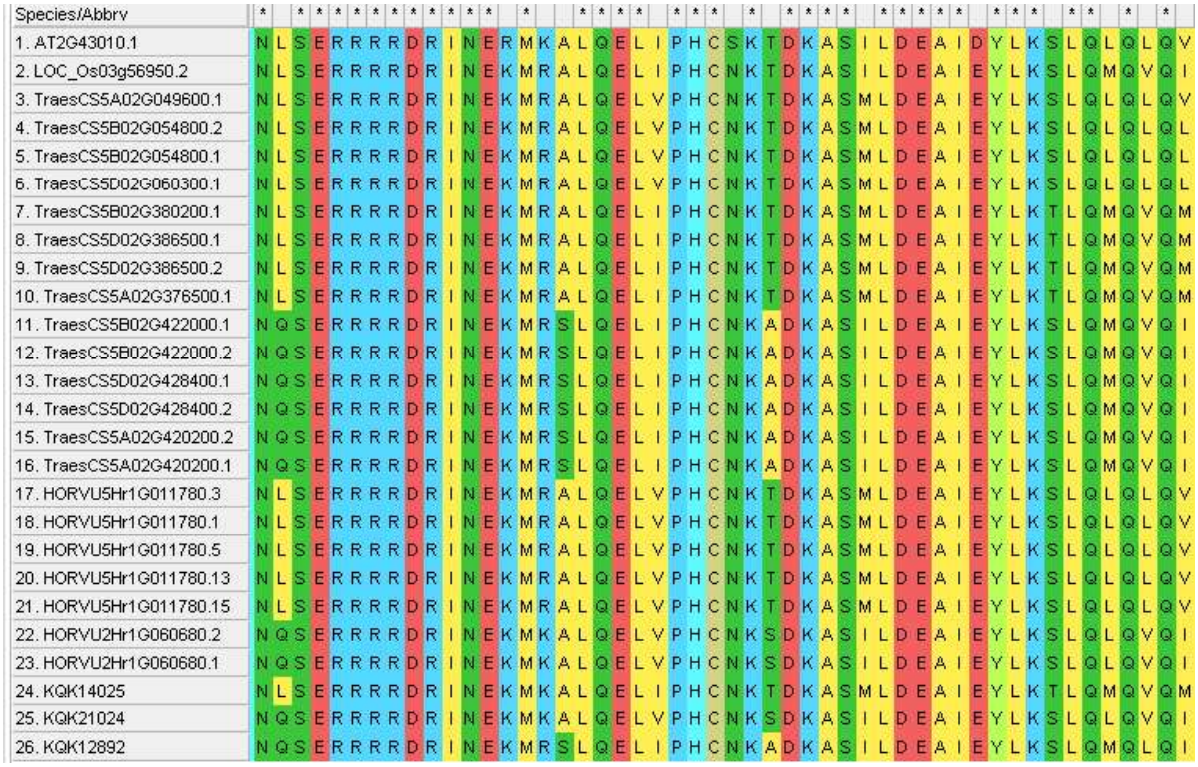


Figure 12: Multiple sequence alignment representing the conserved bHLH domain

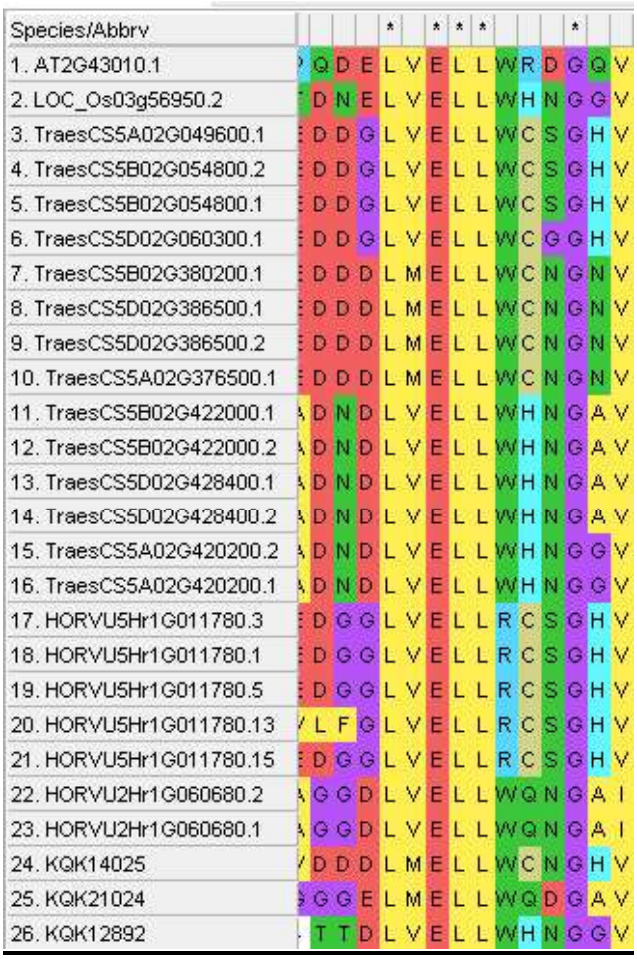


Figure 13: Multiple sequence analysis representing the conserved domain of APB.

A conserved motif APB was seen in all the protein sequences as shown in (Fig. 13). In dicots, PIF4 APB motif has a conserved amino acid 'Q' just after the conserved 'G'. In monocots, it can be seen that it is variable, this can be due to the mutations that occur during the course of evolution. The sequence was also checked to see if APA motif is present, but no conserved amino acid of this motif was found. This confirms that these sequence do not have a APA motif. APB motif is a chracteristic feature of PIF4, which can confirm that these following sequences obtained are candidates of orthologs of PIF4.

4.4.3 Phylogenetic analyis of TaPIFs

A phylogentic analysis was performed, using the peptide sequence of of AtPIF4, OsPIL13 and retrieved peptide sequence of the genes. The tree in (Fig. 14) shows the relationship between the genes. All the sequence of the wheat gene selected and OsPIL13 have a common ancestor with AtPIF4. In the tree TraesCS5B02G422000.2 is seen to have the most common recent ancestry with that of rice PIF. During the

course of evolution, genetic mutation in the sequence of organism has led to the minor change in their sequence of amino acid.

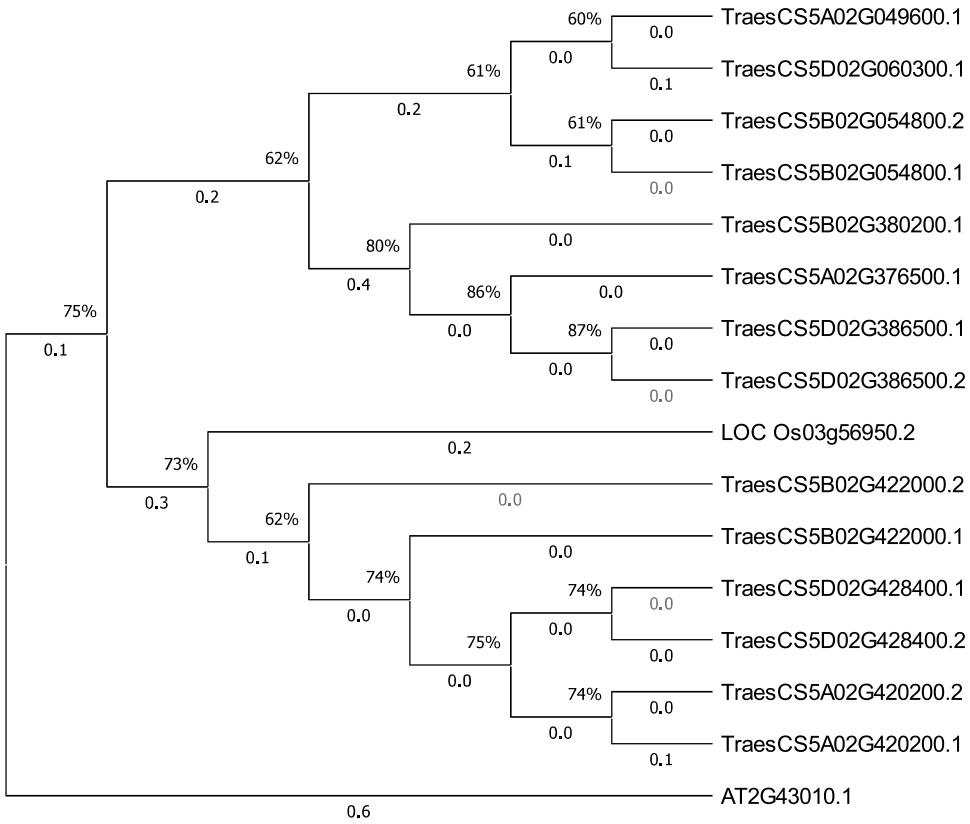


Figure 14: Phylogenetic tree of AtPIF4, OsPIL13 and genes of wheat obtained from blast analysis

4.4.4 Promoter analysis of *TaPIF*s and putative *TaPIF*4

The promoter analysis of wheat genes was performed to find out the elements present in the upstream of transcription initiation site (Table 10). The results showed that a large number of hormone responsive elements and light responsive elements. A total of 103 hormone responsive elements was reported, which are Methyl-jasmonate (TGACG-motif (25), CGTCA-motif (25)), Absiscic acid (ABRE-(39), Gibberellic acid (GARE-motif (4), P-box (2)), Auxin (TGA-motif (4), AuxRR-core (1)), Salicylic acid (TCA-motif (3)). A total of 93 light element were seen, with highest G-box element which was 50 in number, Sp1 (7), GTGGC-motif (1), GATA-motif (2), TCCC-motif (3), TCT-motif (3), AE-box (4), Gap-box (1), I-Box (4), MRE (3), ATCT-motif (1), Box-4 (2), Box-II (1), chs-Unit 1 m1 (1), ATC-motif (1). It is also interesting to note that PIFs show a preferential binding to G-box elements, and each of the genes had a G-box element.

Stress responsive elements were also seen and 9 ARE elements responsible for anaerobic induction, 7 GC- motif for anoxic inducibility, 4 MBS which are MYB binding elements responsible for drought inducibility. 3 LTRs for cold tolerance and 1 TC-rich repeat for defence and stress responsiveness. Apart from these elements 22 development regulatory element were also seen, RY-element (10) involved in seed specific regulation, CAT-box (8) this is an element which is related to meristem expression, circadian (2) involved in circadian regulation of plant. MAS-like element is involved in cell-cycle regulation and MBSI element responsible for flavonoid biosynthesis genes.

Cis-acting regulatory elements play an important role in the regulation of biological process. The promoter analysis of wheat revealed that they have almost all the elements that were seen in *Arabidopsis* and rice. This similarity in the presence of promoter elements in wheat gives a proof that may be the genes of wheat are putative PIF4 genes.

Table 5: Promoter elements present in TaPIFs

Light responsive	No. of elements	Hormone responsive element	No. of elements	stress responsive elements	No. of elements	Development regulatory elements	No. of elements
G-box	50	ABRE	39	ARE	9	RY-elements	10
Sp1	7	TGACG-motif	25	MBS	4	CAT-box	8
GTGGC-motif	1	CGTCA-motif	25	GC-motif	7	MBSI	1
GATA-motif	2	GARE-motif	4	LTR	3	MSA-like	1
TCCC-motif	3	P-box	2	TC-rich repeats	1	circadian	2
TCT-motif	3	TCA-motif	3				
AE-box	4	TGA-motif	4				
Gap-box	1	AuxRR-core	1				
I-box	4						
MRE	3						
ATCT-motif	1						
Box-4	2						
Box-II	1						
chs-Unit 1 m1	1						
ATC-motif	1						

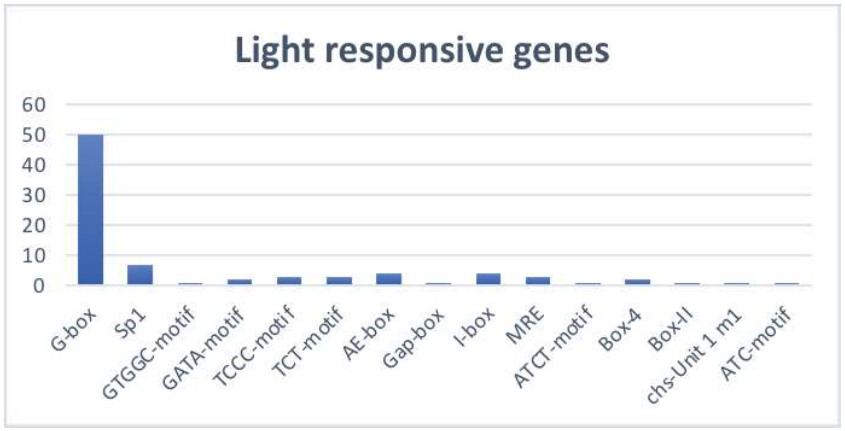


Figure 15: Representing the number of rice responsive elements present in wheat genes.

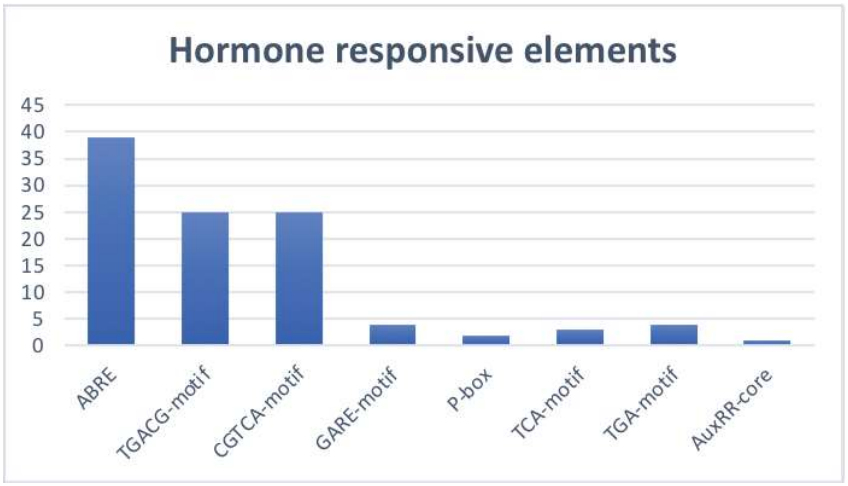


Figure 16: Representing the number of hormone responsive genes present in wheat genes.

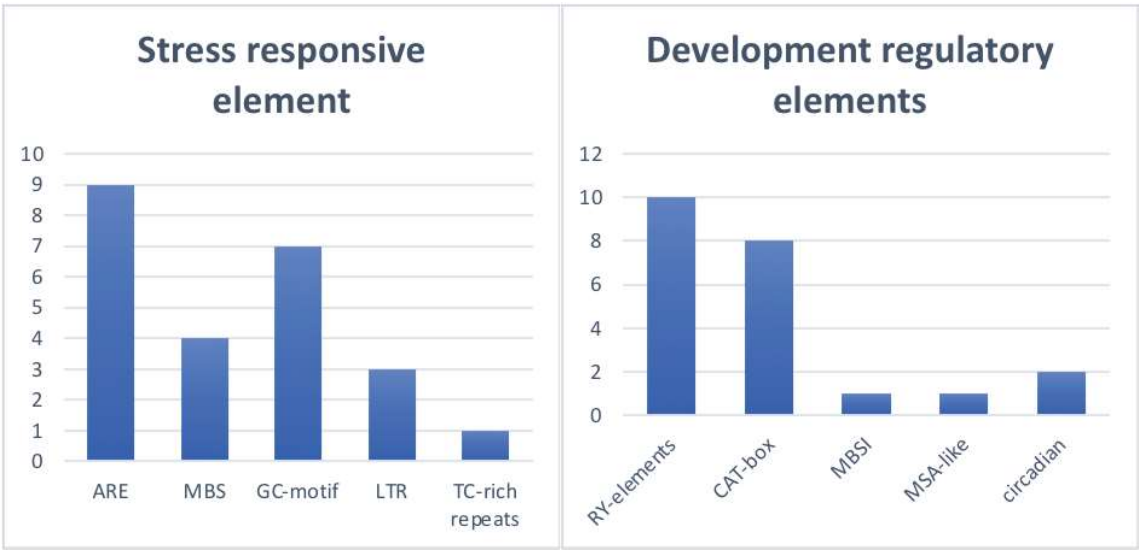


Figure 17: Representing the number of stress responsive and development regulatory elements present in wheat genes.

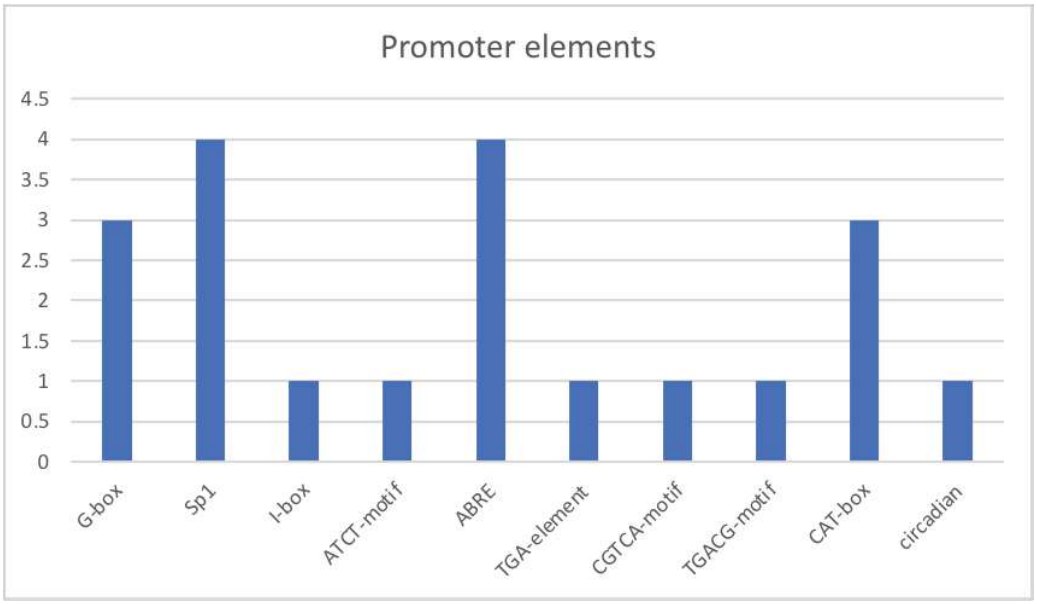


Figure 18: Following represents the number of promoter elements present in wheat putative *TaPIF4*

From the above various analysis TraesCS5B02G422000.2 was selected as the putative PIF4 in wheat. It had the most common recent relative with the rice PIF4 i.e., *OsPIL*13. Promoter analysis performed for the gene, gave the following result. Nine light responsive elements were seen, G-box elements (3), Sp1 (4), I-box (1), ATCT-motif (1). It is also a known fact that G-box elements have a preferential binding to the PIFs. Seven hormone responsive elements were found for abscisic acid (ABRE (4), methyl-jasmonate (CGTCA-motif (1), TGACG-motif (1)), auxin (TGA-element (1)). Development responsive elements were also seen CAT-box (3), *cis*-acting elements related to meristem expression and circadian (1) elements involved in the circadian control of the plant. Similar elements were also found in the PIFs of *Arabidopsis* and rice. The presence of development responsive element for meristem and auxin responsive element are essential component in the thermoresponsive growth of plant in which PIF4 is involved. PIF4 also plays a role in the transcriptional and post-transcriptional regulation of ABA level and the promoter has abscisic acid responsive element. No elements for stress response were seen (Table 6).

Table 6: Promoter elements in putative *TaPIF4*

	Elements	No. Of elements
Light responsive elements	G-box	3
	Sp1	4
	I-box	1
	ATCT-motif	1
Hormone responsive elements	ABRE	4
	TGA-element	1
	CGTCA-motif	1
	TGACG-motif	1
Development regulating elements	CAT-box	3
	circadian	1

5. Conclusion and Future Prospects

The work presented here gives a overall view of the PIFs in the plants. PIFs are bHLH transcription factors that are involved in the various process of development in the plants and studying about their mechanism and action becomes a crucial part in molecular biology to understand the underlying mechanism of plant development or stress tolerance in extreme conditions. The work gives a overview of of the various PIFs known in the plant kingdom and their fuctions in the plant. The phylogenetic analysis carried out gives an idea about the genes which got segregated in the process of evolution and the various functions of the PIFs that are conserved in the differnet plants. Promoter analysis gives an insight into the underlying regulation of PIFs by light, hromone, stress or structural development of plant.

Characterization of *TaPIF4* gives an insight into the presence of PIF4 in wheat. Multiple sequence analysis confirms the presence of APB domain in the sequences. This suggests that the genes obtained can be putative PIF4 genes. However, further investigation needs to be done to confirm the one-to-one interaction of these proteins using yeast-two hybrid and BiFC assay.

6. References:

Alabadí, D., Gil, J., Blázquez, M.A., and García-Martínez, J.L. (2004). Gibberellins repress photomorphogenesis in darkness. Plant Physiol **134**: 1050–1057.

Bae, G. and Choi, G. (2008). Decoding of light signals by plant phytochromes and their interacting proteins. Annu Rev Plant Biol **59**: 281–311.

Bertoni, G. (2020). MYB30 regulates photomorphogenesis via interactions with active phytochromes and PIFs. Plant Cell **32**: 2065–2066.

Botto, J.F., Sanchez, R.A., Whitelam, G.C., and Casal, J.J. (1996). Phytochrome A mediates the promotion of seed germination by very low fluences of light and canopy shade light in *Arabidopsis*. Plant Physiol **110**: 439–444.

Box, M.S. et al. (2015). ELF3 controls thermoresponsive growth in *Arabidopsis*. Curr Biol **25**: 194–199.

Casal, J.J. (2012). Shade avoidance. *Arabidopsis* Book **10**: e0157.

Chen, D., Xu, G., Tang, W., Jing, Y., Ji, Q., Fei, Z., and Lin, R. (2013). Antagonistic basic helix-loop-helix/bZIP transcription factors form transcriptional modules that integrate light and reactive oxygen species signaling in *Arabidopsis*. Plant Cell **25**: 1657–1673.

Cordeiro, A.M., Figueiredo, D.D., Tepperman, J., Borba, A.R., Lourenço, T., Abreu, I.A., Ouwerkerk, P.B.F., Quail, P.H., Margarida Oliveira, M., and Saibo, N.J.M. (2016). Rice phytochrome-interacting factor protein OsPIF14 represses OsDREB1B gene expression through an extended N-box and interacts preferentially with the active form of phytochrome B. Biochim Biophys Acta **1859**: 393–404.

Delker, C. et al. (2014). The DET1-COP1-HY5 pathway constitutes a multipurpose signaling module regulating plant photomorphogenesis and thermomorphogenesis. Cell Rep **9**: 1983–1989.

Favero, D.S., Jacques, C.N., Iwase, A., Le, K.N., Zhao, J., Sugimoto, K., and Neff, M.M. (2016). SUPPRESSOR OF PHYTOCHROME B4-#3 represses genes associated with auxin signaling to modulate hypocotyl growth. Plant Physiol **171**: 2701–2716.

Favero, D.S., Kawamura, A., Shibata, M., Takebayashi, A., Jung, J.-H., Suzuki, T., Jaeger, K.E., Ishida, T., Iwase, A., Wigge, P.A., Neff, M.M., and Sugimoto, K. (2020). AT-Hook transcription factors restrict petiole growth by antagonizing PIFs. Curr Biol **30**: 1454-1466.e6.

Fiorucci, A.-S., Galvão, V.C., Ince, Y.Ç., Boccaccini, A., Goyal, A., Allenbach Petrolati, L., Trevisan, M., and Fankhauser, C. (2020). PHYTOCHROME INTERACTING FACTOR 7 is important for early responses to elevated temperature in *Arabidopsis* seedlings. New Phytol **226**: 50–58.

Franklin, K.A., Lee, S.H., Patel, D., Kumar, S.V., Spartz, A.K., Gu, C., Ye, S., Yu, P., Breen, G., Cohen, J.D., Wigge, P.A., and Gray, W.M. (2011). Phytochrome-interacting factor 4 (PIF4) regulates auxin biosynthesis at high temperature. Proc Natl Acad Sci U S A **108**: 20231–20235.

Franklin, K.A. and Whitelam, G.C. (2005). Phytochromes and shade-avoidance responses in plants. Ann Bot **96**: 169–175.

Gao, Y., Wu, M., Zhang, M., Jiang, W., Ren, X., Liang, E., Zhang, D., Zhang, C., Xiao, N., Li, Y., Dai, Y., and Chen, J. (2018). A maize phytochrome-interacting factors protein ZmPIF1 enhances drought tolerance by inducing stomatal closure and improves grain yield in Oryza sativa. Plant Biotechnol J **16**: 1375–1387.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res **40**: D1178-1186.

Gu, D., Chen, C.-Y., Zhao, M., Zhao, L., Duan, X., Duan, J., Wu, K., and Liu, X. (2017). Identification of HDA15-PIF1 as a key repression module directing the transcriptional network of seed germination in the dark. Nucleic Acids Res **45**: 7137–7150.

Hayes, S., Sharma, A., Fraser, D.P., Trevisan, M., Cragg-Barber, C.K., Tavridou, E., Fankhauser, C., Jenkins, G.I., and Franklin, K.A. (2017). UV-B perceived by the UVR8 photoreceptor inhibits plant thermomorphogenesis. Curr Biol **27**: 120–127.

He, Y., Li, Y., Cui, L., Xie, L., Zheng, C., Zhou, G., Zhou, J., and Xie, X. (2016). Phytochrome B negatively affects cold tolerance by regulating OsDREB1 gene expression through phytochrome interacting factor-like protein OsPIL16 in rice. Front Plant Sci **7**: 1963.

Huq, E., Al-Sady, B., Hudson, M., Kim, C., Apel, K., and Quail, P.H. (2004). Phytochrome-interacting factor 1 is a critical bHLH regulator of chlorophyll biosynthesis. Science **305**: 1937–1941.

Itoh, H., Ueguchi-Tanaka, M., Sato, Y., Ashikari, M., and Matsuoka, M. (2002). The gibberellin signaling pathway is regulated by the appearance and disappearance of SLENDER RICE1 in nuclei. Plant Cell **14**: 57–70.

J, J., K, K., ME, K., HG, K., GS, H., OK, P., YI, P., G, C., and E, O. (2016). Phytochrome and ethylene signaling integration in *Arabidopsis* occurs via the transcriptional regulation of genes co-targeted by PIFs and EIN3. Frontiers in plant science **7**.

Ji, X., Du, Y., Li, F., Sun, H., Zhang, J., Li, J., Peng, T., Xin, Z., and Zhao, Q. (2019). The basic helix-loop-helix transcription factor, OsPIL15, regulates grain size via directly targeting a purine permease gene OsPUP7 in rice. Plant Biotechnol J **17**: 1527–1537.

Jia, Y. et al. (2020). PIFs coordinate shade avoidance by inhibiting auxin repressor ARF18 and metabolic regulator QQS. New Phytol **228**: 609–621.

Jiang, B., Shi, Y., Zhang, X., Xin, X., Qi, L., Guo, H., Li, J., and Yang, S. (2017). PIF3 is a negative regulator of the CBF pathway and freezing tolerance in *Arabidopsis*. Proc Natl Acad Sci U S A **114**: E6695–E6702.

Jiang, Y., Yang, C., Huang, S., Xie, F., Xu, Y., Liu, C., and Li, L. (2019). The ELF3-PIF7 interaction mediates the circadian gating of the shade response in *Arabidopsis*. iScience **22**: 288–298.

Kathare, P.K., Xu, X., Nguyen, A., and Huq, E. (2020). A COP1-PIF-HEC regulatory module fine-tunes photomorphogenesis in *Arabidopsis*. Plant J **104**: 113–123.

Khanna, R., Shen, Y., Marion, C.M., Tsuchisaka, A., Theologis, A., Schäfer, E., and Quail, P.H. (2007). The basic helix-loop-helix transcription factor PIF5 acts on ethylene biosynthesis and phytochrome signaling by distinct mechanisms. Plant Cell **19**: 3915–3929.

Kim, C., Kim, S.J., Jeong, J., Park, E., Oh, E., Park, Y.-I., Lim, P.O., and Choi, G. (2020). High ambient temperature accelerates leaf senescence via PHYTOCHROME-INTERACTING FACTOR 4 and 5 in *Arabidopsis*. Mol Cells **43**: 645–661.

Kudo, M., Kidokoro, S., Yoshida, T., Mizoi, J., Todaka, D., Fernie, A.R., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2017). Double overexpression of DREB and PIF transcription factors improves drought stress tolerance and cell elongation in transgenic plants. Plant Biotechnol J **15**: 458–471.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol **35**: 1547–1549.

Kumar, S.V., Lucyshyn, D., Jaeger, K.E., Alós, E., Alvey, E., Harberd, N.P., and Wigge, P.A. (2012). Transcription factor PIF4 controls the thermosensory activation of flowering. Nature **484**: 242–245.

Lee, N. and Choi, G. (2017). Phytochrome-interacting factor from *Arabidopsis* to liverwort. Curr Opin Plant Biol **35**: 54–60.

Leivar, P., Martín, G., Soy, J., Dalton-Roesler, J., Quail, P.H., and Monte, E. (2020). Phytochrome-imposed inhibition of PIF7 activity shapes photoperiodic growth in *Arabidopsis* together with PIF1, 3, 4 and 5. Physiol Plant **169**: 452–466.

Leivar, P. and Monte, E. (2014). PIFs: systems integrators in plant development. Plant Cell **26**: 56–78.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P., and Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res **30**: 325–327.

Li, K., Yu, R., Fan, L.-M., Wei, N., Chen, H., and Deng, X.W. (2016). DELLA-mediated PIF degradation contributes to coordination of light and gibberellin signalling in *Arabidopsis*. Nat Commun **7**: 11868.

Li, L. et al. (2012). Linking photoreceptor excitation to changes in plant architecture. Genes Dev **26**: 785–790.

Li, X. et al. (2006). Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and *Arabidopsis*. Plant Physiol **141**: 1167–1184.

Liang, S., Gao, X., Wang, Y., Zhang, H., Yin, K., Chen, S., Zhang, M., and Zhao, R. (2020). Phytochrome-interacting factors regulate seedling growth through ABA signaling. Biochem Biophys Res Commun **526**: 1100–1105.

Liu, Z., Zhang, Y., Liu, R., Hao, H., Wang, Z., and Bi, Y. (2011). Phytochrome interacting factors (PIFs) are essential regulators for sucrose-induced hypocotyl elongation in *Arabidopsis*. J Plant Physiol **168**: 1771–1779.

Llorente, B., D'Andrea, L., Ruiz-Sola, M.A., Botterweg, E., Pulido, P., Andilla, J., Loza-Alvarez, P., and Rodriguez-Concepcion, M. (2016). Tomato fruit carotenoid biosynthesis is adjusted to actual ripening progression by a light-dependent mechanism. Plant J **85**: 107–119.

Lorrain, S., Allen, T., Duek, P.D., Whitelam, G.C., and Fankhauser, C. (2008). Phytochrome-mediated inhibition of shade avoidance involves degradation of growth-promoting bHLH transcription factors. Plant J **53**: 312–323.

Lymperopoulos, P., Msanne, J., and Rabara, R. (2018). Phytochrome and phytohormones: working in tandem for plant growth and development. Front Plant Sci **9**: 1037.

Ma, D., Li, X., Guo, Y., Chu, J., Fang, S., Yan, C., Noel, J.P., and Liu, H. (2016). Cryptochrome 1 interacts with PIF4 to regulate high temperature-mediated hypocotyl elongation in response to blue light. Proc Natl Acad Sci U S A **113**: 224–229.

MA, K., L, A., T, A., CA, T., NP, H., GC, W., and KA, F. (2009). High temperature-mediated adaptations in plant architecture require the bHLH transcription factor PIF4. Current biology: CB **19**.

Mo, W., Tang, W., Du, Y., Jing, Y., Bu, Q., and Lin, R. (2020). PHYTOCHROME-INTERACTING FACTOR-LIKE14 and SLENDER RICE1 interaction controls seedling growth under salt stress. Plant Physiol **184**: 506–517.

Monte, E. (2020). Plant Biology: AHL transcription factors inhibit growth-promoting PIFs. Curr Biol **30**: R354–R356.

Monte, E., Tepperman, J.M., Al-Sady, B., Kaczorowski, K.A., Alonso, J.M., Ecker, J.R., Li, X., Zhang, Y., and Quail, P.H. (2004). The phytochrome-interacting transcription factor, PIF3, acts early, selectively, and positively in light-induced chloroplast development. Proc Natl Acad Sci U S A **101**: 16091–16098.

Moon, J., Zhu, L., Shen, H., and Huq, E. (2008). PIF1 directly and indirectly regulates chlorophyll biosynthesis to optimize the greening process in *Arabidopsis*. Proc Natl Acad Sci U S A **105**: 9433–9438.

Nakamura, Y., Kato, T., Yamashino, T., Murakami, M., and Mizuno, T. (2007). Characterization of a set of phytochrome-interacting factor-like bHLH proteins in Oryza sativa. Biosci Biotechnol Biochem **71**: 1183–1191.

Ni, M., Tepperman, J.M., and Quail, P.H. (1998). PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein. Cell **95**: 657–667.

Nusinow, D.A., Helfer, A., Hamilton, E.E., King, J.J., Imaizumi, T., Schultz, T.F., Farré, E.M., and Kay, S.A. (2011). The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. Nature **475**: 398–402.

Oh, E., Kang, H., Yamaguchi, S., Park, J., Lee, D., Kamiya, Y., and Choi, G. (2009). Genome-wide analysis of genes targeted by PHYTOCHROME INTERACTING FACTOR 3-LIKE5 during seed germination in *Arabidopsis*. Plant Cell **21**: 403–419.

Oh, E., Yamaguchi, S., Hu, J., Yusuke, J., Jung, B., Paik, I., Lee, H.-S., Sun, T., Kamiya, Y., and Choi, G. (2007). PIL5, a phytochrome-interacting bHLH protein, regulates gibberellin responsiveness by binding directly to the GAI and RGA promoters in *Arabidopsis* seeds. Plant Cell **19**: 1192–1208.

Oh, E., Yamaguchi, S., Kamiya, Y., Bae, G., Chung, W.-I., and Choi, G. (2006). Light activates the degradation of PIL5 protein to promote seed germination through gibberellin in *Arabidopsis*. Plant J **47**: 124–139.

Oh, E., Zhu, J.-Y., and Wang, Z.-Y. (2012). Interaction between BZR1 and PIF4 integrates brassinosteroid and environmental responses. Nat Cell Biol **14**: 802–809.

Oh, J., Park, E., Song, K., Bae, G., and Choi, G. (2020). PHYTOCHROME INTERACTING FACTOR8 inhibits phytochrome A-mediated far-red light responses in *Arabidopsis*. Plant Cell **32**: 186–205.

Oyama, T., Shimura, Y., and Okada, K. (1997). The *Arabidopsis* HY5 gene encodes a bZIP protein that regulates stimulus-induced development of root and hypocotyl. Genes Dev **11**: 2983–2995.

Paik, I., Chen, F., Ngoc Pham, V., Zhu, L., Kim, J.-I., and Huq, E. (2019). A phyB-PIF1-SPA1 kinase regulatory complex promotes photomorphogenesis in *Arabidopsis*. Nat Commun **10**: 4216.

Park, H.J., Kim, W.-Y., and Yun, D.-J. (2016). A New insight of salt stress signaling in Plant. Mol Cells **39**: 447–459.

Park, Y.-J., Kim, J.Y., Lee, J.-H., Lee, B.-D., Paek, N.-C., and Park, C.-M. (2020). GIGANTEA shapes the photoperiodic rhythms of thermomorphogenic growth in *Arabidopsis*. Mol Plant **13**: 459–470.

Pedmale, U.V., Huang, S.-S.C., Zander, M., Cole, B.J., Hetzel, J., Ljung, K., Reis, P.A.B., Sridevi, P., Nito, K., Nery, J.R., Ecker, J.R., and Chory, J. (2016). Cryptochromes interact directly with PIFs to control plant growth in limiting blue light. Cell **164**: 233–245.

Penfield, S., Josse, E.-M., and Halliday, K.J. (2010). A role for an alternative splice variant of PIF6 in the control of *Arabidopsis* primary seed dormancy. Plant Mol Biol **73**: 89–95.

Qiu, Y. (2020). Regulation of PIF4-mediated thermosensory growth. Plant Sci **297**: 110541.

Qiu, Y., Li, M., Kim, R.J.-A., Moore, C.M., and Chen, M. (2019). Daytime temperature is sensed by phytochrome B in *Arabidopsis* through a transcriptional activator HEMERA. Nat Commun **10**: 140.

Qiu, Y., Li, M., Pasoreck, E.K., Long, L., Shi, Y., Galvão, R.M., Chou, C.L., Wang, H., Sun, A.Y., Zhang, Y.C., Jiang, A., and Chen, M. (2015). HEMERA couples the proteolysis and transcriptional activity of PHYTOCHROME INTERACTING FACTORs in *Arabidopsis* photomorphogenesis. Plant Cell **27**: 1409–1427.

Rosado, D., Gramegna, G., Cruz, A., Lira, B.S., Freschi, L., de Setta, N., and Rossi, M. (2016). Phytochrome Interacting Factors (PIFs) in Solanum lycopersicum: diversity, evolutionary history and expression profiling during different developmental processes. PLoS One **11**: e0165929.

Sakuraba, Y., Jeong, J., Kang, M.-Y., Kim, J., Paek, N.-C., and Choi, G. (2014). Phytochrome-interacting transcription factors PIF4 and PIF5 induce leaf senescence in *Arabidopsis*. Nat Commun **5**: 4636.

Sakuraba, Y., Kim, E.-Y., Han, S.-H., Piao, W., An, G., Todaka, D., Yamaguchi-Shinozaki, K., and Paek, N.-C. (2017). Rice Phytochrome-Interacting Factor-Like1 (OsPIL1) is involved in the promotion of chlorophyll biosynthesis through feed-forward regulatory loops. J Exp Bot **68**: 4103–4114.

Si, L. et al. (2016). OsSPL13 controls grain size in cultivated rice. Nat Genet **48**: 447–456.

Suárez-López, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., and Coupland, G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in *Arabidopsis*. Nature **410**: 1116–1120.

Sun, J., Qi, L., Li, Y., Zhai, Q., and Li, C. (2013). PIF4 and PIF5 transcription factors link blue light and auxin to regulate the phototropic response in *Arabidopsis*. Plant Cell **25**: 2102–2114.

Thines, B.C., Youn, Y., Duarte, M.I., and Harmon, F.G. (2014). The time-of-day effects of warm temperature on flowering time involve PIF4 and PIF5. J Exp Bot **65**: 1141–1151.

Wang, S., Wu, K., Yuan, Q., Liu, X., Liu, Z., Lin, X., Zeng, R., Zhu, H., Dong, G., Qian, Q., Zhang, G., and Fu, X. (2012). Control of grain size, shape and quality by OsSPL16 in rice. Nat Genet **44**: 950–954.

Wigge, P.A. (2013). Ambient temperature signalling in plants. Curr Opin Plant Biol **16**: 661–666.

Wu, G., Zhao, Y., Shen, R., Wang, B., Xie, Y., Ma, X., Zheng, Z., and Wang, H. (2019). Characterization of maize phytochrome-interacting factors in light signaling and photomorphogenesis. Plant Physiol **181**: 789–803.

Xie, Y., Liu, Y., Wang, H., Ma, X., Wang, B., Wu, G., and Wang, H. (2017). Phytochrome-interacting factors directly suppress MIR156 expression to enhance shade-avoidance syndrome in *Arabidopsis*. Nat Commun **8**: 348.

Xu, X., Paik, I., Zhu, L., Bu, Q., Huang, X., Deng, X.W., and Huq, E. (2014). PHYTOCHROME INTERACTING FACTOR1 enhances the E3 ligase activity of CONSTITUTIVE PHOTOMORPHOGENIC1 to synergistically repress photomorphogenesis in *Arabidopsis*. Plant Cell **26**: 1992–2006.

Yan, Y. et al. (2020). MYB30 Is a key negative regulator of *Arabidopsis* photomorphogenic development that promotes PIF4 and PIF5 protein accumulation in the light. Plant Cell **32**: 2196–2215.

Zhang, K., Zheng, T., Zhu, X., Jiu, S., Liu, Z., Guan, L., Jia, H., and Fang, J. (2018). Genome-wide identification of PIFs in grapes (Vitis vinifera L.) and their transcriptional analysis under lighting/shading conditions. Genes (Basel) **9.**

Zhang, Y., Liu, Z., Chen, Y., He, J.-X., and Bi, Y. (2015). PHYTOCHROME-INTERACTING FACTOR 5 (PIF5) positively regulates dark-induced senescence and chlorophyll degradation in *Arabidopsis*. Plant Sci **237**: 57–68.

Zhang, Y., Liu, Z., Wang, L., Zheng, S., Xie, J., and Bi, Y. (2010). Sucrose-induced hypocotyl elongation of *Arabidopsis* seedlings in darkness depends on the presence of gibberellins. J Plant Physiol **167**: 1130–1136.

Zhao, J., Favero, D.S., Peng, H., and Neff, M.M. (2013). *Arabidopsis* thaliana AHL family modulates hypocotyl growth redundantly by interacting with each other via the PPC/DUF296 domain. Proc Natl Acad Sci U S A **110**: E4688-4697.

Zheng, P.-F., Wang, X., Yang, Y.-Y., You, C.-X., Zhang, Z.-L., and Hao, Y.-J. (2020). Identification of phytochrome-interacting factor family members and functional analysis of MdPIF4 in Malus domestica. Int J Mol Sci **21.**

Zhu, L., Bu, Q., Xu, X., Paik, I., Huang, X., Hoecker, U., Deng, X.W., and Huq, E. (2015). CUL4 forms an E3 ligase with COP1 and SPA to promote light-induced degradation of PIF1. Nat Commun **6**: 7245.

Expression analysis of heat stress regulated genes at diverse developmental stages by *in silico* tools and role of Hsp101 in metal stress response by wet laboratory experiments in *Arabidopsis thaliana*

Thesis Submitted to the University of Delhi

in the Partial fulfilment for the Degree of

Master of Science

in

Plant Molecular Biology and Biotechnology

2021



Shitij Gupta

Department of Plant Molecular Biology

University of Delhi South Campus

New Delhi 110021, India

Expression analysis of heat stress regulated genes at diverse developmental stages by *in silico* tools and role of Hsp101 in metal stress response by wet laboratory experiments in *Arabidopsis thaliana*

Thesis Submitted to the University of Delhi

in the Partial fulfilment for the Degree of

Master of Science

in

Plant Molecular Biology and Biotechnology

2021

Shitij Gupta

Department of Plant Molecular Biology

University of Delhi South Campus

New Delhi 110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**Expression analysis of heat stress regulated genes at diverse developmental stages by *in silico* tools and role of Hsp101 in metal stress response by wet laboratory experiments in *Arabidopsis thaliana***" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 28th June 2021

Signature of the Candidate

(SHITIJ GUPTA)

Signature of the Supervisor

(PROF. ANIL GROVER)

Signature of the HOD

Type text here

(PROF. ANIL GROVER

Preface

Around 11,000 worldwide weather-related catastrophic calamities has been reported since 1996 till 2015 (Global Climate Risk Index 2017.). Previous decades report distinctive increase in the median global sea and surface temperature in both the northern and southern hemispheres. Continuous rise in temperature has led to 9.1% deficit in national production of U.S.A.(Lesk et al., 2016). Also the average Indian temperature has been predicted to increase by 1.89°C between 2040 and 2059 (https://climateknowledgeportal.worldbank.org/country/india/climate-data-projections). Continuous shift in climatic conditions along with falling production and rising population has led to immense pressure on the agricultural sector for producing more food. Therefore, the need of the hour is to achieve food security under such uncertain climatic conditions.

Apart from rising temperatures, plants are affected severely by heavy metals worldwide. Owing to our activities, there is a boost in the discharge of pollutants in the environment. Enormous industrial development and urbanization led to a steep rise in the emanation of heavy metals thus polluting the soil. Soil functions as a crucial component for the functioning of terrestrial ecosystems. It serves as the nutrient pool for the growth of plants. Thus, accumulation of toxic substances in the soil has critically affected the quality of crops. Plants play a crucial role in the pioneering stages of the food chain hence exposure with heavy metals may biomagnify and results in fatal disorders in humans.

Variety of these abiotic stress forces us to develop resilient crops to catch up with the ever-increasing population.

In our own country, many states (West Bengal, Punjab, Orissa, etc) are drinking and using water contaminated with heavy metals, usage of this polluted water has resulted in many forms of cancers and human health disorders.

Heat shock proteins and heat shock factors are produced as a common phenomenon in plants upon encountering stress. These proteins help in managing aggregated and misfolded proteins during stress and recovery periods. HSPs and HSFs are well-defined warriors of high temperature stress.

The focus of this thesis was to analyse the tissue-specific and spatio-temporal expression pattern of all HSPs and Hsfs in *Arabidopsis thaliana.* We observed that out of the total family members of HSPs and Hsfs only few of them appears to be of greater importance. Since, cytosolic AtHSP101 is a well-accepted molecule for conferring thermotolerance, we wanted to study its role in other stresses. The latter part of the thesis focussed on the role of HSP101 during metal stress. It is reported as the only chaperone that helps in reactivation of aggregated proteins apart from refolding mis-folded proteins. Also, this protein has been observed to be induced during heavy metal stresses thus pointing towards its role.

Arabidopsis thaliana (Col-0 ecotype) and *hot1-3* (HSP101 null mutant) were employed to characterize the role of HSP101 at phenotypic level with respect to basal seed tolerance during arsenic and cadmium stress. We believe that AtHSP101 has an important role in maintaining root length during As and Cd stress. Responses mediated by this gene under heavy metal stress has to be elucidated for its consideration to be used as a candidate gene for heavy metal stress tolerance.

Acknowledgements

This dissertation would not have been possible without the help of many people in so many ways. It was also the product of a large measure of serendipity and fortuitous encounters with people that have changed the course of my academic career. I would take this opportunity to express my gratitude towards them.

First and foremost, I am highly grateful to my supervisor **Prof. Anil Grover** for his guidance, support, kind advice and caring nature throughout the course of this dissertation. I am grateful for his supervision, time and ideas regarding this thesis and for giving me complete freedom to think and work. I am highly grateful for his patience that he showed in teaching me even the basic concepts. The joy and enthusiasm he has for research are motivational for me and I would like to sincerely thank him for making every difficult task achievable and supporting me in every possible way.

I would like to thank, all my teachers, **Prof. A.K. Tyagi, Prof. J.P. Khurana, Prof. P. Khurana, Prof. Indranil Dasgupta, Prof. Sanjay Kapoor, Prof. Arun K. Sharma, Prof. Girdhar K. Pandey, Prof. Saurabh Raghuvanshi** and **Prof. Surekha Katiyar-Agarwal** for sharing their vast scientific experience and help me develop a scientific temperament.

I wish to express my sincere thanks to **Gayatri Tripathi** for her support, guidance and valuable time. I thoroughly appreciate her positive, patient and generous behaviour in making the working environment very light and easy-going. I sincerely thank her for her support and suggestions at various points of my work especially during the compilation of this thesis.

I am also thankful to **Rinchuila Shimphrui** for her patience and cooperation during the bioinformatic experimentation and compilation of my thesis. I thank her for her support and guidance throughout my stay in the lab.

I also wish to thank **Richa, Lisha, Garima and Manoj** for being exceptionally helpful. I thank **Dr. Chhavi Agrawal** for their very nice and fruitful cooperation during my lab tenure. I would like to thanks to **Dr. Neelam Sarkar** for her immense support and help during compilation of my thesis.

I would like to thank **Aditi** for providing a wonderful environment in the lab to work in and all the fun moments we had. **Brij Kishor** and **Vakil** also deserve special thanks for keeping things in order in the lab and especially providing miscellaneous items on time. I also thank my friends **Apurva, Nitika, Karishma, Sritama, Thanduanlung, Atul, Aashmeen, Gopika**, **Nidhi and Meher;** and my seniors **Rohit Kumar** and **Barkha Ravi** for their supportive nature and all the great times we had together.

I am also thankful to **Pratibha Gour** for her help with the RNA-seq data analysis.

I would like to thank **Sahibjot Kaur** for helping me out with the graphing software and data representation.

Lastly, I thank Almighty and my Family for their constant encouragement, love and support without which this assignment would not have been possible.

List of Abbreviations

Abbreviation	Definition
As	Arsenic
°C	Degree centigrade
Cd	Cadmium
ClpB-C	Caesinolytic protease B/ cytoplasmic form
g	Grams
h	Hour
HS	Heat stress
HSFs	Heat shock transcription factors
HSPs	Heat shock proteins
HSR	Heat stress responses
kDa	Kilodaltons
HM	Heavy metals
L	Litre
µg	Microgram
mg	Milligram
µl	Microlitre
µM	Micromolar
min	Minutes
NTD	N-terminal Domain
OE	Overexpression line
PIFs	Phytochrome interacting factors
RNA	Ribonucleic acids
UTR	Untranslated regions
Zt	Zeitgeber time

Table of Contents

Certificate……i
Preface ii
Acknowledgements.....iii
List of Abbreviations iv
Table of Contents v
Review of literature 1
Material and methods 18
Results and Discussion..... 21
Conclusions and future work..... 77
References 79

Review of literature



The sessile nature of plants makes them prone to environmental stress more than animals. This immobility forces plants to evolve mechanisms to tolerate or/and to fight back this continuous battle with abiotic and biotic agents of the environment. Heat stress and heavy metal(loid)s stress are some of the many known evils affecting crop productivity. As a consequence of our doings, industrialization, urbanization and their related activities such as coal-burning (rapid increment in atmospheric CO_2 levels), use of agricultural additives, etc has resulted in substantial increment in the mean global temperature and rapid emanation of heavy metals in the environment. Average temperature anomalies from 1850 till march 2021 relative to 1961-1990 (HadCRUT version 5.0.1.0, Morice et al., 2021) is represented in the form of a heat map and a median scatter plot (Fig. 1.1). United States Agency of Toxic Substances and Disease Registry (ATSDR) released a 'Special Priority List 2019' which ranked arsenic and cadmium at 1st and 7th position respectively in terms of their potentiality, toxicity and frequency of exposure to humans. (https://www.atsdr.cdc.gov/spl/in dex.html#2019spl).

1.1 Heat Stress

One of the most essential environmental factors is the temperature. It not only affects the diurnal rhythms and seasonal growth but also determines the geographical distribution of plants. Slight change in temperature (i.e., warming stress) results in morphological variations (thermomorphogenesis) whereas a drastic change in the temperature results in cellular damage and ultimately death. Climate change due to global warming has resulted in catastrophic losses in the area of food security and crop production (Lobell et al., 2011; Lesk et al., 2016). Under natural conditions, heat stress (HS) is accompanied by a variety of different stresses such as drought stress and light stress etc. The first and foremost step in tackling the HS is to understand the mechanism by which plants perceive and respond [Heat Stress Response (HSR)] to it. This knowledge can help in engineering plants that are tolerant to heat stress.

The temperature being a physical factor mingles with the cellular membrane and cell interior at the same time. Ergo, perception of high temperature can be anticipated at multiple locations within the cell and results in a plethora of responses (Li et al., 2019). These perception mechanisms differ depending on the magnitude of temperature.

At ambient elevated temperature (up to 30°C), Phytochrome B, PIF7 and ELF3 act as a primary sensor. Temperature-dependent reversal of P_{fr} form to P_r form occurs which releases the repression caused by the P_{fr} form of phytochrome B on PIF4 and PIF7 activity. This warming temperature also activates the translation of PIF7 mRNA via disturbing the 5 UTR hairpin loop structure. At warming temperatures, repression of PIF4 transcription is also relieved via

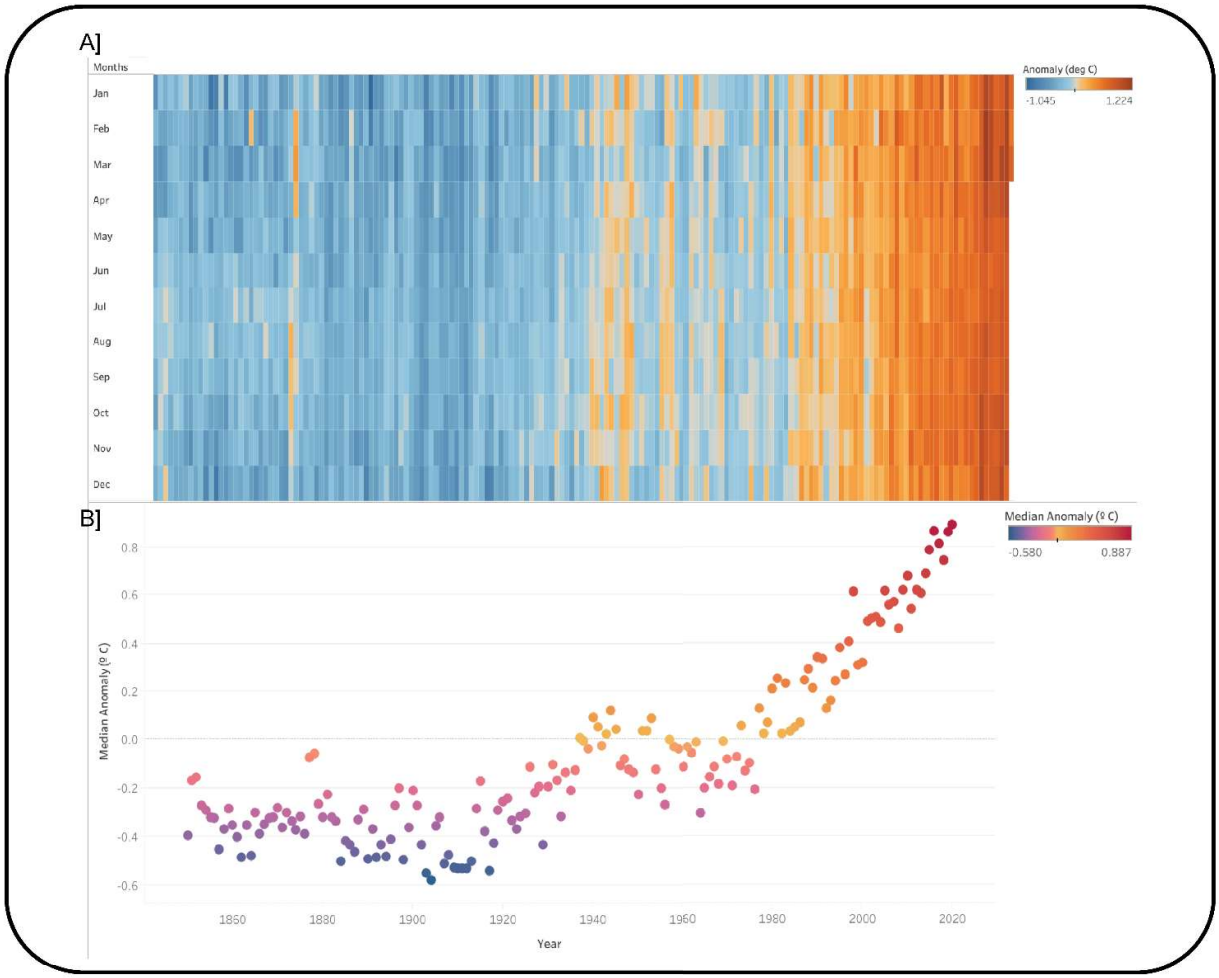


Fig.1.1. Average temperature anomalies from 1850 till march 2021 relative to 1961-1990 (HadCRUT version 5.0.1.0, Morice et al., 2021).A] Heat map and B} Median scatter plot.

speckle formation of ELF3 proteins in evening complex. After release from the repression and translational activation of PIF7, these PIFs activates the transcription of thermomorphogenic genes (Hayes et al., 2020).

During HS, in principle all biomacromolecules 'senses' heat and can act as a thermosensor owing to the large surface area to volume ratio of *Arabidopsis* leaves. But some of these temperature affected molecules can act as a primary heat sensor (Mittler et al., 2012). Sensing occurs at various levels of the cell such as the plasma membrane. HS dependent influx of calcium ions is mediated by the channels such as annexin1 (potentially) or CNGC 16 as in the case of pollen grains (Rahmati Ishka et al., 2018). Due to increment in the fluidity of plasma membrane under HS, membrane-associated adenyl cyclase may get activated to convert ATP to cAMP which leads to the activation of these calcium channels (CNGC 6) to import calcium from the apoplast. Intracellular calcium along with the calmodulin might inhibit the action of CNGCs. Increased concentration of calcium is attributable to activation of lipid rafts associated NADPH oxidase, RBOHD which increase the H_2O_2 concentration near the plasma membrane and it enters symplast via aquaporins. H_2O_2 results in structural changes in the PLDδ enzyme and promotes its binding with calcium. This calcium bound form of PLDδ causes depolymerization of microtubules which in turn activates the MAPK signalling pathways and results in the indirect activation of Heat shock transcription factors (Hsfs). Another lipid-based signalling pathway also shapes the heat stress signalling pathways (reviewed in (Hayes et al., 2020). With increasing temperatures plants starts to acclimatize with the response, and show many different variable sensing mechanisms. Major targets of well-known HS responsive TF genes (Hsfs) are Heat shock proteins (HSPs) and ROS scavenging genes (Ohama et al., 2017).

1.1.1 Warriors of HS

1.1.1.1 Plant heat shock transcription factors –

These transcription factors are the ultimate targets of sensor pathways for many types of stresses (including heat and heavy metal(loid)s stress). They act by binding to the conserved cis-acting heat stress elements (HSEs) found in the promoters of high temperature stress inducible genes (von Koskull-Döring et al., 2007). Like most gene regulatory proteins, these factors also possess modular structure. Plant Hsfs typically comprises of a N-terminal DNA binding domain, oligomerization domain, nuclear localization signal (NLS), nuclear export signal (NES) and a C-terminal activation domain. N-terminal DNA binding domain consists of a helix-turn-helix motif for binding to HSEs in promoter regions of target genes. The oligomerization domain (OD) is connected to the N-terminal DNA binding domain with a flexible linker region. OD consists of the HR-A/B region (repetitive heptad pattern hydrophobic amino acid residues) which forms the coiled-coil domains for protein-protein interactions

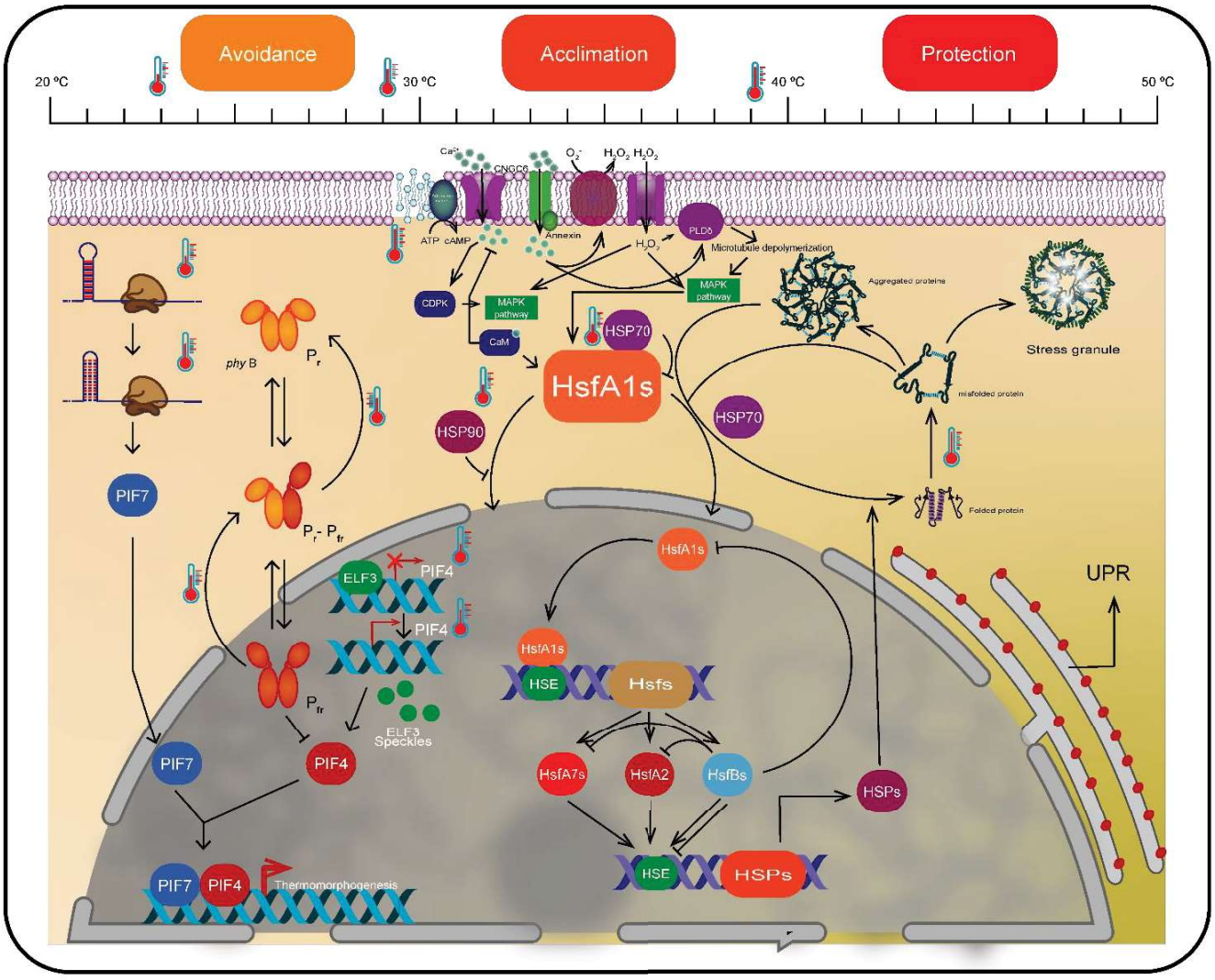


Fig.1.2. Model depicting HSR in different temperature ranges. Modified from Hayes et al., 2020 and Ohama et al., 2017.

This region is followed by the presence of a NLS and NES region. C-terminal activation domain consists of activator motifs (AHA motifs) rich in aromatic, large hydrophobic and acidic amino acid residues (Guo et al., 2016; Scharf et al., 2012; von Koskull-Döring et al., 2007; Nover et al., 2001).

Hsfs are classified into 3 classes, namely- Class A, Class B and Class C Hsfs. This classification is based on the differences in the HR-A/B region of OD and the length of the linker between the NTD and OD. Hsf As shows 21 amino acid insertion in the HR-A/B region resulting in extended OD and AHA motifs in C-terminal activation domains. Hsf C shows 7 amino acid insertion in the HR-A/B region resulting in extended OD. Whereas Hsf B comprises compact OD and repression domain is specific to this class (Scharf et al., 2012). Arabidopsis has 21 known Hsfs, listed in the table1.1.

Table1.1 Hsfs in *Arabidopsis thaliana* (Nover et al., 2001)

Class of Hsfs	**Gene ID**	**Gene Name**
Class A (15)	AT4G17750	HsfA1a
	AT5G16820	HsfA1b
	AT1G32330	HsfA1d
	AT3G02990	HsfA1e
	AT2G26150	HsfA2
	AT4G18880	HsfA4a
	AT5G45710	HsfA4c
	AT4G13980	HsfA5
	AT3G51910	HsfA7a
	AT3G63350	HsfA7b
	AT1G67970	HsfA8
	AT5G03720	HsfA3
	AT5G43840	HsfA6a
	AT3G22830	HsfA6b
	AT5G54070	HsfA9
Class B (5)	AT4G36990	HsfB1
	AT5G62020	HsfB2a
	AT4G11660	HsfB2b
	AT1G46264	HsfB4
	AT2G41690	HsfB3
Class C (1)	AT3G24520	HsfC1

Major physiological functions of these genes are summarized by Yabuta, 2016. Some important functions of Hsfs in plants are as follows-

1. HsfA1s functions as "master regulators" in a variety of stresses. In *Solanum lycopersicum,* the role of HsfA1a is indispensable for conferring thermotolerance. But in *Arabidopsis thaliana*, due to the presence of the many functional redundant forms, master regulator is not defined.
2. Owing to the ability to increase the expression of HSP genes and its heavy accumulation when exposed to heat stress, HsfA2 is reported to function as a signal enhancer in thale cress.

1.1.1.2 Plant heat shock proteins-

The first observation of HSR was reported as changes in expression of proteins due to puffing in the chromosome of *Drosophila melanogaster* on exposure to elevated temperature and various other chemicals such as 2,4-DNP (Ritossa, 1962). But the name heat-shock proteins was given by Tissières et al., 1974. Heat shock proteins are induced by a variety of environmental stresses including metal stress and ethanol treatment (Sanchez et al., 1992). HSPs are generally categorised into different families based on their molecular weights.

sHSPs family-

- The ancient family found in archaea, eubacteria and eukarya.
- It is composed of an N-terminal variable region, α- crystallin domain and a C-terminal extension domain. Organelle specific sHSPs also possess transit peptide at the N-terminal region (Waters, 2013).
- They are divided into 6 classes based on immune-cross reactive reaction, sequence alignment and cellular compartmentalization.
- Localised in the cytosol, chloroplast, mitochondria and endoplasmic reticulum (Yadav et al., 2020).
- Some role of this family are as follows-
    1. This family plays important role in stress tolerance at the seed stage (Avelange-Macherel et al., 2015).

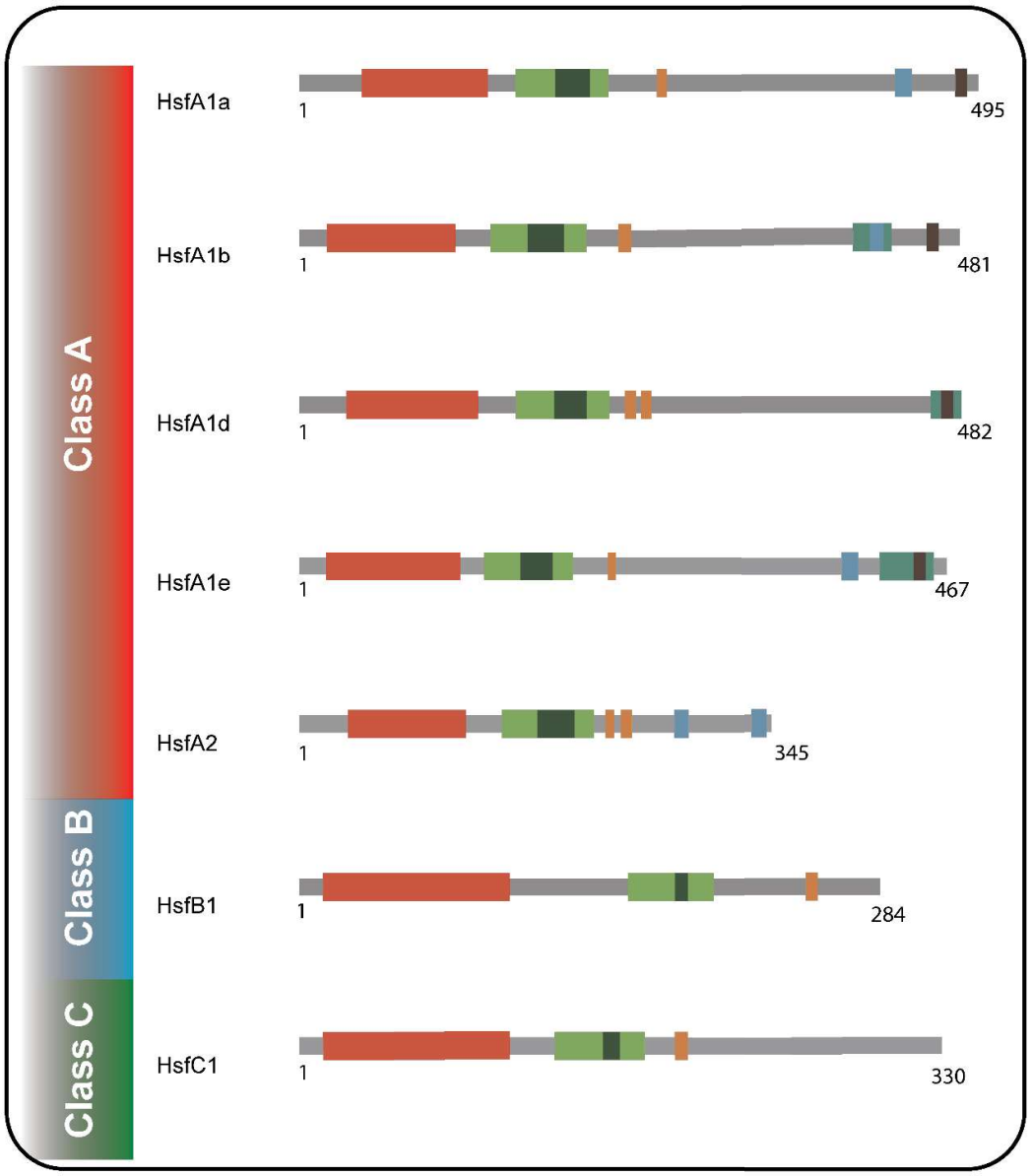


Fig.1.3. Major domains of some members of different classes of Hsfs

2. Overexpression of HSP24.7 enhances seed germination (Waters and Vierling, 2020).

HSP 60 family-

- Based on sequence variability, this family of proteins is divided further into 2 categories: group I and group II. Group I proteins are found in prokaryotes and endosymbiotic organelles and group II are found in archaea and eukaryotes.
- HSP 60 consists of 3 domains: Apical, equatorial and intermediate domain.
- Localised in the cytosol, chloroplast and mitochondria (Yadav et al., 2020).
- some roles of this family –
    1. They are involved in assisting the folding of plastid proteins such as RuBisCO (Usman et al., 2014).
    2. Improvement against salt and osmotic stress response is observed in group II chaperonin transformed *E.coli* cells (Wang et al., 2004).

HSP 70 family-

- Ubiquitously found in both plants and animal systems.
- HSP70s are generally made up of two domains N-terminal ATPase domain and the substrate binding domain which is further divided into two sub-domains, i.e., Substrate binding pocket and C-terminal subdomain (Yadav et al., 2020).
- Localised in the cytosol, chloroplast, mitochondria and endoplasmic reticulum.
- Some roles of this family :
    1. Certain members of this family are constitutively expressed and known as HSC70 (heat shock cognate 70) and assist in protein folding (Wang et al., 2004; Tiwari et al., 2020).
    2. HSC70-1 along with HSP90s negatively regulates the activities of other Hsfs under unstressed conditions (Tiwari et al., 2020).
    3. In *Aponogeton madagascariensis,* leaf morphology-based PCD employs the role of HSP70s (Rowarth et al., 2020).

HSP90 family-

- This family of proteins are ubiquitously present except for archaebacteria (Yadav et al., 2020).
- HSP90 is made up of three domains: N-terminal ATP binding domain followed by a charged region attached with Middle domain and MEEVD amino acid residue rich C-terminal domain (Xu et al., 2012).

- Constitutes around 1-2% of cytoplasmic proteins (Usman et al., 2014).
- Localised in the cytoplasm and endoplasmic reticulum.
- These proteins show constitutive expression but are also stress-inducible (Efeoğlu, 2009).
- Some roles of this family -
    1. Act as a co-chaperone with Hsp70 in a variety of protein-mediated functions and signalling (Pratt and Toft, 2003; Al-Whaibi, 2011).
    2. HSP90 has a role in developmental wiring as its loss of function mutant results in a wide array of variations in the phenotype of *Arabidopsis thaliana* (Queitsch et al., 2002).
    3. Helps in maintaining the level of HSFs in unstressed and stressed conditions (Xu et al., 2012; Ohama et al., 2017).

HSP100 family-

- These proteins belong to the superfamily of **ATPases associated with diverse cellular activity** (AAA^+) which is further divided into 2 classes i.e., Class I and Class II and 8 subdivisions (ClpA, B, C, D, M, N, Y and X) are seen in a total of these two categories (Mishra and Grover, 2016). Class I (ClpA, B, C and D) have two nucleotide-binding domains, hence larger in size (68-110 kDa) and Class II (ClpM, N, Y and X) possess only one nucleotide binding domain therefore, smaller in size (40-50 kDa).
- These genes are generally localized in the cytosol, chloroplast and mitochondria (Yadav et al., 2020).
- These proteins show constitutive as well as stress-inducible expression pattern (Efeoğlu, 2009).
- Some roles of HSP 100 family proteins:
    1. Under severe heat stress, ClpB proteins act as a saviour by maintaining proteostasis (Mishra and Grover, 2019).
    2. Conferring thermotolerance is also one of the important function of this HSP101, and there is a replaceable role of these genes in development and germination under non-stress conditions (Hong and Vierling, 2001).
    3. HSP101 also destroys those proteins which are unable to be resurrected via interaction with proteasomal complex (McLoughlin et al., 2019).
    4. Although mutant lines null for the HSP101 shows distinct morphological features throughout the plant body. (Ritesh Kumar, Ph.D. thesis (2019), DPMB, UDSC).

5. APG6, another member of this family has a significant role in chloroplast development. Mutant of this gene shows seedling lethal phenotype due to impaired development of chloroplast (Myouga et al., 2006).

There exist a regulatory network between these HSPs and HSFs. As mentioned earlier HsfA1s act as master regulators, they modulate the expression of various HS responsive genes (DREB2A, MBF1C, HsfA2, HsfA7a and HsfBs). Under unstressed condition, HsfA1 expression is under the control of protein-protein interaction with HSP90 (via affecting nuclear localization) and HSP70 (via affecting transactivation activity)(Tiwari et al., 2020). Upon heat stress, a possible mechanism for relieving this repression by HSP70 and 90 is the formation of aggregated misfolded proteins which out-titrates HsfA1s for binding with HSP70. This results in the release of HsfA1s from this repression (Hayes et al., 2020). Fine-tuning of HsfA1s is maintained by HsfBs which are upregulated by the HsfA1s and forms an autoregulatory negative feedback loop. HsfA1s are regulated at both transcriptional and protein levels as the modulating factor can be an element directly under the control of the primary sensor. There are different kinds of post-translational modifications such as SUMOylation and phosphorylation or dephosphorylation by CBK3 or PP7 (Ohama et al., 2017). HsfA1s results in the expression of HsfA2 which regulates the expression of HSP101, an important chaperone via binding on the heat shock element present in the promoter of this gene.

1.2. AtHSP101 or ClpB-C

As mentioned earlier, various stresses result in denaturation or misfolding of proteins which can have catastrophic effects on the biology of the plants due to the formation of aggregates (because of the presence of exposed hydrophobic moieties). It is well characterised that denatured or misfolded proteins have two fates - either they can be refolded or if the protein's conformation is too far gone, the last resort is to degrade these proteins for salvation purposes. One of the well-known machinery involved in refolding is the bichaperone machinery consisting of hsp70 and hsp40. Even with an active response of this machinery, if denatured proteins are still not able to refold; they will undergo the ultimate recycling process by degradation (Kumar et al., 2020). The constitutive nature of these mechanisms makes them important for the maintenance of proteostasis in the plants and upregulation is also seen in the operation of this machinery under stressful conditions.

One of the other important acting mechanism for removal of toxicity caused by misfolded aggregated proteins known in plants, yeasts & bacteria is carried upon by **ATP-dependent caseinolytic proteases (Clp)**. These proteases were first discovered in *E.coli* by Alfred Goldberg's group and Michael Maurizi's group in 1987 (Kumar et al., 2020). Clp complex was earlier deduced to be made up of 2 components- regulatory ClpA and proteolytic ClpP. Later on, a variant, ClpB is reported in E. coli. This variant differs from ClpA in the aspect of heat

inducibility and absence of interaction with ClpP protein. Further down the timeline, other regulatory variants such as ClpC, ClpE, etc were also identified. (for more information check (Katiyar-Agarwal et al., 2001).

Clps were also discovered in yeasts. Instigatory work on yeast done by Lindquist and her group along with Sanchez showed that *Sc*HSP104 have sequence homology with EcClpB DNA. The involvement of this protein in the alleviation of proteotoxic aggregation and thermotolerance was confirmed by raising the deletion line of yeasts. Its eukaryotic counterpart was confirmed by the cross-reaction of Schsp104 antisera with EcClpB proteins (Kumar et al., 2020). Sanchez et al., 1992 showed basal thermotolerance associated with constitutive expression and induction of Schsp104 by arsenite and ethanol to maintain proteostasis. Pioneering work in the discovery of this protein in plants was done by Elizabeth Vierling's group, Joe Key, Kazuo Shinozaki's group and Prof. Anil Grover's group. Expression profile of HSP101 and its organellar counterparts during seed imbibition and germination is shown in Fig1.4.

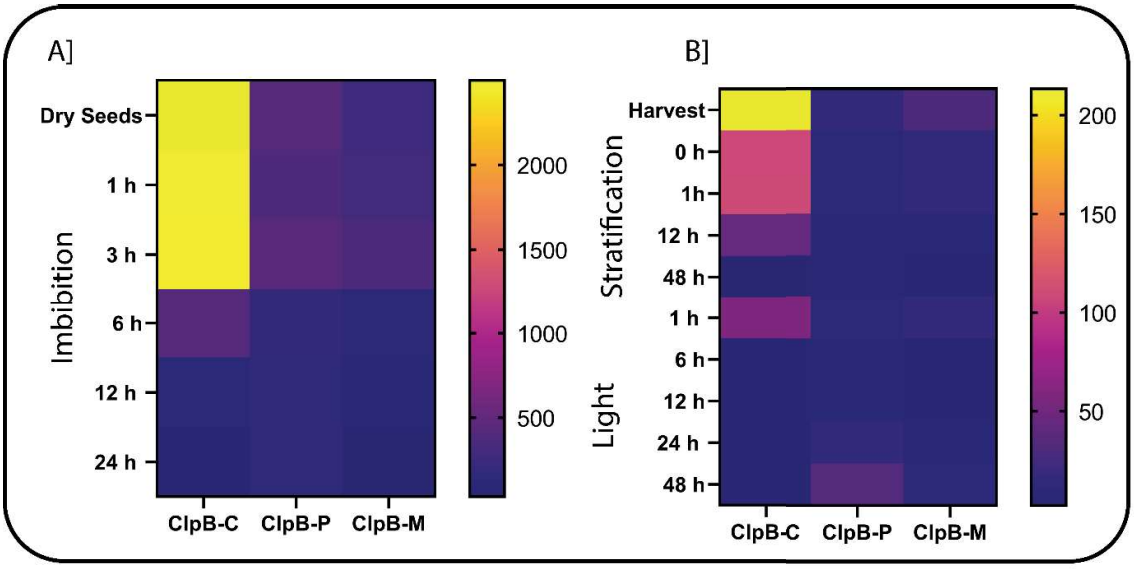


Fig.1.4. Expression profile of ClpB-C, ClpB-P and ClpB-M in imbibed seeds (A) and through stratification and germination (Narsai et al., 2011; Nakabayashi et al., 2005).

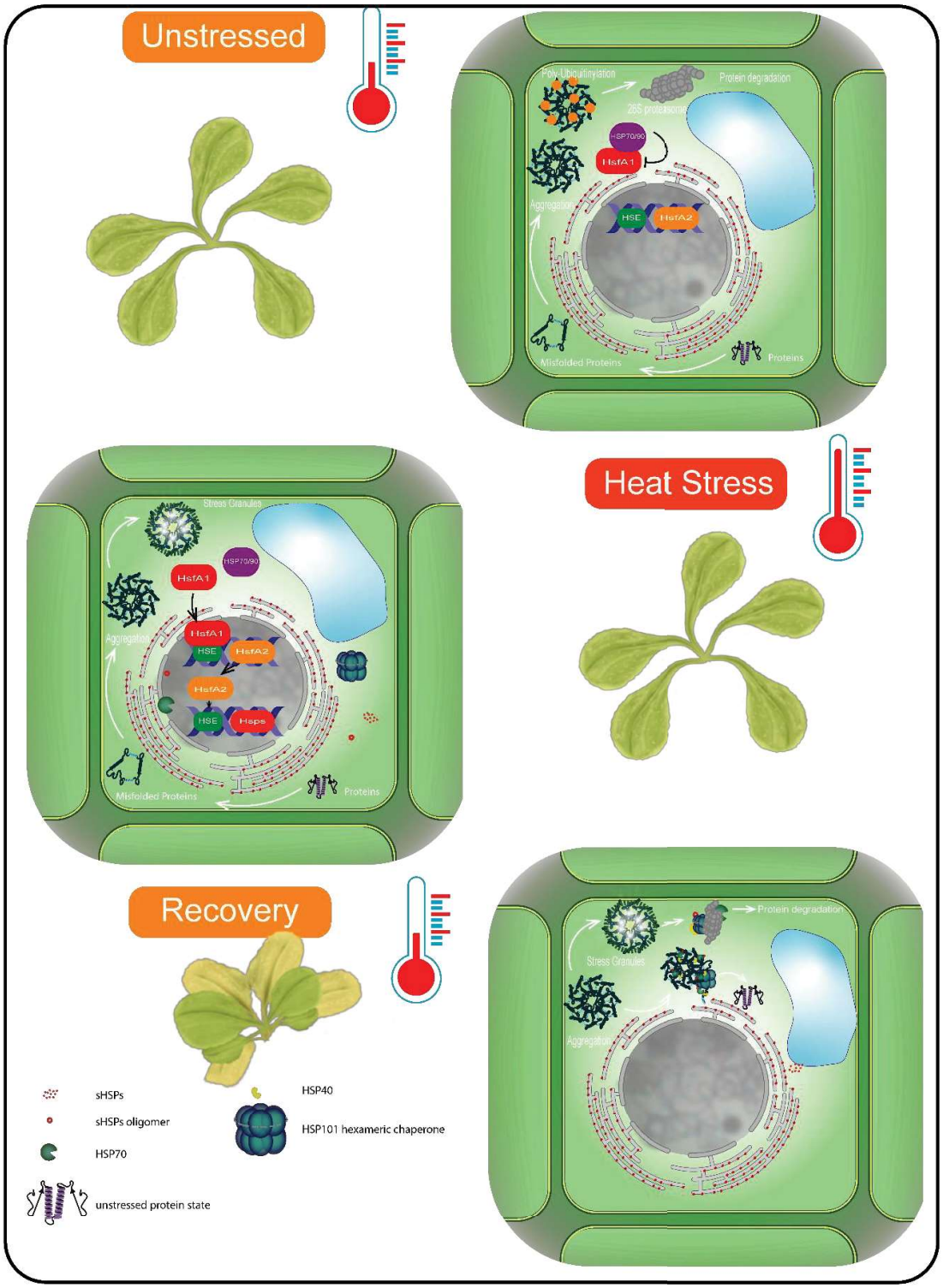


Fig.1.5. Action of HSP101 in unstressed , stresses and recovery phase. Data source : Liu and Charng, 2013; McLoughlin et al., 2019; Ohama et al., 2017 and Ritesh Kumar, Ph.D. thesis 2019, DPMB,UDSC.

1.2.1. Mode of action of HSP101:

Lee et al., 2003 proposed two functions of this protein:

1. Disaggregation of high molecular weight (HMW) aggregates into smaller ones.
2. Resolubilization of smaller aggregates with co-chaperone systems.

This protein exists in the form of a hexamer. It recognizes large aggregates and bind to them using motif 1 and motif 2. ATP driven conformational change provides the necessary force to pull apart the larger aggregates via the swivelling motion. Following disaggregation, the smaller aggregates are recognised by the HSP 70-40 bichaperone machinery along with small HSPs or they can enter through the central cavity of the disaggregase and can be refolded either via "Capture and Release" or "Translocation." Then refolding is mediated by HSP 70-40 bichaperone machinery.

1.2.2. *hot* mutant

Point mutation in HSP101 resulted in the development of hot1 mutant which is thermosensitive in nature. This mutation replaced the conserved glutamate residue to lysine in the second ATP-binding domain. The first *hot1* mutation was discovered in plants which showed leaky expression (Hong and Vierling, 2000). Later on, the *hot1-3* mutant was obtained from a separate T-DNA mutagenized population. It has an insertion in the second exon and shows total knockout of HSP101 (Hong and Vierling, 2001).

1.3. Heavy Metal(loid)s Stress

The term "Heavy Metals" is given to the high density (>5.0 x 10^3 Kg/m^3) members, the d-block family (not all) of the modern periodic table (Singh et al., 2015). But there is growing belief to substitute the term "Heavy Metals" with "Potentially Toxic Elements" while reporting the scientific discoveries and research as this term is used for toxic elements or some elements that do not fall in the category of metal (As) and are neither heavy nor dense (Ba) (Duffus, 2002; Pourret and Bollinger, 2018; Buzea and Pacheco, 2020).

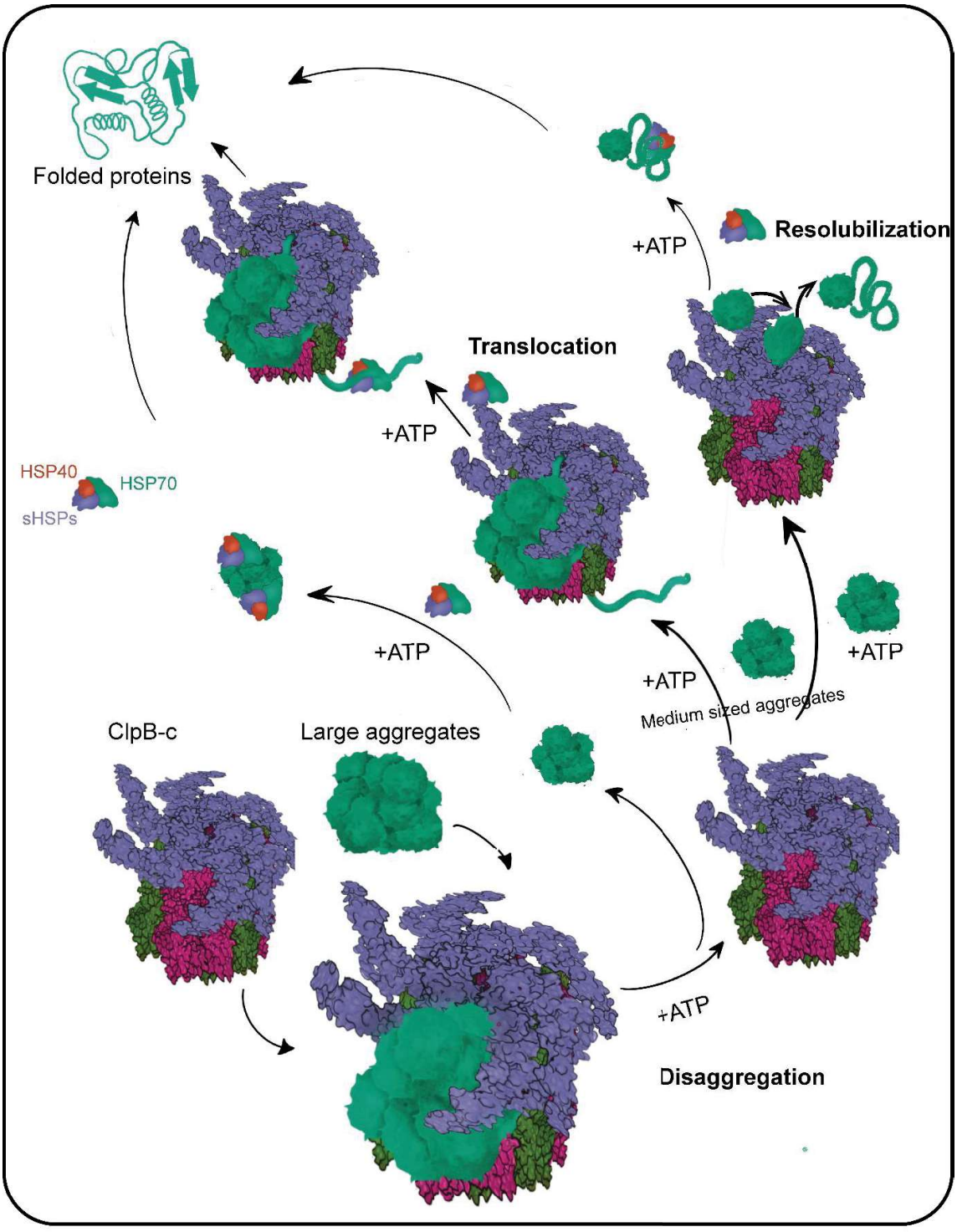


Fig.1.6. Mode of action of ClpB . Remodeled from Lee et al., 2003.

Ruinous effects of heavy metals on both plants and animals are well reported (Tchounwou et al., 2012; Keyster et al., 2020; Garg and Singla, 2011). These heavy metals occur naturally in the Earth's crust (conjugated with minerals such as arsenopyrite). But ever-increasing industrialization (semiconductor industry), migration rate towards urban areas (urbanization) and need for agricultural additives such as insecticides, pesticides and fertilizers results in the rapid emanation of heavy metals in the environment and made them ubiquitous. Due to their prolonged half-life, these heavy metals endure in the environment and accumulates. Hence, they cause toxicity in nature. Major problems are an accumulation of heavy metals in food webs and leaching into the groundwater aquifers (Hasan et al., 2017; Garg and Singla, 2011; Clemens and Ma, 2016). Consumption of plant-based food exposed to heavy metals either by irrigation practices (Meharg, 2004) or direct uptake from soil can potentially become a threat towards human exposure.

This thesis majorly focuses upon Cadmium (Metal) and arsenic (Metalloid), hereafter referred to as heavy metals. Cadmium exposure (Non-occupational) to non-smoking population (smoking population have a different value due to cadmium accumulated in *Nicotiana spp.* leaves) is attributable to more than 90% to a vegetarian diet (such as cereals and vegetables). Cadmium environmental toxicity catches the limelight when the environmental exposure of cadmium results in Itai-Itai syndrome in Japan because of consumption of contaminated rice grains due to unconstrained dumping of cadmium based waste in the Jenzu river basin (irrigation source) (Clemens et al, 2013). Arsenic poisoning is prevalent in South and Southeast Asia. In a country like Bangladesh and West Bengal and Punjab (India), where groundwater is contaminated which is attributable for the major exposure.

1.3.1. Occurrence and Sources of Heavy metals

As mentioned earlier, heavy metals occur naturally in Earth's crust (Arsenic > Cadmium). Major sources of cadmium emanation are due to human activities such as mining, smelting or agricultural additives. Inorganic cadmium exists in the oxidation of '+2', more prevalent in aerobic soils. Whereas arsenic occurs in a variety of different forms depending on the pH of soils. In its inorganic form arsenic occurs in 2 oxidative states,i.e., Arsenate ('+5') and Arsenite ('+3') and it also occurs in methylated arsenic-based acids [Monomethylarsenic acid (MMA) and dimethylarsinic acid (DMA)] by the action of rhizosphere bacteria (Clemens and Ma, 2016; Garg and Singla, 2011).

1.3.2 Effects of heavy metals on plants

Mechanisms by which heavy metals induce toxicity are as follows (Singh et al., 2015):

- Heavy metals are non-essential cations, they enter the plasma membrane as biological hitchhikers via transporters for essential ions hence results in the competition for

absorption with essential ions (this relation is observed between As and P; Cd and Zn)(Clemens and Ma, 2016; Mendoza-Cózatl et al., 2011).

- Heavy metals act by disturbing the working of proteins by binding or interacting with free sulfhydryl/thiol or any other functional group.
- Some heavy metals act by collapsing the function of folded proteins via displacing the vital cation in metalloproteins (Tamás et al., 2014; Jalmi et al., 2018).
- Heavy metal imposes oxidative stress upon accumulation via the production of reactive oxygen species (ROS). Heavy metals can employ two ways to act via redox pathway: direct or indirect. Redox-active metal such as Fe and Cu can generate ROS directly whereas other heavy metals employ indirect ways such as activation of NADPH oxidases or inhibiting the action of enzymes via the above-mentioned strategies (Jalmi et al., 2018).

1.3.3 Tolerance strategies of plants

Exposure to heavy metal leads to a plethora of responses at various levels of organizations. Strategies employed by plants to tackle heavy metal stress are as follows:

- Chelation of heavy metals (Hasan et al., 2017; Joshi et al., 2016).
- Vacuolar sequestration (Yang and Chu, 2011).
- Regulation of Metal intake (Caetano et al., 2015).
- A surge in antioxidative mechanisms(Marques et al., 2019)
- Accumulation of osmolytes (Ashfaque et al., 2020; Garg and Singla, 2011).
- Mycorrhization (Garg and Singla, 2011).
- Expression of heat shock proteins (Hasan et al., 2017).

1.4 Work carried out in this dissertation aimed at-

1. In-silico analysis of HSFs and HSPs (HSP40, HSP60, HSP70, HSP90, HSP100 and small HSPs) under heat stress in *Arabidopsis thaliana*.
2. Analysis of HSP101 in seed germination under cadmium and arsenic stress in *Arabidopsis thaliana*.

Material and methods



2.1 Experimental Materials

Experimental materials and chemicals used in the present work were obtained from diverse sources as shown in Table 2.1.

Table 2.1. Sources of various experimental materials and chemicals used.

Material	Source
Arabidopsis seeds (Col-0, *hot1-3* mutant)	Professor Anil Grover's laboratory stocks, DPMB, UDSC
Constructs, development and molecular characterization of OE line	Developed by Richa Babbar, Ph.D. thesis (2019), Prof. Anil Grover's laboratory, DPMB, UDSC
Other fine chemicals	Diverse sources including Qiagen (Germany), Sigma-Aldrich (Germany), Roche (US), Merck (Germany), Invitrogen (US) and Fermentas (US).

2.2 *Arabidopsis* growth conditions

2.2.1 Seed sterilization and plating

Arabidopsis seeds were first surface sterilized with 70% ethanol for 30s followed by treatment with 2% (v/v) sodium hypochlorite and 1 drop of Tween-20 for 5min. The seeds were then washed with autoclaved distilled water 3-4 times until no frothing was observed and finally suspended in 0.1% (w/v) agar. The seeds were plated onto MS medium (½MS, 1% sucrose and 0.8% agar, pH 5.8-5.85). The plates were placed in the cold room (4°C) for 36h for stratification and then transferred to the culture room maintained at 22±1°C with 16h light/8h dark conditions and light intensity of 100-125 $\mu molm^{-2}\ s^{-1}$.

2.2.2 Maintenance of mature plants

Arabidopsis thaliana ecotype Col-0 and *hot1-3* were placed in pots containing Soilrite MixTM (Kelperlite, Banglore, India) saturated with OS medium (Somerville and Ogren, 1982). The pots were kept in the cold room (4°C) for 36h for stratification and then transferred to culture room maintained at 22±1°C with 16h light/ 8h dark conditions with a light intensity of 100-125 $\mu molm^{-2}s^{-1}$ for raising the full-grown plants.

2.3 High temperature stress

2.3.1 Phenotypic assays

Five-day old intact seedlings obtained were stressed on 0.5X MS medium plates for high temperature regimes as showed in results, in a water bath after sealing the plates using petri-seal. The plates were then transferred to the growth room for recovery under control conditions.

2.4 Metal stress assay

2.4.1 Phenotypic assay and fresh weight analysis

For fresh weight analysis, seeds were grown on 0.5X MS medium containing different concentration of heavy metals (20µM As and 50µM Cd) and then the fresh weight was recorded 10d post-stress. For the phenotypic assay, seeds were grown on 0.5X MS medium containing different concentration of As (20µM, 50µM, 100µM) and Cd (50 µM, 100µM, 200µM) for 10d. Microscopic analysis was done using Leica S8APO stereo zoom systems available at DPMB, UDSC.

2.4.2 Root length analysis

Seeds were plated on 0.5X MS medium containing different concentration of heavy metals As (20µM, 50µM) and Cd (50µM, 100µM, 200µM). Plates were kept vertically for 10d to analyze the root length.

2.4.3 RNA-seq data mining of HSPs and HSFs in *Arabidopsis thaliana* under heat stress

RNA-seq data was retrieved from the supplementary material from the selected publications (Table 3.1.1.) and data generated at the laboratory of Prof. Anil Grover. Mined data were then classified into 4 categories (mentioned later). A list of HSPs and HSFs genes was provided at the laboratory. Differentially expressing genes were filtered at the criteria, $-1 \geq \text{Log}_2(\text{Fold Change}) \geq 1$ (Foldchange is greater than or equal to 2 and less than or equal to -2), Adjusted *p*-value ≤ 0.05. Filtered genes were sorted in the form of a table using Jvenn online tool (Bardou et al., 2014) and these results were graphed using the Tableau (https://www.tableau.com/) desktop version.

2.4.4 Expression profiling of *At*HSP101 under non-stress conditions.

Expression data (at default settings) was retrieved and visualized from *Arabidopsis* efp browser (Winter et al., 2007) and ePlant browser (Waese et al., 2017) available at Bio-analytic resource for plant biology website (http://bar.utoronto.ca/).

Results and Discussion



Objective 1 In-silico analysis of HSFs and HSPs (HSP40, HSP60, HSP70, HSP90, HSP100 and small HSPs) under heat stress in *Arabidopsis thaliana*.

Owing to the technical advancement, ease in availability of technologies and a compulsion imposed by journals to deposit expression data, there is a surge in the availability of quality gene expression data (Wang et al., 2019).

This section presents a comparative transcriptome analysis of various HSFs and HSPs genes in different tissues of *Arabidopsis thaliana* ecotype Col-0. Transcriptome provides information via space and time affected by a variety of factors such as age, developmental stage, environmental cues etc. Such comparative studies provide insights into altered transcription and its implications in a developmental stage, tissue, cell, genotype-specific or as in this case effect of stress (Wan and Li, 2019).

Processed RNA-seq data was retrieved from supplementary material from the mentioned studies (Table.3.1.1). Retrieved data were categorised into 4 categories, namely category 1, category 2, category 3 and category 4. The stress regime for each category is mentioned in Fig. 3.1.1 and Fig. 3.1.2.

Table.3.1.1 List of sources of processed RNA-seq dataset			
S.No.	**Category**	**Category Name**	**Studies**
1	**1**	Embryonic and seed tissues	**Malabarba, J., Windels, D., Xu, W., and Verdier, J.** (2021). Regulation of DNA (de)Methylation Positively Impacts Seed Germination during Seed Development under Heat Stress. Genes (Basel) **12.**
2	**2**	Seedlings subjected to heat stress at different time of the day	In-Lab data generated at the laboratory of **Prof. Anil Grover**
3			**Blair, E.J., Bonnot, T., Hummel, M., Hay, E., Marzolino, J.M., Quijada, I.A., and Nagel, D.H.** (2019). Contribution of time of day and the circadian clock to the heat stress-responsive transcriptome in Arabidopsis. Sci Rep **9**: 4814.
4			**Grinevich, D.O., Desai, J.S., Stroup, K.P., Duan, J., Slabaugh, E., and Doherty, C.J.** (2019). Novel transcriptional responses to heat revealed by turning up the heat at night. Plant Mol Biol **101**: 1**–**19
5	**3**	Reproductive and vegetative tissues	**Zhang, S.-S., Yang, H., Ding, L., Song, Z.-T., Ma, H., Chang, F., and Liu, J.-X.** (2017). Tissue-Specific Transcriptomics Reveals an Important Role of the Unfolded Protein Response in Maintaining Fertility upon Heat Stress in Arabidopsis. Plant Cell **29**: 1007**–**1023.
6			**Rahmati Ishka, M., Brown, E., Weigand, C., Tillett, R.L., Schlauch, K.A., Miller, G., and Harper, J.F.** (2018). A comparison of heat-stress transcriptome changes between wild-type Arabidopsis pollen and a heat-sensitive mutant harboring a knockout of cyclic nucleotide-gated cation channel 16 (cngc16). BMC Genomics **19**: 549
7	**4**	Seedlings subjected to different durations of heat stress.	**Li, B., Gao, Z., Liu, X., Sun, D., and Tang, W.** (2019). Transcriptional Profiling Reveals a Time-of-Day-Specific Role of REVEILLE 4/8 in Regulating the First Wave of Heat Shock-Induced Gene Expression in Arabidopsis. Plant Cell **31**: 2353**–**2369.

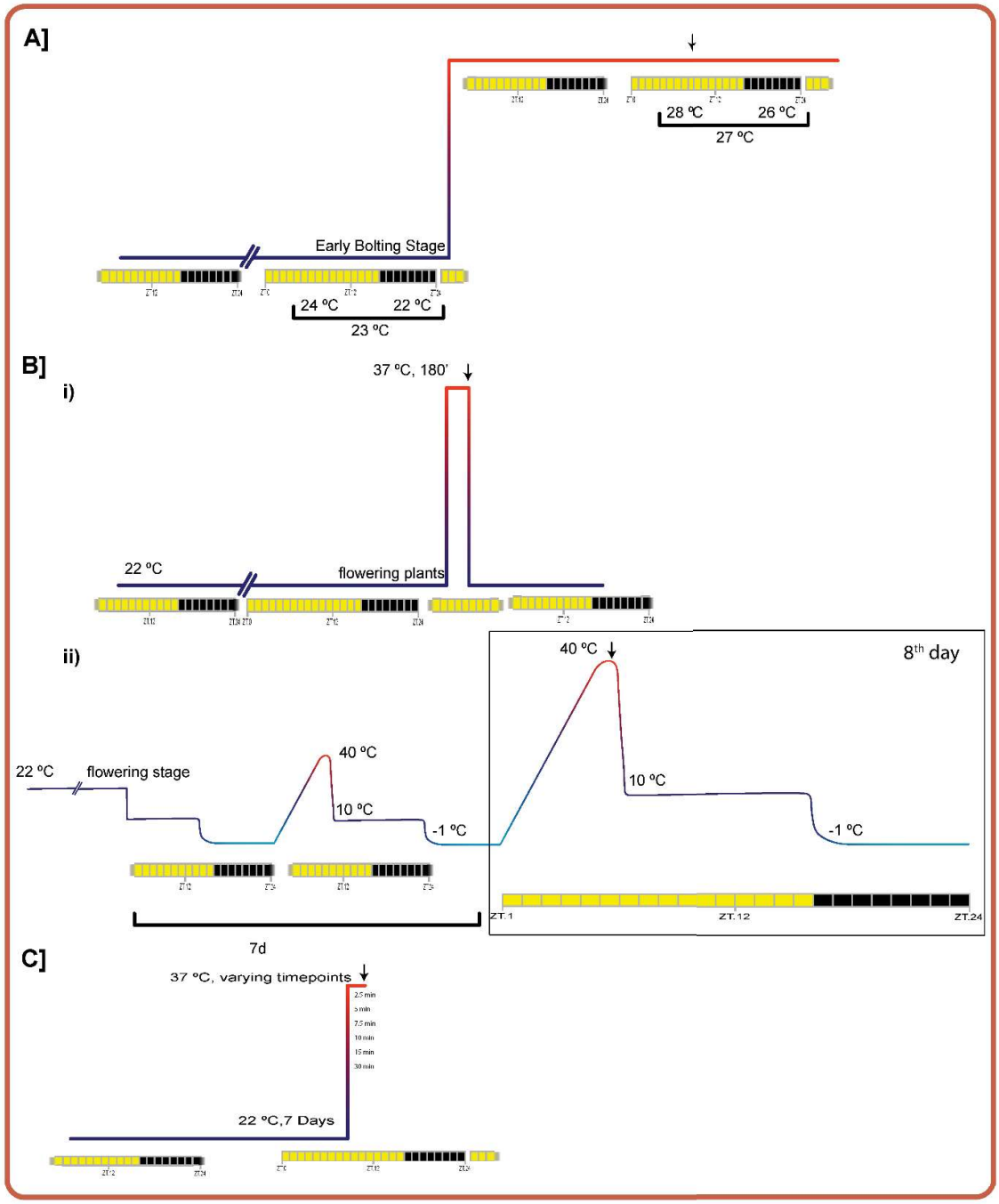


Fig.3.1.1. Followed stress regime for category 1, A](Malabarba et al., 2021.); category 3, B] i) for rosette leaves, early and late flowers (Zhang et al., 2017) , B] ii) for pollen grains(Rahmati Ishka et al., 2018) and category 4, C] ((Li et al., 2019) arrow heads depicts sample harvesting.

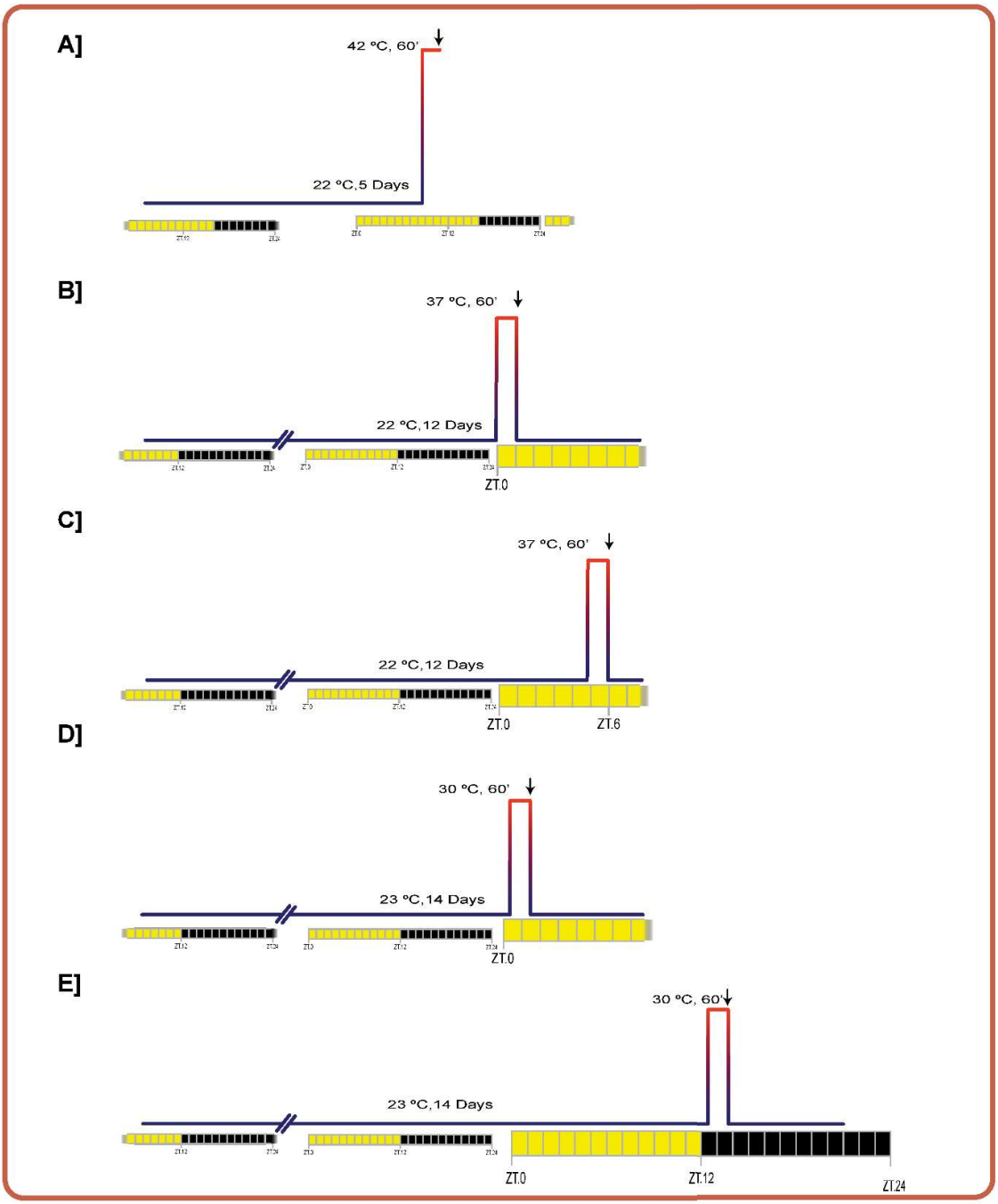


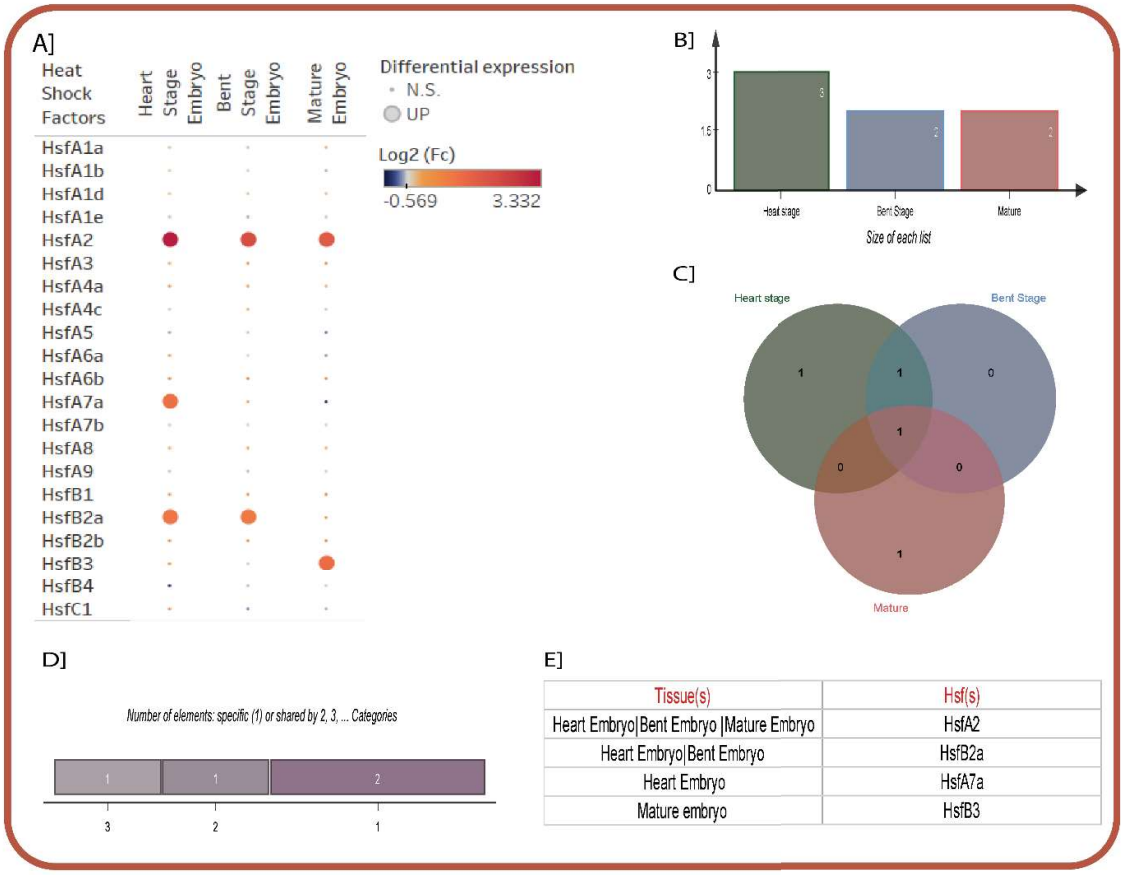
Fig.3.1.2. Followed stress regime for category 2, A] Seedling type A (in lab data generated at the laboratory of Prof. Anil Grover,DPMB,UDSC); B] Seedling type B (Blair et al., 2019); C] Seedling type C (Blair et al., 2019); D] Seedling type D (Grinevich et al., 2019) and E] Seedling type E (Grinevich et al., 2019) . Arrow heads depicts sample harvesting.

3.1.1. Expression profile of HSFs in seed and embryonic tissue (category 1).

Data for this study (3 embryonic/seed stages- heart embryo stage seeds, bent embryo stage seeds and mature seeds) was retrieved from the additional data provided by Malabarba et al., 2021.

Upregulated and downregulated genes were represented as red-orange large circles and blue small circles respectively in the dot matrix plot (Fig. 3.1.3A). A total of four HSF genes (HsfA2, HsfB2a, HsfA7a and HsfB3) were found to be differentially upregulated in at least one of these three seed developmental stages. The number of specific and shared HSF genes between the three seed developmental stages were represented in the form of a Venn diagram (Fig. 3.1.3C) and their proportion was represented in the form of a horizontal bar (Fig. 3.1.3D). The sum of numbers mentioned on the ellipse or numbers on the horizontal bar conveys the total no. of transcripts in various combinations and permutations. Expression of these differentially expressed genes were summarised in the form of a table (Fig. 3.1.3E).

Upon heat stress, HsfA2 was upregulated in all the three embryonic stages. HsfB2a showed upregulation in heart and bent stage embryo but not in mature seeds. HsfA7a and HsfB3 were specifically upregulated in heart and bent stage embryo respectively. Significantly upregulated genes in each embryo category are represented in the form of a bar graph (Fig. 3.1.3B). The highest $\log_2$ (fold change) for HsfA2 was observed in the heart stage embryo, followed by its reduction in the consecutive stages. No significant downregulation of any HSF genes was observed in this category.



Tissue(s)	Hsf(s)
Heart Embryo|Bent Embryo |Mature Embryo	HsfA2
Heart Embryo|Bent Embryo	HsfB2a
Heart Embryo	HsfA7a
Mature embryo	HsfB3

Fig.3.1.3 Expression profile of Heat shock Factors in Category 1. A] Dot-matrix plot representing expression of heat shock factors in 3 stages of Seed Development in *Arabidopsis thaliana*. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock factor genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator.

* No significant downregulation of heat shock factor genes was observed.

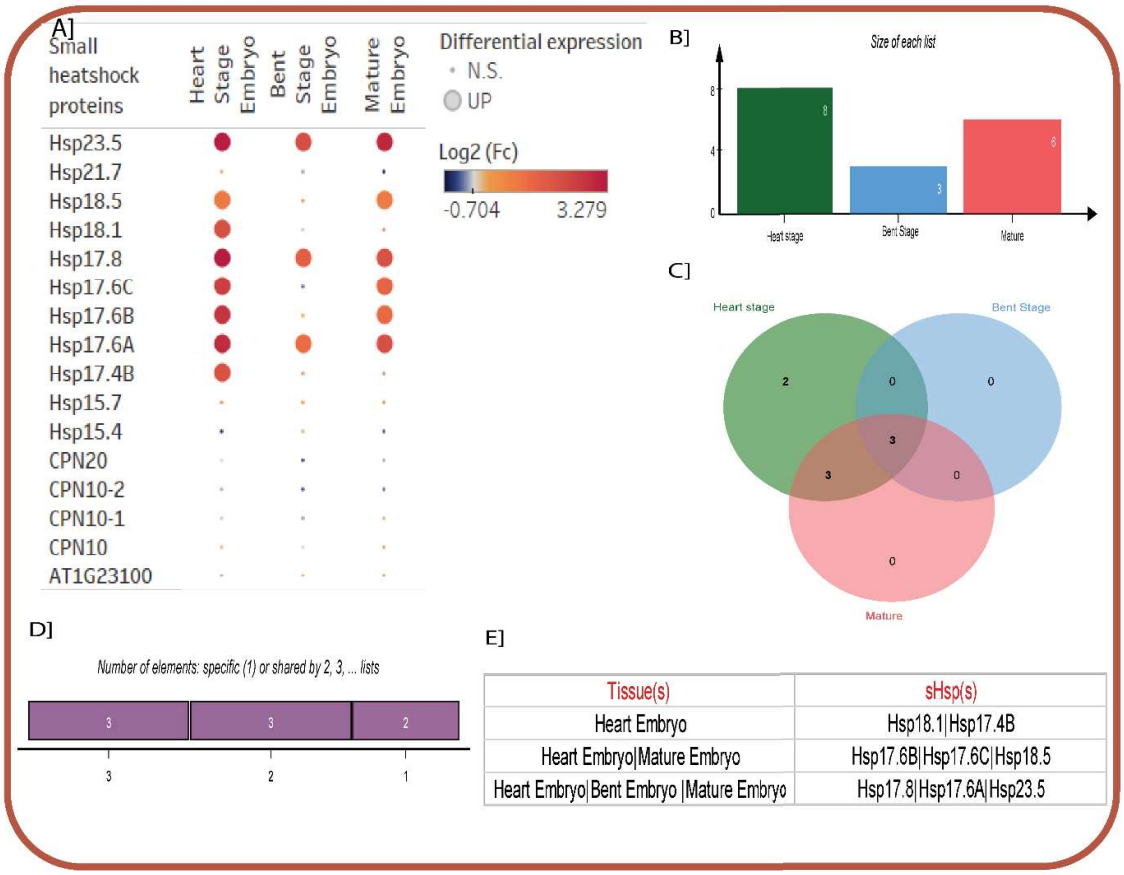
** Seed tissue was also taken with the embryonic tissue

3.1.2. Expression profile of HSPs in seed and embryonic tissue (category 1).

Data for this study was mined from the supplementary data provided by Malabarba et al., 2021 (heart embryo stage seeds, bent embryo stage seeds and mature seeds).

Upregulated and downregulated genes were represented as red-orange large circles and blue small circles respectively in the dot matrix plot- small HSPs (Fig. 3.1.4A) , HSP40s (Fig. 3.1.5A), HSP60s (Fig. 3.1.6A), HSP70s (Fig. 3.1.7A), HSP90s (Fig. 3.1.8 A) and HSP100s (Fig. 3.1.9A). Significant upregulation of HSPs in embryo were represented in the form of a bar graph [sHSPs (Fig. 3.1.4B), HSP40s (Fig. 3.1.5B), HSP70s (Fig. 3.1.7 B)].Significant downregulation of HSPs was observed in the case of HSP70s where bent embryo stage specific downregulation of MED37B was observed. A total of six HSP40 genes (AT1G02650, AT2G47440, AT4G19570, ERDJ3A, AT362570 & AT2G20560), one HSP90 gene (HSP90-1), one HSP100 gene (CLPB1), two HSP70 genes (HSP70-8 & MED37C) and eight of sHSP genes (Hsp18.1, Hsp17.4B, Hsp17.6B, Hsp17.6C, Hsp18.5, Hsp17.8, Hsp17.6A and Hsp23.5) were found differentially upregulated in at least one of the three seed developmental stages. Number of specific (Unique) and shared genes between the three embryo developmental stages were represented in the form of a Venn diagram [sHSPs (Fig. 3.1.4C), HSP40s (Fig. 3.1.5C), HSP70s (Fig. 3.1.7C)] and their proportion were represented in the form of horizontal bar [sHSPs (Fig. 3.1.4D), HSP40s (Fig. 3.1.5D), HSP70s (Fig. 3.1.7D)]. Sum of numbers mentioned on the ellipse or numbers on the horizontal bar conveys the total no. of transcripts in various combinations and permutations. Expression of differentially upregulated genes were summarised in form of a table [sHSPs (Fig. 3.1.4E), HSP40s (Fig. 3.1.5E), HSP70s (Fig. 3.1.7E), HSP90s (Fig. 3.1.8B) and HSP100s (Fig. 3.1.9B)].

Upon heat stress, various upregulated HSPs were observed in all the three embryo developmental stages such as HSP40s (AT2G20560), HSP70s (MED37C), HSP90s (Hsp90-1) and sHSPs (Hsp17.8, Hsp17.6A, Hsp23.5). Stage-specific upregulation of HSPs were also observed. Heart stage-specific upregulated sHSPs were Hsp18.1 and Hsp17.4B and HSP40 were AT1G02650. Upregulation of HSP40s such as AT2G47440 and AT4G19570 were observed in the bent embryo in a stage-specific fashion. Mature seed stage-specific upregulation was observed in HSP40s (ERDJ3A) and HSP70s (Hsp70-8). HSP100s such as CLPB1 (*At*HSP101) was upregulated in all cases of this category except for the mature seed stage. No significant differential expression was observed in any of the HSP60 genes in any of the aforementioned seed developmental stages.

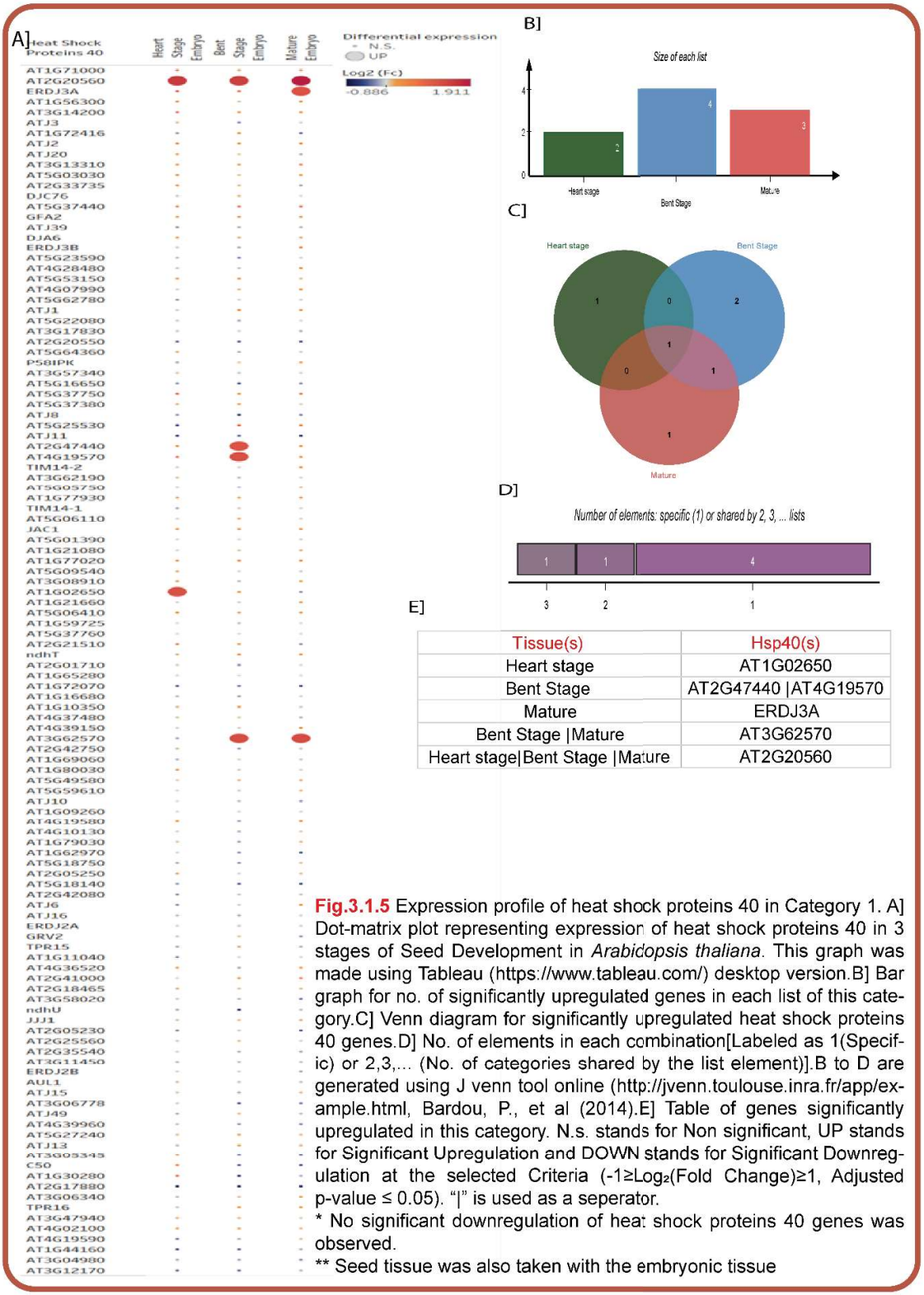


Tissue(s)	sHsp(s)
Heart Embryo	Hsp18.1|Hsp17.4B
Heart Embryo|Mature Embryo	Hsp17.6B|Hsp17.6C|Hsp18.5
Heart Embryo|Bent Embryo |Mature Embryo	Hsp17.8|Hsp17.6A|Hsp23.5

Fig.3.1.4 Expression profile of sHSPs in Category 1. A] Dot-matrix plot representing expression of small heat shock proteins in 3 stages of Seed Development in *Arabidopsis thaliana*. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated small heat shock protein genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator.

* No significant downregulation of small heat shock protein genes was observed.

** Seed tissue was also taken with the embryonic tissue



Tissue(s)	Hsp40(s)
Heart stage	AT1G02650
Bent Stage	AT2G47440 |AT4G19570
Mature	ERDJ3A
Bent Stage |Mature	AT3G62570
Heart stage|Bent Stage |Mature	AT2G20560

Fig.3.1.5 Expression profile of heat shock proteins 40 in Category 1. A] Dot-matrix plot representing expression of heat shock proteins 40 in 3 stages of Seed Development in *Arabidopsis thaliana*. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock proteins 40 genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator.

* No significant downregulation of heat shock proteins 40 genes was observed.

** Seed tissue was also taken with the embryonic tissue

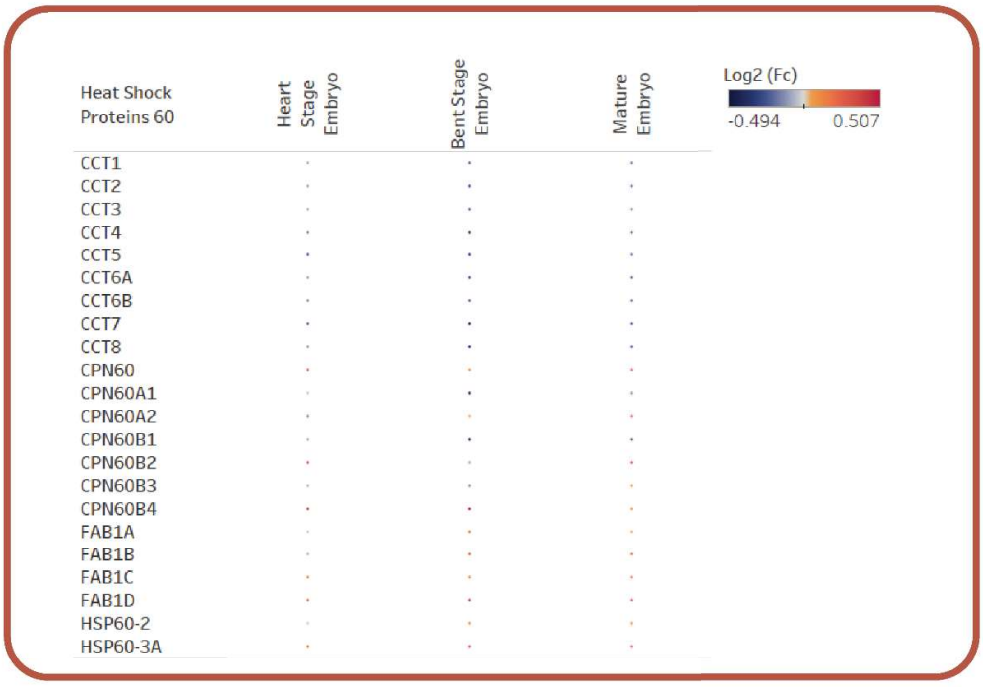
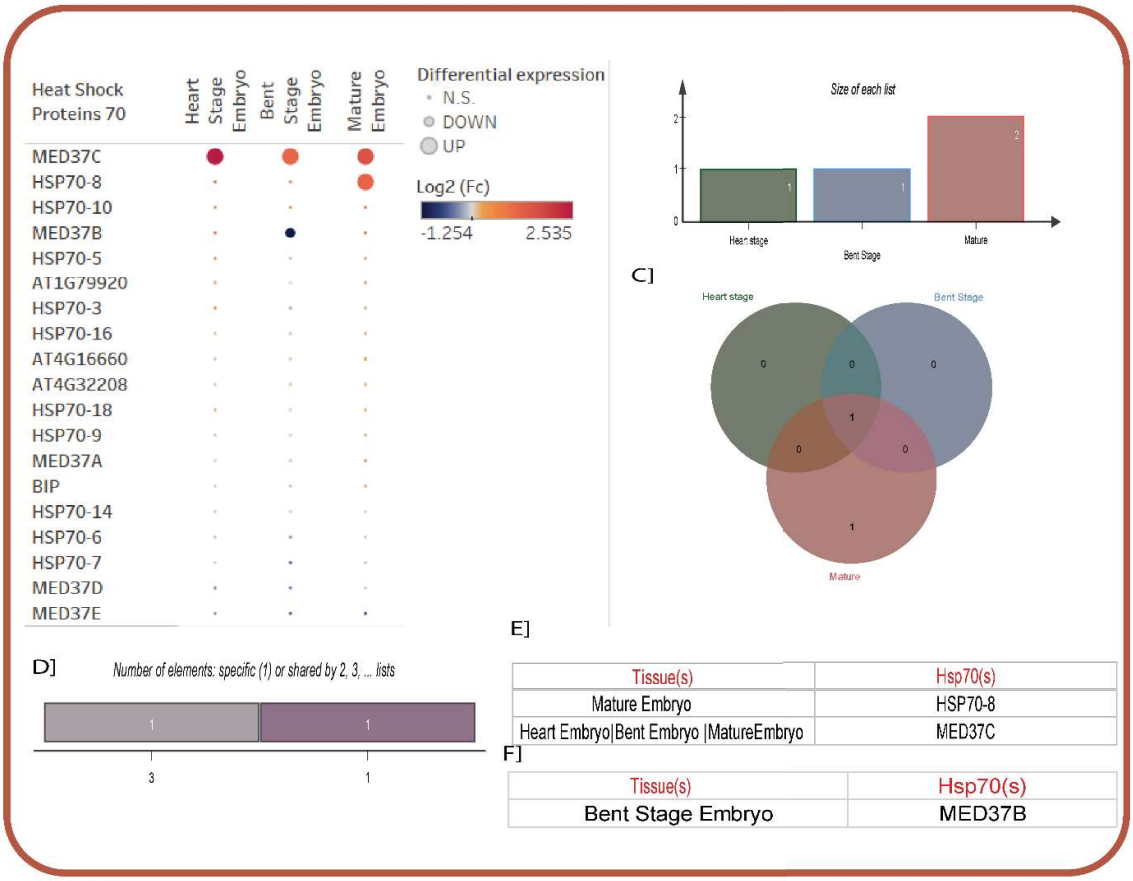


Fig.3.1.6 Expression profile of heat shock proteins 60. Dot-matrix plot representing expression of heat shock proteins 60 in 3 stages of Seed Development in *Arabidopsis thaliana* (Category 1).N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). No significant downregulation or upregulation of heat shock proteins 60 genes was observed.



Tissue(s)	Hsp70(s)
Mature Embryo	HSP70-8
Heart Embryo|Bent Embryo |MatureEmbryo	MED37C

Tissue(s)	Hsp70(s)
Bent Stage Embryo	MED37B

Fig.3.1.7 Expression profile of heat shock proteins 70 in Category 1. A] Dot-matrix plot representing expression of sheat shock proteins 70 in 3 stages of Seed Development in *Arabidopsis thaliana*. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock proteins 70 genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly down-regulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05)."|" is used as a seperator.

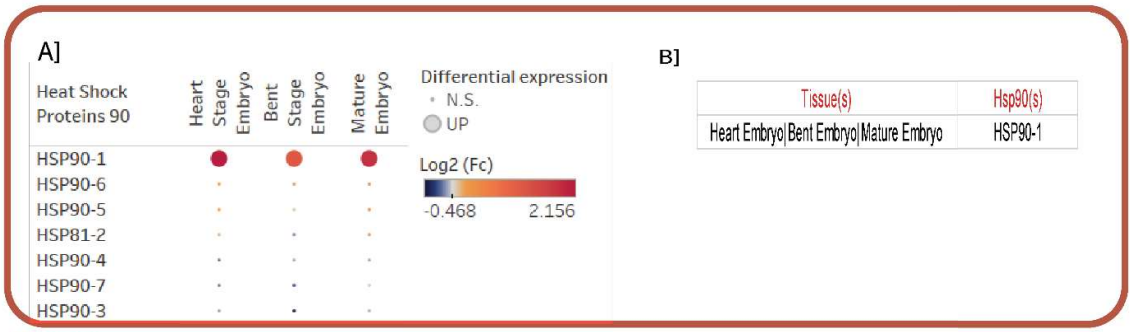
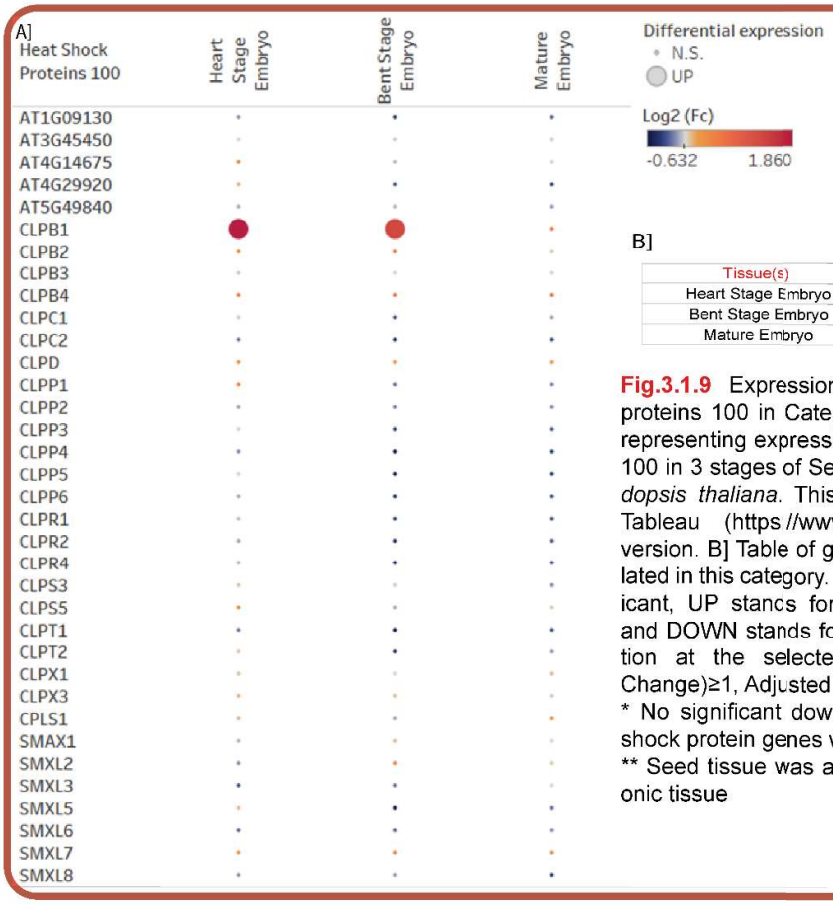


Fig.3.1.8 Expression profile of heat shock proteins 90 in Category 1. A] Dot-matrix plot representing expression of heat shock proteins 90 in 3 stages of Seed Development in *Arabidopsis thaliana*. This graph was made using Tableau (https://www.tableau.com/) desktop version. B] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Down-regulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05).
* No significant downregulation of small heat shock protein genes was observed.
** Seed tissue was also taken with the embryonic tissue

A]



B]

Tissue(s)	Hsp100(s)
Heart Stage Embryo	CLPB1
Bent Stage Embryo	CLPB1
Mature Embryo	No Significant Upregulation

Fig.3.1.9 Expression profile of heat shock proteins 100 in Category 1. A] Dot-matrix plot representing expression of heat shock proteins 100 in 3 stages of Seed Development in *Arabidopsis thaliana*. This graph was made using Tableau (https //www.tableau.com/) desktop version. B] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stancs for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05).

* No significant downregulation of small heat shock protein genes was observed.

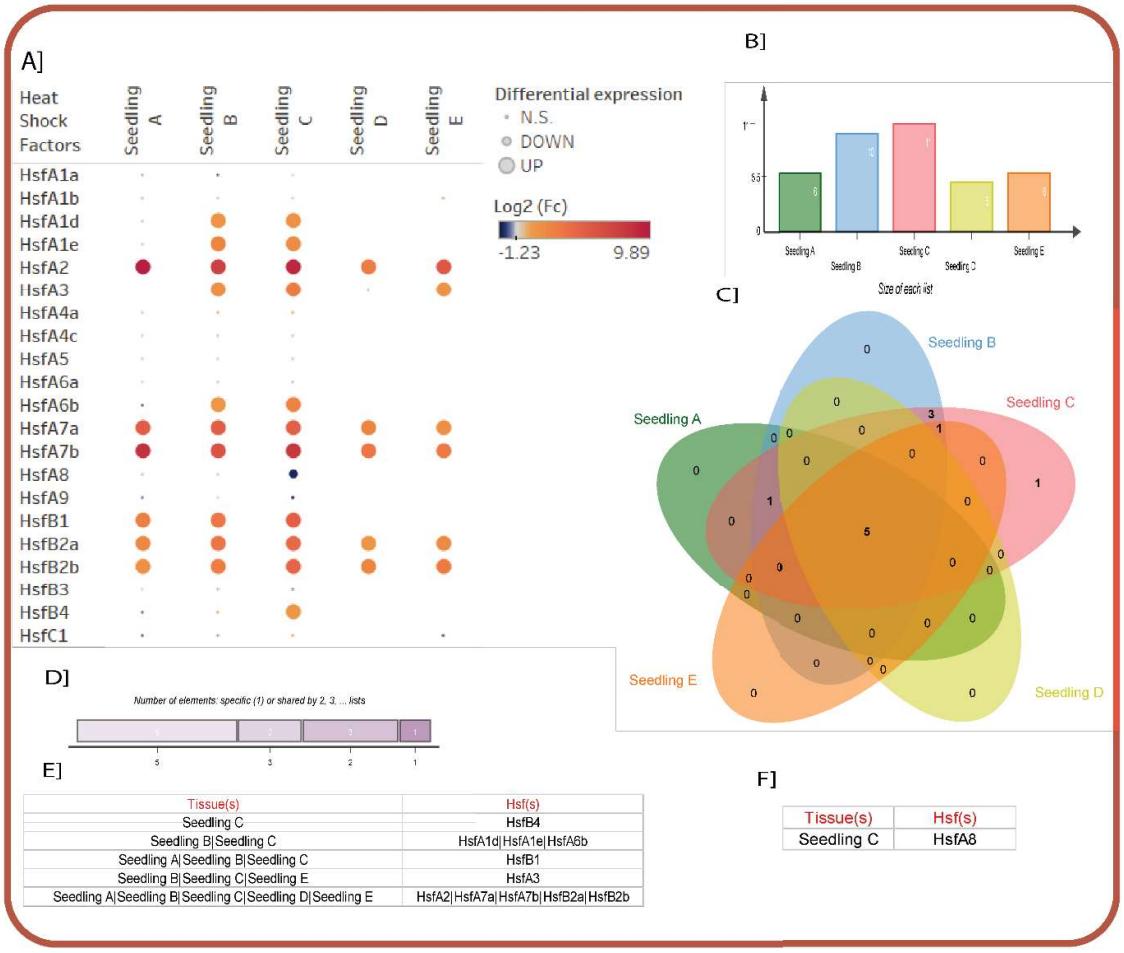
** Seed tissue was also taken with the embryonic tissue

3.1.3 Expression profile of HSFs in seedlings under different stress regime (Category 2).

This category examines the temporal (time of day) effects of heat stress on seedlings and their comparison with in-lab RNA-seq data generated at the laboratory of Prof. Anil Grover, Department of Plant Molecular Biology, UDSC. The publicly available RNA-seq dataset (Blair et al., 2019 and Grinevich et al., 2019) was categorized into five seedling types (A, B, C, D & E) . These seedlings have been given stress at a different time of the day (Only In-lab data has no specified zeitgeber time (Zt) for heat temperatures).

Upregulated and downregulated HSFs were represented as red-orange large circles and blue small circles respectively in the dot matrix plot (Fig. 3.1.10A). A total of 11 HSF genes were found to be differentially upregulated in at least one of these five seedling types. The number of unique and shared HSF genes between five seedlings type were represented in the form of a Venn diagram (Fig. 3.1.10C) and their proportion of genes in either specific or shared categories in each category is represented in the form of a horizontal bar (Fig. 3.1.10D). The sum of numbers mentioned on the ellipse or numbers on the horizontal bar conveys the total no. of transcripts in various combinations and permutations. Expression of these differentially expressed genes are summarised in form of a table (Fig. 3.1.10E and F).

Upon heat stress, HsfA2, HsfA7a, HsfA7b, HsfB2a and HsfB2b were upregulated in all five seedling types. Type C seedling showed upregulation of HsfB4 and downregulation of HsfA8 under above mentioned conditions. Significant upregulation of genes in each seedling category were represented in the form of a bar graph (Fig. 3.1.10B).



Tissue(s)	Hsf(s)
Seedling C	HsfB4
Seedling B|Seedling C	HsfA1d|HsfA1e|HsfA6b
Seedling A|Seedling B|Seedling C	HsfB1
Seedling B|Seedling C|Seedling E	HsfA3
Seedling A|Seedling B|Seedling C|Seedling D|Seedling E	HsfA2|HsfA7a|HsfA7b|HsfB2a|HsfB2b

Tissue(s)	Hsf(s)
Seedling C	HsfA8

Fig.3.1.10 Expression profile of Heat shock Factors in Category 2. A] Dot-matrix plot representing expression of heat shock factors in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock factor genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly down-regulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator.

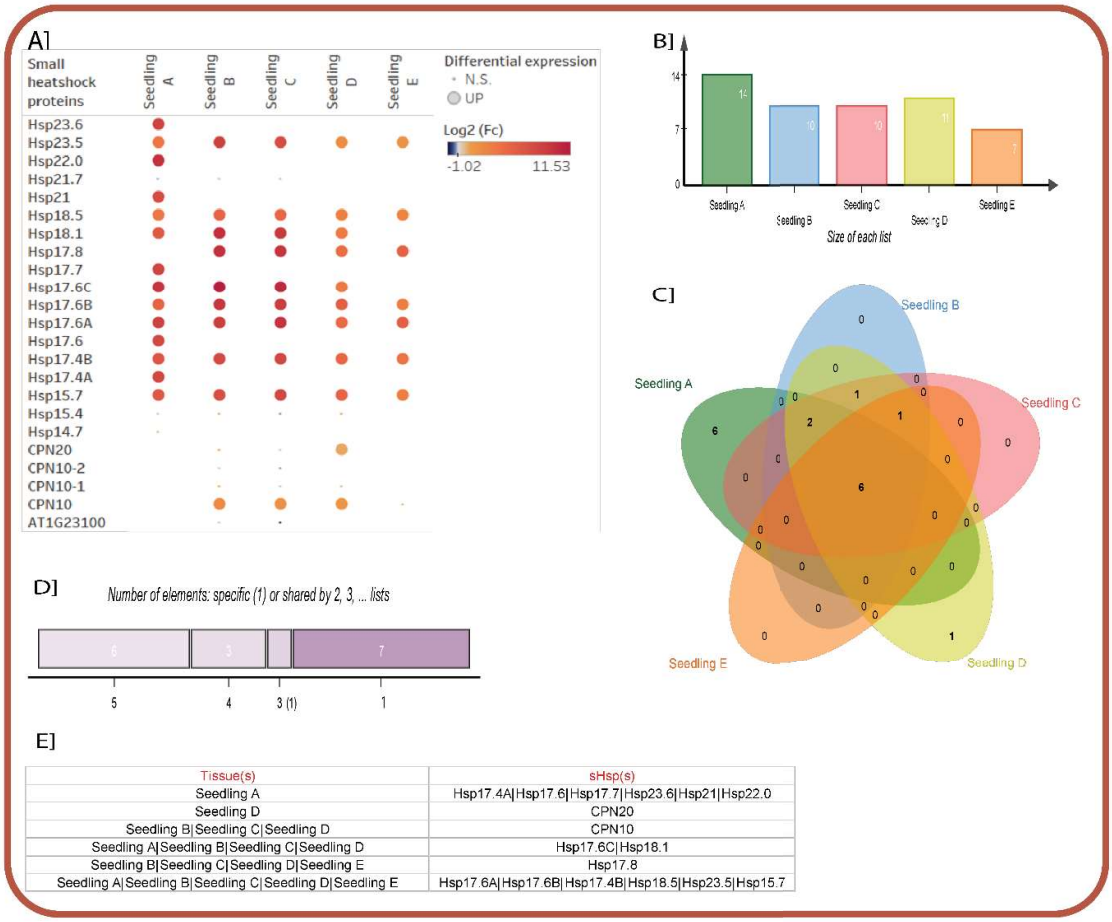
3.1.4 Expression profile of HSPs in seedlings under different stress regime (Category 2).

As mentioned earlier, this subsection aim to examine the effects of heat stress on seedlings under different clock timings and its comparison with the In-Lab data (Prof. Anil Grover, DPMB, UDSC).

Upregulated and downregulated HSPs are represented in red-orange large circles and blue small circles respectively in the dot matrix plot [small HSPs (Fig. 3.1.11A) , HSP40s (Fig. 3.1.12A), HSP60s (Fig. 3.1.13A), HSP70s (Fig. 3.1.14A), HSP90s (Fig. 3.1.15A) and HSP100s (Fig. 3.1.16A)]. Significantly upregulated genes in each seedling category are represented in the form of a bar graph [small HSPs (Fig. 3.1.11B) , HSP40s (Fig. 3.1.12B), HSP60s (Fig. 3.1.13B), HSP70s (Fig. 3.1.14B), HSP90s (Fig. 3.1.15B) and HSP100s (Fig. 3.1.16B)]. Significant downregulation of genes are represented in the form of bar graph [HSP40s (Fig. 3.1.12E). A total of 28 HSP40 genes, 7 HSP90 genes, 5 HSP100 genes, 5 HSP60 genes, 16 HSP70 genes and 17 of sHSPs genes were found to be differentially UP in at least one of the five seedling types. Number of specific (unique) and shared genes between five seedling types are represented in the form of a Venn diagram [sHSPs (Fig. 3.1.11C), HSP40s (Fig. 3.1.12C & F), HSP60s (Fig. 3.1.13C), HSP70s (Fig. 3.14.C), HSP90s (Fig. 3.1.15C) and HSP100s (Fig. 3.1.16C)] and their proportion of genes in the either specific or shared categories were represented in the form of horizontal bar [sHSPs (Fig. 3.1.11D), HSP40s (Fig. 3.1.12D & G), HSP60s (Fig. 3.1.13D), HSP70s (Fig. 3.1.14 D), HSP90s (Fig. 3.1.15D) and HSP100s (Fig. 3.1.16D)]. Sum of numbers mentioned on the ellipse or numbers on the horizontal bar conveys the total no. of transcripts in various combinations and permutations. Expression of these differentially expressed genes are summarised in form of a table [sHSPs (Fig. 3.1.11E), HSP40s (Fig. 3.1.12H & I), HSP60s (Fig. 3.1.13E), HSP70s (Fig. 3.1.14E), HSP90s (Fig. 3.1.15B) and HSP100s (Fig. 3.1.16 B)].

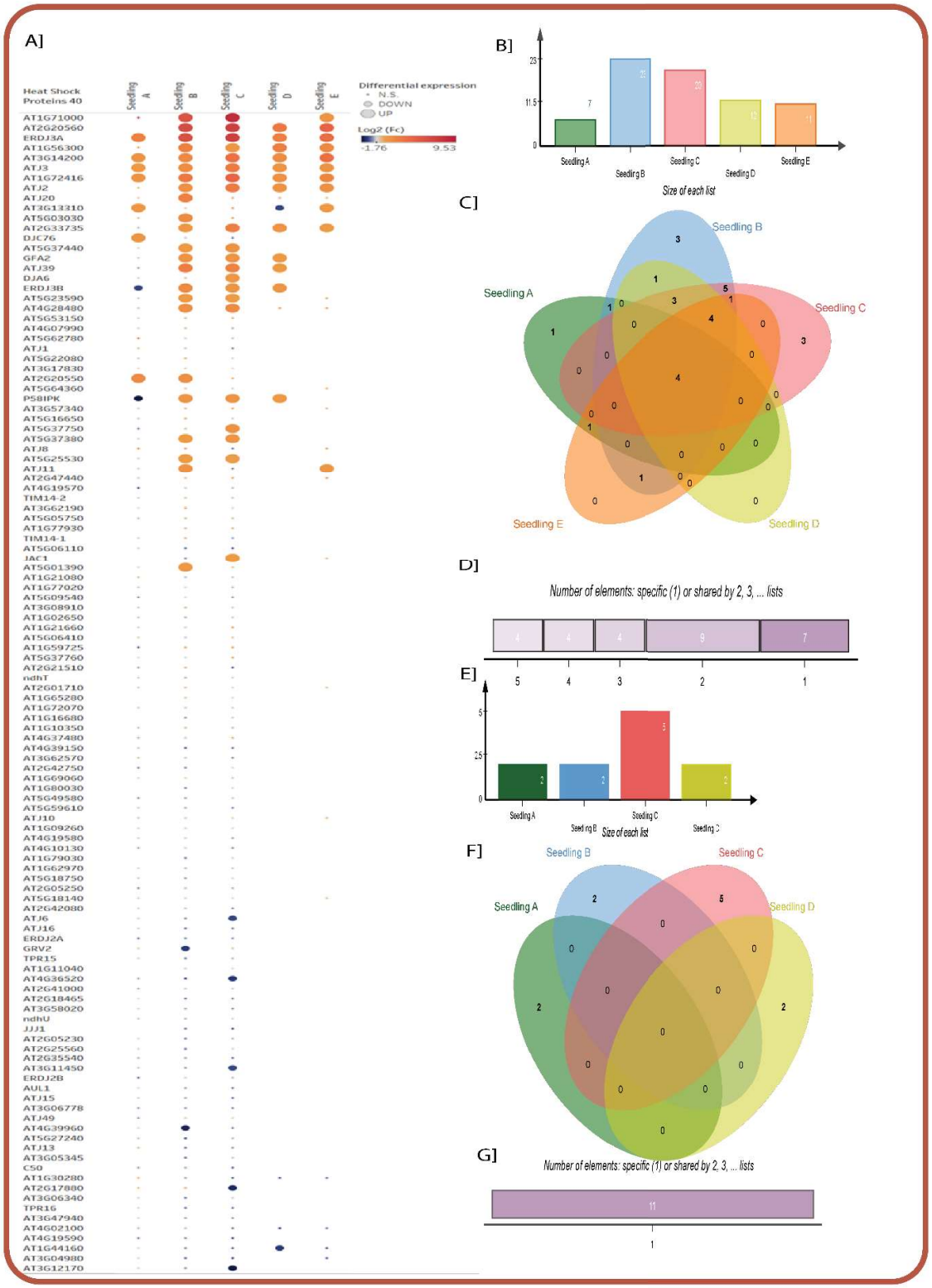
Upon heat stress, various HSPs were upregulated in all five seedling types such as HSP40s (AT1G72416, ERDJ3A, AT3G14200 & ATJ3), HSP70s (HSP70-5, AT1G79920, HSP70-8, HSP70-3, MED37C, MED37D & HSP70-10), HSP90s (Hsp90-1), HSP100s (CLPB1 & CLPB4) and sHSPs (Hsp17.6B,Hsp17.6A,Hsp17.4B,Hsp18.5, Hsp15.7 and Hsp23.5). Condition-specific upregulation of HSPs was also observed. Type C seedling specific upregulated HSP100s were SMXL6 and SMXL3; HSP70s was HSP70-14 and HSP40s were JAC1, DJA6 & AT5G37750. Similarly, type A seedling showed specific upregulation of HSP40s (DJC76) and sHSPs (Hsp17.4A, Hsp17.6, Hsp17.7, Hsp23.6, Hsp21 and Hsp22); type D seedling showed upregulation of sHSPs (CPN20); type B seedling showed upregulation of only HSP40s (ATJ20, AT5G03030 and AT5G01390). Condition specific downregulation was also

observed in every seedling type/condition. HSP40s (ERDJ3B & P58IPK), HSP70s (MED37B) and HSP90s (HSP90-7) were specifically downregulated in type A seedling. Distinctive downregulation of HSP40s (GRV2 & AT4G39960) and HSP60s (FAB1C & FAB1B) were observed in type B seedling. Type C seedling showed distinctive downregulation of HSP40s (AT2G17880, AT3G11450, AT3G12170, AT4G36520 and ATJ6). Type D seedling showed exclusive downregulation of HSP40s (AT1G44160 & AT3G13310) and type E seedling showed downregulation of HSP100s (AT4G29920). HSP100s such as CLPB1 (*At*HSP101), CLB3 & CLPB4 were upregulated in all seedling types except for type E seedling, which showed no significant differential expression of CLPB3 (Chloroplastic form CLPB protein).



Tissue(s)	sHsp(s)
Seedling A	Hsp17.4A|Hsp17.6|Hsp17.7|Hsp23.6|Hsp21|Hsp22.0
Seedling D	CPN20
Seedling B|Seedling C|Seedling D	CPN10
Seedling A|Seedling B|Seedling C|Seedling D	Hsp17.6C|Hsp18.1
Seedling B|Seedling C|Seedling D|Seedling E	Hsp17.8
Seedling A|Seedling B|Seedling C|Seedling D|Seedling E	Hsp17.6A|Hsp17.6B|Hsp17.4B|Hsp18.5|Hsp23.5|Hsp15.7

Fig.3.1.11 Expression profile of small heat shock proteins in Category 2. A] Dot-matrix plot representing expression of small heat shock proteins in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated small heat shock proteins genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05)."|" is used as a seperator
* No significant downregulation of small heat shock proteins genes was observed.


A]
Heat Shock Proteins 40
Seedling A
Seedling B
Seedling C
Seedling D
Seedling E
Differential expression
N.S.
DOWN
UP
Log2 (Fc)
-1.76
9.53
AT1G71000
AT2G20560
ERDJ3A
AT1G56300
AT3G14200
ATJ3
AT1G72416
ATJ2
ATJ20
AT3G13310
AT5G03030
AT2G33735
DJC76
AT5G37440
GFA2
ATJ39
DJA6
ERDJ3B
AT5G23590
AT4G28480
AT5G53150
AT4G07990
AT5G62780
ATJ1
AT5G22080
AT3G17830
AT2G20550
AT5G64360
P58IPK
AT3G57340
AT5G16650
AT5G37750
AT5G37380
ATJ8
AT5G25530
ATJ11
AT2G47440
AT4G19570
TIM14-2
AT3G62190
AT5G05750
AT1G77930
TIM14-1
AT5G06110
JAC1
AT5G01390
AT1G21080
AT1G77020
AT5G09540
AT3G08910
AT1G02650
AT1G21660
AT5G06410
AT1G59725
AT5G37760
AT2G21510
ndhT
AT2G01710
AT1G65280
AT1G72070
AT1G16680
AT1G10350
AT4G37480
AT4G39150
AT3G62570
AT2G42750
AT1G69060
AT1G80030
AT5G49580
AT5G59610
ATJ10
AT1G09260
AT4G19580
AT4G10130
AT1G79030
AT1G62970
AT5G18750
AT2G05250
AT5G18140
AT2G42080
ATJ6
ATJ16
ERDJ2A
GRV2
TPR15
AT1G11040
AT4G36520
AT2G41000
AT2G18465
AT3G58020
ndhU
JJJ1
AT2G05230
AT2G25560
AT2G35540
AT3G11450
ERDJ2B
AUL1
ATJ15
AT3G06778
ATJ49
AT4G39960
AT5G27240
ATJ13
AT3G05345
C50
AT1G30280
AT2G17880
AT3G06340
TPR16
AT3G47940
AT4G02100
AT4G19590
AT1G44160
AT3G04980
AT3G12170
B]
Seedling A
Seedling B
Seedling C
Seedling D
Seedling E
Size of each list
C]
Seedling A
Seedling B
Seedling C
Seedling D
Seedling E
D]
Number of elements: specific (1) or shared by 2, 3, ... lists
E]
Seedling A
Seedling B
Seedling C
Seedling D
Size of each list
F]
Seedling A
Seedling B
Seedling C
Seedling D
G]
Number of elements: specific (1) or shared by 2, 3, ... lists

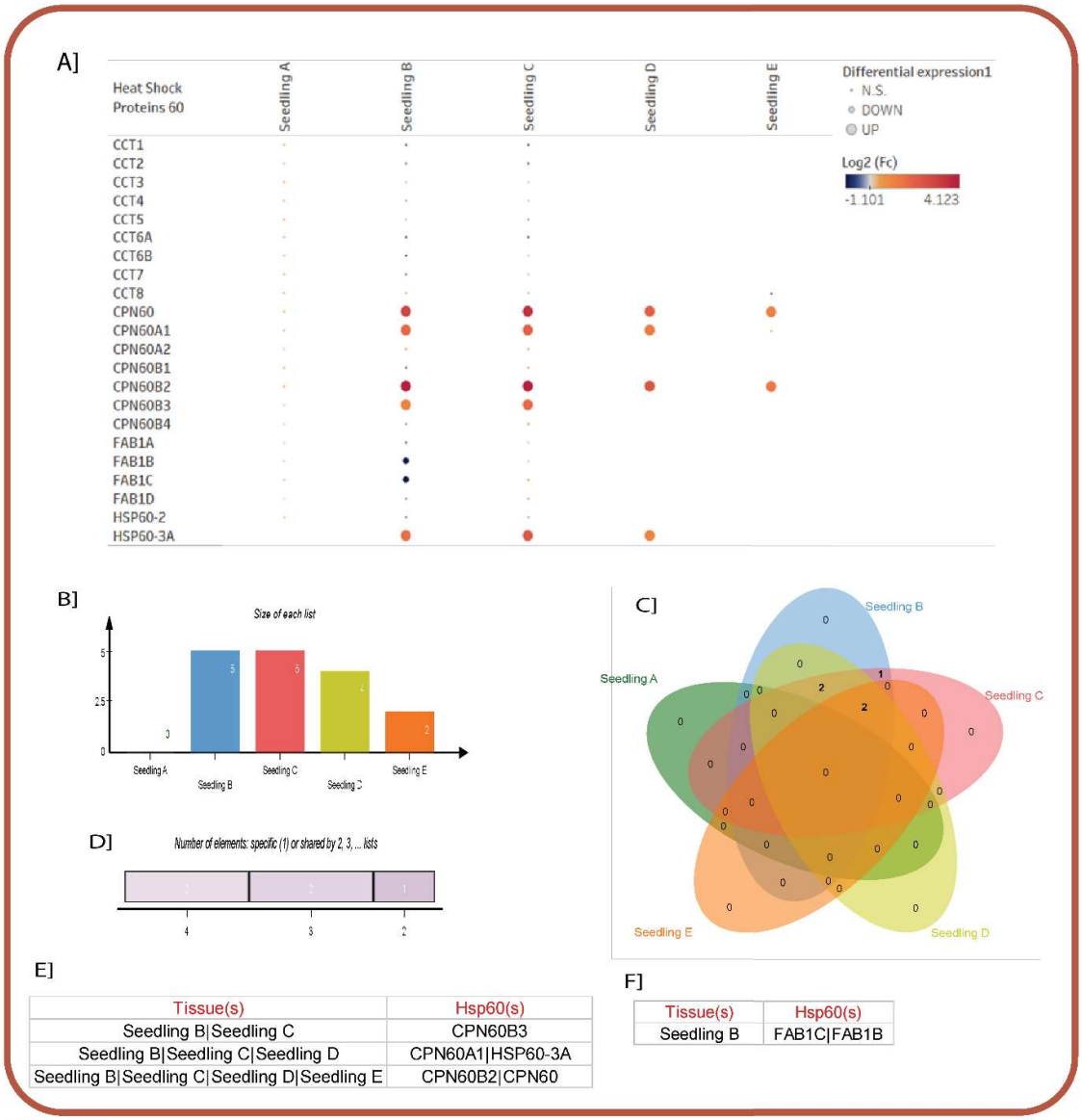
H]

Tissue(s)	Hsp40(s)
Seedling A	DJC76
Seedling B	ATJ20|AT5G01390|AT5G03030
Seedling C	JAC1|DJA6|AT5G37750
Seedling A|Seedling B	AT2G20550
Seedling A|Seedling E	AT3G13310
Seedling B|Seedling C	AT4G28480|AT5G23590|AT5G25530|AT5G37380
Seedling B|Seedling D	GFA2
Seedling B|Seedling E	ATJ11
Seedling B|Seedling C|Seedling D	ATJ39|ERDJ3B|P58IPK|
Seedling B|Seedling C|Seedling E	AT1G71000
Seedling B|Seedling C|Seedling D|Seedling E	AT1G56300|AT2G20560|AT2G33735|ATJ2
Seedling A|Seedling B|Seedling C|Seedling D|Seedling E	AT1G72416|ERDJ3A|AT3G14200|ATJ3

I]

Tissue(s)	Hsp40(s)
Seedling A	ERDJ3B|P58IPK
Seedling B	GRV2|AT4G39960
Seedling C	AT2G17880|AT3G11450|AT3G12170|AT4G36520|ATJ6
Seedling D	AT1G44160|AT3G13310

Fig.3.1.12 Expression profile of heat shock proteins 40 in Category 2. A] Dot-matrix plot representing expression of heat shock proteins 40 in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock proteins 40 genes.D] No. of elements in each combination in C[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].E] Bar graph for no. of significantly downregulated genes in each list of this category.F] Venn diagram for significantly downregulated heat shock proteins 40 genes. G]No. of elements in each combination in F[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)]. B to G are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).H] Table of genes significantly upregulated in this category.I] Table of genes significantly downregulated in this category N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator



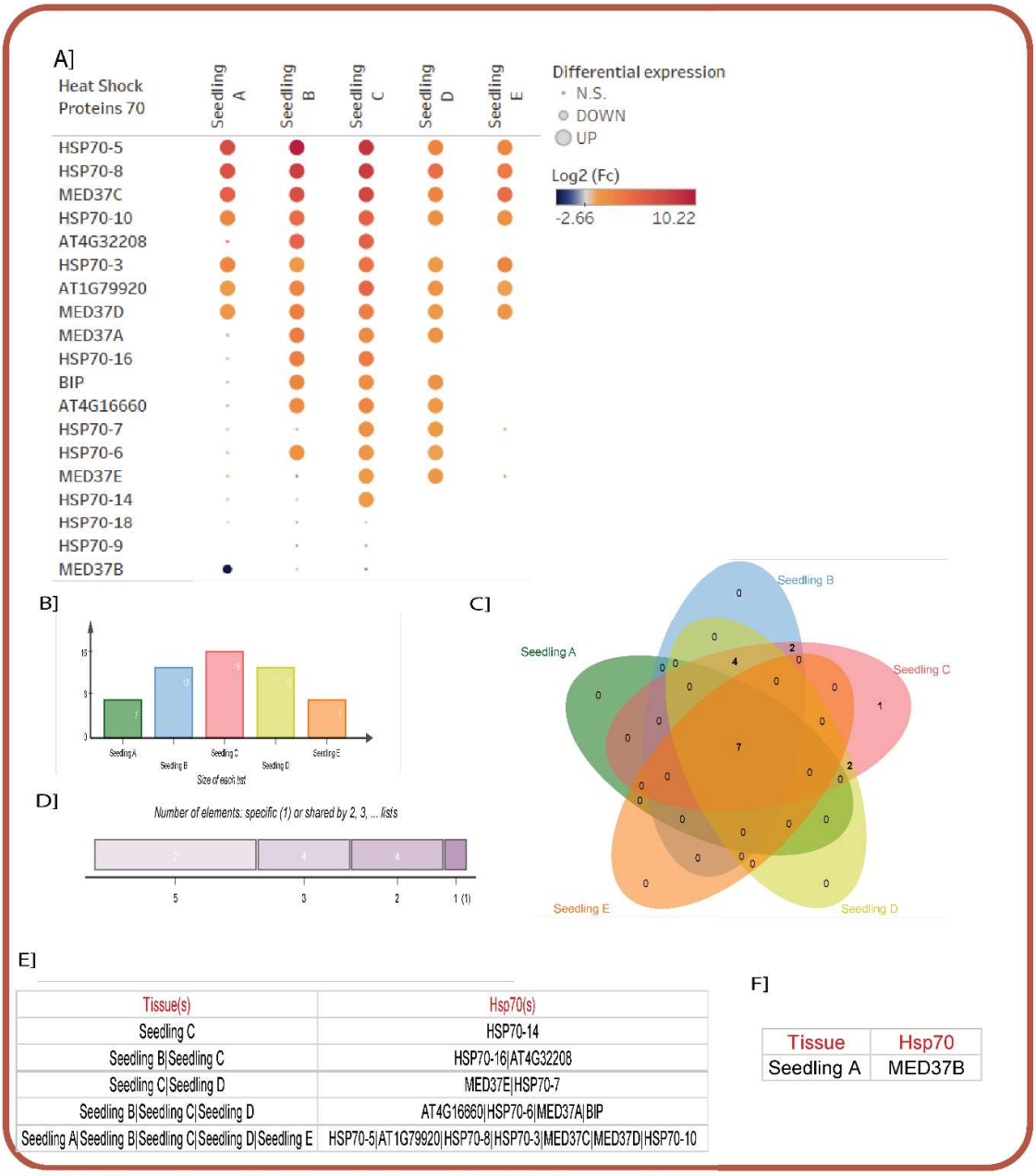
E]

Tissue(s)	Hsp60(s)
Seedling B|Seedling C	CPN60B3
Seedling B|Seedling C|Seedling D	CPN60A1|HSP60-3A
Seedling B|Seedling C|Seedling D|Seedling E	CPN60B2|CPN60

F]

Tissue(s)	Hsp60(s)
Seedling B	FAB1C|FAB1B

Fig.3.1.13 Expression profile of heat shock protein 60(s) in Category 2. A] Dot-matrix plot representing expression of heat shock protein 60(s) in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 60(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly downregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05)."|" is used as a seperator .



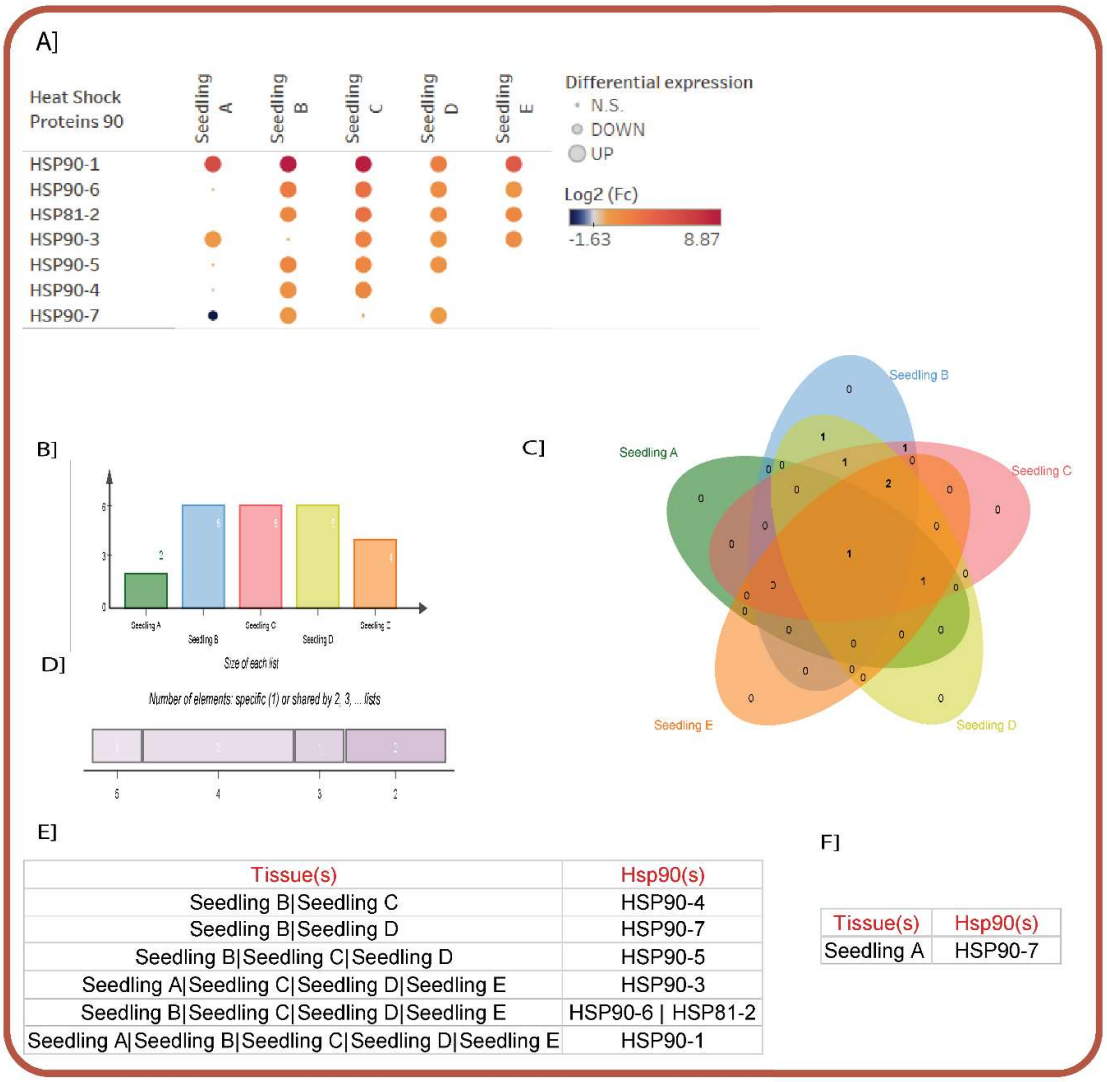
E]

Tissue(s)	Hsp70(s)
Seedling C	HSP70-14
Seedling B|Seedling C	HSP70-16|AT4G32208
Seedling C|Seedling D	MED37E|HSP70-7
Seedling B|Seedling C|Seedling D	AT4G16660|HSP70-6|MED37A|BIP
Seedling A|Seedling B|Seedling C|Seedling D|Seedling E	HSP70-5|AT1G79920|HSP70-8|HSP70-3|MED37C|MED37D|HSP70-10

F]

Tissue	Hsp70
Seedling A	MED37B

Fig.3.1.14 Expression profile of heat shock protein 70(s) in Category 2. A] Dot-matrix plot representing expression of heat shock protein 70(s) in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 70(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly downregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator .



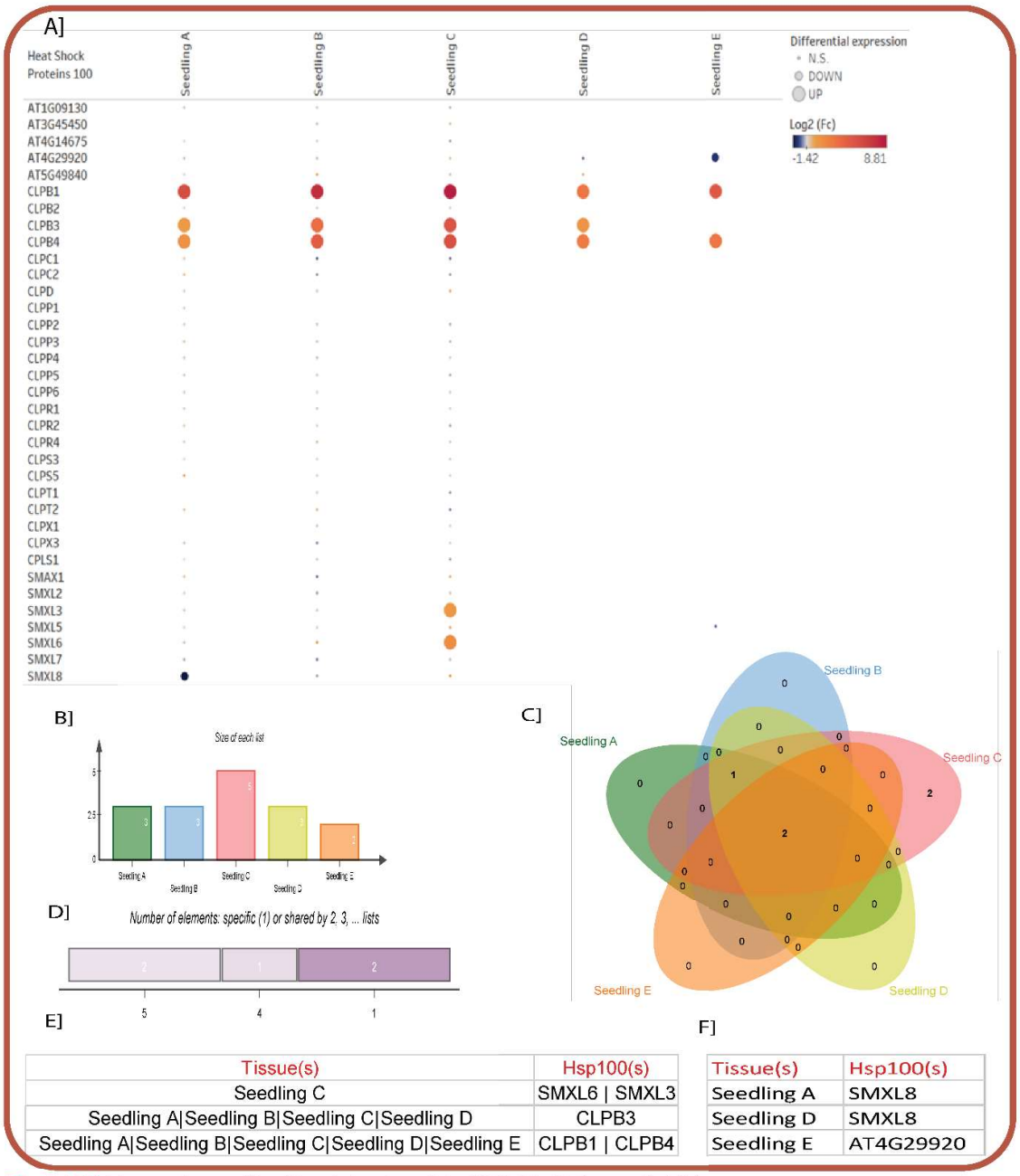
E]

Tissue(s)	Hsp90(s)
Seedling B|Seedling C	HSP90-4
Seedling B|Seedling D	HSP90-7
Seedling B|Seedling C|Seedling D	HSP90-5
Seedling A|Seedling C|Seedling D|Seedling E	HSP90-3
Seedling B|Seedling C|Seedling D|Seedling E	HSP90-6 | HSP81-2
Seedling A|Seedling B|Seedling C|Seedling D|Seedling E	HSP90-1

F]

Tissue(s)	Hsp90(s)
Seedling A	HSP90-7

Fig.3.1.15 Expression profile of heat shock protein 90(s) in Category 2. A] Dot-matrix plot representing expression of heat shock protein 90(s) in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 90(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly downregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator .



Tissue(s)	Hsp100(s)
Seedling C	SMXL6 | SMXL3
Seedling A|Seedling B|Seedling C|Seedling D	CLPB3
Seedling A|Seedling B|Seedling C|Seedling D|Seedling E	CLPB1 | CLPB4

Tissue(s)	Hsp100(s)
Seedling A	SMXL8
Seedling D	SMXL8
Seedling E	AT4G29920

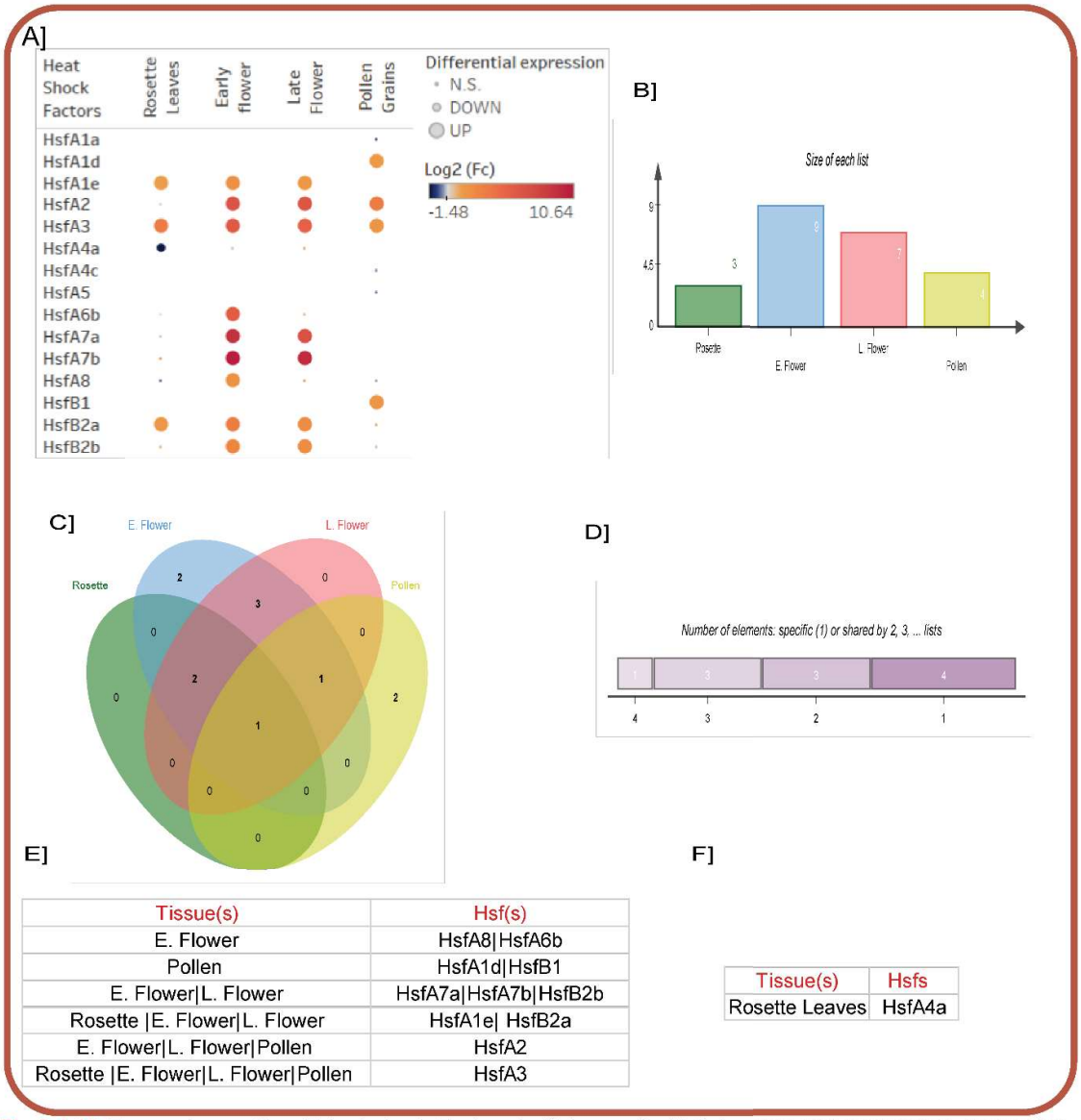
Fig.3.1.16 Expression profile of heat shock protein 100(s) in Category 2. A] Dot-matrix plot representing expression of heat shock protein 100(s) in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 100(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly downregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05)."|" is used as a seperator.

3.1.5 Expression profile of HSFs in vegetative and reproductive tissues (Category 3).

This category examines the tissue-specific effects of heat stress on vegetative and reproductive tissues of *A. thaliana*. The publicly available RNA-seq dataset (Rahmati Ishka et al., 2018; Zhang et al., 2017) was categorized into four tissue types (rosette leaves, early flowers, late flowers and pollen grains) .

Upregulated and downregulated HSFs are represented as red-orange large circles and blue small circles respectively in the dot matrix plot (Fig. 3.1.17A). Significant upregulation of genes in each seedling category are represented in the form of a bar graph (Fig. 3.1.17B). A total of 11 HSF genes were found to be differentially upregulated in at least one of the mentioned tissues. The number of unique and shared HSF genes between four tissue type are represented in the form of a Venn diagram (Fig. 3.1.17C) and their proportion of genes in the either specific or shared list in each category is represented in the form of a horizontal bar (Fig. 3.1.17D). The sum of numbers mentioned on the ellipse or numbers on the horizontal bar conveys the total no. of transcripts in various combinations and permutations. Expression of these differentially expressed genes are summarised in form of a table (Fig. 3.1.17E and F).

Upon heat stress, HsfA3 was found to be exclusively upregulated in all four tissue types. Only studied reproductive tissues showed upregulation of HsfA2. Flower (both early and late) specific upregulated HSF genes were also observed (HsfA7a, HsfA7b and HsfB2B). Distinctively upregulation of HsfA8 and HsfA6 were observed in early flowers. Pollen specific upregulated genes were also observed (HsfA1d and HsfB1). Rosette leaves showed exclusive downregulation of HsfA4a genes.



Tissue(s)	Hsf(s)
E. Flower	HsfA8|HsfA6b
Pollen	HsfA1d|HsfB1
E. Flower|L. Flower	HsfA7a|HsfA7b|HsfB2b
Rosette |E. Flower|L. Flower	HsfA1e| HsfB2a
E. Flower|L. Flower|Pollen	HsfA2
Rosette |E. Flower|L. Flower|Pollen	HsfA3

Tissue(s)	Hsfs
Rosette Leaves	HsfA4a

Fig.3.1.17 Expression profile of Heat shock Factors in Category 3. A] Dot-matrix plot represnting expression of heat shock factors in vegetative and reproductive of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock factor genes.D] No. of elements in each combina-tion[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly downregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria (-1≥Log2(Fold Change)≥1, Adjusted p-value ≤ 0.05). "|" is used as a seperator.

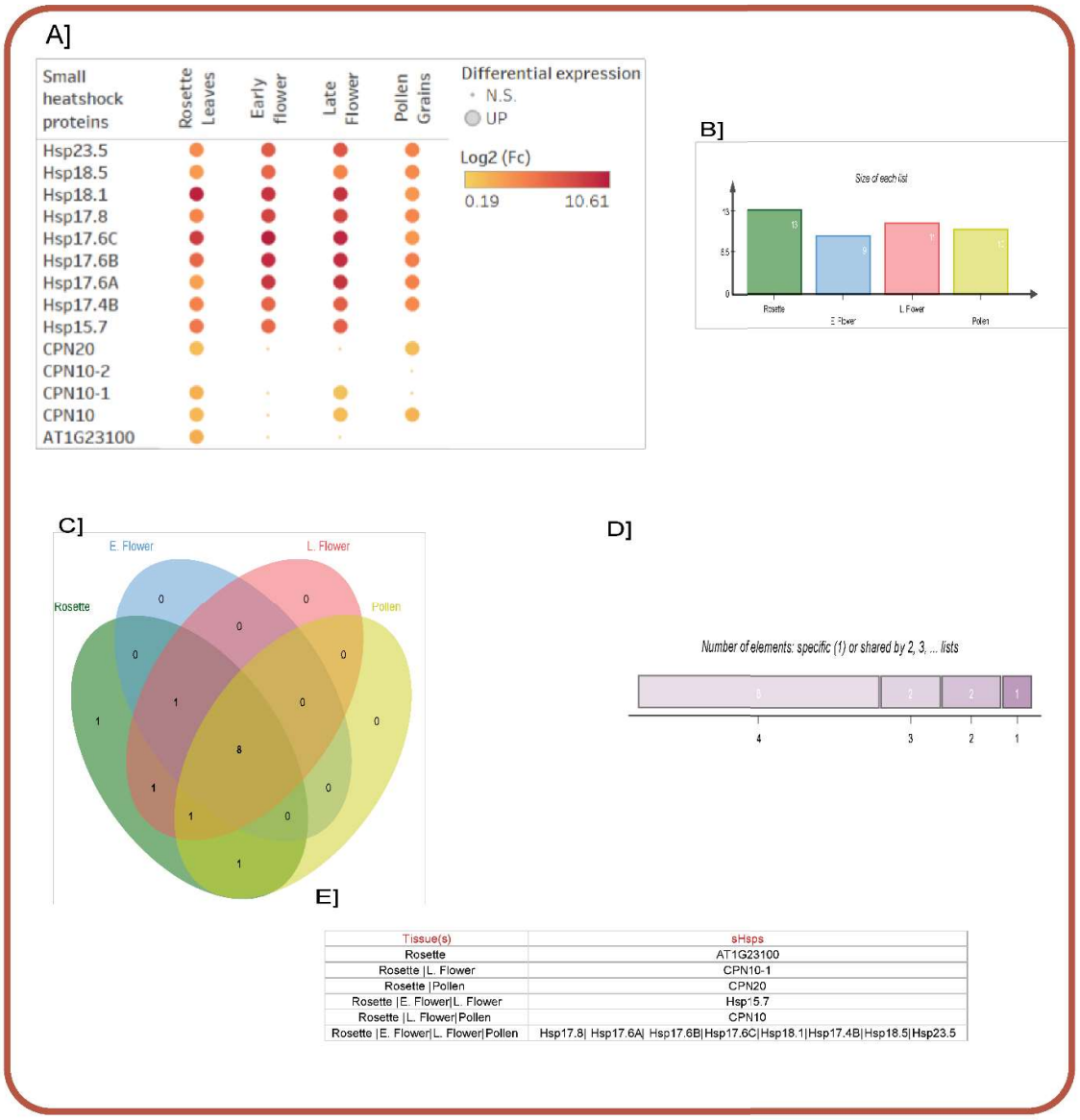
3.1.6 Expression profile of HSPs in vegetative and reproductive tissues (Category 3).

As mentioned earlier, this subsection proposes to examine the effects of heat stress on the transcriptomic profile of HSP genes in vegetative and reproductive tissues of *A. thaliana.*

Upregulated and downregulated HSPs are represented in red-orange large circles and blue small circles respectively in the dot matrix plot [small HSPs (Fig. 3.1.18A) , HSP40s (Fig. 3.1.19A), HSP60s (Fig. 3.1.20A), HSP70s (Fig. 3.1.21A), HSP90s (Fig. 3.1.22A) and HSP100s (Fig. 3.1.23A)]. Significant upregulation of genes in each tissue type were represented in the form of a bar graph [small HSPs (Fig. 3.1.18B) , HSP40s (Fig. 3.1.19 B), HSP60s (Fig. 3.1.20B), HSP70s (Fig. 3.1.21B), HSP90s (Fig. 3.1.22B) and HSP100s (Fig. 3.1.23B)]. Significant downregulation of genes are also represented in the form bar graph [HSP40s (Fig. 3.1.19E)]. A total of thirty five HSP40 genes, six HSP90 genes, ten HSP100 genes, eleven HSP60 genes, fifteen HSP70 genes and thirteen of sHSPs genes were found to be differentially upregulated in at least one of the tissue types. Number of specific (unique) and shared genes between tissue types are represented in the form of a Venn diagram [sHSPs (Fig. 3.1.18C), HSP40s (Fig. 3.1.19C & F), HSP60s (Fig. 3.1.20C), HSP70s (Fig. 3.21 C), HSP90s (Fig. 3.1.22C) and HSP100s (Fig. 3.1.23C)] and the distribution of genes with respect to the shared elements as well as the specifically expressed genes are represented in the form of a horizontal bar [sHSPs (Fig. 3.1.18D), HSP40s (Fig. 3.1.19D & G), HSP60s (fig. 3.1.20D), HSP70s (Fig. 3.1.21D), HSP90s (Fig. 3.1.22D) and HSP100s (Fig. 3.1.23D)]. Sum of numbers mentioned on the ellipse or numbers on the horizontal bar conveys the total no. of transcripts in various combinations and permutations. Expression of these differentially expressed genes are summarised in form of a table [sHSPs (Fig. 3.1.18 E), HSP40s (Fig. 3.1.19H & I), HSP60s (Fig. 3.1.20E), HSP70s (Fig. 3.1.21E), HSP90s (Fig. 3.1.22B) and HSP100s (Fig. 3.1.23B)].

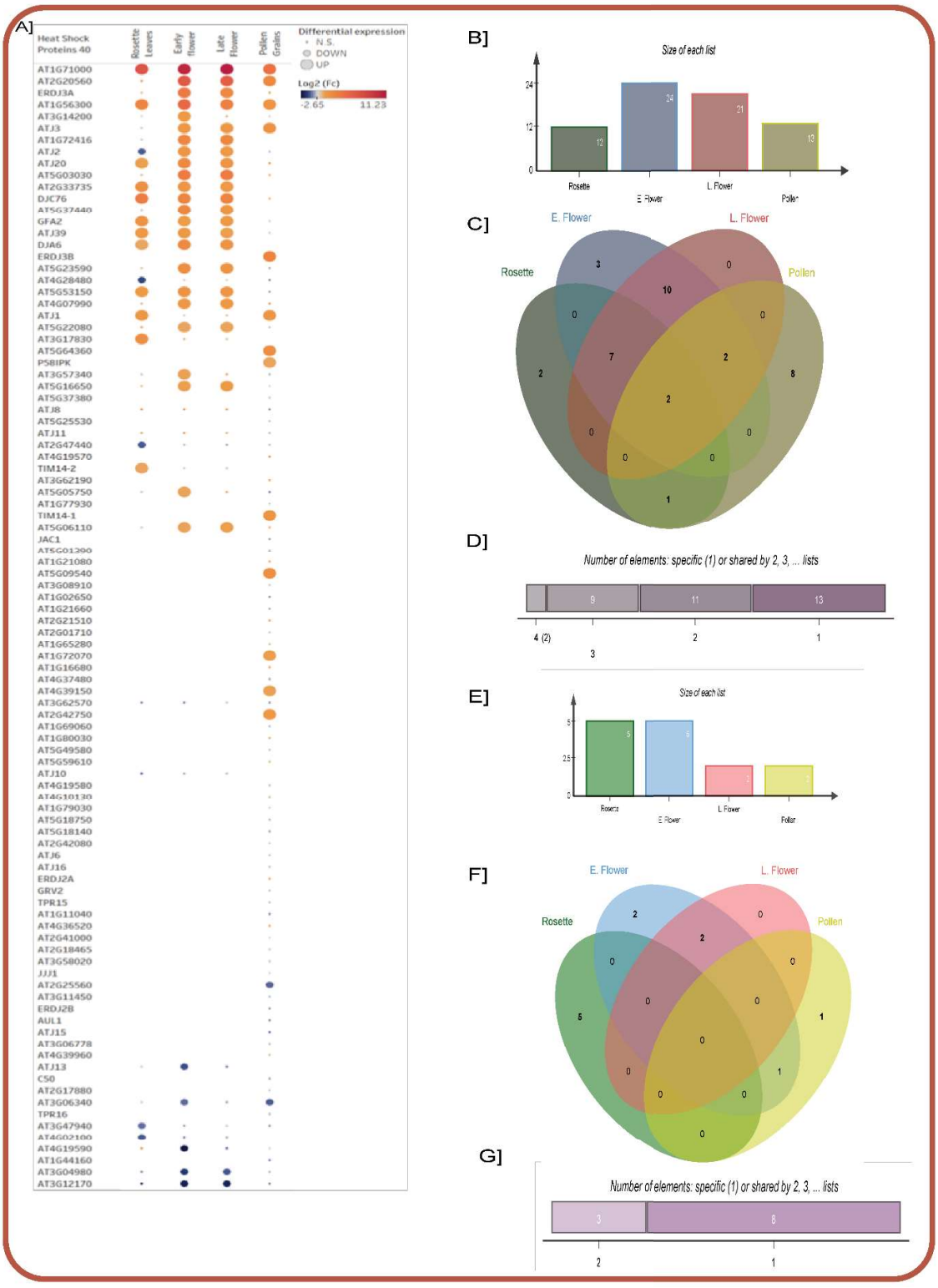
Upon heat stress, various HSPs were upregulated in all tissue types such as HSP40s (AT1G56300 & AT1G71000), HSP70s (HSP70-5, HSP70-8 & HSP70-6), HSP90s (Hsp90-5), HSP100s (CLPB4) and sHSPs (Hsp17.6B,Hsp17.6A,Hsp17.8,Hsp17.6C, Hsp18.1,Hsp18.5, Hsp23.5 and Hsp17.4B). Tissue/cells specific upregulated HSPs were also observed. Rosette leaf specific upregulated HSP100s was CLPC; HSP60s was HSP60-2; sHSP was AT3G23100 and HSP40s were TIM14-2 & AT3G17830. Similarly, early stage flowers showed specific upregulation of HSP40s (AT3G14200, AT3G57340 and AT5G05750), HSP60 (FAB1D) and HSP100s (SMXL6, CLPC1, CLPD and SMAX1); Late flowers showed upregulation of HSP60s (FAB1C) and HSP70s (HSP70-16); Pollen grains showed upregulation of HSP40s (AT1G72070, TIM14-1, AT2G42750 AT4G39150, AT5G09540,

AT5G64360, ERDJ3B & P58IPK), HSP60 (CCT5) HSP70s (MED37B, MED37D, MED37E,MED37A and BIP) , HSP90s (HSP90-7 and HSP90-3) and HSP100(CLPP3). Condition-specific downregulated genes were also observed in every tissue /cells except for late flowers. HSP40s (AT2G47440, AT3G47940, AT4G02100, AT4G28480 and ATJ2), HSP70s (MED37D) and HSP100s (AT4G29920 and SMAX1) were specifically downregulated in rosette leaves. Distinctive downregulation of HSP40s (AT2G25560), HSP100s (AT1G09130 and SMXL7) and HSP60s (CCT1 and CCT2) were observed in pollen grains. Early flowers showed downregulation of HSP40s (AT4G19590 and ATJ13). HSP100s such as CLPB1 (*At*HSP101) and CLPB3 were upregulated in all tissue types except for rosette leaves, which showed no significant differential expression of these genes.



Tissue(s)	sHsps
Rosette	AT1G23100
Rosette |L. Flower	CPN10-1
Rosette |Pollen	CPN20
Rosette |E. Flower|L. Flower	Hsp15.7
Rosette |L. Flower|Pollen	CPN10
Rosette |E. Flower|L. Flower|Pollen	Hsp17.8| Hsp17.6A| Hsp17.6B|Hsp17.6C|Hsp18.1|Hsp17.4B|Hsp18.5|Hsp23.5

Fig.3.1.18 Expression profile of small heat shock proteins in Category 3. A] Dot-matrix plot representing expression of small heat shock proteins in Vegetative and reproductive tissue of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated small heat shock proteins genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator.
* No significant downregulation of small heat shock proteins genes was observed.


A]
Heat Shock Proteins 40
Rosette Leaves
Early flower
Late Flower
Pollen Grains
Differential expression
N.S.
DOWN
UP
Log2 (Fc)
-2.65
11.23
AT1G71000
AT2G20560
ERDJ3A
AT1G56300
AT3G14200
ATJ3
AT1G72416
ATJ2
ATJ20
AT5G03030
AT2G33735
DJC76
GFA2
ATJ39
DJA6
ERDJ3B
AT5G23590
AT4G28480
AT5G53150
AT4G07990
ATJ1
AT5G22080
AT3G17830
AT5G64360
PSBIPK
AT3G57340
AT5G16650
AT5G37380
ATJ8
AT5G25530
ATJ11
AT2G47440
AT4G19570
TIM14-2
AT3G62190
AT5G05750
AT1G77930
TIM14-1
AT5G06110
JAC1
AT1G21080
AT5G09540
AT3G08910
AT1G02650
AT1G21660
AT2G21510
AT2G01710
AT1G65280
AT1G72070
AT1G16680
AT4G37480
AT4G39150
AT3G62570
AT2G42750
AT1G69060
AT1G80030
AT5G49580
AT5G59610
ATJ10
AT4G19580
AT1G79030
AT5G18750
AT5G18140
AT2G42080
ATJ6
ATJ16
ERDJ2A
GRV2
TPR15
AT1G11040
AT4G36520
AT2G41000
AT2G18465
AT3G58020
JJJ1
AT2G25560
AT3G11450
ERDJ2B
AUL1
ATJ15
AT3G06778
AT4G39960
ATJ13
C50
AT2G17880
AT3G06340
TPR16
AT3G47940
AT4G19590
AT1G44160
AT3G04980
AT3G12170
B]
Size of each list
24
12
0
Rosette
E. Flower
L. Flower
Poller
C]
E. Flower
L. Flower
Rosette
Pollen
D]
Number of elements: specific (1) or shared by 2, 3, ... lists
9
11
13
4 (2)
2
1
3
E]
Size of each list
5
2.5
0
Rosette
E. Flower
L. Flower
Poller
F]
E. Flower
L. Flower
Rosette
Pollen
G]
Number of elements: specific (1) or shared by 2, 3, ... lists
3
6
2
1

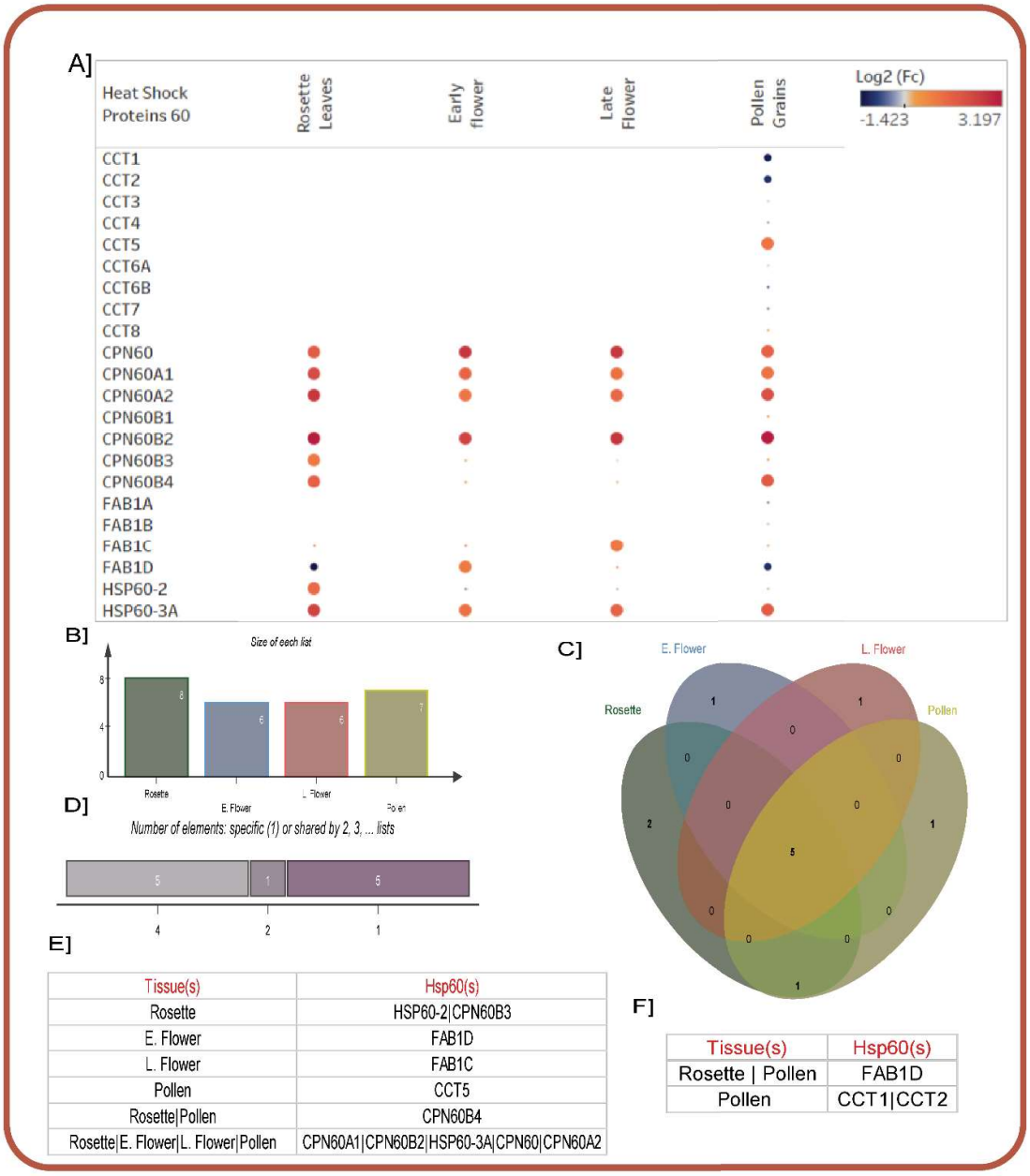
H]

Tissue(s)	Hsp40(s)
Rosette	TIM14-2|AT3G17830
E. Flower	AT3G14200|AT3G57340|AT5G05750
Pollen	AT1G72070|TIM14-1|AT2G42750|ERDJ3B|AT4G39150|P58IPK|AT5G09540|AT5G64360
Rosette |Pollen	ATJ1
E. Flower|L. Flower	AT1G72416|ERDJ3A|AT4G07990|AT5G03030|AT5G06110|AT5G16650|ATJ2|AT5G22080|AT5G23590|AT5G37440
Rosette |E. Flower|L. Flower	ATJ39|DJA6|AT2G33735|ATJ20|DJC76|GFA2|AT5G53150
E. Flower|L. Flower|Pollen	AT2G20560|ATJ3
Rosette |E. Flower|L. Flower|Pollen	AT1G56300|AT1G71000

I]

Tissue(s)	Hsp40(s)
Rosette	AT2G47440|AT3G47940|AT4G02100|AT4G28480|ATJ2
E. Flower	ATJ13|AT4G19590
Pollen	AT2G25560
E. Flower|L. Flower	AT3G04980|AT3G12170
E. Flower|Pollen	AT3G06340

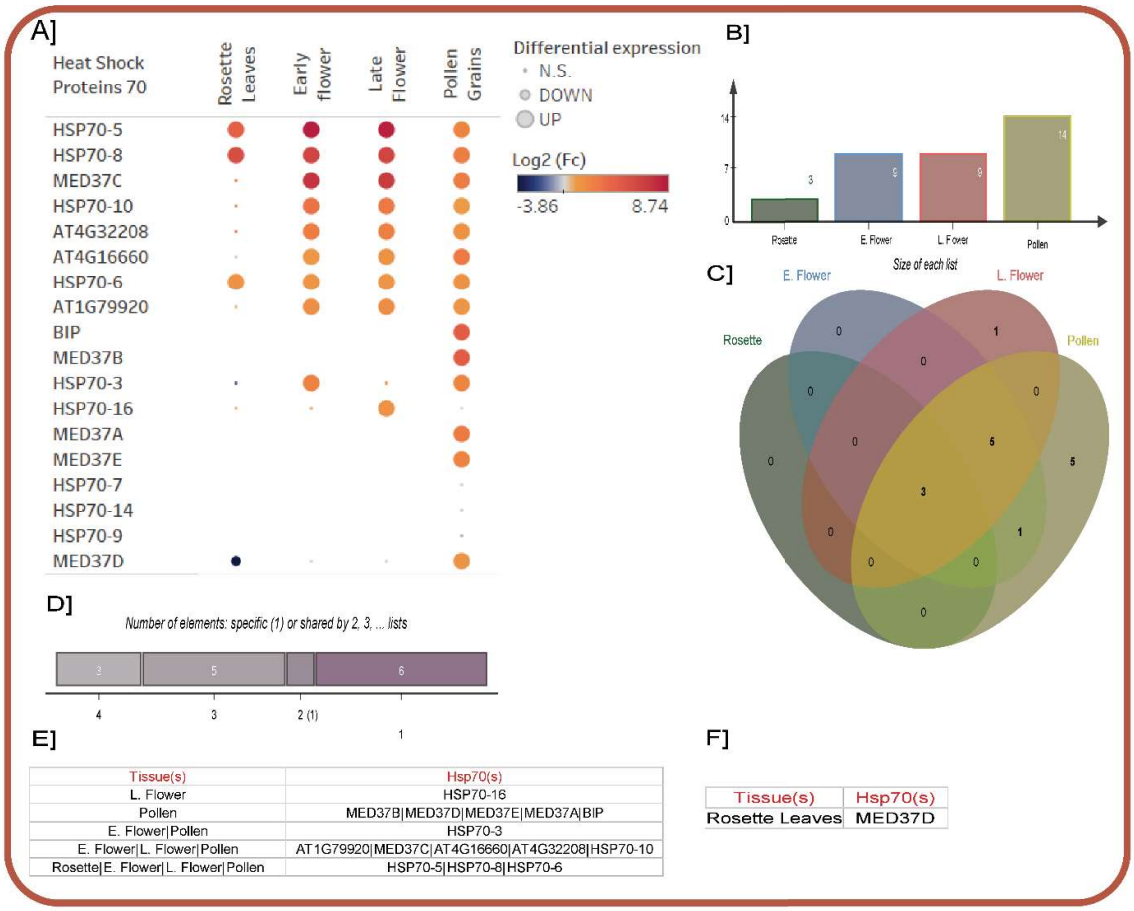
Fig.3.1.19 Expression profile of heat shock proteins 40 in Category 3. A] Dot-matrix plot representing expression of heat shock proteins 40s in vegetative and reproductive tissue of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 40 genes.D] No. of elements in each combination in C[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].E] Bar graph for no. of significantly downregulated genes in each list of this category.F] Venn diagram for significantly downregulated heat shock proteins 40 genes. G]No. of elements in each combination in F[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)]. B to G are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).H] Table of genes significantly upregulated in this category.I] Table of genes significantly downregulated in this category N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Down-regulation at the selected Criteria (-1≥Log2(Fold Change)≥1, Adjusted p-value ≤ 0.05). "|" is used as a seperator.



Tissue(s)	Hsp60(s)
Rosette	HSP60-2|CPN60B3
E. Flower	FAB1D
L. Flower	FAB1C
Pollen	CCT5
Rosette|Pollen	CPN60B4
Rosette|E. Flower|L. Flower|Pollen	CPN60A1|CPN60B2|HSP60-3A|CPN60|CPN60A2

Tissue(s)	Hsp60(s)
Rosette | Pollen	FAB1D
Pollen	CCT1|CCT2

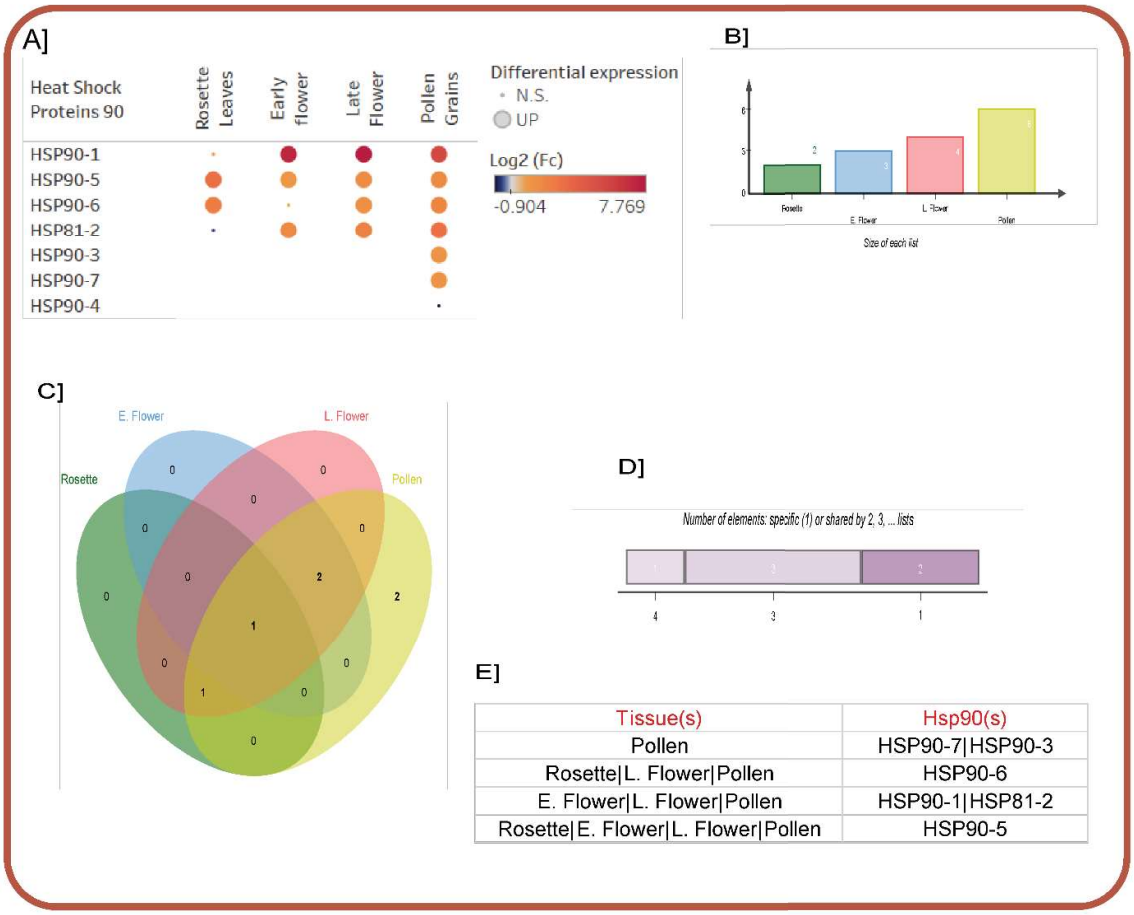
Fig.3.1.20 Expression profile of heat shock protein 60(s) in Category 3. A] Dot-matrix plot representing expression of heat shock protein 60(s) in vegetative and reproductive tissues of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 60(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly downregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as an seperator.



Tissue(s)	Hsp70(s)
L. Flower	HSP70-16
Pollen	MED37B|MED37D|MED37E|MED37A|BIP
E. Flower|Pollen	HSP70-3
E. Flower|L. Flower|Pollen	AT1G79920|MED37C|AT4G16660|AT4G32208|HSP70-10
Rosette|E. Flower|L. Flower|Pollen	HSP70-5|HSP70-8|HSP70-6

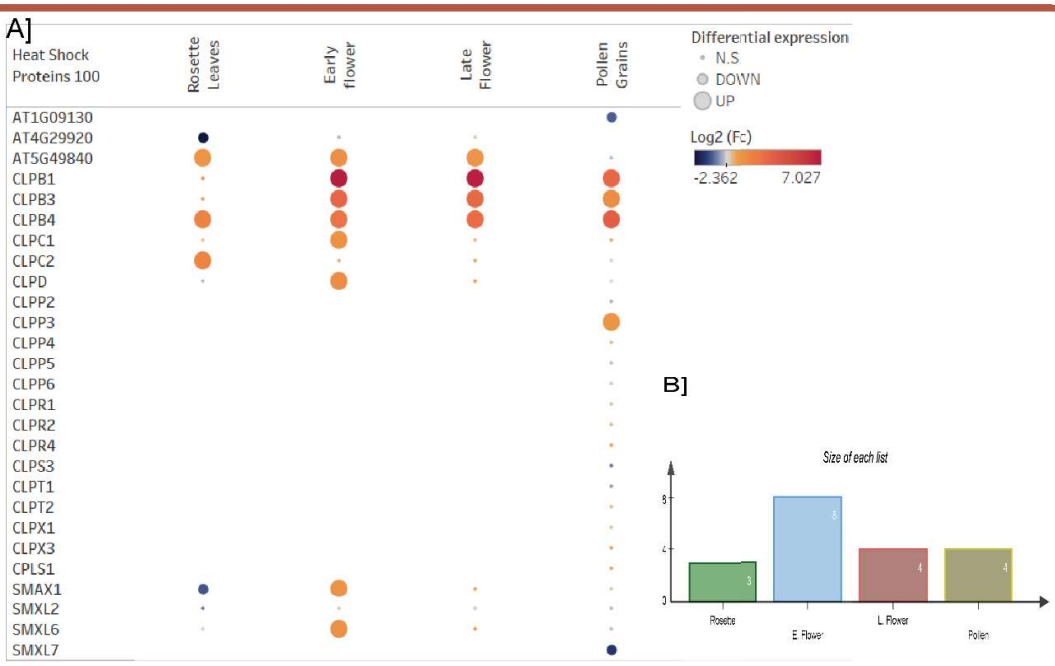
Tissue(s)	Hsp70(s)
Rosette Leaves	MED37D

Fig.3.1.21 Expression profile of heat shock protein 70(s) in Category 3. A] Dot-matrix plot representing expression of heat shock protein 70(s) in vegetative and reproductive tissues of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 70(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly downregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as an seperator.



Tissue(s)	Hsp90(s)
Pollen	HSP90-7|HSP90-3
Rosette|L. Flower|Pollen	HSP90-6
E. Flower|L. Flower|Pollen	HSP90-1|HSP81-2
Rosette|E. Flower|L. Flower|Pollen	HSP90-5

Fig.3.1.22 Expression profile of heat shock protein 90(s) in Category 3. A] Dot-matrix plot representing expression of heat shock protein 90(s) in vegetative and reproductive tissues of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 90(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as an seperator
*No significant downregulation was observed.



E]

Tissue(s)	Hsp100(s)
Rosette	CLPC2
E. Flower	SMXL6|CLPC1|CLPD|SMAX1
Pollen	CLPP3
Rosette |E. Flower|L. Flower	AT5G49840
E. Flower|L. Flower|Pollen	CLPB1|CLPB3
Rosette |E. Flower|L. Flower|Pollen	CLPB4

F]

Tissue(s)	Hsp100(s)
Rosette	AT4G29920|SMAX1
Pollen	AT1G09130|SMXL7

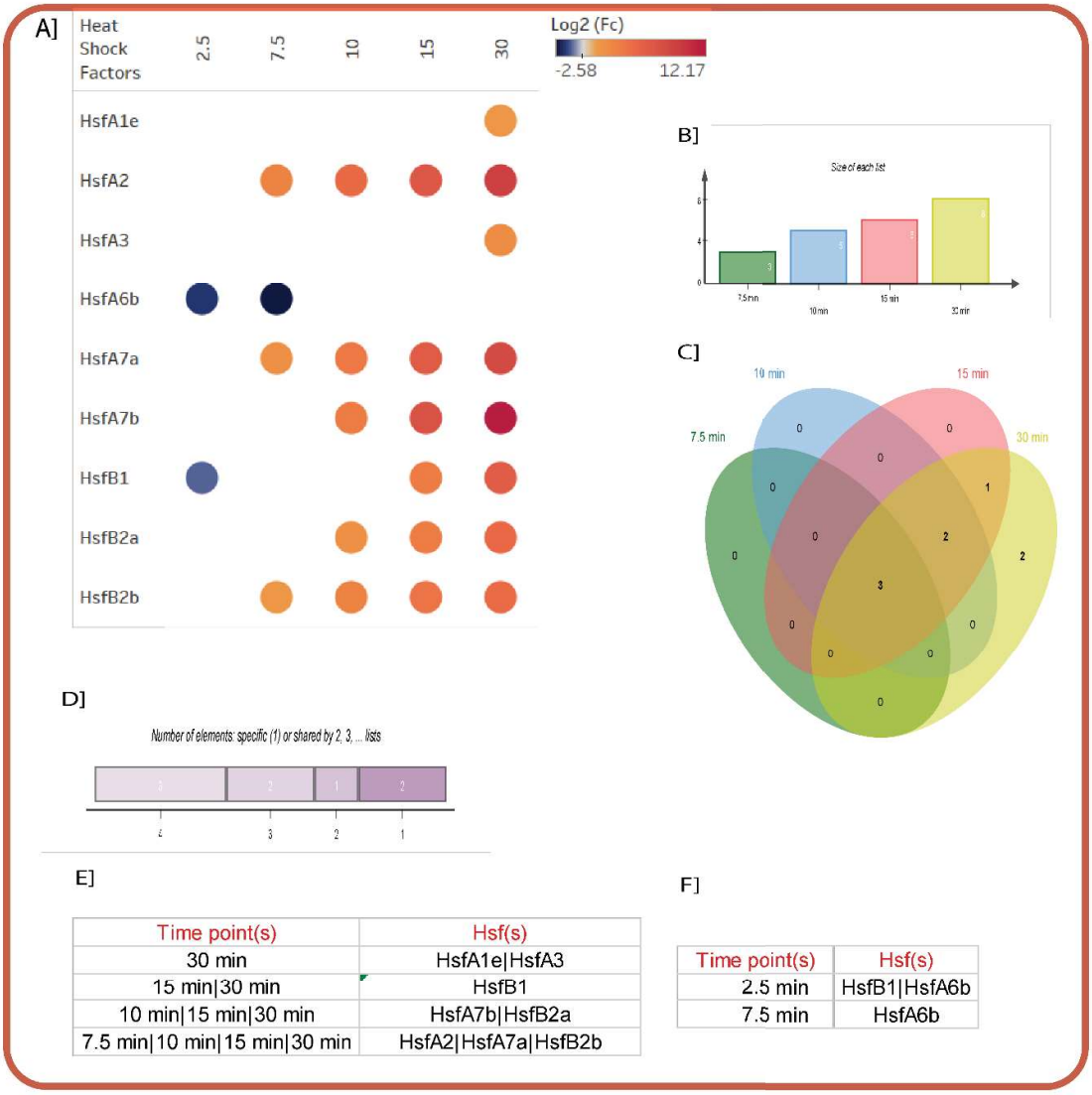
Fig.3.1.23 Expression profile of heat shock protein 100(s) in Category 3. A] Dot-matrix plot representing expression of heat shock protein 100(s) in vegetative aand reproductive tissues of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 100(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly downregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as an seperator.

3.1.7 Expression profile of HSFs in seedlings subjected to different durations of heat stress (Category 4).

This subsection examines the effects of different duration of heat stress on 7 days old seedlings of *A. thaliana*. Publicly available RNA-seq dataset (Li et al., 2019) was used to find (if any) HSFs in early heat stress responses.

Upregulated and downregulated HSFs are represented as red-orange large circles and blue small circles respectively in the dot matrix plot (Fig. 3.1.24A). Significant upregulation of genes in each seedling category are represented in the form of a bar graph (Fig. 3.1.24 B). Eight HSF genes were found to be differentially upregulated in at least one of the time point. The number of unique and shared HSF genes between different time points are represented in the form of a Venn diagram (Fig. 3.1.24C]) and their proportion of genes in the either specific or shared list in each category is represented in the form of a horizontal bar (Fig. 3.1.24D). The sum of numbers mentioned on the ellipse or numbers on the horizontal bar conveys the total no. of transcripts in various combinations and permutations. Expression of these differentially expressed genes are summarised in form of a table (Fig. 3.1.24E and F).

Upon heat stress, no differential expression was observed post 2.5 and 5 min induction. HsfA2, HsfA7a and HsfB2b were upregulated at all-time points except for above mentioned time points. Number of upregulated genes increased with time and remained upregulated for consecutive time points. After 7.5 min of induction HsfA2, HsfA7a and HsfB2a were found to be upregulated. After 10 min of heat stress, HsfA7b and HsfB2a were upregulated. Specific upregulation of HsfB1was observed after 15 min under heat stress. After 30 min of heat stress, upregulation of HsfA1e and HsfA3 was observed. After 2.5 min and 5 min induction, exclusive downregulation of HsfA6b was observed. Downregulation of HsfB1 was unique after 2.5 min induction. After these two time points, downregulation of genes was not observed.



E]

Time point(s)	Hsf(s)
30 min	HsfA1e|HsfA3
15 min|30 min	HsfB1
10 min|15 min|30 min	HsfA7b|HsfB2a
7.5 min|10 min|15 min|30 min	HsfA2|HsfA7a|HsfB2b

F]

Time point(s)	Hsf(s)
2.5 min	HsfB1|HsfA6b
7.5 min	HsfA6b

Fig.3.1.24 Expression profile of Heat shock Factors in Category 4. A] Dot-matrix plot representing expression of heat shock factors in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock factor genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly down-regulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05).Time points are in minutes."|" is used as a seperator .

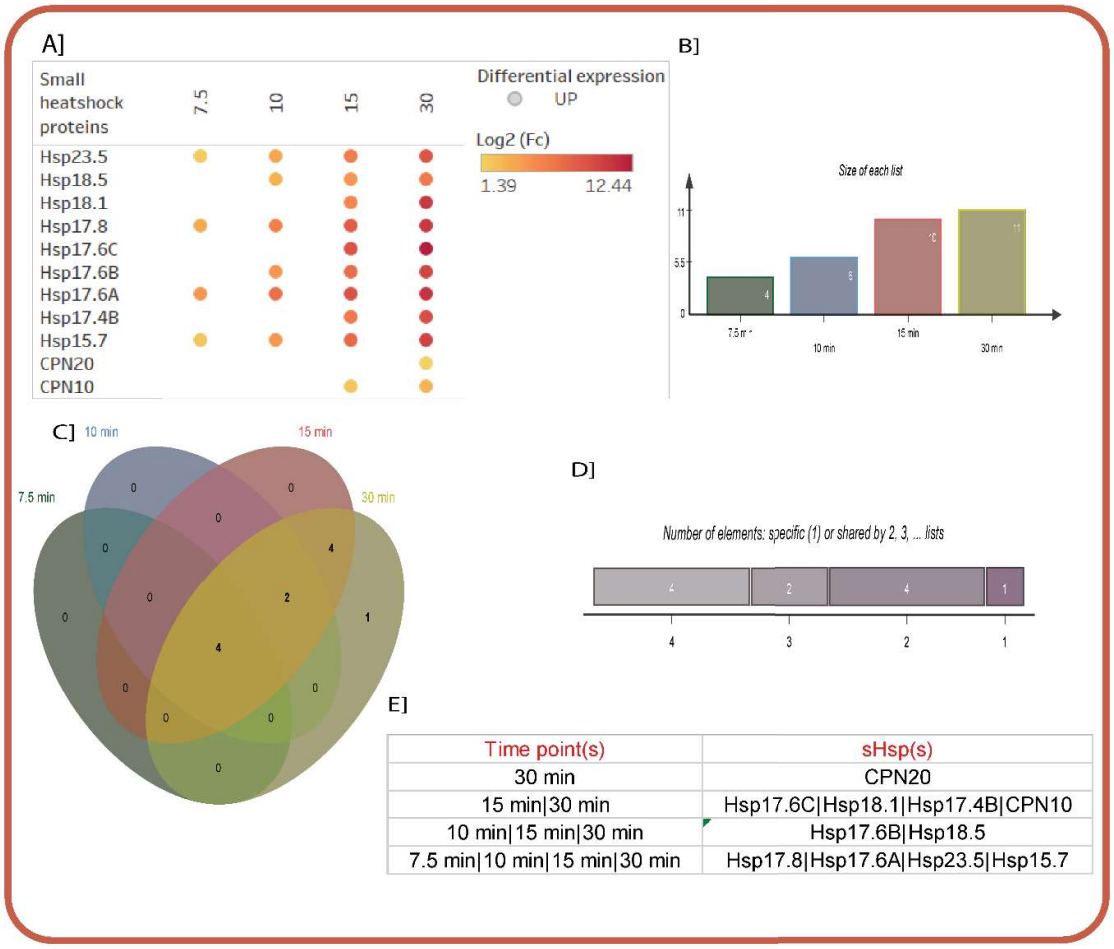
3.1.8 Expression profile of HSPs in seedlings subjected to different durations of heat stress (Category 4).

As mentioned earlier, this subsection proposes to examine the effects of duration of heat stress on 7 days old seedlings of *A. thaliana.*

Upregulated and downregulated HSPs are represented in red-orange large circles and blue small circles respectively in the dot matrix plot [small HSPs (Fig. 3.1.25A) , HSP40s (Fig. 3.1.26A), HSP60s (Fig. 3.1.27A), HSP70s (Fig. 3.1.28A), HSP90s (Fig. 3.1.29A) and HSP100s (Fig. 3.1.30A)]. Significant upregulation of genes at each time point were represented in the form of a bar graph [small HSPs (Fig. 3.1.25B) , HSP40s (Fig. 3.1.26 B), HSP60s (Fig. 3.1.27B), HSP70s (Fig. 3.1.28B), HSP90s (Fig. 3.1.29B) and HSP100s (Fig. 3.1.30B)]. A total of ten HSP40 genes, five HSP90 genes, six HSP100 genes, four HSP60 genes, fourteen HSP70 genes and eleven of sHSPs genes are found to be differentially UP in at least one of the time point. Number of specific (unique) and shared genes between different time points are represented in the form of a Venn diagram [sHSPs (Fig. 3.1.25C), HSP40s (Fig. 3.1.26C), HSP60s (Fig. 3.1.27C), HSP70s (Fig. 3.1.28C), HSP90s (Fig. 3.1.29C) and HSP100s (Fig. 3.1.30C)] and the distribution of genes with respect to the shared elements as well as the specifically expressed genes are represented in the form of a horizontal bar [sHSPs (Fig. 3.1.25D), HSP40s (Fig. 3.1.26D), HSP60s (Fig. 3.1.27D), HSP70s (Fig. 3.1.28D), HSP90s (Fig. 3.1.29D) and HSP100s (Fig. 3.1.30 D)]. Sum of numbers mentioned on the ellipse or numbers on the horizontal bar conveys the total no. of transcripts in various combinations and permutations. Expression of these differentially expressed genes are summarised in form of a table [sHSPs (Fig. 3.1.25E), HSP40s (Fig. 3.1.26E and F), HSP60s (Fig. 3.1.27E), HSP70s (Fig. 3.1.28E), HSP90s (Fig. 3.1.29E) and HSP100s (Fig. 3.1.30E and F)].

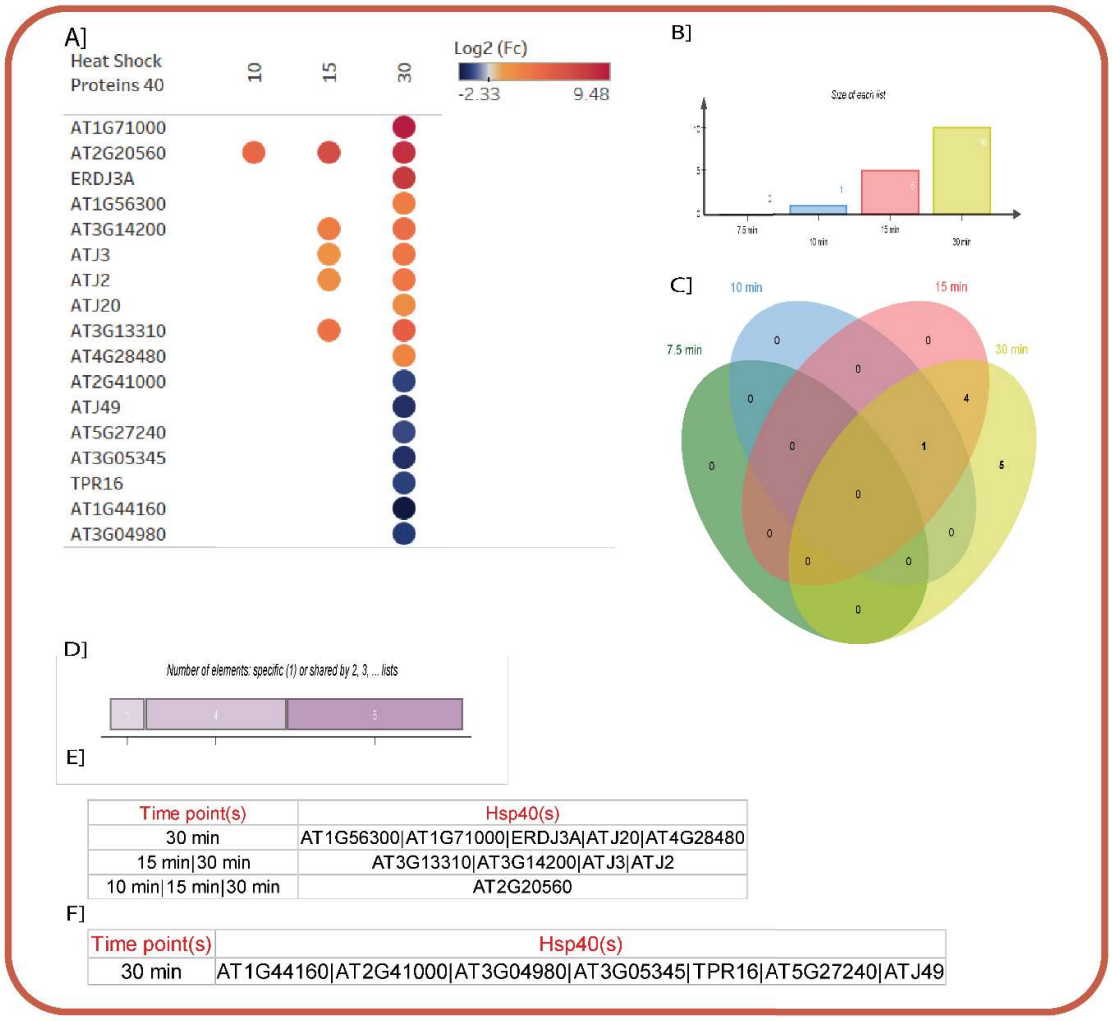
Upon heat stress, no differential expression was observed after 2.5 and 5 min induction. Number of upregulated genes increased with time and remained upregulated for consecutive time points. Only sHSPs genes (Hsp17.8, Hsp17.6A, Hsp23.5 and Hsp15.7) were upregulated just after 7.5 min of heat stress at 37°C. 10 min heat stress response showed specific upregulation of sHSPs (Hsp17.6B and Hsp18.5), HSP40 (AT2G20560), HSP70 (HSP70-8 and MED37C) and HSP 100 (CLPB1). After 15 min of stress, specifically upregulated genes were HSP100s (CLPB4 and CLPB3), HSP90s (HSP90-3 and HSP81-2), HSP70s (HSP70-5, HSP70-3, AT1G79920, MED37D and HSP70-10), HSP60s (CPN60B2 and CPN60), HSP40s (AT3G13310, AT3G14200, ATJ3 and ATJ2) and sHSPs (Hsp17.6C, Hsp18.1, Hsp17.4B and CPN10). sHSP (CPN20) , HSP40s (AT1G56300, AT1G71000, ERDJ3A, ATJ20 and AT4G28480), HSP60s (CPN60A1 and HSP60-3A), HSP70s (HSP70-

16, HSP70-6, MED37E, MED37A, BIP and HSP70-7), HSP90s (HSP90-5, HSP90-6 and HSP90-1) and HSP100s (AT1G09130, CLPR2 and CLPT1). Some of the HSP100 genes such as SMXL7 and HSP40s (AT1G44160, AT2G41000, AT3G04980, AT3G05345, TPR15, AT5G27240 and ATJ49) were observed to be downregulated after 30 min of stress.



Time point(s)	sHsp(s)
30 min	CPN20
15 min|30 min	Hsp17.6C|Hsp18.1|Hsp17.4B|CPN10
10 min|15 min|30 min	Hsp17.6B|Hsp18.5
7.5 min|10 min|15 min|30 min	Hsp17.8|Hsp17.6A|Hsp23.5|Hsp15.7

Fig.3.1.25 Expression profile of small heat shock proteins in Category 4. A] Dot-matrix plot representing expression of small heat shock proteins in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated small heat shock proteins genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05).Time points are in minutes."|" is used as a seperator
* No significant downregulation of small heat shock proteins genes was observed.

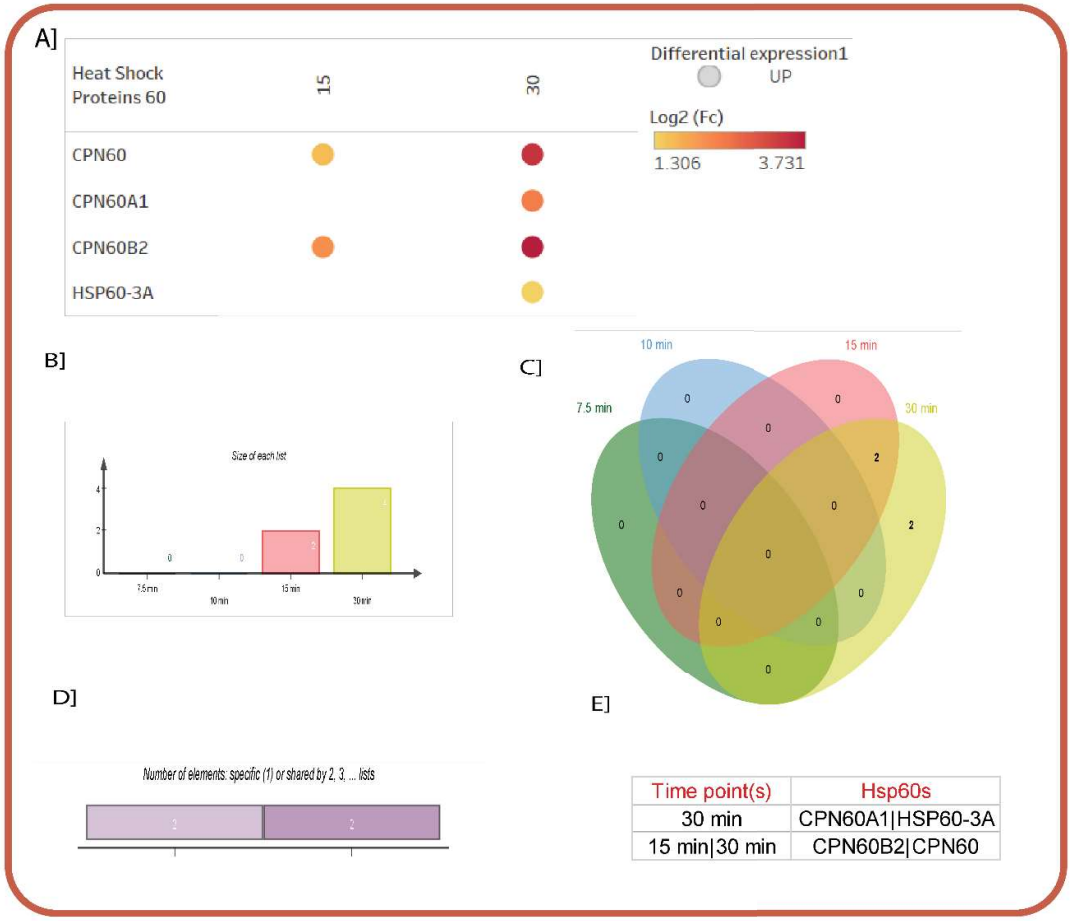


Time point(s)	Hsp40(s)
30 min	AT1G56300|AT1G71000|ERDJ3A|ATJ20|AT4G28480
15 min|30 min	AT3G13310|AT3G14200|ATJ3|ATJ2
10 min|15 min|30 min	AT2G20560

F]

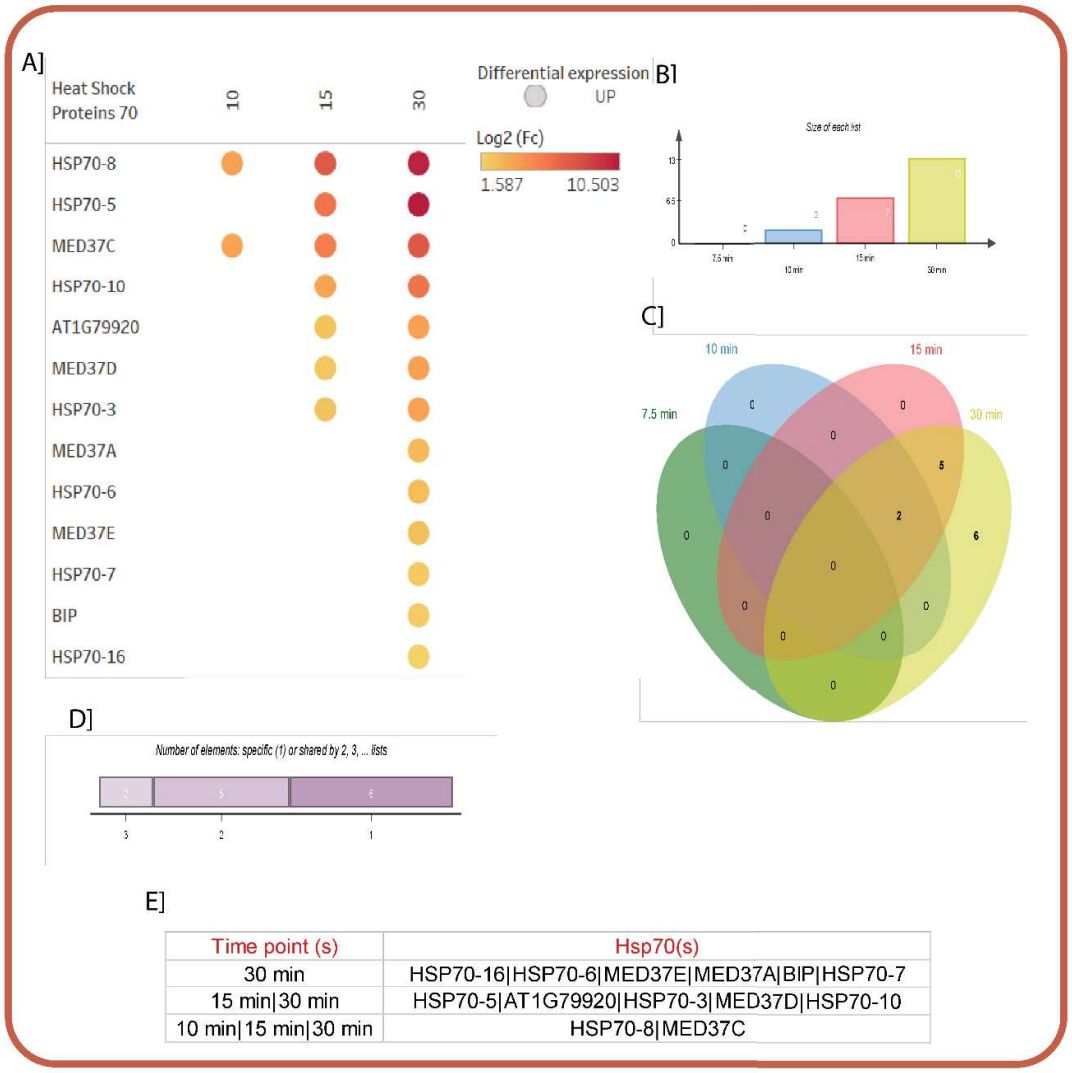
Time point(s)	Hsp40(s)
30 min	AT1G44160|AT2G41000|AT3G04980|AT3G05345|TPR16|AT5G27240|ATJ49

Fig.3.1.26 Expression profile of heat shock proteins 40 in Category 4. A] Dot-matrix plot representing expression of heat shock proteins 40 in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock proteins 40 genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. F] Table of genes significantly downregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05)."|" is used as a seperator .



Time point(s)	Hsp60s
30 min	CPN60A1|HSP60-3A
15 min|30 min	CPN60B2|CPN60

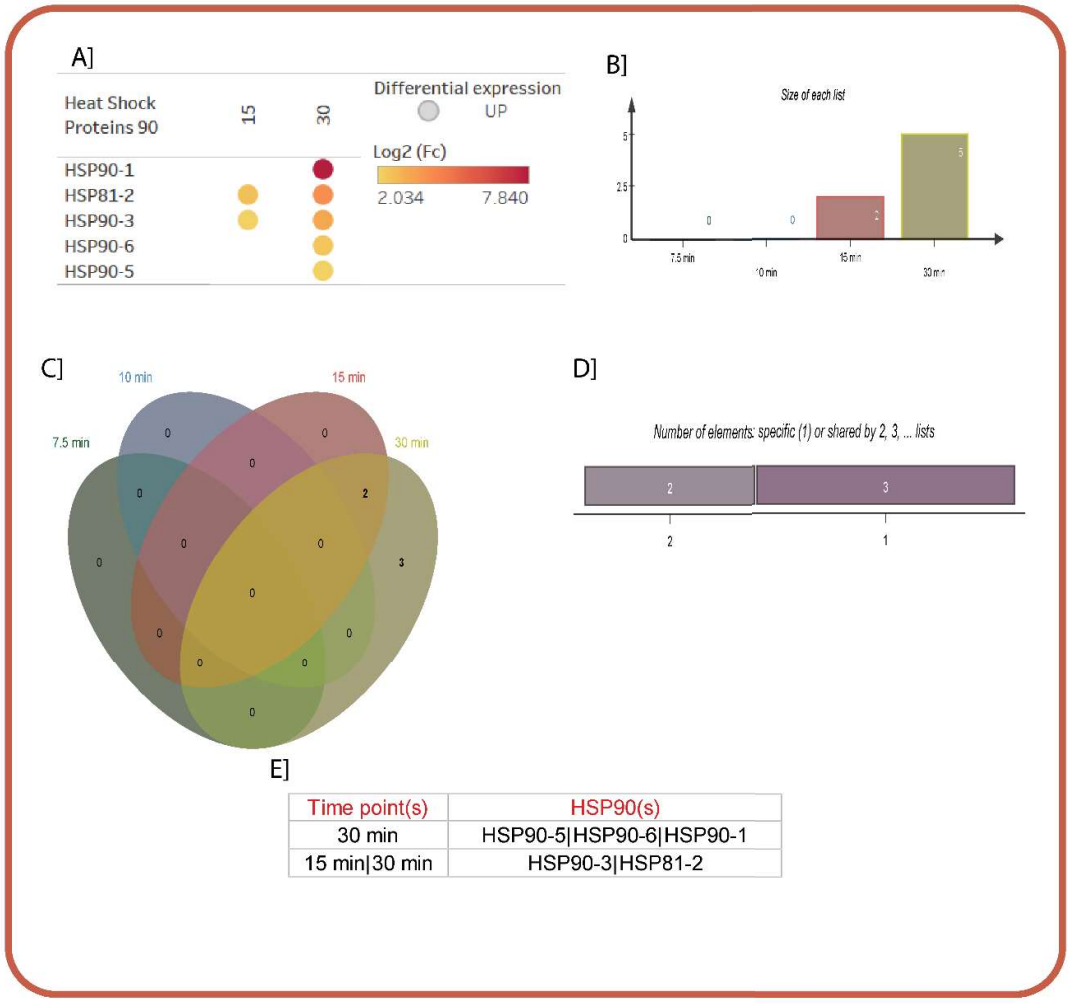
Fig.3.1.27. Expression profile of heat shock protein 60(s) in Category 4. A] Dot-matrix plot representing expression of heat shock protein 60(s) in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 60(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://-jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05)."|" is used as a seperator.
*No significant downregulation was observed in this category.



E]

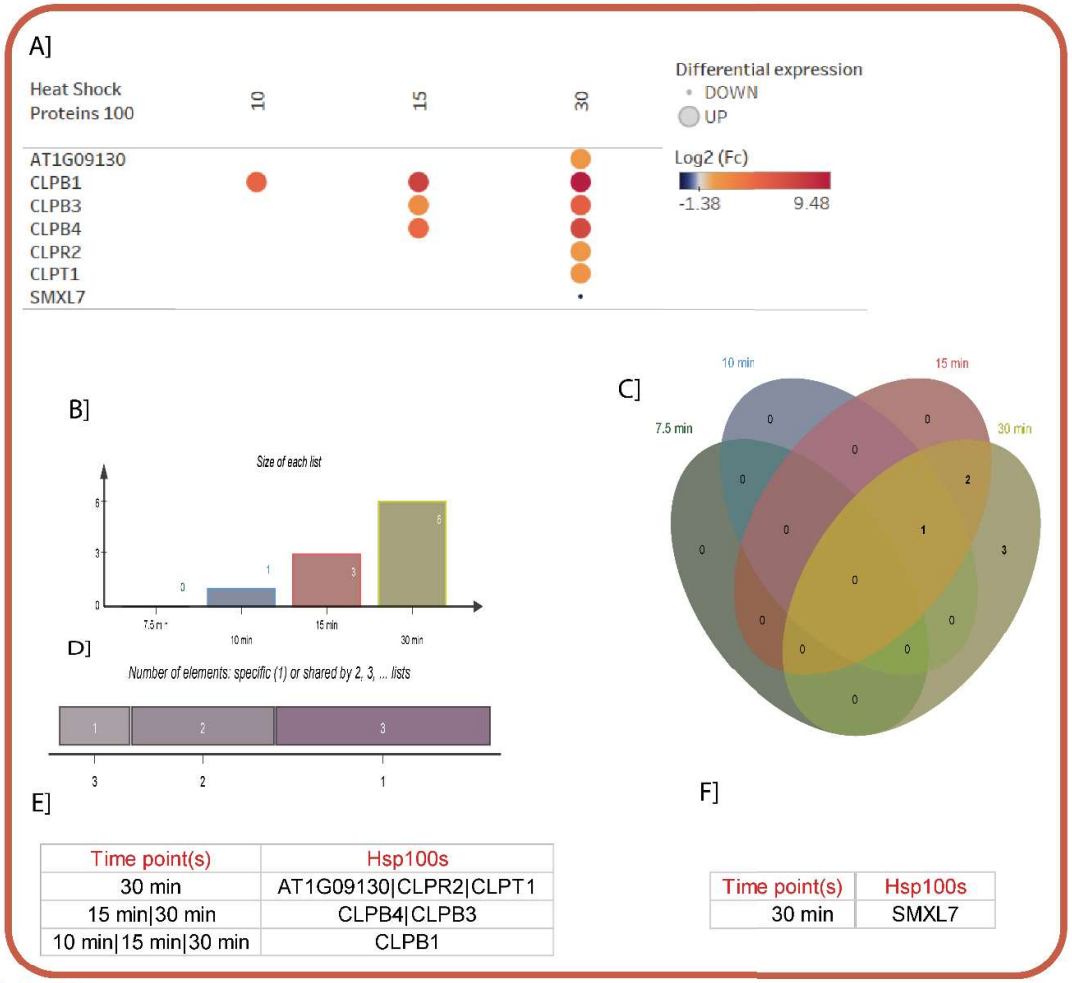
Time point (s)	Hsp70(s)
30 min	HSP70-16|HSP70-6|MED37E|MED37A|BIP|HSP70-7
15 min|30 min	HSP70-5|AT1G79920|HSP70-3|MED37D|HSP70-10
10 min|15 min|30 min	HSP70-8|MED37C

Fig.3.1.28 Expression profile of heat shock protein 70(s) in Category 4. A] Dot-matrix plot representing expression of heat shock protein 70(s) in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 70(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator .
*No significant downregulation was observed in this category.



Time point(s)	HSP90(s)
30 min	HSP90-5|HSP90-6|HSP90-1
15 min|30 min	HSP90-3|HSP81-2

Fig.3.1.29 Expression profile of heat shock protein 90(s) in Category 4. A] Dot-matrix plot representing expression of heat shock protein 90(s) in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 90(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator .
*No significant downregulation was observed in this category.



Time point(s)	Hsp100s
30 min	AT1G09130|CLPR2|CLPT1
15 min|30 min	CLPB4|CLPB3
10 min|15 min|30 min	CLPB1

Time point(s)	Hsp100s
30 min	SMXL7

Fig.3.1.30 Expression profile of heat shock protein 100(s) in Category 4. A] Dot-matrix plot representing expression of heat shock protein 100(s) in Seedlings of *Arabidopsis thaliana* under different stress regime. This graph was made using Tableau (https://www.tableau.com/) desktop version.B] Bar graph for no. of significantly upregulated genes in each list of this category.C] Venn diagram for significantly upregulated heat shock protein 100(s) genes.D] No. of elements in each combination[Labeled as 1(Specific) or 2,3,... (No. of categories shared by the list element)].B to D are generated using J venn tool online (http://jvenn.toulouse.inra.fr/app/example.html, Bardou, P., et al (2014).E] Table of genes significantly upregulated in this category.F] Table of genes significantly downregulated in this category. N.s. stands for Non significant, UP stands for Significant Upregulation and DOWN stands for Significant Downregulation at the selected Criteria ($-1 \geq Log_2(\text{Fold Change}) \geq 1$, Adjusted p-value ≤ 0.05). "|" is used as a seperator .

Objective 2 Analysis of HSP101 in seed germination under cadmium and arsenic stress in *Arabidopsis thaliana*.

Another objective of this dissertation is to analyze the role of HSP101 in *Col*-0 and *hot 1-3* (null for *At*Hsp101) during heavy metal stress (MS). Proteins are the functional workhorses of an organism. It is known in the literature that heavy metals interact and interfere with the natively folded proteins. Plants use various strategies to combat MS including the production of HSPs. As mentioned earlier, the interaction of heavy metals with proteins (displacement of important ions in metalloenzymes, oxidation of amino acid residues and interaction with functional groups) results in an imbalance in maintaining proteostasis (Hasan et al., 2017; Tamás et al., 2014). This may suggest the induction of HSPs as a safeguard mechanism for protecting cells against proteotoxic stress. Induction of HSP101 under cadmium and arsenite stress at the seed stage has been established previously (Ratnesh Chandra Mishra, Ph.D. thesis, 2014, DPMB, UDSC and Nidhi Yadav, M.Sc. thesis, 2019, DPMB, UDSC).

The susceptibility of *hot1-3* to heat stress (HS) is already known (Hong and Vierling, 2001). So, before analyzing the response of col-0, HSP101 overexpression lines (OE) and *hot1-3* to MS we analyzed its HS response.

3.2.1. Phenotypic analysis of col-0 and *hot1-3* mutant under HS

HS response of col-0*,* AtHsp101 overexpression line (hereby referred to as OE) and *hot1-3* was analyzed via subjecting five-day-old *A. thaliana* Col-0*,* OE and *hot1-3* seedlings to HS at 42°C for 60 minutes. Plants were then allowed to recover for 4 and 8 days at 22°C, 16/8 hours day/night conditions. Excellent recovery was observed in the case of OE seedlings. Col-0 also recovered around 8 days in recovery whereas *hot1-3* died completely. (Fig. 3.2.1).

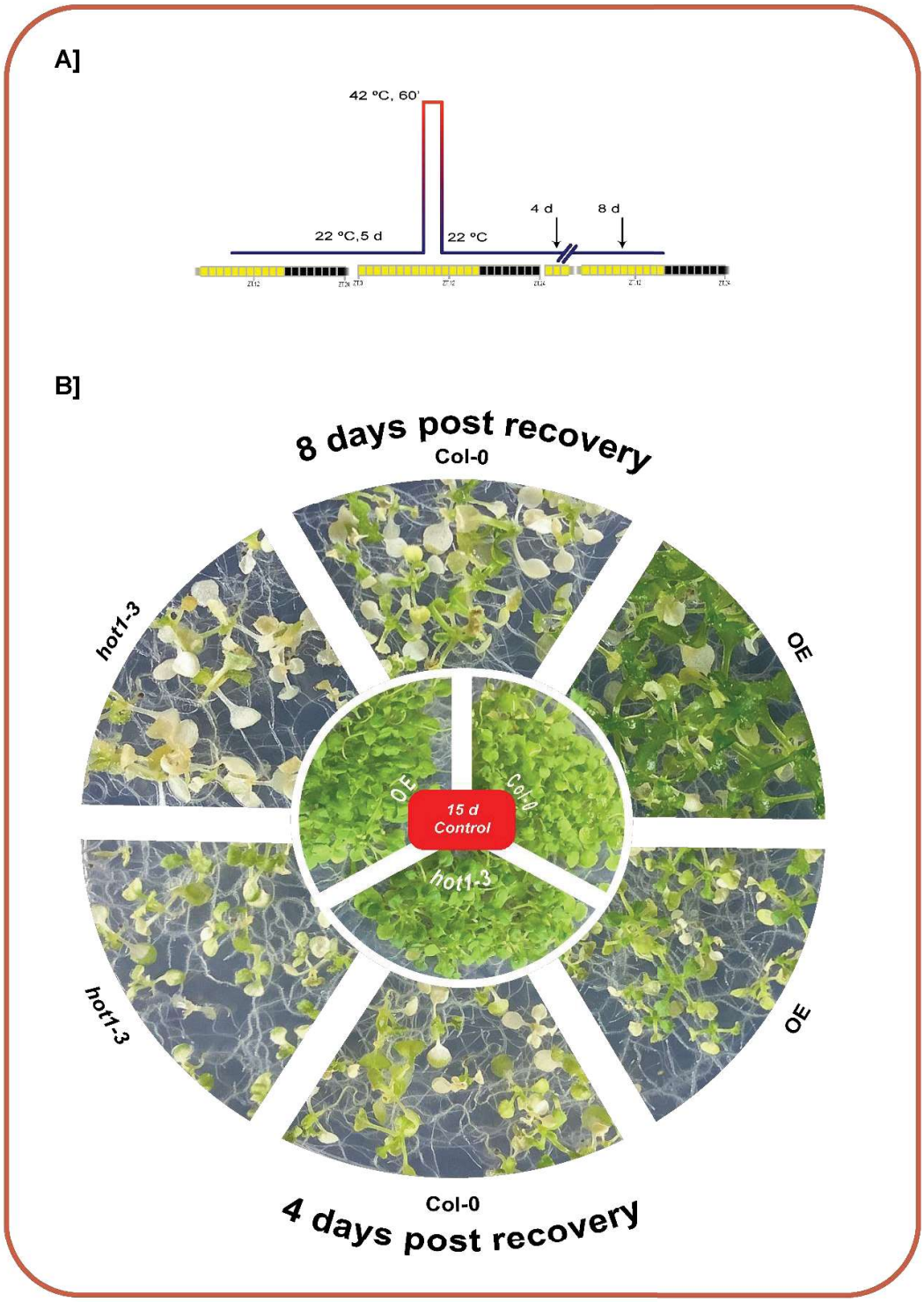


Fig.3.2.1 Heat stressed phenotype of A. thaliana Col-0, OE and *hot1-3* mutant plants.A]Followed stress regime. B] Basal thermotolerance of Col-0,OE and *hot1-3* subjected to lethal heat stress (42°C) for 60 minutes, followed by 4 and 8 days in recovery.

3.2.2. Phenotypic analysis of Col-0 and *hot1-3* mutant under metal stress

Heavy metal stress response was examined via exposing seeds of *A. thaliana* Col-0 and hot1-3 mutant plants with ½ MS nutrient media appended with different concentrations of cadmium and arsenite for 10 days.

Cadmium stress response

For elucidation of the hypothesis that 'whether AtHsp101 is a member of clan against metal stress', seeds of col-0 and *hot1-3* were plated on ½ MS nutrient media supplemented with heavy metal (Basal seed stress). Cadmium was added to the nutrient media in varying concentrations (50 µM, 100 µM and 200 µM). No difference in the growth profile of Col-0 and null mutant was observed under control conditions. However, under cadmium stress, *hot1-3* performs distinctively better than col-0. Quantification of root length and fresh weight post 10 days of basal seed stress also showed significant differences between col-0 and hot1-3 at most of the concentrations. *hot1-3* performed better at all the three cadmium concentrations (Fig. 3.2.2).

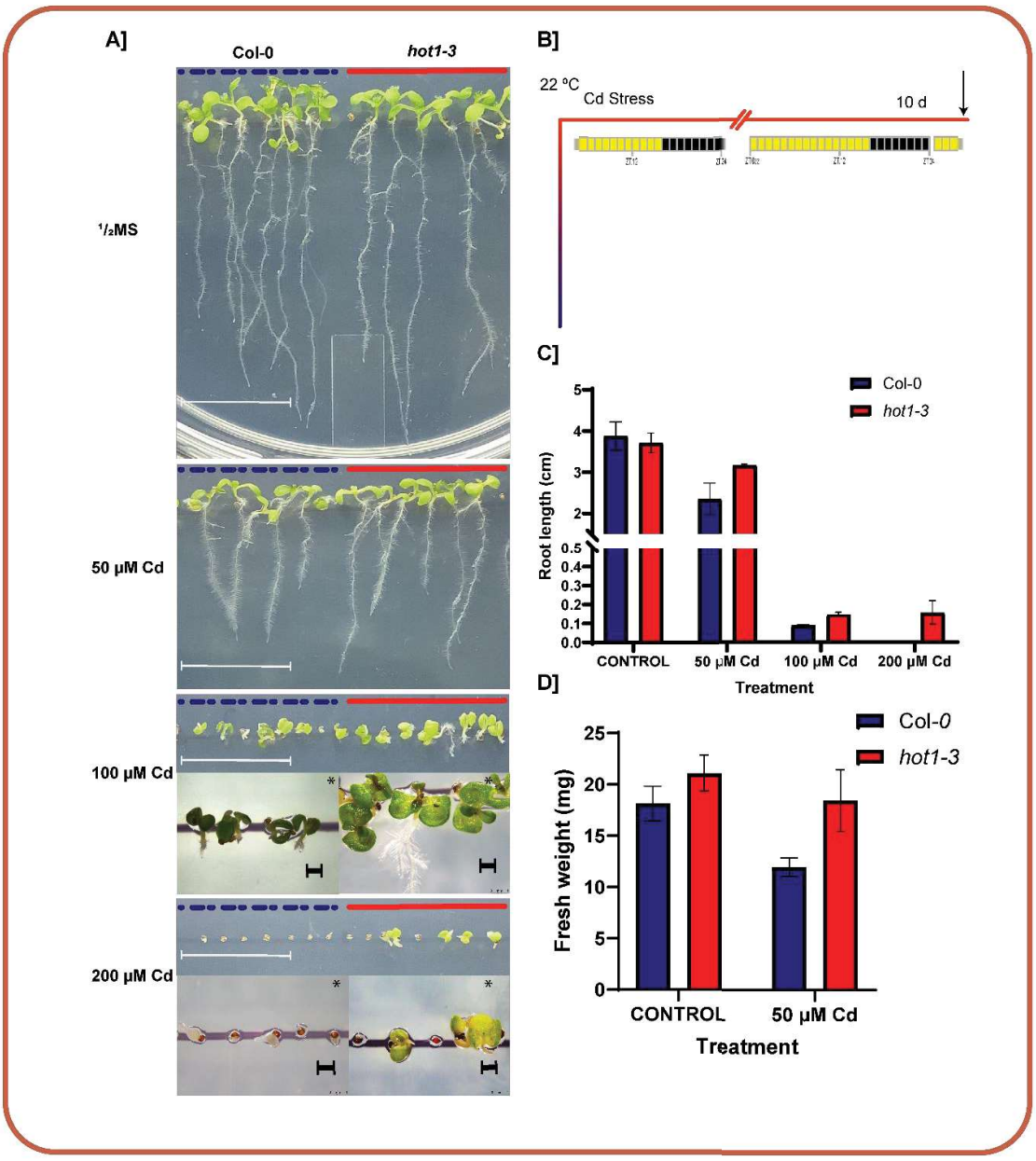


Fig.3.2.2. Cadmium stress response of 10 days old *A. thaliana* Col-0 and *hot1-3* mutant plants.A] Basal seed stress phenotype due to exposure of varying concentration of cadmium for 10 days,White scale bar (= 1.5 cm) and black scale bar (= 1 mm).B] Followed stress regime. Graphical representation of Root length (C]) in cm and fresh weight (D]) in mg, error bars represent means with standard error of means and graphs were plotted using Graphpad version 9.1.2. Root length was taken with help of ImageJ software (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.nih.gov/ij/, 1997-2018).
* magnified 1x

Arsenite stress response

Similarly, seeds of col-0 and null mutant (*hot1-3*) were plated on ½ MS nutrient media supplemented with heavy metal (Basal seed stress). Arsenite was added to the nutrient media in varying concentrations (20 µM, 50 µM and 100 µM). No difference in the growth profile of wildtype and null mutant was observed under control conditions. Whereas under stressed conditions, *hot1-3* performs evasively better than col-0. Quantification of root length and fresh weight post 10 days of stress showed an evasive difference between col-0 and *hot1-3* at most of the concentrations (Fig. 3.3.3).

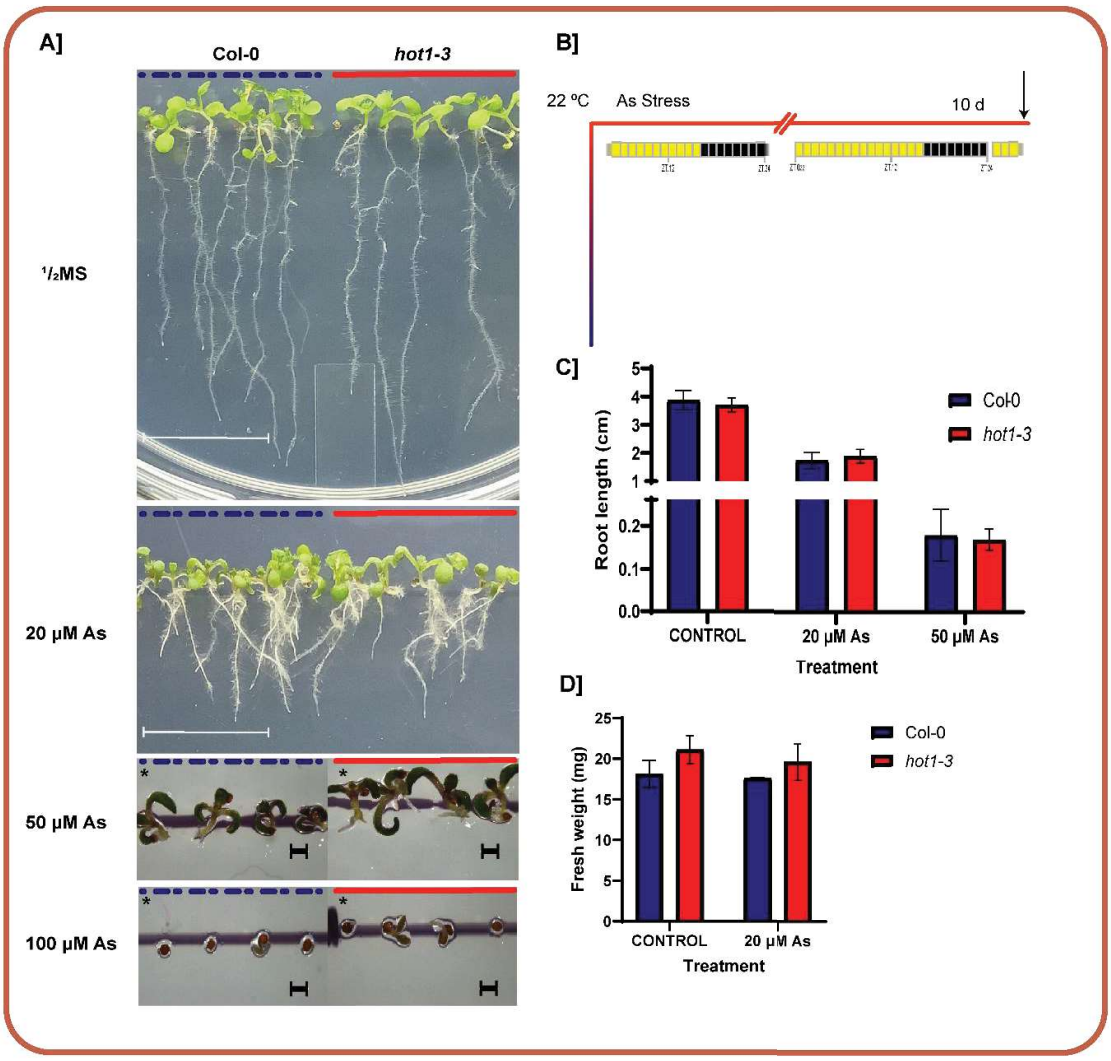


Fig.3.2.3. Arsenite stress response of 10 days old *A. thaliana* Col-0 and *hot1-3* mutant plants.A] Basal seed stress phenotype due to exposure of varying concentration of arsenite for 10 days,White scale bar (= 1.5 cm) and black scale bar (= 1 mm).B] Followed stress regime. Graphical representation of Root length (C]) in cm and fresh weight (D]) in mg, error bars represent means with standard error of means and graphs were plotted using Graphpad version 9.1.2. Root length was taken with help of ImageJ software (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.nih.gov/ij/, 1997-2018).
* magnified 1x

3.3 Discussion-

High surface area to volume ratio imposes simultaneous perception of HS to all the biomacromolecules of the cells such as protein supramolecular complexes. (Mittler et al., 2012). Proteins are the major macromolecules to be affected by both heat stress and heavy metals. These two stressors interfere with native proteins and negatively effects their folding resulting in proteotoxic aggregation. To maintain the homeostasis in cell physiome, expression profile is modulated and genes such as HSPs and HSFs are induced (Hasan et al., 2017). HSPs assist in the folding of misfolded proteins which otherwise can be toxic to cell via aggregation. Numerous classes of HSPs has been analysed and categorised based on molecular weights.

Here we have analyzed comparative transcriptome changes in different tissues, the effect of time of day and duration of heat stress in *Arabidopsis thaliana*. Analysis of this transcriptomic study showed that there are few HSFs and HSPs playing major role under HS. For instance, HsfA2 showed upregulation in all tissues under every stress regime studied. This observation is in conjunction with the profile of HSP101 which showed upregulation in all tissues examined. No upregulation of HSP101 was observed in mature seeds and this might be due to the already high absolute expression in mature seeds. Plastidial and mitochondrial form of HSP101 showed no differential expression in any of the seed stages examined. Interestingly, the plastidial form CLPB3 showed no upregulation in seedlings given heat stress after dusk.

Li et al., 2019 provided an extensive data showing the effect of brief HS in *Arabidopsis thaliana.* In this study downregulation of HsfA6b and HsfB1 was observed just after 2.5 minutes of HS followed by upregulation of HsfB1 after 15 minutes of exposure. This pattern of expression correlates with the expression of HSP101/CLPB1 which induces after 10 mins of exposure. It is also reported that, HSP101 transcript upregulates in accordance with the expression of HsfA2 (Liu and Charng, 2013). HsfB1 negatively regulates the expression of HSP101 via HsfA1s expression modulation and its amplifier HsfA2 (Induction of this gene was observed after 7.5 min exposure). After longer exposure to HS HsfB1 forms a negative feedback loop with HsfA1. Plastidial and mitochondrial form of HSP101 upregulates after 15 mins of treatment. Various differential expressing genes belonging to different classes of HSPs has been observed and it requires validation via RT-PCR. It is quite possible that some of the differences in the expression pattern observed are because of the different stages of the plant and variable stress regimes that were used.

Another objective of this study was to analyse the role of HSP101 in MS. Heat sensitivity of *hot1-3* is well established in the literature (Hong and Vierling, 2001). Here we analyzed the As and Cd metal stress response of Col-0 and *hot1-3* mutant seeds of A. thaliana by subjecting

them to varying concentrations of these heavy metals. Surprisingly, under stressful conditions, *hot1-3* showed better growth and longer roots. But these two genotypes show the same phenotype under control conditions. No difference in root and hypocotyl length has been observed under control conditions in Col-0 and *hot1-3* (Lalit Dev Tiwari, Ph.D. thesis 2018, DPMB, UDSC).

Based on the limited time in the laboratory due to precautionary measures taken amidst the Covid-19 pandemic further experiments that were planned could not be completed. However, literature survey guided us in understanding the possible explanations for the phenotype observed.

Heavy metals like Cd and As disrupts root architecture in *Arabidopsis* via NO signalling (Piacentini et al., 2020). During unfavourable conditions, plants allocate their resources towards stress response (upregulation of stress responsive genes) thereby slowing down growth and developmental processes. HSPs are induced as a stress response in multiple stresses. The promoter of AtHSP101 harbours metal responsive elements such as AP1 (animal proto-oncogene1) and stress-responsive elements (STREs) and shows metal (As and Cd) inducibility (Mishra and Grover, 2014). Induction of AtHSP101 during metal stress points that apart from heat stress it also plays a role during MS.

The contrasting phenotype of Col-0 vs *hot1-3* under MS draw parallel with the role of *Sc*HSP104 (homologue of HSP101 in yeast) in the propagation of [*PSI*$^{+}$] prion. In the absence or low level of this chaperone, partial oligomerization of prion proteins (capable of nucleation for fibrillization) results in functional propagating prion whereas in overexpressing lines complete disassembly takes place (Chernoff et al., 1995; López-Frías et al., 2011).

In conjunction with these results and also nodal root inhibition in maize (López-Frías et al., 2011), we may propose that HSP101 by virtue of its disaggregase activity might function in the resurrection of primary root growth inhibitor which then negatively impacts the primary root growth. In *hot1-3* mutant lines, due to the absence of this disaggregase protein, resurrection of inhibitor protein does not take places, hence longer roots under stress.

A paradoxical result was observed in our case, as the overexpression line showed comparable results with the wild type (not shown here). In the light of the above results, we can say that the presence of this disaggregase though in higher amount results in the resurrection of inhibitor protein hence, the phenotype.

Another reasonable explanation for this phenotype is the regulation of root length involving co-chaperones of HSP101 in the disaggregation of aggregated proteins. This hypothesis revolves around the fact that the absence of HSP101 may result in impairment of

disaggregation machinery, thus all cellular proteins including the root growth inhibitor are not able to properly function, hence the longer roots. In wildtype plants, HSP101 functions as in principle resulting in functional primary root growth inhibitor which represses the root length under stress.

Another possibility relates to the role of HSP101 as a translational regulator (Wells et al., 1998). Translation of this primary root growth inhibitor might be HSP101 dependent and hence longer roots are observed only in case of *hot1-3*.

Under natural conditions, it has been observed in maize that the null mutant for HSP101 shows longer primary roots (Nieto-Sotelo et al., 2002) In maize, it was reported that in HSP101 null mutant, tissue-specific differential proteomic expression of auxin and cytokinin metabolism was there which correlates with primary root growth. Relation between auxin levels and HSP101 induction in roots of maize has been reported (Martínez-de la Cruz et al., 2015). This knockout mutant also shows an enhanced actin level (Lázaro-Mixteco et al., 2012) and the important role of actin in mediating root growth and germination has been reported in *Arabidopsis* (Du et al., 2020; Rahman et al., 2007). Although most of these studies are based on maize but can be extrapolated for *Arabidopsis.* Further investigation is required to unravel the mechanism of the relationship between root growth and actions of HSP101 under heavy metal stress.

In accordance with these reports, we may propose the impairment of the sensing mechanism for these metal(loid)s in *hot1-3.* This aberration in sensing may affect the reallocation of resources towards stress response and growth continues hence, longer roots.

At this juncture, we can only propose that *At*HSP101 has a role in As and Cd metal stress response. Basal seed expression of HSP101 in Col-0 seeds and its absence in *hot1-3* might be responsible for the difference in seed germination and root length phenotype. Due to paucity of time, further experiments could not be performed and results thus obtained should be repeated for further validation.

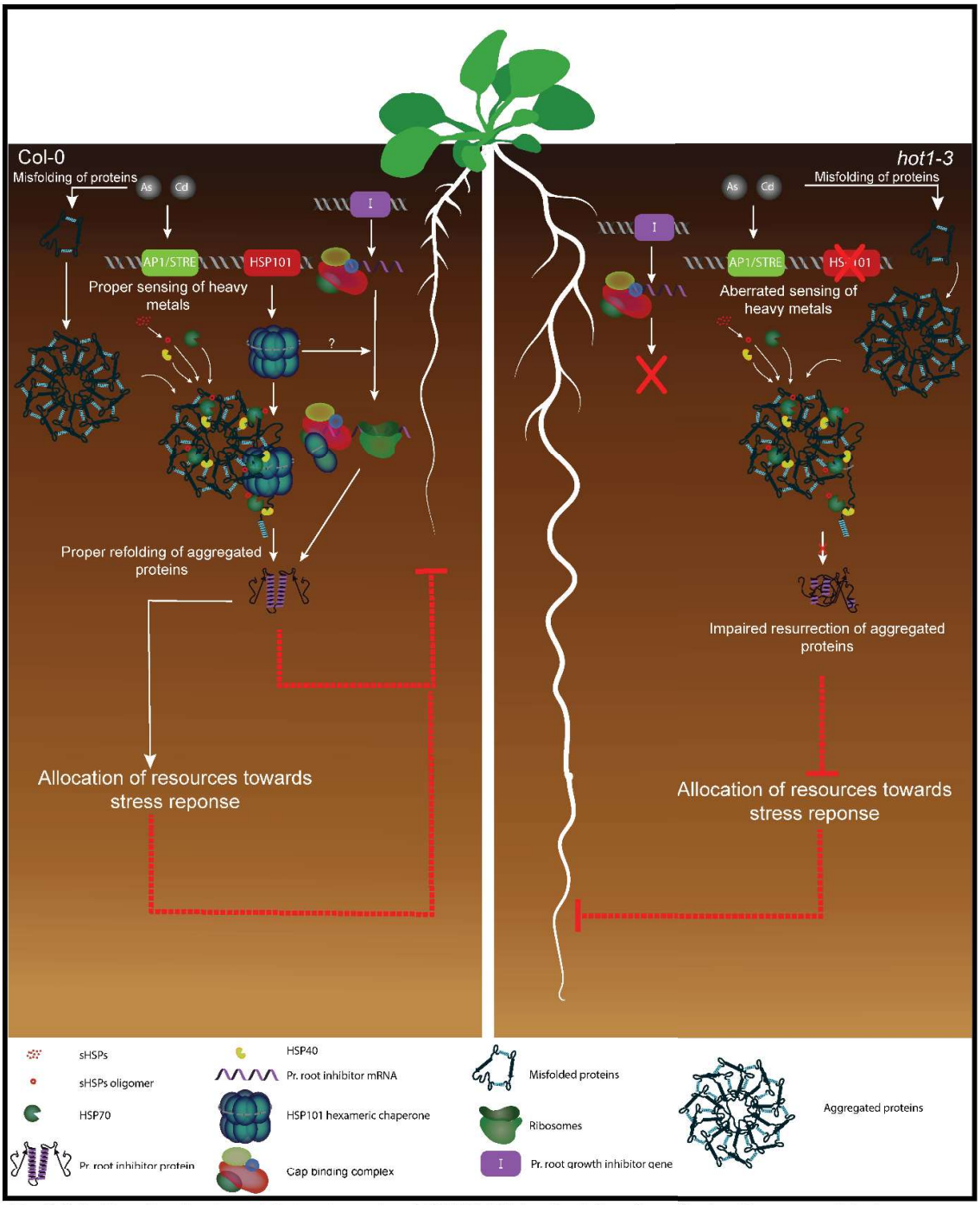


Fig.3.3.1. Hypothesized model showing role of *At*HSP101 in alleviating the effects of heavy metal stress.

Conclusions and future work



Due to unsustainable lifestyle, industrialization, urbanization, mismanagement of agricultural additives and improper waste disposal system has led to environmental problems such as pollution of arable soil and aquifers; and global warming etc. As a repercussion of this, plants has to fight daily for its survival with the variable environmental factors (as in this study, heavy metals and heat stress). Both these stresses induce proteotoxicity in plants resulting in stunted growth and retarded development. As an effect of these stresses, protein structure and folding is disrupted and results in misfolded proteins which with their exposed hydrophobic domains may lead to the formation of toxic aggregates. HSPs are induced as a result of these stresses.

In this dissertation work, we provide the tissue-specific and spatio-temporal expression profile of HSPs and HSFs genes. Different classes of HSPs and HSFs were found to be differentially expressed in a specific or constitutive manner. These expression profiles may help in the discovery of novel factors by which these stress-induced genes can mediate stress tolerance. Results obtained by this analysis can be validated via RT-PCR. These stress tolerance mechanisms can be exploited for the development of climate-resilient crops via transgenic approach.

Another theme that is described in this thesis is regarding the effects shown by cadmium and arsenite on *Arabidopsis thaliana*. Previously, induction of *At*HSP101 was reported on exposure to these ions. The growth of *hot1-3* (null for HSP101) is better under stressed conditions as compared with wild-type plants. No difference was observed under unstressed conditions. In the essence of this observation, we suggest the role of AtHSP101 in root development under cadmium and arsenite stress.

In the future course of work, there is a need to validate the observed phenotype of *hot1-3* by comparing with the complemented line of this gene as *hot1-3* mutant being a T-DNA insertion line. Expression profiling at different time points will provide insights into HSP101 at various stages of development under basal seed stress. Similar profiling can be done for the action of this gene during recovery periods. Phenotyping and expression analysis can also be performed by subjecting seedlings to basal metal stress.By understanding the role of this disaggregase protein in the heavy metal stress responses and the insights gained from this dissertation can be exploited to develop tolerant crops with enhanced phytoremediation capacity to treat contaminated soils.

References



Al-Whaibi, M.H. (2011). Plant heat-shock proteins: A mini review. Journal of King Saud University - Science **23**: 139–150.

Ashfaque, F., Farooq, S., Chopra, P., Chhillar, H., and R, K.M.I. (2020). Improving Heavy Metal Tolerance through Plant Growth Regulators and Osmoprotectants in Plants. In Improving Abiotic Stress Tolerance in Plants (CRC Press).

Avelange-Macherel, M.-H., Payet, N., Lalanne, D., Neveu, M., Tolleter, D., Burstin, J., and Macherel, D. (2015). Variability within a pea core collection of LEAM and HSP22, two mitochondrial seed proteins involved in stress tolerance. Plant Cell Environ **38**: 1299–1311.

Bardou, P., Mariette, J., Escudié, F., Djemiel, C., and Klopp, C. (2014). jvenn: an interactive Venn diagram viewer. BMC Bioinformatics **15**: 293.

Blair, E.J., Bonnot, T., Hummel, M., Hay, E., Marzolino, J.M., Quijada, I.A., and Nagel, D.H. (2019). Contribution of time of day and the circadian clock to the heat stress responsive transcriptome in Arabidopsis. Sci Rep **9**: 4814.

Buzea, C. and Pacheco, I. (2020). Heavy Metals: Definition, Toxicity, and Uptake in Plants. In Cellular and Molecular Phytotoxicity of Heavy Metals, M. Faisal, Q. Saquib, A.A. Alatar, and A.A. Al-Khedhairy, eds, Nanotechnology in the Life Sciences. (Springer International Publishing: Cham), pp. 1–17.

Caetano, S.M., Menezes, R., Amaral, C., Rodrigues-Pousada, C., and Pimentel, C. (2015). Repression of the Low Affinity Iron Transporter Gene FET4: A NOVEL MECHANISM AGAINST CADMIUM TOXICITY ORCHESTRATED BY YAP1 VIA ROX1 *. Journal of Biological Chemistry **290**: 18584–18595.

Chernoff, Y.O., Lindquist, S.L., Ono, B., Inge-Vechtomov, S.G., and Liebman, S.W. (1995). Role of the chaperone protein Hsp104 in propagation of the yeast prion-like factor [psi+]. Science **268**: 880–884.

Clemens, S. and Ma, J.F. (2016). Toxic Heavy Metal and Metalloid Accumulation in Crop Plants and Foods. Annu Rev Plant Biol **67**: 489–512.

Du, M., Wang, Y., Chen, H., and Han, R. (2020). Actin filaments mediated root growth inhibition by changing their distribution under UV-B and hydrogen peroxide exposure in Arabidopsis. Biological Research **53**: 54.

Duffus, J. (2002). Heavy metals" a meaningless term? (IUPAC Technical Report). Pure and Applied Chemistry **74**: 793–807.

Efeoğlu, B. (2009). Heat shock proteins and heat shock response in plants. Gazi University Journal of Science **22**: 67–75.

Garg, N. and Singla, P. (2011). Arsenic toxicity in crop plants: Physiological effects and tolerance mechanisms. Environmental Chemistry Letters **9**: 303–321.

Grinevich, D.O., Desai, J.S., Stroup, K.P., Duan, J., Slabaugh, E., and Doherty, C.J. (2019). Novel transcriptional responses to heat revealed by turning up the heat at night. Plant Mol Biol **101**: 1–19.

Guo, M., Liu, J.-H., Ma, X., Luo, D.-X., Gong, Z.-H., and Lu, M.-H. (2016). The Plant Heat Stress Transcription Factors (HSFs): Structure, Regulation, and Function in Response to Abiotic Stresses. Front. Plant Sci. **7**.

Hasan, M.K., Cheng, Y., Kanwar, M.K., Chu, X.-Y., Ahammed, G.J., and Qi, Z.-Y. (2017). Responses of Plant Proteins to Heavy Metal Stress-A Review. Front Plant Sci **8**: 1492.

Hayes, S., Schachtschabel, J., Mishkind, M., Munnik, T., and Arisz, S.A. (2020). Hot topic: Thermosensing in plants. Plant Cell Environ.

Hong, S.W. and Vierling, E. (2001). Hsp101 is necessary for heat tolerance but dispensable for development and germination in the absence of stress. Plant J **27**: 25–35.

Hong, S.-W. and Vierling, E. (2000). Mutants of Arabidopsis thaliana defective in the acquisition of tolerance to high temperature stress. PNAS **97**: 4392–4397.

Jalmi, S.K., Bhagat, P.K., Verma, D., Noryang, S., Tayyeba, S., Singh, K., Sharma, D., and Sinha, A.K. (2018). Traversing the Links between Heavy Metal Stress and Plant Signaling. Front Plant Sci **9**: 12.

Joshi, R., Pareek, A., and Singla-Pareek, S.L. (2016). Chapter 9 - Plant Metallothioneins: Classification, Distribution, Function, and Regulation. In Plant Metal Interaction, P. Ahmad, ed (Elsevier), pp. 239–261.

Katiyar-Agarwal, S., Agarwal, M., Gallie, D.R., and Grover, A. (2001). Search for the Cellular Functions of Plant Hsp100/Clp Family Proteins. Critical Reviews in Plant Sciences **20**: 277–295.

Keyster, M., Niekerk, L.-A., Basson, G., Carelse, M., Bakare, O., Ludidi, N., Klein, A., Mekuto, L., and Gokul, A. (2020). Decoding Heavy Metal Stress Signalling in Plants: Towards Improved Food Security and Safety. Plants (Basel) **9**.

von Koskull-Döring, P., Scharf, K.-D., and Nover, L. (2007). The diversity of plant heat stress transcription factors. Trends Plant Sci **12**: 452–457.

Kumar, R., Khungar, L., Shimphrui, R., Tiwari, L.D., Tripathi, G., Sarkar, N.K., Agarwal, S.-K., Agarwal, M., and Grover, A. (2020). AtHsp101 research sets course of action for the genetic improvement of crops against heat stress. J. Plant Biochem. Biotechnol. **29**: 715–732.

Lázaro-Mixteco, P.E., Nieto-Sotelo, J., Swatek, K.N., Houston, N.L., Mendoza-Hernández, G., Thelen, J.J., and Dinkova, T.D. (2012). The absence of heat shock protein HSP101 affects the proteome of mature and germinating maize embryos. J Proteome Res **11**: 3246–3258.

Lee, S., Sowa, M.E., Watanabe, Y., Sigler, P.B., Chiu, W., Yoshida, M., and Tsai, F.T.F. (2003). The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. Cell **115**: 229–240.

Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. Nature **529**: 84–87.

Li, B., Gao, Z., Liu, X., Sun, D., and Tang, W. (2019). Transcriptional Profiling Reveals a Time-of-Day-Specific Role of REVEILLE 4/8 in Regulating the First Wave of Heat Shock-Induced Gene Expression in Arabidopsis. Plant Cell **31**: 2353–2369.

Liu, H. and Charng, Y. (2013). Common and distinct functions of Arabidopsis class A1 and A2 heat shock factors in diverse abiotic stress responses and development. Plant Physiol **163**: 276–290.

Lobell, D.B., Schlenker, W., and Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. Science **333**: 616–620.

López-Frías, G., Martínez, L.M., Ponce, G., Cassab, G.I., and Nieto-Sotelo, J. (2011). Role of HSP101 in the stimulation of nodal root development from the coleoptilar node by light and temperature in maize (Zea mays L.) seedlings. J Exp Bot **62**: 4661–4673.

Marques, D.N., Amaral Carvalho, M.E., Piotto, F.A., Batagin-Piotto, K.D., Nogueira, M.L., Gaziola, S.A., and Azevedo, R.A. (2019). Chapter 16 - Antioxidant Defense Response in Plants to Cadmium Stress. In Cadmium Tolerance in Plants, M. Hasanuzzaman, M.N. Vara Prasad, and K. Nahar, eds (Academic Press), pp. 423–461.

Martínez-de la Cruz, E., García-Ramírez, E., Vázquez-Ramos, J.M., Reyes de la Cruz, H., and López-Bucio, J. (2015). Auxins differentially regulate root system architecture and cell cycle protein levels in maize seedlings. Journal of Plant Physiology **176**: 147–156.

McLoughlin, F., Kim, M., Marshall, R.S., Vierstra, R.D., and Vierling, E. (2019). HSP101 Interacts with the Proteasome and Promotes the Clearance of Ubiquitylated Protein Aggregates. Plant Physiol **180**: 1829–1847.

Meharg, A.A. (2004). Arsenic in rice--understanding a new disaster for South-East Asia. Trends Plant Sci **9**: 415–417.

Mendoza-Cózatl, D.G., Jobe, T.O., Hauser, F., and Schroeder, J.I. (2011). Long-distance transport, vacuolar sequestration, tolerance, and transcriptional responses induced by cadmium and arsenic. Curr Opin Plant Biol **14**: 554–562.

Mishra, R. and Grover, A. (2019). Voyaging around ClpB/Hsp100 proteins and plant heat tolerance.

Mishra, R.C. and Grover, A. (2016). ClpB/Hsp100 proteins and heat stress tolerance in plants. Crit Rev Biotechnol **36**: 862–874.

Mishra, R.C. and Grover, A. (2014). Intergenic sequence between Arabidopsis caseinolytic protease B-cytoplasmic/heat shock protein100 and choline kinase genes functions as a heat-inducible bidirectional promoter. Plant Physiol **166**: 1646–1658.

Mittler, R., Finka, A., and Goloubinoff, P. (2012). How do plants feel the heat? Trends Biochem Sci **37**: 118–125.

Morice, C.P., Kennedy, J.J., Rayner, N.A., Winn, J.P., Hogan, E., Killick, R.E., Dunn, R.J.H., Osborn, T.J., Jones, P.D., and Simpson, I.R. (2021). An Updated

Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set. Journal of Geophysical Research: Atmospheres **126**: e2019JD032361.

Myouga, F., Motohashi, R., Kuromori, T., Nagata, N., and Shinozaki, K. (2006). An Arabidopsis chloroplast-targeted Hsp101 homologue, APG6, has an essential role in chloroplast development as well as heat-stress response. Plant J **48**: 249–260.

Nakabayashi, K., Okamoto, M., Koshiba, T., Kamiya, Y., and Nambara, E. (2005). Genome-wide profiling of stored mRNA in Arabidopsis thaliana seed germination: epigenetic and genetic regulation of transcription in seed. Plant J **41**: 697–709.

Narsai, R., Law, S.R., Carrie, C., Xu, L., and Whelan, J. (2011). In-Depth Temporal Transcriptome Profiling Reveals a Crucial Developmental Switch with Roles for RNA Processing and Organelle Metabolism That Are Essential for Germination in Arabidopsis. Plant Physiology **157**: 1342–1362.

Nieto-Sotelo, J., Martínez, L.M., Ponce, G., Cassab, G.I., Alagón, A., Meeley, R.B., Ribaut, J.-M., and Yang, R. (2002). Maize HSP101 plays important roles in both induced and basal thermotolerance and primary root growth. Plant Cell **14**: 1621–1633.

Nover, L., Bharti, K., Döring, P., Mishra, S.K., Ganguli, A., and Scharf, K.D. (2001). Arabidopsis and the heat stress transcription factor world: how many heat stress transcription factors do we need? Cell Stress Chaperones **6**: 177–189.

Ohama, N., Sato, H., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2017). Transcriptional Regulatory Network of Plant Heat Stress Response. Trends Plant Sci **22**: 53–65.

Piacentini, D., Corpas, F.J., D'Angeli, S., Altamura, M.M., and Falasca, G. (2020). Cadmium and arsenic-induced-stress differentially modulates Arabidopsis root architecture, peroxisome distribution, enzymatic activities and their nitric oxide content. Plant Physiol Biochem **148**: 312–323.

Pourret, O. and Bollinger, J.-C. (2018). "Heavy metal" - what to do now: To use or not to use? (Letter to the Editor). Science of The Total Environment **610–611**: 419–420.

Pratt, W.B. and Toft, D.O. (2003). Regulation of signaling protein function and trafficking by the hsp90/hsp70-based chaperone machinery. Exp Biol Med (Maywood) **228**: 111–133.

Queitsch, C., Sangster, T.A., and Lindquist, S. (2002). Hsp90 as a capacitor of phenotypic variation. Nature **417**: 618–624.

Rahman, A., Bannigan, A., Sulaman, W., Pechter, P., Blancaflor, E.B., and Baskin, T.I. (2007). Auxin, actin and growth of the Arabidopsis thaliana primary root. Plant J **50**: 514–528.

Rahmati Ishka, M., Brown, E., Weigand, C., Tillett, R.L., Schlauch, K.A., Miller, G., and Harper, J.F. (2018). A comparison of heat-stress transcriptome changes between wild-type Arabidopsis pollen and a heat-sensitive mutant

harboring a knockout of cyclic nucleotide-gated cation channel 16 (cngc16). BMC Genomics **19**: 549.

Ritossa, F. (1962). A new puffing pattern induced by temperature shock and DNP in drosophila. Experientia **18**: 571–573.

Rowarth, N.M., Dauphinee, A.N., Denbigh, G.L., and Gunawardena, A.H. (2020). Hsp70 plays a role in programmed cell death during the remodelling of leaves of the lace plant (Aponogeton madagascariensis). Journal of Experimental Botany **71**: 907–918.

Sanchez, Y., Taulien, J., Borkovich, K.A., and Lindquist, S. (1992). Hsp104 is required for tolerance to many forms of stress. EMBO J **11**: 2357–2364.

Scharf, K.-D., Berberich, T., Ebersberger, I., and Nover, L. (2012). The plant heat stress transcription factor (Hsf) family: structure, function and evolution. Biochim Biophys Acta **1819**: 104–119.

Singh, S., Parihar, P., Singh, R., Singh, V.P., and Prasad, S.M. (2015). Heavy Metal Tolerance in Plants: Role of Transcriptomics, Proteomics, Metabolomics, and Ionomics. Front Plant Sci **6**: 1143.

Tamás, M.J., Sharma, S.K., Ibstedt, S., Jacobson, T., and Christen, P. (2014). Heavy metals and metalloids as a cause for protein misfolding and aggregation. Biomolecules **4**: 252–267.

Tchounwou, P.B., Yedjou, C.G., Patlolla, A.K., and Sutton, D.J. (2012). Heavy Metals Toxicity and the Environment. EXS **101**: 133–164.

Tissières, A., Mitchell, H.K., and Tracy, U.M. (1974). Protein synthesis in salivary glands of Drosophila melanogaster: relation to chromosome puffs. J Mol Biol **84**: 389–398.

Tiwari, L.D., Khungar, L., and Grover, A. (2020). AtHsc70-1 negatively regulates the basal heat tolerance in Arabidopsis thaliana through affecting the activity of HsfAs and Hsp101. Plant J **103**: 2069–2083.

Usman, M.G., Rafii, M., Ismail, M., Malek, M., A., L., and Yusuff, O. (2014). Heat Shock Proteins: Functions And Response Against Heat Stress In Plants. INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH **3**: 204–218.

Waese, J. et al. (2017). ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. The Plant Cell **29**: 1806–1821.

Wan, X. and Li, Z. (2019). Plant Comparative Transcriptomics Reveals Functional Mechanisms and Gene Regulatory Networks Involved in Anther Development and Male Sterility, Transcriptome Analysis, Miroslav Blumenberg, IntechOpen.

Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci **9**: 244–252.

Wang, Z., Lachmann, A., and Ma'ayan, A. (2019). Mining data and metadata from the gene expression omnibus. Biophys Rev **11**: 103–110.

Waters, E.R. (2013). The evolution, function, structure, and expression of the plant sHSPs. J Exp Bot **64**: 391–403.

Waters, E.R. and Vierling, E. (2020). Plant small heat shock proteins – evolutionary and functional diversity. New Phytologist **227**: 24–37.

Wells, D.R., Tanguay, R.L., Le, H., and Gallie, D.R. (1998). HSP101 functions as a specific translational regulatory protein whose activity is regulated by nutrient status. Genes Dev **12**: 3236–3251.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J. (2007). An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets. PLOS ONE **2**: e718.

Xu, Z.-S., Li, Z.-Y., Chen, Y., Chen, M., Li, L.-C., and Ma, Y.-Z. (2012). Heat Shock Protein 90 in Plants: Molecular Mechanisms and Roles in Stress Responses. Int J Mol Sci **13**: 15706–15723.

Yabuta, Y. (2016). Functions of heat shock transcription factors involved in response to photooxidative stresses in Arabidopsis. Biosci Biotechnol Biochem **80**: 1254–1263.

Yadav, A., Singh, J., Ranjan, K., Kumar, P., Khanna, S., Gupta, M., Kumar, V., Wani, S.H., and Sirohi, A. (2020). Heat Shock Proteins. In Heat Stress Tolerance in Plants (John Wiley & Sons, Ltd), pp. 189–211.

Yang, Z. and Chu, C. (2011). Towards Understanding Plant Response to Heavy Metal Stress. In.

Zhang, J., Li, X.-M., Lin, H.-X., and Chong, K. (2019). Crop Improvement Through Temperature Resilience. Annual Review of Plant Biology **70**: 753–780.

Zhang, S.-S., Yang, H., Ding, L., Song, Z.-T., Ma, H., Chang, F., and Liu, J.-X. (2017). Tissue-Specific Transcriptomics Reveals an Important Role of the Unfolded Protein Response in Maintaining Fertility upon Heat Stress in Arabidopsis. Plant Cell **29**: 1007–1023.

A comparative study of differentially expressed mRNAs in *Oryza sativa* and *Nicotiana benthamiana* upon viral infection

Thesis Submitted to the University of Delhi

in the Partial fulfilment for the Degree of

Master of Science

in

Plant Molecular Biology and Biotechnology

2021

Aashmeen Kaur

Department of Plant Molecular Biology

University of Delhi South Campus

New Delhi 110021, India

University of Delhi South Campus

DEPARTMENT OF PLANT MOLECULAR BIOLOGY

CERTIFICATE

This is to certify that the research work embodied in this M.Sc. dissertation entitled, "**A comparative study of differentially expressed mRNAs in *Oryza sativa* and *Nicotiana benthamiana* upon viral infection**" was carried out during semester IV in the Department of Plant Molecular Biology, University of Delhi South Campus and is in partial fulfilment for the degree of Master of Science in Plant Molecular Biology and Biotechnology. The work presented here demonstrates that the candidate has been trained in critical reading and summarizing research articles as well as planning, execution and presentation aspects of the experiments included in this dissertation. The candidate ensures the originality of the work as well as presentation.

Date: 15.06.2021

Signature of the Candidate

(Aashmeen Kaur)

Signature of the Supervisor

(Prof. Indranil Dasgupta)

Signature of the HOD

for (Prof. Anil Grover)

Preface

Plant viruses are responsible for loss of agricultural produce throughout the world. This has created a need for development of economical and uniform antiviral strategies. Along with the involvement of resistance genes, small RNA mediated antiviral defence has gained enormous attention in the past decade. The extent of viral infection and plant response are consequences of contributions by both endogenous small RNAs and virus derived small RNAs. In addition, viruses employ suppressors to interfere with the small RNA mediated silencing. Whether or not the host plants devise an effective and timely counter strategy, determines their survivability. However, the basic underlying cause of infection during plant-virus interactions, is transcriptional programming of a specific set of genes.

This study aims at investigating the change in expression of selected mRNAs upon Rice tungro bacilliform virus (RTBV) infection in comparison to RTBV and Rice tungro spherical virus (RTSV) infection in the monocot model plant; rice. The selected mRNAs are targets for previously characterised miRNAs in rice enabling an indirect analysis for the corresponding miRNAs upon RTBV infection. Another aim of this study was to analyze the transcriptional regulation in the dicot model plant; *Nicotiana benthamiana* upon infection with closely related geminiviruses.

Chapter 1 gives a brief introduction to the two model plants; rice and *N. benthamiana* as well their respective disease under study; rice tungro disease and cassava mosaic disease.

Chapter 2 reviews the molecular biology associated with the disease-causing viruses and the role of miRNAs in plant virus interactions.

Chapter 3 describes the experiments performed to generate rice plants infected with RTBV and *N. benthamiana* infected with Sri Lankan cassava mosaic virus (SLCMV) and tomato leaf curl New Delhi virus (ToLCNDV) in preparation of the differential expression study. However, due to restrictions in force to deal with COVID pandemic, the experiments could not be performed and the tissues were stored frozen for possible analysis in the future.

Chapter 4 highlights the existence of a similar pattern of transcriptional reprogramming in host plants upon virus infections. Despite the taxonomical differences in viruses as well as the host plants, the differentially expressed genes can be classified into a common set due to the conserved nature of plant defence-related pathways. Further analysis of the commonalities in differentially expressed genes in diverse plants due to infection by diverse viruses can lead to formulation of a uniform and comprehensive antiviral strategies.

Acknowledgements

I sincerely praise and thank the almighty God for bestowing countless blessings and providing me with the opportunity to pursue this dissertation. I feel indebted for the strength you give me each day and for the people around me who make life more meaningful.

First and foremost, I would like to express my sincere gratitude to my teacher and supervisor, Prof. Indranil Dasgupta for his guidance, untiring cooperation and indispensable advise without which this work would not have been possible. I am fortunate to have got the chance to work under his astute counsel and learn that, "dreams begin with a supervisor that impels you to overcome hurdles and leads you to the next plateau". I am truly thankful for his immense patience and critical insight for steering me to always improve during this work.

I am beholden to all the faculty members, Prof. Jitendra P. Khurana, Prof. Akhilesh K. Tyagi, Prof. Anil Grover, Prof. Paramjit Khurana, Prof Arun Sharma, Prof. Girdhar K. Pandey, Prof Sanjay Kapoor, Prof. Saurabh Raghuvanshi, and Prof. Surekha K. Aggarwal for upskilling me and motivating me to expand my capabilities for a better understanding of molecular biology. I am also very appreciative of the help offered by Mrs. Neelam and Mrs. Renu for the past two years.

I truly want to thank Dr. Gaurav Kumar and Mr. Rohit Kumar for being such supportive mentors throughout my dissertation. I greatly admire the patience and time invested by them in my work. Their proficiency and hardworking outlook have inspired me to rectify mistakes with the intention to constantly upgrade my knowledge. I acknowledge Dr. Zainul Abdeen Khan, Miss Kanika Gupta, Miss Kamlesh Kumari and Miss Neelam Jagram for their advice, aid and for making the working environment effervescent. I am especially grateful to Mr. Trilok, Mr. Brij Kishore for their unquestionable help during day-to-day lab organisation as well as for maintenance of the plant growth room.

I feel blessed for such thoughtful and cooperative batchmates who have been the cure for my homesickness for the past two years.

Finally, I offer my profound appreciation to my family for being the anchor that connects me to my roots and yet giving me the freedom to question and learn. Thank you for teaching me to be considerate about my privileges and to nullify them by merit alone. I lift up my heart to my best friends for being my support whenever I felt irresolute towards my potential. I cannot thank Elsa enough for being the most patient pet and for being the source of comfort at the end of a long tiring day.

I want to dedicate this work to all the students and teachers that are working at their full capacity to contribute to science in spite of the curtailment of resources due to the pandemic.

List of commonly used abbreviations

bp	Base pairs
CABI	Commonwealth Agricultural Bureaux International
CMB	Cassava mosaic begomoviruses
CMD	Cassava mosaic disease
CMV	Cassava mosaic virus
CP	Coat protein
CR	Common region
DNA	Deoxyribonucleic acid
dNTP	Deoxy nucleoside triphosphate
ds	Double stranded
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EtBr	Ethidium bromide
FAOSTAT	Food and Agriculture OrganizationCorporate Statistical Database
g	Gram
hr	Hour
kb	Kilo base pair
LB	Luria Bertani
lncRNAs	long non-coding RNAs
MCT	Microcentrifuge tube
miRNA	MicroRNA
min	Minutes
mM	Millimolar
MP	Movement protein
ng	Nanogram
nM	Nanometer
nmoles	Nanomoles
NSP	Nuclear shuttle protein
nt	Nucleotide
OD	Optical density
ORF	Open reading frame
PCR	Polymerase chain reaction
piRNAs	Piwi-interacting RNAs
Rep	Replication associated protein
RGM	Rice growth media
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolution per minute
RT	Room temperature
RTBV	Rice tungro bacilliform virus

RTSV	Rice tungro spherical virus
s	Second
SDS	Sodium dodecyl sulphate
siRNAs	Small interfering RNAs
SLCMV	Sri Lankann cassava mosaic virus
ss	Single stranded
ta-siRNAs	Trans acting siRNAs
TBE	Tris-borate EDTA
TE	Tris EDTA
ToLCD	Tomato leaf curl disease
ToLCNDV	Tomato leaf curl New Delhi virus
ToLCV	Tomato leaf curl virus
Tris	Tris (hydroxymethyl) aminomethane
UV	Ultraviolet
v/v	Volume/volume
VIGS	Virus induced gene silencing
v-siRNAs	Virally encoded siRNAs
VSRs	Viral suppressors of RNA silencing
µg	Microgram
µl	Microlitre
µm	Micrometre
µM	Micromolar

Table of contents

1. GENERAL INTRODUCTION……..1-10

1.1 GENERAL INTRODUCTION

1.2 RICE (*Oryza sativa* L)

1.3 RICE TUNGRO DISEASE

1.4 *Nicotiana benthamiana*

1.5 CASSAVA MOSAIC DISEASE

1.6 TOMATO LEAF CURL DISEASE

1.7 OBJECTIVES

2. REVIEW OF LITERATURE……11-24

2.1 MOLECULAR BIOLOGY ASSOCIATED WITH VIRUSES

2.1.1 RICE TUNGRO BACILLIFORM VIRUS

2.1.2 SRI LANKAN CASSAVA MOSAIC VIRUS

2.1.3 TOMATO LEAF CURL NEW DELHI VIRUS

2.2. MICRO-RNAs- INTRODUCTION AND BIOGENESIS

2.3 MICRO-RNA AND PLANT VIRUS INTERACTIONS

3.MATERIAL AND METHODS……25-32

3.1. For rice

3.2. For *Nicotiana benthamiana*

4.RESULTS AND DISCUSSION……33-50

4.1. RESULTS (RICE)

4.2 RESULTS (*N. benthamiana*)

4.3 DISCUSSION

5. CONCLUSIONS AND FUTURE WORK……51-52

5.1 CONCLUSIONS

5.2 FUTURE WORK

6. BIBLIOGRAPHY……53-64

CHAPTER 1
GENERAL INTRODUCTION

1. GENERAL INTRODUCTION

1.1 GENERAL INTRODUCTION

Plant viruses cause significant loss of agricultural produce every year. Plants and viruses have always struggled to outcompete each other by evolving new defense strategies with time. The limited knowledge about the underlying mechanisms of viral infection and its dynamic nature has created constraints in devising a single efficient strategy for virus resistance in crop plants. However, human intervention has led to increased deployment of techniques such as coat protein (CP) mediated resistance and antisense technology to manage plant viral diseases. CP-mediated resistance was first demonstrated by Beachy and co-workers (1986) against Tobacco Mosaic virus by expressing the viral coat protein in transgenic tobacco plants. One of the best examples of commercialization of this technology is the transgenic papaya resistant to Papaya Ringspot Potyvirus, developed by Gonsalves and co-workers in 1991. On the other hand, antisense technology targets a part of viral genome by production of a complementary RNA which hybridizes with the target to form a double stranded molecule. This dsRNA is then degraded by RNA silencing machinery.

Gene silencing, a phenomenon commonly associated with viral infection, has been elucidated after the discovery of three major classes of small RNAs; microRNAs (miRNAs), small interfering RNAs (siRNAs) and piwi-interacting RNAs (piRNAs). They regulate the gene expression by transcriptional gene silencing, post-transcriptional gene silencing as well as activation of gene expression.

In addition, small RNAs previously known to contribute to several life processes in plants, are now being studied for their immense role in antiviral defense. Comprehensive studies targeting the biogenesis and mode of action of miRNA have already laid the foundation of biotechnological approaches such as virus induced gene silencing (VIGS) and RNA interference to produce virus resistant plants. The potential of this technology is however limited by the non-conserved nature of viral miRNA, their abiding effects and their inadequate response to severe viral infections under field conditions. Despite this, it is one of the fastest growing fields with respect to crop improvement.

Micro RNAs have been extensively studied in monocot as well as dicot model plants with respect to plant-virus interactions. For example, in rice, the expression of 53 previously characterized miRNAs was analyzed upon Rice tungro spherical virus (RTSV) and Rice tungro bacilliform virus (RTBV) infection which helps to unfold the regulation of target genes that directly correspond to symptom development as well as plant defense responses (Zarreen *et al*., 2018). Similarly, dicot model plants like *Nicotiana benthamiana* have been studied for transcriptional reprogramming during plant virus interactions.

In a study conducted by Wu *et al*, 2019, several mRNAs correlating with Tomato yellow leaf curl virus infection in *N. benthamiana* were identified and their expression was also analyzed. In order to check, if the differential expression of these mRNAs was reproducible in *N. benthamiana* plants upon infection with related viruses like SLCMV

and ToLCNDV, similar set of genes were selected to conduct a variation study. Hence, in this dissertation, the differential regulation of plant genes by viral infections has been studied to get a broad idea of the changes taking place under such conditions.

1.2 RICE (*Oryza sativa* L)

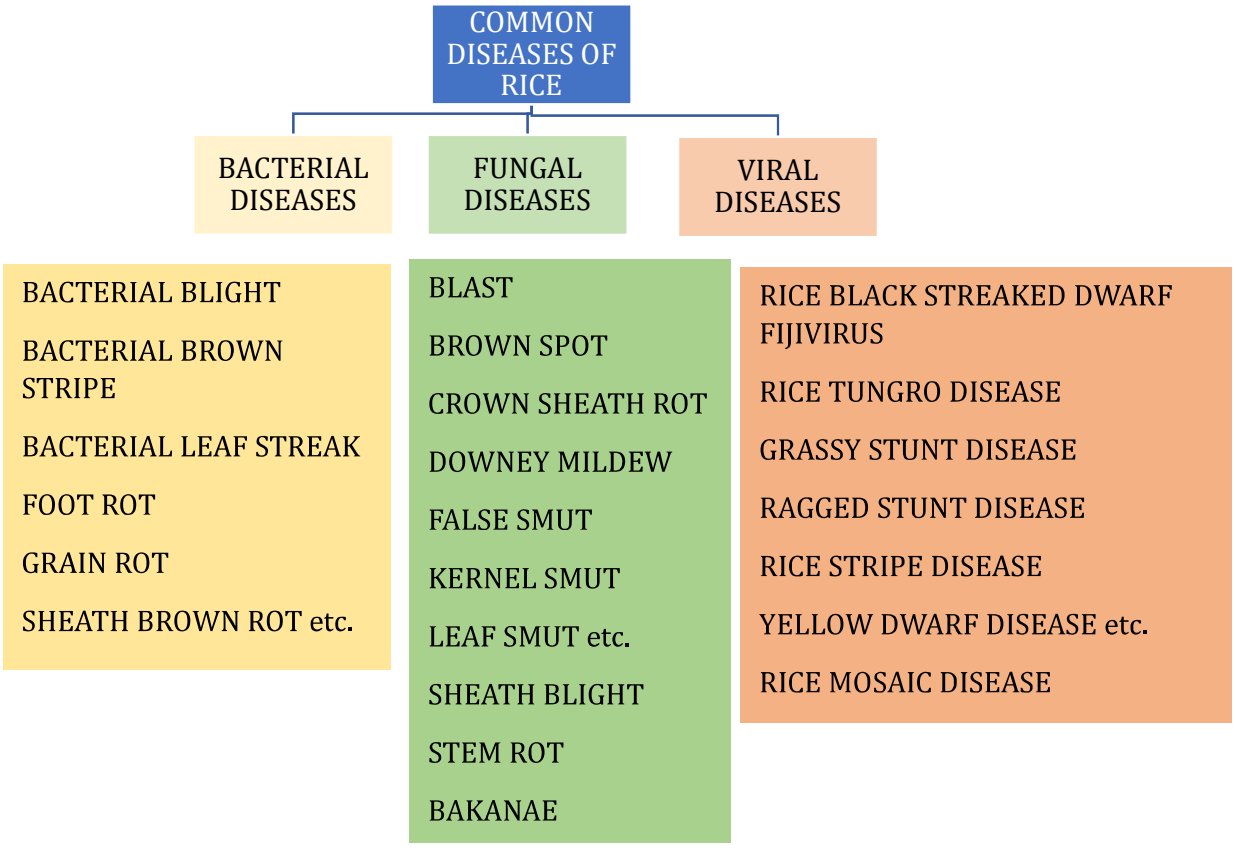
Rice (family; *Poaceae*) is one of the major cereals and its production has been recorded to be 513.2 million tonnes as per the statistics recorded till March, 2021 (FAO forecast, 04.03.2021). In order to meet the high demand, several strategies for increasing the production have been adapted while some are still in formulation.

WORLD RICE MARKET

	Production [1/]	Supply [2/]	Utilization	Trade [3/]	Ending stocks [4/]	World stock-to-use ratio	Major exporters' stock-to-disappearance ratio [5/]
	(million tonnes)					(percent)	
2011/12	480.0	609.3	462.0	40.6	146.8	31.2	25.2
2012/13	485.1	631.9	469.3	40.2	162.4	33.8	28.0
2013/14	490.9	653.3	480.4	45.5	171.9	35.3	29.0
2014/15	490.5	662.4	487.1	45.1	173.9	35.4	24.6
2015/16	489.0	662.9	490.7	41.4	172.2	34.8	19.7
2016/17	497.1	669.3	494.7	48.4	173.4	34.8	18.9
2017/18	499.9	673.3	497.8	48.6	176.6	35.3	18.1
2018/19	507.3	683.9	500.8	44.2	185.4	36.8	22.6
2019/20	502.8	688.1	504.2	45.1	182.2	35.4	24.3
2020/21	513.2	695.4	514.4	48.2	182.5	35.1	24.7

Fig 1.1: world Rice market (FAOSTAT, 2021)

Susceptibility of rice to attack of certain pests and diseases have adversely affected its production. Some of the major common diseases of rice are:



1.3 RICE TUNGRO DISEASE

Rice Tungro Disease is caused by the combination *Rice tungro spherical virus* and *Rice tungro bacilliform virus* (Hibino *et al*., 1978). It is prevalent in the major rice producing South Asian countries as depicted in figure 1.2.

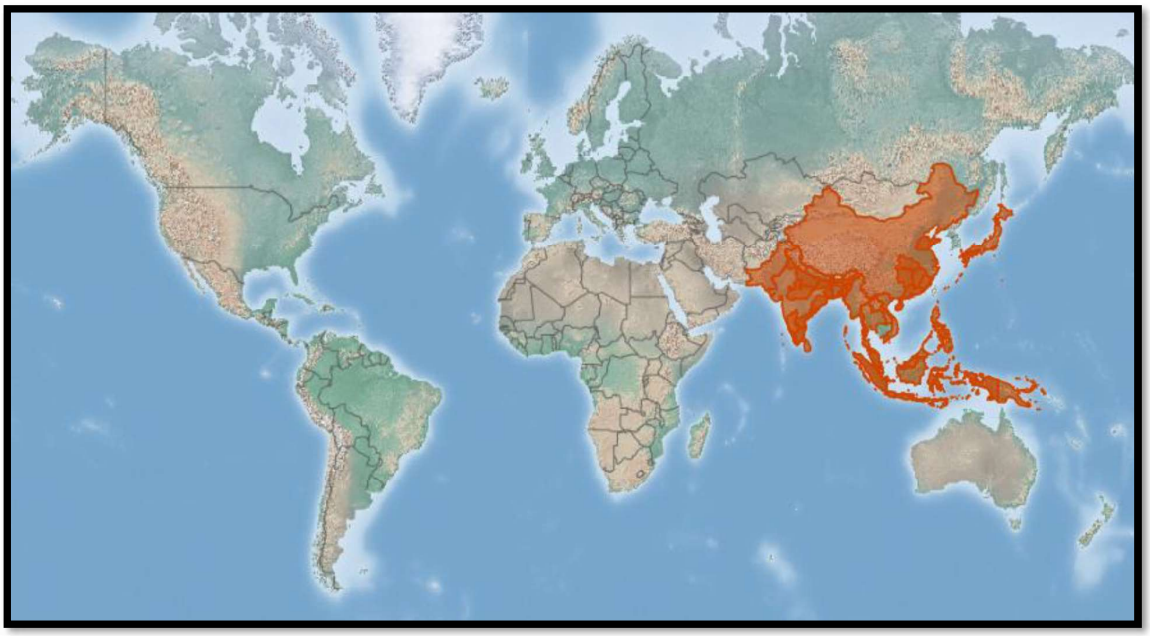


Fig1.2. Distribution of Rice Tungro Disease **(source: CABI Summary data, 2021)**

The virus complex is transmitted by green leafhopper (*Nephotettix virescens*). RTSV can be transmitted alone by leafhopper but it induces fewer symptoms. On the contrary, RTBV which is not transmitted alone by leafhopper can cause severe symptoms. However, the combination of both the viruses can be transmitted by the leafhopper and induces extreme symptoms (Hibino *et al*., 1983).



Fig 1.3. *Nephotettix virescens* (Green leafhopper) (source: https://www.plantwise.org)

The disease symptoms include yellow to orange coloration of leaves, abnormal leaf shape, irregular dark brown spots on leaves, reduced tillering, discoloration of panicle, sterile panicles, seed discoloration, stunting of stem, and retarded or stunted growth. These symptoms can easily be associated with other biotic as well as abiotic factors. Therefore, efficient methods of diagnosis including serological (for example, ELISA) as well as PCR-based (RT-PCR) techniques are used. Multiplex RT-PCR (Periasamy *et al*., 2003) and SYBR

green based RT-PCR (Dasgupta *et al*., 2012) have been widely used for detection of RTBV and RTSV.

1.4 *Nicotiana benthamiana*

Nicotiana benthamiana (Family; *Solanaceae*) is widely used as a model plant in plant virology due its easy manipulation for studying various host-pathogen interactions. It was first discovered in Australia by Benjamin Bynoe. It has been predominantly employed in virus induced gene silencing (VIGS) as well as protein expression studies due to its broad range of susceptibility for viruses as well as other pathogens like bacteria and fungi. It In addition, it is quite easy to carry out agroinfiltration in *N. benthamiana* unlike in case of other model plants like *Arabidopsis thaliana* due to the lack of an efficient high throughput transient system in the latter. Hence, it has contributed immensely to supplement the plant biology related research in other model plants as well (Goodin *et al*., 2008)

1.5 CASSAVA MOSAIC DISEASE

One of the prominent staple foods of Africa is cassava (*Manihot esculenta*). Its productivity is greatly reduced by Cassava mosaic disease (CMD). The causal agent of CMD is *cassava mosaic* begomoviruses (CMBs) which belongs to the family, *Geminiviridae*. CMD shows a wide distribution covering Africa, Indian Ocean Islands, India and Sri Lankan. (CABI, 2019).

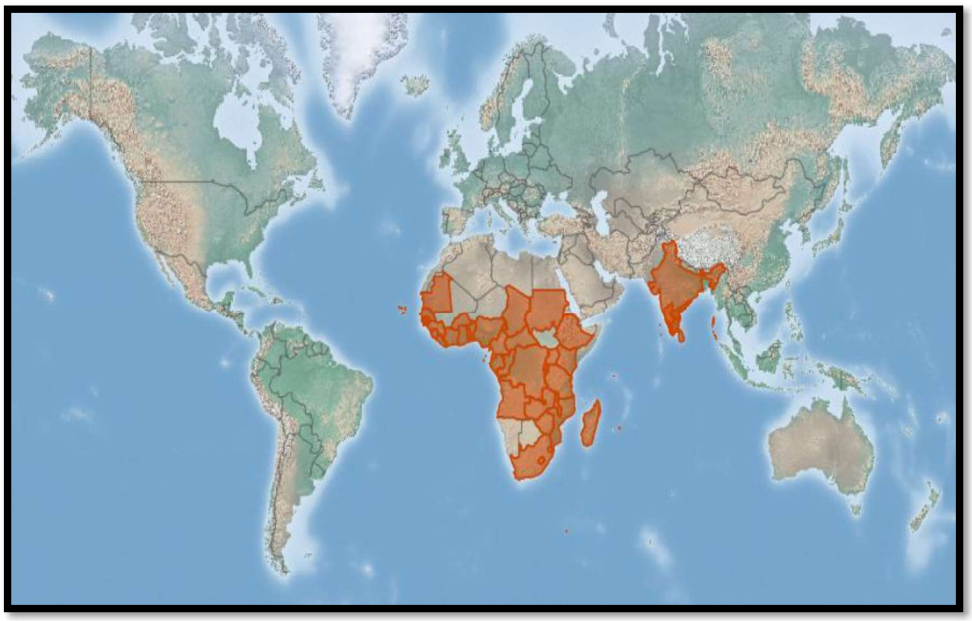


Fig 1.4. Distribution of Cassava mosaic disease (Source: CABI summary data, 2019)

Table1.1: List of major cassava mosaic Geminiviruses (CMGs)

Commonly occurring CMGs in Africa.
African cassava mosaic virus (ACMV)
East African cassava mosaic virus (EACMV)
East African cassava mosaic Cameroon virus
East African cassava mosaic Kenya virus (EACMKV)
East African cassava mosaic Malawi virus
East African cassava mosaic Zanzibar virus
South African cassava mosaic virus (SACMV)

Commonly occurring CMGs in India
Indian cassava mosaic virus (ICMV)
Sri Lankann cassava mosaic virus (SLCMV)

Whitefly (*Bemisia tabaci,* order*: Hemiptera,* family: *Aleyrodidae*) is the primary transmission vector for the virus however, cassava mosaic begomoviruses can also spread through infected plant material as suggested by Harrison (1987).



Fig 1.5. Whitefly (*Bemisia tabaci)* (source: https://irac-online.org/pests/bemisia-tabaci/)

The disease symptoms show a mosaic pattern of leaf chlorosis, abnormal leaf shape and pattern and retarded growth. Diagnosis of CMD is polyclonal as well as monoclonal antibody based. Sandwich ELISA (enzyme-linked immunosorbent assay) is a commonly used diagnostic tool. In addition, PCR- based as well as hybridization-based methods are frequently employed for detection.

1.6 TOMATO LEAF CURL DISEASE

Tomato (*Solanum lycopersicon* L.) is the second most important economically cultivated vegetable crop in the world. China and India rank first and second respectively, in the world in terms of area covered and production. Tomato leaf curl disease (ToLCD) is one of the major biotic factors affecting the yield. Tomato leaf curl virus (ToLCV), the causal agent of the disease, belongs to the genus *Begomovirus* (Family: *Geminiviridae*) This virus, having originated in the Middle East, shows a wide geographical distribution covering several countries of Africa, Asia, Europe, North America, Oceana and South America (Fig. 1.6)

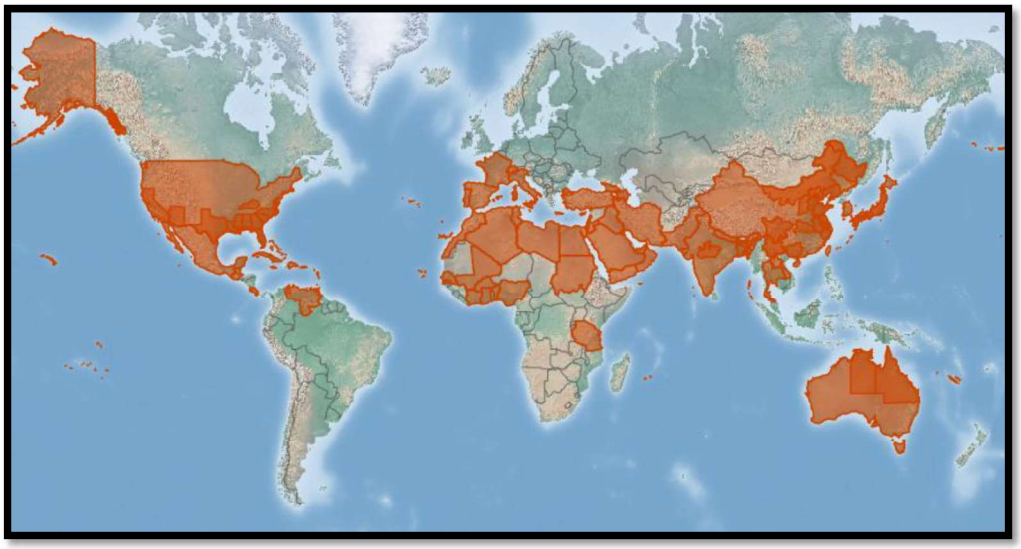


Fig 1.6. Worldwide distribution of tomato yellow leaf curl disease (source: CABI summary data, 2020)

In India, Leaf curl disease in tomato is caused by monopartite as well as bipartite begomoviruses.

Table 1.2. Monopartite species of ToLCD causing viruses in India (source: Kumar *et al*., 2018)

Name of the virus species
Tomato leaf curl Kerala virus (ToLKeV)
Tomato leaf curl Ranchi virus (ToLCRnV)
Tomato leaf curl Patna virus (ToLCPaV)
Tomato leaf curl Rajasthan virus (ToLCRV)
Tomato leaf curl Pune virus (ToLCPuV)
Tomato leaf curl Bangalore virus (ToLCBV)
Tomato leaf curl Karnataka virus (ToLCKV)
Tomato leaf curl Joydebpur virus (ToLCJoV)

Table 1.3. Bipartite species of ToLCD causing viruses in India (source: Kumar *et al*., 2018)

Name of the virus species
Tomato leaf curl New Delhi virus (ToLCNDV)
Tomato leaf curl Palampur virus (ToLCPalV)
Tomato leaf curl Gujarat virus (ToLCGV)

The virus is transmitted by whitelfly, *Bemisia tabaci.* Disease symptoms include abnormal coloration, forms or patterns of leaves with stunted plant growth. The seedlings appear bushy and fruiting is reduced if virus infects prior to flowering and the margins of the leaves are rolled upwards. Diagnosis of infection includes a wide range of techniques like Hybridization based; Southern hybridization and tissue print hybridization (Navot *et al*., 1989), polymerase chain reaction, rolling circle amplification, microarrays, immunology based; ELISA, dot-immuno binding assay and immune electron microscopy.

1.7 OBJECTIVES

There have been several reports of miRNA-mediated plant-virus interactions in rice as well as *N. benthamiana* some of which have been discussed further in chapter 2. A study conducted by Zarreen *et al*. (2018) exemplified the role of previously characterized miRNAs during RTBV and RTSV infection in rice. These miRNAs targeted the essential genes for plant defense responses, hormone signaling and various developmental aspects and displayed differential expression upon viral infection. The first objective of this work relates to the study of the effect of RTBV infection on a selected set of genes targeted by previously known miRNAs. The aim was to draw a comparison between miRNA-mediated plant-virus interaction upon viral complex (RTBV+RTSV) infection and RTBV infection.

Similarly, another study was conducted by Wu *et al*, (2019) to study the transcriptomic changes in *N. benthamina* upon infection by tomato yellow leaf curl virus (TYLCV). Virus infection led to differential expression of genes associated with signaling of hormones like auxin and ethylene. The second objective was undertaken to check whether a similar

pattern of transcriptional reprogramming takes place during infection by related geminiviruses (SLCMV and ToLCNDV) in the model plant *N. benthamiana*.

The objectives are as follows:

1. Indirect analysis of previously characterized miRNAs by studying the change in expression of their respective mRNA targets upon RTBV infection in rice using RT-PCR.
2. Variation study to investigate the differential expression of selected genes in *N. benthamiana* upon SLCMV and ToLCNDV infection using RT-PCR.

CHAPTER 2
REVIEW OF LITERATURE

2. REVIEW OF LITERATURE

2.1 MOLECULAR BIOLOGY ASSOCIATED WITH VIRUSES

2.1.1 RICE TUNGRO BACILLIFORM VIRUS

As the name suggests, RTBV is a bacilliform particle (130nm X 130nm). The genome comprises of a double stranded DNA. Being a pararetrovirus, it replicates via reverse transcription. The genome size ranges from 7907-8016 bp. The transcription is asymmetrical such that one strand consists of four open reading frames (ORFs):

Table 2.1. Proteins encoded by ORFs of RTBV and their potential roles.

ORF	Protein	Role	References
I	P24	Viral particle assembly	Hay *et al*. (1994)
II	P12	Interaction with coat protein for capsid shell assembly.	Herzog *et al*. (2000)
III	P194	Polyprotein having motifs for viral coat protein, aspartate protease, reverse transcriptase and ribonuclease H.	Guerra-Peraza *et al*. (2005) Hagen *et al*. (1993) Hull *et al*. (1996) Marmey *et al*. (1999) Marmey *et al*. (2005) Laco *et al*. (1994)
IV	P46	RNAi silencing suppressor	Rajeswaran *et al*. (2014)

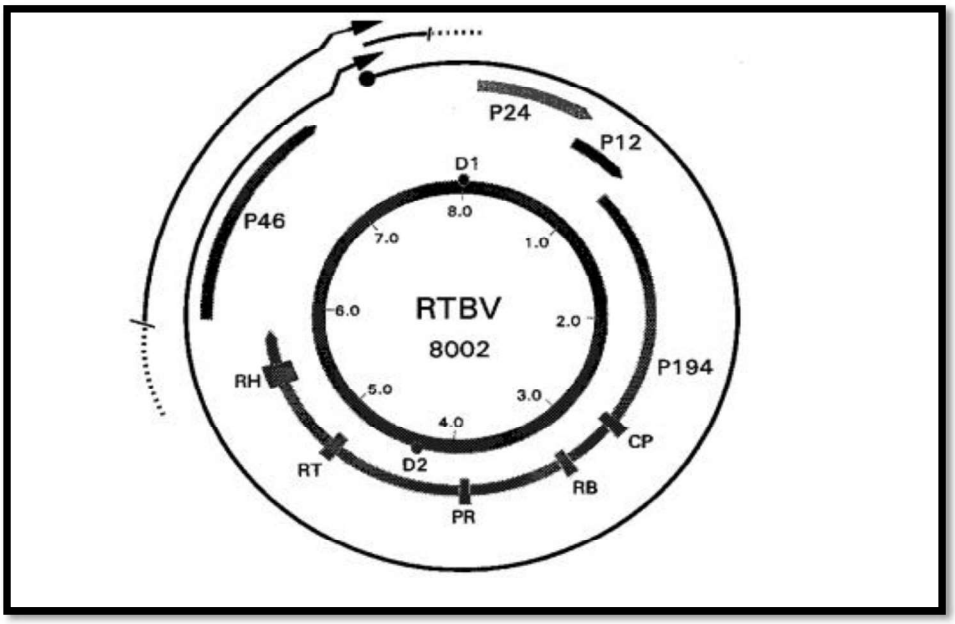


Fig 2.1. Structural organization of RTBV genome (CP=Coat protein, PR=Aspartate protease, RT=Reverse transcriptase, RH=RNaseH) (Source: Hull *et al*., 1996)

2.1.2 SRI LANKAN CASSAVA MOSAIC VIRUS

SLCMV shows a typical geminate partial icosahedral morphology. The genome of SLCMV comprises of single stranded DNAs; DNA A and DNA B (characteristic of begomoviruses). DNA A encodes for seven genes; AC1, AC2, AC3, AC4, AC5, AV1 and AV2 whereas, DNA B encodes for two genes; BC1 and BV1. The replication takes place by the mechanism of rolling circle amplification.

Table 2.2. proteins encoded by ORFs of SLCMV and their potential roles.

ORFs	Protein	Role	References
AC1	Replication Associated Protein, Rep	Initiation of viral replication	Lazarowitz *et al*. (1992)
AC2	Transcriptional Activator Protein, TrAP	Gene activation, pathogenicity, suppression of RNAi	Hanley-Bowdoin *et al*. (1999) Saunders and Stanley. (1995) Sunter *et al*. (1994)
AC3	Replication Enhancer Protein, Ren	Enhances DNA replication	Settlage *et al*. (2005) Pasumarthy *et al*. (2010)

AC4	C4	Symptom development and infection. Potential suppressor of RNAi	Fondong (2013)
AC5	-	Suppression of post transcriptional gene silencing.	Li *et al.* (2015)
AV1	Coat Protein	Packaging of genome, virus transmission	Briddon *et al.* (1990)
AV2	Pre-coat Protein	Cell-to-cell movement and pathogenicity	Fondong (2013)
BC1	Movement Protein, MP	Cell-to-cell as well as long distance transmission of virus	Fondong (2013)
BV1	Nuclear Shuttle Protein, NSP	Transports the viral ssDNA between nucleus and cytoplasm.	Ward and Lazarowitz. (1999)

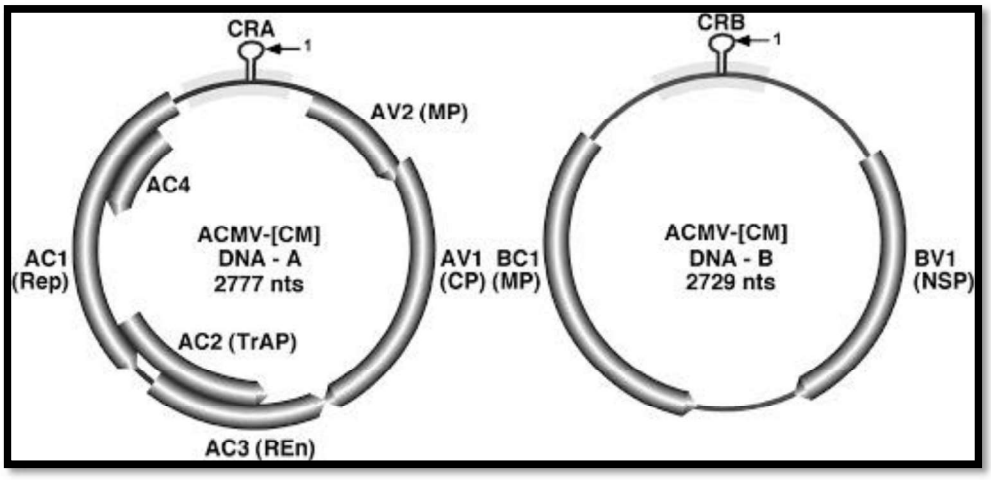


Fig 2.2. Structural organization of SLCMV genome. (source: Chellappan *et al.*, 2005)

2.1.3 TOMATO LEAF CURL NEW DELHI VIRUS (ToLCNDV)

Similar to SLCMV, ToLCNDV also belongs to the family *Geminiviridae* and genus *Begomovirus*. The virus exhibits monopartite as well as bipartite genome structure. The bipartite genome of ToLCNDV comprises of DNA A and DNA B. The ORFs encode for similar proteins as in SLCMV.

However, the monopartite ToLCNDV consists of DNA A and β- satellite molecule (important for inducing symptoms). The DNA molecules share a common region (CR)

which acts as the origin of replication. The virus exhibits the characteristic rolling circle replication of geminiviruses.

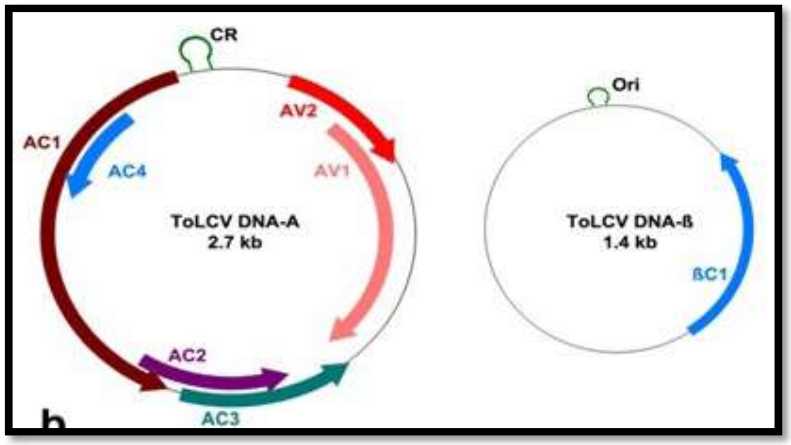


Fig 2.3. Genome organization of monopartite ToLCNDV (source: Pradhan *et al.*, 2017)

2.2. MICRO-RNAs- INTRODUCTION AND BIOGENESIS

MicroRNAs are (avg. 22 nucleotide) long single stranded RNA transcribed from non-protein encoding RNAs. Advancement in technology has led to the elucidation of their role in basic developmental processes in plants as well as animals. Their mechanism of action is based on transcriptional and post-transcriptional gene silencing by degradation of the target (complementary) genes. However, the miRNA dependent target cleavage may not be a consequence of perfect complementarity both in plants and animals.

This type of gene regulation makes miRNA a target applicant for the strategies for controlling the host's genomic machinery by viruses**.** In fact, it would not be improper to presume that the primary role of miRNA mediated regulation was to counter viral infections and the other secondary roles developed later during evolution. Further understanding of their diversified roles during plant viral infection can open avenues for unveiling the mechanisms of plant immunity as well as for generating viral resistant transgenic plants.

MicroRNAs were previously considered to have evolved after multi-cellularization but extensive studies have now revealed that their origin is as primitive as unicellular organisms (*Chlamydomonas* sp). And their conserved nature across species indicates the presence of a common ancestor. Despite this, among the plant species, their regulation and function remain diverse.

Their evolution could be potentially due to several events. In addition, the evolution of viruses has greatly contributed to evolution of small RNA mediated silencing, to keep up with these infiltrators. In fact, not only endogenous but virally encoded miRNA play an important role in mediating the host-pathogen interactions.

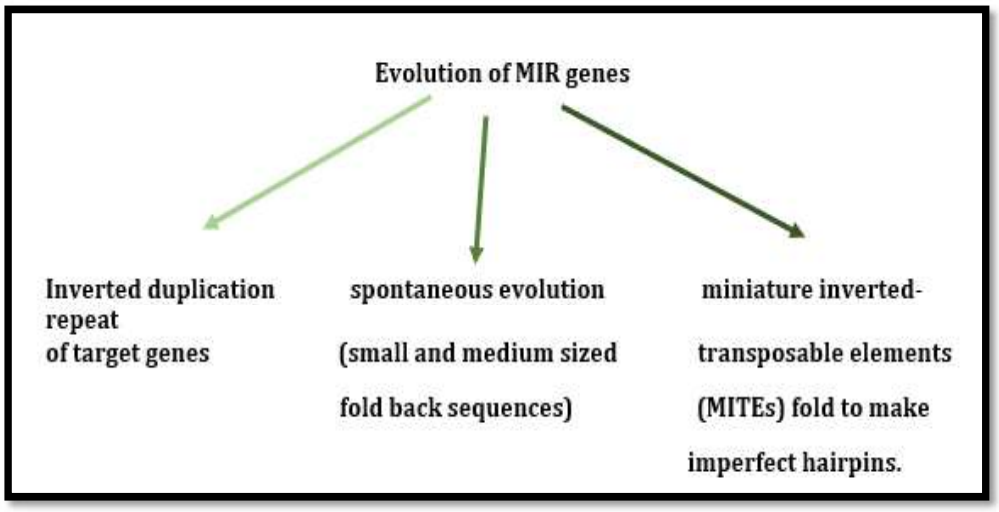


Fig 2.4. Factors responsible for evolution of MIR genes (adapted from Liu *et al.*, 2017)

In addition, the targeted genes also remain conserved. So, it would be suitable to convey the idea of different mechanisms of regulation associated with miRNA. The conservation however, is absent between plants and animals. The conserved biogeny of the small RNA can be correlated with the functional aspects of small RNA which are highly crucial for plant development and defense.

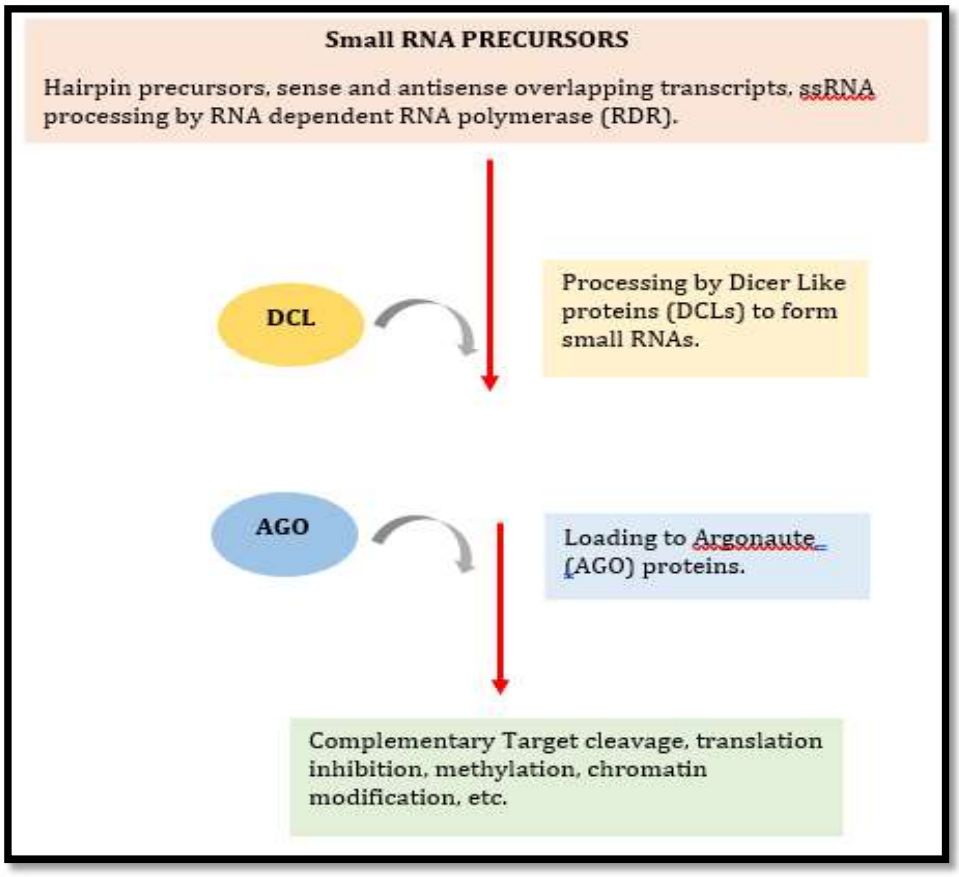


Fig 2.5. Overview of small RNA biogenesis.

Stem-loop precursors of microRNAs are transcribed from microRNA (MIR) genes via RNA polymerase II. Dicer-like proteins (DCLs) in association with dsRNA-binding protein and HYPONASTIC LEAVES1 (HYL1), cleave the primary transcript and leads to generation of partially double stranded precursor miRNAs (pre-miRNAs) in the nucleus. DCL1 and HYL1 then cleave the passenger strand and HEN1 methylates the mature strand. This strand is then transported to cytoplasm by Exportin-5 ortholog HASTY. RNA-induced silencing complex (RISC) is formed when association between mature strand and Argonaute 1 (AGO1) takes place. Hsc70/Hsp90 mediates this process in an ATP-dependent manner. RISC targets the complementary mRNA and induces target suppression by either translational repression or target cleavage.

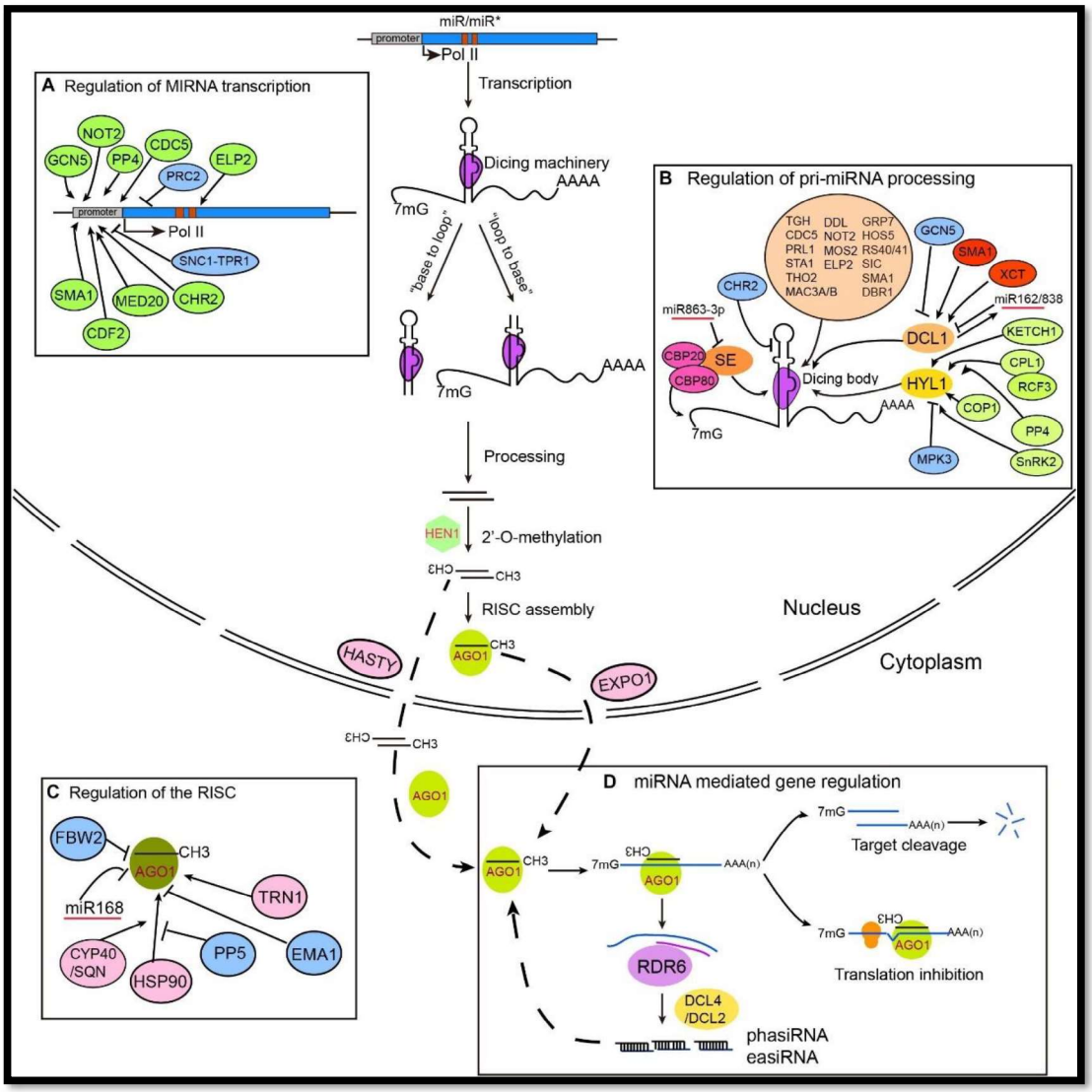


Fig 2.6. Process of microRNA biogenesis in plants (source: Wang *et al.*, 2019).

2.3 MICRO-RNA AND PLANT VIRUS INTERACTIONS

The non-coding endogenous RNAs play a vital role in plant's life cycle. Upon viral infection, they are the first to draw a line of defense. Recently miRNAs, virally encoded siRNAs (v-siRNAs), trans acting siRNAs (ta-siRNAs) and long non-coding RNAs (lncRNAs) have gained a lot of attention for their roles in plant virus interactions. Several miRNAs have been characterized which can either be downregulated; miR171b in rice during RSV infection (Tong *et al.*,2017) or upregulated; miR168 due to p19 suppressor (Várallyay *et al.*, 2010) upon viral infection. Virus activated siRNAs on the other hand are dependent on RDR1 and DCL4 (Garcia-Ruiz *et al.*, 2010) for genesis, hence resembling vi-siRNAs. LncRNAs can be both endogenous and virally encoded and sometimes even fusion of both.

Both viral and host specific miRNA participate in the race to survival. Viruses incapacitate the plant defense system with the help of viral encoded large T antigens as well as by affecting the host's silencing machinery. Several miRNAs have been identified and characterized across a wide range of plant species as they are differentially regulated upon virus invasion in contrast to animals where siRNAs show a rapid increase upon infection.

Also, to retaliate, viruses make use of viral suppressors of RNA silencing (VSRs) by following ways:

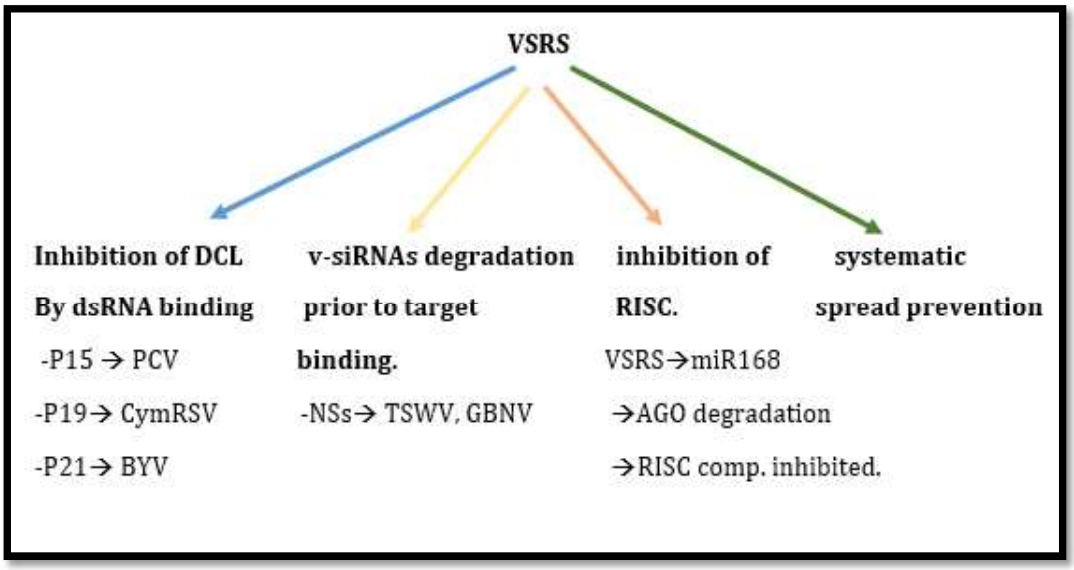


Fig 2.7. Mode of action of VSRs (adapted from Prasad *et al*., 2019)

In fact, there have been several reports where the differentially regulated miRNAs upon virus infection, can be directly correlated with the disease symptoms. One such example is the downregulation of osa-miR171b in rice upon RSV infection. The targets of osa-miR171b include GRAS (Gibberellic-Acid Insensitive, Resistance gene analogs, SCARECROW) family genes; *OsSCL6-IIa, OsSCL6-IIb* and *OsSCL6-IIc* (inhibitors chlorophyll synthesis) and *Early heading date genes; Ehd1, Ehd2, Ehd3, Ehd4* and *Heading date 3a; Hd3a (*positive regulators of heading in rice). These genes are involved in gibberellin signalling in plants. Therefore, the down-regulation of osa-microRNA171b during RSV infection could be responsible for the development of symptoms like yellowing of leaves, stunted growth or reduced height, delayed heading, etc (Tong *et al*., 2017).

On the contrary sequestration of miRNAs like miR528 have a positive effect on RSV infection. RSV infection leads to increase in expression of AGO18 which competes with AGO1 and preferentially binds to miR528 and trapping it. This prevents the degradation of target ascorbate oxidase mRNA, further preventing the production of Reactive Oxygen Species (ROS). The decrease in ROS levels increases vulnerability to viral infection (Wu *et al*., 2017).

The contribution of miRNAs in plant-virus interaction can be extended to other monocots like maize as well. miR159, miR166, miRNA168, miR390, miR397, miR408, miR528, miR529 and miR827 have been reported to be differentially regulated in maize upon *Sugarcane mosaic virus* infection (Xia *et al*., 2018). This regulation directly correlates with their targets which ultimately contribute to the antiviral response as summarized in table 2.3.

Table 2.3. Correlation between identified miRNAs and their targets to understand their role in virus infection (adapted from Xia *et al*., 2018.

MicroRNA family	Expression change	Target	Role of target in antiviral response
miR168	Upregulated	Argonaute 1	Homeostasis and feedback regulation of miRNA.
miR528	Upregulated	Ascorbate oxidase	ROS; signal transduction.
miR159	Downregulated	GAMYB (Gibberellin-Myeloblastosis)	Symptom development.
miR397	Downregulated	Laccase	Copper homeostasis
miR827	Downregulated	SPX-MFS (SYG1, PHO81, XPR1-Major Facility Superfamily	Phosphate homeostasis
miR408	Upregulated	Copper containing proteins	Abiotic stress response

Another rice infecting virus that hinders its production is a complex of RTBV and RTSV. Deep sequencing analysis in rice upon both RTBV and RTSV infection has shown the predominant production of siRNAs as well as the differential expression of 54 previously characterized miRNAs (Zarreen *et al*., 2018). In the present study, an indirect analysis of differential expression of these previously characterized miRNAs was performed by studying the variation in expression of their target mRNAs upon only RTBV infection. This can further elaborate the differences in miRNA expression profile during the infection by the virus complex and the infection caused by RTBV alone.

A similar pattern of miRNA-based regulation of viral infection has been observed in the dicot model plant *N. benthamiana*. Several virus infections in *N. benthamiana* have been reported to be a consequence of miRNA-based silencing. For example, the effect of *Tobacco curly shoot Virus* (TbCSV) and *Tobacco curly shoot betasatellite* (TbCSB) infection in *N. benthamiana* plants on endogenous miRNA expression has been analyzed by small RNA library construction which led to discovery of several novel miRNAs (Du *et al*., 2019).

Table 2.4. Expression profile of miRNAs and their target genes in *Nicotiana benthamiana* upon *Tobacco curly shoot Virus* and *Tobacco curly shoot betasatellite* infection (adapted from Du *et al*., 2019).

miRNAs	Expression change	Predicted target	Expression change
miR156a	Upregulated	SPL	Downregulated
miR482a	Upregulated	Disease resistance protein	Downregulated
miR169c	Upregulated	Nuclear transcription factor Y subunit	Downregulated
miR4376	Upregulated	Calcium transporting ATPase.	Downregulated
miR4376	Upregulated	Conserved hypothetical protein	Downregulated
miR159	Upregulated	MYB-like transcription factor	Downregulated
Novel-121	Upregulated	Heavy metal transport	Downregulated
Novel-71	Upregulated	N-Acetyltransferase	Downregulated
miR171	Downregulated	GRAS transcription factor family	Upregulated
miR164a	Downregulated	NAC-domain containing protein	Upregulated
Novel-70	Downregulated	Transcription factor	Upregulated
Novel-94	Downregulated	1-aminocyclopropane-1-carboxylate oxidase.	Upregulated

In addition, the interaction of *Potato virus Y* (PVY) and tobacco plant has been analyzed by combining data from high throughput sequencing of miRNAs as well as mRNAs upon viral infection. It was observed that a total of 81 miRNAs showed significant alterations upon viral infection (21 miRNAs downregulated and 57 miRNAs upregulated). Further, it was observed that miRNA targeting translationally controlled tumor protein (NtTCTP) gene was upregulated to decrease the expression of this target gene which conferred resistance in plants against PVY (Guo *et al*., 2017). This helped in further understanding the mechanism employed by the virus to weaken the plant's immune response**.** Another study, was aimed to investigate the efficiency *of Potato virus X* (PVX) based vectors in restraining the miRNA-based gene silencing using short tandem target mimic (STTM) technology in *N. benthamiana* plants. Further, the predicted role of specific miRNAs was verified based on phenotypic evaluation and real time RT-PCR. This gave an elaborate account of virus-based miRNA suppression (VbMS) of previously characterized miRNAs such as miR165/166 and miR159 via STTMs using PVX vectors (Zhao *et al*., 2016).

Table 2.5. Elucidation of roles of miRNAs based on their suppression by transformation with Potato virus X-based Short Tandem Target Mimics.

miRNA	Target	Techniques employed	Remarks
miR165/166	Homeodomain-Leucine Zipper (HDZip) family involved in leaf development	Transformation with PVX-based STTM165/166. RT-PCR	miR165/166 was responsible for down-regulation of HD-Zip genes which could express when the miRNAs were sequestered by their complement STTMs.
miR159	MYB-transcription factors	PVX-STTM159 was used to transform *N. benthamiana* plants. RT-PCR	miR159 is responsible for targeted cleavage of MYB-transcription factors but due to its binding with STTMs, the expression of these genes remains unhindered.

Plant viruses have been reported to employ counter strategies against miRNA-related plant defense response. In order to study the tactics utilized by the viruses, various *Plum Pox virus* chimeras containing target sites as well as mutated target sites of miRNAs (miR159, miR167 and miR171) were used to infect *Nicotiana clevelandii, N. benthamiana* and *A. thaliana* plants. The presence of unaltered as well as mutated target sites decreased the virus infectivity. In the constructs with unaltered target sites, acquisition of mutations was observed as an escape strategy from miRNA-based defense. This could also be the counter strategy exploited by the plant viruses against miRNA-mediated silencing. (Mateo *et al*, 2006).

In order to understand the symptoms associated with certain begomoviruses like *African cassava mosaic virus* (ACMV), *Cabbage leaf curl virus* (CbLCuV), *Cotton leaf curl Multan virus/ Cotton leaf curl betasatellite* (CLCuV/CLCuMB) and *Tomato yellow leaf curl virus* (TYLCV), miRNA analysis have been performed on *N. benthamiana* plants infected with the individual viruses. The levels of a specific set of miRNAs varied corresponding to infections with several begomoviruses which could be correlated with the respective disease symptoms (Amin *et al*., 2011).

Table 2.6. Differential expression of miRNAs in *N. benthamiana* following infection with begomoviruses (adapted from Amin *et al.*, 2011)

miRNA	**Potato virus X (PVX) as control**	***African cassava mosaic virus* (ACMV)**	***Cabbage leaf curl virus* (CbLCuV)**	***Cotton leaf curl Multan virus/ Cotton leaf curl betasatellite* (CLCuV/CLCuMB)**	***Tomato yellow leaf curl virus* (TYLCV)**	**TARGET**
miRNA156	Increase	Decrease	Decrease	Increase	Increase (highest)	transcription factor, SQUAMOSA PROMOTER BINDING PROTEIN (SPL)
miRNA159	Increase	Increase	Increase (highest)	Increase	Increase	MYB proteins
miRNA160	Slight decrease	Decrease	Slight increase	Slight increase	Increase (highest)	Auxin response factor
miRNA164	increase	Increase	Increase (highest)	Increase	Increase	CUP-SHAPED COTYLEDON 1 (CUC1) and CUC2 genes which are NAC like transcription factors
miRNA165/166	Slight Increase	Increase	Increase (second highest)	Increase (highest)	Increase	HD-ZIPIII transcription family (PHABULOSA, PHAVOLUTA)
miRNA167	Slight Increase	Increase	Increase (second highest)	Increase (highest)	Increase	Auxin response factors like ARF 6 and ARF 8
miRNA168	Increase	Increase	Increase (second highest)	Increase	Increase (highest)	-
miRNA169	Slight Increase	Non-significant decrease	Slight Increase	Slight Increase	Slight Increase	act as P and N-dependent signalling molecules.
miRNA170	Slight Increase	Non-significant increase	Increase (highest)	Increase	Non-significant increase	-

There are several reports documenting the differential expression of miRNAs following infection with various begomoviruses in *N. benthamiana*. These miRNAs target some essential genes for plant development and growth. A study was conducted to analyze the reorganized transcriptome upon TYLCV infection in *N. benthamiana* and it was found that

a drastic change in the expression level of several developmentally important genes takes place (Wu *et al*, 2019). To determine whether a similar change is also seen following infections with SLCMV and ToLCNDV in *N. benthamiana*, this study is conceived. Based on BLAST analysis, it was observed that TYLCV has 74.37% nucleotide sequence identity with ToLCNDV and 74.4% nucleotide sequence identity with SLCMV. So, the present study aimed to analyze the differential expression of the mRNAs of 12 selected genes in *N. benthamiana* upon ToLCNDV and SLCMV infection and to derive a correlation with the results obtained by Wu *et al*. (2019).

CHAPTER 3
MATERIAL AND METHODS

3.MATERIAL AND METHODS

3.1. RICE

3.1.1. PLANT GROWTH

Seeds of rice varieties; Pusa Basmati 1, Taichung Native 1, backcrossed RTBV resistant varieties of ASD-16, BPT-5204 were soaked in water (RO) for 24 hours followed by their transfer onto muslin cloth soaked in 1X Rice Growth Medium (RGM; Yoshida *et al* 1972). The seeds were allowed to germinate in the plant growth chamber with a 12 hr light period and a temperature of 28-32^{0}C

Table 3.1 Composition of 5X Rice growth medium (RGM)

Stock no.	Chemical name	Chemical formula	500 ml stock	Working solution 1L
1	Ammonium Nitrate	NH_4NO_3	45.7 g	2 ml
2	Sodium Dihydrogen Phosphate	$NaH_2PO_4.2H_2O$	20.15 g	1 ml
3	Potassium Sulphate	K_2SO_4	35.7 g	1 ml
4	Calcium Chloride	$CaCl_2$	44.3 g	1 ml
5	Magnesium Sulphate	$MgSO_4.7H_2O$	162.0 g	1 ml
6-A	Manganese Chloride	$MnCl_2.4H_2O$	0.75 g	1 ml
6-B	Ammonium Molybdate	$(NH_4)_6Mo_7O_{24}.4H_2O$	0.037 g	
6-C	Boric Acid	H_3BO_3	0.467 g	
6-D	Zinc Sulphate	$ZnSO_4.7H_2O$	0.0175 g	
6-E	Copper Sulphate	$CuSO_4.5H_2O$	0.0155 g	
6-F	Ferric Chloride	$FeCl_3.6H_2O$	2.31 g	
6-G	Citric Acid		5.95 g	
6-H	Sulphuric Acid	H_2SO_4	25 ml	
	Sodium Hydroxide pellet	NaOH	40.0 g	Adjust pH to 4.5-5.0

3.1.2. PLASMID (pRTBV-Inf) ISOLATION.

Plasmid isolation was done using alkaline lysis method (Birnboim and Doly., 1979). Culture of volume 1.5 ml was obtained in an MCT and the cells were harvested by centrifugation at 6000 x g for 10 minutes at 4^{0}C. The pellet was washed in 300 µl of MQ

and further centrifugation was carried out at 6000 x g (5 minutes at 4^0C). After inverting the tube gently, 200 µl of Lysis I solution and 2 µl of RNase (10mg/ml) was added. The pellet was dissolved and then incubated in ice for 5 minutes. Freshly prepared Lysis II solution (200 µl) was added and mixed by slow inverting. After, addition of ice-cold Lysis III solution (300 µl), the samples were incubated on ice for 10 minutes. Supernatant (500 µl) was then collected in fresh tube and 0.8 volume of isopropanol was added. Centrifugation was carried out at 16060 x g for 15 minutes at 4^0C. the pellet was air dried for 15 minutes at 37^0C. The pellet was then washed with 500 µl of 70% ethanol. Final centrifugation was carried out at 16060 x g for 5 minutes. The pellet was aspirated at 37^0C for 20 minutes and then dissolved in 20 µl of MQ.

Table 3.2. Buffer composition for plasmid isolation via alkaline lysis method.

Buffer	Components
Lysis I solution	50 mM Tris, 10mM EDTA (pH 8.0), 10 mg/ml stock of RNase
Lysis II solution	0.2 N NaOH, 1 % SDS
Lysis III solution	(5 M potassium acetate, glacial acetic acid and distilled water)

3.1.3. VERIFYING THE INTEGRITY OF pRTBV-Inf BY RESTRICTION DIGESTION

Quantification of the freshly as well as previously isolated plasmid was carried out for different dilutions using Nanodrop®. The freshly isolated plasmid was then analysed using restriction digestion by KpnI and *Nco*I (NEB). The reaction set-up components are given in table 3.2. Prepared reaction was kept at 37^0 C for two hours for digestion and then checked on agarose gel (1%).

Table 3.3. Composition of reaction mixture for Restriction digestion

Component	For 2 µl DNA template	For 5 µl DNA template	For 10 µl DNA template
MQ	15 µl	11.5 µl	6 µl
Restriction enzyme	1 µl	1.5 µl	2 µl
Buffer	2 µl	2 µl	2 µl
DNA template	2 µl	5 µl	10 µl

3.1.4. AGROINOCULATION OF RICE WITH RTBV-Inf (FULL LENGTH INFECTIOUS CLONE OF RTBV) AND pRTBV-mVIGS AND THEIR RESPECTIVE *E. coli* CULTURES (CONTROL).

Agrobacterium stock was obtained and streaked on a plate (containing rifampicin + kanamycin) and *E. coli* stock was streaked on a plate having kanamycin and nalidixic acid. The plates were kept for two days at 28^0C in growth chamber. A single colony was picked and dipped in 5ml broth (containing rifampicin + kanamycin). It was then kept for 36 hours on shaker at 28^0C, 200 rpm. Broth (300 ml) was prepared and autoclaved. The broth was then divided into a large (280 ml) and a small lot (20 ml). Agrobacterium primary culture was added to large lot (1%). *E. coli* was added to smaller lot by scrapping

it with toothpick and inserted in 20 ml culture. It was then kept overnight in shaker (Agrobacterium at 28^{0}C, 200 rpm; *E. coli* at 37^{0}C, 200 rpm) for secondary culture preparation. Next day, the culture was divided equally in SS34 tube and centrifuged at 2380 x g at room temperature for 5 mins. The supernatant was discarded and the pellet was dissolved in 2ml resuspension buffer. The solutions were transferred in 2 ml MCTs and dilutions were made to set the OD of 0.6-0.8 at 595 nm. Replicates were made for solutions with appropriate OD. The solution was injected (50 µl) with 1 ml syringe in the crown part of each seedling (2 weeks old). After agroinoculation, the seedlings were kept on the assembly having Whatman paper soaked with RGM in a glass tray. The assembly was kept in growth chamber after covering it with kiln wrap to create a humidity chamber. The plants were incubated for 24 hours. The plants were then washed and transferred to test tubes with fresh RGM and incubated for 15 days.

Table3.4. Composition of resuspension buffer for agroinoculation of rice

Chemical name	**Chemical formula**	**Working concentration**	**50 ml resuspension buffer**
2-(N-morpholino)-ethanesulfonic acid (MES)	$C_6H_{13}NO_4S$	10 mM	500µL from stock
Magnesium Chloride	$MgCl_2$	10 mM	500µL from stock
Acetosyringone	$C_{10}H_{12}O_4$	200 µM	40µL from stock

3.1.5. DNA EXTRACTION FROM AGRO-INOCULATED RICE PLANTS.

DNA extraction was carried out by Puregene® method (Gentra Systems (2004) Puregene® genetic DNA purification kit instructions.).

Approximately 5 cm of leaf was ground using liquid nitrogen. Extraction buffer (600 µl) was added to the powder and thawed. It was then transferred to a microcentrifuge tube or MCT and vortexed. The mixture was then incubated at 65^{0}C for 20 minutes. It was then cooled on ice for 5 minutes. To the tube, 200 µl of 6M ammonium acetate was added and vortexed. Incubation was carried out on ice for 30 minutes. The samples were then centrifuged in Biofuge Fresco benchtop centrifuge at 16060 x g for 10 minutes at 4^{0}C to pellet out the debris. The supernatant was again centrifuged for the second time under similar conditions to pellet any remaining debris. Clean supernatant was transferred into a fresh MCT containing 600 µl of isopropanol to facilitate the precipitation of DNA. DNA precipitation was carried out at -80^{0}C for 30 minutes. Centrifugation was carried out at 16060 x g for 15 minutes at 4^{0}C to obtain a pellet. The pellet thus obtained, was washed with 70% ethanol (500 µl) for 15-20 minutes at 37^{0}C. Centrifugation at 16060 x g for 10 minutes was carried out. Decanting was done slowly to remove the supernatant and the pellet was air dried till no traces of ethanol were left. The air-dried pellet was then suspended in 20 µl of MQ.

3.1.6. CONVENTIONAL END-POINT PCR TO CHECK THE PRESENCE OF VIRAL INFECTION.

Checking the presence of viral infection in rice plants was carried out at 15 days post inoculation (dpi).

Amplification of DNA was carried out using Phusion™ High-Fidelity DNA polymerase using the primers designed specific to Movement Protein (MP). (The details of the primers used are given in appendices). The annealing temperature was optimized according to the melting temperature of the primer pair. The PCR amplicons were checked on 1% agarose gel for amplicon size of 940 bp. The PCR conditions and composition of the reaction are given below.

Table 3.5. Composition of reaction mixture used for PCR amplification.

Components	Volume (µl) required for 1X reaction
DNA template	1
10x Taq buffer	1.5
Forward primer	0.75
Reverse primer	0.75
10 mM dNTPs	0.45
MQ water	10.35
Taq polymerase	0.2
Total reaction volume	15 µl

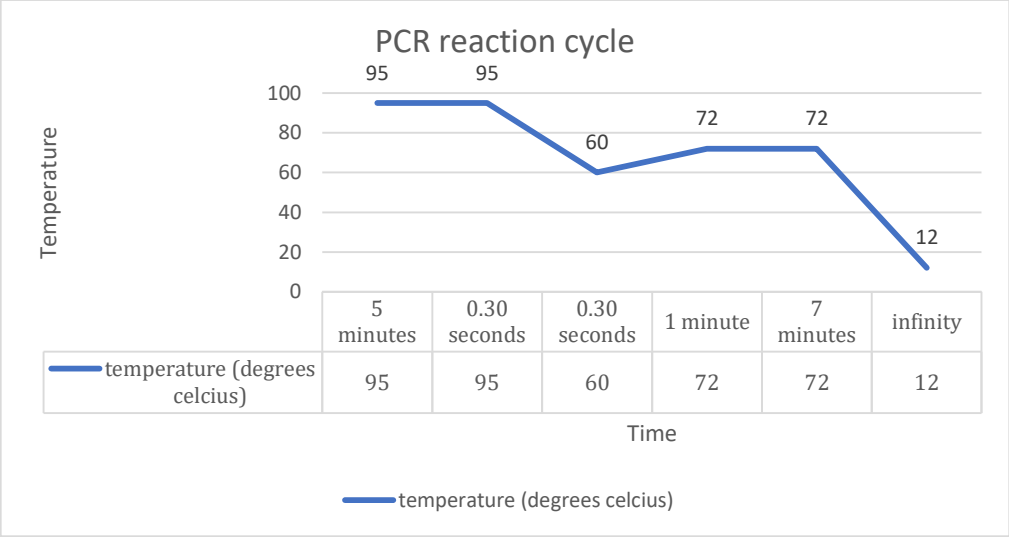


Fig 3.1. The thermal cycle for PCR amplification.

3.2. *N. benthamiana*

3.2.1. PLANT GROWTH

Seeds of *N. benthamiana* were sown in a tray consisting of vermicompost and cocopeat in the ratio 1:3. The tray was covered with kiln wrap and incubated in plant growth room (24 ± 1^0C, 14/10-h light-dark cycle). The seedlings were then transferred to individual pots after 1-2 weeks. Plants in the four-leaf stage were further used for Agroinfiltration.

3.2.2. AGROINFILTRATION OF *N. benthamiana* LEAVES WITH SLCMV and ToLCNDV.

Agrobacterium stocks harboring the plasmids consisting of fragments of SLCMV DNA A (Mittal *et al*., 2008), SLCMV DNA B (Karthikeyan *et al*., 2016). ToLCNDV DNA A and ToLCNDV DNA B (Chakraborty *et al.,* 2003) was obtained and streaked on a plate (containing rifampicin + kanamycin). The plates were kept for two days at 28^0C in growth chamber. A single colony was picked and dipped in 5ml broth (containing rifampicin + kanamycin). It was then kept for 36 hours on shaker at 28^0C, 200 rpm. Broth (400 ml) was prepared and autoclaved. The broth was then divided into four 100 ml flasks. Agrobacterium primary culture for DNA A and DNA B clone was added to one flask each (100 ml) for both SLCMV and ToLCNDV. It was then kept overnight in shaker (28^0C, 200 rpm) for secondary culture preparation. Next day, the culture was divided equally in SS34 tube and centrifuged at 2380 x g at room temperature for 5 mins. The supernatant was discarded and the pellet was dissolved in resuspension buffer. The solutions were transferred in fresh tubes such that four solutions of DNA A, DNA B, DNA A+B, mock (resuspension buffer only) were prepared for SLCMV and ToLCNDV. Dilutions were made to set the OD of 0.6-0.8 at 595 nm and replicates were made for solutions with appropriate OD. The solution was in injected (50 µl) with 1 ml syringe in the abaxial part of at least three leaves in each seedling (4 leaf stage). After Agroinfiltration, the seedlings were kept in the plastic tray and kept in growth room.

Table 3.6. Composition of resuspension buffer for agroinfiltration of *Nicotiana benthamiana*

Chemical name	**Chemical formula**	**Working concentration**	**200 ml resuspension buffer**
2-(N-morpholino)-ethanesulfonic acid (MES)	$C_6H_{13}NO_4S$	10 mM	2 mL from stock
Magnesium Chloride	$MgCl_2$	10 mM	2 mL from stock
Acetosyringone	$C_{10}H_{12}0_4$	100 µM	200µL from stock

Phenotypic evaluation was done after 2 weeks when the symptomatic plants were clearly distinguishable from the non-symptomatic as well as the mock inoculated plants. Pictures were taken for both SLCMV and ToLCNDV infected plants.

3.2.3 DNA ISOLATION FROM *N. benthamiana* LEAVES.

For systemic tissue, leaves from the apical part of the plant were used for isolation of total genomic DNA. Approximately 100 mg of leaf tissue was crushed to fine powder in a mortar-pestle with liquid nitrogen. The powder was transferred to an MCT containing 0.5 ml of extraction buffer. The MCT was then incubated at 65^{0}C for 10 minutes. At the end of incubation period, the MCT was cooled to room temperature and about 200µl of solution III or 5M potassium acetate, pH 5.5, was added and mixed by inverting. The supernatant was collected after centrifuging at 16060 x g, RT for 15 minutes. To this 500µl phenol: chloroform: isoamyl alcohol (PCI) was added and centrifugation was carried out at 16060 x g for 20 minutes. DNA was pelleted by adding 400µl of absolute ethanol followed by centrifugation at 16060 x g. The pellet was washed with 70% ethanol to remove the salts and finally resuspended in 20µl of sterile MQ (Sambrook *et al.*, 1989).

3.2.4. CONVENTIONAL END-POINT PCR TO CHECK THE PRESENCE OF SLCMV INFECTION.

Checking the presence of viral infection in *N. benthamiana* plants was also carried out at 15 days post inoculation (dpi).

Amplification of DNA was carried out using Phusion™ High-Fidelity DNA polymerase using the primers designed specific to AC4 (C4 protein) for detection of SLCMV infection. (The details of the primers used are given in appendices). The annealing temperature was optimized according to the melting temperature of the primer pair. The PCR amplicons were checked on 1% agarose gel for amplicon size of 302 bp. The PCR conditions and composition of the reaction are given below.

Table 3.7. Composition of reaction mixture used for PCR amplification.

Components	Volume (µl) required for 1X reaction
DNA template	1
10x Taq buffer	1.5
Forward primer	0.8
Reverse primer	0.8
10 mM dNTPs	0.3
MQ water	7.5
Taq polymerase	0.2
Total reaction volume	15 µl

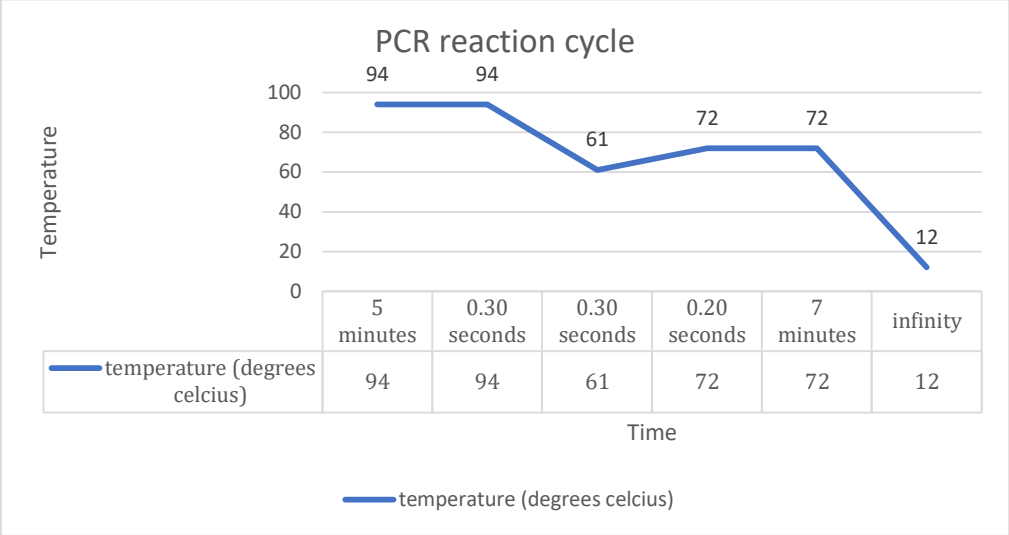


Fig 3.2. The thermal cycle for PCR amplification.

CHAPTER 4
RESULTS AND DISCUSION

4.RESULTS AND DISCUSSION

4.1. RESULTS (RICE)

4.1.1. QUANTIFICATION OF ISOLATED pRTBV-Inf

After plasmid isolation, quantification of DNA was carried out using Nanodrop®. The previously as well as the freshly isolated plasmid was diluted 10 times and the readings were obtained as triplicates.

Table 4.1.1. Results of DNA quantification using Nanodrop®.

Sample	Concentration in ng/µl	Average concentration in ng/ µl	$A_{260/280}$ ratio	$A_{260/230}$ ratio
pRTBV-Inf	4217.3	4827.7	2.22	2.40
	5351.3		2.25	2.41
	4914.5		2.25	2.41
pRTBV-Inf (1/10 dilution)	637.0	638	2.10	2.40
	636.6		2.09	2.40
	640.4		2.09	2.36
pRTBV (previously isolated)	3378.1	3219.93	2.17	2.34
	3104.3		2.17	2.36
	3177.4		2.19	2.37
pRTBV (previously isolated)-1/10 dilution.	476.4	1421.1	2.07	2.27
	469.8		2.07	2.39
	474.9		2.07	2.39

As per the readings obtained, freshly isolated pRTBV-Inf was further used for PCR amplification and restriction digestion.

4.1.2. PCR TO CONFIRM RTBV-Inf CLONE

PCR amplification was carried out to confirm cloned RTBV-Inf at original as well as $1/10^{th}$ dilution in comparison to previously isolated plasmid. Movement protein specific primers were used as per the reaction conditions mentioned in Table 3.4 and figure 3.1. The presence of pRTBV-Inf is indicated by the presence of a band corresponding to 940 bp. This is the amplified product obtained by using the forward and reverse primer specific to RTBV movement protein. The band obtained in Lane3-6 confirms pRTBV-Inf.

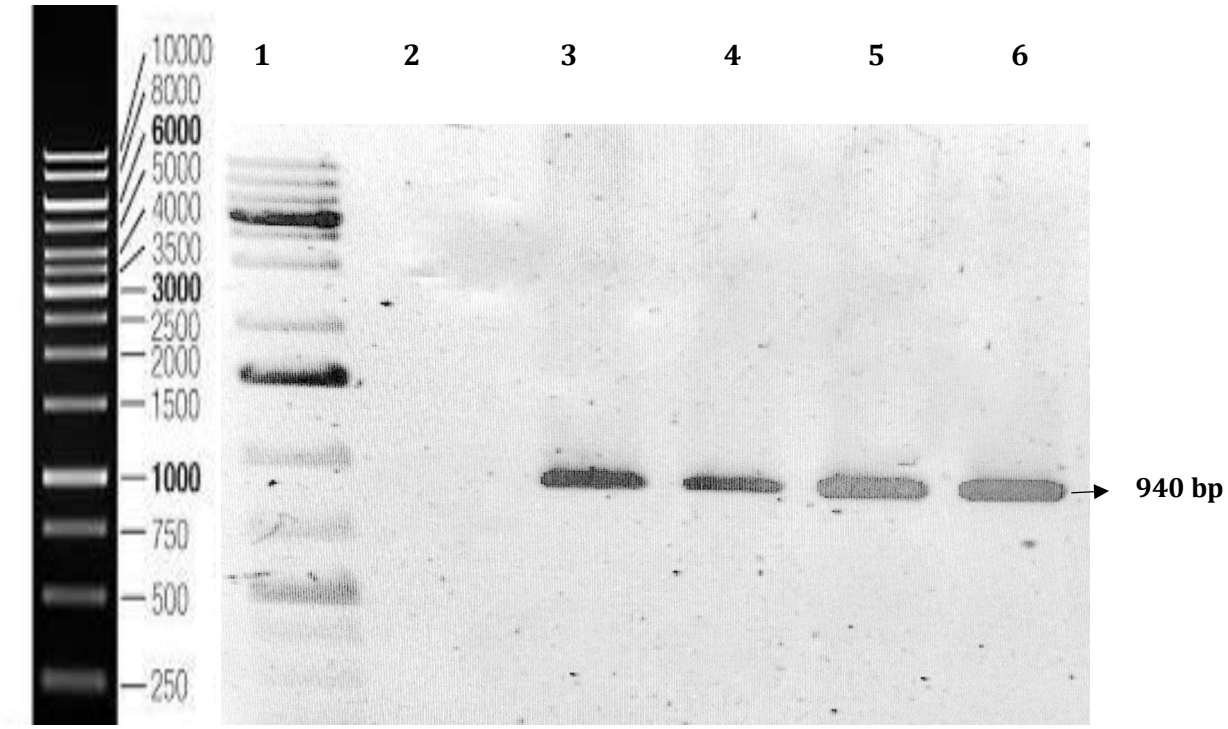


Fig 4.1. Gel profile of PCR-amplified product using movement protein (940 bp) specific primers. Lane 1: GeneRuler™ 1kb ladder, lane 2: Negative control, Lane 3: pRTBV-Inf (freshly isolated), Lane 4: pRTBV-Inf (freshly isolated)-1/10th dilution, Lane 5: pRTBV-Inf (previously isolated), Lane 6: pRTBV-Inf (freshly isolated)-1/10th dilution.

4.1.3. VERIFYING THE INTEGRITY OF pRTBV-Inf BY RESTRICTION DIGESTION.

In order to verify the complete and correct orientation of RTBV-Inf, restriction digestion with KpnI and *Nco*I was performed. The results obtained were in accordance with the restriction map obtained by Fauzia Zareen, PhD thesis, 2015. Therefore, it was concluded that the orientation for the cloned RTBV-Inf was correct and the clone was functional for further experimentation.

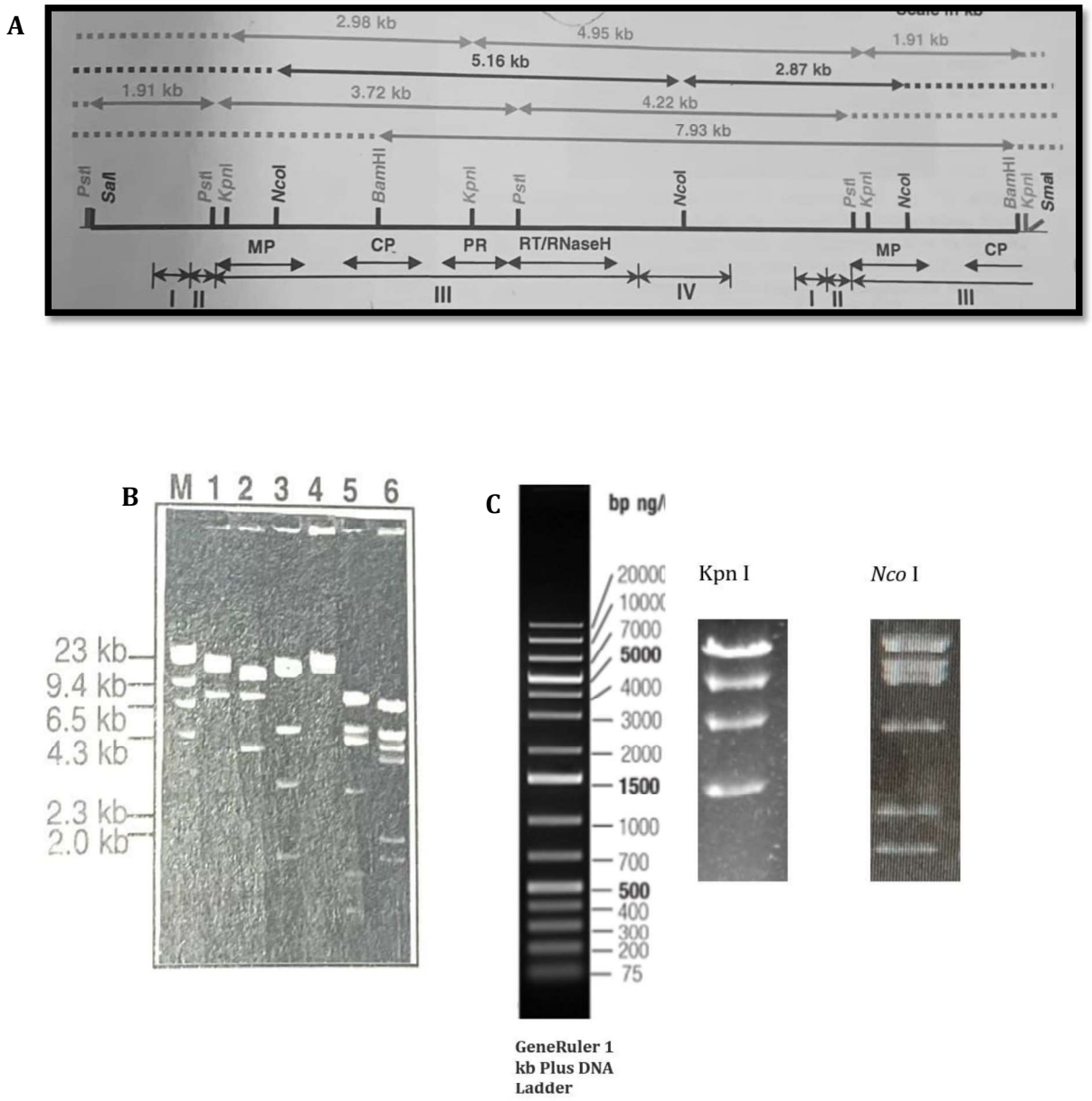


Fig 4.2. A. positions of different restriction enzymes sites in pRTBV-Inf (source: Fauzia Zareen, PhD thesis, 2015). B. Gel profile of pRTBV-Inf for checking the orientation and completeness of the clone; Lane M: λ *Hind*III marker, Lane 1: Digested product from *Bam*HI, Lane 2: *Bam*HI+ *Sal*I, Lane 3: KpnI, Lane 4: *Sma*I, Lane 5: *Nco*I and Lane 6: *Pst*I (source: Fauzia Zareen, PhD thesis, 2015). C. Results from restriction digestion with KpnI and *Nco*I in comparison with GeneRuler™ 1kb plus DNA ladder.

4.1.4. CONVENTIONAL END-POINT PCR TO CHECK THE PRESENCE OF VIRAL INFECTION.

DNA was isolated from the rice plants agro-inoculated with pRTBV-Inf followed by PCR amplification using movement protein specific primers. PCR amplification was

performed as three technical replicates and the results were reproducible. A total of five plants tested positive for the presence of RTBV infection as indicated by the band corresponding to 940 bp.

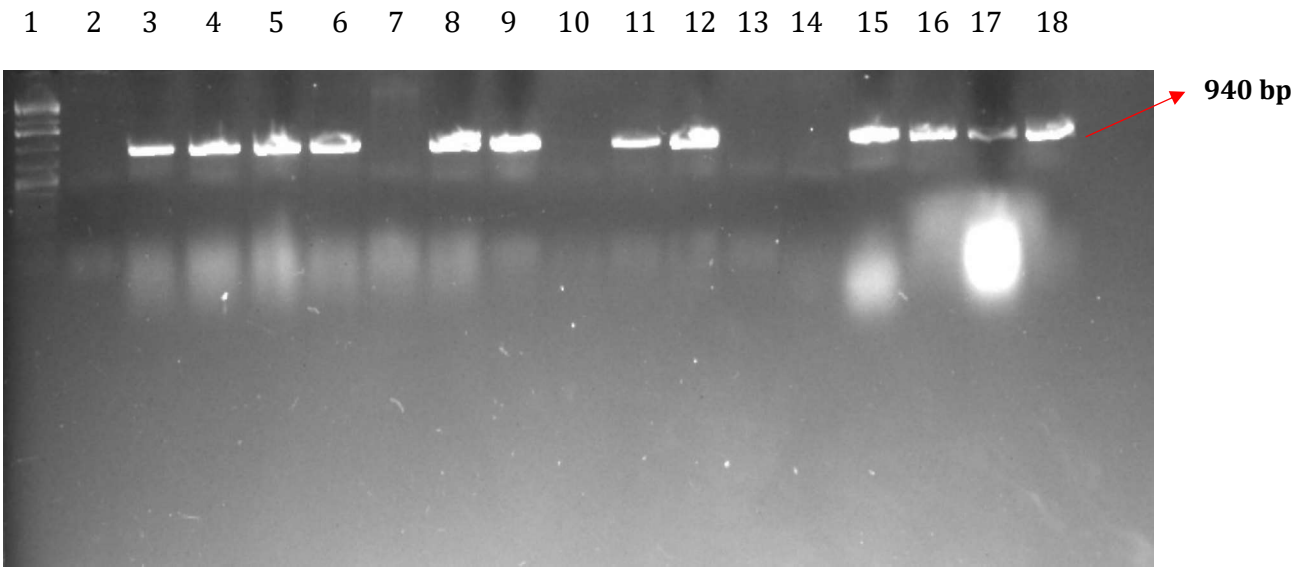


Fig 4.3. Gel profile of PCR amplification of DNA isolated from Agro-inoculated rice plants for checking the presence of RTBV infection. Lane 1; GeneRuler™ 1kb ladder, Lane 2; negative control, Lane 3; sample 1, Lane 4; sample 3, Lane 5; Sample 8, Lane 6; sample 10, Lane 7; sample 18, Lane 8; sample 20, Lane 9: sample 1(1/10th dilution), Lane 10; sample 3(1/10th dilution), Lane 11; Sample 8(1/10th dilution), Lane 12; sample 10(1/10th dilution), Lane 13; sample 18(1/10th dilution), Lane 14; sample 20(1/10th dilution), Lane 15; positive control (freshly isolated pRTBV-Inf as DNA template) Lane 16; positive control (freshly isolated pRTBV-Inf (1/10th dilution)) Lane 17; positive control (previously isolated pRTBV-Inf) Lane 18; positive control (previously isolated pRTBV-Inf (1/10th dilution)) .

The plant tissue from the positive plants was harvested and stored in liquid nitrogen as it can act as a valuable resource for continuing this work.

4.2 RESULTS (*N. benthamiana*)

4.2.1. PHENOTYPIC ANALYSIS OF *N. benthamiana* PLANTS INFECTED WITH SLCMV.

Pictures of mock as well as SLCMV inoculated plants were taken around 2 weeks after agroinoculation. The plants inoculated with SLCMV DNA-A and DNA-B demonstrated reduced height (fig 4.4) and retarded growth similar to plants inoculated with SLCMV DNA-A. It was evident that DNA-A fragment is sufficient for disease development. On the other hand, mock inoculated plants followed the normal pattern of growth and development.

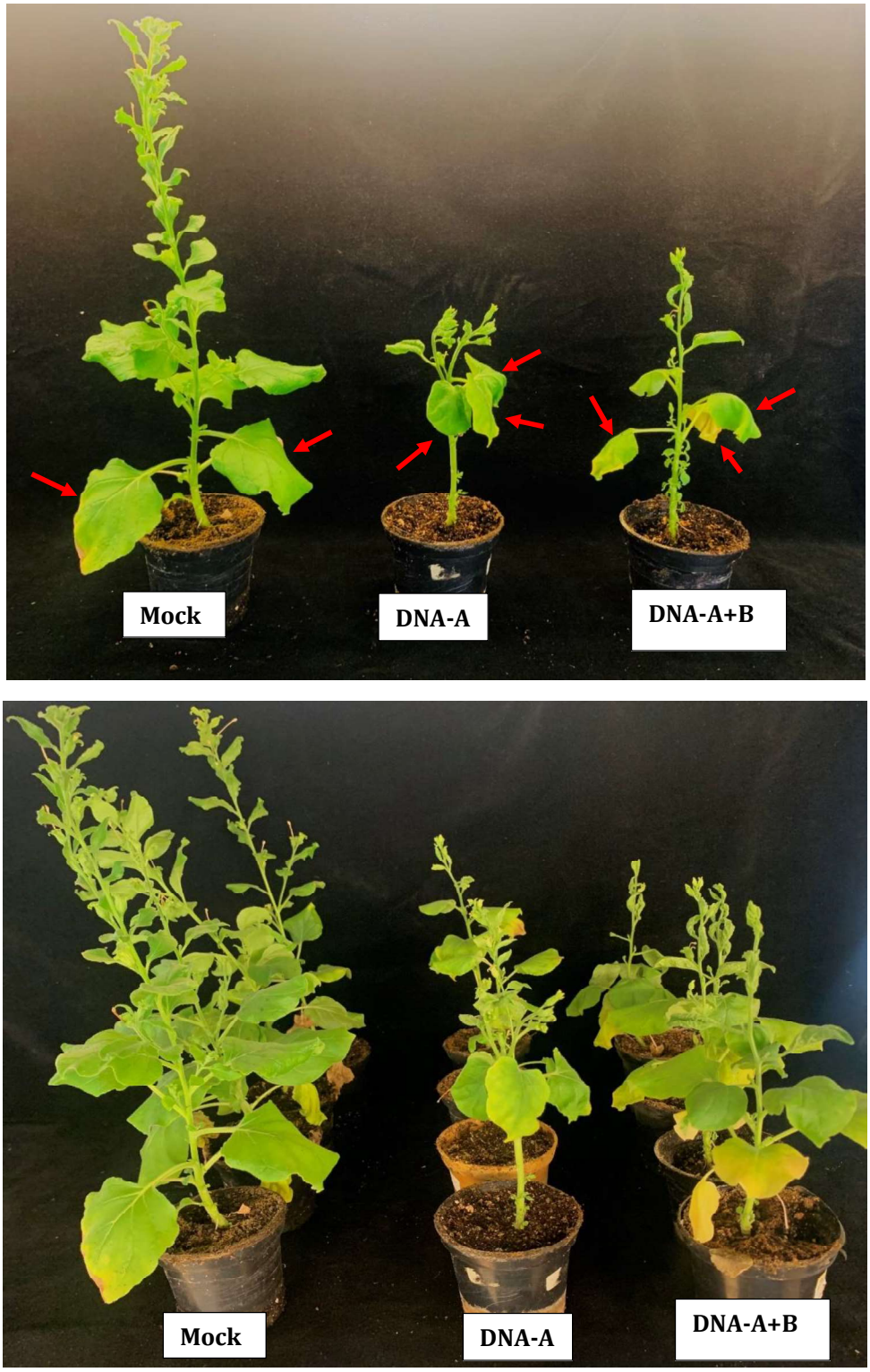


Fig 4.4. Phenotypic evaluation for height in *N. benthamiana* plants upon Sri Lankan Cassava Mosaic Virus infection. Three-week-old plants were agro-inoculated and the pictures were taken fourteen days post inoculation. Red arrow indicates the leaves targeted for inoculation.

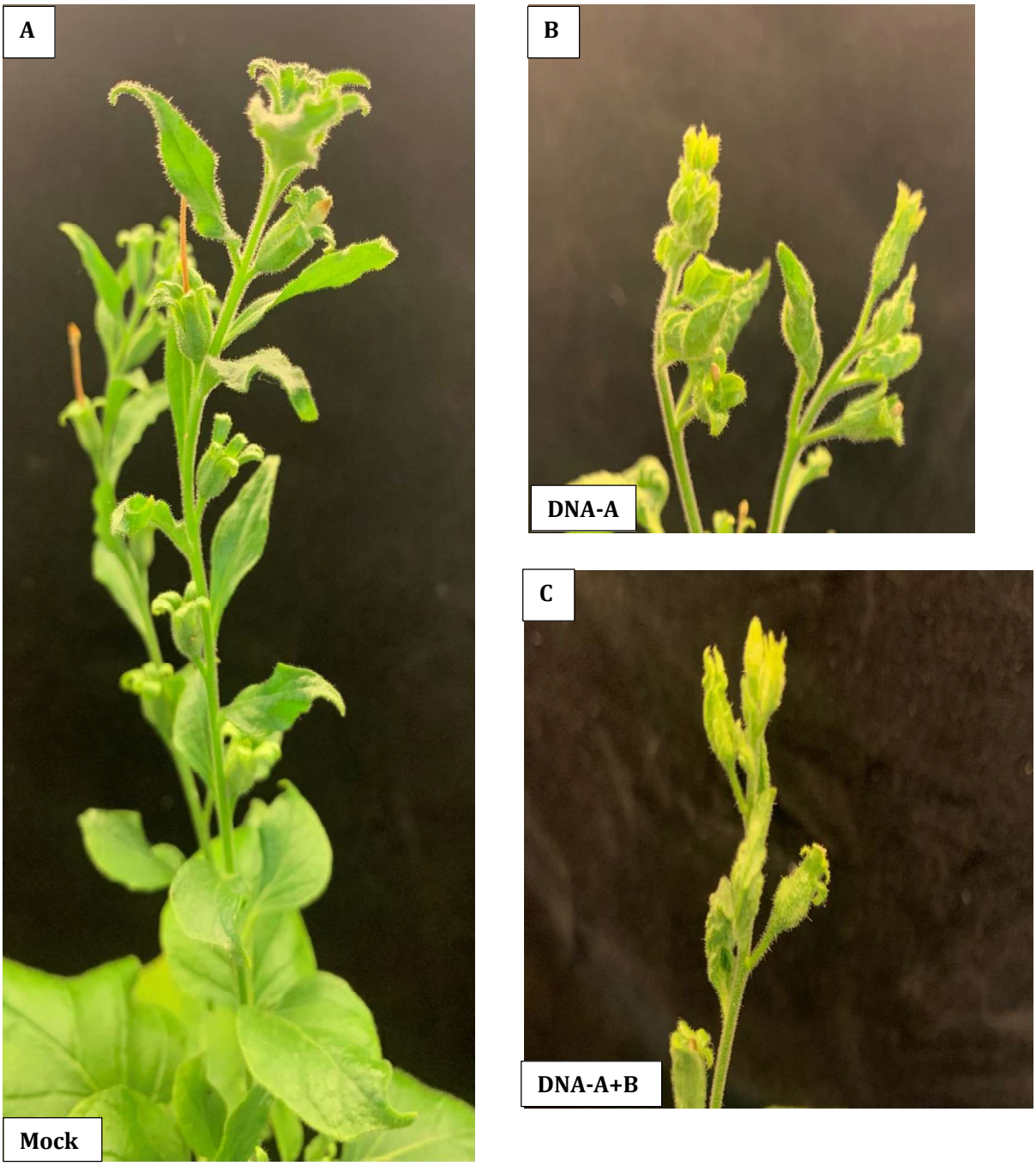


Fig 4.5. Comparison of symptom development in *N. benthamiana* plants upon Sri Lankan Cassava Mosaic Virus infection and mock inoculated plants. Three-week-old plants were agro-inoculated and the pictures were taken fourteen days post inoculation.

Table 4.2.1. List of numbers of plants inoculated with each combination of DNA (SLCMV) as well as mock and numbers showing symptoms after 14 dpi.

DNA combination	No. of inoculated plants	No. of symptomatic plants
Mock inoculated	15	-
DNA-A	15	8
DNA-A+B	15	11

The symptomatic plants could be clearly differentiated from non-symptomatic plants seven days post inoculation as also observed by Mittal *et al*, 2008. Upward as well as downward leaf rolling was observed in plants inoculated with DNA A and DNA B (fig 4.5.C). However, plants inoculated with only DNA-A (fig 4.5.B) dominantly demonstrated the upward leaf rolling pattern.

4.2.2. CONVENTIONAL END-POINT PCR TO CHECK THE PRESENCE OF SLCMV INFECTION.

Phenotypic evaluation revealed that not all the leaves developed symptoms in infected plants. So, DNA isolation was carried out for symptomatic as well as non-symptomatic leaves from the same plant inoculated with SLCMV DNA-A as well as SLCMV DNA-A+B. The results from PCR amplification using DNA-A specific primers (AC4) clearly indicated the presence of viral DNA in both, symptomatic and non-symptomatic leaves as shown in fig 4.6. Therefore, phenotypic evaluation alone is not sufficient for detection of SLCMV infection.

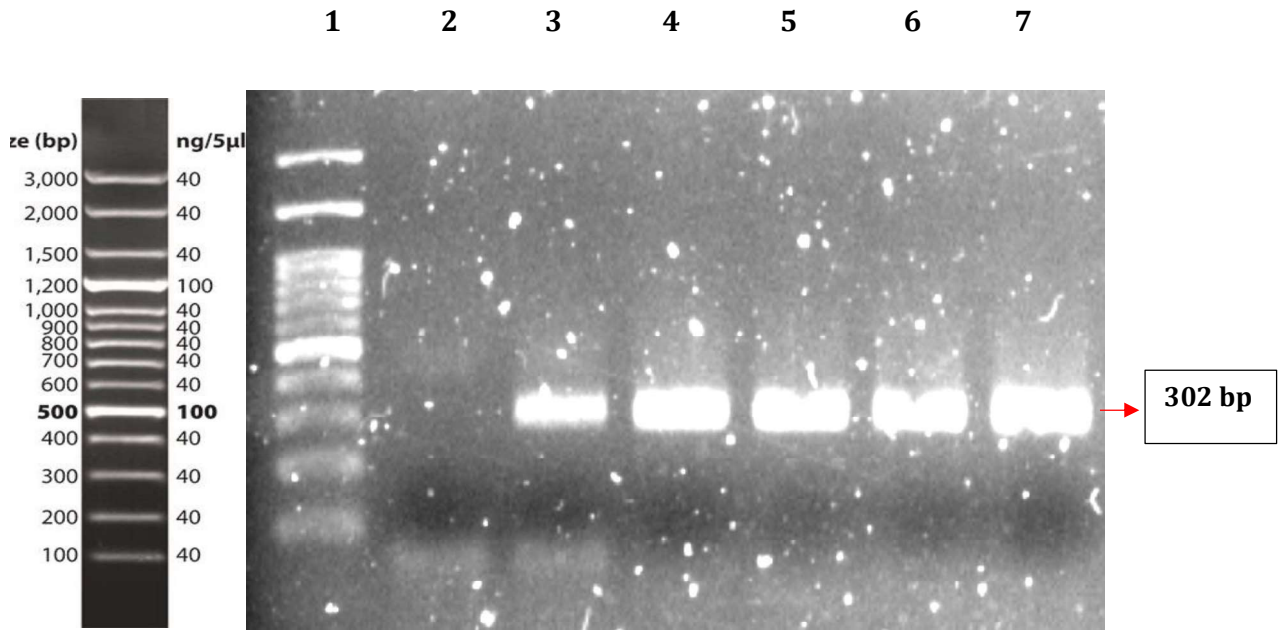


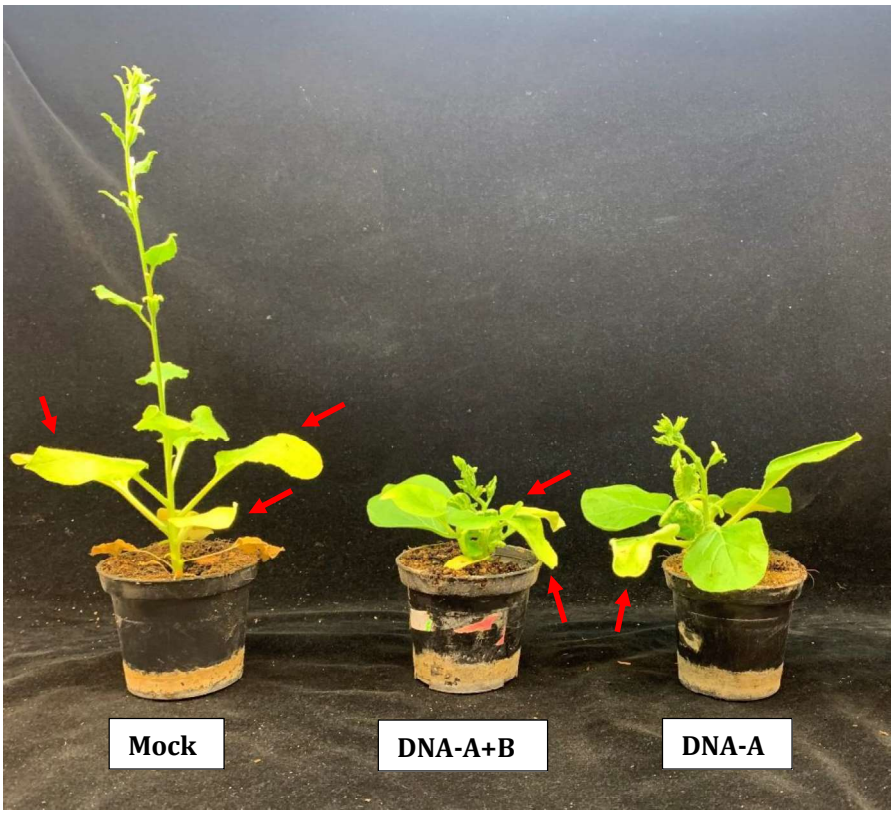
Fig 4.6. Gel profile of PCR amplification to check the presence of SLCMV infection *N. benthamiana* plants. Lane 1; 100 bp ladder, Lane 2; negative control, Lane 3; DNA template from symptomatic plant infected with only Sri Lankan Cassava Mosaic Virus DNA-A, Lane 4; DNA template from non-symptomatic plant infected with only DNA-A, Lane 5; DNA template from symptomatic plant infected with DNA-A and DNA-B, Lane 6; DNA template from non-symptomatic plant infected with DNA-A and DNA-B, Lane 7; positive control.

The leaf tissue from the positive tested plants for SLCMV infection were harvested and stored in liquid nitrogen for future use.

4.2.3. PHENOTYPIC ANALYSIS OF *N. benthamiana* PLANTS INFECTED WITH ToLCNDV.

The phenotypic analysis for mock inoculated and ToLCNDV inoculated plants was also carried out two weeks after agroinoculation. All the plants inoculated with DNA-A and DNA-A+B developed severe disease symptoms within 7 days of inoculation. In addition, the infectivity was similar in both the combinations indicating that ToLCNDV-A fragments can cause the viral infection individually as well as in combination with ToLCNDV-B (fig 4.7). The plants exhibited the characteristic symptoms of upward leaf rolling and stunted growth as shown in fig 4.8.

A



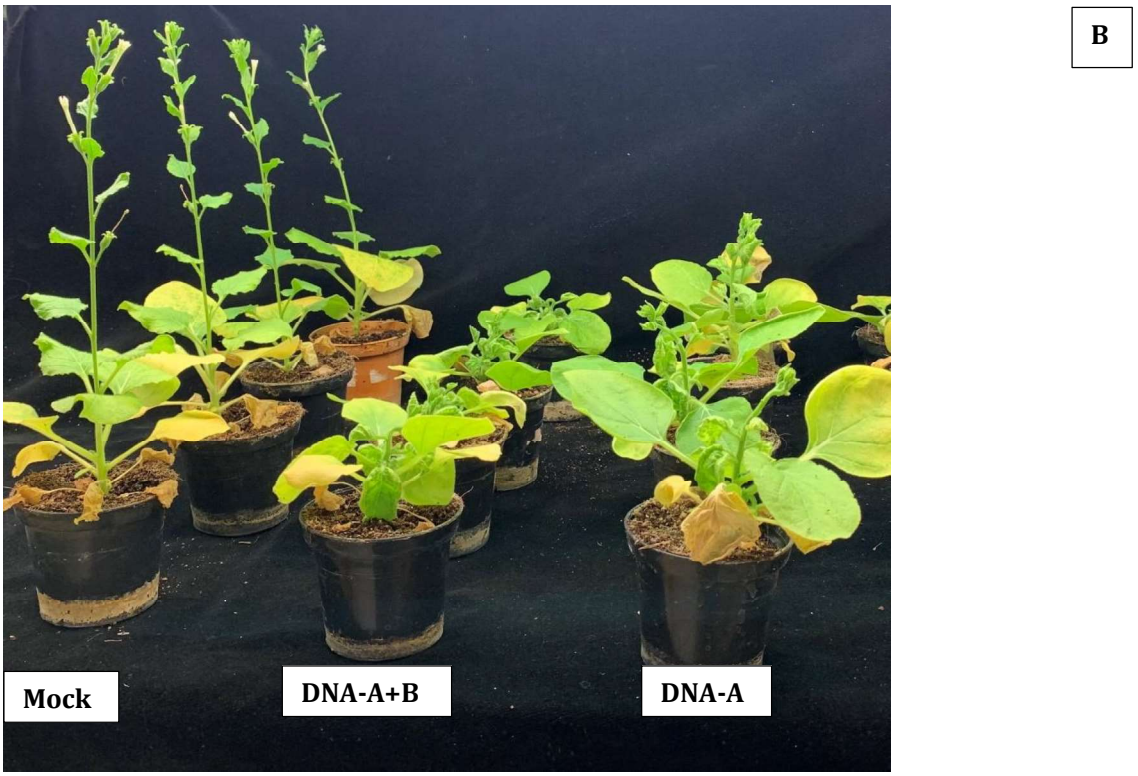


Fig 4.7. Phenotypic evaluation for height in *N. benthamiana* plants upon Tomato Leaf Curl New Delhi Virus infection. Three-week-old plants were agro-inoculated and the pictures were taken fourteen days post inoculation. Red arrow indicates the leaves targeted for inoculation.

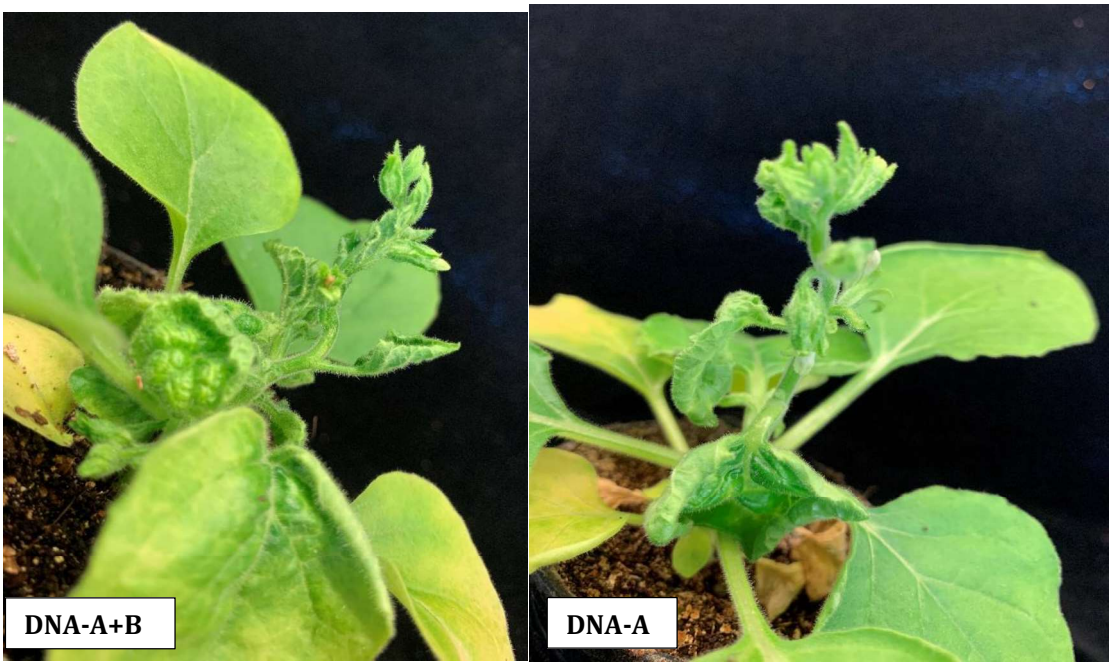


Fig 4.8. Symptom development in *N. benthamiana* plants upon Tomato Leaf Curl New Delhi Virus infection after fourteen days post inoculation.

Table 4.2.2. List of numbers of plants inoculated with each combination of DNA (ToLCNDV) as well as mock and numbers showing symptoms after 14 dpi.

Experiment number	DNA combination	No. of inoculated plants	No. of symptomatic plants
Experiment 1	Mock inoculated	15	-
	DNA-A	15	15
	DNA-A+B	15	15
Experiment 2	Mock inoculated	15	-
	DNA-A	15	13
	DNA-A+B	15	15

Leaves were harvested from mock inoculated and symptomatic plants for further confirmation of ToLCNDV infection via conventional end-point PCR. Plants testing positive for infection were used a source of leaf tissue for future work.

4.3. DISCUSSION

Based on the phenotypic analysis for *N. benthamiana* plants infected with different combinations of ToLCNDV DNA, it was observed that all the plants infected with DNA-A as well as DNA-A+B exhibited disease symptoms. However, the infection can be further confirmed by conventional end-point PCR using ToLCNDV specific primers. Although it would not be improper to expect viral infection in plants infected with DNA-A, as it has been previously reported (Jyothsna *et al*., 2013) that similar to DNA-A fragment in SLCMV, the DNA-A fragment of ToLCNDV alone, is capable of causing infection in *N. benthamiana.* On the other hand, RTBV infection in rice and SLCMV infection in *N. benthamiana* were confirmed by conventional end-point PCR using movement protein and AC4 specific primers respectively. Further confirmation of viral infection can be done using RT-PCR which would also provide a considerable correlation between the infection and virus titres.

There are several reports on the transcriptomic changes in plants upon viral infections. The differentially regulated mRNAs can be the potential targets for miRNA and therefore a direct correlation between the two exists. In a study conducted by Zarreen *et al*. (2018), a total of previously known 17 osa-miRNAs were analyzed using qRT-PCR upon RTBV and RTSV infection in rice. The aim of this study was to analyze the differential expression of the 17 target mRNAs:

Table 4.3.1. List of the miRNAs and their corresponding target mRNAs to be analyzed by qRT-PCR (source: Zarreen *et al*. (2018))

miRNA	Target gene name
MIR159e	TCP family transcription factor containing protein, expressed
MIR159e	UDP-glucoronosyl and UDP-glucosyl transferase family protein
MIR167h	HEAT repeat family protein, expressed

MIR164f	BURP domain containing protein, expressed
MIR815b	OsWAK129a - OsWAK receptor-like protein kinase
MIR1439	NAD dependent epimerase/dehydrate
MIR1439	CUE domain containing protein
MIR1439	KHdomain containing protein
MIR171b	SCARECROW gene regulator, putative, expressed
miR818c	OsWRKY34 - Superfamily of rice TFs having WRKY and zinc finger domains, expressed
miR818c	Myb-like DNA-binding domain, SHAQKYF class family protein, expressed
miR444e	Beige/BEACH domain containing protein, expressed
miR812e	Protein phosphatase 2C containing protein
miR156a	SBP domain containing protein
miR172a	AP2 domain containing protein
miR815a	Exocyst complex subunit Sec15-like family protein
miR408	OsIAA6 - Auxin-responsive Aux/IAA gene family member

A common set of target genes have been reported to be differentially expressed upon virus infection in rice. A similar study was conducted by Ahmed *et al* (2017), which illustrated the upregulation of rice wall-associated kinase (*OsWAK32*) upon Rice black-streaked dwarf virus (RBSDV) infection in rice. The wall-associated kinase (WAK) gene family plays an important role in pathogen resistance and its regulation can have a direct consequence on plant-virus interactions. Similarly, differentially regulated genes upon RSV infection in rice can be divided into 13 major categories (Satoh *et al*., 2009):

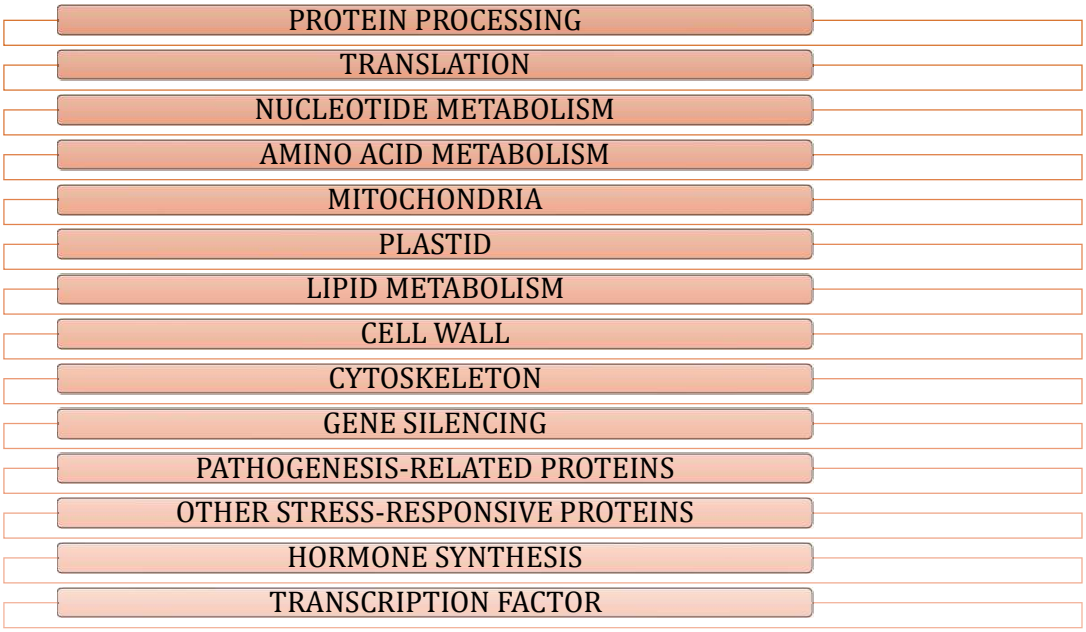


Fig 4.3.1. List of categories of differentially regulated genes upon RSV infection (adapted from Satoh *et al*., 2009).

This categorization is based on the conserved developmental pathways in plants which are affected during the viral attack. Various components of these pathways are regulated to meet the high metabolic demand during a pathogenic attack. In addition, transcription factors have also been reported to be adversely affected causing alterations in the plant developmental processes. For example, MYB transcription factors are induced in rice plants upon RBSDV infection (Ahmed *et al*., 2017) as well as upon RSV infection (Satoh *et al*., 2009). Other genes related to the process of transcription like DNA binding protein have also been reported to be upregulated in plants like Arabidopsis during Tobacco mosaic virus (TMV) infection (Golem *et al*., 2003) and also during SLCMV infection (Pierce *et al*., 2013). This can directly influence the transcription of various essential genes.

On the other hand, chloroplast and mitochondria related genes are involved in the major energy yielding pathways of plants for example, genes involved in TCA cycle, oxidative phosphorylation, photosynthesis, Porphyrin and chlorophyll metabolism. To cope up with the virus infiltration, these genes are usually upregulated as was observed in rice upon RSV infection (Satoh *et al*., 2009). In fact, the same virus infecting a different host exhibits similar pattern of transcriptome regulation. Arabidopsis plants infected with RSV followed a similar pattern of transcriptional reprogramming. Highest upregulation was observed in case of metabolism related genes whereas, some energy related genes were either downregulated or upregulated. In addition, several genes differentially expressed in rice upon RSV infection were also identified in Arabidopsis plants infected with the same virus. These genes were involved in defence signalling, secondary metabolism and post translational modifications. For example, the defence related transcript, WAK1 (cell wall-associated kinase) was upregulated during the initial period of RSV infection in Arabidopsis (Sun *et al*., 2016). An indistinguishable pattern of transcriptional repression has also been reported for other dicot model plants such as *N. benthamiana* for other defence related genes involved in terpenoid biosynthesis upon

TYLCV (Wu *et al.*, 2019), Tobacco curly shoot virus (TbCSV) (*Li et al., 2018)* and Tomato yellow leaf curl China virus infection (TYLCCNV) (Luan *et al*., 2013).

Table 4.3.2. Suppression of genes involved in terpenoid biosynthesis upon infection with different viruses in tobacco plant.

Plant system	Infecting virus	Gene description	Reference
N. benthamiana	TbCSV	- 1-deoxy-D-xylulose-5-phosphate reductoisomerase	Li *et al*., 2018
N. benthamiana	TYLCV	- 1-deoxy-D-xylulose-5-phosphate reductoisomerase - 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase - 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase - Geranylgeranyl diphosphate synthase - 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase - Isopentenyl-diphosphate Delta-isomerase II	Wu *et al*., 2019
N. tabacum	TYLCCNV	- 1-deoxy-D-xylulose-5-phosphate reductoisomerase - 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase - Isopentenyl diphosphate isomerase 1 - Monoterpene synthase 2 - 5-epi-aristolochene synthase	Luan *et al*., 2013

The symptoms associated with most plant infecting viruses include deformity and stunted growth. This could be due to downregulation of cellular structure and cell elongation-related genes (specifically the cytoskeleton related genes). In fact, plant viruses belonging to different genera can lead to alterations of a particular set of genes to induce similar symptoms in host. This claim has been validated in tomato upon TYLCV and tomato chlorosis virus (ToCV) infection. In particular, both the viral infections led to the downregulation of differentially expressed genes related to photosynthesis, thylakoid and thylakoid membrane. Transcriptional repression of these genes has been anticipated to be responsible for chlorosis induced by ToCV and yellowing caused by TYLCV infection. But this observation is not absolute as leaf senescence is also a common symptom of both these viral infections however, it is induced by regulation of different mechanisms during TYLCV (majorly by enrichment of hormone signalling) and ToCV infection (Seo *et al*., 2018).

Table. 4.3.3. List of some important photosynthesis related genes, downregulated in tomato upon ToCV infection (source: Seo *et al*., 2018).

Gene description	**Role**
Carbonic anhydrase 1	Essential for photosynthesis as it reversibly coverts CO_2 into HCO^{3-}. It also regulates closing of stomata.
Photosystem 1 light harvesting complex gene 3	Downregulation of chloroplast- and photosynthesis-related genes induces chlorosis by activation of retrograde signalling pathway.
Light-harvesting chlorophyll B-binding protein 3	
Photosystem II light harvesting complex gene 2.1	
NADH-dependent cyclic electron flow 1	
Photosystem II 5kD protein	
Chlorophyll A/B binding protein 1	Chlorophyll binding and light harvesting.

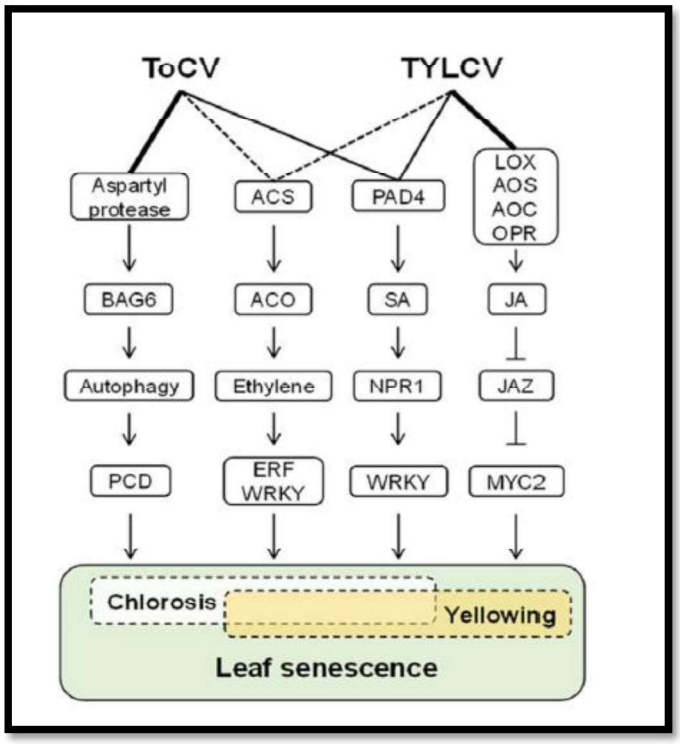


Fig. 4.3.2. Pathways indicating a link between senescence, chlorosis and yellowing in tomato upon TYLCV and ToCV (source: Seo *et al*., 2018).

Table. 4.3.4. List of some important hormone signalling related genes, regulated in tomato upon ToCV and TYLCV infection (source: Seo *et al*., 2018).

Gene description	Infecting virus	Regulation	Role
Tomato homolog of cytokinin oxidase	TYLCV	Upregulated	It catalyses degradation of cytokinin which is necessary for plant growth and differentiation leading to senescence.
1-aminoacyclopropane 1-carboxylate synthase (ACC synthase 1)	TYLCV and ToCV	Downregulated	It catalyses rate limiting step in ethylene biosynthesis.
WRKY transcription factor genes	TYLCV and ToCV	Upregulated	Ethylene synthesis

As gene silencing and hormone signalling are the key regulators of plant defence, the genes associated with these processes are universally affected in plants upon viral attack.

Table 4.3.6. List of some important genes related to gene silencing which show alteration in expression upon RSV infection in rice (Satoh *et al*., 2009).

Gene description	No of genes
Argonaute gene family	17
Dicer-like group	5
Double-stranded RNA-binding protein Group	5
HUA enhancer	1
RNA polymerase IV large subunit	2
RNA polymerase IV small subunit	2
RNA-dependent RNA polymerases	5
Suppressor of gene silencing	2

The hormone signalling genes can either be upregulated as in case of auxin signalling genes in *N. benthamiana* during TYLCV infection (Wu *et al.*, 2019) or downregulated, as in case of ethylene signaling related genes in rice upon RSV infection (Satoh *et al*., 2009) and in *N. benthamiana* during TYLCV infection (Wu *et al.*, 2019).

In order to study the regulation of other plant defense genes during viral infection, comparative transcriptome analysis has been carried out in resistant (CLN2777A) as well as susceptible (TMXA48-4-0) tomato plant varieties against TYLCV infection by whole transcriptome sequencing. A significant downregulation of defense genes like WRKY transcriptional factors, R genes, protein kinases and receptor (-like) kinases has been observed in susceptible (S) tomato line upon TYLCV infection. This is the probable reason for the severe infection in S line in comparison to the resistant line which exhibits either upregulation or unaltered expression of these genes upon viral attack (Chen *et al*., 2013).

Table 4.3.6. List of putative defence related genes from the up-regulated DEGs in the R line.

Gene description	log2 fold change	Function	Reference
Ethylene response factor 4	5.97883	Defence response	Mase *et al*., 2013. Sahu *et al*., 2012. Son *et al*., 2012.
Ethylene response factor 5	4.08545	Defence response	Mase *et al*., 2013. Sahu *et al*., 2012.

			Son *et al*., 2012.
Ethylene-responsive transcription factor erf061-like	4.697	Defence response	Mase *et al*., 2013. Sahu *et al*., 2012. Son *et al*., 2012.
Cellulose synthase-like protein g3-like	4.37788	Cell wall formation	Schardl *et al*., 2013.
Bahd acyltransferase dcr-like	6.88485	Pathogen defense	Zheng *et al*., 2009.
F-box protein pp2-b15-like	4.36437	Ubiquitination during plant immune signaling	Marino *et al*., 2012. Furlan *et al*., 2012.
Flavonol synthase flavanone 3	4.68864	Soybean mosaic virus resistance	Cheng *et al*., 2010.
Hydroxyproline-rich glycoprotein family protein	4.49604	Plant defense against pathogen infection	Corbin *et al*., 1987.
Nac transcription factor 29-like	7.82817	Regulating jasmonic acid-signaled defense responses	Jiang *et al*., 2009.
Osmotin-like protein	6.85817	Blocking the growth of *Phomopsis viticola* and *Botrytis cinerea* mycelia	Monteiro *et al*., 2003.
Probable inactive purple acid phosphatase 27-like	4.18825	Basal resistance against *Pseudomonas syringae*	Ravichandran *et al*., 2013.
Small heat-shock	6.20304	Specific restriction of tobacco Etch virus	Whitham *et al*., 2000.
Tyramine n-feruloyltransferase 4 11-like	6.56522	Synthesis of phenolic compounds	von Roepenack-Lahaye *et al*., 2003.

Based on the numerous reports exhibiting a similar transcriptional regulation in host upon viral infection, it is not improper to conclude that diverse viruses can infect diverse plant hosts as a consequence of transcriptional reprogramming of a common set of genes. Despite the taxonomical differences in viruses as well as the host plants, the differentially expressed genes can be categorised into following major functions; protein synthesis system, transcription related genes, organellar function, cell structure, defence (hormone signalling and gene silencing), metabolic processes (plastid and mitochondria related genes), pathogenesis related genes, etc. This classification is based on the conservation of pathways responsible for defence response in plants against various viruses.

CHAPTER 5
CONCLUSIONS AND FUTURE WORK

5.1. CONCLUSIONS

This dissertation is aimed at further validating the hypothesis that there are evident commonalities between the differentially expressed genes in plants upon viral infection due to conservation in defence related pathways. The major objective was to derive a correlation between the expression levels of 17 previously known miRNA-targets (table 4.3.1) upon infection by rice tungro virus complex (RTBV+RTSV) (Zarreen *et al*., 2018), and infection by RTBV alone in rice. This would have demonstrated the similarities as well as the differences in transcriptional regulation during RTBV+RTSV and RTBV infection. In addition, indirect analysis of previously characterized miRNAs could be achieved by analyzing the change in expression of their target genes.

This study also aimed at investigating whether a similar pattern of transcriptional reprogramming takes place by related geminiviruses (SLCMV and ToLCNDV) with respect to the results obtained by Wu *et al.* (2019), in *N. benthamiana* upon TYLCV infection. The rice plants were successfully agroinoculated with RTBV-inf clone and the *N. benthamiana* plants were agroinfiltrated with SLCMV and ToLCNDV. The viral infection was confirmed using conventional end-point PCR. The positive-tested plant samples were stored in liquid nitrogen for further use.

5.2. FUTURE WORK

In the future, after harvesting the plant samples tested positive for viral infections, RNA isolation may be carried out using one-step RNA reagent. Subsequently, the expression levels of the selected mRNAs upon RTBV infection in rice as well as SLCMV and ToLCNDV infection in *N. benthamiana* could be analysed using qRT-PCR in correlation with the study conducted by Zarreen *et al*. (2018) and Wu *et al.* (2019). The hypothesis of similarities in the expression levels of a specific set of genes during plant-virus interactions needs to be tested and future work in this direction may be able to establish this possibility.

CHAPTER 6
BIBLIOGRAPHY

6. BIBLIOGRAPHY

Ahmed, M. M., Ji, W., Wang, M., Bian, S., Xu, M., Wang, W. & Yu, H. (2017). Transcriptional changes of rice in response to rice black-streaked dwarf virus. *Gene*, ***628***, 38-47.

Amin, I., Patil, B. L., Briddon, R. W., Mansoor, S., & Fauquet, C. M. (2011). A common set of developmental miRNAs are upregulated in *Nicotiana benthamiana* by diverse begomoviruses. *Virology Journal*, ***8***(1), 1-9.

Bally, J., Jung, H., Mortimer, C., Naim, F., Philips, J. G., Hellens, R., Waterhouse, P. M. (2018). The rise and rise of *Nicotiana benthamiana*: a plant for all reasons. *Annual review of phytopathology*, **56**:405-426.

Birnboim, H. C., & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids research*, ***7***(6), 1513-1523.

Briddon, R. W., Pinner, M. S., Stanley, J., & Markham, P. G. (1990). Geminivirus coat protein gene replacement alters insect specificity. *Virology*, ***177***(1), 85-94.

Chakraborty, S., Pandey, P. K., Banerjee, M. K., Kalloo, G., & Fauquet, C. M. (2003). Tomato leaf curl Gujarat virus, a new begomovirus species causing a severe leaf curl disease of tomato in Varanasi, India. *Phytopathology*, ***93***(12), 1485-1495.

Chellappan, P., Vanitharani, R., Ogbe, F., & Fauquet, C. M. (2005). Effect of temperature on geminivirus-induced RNA silencing in plants. *Plant Physiology*, ***138***(4), 1828-1841.

Chen, T., Lv, Y., Zhao, T., Li, N., Yang, Y., Yu, W. & Zhang, B. (2013). Comparative transcriptome profiling of a resistant vs. susceptible tomato (*Solanum lycopersicum*) cultivar in response to infection by tomato yellow leaf curl virus. *PloS one*, ***8***(11), e80816.

Cheng, H., Yang, H., Zhang, D., Gai, J., & Yu, D. (2010). Polymorphisms of soybean isoflavone synthase and flavanone 3-hydroxylase genes are associated with soybean mosaic virus resistance. *Molecular Breeding*, ***25***(1), 13-24.

Corbin, D. R., Sauer, N. O. R. B. E. R. T., & Lamb, C. J. (1987). Differential regulation of a hydroxyproline-rich glycoprotein gene family in wounded and infected plants. *Molecular and Cellular Biology*, ***7***(12), 4337-4344.

Du, J., Wu, G., Zhou, Z., Zhang, J., Li, M., Sun, M., Qing, L. (2019). Identification of microRNAs regulated by tobacco curly shoot virus co-infection with its betasatellite in *Nicotiana benthamiana*. *Virology journal*, ***16***(1), 1-12.

Fattash, I., Voß, B., Reski, R., Hess, W. R., & Frank, W. (2007). Evidence for the rapid expansion of microRNA-mediated regulation in early land plant evolution. *BMC plant biology*, ***7***(1), 1-19.

Fondong, V. N. (2013). Geminivirus protein structure and function. *Molecular plant pathology*, ***14***(6), 635-649.

Furlan, G., Klinkenberg, J., & Trujillo, M. (2012). Regulation of plant immune receptors by ubiquitination. *Frontiers in plant science*, ***3***, 238.

Gan, Q. H., Chi, X. Y., & Qin, S. (2008). Roles of microRNA in plant defense and virus offense interaction. *Plant cell reports*, ***27***(10), 1571-1579.

Garcia-Ruiz, H., Takeda, A., Chapman, E. J., Sullivan, C. M., Fahlgren, N., Brempelis, K. J., & Carrington, J. C. (2010). Arabidopsis RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip Mosaic Virus infection. *The Plant Cell*, ***22***(2), 481-496.

Gentra Systems (2004) Puregene® genetic DNA purification kit instructions.

Golem, S., & Culver, J. N. (2003). Tobacco mosaic virus induced alterations in the gene expression profile of *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions*, ***16***(8), 681-688.

Goodin, M. M., Zaitlin, D., Naidu, R. A., & Lommel, S. A. (2008). *Nicotiana benthamiana*: its history and future as a model for plant–pathogen interactions. *Molecular plant-microbe interactions*, ***21***(8), 1015-1026.

Guerra-Peraza, O., Kirk, D., Seltzer, V., Veluthambi, K., Schmit, A. C., Hohn, T., & Herzog, E. (2005). Coat proteins of Rice tungro bacilliform virus and Mungbean yellow mosaic virus contain multiple nuclear-localization signals and interact with importin α**.** *Journal of general virology*, ***86***(6), 1815-1826.

Guo, Y., Jia, M. A., Yang, Y., Zhan, L., Cheng, X., Cai, J., Shamsi, I. H. (2017). Integrated analysis of tobacco miRNA and mRNA expression profiles under PVY infection provids insight into tobacco-PVY interactions. *Scientific reports*, ***7***(1), 1-10.

Hagen, L. S., Jacquemond, M., Lepingle, A., Lot, H., & Tepfer, M. (1993). Nucleotide sequence and genomic organization of cacao swollen shoot virus. *Virology*, ***196***(2), 619-628.

Hanley-Bowdoin, L., Settlage, S. B., Orozco, B. M., Nagar, S., & Robertson, D. (1999). Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. *Critical Reviews in Plant Sciences*, ***18***(1), 71-106.

Hay, J., Grieco, F., Druka, A., Pinner, M., Lee, S. C., & Hull, R. (1994). Detection of rice tungro bacilliform virus gene products in vivo. *Virology*, ***205***(2), 430-437.

Herzog, E., Guerra-Peraza, O., & Hohn, T. (2000). The rice tungro bacilliform virus gene II product interacts with the coat protein domain of the viral gene III polyprotein. *Journal of virology*, ***74***(5), 2073-2083.

Hibino, H. (1983). Transmission of two rice tungro-associated viruses and rice waika virus from doubly or singly infected source plants by leafhopper vectors. *Plant Disease*, ***67***(7), 774-777.

Hibino, H., Roechan, M., & Sudarisman, S. (1978). Association of two types of virus particles with penyakit habang (tungro disease) of rice in Indonesia. *Phytopathology*, ***68***(10), 1412-1416.

Hull, R. (1996). Molecular biology of rice tungro viruses. *Annual review of phytopathology*, ***34***(1), 275-297.

Jiang, H., Li, H., Bu, Q., & Li, C. (2009). The RHA2a-interacting proteins ANAC019 and ANAC055 may play a dual role in regulating ABA response and jasmonate response. *Plant signaling & behavior*, ***4***(5), 464-466.

Jyothsna, P., Haq, Q. M. I., Singh, P., Sumiya, K. V., Praveen, S., Rawat, R., & Malathi, V. G. (2013). Infection of tomato leaf curl New Delhi virus (ToLCNDV), a bipartite begomovirus with betasatellites, results in enhanced level of helper virus components and antagonistic interaction between DNA B and betasatellites. *Applied microbiology and biotechnology*, ***97***(12), 5457-5471.

Karthikeyan, C., Patil, B.L., Borah, B.K., Resmi, T.R., Turco, S., Pooggin, M.M., Hohn, T. and Veluthambi, K., (2016). Emergence of a latent Indian cassava mosaic virus from cassava which recovered from infection by a non-persistent Sri Lankann cassava mosaic virus. *Viruses*, ***8***(10), p.264.

Kumar, P., & Kumar, M. (2018). Leaf curl disease: a significant constraint in the production of tomato in India. *Advances in Plant Pathology*, **93**.

Laco, G. S., & Beachy, R. N. (1994). Rice tungro bacilliform virus encodes reverse transcriptase, DNA polymerase, and ribonuclease H activities. *Proceedings of the National Academy of Sciences*, ***91***(7), 2654-2658.

Lazarowitz, S. G., & Beachy, R. N. (1999). Viral movement proteins as probes for intracellular and intercellular trafficking in plants. *The Plant Cell*, ***11***(4), 535-548.

Lazarowitz, S. G., & Shepherd, R. J. (1992). Geminiviruses: genome structure and gene function. *Critical Reviews in Plant Sciences*, ***11***(4), 327-349.

Li, F., Xu, X., Huang, C., Gu, Z., Cao, L., Hu, T., & Zhou, X. (2015). The AC 5 protein encoded by Mungbean yellow mosaic India virus is a pathogenicity determinant that suppresses RNA silencing-based antiviral defenses. *New Phytologist*, ***208***(2), 555-569.

Li, K., Wu, G., Li, M., Ma, M., Du, J., Sun, M., & Qing, L. (2018). Transcriptome analysis of *Nicotiana benthamiana* infected by Tobacco curly shoot virus. *Virology journal*, ***15***(1), 1-15.

Liu, S. R., Zhou, J. J., Hu, C. G., Wei, C. L., & Zhang, J. Z. (2017). MicroRNA-mediated gene silencing in plant defense and viral counter-defense. *Frontiers in microbiology*, ***8***, 1801.

Lu, S., Sun, Y. H., Shi, R., Clark, C., Li, L., & Chiang, V. L. (2005). Novel and mechanical stress–responsive microRNAs in *Populus trichocarpa* that are absent from Arabidopsis. *The Plant Cell*, ***17***(8), 2186-2203.

Luan, J. B., Yao, D. M., Zhang, T., Walling, L. L., Yang, M., Wang, Y. J., & Liu, S. S. (2013). Suppression of terpenoid synthesis in plants by a virus promotes its mutualism with vectors. *Ecology Letters*, ***16***(3), 390-398.

Marino, D., Peeters, N., & Rivas, S. (2012). Ubiquitination during plant immune signaling. *Plant physiology*, ***160***(1), 15-27.

Marmey, P., Bothner, B., Jacquot, E., de Kochko, A., Ong, C. A., Yot, P., Fauquet, C. M. (1999). Rice tungro bacilliform virus open reading frame 3 encodes a single 37-kDa coat protein. *Virology*, ***253***(2), 319-326.

Marmey, P., Rojas-Mendoza, A., De Kochko, A., Beachy, R. N., & Fauquet, C. M. (2005). Characterization of the protease domain of Rice tungro bacilliform virus responsible for the processing of the capsid protein from the polyprotein. *Virology journal*, ***2***(1), 1-12.

Mase, K., Ishihama, N., Mori, H., Takahashi, H., Kaminaka, H., Kodama, M., & Yoshioka, H. (2013). Ethylene-responsive AP2/ERF transcription factor MACD1 participates in phytotoxin-triggered programmed cell death. *Molecular plant-microbe interactions*, ***26***(8), 868-879.

Mittal, D., Borah, B. K., & Dasgupta, I. (2008). Agroinfection of cloned Sri Lankann cassava mosaic virus DNA to Arabidopsis thaliana, Nicotiana tabacum and cassava. *Archives of virology*, ***153***(11), 2149-2155.

Monteiro, S., Barakat, M., Piçarra-Pereira, M. A., Teixeira, A. R., & Ferreira, R. B. (2003). Osmotin and thaumatin from grape: a putative general defense mechanism against pathogenic fungi. *Phytopathology*, ***93***(12), 1505-1512.

Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., Jones, J. D. (2006). A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science*, ***312***(5772), 436-439.

Navot, N., Ber, R., & Czosnek, H. (1989). Rapid detection of tomato yellow leaf curl virus in squashes of plants and insect vectors. *DNA*, ***79***(562), 568.

Pasumarthy, K. K., Choudhury, N. R., & Mukherjee, S. K. (2010). Tomato leaf curl Kerala virus (ToLCKeV) AC3 protein forms a higher order oligomer and enhances ATPase activity of replication initiator protein (Rep/AC1). *Virology Journal*, ***7***(1), 1-8.

Periasamy, M., Niazi, F. R., & Malathi, V. G. (2006). Multiplex RT-PCR, a novel technique for the simultaneous detection of the DNA and RNA viruses causing rice tungro disease. *Journal of virological methods*, ***134***(1-2), 230-236.

Pierce, E. J., & Rey, M. C. (2013). Assessing global transcriptome changes in response to South African cassava mosaic virus [ZA-99] infection in susceptible *Arabidopsis thaliana. PloS one*, ***8***(6), e67534.

Pradhan, B., Van Tien, V., Dey, N., & Mukherjee, S. K. (2017). Molecular biology of geminivirus DNA replication. *Viral Replication*, **2**-35.

Prasad, A., Sharma, N., Muthamilarasan, M., Rana, S., & Prasad, M. (2019). Recent advances in small RNA mediated plant-virus interactions. *Critical reviews in biotechnology*, ***39***(4), 587-601.

Rajeswaran, R., Golyaev, V., Seguin, J., Zvereva, A. S., Farinelli, L., & Pooggin, M. M. (2014). Interactions of Rice tungro bacilliform pararetrovirus and its protein P4 with plant RNA-silencing machinery. *Molecular plant-microbe interactions*, ***27***(12), 1370-1378.

Ravichandran, S., Stone, S. L., Benkel, B., & Prithiviraj, B. (2013). Purple Acid Phosphatase5 is required for maintaining basal resistance against *Pseudomonas syringae* in Arabidopsis. *BMC plant biology*, ***13***(1), 1-12.

Sahu, P. P., Puranik, S., Khan, M., & Prasad, M. (2012). Recent advances in tomato functional genomics: utilization of VIGS. *Protoplasma*, ***249***(4), 1017-1027.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning: a laboratory manual* (No. Ed. 2). Cold spring harbor laboratory press.

Satoh, K., Kondoh, H., Sasaya, T., Shimizu, T., Choi, I. R., Omura, T., & Kikuchi, S. (2010). Selective modification of rice (*Oryza sativa*) gene expression by rice stripe virus infection. *Journal of General Virology*, ***91***(1), 294-305.

Saunders, K., & Stanley, J. (1995). Complementation of African cassava mosaic virus AC2 gene function in a mixed bipartite geminivirus infection. *Journal of general virology*, ***76***(9), 2287-2292.

Schardl, C. L., Florea, S., Pan, J., Nagabhyru, P., Bec, S., & Calie, P. J. (2013). The epichloae: alkaloid diversity and roles in symbiosis with grasses. *Current opinion in plant biology*, ***16***(4), 480-488.

Seo, J. K., Kim, M. K., Kwak, H. R., Choi, H. S., Nam, M., Choe, J., & Jung, C. (2018). Molecular dissection of distinct symptoms induced by tomato chlorosis virus and tomato yellow leaf curl virus based on comparative transcriptome analysis. *Virology*, ***516***, 1-20.

Settlage, S. B., See, R. G., & Hanley-Bowdoin, L. (2005). Geminivirus C3 protein: replication enhancement and protein interactions. *Journal of virology*, ***79***(15), 9885-9895.

Sharma, S., & Dasgupta, I. (2012). Development of SYBR Green I based real-time PCR assays for quantitative detection of Rice tungro bacilliform virus and Rice tungro spherical virus. *Journal of Virological Methods*, ***181***(1), 86-92.

Simón-Mateo, C., & García, J. A. (2006). MicroRNA-guided processing impairs Plum pox virus replication, but the virus readily evolves to escape this silencing mechanism. *Journal of virology*, ***80***(5), 2429-2436.

Son, G. H., Wan, J., Kim, H. J., Nguyen, X. C., Chung, W. S., Hong, J. C., & Stacey, G. (2012). Ethylene-responsive element-binding factor 5, ERF5, is involved in chitin-induced innate immunity response. *Molecular Plant-Microbe Interactions*, ***25***(1), 48-60.

Sun, F., Fang, P., Li, J., Du, L., Lan, Y., Zhou, T., & Zhou, Y. (2016). RNA-seq-based digital gene expression analysis reveals modification of host defense responses by rice stripe virus during disease symptom development in Arabidopsis. *Virology journal*, ***13***(1), 1-13.

Sunter, G., Stenger, D. C., & Bisaro, D. M. (1994). Heterologous complementation by geminivirus AL2 and AL3 genes. *Virology*, ***203***(2), 203-210.

Tong, A., Yuan, Q., Wang, S., Peng, J., Lu, Y., Zheng, H., Yan, F. (2017). Altered accumulation of osa-miR171b contributes to rice stripe virus infection by regulating disease symptoms. *Journal of experimental botany*, ***68***(15), 4357-4367.

Várallyay, É., Válóczi, A., Ágyi, Á., Burgyán, J., & Havelda, Z. (2010). Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. *The EMBO journal*, ***29***(20), 3507-3519.

von Roepenack-Lahaye, E., Newman, M. A., Schornack, S., Hammond-Kosack, K. E., Lahaye, T., Jones, J. D., & Dow, J. M. (2003). p-Coumaroylnoradrenaline, a novel plant metabolite implicated in tomato defense against pathogens. *Journal of Biological Chemistry*, ***278***(44), 43373-43383.

Wang, J., Mei, J., & Ren, G. (2019). Plant microRNAs: biogenesis, homeostasis, and degradation. *Frontiers in plant science*, ***10***, 360.

Whitham, S. A., Anderberg, R. J., Chisholm, S. T., & Carrington, J. C. (2000). Arabidopsis RTM2 gene is necessary for specific restriction of tobacco etch virus and encodes an unusual small heat shock–like protein. *The Plant Cell*, ***12***(4), 569-582.

Wu, J., Yang, R., Yang, Z., Yao, S., Zhao, S., Wang, Y., Li, Y. (2017). ROS accumulation and antiviral defence control by microRNA528 in rice. *Nature plants*, ***3***(1), 1-7.

Wu, M., Ding, X., Fu, X., & Lozano-Duran, R. (2019). Transcriptional reprogramming caused by the geminivirus Tomato yellow leaf curl virus in local or systemic infections in *Nicotiana benthamiana*. *BMC genomics*, ***20***(1), 1-17.

Xia, Z., Zhao, Z., Li, M., Chen, L., Jiao, Z., Wu, Y., Fan, Z. (2018). Identification of miRNAs and their targets in maize in response to Sugarcane mosaic virus infection. *Plant Physiology and Biochemistry*, ***125***, 143-152.

Yoshida, S. (1972). Physiological aspects of grain yield. *Annual review of plant physiology*, ***23***(1), 437-464.

Zarreen, F., Kumar, G., Johnson, A. A., & Dasgupta, I. (2018). Small RNA-based interactions between rice and the viruses which cause the tungro disease. *Virology*, ***523***, 64-73.

Zheng, Z., Qualley, A., Fan, B., Dudareva, N., & Chen, Z. (2009). An important role of a BAHD acyl transferase-like protein in plant innate immunity. *The Plant Journal*, ***57***(6), 1040-1053.