

**Investigating the molecular role of bryophyte-specific protein
SHORT-LEAF (SHLF) in gametophore development of moss
(*Physcomitrium patens*)**

A thesis submitted in partial fulfillment of the requirement of the degree
of Doctor of Philosophy by

Shirsa Palit

20173510



**INDIAN INSTITUTE OF SCIENCE EDUCATION AND
RESEARCH, PUNE**

2023

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

डॉ. होमी भाभा मार्ग, पुणे 411008, महाराष्ट्र, भारत | Dr. Homi Bhabha Road, Pune 411008, Maharashtra, India

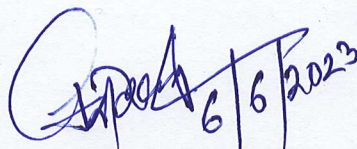
T +91 20 2590 8001 W www.iiserpune.ac.in



CERTIFICATE

Certified that the work incorporated in the thesis entitled “**Investigating the molecular role of bryophyte-specific protein SHORT-LEAF (SHLF) in gametophore development of moss (*Physcomitrium patens*)**” submitted by **Shirsa Palit** (Reg. no. 20173510) was carried out by the candidate, under my supervision. The work presented here or part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

Date: 6th June, 2023


Prof. A. K. Banerjee

(Thesis supervisor)

प्रो. अंजन के. बॅनर्जी/Prof. Anjan K. Banerjee
प्राध्यापक, जीवशास्त्र कार्यक्रम
Professor, Biology Program
भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान
Indian Institute of Science Education & Research
पुणे, भारत /Pune- 411 008, India

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

डॉ. होमी भाभा मार्ग, पुणे 411008, महाराष्ट्र, भारत | Dr. Homi Bhabha Road, Pune 411008, Maharashtra, India

T +91 20 2590 8001 W www.iiserpune.ac.in



DECLARATION

I declare that this written submission solely represents my original ideas reflected in my own words and also that places where others' ideas are included, I have adequately cited and referenced the original sources. I also declare that I have adhered to the principles of academic honesty, integrity and transparency and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will warrant disciplinary action by the Institute and may also evoke penal action from the sources that have not been properly cited or from whom proper permission has not been taken when necessary.

Date: 6th June 2023

Shirsa Palit

Shirsa Palit

Reg. No. 20173510

Acknowledgements

I am deeply grateful to Prof. Anjan Banerjee, my PhD supervisor. His unwavering support and guidance have shaped me as a researcher and individual. His motivation, insights, and scientific acumen have inspired me, while his encouragement to "stick to the bench till it's done!" and relentless pursuit of science have left me truly inspired. I am forever indebted to him for my personal and professional growth.

I am sincerely grateful to my RAC committee members, Prof. Utpal Nath, Prof. Aurnab Ghose and Dr Deepak Barua for their insightful comments and encouragement which have been invaluable in shaping my research progress over the past five years. Their critical feedback has helped me broaden my perspectives and improve my work. I am humbled and honoured to have been the recipient of their guidance and support.

I am deeply grateful to my host institute, IISER Pune, for the opportunities and resources they have provided throughout my research journey. Their multidisciplinary platform and world-class facilities have been instrumental in my growth as a researcher. I would also like to express my sincere gratitude to the technical staff and the members of the bio-office for their invaluable support during my thesis work. I am especially indebted to Mr Vijay Vittal and Dr Santosh Poddar (Microscopy facility), Mr Saddam Sheikh (Proteomic facility) and Dr Sagar Tatale (Animal house facility), for their help and support during the execution of my experiments.

My sincere thanks to the plant research community for sharing many biological reagents and transgenic lines which enabled me to complete my PhD thesis. My sincere thanks to Prof. Magdalena Bezanilla, University of Dartmouth, USA for providing the Gateway-based overexpression constructs, which have been used in this study. I would like to thank Prof. Mitsuyasu Hasebe, NIBB, Japan for sharing the moss knockout constructs.

I am thankful to the Council of Scientific & Industrial Research (CSIR), Ministry of Science and Technology, Government of India for awarding me the Junior and Senior Research Fellowships (CSIR-JRF/ SRF). I am also grateful to the American Society of Plant Biology and the Infosys foundation for granting me travel awards to attend Plant Biology 2022 at Oregon USA which provided me with an opportunity to present my PhD research work before peers and colleagues in plant biology. I am also grateful to have the recipient of the Best Poster award by EMBO, at Insights into early land plant evolution, 2022 held at NISER Bhubaneswar and the best oral talk award at the Current Trends and Future Prospects of Plant Biology 2023, held at the University of Hyderabad.

I am deeply grateful to Dr Boominathan Mohanasundaram and Dr Amey Jayant Bhide for their critical input, coaching and guidance. I have been blessed to be able to be coached by such seniors who not only encouraged and motivated me at every step of my research but provided an extremely conducive environment for scientific curiosity and discussion. They have gone above and beyond to ensure both

my professional and personal growth. Over time I have come to consider them as family. They have always been my sounding board and source of strength.

I would like to thank Gargi Chaturdevi, Madhusmita Pala, Kanishka Rajoria, Sanskar Agrawal, Payal Girigosavi, Srijan Das, Elakiya Venkatesan, Amlan Naskar, Amisha Gupta and Priyesh Agrawal for being the members of the moss team. I am grateful to have received their help, support and companionship during my PhD. I want to warmly thank the past members of the lab, Kirti Kumar Kondhare, Bhavani Natarajan, Ravi Devani, Amit Kumar, Harpreet S. Kalsi Anindita Karkhanis, and all current members including Kishan Saha, Nilam Malankar, Nikita Patil, Shikha Singh, Mohit Mantri and Debasmita Pal, for making the process joyful and easing the struggle of this journey.

I have heavily relied on the support of my friends, Dilsha C, Rajeshwari BR, Shruti Marathe, Firdousi Parvez, Gauri Binayak, Rituparna Ghosh, Rutwik Bardapurkar and Himani Khurana, Shivani Bodas, Anuja Patwardhan and Swapna Joshi throughout my PhD journey. I am also deeply grateful to Manish Kumar for his unwavering support and companionship throughout this journey. I am grateful to have Arunima Mukherjee and Shriya Kushwaha as my friends. I have truly cherished their love, support and companionship over the past decade. I will forever cherish all the conversations, laughs, and occasional tears we shared.

I am grateful for the inputs in learning basic research techniques, laboratory hacks, social skills, and motivational pep talks from senior colleagues Mr Mukul Rawat, Mr Abhishek Kanyal, Ms Ankita Sharma and Dr Saurabh Pradhan, my PhD journey wouldn't have been this comfortable and enjoyable.

I am obliged to my past teachers and mentors who have supported and encouraged me throughout my journey, particularly Mr Mrinal Krishna Deb and Mrs Ishita Dasgupta. Their trust in me has been instrumental in boosting my self-confidence and reaching my goals.

I would like to extend my heartfelt appreciation to BTS, whom I discovered during the COVID-19 lockdown in 2020. Their music and comforting messages have helped me persevere especially during the lockdown when I was away from my lab and the work that I love so much.

I dedicate this thesis to my family for their unwavering support and encouragement throughout my academic journey. My heartfelt and humble gratitude to my parents, Dr Anup Palit, and Mrs Mousumi Palit, and my grandparents Mrs Sabita Palit and Mrs Shila Dey, for their love, trust, understanding, and constant moral support. Their encouragement, patience, and understanding have been instrumental in keeping me motivated and focused during the ups and downs of this challenging journey. I am grateful to my family, who have always believed in me and provided me with the resources and opportunities to pursue my dreams. I have been loved, cared for, tolerated, reprimanded, cherished, and encouraged and have persevered in the darkest of times all owing to my family. They are my pillars of strength and I owe them everything.

Shirsa Palit

June 2023

TABLE OF CONTENTS

List of figures	viii
List of tables.....	x
Abbreviations.....	xi
Synopsis.....	xv
1. INTRODUCTION	2
1.1. Terrestrialization and land plant evolution	2
1.2. Adaptations which facilitated plant evolution	4
1.2.1. Genetic adaptations driving plant terrestrialization and evolution.....	5
1.3. Significance of lineage-specific genes in land plant evolution.....	7
1.4. Evolution and significance of repeat-containing proteins in plant development....	9
1.5. Evolution and function of phytohormones in land plants.....	11
1.5.1. Roles of hormones in plant development	11
1.6. Role of auxin in plant development	16
1.6.1. Origin and evolution of auxin transport.....	16
1.6.2. Evolutionary conservation of polar auxin transport across land plant lineage..	17
1.6.3. Regulation of PAT by signalling peptides.....	20
1.6.4. Role of flavonoids in regulating PAT.....	21
1.6.1. Effect of reactive oxygen species in the regulation of PAT.....	22
1.7. Functional conservation of non-polar auxin transport across land plant lineage .	25

1.8.	Significance of bryophytes in the evolutionary history of land plants	27
1.9.	The model moss <i>Physcomitrium patens</i>	30
1.10.	Hypothesis and objectives.....	32
2.	IDENTIFICATION OF CRUCIAL SHLF DOMAINS REGULATING GAMETOPHORE DEVELOPMENT	35
2.1.	Introduction	35
2.1.1.	Lineage specific genes in plants	35
2.1.2.	Plant repeat containing protein families and their post translational modifications.....	36
2.2.	Materials and methods	38
2.2.1.	Plant tissue culture and maintenance.....	38
2.2.2.	Cloning and plant transformation	38
2.2.3.	PEG mediated protoplast transformation	39
2.2.4.	Phenotypic characterization.....	40
2.2.5.	Statistical analyses	40
2.2.6.	Confocal Microscopy.....	41
2.2.7.	Generation of polyclonal SHLF antibody	41
2.2.8.	Western blot analyses	42
2.2.9.	Mass spectrometric analysis	42
2.2.10.	RT-qPCR analysis	43
2.2.11.	In silico protein analyses	43
2.1.	Results.....	43
2.1.1.	N-terminal and the highly conserved 2 nd TDR of SHLF are essential for moss gametophyte development	43

2.1.2. Subcellular localization of SHLF is dependent on both N- and C-terminal domains	48
2.1.3. SHLF undergoes post-translational processing to generate cryptic cleavage products	50
2.2. Discussion	58
2.2.1. N-terminal mediated ER trafficking is essential for SHLF's function.....	58
2.2.2. The C-terminal tail of SHLF may be functionally redundant	59
2.2.3. Signal peptide along with the first two TDRs constitute the miniSHLF.....	60
2.2.4. SHLF undergoes cleavage <i>in planta</i>	61
3. STUDYING THE SECRETORY NATURE OF SHLF AND THE ROLE OF ITS CLEAVAGE PRODUCTS IN GAMETOPHORE DEVELOPMENT	64
3.1. Introduction	64
3.1.1. The plant apoplastic space and secretome	64
3.1.2. Studying the components of plant secretome	64
3.1.3. Secretory proteins and peptides in plant development and stress response	65
3.2. Materials and methods	66
3.2.1. Plant tissue culture and maintenance	66
3.2.2. Secretome and extracellular peptidome isolation	66
3.2.3. Secretome and peptidome supplementation assays	67
3.2.4. Synthetic SHLF peptide supplementation	67
3.2.5. Western blot analyses	67
3.2.6. Sample preparation for MS/MS analyses	67
3.2.7. LC-MS/MS analysis and peptide identification	68
3.2.8. RT-qPCR analysis.....	69
3.2.9. Phenotypic characterization.....	69

3.2.10. Statistical analyses.....	69
3.2.11. GUS staining	70
3.3. Results.....	70
3.3.1. A truncated version of SHLF is an apoplast resident protein.....	70
3.3.2. SHLF acts as a secretory protein via functional cryptic peptides	72
3.3.3. SHLFpep 3 affects auxin distribution pattern and gametophore development in a dosage dependent manner	76
3.4. Discussion.....	80
3.4.1. SHLF undergoes cleavage <i>in planta</i> and functions as a secretory protein to regulate gametophore development.	80
3.4.2. Secretory SHLF peptides affect auxin distribution pattern and gametophore development in a dosage-dependent manner	81
 4. A TRANSCRIPTOMIC AND METABOLOMIC APPROACH TO ELUCIDATE THE MODE OF SHLF FUNCTION IN REGULATING MOSS DEVELOPMENT.....	 85
4.1. Introduction	85
4.1.1. Auxin transport in bryophytes	85
4.1.2. Interplay of flavonoids and ROS in auxin transport.....	86
4.2. Materials and methods	87
4.2.1. RNA Seq analysis	87
4.2.2. RT-qPCR analysis.....	88
4.2.3. Toluidine blue staining	88
4.2.4. DAB staining	88
4.2.5. ROS quantification using carboxy H ₂ DCFDA assay	89
4.2.6. Total Chlorophyll estimation	89

4.2.7. UHPLC-MS/MS analysis	89
4.2.8. Analyzing plasmodesmata (PD) permeability of the mutant using Dendra2	90
4.2.9. Co-immunoprecipitation analyses	90
4.2.10. Mass spectrometric analyses of the interacting partners.....	91
4.3. Results.....	91
4.3.1. Several stress responsive genes are differentially regulated in <i>shlf</i>	91
4.3.2. <i>SHLF</i> modulates the levels of metabolites regulating auxin distribution	100
4.3.3. <i>shlf</i> exhibits increased PD permeability for protein diffusion.....	106
4.3.4. Identification of SHLF interactors by Co-IP using anti-SHLF	108
4.4. Discussion	110
5. SUMMARY AND FUTURE DIRECTIONS	116
6. REFERENCES	126
7. ANNEXURES	150

List of figures

Figure 1.1	Streptophyte colonization of terrestrial habitats by extant plants.	3
Figure 1.2	Adaptations underlying land plant evolution.	4
Figure 1.3	Dendrogram depicting the evolution of genetic pathways during early plant evolution ...	6
Figure 1.4	Tertiary structures of plant repeat proteins.....	10
Figure 1.5	Major roles of phytohormones in plant development and defense.....	12
Figure 1.6	Peptide hormones in plant development.....	14
Figure 1.7	A schematic representation of non-polar and polar auxin transport in Arabidopsis	17
Figure 1.8	Auxin transport pathways in the evolution of branching forms.	18
Figure 1.9	Proposed regulatory network of auxin and flavonoids.	23
Figure 1.10	A diagrammatic representation of how redox processes regulate auxin signalling.	24
Figure 1.11	The key components of auxin transport are represented.	26
Figure 1.12	Life cycle of <i>Physcomitrium patens</i>	30
Figure 1.13	Chronological accounts of <i>P. patens</i> research.....	32
Figure 2.1	Schematic representation of the SHLF domain deletion constructs.....	39
Figure 2.2	Overexpression of domain deleted SHLF constructs in <i>shlf</i>	44
Figure 2.3	PCR based molecular confirmation of SHLF domain-deleted overexpression lines.	45
Figure 2.4	The 2- Δ Ct plot showing the level of overexpression of the domain-deleted constructs..	46
Figure 2.5	Comparative phenotypic analyses of the stable overexpression lines in <i>shlf</i>	47
Figure 2.6	Domain deletion of SHLF affects subcellular localization pattern	49
Figure 2.7	Western blotting with anti-GFP to show that SHLF undergoes cleavage.....	50
Figure 2.8	Alphafold2 based structure predictions reveal a linker-like region at SHLF TDRs	52
Figure 2.9	Hydropathy index calculation for SHLF protein using the scale Hydropath.	53
Figure 2.10	Cloning and generation of anti-SHLF primary antibody in rabbit.....	54
Figure 2.11	In-house generation and validation of anti-SHLF primary antibody in rabbit.....	55
Figure 2.12	Western blotting with anti-SHLF to show that SHLF undergoes cleavage in-planta	56
Figure 3.1	A schematic representation of isolation and identification of total extracellular proteins .	70
Figure 3.2	Immunoblotting with polyclonal SHLF antibody secretome	71
Figure 3.3	A schematic representation the WT SHLF protein and secretory peptides.....	72

Figure 3.4	A schematic representation of protocol for supplementation of secretome to <i>shlf</i>	73
Figure 3.5	Independent supplementation of secretome and extracellular peptidome.....	74
Figure 3.6	Phenotypic analyses of the supplemented <i>shlf</i> lines.....	75
Figure 3.7	A schematic representing the sites of origin of SHLF peptides (SHLFpep1-3).....	76
Figure 3.8	SHLFpep1 and SHLFpep3 shows dosage dependent effects	77
Figure 3.9	Phenotypic characterization of the SHLF peptide supplementation in WT and <i>shlf</i>	78
Figure 3.10	Schematic of the phenotypic effects of SHLFpep supplementation to WT and <i>shlf</i>	79
Figure 4.1	Workflow for the RNASeq analyses of WT and <i>shlf</i>	92
Figure 4.2	Comparative analyses of WT and <i>shlf</i> transcriptomes.	93
Figure 4.3	Comparative analyses of WT and <i>shlf</i> transcriptomes	94
Figure 4.4	Word cloud (generated using python) representing top 20 DEGs	95
Figure 4.5	Heatmap showing hierarchical clustering analysis of top 20 DEGs	96
Figure 4.6	The bar plot shows log ₂ Fold change in levels of stress responsive genes in <i>shlf</i>	97
Figure 4.7	<i>shlf</i> gametophores exhibit elevated stress response.....	98
Figure 4.8	Panel represents the DAB-stained gametophores and the relative ROS levels.....	99
Figure 4.9	Heatmap representing the mean 2 ^{-ΔCt} values of the top 10 DEGs.....	100
Figure 4.10	Schematic of flavonoid biosynthesis pathway and 2 ^{-ΔCt} values of their genes	101
Figure 4.11	Hierarchical clustering of differential metabolites between WT and <i>shlf</i>	101
Figure 4.12	Barplots represent six bio-replicates depicting the relative abundance of metabolites..	103
Figure 4.13	Graphs representing the fragmentation pattern of each metabolite.....	104
Figure 4.14	Time-lapse imaging of Dendra2 movement through PD in WT and <i>shlf</i>	106
Figure 4.15	Time-course of relative fluorescence intensity in cell 0 , +1 and -1 in WT and <i>shlf</i>	107
Figure 4.16	Western blot representing the anti-SHLF immunoprecipitated WT total protein	108
Figure 4.17	A schematic representation of the D-box element present within the SHLF TDRs	109
Figure 4.18	Proposed pathway for SHLF function.....	113

List of tables

Table 2.1: List of peptides detected upon MS/MS analyses of the bands in anti-SHLF blot.....	61
Table 4.1 List of SHLF interactors identified from whole tissue lysate.....	111
Table 4.2 List of DEGs associated with polar auxin transport (PAT).....	116
Table 5.0 List of primers used in the current study.....	123
Table 7.1. List of peptides detected upon MS/MS analyses of the bands in anti-GFP blot.....	150
Table 7.2 List of secretory proteins identified from WT secretome.....	154
Table 7.3 List of secretory proteins identified from <i>shlf</i> secretome.....	157
Table 7.4 List of secretory SHLF peptides identified from moss lines used in this study.....	160
Table 7.5 List of most up-regulated genes between WT and <i>shlf</i> at p value <0.01.....	163
Table 7.6 List of most down-regulated genes between WT and <i>shlf</i> at p value <0.01.....	165
Table 7.7 List of all metabolites differentially regulated between WT and <i>shlf</i>	167

ABBREVIATIONS

1-Naphthaleneacetic acid	NAA
1-N-Naphthylphthalamic acid	NPA
2,4-Dichlorophenoxyacetic acid	2,4-D
2',7'-dichlorodihydrofluorescein diacetate	H2DCFDA
2-Deoxy-D-Glucose	DDG
6-Benzylaminopurine	BAP
ABC transporter	ABCB
abscisic acid	ABA
amino acid	aa
<i>Arabidopsis thaliana</i>	<i>A. thaliana</i>
Auxin Response Elements	AREs
AUXIN RESPONSE FACTOR 3	ARF3
AUXIN RESPONSE FACTOR 4	ARF4
Auxin Responsive Endogenous Polypeptide 1	AREP1
AUXIN/INDOLE-3-ACETIC ACID	AUX/IAA
beta-1,3-glucanase	PdGB
Brassinosteroids	BR
Callose synthase	CalS
CHALCONE SYNTHASE	CHS
CLAVATA3/Endosperm surrounding region-related	CLE
CLE-LIKE	CLEL
Constitutive Expresser of PR Genes 5	CPR5
cyclic nucleotide-gated channels	CNGC
cytokinin	CK
Diaminobenzidine	DAB
DIOXYGENASE for AUXIN OXIDATION1	DAO1
ELONGATED HYPOCOTYL	HY
endoplasmic reticulum	ER
EPIDERMAL PATTERNING FACTOR-LIKE 2 (EPFL2)	EPFL2

epidermal patterning factors	EPF
Ethylene	ETH
ETTIN	ETT
Gibberellic Acid	GA
GIBBERELIC-ACID INSENSITIVE (GAI), REPRESSOR of GAI (RGA) and SCARECROW (SCR)	GRAS
gibberellin	GA
GLOVEN	GLV
Glucan synthase Like	GSL
Glutathione	GSH
herpes simplex virus VP16-associated	HCF
Homeodomain-leucine zipper	HD Zip
INDEHISCENT	IHD
Indole acetic acid	IAA
indole-3-pyruvic acid	IPyA
Jasmonic acid	JA
Karrikin	KAR
KNOTTED-like homeobox	KNOX
Knotted-like homeobox	KNOX
Lineage specific genes	LSGs
MADS-Protein Complexes	MIKC
MAP Kinase	MAPK
<i>medio-lateral</i>	<i>M-L</i>
MINICHROMOSOME MAINTENANCE FACTOR 1, AGAMOUS, DEFICIENS and Serum Response Factor (SRF)	MADS
minimal SHLF	miniSHLF
monothiol glutaredoxin	GrxS
Multi Drug Resistant	MDR
Myeloblastosis viral oncogene homolog	MYB
N-1-Naphthylphthalamic Acid	NPA
Naphthoxyacetic acid	NOA
nicotinamide adenine dinucleotide phosphate thioredoxin reductases	NTR
Oryza sativa defense-responsive gene 10	OsDR10

PAMP-induced secreted peptides	PIPs
PAT inhibitor	PAT-I
PD callose-binding proteins	PDCB
PD-localized proteins	PDLPs
PEP receptor	PEP-EPR
P-glycoprotein	PGP
<i>Physcomitrium patens</i>	<i>P. patens</i>
<i>PINFORMED1</i>	<i>PINI</i>
PIN-LIKES	PILS
PINOID	PID
PIP-LIKES	PILS
Plant elicitor peptides	PEPs
Plasmodesmata	PD
Plasmodesmata associated callose	PDAC
polar auxin transport	PAT
POLARIS	PLS
polymerase chain reaction	PCR
Primer Binding Site	PBS
Qua Quine Starch	QQS
quiescent center	QC
Rapid Alkalinization Factor	RALF
Reactive oxygen species	ROS
Regulatory particle triple-A ATPase 4	RPT4
REPLUMLESS	RPL
respiratory burst oxidase homologs	RBOH
restriction enzyme	RE
Root Growth Factor	RGF
ROOT GROWTH FACTOR	RGF
ROOT HAIR DEFECTIVE SIXLIKE 4	RSL4
Salicylic acid	SA
SECONDARY WALL THICKENING PROMOTING FACTOR/secondary wall-associated NAC domain protein	NST/SND

Shoot apical meristem	SAM	
SHORT-LEAF	SHLF	
short-leaf	shlf	
Skp, Cullin, F-box-containing complex	SCF	
SOMBRERO	SMB	
Strigolactones	SLs	
Tandem direct repeat	TDR	
TARGET OF LBD SIXTEEN 2	TOLS2	
TOPLESS	TPL	
transcription factor	TF	
Transmembrane Kinases	TMKs	
<i>transparent testa 4</i>	tt4	
TRANSPORT INHIBITOR RESPONSE1	TIR1	
Triple mutant of TRX reductase (NTRA,NTRB) and glutathione biosynthesis gene cadmiumsensitive2 (CAD2)	ntra cad2	<i>ntrb</i>
<i>TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS</i>	TAA	
VASCULAR-RELATED NAC-DOMAIN	VND	
Whole genome sequencing	WGS	
wild-type	WT	
YUCCA	YUC	
β -glucuronidase	GUS	

SYNOPSIS

SYNOPSIS

Investigating the molecular role of bryophyte-specific protein SHORT-LEAF (SHLF) in gametophore development of moss (*Physcomitrium patens*)

Name: Shirsu Palit

Registration No: 20173510

Name of the supervisor: Prof. Anjan K. Banerjee

Department: Biology

Date of registration: 1st August 2017

Institute: Indian Institute of Science Education and Research (IISER), Pune, India.

Introduction

Terrestrialization can be heralded as the driving force behind land plant evolution. To ensure survival in dynamic environmental conditions, plants have acquired a plethora of adaptations, mediated by the intricate interplay of genes and phytohormones. Interestingly, though the molecular and physiological roles of phytohormones appear to be evolutionarily conserved, the genetic components governing plant development appear to be varied across land plants. Significant part of land plant genomes encodes lineage-specific genes (LSGs) having unknown function. Most of the functionally annotated LSGs are shown to be necessary for survival in a specific ecological niche and have been identified from angiosperms. These genes are mostly involved in regulating plant development by modulating stress response, for example Qua Quine Starch (QQS) (Li *et al.*, 2009; Arendsee, Li and Wurtele, 2014; Li and Wurtele, 2015), the Constitutive Expresser of PR Genes 5 (CPR5) (Jing *et al.*, 2007) in *Arabidopsis* and the rice-specific gene, OsDR10 (*Oryza sativa* defense-responsive gene 10)

(Xiao *et al.*, 2009). However, the evolution and function of LSGs in non-flowering plant lineages remains largely unknown.

Bryophytes, being one of the earliest colonizers of land, are interesting models to explore the evolution of plant adaptations underlying terrestrialization. Specifically, the haploid moss gametophore bears a resemblance to the diploid stem in angiosperms, however, the molecular machinery underlying the development of body plan appears to be ploidy-specific. This makes it essential to study haploid-specific genes governing the development of moss gametophore. Reverse genetic approaches in model moss *Physcomitrium patens*, have resulted in the functional characterization of many evolutionarily conserved genes (Sakakibara *et al.*, 2008; Coudert *et al.*, 2015). However, not much effort has been made to identify the bryophyte-specific genes that are indispensable for gametophore development.

Recently, a Tnt1 based forward genetic screen resulted in the identification of a bryophyte-specific gene, *SHORT-LEAF (SHLF)* to be the genetic cause of a *short-leaf (shlf)* mutant with altered leaf dimensions, plasmodesmata frequency and increased auxin accumulation at sites of synthesis (Mohanasundaram *et al.*, 2021). Primary protein structure analysis revealed that SHLF encodes an N-terminal endoplasmic reticulum (ER) signal peptide, four highly similar Tandem Direct Repeats (TDRs 1-4) and a short C-terminal tail. To the best of our knowledge, the SHLF TDRs are the longest known TDRs reported till date. The presence of ER signal peptide hints at protein trafficking into the conventional secretory pathway and SHLF is has been shown to traffick to ER and undergo cleavage.

One of the striking features of the *shlf* mutant is its altered auxin distribution pattern, where auxin gets restricted to its site of synthesis. As auxin is a master regulator which governs almost all aspects of plant development and stress response, it is plausible that the mutant phenotype could be a readout of multiple impaired pathways. Auxin being a small molecule, can be transported across membranes both by carrier-mediated active transport (polar auxin transport; PAT) and plasmodesmata-mediated passive diffusion (Petrasek and Friml, 2009). In bryophytes, PAT is involved in phyllid and gametophore stem development, however, auxin movement in gametophore stems has been shown to be mostly symplastic (Coudert *et al.*, 2015). The exact tissue-specific interplay of polar and non-polar auxin

transport mechanisms remains to be established. Additionally, the role of endogenous PAT regulators (such as flavonoids) has not yet been reported in bryophytes, making it interesting to study the evolutionary significance of polar vs non-polar auxin transport dynamics and PAT regulation across plant lineages.

As SHLF appears to be involved in regulating auxin distribution through yet unknown mechanisms, it is necessary to unravel the mechanistic basis of the protein function in the regulation of moss gametophore development. However, the absence of any known conserved domains in SHLF made it difficult to decipher the molecular role of SHLF. Hence the following objectives were proposed for the following study:

1. Identification of crucial SHLF domains involved in gametophore development
2. Studying the secretory nature of SHLF and the role of its cleavage products in gametophore development
3. A transcriptomic and metabolomic approach to elucidate the mode of SHLF function in regulating moss development.

Chapter 2: Identification of crucial SHLF domains regulating gametophore development

As SHLF lacks any known conserved domains, it is imperative to first identify the functional domains which are necessary for protein function. Literature suggests that TDRs can either act individually or in concert with each other to mediate protein function (Schaper and Anisimova, 2015). The high sequence similarity between the SHLF TDRs prompted us to identify possible functional redundancy between them. Here, we attempted sequential domain deletion analyses, where the deletion constructs were independently overexpressed in the mutant background (*shlf*). The transgenic plants were scored based on phenotype reversal and protein localization to identify the role of each SHLF domain in protein localization and overall gametophore development. Upon overexpression of a construct lacking the N-terminal signal peptide, the transgenic plants failed to recover the phenotype and the protein did not traffick to the ER. Additionally, immunoblot analyses using polyclonal GFP antibody revealed that inhibition of ER trafficking also interfered with the proteolytic processing of the protein, indicating that N-terminal mediated ER trafficking is indispensable for protein

function. Amongst the four SHLF TDRs, the 1st TDR bears 90% similarity to the rest (TDR2-4), which are 99% similar to each other. Overexpression analyses using minimal SHLF constructs, in an attempt to identify functional redundancy between the TDRs revealed that the slightly divergent 1st TDR is not sufficient for protein function. Rather a minimal construct comprising the signal peptide along with the divergent 1st and conserved 2nd TDR comprises the minimal functional protein (miniSHLF). C-terminal tails in proteins are generally deemed to be essential for protein-protein interaction and for conferring post-translational modifications on the protein. Upon overexpression of a construct lacking the C-terminal tail, the mutant failed to recover the phenotype. This observation was quite curious as miniSHLF which lacks a C-terminal tail is sufficient to behave as a functional protein. This prompted us to carry out protein structural analyses using AlphaFold2 which revealed the presence of a 9 amino acid linker-like region between each TDR, which may bear functional redundancy to the C-terminal region. and could possibly explain the restoration of protein activity in the miniSHLF overexpression lines.

To detect the products of SHLF cleavage and to identify the role of SHLF domains in the proteolytic processing of the protein, a polyclonal SHLF-specific antibody was generated. Immunoblot analyses with these lines revealed that SHLF was indeed cleaved and that the cleavage products were comparable across WT, *shlf* and the overexpression lines,. Overall, here we identified the role of each SHLF domain in gametophore development and identified a minimal functional protein (miniSHLF) which is sufficient for SHLF activity.

Chapter 3: Studying the secretory nature of SHLF and the role of its cleavage products in gametophore development

As ER trafficking appears to be crucial for SHLF processing and function, we attempted to investigate the significance of secretion on protein activity. The plant apoplastic space and secretome constitute the primary distinguishing barrier from the external environment. *P. patens* being ectohydric and comprising uni-layered phyllid cells, is ideal to grow in liquid cultures (Lehtonen *et al.*, 2014). We conducted immunoblotting and MS/MS-based analyses of secretomes from WT, *shlf* and the transgenic lines to uncover the nature of secreted SHLF, the significance of secretion on protein function and the role of SHLF domains

in protein secretion. Upon immunoblotting with SHLF specific antibody, we detected a truncated secretory version of SHLF from the WT secretome, which corresponded to ~ 45 kDa, in contrast to the ~ 75 kDa of the total protein. In-gel digestion and MS/MS of the ~45, kDa revealed tryptic peptides originating from the divergent 1st TDR and the conserved TDRs, further highlighting the significance of the conserved TDRs in SHLF function. No SHLF-specific band was observed in the mutant secretome. Upon supplementing the mutant with WT and miniSHLF secretomes, a complete phenotypic recovery was observed, whereas peptidome supplementation resulted in only partial recovery, suggesting a dosage-dependent function of secretory SHLF.

Secretory peptides have been demonstrated to act at the confluence of genetic and phytohormone pathways to regulate plant development. Recently, the roles of evolutionarily conserved peptides like CLE and RALF have been identified in moss (Campbell *et al.*, 2007; Campbell and Turner, 2017; Whitewoods *et al.*, 2018; Mamaeva *et al.*, 2022). As lineage-specific genes are known to be involved in mediating plant survival in specific niches, it is interesting to explore the evolutionary significance of SHLF peptides in regulating moss growth development (Lespinet *et al.*, 2002; Hanada *et al.*, 2008). Our MS/MS analyses revealed that several SHLF peptides were detected in WT and in the miniSHLF lines, but were absent in the mutant as well as the domain-deleted overexpression lines which did not exhibit phenotypic recovery. Synthetic peptide supplementation assays revealed that SHLFpep3, a peptide originating from the conserved TDRs acts in a dosage-dependent manner to completely recover both the mutant phenotype and auxin distribution pattern. Interestingly, SHLFpep3 and SHLFpep1 supplementation in WT revealed that these peptides act in a dosage-dependent manner to affect gametophore phenotypes (including plant height and leaf dimensions) and overall auxin response. Overall, our analyses revealed that SHLF acts like a typical secretory protein, forms a pool of secretory peptides and that functional SHLFpep1/3 act in a dosage-dependent manner to regulate gametophore development and auxin response in moss. In future, identification of the SHLFpep3 receptor and its downstream signalling will further shed light on the molecular dynamics of the SHLFpep3 activity in bryophytes.

Chapter 4: A transcriptomic and metabolomic approach to elucidate the mode of SHLF function in regulating moss development

Given the lack of conserved domains in SHLF, elucidating the pathways affected in the *shlf* mutant was imperative to determine the underlying molecular mechanism of SHLF function in plants. Here, we took a multi omics approach to investigate the affected developmental pathways leading to the reduced leaf dimensions and altered auxin distribution pattern of *shlf*.

Comparative transcriptomic analyses of the mutant and the WT revealed the differential regulation of several stress-responsive genes, including those responsible for maintaining ROS (reactive oxygen species) levels and those involved in the phenylpropanoid biosynthesis pathway. Further analyses of the ROS levels using DAB staining and H₂DCFDA assay revealed that the mutant shows a high accumulation of ROS. The mutant transcriptome also showed differential regulation of genes involved in abiotic stress response including dehydrins and cell wall integrity maintenance genes.

Further, non-targeted metabolomic profiling of the mutant also revealed a high accumulation of metabolite intermediates of the phenylpropanoid pathway, including antioxidants like ascorbic acid and flavonoids such as quercetin and kaempferol which act as endogenous inhibitors of PAT in angiosperms (Brown *et al.*, 2001). Interestingly, the miniSHLF overexpression lines and SHLFpep3 supplementation were able to successfully recover the ROS levels and the differentially accumulated stress-responsive genes and metabolites in the mutant. This is indicative of the fact that SHLFpep3 is the minimal functional unit of SHLF and that this bryophyte-specific peptide affects the auxin distribution pattern via a flavonoid-ROS loop.

We also carried out co-immunoprecipitation assays using anti-SHLF, followed by MS/MS analyses to identify the SHLF interacting partners. Most of the SHLF interactors detected were appear to have dual roles in plant development and stress response. Notably, the interactor RPT4, which is also differentially regulated in *shlf*, has been shown to be responsible for modulation in stress response in plants (Han *et al.*, 2008). The implication of

SHLF-RPT4 interaction will indeed be an interesting direction for future investigation. Overall, our multi omics approaches coupled with peptide supplementation assays revealed that *shlf* exhibits elevated stress response and that SHLFpep3 is involved in regulating the auxin-ROS-flavonoid loop in bryophytes.

Summary

Land colonization by plants was mediated by a plethora of adaptations which enabled them to survive the harsh terrestrial environments. It is intriguing to study the underlying genetic and hormonal factors involved in regulation of the overall body plan of the early land colonizers. Plants like mosses which are positioned at the evolutionary crossroads between charophycean algae and flowering plants are the perfect models to trace the evolutionary trajectories of these players (Prigge *et al.*, 2010). These extant plants may also harbor certain clade-specific genes which are essential for their survival in specific ecological niches.

Previously, we reported a unique gene family called SHORT-LEAF, which is specific to bryophytes and regulates gametophore development in moss. The disruption of this gene by Tnt1 insertion resulted in the *short-leaf* (*shlf*) mutant, defective in gametophore development. SHLF contains an N terminal predicted signal peptide, 4 TDRs (Tandem Direct Repeats), which share 90-99% sequence similarity; and a C terminal tail. In this thesis, we have validated the secretory role of the signal peptide, carried out a mass spectrometric analysis of wild-type (WT) secretome and identified several SHLF peptides. Overexpression of a construct lacking signal peptide failed to recover the *shlf* mutant phenotype and SHLF peptides was absent in the secretome, indicating that secretion is necessary for SHLF function. Supplementation of both WT peptidome and secretome to *shlf* mutant exhibited partial and full phenotypic recovery respectively, without any changes in the SHLF transcript levels. In a quest to find a minimal functional SHLF, we also discovered the importance of the 2nd TDR in SHLF function and identified several secretory peptides derived from it. Supplementation of 2nd TDR-specific synthetic peptides (SHLFpep/s) rescued the *shlf* phenotypes, including leaf dimensions, auxin distribution, internodal distance and apical dominance. These data suggest that SHLF peptides may function in tandem with auxin to regulate moss gametophore development. Further, RNA-seq analysis and metabolomic profiling of *shlf* showed the up-

regulation of the levels of auxin-regulated stress responsive genes and metabolites, including intermediates of the phenylpropanoid pathway like flavonoids and chalcones, which are rescued upon SHLFpep/s supplementation. We have also identified putative key SHLF interactors which appear to play dual roles in plant development and defense.

In summary, we have demonstrated that the conserved SHLF TDRs forms functional secretory peptides which affect the auxin distribution pattern, stress response and gametophore development of moss. We propose that SHLF peptides may be useful in studying the evolutionary significance of lineage-specific peptide-hormone cross-talk in land plants.

List of publications

- **Palit S.**, Bhide A J., Mohanasundaram B., Pala M., & Banerjee, A. K. (2022) Secretory peptides from conserved tandem repeats of SHORT-LEAF regulate gametophore development in moss. (Under revision in **Plant Physiology**)
- Mohanasundaram, B., **Palit, S.**, Bhide, A. J., Girigosavi P & Banerjee, A. K (2021). SCARECROW regulated leaf blade and lamina development in the moss (*Physcomitrium patens*) (Under revision in **Plant Molecular Biology**)
- Mohanasundaram, B., Bhide, A. J., **Palit, S.**, Chaturvedi, G., Lingwan, M., Masakapalli, S. K., & Banerjee, A. K. (2021). The unique bryophyte-specific repeat-containing protein SHORT-LEAF regulates gametophore development in moss. **Plant Physiology**, 187(1), 203-217.

References

- Brown, D. E. et al. (2001) 'Flavonoids act as negative regulators of auxin transport in vivo in Arabidopsis', Plant physiology. American Society of Plant Biologists, 126(2), pp. 524–535.
- Campbell, L. and Turner, S. R. (2017) 'A comprehensive analysis of RALF proteins in green plants suggests there are two distinct functional groups', Frontiers in plant science. Frontiers Media SA, 8, p. 37.
- Campbell, M. A. et al. (2007) 'Identification and characterization of lineage-specific genes within the Poaceae', Plant physiology. American Society of Plant Biologists, 145(4), pp. 1311–1322.

- Coudert, Y. et al. (2015) 'Three ancient hormonal cues co-ordinate shoot branching in a moss', *Elife*. eLife Sciences Publications Limited, 4, p. e06808.
- Gomes, G. L. B. and Scortecci, K. C. (2021) 'Auxin and its role in plant development: structure, signalling, regulation and response mechanisms', *Plant Biology*. Wiley Online Library, 23(6), pp. 894–904.
- Han, Y. et al. (2008) 'Rice ROOT ARCHITECTURE ASSOCIATED1 binds the proteasome subunit RPT4 and is degraded in a D-box and proteasome-dependent manner', *Plant physiology*. American Society of Plant Biologists, 148(2), pp. 843–855.
- Hanada, K. et al. (2008) 'Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli', *Plant physiology*. American Society of Plant Biologists, 148(2), pp. 993–1003.
- Jing, H.-C. et al. (2007) 'Arabidopsis CPR5 is a senescence-regulatory gene with pleiotropic functions as predicted by the evolutionary theory of senescence', *Journal of experimental botany*. Oxford University Press, 58(14), pp. 3885–3894.
- Lehtonen, M. T. et al. (2014) 'Protein secretome of moss plants (*Physcomitrella patens*) with emphasis on changes induced by a fungal elicitor', *Journal of Proteome Research*. ACS Publications, 13(2), pp. 447–459.
- Lespinet, O. et al. (2002) 'The role of lineage-specific gene family expansion in the evolution of eukaryotes', *Genome research*. Cold Spring Harbor Lab, 12(7), pp. 1048–1059.
- Li, L. and Wurtele, E. S. (2015) 'The QQS orphan gene of Arabidopsis modulates carbon and nitrogen allocation in soybean', *Plant biotechnology journal*. Wiley Online Library, 13(2), pp. 177–187.
- Li, L. et al. (2009) 'Identification of the novel protein QQS as a component of the starch metabolic network in Arabidopsis leaves', *The Plant Journal*. Wiley Online Library, 58(3), pp. 485–498.
- Mamaeva, A. et al. (2022) 'RALF peptides modulate immune response in the moss *Physcomitrium patens*', *bioRxiv*. Cold Spring Harbor Laboratory.
- Mohanasundaram, B. et al. (2021) 'The unique bryophyte-specific repeat-containing protein SHORT-LEAF regulates gametophore development in moss'. Oxford University Press.
- Petrasek, J. and Friml, J. (2009) 'Auxin transport routes in plant development'. Oxford University Press for The Company of Biologists Limited.
- Prigge, M. J. et al. (2010) '*Physcomitrella patens* auxin-resistant mutants affect conserved elements of an auxin-signaling pathway', *Current Biology*. Elsevier, 20(21), pp. 1907–1912.

Sakakibara, K. et al. (2008) 'Class 1 KNOX genes are not involved in shoot development in the moss *Physcomitrella patens* but do function in sporophyte development', *Evolution & development*. Wiley Online Library, 10(5), pp. 555–566.

Schaper, E. and Anisimova, M. (2015) 'The evolution and function of protein tandem repeats in plants', *New Phytologist*. Wiley Online Library, 206(1), pp. 397–410.

Whitewoods, C. D. et al. (2018) 'CLAVATA was a genetic novelty for the morphological innovation of 3D growth in land plants', *Current Biology*. Elsevier, 28(15), pp. 2365–2376.

Xiao, W. et al. (2009) 'A rice gene of de novo origin negatively regulates pathogen-induced defense response', *PloS one*. Public Library of Science San Francisco, USA, 4(2), p. e4603.

CHAPTER 1:

Introduction

1. Introduction

1.1. Terrestrialization and land plant evolution

Life on earth originated in an aquatic environment. Terrestrialization of the primitive aquatic life was the driving force behind shaping earth's biosphere. Though the exact reason why terrestrialization occurred remains debatable to date, the availability of limited nutrient sources with ever-increasing space constraints may have favoured land colonization. The transition of life from water to land is considered as one of the landmark events in the evolution of the photosynthetic green lineage (Figure 1.1).

Approximately a billion years ago, the green lineage (Chloroplastida) divided into chlorophytes and streptophytes. The streptophytes consist of organisms belonging to both the embryophytes (land plants) and the streptophyte algae, which can be further divided into the basal-branching KCM grade (Klebsormidiophyceae, Chlorokybophyceae, and Mesostigmatophyceae) and the higher-branching ZCC grade (Zygnematophyceae, Coleochaetophyceae, and Charophyceae) (De Vries and Archibald, 2018) (Figure 1.1). Chlorophytes are found in various environments, including marine, freshwater, and terrestrial habitats, while streptophyte algae are found in freshwater and terrestrial habitats such as wet soil, rock surfaces, lake and stream sediment, and as epiphytes on other algae (Delwiche and Cooper, 2015).

Recent studies have established that land plants share their last common ancestor with the Zygnematophyceae, a group comprising of both unicellular and filamentous algae found in aquatic and terrestrial ecosystems (Becker and Marin, 2009). The closest relative to early aquatic land plants which evolved from unicellular haplontic algae like *Chlamydomonas*, possess flagella and are motile (Bold, 1949). This helped them to transcend locomotive barriers and move from one place to another in search of food. Over the course of time, the multicellular Charophycean algae started colonizing terrestrial habitats (Graham, 1985). This new environmental niche posed many challenges for their survival which led to the evolution of land plants. During this process they lost their flagella and formed symbiotic relationship with fungal partners. The ancestors to land plants then gained rhizoids (Bryophytes) or root (Tracheophytes) for anchorage and became sessile (Harris *et al.*, 2022).

These ancestral plants had to survive a range of challenging conditions on land, including desiccation, limited nutrient supply, abiotic stressors like high temperature and UV radiation and harmful microbes. They responded to these stressors by arming themselves with a plethora of adaptations, which ensured their survival on land (Figure 1.2) (Arteaga-Vazquez, 2016; Harrison, 2017).

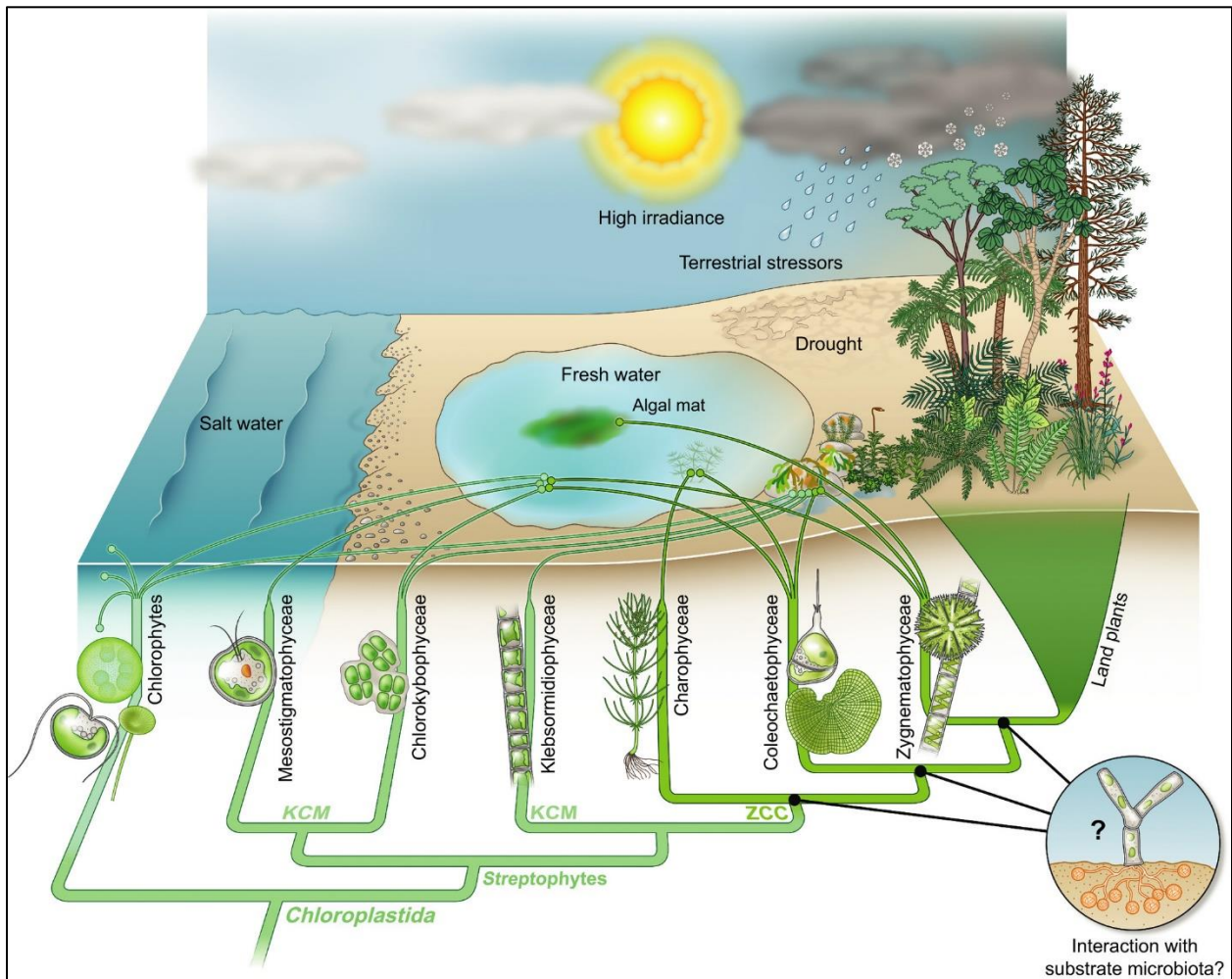


Figure 1.1 Streptophyte terrestrialization and the colonization of terrestrial habitats by extant plants (De Vries and Archibald, 2018). *Reproduced with permission from John Wiley and Sons.*

1.2. Adaptations which facilitated plant evolution

Being a dynamic process, it would be incorrect to think of evolution as merely the incorporation of plant adaptations in a sequential manner over time. Plant evolution entails both adaptation and exaptation of traits and functions to enable survival in specific niches.

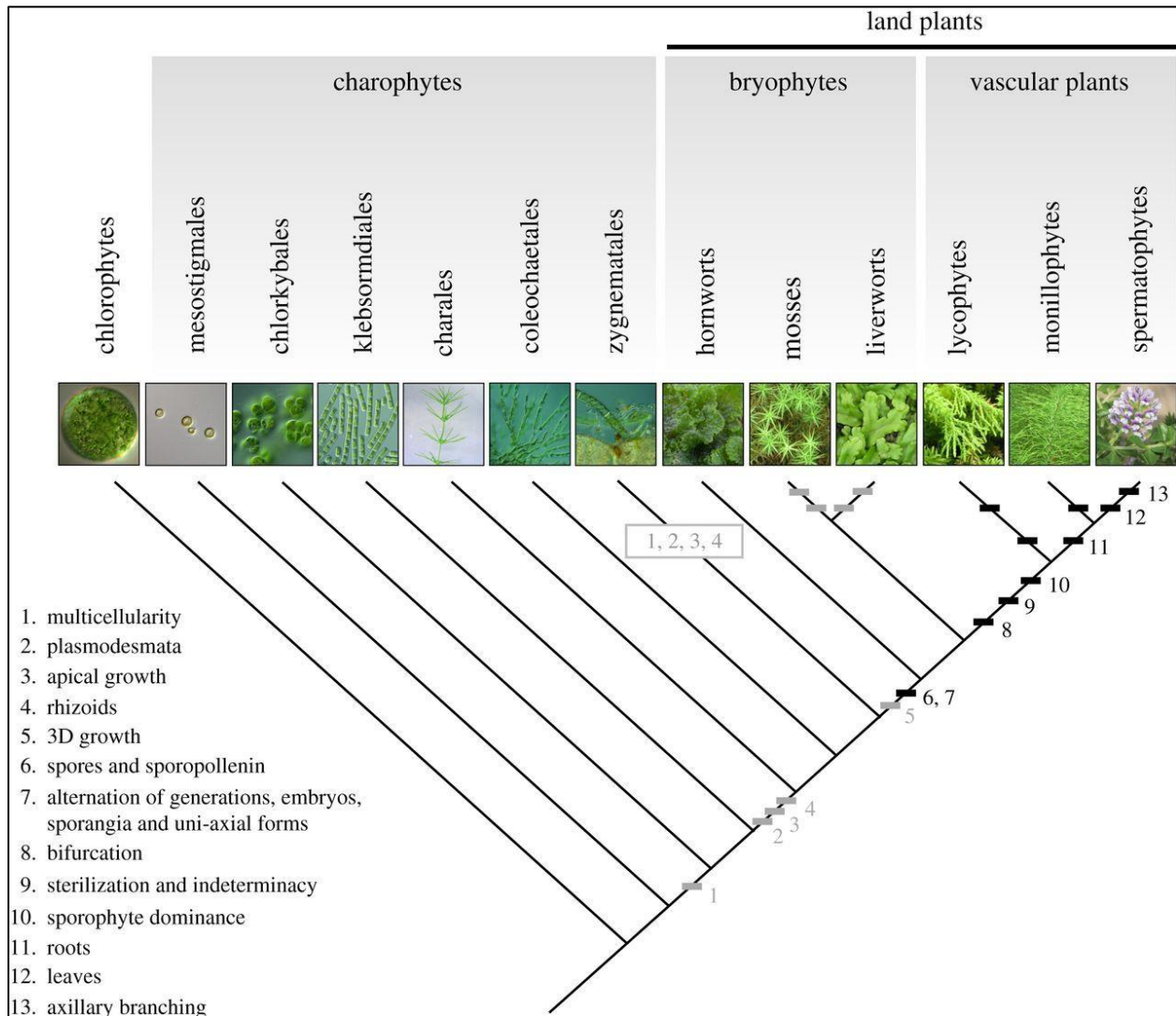


Figure 1.2 Adaptations underlying land plant evolution Gametophytic (grey bars) and sporophytic (black bars) innovations in the radiation of plant body plans. The earliest plant forms were unicellular freshwater algae, and land plants emerged from a grade of charophyte algae. Photos from left to right: *Erymosphaera sp.*, *Mesostigma viride*, *Chlorokybus atmophyticus*, *Klebsormidium flaccidum*, *Chara braunii*, *Coleochaete pulvinata*, *Spirogyra sp.* (Harrison, 2017). Reproduced with permission from Creative commons.

The evolution of land plants was preceded by several innovations adapted from their ancestors, including multicellularity, intercellular symplastic connections via plasmodesmata, specialized apical cell fates, and rhizoids (Figure 1.2). The simultaneous evolution of 3D apical growth also played a role in the origin of land plants. Reports also suggest that the evolution of spores and spores with desiccation-tolerant coats occurred before the advent of multicellular sporophytes. The earliest sporophytes were single-axial and ended in the formation of sporangia, but later formed bifurcations. This change came about much before the development of indeterminate axial forms in vascular plants. An axillary branching pattern emerged in both the precursors of spermatophytes and in liverwort and moss gametophytes. The evolution of vascular plants marked a shift towards a sporophyte-dominated life cycle, and roots and leaves evolved independently in different vascular plant lineages. Incorporation of such significant body plan transformations over time raises questions about the genetic mechanisms underlying these developmental transitions.

1.2.1. Genetic adaptations driving plant terrestrialization and evolution

The control of gene expression is a crucial aspect for the advancement of morphological diversity, playing a key role in the development and maintenance of developmental processes that allow plants to adapt to life on land. The availability of sequenced and annotated genomes from several reference streptophyte genomes including Streptophyta lineages, including basal angiosperms, gymnosperms, lycophytes, bryophytes, and Charophyta algae have heralded rapid identification, characterization and evolutionary conservation of the underlying genetic toolkit of plant evolution.

Most of the genes encoding transcription factors follow the Gene (Dosage) Balance hypothesis, according to which genic preservation only occurs if the rest of the genes involved in that particular pathway are also conserved (Birchler and Veitia, 2010). As we move across the green plant lineage, it is interesting to note that while some gene families were expanded over time, others were retained as distinct genes with minimal changes, as a result of strong selection pressure (Rensing, 2020) (Figure 1.3). Transcriptional regulators, such as KNOX/BELL genes and bHLH genes, are conserved across the green lineage (Bowman *et al.*, 2016; Chater *et al.*, 2016; Horst *et al.*, 2016). Contrastingly, transcription factors, such as Class III HD-Zip, Class IV HD-

Zip, WOX-type homeodomain proteins, GRAS, and MIKC-type MADS-box proteins, are found in both land plants and some forms of Charophyta algae but are absent in Chlorophytes (Romani and Moreno, 2021). Additionally, subfamilies of transcription factor genes, including VNS (VND-, NST/SND-, SMB-related), have not been detected in any of the sequenced Charophyta genomes and are exclusive to land plants (Romani and Moreno, 2021). This suggests that the acquisition of these new regulatory genes may be linked to the process of terrestrialization and the related changes in body plan.

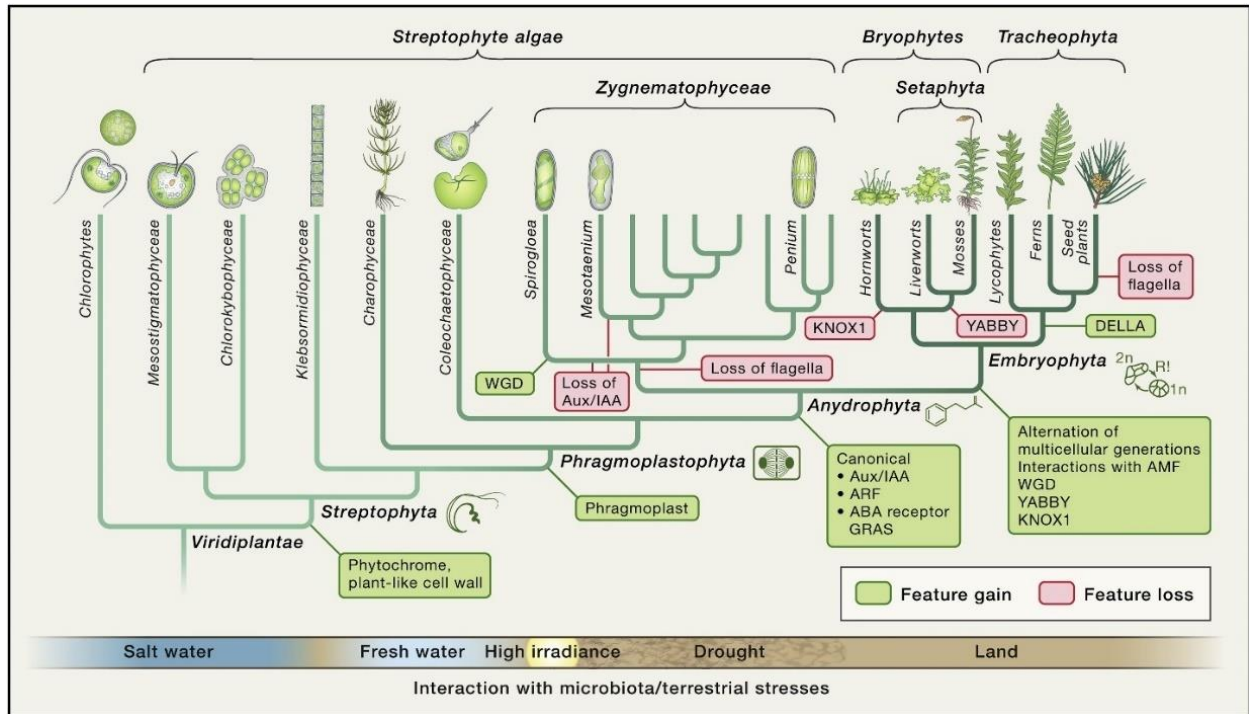


Figure 1.3 Dendrogram depicting the evolution of genetic pathways during the formative events of early plant evolution (Rensing, 2020). Reproduced with permission from Elsevier.

Interestingly, though transcription factors are evolutionarily conserved, it has been noted that their function may or may not be conserved between vascular and non-vascular plants. For example, the class I KNOX family genes play important roles in the meristem maintenance of angiosperms (Jasinski *et al.*, 2005; Yanai *et al.*, 2005). However, Sakakibara *et al.*, showed that the three KNOX1 orthologs in the non-vascular moss *P. patens* have shown that all three genes are expressed in the apical cell and meristematic region of the sporophyte and that they regulate the development of the diploid sporophyte, however, they play no roles in the development of the

gametophytic meristem (Sakakibara *et al.*, 2008). This suggests that the KNOX1 gene regulatory network for the sporophytic apical meristem is conserved among land plants and was most likely established early in the evolution of Embryophyta.

The loss-of-function mutants of KNOX2 in *A. thaliana* showed that KNOX2 acts in opposition to KNOX1 and modulates the differentiation of all aerial organs in the sporophyte body of *A. thaliana* (Smith and Hake, 2003). Additionally, disruption of the KNOX2 genes in *P. patens*, caused apospory, i.e., the formation of gametophyte-like structures from sporophytes without meiosis (Sakakibara *et al.*, 2013). This further highlights the critical role of KNOX2 genes in suppressing gametophytic developmental programming within the diploid sporophyte. The KNOX1 and KNOX2 subfamilies most likely arose from a single KNOX gene in an early Streptophyte alga through gene duplication, which allowed for the development and further evolution of more complex gene regulatory networks in the multicellular diploid sporophyte (Bharathan *et al.*, 1999; Reiser, Sánchez-Baracaldo and Hake, 2000; Furumizu *et al.*, 2015). This also highlights the fact that the development of a dominant indeterminate sporophytic apical meristem was a key innovation in vascular plants and that the existing toolkit of non-vascular plants was repurposed to enable this. It is evident from these comparative genetic analyses among angiosperms and bryophytes that the role of conserved genetic players appears to be ploidy specific. This underscores the crucial significance of investigating haploid-specific genes in bryophytes as a means to unravel the developmental pathways governing the regulation of gametophytic body plan development.

1.3. Significance of lineage-specific genes in land plant evolution

Most plant genomes share a common genetic toolkit, which governs multiple aspects of growth and development. However, some clade-specific or lineage-specific genes (LSGs) are the exceptions to this rule. Wu and Lambert have postulated that these LSGs can be formed in many ways and can range from being functionally similar to the parent locus from which they duplicated or may consist of novel sequences originated *de-novo* to govern lineage-specific processes (Wu and Lambert, 2022). Some common genomic features that have been recognized in LSGs include a (i) short length, (ii) lesser number of introns, (iii) an unusual GC content, and (iv) enhanced evolutionary rates. Gene duplications are one of the drivers behind generating evolutionary

innovations. They can sometimes also cause partial truncations and domain shuffling, and frameshift mutations inside the coding sequence leading to the generation of novel proteins with clade-specific functions. LSGs may also emerge de novo from non-coding sequences or through a two-stage evolutionary process involving an initial rapid phase followed by a gradual sequence evolution of genes specific to outgroup lineages. Cross-species lateral gene transfer and subsequent neo-functionalization may also result in the evolution of clade-specific genes. In this case, partial sequence homology may be observed in other species. It has long been hypothesized that LSGs are quite likely to be involved in regulating some kinds of clade-specific morphological and metabolic functions (Weisman, Murray and Eddy, 2020; Wu and Lambert, 2022).

When complex new traits emerged in the plant lineage, likely, a gene that originated and became established at the same evolutionary moment may be responsible, because the trait and the establishment event occurred around the same period. This is also applicable and perhaps more accurate for genes which seem to have arisen spontaneously, compared to those specific to a clade. When a pre-existing gene duplicates, it may be used for multiple processes because it is a functional protein with existing molecular interactions. Conversely, the emergence of entirely novel protein sequences is more likely to coincide with the development of complex traits that necessitate new molecular functionalities. This may also explain the loss of clade-specific genes in plant lineages where the particular function being performed by the LSGs is no longer required or may be performed by other conserved genes of the toolkit (Weisman, Murray and Eddy, 2020; Wu and Lambert, 2022).

Most of the LSGs reported in both mammalian, fungal and fungal lineages belong to either structural proteins, enzymes involved in responding to environmental stress and pathogens, or specific signalling pathways like ubiquitin ligase E3 subunits and transcription factors. In nematodes and flies, LSGs have been identified to function as small-molecule kinases and methylases (Lespinet *et al.*, 2002). Recent years have seen the functional characterization of many LSGs in flowering plants. Donoghue *et al.*, demonstrated that in *Arabidopsis*, as much as ~13% of genes encode proteins having no conserved protein domains and may have lineage-specific functions (Donoghue *et al.*, 2011). For example, the Qua Quine Starch (QQS, AT3G30720) has been shown to control metabolic responses to internal and environmental stresses (Li *et al.*, 2009; Arendsee, Li and Wurtele, 2014; Li and Wurtele, 2015). Similarly, the Constitutive Expresser of

PR Genes 5 (CPR5) (Jing *et al.*, 2007) and others (Horan *et al.*, 2008; Mentzen and Wurtele, 2008; Luhua *et al.*, 2013) are involved in ensuring plant survival in response to biotic and abiotic stressors. Also, a rice-specific gene, OsDR10 (*Oryza sativa* defense-responsive gene 10), has been reported to be involved in suppressing the Salicylic acid dependent pathway and up-regulating the Jasmonic acid signalling to regulate the differences in the plant's response to pathogenic infections compared to other plant species (Xiao *et al.*, 2009).

Recently, a bryophyte-specific gene *SHORT-LEAF (SHLF)*, isolated from the moss *Physcomitrium patens*, has been reported to govern moss gametophytic development by influencing the auxin distribution pattern in moss stems (Mohanasundaram *et al.*, 2021). The subsequent loss of SHLF in pteridophytes and angiosperms implies its involvement in mediating clade-specific functions. SHLF is a protein characterized by a tandem direct repeat (TDR) structure, lacking any known conserved domains, which poses challenges in deciphering its molecular role in moss. Given that repeat-containing proteins often act as catalyst hubs, unraveling the nature and function of SHLF TDRs could provide insights into the mechanistic underpinnings of SHLF's function in moss.

1.4. Evolution and significance of repeat-containing proteins in plant development

The presence of a large genome provides plasticity and helps plants to survive the unfavorable conditions. It has also been reported that many plant genomes including the moss (*P. patens*) have undergone several whole genome duplications, which has helped to expand gene families by providing a repertoire of genes with possibly redundant and overlapping functions (Rensing *et al.*, 2008). Multigene repeat families in plants helps them to adapt to harsh environmental conditions. Sharma and Pandey put forward the hypothesis that the high propensity of repeat proteins in plants is thought to be the result of a high frequency of internal tandem duplication (Sharma and Pandey, 2016).

These repeats may be the result of multiple sequential duplication events which can cause variation in the number and sequence of repeats, even among related genes. These internal tandem repeats (TR) may be further divided into three categories based on the length of the repeats and

their functions (Katti, Ranjekar and Gupta, 2001). Shorter repeats of about 2-20 amino acid residues cannot exist as standalone structural units and only help in enabling repeat interactions.

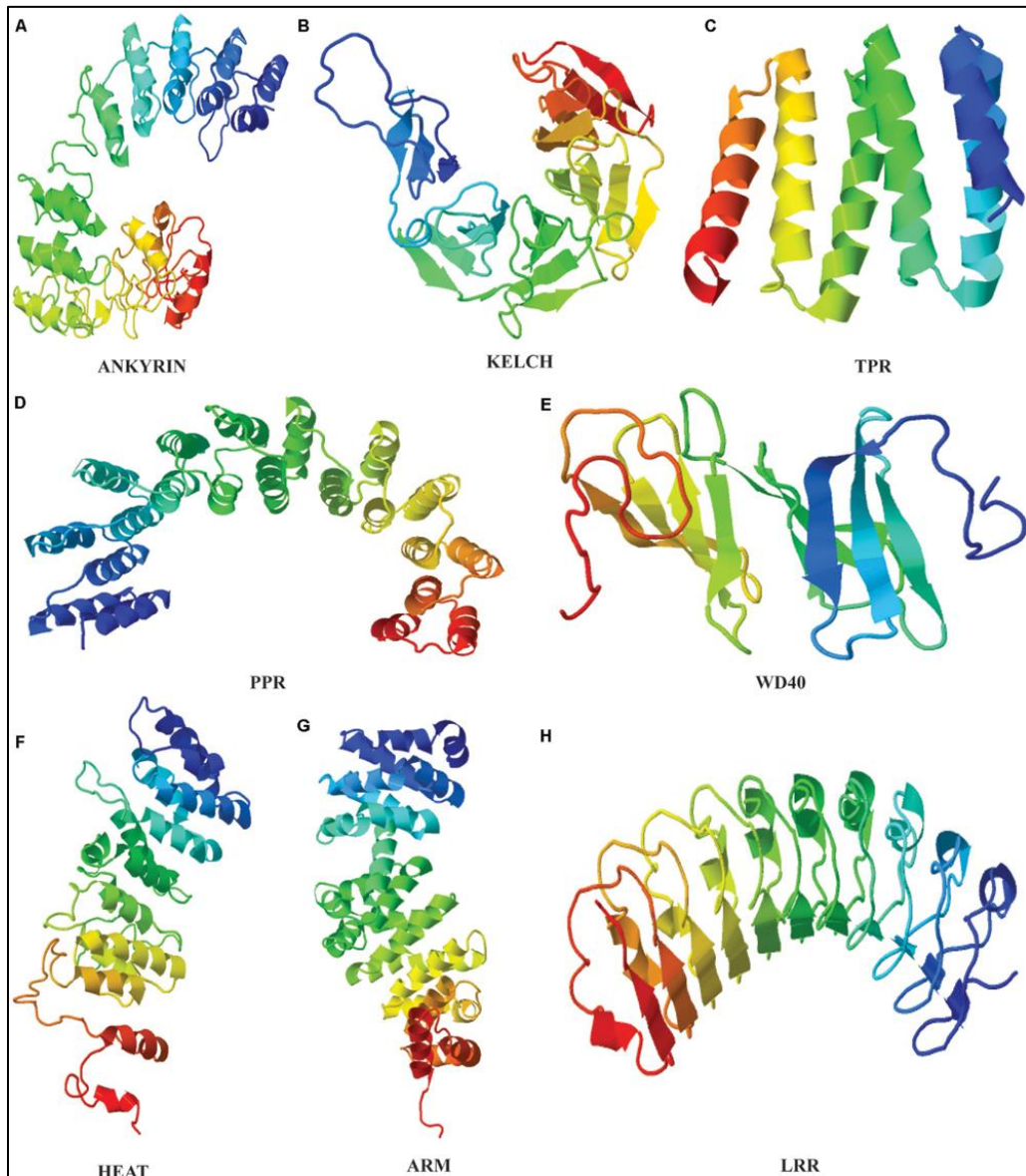


Figure 1.4 Tertiary structures of plant repeat proteins. Various domains such as the ankyrin repeat domain of BDA1, the Kelch repeat domain of BSU1, the TPR repeat domain of AT1G01320, the PPR domain of PPR1, the WD40 domain of ATG18, the HEAT domain of ILITHIYA, the ARM domain of Arabidillo-1, and the LRR domain of RLK7 (Receptor-like-Kinase 7) (Sharma and Pandey, 2016) *Reproduced with permission from Creative commons.*

Slightly longer repeats, with around 20-40 amino acid residues, may function independently and can form three-dimensional structure to facilitate interactions between proteins.

Longer repeats consisting of more than 100 amino acids have the ability to both exist independently and perform autonomously at the structural and functional levels (Groves and Barford, 1999). Several repeat containing protein families have been reported to be functional in plants, including Leucine rich repeats (LRR), WD40, HEAT, Kelch-like repeats, Armadillo (ARM), Ankyrin (ANK), Tetratricopeptide (TPR), and Pentatricopeptide repeats (PPR). These repeat containing domains play crucial roles in plant cell physiology, stress response, and development (Figure 1.4). The repeats act either as scaffolds to form functional multi-protein complexes or may associate with other domains to mediate protein function (Sharma and Pandey, 2016; Jain and Pandey, 2018). To the best of our knowledge, there are no report in plants, demonstrating proteolytic cleavage within the protein repeat region to generate the functional products.

1.5. Evolution and function of phytohormones in land plants

Interestingly, genes encoding components of phytohormone signalling and the plant's secondary metabolism are mostly conserved. This may be because of the plant's need to adapt to rapid environmental changes (Bowman *et al.*, 2019).

1.5.1. Roles of hormones in plant development

Hormones are the backbone of plants and regulate multiple aspects of both plant development and stress response. They can be mostly classified into two categories, phytohormones and peptide hormones. The common unifying trait in both phytohormone and peptide hormone signalling is the fact that they need to bind to specific receptors to activate their downstream signalling. In fact, proper plant development is mediated by an intricate interplay of multi-various combinations of phytohormones and peptide hormones.

1. **Phytohormones** – comprising of structurally un-related diverse group of small signalling molecules including the “classical” hormones like auxin (AUX), abscisic acid (ABA), cytokinin (CK), gibberellin (GA) and ethylene (ETH) along with the recently identified brassinosteroids (BR), Jasmonate (JA), salicylic acid (SA), nitric oxide (NO) and strigolactones (SL) (Figure 1.5) (Santner and Estelle, 2009).

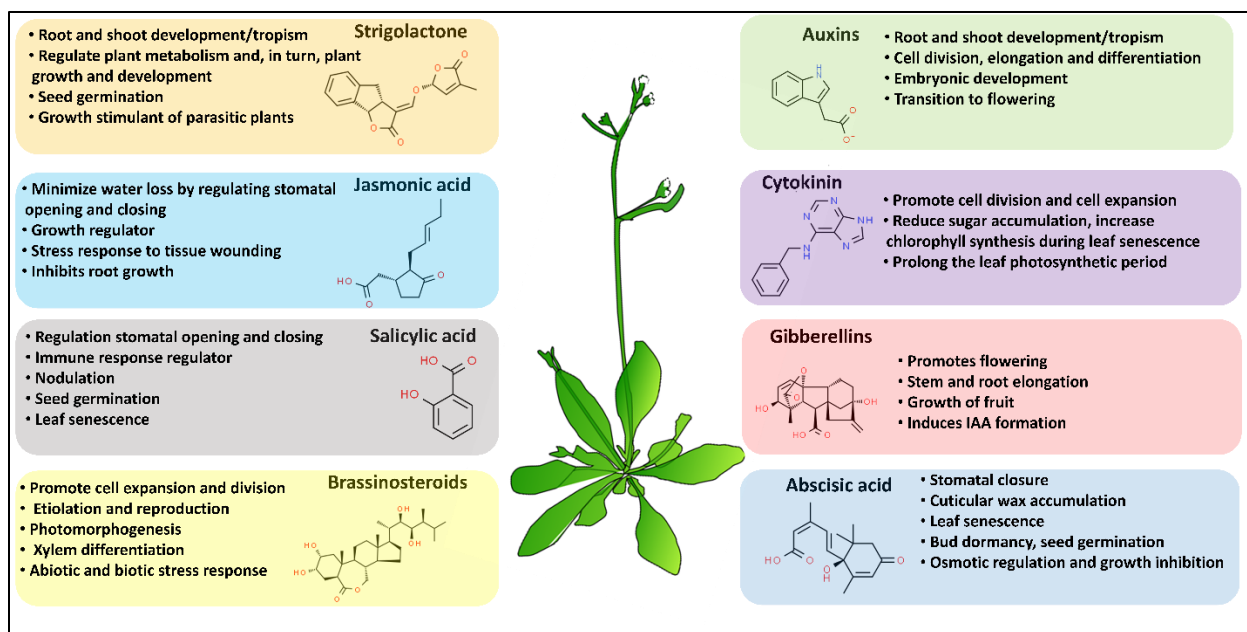


Figure 1.5 Major roles of phytohormones in plant development and defense

Auxin plays a central role in almost all aspects of plant growth and development (Leyser, 2018). Gibberellins (GAs) and cytokinins (CKs) regulate cell proliferation, differentiation, senescence, and complexity of leaves (Achard *et al.*, 2009; Kieber and Schaller, 2014). Abscisic acid (ABA), promotes stress resistance in plants during early stages of development, helping them survive adverse conditions (Sakata, Komatsu and Takezawa, 2014). Auxin, GAs, and CKs are typically associated with promoting differentiation, as they trigger developmental processes and regulate the final formation of plant organs. Contrastingly, ABA and ethylene (ETH) often act in concert to inhibit plant growth and induce maturity and aging. Additionally, research has shed light on the functions of other small molecule phytohormones, including brassinosteroids (BRs), jasmonic acid (JA), salicylic acid (SA), strigolactones (SLs) and karrkins (KARs). BRs are involved in almost every aspect of plant growth and adaptation to the environment (Nolan *et al.*, 2020), while JA and SA play critical roles in mediating plant stress response (Hassoon and Abduljabbar, 2019; J. Wang *et al.*, 2020). SLs and KARs are known to function in governing plant responses to abiotic stress. These hormone signalling pathways can be segregated into the following categories; firstly, the F-box mediated pathways for hormones like auxin, jasmonate, GA, SL and KARs, which is dependent on ubiquitin mediated repressor degradation, secondly, the two-component signalling pathways which is mediated by protein kinases and a phospho-relay

signal from the membrane receptors to downstream target genes, and lastly, the independent signalling pathways followed by abscisic acid and brassinosteroids (Santner and Estelle, 2009; Bowman *et al.*, 2019; Blázquez, Nelson and Weijers, 2020).

The signaling pathways of AUX, CK, and SL can be traced back to their origin in charophytes (Wang *et al.*, 2015; Bowman *et al.*, 2019). On the other hand, the ABA, JA, and SA signaling pathways evolved only in the most recent common ancestor of land plants (Wang *et al.*, 2015). GA signaling emerged after the divergence of bryophytes from land plants (Bennett, 2020). The canonical BR signaling originated in the last common ancestor of angiosperms, likely after the split between gymnosperms and angiosperms. The canonical ETH signaling emerged shortly after the emergence of angiosperms but before the split of monocots and eudicots (Wang *et al.*, 2015). Notably, while various plant hormone signaling mechanisms developed at different stages in the evolutionary history of plants, several signaling components were already encoded in the genomes of algae. This thesis aims to unravel the molecular intricacies of the bryophyte-specific protein SHORT-LEAF, a key regulator of auxin distribution in moss. To maintain a clear focus on the main objective, this thesis deliberately omits an in-depth exploration of the molecular roles of other phytohormones.

2. **Peptide hormones** – They are initially translated as pre-peptides which are further processed through proteolytic cleavage to form mature peptides. These can be subdivided into small post-translationally modified peptides (approximately 5–20 amino acids) and cysteine-rich peptides having an even number of cysteine residues (typically 6 or 8) which help in the formation of intramolecular disulfide bonds. Examples of signalling peptides include systemin, PSK (phytosulfokine), HypSys (hydroxyproline-rich glycopeptide systemin), Pep1, CLE (CLAVATA3/EMBRYO SURROUNDING REGION-related)/TDIF (tracheary element differentiation inhibitory factor), PSY (plant peptide containing sulfated tyrosine), CEP (C-terminally encoded peptide), RGF/CLEL/GLV (root meristem growth factor/CLE-like/GOLVEN), PIP (PAMP-INDUCED PEPTIDE), IDA (INFLORESCENCE DEFICIENT IN ABSCISSION) and CIF (Casparian strip integrity factor) subclasses (Figure 1.6) (Gancheva *et al.*, 2019) (Pearce *et al.*, 1991, 2001; Matsubayashi and Sakagami, 1996; Huffaker, Pearce and Ryan, 2006; Ito *et al.*, 2006; Ohyama, Ogawa and Matsubayashi, 2008; Ohyama *et al.*, 2009; Matsuzaki

et al., 2010; Okamoto *et al.*, 2013; Hou *et al.*, 2014; Schardon *et al.*, 2016; Doblus *et al.*, 2017; Nakayama *et al.*, 2017).

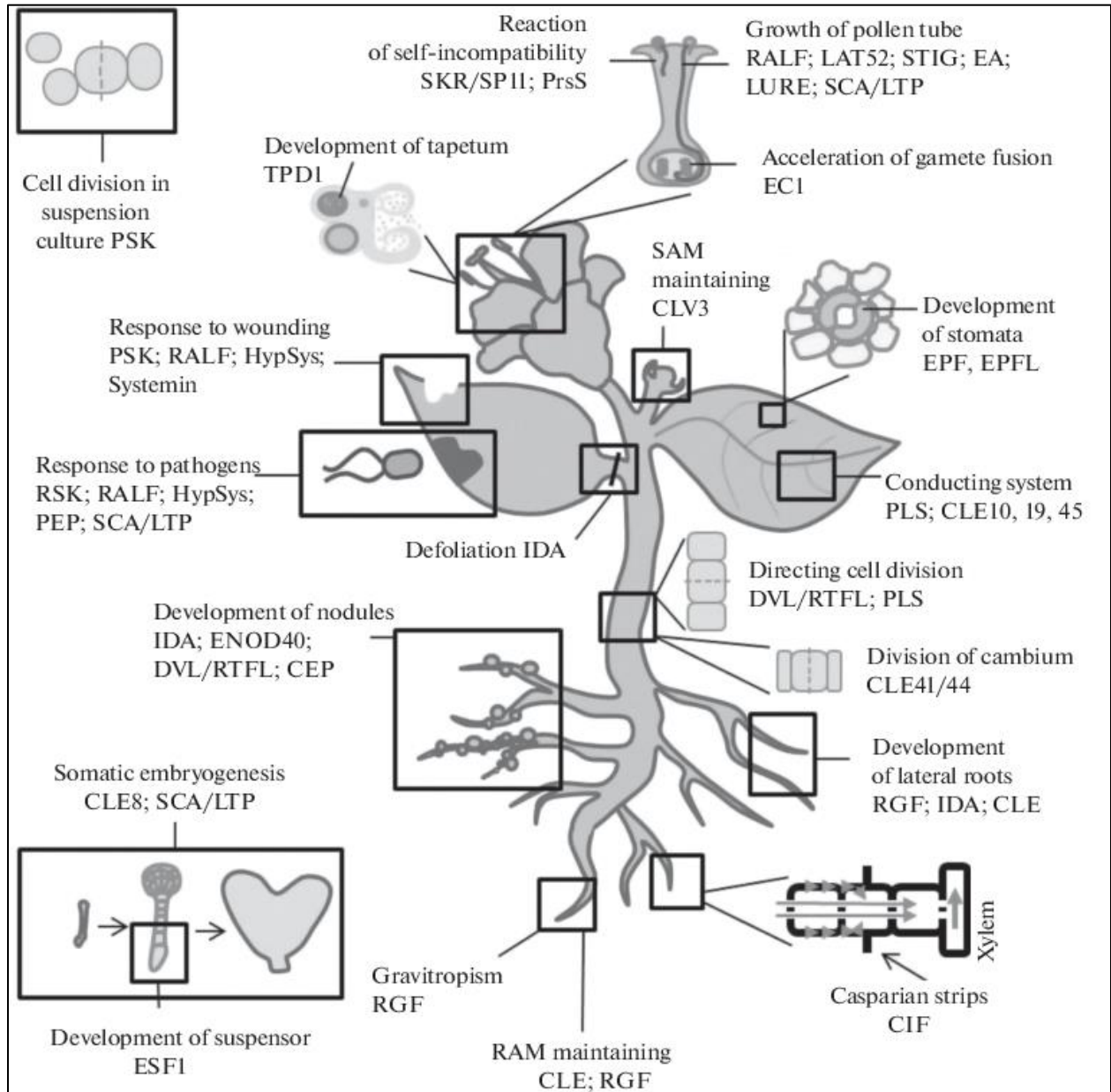


Figure 1.6 Peptide hormones in plant development (Gancheva *et al.*, 2019) *Reproduced with permission from Springer nature.*

The small signalling peptides are encoded as precursor proteins which undergo post-translational processing to form mature peptides. This is mediated through proteolytic cleavage by

proteases (Tamaki *et al.*, 2013; Engineer *et al.*, 2014; Schardon *et al.*, 2016). In cases, where the peptide is located at the very end of the pro-protein, as in the case of certain CLE and RGF family of peptides, a single proteolytic event is sufficient to cleave the N-terminus and release the mature peptide. On the other hand, when the peptides are encoded near the C-terminus (and not at the very end of the precursor) at least two cleavage events are required for peptide generation. There have also been instances of additional processing events observed in the variable region of the precursors, and even though these cuts do not mark either the N- or C-terminus of the mature peptide, they can be necessary for peptide formation (Ghorbani *et al.*, 2016). Such cleavage events are referred to as pre-processing or, when required for maturation, as pre-activation steps. Additionally, these peptides can also be modified by the addition of extra groups including tyrosine sulfation, proline hydroxylation, and hydroxyproline arabinosylation of specific residues by modifying enzymes (Hieta and Myllyharju, 2002; Tiainen, Myllyharju and Koivunen, 2005; Yuasa *et al.*, 2005; Komori *et al.*, 2009; Ogawa-Ohnishi, Matsushita and Matsubayashi, 2013).

With an ever-increasing number of plant genomes being sequenced, the community has gained valuable insights into the evolutionary conservation of these pathways. For instance, the genome of the extant moss *P. patens* genome, has genes which encodes proteins regulating auxin, abscisic acid, and cytokinin signaling, which are missing from the green algal genomes not, indicating that these pathways developed during land colonization by plants. On the other hand, components of gibberellin, ethylene, and brassinosteroids signaling pathways did not seem to have evolved until after moss and vascular plants split in the evolutionary timescale. Though peptide hormones such as PSK (PHYTOSULPHOKINE) (Tost *et al.*, 2021), CIF (CASPARIAN STRIP INTEGRITY FACTORS) (Furumizu *et al.*, 2021), RALFs (RAPID ALKALIZATION FACTOR) and CLE (CLAVATA3/EMBRYO SURROUNDING REGION-related) peptides have been found in the genomes of non-vascular plants such as *P. patens*, most of their physiological roles are yet to be properly understood. (Furumizu *et al.*, 2021). The role of CLE peptides in mediating 2D to 3D growth transition in moss has been recently established (Whitewoods *et al.*, 2018), while RALF peptides have been linked to both protonemal tip growth and immune response in moss (Ginanjar, Teh and Fujita, 2021; Mamaeva *et al.*, 2022). Additionally, given the evolutionary gap between vascular and non-vascular plants as well as the predisposition for lack of functional conservation of genes governing haploid development in bryophytes, it may be assumed that certain clade-

specific peptide hormones may also be involved in the regulation of certain non-vascular plant specific processes.

1.6. Role of auxin in plant development

Auxin regulates various aspects of plant growth and development in various environmental conditions. Even in small amounts, auxins can control gene expression by interacting with specific proteins and transcription factors that respond to environmental changes in a signaling pathway. Rapidly dividing cells are active sites for auxin synthesis which is subsequently transported by specific proteins that regulate the movement of the hormone into and out of cells. The physiological function of auxin is spatiotemporally regulated through three major regulatory measures: auxin biosynthesis, gradient-directed transport, and signal transduction (Gomes and Scortecci, 2021).

1.6.1. Origin and evolution of auxin transport

Plant development is mediated by the establishment of necessary phytohormone gradients through hormone transport from sites of synthesis to places where they are required. This occurs by regulating a tight control over cellular auxin influx and efflux. Auxin transport in plants mostly follows two pathways; namely, the polar auxin transport (PAT) and non-polar symplastic diffusion (Figure 1.7) (Reviewed in details in (Vosolsobě, Skokan and Petrášek, 2020). PAT is responsible for transporting auxin over long distances, where cell-to-cell transport via plasmodesmata transports auxin over shorter distances (Liu, Xu and Chua, 1993, Steinmann *et al.*, 1999; Reinhardt, Mandel and Kuhlemeier, 2000). Plants in which genes regulating PAT are either mutated or inhibited by chemicals, show a plethora of developmental defects, resulting in a pin-like shoots lacking lateral organs because of absence of proper polar transport of auxin resulting in a defective auxin gradient at the shoot apex (Okada *et al.*, 1991; Fischer *et al.*, 1997; Galweiler *et al.*, 1998; Noh, Murphy and Spalding, 2001; Benjamins, Malenica and Luschnig, 2005). Indeed, auxin acts as a morphogen in a concentration dependent manner to regulate plant development (Swarup and Bennett, 2003; Tanaka *et al.*, 2006).

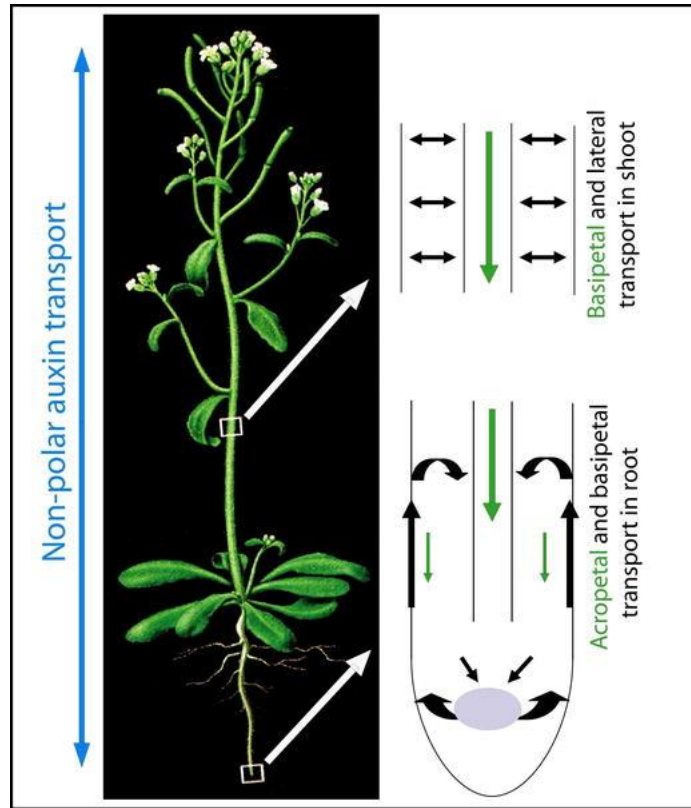


Figure 1.7 A schematic representation of non-polar and polar auxin transport in *Arabidopsis thaliana* (Michniewicz, Brewer and Friml, 2007). Reproduced with permission from American Society of Plant Biologists.

1.6.2. Evolutionary conservation of polar auxin transport across land plant lineage

Polar auxin transporters can be divided into four major categories including AUXIN-RESISTANT1/LIKEAUX1 (AUX1/LAX), the PIN-FORMED (PIN), and the B subgroup of ABC transporter (ABCB) families and PIN-LIKES (PILS) (Zazimalova *et al.*, 2010; Hammes, Murphy and Schwechheimer, 2022). Most members of the first three families localize to plasma membrane (Mravec *et al.*, 2009; Ding *et al.*, 2012), while short PINs and PILS proteins are bound to the ER membrane (Schwuchow, Michalke and Hertel, 2001; Barbez and Kleine-Vehn, 2013). AUX1/LAX isoforms function as importers, while long PINs export auxin (Yang *et al.*, 2006). Most of the ABCBs are auxin efflux carriers, however some members have also been reported to act as importers (Figure 1.7) (Geisler *et al.*, 2005; Santelia *et al.*, 2005; Terasaka *et al.*, 2005; Kamimoto *et al.*, 2012; Ofori *et al.*, 2018; Zhang *et al.*, 2018). Additionally, certain promiscuous “moonlighting” transporters like NRT1.1/ NPF6.3/ CHL1 WAT1/UmamiT5 act as multiple

transporters for auxin and other substrates (Jefferson *et al.*, 2009; Beeckman and Friml, 2010; Krouk *et al.*, 2010; Ranocha *et al.*, 2013).

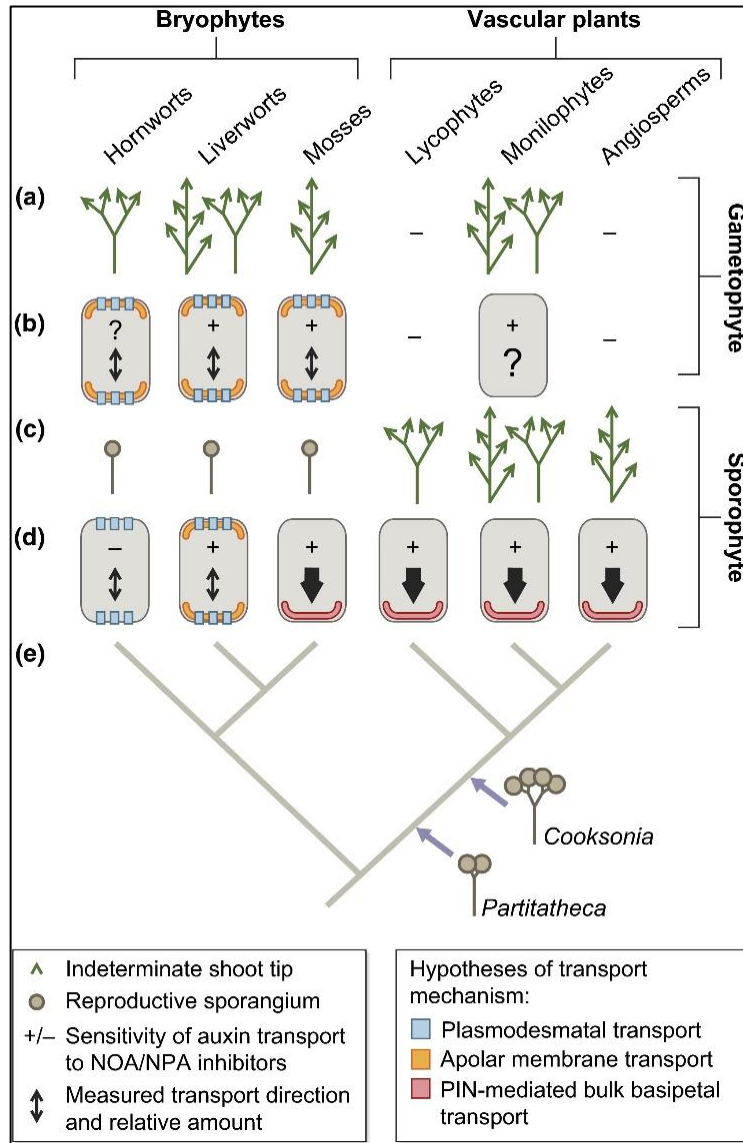


Figure 1.8 Auxin transport pathways in the evolution of branching forms Transport patterns in moss gametophores suggest bidirectional transport via membrane transporters or plasmodesmata, with evidence linking auxin transport to branching in *Physcomitrium*. While there is some PIN-mediated auxin transport in gametophores, it has a minor role in regulating branching compared to the PD mediated symplastic mechanism. In moss sporophytes, bulk basipetal auxin transport is PAT inhibitor (PAT I)- and NOA (naphthoxyacetic acid)-sensitive, with disruption of PIN function leading to sporophytic abnormalities such as bifurcation (Harrison, 2017). *Reproduced with permission from John Wiley and Sons.*

Studies on the evolutionary trajectory of the auxin transporters have shown that while ABCBs and PILS are ancient auxin transporters, PINs and AUX1/LAXs are more recent, with PINs being selectively more present in charophytes than chlorophytes (Vosolsobě, Skokan and Petrášek, 2020). As AUX1/LAX transporters only occur in some charophytes and chlorophytes, their evolutionary timeline is not clearly understood (Viaene *et al.*, 2014; Skokan *et al.*, 2019; Zhang *et al.*, 2019). Phylogenetic studies show that some basal charophytes seem to have only ABCBs and lack all other transporters, with the exception of *Chara*, which contains only PIN and ABCB-type transporters (Vosolsobě, Skokan and Petrášek, 2020). Overall, it is evident that the four transporter classes have all evolved independently and with the exception of *Klebsormidium* sp, do not occur together in a single algal lineage

. In non-flowering plants like mosses, auxin has been shown to affect developmental processes like apical dominance, tropic response, stem cell reprogramming etc., to regulate both protonemal and gametophyte (Eklund *et al.*, 2010; Prigge *et al.*, 2010; Lavy *et al.*, 2012; Pires *et al.*, 2013). Moreover, though polar auxin movement has also been detected in the sporophyte seta (Poli, Jacobs and Cooke, 2003; Fujita *et al.*, 2008), its mode of action in mediating diploid stage development is unknown. Additionally, it has been shown that bulk basipetal polar auxin transport is absent in moss gametophores which mostly relies on symplastic diffusion for development (Figure 1.8) (Harrison, 2017) (Fujita *et al.*, 2008; Coudert *et al.*, 2015). This is in stark contrast to the developmental programs of flowering plants, further highlighting the convergent evolution of plant shoots (Figure 1.8) (Fujita *et al.*, 2008; Donoghue *et al.*, 2021). Though bulk flow of auxin through PAT has not been detected in gametophores, it has been hypothesized that to drain auxin its sites of synthesis (leaves and meristem) auxin transport mostly occur in a localized manner in moss gametophores. Additionally, the bulk basipetal transport in mosses is PIN-mediated, and changes in PIN function are hypothesized to have contributed to the development of sporophytic branching forms (Harrison, 2017). It may also be possible that *Physcomitrium* PINs are able to distribute auxin mostly in the epidermal layers thereby lowering the overall auxin levels to below detection limit. In stomatophytes like *Physcomitrium*, PIN-mediated auxin transport has been shown to regulate a variety of developmental programs including, such as asymmetric cell division, gametophyte and sporophyte development, protonema and leaf development etc., (Bennett *et al.*, 2014; Viaene *et al.*, 2014). Interestingly, given the presence of long-distance auxin

transport in algae like *Chara*, it can be hypothesized that auxin transport, a key regulator of plant development, could have been recruited from the gametophyte to the sporophyte during land plant evolution.

1.6.3. Regulation of PAT by signalling peptides

Plant developmental pathways are influenced by the interplay of auxin response genes and peptide hormones. For example, peptide hormones such as PLS, AREP1, RGF/CLEL/GLV, and the Aux/IAA proteins such as IAA9, both act in concert to regulate root growth (Casson *et al.*, 2002; Matsuzaki *et al.*, 2010; Fernandez *et al.*, 2013; Yang *et al.*, 2014). Additionally, the CLE peptides act in an auxin- dependent manner to regulate vascular proliferation. Interestingly, many plant peptides have also been shown to cross-talk with auxin (Whitford *et al.*, 2008; Meng *et al.*, 2012). The expression of PLS, AREP1, and RGF/CLEL/GLV genes is induced by auxin, while in turn the PLS and RGF/CLEL/GLV peptides regulate auxin transport (Chilley *et al.*, 2006; Whitford *et al.*, 2008; Meng *et al.*, 2012). The members of the PAMP-INDUCED SECRETED PEPTIDES (PIPs) and PIP-LIKEs (PIPLs) modulate immunity, while PIP2 and PIPL3/TARGET OF LBD SIXTEEN 2(TOLS2) have been shown to be auxin responsive and regulate lateral root formation (Hou *et al.*, 2014; Vie *et al.*, 2015; Toyokura *et al.*, 2019). The POLARIS peptide of Arabidopsis has been shown to regulate both auxin transport and lateral root formation (Chilley *et al.*, 2006). Auxin has also been shown to act in a negative loop to suppress the expression of EPIDERMAL PATTERNING FACTOR-LIKE 2 (EPFL2) peptide and its receptor ERECTA to regulate the leaf tooth growth (Tameshige *et al.*, 2016). Additionally, the damage associated peptide Plant elicitor peptides (Pep) and their receptor kinases, i.e. the Pep1-PEPR have been shown to cross-talk with auxin to regulate both cell expansion and root cell differentiation in response to biotic and abiotic stress (Jing *et al.*, 2019). The majority of studies on auxin-peptide cross-talk have focused on flowering plants, leaving the dynamics of this interaction in non-flowering plants largely unexplored. Although CLE peptides have been suggested to be involved in auxin signaling in moss (Whitewoods *et al.*, 2018; Nemeč-Venza *et al.*, 2022), further investigation is needed in order to fully understand the role of auxin-peptide cross-talk in non-flowering plants.

1.6.4. Role of flavonoids in regulating PAT

The primary and secondary metabolic pathways of plants are highly interconnected and work together to mediate plant development. Polar auxin transport (PAT) in flowering plants has been shown to be directly affected by the phenylpropanoid biosynthesis pathway (Jacobs and Rubery, 1988; Pharis and Rood, 2012). Flavonoids, a group of phenylpropanoids, are the natural scavengers of reactive oxygen species (ROS) and can affect numerous developmental processes including auxin transport in plants (Nascimento and Tattini, 2022). They can also influence the activity of many protein families including auxin efflux carriers (Teale *et al.*, 2021). Flavonoids such as the Ortho-dihydroxy B-ring-substituted quercetin, affect the activity of serine–threonine PINOID (PID) proteins which regulate PIN localization (Peer and Murphy, 2006, 2007; Michniewicz *et al.*, 2007; Adamowski and Friml, 2015; Brunetti *et al.*, 2018). The fact that flavonoids can act as negative regulators of auxin transport has been well established with the *Arabidopsis transparent testa (tt4)* mutants having a mutation in *CHALCONE SYNTHASE (CHS)*, which is the gene encoding the first enzyme in flavonoid biosynthesis (Brown *et al.*, 2001). In comparison to WT plants, the *tt4* mutants had pleiotropic phenotypes, including three times as many secondary inflorescence stems, reduced plant height, decreased stem diameter, and increased secondary root development. Additionally, it was seen that WT *Arabidopsis* plants grown on naringenin (a biosynthetic precursor to those flavonoids like quercetin and kaempferol with auxin transport inhibitor activity), leads to a reduction in root growth and gravitropism, similar to the effects of synthetic auxin transport inhibitors like NPA (N-1-naphthylphthalamic acid) (Brown *et al.*, 2001). It was also seen that auxin transport was higher in the *tt4* mutant and was reversed when the plants were grown on the flavonoid precursor, naringenin. These results indicate that flavonoids act as endogenous regulators of PAT in flowering plants. Although there is limited proof of PIN-flavonoid regulation affecting non-flowering plant development, recent studies indicate that PINA in *P. patens* possess dual ER and PM localization, a characteristic of the 'ancestral' PIN6 protein (Friml and Jones, 2010; Simon *et al.*, 2016). Moreover, the *P. patens* PINA were found to be highly reactive to naringenin and could influence shoot development (Bennett *et al.*, 2014). However, the molecular players modulating this cross-talk is not well established in the bryophytes.

1.6.1. Effect of reactive oxygen species in the regulation of PAT

Reactive oxygen species (ROS) have been shown to act in a feedback loop with auxin and govern root development (Tognetti, Bielach and Hrtyan, 2017). Auxin can also induce the expression of ROS-related genes by activating ROOT HAIR DEFECTIVE SIXLIKE 4 (RSL4), which governs the elongation of root hair cells (Mangano *et al.*, 2017). Recent studies have also shown that the interplay of auxin and ROS together can induce the formation of lateral root primordia by controlling the polarity of PINs which causes changes in the pattern of cell division and affects root development. Additionally, it is known that the growth of Arabidopsis roots is influenced by glutathione (GSH)-dependent redox regulation, which is associated with changes in the expression of PINs (Koprivova, Mugford and Kopriva, 2010). Lastly, the Arabidopsis triple mutant *ntra ntrb cad2*, which lacks cytosolic reduced nicotinamide adenine dinucleotide phosphate thioredoxin reductases (NTRA and NTRB) as well as GSH biosynthesis, exhibited reduced expression of genes related to auxin transport and response (Eckardt, 2010). Increase in auxin levels in *Solanum lycopersicum* root tips has shown to induce the accumulation of H₂O₂, which subsequently caused inhibition of elongation of root cells and hindered root growth (Ivanchenko *et al.*, 2013). Figure 1.10 is a diagrammatic representation of how redox processes regulate auxin transport and signaling (Adapted from Considine and Foyer, 2014)

Flavonoids may also act as buffers of cellular ROS levels, allowing plants to respond promptly to environmental changes. Under stress, high levels of ROS can activate specific MAPK (ANP1 kinase in Arabidopsis and NPK1 in tobacco), which represses auxin signaling while transducing oxidative stress signaling (Figure 1.9) (Kovtun *et al.*, 2000; Gayomba, Watkins and Muday, 2017, Brunetti *et al.*, 2018). The massive generation of H₂O₂ may trigger ANP1-mediated MAPK cascade and help stressed plants to prioritize stress protection over auxin-related activities (Kovtun *et al.*, 2000). However, the specific ways in which flavonoids regulate cytoplasmic ROS levels involved in IAA-oxidation and IAA-signaling are not yet fully understood.

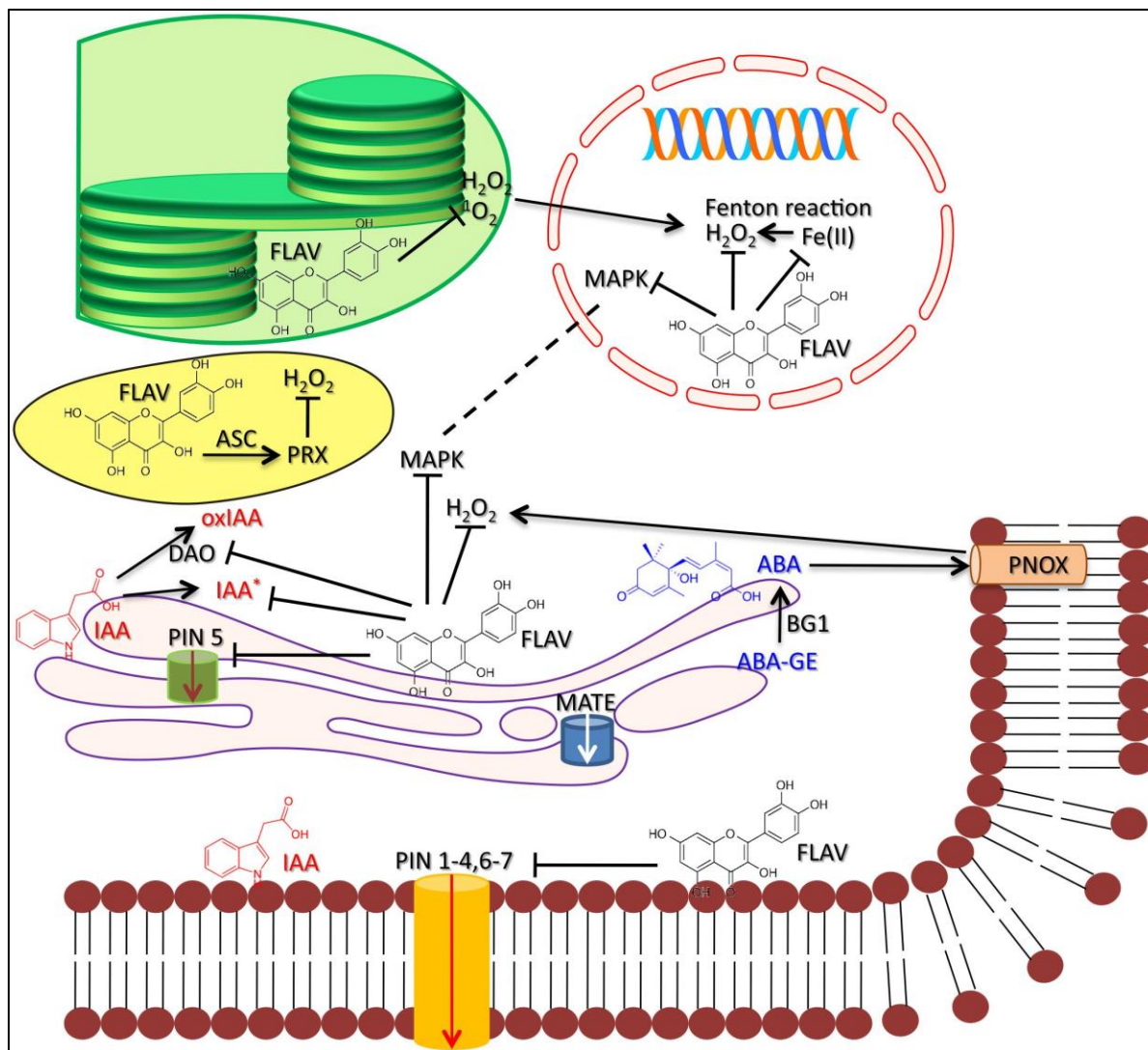


Figure 1.9 Proposed regulatory network of auxin and flavonoids Under stress (like high light) conditions, IAA biosynthesis is activated, leading to an increase in flavonoid biosynthesis. IAA is synthesized at the same site as flavonoid biosynthesis. The bZIP transcription factor HY5 is thought to play a role in IAA- induced flavonol biosynthesis by activating the expression of MYB12. Flavonols located in different cellular compartments regulate IAA signaling. ER-located flavonoids may inhibit the activity of the auxin transport protein PIN5 and possibly PIN6 and PIN8, which transport auxin into the ER lumen. Flavonols are transported to the plasma membrane, where they inhibit the cell-to-cell movement of auxin by acting on 'long' PINs and PIN6. They may also alter the catabolism of auxin and limit the generation of IAA radicals. (Brunetti *et al.*, 2018). *Reproduced with permission from Creative commons.*

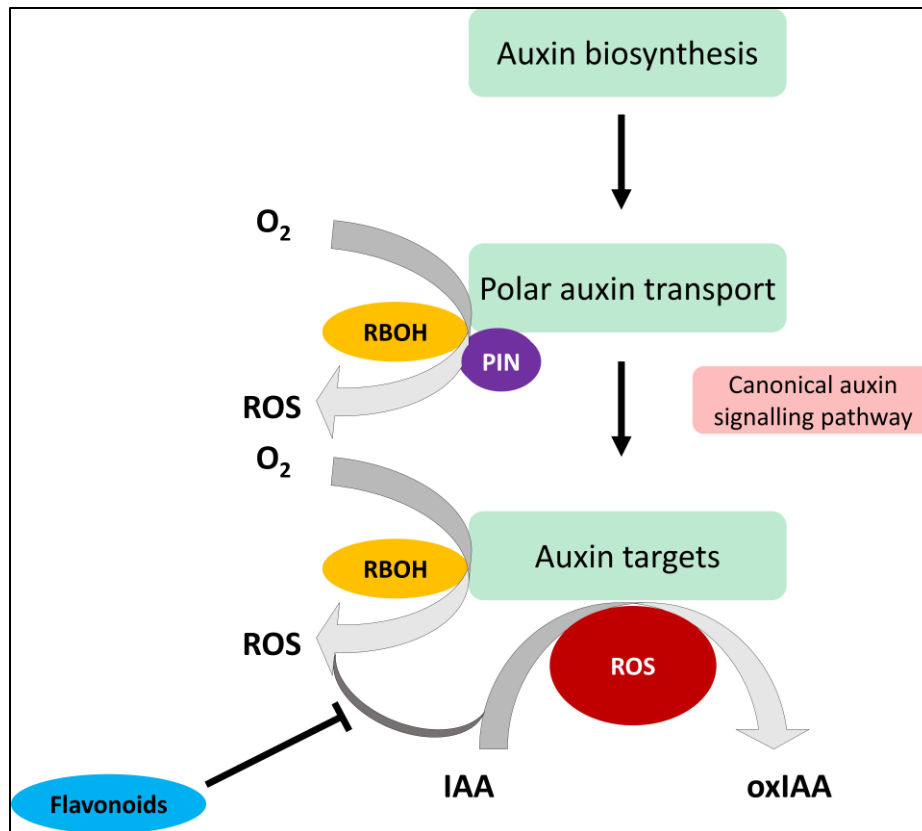


Figure 1.10 A diagrammatic representation of how redox processes regulate the signaling of auxin (IAA). The transport of auxin between cells occurs in a polar manner through efflux transporters like PIN proteins, and its distribution across tissues determines plant growth and development. The accumulation of auxin is linked to changes in cellular redox status, which involve the production of reactive oxygen species (ROS) by NADPH oxidases (also known as respiratory burst oxidase homologs [RBOH]) and redox components like the monothiol glutaredoxin, AtGrxS17. The resulting oxidation of IAA leads to the formation of oxIAA, which weakens the signaling of auxin. In this context, IAA refers to indole-3-acetic acid, oxIAA stands for 2-oxindole-3-acetic acid, and PIN denotes PIN-formed [adapted from (Considine and Foyer, 2014)].

All these studies have proved the existence of a direct relationship between plant auxin signaling and cellular redox status. In bryophytes however, unlike flowering plants, though evolutionary and functional conservation of ROS, flavonoids and auxin transporters has been reported, the molecular nature of the dynamics of ROS and auxin signalling, is not well established.

1.7. Functional conservation of non-polar auxin transport across land plant lineage

Besides active polar auxin transport, passive diffusion through plasmodesmata also contributes to multiple developmental paradigms such as phyllotaxis and SAM maintenance in flowering plants (Guenot *et al.*, 2012; Verna *et al.*, 2019; Mellor *et al.*, 2020). Plasmodesmata (PD) plasma membrane-lined channels entrapping a thin sheet of appressed ER called the desmotubule (Robards and Lucas, 1990; Faulkner, 2018; Sager and Lee, 2018; Li *et al.*, 2021). PD connect almost all the cells of a plant body via a common symplast and mediate intercellular communication (Sevilem, Yadav and Helariutta, 2015). The space between the desmotubule and the plasma membrane forms the cytoplasmic sleeve which enables symplastic transport (Li *et al.*, 2021).

Auxin molecules being small enough can easily pass through the PD (Rutschow, Baskin and Kramer, 2011; Han *et al.*, 2014). In fact, the flux of PIN mediated auxin transport and PD mediated diffusion were found to be nearly comparable (Kramer, 2004; Kramer, Rutschow and Mabie, 2011; Rutschow, Baskin and Kramer, 2014). Unlike the directed gradient establishment by polar PINs symplastic diffusion can distribute in all directions (Band, 2021). The plasmodesmal diffusion process reduces the auxin velocity produced by the PINs, allowing auxin to move between adjacent cell files and regulate development (Figure 1.11) (Mitchison, 1980; Rutschow, Baskin and Kramer, 2011). Additionally, it has been shown that PD mediated diffusion and influx carriers like AUX/LAX1 in root tips can act concertedly to create a reflux loop which effectively increases the total auxin content (Mellor *et al.*, 2020). Thus, it can be concluded that the overall auxin gradient and flux in a plant is established by the combined effects of both polar and symplastic auxin transport (Figure 1.11) (Band, 2021). In fact, a certain degree of cross-talk occurs between these two modes. Proteins that control plasmodesmata permeability are influenced by auxin levels, while auxin biosynthesis and transport are affected by intercellular communication (Paterlini, 2020). In terms of evolution, transporter-driven cell-cell auxin movement and plasmodesmata appear to have evolved concurrently in the green lineage, indicating their early coexistence and potential functional specificity. However, not much is known about the role and regulation of symplastic auxin diffusion in the development of non-vascular plants.

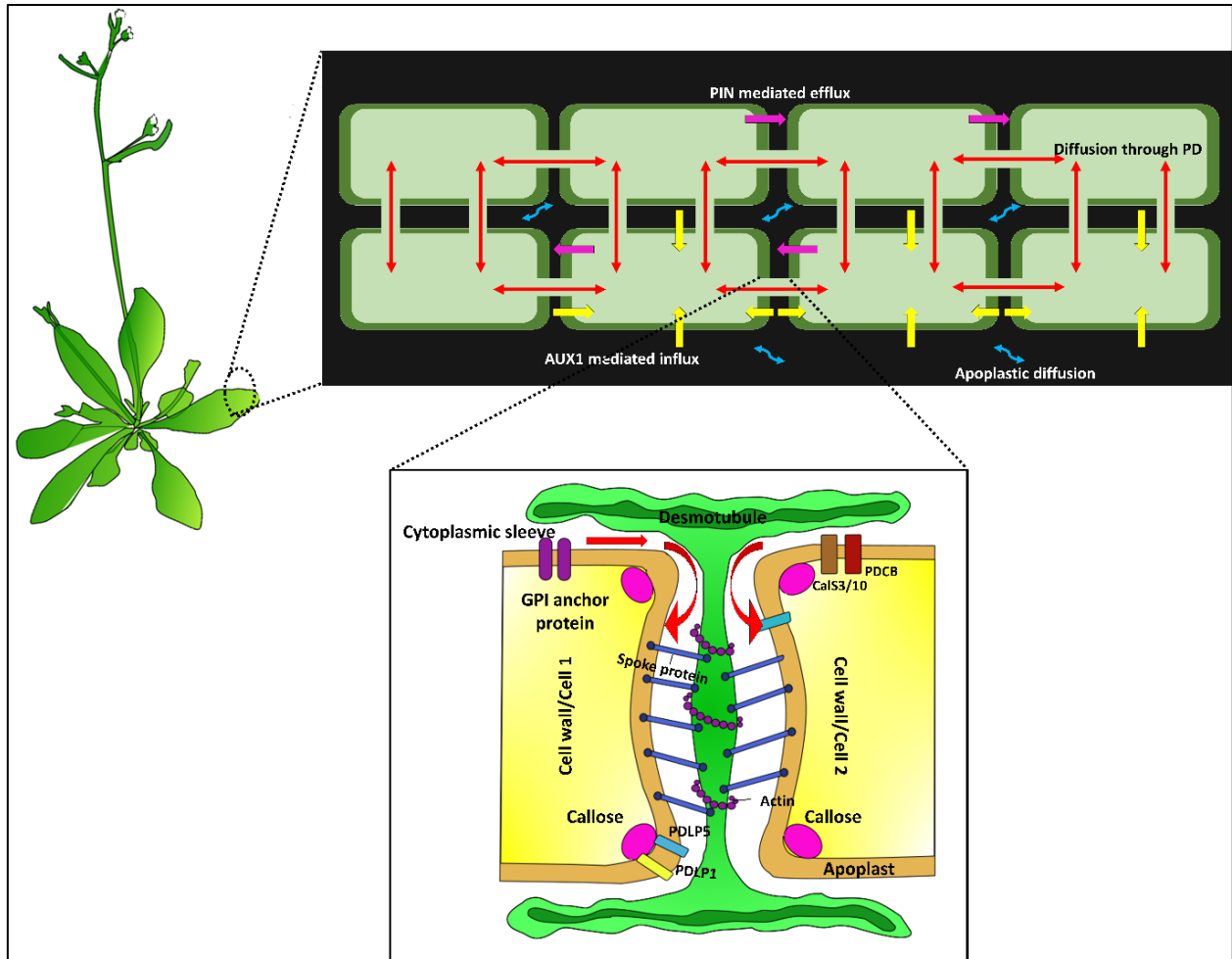


Figure 1.11 The key components of auxin transport are represented, namely polar auxin transport; with magenta and yellow arrows denoting PIN mediated efflux and AUX1 mediated influx. Non-polar auxin transport has been denoted as red and blue arrows showing PD mediated symplastic diffusion and apoplastic diffusion respectively. Plasma membrane-lined pores called plasmodesmata (PD) provide symplastic connectivity between neighboring cells. Compressed endoplasmic reticulum named desmotubule traverses through the pore and is associated with spoke proteins and cytoskeletal proteins (actin, myosin etc). Callose turnover at the neck region of PD, mediated by callose-binding proteins (PDCB) and callose synthase (CalS3/10) regulates the channel aperture. Various PD-localized such as GPI-anchor proteins and PD-localized proteins (PDLPs) have also been shown to modulate PD aperture and affect auxin transport.

1.7.1. Regulation of symplastic auxin transport

The passive transport of auxin through the plasmodesmata (PD) is under tight control through the regulation of plasmodesmal flux through changes in PD-associated callose dynamics, which physically restricts the pore size of PD (Figure 1.11) (Amsbury, Kirk and Benitez-Alfonso, 2018). As callose has a high turnover rate (Jaffe and Leopold, 1984), it is quite effective to controlling PD fluxes. Multiple genes in plant callose metabolism have been identified from flowering plants including callose biosynthesis genes including GLUCAN SYNTHASE LIKE 8 (GSL8) and CALLOSE SYNTHASE 3 (CALS3), (Chen *et al.*, 2009; Vatén *et al.*, 2011; Han *et al.*, 2014) and genes encoding callose degrading enzymes such as beta-1,3-glucanase (PdBG), (Iglesias and Meins Jr, 2000; Levy *et al.*, 2007); PLASMODESMATA CALLOSE BINDING1 (PDCB1), which binds callose in the cell wall around the plasmodesmata (Simpson *et al.*, 2009); and PDLP5, which stimulates callose deposition (Lee *et al.*, 2011; Wang *et al.*, 2013). Mutations in these genes have been shown to modify symplastic diffusion (Simpson *et al.*, 2009; Rutschow, Baskin and Kramer, 2011) and affect auxin distribution (Han *et al.*, 2014; Sager and Lee, 2018; Gao *et al.*, 2020; Mellor *et al.*, 2020). Interestingly, recent studies have also reported that auxin feeds back on its own transport, by increasing callose levels to restrict its own movement across PD. For example, on one hand in the root meristem, a reduction in auxin biosynthesis genes was observed upon disrupting symplastic transport into the quiescent center (QC) (Liu *et al.*, 2017), whereas on the other hand, ectopic PIN2 expression was detected upon induction of callose biosynthesis in root endodermis (Wu *et al.*, 2016).

Though these findings clearly show that PD mediated auxin diffusion through plasmodesmata would does not dissipate the auxin maxima, the exact mechanisms of this feedback remain unexplored. Additionally, it would be interesting to study the dynamics of auxin feedback on PD permeability in systems of plants like the moss *P. patens*, where symplastic auxin transport is the major route for auxin distribution in the gametophore stem.

1.8. Significance of bryophytes in the evolutionary history of land plants

Land plants evolved from the charophyte lineage, which includes unicellular ancestral forms and life cycles where meiosis occurs immediately after zygote formation (Ishizaki, 2017).

Over time, there was a trend towards the development of more complex multicellular algal forms with specialized cells and tissues, but the diploid life cycle stage remained unchanged (Vanderpoorten and Goffinet, 2009). The life cycle of land plants, on the other hand, features alternating multicellular haploid (gametophyte) and diploid (sporophyte) phases. The dominance of each phase changed from the gametophyte in bryophytes to the sporophyte in vascular plants as land plants evolved (Vanderpoorten and Goffinet, 2009). The diverse forms of land plants developed along separate paths in each life cycle stage. The major existing lineages of land plants, including hornworts, liverworts, mosses, lycophytes, monilophytes, and spermatophytes, were established around 360 million years ago. Their evolution helped shape soil formation, increased primary productivity, and had a significant impact on weathering and global climates (Vanderpoorten and Goffinet, 2009).

The term 'bryophyte' is derived from the Greek language and refers to plants that expand when they absorb water. It is a general name given to plants that have a life cycle consisting of alternating haploid and diploid generations with a predominant gametophyte. Bryophytes are unique among land plants in having a dominant and branching gametophyte, with a range of morphologies not found in tracheophytes (Crum, 2001). Bryophytes are reportedly monophyletic and the three main lineages of bryophytes include liverworts, mosses, and hornworts (Bold *et al.*, 1949; Nishiyama *et al.*, 2004; Su *et al.*, 2021). They all share some features, such as the presence of an embryo, which led to their classification as 'embryophytes' (Gerrienne and Gonez, 2011; Delwiche and Cooper, 2015; Harholt, Moestrup and Ulvskov, 2016). However, they also differ in several aspects, such as the architecture of the gametophyte and the sporophyte, making them easily distinguishable in the field. Unlike other land plants, the gametophyte in bryophytes does not have stomata (Harris *et al.*, 2020).

Bryophytes, which have been present for hundreds of millions of years and are considered the closest living relatives of the first terrestrial plants, played a critical role in the evolution of land plants. These small, often overlooked members of the plant world serve as a link between seed and vascular plants and their algal ancestors. The transition to land and the origin of vascular plants can be traced back to bryophytes, making them an important piece of the evolutionary history of land plants (Proctor, 2000). The evolutionary position of bryophytes in which they share their last common ancestor with flowering plants, makes them interesting candidates to study the

evolution of plant adaptations in the green plant lineage (Proctor, 2000; Vanderpoorten and Goffinet, 2009). However, it would be unwise to think of them to be the intermediates between algae and flowering plants. This viewpoint assumes that bryophytes to be somewhat frozen in time, which is quite misleading given the immense diversity and high biological variation within bryophytes, which like their sister groups have been subjected to the same evolutionary forces over time (McDaniel, 2021).

Bryophytes are mostly classified into three categories:

- The Marchantiophyta are the most structurally simple amongst embryophytes bearing a thalloid or leafy stem-based gametophyte, with parallelly arranged “leaves” or phyllids. Some taxa have specialized water conducting cells for endohydric transport in the gametophyte which are absent in the sporophyte (Edwards, Axe and Duckett, 2003).
- The Bryophyta (mosses) have a stem-based vegetative body with spiral rows of leaves. Axial water conducting strands occur in both generations of many taxa (Goffinet, Buck and Shaw, 2009). The seta is formed from an intercalary meristem and is usually unbranched. The sporangium is terminal and sheds an operculum. Stomata may occur on the capsule wall, but not the seta. The columella extends beyond the sporogenous layer. Spore germination results in a filamentous or thalloid sporeling, which develops into one to several gametophytes (Zechmeister, Grodzińska and Szarek-Łukaszewska, 2003).
- The Anthocerotophyta (hornworts) have a thalloid vegetative gametophyte but never bear leaves (Renzaglia *et al.*, 2009; Frangedakis *et al.*, 2021). Water conducting cells are absent in both generations. The sporophyte is linear, lacks a seta and matures basipetally. Dehiscence follows two longitudinal lines exposing a spore mass surrounding an axial columella.

The moss *Phycomitrium patens* (Hedw.) and the liverwort *Marchantia polymorpha* L. and are among the extant bryophytes that have genomes equipped with the traits necessary for ancestral land plants to endure terrestrial stressors and are being widely investigated to study the evolution of major developmental pathways in plants. (Rensing *et al.*, 2008; Bowman *et al.*, 2017).

1.9. The model moss *Physcomitrium patens*

A typical life cycle of moss passes through a variety of tissue types, which alternates between a dominant haploid (the gametophyte) and a transient diploid phase (the sporophyte). Microscopic observations reveal that the capsule of lone diploid tissue – the sporophyte, houses thousands of haploid spores (Figure 1.12). Spores, which are the beginning of the gametophytic generation, germinate and develop into protonemal apical cell which is committed to form two dimensional filaments, comprising of the chlorophyll-rich chloronema and the fast-growing caulonema (Reski and Abel, 1985). Only about five percent of caulonemal branch initial cells divide at markedly different planes to form the gametophore apical cell which divides further to form stem, leaf apical cell and gametangia apical cell (Kofuji and Hasebe, 2014).

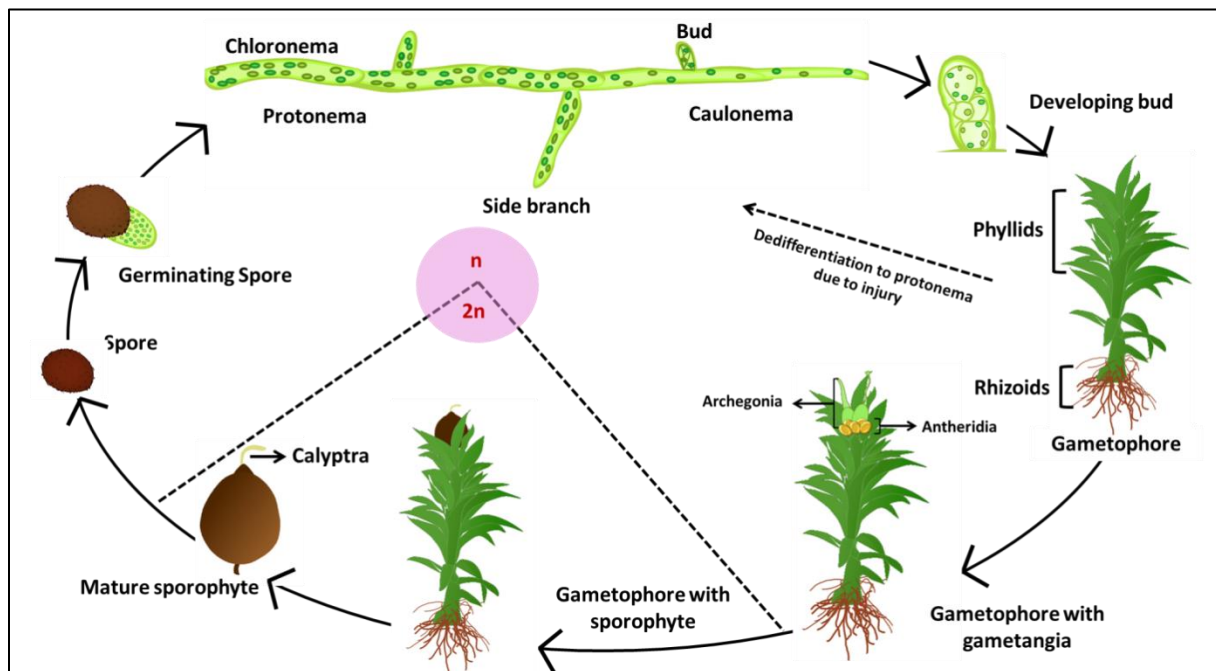


Figure 1.12 Life cycle of *Physcomitrium patens* Moss begins its life cycle with a haploid spore which germinates to form the filamentous 2D protonema. The protonema branches to form the bud which undergoes asymmetric cell divisions to form the gametophore. Under cool short-day condition, gametangiogenesis occurs in the monoicous *P. patens*. Upon successful fertilization of the gametes, a diploid sporophyte is formed which dehisces upon maturation to form the haploid spore, thereby completing the life cycle.

This developmental plan results in the formation of gametophore stem with spirally arranged leaves and gametangia. With the guidance of hormonal gradients, certain epidermal cells of the stem de-differentiate to form rhizoid apical cells, resulting in the formation of filamentous rhizoids (Sakakibara *et al.*, 2003; Jang *et al.*, 2011).

Upon environmental induction, i.e., 15°C short day condition, gametophore apical cell terminally differentiates to form gametangia apical cells, which develop to form the male (antheridia) and the female (archegonia) sexual organs at the gametophore apex (Hohe *et al.*, 2002). Water facilitates the motile biflagellate sperm from antheridia to reach and fertilize the egg cell present in archegonia to form the zygote (Renzaglia and Garbary, 2001). The sporophyte develops from the zygote and dehisces upon maturation to release spores, thereby completing the life cycle within merely 3 - 4 months. In the event of an injury, moss can still complete its life cycle vegetatively by reprogramming the injured differentiated cells to form protonemal filaments (Sakakibara *et al.*, 2014). Mosses, one of the earliest colonizers of land have long fascinated scientists with their diversity, adaptability to new environments and exceptional physiological characters. *Syntrichia caninervis* of cold deserts, *Polytrichum commune* of acidic bogs, the carbon capturing *Sphagnum* species of marshes and the concrete dwelling species you spotted growing on your wall are, but a few examples of the robustness exhibited by mosses. Despite their small stature, mosses have equally complex physiology like flowering plants, but they occupy the base of land plant phylogenetic tree, a strategic position which comes with the advantage of low genetic redundancy (Rensing *et al.*, 2008), making them lucrative models to study complex physiology of plants.

To conduct intricate plant molecular studies and unravel the complex physiology of plants, researchers have established and sequenced the genome of a model moss. *P. patens* (class Funariales ; phylum Bryopsida) was chosen for this purpose due to its short life cycle (3 - 4 months), small genome (512 Million base pairs ; 27 chromosomes), amenability to genetic transformation via PEG-mediated (Liu and Vidali, 2011) or biolistic methods (Šm\`idková, Hala and Angelis, 2010) a high homologous recombination frequency (~ 5- 20%) and its suitability for microscopic observations. A brief time-line of the landmark studies have been represented in Figure 1.13.

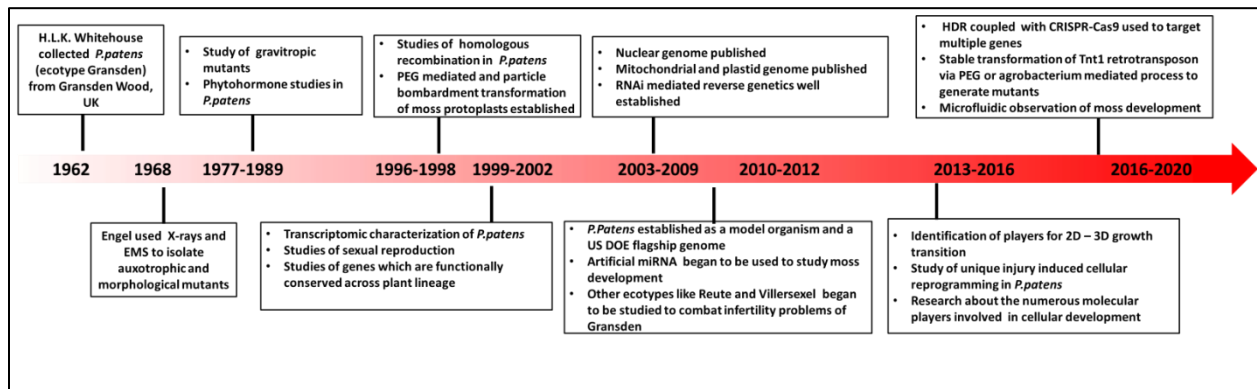


Figure 1.13 Chronological accounts of *P. patens* research A brief timeline of the landmark studies which shaped *P. patens* research

The genome of moss *Physcomitrium patens* consists of ~78% annotated and evolutionarily conserved genes, with ~52% of the genes having no annotated Pfam domains. Approximately 0.9% of them are moss-specific and lack Gene Ontology assignment, suggesting the presence of lineage-specific genes of unknown function (Nishiyama *et al.*, 2000; Lang *et al.*, 2005; Rensing *et al.*, 2008; Zimmer *et al.*, 2013). With its distinctive evolutionary status, regenerative potential, filamentous protonema, single-layered leaves, and high susceptibility to genetic manipulations, this moss model, characterized by a haploid-dominant life cycle, provides an exceptional opportunity to investigate the intricacies and roles of haploid-specific genetic pathways in plant development.

1.10. Hypothesis and objectives

Previously, we have reported a bryophyte-specific gene family, *SHORT-LEAF* (*SHLF*), involved in auxin distribution pattern and gametophore development in moss (Mohanasundaram *et al.*, 2021). The Tnt1 insertional *short-leaf* (*shlf*) moss mutant showed growth retardation phenotypes reminiscent of Arabidopsis mutants affected in auxin distribution pattern, including shorter leaves, increased internodal distance, reduced apical dominance and high auxin accumulation at the sites of synthesis (Ruegger *et al.*, 1997; Noh, Murphy and Spalding, 2001; Viaene *et al.*, 2014). A unique protein *SHORT-LEAF* (*SHLF*) is the genetic cause of the *shlf* phenotypes. *SHLF* has no known conserved domains and encodes a N-terminal signal peptide,

four highly similar longest known tandem direct repeats (TDRs) and a C-terminal tail. SHLF follows the conventional secretory pathway and undergoes proteolytic cleavage. Nonetheless, the absence of conserved domains poses a significant challenge in unraveling the precise molecular characteristics of the protein. Reports from flowering plants have shown that repeat containing domains in have evolved through internal tandem duplication and are essential for rapid adaptation to dynamic environmental conditions by regulating primary and secondary metabolic pathways (Hanada *et al.*, 2008; Schaper and Anisimova, 2015; Sharma and Pandey, 2016). To fully grasp the role of SHLF in the regulation of moss development, it is essential to elucidate the molecular nature of SHLF and its associated regulatory pathways, considering its pivotal role as a clade-specific player in influencing auxin distribution patterns.

Hence the following objectives were proposed for the following study:

- Identification of crucial SHLF domains regulating gametophore development (Chapter 2)
- Studying the secretory nature of SHLF and the role of its cleavage products in gametophore development (Chapter 3)
- A transcriptomic and metabolomic approach to elucidate the mode of SHLF function in regulating moss development (Chapter 4)

CHAPTER 2:

Identification of crucial SHLF domains regulating gametophore development

2. Identification of crucial SHLF domains regulating gametophore development

2.1. Introduction

The *short-leaf* mutant, isolated from a Tnt1 retrotransposon based forward genetic screen in moss, exhibits pleiotropic phenotypes including two-fold shorter leaves, reduced apical dominance, increased internodal distance, reduced plasmodesmata frequency and altered auxin distribution pattern. A tandem direct repeat (TDR) containing bryophyte-specific gene *SHORT-LEAF* (*SHLF*), having no known functional domains was identified to be the underlying genetic cause of this mutant (Mohanasundaram *et al.*, 2021). The authors demonstrated that *SHLF* traffics to the endoplasmic reticulum (ER) via its N-terminal signal peptide and undergoes cleavage. However, the absence of known conserved domains in SHLF made it difficult to analyze its molecular role, making it imperative to explore the significant functional domains in the protein.

Proteins with domains of unknown function (DUFs), are present in life forms and conserved DUFs have long since been linked to perform essential functions for survival (Goodacre, Gerloff and Uetz, 2014). However, many of the proteins containing DUFs may either behave as modular proteins or may bear structural similarity to other functionally annotated proteins, making it easy to assess their molecular function by using *in silico* functional annotation tools (Koshy *et al.*, no date). Domain deletion or loss of function studies conducted with protein DUFs have led to the characterization of several plant DUFs, many of which have been found to be associated with both development and stress response in plants (Lv *et al.*, 2023). Given the lack of structural similarity of SHLF with any known proteins, it is prudent to take a domain deletion approach to assess the function and possible modular nature of SHLF domains.

2.1.1. Lineage specific genes in plants

Lineage specific genes (LSGs) are representative of the genetic novelty in each clade and are necessary for survival in specific ecological niches (Lespinet *et al.*, 2002; Cai *et al.*, 2006). Generally, they bear low to no sequence similarity with genes from other lineages and the origin of many of these genes may be attributed to gene duplication (Lespinet *et al.*, 2002) and neo-

functionalization (Tautz and Domazet-Lošo, 2011) or homology detection failure (Weisman, Murray and Eddy, 2020). Section 1.3 of this thesis provides comprehensive and intricate details about Lineage Specific Genes (LSGs). LSGs in plants have been previously identified from angiosperms including Poaceae (Bold *et al.*, 2004), *Oryza sativa* (Koide *et al.*, 2018), Triticeae (Ma *et al.*, 2020) *Arabidopsis* (Campbell *et al.*, 2007; Donoghue *et al.*, 2011), *Populus* (Yang *et al.*, 2009), sweet orange (Xu *et al.*, 2015), tea plant (*Camellia sinensis*) (Zhao and Ma, 2021) and Caryophyllales (Brockington *et al.*, 2015). Functional characterization of plant LSGs has revealed that many of them are involved in signalling pathways involved in mediating plant response to dynamic environmental cues (Lespinet *et al.*, 2002).

SHLF is a LSG exclusive to bryophytes, characterized by its distinct composition of repeats. The N-terminal region of SHLF has been well-established as a signal peptide, mediating protein trafficking to ER (Mohanasundaram *et al.*, 2021). The four tandem repeat domains (TDRs) of SHLF and the C-terminal region can be categorized as Domain of Unknown Function (DUFs), showcasing their unique nature.

2.1.2. Plant repeat containing protein families and their post translational modifications

Previous reports from flowering plants have shown that repeat-containing domains in proteins have evolved through internal tandem duplication and are essential for rapid adaptation to dynamic environmental cues (Sharma and Pandey, 2016). Repeat-containing protein families reported in plants are known to be involved in both primary and secondary metabolic pathways to regulate plant development (Schaper and Anisimova, 2015). Longer tandem repeats in protein are involved in the evolution of neo-functionalized proteins and are also known to be hubs of catalytic activity and may serve as scaffolds for protein interaction and/or stability (Rajathei, Parthasarathy and Selvaraj, 2019).

Though plants proteins containing tandem direct repeats may be post-translationally processed, not much is known about the internal processing of the repeat themselves and the roles of their cleavage products in plant development. Few studies reported that repeat containing proteins undergoing site specific cleavage to form functional cleavage products. These include the Potato type-II Inhibitor family Protease Inhibitors (Pin-II type PIs) which constitute essential plant defense molecules, and are proteolytically processed at linker regions in between repeats to

generate serine protease inhibitors (Yadav, Saikhedkar and Giri, 2021). Proteins from other systems are also shown to exhibit internal repeat specific cleavage. Reelin, a well-studied secretory glycoprotein found in brain is known to undergo cleavage at two internal sites within the repeats to regulate neural development (Koie *et al.*, 2014). The herpes simplex virus VP16-associated protein HCF encodes a set of six near perfect repeats which undergo internal cleavage and the processed products can mediate the protein function (Wilson, Peterson and Herr, 1995). To the best of our knowledge, plant specific repeat containing proteins exhibiting internal cleavage and generation of functional products from within the repeats are not reported till date.

Bryophytes being amongst the earliest colonizers of land, had to deal with several adverse conditions to ensure survival. They also share their last common ancestors with flowering plants, making them interesting candidates to investigate the evolution of plant response of environmental cues by studying bryophyte-specific LSGs. Previously, we have reported the first bryophyte-specific gene, *SHORT-LEAF (SHLF)*, involved in moss gametophore development (Mohanasundaram *et al.*, 2021). *SHLF* encodes a protein (SHLF) comprises of a N-terminal signal peptide, followed by four highly similar TDRs (171 amino acids) and a short C-terminal tail. The unique characteristics of SHLF including high gene expressed levels, presence of the longest reported TDRs, lack of introns and absence of any known conserved domains, makes it a very interesting candidate to explore multifaceted aspects of this protein in moss.

In this study, we attempted to functionally characterize the crucial SHLF domains, through a series of sequential domain deletion assays. The transgenic plants generated upon overexpression of the domain deletion constructs in the mutant background were scored for phenotypic reversal. We also attempted to study the role of SHLF domains for its ER trafficking and post translational processing. The functionality of the SHLF cleavage products was also explored in the present study. Additionally, the high sequence similarity in the SHLF TDRs also raises the possibility that either the TDRs are functionally redundant or that this protein may function as a modular protein. In light of this, here, we attempted to identify a minimal functional protein which can potentially be sufficient for function.

2.2. Materials and methods

2.2.1. Plant tissue culture and maintenance

The moss *Physcomitrium patens* (Ecotype Gransden) was grown in BCDAT media with 0.8% agar at 24 °C under 16 h photoperiod conditions for all experiments, as per the protocol of Cove *et al* (Cove *et al.*, 2009).

2.2.2. Cloning and plant transformation

The Δ N SHLF-eGFP cassette was amplified from SHLF-eGFP in pTK-Ubi-Gate (Mohanasundaram *et al.*, 2021) using the primer pair attB1dN_SHLF_F and attB2_eGFP_R. Δ C SHLF cassette was amplified from pTK-Ubi-Gate containing SHLF-eGFP using the primer pair attB1_SHLF_F and attB2_dC_SHLF_R. The above cassettes were used for one step recombination-based BP-LR cloning using entry vector, pDONR 201 (P1P2) and the destination vector pTK-Ubi-Gate, respectively. Δ C SHLF – eGFP was amplified from the destination vector pTK-Ubi-Gate containing SHF-eGFP using the primer pairs attB1_SHLF_F and attB5r_dC_SHLF_R and used for recombination-based BP cloning into pDONR 221 (P1P5r). The positive clone obtained was recombined with pENTR L5L2 and pTK-Ubi-Gate using LR cloning (Invitrogen). The Δ N Δ C SHLF – eGFP cassette was amplified from Δ C SHLF – eGFP in pTK-Ubi-Gate using attB1dN_SHLF_F and attB2_eGFP_R. Similarly, the 1st TDR of SHLF was amplified using attB1dN_SHLF_F and attB2_dC_SHLF_R. Both these cassettes were used for one step recombination-based BP-LR cloning using entry vector, pDONR 201 (P1P2) and the destination vector pTK-Ubi-Gate, respectively. The N-TDR1-C- eGFP cassette was amplified from SHLF-eGFP in pTK-Ubi-Gate using the primer pair attB1_SHLF_F and attB5r_SHLF_R. The N-TDR1-2 – eGFP and the N terminal regions cassettes were amplified from SHLF – eGFP using the primer pairs attB1_SHLF_F and attB5r_2_SHLF_R and ttB1 SHLF F and attB5r N_SHLF R respectively. Both these cassettes were then cloned into pTK-Ubi Gate via BP-LR cloning mediated by the entry vectors pENTR 221 (P1P5r) and pENTR L5L2. The 7 different destination constructs generated were used for PEG mediated protoplast transformation of *shlf* as per section 2.2.2. The regenerated protoplasts were selected on BCDAT media containing 20 mg/L G418. Putative moss clones surviving two rounds of antibiotic selection were subjected to

molecular analysis for further verification. The schematic representation of the deletion constructs used in this study are provided in Figure 2.1.

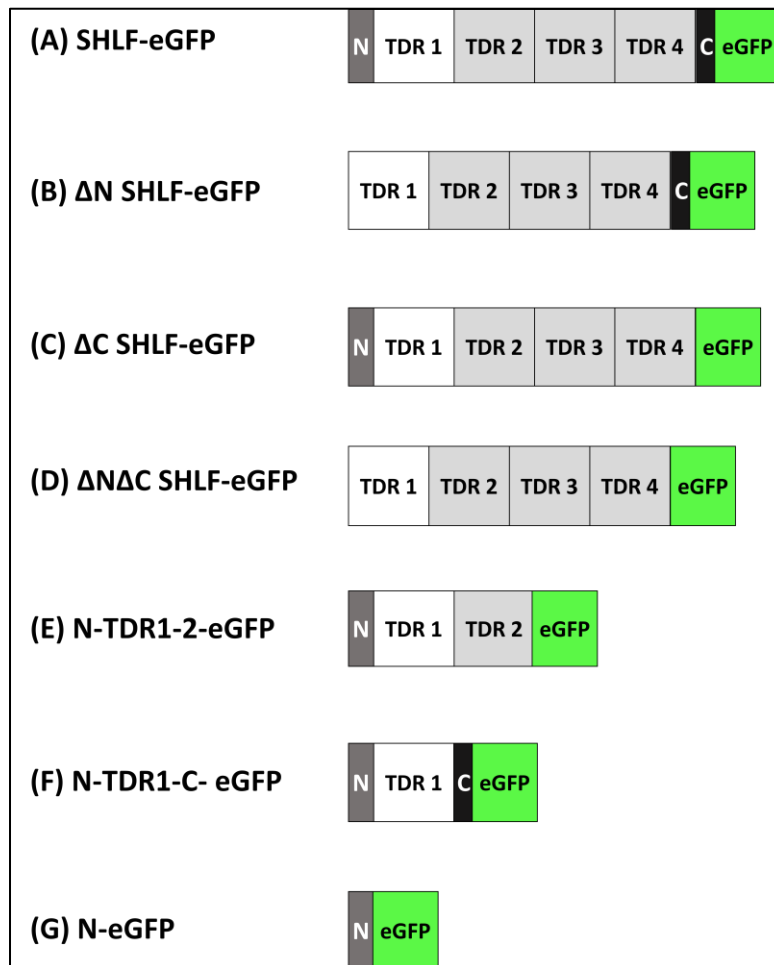


Figure 2.1 Schematic representation of the SHLF domain deletion constructs overexpressed in *shlf* background

2.2.3. PEG mediated protoplast transformation

PEG (Polyethyleneglycol) mediated protoplast transformation was performed as per the protocol of Nishiyama *et al* (Nishiyama *et al.*, 2000). 5–7-day old moss protonema was digested using 1% (w/v) Driselase made in 8% mannitol. The protoplasts released were washed twice using 8% mannitol following transfer to a 2% PEG solution containing MgCl₂, Ca(NO₃) and 7-8 μg of the desired DNA for transformation. This protoplast mixture was subjected to heat shock at 45°C for 5 minutes to enable uptake of DNA, followed by cooling at room temperature for 10 minutes

to enable closure of the cell wall pores. The osmolarity of the protoplasts was adjusted using protoplast regeneration media (containing BCDAT, 8% mannitol and CaCl₂) and they were incubated in the dark to enable cell wall regeneration. After five days of growth, the protoplasts were transferred to selection media containing 20 mg/L G418 and incubated for 2 weeks. The colonies which survived the 1st selection were transferred to a relaxation media, lacking any antibiotic for a period of two weeks. This was followed by a secondary selection using 20 mg/L G418 for two weeks. Surviving colonies were then subjected to molecular confirmation analyses using PCR and RT-qPCR techniques for further verification

2.2.4. Phenotypic characterization

Gametophores (Three-week-old) from two independent stable transgenic lines generated from each of the constructs described in above section (Section 2.2.3) were subjected to detailed phenotypic analyses. For each line, length and width of the 9th leaf from the apex and internodal distance of gametophores (n=20) were recorded using Olympus stereomicroscope and measured using ImageJ (Fiji). Silhouettes for the first nine leaves of each line was traced using Inkscape (Version 0.92.3-1) from the images obtained on Leica stereomicroscope.

2.2.5. Statistical analyses

Boxplots were generated from the datasets using the ggplot2 package with the whiskers representing the closest data-points with the 1.5x interquartile range following Tukey's style, above the first quartile and below the third quartile respectively. Outliers were also included in the analyses and plotted along the boxplots. The bar plots were plotted with error bars depicting the standard deviation between the samples. For datasets having normal distribution and equal variance, Student's *t* test was performed. *P*-value < 0.05, <0.001, and <0.0001 were marked as *, **, and ***, respectively. In case of samples having non-normal distribution, Levene's test was conducted to check for homogeneity in the population (Levene, 1960). Samples having p-value <0.05 were considered non-homogenous and Kruskal–Wallis test was performed with *post hoc* Dunn test having Bonferroni corrected p-values (Bonferroni, 1936; Dunn, 1964; McKight and Najab, 2010). The letters on top of the graphs depict the similarity between two data sets. Graphs were generated using GraphPad prism (Version 9.5.0) and R (Version 4.2.3).

2.2.6. Confocal Microscopy

To observe the sub-cellular localization of SHLF in the various transgenic moss lines, leaves plucked from three-week-old moss gametophores, mounted with liquid BCDAT media and imaged in Leica TCS SP8 confocal laser scanning microscope under 488-nm excitation from an argon laser (30%) using the emission range 493–527 (for eGFP) and 630–761 nm (for chlorophyll). Similar conditions were used for the visualization of GFP signals in the transient agroinfiltrated leaves of *N. benthamiana* using 35S: SHLF-eGFP. In case of the 35S: TVCV MP-mCherry agroinfiltrated leaves, an excitation laser of 543 nm was used with the emission range 600-6890 nm to visualize m-Cherry and 630-761 for chlorophyll respectively. Signal Image analysis was carried out using ImageJ (Fiji).

2.2.7. Generation of polyclonal SHLF antibody

The sequences coding for N terminal, 1st TDR and C-terminal region (N-TDR1-C) of SHLF were cloned into pET-28a vector (Merck Millipore, MA, USA) containing 6X Histidine tags at both N and C terminal. This construct was then transformed into *E. coli* SHuffle competent cells (Merck Millipore, MA, USA). The cells were grown in batches of 500 ml LB Broth (Hi-Media, India) at 30 °C till OD 0.5. Induction was carried out using 0.5 mM IPTG (Sigma, MA, USA) and the culture was grown at 16 °C for 18 h. Cells were then pelleted at high speed for 20 minutes at 4 °C and lysed in sonication buffer (10mM Tris, pH 8, 300mM NaCl, 10% glycerol; 10mM Imidazole and 1X PMSF) on ice by sonication for 15 minutes with 2s on and 5s off cycles at 65% amplitude. Lysed cells were pelleted at maximum speed for 20 minutes at 4 °C.

The protein from the supernatant obtained after cell lysis was bound to Ni-NTA pre-equilibrated resin (Invitrogen, CA, USA) at 4 °C for 4 h. The column was washed in buffers (10mM Tris pH 8, 300mM NaCl, 10% Glycerol, 1X PMSF) with increasing concentration of Imidazole from 20 mM to 100 mM and finally the protein was eluted with 300 mM Imidazole. The purity was checked in 12.5% PAGE gel followed by a Bradford analysis to determine the protein concentration. The gel slices containing 3 mg of 27 kDa SHLF protein was crushed in liquid nitrogen and resuspended in 500 µL of 1X PBS. This slurry was mixed with 500 µL of Freund's complete adjuvant (Sigma) to form 1 mL emulsion, which was injected subcutaneously into rabbit

for antisera production. Subsequent booster doses, prepared using Freund's incomplete adjuvant were given to the animal at every two-week interval, and antibody titer was checked. Thirty milliliter of final bleed was collected and used for antisera preparation.

SHLF polyclonal antibody was precipitated out using 100% ammonium sulphate and repeatedly washed with 1x PBS using 3kDa cutoff column (The protein protocols handbook, JM Walker). The bands obtained on western blot by probing total WT tissue with 1:20,000 dilution anti-SHLF and pre-immune sera were subjected to mass spectrometric analysis for protein specificity.

2.2.8. Western blot analyses

Three-week-old moss gametophores were crushed in liquid nitrogen and resuspended in PBS extraction buffer (20 mM, pH 7.4) containing 2X SDS and 1X protease inhibitor cocktail (Sigma, USA). The resuspended tissues were vortexed and incubated in a water bath sonicator for 10 minutes and centrifuged at 12,000 g for 10 min at 4°C and equal amounts of protein for all transgenic lines were loaded on a 12.5% SDS-PAGE gel and ran for 1.5 h at 110 V. The protein samples were then transferred to a PVDF membrane in phosphate transfer buffer (pH 7.4) at 4°C for 3 h. The samples were probed using rabbit anti-SHLF at a dilution of 1:20,000 using goat anti-rabbit secondary antibody at a dilution of 1:10,000. SHLF protein (27 kDa) used for antibody generation and total protein from *SHLF_KO* lines were used as positive and negative controls respectively. Equal amounts of total protein from transgenic lines (determined from Ponceau-S stained blots by band densitometry) expressing recombinant GFP were probed using polyclonal anti-GFP antibody at a dilution of 1:500.

2.2.9. Mass spectrometric analysis

Corresponding SHLF bands were cut from an identical 12.5% SDS-PAGE gel and used for in-gel digestion using trypsin (Promega). The digested peptides were extracted from the gel pieces and subjected to MS/MS analysis using ABSciex Triple TOF 6600 system (Kelkar *et al.*, 2019). The results were interpreted in an IDA mode using Protein Pilot software (Version 4.0).

2.2.10. RT-qPCR analysis

Three-week-old moss gametophores were subjected to total RNA extraction using RNAiso Plus (Takara Bio USA Inc.) Two micrograms (2 µg) of RNA samples were reverse-transcribed using SS-IV reverse transcriptase (Invitrogen) and oligo dT primers. SYBR Premix Ex Taq II (Tli RNaseH Plus) was used for the relative quantification of transcripts via Bio-Rad CFX96 Touch Real-Time PCR Detection System (Bio-Rad). Moss *β-Actin* and eIF2 were used as a reference gene for the quantification of target gene expression levels in all samples. The fold-change values (sample value/reference value) and $2^{-\Delta\Delta Ct}$ were calculated using the $2^{-\Delta\Delta Ct}$ method, according to calculated according to Schmittgen *et al.*, 2008 (Schmittgen and Livak, 2008). All the primers used for RT-qPCR analysis are listed in Table 5.

2.2.11. In silico protein analyses

AlphaFold2 was executed on the full length and truncated protein sequences using the default prediction mode (Jumper *et al.*, 2021). The resulting models were generated in the Protein Data Bank (PDB) format. The predictions were further refined using the ProSA-web server. The hydrophobicity of N-TDt1-C used for antibody generation was predicted using the GRAVY Hydropathy index calculator (<https://www.gravy-calculator.de/>)

2.1. Results

2.1.1. N-terminal and the highly conserved 2nd TDR of SHLF are essential for moss gametophyte development

The *short-leaf* (*shlf*) mutant, generated by Tnt1 retrotransposon insertion in the unique bryophyte-specific gene SHORT-LEAF (SHLF), has scorable pleiotropic phenotypes, such as reduced phyllid length, width (both at the center and base) and increased internodal distance. Previously, we have demonstrated the genetic basis for the *shlf* mutant and also reported the primary protein structure of SHLF (Mohanasundaram *et al.*, 2021). SHLF consists of an N-terminal endoplasmic reticulum (ER) specific signal peptide, followed by four near-identical tandem direct repeats (TDRs) and a C-terminal tail. The slightly divergent 1st TDR bears 90% sequence similarity to the rest of the highly conserved TDRs (2nd, 3rd and 4th), which are 99%

similar. The presence of longest known TDRs and the absence of any known conserved domains make the SHLF protein unique. In light of this, it is imperative to elucidate the mechanistic basis of the functional domain/s of SHLF necessary for its role in gametophore development.

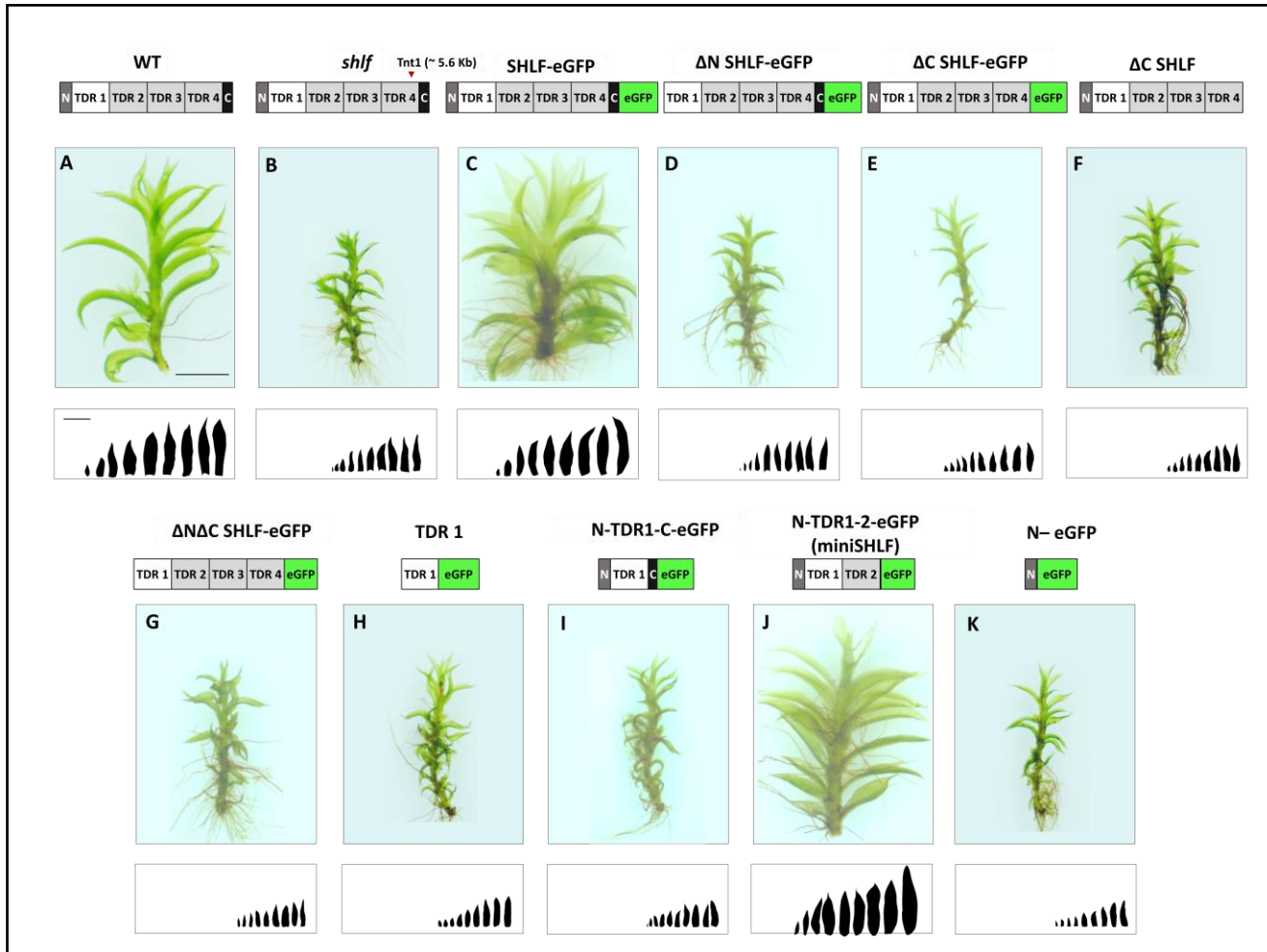


Figure 2.2 Overexpression of domain deleted SHLF constructs in *shlf* Panels (A-K) are representative gametophore images of (A) WT, (B) *shlf*, and the 10 different overexpression cassettes in *shlf* (C) SHLF-eGFP, (D) ΔN SHLF-eGFP (E) ΔC SHLF-eGFP, (F) ΔC SHLF, (G) ΔN ΔC SHLF-eGFP, (H) TDR1, (I) N-TDR1-C-eGFP, (J) N-TDR1-2-eGFP (miniSHLF), (K) N-eGFP. The schematic above each panel represents the respective SHLF overexpression cassettes. Red arrow in (B) represents the Tnt1 insertion site in *shlf*. Three-week-old gametophores of the stable overexpression lines are used for representation, Scale = 1mm. The first nine leaves from the apex are represented as silhouettes in the panel below, Scale = 1mm.

Complementation studies in *shlf* background, using various sequential domain deletion constructs, were carried out (Figure 2.2 A-K), and the phyllid length, width and internodal

distances were scored in the stable transgenic lines (Figure 2.5 A- D). The stable transgenic lines generated upon overexpression of the various domain deletion constructs were subjected to PCR-based molecular confirmation. The ~ 360bp kanamycin resistance cassette was amplified from the gDNA of these lines and the positive lines were used for further analyses (Figure 2.3).

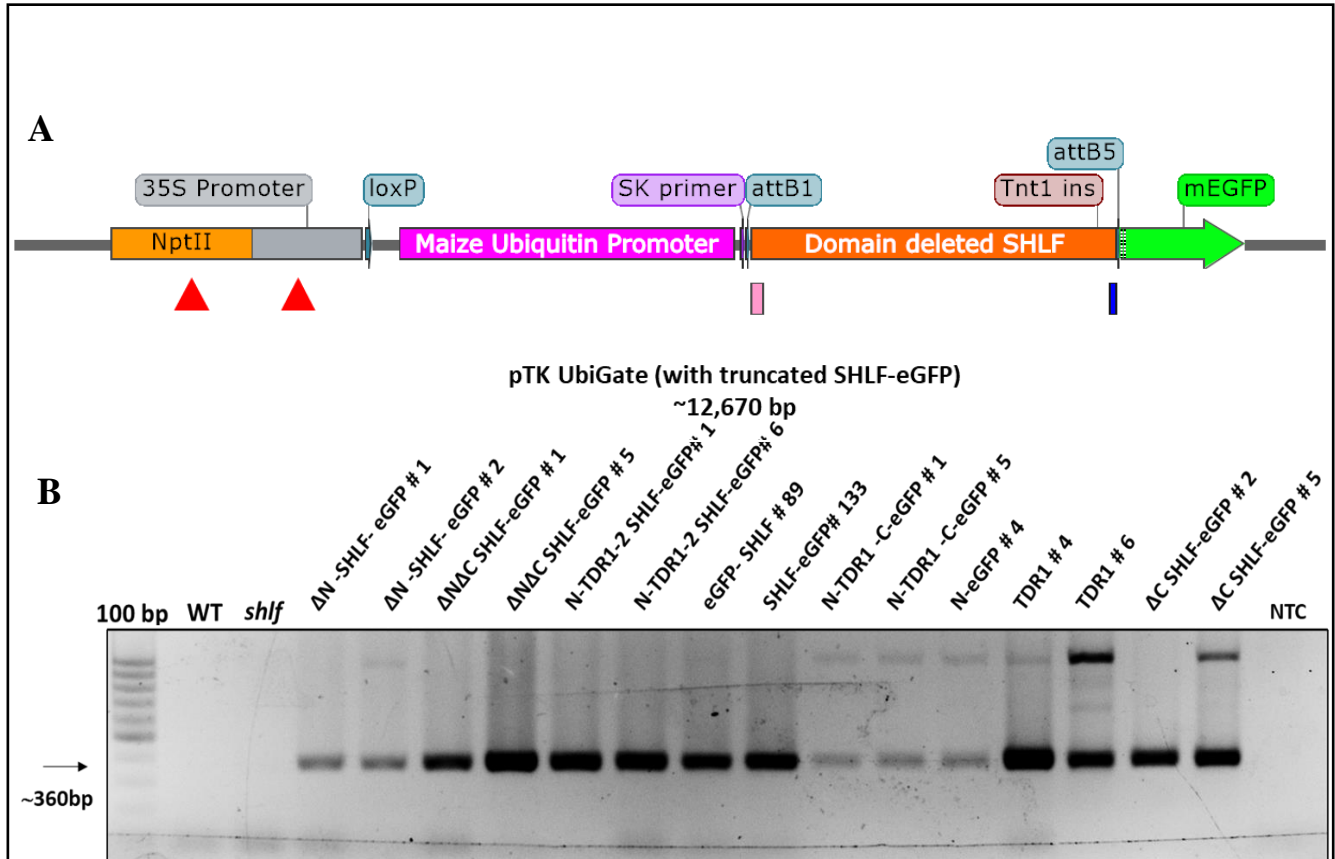


Figure 2.3 PCR based molecular confirmation of SHLF domain-deleted overexpression lines (A) Primer pair SbKan_F and SbKan_R was used to amplify a segment of the kanamycin resistance gene (360 bp), marked by red arrows in the schematics of final destination vector pTK Ubi Gate-Truncated SHLF. (B) Amplification of part of Kanamycin cassette (the area highlighted between the red arrows in panel A) from the genomic DNA of the stable domain-deleted SHLF overexpression lines.

RT-qPCR analyses was also carried out using the cDNA from the stable lines using four different sets of primer pairs. The $2^{-\Delta Ct}$ values of lines were analyzed using the primer pairs represented in Figure 2.4B and only those showing significant level of overexpression of SHLF were carried forward for phenotypic characterization (Figure 2.4).

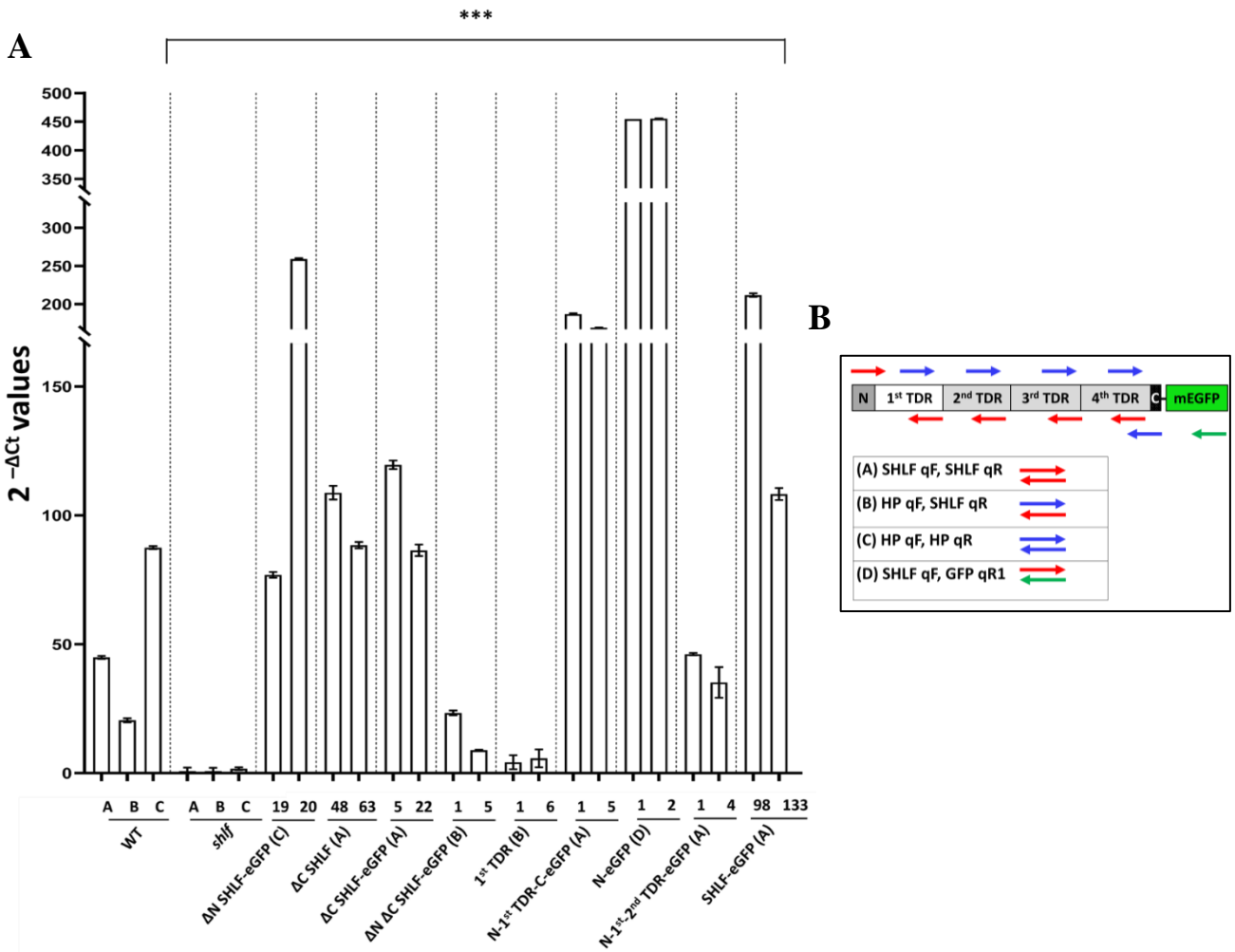


Figure 2.4 The $2^{-\Delta Ct}$ plot showing the level of overexpression of the domain-deleted SHLF constructs in *shlf* background (A) The $2^{-\Delta Ct}$ value depicting the overexpression of each SHLF domain-deleted overexpression construct was plotted for three independent bio-replicates. Moss β -Actin was used as a control (B) Schematics of primer binding sites for the four primer pairs (labelled A, B, C and D) used in the analysis. Primer pairs A and C are specific to the N and C-terminal respectively and primer pair B binds to the four TDRs. Primer pair D was used to amplify N-terminal with eGFP.

Overexpression of the full-length SHLF, tagged with C-terminal GFP (SHLF-eGFP) reverted the mutant phenotypes (Figure 2.2 C). However, overexpression of SHLF without the N-terminal (Δ N SHLF-eGFP) (Figure 2.2 D), or only the N-terminal without the rest of the protein (Figure 2.2 K), failed to recover the *shlf* phenotypes, indicating that the N-terminal may be essential for directing SHLF into ER and for phenotypic recovery. The complementation of the mutant with

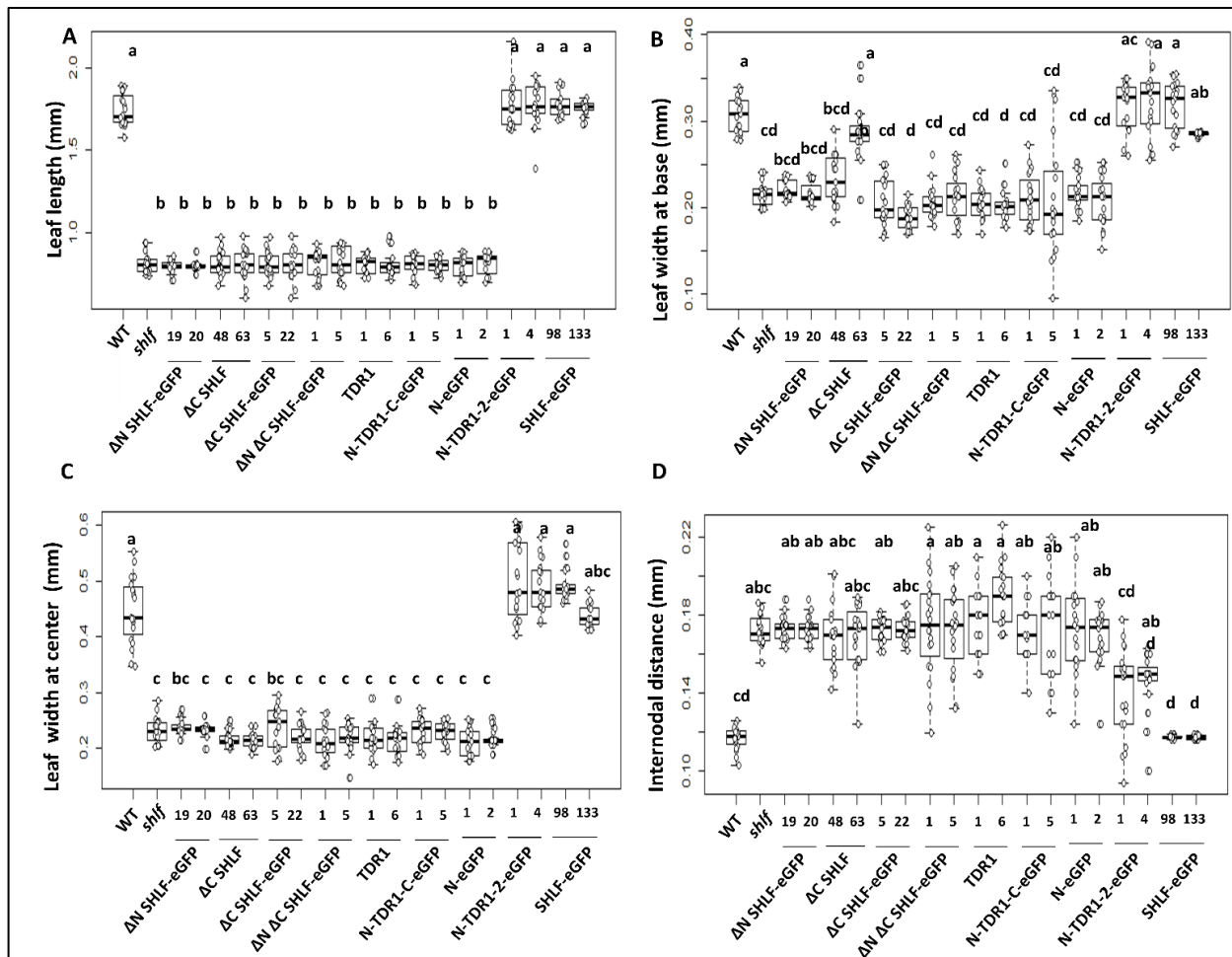


Figure 2.5 Comparative phenotypic analyses of the stable overexpression lines in *shlf* background (A) Leaf length, (B) leaf width at the base, (C) at center and (D) the internodal distance were measured for the 9th leaves of two independent overexpression lines for each genotype along with WT and *shlf* (n = 20). Shapiro-wilk normality test was performed on the datasets. Tukey's test with Dunn's post-hoc test and Bonferroni p-value corrections were performed to check for the statistically significant difference between the sample sets. The letters indicate the significant pairwise comparisons between the sample sets, as per the R statistical analyses software.

SHLF constructs lacking the C-terminal tail (Δ C SHLF-eGFP and Δ C SHLF) did not rescue the mutant phenotype (Figure 2.2 E and F). Subsequently, plants generated by overexpression of truncated constructs containing four TDRs (Δ N Δ C SHLF-eGFP) (Figure 2.2 G) or the 1st TDR only (lacking both N- and C-terminal) were unable to recover the mutant phenotypes (Figure 2.2 H). The presence of highly similar TDRs in SHLF prompted us to carry out complementation

studies to identify the minimal SHLF. Constructs having only the N-terminal, 1st TDR and the C-terminal (N-TDR1-C) failed to recover *shlf* phenotypes (Figure 2.2I). The successful complementation of *shlf* was only observed by the overexpression of a construct having the N-terminal, 1st and 2nd TDR (N-TDR1-2) but lacking a C-terminal tail, which resulted in complete rescue of the mutant phenotypes (Figure 2.2 J). Additionally, the internodal distance of these plants was observed to be significantly longer than WT (Figure 2.5 D). Thus, complementation studies revealed that the minimal SHLF containing N-terminal signal peptide, 1st and the 2nd TDRs (miniSHLF) is sufficient for SHLF function in regulating moss gametophore development.

2.1.2. Subcellular localization of SHLF is dependent on both N- and C-terminal domains

Phenotypic analyses revealed that moss gametophore development relies on the presence of both N- and C-terminal domains. Our previous report showed that SHLF enters the conventional secretory pathway aided by the N-terminal signal peptide and may be post-translationally processed (Mohanasundaram *et al.*, 2021). As proper subcellular localization and post-translational modifications (PTMs) are imperative for protein function, we further investigated the SHLF localization pattern in the domain deletion lines.

The 35S:eGFP and the eGFP-HDEL lines served as positive controls for ubiquitous and ER specific localization pattern of eGFP respectively. (Figure 2.6 A and B). All the stable transgenic lines containing an intact N-terminal showed GFP signal in the ER, which was comparable to the eGFP-HDEL line (Figure 2.6). The overexpression lines of SHLF-eGFP, N-eGFP, N-TDR1-C-eGFP, and N-TDR1-2-eGFP (miniSHLF) category showed eGFP localization in the reticulate cortical ER originating from the nuclear envelope (Figure 2.6 C, D, H and I). The Δ N SHLF-eGFP lines showed cytosolic and nuclear localization of GFP because of the absence of the signal peptide (Figure 2.6 B). This pattern was reminiscent of the 35S:GFP line, which ubiquitously expresses GFP in the plant (Figure 2.6A).

Interestingly, the Δ C SHLF-eGFP and Δ N Δ C SHLF-eGFP lines, which lacked the C-terminal, did not show any GFP signal, hinting at the importance of C-terminal in SHLF stability (Figure 2.6 F and G). It was also intriguing that the miniSHLF, despite lacking the C-terminal tail, was able to traffic to ER, indicating a possible functional redundancy of the C-terminal, which is

fulfilled by the miniSHLF. The subcellular localization patterns of the domain deletion lines indicate that both the N- and C-terminals are essential for SHLF to enter the secretory pathway, failing which SHLF is unable to contribute to moss development

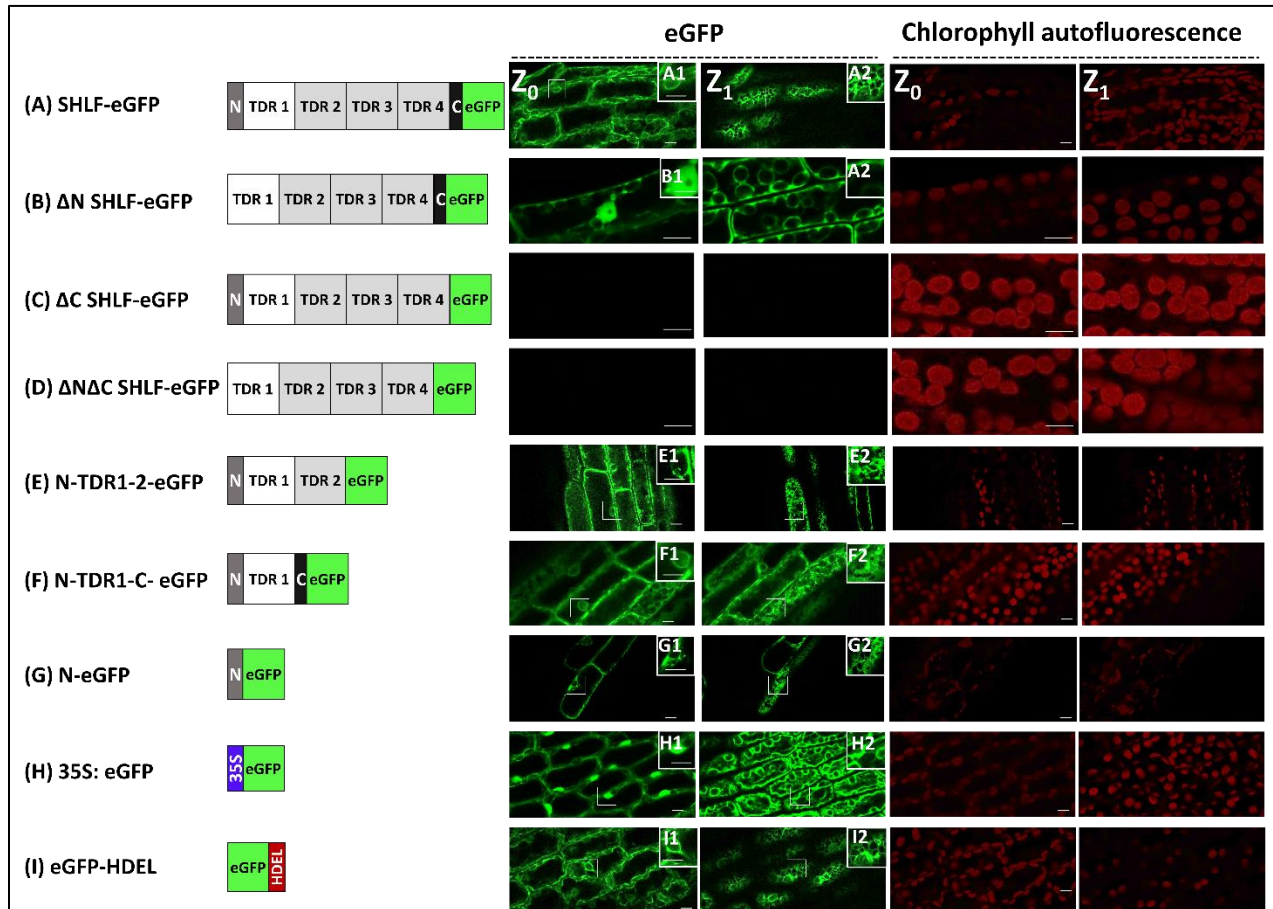


Figure 2.6 Domain deletion of SHLF affects subcellular localization pattern (A-I) are row-wise representation of the various domain deletion constructs including controls used for confocal imaging. (A) In SHLF-eGFP lines, SHLF localized to the cortical ER as seen in the inset (A1-2); similar to the N-eGFP lines and (G) the ER marker line eGFP-HDEL (I), evident from the eGFP signal in nuclear envelope and mesh-like network in the insets (A1-2, G1-2 and I1-2). (B) Absence of signal peptide in ΔN SHLF-eGFP lines shows the protein localization in the nucleus and its ubiquitous presence throughout the cytoplasm (B1-2), akin to the signal observed in the 35S-eGFP line (H). (C-D) No eGFP signal is observed in overexpression lines (C) ΔC SHLF-eGFP and (D) $\Delta N \Delta C$ SHLF-eGFP lacking a C-terminal tail; (E-F) Both N-TDR1-2-eGFP (E) and N-TDR1-C-eGFP (F) showed protein localization in the ER. For all lines, two optical sections along the Z plane, labeled as Z_0 and Z_1 , represents nucleus and/or nuclear membrane with cortical ER network respectively. In parallel, chlorophyll auto-fluorescence and bright field images for respective lines are also provided. All scale bars are of 10 μm , except for the insets which are of 1 μm .

2.1.3. SHLF undergoes post-translational processing to generate cryptic cleavage products

SHLF, a highly expressed gene (>16 fold higher expression than moss β -Actin) encodes the longest known TDRs (TDR1-4) having no introns and undergoes post translational processing (Mohanasundaram *et al.*, 2021).

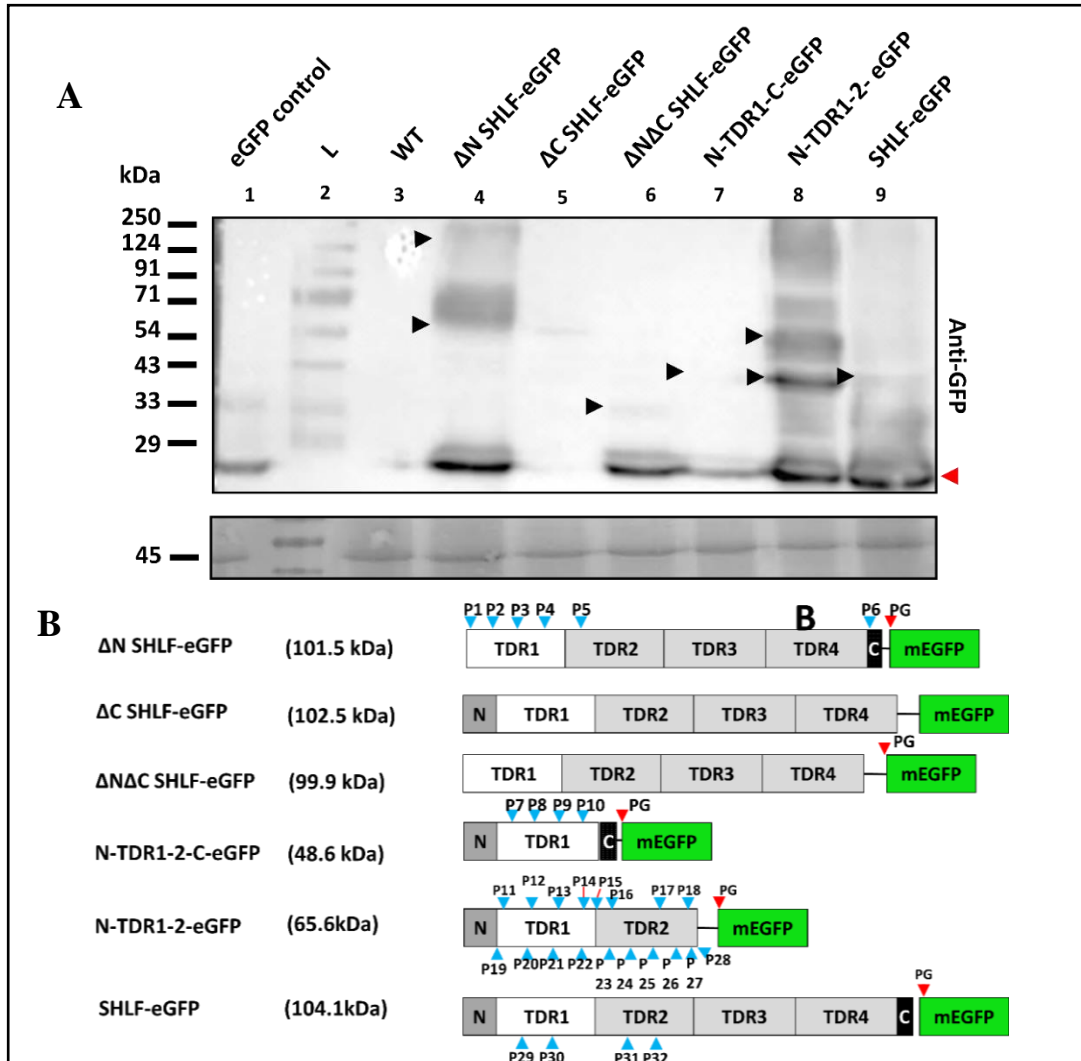


Figure 2.7 Western blotting with anti-GFP to show that SHLF undergoes cleavage in-planta to form smaller protein fragments b

Here, we examined the role of each domain in proteolytic processing of SHLF by carrying out western blot and MS/MS analyses. Immunoblot analyses of the domain deletion lines with the anti-GFP antibody revealed differential banding patterns (Figure 2.7 A). Although the Δ N SHLF-

eGFP lines showed a diffused banding pattern between 54 kDa to 124 kDa, through in-gel digestion and MS/MS analysis, we could through in-gel digestion and MS/MS analysis, we could detect a partially processed SHLF indicating partial cleavage inside the TDRs (Figure 2.7 B). This indicates that the trafficking of SHLF to the ER is essential for *in-planta* proteolytic processing. The ΔC SHLF-eGFP lines did not show any GFP bands in the western blot (Figure 2.7 A). The $\Delta N\Delta C$ SHLF-eGFP lines showed an eGFP 27 kDa band and a faint band at ~ 33 kDa, whose tryptic digests contained peptides from the 4th TDR (Figure 2.6 B). It was interesting to note that the $\Delta N\Delta C$ SHLF-eGFP lines showed minimal processing in SHLF TDRs despite both the ΔC SHLF-eGFP and $\Delta N\Delta C$ SHLF-eGFP lines not exhibiting any GFP fluorescence (Figure 2.6 F and G). *In silico* protein structure prediction using AlphaFold2 Colab notebook (Jumper *et al.*, 2021; Mirdita *et al.*, 2022) showed that the C-terminal domain may act as a linker between the 4th TDR and the eGFP reporter in SHLF-eGFP lines (Figure 2.8). The list of peptides detected upon MS/MS analyses of the bands observed in anti-GFP western blot are mentioned in Table 7.1

These analyses also revealed that a linker-like stretch of 9 residues (ILESNNLQE) is present at the end of the 1st, 2nd and 3rd TDRs (but not the 4th) and may act as a linker between the 2nd TDR and eGFP in N-TDR1-2 eGFP lines (Figure 2.8). MS/MS of the lower and upper (~ 45 kDa and ~ 54 kDa) bands from N-1st-2nd TDR eGFP lines resulted in peptides common between 1st and 2nd TDRs and also 2nd TDR specific peptides (Figure 2.7 B). The N-TDR1-C-eGFP lines showed a ~ 43 kDa band resulting from 1st TDR and C-terminal specific peptides. Our findings suggest the domain-dependent proteolytic processing of SHLF (Figure 2.7 B).

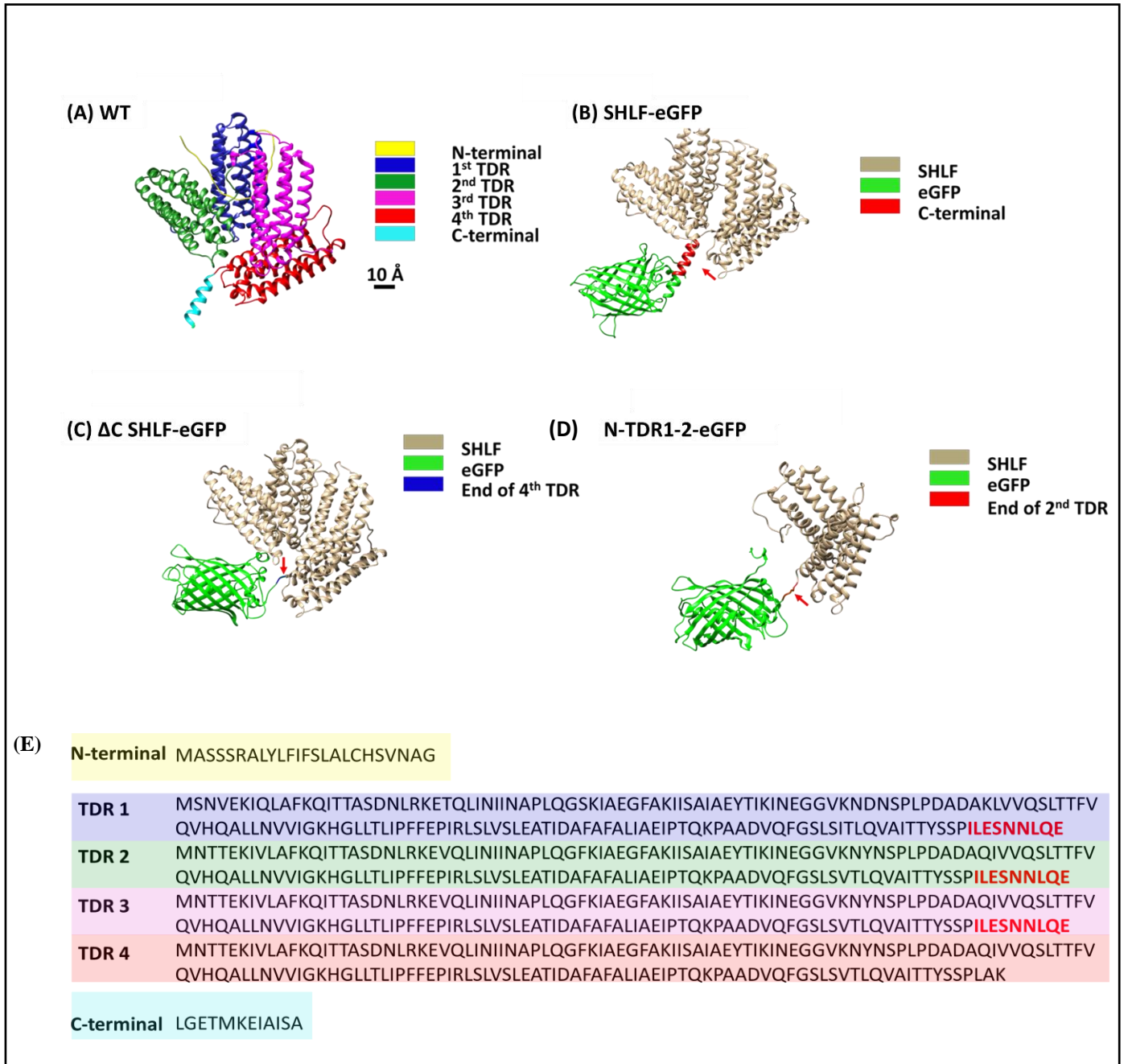


Figure 2.8 AlphaFold2 based protein structure predictions reveal a linker-like region at the end of SHLF TDRs (A) WT (B) SHLF-eGFP structure analyses showed alpha coiled helices interconnected by unstructured linker like regions. The C-terminal at the end of 4th TDR appears to separate the SHLF and eGFP proteins. (C) Δ C SHLF-eGFP fusion protein structure showed close proximity of the truncated SHLF and eGFP. (D) The N-TDR1-2-eGFP fusion protein has a spatial separation of the truncated SHLF and the eGFP, which may be due to the nine amino acid stretch (highlighted in bold red letters) present at the end of TDRs 1-3 (TDRs), but not in TDR4, which terminates in the C-terminal tail. The red arrows indicate the spatial separation between

SHLF and eGFP. (E) Protein sequence of SHLF TRDs with the TDR linker-like region highlighted in red.

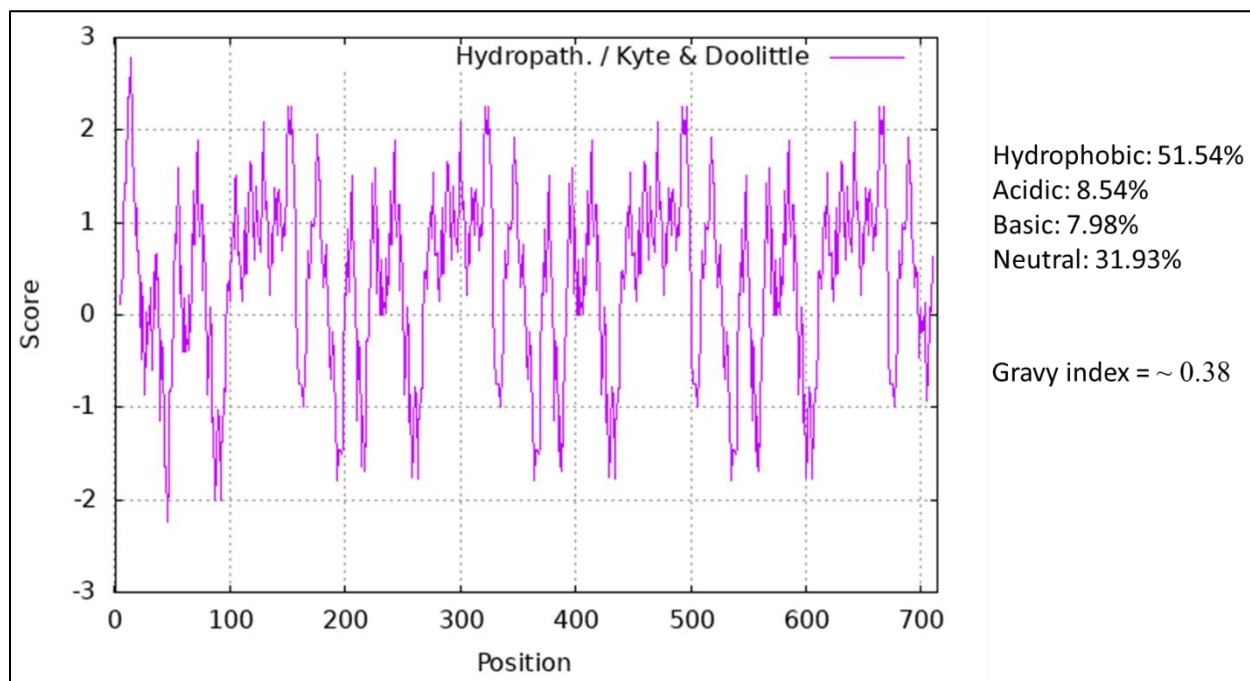


Figure 2.9 Hydropathy index calculation for SHLF protein using the scale Hydropath.

As SHLF undergoes cleavage at multiple sites, it was essential to detect all the cleavage products. To achieve this, we attempted to generate anti-SHLF antibody.

SHLF is a highly hydrophobic protein (~51.54% of amino acids in SHLF are hydrophobic), with a GRAVY index of ~0.38 (Kyte and Doolittle, 1982) (Figure 2.9), it would be challenging to express and purify the 77 kDa SHLF from heterologous bacterial systems. Given the sequence similarity between the four SHLF TDRs, a truncated version of SHLF comprising of the N-terminal, 1st TDR and the C-terminal was cloned in pET 28a using the primer pair and used for protein expression in *E. coli* Shuffle T7 cells (Figure 2.9 A), which confers a 6X Histidine tag to both the N and C-terminal of the protein. Upon western blotting of the induced protein eluted from the Ni-NTA beads, using anti-His antibody, the expected 27 kDa band was observed. This band

was then cut out, gel eluted and analyzed using MS/MS to reveal that the protein obtained upon overexpression was indeed SHLF (Figure 2.10 B-D).

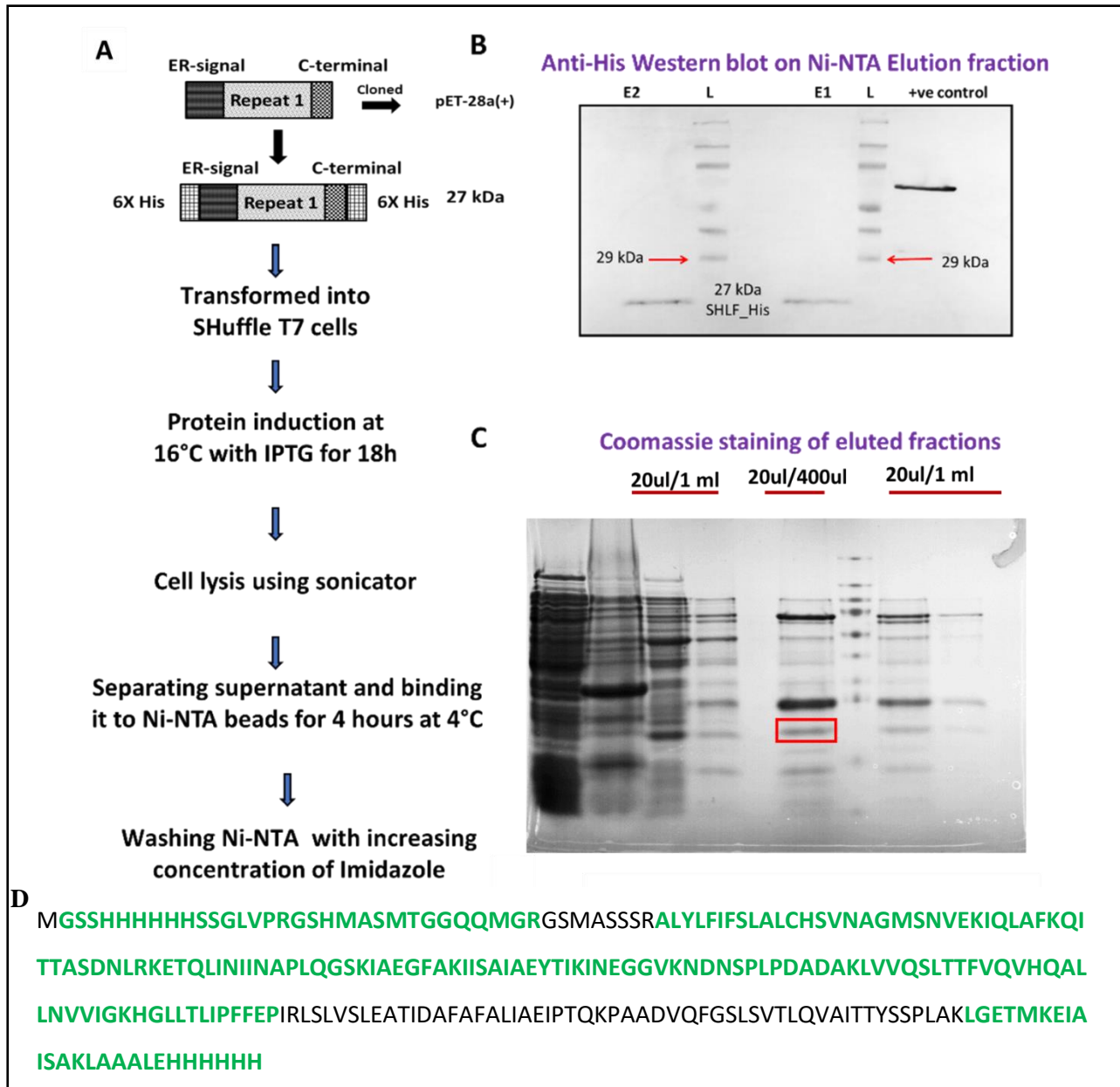


Figure 2.10 Cloning and generation of anti-SHLF primary antibody in rabbit (A) Cloning of the N -TDR1-C terminal of SHLF into pET28a and transformation into *E.coli* SHuffle for protein induction at 16°C (B) Anti-His western blot for confirmation of expressed protein (C) In gel digestion and (D) mass spectrometric analysis of the 27 kDa band for confirmation of expressed SHLF sequence.

0.3 mg of the protein obtained upon overexpression in bacteria was used for primary injections in rabbit, which was followed by booster injections with and 0.15 mg protein. The final bleed obtained from the rabbit showed that the antibody generated was specific to SHLF and the bands obtained upon probing total WT protein with this antibody were absent in a duplicate blot probed with pre-immune sera (Figure 2.11)

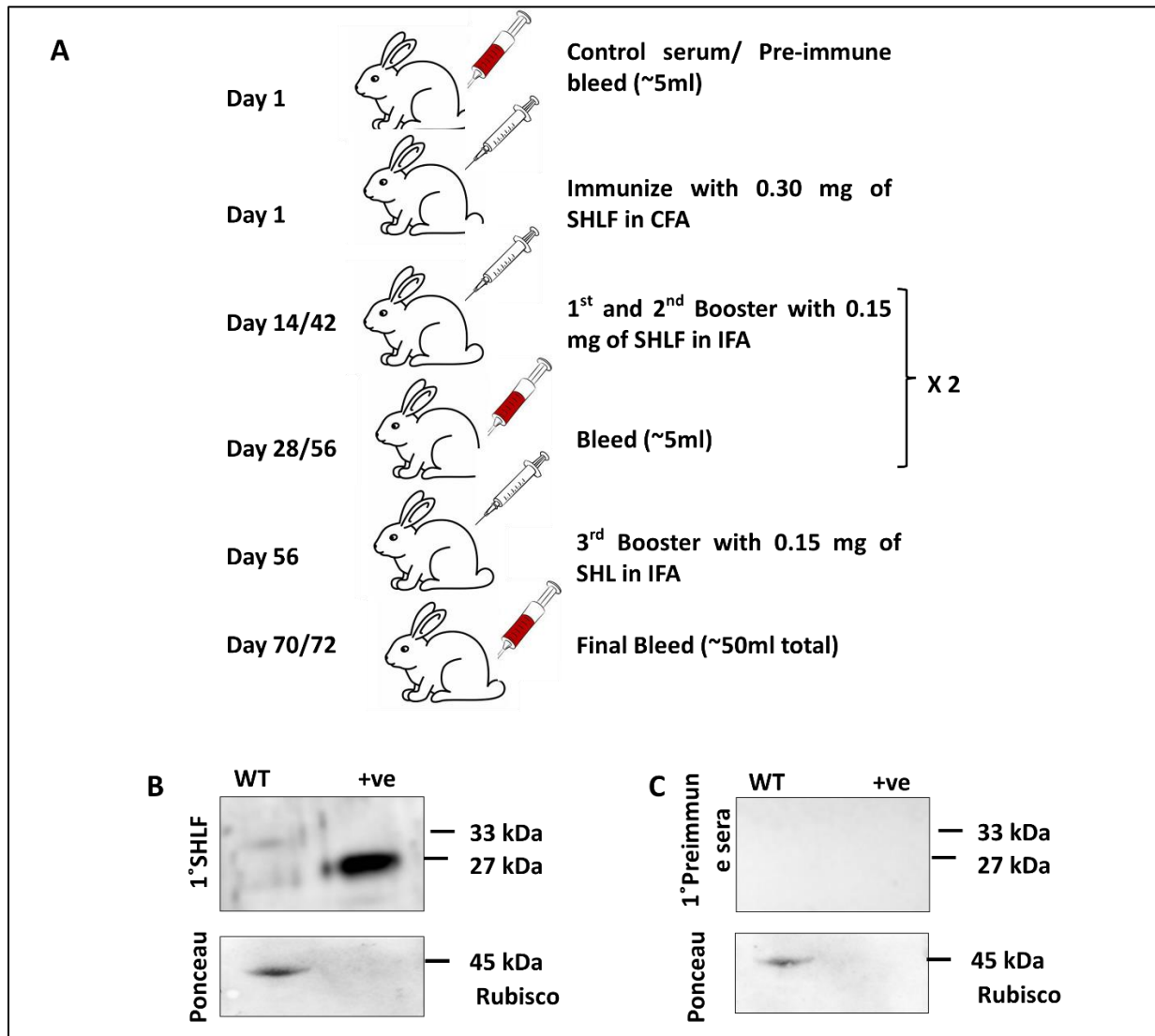


Figure 2.11 In-house generation and validation of anti-SHLF primary antibody in rabbit (A) Injection of an emulsified 27 kDa SHLF protein band (0.3 mg) with Freund's complete adjuvant for antibody generation. Subsequent immunizations were done with booster doses of 0.15 mg protein and Freund's incomplete adjuvant and the final sera collected was probed using anti-SHLF antibody. (B-C) 20 µg of total protein from WT and 30 µg from positive control probed with diluted (1:10000) 1° anti-SHLF antibody (B) and 1° pre-immune sera (as a control) (C) showed

distinct SHLF bands only in the test sample. Rubisco protein (45 kDa) was used as a loading control.

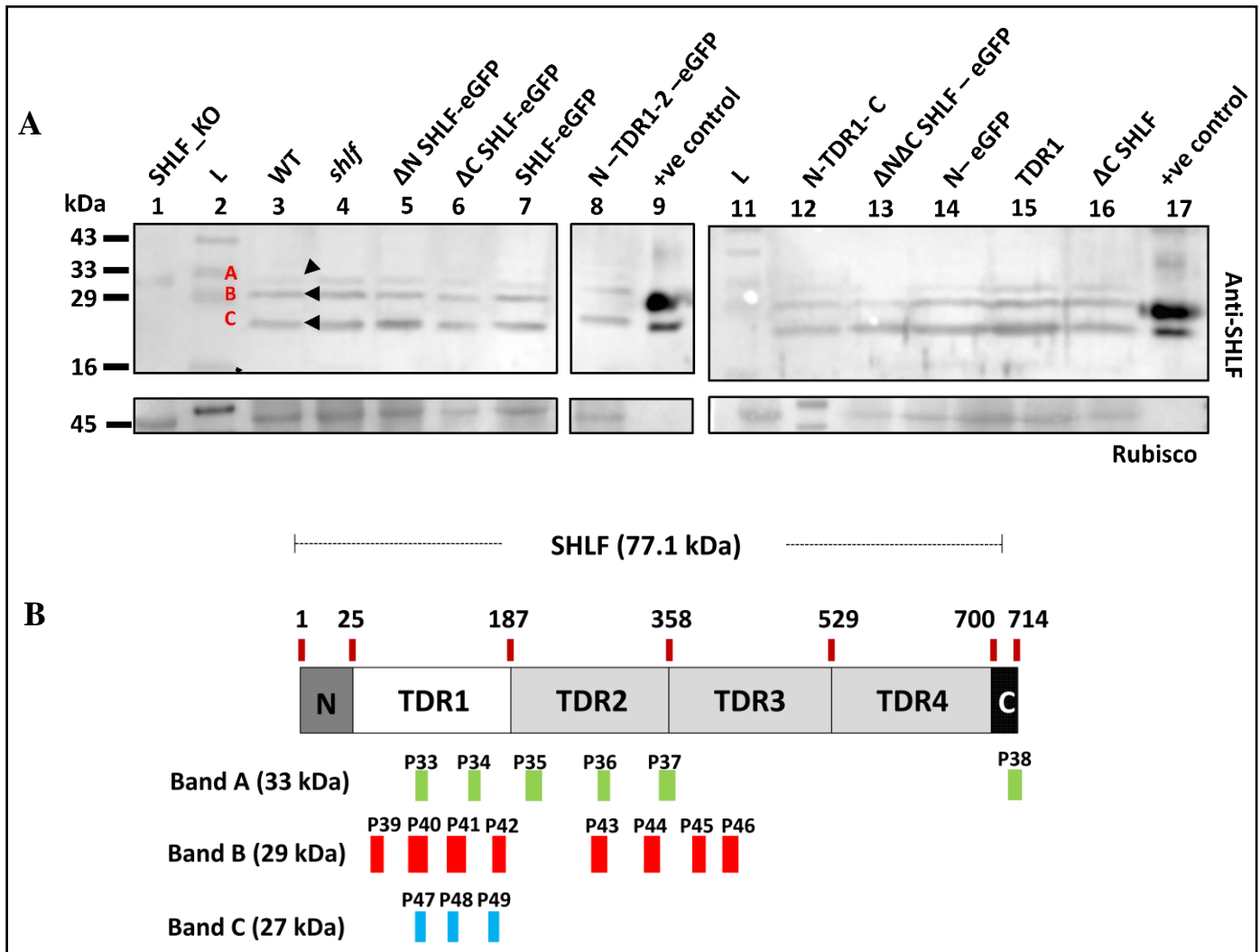


Figure 2.12 Western blotting with anti-SHLF to show that SHLF undergoes cleavage in-planta to form smaller protein fragments (A) Western blots of equal amounts total protein from three week old gametophores with anti-SHLF shows that SHLF undergoes proteolytic cleavage; bands A, B and C (~42 kDa, ~33 kDa and ~27 kDa) are marked by black arrowheads. SHLF_KO protein and N-1st TDR-C heterologously expressed SHLF are used as negative and positive controls respectively (The upper ~ 33 kDa band is a non-specific bacterial protein) (B) Schematic representing the peptides detected upon MS/MS analyses of the three SHLF bands (Origin sites of tryptic peptides; 33 kDa-Green rectangles, 29 kDa-Red rectangles, 28 kDa-Blue rectangles)

Immunoblotting with polyclonal anti-SHLF antibody resulted in three bands (labeled as A, B, C) across WT, *shlf*, and the domain deletion lines (Figure 2.12 A). SHLF_KO and the 27 kDa N-

TDR1-C protein expressed in bacteria were used as negative and positive controls, respectively. MS-MS of band A (~33 kDa) showed tryptic peptides common to all the four TDRs (Figure 2.12 B). Since the TDR2, 3 and 4 are near-identical, their tryptic-peptides cannot be uniquely assigned to any one TDR. In addition, the band B (~29kDa) and C (~27kDa) had peptides common to all the four TDRs (Figure 2.12 B). The list of tryptic peptides observed in all the aforementioned lines are given in Table 2.1.

Table 2.1: List of peptides detected upon MS/MS analyses of the bands observed in anti-SHLF western blot

Band Size	Sequence	TDR number	Peptide number
33 KDa	AGMSNVEKIQLAFK	1st TDR	P33
	PLQGSKIAEGFAK	1st TDR	P34
	NIINAPLQGSK	1st TDR	P35
	ALLNVVIGKHGLL	All TDR	P36
	QITTASDNLRK	All TDR	P37
	LGETMKEIAISA	C terminal	P38
29 KDa	SLALCHSVNAGMSNVEKIQLAF	N and 1st TDR	P39
	ETQLINIINAPLQGSKIAEGFAK	1st TDR	P40
	INEGGVKNDNSPLPDADAK	1st TDR	P41
	NDNSPLPDADAK	1st TDR	P42
	IVLAFKQITTASDNLR	2,3,4 TDR	P43
	QITTASDNLRK	All TDR	P44
	IISAIAEYTIKI	All TDR	P45
	HGLLTLIPFFPIR	All TDR	P46
27 KDa	QITTASDNLRK	All TDR	P47
	IISAIAEYTIK	All TDR	P48
	TFVQVHQALLNVVIGK	All TDR	P49

2.2. Discussion

The SHLF protein family is unique owing to the distinctive long near-perfect tandem direct repeats (TDRs) that do not contain any recognized conserved domains. SHLF in *P. patens* encodes a signal peptide containing N terminal region, followed by four highly similar TDRs and a small C terminal tail. Our previous report had demonstrated that the full length SHLF traffics to the ER via a N terminal signal peptide, where it was presumed to be subsequently cleaved (Mohanasundaram *et al.*, 2021). SHLF TDRs 2-4 are ~99% similar to each other at amino acid level, whereas TDR1 is slightly divergent with a ~90% sequence similarity to the rest of the TDRs. The absence of any known conserved domains in SHLF combined with high sequence similarity within the repeats and the fact that SHLF undergoes cleavage promoted us to question about the importance of SHLF domains in its post translational processing and the function of its cleavage products. To address this, we conducted sequential deletion analysis of the protein to elucidate its functional domains.

2.2.1. N-terminal mediated ER trafficking is essential for SHLF's function in gametophore development

N-terminal signal peptides are short sequences (~15 to 30 amino acids) which act as localization signals to deliver the proteins into its target organelles, and are then subsequently cleaved off. For most proteins secreted via the conventional secretory pathway, an N-terminal ER signal peptide is necessary for proper protein function and secretion (Walter, Gilmore and Blobel, 1984; Rothman, 2014). SHLF possesses a short N-terminally encoded ER signal peptide (24 amino acids), which is found to be essential for protein localization and function as evident from the mutant lines complemented with a Δ N SHLF-eGFP construct (Figure 2.2 D, 2.3, 2.4), which failed to recover the phenotype. Our study shows that ER trafficking is the first crucial step of SHLF's post-translational processing and function, as indicated by the lesser SHLF cleavage products detected in the Δ N SHLF-eGFP lines (Figure 2.7A) and absence of any secretory SHLF peptides (Table 7.4). Since limited cleavage of SHLF is observed in the Δ N SHLF-eGFP lines (Figure 2.6, lane 4), where the protein does not trafficks to the ER, it is plausible that SHLF may also be subjected to proteolytic processing outside the ER. This also raises interesting questions about potential proteolytic processing of SHLF outside the ER, highlighting the need for further research

in this area. However, given the absence of any known domains and high repeat similarity, it is currently challenging to comment on the nature of the protease family responsible for the processing of SHLF.

2.2.2. The C-terminal tail of SHLF may be functionally redundant

C-terminal sequences in proteins have been shown to be associated with protein folding and stability (Björnsson, Mottagui-Tabar and Isaksson, 1996; Sharma *et al.*, 2016). (Wallace *et al.*, 1982; Fortelny *et al.*, 2015; Tanco, Gevaert and Van Damme, 2015; Sharma *et al.*, 2016). The SHLF TDRs terminate in a short non-conserved C-terminal tail (15 amino acids - LGETMKEIAISA), which appears to protrude outside the protein structure and may be exposed to the solvent phase (Figure 2.8) to facilitate the proteolytic processing of SHLF. Interestingly, as the C-terminal sequence was not detected in the three bands observed in the western blot upon probing the WT protein with polyclonal SHLF antibody, this indicates that the C-terminal tail may get lost upon proteolytic processing of SHLF (Figure 2.12, Lane 3). This observation aligns with existing knowledge in the field, which indicates that the exposed C-termini of pro-peptides are susceptible to proteolytic cleavage. These regions often serve as sites for both exo- and endo-peptidase activity, resulting in the release of biologically active peptides.

We hypothesize that the C-terminal tail is essential for protein stability and processing. This is highlighted by the fact that Δ C SHLF-eGFP lines failed to recover the phenotype (Figure 2.1E-G, L-O) and also did not exhibit any GFP signal (Figure 2.6 F-G; 2.7A), despite showing appropriate construct overexpression (Figure 2.3). The significance of the C-terminal region in protein processing is underscored by the absence of SHLF peptides in the secretome, despite the presence of signal peptides in these lines (Table 7.4), emphasizing its crucial role in peptide release and secretion.

Interestingly, our data also suggests that the C-terminal sequence may bear functional redundancy with the TDR linker regions. AlphaFold2 based protein 3D structure prediction showed a probable linker region (9 amino acids - ILESNNLQE) between the TDRs (TDR1-2, TDR2-3 and TDR3,4) but is absent at the end of the TDR4 which terminates in the C-terminal tail. (Figure 2.8) Hence, in the eGFP-fusion protein overexpression lines, the role of C-terminal as a

linker at the end of TDR4 to spatially separate SHLF and eGFP might be crucial. Even in the absence of any fusion protein, C-terminal deletion renders SHLF non-functional (Δ C SHLF) (Figure 2.2 F). This is in contrast to the miniSHLF overexpression lines, where despite the absence of a C-terminal tail, the truncated construct comprising of N terminal and TDR1-2 (terminating in the linker between TDR 2 and 3) is sufficient for protein function. As TDR linkers are well known to function as hubs for proteolytic activity, protein stability and function (Dice, 1990; Kloss and Barrick, 2009; Yadav, Saikhedkar and Giri, 2021), we propose that SHLF TDR linkers may behave similarly.

We also note that, the *shlf* mutant does not express the C-terminal region, as a result of the Tnt1 insertion in TDR4 of SHLF, but is still able to undergo proteolytic cleavage, as evident from the western blot (Figure 2.7, Lane 5,6). This suggests that the SHLF C-terminal sequence may not be indispensable for function, but may function redundantly as a TDR terminating “minimotif” as a scaffold for protease interaction and proteolytic activity (Sharma and Schiller, 2019).

2.2.3. Signal peptide along with the first two TDRs constitute the minimal functional SHLF (miniSHLF)

Near perfect tandem repeats (TRs) participate in several physiological processes in plants. Though TRs are hotspots of dynamic evolution, the copy number of TRs are under stringent control in all living systems. For example, aberrant repeat expansion or contraction leads to several pathological diseases in human (Ryan, 2019). In fact, deletion in TRs can have adverse impacts on protein function (Puopolo *et al.*, 2001). For example, long TRs in fibrous proteins such as collagen or α -helical coiled coil proteins require all their repetitive units to maintain structure (Kajava, 2012).

Given the remarkably high degree of sequence conservation in the tandem repeat domains (TDRs) of SHLF, we postulated that this protein may exhibit modular behavior. Our findings support this hypothesis and reveal functional redundancy between the TDRs. In our initial attempt to overexpress the N-TDR1-C-eGFP construct, we were unable to recover the mutant phenotype despite observing successful protein trafficking to the ER and proteolytic cleavage (Figure 2.2 I, L-O; Figure 2.6 H; Figure 2.7 A and B). These results suggest that the divergent TDR1 alone is

insufficient for SHLF function and that specific sequences within the conserved TDRs (TDR2-4) may be necessary. In contrast, our second construct, N-TDR1-2-eGFP (miniSHLF), not only exhibited protein trafficking to the ER and proteolytic cleavage, but also fully restored *shlf* phenotypes (Figure 2.2 J, L-O; Figure 2.6 I; Figure 2.7 A and B). The functionality and stability of the miniSHLF (despite the terminal deletion of the TDR3-4) closely resembles the deletion tolerance of modular repeats containing proteins like Ankyrin and poly-RNA polymerase II, (Nonet, Sweetser and Young, 1987; Tripp and Barrick, 2004). We propose that SHLF behaves as a modular protein with the conserved TDRs serving as the composite parts of the total protein structure.

2.2.4. SHLF undergoes cleavage *in planta*

Tandem repeats (TRs) are known to be enriched in extracellular proteins but not much is known about their proteolytic processing and secretion (Rothberg *et al.*, 1990; Schaper and Anisimova, 2015; Schoelmerich *et al.*, 2023). Our previous report indicated that SHLF undergoes proteolytic cleavage, (Mohanasundaram *et al.*, 2021) but all cleavage products could not be detected. In this study, we detected all the SHLF products (~ 27, ~ 29 and ~ 33 kDa) which were surprisingly comparable between the WT, *shlf* and the overexpression lines but were absent in the SHLF knockout lines (SHLF_KO) (Figure 2.12). Additionally, our observations did not reveal any truncated protein in the mutant (Figure 2.12.).

As the *shlf* locus has a Tnt1 insertion at the end of TDR4, this results in the suppression of *SHLF* transcript level (Mohanasundaram *et al.*, 2021), possibly by the epigenetic suppression of transcription by the retrotransposon (Saze and Kakutani, 2011; Hernández-Pinzón *et al.*, 2012). Accordingly, SHLF protein could still be present in the mutant, albeit in low quantity and this nominal amount of protein in the mutant may not be sufficient to be functional. Hence, SHLF may be subjected to threshold level genic regulation as previously reported in other systems (Kobiler *et al.*, 2005; Little, 2005). Furthermore, the high expression level of SHLF in the wild-type (WT) provides additional evidence supporting its significance in establishing the minimal functional threshold.

In summary, we functionally characterized the nature and function of SHLF domains. We found that the N-terminal region is crucial for ER trafficking, while signal peptide-mediated trafficking is essential for protein function. The first TDR alone is insufficient, requiring the second TDR for activity. The minimal functional form, miniSHLF, includes the N-terminal, 1st, and 2nd TDR. The C-terminal domain is essential for function and stability, possibly overlapping with a TDR linker. We detected internal cleavage sites within TDRs, and the N-terminal deletion hindered cleavage, while the C-terminal deletion affected protein stability. Upon Immunoblotting with a SHLF-specific antibody we confirmed proteolytic cleavage in SHLF and were able to detect the truncated products.

Overall, our study provides insights into SHLF domain roles in trafficking, proteolysis, and gametophore development. Here, we identified a minimal form of the protein (miniSHLF) which is sufficient and necessary for gametophore development. Given the fact that SHLF traffics to ER and undergoes cleavage, we hypothesize that it may function as a secretory protein.

CHAPTER 3:

Studying the secretory nature of SHLF and the role of its cleavage products in gametophore development

3. Studying the secretory nature of SHLF and the role of its cleavage products in gametophore development

3.1. Introduction

3.1.1. The plant apoplastic space and secretome

Plants being sessile organisms respond to dynamic external cues by efficient intra- and intercellular mechanism. As plant cells are spatially separated from each other by a cellulosic cell wall, intercellular communication is mediated by either by exchange of cues across the symplasm via the plasmodesmata or by secretion of cues (hormones and metabolites) into the plant apoplast. The apoplast functions as a highway for the transduction of mobile signals throughout the plant, connecting almost all the plant cells. The apoplast has the following major functions (Sakurai, 1998):

1. Shaping cellular dimensions through apoplastic enzymes by regulating deposition and cross-linking of cell wall materials
2. Defense against invading biotic pathogens such as bacteria and fungi
3. Regulates abiotic stress response by controlling the pH and the levels of reactive oxygen species (ROS) in the plants.
4. Transport of proteins, peptides, hormones and metabolites

The term “secretome” mostly refers to describe only the secreted proteins and peptides component of the apoplastic space (Agrawal *et al.*, 2010). Most of the plant secretory proteins have been implicated in a plethora of biological processes including the reduction of developmental paradigms such as cell expansion, intercellular crosstalk and also for responding to biotic and abiotic stresses. Many of these secretory proteins can act as ligand to interact with cellular receptors and trigger down-stream signalling (Agrawal *et al.*, 2010).

3.1.2. Studying the components of plant secretome

A majority of the secretory proteins which follow the conventional secretory pathway contain an N-terminal signal peptide and lack transmembrane or membrane anchored domains (Agrawal *et al.*, 2010). The protein trafficks to ER via its signal peptide, which is cleaved off by an ER-resident signal peptidase. The protein is further subjected to post-translational processing such as proteolytic cleavage or addition of modifications like acetyl, phosphoryl,

glycosyl and methyl groups and is ultimately secreted after passing through the trans-Golgi network (Matsubayashi, 2014).

As most plant cells have rigid cellulosic cell walls, the majority of previous attempts made to study plant secretome were carried out using undifferentiated suspension cell/tissue culture. However, as the true nature of plant secretomes can only be revealed through *in planta* studies, attempts were made to isolate secretory proteins and peptides by vacuum infiltration techniques in vascular plants such as *A. thaliana*, tobacco (*Nicotiana tabacum*), rice (*Oryza sativa*), *Medicago truncatula*, *Medicago sativa* and pea (*Pisum sativum*) (Agrawal *et al.*, 2010; Alexandersson *et al.*, 2013). *In planta* secretome studies have also been attempted for plants such as tomato (*Solanum lycopersicum*) (Konozy *et al.*, 2013), sunflower (*Helianthus annuus*) (Pinedo *et al.*, 2012) and the seabuckthorn (*Hippophae rhamnoides*) (Gupta and Deswal, 2012).

Non-vascular plants such as the model moss *Physcomitrium patens*, offers several advantageous characteristics to aid secretome studies. This predominantly haploid bryophyte starts its life cycle by forming filamentous protonema, which later differentiate to give rise to a gametophore with single cell layered leaves and having structurally simple cell walls. Lehtonen *et al.*, demonstrated that these plants can be cultured as such in liquid cultures which encourages protein secretion into the media because of the close contact of the plant surface with the liquid (Lehtonen *et al.*, 2014). The secreted proteins can then be subjected to 2D gel electrophoresis and mass spectrometric analyses for identification.

3.1.3. Secretory proteins and peptides in plant development and stress response

Secretory proteins and peptides play major roles in controlling cell growth and also form the first line of defense against pathogenic invaders (Chung and Zeng, 2017). Plant proteases are present both in intracellular compartments and may also be secreted into the extracellular space (Antalis and Buzza, 2016). Their major role is to cleave the pathogenic effector proteins and mediate plant defense, however not much is known about the role of extracellular proteases cleaving secretory plant proteins to form functional signalling peptide. Small post translationally modified peptides (PTMPs), mostly generated from longer inactive pro-peptides inside cellular compartments and are known to influence both development and stress response in flowering plants (Matsubayashi and Sakagami, 2006; Katsir *et al.*, 2011). Many of the known PTMPs are also known to cross-talk with phytohormones like auxin to govern plant development (Detailed in section 1.6.3 in this thesis).

In the past few years, there has been a concerted effort to unravel the precise molecular functions of PTMPs in non-flowering plants (Furumizu *et al.*, 2021). Recent studies have reported the presence of small PTMPs including of PSK (PHYTOSULPHOKINE) (Tost *et al.*, 2021) and CIF (CASPARIAN STRIP INTEGRITY FACTORS) (Furumizu *et al.*, 2021) in *P. patens*. Whitewoods *et al.*, showed that CLE peptides mediate the crucial 2D to 3D growth transition in moss (Whitewoods *et al.*, 2018). On the other hand, RALF peptides in moss have been implicated both in moss protonemal tip growth and immune response (Ginanjari, Teh and Fujita, 2021; Mamaeva *et al.*, 2022).

Our previous results showed that SHLF trafficks to ER via its N-terminal signal peptide and undergoes cleavage. Given the significance of the signal peptide on ER trafficking and function, we hypothesized that SHLF may follow the conventional secretory pathway and function as a secretory protein. Here, we have attempted to test the secretory nature of SHLF and its cleavage products. We also identified and characterized the functional secretory peptides (SHLFpeps) generated from its TDRs and investigated their role in auxin distribution and gametophore development in moss.

3.2. Materials and methods

3.2.1. Plant tissue culture and maintenance

The moss *Physcomitrium patens* (Ecotype Gransden) was grown in BCDAT media at 24 °C under 16 h photoperiod conditions for all experiments, as per section 2.2.1. For moss secretome preparation, freshly homogenized protonemal tissues were grown on BCDAT media with 0.8 % agar, overlaid with cellophane membranes for a week, followed by transfer to BCD media with 0.8 % agar and were incubated for a week to enhance the growth of gametophores. The plants were then transferred to 50 mL BCD liquid media containing 5mM ammonium tartarate and 1% glucose and incubated for a week at 24 °C, 80 rpm under 16:8 h photoperiod.

3.2.2. Secretome and extracellular peptidome isolation

The moss secretome was filtered through a 2 µm Whatman filters (Sigma, GMF 2UM) to remove all plant material and lyophilized at -80 °C. Total secretome protein was precipitated using 20% TCA/Acetone. The protein pellet was washed sequentially with 100% and 80% acetone and dried before re-suspending in MilliQ water for further use. To isolate extracellular peptides from the secretome, the moss secretome was filtered through 0.2 µm filter (Whatman filter), lyophilized and resuspended in 500 µL 5% ACN/0.1% TFA. Further extraction of

peptides were carried out using DSC18 cartridges (Discovery DSC-18, Supelco, USA) (Fesenko *et al.*, 2019). The peptide pool was desalted using C-18 cartridges (Sigma, USA) and subjected to further mass spectrometric analysis.

3.2.3. Secretome and peptidome supplementation assays

Secretome harvested from 100 ml liquid cultures of WT, *shlf* and miniSHLF was lyophilized and precipitated using 20% TCA/Acetone. The protein pellet isolated from each line was resuspended in 100 μ L BCDAT liquid media and supplied to freshly homogenized WT and *shlf* protonema. For peptidome, the protocol of was followed and supplemented as stated above (Fesenko *et al.*, 2019). The cultures were incubated for a month in the aforementioned conditions (Section 2.2.1). Gametophores were then harvested for phenotypic and RT-qPCR analysis.

3.2.4. Synthetic SHLF peptide supplementation

Synthetic SHLF peptides of >95% purity (BiotechDesk, India) were dissolved in 1:1 (v/v) ratio of ACN: H₂O to a stock concentration of 100 mM. Protonema-stage moss plants were cultivated on BCDAT media, enriched with diluted concentrations ranging from 1 to 1000 μ M (1 μ M, 10 μ M, 25 μ M, 50 μ M, 75 μ M, 100 μ M and 1000 μ M) and grown for a month as per section 2.2.1.

3.2.5. Western blot analyses

Equal amounts of protein from the isolated secretomes were loaded on a 12.5% SDS-PAGE gel and run for 1.5 h at 110 V. The protein samples were then transferred to a PVDF membrane in phosphate transfer buffer (pH 7.4) at 4°C for 3 h. The samples were probed using rabbit anti-SHLF at a dilution of 1:20,000 followed by goat anti-rabbit secondary antibody at a dilution of 1:10,000. Purified SHLF protein (27 kDa; N-TDR1-C) and total protein from *SHLF_KO* lines were used as positive and negative controls respectively.

3.2.6. Sample preparation for MS/MS analyses

To identify the sequence of the protein bands detected in the western blots, an identical SDS-PAGE 12.5 % gel was run for 1.5 h at 110V. The bands of interest were then cut from the PAGE gel and used for in-gel trypsin digestion and peptide extraction (Kelkar *et al.*, 2019). The samples were then desalted using ZipTips (Merck Millipore) subjected to LC-MS/MS based peptide identification. The data generated from IDA (Information-

dependent acquisition) mode were interpreted using Protein Pilot software. For moss extracellular peptide extraction, the protocol of Fesenko *et al*, 2019 was followed with minor modifications (Fesenko *et al.*, 2019). Forty (40) mL of liquid secretome of WT, *shlf* and the transgenic lines were used for the isolation of extracellular peptidome as per Fesenko *et al*, 2019. The extracellular peptides were then desalted, resuspended in 0.1% TFA and used for LC-MS/MS analyses and peptide identification as detailed below.

3.2.7. LC-MS/MS analysis and peptide identification

We performed mass spectrometry analysis using a Sciex TripleTOF6600 mass spectrometer coupled with an Eksigent nano-LC 425. We used both tryptic and extracellular peptides ($\geq 1 \mu\text{g}$) and conducted six independent biological and three technical repeats. Samples were loaded onto an Eksigent C18 trap (5 μg capacity) and eluted using an Eksigent C18 analytical column (15 cm x 75- μm internal diameter) with a linear acetonitrile gradient. The LC run lasted 4 hours at a constant flow rate of 300 nL/min, with solvent A containing water and 0.1% formic acid and solvent B containing acetonitrile. The gradient schedule involved a 10-minute run with 5% (vol/vol) B, followed by a linear gradient of B from 0% to 80% (v/v) over 80 minutes, 15 minutes with 80% (v/v) B, and 15 minutes of equilibration with 5% (v/v) B. We acquired data in information-dependent acquisition (IDA) mode over a mass range of 100-3,000 m/z, with each full MS survey scan followed by MS/MS of the 20 most intense peptides. Dynamic exclusion was enabled with a repeat count of 1 and exclusion duration of 6 seconds. Peptide identification and quantification were conducted using Protein Pilot, with the *Physcomitrium patens* protein database containing the protein sequences from Phytozome v12.0 used for tryptic peptide identification. The searches were conducted using the Protein Pilot algorithm, with +57.0215 m/z iodoacetamide alkylation of cysteine as a static modification. Precursor ion and MS/MS mass tolerances were set at 20 and 40 ppm, respectively, for peptide searches. For reductive-dimethylation-based quantification of tryptic peptides, we specified the N-terminal and lysine dimethyl labels as fixed modifications in the search algorithm. We calculated the PSM-level false-positive discovery rate (FDR) with decoy database search and filtered peptides and proteins for <1% FDR.

For the extracellular peptide analysis, mass-spectrometry data were searched against a database containing protein sequences from Phytozome v12.0 using Protein Pilot. The peptide FDR was set to 1%, while the Protein Pilot protein FDR filter was disabled. The “Digestion Mode” was set to “unspecific”, and no modifications were allowed. All other parameters were

kept at their default values. PSM features, such as length, intensity, number of spectra, intensity coverage, and peak coverage, were extracted from Protein Pilot's .xlsx files.

3.2.8. RT-qPCR analysis

Three-week-old moss gametophores were subjected to total RNA extraction using RNAiso Plus (Takara Bio USA Inc.) Two micrograms (2 µg) of RNA samples were reverse-transcribed using SS-IV reverse transcriptase (Invitrogen, USA) and oligo dT primers. SYBR Premix Ex Taq II (Tli RNaseH Plus) was used for the relative quantification of transcripts via Bio-Rad CFX96 Touch Real-Time PCR Detection System (Bio-Rad, USA). Moss *β-Actin* and *eIF2* were used as reference genes for the quantification of target gene expression levels in all samples. The fold-change values (sample value/reference value) and $2^{-\Delta\text{Ct}}$ were calculated according to Schmittgen *et al.*, 2008. (Schmittgen and Livak, 2008). All the primers used for RT-qPCR analysis are listed in Table 5.

3.2.9. Phenotypic characterization

For detailed phenotypic analyses, gametophores at the three-week-old stage derived from lines supplemented with secretome, peptidome, or synthetic peptides were utilized. For each line, length and width of the 9th leaf from the apex and internodal distance of gametophores (n=20) were recorded using Olympus stereomicroscope and measured using ImageJ (Fiji). Silhouettes for the first nine leaves of each line were traced using Inkscape (Version 0.92.3-1) from the images obtained on Leica stereomicroscope.

3.2.10. Statistical analyses

Boxplots were generated from the datasets using the ggplot2 package with the whiskers representing the closest data-points with the 1.5x interquartile range following Tukey's style (Tukey, 1949) above the first quartile and below the third quartile respectively. Outliers were also included in the analyses and plotted along the boxplots. The bar plots were plotted with error bars depicting the standard deviation between the samples. For datasets having normal distribution and equal variance, Student's *t* test was performed. *P*-value < 0.05, <0.001, and <0.0001 were marked as *, **, and ***, respectively. In case of samples having non-normal distribution, Levene's test was conducted to check for homogeneity in the population (Levene, 1960). Samples having *p*-value <0.05 were considered non-homogenous and Kruskal–Wallis test was performed with *post hoc* Dunn test having Bonferroni corrected *p*-values (Bonferroni, 1936; Dunn, 1964; McKight and Najab, 2010). The letters on top of the

graphs depict the similarity between two data sets. Graphs were generated using GraphPad prism (version 9.5.0) and R (version 4.2.3).

3.2.11. GUS staining

Three-week-old *P. patens* colonies grown on BCDAT medium with 0.8% agar was used for glucuronidase (GUS) staining assay. The plants were transferred to GUS staining solution and incubated at 37 °C overnight in dark (Jefferson *et al.*, 1987). The plants were dehydrated on the following day using a series of ethanol washes (30%, 50%, 80% and 100%), followed by imaging using a Leica S8 APO stereomicroscope (Leica, Wetzlar, Germany)

3.3. Results

3.3.1. A truncated version of SHLF is an apoplast resident protein

As SHLF trafficks to ER via a secretory signal peptide (Mohanasundaram *et al.*, 2021), we hypothesized that SHLF could function as a secretory protein

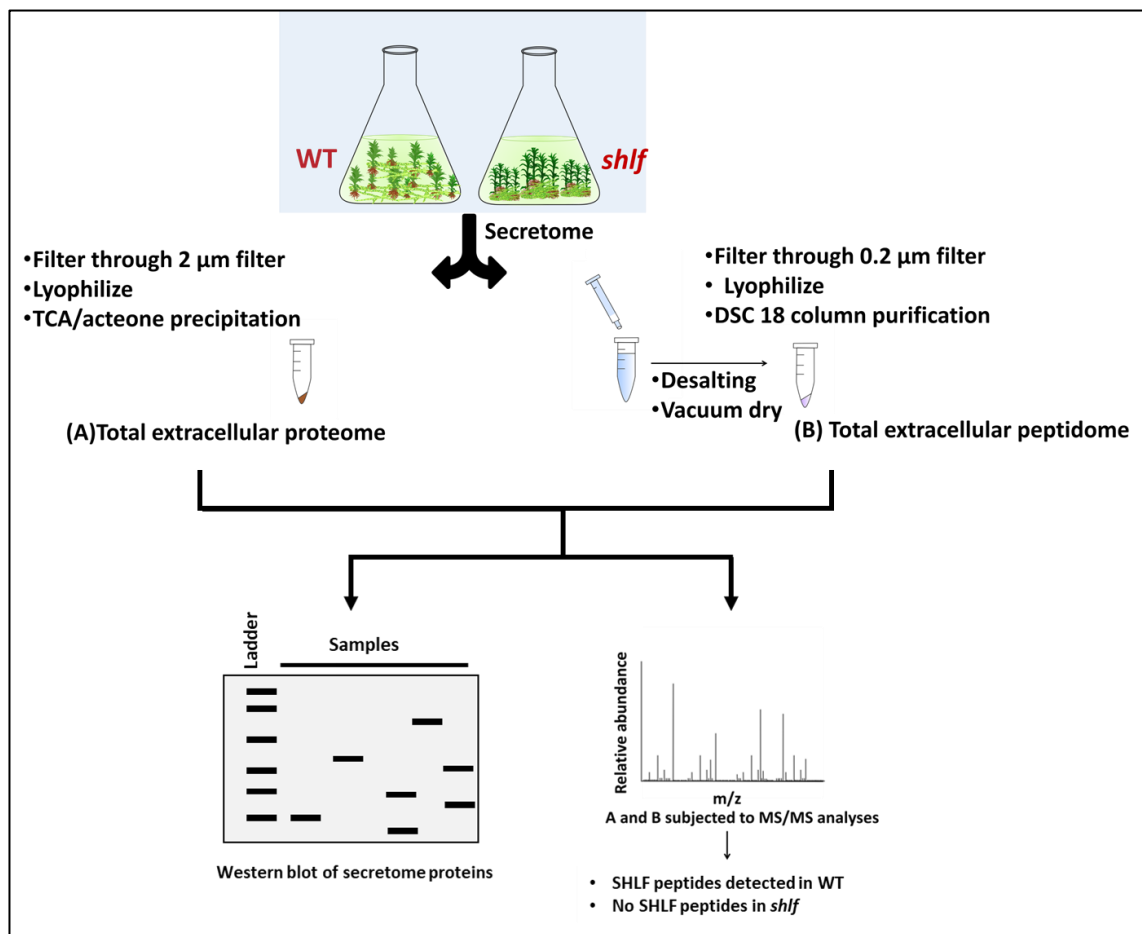


Figure 3.1 A schematic representation of isolation and identification of total extracellular proteins and peptides from WT and *shlf*

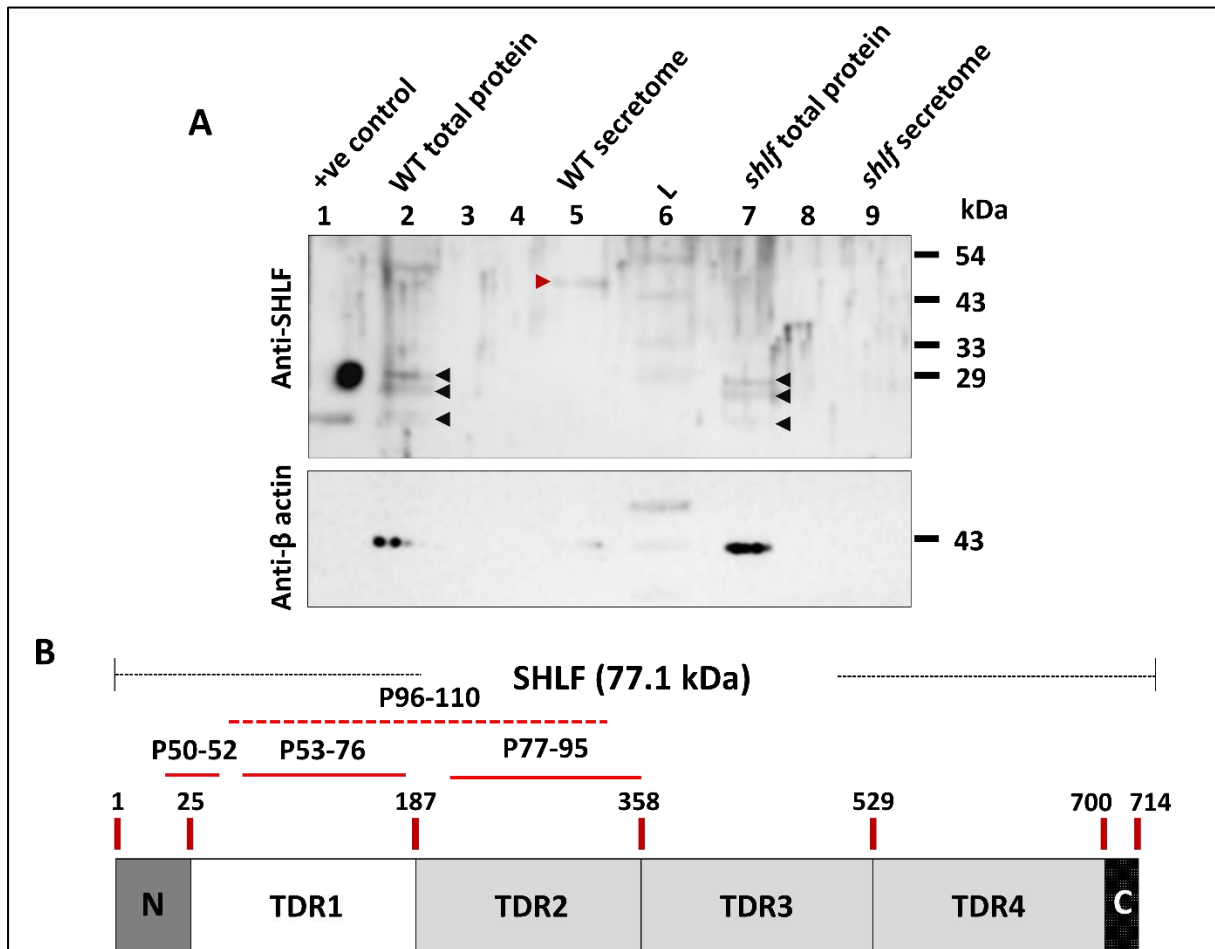


Figure 3.2 Immunoblotting with polyclonal SHLF antibody (A) Immunoblotting using anti-SHLF against the three-week-old WT and *shlf* secretomes show a ~ 43 kDa band in WT, with no distinct bands in *shlf* secretome. This shows that a truncated form of SHLF exists in the WT secretome. (B) Schematic represents the numbered tryptic peptides detected upon MS/MS analyses of the SHLF bands in WT secretome (red). Non-specific bands observed in some of the lines were also analyzed by MS/MS and found to be a result of overexposure

The WT and mutant secretome was isolated (Lehtonen *et al.*, 2014) (Figure 3.1) and the total secretome protein were probed with anti-SHLF specific antibody (Figure 3.2).

We detected a ~ 45 kDa band of a truncated extracellular SHLF in the WT secretome, which was absent in the mutant (Figure 3.2). MS/MS analyses showed that this band consists of N-terminal, 1st and 2nd TDRs-specific tryptic peptides, whereas no bands were detected in the *shlf* secretome (Figure 3.2). Further, we also carried out MS/MS analysis of the extracellular peptidome from the domain deletion lines to investigate the importance of SHLF domains in its secretion (Figure 3.3, 3.4 and Table 7.4).

The WT and SHLF-eGFP extracellular peptidomes contain several SHLF peptides generated from the four TDRs and the C-terminal. The secretome of N-TDR1-2- eGFP (mini SHLF) lines also showed multiple secreted SHLF peptides, originating from both the 1st and 2nd TDR. SHLF is the most abundant protein in the WT extracellular peptidome (Table 7.2, 7,4), whereas the mutant peptidome lacks any SHLF-specific peptides (Table 7.3). Interestingly, the extracellular peptidomes of non-reverted deletion lines (Δ N SHLF-eGFP, Δ C SHLF-eGFP, TDR1, Δ N Δ C SHLF-eGFP, N-eGFP and N-TDR1-C-eGFP) also had no SHLF-specific peptides (Table 7.4). Based on these findings, we conclude that the N- and C-terminal domains are necessary for SHLF trafficking to the ER, post-translational processing and secretion Figure 3.3 details the secretory SHLF peptides (and their sites of origin) detected from WT secretome.

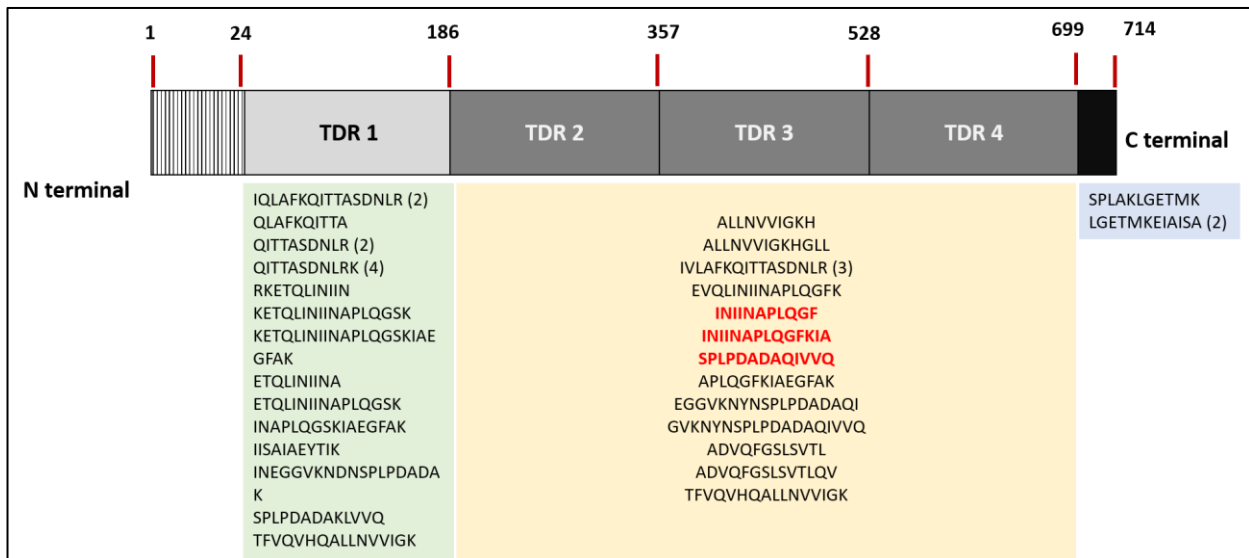


Figure 3.3 A schematic representation the WT SHLF protein and the number and sequence of the secretory SHLF peptides derived from the TDRs and the C-terminal tail

3.3.2. SHLF acts as a secretory protein via functional cryptic peptides

To test the functional importance of secreted SHLF in gametophore development, the mutant was supplemented independently with total secretome and peptidome of WT and miniSHLF (N-TDR1-2-eGFP) (Figure 3.4, 3.5). The level of other housekeeping secretome proteins in WT and *shlf* were mostly comparable except for the absence of SHLF in the mutant (Table 3.1). Supplementation of *shlf* secretome and peptidome served as negative controls

(Figure 3.4, 3.5). The leaf phenotypes observed during the supplementation assays are visually represented using silhouettes (Figure 3.5B).

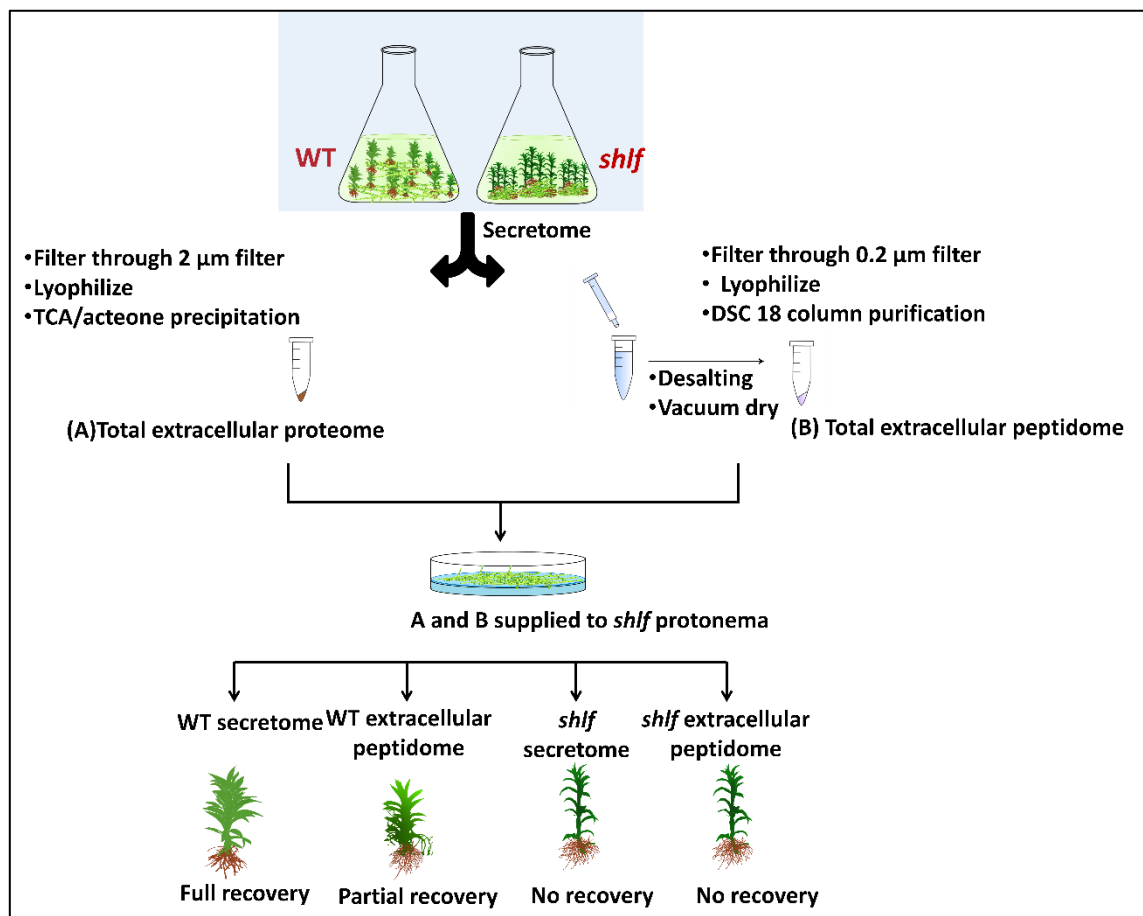
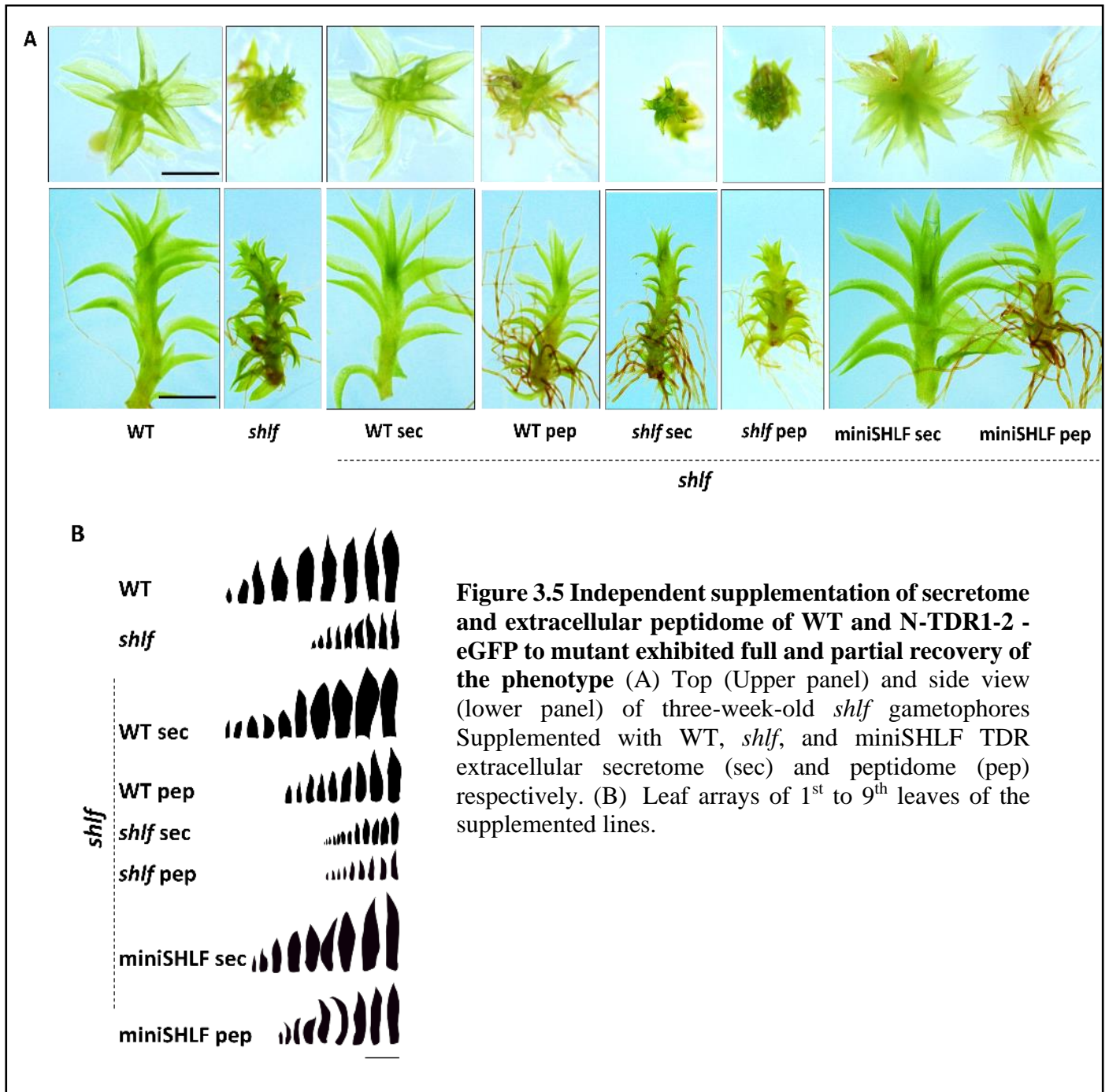
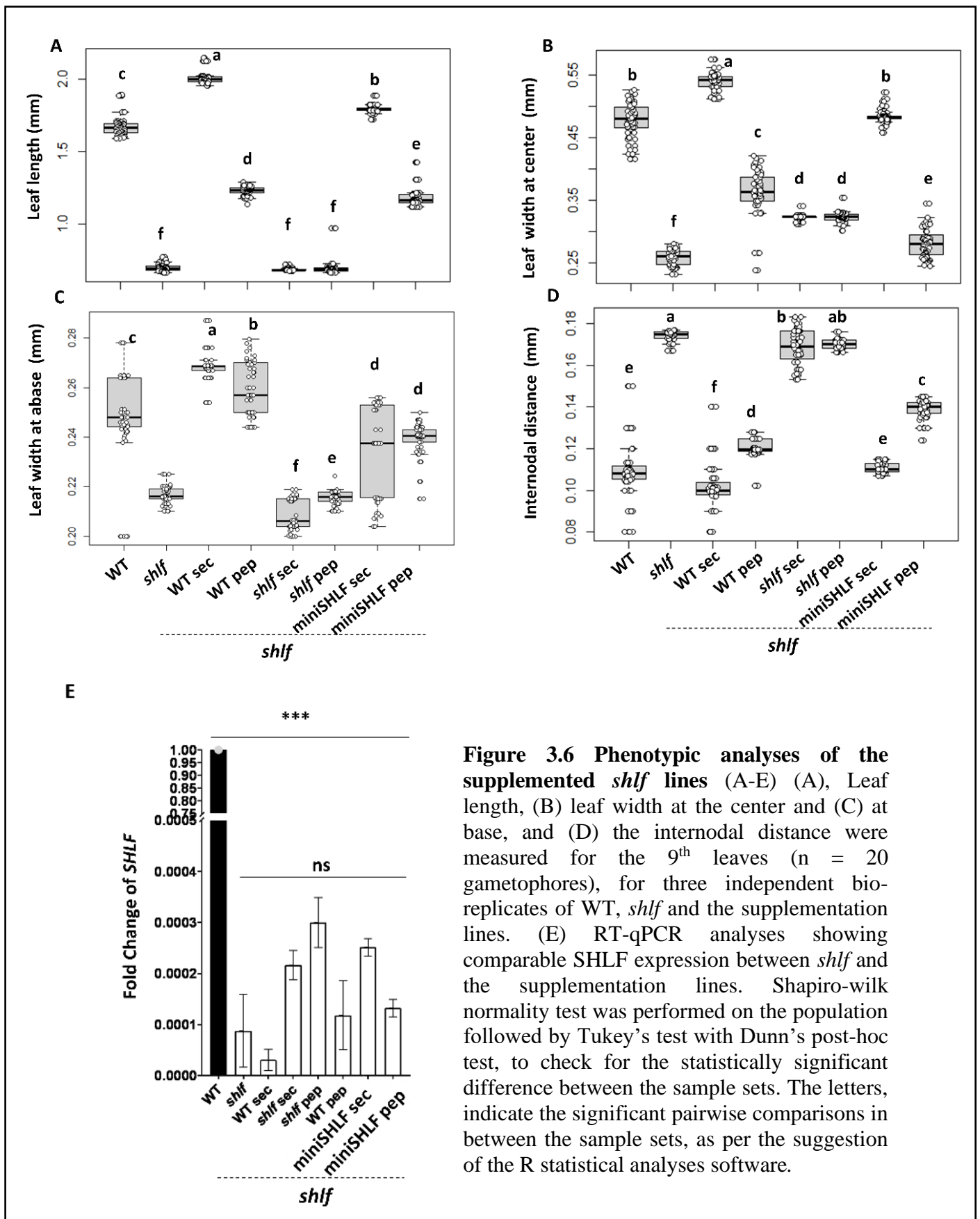


Figure 3.4 A schematic representation of protocol for supplementation of total extracellular proteins and peptides from WT, miniSHLF and *shlf* to the mutant. Supplementation of WT but not *shlf* extracellular proteome and peptidome to *shlf* resulted in complete and partial recovery of mutant phenotypes respectively.

We noted that the independent supplementation of WT and miniSHLF secretomes resulted in the complete recovery of all the mutant phenotypes (Figure 3.5, 3.6 A-D). However, supplementation of the purified WT peptidome to *shlf* showed a partial recovery of leaf length (~73%) and width at the center (~75%) but fully recovered the leaf width at the base and the internodal distance (Figure 3.6 C-D). The miniSHLF peptidome-supplemented gametophores showed a partial recovery of the phenotypes, including leaf length (~70%), leaf width at the center (~60%), width at the base (~96%) and the internodal distance (~82%). Interestingly, no significant difference in *SHLF* transcript levels was observed in any of these supplementation lines (Figure 3.6 E).





3.3.3. SHLFpep 3 affects auxin distribution pattern and gametophore development in a dosage dependent manner

The complete recovery of *shlf* upon overexpression of N-TDR1-2-eGFP (miniSHLF) highlights the importance of conserved TDRs and justifies the need to identify functional peptide/s generated from them. Additionally, the partial recovery of *shlf* upon supplementation of peptidome (from WT and miniSHLF) (Figure 3.6), suggests that SHLF peptides may function in a dosage dependent manner to regulate moss development.

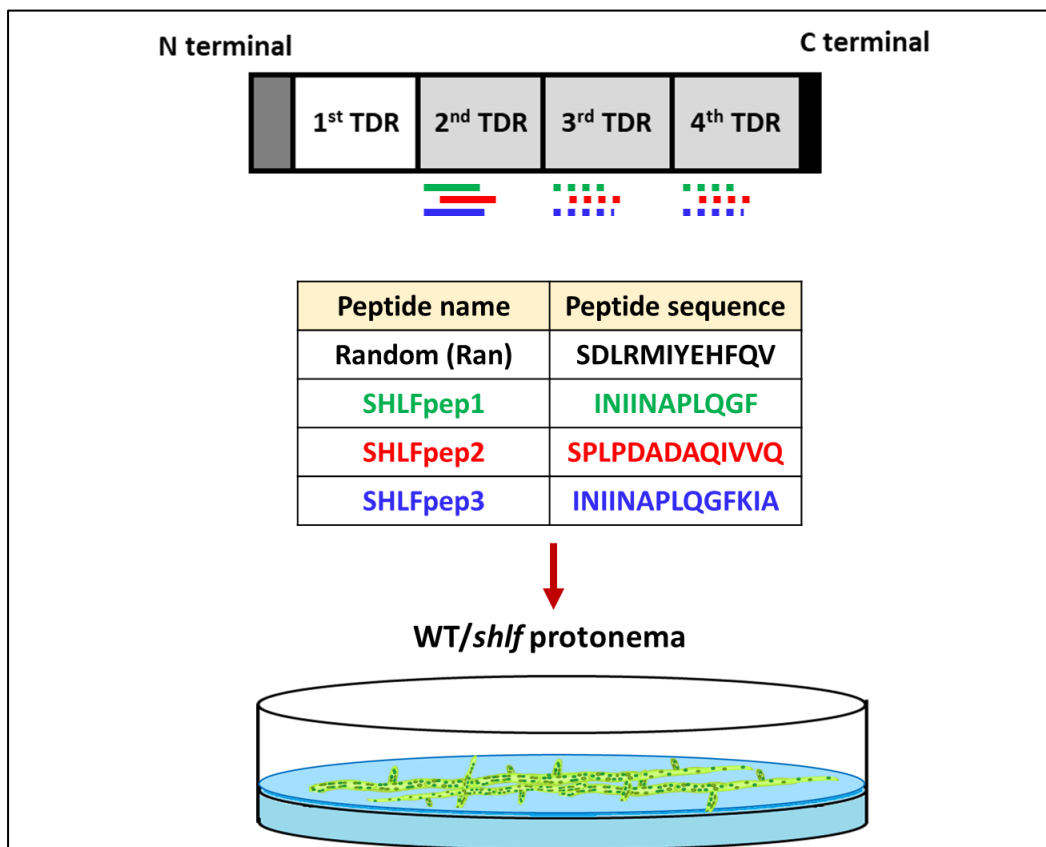


Figure 3.7 A schematic representing the sites of origin of SHLF peptides (SHLFpep1-3) used for peptide synthesis and supplementation to the WT and *shlf* protonema SHLF pep1-3 are all derived specifically from the highly conserved TDRs (TDR 2-4) of SHLF. A random peptide (Ran) was used as a negative control.

In order to identify the functional SHLF peptides, conserved TDR specific synthetic peptides; SHLFpep1 (INIINAPLQGF), SHLFpep2 (SPLPDADAQIVVQ), SHLFpep3 (INIINAPLQGFKIA), and a random peptide (Ran) (SDLRMIYEHFQV) were supplemented

to wild-type (WT) and mutant (*shlf*) protonema in increasing concentrations (1 μ M – 1mM) (Figure 3.7)

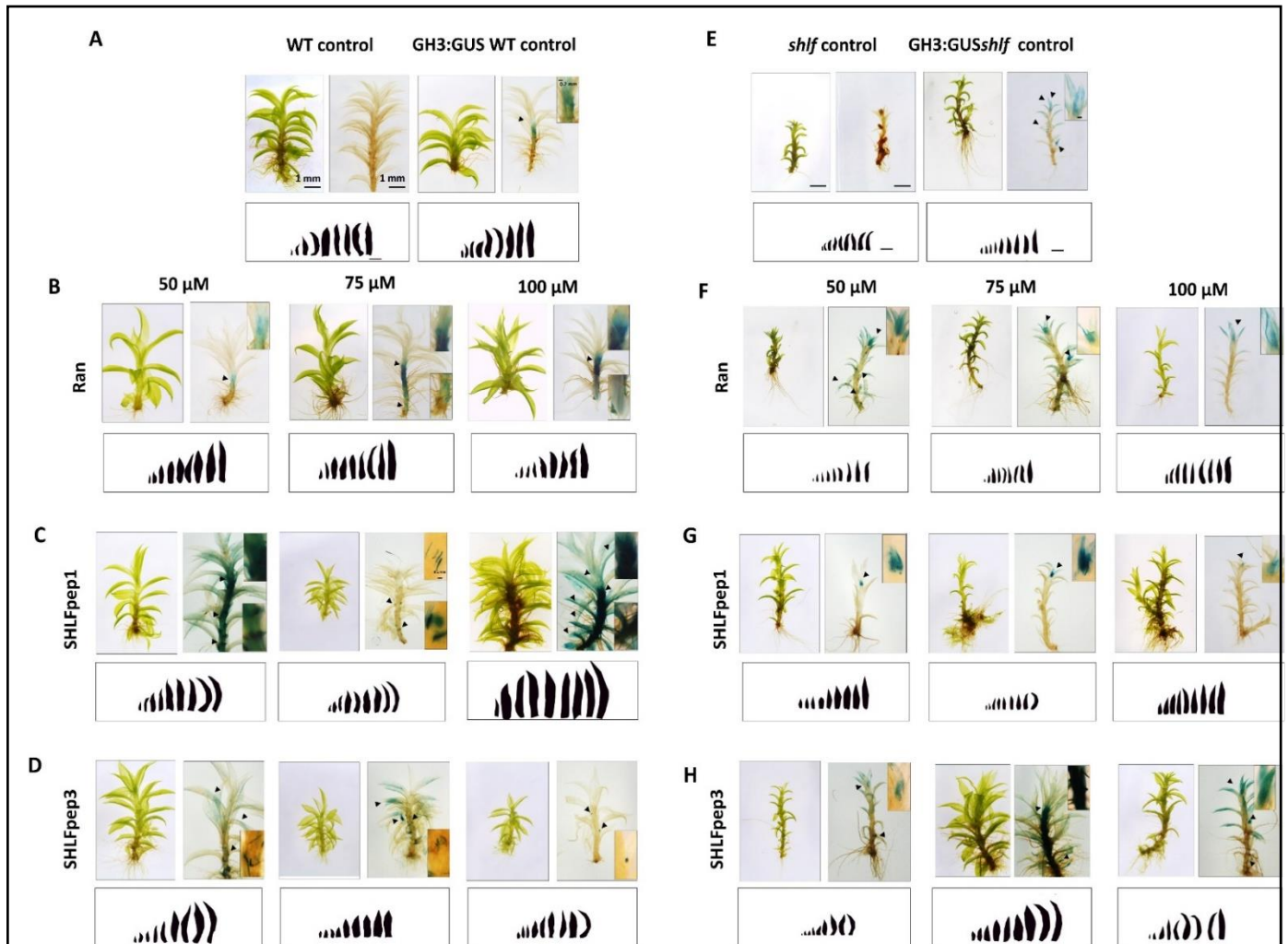


Figure 3.8 SHLFpep1 and SHLFpep3 supplementation shows dosage dependent effects on auxin distribution pattern and gametophore development in moss Random (Ran), SHLFpep1 and SHLFpep3 supplementations to both (B-D) GH3:GUS WT and (F-H) GH3:GUS *shlf* at 50, 75 and 100 μ M concentrations, showing plant phenotype and auxin distribution patterns (black arrow heads). Panels A and B represent the GH3: GUS WT and GH3:GUS *shlf* untreated controls respectively. Leaf arrays of 1st to 9th leaves of the supplemented lines are represented in the bottom panel.

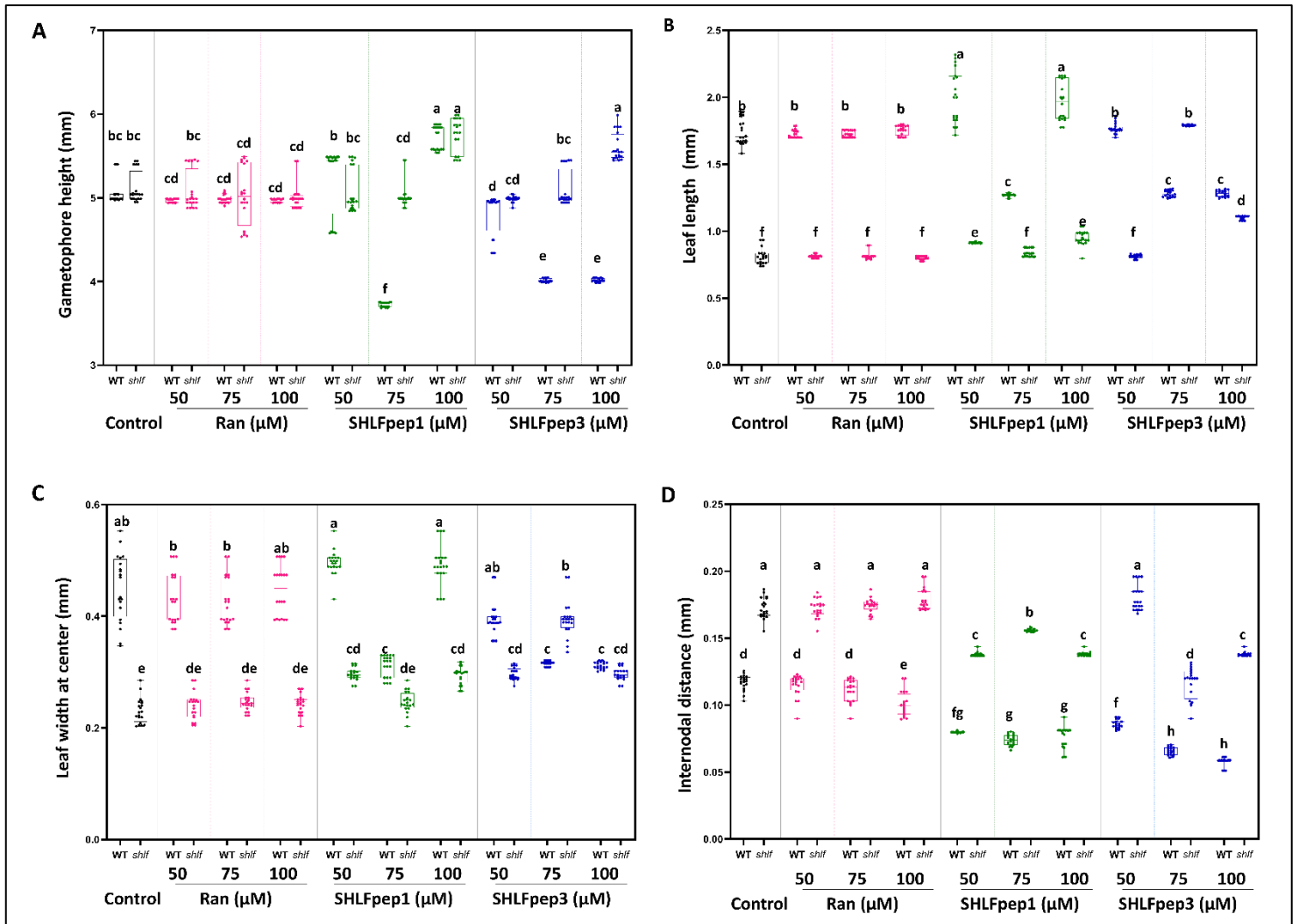


Figure 3.9 Phenotypic characterization of the SHLF peptide supplementation in WT and *shlf* (A-D) (A) Gametophore height, (B) Leaf length, (C) leaf width at the center (of the medio-lateral axis) and (D) the internodal distance were measured for the 9th leaves of (n = 20 gametophores each) for three independent bio-replicates of WT, *shlf* and all the synthetic peptide supplemented lines and the untreated control (WT and *shlf*). Scale = 1 mm for all lines, scale in insets is 0.2 mm. Shapiro-wilk normality test was performed on the datasets. Tukey's test with Dunn's post-hoc test and Bonferroni p-value corrections were performed to check for the statistically significant difference between the sample sets. The letters indicate the significant pairwise corrections in between the sample sets.

As SHLF exhibits altered auxin distribution pattern (with increased auxin accumulation at the apex, branch initials, young leaves and leaf tips), in contrast to the sub-apical auxin distribution pattern in WT gametophores (Mohanasundaram *et al.*, 2021), we wanted to investigate if the SHLFpeps supplementation has any influence on the auxin distribution pattern in gametophore. Synthetic SHLFpeps (1 μM – 1mM) were independently supplemented to WT and *shlf* expressing the auxin responsive GH3:GUS construct (Figure 3.8,3.9). Partial and complete phenotypic recovery was observed in *shlf* with SHLFpep1 (at 50 and 100 μM)

and SHLFpep3 (75 μ M only) supplementation, respectively, whereas SHLFpep2 and Ran peptides had no effect (Figure 3.8).

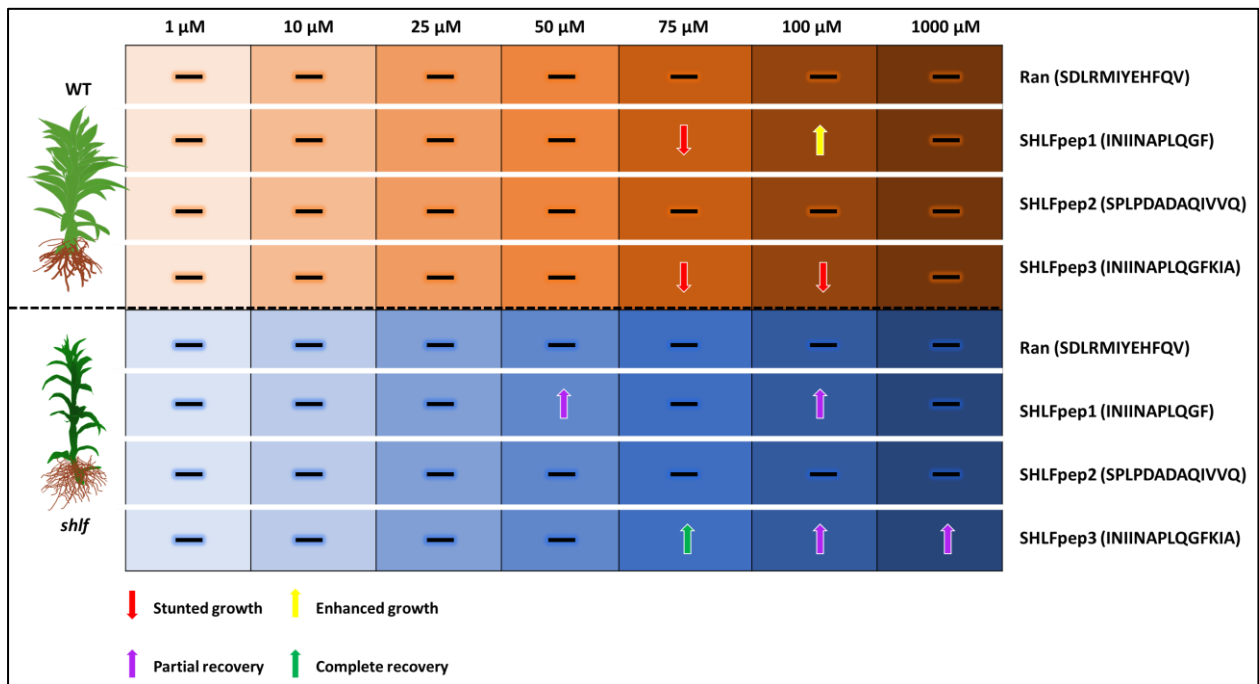


Figure 3.10 Schematic representation of the phenotypic effects of SHLFpep supplementation to WT and *shlf*. Representation of phenotypic effects on WT (Top panel in Brown) and *shlf* (Bottom panel in Blue) supplemented with increasing concentration (1-1000 μ M) of synthetic Random and SHLF peptides. Index panel at the bottom represents the effect on gametophore phenotypes.

In general, supplementation of SHLFpep1, -3 but not SHLFpep2, showed varied effects on gametophore phenotypes (gametophore height and internodal distance; leaf length and width) and on auxin distribution pattern. Figure 3.10 provides a comprehensive schematic representation of the different concentrations tested and their corresponding effects on both the wild-type (WT) and *shlf* genotypes. SHLFpep1 at 50 μ M resulted in partial recovery of *shlf* and a reduced auxin response at the apex but showed no effect on WT (Figure 3.8, 3.9). At 75 μ M however, SHLFpep1 had no influence on *shlf*, but in WT, exhibited stunted gametophores with reduced leaf dimensions and internodal distance. These plants exhibited a reduced auxin response in the stem and increased GUS activity in the axillary hairs. SHLFpep1 at 100 μ M, caused taller gametophores with increased leaf dimensions along with intense GUS activity throughout the stem and also in the leaf blades in WT, whereas, it showed a partial recovery in *shlf*. We noticed that SHLFpep3 at all concentration tested, resulted in an increased GUS activity in axillary hairs of WT (Figure 3.8). Interestingly, at 75 μ M, SHLFpep3 completely rescued all the mutant phenotypes and showed a WT like auxin distribution pattern evident

from the sub-apical GUS activities. We noticed that WT develops stunted gametophores at 75 and 100 μM of SHLFpep3. This peptide at 100 and 1000 μM caused partial phenotypic recoveries and reduced GUS staining at the apex in the mutant (Figure 3.8). In contrast to SHLFpep1 and pep3, the SHLFpep2 did not produce any noticeable effects on WT and *shlf*. The phenotypic changes observed upon peptide supplementation are represented in Figure 3.9 and 3.10. Our data indicates that the SHLFpep1 and SHLFpep3 act in a dosage dependent manner to regulate auxin response and gametophore development in moss gametophores. Additionally, our findings indicate that SHLFpep3 (INIINAPLQGFKIA), derived from the conserved TDRs, serves as the minimal functional unit of SHLF. Remarkably, this peptide alone is sufficient to regulate gametophore development and exert influence on the auxin distribution pattern in moss.

3.4. Discussion

3.4.1. SHLF undergoes cleavage *in planta* and functions as a secretory protein to regulate gametophore development.

The high abundance of SHLF peptides in the WT secretome (Table 7.4) points towards active trafficking and secretion of the protein outside the cell. It is plausible that SHLF behaves primarily as a secretory protein and that the WT and overexpression lines may only have a basal level of intercellular SHLF comparable to that of mutant. Secretory proteins containing tandem repeats have been reported from multiple plant and animal pathogens (Kaur *et al.*, 2012; Yu *et al.*, 2017). However, not much is known about such proteins in plants. The few reported secretory repeat containing proteins in plants include Ginkbilobin2 and Root Meander Curling protein from *O. sativa* (OsRMC) which confer biotic and abiotic stress tolerance in plants (Zhang *et al.*, 2009; Miyakawa *et al.*, 2014). Since SHLF trafficks to ER and undergoes cleavage *in planta*, we speculated that it may function as a secretory protein. In line with this, we detected a ~ 45 kDa band in the WT secretome, which was absent in the mutant secretome (Figure 3.2 and 3.3). This observation aligns with the findings reported by Lehtonen *et al.*, who identified the presence of Pp3c14_22870 (SHLF) in the moss secretome. The coexistence of a truncated form of SHLF (~ 44 kDa, Figure 3.2) in the secretome of the WT, alongside the presence of a pool of secretory SHLF peptides (Table 7.4), suggests that extracellular proteases might be involved in processing this secretory protein. However, due to the absence of conserved domains and the high sequence similarity within the SHLF TDRs, identifying the specific proteases responsible for this cleavage presents a challenge.

The presence of SHLF in the secretomes of reverted overexpression (OE) lines (miniSHLF) but not in the non-reverted OE lines, indicate that that secretion is essential for SHLF function (Table 7.4). Interestingly, the secretory form of SHLF detected in the WT secretome (Figure 3.2) exhibits a size remarkably similar to the previously identified miniSHLF (N-TDR1-2). This intriguing finding further emphasizes the critical role of the conserved TDRs. Furthermore, the presence of secretory SHLF peptides in the miniSHLF lines, but their absence in the N-TDR1-C line, indicates that while TDR1 may facilitate protein secretion, it alone is insufficient for proper protein function (Table 7.4). These observations align with the fact that SHLFpep3, the functional secretory peptide responsible for protein function, is derived from the conserved TDRs (Figure 3.8). Our consolidated results prove that secretion is essential for SHLF function. Furthermore, our study reveals that supplementation of secretome from WT or miniSHLF leads to complete recovery of the *shlf* phenotypes, indicating the crucial role of secretory SHLF in gametophore development (Figure 3.5 and 3.6). This finding strengthens our hypothesis that SHLF functions as a secretory protein. The results of our supplementation assays, combined with immunoblot analyses, provide strong evidence that secretory SHLF plays a vital role in the development of gametophores.

We speculate that being an apoplast resident protein, secretory SHLF could serve as a surveillance system to detect dynamic environmental cues. As mosses are known to inhabit diverse ecosystems, the ability to detect and respond to environmental changes is vital for their survival. The secretory SHLF protein could play a significant role in monitoring environmental changes and activating signaling pathways that lead to adaptive responses. Further studies could explore this hypothesis and shed light on the role of secretory SHLF in the survival of mosses in diverse habitats.

3.4.2. Secretory SHLF peptides affect auxin distribution pattern and gametophore development in a dosage-dependent manner

Secretory peptides in plants orchestrate multifaceted aspects of growth, development and stress response by regulating various signaling cascades and phytohormone pathways (Katsir *et al.*, 2011; Whitford *et al.*, 2012; De Coninck and De Smet, 2016; Wang, Zhang and Wu, 2016; Roy *et al.*, 2018; Chen *et al.*, 2020). In flowering plants, secretory peptides have been shown to regulate meristem maintenance (Hazak and Hardtke, 2016; Yamaguchi, Ishida and Sawa, 2016), stomatal development (Hara *et al.*, 2007) shoot development (Fletcher *et al.*, 1999) and reproductive processes (Bircheneder and Dresselhaus, 2016). Most of the secretory peptides reported till date in plants are generated from inactive precursor proteins with an N-

terminal signal peptide which directs the protein to the secretory pathway (Matsubayashi, 2014; Tavormina *et al.*, 2015). However, there are no reports in plants which describes that secretory peptides could be generated from protein repeats and involved in regulation of developmental programs.

SHLF forms a pool of secretory peptides derived from its TDRs (Table 7.4). Upon supplementation of the mutant with the extracellular peptidome of WT or miniSHLF, we observed a partial recovery of *shlf* (Figure 3.5, 3.6). This may be due to the sub-optimal availability of functional SHLF peptides in the peptidome. Independent supplementation with synthetic peptides (SHLFpep1, SHLFpep2, SHLFpep3) originating from conserved TDRs (TDR 2-4) to mutant revealed that phenotypic recovery is dependent on both specific peptide sequence as well as concentration (Figure 3.8, 3.9). Amongst all the peptides and concentrations tested, SHLFpep3 at 75 μ M was able to fully recover the mutant phenotypes (Figure 3.8). Additionally, SHLFpep1 and SHLFpep3 also affected gametophore development in a dosage dependent manner, whereas the SHLFpep2 had no effect, justifying the importance of sequence specificity and dosage necessary for peptide function (Figure 3.8-3.9). *shlf* exhibits an increased auxin accumulation at the sites of synthesis (shoot apex and leaf tips) in *shlf*GH3:GUS line, unlike the WT GH3:GUS lines, which showed a diffused auxin distribution pattern (Mohanasundaram *et al.*, 2021). Based on our findings, where SHLFpep3 at 75 μ M successfully restored both the mutant phenotype and the auxin distribution pattern (Figure 3.8-3.9), we hypothesized that SHLFpep3 plays a significant role in mediating auxin transport. Remarkably, our results reinforced our speculation. Upon supplementing SHLFpep3 to *shlf* plants, we observed a complete recovery in the auxin distribution pattern of the mutant. The supplemented plants exhibited a diffused auxin distribution pattern along the stem, closely resembling the pattern observed in WT plants. This outcome further strengthens the notion that SHLFpep3 plays a pivotal role in influencing auxin distribution in moss. Interestingly, amongst the ~40 secretory SHLF peptides, only SHLFpep3 was recently reported to play a crucial role in plant defense against microbes (Fesenko *et al.*, 2019). It is intriguing to observe that numerous secretory peptides in plants serve dual roles in both defense and development (Gancheva *et al.*, 2019). We hypothesize that SHLFpep3 may exhibit analogous behavior, suggesting its potential involvement in both defense mechanisms and developmental processes

In summary, we demonstrated the secretory nature of SHLF and the role of its domains on protein secretion. Our study employed immunoblotting analyses using anti-SHLF to examine the WT secretome. We detected a truncated secretory form of SHLF (~45 kDa, absent in the mutant), while the total protein exhibited a size of ~75 kDa. In-gel digestion and MS/MS

analysis of the ~45 kDa band resulted tryptic peptides originating from the divergent 1st TDR and the conserved TDRs. Furthermore, MS/MS analysis of secretomes revealed SHLF as a major component in the WT secretome. Complete phenotypic recovery of the mutant was achieved upon supplementation of WT and miniSHLF secretomes, while only partial recovery was observed upon peptidome supplementation, indicating the dosage dependent nature of SHLF. Notably, MS/MS analysis detected ~40 SHLF peptides in WT and miniSHLF secretomes, while no peptides were found in the mutant or in the non-reverted domain-deleted overexpression lines. Subsequent synthetic peptide supplementation assays identified SHLFpep1 (INIINAPLQGF) and SHLFpep3 (INIINAPLQGFKIA), originating from the conserved TDRs, as dosage-dependent regulators of moss gametophore development and auxin distribution. In addition, our investigation led to the identification of SHLFpep3 as the minimal functional unit of SHLF, demonstrating its sufficient capacity to regulate moss development and protein function. Overall, our findings highlight SHLF's role as a secretory protein, its presence in the secretome, and the functional significance of SHLFpeps in moss gametophore development and auxin distribution pattern.

CHAPTER 4:

A transcriptomic and metabolomic approach to elucidate the mode of SHLF function in regulating moss development

4. A transcriptomic and metabolomic approach to elucidate the mode of SHLF function in regulating moss development

4.1. Introduction

SHORT-LEAF is a bryophyte specific protein which forms peptides and regulates auxin response and development in moss gametophores. A Tnt1 insertion in TDR4 of SHLF is the cause for the *shlf* mutant having shorter leaves, impaired apical dominance, reduced PD frequency and impaired auxin distribution (Mohanasundaram *et al.*, 2021). However, due to the lack of conserved domains in SHLF, unraveling the precise molecular role of SHLF in regulating auxin distribution of moss gametophores poses a significant challenge. Consequently, it becomes imperative to thoroughly explore the molecular landscape of the mutant in order to acquire a comprehensive understanding of SHLF's function in moss.

In this study, we adopted a multi-omics approach to identify the underlying genetic pathways affected by SHLF and also attempted to identify its interacting partners, to further establish the molecular role of this clade specific protein.

4.1.1. Auxin transport in bryophytes

Phytohormones are known to be functionally conserved across the plant lineage. Auxin is one of the major players regulating almost all aspects of plant growth. The establishment of a proper auxin gradient is essential for plant development and is mediated by the dynamic interplay between external cues and cellular response. Auxin response in a plant mainly comprises of signalling by the most common naturally occurring auxin; Indole acetic acid (IAA), auxin metabolism (biosynthesis and degradation) and transport. Auxin transport in plants can be either via polar auxin transport (PAT) or by plasmodesmata (PD) mediated symplastic diffusion. Directional auxin transport is essential for plant development and is mostly dictated by the auxin transporters. Influx carriers such as the AUX1/LAX family of PM permeases transport auxin into cells via their H⁺ symport activity. Efflux carriers such as the primarily polar localized PINs and the majorly apolar ATP-binding cassette (ABC) superfamily of transporters, mainly the B type (ABCB/multidrug resistance [MDR]/phosphoglycoprotein [PGP]) (Michniewicz, Brewer and Friml, 2007). PIN proteins can be further subdivided into

the Plasma membrane (PM) localized long PINs and the short PINs which are localized in cellular organelles (Adamowski and Friml, 2015).

Auxin plays crucial roles in development of the moss *P. patens*, by regulating both the filamentous protonema (Johri and DESAI, 1973; Ashton, Grimsley and Cove, 1979; Sood and Hackenberg, 1979; Thelander, Landberg and Sundberg, 2018) and the leafy shoot-like gametophores (Ashton, Grimsley and Cove, 1979; Nyman and Cutter, 1981; Sakakibara *et al.*, 2003; Harrison *et al.*, 2009; Bennett *et al.*, 2014; Kofuji and Hasebe, 2014; Lavy *et al.*, 2016). The key players regulating auxin biosynthesis, metabolism, transport, and signaling, are evolutionarily conserved and affect moss development (Thelander, Landberg and Sundberg, 2018). Though auxin efflux carriers like the PINs are present in moss, PAT does not occur in moss gametophore stems and is present only in the sporophytes (Viaene *et al.*, 2013, 2014; Bennett *et al.*, 2014). Instead, nonpolar symplastic auxin diffusion occurs in the gametophore stem, which is majorly regulated by the levels of callose at the PD (Poli, Jacobs and Cooke, 2003; Fujita *et al.*, 2008; Han *et al.*, 2014; Coudert *et al.*, 2015). Further, both PD and symplastic auxin transport appear to have parallelly evolved in the green plant lineage, highlighting the significance of this passive mode of auxin transport (Liu *et al.*, 2017). Additionally, though PAT is absent in gametophore stem, it appears to still be involved in leaf and stem development in moss (Bennett *et al.*, 2014; Viaene *et al.*, 2014; Thelander, Landberg and Sundberg, 2018).

4.1.2. Interplay of flavonoids and ROS in auxin transport

In plants, the regulation of auxin transport is of utmost importance in ensuring the appropriate development and growth (Muday and Murphy, 2002; Armengot, Marquès-Bueno and Jaillais, 2016). Phenylpropanoids such as cinnamic acid act as natural inhibitors of the PIN family of proteins and affect cellular auxin influx (Steenackers *et al.*, 2017). The widely occurring ABC transporters have high substrate specificity and may work in parallel with PINs to mediate auxin efflux. Certain members of the family have also been reported to be involved in governing auxin import (AtPGP4) (Geisler *et al.*, 2005). The ATPase activity of the ABC transporter family is inhibited by flavonoids, hence making them more susceptible to flavonoid levels than PIN proteins (Peer and Murphy, 2007).

Naturally occurring phenolic compounds including flavonoids such as quercetin and kaempferol, have been long reported to inhibit auxin efflux carriers, thereby blocking polar auxin transport (STENLID, 1976; MARIGO and BOUDET, 1977; Jacobs and Rubery, 1988;

Fischer *et al.*, 1997; Murphy, Peer and Taiz, 2000). Recent studies using the *transparent testa 4 (tt4)* Arabidopsis mutant, defective in the Chalcone Synthase (CHS) have further established that flavonoids act as negative regulators of auxin transport which causes pleiotropic growth defects observed in the mutant including altered apical dominance, root growth and gravitropism (Brown *et al.*, 2001). In general, flavonoids serve as natural antioxidants synthesized in response to altered redox states, to scavenge the ROS (Pietta, 2000; Brunetti *et al.*, 2013). The feedback loop between flavonoids, ROS and auxin is quite intriguing. Flavonoid accumulation has been reported in the absence of auxin efflux carriers or upon their inhibition (Peer *et al.*, 2004). Further, some studies report that flavonoid synthesis can occur in response to auxin accumulation to regulate the ROS levels in the plant (Peer and Murphy, 2006). Additionally, flavonoids such as kaempferol regulate the callose deposition at PD, thereby affecting PD permeability but not much is known about the effect of this increased permeability on symplastic auxin transport. Recently, Peer and Murphy proposed that questions regarding the specificity of flavonoid regulation on auxin and ROS signalling may be studied using the genetically simple bryophytes as model systems, to decipher the complexities of flavonoid biosynthesis and function (Peer and Murphy, 2007).

The genome of *P. patens* harbors homologs of genes responsible for phenylpropanoid biosynthesis, which have been identified in flowering plants (Naoumkina *et al.*, 2010). Similar to vascular plants, flavonoids regulate ROS levels in moss to combat abiotic stress (H. Wang *et al.*, 2020; Wang *et al.*, 2022). These phenylpropanoids also affect moss shoot development in a PIN protein dependent manner (Bennett *et al.*, 2014). However, the molecular players involved in regulating the auxin-flavonoid cross-talk have not been identified in bryophytes.

4.2. Materials and methods

4.2.1. RNA Seq analysis

Three-week-old moss colonies (WT and *shlf*) were grown under at 16:8 light and dark regime at 24°C and harvested in three bio-replicates for total RNA isolation using RNAiso Plus (DSS Takara). Total RNA (3 µg) was used for sequencing libraries using NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs), following which index codes were assigned to each sample. The library quality was assessed using a Bioanalyzer 2100 system (Agilent Technologies). The clustering of the index-coded samples was performed on a cBot Cluster Generation System using a TruSeq PE Cluster Kit v3-cBot-HS (Illumina). After cluster generation, the library preparations were sequenced on an Illumina HiSeq platform and 150-bp

paired-end reads were generated. FastQC analysis was performed on the raw reads to estimate the quality of the data. PCA analysis was carried out in R (Version 4.2.3) to assess the closeness of each replicate. The raw reads were indexed and mapped onto the *Physcomitrium patens* genome (Ref. v 3.3) using the mapping mode of Salmon (Patro *et al.*, 2017). Differentially expressed genes were identified using the DESeq2 package in R. GO analysis was performed using the BinGO plugin of Cytoscape (Version 3.7.1). Out of the differentially regulated genes, top 20 up- and down- regulated genes were used for the generation of the heatmap and word clouds using R (version 4.2.3).

4.2.2. RT-qPCR analysis

Three-week-old moss gametophores were subjected to total RNA extraction using RNAiso Plus (Takara Bio USA Inc.). Two micrograms of RNA samples were reverse-transcribed using SS-IV reverse transcriptase (Invitrogen, Thermo Fisher Scientific - US) and oligo dT primers. SYBR Premix Ex Taq II (Tli RNaseH Plus) was used for the relative quantification of transcripts via Bio-Rad CFX96 Touch Real-Time PCR Detection System (Bio-Rad). Moss β -Actin and eIF2 were used as a reference gene for the quantification of target gene expression levels in all samples. The fold-change values (sample value/reference value) and $2^{-\Delta Ct}$ were calculated (Schmittgen and Livak, 2008). All the primers used for RT-qPCR analysis are listed in Table 5.

4.2.3. Toluidine blue staining

Three-week-old (WT and *shlf*) plants were weighed and their fresh weight was recorded. The gametophores were then stained with 0.05% (w/v) Toluidine Blue (Sigma-Aldrich) prepared in 50mM citrate buffer (pH 3.5) containing 50% ethanol. Staining, imaging and stain quantification was carried out as per established protocols (Bressendorff *et al.*, 2016) with minor modifications. To quantify the amount of stain retained by the plants, data from three independent bio-replicates was obtained and plotted as barplots representing the OD₆₀₀/g fresh weight. Error bars were assigned for standard deviation between the sample sets.

4.2.4. DAB staining

DAB (3,3'-Diaminobenzidine) staining solution was freshly prepared (Daudi and O'Brien, 2012), and three-week-old WT and *shlf* gametophores were fully immersed in the solution and stained for 2 h at RT with gentle shaking. Gametophores were then bleached at

90°C with 200 µl of Methanol:Acetic acid:Glycerol (3:1:1) solution. The samples were cooled at RT for an hour followed by imaging under Leica Stereomicroscope.

4.2.5. ROS quantification using carboxy H₂DCFDA assay

Forty milligrams of tissues from three-week-old WT and *shlf* gametophores were weighed, crushed in liquid nitrogen and resuspended in 200 µl of 10mM Tris-Cl (pH 7.3). This suspension was centrifuged at 12,000 rpm for 10 min, and the volume of the supernatant was made up to 300µl. The protocol of Juarez *et al.*, was followed for ROS quantification, with minor modifications (Juárez, Mangano and Estevez, 2015). Briefly, 10 µl of plant extract was added to 10 µl of CFDA (100 mM) along with 90 µl of 10mM Tris-Cl (pH 7.3) for quantification and fluorescence intensity was measured at 494 nm excitation and 517 nm emission spectra, with a bandwidth of 12 nm. Same amount of plant extract was used for estimation of total protein concentration using Bradford Assay. Plant extract and buffer were used as a buffer control, while buffer and dye together were used as dye control.

4.2.6. Total Chlorophyll estimation

WT and *shlf* plants (three-week-old) were weighed to generate 100 mg of fresh tissue. The plants were crushed in liquid nitrogen and total chlorophyll was extracted in 80% (v/v) acetone. Chlorophyll absorbance was measured at 645 and 663 nm and total chlorophyll content was calculated in mg/g (Bhattacharjee and Sharma, 2012).

Total chlorophyll content (mg/g) = $[20.2 (A_{645}) + 8.02 (A_{663})] V / (1000 \times W)$, Where, A = Absorbance at 645 and 663 nm respectively; V = Final volume of total chlorophyll in 80% acetone; W = Fresh weight of the WT and *shlf* gametophores.

4.2.7. UHPLC-MS/MS analysis

Three-week-old WT and *shlf* gametophores (200 mg) were crushed in liquid nitrogen and resuspended in 1 ml of 70% Methanol containing 400 ng/ml of Formononetin. The total metabolome of the plants was extracted by two rounds of vortexing followed by high speed centrifugation for 10 min. The supernatant was stored at -80°C overnight and re-centrifuged on the following day to reduce the lipid contents in the samples. The final extract was then injected (20 µL) into X500R U(H)PLC coupled with ESI-QTOF mass spectrometer (AB SCIEX Pvt. Ltd.) A Phenomenex Gemini® C18 column (50 mm x 4.6 mm, 5 µm, 110 Å) was used for metabolite separation.

Mobile phase used was a mixture of solvent A (MilliQ water) and solvent B (100% (v/v) methanol, containing 0.1% (v/v) formic acid), with flow rate of 0.5 mL/min and 30°C column oven temperature. A gradient was set using solvent B with 5% (0 min), 10% (3 min), 50% (6 min), 70% (7 min), 95% (10 min), 95% (12 min), 5% (12.5 min), 5% (16 min). MS1 scan was performed between 100 to 1000 Da in negative ionization mode, with 4500 V spray voltage, 400°C curtain gas temperature and 10 V collision energy. Fragmentation was performed using 35 V collision energy (with a CE spread of 10 V). Fragment masses between 40 and 1000 Da were scanned using TOF mass analyzer. All data were processed using MS-DIAL 4.90 (<http://prime.psc.riken.jp>) and the concentrations were calculated using internal standards (Formononetin). Raw data files (.wiff2) were directly imported as profile data. File deconvolution parameters were set as: MS1 tolerance of 0.01 Da; MS2 tolerance of 0.025 Da; peak detection of 1000 amplitudes (minimum peak height); mass slice width of 0.1 Da; respectively. Metabolites were identified from all public MS/MS data available on MSP Spectral Database (<http://prime.psc.riken.jp/compms/msdial/main.html#MSP>) as [M-H]⁻ adducts. Identification parameters were set as; MS1 tolerance of 0.01 Da; MS2 tolerance of 0.05 Da; and 80% identification score cut-off. Alignment parameters were set as; retention time tolerance of 0.05 min and a MS1 tolerance of 0.01 Da. A minimum of two reference MS2 fragment matches were set as cut-off for confirmation of identity of suggested metabolite.

4.2.8. Analyzing plasmodesmata (PD) permeability of the mutant using photoactivable Dendra2

Stable transgenic lines of WT and *shlf* overexpressing Dendra2 were grown in 35mm Petri plates with glass windows (22 mm coverslips) for a week under red light ,with and without 10 μM Di-deoxy Glucose (DDG) (Kitagawa and Fujita, 2013) and imaged using the FRAP mode of Leica Sp8. The desired protonemal filament was located, the cell of interest was selected and excited using the 405 laser (2.5% laser power). Images were captured using the 561 laser (15% power) at the emission range of 575–615 nm. Images were obtained at an interval of 5 min for a period of 30 min and processed using ImageJ. Graphs were plotted in excel to detect the rate of diffusion of the photoactivated Dendra2 as per established protocols (Kitagawa and Fujita, 2013).

4.2.9. Co-immunoprecipitation analyses

One month old WT and *shlf* (1 g) gametophores was crushed in liquid nitrogen, 3 mL of lysis buffer was added and incubated overnight on a disc rotator at 8 RPM at 4°C. The

supernatant obtained upon centrifugation of the sample at 14,000 RPM for 20 min at 4°C, was collected. This served as the whole tissue lysate.

800 µL of 0.1 M sodium phosphate buffer (pH 8) was added to 50 µL of Protein A Dynabeads (Invitrogen) and incubated for 15 min at room temperature on a disc rotator. 7 µL of anti-SHLF and 4 µL of anti-IgG was added to 700 µL of the washed bead solution, followed by washing with 800 µL of 0.1 M sodium phosphate buffer. The beads were then incubated with 0.2 M triethanolamine (TEA) at 12 RPM for 30 min with a crosslinking solution (1 mL TEA and 0.0054 g DMP) for 1 hour at 12 RPM. Then 1 mL of 50 mM Tris-Cl (pH 7.5) was added to this solution to stop crosslinking which was incubated for 30 min at 20 RPM. The antibody-bound beads were then washed three times with 1 mL of 0.1 M sodium phosphate buffer for 10 min at 25 RPM and incubated with 1 mL of whole tissue lysate for 13 hours at 4°C on a disc rotator at 8 RPM. The final step involved washing the antibody-antigen bound beads with 800 µL of 0.1 M sodium phosphate buffer for 10 min at 25 RPM. This step was repeated six times and the interacting protein partners were eluted using 0.2 M acidic glycine (pH 3). This sample was then probed using anti-SHLF for detection of the interacting partners.

4.2.10. Mass spectrometric analyses of the interacting partners

The interactors of SHLF were identified using in gel digestion and subsequent MS/MS analyses as mentioned in Chapter 2, Section 2.2.9.

4.3. Results

4.3.1. Several stress responsive genes are differentially regulated in *shlf*

To elucidate the effects of *SHLF* loss of function in the mutant, and to gain insights into the molecular role of SHLF, a comparative transcriptomic analysis was performed in *shlf* and WT backgrounds, as per the workflow detailed in Figure 4.1.

Our analyses revealed a set of 19,235 genes commonly present between WT and *shlf* with a set of 2154 and 2044 genes were differentially expressed with a p value of <0.01 (Figure 4.2 A). The heatmap (Figure 4.5) represents the top (~40) differentially regulated genes in the WT and *shlf* transcriptomes, in which several stress responsive genes were detected. Principal Component Analysis (PCA) of WT and *shlf* samples using the first 2 principal components, reflected the closeness of the bio-replicates in each of the sample sets (Figure 4.2 B). Further, gene ontology analyses revealed that majority of the DEGs belonged to stress response and

organism developmental processes (Figure 4.3 C). The list of all top most differentially expressed genes (DEGs) are provided in Tables 7.5 and 7.6.

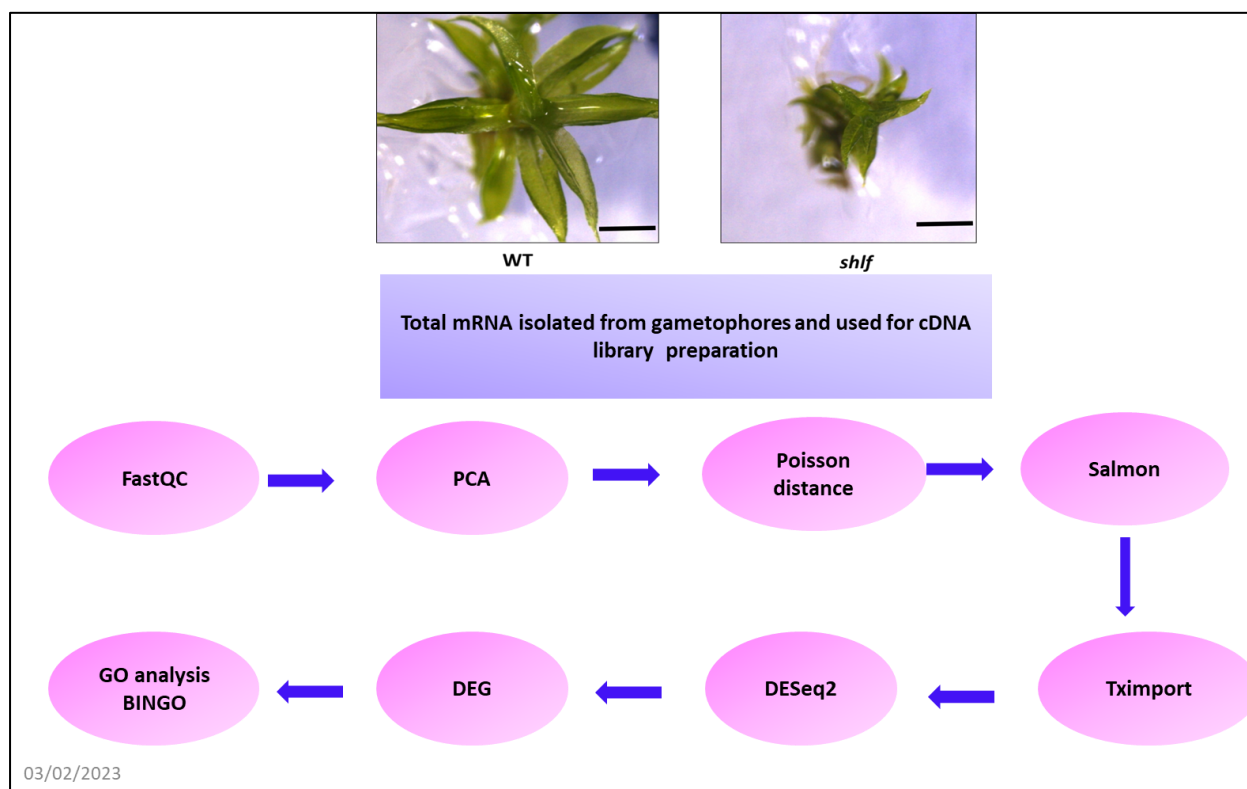


Figure 4.1 Workflow for the RNASeq analyses of WT and *shlf* Total mRNA isolated from three independent bio-replicates of WT and *shlf* were subjected to cDNA library preparation using the Illumina HiSeq platform. The adapter-ligated reads generated were screened for quality control using FastQC, PCA and poisson distribution. This was followed by further analyses of the DEGs using SALMON, DESeq2 package of R and GO analyses using BinGo.

As expected, *SHLF* was one of the most down regulated genes in the mutant. Gene Ontology analyses revealed relative abundance of transcripts related to several cellular, molecular and biological processes in the mutant. Majority of the DEGs appear to be involved in intracellular processes, catalytic activity and metabolic processes (Figure 4.3). Further, word-clouds (Figure 4.4) provide a visual representation of the DEGs in *shlf*, whose heatmap cluster with representative Z -values are represented in Figure 4.5

Among the top up-regulated DEGs (Table 7.5) were several homologs of the EARLY LIGHT INDUCED PROTEIN 1 (Pp3c11_7280, Pp3c2_31380, Pp3c24_9670), which are known to be up-regulated during high light stresses; A Rubredoxin-like domain containing protein (Pp3c2_6110), known to regulate ROS levels under high light stress.

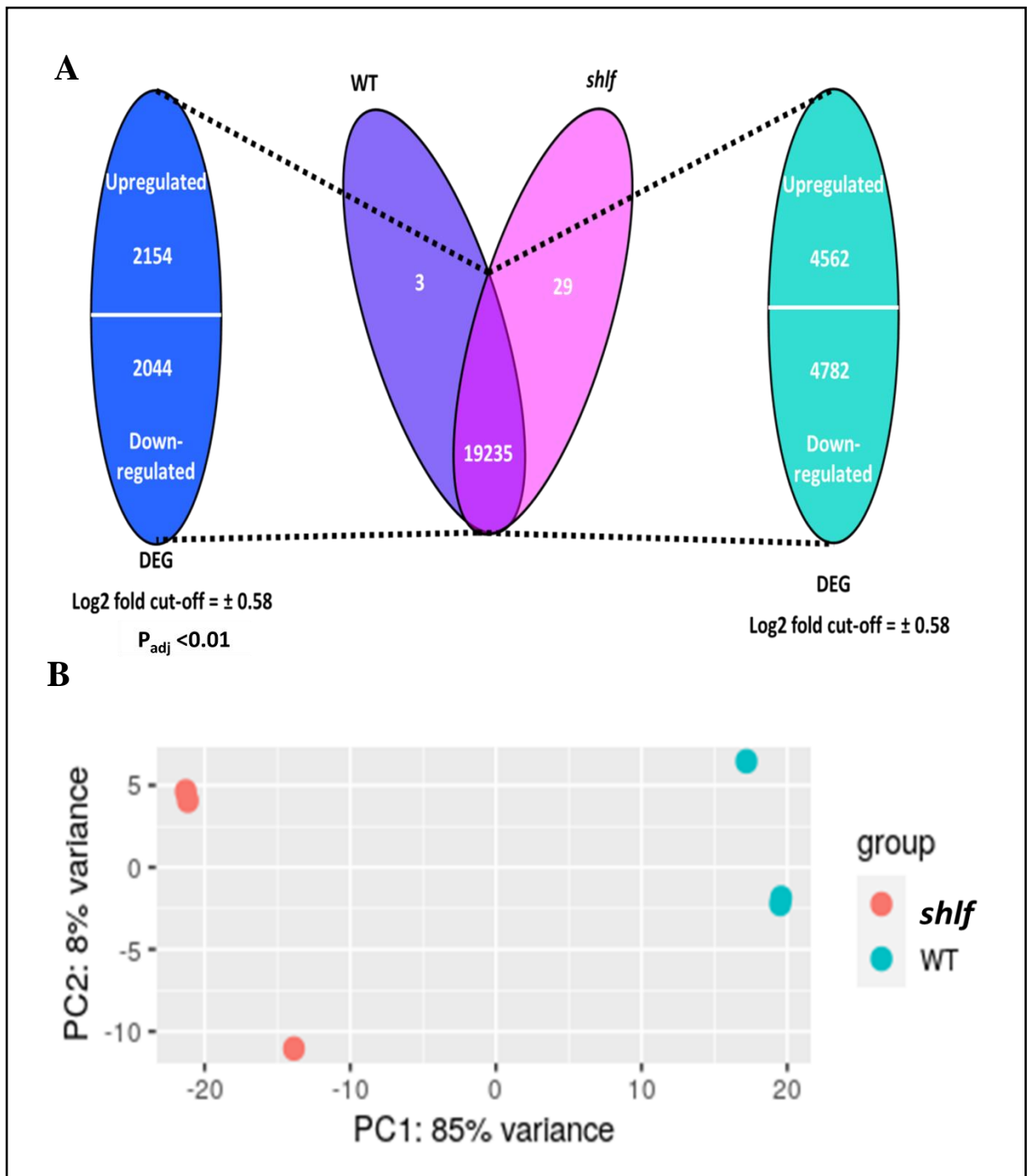


Figure 4.2 Comparative analyses of WT and *shlf* transcriptomes (A) Venn diagram representing differentially expressed genes in WT and *shlf* (B) Principal Component Analysis (PCA) plot of three independent bio-replicates of WT and *shlf*, built using 2 first principal components.

Among the other up-regulated genes were a heavy metal transport/detoxification domain-containing protein, (Pp3c6_5150), known to get up-regulated during drought stress; an AP2 domain containing transcription factor (Pp3c17_13620), differentially regulated upon drought stress; a Vaccinia Virus protein VP39 (Pp3c3_14410), involved in drought tolerance response in plants; late embryogenesis abundant (LEA) domain-containing proteins (Pp3c4_18020,

Pp3c16_20740) known to confer tolerance to abiotic stresses such as drought, salinity, and high/low temperatures

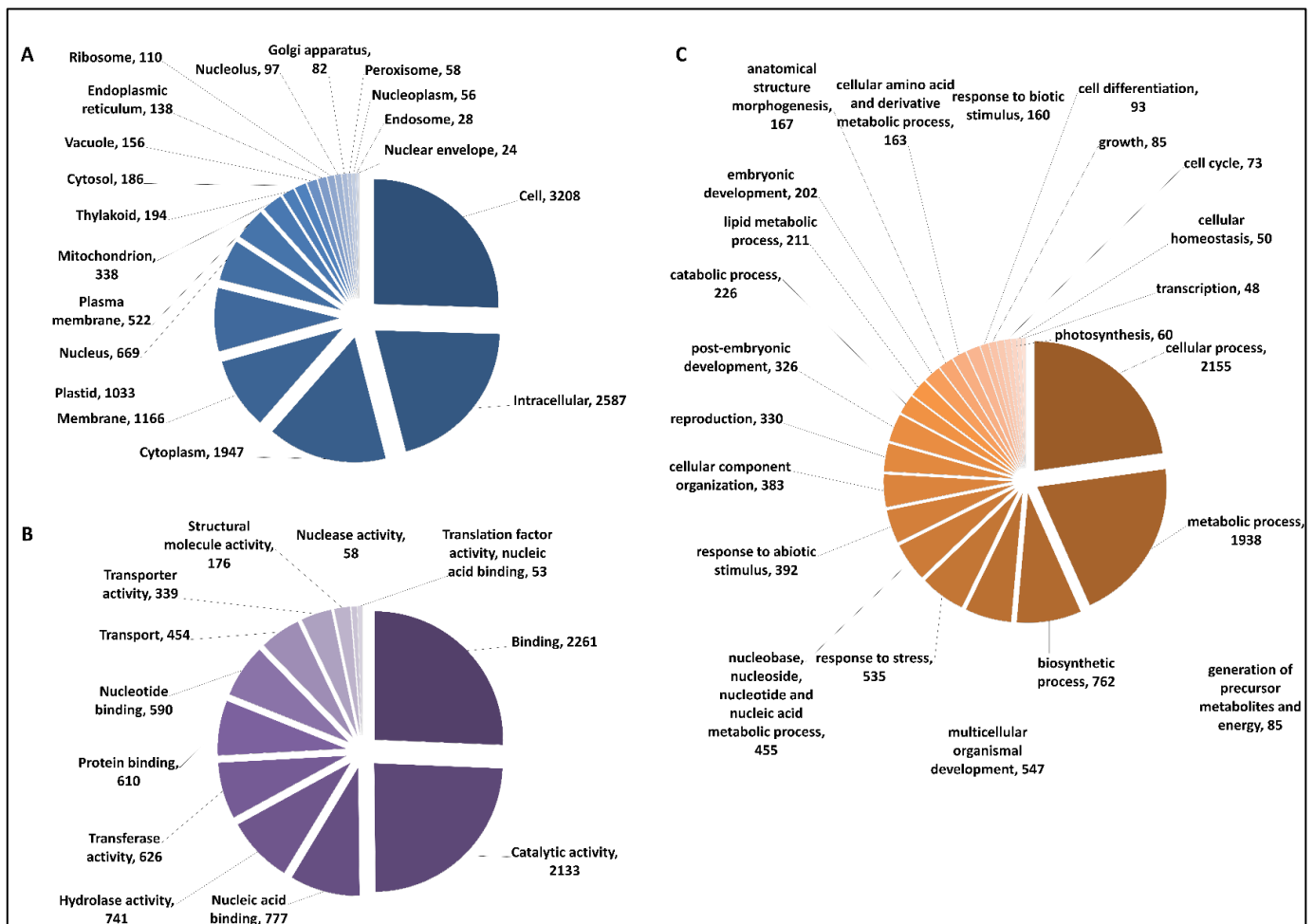


Figure 4.3 Comparative analyses of WT and *shlf* transcriptomes Genes were majorly divided into Cellular Component (A), Molecular Function (B) and Biological Process (C) categories by BINGO analysis in Cytoscape. The number attributed to the components in each category represents the relative abundance of each component in each category.

an SMI1/KNR4-like protein (Pp3c17_17450) responsible for regulation of cell wall composition during pant immune response; a chloride intracellular channel, isoform protein (Pp3c15_21480), a critical antiporter for nitrate transport into the vacuoles during resistance to biotic and abiotic stress, a chitinase-related protein (Pp3c4_3600), which confers resistance to fungal pathogens, a Chitin recognition protein (Chitin_bind_1) (Pp3c11_1420), which plays important roles in plant-fungal interactions and a flotillin-related protein (Pp3c3_21910), involved in plant defense response against infection from symbiotic bacteria.

Up-regulated genes

SMI1/KNR4-like
Uncharacterized bryophyta specific protein
Uncharacterized protein
Early light-induced protein 1
Alcohol dehydrogenase-related
Uncharacterized bryophyta-specific protein
Rubredoxin-like protein
Heavy metal transport domain protein
Uncharacterized moss-specific protein
Uncharacterized orphan gene
F-BOX Associated Ubiquitination effector family-related
Uncharacterized bryophyte specific protein
F-BOX protein
Glutathione dehydrogenase
AP2/ERF domain containing protein
AP2 domain (AP2) protein
Uncharacterized bryophyte-specific protein
Cytoplasmic protein tyrosine kinase
Flotillin-related

Down-regulated genes

AP2/ERF transcription factor
Pathogenesis-related protein
NADH-ubiquinone reductase complex 1 MLRQ subunit
Uncharacterized orphan protein
Revertase
Hypoxia-inducible factor 1 related
Uncharacterized protein
Chlorophyll A/B binding protein
Oxoglutarate/iron-dependent dioxygenase
DnaJ domain protein
BURP domain protein
Copper transporters
AP2-like transcription factor
Superoxide dismutase
Uncharacterised orphan gene
Very-long-chain beta-ketoacyl-CoA synthase
KDA class 1 heat shock protein
Short-leaf
CUPIN domain protein

Figure 4.4 Word cloud (generated using python) representing top 20 up- and down regulated genes in *shlf* where font size corresponds to Log_2 changes in transcript abundance ($\text{log}_2 \text{foldchange}_{\text{mutant/wild type}}$)

Among the top down-regulated DEGs (Table 7.6) were a superoxide dismutase (Pp3c17_14510), which provides protection against ROS; a BURP domain containing drought responsive protein (Pp3c13_16030), responsible for cell expansion and cell wall modifications; an ethylene responsive SHN SHINE transcription factor, (Pp3c13_3830), involved in drought responsive regulation of cutin biosynthesis; abiotic stress responsive proteins like a NADH-ubiquinone reductase complex 1 MLRQ subunit (B12D) (Pp3c3_14700) and PROKAR_LIPOPROTEIN (Pp3c5_3169); a 17.6 kDa class heat shock protein (Pp3c6_3710), which is a hypoxia responsive protein and regulated ROS levels in plants; a DnaJ domain (DnaJ) protein (Pp3c27_1590) which regulates both plant growth and stress response and a chlorophyll a/b binding protein (Pp3c5_22920), which effects ROS homeostasis and ABA response in plants.

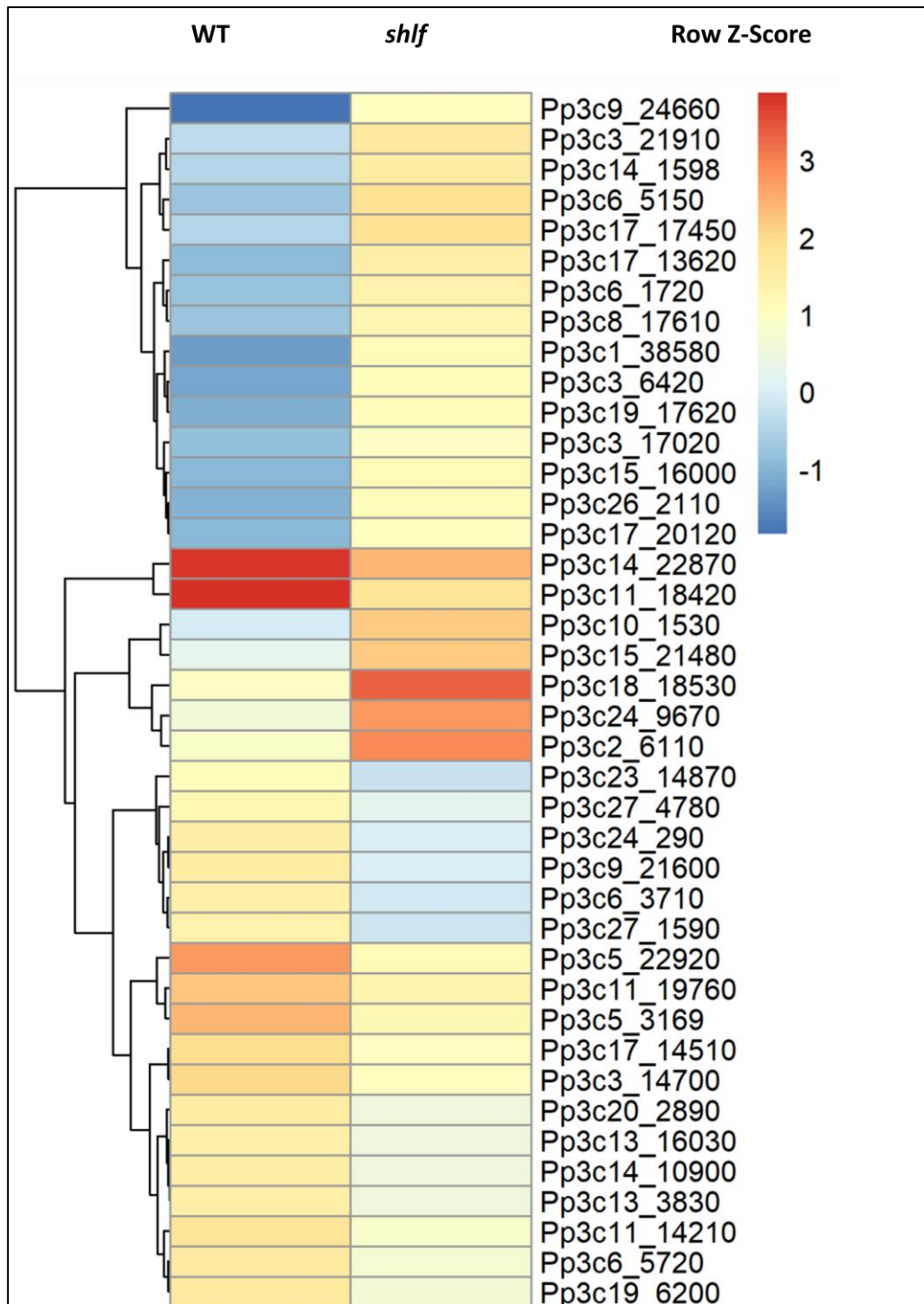


Figure 4.5 Heatmap showing hierarchical clustering analysis of top 20 up and downregulated genes with the row z-score denoting the similarity in expression patterns between the sample sets.

. Further, the RNA-Seq dataset of *shlf* also revealed the up-regulation of several common markers of stress including *PNENYLALANINE AMMONIA LYASE (PAL)*, *ALLENE OXIDE SYNTHASE (AOS)*, *CHALCONE SYNTHASE (CHS)*, α -*DIOXYGENASE (DOX)*, *PATHOGENESIS RELATED PROTEIN 10* and both abiotic stress responsive *DEHYDRIN A*, *DEHYDRIN C* and the biotic stress responsive *DEHYDRIN B* genes respectively (Figure 4.6).

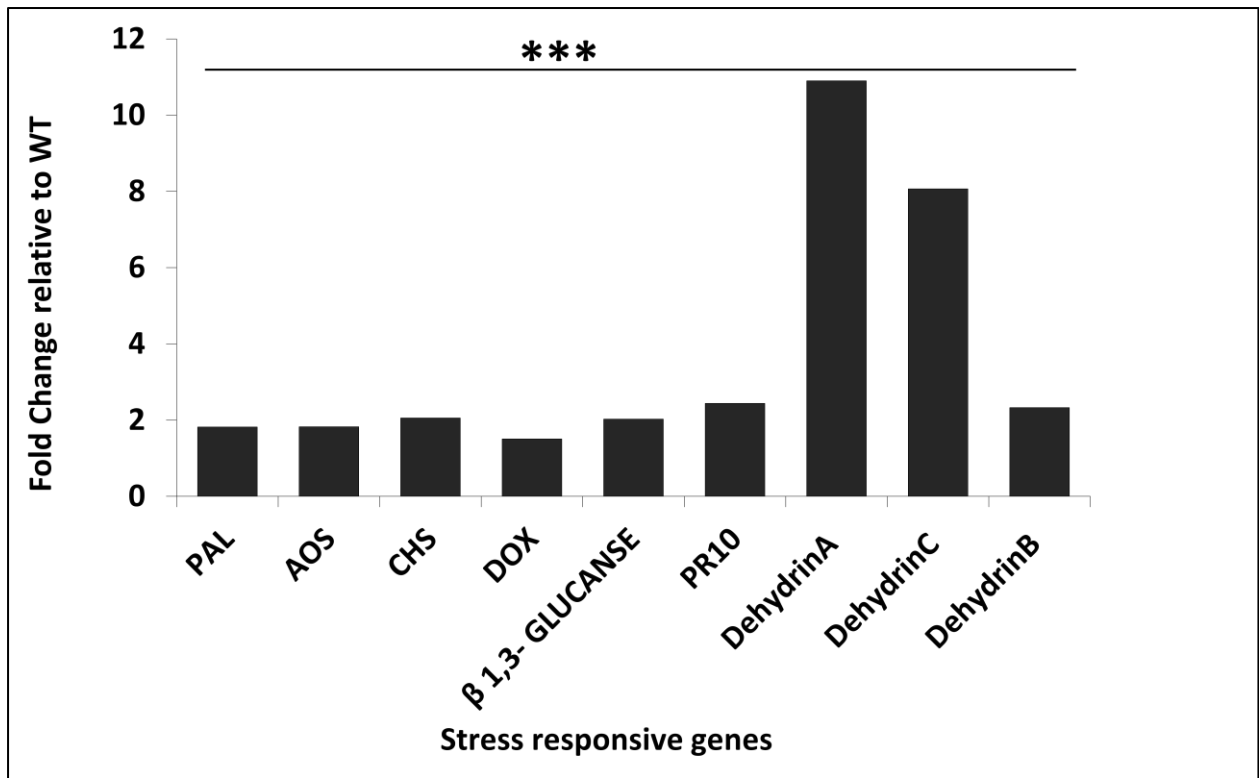


Figure 4.6 The bar plot shows log₂Fold change in transcript levels of key stress responsive genes in the *shlf* transcriptome including PNENYLALANINE AMMONIA LYASE (PAL), ALLENE OXIDE SYNTHASE (AOS), CHALCONE SYNTHASE (CHS), α -DIOXYGENASE (DOX), PATHOGENESIS RELATED PROTEIN 10 (PR 10) and DEHYDRIN A, DEHYDRIN B and DEHYDRIN C. Student's t test was performed and the P-values < 0.05, <0.001, and <0.0001 were represented as *, **, and ***, respectively.

In line with this, DAB staining and total ROS quantification of the mutant also showed high ROS levels in the mutant as compared to WT (Figure 4.7). Further, upon estimation of total chlorophyll content in the plants, *shlf* showed a higher accumulation of total chlorophyll (Figure 4.7). Interestingly, as the *RNA-Seq* revealed down-regulation of cell wall biogenesis/modification genes and cutin biosynthesis genes, we performed toluidine blue (TB) staining of the mutant to determine the cell wall porosity of *shlf*. The mutant showed a higher retention of TB stain as compared to WT and was more easily stained (Figure 4.7).

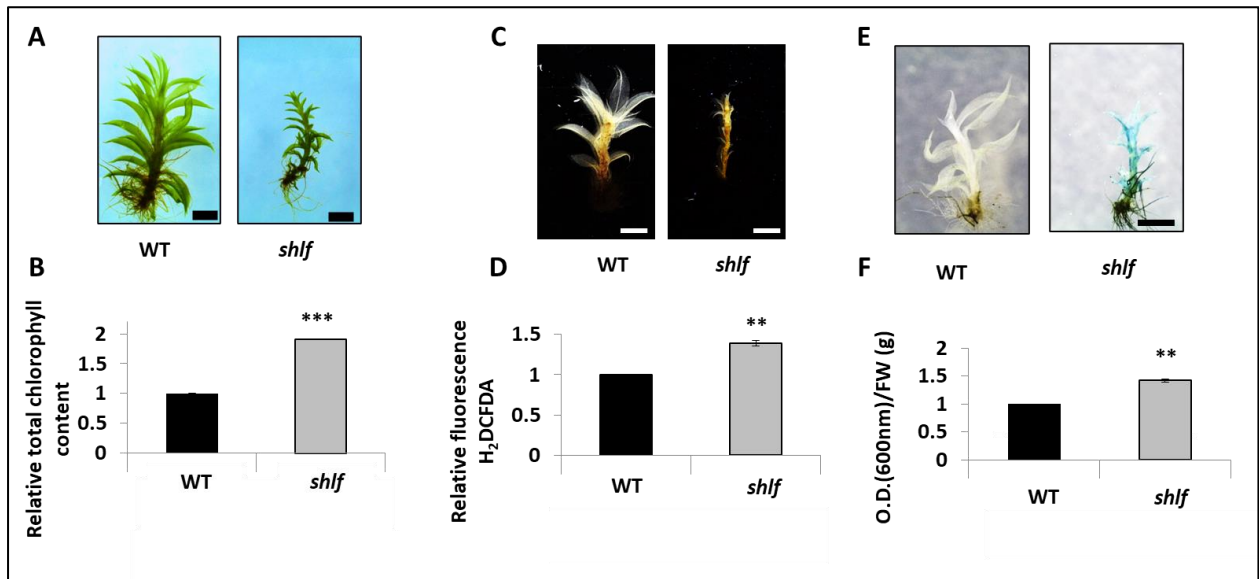


Figure 4.7 *shlf* gametophores exhibit elevated stress response. (A-B) *shlf* gametophores have higher levels of chlorophyll compared to WT. (B) Bar chart depicting 3 bio-replicates of relative chlorophyll quantification in *shlf* compared to WT. (C-D) *shlf* gametophores have elevated ROS levels as depicted by DAB staining (C) and relative H₂DCFDA fluorescence levels (D). (E-F) Toluidine blue staining of moss gametophores shows that *shlf* gametophores has higher stain penetration (E) and retention (F). Toluidine blue stain retention was plotted as O.D at 600nm per fresh weight in grams of tissue. Shapiro-wilk normality test was performed on the populations. Kruskal Wallis test with Dunn’s post-hoc test Bonferroni p-value corrections were performed to check for the statistically significant difference between the sample sets. ‘***’ and ‘**’ represent *p* value < 0.001 and < 0.01 respectively. Scale bar = 1 mm in all images.

The elevated stress response observed in *shlf* is restored upon SHLFp3 supplementation

As both miniSHLF and SHLFp3 (75 μ M) was able to successfully recover the mutant phenotypes including auxin distribution pattern, we attempted to study their effect on the mutant’s stress response. Our DAB staining and ROS quantification analyses using H₂DCFDA showed that a complete rescue was observed in the ROS levels of the mutant both upon miniSHLF overexpression and SHLFp3 supplementation (Figure 4.8). This shows that SHLF plays a role in mediating the cellular redox levels of moss. Supplementation of the random peptide (75 μ M), served as the negative control and showed comparable ROS levels (Figure 4.9).

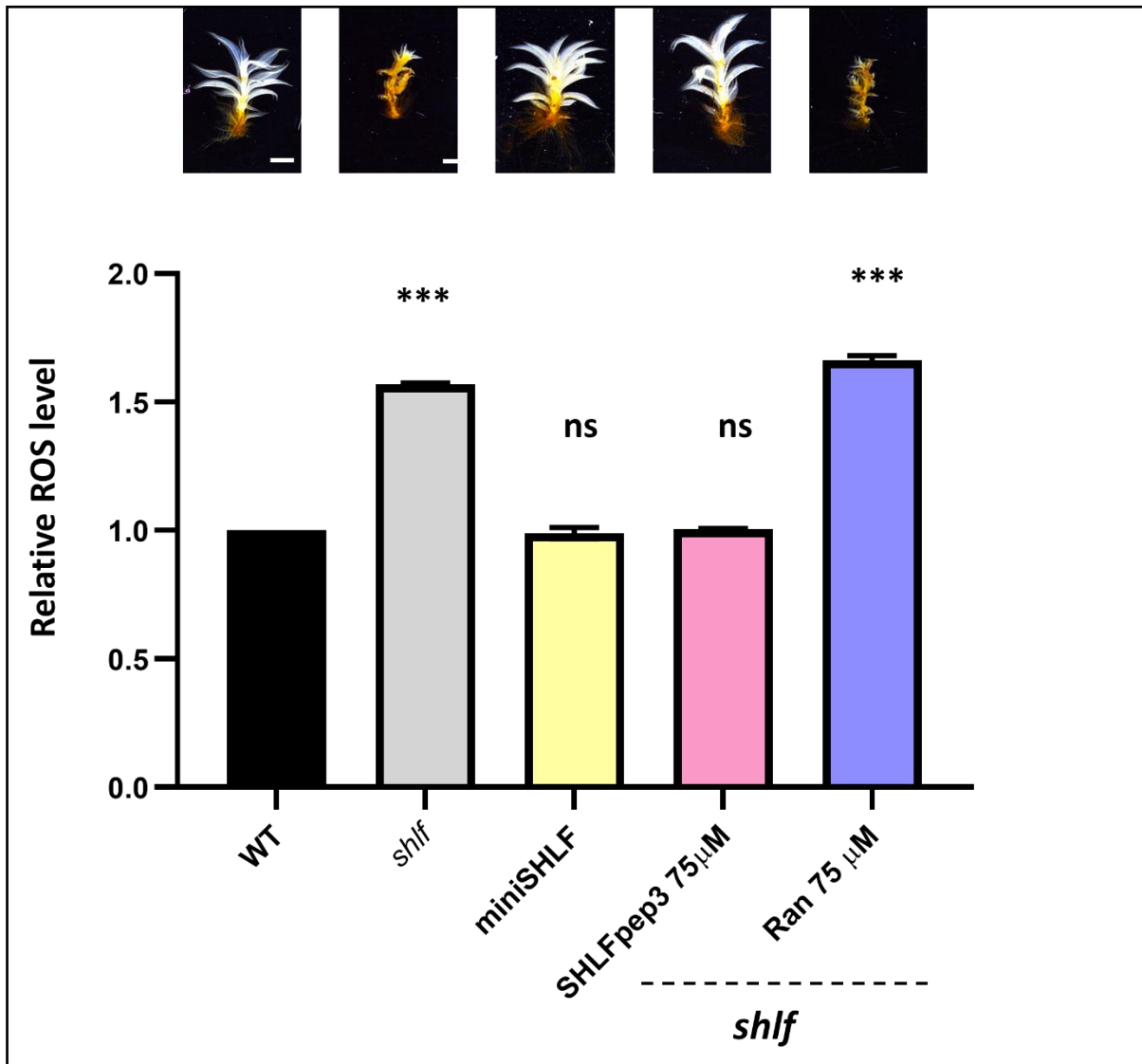


Figure 4.8 Panel represents the DAB-stained gametophores and the relative ROS levels in WT, *shlf*, N-TDR1-2 (miniSHLF) and *shlf* supplemented with SHLFpep3 and Ran (75µM)

Further, RT-qPCR validation of the stress responsive up-regulated in *shlf*, interestingly revealed that miniSHLF (N-TDR1-2-eGFP) and the SHLFpep 3 (75 µM) are able fully recover the level of these genes (Figure 4.9). In light of these analyses, we suggest that *shlf* has an impaired stress response, which is fully recovered upon supplementation of minimal SHLF and SHLFpep3.

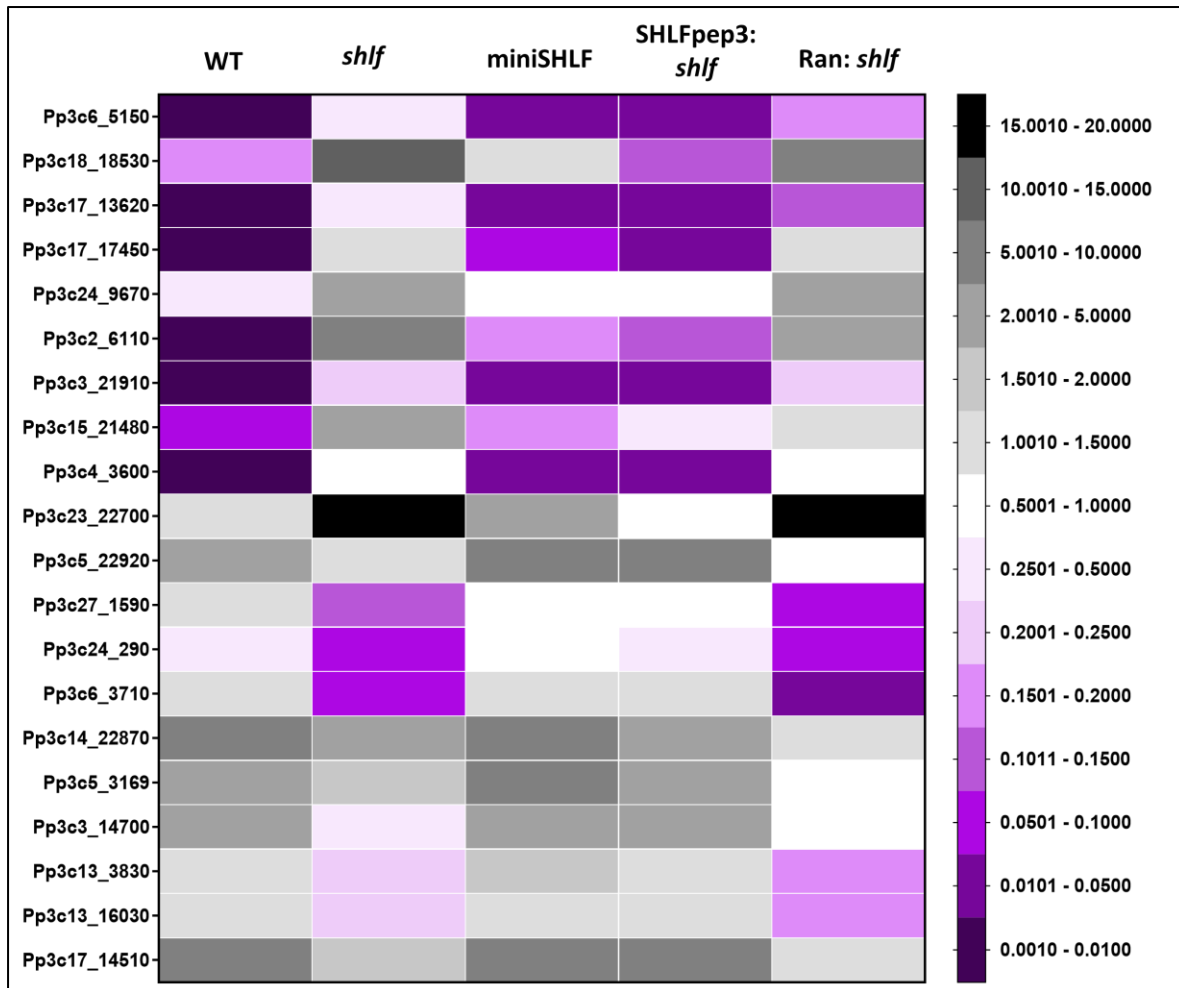


Figure 4.9 Heatmap representing the mean $2^{-\Delta C_t}$ values of the top 10 up- and down-regulated genes in WT, *shlf*, miniSHLF (N-TDR1-2), SHLFpep3 supplemented *shlf* gametophores, by using moss β -Actin and eIF2 as reference genes. Random peptide (Ran 75 μ M) supplemented *shlf* gametophores were used as a negative control. The color code in the heatmap depict the transcript abundance in each line.

4.3.2. SHLF modulates the levels of metabolites known to regulate auxin distribution pattern

The *shlf* transcriptome showed the up-regulation of flavonoid biosynthesis genes including *CHALCONE SYNTHASE*, *FLAVONOID SYNTHASE (FLS)* AND *FLAVNONE 3-HYDROXYLASE (F3H)*, which was validated using RT-qPCR analysis. Our analysis also revealed that the level of these genes was recovered upon overexpression of miniSHLF (N-TDR1-2-eGFP) and upon SHLFpep 3 (75 μ M) supplementation in *shlf* background (Figure 4.10).

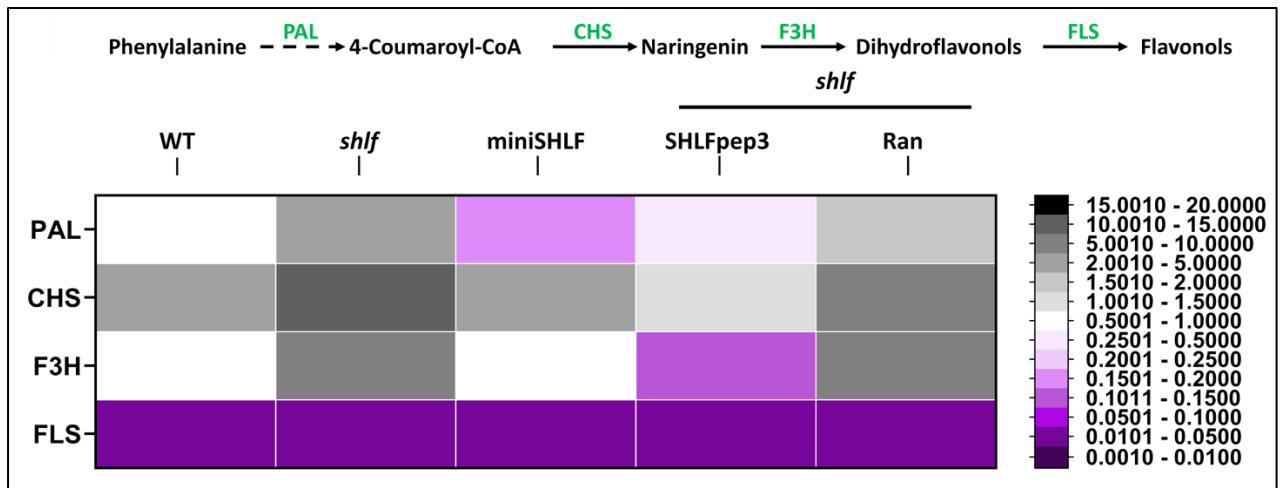


Figure 4.10 Schematic of flavonoid biosynthesis pathway with heatmap representing the mean $2^{-\Delta C_t}$ values of flavonoid biosynthesis genes including CHALCONE SYNTHASE, FLAVONOID SYNTHASE (FLS) AND FLAVNONE 3-HYDROXYLASE (F3H), in WT, *shlf*, miniSHLF, SHLFpep3 (75 μ M) and Ran (75 μ M) supplemented *shlf* gametophores, by using moss β -Actin and eIF2 as reference genes. The color code in the heatmap depict the transcript abundance in each line.

To further assess the possible involvement of SHLF in secondary metabolic pathways, a comparative metabolic profiling was attempted between *shlf* and WT (Figure 4.11A). Additionally, fold enrichment analyses revealed that most of the differentially accumulated metabolites in *shlf* belonged to the phenylpropanoid and flavonoid biosynthesis pathway (Figure 4.11B). The complete list of differentially expressed metabolites and their relative abundance in *shlf* are listed in Table 7.7.

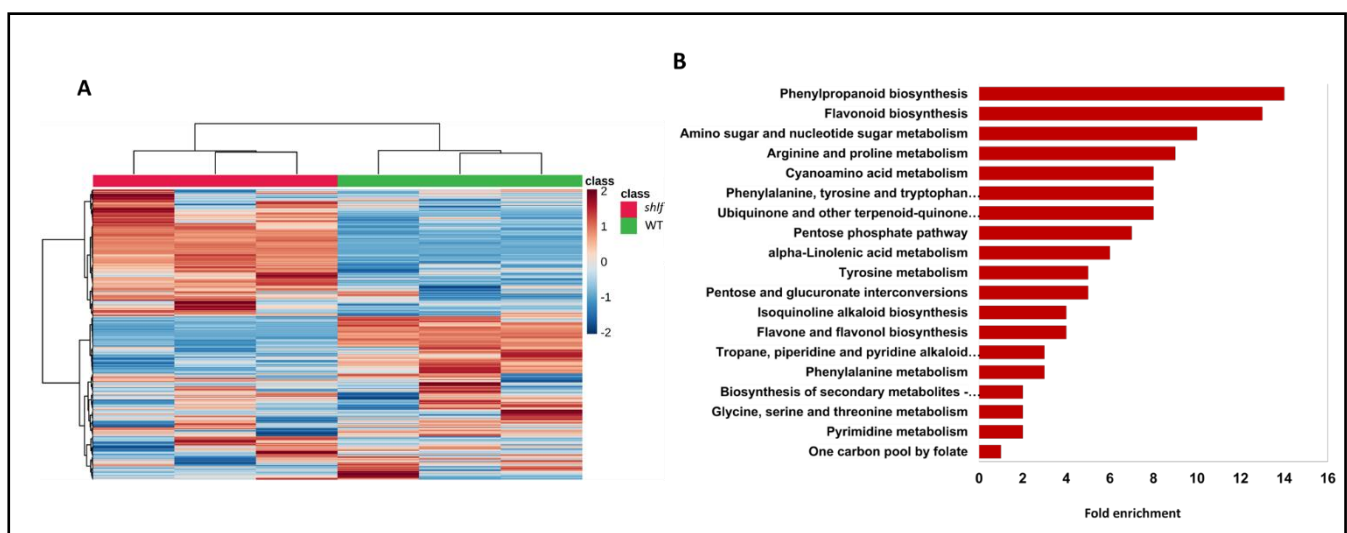
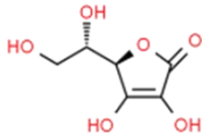
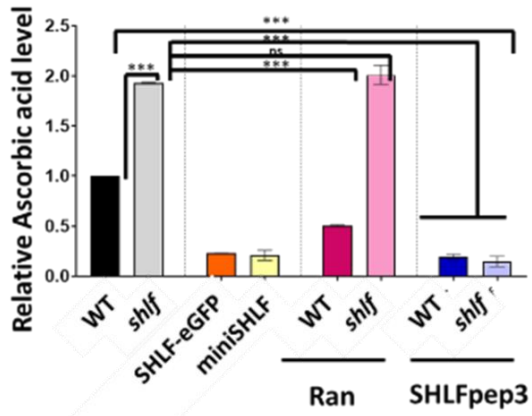


Figure 4.11 (A) Hierarchical clustering of differential metabolites between WT and *shlf*. (B) Fold enrichment depicting the pathways containing the differentially accumulated metabolites.

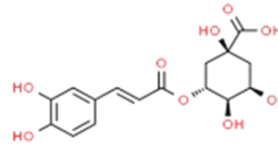
A Ascorbic Acid



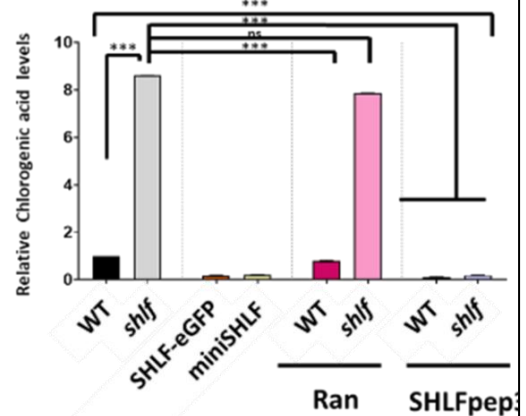
Formula	C ₆ H ₈ O ₆
RT (min)	1.52
[M-H] ⁻	175.0248



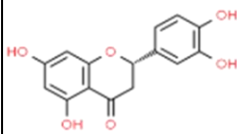
B Chlorogenic acid



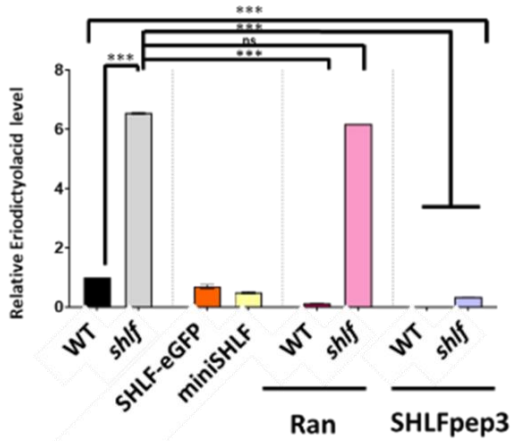
Formula	C ₉ H ₈ O ₄
RT (min)	7.53
[M-H] ⁻	179.0350



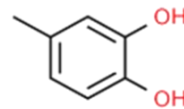
C Eriodictyol



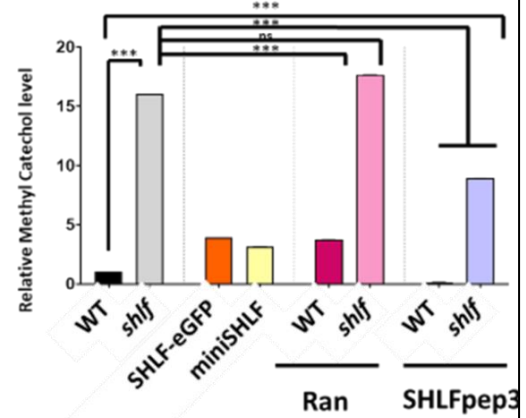
Formula	C ₁₅ H ₁₂ O ₆
RT (min)	9.69
[M-H] ⁻	287.056



D Methyl Catechol



Formula	C ₇ H ₈ O ₂
RT (min)	7.77
[M-H] ⁻	123.045



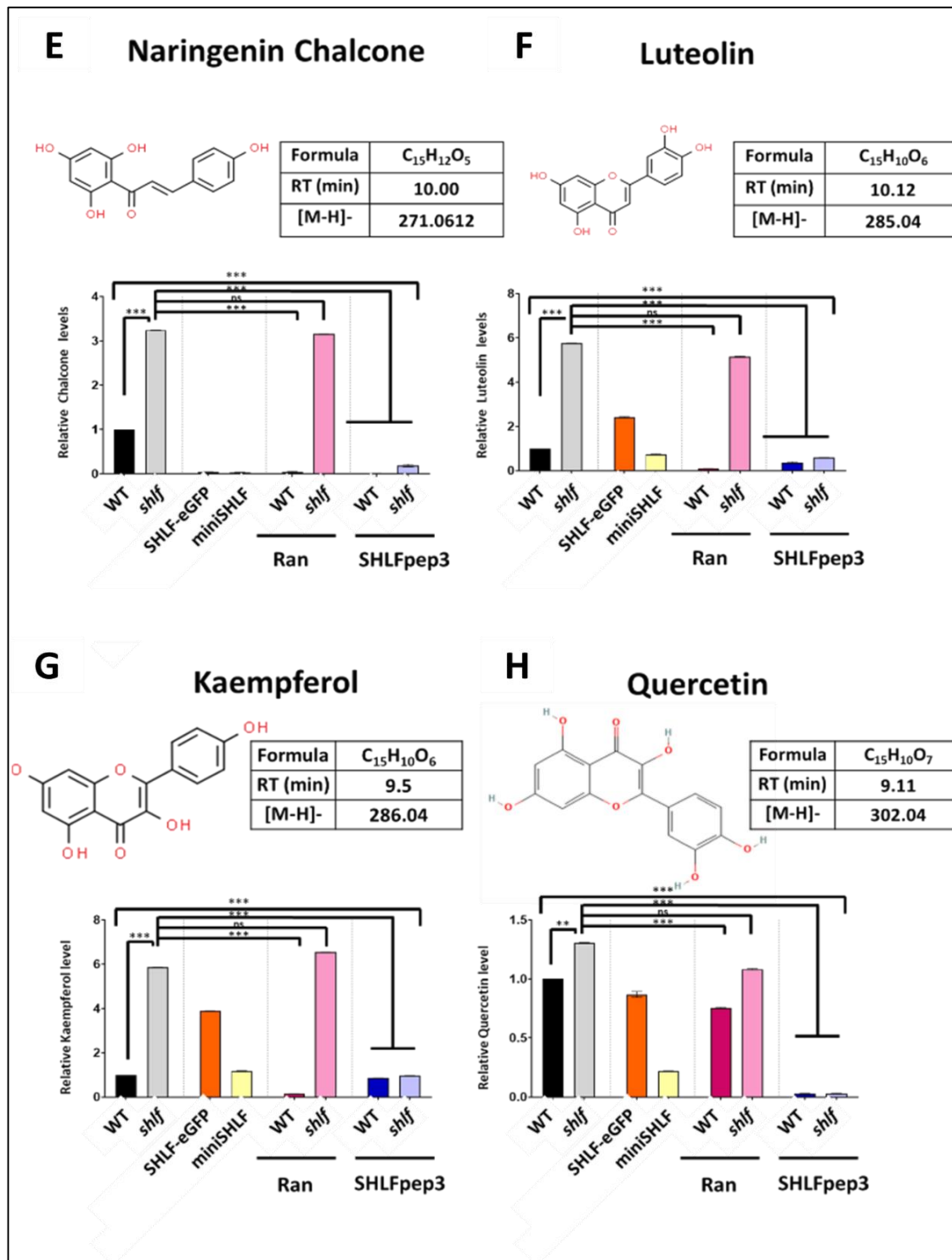


Figure 4.12 Barplots represent six bio-replicates depicting the relative abundance of metabolites (A-H) (A) ascorbic acid, (B) chlorogenic acid (CGA), stress responsive metabolites such as (C) Eriodyctiol and (D) Methyl catechol (E) naringenin chalcone and flavonoids, such as (F) Luteolin, (G) Kaempferol and (H) Quercetin; in WT, *shlf*, SHLF-eGFP, N-TDR1-2-eGFP, SHLFpep3 (75 μ M) and Ran (75 μ M) as a negative control. All the metabolites were identified in the negative ionization mode. Retention times, molecular weight and formula of each compound is represented in the insets. Shapiro-wilk normality test, followed by Student's t test was performed and the p-values < 0.05, < 0.001, and < 0.0001 were marked as *, **, and ***, respectively

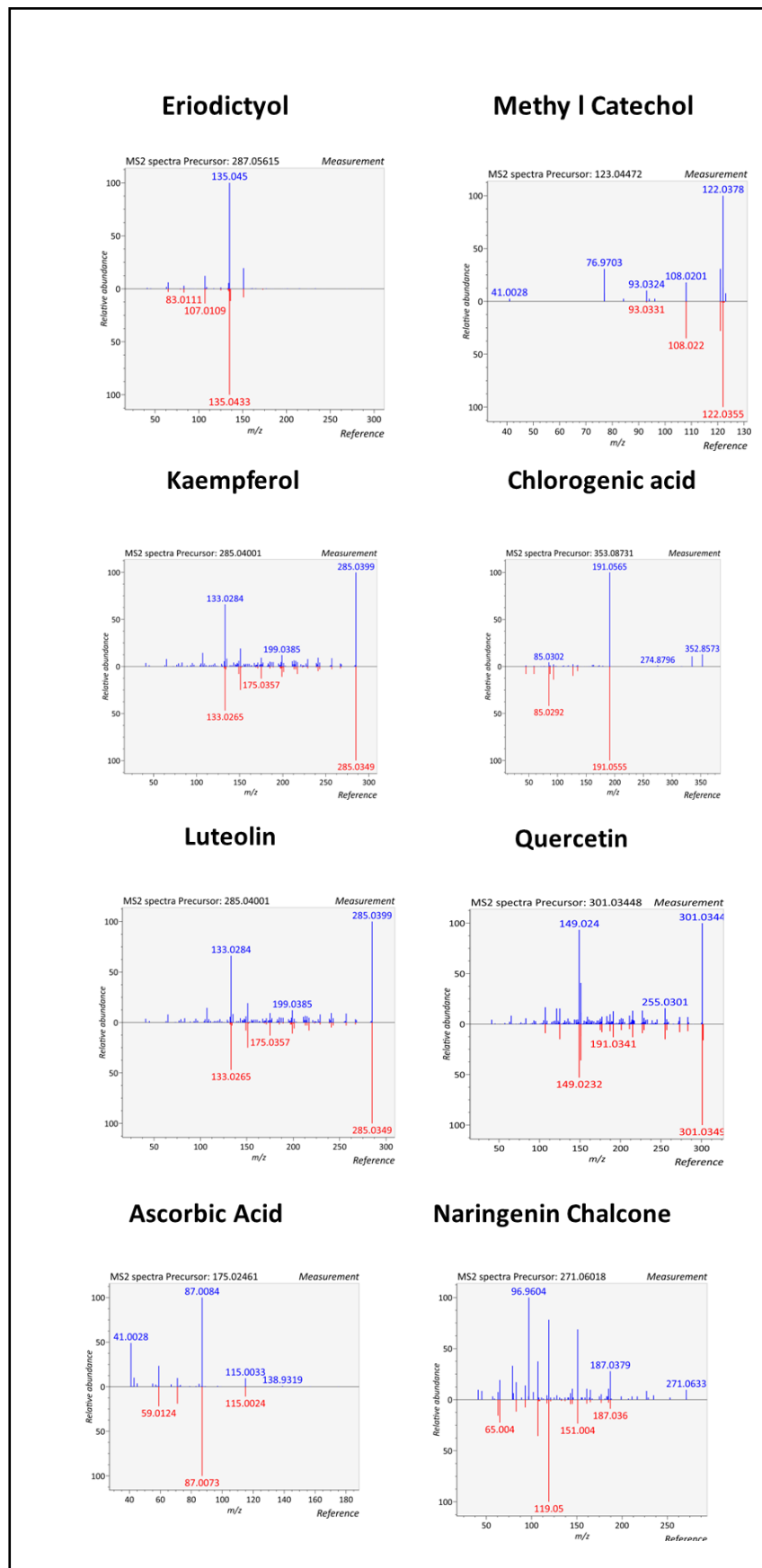


Figure 4.13 Graphs representing the fragmentation pattern of each metabolite quantified in the study matched with the MS-DIAL reference database.

The *shlf* metabolome also revealed the up-regulation of stress metabolites including several antioxidants, such as ascorbic acid, chlorogenic acid etc., which are known precursors of the phenylpropanoid biosynthesis pathway. Intermediates of the pathway like naringenin chalcone and flavonoids, including luteolin, eriodyctiol, methyl catechol, kaempferol and quercetin, were highly up-regulated in the mutant (Figure 4.12). The fragmentation pattern of each metabolite quantified in the study matched with the MS-DIAL reference database (Figure 4.13).

It is interesting to note that kaempferol and quercetin are known modulators of polar auxin transport (PAT) in flowering plants. In moss gametophores, the components of PAT are present and functional only in the leaves but not in stem. To study the effect of SHLF in regulating flavonoid levels in moss, we analyzed the metabolomes of miniSHLF lines and the SHLF_{pep3} (75 μ M) supplemented WT and *shlf* plants. The miniSHLF and SHLF_{pep3} *shlf* plants showed a complete rescue in the levels of the aforementioned metabolites with their levels being comparable to WT, while the Ran (75 μ M) supplemented *shlf* plants had similar metabolite levels as *shlf* (Figure 4.13 D-K). Our metabolomic analyses revealed that loss of function of *SHLF* causes increased accumulation of metabolites involved in the maintenance of auxin distribution pattern, which can be rescued by the miniSHLF (N-TDR1-2-eGFP) an SHLF_{pep3} supplementation.

4.3.3. *shlf* exhibits increased PD permeability for protein diffusion

The *shlf* mutant exhibits increased auxin accumulation at sites of synthesis combined with a reduce plasmodesmata frequency.

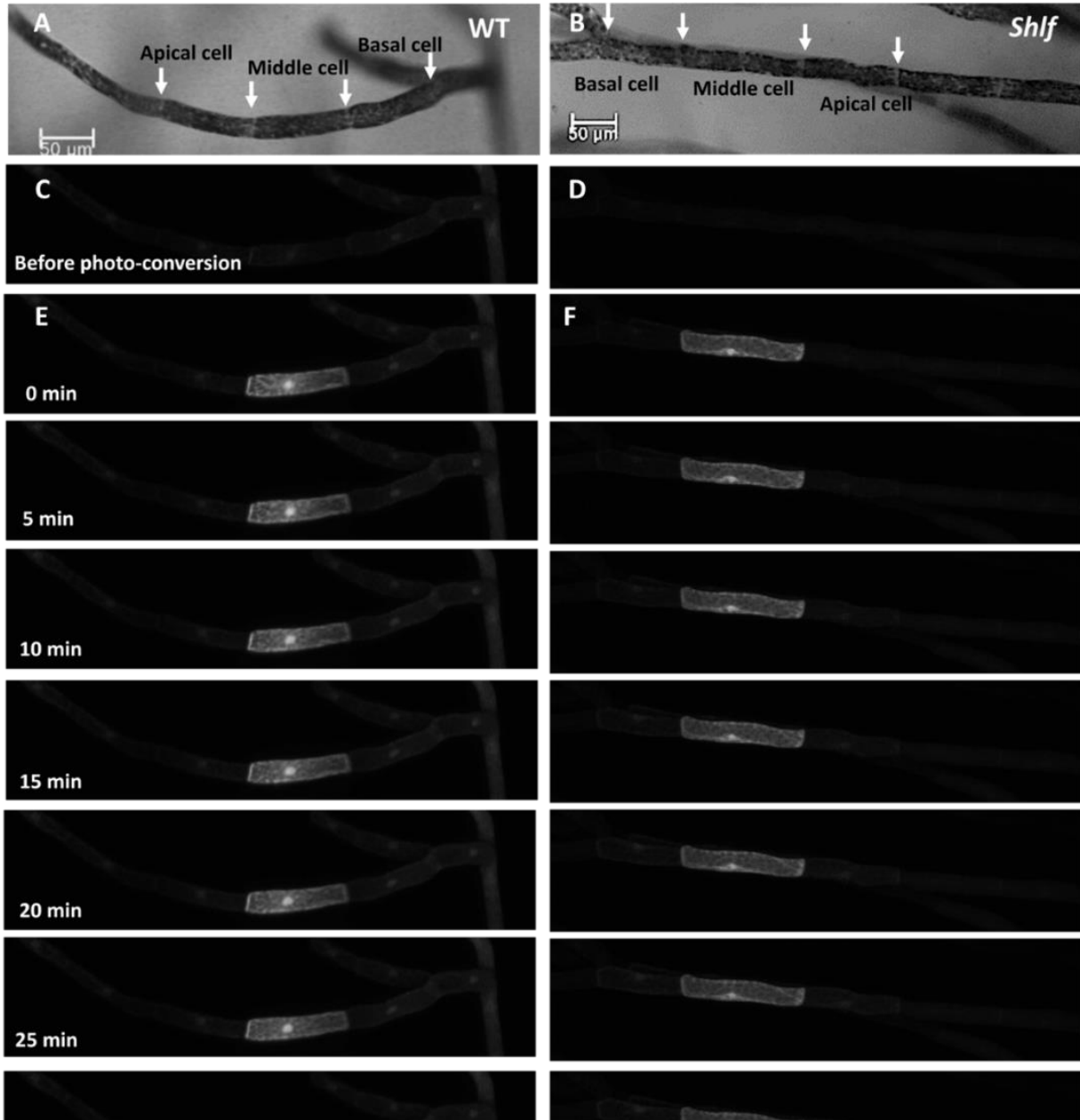


Figure 4.14 Time-lapse imaging of differential Dendra2 movement through PD in WT and *shlf* (A, B) Representative bright field images of photoconverted cells (cell 0) and their neighboring cells (cell +1 and -1) Arrowheads indicate the septa between cells. (C – E) Time lapse images of red fluorescence channel for WT and *shlf* protonema before (C, D) and after (E, F) photoconversion, showing the differential intercellular movement of Dendra2 from cell 0 to cell +1 (apical) and -1 (basal). The time-lapse imaging was started immediately after photoconversion and was recorded up to 30 min post with a 5 min interval (as shown). Scale bars = 50 μm.

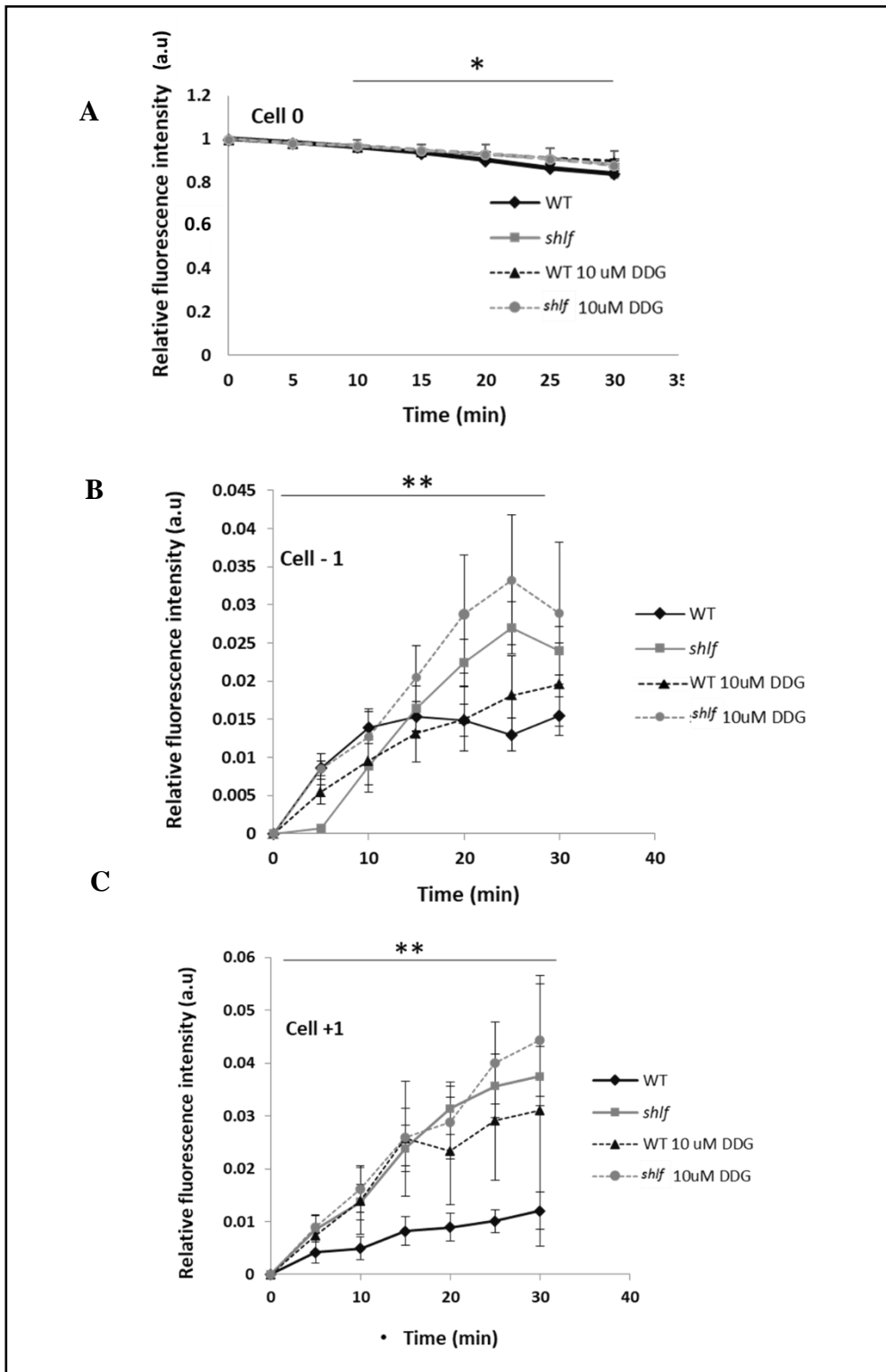


Figure 4.15 The time-course of relative fluorescence intensity of cell 0 (A), cell +1 (B) and cell -1 (C) in WT and *shlf*, with and without 24 hours of 10 μ M DDG treatment. The mean fluorescence intensity per pixel of cells +1, 0 and -1 was measured using Image J and was normalized to the mean fluorescence intensity per pixel of cell 0 right after photoconversion. Data sets represent the mean \pm SD. N = 14 for all sample sets. The symbols *, ** and *** have been used to represent p values of < 0.05 , < 0.01 and < 0.001 , respectively

Our results show that the permeability of the mutant PD was twice more than the WT. This result also corroborates the reduction of PD associated callose previously observed in the mutant (Figure 4.14). Further upon supplementation of a callose synthesis inhibitor, DDG (Di-deoxy glucose), the PD permeability of the mutant was further increased, though this increase was also observed in WT (Figure 4.15). It is curious to note that despite the increased permeability of the *shlf* OD, the symplastic auxin transport is still impaired in the mutant (Figure 4.15)

4.3.4. Identification of SHLF interactors by Co-IP using anti-SHLF

In an attempt to identify the interacting partners of SHLF, we performed Co-IP analyses using the total WT protein and the polyclonal anti-SHLF. Upon probing the elution fraction of the pull-down using anti-SHLF, a band of high molecular weight (between ~ 124 and ~ 250 kDa) was observed. Interestingly, this band was also observed in the washes and in the flowthrough fractions (Figure 4.15 A). However, this band was not seen upon probing the Co-IP fractions of an anti-IgG pull-down with anti-SHLF antibody. Instead, a ~50 kDa band was observed which was common between in the 3rd elution fraction of both A and B, which could be a non-specific protein (Figure 4.15 B).

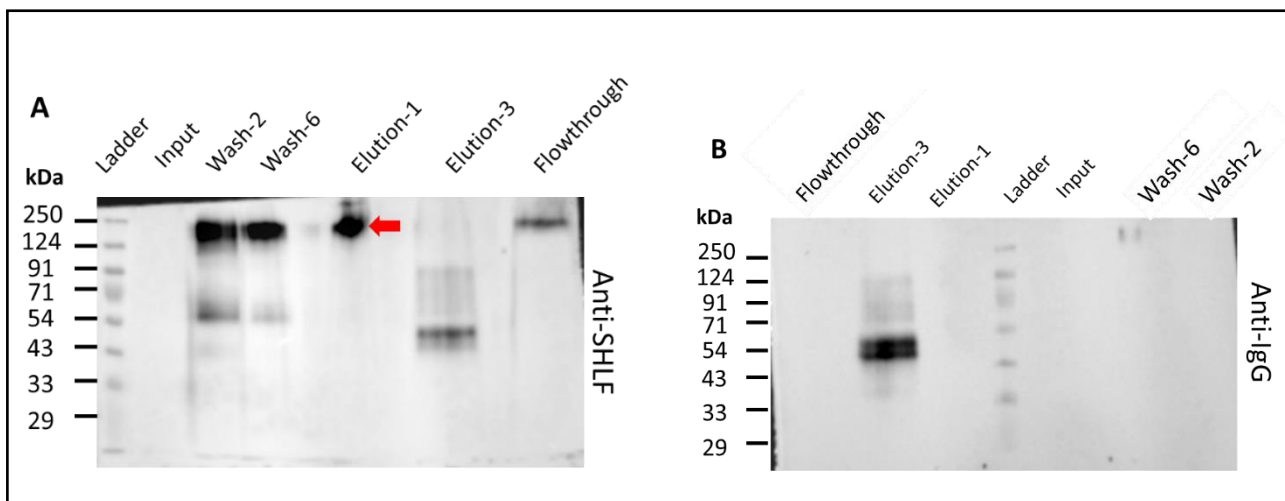


Figure 4.16 Western blot representing the anti-SHLF immunoprecipitated WT total protein probed with anti-SHLF (A) and anti-IgG (B).

Upon MS/MS analyses of the samples after Co-IP, the following list of SHLF specific interactors were detected at ~99.9% confidence (Table 4.1). Most of these proteins appear to play significant roles in both development and stress response. Out of them, two proteins

(Pp3c1_32560 and Pp3c10_9270) were designated as orphans as they are specifically present only in *P. patens* (Table 4.1).

Table 4.1 List of SHLF interactors identified from whole tissue lysate

Gene ID	Gene name	Function in plant development	Function in
Pp3c2_24240	Phosphoribosylaminoimidazole	Nucleotide/purine biosynthesis	Defense
Pp3c7_7230	Ribose-phosphate	enzyme that converts ribose 5-	Plant defense
Pp3c14_21530	Glutamate decarboxylase / L-	catalyzes the synthesis of GABA	Pathogen
Pp3c21_15520	Nitrilase related	Catalyzes the hydrolysis of nitrile	Abiotic and
Pp3c15_13020	Arachidonate 12-lipoxygenase /	promotes stress (both oxidative	
Pp3c26_6740	CHLOROPHYLL A/B BINDING	Serves as antenna complex	Modulates plant
Pp3c2_24270	Glyceraldehyde-3-phosphate	ubiquitous proteins that play	
Pp3c3_26590	MEDIATOR OF RNA	RNA transcription	Abiotic stress
Pp3c18_15811	Nucleosome associated protein 1	Stress induced chromatin	
Pp3c5_3820	Glucose 6 P isomerase	Oxidative enzyme of pentose	ROS elimination
Pp3c1_32560	Orphan		
Pp3c5_25640	Dihydrolipoyllysine-residue	catalyzes the overall conversion	SA binding
Pp3c10_9270	Orphan		
Pp3c13_13170	26S proteasome regulatory	Protein degradation	cellular response

N-terminal MASSSRALYLFIFSLALCHSVNAG

TDR 1 MSNVEKIQLAFKQITTASDNLKQQLINIINAPLQGSKIAEGFAKIIISAI AEYTIKINEGGVKNDNSPLPDADAKLVVQSLTTFV
 QVHQALLNVVIGKHGLLTLPFFPEPI**RLSLVLSLEA**TIDAFAFALIAEIPQKPAADVQFGSLSITLQVAITTYSSPILESNNLQE

TDR 2 MNTTEKIVLAFKQITTASDNLKQQLINIINAPLQGFKIAEGFAKIIISAI AEYTIKINEGGVKNYNSPLPDADAQIVVQSLTTFV
 QVHQALLNVVIGKHGLLTLPFFPEPI**RLSLVLSLEA**TIDAFAFALIAEIPQKPAADVQFGSLSVTLQVAITTYSSPILESNNLQE

TDR 3 MNTTEKIVLAFKQITTASDNLKQQLINIINAPLQGFKIAEGFAKIIISAI AEYTIKINEGGVKNYNSPLPDADAQIVVQSLTTFV
 QVHQALLNVVIGKHGLLTLPFFPEPI**RLSLVLSLEA**TIDAFAFALIAEIPQKPAADVQFGSLSVTLQVAITTYSSPILESNNLQE

TDR 4 MNTTEKIVLAFKQITTASDNLKQQLINIINAPLQGFKIAEGFAKIIISAI AEYTIKINEGGVKNYNSPLPDADAQIVVQSLTTFV
 QVHQALLNVVIGKHGLLTLPFFPEPI**RLSLVLSLEA**TIDAFAFALIAEIPQKPAADVQFGSLSVTLQVAITTYSSPLAK

C-terminal LGETMKEIAISA

Figure 4.17 A schematic representation of the D-box element present within the SHLF TDRs

We noted, out of the list of interactors, two of the proteins also had differential accumulation of transcripts in the mutant, as observed in our transcriptomic analyses. Pp3c26_6740 which encodes a CHLOROPHYLL A/B BINDING PROTEIN, was downregulated ~ 0.6-fold in *shlf*. Quite interestingly, Pp3c13_13170 which encodes the 26S proteasome regulatory complex, RP Triple-A ATPase 4 (RPT4), was up-regulated ~1.45 fold in the mutant (Table 7.5). Additionally, SHLF contains four D-box motifs (RLSLVSLEA), one in each TDR, which is recognized by the RPT4 protein family members for precise protein cleavage and/or degradation (Figure 4.16). As RPT4 proteins have been previously linked to regulating both plant development and stress response, this could serve crucial insights in the cleavage and regulation of SHLF for moss gametophore development.

4.4. Discussion

Auxin transport in *P. patens* gametophore stem has been proposed to be primarily symplastic (through PD), however, all the components of polar auxin transport (PAT) are reported to be present and influence gametophore development and branching (Bennett *et al.*, 2014; Viaene *et al.*, 2014; Coudert *et al.*, 2015). Most of the studies conducted in bryophytes to understand the evolutionary and functional conservation of auxin signalling components have employed reverse genetics approaches. The ploidy specific nature of molecular regulators makes it interesting to study the role of haploid specific genes in haploid plants. Given the haploid nature and less genetic redundancy of mosses, forward genetic studies can provide clues about novel clade-specific players involved in auxin signalling in bryophytes.

The *shlf* mutant has increased auxin accumulation at sites of synthesis, reduced plasmodesmata frequency and PD associated callose. Hence, we hypothesized that this phenotype could be due to the reduction in the transcript level of genes involved in PAT or due to reduced PD mediated symplastic transport. Our PD permeability assays using pDendra2 shows that though the mutant has a reduced PD frequency, these PD are twice as permeable to protein diffusion as compared to WT (Figure 4.14). Additionally, upon treatment of DDG, there is reduction in the levels of PD associated callose which further increases PD permeability (Figure 4.14). This is also corroborated by the results of previous experiments in GH3:GUS *shlf* lines, wherein upon treatment with DDG, an increased GUS accumulation was observed along the gametophore stem of *shlf* in contrast to the apex restricted GUS accumulation observed in the control GH3:GUS *shlf* lines (Mohanasundaram *et al.*, 2021). Though increased symplastic transport caused partial recovery in the auxin accumulation phenotype of the mutant, no significant phenotypic recovery was observed. This shows that reduced symplastic

diffusion of auxin in the mutant may not be the major contributing cause for the mutant phenotype. The increased kaempferol levels and the reduced PD associated callose in *shlf* (Figure 4.12, 4.14) appear to be reminiscent of the *Arabidopsis cad1-3* mutant, which is deficient in the PHYTOCHELATIN SYNTHASE gene (De Benedictis *et al.*, 2018). This indicates that akin to vascular plants, flavonoids could also play a role in regulating plasmodesmal aperture and permeability in mosses.

Additionally, we did not observe any transcript-level regulation of the previously characterized PAT-associated genes (PINA-D, YUCCA, ARF 12, TIR/AFB, PpIAA1a, PpIAA1b, PpIAA12, CBS, RSL4, DIAGEOTROPICA.) in the *shlf* transcriptome (Table 7.5). However, we did observe that certain uncharacterized members of the PIN superfamily (Pp3c24_2970 and Pp3c23_10200) were down-regulated in *shlf* (Table 7.6). A putative ABC transporter and CHALCONE SYNTHASE (CHS), were found to be up-regulated in the mutant. In flowering plants, PAT is influenced by flavonoids like quercetin and kaempferol, which exerts inhibitory effects on PINs (PIN -FORMED) and PGPs (P-glycoprotein/ABCB carriers) (Jacobs and Rubery, 1988; Fischer *et al.*, 1997; Murphy, Peer and Taiz, 2000; Brown *et al.*, 2001; Peer and Murphy, 2007). In addition, flavonoids have a regulatory role in mediating oxidative stress across plant lineages, including moss (Stevenson *et al.*, 2016; Wang *et al.*, 2022). Recently, high auxin accumulation in cells has also been shown to trigger ROS production, which in turn stimulates flavonoid biosynthesis, suggesting a feedback loop amongst the three molecules (Peer, Cheng and Murphy, 2013).

As the *shlf* gametophores have high ROS and flavonoid accumulation (Figure 4.8 and Figure 4.12), we speculate that a similar feedback-loop among auxin, ROS, and flavonoids may have caused high auxin accumulation at the gametophore apex and young leaves, elevated ROS levels and flavonoids accumulation (Quercetin, Kaempferol) in *shlf* gametophores (Figure 4.12 G and H). Notably, we observed that both miniSHLF overexpression and SHLF_{pep3} supplementation to *shlf* recovered the auxin, ROS and flavonoid accumulation to WT levels (Figure 4.8, 4.10-4.13), implicating a possible role of SHLF in regulating this feedback loop.

To our knowledge, this is the first report of a bryophyte-specific peptide affecting auxin distribution pattern via a flavonoid-ROS loop. Recent reports from *Arabidopsis* have shown that small miRNA encoding peptides regulates flavonoid levels and auxin signalling to modulate plant development (Sharma *et al.*, 2020). However, the link between this peptide with cellular redox levels and flavonoid concentration has not been established.

As SHLF does not encode for any known conserved domains and also undergoes cleavage, the nature and role of its interactors was challenging. Upon Co-IP analyses using the anti-SHLF specific polyclonal antibody, we found a list of possible interactors, which await functional validation (Table 4.1). It is interesting that the 26S Proteasome Subunit RPT4, is both an interactor and is differentially regulated in the mutant. The targets of RPT4 are usually recognized by a “destruction box” termed as the D-box motif (RXXLXX[L/I]XN) , in their protein sequence (and are degraded in a 26S proteasome dependent manner (Han *et al.*, 2008). SHLF encodes the D-box motif (RLSLVSLEA) in its TDRs and may thus be targeted for degradation by RPT4 (Figure 4.16). As SHLF is a highly abundant protein, this degradation machinery may serve as a control mechanism to monitor the levels of SHLF in the plant at any given time.

However, it is also plausible that the ubiquitin ligase RPT4 may associate with SHLF and act in a non-proteolytic way to regulate the protein function. Recent reports suggest that proteins such as E3 ubiquitin ligases RHA2 could function via the ubiquitin–26S proteasome pathway to regulate flavonoids biosynthesis (Lu *et al.*, 2021). However, the role of the 26S family members in regulating the auxin-flavonoid feedback loop has not been yet established. Given the limited knowledge of the key players involved in mediating the dynamics between auxin and flavonoid accumulation in bryophytes, it is conceivable that clade-specific genes such as SHLF and their interacting partners may provide crucial evolutionary insights in this conundrum.

The mutant, characterized by reduced SHLF expression, displays an elevated stress response even in the absence of external stressors. This suggests a potential link between the lack of secretory SHLF peptides and the initiation of autoimmunity in *shlf*. Interestingly, the complete rescue of stress response in the mutant upon supplementation with SHLFpep3 (as shown in Figure 4.8-4.10 and Figure 4.12) indicates that SHLFpeps act upstream of the ROS and flavonoid biosynthesis pathways. The absence of SHLFpeps influences the activity of auxin transporters by modulating flavonoid levels, leading to restricted auxin mobility from its sites of synthesis. Consequently, the overall increase in flavonoids in the plant results in the blockade of auxin transporters, disrupting proper auxin transport. Additionally, the reduced frequency of plasmodesmata (PD) observed in the mutant may be a secondary effect of the heightened stress response in *shlf*. As a response to elevated flavonoid levels, the plant reduces PD-associated callose. However, this adjustment alone proves insufficient for enabling

adequate auxin transport due to the presence of blocked auxin transporters at the sites of synthesis.

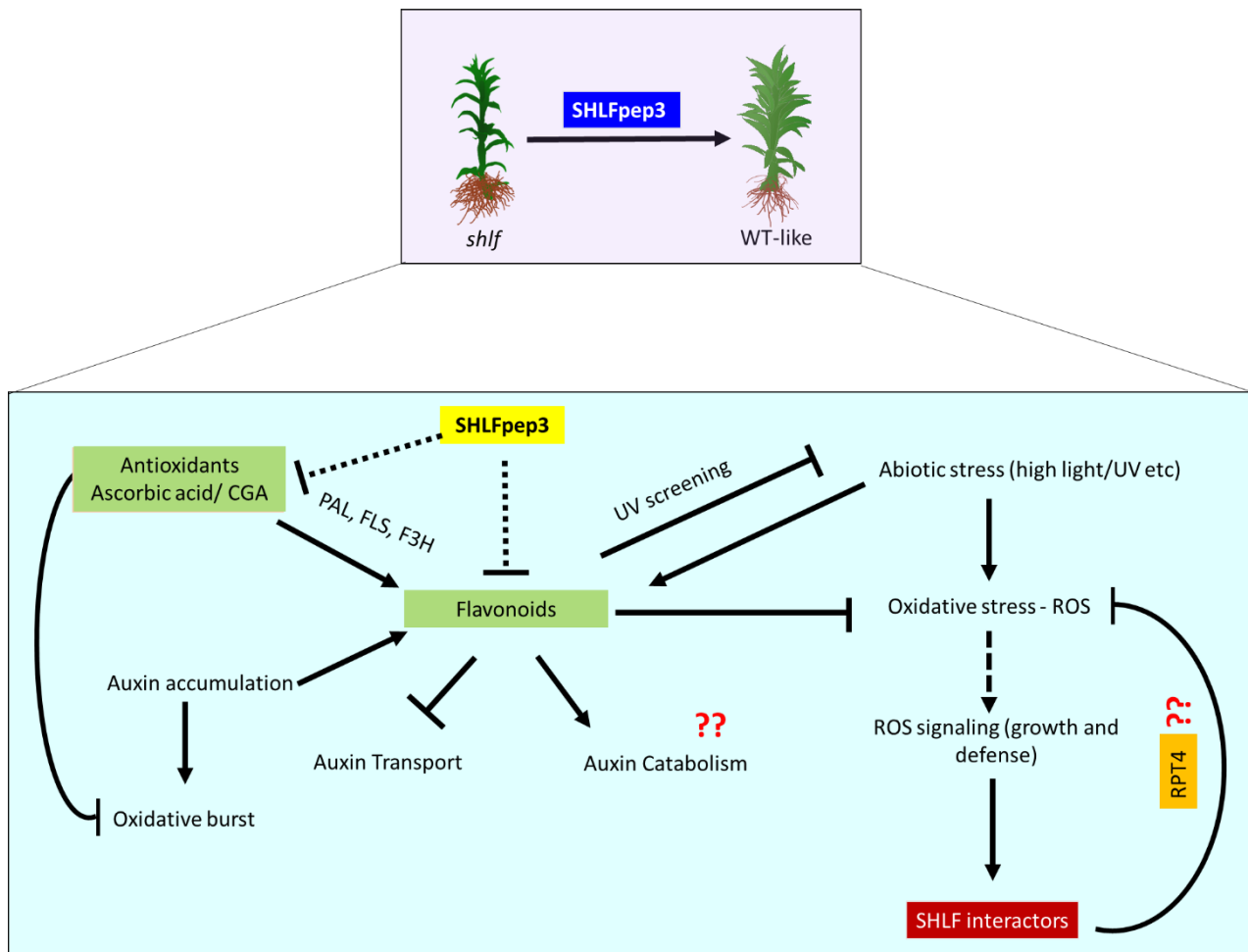


Figure 4.18 Proposed pathway for SHLF function

The elevated levels of flavonoids and ROS in the *shlf* mutant closely resemble the defense responses triggered by the perception of biotic or abiotic stressors. It is worth noting that the interaction between SHLF and its interactors, such as RPT4, likely contributes to maintaining stress levels when SHLF is present. In the absence of SHLF, these interactors might initiate a defense response. Overall, our proposed explanation suggests that the mutant exhibits both autoimmunity and auxin-related developmental phenotypes due to the absence of SHLF peptides. However, the introduction of SHLFpep3 supplementation restores the mutant to a healthy state, indicating the crucial role of SHLF peptides in maintaining proper plant function. The proposed pathway for SHLF function is detailed in Figure 4.17

Table 4.2 List of DEGs associated with polar auxin transport (PAT)

<i>P.patens</i> Gene ID	<i>A. thaliana</i> Gene ID	Gene function_Phytozome	log2FoldChange
Pp3c24_2970	AT1G70940	Auxin efflux carrier	-1.597952765
Pp3c23_10200	AT1G70940	Auxin efflux carrier	-1.089853045
Pp3c10_20720	AT5G55540	COR domain	0.88893991
Pp3c6_27380	AT5G19530	polyamine biosynthesis	0.955837903
Pp3c24_13120	AT5G13930	Chalcone synthase	1.039759671
Pp3c6_7260	AT3G28860	ATPase/ABC transporter	1.106131014

In summary, we report the molecular pathways regulated by SHLF and its peptides. The *shlf* gametophores display increased ROS levels, elevated chlorophyll accumulation, and enhanced stain retention, indicating cell wall damage. RNA-Seq analysis revealed the up-regulation of stress-responsive genes in *shlf*, including hallmark genes associated with both biotic and abiotic stress responses. Genes involved in phenylpropanoid biosynthesis, crucial for plant defense, were also up-regulated in *shlf*. Metabolomic profiling of the mutant identified the up-regulation of flavonoids like quercetin and kaempferol, known inhibitors of PAT (polar auxin transport). Supplementation with SHLFpep3 restored the heightened stress response, normalized levels of stress-responsive genes and metabolites. Co-IP assays identified SHLF interactors involved in plant development and stress response, including RPT4, a 26S proteasomal subunit that likely binds to SHLF's D-box domain to regulate protein cleavage and degradation. Our findings indicate that *shlf* exhibits an elevated stress response, which can be restored by supplementation with bryophyte-specific SHLFpep3. Our study reveals the involvement of SHLF in modulating the interactions among auxin transport, flavonoid accumulation, and ROS levels in bryophytes.

Results from Chapter 2, 3 and a part of Chapter 4 have been compiled in the following manuscript:

Palit S., Bhide A J., Mohanasundaram B., Pala M., & Banerjee, A. K. (2023) Secretory peptides from conserved tandem repeats of SHORT-LEAF regulate gametophore development in moss. (Under revision in **Plant Physiology**)

Summary and future directions

5. Summary and future directions

Terrestrialization by the ancestors of land plants prompted them to undergo rapid evolution to ensure survival (Lüttge, 2020). The last common ancestors of bryophytes and angiosperms were amongst the earliest land colonizers and garnered a plethora of adaptations to survive harsh conditions on land (Bennici, 2008). Plant evolution has been influenced by a variety of factors, including interactions with other organisms, environmental conditions, and genetic events such as gene duplication and horizontal gene transfer (De Vries and Archibald, 2018). Evolutionary processes have been brought about by changes in plant genomes, which are versatile enough to have significant genetic variation and plasticity to respond to dynamic environmental conditions. One such unique feature of plant genomes are Lineage-Specific Genes (LSGs), which are unique to specific clades and resulted from the amalgamation of evolutionary processes such as gene duplication, loss, and horizontal gene transfer (Zhang and Yin, 2015). LSGs play crucial roles in the evolution and adaptation of organisms, as they can undergo neo-functionalization to perform new functions or enhance existing ones (Lespinet *et al.*, 2002). Some examples of LSGs in plants include genes involved in C4 and CAM pathway, seed development and secondary metabolism contributing to plant development and defense (Lespinet *et al.*, 2002; Luna and Chain, 2021).

In recent years, there has been a remarkable progress in our understanding of genetic network evolution through the study of bryophyte model systems. Comparative genetic studies have shed light on the convergent evolution of shoot development across different plant lineages (Sakakibara *et al.*, 2008; Coudert *et al.*, 2015). While investigations have explored the conservation of gene networks governing shoot development in flowering plants and bryophyte gametophores, the role of bryophyte-specific genes remains unknown. Being one of the earliest colonizers of land, bryophytes such as mosses are ideal candidates to explore the contribution of LSGs in the evolution of unique adaptations. Recent studies have identified the presence of LSGs in the moss *Physcomitrium patens* (Xiao *et al.*, 2012) and in liverwort *Marchantia polymorpha*, (Tan *et al.*, 2023) and have linked their involvement in diverse biological processes including signal transduction, stress response and phytohormone signaling. However, the detailed functional characterization of LSGs in bryophytes have not been conducted till date.

We have recently reported isolation and characterization of the *short-leaf* (*shlf*) mutant from retrotransposon based forward genetic screen. The *shlf* mutant exhibits distinct traits, including a short-leaf phenotype, altered auxin distribution pattern, decreased apical

dominance, and a lower frequency of plasmodesmata (PD) (Mohanasundaram *et al.*2021,). These characteristics result from the insertion of a Tnt1 element in a previously unexplored gene family, which encodes a tandem direct repeat containing bryophyte-specific protein known as SHORT-LEAF (SHLF), which lacks any known conserved domains. The distinctive characteristics of SHLF may offer valuable insights into its role in specific cellular mechanisms unique to bryophytes.

To investigate the mechanistic basic for the molecular role of SHLF in moss gametophore development, we laid out the following objectives for this study.

1. Identification of crucial SHLF domains involved in gametophore development
2. Studying the secretory nature of SHLF and the role of its cleavage products in gametophore development
3. A transcriptomic and metabolomic approach to elucidate the mode of SHLF function in regulating moss development

Chapter 1: Introduction

SHORT-LEAF (SHLF), a protein found in moss, is a LSG (lineage-specific gene) with notable and distinctive attributes. These include possessing the longest known tandem direct repeats (TDRs), lacking conserved domains, localizing in the endoplasmic reticulum, and exhibiting a unique proteolytic cleavage pattern. Collectively, these attributes strongly suggest SHLF's role in specific cellular processes unique to bryophytes. However, given the absence of any known conserved domains in the protein, the molecular role of SHLF in moss gametophore development was unknown.

To gain insight into the evolutionary significance of SHLF in moss, we conducted an extensive literature survey exploring the evolution and function of clade-specific genes across diverse plant lineages. Additionally, considering that SHLF is a protein containing tandem direct repeats (TDRs) but lacks known conserved domains, we delved into the significance of TDRs and investigated other repeat-containing proteins in plants. We also examined the role of proteins containing domains of unknown functions (DUFs). This comprehensive survey allowed us to deepen our understanding of SHLF and its unique characteristics in the context of plant evolution.

The diminished leaf size, modified auxin distribution pattern, and low plasmodesmata (PD) frequency observed in the *shlf* mutant prompted us to investigate the evolution of auxin

distribution and transport in plants, with a particular focus on bryophytes. Through an extensive literature survey, we discovered that while endogenous regulators like flavonoids and ROS have been shown to influence auxin transport in flowering plants, limited knowledge exists regarding the genes involved in regulating the interplay between auxin, flavonoids, and ROS in both flowering and non-flowering plants. This knowledge gap highlights the need to explore the molecular mechanisms underlying auxin-flavonoid-ROS dynamics across plant species.

To shed light on the role of SHLF as a bryophyte-specific gene involved in moss development, we have formulated three objectives (as mentioned earlier) to identify the crucial SHLF domains and their mode of action. We aimed to unravel the molecular mechanisms underlying the involvement of SHLF in the developmental processes of moss gametophores.

Chapter 2: Identification of crucial SHLF domains regulating gametophore development

SHLF, a gene specific to the lineage, encodes an N-terminal signal peptide, four near-perfect tandem direct repeats (TDRs), and a C-terminal tail. Despite lacking any known functionally conserved domains, the crucial domains essential for SHLF function were identified through sequential domain deletion analyses. Additionally, we examined the effect of domain deletion on the sub-cellular localization pattern of SHLF and conducted immunoblotting analyses to understand the impact of SHLF domains on its post-translational processing. The study yielded the following significant findings:

1. We noticed that the N-terminal region of the protein encoding the signal peptide plays a critical role in facilitating ER trafficking, which is essential for protein function. Transgenic overexpression lines lacking the signal peptide (Δ N SHLF-eGFP) failed to recover the mutant phenotype (Fig. 2.2 and Fig. 2.4-2.6).
2. We showed that the divergent TDR1 exhibits limited functionality on its own, but it necessitates the presence of the conserved TDR2 for proper protein function, highlighting the modular nature of SHLF (Fig.2.2 and Fig 2.6)
3. Here, we report that a minimal SHLF (miniSHLF) comprising of the N-terminal, TDR1-2 is necessary and sufficient for protein function (Fig. 2.1 and 2.2).
4. Immunoblotting with GFP antibody revealed internal cleavage sites within the SHLF TDRs. Deletion of N-terminal was found to interfere with proteolytic protein cleavage. No GFP-tagged SHLF was detected in the C-terminal deletion lines, hinting at the role of the C-terminal in protein stability (Fig. 2.7).

5. Though the C-terminal domain is essential for protein function and stability but AlphaFold2-based structural predictions showed that this domain may have functional overlaps with the 9 amino acid inter TDR linker (ILESNNLQE) (Fig.2.8).
6. We generated a polyclonal SHLF-specific antibody (in rabbit) to detect all SHLF cleavage products (Fig. 2.10)
7. Immunoblotting with anti-SHLF showed that SHLF is indeed cleaved *in-planta*. The SHLF cleavage products were comparable across WT, *shlf* and the transgenic lines, which may be due to the basal level of SHLF protein present inside the cell (Fig.2.12).

Overall, our findings revealed the role of each SHLF domain in protein trafficking, proteolytic activity and in gametophore development. We also reported a minimal protein (miniSHLF) which was found to be necessary and sufficient for SHLF function.

Chapter 3: Studying the secretory nature of SHLF and the role of its cleavage products in gametophore development

Since the trafficking of SHLF to the ER via its signal peptide was proven to be crucial for its processing and function, we then endeavored to explore the necessity of secretion on protein function. For this purpose, we grew the plants in liquid cultures in an attempt to isolate total extracellular protein and peptides, which were then subjected to immunoblotting analyses and MS/MS-based identification to detect SHLF secretory peptides. We carried out supplementation assays using the WT and miniSHLF secretomes and peptides to identify the possible role of secretory SHLF. Further, we conducted synthetic peptide-based supplementation assays to identify the functional secretory peptide sufficient for auxin distribution and gametophore development in moss. The following were the important findings from this study:

1. Immunoblotting analyses using anti-SHLF against the WT secretome resulted in the detection of a truncated secretory version of SHLF, which was found to be absent in the mutant. This band corresponded to ~ 45 kDa, in contrast to the ~ 75 kDa of the total protein (Fig. 3.2)
2. In-gel digestion and MS/MS of the ~45, kDa revealed tryptic peptides originating from the divergent TDR1 and the conserved TDRs 2-4 (Fig. 3.2). The identification of this truncated version, along with the prior report of miniSHLF, highlights the significance of conserved tandem direct repeats (TDRs) and underscores the modular nature of

SHLF (Fig.2.1). This observation reinforces the notion that the conserved TDRs play a crucial role in the functionality of SHLF.

3. Our MS/MS analyses of secretomes revealed that SHLF is a major component of the WT secretome (Table 7.2 and Table 7.4).
4. We observed a complete phenotypic recovery upon supplementing the mutant with WT and miniSHLF secretomes, whereas, peptidome supplementation resulted in only partial recovery, hinting at the dosage dependent function of SHLF (Fig.3.5 and Fig. 3.6).
5. MS/MS analyses detected ~40 SHLF peptides in WT and in the miniSHLF secretomes but no peptides were seen in the mutant as well as in the domain-deleted overexpression lines which failed to recover the mutant phenotypes (Table 7.4).
6. Synthetic peptide supplementation assays revealed that SHLFpep1 (INIINAPLQGF) and SHLFpep3 (INIINAPLQGFKIA), peptides originating from the conserved TDRs act in a dosage-dependent manner to regulate moss gametophore development and auxin distribution pattern (Fig. 3.8 and Fig. 3.9).
7. Finally, SHLFpep3 was identified as the sole peptide that is both necessary and sufficient for establishing the appropriate auxin distribution pattern and facilitating gametophore development in moss (Fig. 3.8 and Fig. 3.9).

Taken together, our analyses revealed that SHLF acts like a typical secretory protein, by trafficking to ER and forming a pool of secretory peptides. We also identified the functional SHLFpeps regulating gametophore development and auxin response in moss.

Chapter 4: A transcriptomic and metabolomic approach to elucidate the mode of SHLF function in regulating moss development

Despite the indication from the altered auxin distribution pattern of the mutant about the role of SHLF in auxin transport, the absence of known conserved domains in SHLF makes it challenging to identify the exact molecular role of this protein. Here, we carried out further characterization of the mutant to identify the impact of altered auxin distribution on its stress response. We then carried out a transcriptomic and metabolomic-based analysis to identify the affected pathways in the mutant. Further, co-immunoprecipitation assays using SHLF-specific antibody were performed to identify the SHLF interactors. Based on the findings of the previous objective, we established that SHLFpep3 is the minimal functional unit of SHLF, hence, here we attempted to study the impact of SHLFpep3 supplementation on the altered pathways in the mutant. The following were the important findings from this study:

1. We report that *shlf* gametophores exhibit elevated ROS levels, higher chlorophyll accumulation and increased stain retention indicative of cell wall damage (Fig. 4.7).
2. RNA-Seq analysis of *shlf* showed the up-regulation of several stress-responsive genes including the hallmark genes of both biotic and abiotic stress response (Fig.4.3, 4.6 and Fig. 4.9).
3. Genes encoding the enzymes necessary for phenylpropanoid biosynthesis were up-regulated in *shlf* (Fig. 4.10)
4. Metabolomic profiling of *shlf* revealed the up-regulation of flavonoids such as quercetin and kaempferol, which are known classical endogenous inhibitors of polar auxin transport (Fig. 4.12)
5. SHLF_{pep3} supplementation recovers the elevated stress response of the mutant and restores the levels of stress-responsive genes and metabolites (Fig. 4.12)
6. Co-IP assays resulted in a list of SHLF interactors, which have been predicted to play dual roles in plant development and stress response (Fig. 4.16).
7. Notably, a 26S proteasomal subunit-encoding gene RPT4 was identified as an interactor and presumably binds to SHLF at its D-box domain to regulate protein cleavage and degradation (Fig. 4.17)

In summary, our results demonstrate that *shlf* exhibits an elevated stress response, which can be rescued upon supplementation of the bryophyte-specific SHLF_{pep3}. Here, we demonstrate that SHLF_{pep3} mediates the dynamics between auxin transport, flavonoid accumulation and ROS levels, and regulates gametophore development and stress response in moss (Fig. 4.18).

Future directions

Our study highlights the crucial role of secretion in the functionality of SHLF. We have demonstrated that SHLF functions primarily as a secretory protein, and through cleavage in plants, it generates a pool of secretory peptides. Specifically, SHLFpep3 derived from the conserved TDR plays a crucial role in regulating auxin distribution and gametophore development in moss. Moreover, our findings reveal that the *shlf* mutant displays an impaired immune response, increased production of flavonoids and ROS, which can be fully restored by SHLFpep3. These results suggest that SHLF peptides work in conjunction with primary and secondary metabolic pathways to elicit bryophyte-specific responses and regulate moss gametophore development.

Our investigation has paved the way for new avenues of research to uncover the clade-specific implications of SHLF and its role in mediating auxin distribution through endogenous regulators. By addressing the following questions in future studies, we may be able to further enhance our understanding of SHLF's significance and shed light on its specific contributions to auxin distribution in different plant lineages.

1. What are the SHLFpep receptors and the down-stream signalling involved in mediating gametophore development?
2. What is the mechanistic basis for the elevated stress response in the *shlf* mutant? Does it exhibit autoimmunity?
3. What is the role of phenylpropanoids in modulating development and defense mechanisms in bryophytes?
4. What is the evolutionary significance for clade specificity of SHLF?
5. Does SHLF act through conserved genetic pathways? If so, what could be the potential effects of SHLF in the development paradigms of other plant species?

Table 5: List of primers used in the current study.

Cloning primers	
attB1dN_SHLF_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAGCAACGTCGAGAAAATTCAG
attB2_eGFP_R	GGGGACCACTTTGTACAAGAAAGCTGGGTTAAGATCTTTACTTGTACAGCTCGTC
attB1_SHLF_F	GGGGACAACCTTTGTATACAAAGTTGTCCCGCATTGACTGAGTGACATAGTG
attB2_dC_SHLF_R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTCATCGTCTCCCGAGCTAGGCGAG
attB5r_dC_SHLF_R	GGGGACAACCTTTGTATACAAAGTTGTTGGCGAGGGGCGACGAGTAGGTGGTG
attB5r_SHLF_R	GGGGACCACTTTGTACAAGAAAGCTGGGTTCTAAGCGGATATCGCAATTTCTTT
attB5r_2_SHLF_R	GGGGACAACCTTTGTATACAAAGTTGTGTTGTTTCGATTCCAATATGGGCG
and attB5r_N_SHLF R	GGGGACAACCTTTGTATACAAAGTTGTCCCGCATTGACTGAGTGACATAGTG
Line confirmation primers	
Sb Kan_F	TCCTCCTGCCGAGAAAGTAT
Sb Kan_R	AATATCACGGGTAGCCAACG
RT-qPCR primers	
Actin qF	ACCGAGTCCAACATTCTACC
Actin qR	GTCCACATTAGATTCTCGCA
GFR qF1	TACCCCGACCACATGAAGCAGCAC
GFP qR1	TGCCGTCCTCCTTGAAGTCGATGCC
GFP qF2	CAACCACTACCTGAGCACCCAGTCC
GFP qR2	AGCTCGTCCATGCCGAGAGTGATCC
SHLF qF	GCGTCCAGCTCCAGGGCCTT
SHLF qR	CTGCGTTTCCTTGCGGAGG
HP qF	CTGGAAGCAACCATAGACGCCT
HP qR	TCATCGTCTCCCGAGCTTGGC
dNSHLF qF	ATGAGCAACGTCGAG
dCSHLFqR	GGGCGACGAGTAGGT
CHS qF	GGCCAAGGATCTAGCCGAAA
CHS qR	TGGCTTCTCTACCTCAGGCT
PAL qF	GCGGATGGTTGCTTACCAAA
PAL qR	CCGCCCATGATCCTATCGAG
FLS qF	CAGTCGAGCCACCGTTCTTC
FLS qR	CCGTGGTTGAGCACCTGAAA
F3H qF	TGGTTATGGCAACACGGGAA
F3H qR	TCGCACAGCCTTTTCCACAG
RNA-seq validation primers	
Pp3c6_5150 qF	GGGAATTCACCGAGGACGGA
Pp3c6_5150 qR	ATCGTGCTTGTCCGCATAGC
Pp3c18_18530 qF	GGCATCTCTACACCAGGCCA

Pp3c18_18530 qR	CTCCTGATGGCACTGTCACG
Pp3c17_13620 qF	AGGAATCAGACCACCGAGCA
Pp3c17_13620 qR	CTGCGGAGGTGAAAAGCAGT
Pp3c17_17450 qF	CAGAGGGATGGGCATCATGG
Pp3c17_17450 qR	ATGCAGACAAAGCTTCCCGT
Pp3c24_9670 qF	ACTCCACTTCCCGGAACGAA
Pp3c24_9670 qR	ATCAGGCTACTGGTGTCTGGT
Pp3c2_6110 qF	CGCTTGGCGTAATCTCCTCC
Pp3c2_6110 qR	CTTGCACTGAGGGCAGTTGT
Pp3c3_21910 qF	CGTCCCACGACAAATCAGGG
Pp3c3_21910 qR	GCTCCAATTGCACCTTGTCA
Pp3c15_21480 qF	CTACGCCACACTTTGTCCGT
Pp3c15_21480 qR	TTCCAAATCCAAGCCAGCGT
Pp3c4_3600 qF	ACGAGTCTGGAGGTTTGAAGC
Pp3c4_3600 qR	AATCGCCACAAGCACCGTAG
Pp3c23_22700 qF	CGCGGTACAGTTGGCACATT
Pp3c23_22700 qR	ACGGCATCCCACCAGATCC
Pp3c5_22920 qF	GTCAATGGCGATCCAGGGTG
Pp3c5_22920 qR	CACTTGGTGCCTGTTGGACA
Pp3c27_1590 qF	ACGATGTGAAGCGCGTTGAA
Pp3c27_1590 qR	ACACAGTGCGATTTCGTCCAC
Pp3c24_290 qF	CCTTGTGGTGTCAACGGAGG
Pp3c24_290 qR	ATCTGTCCGCTCTGCGATTG
Pp3c6_3710 qF	ACACCTTCAAGCTTCGCCTG
Pp3c6_3710 qR	TCTTTGCACTGGGAGTCGGA
Pp3c14_22870 qF	CTCCGCTCCAGGGTTCCA
Pp3c14_22870 qR	ATTGCACGACGAGCTTAGCG
Pp3c5_3169 qF	CCAAACCTTGACGCTGCTTG
Pp3c5_3169 qR	AGGGCACTTCTTCCCAAAGC
Pp3c3_14700 qF	GCGGCAAAGATCCCTTCTTA
Pp3c3_14700 qR	TCCTGCTTACGATCCTGCT
Pp3c13_3830 qF	TTAACACAATTTCCGGCGGC
Pp3c13_3830 qR	GCATCATGCAACCTTGTAGAGTC
Pp3c13_16030 qF	GGCTGGATACTGGACCTCCT
Pp3c13_16030 qR	TTCTCTTCAACGCTTCGGCG
Pp3c17_14510 qF	CCTTACGCTTTGGATGCGCT
Pp3c17_14510 qR	TGTCCGTTGTTGTAAGTTGTCTTT

References

6. References

- Achard, P. *et al.* (2009) ‘Gibberellin signaling controls cell proliferation rate in Arabidopsis’, *Current biology*. Elsevier, 19(14), pp. 1188–1193.
- Adamowski, M. and Friml, J. (2015) ‘PIN-dependent auxin transport: action, regulation, and evolution’, *The Plant Cell*. American Society of Plant Biologists, 27(1), pp. 20–32.
- Agrawal, G. K. *et al.* (2010) ‘Plant secretome: unlocking secrets of the secreted proteins’, *Proteomics*. Wiley Online Library, 10(4), pp. 799–827.
- Alexandersson, E. *et al.* (2013) ‘Plant secretome proteomics’, *Frontiers in plant science*. Frontiers Media SA, 4, p. 9.
- Amsbury, S., Kirk, P. and Benitez-Alfonso, Y. (2018) ‘Emerging models on the regulation of intercellular transport by plasmodesmata-associated callose’, *Journal of experimental botany*. Oxford University Press UK, 69(1), pp. 105–115.
- Antalis, T. M. and Buzza, M. S. (2016) ‘Extracellular: Plasma Membrane Proteases--Serine Proteases’, *Encyclopedia of Cell Biology*. Elsevier, p. 650.
- Arendsee, Z. W., Li, L. and Wurtele, E. S. (2014) ‘Coming of age: orphan genes in plants’, *Trends in plant science*. Elsevier, 19(11), pp. 698–708.
- Armengot, L., Marquès-Bueno, M. M. and Jaillais, Y. (2016) ‘Regulation of polar auxin transport by protein and lipid kinases’, *Journal of experimental botany*. Oxford University Press UK, 67(14), pp. 4015–4037.
- Arteaga-Vazquez, M. A. (2016) ‘Land plant evolution: listen to your elders’, *Current Biology*. Elsevier, 26(1), pp. R26--R29.
- Ashton, N. W., Grimsley, N. H. and Cove, D. J. (1979) ‘Analysis of gametophytic development in the moss, *Physcomitrella patens*, using auxin and cytokinin resistant mutants’, *Planta*. Springer, 144(5), pp. 427–435.
- Band, L. R. (2021) ‘Auxin fluxes through plasmodesmata’, *New Phytologist*. Wiley Online Library, 231(5), pp. 1686–1692.
- Barbez, E. and Kleine-Vehn, J. (2013) ‘Divide Et Impera—cellular auxin compartmentalization’, *Current opinion in plant biology*. Elsevier, 16(1), pp. 78–84.
- Becker, B. and Marin, B. (2009) ‘Streptophyte algae and the origin of embryophytes’, *Annals of botany*. Oxford University Press, 103(7), pp. 999–1004.
- Beeckman, T. and Friml, J. (2010) ‘Nitrate contra auxin: nutrient sensing by roots’, *Developmental cell*. Elsevier, 18(6), pp. 877–878.
- De Benedictis, M. *et al.* (2018) ‘The Arabidopsis thaliana knockout mutant for phytochelatin synthase1 (cad1-3) is defective in callose deposition, bacterial pathogen defense and auxin content, but shows an increased stem lignification’, *Frontiers in Plant Science*. Frontiers Media SA, 9, p. 19.

- Benjamins, R., Malenica, N. and Luschnig, C. (2005) 'Regulating the regulator: the control of auxin transport', *Bioessays*. Wiley Online Library, 27(12), pp. 1246–1255.
- Bennett, T. (2020) 'The evolution of hormonal signalling in plant development', in *Seminars in Cell & Developmental Biology*.
- Bennett, T. A. *et al.* (2014) 'Plasma membrane-targeted PIN proteins drive shoot development in a moss', *Current Biology*. Elsevier, 24(23), pp. 2776–2785.
- Bennici, A. (2008) 'Origin and early evolution of land plants: Problems and considerations', *Communicative & Integrative Biology*. Taylor & Francis, 1(2), pp. 212–218.
- Bharathan, G. *et al.* (1999) 'Phylogenetic relationships and evolution of the KNOTTED class of plant homeodomain proteins.', *Molecular Biology and Evolution*. Oxford University Press, 16(4), pp. 553–563.
- Bhattacharjee, S. and Sharma, G. D. (2012) 'Effect of dual inoculation of arbuscular mycorrhiza and rhizobium on the chlorophyll, nitrogen and phosphorus contents of pigeon pea (*Cajanus cajan* L.)'. Scientific Research Publishing.
- Bircheneder, S. and Dresselhaus, T. (2016) 'Why cellular communication during plant reproduction is particularly mediated by CRP signalling', *Journal of Experimental Botany*. Oxford University Press UK, 67(16), pp. 4849–4861.
- Birchler, J. A. and Veitia, R. A. (2010) 'The gene balance hypothesis: implications for gene regulation, quantitative traits and evolution', *New Phytologist*. Wiley Online Library, 186(1), pp. 54–62.
- Björnsson, A., Mottagui-Tabar, S. and Isaksson, L. A. (1996) 'Structure of the C-terminal end of the nascent peptide influences translation termination.', *The EMBO journal*, 15(7), pp. 1696–1704.
- Blázquez, M. A., Nelson, D. C. and Weijers, D. (2020) 'Evolution of plant hormone response pathways', *Annual Review of Plant Biology*. Annual Reviews, 71, pp. 327–353.
- Bold, H. C. (1949) 'The morphology of *Chlamydomonas chlamydogama*, sp. nov.', *Bulletin of the Torrey Botanical Club*. JSTOR, pp. 101–108.
- Bold, H. C. *et al.* (1949) 'The Origin of the Life Cycle of Land Plants: A simple modification in the life cycle of an extinct green alga is the likely origin of the first land plants', *Bulletin of the Torrey Botanical Club*. JSTOR, 30(2), pp. 178–186.
- Bold, H. C. *et al.* (2004) 'Computational identification and characterization of novel genes from legumes', *Bulletin of the Torrey Botanical Club*. American Society of Plant Biologists, 135(3), pp. 1179–1197.
- Bonferroni, C. (1936) 'Teoria statistica delle classi e calcolo delle probabilita', *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, 8, pp. 3–62.
- Bowman, J. L. *et al.* (2016) 'Evolution in the cycles of life', *Annual review of genetics*. Annual Reviews, 50, pp. 133–154.

- Bowman, J. L. *et al.* (2017) ‘Insights into land plant evolution garnered from the *Marchantia polymorpha* genome’, *Cell*. Elsevier, 171(2), pp. 287–304.
- Bowman, J. L. *et al.* (2019) ‘Something ancient and something neofunctionalized—evolution of land plant hormone signaling pathways’, *Current Opinion in Plant Biology*. Elsevier, 47, pp. 64–72.
- Bressendorff, S. *et al.* (2016) ‘An innate immunity pathway in the moss *Physcomitrella patens*’, *The Plant Cell*. American Society of Plant Biologists, 28(6), pp. 1328–1342.
- Brockington, S. F. *et al.* (2015) ‘Lineage-specific gene radiations underlie the evolution of novel betalain pigmentation in *Caryophyllales*’, *New Phytologist*. Wiley Online Library, 207(4), pp. 1170–1180.
- Brown, D. E. *et al.* (2001) ‘Flavonoids act as negative regulators of auxin transport in vivo in *Arabidopsis*’, *Plant physiology*. American Society of Plant Biologists, 126(2), pp. 524–535.
- Brunetti, C. *et al.* (2013) ‘Flavonoids as antioxidants and developmental regulators: relative significance in plants and humans’, *International journal of molecular sciences*. MDPI, 14(2), pp. 3540–3555.
- Brunetti, C. *et al.* (2018) ‘Modulation of phytohormone signaling: A primary function of flavonoids in plant--environment interactions’, *Frontiers in Plant science*. Frontiers Media SA, 9, p. 1042.
- Cai, J. J. *et al.* (2006) ‘Accelerated evolutionary rate may be responsible for the emergence of lineage-specific genes in ascomycota’, *Journal of Molecular Evolution*. Springer, 63(1), pp. 1–11.
- Campbell, L. and Turner, S. R. (2017) ‘A comprehensive analysis of RALF proteins in green plants suggests there are two distinct functional groups’, *Frontiers in plant science*. Frontiers Media SA, 8, p. 37.
- Campbell, M. A. *et al.* (2007) ‘Identification and characterization of lineage-specific genes within the *Poaceae*’, *Plant physiology*. American Society of Plant Biologists, 145(4), pp. 1311–1322.
- Casson, S. A. *et al.* (2002) ‘The POLARIS gene of *Arabidopsis* encodes a predicted peptide required for correct root growth and leaf vascular patterning’, *The Plant Cell*. American Society of Plant Biologists, 14(8), pp. 1705–1721.
- Chater, C. C. *et al.* (2016) ‘Origin and function of stomata in the moss *Physcomitrella patens*’, *Nature plants*. Nature Publishing Group, 2(12), pp. 1–7.
- Chen, X.-Y. *et al.* (2009) ‘The *Arabidopsis* callose synthase gene *GSL8* is required for cytokinesis and cell patterning’, *Plant physiology*. American Society of Plant Biologists, 150(1), pp. 105–113.
- Chen, Y.-L. *et al.* (2020) ‘The role of peptides cleaved from protein precursors in eliciting plant stress reactions’, *New Phytologist*. Wiley Online Library, 225(6), pp. 2267–2282.
- Chilley, P. M. *et al.* (2006) ‘The POLARIS peptide of *Arabidopsis* regulates auxin transport

and root growth via effects on ethylene signaling', *The Plant Cell*. American Society of Plant Biologists, 18(11), pp. 3058–3072.

Chung, K. P. and Zeng, Y. (2017) 'An overview of protein secretion in plant cells', *Plant Protein Secretion: Methods and Protocols*. Springer, pp. 19–32.

De Coninck, B. and De Smet, I. (2016) 'Plant peptides--taking them to the next level', *Journal of Experimental Botany*. Oxford University Press UK, pp. 4791–4795.

Considine, M. J. and Foyer, C. H. (2014) 'Redox regulation of plant development', *Antioxidants & redox signaling*. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, 21(9), pp. 1305–1326.

Coudert, Y. *et al.* (2015) 'Three ancient hormonal cues co-ordinate shoot branching in a moss', *Elife*. eLife Sciences Publications Limited, 4, p. e06808.

Cove, D. J. *et al.* (2009) 'Culturing the moss *Physcomitrella patens*', *Cold Spring Harbor Protocols*. Citeseer, 2009(2), p. pdb--prot5136.

Crum, H. A. (2001) *Structural diversity of bryophytes*. University of Michigan Press.

Daudi, A. and O'Brien, J. A. (2012) 'Detection of hydrogen peroxide by DAB staining in *Arabidopsis* leaves', *Bio-protocol*, 2(18), pp. e263--e263.

Delwiche, C. F. and Cooper, E. D. (2015) 'The evolutionary origin of a terrestrial flora', *Current Biology*. Elsevier, 25(19), pp. R899--R910.

Dice, J. F. (1990) 'Peptide sequences that target cytosolic proteins for lysosomal proteolysis', *Trends in biochemical sciences*. Elsevier, 15(8), pp. 305–309.

Ding, Z. *et al.* (2012) 'ER-localized auxin transporter PIN8 regulates auxin homeostasis and male gametophyte development in *Arabidopsis*', *Nature communications*. Nature Publishing Group UK London, 3(1), p. 941.

Doblas, V. G. *et al.* (2017) 'Root diffusion barrier control by a vasculature-derived peptide binding to the SGN3 receptor', *Science*. American Association for the Advancement of Science, 355(6322), pp. 280–284.

Donoghue, M. T. A. *et al.* (2011) 'Evolutionary origins of Brassicaceae specific genes in *Arabidopsis thaliana*', *BMC evolutionary biology*. Springer, 11(1), pp. 1–23.

Donoghue, P. C. J. *et al.* (2021) 'The evolutionary emergence of land plants', *Current Biology*. Elsevier, 31(19), pp. R1281--R1298.

Dunn, O. J. (1964) 'Multiple comparisons using rank sums', *Technometrics*. Taylor & Francis Group, 6(3), pp. 241–252.

Eckardt, N. A. (2010) 'Redox regulation of auxin signaling and plant development'. American Society of Plant Biologists.

Edwards, D., Axe, L. and Duckett, J. G. (2003) 'Diversity in conducting cells in early land plants and comparisons with extant bryophytes', *Botanical Journal of the Linnean Society*.

Oxford University Press, 141(3), pp. 297–347.

Eklund, D. M. *et al.* (2010) ‘Homologues of the *Arabidopsis thaliana* SHI/STY/LRP1 genes control auxin biosynthesis and affect growth and development in the moss *Physcomitrella patens*’, *Development*. Oxford University Press for The Company of Biologists Limited, 137(8), pp. 1275–1284.

Engineer, C. B. *et al.* (2014) ‘Carbonic anhydrases, EPF2 and a novel protease mediate CO₂ control of stomatal development’, *Nature*. Nature Publishing Group UK London, 513(7517), pp. 246–250.

Faulkner, C. (2018) ‘Plasmodesmata and the symplast’, *Current Biology*. Elsevier, 28(24), pp. R1374–R1378.

Fernandez, A. *et al.* (2013) ‘Transcriptional and functional classification of the GOLVEN/ROOT GROWTH FACTOR/CLE-like signaling peptides reveals their role in lateral root and hair formation’, *Plant physiology*. American Society of Plant Biologists, 161(2), pp. 954–970.

Fesenko, I. *et al.* (2019) ‘Phytohormone treatment induces generation of cryptic peptides with antimicrobial activity in the Moss *Physcomitrella patens*’, *BMC plant biology*. BioMed Central, 19(1), p. 9.

Fischer, C. *et al.* (1997) ‘Induction of zygotic polyembryos in wheat: influence of auxin polar transport.’, *The Plant Cell*. American Society of Plant Biologists, 9(10), pp. 1767–1780.

Fletcher, J. C. *et al.* (1999) ‘Signaling of cell fate decisions by CLAVATA3 in *Arabidopsis* shoot meristems’, *Science*. American Association for the Advancement of Science, 283(5409), pp. 1911–1914.

Fortelny, N. *et al.* (2015) ‘Proteome TopFIND 3.0 with TopFINDER and PathFINDER: database and analysis tools for the association of protein termini to pre-and post-translational events’, *Nucleic acids research*. Oxford University Press, 43(D1), pp. D290–D297.

Frangedakis, E. *et al.* (2021) ‘The hornworts: morphology, evolution and development’, *New Phytologist*. Wiley Online Library, 229(2), pp. 735–754.

Friml, J. and Jones, A. R. (2010) ‘Endoplasmic reticulum: the rising compartment in auxin biology’, *Plant physiology*. American Society of Plant Biologists, 154(2), pp. 458–462.

Fujita, T. *et al.* (2008) ‘Convergent evolution of shoots in land plants: lack of auxin polar transport in moss shoots’, *Evolution & development*. Wiley Online Library, 10(2), pp. 176–186.

Furumizu, C. *et al.* (2015) ‘Antagonistic roles for KNOX1 and KNOX2 genes in patterning the land plant body plan following an ancient gene duplication’, *PLoS Genetics*. Public Library of Science San Francisco, CA USA, 11(2), p. e1004980.

Furumizu, C. *et al.* (2021) ‘The sequenced genomes of nonflowering land plants reveal the innovative evolutionary history of peptide signaling’, *The Plant Cell*. Oxford University Press, 33(9), pp. 2915–2934.

- Galweiler, L. *et al.* (1998) 'Regulation of polar auxin transport by AtPIN1 in Arabidopsis vascular tissue', *Science*. American Association for the Advancement of Science, 282(5397), pp. 2226–2230.
- Gancheva, M. S. *et al.* (2019) 'Plant peptide hormones', *Russian Journal of Plant Physiology*. Springer, 66, pp. 171–189.
- Gao, C. *et al.* (2020) 'Directionality of plasmodesmata-mediated transport in Arabidopsis leaves supports auxin channeling', *Current Biology*. Elsevier, 30(10), pp. 1970–1977.
- Gayomba, S. R., Watkins, J. M. and Muday, G. K. (2017) 'Flavonols regulate plant growth and development through regulation of auxin transport and cellular redox status', *Recent Advances in Polyphenol Research*. Wiley Online Library, pp. 143–170.
- Geisler, M. *et al.* (2005) 'Cellular efflux of auxin catalyzed by the Arabidopsis MDR/PGP transporter AtPGP1', *The Plant Journal*. Wiley Online Library, 44(2), pp. 179–194.
- Gerrienne, P. and Genez, P. (2011) 'Early evolution of life cycles in embryophytes: a focus on the fossil evidence of gametophyte/sporophyte size and morphological complexity', *Journal of Systematics and Evolution*. Wiley Online Library, 49(1), pp. 1–16.
- Ghorbani, S. *et al.* (2016) 'The SBT6. 1 subtilase processes the GOLVEN1 peptide controlling cell elongation', *Journal of Experimental Botany*. Oxford University Press UK, 67(16), pp. 4877–4887.
- Ginanjar, E. F., Teh, O. K. and Fujita, T. (2021) 'Characterisation of Rapid Alkalinisation Factors (RAFLs) in *Physcomitrium patens* Reveals Functional Conservation in Tip Growth.', *The New Phytologist*.
- Goffinet, B., Buck, W. R. and Shaw, A. J. (2009) 'Morphology, anatomy, and classification of the Bryophyta', *Bryophyte biology*. Cambridge University Press Cambridge, 2, pp. 55–138.
- Gomes, G. L. B. and Scortecci, K. C. (2021) 'Auxin and its role in plant development: structure, signalling, regulation and response mechanisms', *Plant Biology*. Wiley Online Library, 23(6), pp. 894–904.
- Goodacre, N. F., Gerloff, D. L. and Uetz, P. (2014) 'Protein domains of unknown function are essential in bacteria', *MBio*. Am Soc Microbiol, 5(1), pp. e00744--13.
- Graham, L. E. (1985) 'The Origin of the Life Cycle of Land Plants: A simple modification in the life cycle of an extinct green alga is the likely origin of the first land plants', *American Scientist*. JSTOR, 73(2), pp. 178–186.
- Groves, M. R. and Barford, D. (1999) 'Topological characteristics of helical repeat protein', *Current opinion in structural biology*. Elsevier, 9(3), pp. 383–389.
- Guenot, B. *et al.* (2012) 'Pin1-independent leaf initiation in Arabidopsis', *Plant Physiology*. American Society of Plant Biologists, 159(4), pp. 1501–1510.
- Gupta, R. and Deswal, R. (2012) 'Low temperature stress modulated secretome analysis and purification of antifreeze protein from *Hippophae rhamnoides*, a Himalayan wonder plant', *Journal of Proteome Research*. ACS Publications, 11(5), pp. 2684–2696.

- Hammes, U. Z., Murphy, A. S. and Schwechheimer, C. (2022) ‘Auxin transporters—a biochemical view’, *Cold Spring Harbor Perspectives in Biology*. Cold Spring Harbor Lab, 14(2), p. a039875.
- Han, X. *et al.* (2014) ‘Auxin-callose-mediated plasmodesmal gating is essential for tropic auxin gradient formation and signaling’, *Developmental cell*. Elsevier, 28(2), pp. 132–146.
- Han, Y. *et al.* (2008) ‘Rice ROOT ARCHITECTURE ASSOCIATED1 binds the proteasome subunit RPT4 and is degraded in a D-box and proteasome-dependent manner’, *Plant physiology*. American Society of Plant Biologists, 148(2), pp. 843–855.
- Hanada, K. *et al.* (2008) ‘Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli’, *Plant physiology*. American Society of Plant Biologists, 148(2), pp. 993–1003.
- Hara, K. *et al.* (2007) ‘The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule’, *Genes & development*. Cold Spring Harbor Lab, 21(14), pp. 1720–1725.
- Harholt, J., Moestrup, Ø. and Ulvskov, P. (2016) ‘Why plants were terrestrial from the beginning’, *Trends in Plant Science*. Elsevier, 21(2), pp. 96–101.
- Harris, B. J. *et al.* (2020) ‘Phylogenomic evidence for the monophyly of bryophytes and the reductive evolution of stomata’, *Current Biology*. Elsevier, 30(11), pp. 2001–2012.
- Harris, B. J. *et al.* (2022) ‘Divergent evolutionary trajectories of bryophytes and tracheophytes from a complex common ancestor of land plants’, *Nature Ecology & Evolution*. Nature Publishing Group UK London, 6(11), pp. 1634–1643.
- Harrison, C. J. *et al.* (2009) ‘Local cues and asymmetric cell divisions underpin body plan transitions in the moss *Physcomitrella patens*’, *Current Biology*. Elsevier, 19(6), pp. 461–471.
- Harrison, C. J. (2017) ‘Auxin transport in the evolution of branching forms’, *New phytologist*. Wiley Online Library, 215(2), pp. 545–551.
- Hassoon, A. S. and Abduljabbar, I. A. (2019) ‘Review on the role of salicylic acid in plants’, *Sustainable crop production*. IntechOpen, pp. 61–64.
- Hazak, O. and Hardtke, C. S. (2016) ‘CLAVATA 1-type receptors in plant development’, *Journal of experimental botany*. Oxford University Press UK, 67(16), pp. 4827–4833.
- Hernández-Pinzón, I. *et al.* (2012) ‘The Tnt1 retrotransposon escapes silencing in tobacco, its natural host’, *PLoS One*. Public Library of Science San Francisco, USA, 7(3), p. e33816.
- Hieta, R. and Myllyharju, J. (2002) ‘Cloning and characterization of a low molecular weight prolyl 4-hydroxylase from *Arabidopsis thaliana*: effective hydroxylation of proline-rich, collagen-like, and hypoxia-inducible transcription factor α -like peptides’, *Journal of Biological Chemistry*. ASBMB, 277(26), pp. 23965–23971.
- Hohe, A. *et al.* (2002) ‘Day length and temperature strongly influence sexual reproduction and expression of a novel MADS-box gene in the moss *Physcomitrella patens*’, *Plant biology*. Georg Thieme Verlag Stuttgart·New York, 4(05), pp. 595–602.

- Horan, K. *et al.* (2008) ‘Annotating genes of known and unknown function by large-scale coexpression analysis’, *Plant physiology*. American Society of Plant Biologists, 147(1), pp. 41–57.
- Horst, N. A. *et al.* (2016) ‘A single homeobox gene triggers phase transition, embryogenesis and asexual reproduction’, *Nature Plants*. Nature Publishing Group, 2(2), pp. 1–6.
- Hou, S. *et al.* (2014) ‘The secreted peptide PIP1 amplifies immunity through receptor-like kinase 7’, *PLoS pathogens*. Public Library of Science San Francisco, USA, 10(9), p. e1004331.
- Huffaker, A., Pearce, G. and Ryan, C. A. (2006) ‘An endogenous peptide signal in Arabidopsis activates components of the innate immune response’, *Proceedings of the National Academy of Sciences*. National Acad Sciences, 103(26), pp. 10098–10103.
- Iglesias, V. A. and Meins Jr, F. (2000) ‘Movement of plant viruses is delayed in a β -1, 3-glucanase-deficient mutant showing a reduced plasmodesmatal size exclusion limit and enhanced callose deposition’, *The Plant Journal*. Wiley Online Library, 21(2), pp. 157–166.
- Ishizaki, K. (2017) ‘Evolution of land plants: insights from molecular studies on basal lineages’, *Bioscience, biotechnology, and biochemistry*. Taylor & Francis, 81(1), pp. 73–80.
- Ito, Y. *et al.* (2006) ‘Dodeca-CLE peptides as suppressors of plant stem cell differentiation’, *Science*. American Association for the Advancement of Science, 313(5788), pp. 842–845.
- Ivanchenko, M. G. *et al.* (2013) ‘Auxin increases the hydrogen peroxide (H₂O₂) concentration in tomato (*Solanum lycopersicum*) root tips while inhibiting root growth’, *Annals of Botany*. Oxford University Press, 112(6), pp. 1107–1116.
- Jacobs, M. and Rubery, P. H. (1988) ‘Naturally occurring auxin transport regulators’, *Science*. American Association for the Advancement of Science, 241(4863), pp. 346–349.
- Jaffe, M. J. and Leopold, A. C. (1984) ‘Callose deposition during gravitropism of *Zea mays* and *Pisum sativum* and its inhibition by 2-deoxy-D-glucose’, *Planta*. Springer, 161, pp. 20–26.
- Jain, B. P. and Pandey, S. (2018) ‘WD40 repeat proteins: signalling scaffold with diverse functions’, *The protein journal*. Springer, 37(5), pp. 391–406.
- Jang, G. *et al.* (2011) ‘RSL genes are sufficient for rhizoid system development in early diverging land plants’, *Development*. Oxford University Press for The Company of Biologists Limited, 138(11), pp. 2273–2281.
- Jasinski, S. *et al.* (2005) ‘KNOX action in Arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities’, *Current Biology*. Elsevier, 15(17), pp. 1560–1565.
- Jefferson, R. A. *et al.* (2009) ‘Local cues and asymmetric cell divisions underpin body plan transitions in the moss *Physcomitrella patens*’, *Development*. Vienna, Austria: Oxford University Press, 6(1), pp. 177–188. Available at: <https://www.r-project.org/>.
- Jing, H.-C. *et al.* (2007) ‘Arabidopsis CPR5 is a senescence-regulatory gene with pleiotropic functions as predicted by the evolutionary theory of senescence’, *Journal of experimental botany*. Oxford University Press, 58(14), pp. 3885–3894.

- Jing, Y. *et al.* (2019) ‘Danger-associated peptides interact with PIN-dependent local auxin distribution to inhibit root growth in Arabidopsis’, *The Plant Cell*. American Society of Plant Biologists, 31(8), pp. 1767–1787.
- Johri, M. M. and DESAI, S. (1973) ‘Auxin regulation of caulonema formation in moss protonema’, *Nature New Biology*. Nature Publishing Group UK London, 245(146), pp. 223–224.
- Juárez, S. P. D., Mangano, S. and Estevez, J. M. (2015) ‘Improved ROS measurement in root hair cells’, *Plant Cell Expansion: Methods and Protocols*. Springer, pp. 67–71.
- Jumper, J. *et al.* (2021) ‘Highly accurate protein structure prediction with AlphaFold’, *Nature*. Nature Publishing Group, 596(7873), pp. 583–589.
- Kajava, A. V (2012) ‘Tandem repeats in proteins: from sequence to structure’, *Journal of structural biology*. Elsevier, 179(3), pp. 279–288.
- Kamimoto, Y. *et al.* (2012) ‘Arabidopsis ABCB21 is a facultative auxin importer/exporter regulated by cytoplasmic auxin concentration’, *Plant and Cell Physiology*. Oxford University Press, 53(12), pp. 2090–2100.
- Katsir, L. *et al.* (2011) ‘Peptide signaling in plant development’, *Current Biology*. Elsevier, 21(9), pp. R356--R364.
- Katti, M. V, Ranjekar, P. K. and Gupta, V. S. (2001) ‘Differential distribution of simple sequence repeats in eukaryotic genome sequences’, *Molecular biology and evolution*. Oxford University Press, 18(7), pp. 1161–1167.
- Kaur, S. J. *et al.* (2012) ‘TolC-dependent secretion of an ankyrin repeat-containing protein of Rickettsia typhi’, *Journal of bacteriology*. Am Soc Microbiol, 194(18), pp. 4920–4932.
- Kelkar, D. S. *et al.* (2019) ‘A chemical--genetic screen identifies ABHD12 as an oxidized-phosphatidylserine lipase’, *Nature chemical biology*. Nature Publishing Group, 15(2), pp. 169–178.
- Kieber, J. J. and Schaller, G. E. (2014) ‘Cytokinins’, *The Arabidopsis Book/American Society of Plant Biologists*. American Society of Plant Biologists, 12.
- Kitagawa, M. and Fujita, T. (2013) ‘Quantitative imaging of directional transport through plasmodesmata in moss protonemata via single-cell photoconversion of Dendra2’, *Journal of plant research*. Springer, 126(4), pp. 577–585.
- Kloss, E. and Barrick, D. (2009) ‘C-terminal deletion of leucine-rich repeats from YopM reveals a heterogeneous distribution of stability in a cooperatively folded protein’, *Protein Science*. Wiley Online Library, 18(9), pp. 1948–1960.
- Kobiler, O. *et al.* (2005) ‘Quantitative kinetic analysis of the bacteriophage λ genetic network’, *Proceedings of the National Academy of Sciences*. National Acad Sciences, 102(12), pp. 4470–4475.
- Kofuji, R. and Hasebe, M. (2014) ‘Eight types of stem cells in the life cycle of the moss *Physcomitrella patens*’, *Current opinion in plant biology*. Elsevier, 17, pp. 13–21.

- Koide, Y. *et al.* (2018) 'Lineage-specific gene acquisition or loss is involved in interspecific hybrid sterility in rice', *Proceedings of the National Academy of Sciences*. National Acad Sciences, 115(9), pp. E1955--E1962.
- Koie, M. *et al.* (2014) 'Cleavage within Reelin repeat 3 regulates the duration and range of the signaling activity of Reelin protein', *Journal of Biological Chemistry*. ASBMB, 289(18), pp. 12922–12930.
- Komori, R. *et al.* (2009) 'Identification of tyrosylprotein sulfotransferase in Arabidopsis', *Proceedings of the National Academy of Sciences*. National Acad Sciences, 106(35), pp. 15067–15072.
- Konozy, E. H. E. *et al.* (2013) 'Proteomic analysis of tomato (*Solanum lycopersicum*) secretome', *Journal of plant research*. Springer, 126, pp. 251–266.
- Koprivova, A., Mugford, S. T. and Kopriva, S. (2010) 'Arabidopsis root growth dependence on glutathione is linked to auxin transport', *Plant cell reports*. Springer, 29, pp. 1157–1167.
- Koshy, C. M. *et al.* (no date) 'In silico Structural and Functional Characterization of a Hypothetical Protein from *Stenotrophomonas maltophilia* SRM01'.
- Kovtun, Y. *et al.* (2000) 'Functional analysis of oxidative stress-activated mitogen-activated protein kinase cascade in plants', *Proceedings of the National Academy of Sciences*. National Acad Sciences, 97(6), pp. 2940–2945.
- Kramer, E. M. (2004) 'PIN and AUX/LAX proteins: their role in auxin accumulation', *Trends in plant science*. Elsevier, 9(12), pp. 578–582.
- Kramer, E. M., Rutschow, H. L. and Mabie, S. S. (2011) 'AuxV: a database of auxin transport velocities', *Trends in plant science*. Elsevier, 16(9), pp. 461–463.
- Krouk, G. *et al.* (2010) 'Nitrate-regulated auxin transport by NRT1. 1 defines a mechanism for nutrient sensing in plants', *Developmental cell*. Elsevier, 18(6), pp. 927–937.
- Kyte, J. and Doolittle, R. F. (1982) 'A simple method for displaying the hydropathic character of a protein', *Journal of molecular biology*. Elsevier, 157(1), pp. 105–132.
- Lang, D. *et al.* (2005) 'Representation and high-quality annotation of the *Physcomitrella patens* transcriptome demonstrates a high proportion of proteins involved in metabolism in mosses', *Plant Biology*. Wiley Online Library, 7(3), pp. 238–250.
- Lavy, M. *et al.* (2012) 'The cyclophilin DIAGEOTROPICA has a conserved role in auxin signaling', *Development*. Company of Biologists, 139(6), pp. 1115–1124.
- Lavy, M. *et al.* (2016) 'Constitutive auxin response in *Physcomitrella* reveals complex interactions between Aux/IAA and ARF proteins', *Elife*. eLife Sciences Publications, Ltd, 5, p. e13325.
- Lee, J.-Y. *et al.* (2011) 'A plasmodesmata-localized protein mediates crosstalk between cell-to-cell communication and innate immunity in Arabidopsis', *The Plant Cell*. Am Soc Plant Biol, 23(9), pp. 3353–3373.

- Lehtonen, M. T. *et al.* (2014) ‘Protein secretome of moss plants (*Physcomitrella patens*) with emphasis on changes induced by a fungal elicitor’, *Journal of Proteome Research*. ACS Publications, 13(2), pp. 447–459.
- Lespinet, O. *et al.* (2002) ‘The role of lineage-specific gene family expansion in the evolution of eukaryotes’, *Genome research*. Cold Spring Harbor Lab, 12(7), pp. 1048–1059.
- Levene, H. (1960) ‘Robust tests for equality of variances. Palo Alto’. CA: Stanford University Press.
- Levy, A. *et al.* (2007) ‘A plasmodesmata-associated β -1, 3-glucanase in *Arabidopsis*’, *The Plant Journal*. Wiley Online Library, 49(4), pp. 669–682.
- Leyser, O. (2018) ‘Auxin signaling’, *Plant physiology*. American Society of Plant Biologists, 176(1), pp. 465–479.
- Li, L. *et al.* (2009) ‘Identification of the novel protein QQS as a component of the starch metabolic network in *Arabidopsis* leaves’, *The Plant Journal*. Wiley Online Library, 58(3), pp. 485–498.
- Li, L. and Wurtele, E. S. (2015) ‘The QQS orphan gene of *Arabidopsis* modulates carbon and nitrogen allocation in soybean’, *Plant biotechnology journal*. Wiley Online Library, 13(2), pp. 177–187.
- Li, Z. P. *et al.* (2021) ‘Intercellular trafficking via plasmodesmata: molecular layers of complexity’, *Cellular and Molecular Life Sciences*. Springer, 78(3), pp. 799–816.
- Little, J. W. (2005) ‘Threshold effects in gene regulation: When some is not enough’, *Proceedings of the National Academy of Sciences*. National Acad Sciences, 102(15), pp. 5310–5311.
- Liu, C. M., Xu, Z. H. and Chua, N. (1993) ‘Auxin polar transport is essential for the establishment of bilateral symmetry during early plant embryogenesis.’, *The Plant Cell*. American Society of Plant Biologists, 5(6), pp. 621–630.
- Liu, Y.-C. and Vidali, L. (2011) ‘Efficient polyethylene glycol (PEG) mediated transformation of the moss *Physcomitrella patens*’, *JoVE (Journal of Visualized Experiments)*, (50), p. e2560.
- Liu, Y. *et al.* (2017) ‘Symplastic communication spatially directs local auxin biosynthesis to maintain root stem cell niche in *Arabidopsis*’, *Proceedings of the National Academy of Sciences*. National Acad Sciences, 114(15), pp. 4005–4010.
- Lu, S. *et al.* (2021) ‘Integrative analyses of metabolomes and transcriptomes provide insights into flavonoid variation in grape berries’, *Journal of Agricultural and Food Chemistry*. ACS Publications, 69(41), pp. 12354–12367.
- Luhua, S. *et al.* (2013) ‘Linking genes of unknown function with abiotic stress responses by high-throughput phenotype screening’, *Physiologia plantarum*. Wiley Online Library, 148(3), pp. 322–333.
- Luna, S. K. and Chain, F. J. J. (2021) ‘Lineage-specific genes and family expansions in dictyostelid genomes display expression bias and evolutionary diversification during

development', *Genes*. MDPI, 12(10), p. 1628.

Lüttge, U. (2020) 'Terrestrialization: the conquest of dry land by plants', in *Progress in Botany Vol. 83*. Springer, pp. 65–89.

Lv, P. *et al.* (2023) 'Unraveling the Diverse Roles of Neglected Genes Containing Domains of Unknown Function (DUFs): Progress and Perspective', *International Journal of Molecular Sciences*. MDPI, 24(4), p. 4187.

Ma, S. *et al.* (2020) 'Identification, characterization and expression analysis of lineage-specific genes within Triticeae', *Genomics*. Elsevier, 112(2), pp. 1343–1350.

Mamaeva, A. *et al.* (2022) 'RALF peptides modulate immune response in the moss *Physcomitrium patens*', *bioRxiv*. Cold Spring Harbor Laboratory.

Mangano, S. *et al.* (2017) 'Molecular link between auxin and ROS-mediated polar growth', *Proceedings of the National Academy of Sciences*. National Acad Sciences, 114(20), pp. 5289–5294.

MARIGO, G. and BOUDET, A. M. (1977) 'Relations polyphénols—croissance: Mise en évidence d'un effet inhibiteur des composés phénoliques sur le transport polarisé de l'auxine', *Physiologia Plantarum*. Wiley Online Library, 41(3), pp. 197–202.

Matsubayashi, Y. (2014) 'Posttranslationally modified small-peptide signals in plants', *Annual review of plant biology*. Annual Reviews, 65, pp. 385–413.

Matsubayashi, Y. and Sakagami, Y. (1996) 'Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L.', *Proceedings of the National Academy of Sciences*. National Acad Sciences, 93(15), pp. 7623–7627.

Matsubayashi, Y. and Sakagami, Y. (2006) 'Peptide hormones in plants', *Annu. Rev. Plant Biol.* Annual Reviews, 57, pp. 649–674.

Matsuzaki, Y. *et al.* (2010) 'Secreted peptide signals required for maintenance of root stem cell niche in *Arabidopsis*', *Science*. American Association for the Advancement of Science, 329(5995), pp. 1065–1067.

McDaniel, S. F. (2021) 'Bryophytes are not early diverging land plants', *New Phytologist*. Wiley Online Library, 230(4), pp. 1300–1304.

McKight, P. E. and Najab, J. (2010) 'Kruskal-wallis test', *The corsini encyclopedia of psychology*. Wiley Online Library, p. 1.

Mellor, N. L. *et al.* (2020) 'Auxin fluxes through plasmodesmata modify root-tip auxin distribution', *Development*. The Company of Biologists Ltd, 147(6), p. dev181669.

Meng, L. *et al.* (2012) 'A putative nuclear CLE-like (CLEL) peptide precursor regulates root growth in *Arabidopsis*', *Molecular plant*. Elsevier, 5(4), pp. 955–957.

Mentzen, W. I. and Wurtele, E. S. (2008) 'Regulon organization of *Arabidopsis*', *BMC plant biology*. BioMed Central, 8(1), pp. 1–22.

- Michniewicz, M. *et al.* (2007) 'Antagonistic regulation of PIN phosphorylation by PP2A and PINOID directs auxin flux', *Cell*. Elsevier, 130(6), pp. 1044–1056.
- Michniewicz, M., Brewer, P. B. and Friml, J. (2007) 'Polar auxin transport and asymmetric auxin distribution', *The Arabidopsis Book/American Society of Plant Biologists*. American Society of Plant Biologists, 5.
- Mirdita, M. *et al.* (2022) 'ColabFold: making protein folding accessible to all', *Nature Methods*. Nature Publishing Group, pp. 1–4.
- Mitchison, G. J. (1980) 'The dynamics of auxin transport', *Proceedings of the Royal Society of London. Series B. Biological Sciences*. The Royal Society London, 209(1177), pp. 489–511.
- Miyakawa, T. *et al.* (2014) 'A secreted protein with plant-specific cysteine-rich motif functions as a mannose-binding lectin that exhibits antifungal activity', *Plant physiology*. American Society of Plant Biologists, 166(2), pp. 766–778.
- Mohanasundaram, B. *et al.* (2021) 'The unique bryophyte-specific repeat-containing protein SHORT-LEAF regulates gametophore development in moss'. Oxford University Press.
- Mravec, J. *et al.* (2009) 'Subcellular homeostasis of phytohormone auxin is mediated by the ER-localized PIN5 transporter', *Nature*. Nature Publishing Group UK London, 459(7250), pp. 1136–1140.
- Muday, G. K. and Murphy, A. S. (2002) 'An emerging model of auxin transport regulation', *The Plant Cell*. American Society of Plant Biologists, 14(2), pp. 293–299.
- Murphy, A., Peer, W. A. and Taiz, L. (2000) 'Regulation of auxin transport by aminopeptidases and endogenous flavonoids', *Planta*. Springer, 211(3), pp. 315–324.
- Nakayama, T. *et al.* (2017) 'A peptide hormone required for Casparian strip diffusion barrier formation in Arabidopsis roots', *Science*. American Association for the Advancement of Science, 355(6322), pp. 284–286.
- Naoumkina, M. A. *et al.* (2010) 'Genome-wide analysis of phenylpropanoid defence pathways', *Molecular plant pathology*. Wiley Online Library, 11(6), pp. 829–846.
- Nascimento, L. B. dos S. and Tattini, M. (2022) 'Beyond photoprotection: The multifarious roles of flavonoids in plant terrestrialization', *International Journal of Molecular Sciences*. MDPI, 23(9), p. 5284.
- Nemec-Venza, Z. *et al.* (2022) 'CLAVATA modulates auxin homeostasis and transport to regulate stem cell identity and plant shape in a moss', *New Phytologist*. Wiley Online Library, 234(1), pp. 149–163.
- Nishiyama, T. *et al.* (2000) 'Tagged mutagenesis and gene-trap in the moss, *Physcomitrella patens* by shuttle mutagenesis', *DNA research*. Oxford University Press, 7(1), pp. 9–17.
- Nishiyama, T. *et al.* (2004) 'Chloroplast phylogeny indicates that bryophytes are monophyletic', *Molecular Biology and Evolution*. Oxford University Press, 21(10), pp. 1813–1819.

- Noh, B., Murphy, A. S. and Spalding, E. P. (2001) 'Multidrug resistance--like genes of Arabidopsis required for auxin transport and auxin-mediated development', *The Plant Cell*. American Society of Plant Biologists, 13(11), pp. 2441–2454.
- Nolan, T. M. *et al.* (2020) 'Brassinosteroids: multidimensional regulators of plant growth, development, and stress responses', *The Plant Cell*. American Society of Plant Biologists, 32(2), pp. 295–318.
- Nonet, M., Sweetser, D. and Young, R. A. (1987) 'Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II', *Cell*. Elsevier, 50(6), pp. 909–915.
- Nyman, L. P. and Cutter, E. G. (1981) 'Auxin--cytokinin interaction in the inhibition, release, and morphology of gametophore buds of *Plagiomnium cuspidatum* from apical dominance', *Canadian Journal of Botany*. NRC Research Press Ottawa, Canada, 59(5), pp. 750–762.
- Ofori, P. A. *et al.* (2018) 'Tomato ATP-binding cassette transporter SlABC4 is involved in auxin transport in the developing fruit', *Plants*. MDPI, 7(3), p. 65.
- Ogawa-Ohnishi, M., Matsushita, W. and Matsubayashi, Y. (2013) 'Identification of three hydroxyproline O-arabinosyltransferases in *Arabidopsis thaliana*', *Nature chemical biology*. Nature Publishing Group US New York, 9(11), pp. 726–730.
- Ohyama, K. *et al.* (2009) 'A glycopeptide regulating stem cell fate in *Arabidopsis thaliana*', *Nature chemical biology*. Nature Publishing Group US New York, 5(8), pp. 578–580.
- Ohyama, K., Ogawa, M. and Matsubayashi, Y. (2008) 'Identification of a biologically active, small, secreted peptide in *Arabidopsis* by in silico gene screening, followed by LC-MS-based structure analysis', *The Plant Journal*. Wiley Online Library, 55(1), pp. 152–160.
- Okada, K. *et al.* (1991) 'Requirement of the auxin polar transport system in early stages of *Arabidopsis* floral bud formation.', *The Plant Cell*. American Society of Plant Biologists, 3(7), pp. 677–684.
- Okamoto, S. *et al.* (2013) 'Root-derived CLE glycopeptides control nodulation by direct binding to HAR1 receptor kinase', *Nature communications*. Nature Publishing Group UK London, 4(1), p. 2191.
- Paterlini, A. (2020) 'Uncharted routes: exploring the relevance of auxin movement via plasmodesmata', *Biology Open*. The Company of Biologists Ltd, 9(11), p. bio055541.
- Patro, R. *et al.* (2017) 'Salmon provides fast and bias-aware quantification of transcript expression', *Nature methods*. Nature Publishing Group, 14(4), pp. 417–419.
- Pearce, G. *et al.* (1991) 'A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins', *Science*. American Association for the Advancement of Science, 253(5022), pp. 895–897.
- Pearce, G. *et al.* (2001) 'RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development', *Proceedings of the National Academy of Sciences*. National Acad Sciences, 98(22), pp. 12843–12847.
- Peer, W. A. *et al.* (2004) 'Variation in expression and protein localization of the PIN family of

- auxin efflux facilitator proteins in flavonoid mutants with altered auxin transport in *Arabidopsis thaliana*', *The Plant Cell*. American Society of Plant Biologists, 16(7), pp. 1898–1911.
- Peer, W. A., Cheng, Y. and Murphy, A. S. (2013) 'Evidence of oxidative attenuation of auxin signalling', *Journal of experimental botany*. Oxford University Press UK, 64(9), pp. 2629–2639.
- Peer, W. A. and Murphy, A. S. (2006) 'The science of flavonoids', *Flavonoids as signal molecules: targets of flavonoid action*. New York: Springer, pp. 239–268.
- Peer, W. A. and Murphy, A. S. (2007) 'Flavonoids and auxin transport: modulators or regulators?', *Trends in plant science*. Elsevier, 12(12), pp. 556–563.
- Petrasek, J. and Friml, J. (2009) 'Auxin transport routes in plant development'. Oxford University Press for The Company of Biologists Limited.
- Pharis, R. P. and Rood, S. B. (2012) *Plant growth substances 1988*. Springer Science & Business Media.
- Pietta, P.-G. (2000) 'Flavonoids as antioxidants', *Journal of natural products*. ACS Publications, 63(7), pp. 1035–1042.
- Pinedo, M. *et al.* (2012) 'Extracellular sunflower proteins: evidence on non-classical secretion of a jacalin-related lectin', *Protein and peptide letters*. Bentham Science Publishers, 19(3), pp. 270–276.
- Pires, N. D. *et al.* (2013) 'Recruitment and remodeling of an ancient gene regulatory network during land plant evolution', *Proceedings of the National Academy of Sciences*. National Acad Sciences, 110(23), pp. 9571–9576.
- Poli, D., Jacobs, M. and Cooke, T. J. (2003) 'Auxin regulation of axial growth in bryophyte sporophytes: its potential significance for the evolution of early land plants', *American Journal of Botany*. Wiley Online Library, 90(10), pp. 1405–1415.
- Prigge, M. J. *et al.* (2010) 'Physcomitrella patens auxin-resistant mutants affect conserved elements of an auxin-signaling pathway', *Current Biology*. Elsevier, 20(21), pp. 1907–1912.
- Proctor, M. C. F. (2000) 'Mosses and alternative adaptation to life on land', *The New Phytologist*. Cambridge University Press, 148(1), pp. 1–6.
- Puopolo, K. M. *et al.* (2001) 'Tandem repeat deletion in the alpha C protein of group B streptococcus is recA independent', *Infection and immunity*. Am Soc Microbiol, 69(8), pp. 5037–5045.
- Rajathej, D. M., Parthasarathy, S. and Selvaraj, S. (2019) 'Identification and analysis of long repeats of proteins at the domain level', *Frontiers in bioengineering and biotechnology*. Frontiers Media SA, 7, p. 250.
- Ranocha, P. *et al.* (2013) 'Arabidopsis WAT1 is a vacuolar auxin transport facilitator required for auxin homeostasis', *Nature communications*. Nature Publishing Group UK London, 4(1), p. 2625.

- Reinhardt, D., Mandel, T. and Kuhlemeier, C. (2000) 'Auxin regulates the initiation and radial position of plant lateral organs', *The Plant Cell*. American Society of Plant Biologists, 12(4), pp. 507–518.
- Reiser, L., Sánchez-Baracaldo, P. and Hake, S. (2000) 'Knots in the family tree: evolutionary relationships and functions of knox homeobox genes', *Plant Molecular Evolution*. Springer, pp. 151–166.
- Rensing, S. A. *et al.* (2008) 'The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants', *Science*. American Association for the Advancement of Science, 319(5859), pp. 64–69.
- Rensing, S. A. (2020) 'How plants conquered land', *Cell*. Elsevier, 181(5), pp. 964–966.
- Renzaglia, K. S. *et al.* (2009) 'New insights into morphology, anatomy, and systematics of hornworts', *Bryophyte biology*. Cambridge University Press Cambridge, 2, pp. 139–171.
- Renzaglia, K. S. and Garbary, D. J. (2001) 'Motile gametes of land plants: diversity, development, and evolution', *Critical reviews in plant sciences*. Taylor & Francis, 20(2), pp. 107–213.
- Reski, R. and Abel, W. O. (1985) 'Induction of budding on chloronemata and caulonemata of the moss, *Physcomitrella patens*, using isopentenyladenine', *Planta*. Springer, 165(3), pp. 354–358.
- Robards, A. W. and Lucas, W. J. (1990) 'Plasmodesmata', *Annual review of plant biology*. Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, 41(1), pp. 369–419.
- Romani, F. and Moreno, J. E. (2021) 'Molecular mechanisms involved in functional macroevolution of plant transcription factors', *New Phytologist*. Wiley Online Library, 230(4), pp. 1345–1353.
- Rothberg, J. M. *et al.* (1990) 'slit: an extracellular protein necessary for development of midline glia and commissural axon pathways contains both EGF and LRR domains.', *Genes & development*. Cold Spring Harbor Lab, 4(12a), pp. 2169–2187.
- Rothman, J. E. (2014) 'The principle of membrane fusion in the cell (Nobel lecture).', *Angewandte Chemie (International Ed. in English)*, 53(47), pp. 12676–12694.
- Roy, S. *et al.* (2018) 'Small and mighty: peptide hormones in plant biology', *Plant Cell*, 30(tc118tt0718).
- Ruegger, M. *et al.* (1997) 'Reduced naphthylphthalamic acid binding in the tir3 mutant of *Arabidopsis* is associated with a reduction in polar auxin transport and diverse morphological defects.', *The Plant Cell*. American Society of Plant Biologists, 9(5), pp. 745–757.
- Rutschow, H. L., Baskin, T. I. and Kramer, E. M. (2011) 'Regulation of solute flux through plasmodesmata in the root meristem', *Plant physiology*. American Society of Plant Biologists, 155(4), pp. 1817–1826.
- Rutschow, H. L., Baskin, T. I. and Kramer, E. M. (2014) 'The carrier AUXIN RESISTANT

(AUX1) dominates auxin flux into Arabidopsis protoplasts', *New Phytologist*. Wiley Online Library, 204(3), pp. 536–544.

Ryan, C. P. (2019) 'Tandem repeat disorders', *Evolution, Medicine, and Public Health*. Oxford University Press, 2019(1), p. 17.

Sager, R. E. and Lee, J.-Y. (2018) 'Plasmodesmata at a glance', *Journal of cell science*. The Company of Biologists Ltd, 131(11), p. jcs209346.

Sakakibara, K. *et al.* (2003) 'Involvement of auxin and a homeodomain-leucine zipper I gene in rhizoid development of the moss *Physcomitrella patens*', *Development*. The Company of Biologists Ltd, 130(20), pp. 4835–4846.

Sakakibara, K. *et al.* (2008) 'Class 1 KNOX genes are not involved in shoot development in the moss *Physcomitrella patens* but do function in sporophyte development', *Evolution & development*. Wiley Online Library, 10(5), pp. 555–566.

Sakakibara, K. *et al.* (2013) 'KNOX2 genes regulate the haploid-to-diploid morphological transition in land plants', *Science*. American Association for the Advancement of Science, 339(6123), pp. 1067–1070.

Sakakibara, K. *et al.* (2014) 'WOX13-like genes are required for reprogramming of leaf and protoplast cells into stem cells in the moss *Physcomitrella patens*', *Development*. Oxford University Press for The Company of Biologists Limited, 141(8), pp. 1660–1670.

Sakata, Y., Komatsu, K. and Takezawa, D. (2014) 'ABA as a universal plant hormone', *Progress in Botany: Vol. 75*. Springer, pp. 57–96.

Sakurai, N. (1998) 'Dynamic function and regulation of apoplast in the plant body', *Journal of Plant Research*. Springer, 111, pp. 133–148.

Santelia, D. *et al.* (2005) 'MDR-like ABC transporter AtPGP4 is involved in auxin-mediated lateral root and root hair development', *FEBS letters*. Elsevier, 579(24), pp. 5399–5406.

Santner, A. and Estelle, M. (2009) 'Recent advances and emerging trends in plant hormone signalling', *Nature*. Nature Publishing Group UK London, 459(7250), pp. 1071–1078.

Saze, H. and Kakutani, T. (2011) 'Differentiation of epigenetic modifications between transposons and genes', *Current opinion in plant biology*. Elsevier, 14(1), pp. 81–87.

Schaper, E. and Anisimova, M. (2015) 'The evolution and function of protein tandem repeats in plants', *New Phytologist*. Wiley Online Library, 206(1), pp. 397–410.

Schardon, K. *et al.* (2016) 'Precursor processing for plant peptide hormone maturation by subtilisin-like serine proteinases', *Science*. American Association for the Advancement of Science, 354(6319), pp. 1594–1597.

Schmittgen, T. D. and Livak, K. J. (2008) 'Analyzing real-time PCR data by the comparative C T method', *Nature protocols*. Nature Publishing Group, 3(6), p. 1101.

Schoelmerich, M. C. *et al.* (2023) 'Tandem repeats in giant archaeal Borg elements undergo rapid evolution and create new intrinsically disordered regions in proteins', *Plos Biology*.

Public Library of Science San Francisco, CA USA, 21(1), p. e3001980.

Schwuchow, J., Michalke, W. and Hertel, R. (2001) 'An auxin transport inhibitor interferes with unicellular gravitropism in protonemata of the moss *Ceratodon purpureus*', *Plant Biology*. Georg Thieme Verlag Stuttgart·New York, 3(04), pp. 357–363.

Sevilem, I., Yadav, S. R. and Helariutta, Y. (2015) 'Plasmodesmata: channels for intercellular signaling during plant growth and development', *Plasmodesmata: methods and protocols*. Springer, pp. 3–24.

Sharma, A. *et al.* (2020) 'Primary transcript of miR858 encodes regulatory peptide and controls flavonoid biosynthesis and development in *Arabidopsis*', *Nature Plants*. Nature Publishing Group UK London, 6(10), pp. 1262–1274.

Sharma, M. and Pandey, G. K. (2016) 'Expansion and function of repeat domain proteins during stress and development in plants', *Frontiers in plant science*. Frontiers Media SA, 6, p. 1218.

Sharma, S. *et al.* (2016) 'The functional human C-terminome', *PLoS One*. Public Library of Science San Francisco, CA USA, 11(4), p. e0152731.

Sharma, S. and Schiller, M. R. (2019) 'The carboxy-terminus, a key regulator of protein function', *Critical reviews in biochemistry and molecular biology*. Taylor & Francis, 54(2), pp. 85–102.

Simon, S. *et al.* (2016) 'PIN6 auxin transporter at endoplasmic reticulum and plasma membrane mediates auxin homeostasis and organogenesis in *Arabidopsis*', *New Phytologist*. Wiley Online Library, 211(1), pp. 65–74.

Simpson, C. *et al.* (2009) 'An *Arabidopsis* GPI-anchor plasmodesmal neck protein with callose binding activity and potential to regulate cell-to-cell trafficking', *The Plant Cell*. Am Soc Plant Biol, 21(2), pp. 581–594.

Skokan, R. *et al.* (2019) 'PIN-driven auxin transport emerged early in streptophyte evolution', *Nature Plants*. Nature Publishing Group UK London, 5(11), pp. 1114–1119.

Šmídová, M., Hola, M. and Angelis, K. J. (2010) 'Efficient biolistic transformation of the moss *Physcomitrella patens*', *Biologia plantarum*. Springer, 54, pp. 777–780.

Smith, H. M. S. and Hake, S. (2003) 'The interaction of two homeobox genes, BREVIPEDICELLUS and PENNYWISE, regulates internode patterning in the *Arabidopsis* inflorescence', *The Plant Cell*. American Society of Plant Biologists, 15(8), pp. 1717–1727.

Sood, S. and Hackenberg, D. (1979) 'Interaction of auxin, antiauxin and cytokinin in relation to the formation of buds in moss protonema', *Zeitschrift für Pflanzenphysiologie*. Elsevier, 91(5), pp. 385–397.

Steenackers, W. *et al.* (2017) 'cis-Cinnamic acid is a novel, natural auxin efflux inhibitor that promotes lateral root formation', *Plant Physiology*. American Society of Plant Biologists, 173(1), pp. 552–565.

Steinmann, T. *et al.* (1999) 'Coordinated polar localization of auxin efflux carrier PIN1 by

GNOM ARF GEF', *Science*. American Association for the Advancement of Science, 286(5438), pp. 316–318.

STENLID, G. (1976) 'Effects of flavonoids on the polar transport of auxins', *Physiologia Plantarum*. Wiley Online Library, 38(4), pp. 262–266.

Stevenson, S. R. *et al.* (2016) 'Genetic analysis of *Physcomitrella patens* identifies ABSCISIC ACID NON-RESPONSIVE, a regulator of ABA responses unique to basal land plants and required for desiccation tolerance', *The Plant Cell*. Am Soc Plant Biol, 28(6), pp. 1310–1327.

Su, D. *et al.* (2021) 'Large-scale phylogenomic analyses reveal the monophyly of bryophytes and Neoproterozoic origin of land plants', *Molecular Biology and Evolution*. Oxford University Press, 38(8), pp. 3332–3344.

Swarup, R. and Bennett, M. (2003) 'Auxin transport: the fountain of life in plants?', *Developmental cell*. Elsevier, 5(6), pp. 824–826.

Tamaki, T. *et al.* (2013) 'SUPPRESSOR OF LLP 1 1-mediated C-terminal processing is critical for CLE 19 peptide activity', *The Plant Journal*. Wiley Online Library, 76(6), pp. 970–981.

Tameshige, T. *et al.* (2016) 'A secreted peptide and its receptors shape the auxin response pattern and leaf margin morphogenesis', *Current Biology*. Elsevier, 26(18), pp. 2478–2485.

Tan, Q. W. *et al.* (2023) 'Cross-stress gene expression atlas of *Marchantia polymorpha* reveals the hierarchy and regulatory principles of abiotic stress responses', *Nature Communications*. Nature Publishing Group UK London, 14(1), p. 986.

Tanaka, H. *et al.* (2006) 'Spatiotemporal asymmetric auxin distribution: a means to coordinate plant development', *Cellular and Molecular Life Sciences CMLS*. Springer, 63, pp. 2738–2754.

Tanco, S., Gevaert, K. and Van Damme, P. (2015) 'C-terminomics: Targeted analysis of natural and posttranslationally modified protein and peptide C-termini', *Proteomics*. Wiley Online Library, 15(5–6), pp. 903–914.

Tautz, D. and Domazet-Lošo, T. (2011) 'The evolutionary origin of orphan genes', *Nature Reviews Genetics*. Nature Publishing Group, 12(10), pp. 692–702.

Tavormina, P. *et al.* (2015) 'The plant peptidome: an expanding repertoire of structural features and biological functions', *The Plant Cell*. American Society of Plant Biologists, 27(8), pp. 2095–2118.

Teale, W. D. *et al.* (2021) 'Flavonol-mediated stabilization of PIN efflux complexes regulates polar auxin transport', *The EMBO journal*, 40(1), p. e104416.

Terasaka, K. *et al.* (2005) 'PGP4, an ATP binding cassette P-glycoprotein, catalyzes auxin transport in *Arabidopsis thaliana* roots', *The Plant Cell*. American Society of Plant Biologists, 17(11), pp. 2922–2939.

Thelander, M., Landberg, K. and Sundberg, E. (2018) 'Auxin-mediated developmental control in the moss *Physcomitrella patens*', *Journal of Experimental Botany*. Oxford University Press UK, 69(2), pp. 277–290.

- Tiainen, P., Myllyharju, J. and Koivunen, P. (2005) 'Characterization of a second *Arabidopsis thaliana* prolyl 4-hydroxylase with distinct substrate specificity', *Journal of Biological Chemistry*. ASBMB, 280(2), pp. 1142–1148.
- Tognetti, V. B., Bielach, A. and Hrtyan, M. (2017) 'Redox regulation at the site of primary growth: auxin, cytokinin and ROS crosstalk', *Plant, Cell & Environment*. Wiley Online Library, 40(11), pp. 2586–2605.
- Tost, A. S. *et al.* (2021) 'The PSY peptide family—expression, modification and physiological implications', *Genes*. MDPI, 12(2), p. 218.
- Toyokura, K. *et al.* (2019) 'Lateral inhibition by a peptide hormone-receptor cascade during *Arabidopsis* lateral root founder cell formation', *Developmental cell*. Elsevier, 48(1), pp. 64–75.
- Tripp, K. W. and Barrick, D. (2004) 'The tolerance of a modular protein to duplication and deletion of internal repeats', *Journal of molecular biology*. Elsevier, 344(1), pp. 169–178.
- Tukey, J. W. (1949) 'Comparing individual means in the analysis of variance', *Biometrics*. JSTOR, pp. 99–114.
- Vanderpoorten, A. and Goffinet, B. (2009) *Introduction to bryophytes*. Cambridge University Press.
- Vatén, A. *et al.* (2011) 'Callose biosynthesis regulates symplastic trafficking during root development', *Developmental cell*. Elsevier, 21(6), pp. 1144–1155.
- Verna, C. *et al.* (2019) 'Coordination of tissue cell polarity by auxin transport and signaling', *Elife*. eLife Sciences Publications, Ltd, 8, p. e51061.
- Viaene, T. *et al.* (2013) 'Origin and evolution of PIN auxin transporters in the green lineage', *Trends in plant science*. Elsevier, 18(1), pp. 5–10.
- Viaene, T. *et al.* (2014) 'Directional auxin transport mechanisms in early diverging land plants', *Current Biology*. Elsevier, 24(23), pp. 2786–2791.
- Vie, A. K. *et al.* (2015) 'The IDA/IDA-LIKE and PIP/PIP-LIKE gene families in *Arabidopsis*: phylogenetic relationship, expression patterns, and transcriptional effect of the PIPL3 peptide', *Journal of Experimental Botany*. Oxford University Press UK, 66(17), pp. 5351–5365.
- Vosolsobě, S., Skokan, R. and Petrášek, J. (2020) 'The evolutionary origins of auxin transport: what we know and what we need to know', *Journal of experimental botany*. Oxford University Press UK, 71(11), pp. 3287–3295.
- De Vries, J. and Archibald, J. M. (2018) 'Plant evolution: landmarks on the path to terrestrial life', *New Phytologist*. Wiley Online Library, 217(4), pp. 1428–1434.
- Wallace, E. F. *et al.* (1982) 'Carboxypeptidase B activity from adrenal medulla—Is it involved in the processing of proenkephalin?', *Life sciences*. Elsevier, 31(16–17), pp. 1793–1796.
- Walter, P., Gilmore, R. and Blobel, G. (1984) 'Protein translocation across the endoplasmic reticulum', *Cell*. Elsevier, 38(1), pp. 5–8.

- Wang, C. *et al.* (2015) ‘Insights into the origin and evolution of the plant hormone signaling machinery’, *Plant physiology*. American Society of Plant Biologists, 167(3), pp. 872–886.
- Wang, G., Zhang, G. and Wu, M. (2016) ‘CLE peptide signaling and crosstalk with phytohormones and environmental stimuli’, *Frontiers in plant science*. Frontiers Media SA, 6, p. 1211.
- Wang, H. *et al.* (1987) ‘GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants.’, *Journal of experimental botany*. Oxford University Press UK, 6(9), pp. 2629–2639.
- Wang, H. *et al.* (2020) ‘The moss flavone synthase I positively regulates the tolerance of plants to drought stress and UV-B radiation’, *Plant Science*. Elsevier, 298, p. 110591.
- Wang, H. *et al.* (2022) ‘A Moss 2-Oxoglutarate/Fe (II)-Dependent Dioxygenases (2-ODD) Gene of Flavonoids Biosynthesis Positively Regulates Plants Abiotic Stress Tolerance’, *Frontiers in Plant Science*. Frontiers Media SA, 13.
- Wang, J. *et al.* (2020) ‘Functions of jasmonic acid in plant regulation and response to abiotic stress’, *International Journal of Molecular Sciences*. MDPI, 21(4), p. 1446.
- Wang, X. *et al.* (2013) ‘Salicylic Acid Regulates Plasmodesmata Closure during Innate Immune Responses in Arabidopsis’, *The Plant Cell*. American Society of Plant Biologists, 25(6), pp. 2315–2329. doi: 10.1105/tpc.113.110676.
- Weisman, C. M., Murray, A. W. and Eddy, S. R. (2020) ‘Many, but not all, lineage-specific genes can be explained by homology detection failure’, *PLoS biology*. Public Library of Science San Francisco, CA USA, 18(11), p. e3000862.
- Whitewoods, C. D. *et al.* (2018) ‘CLAVATA was a genetic novelty for the morphological innovation of 3D growth in land plants’, *Current Biology*. Elsevier, 28(15), pp. 2365–2376.
- Whitford, R. *et al.* (2008) ‘Plant CLE peptides from two distinct functional classes synergistically induce division of vascular cells’, *Proceedings of the National Academy of Sciences*. National Acad Sciences, 105(47), pp. 18625–18630.
- Whitford, R. *et al.* (2012) ‘GOLVEN secretory peptides regulate auxin carrier turnover during plant gravitropic responses’, *Developmental cell*. Elsevier, 22(3), pp. 678–685.
- Wilson, A. C., Peterson, M. G. and Herr, W. (1995) ‘The HCF repeat is an unusual proteolytic cleavage signal.’, *Genes & development*. Cold Spring Harbor Lab, 9(20), pp. 2445–2458.
- Wu, L. and Lambert, J. D. (2022) ‘Clade-specific genes and the evolutionary origin of novelty; new tools in the toolkit’, in *Seminars in Cell & Developmental Biology*.
- Wu, S. *et al.* (2016) ‘Symplastic signaling instructs cell division, cell expansion, and cell polarity in the ground tissue of Arabidopsis thaliana roots’, *Proceedings of the National Academy of Sciences*. National Acad Sciences, 113(41), pp. 11621–11626.
- Xiao, L. *et al.* (2012) ‘Transcriptome of protoplasts reprogrammed into stem cells in Physcomitrella patens’, *PloS one*. Public Library of Science San Francisco, USA, 7(4), p. e35961.

- Xiao, W. *et al.* (2009) 'A rice gene of de novo origin negatively regulates pathogen-induced defense response', *PloS one*. Public Library of Science San Francisco, USA, 4(2), p. e4603.
- Xu, Y. *et al.* (2015) 'Identification, characterization and expression analysis of lineage-specific genes within sweet orange (*Citrus sinensis*)', *BMC genomics*. BioMed Central, 16(1), pp. 1–10.
- Yadav, N. K., Saikhedkar, N. S. and Giri, A. P. (2021) 'PINIR: a comprehensive information resource for Pin-II type protease inhibitors', *BMC plant biology*. BioMed Central, 21(1), pp. 1–14.
- Yamaguchi, Y. L., Ishida, T. and Sawa, S. (2016) 'CLE peptides and their signaling pathways in plant development', *Journal of Experimental Botany*. Oxford University Press UK, 67(16), pp. 4813–4826.
- Yanai, O. *et al.* (2005) 'Arabidopsis KNOXI proteins activate cytokinin biosynthesis', *Current Biology*. Elsevier, 15(17), pp. 1566–1571.
- Yang, F. *et al.* (2014) 'An auxin-responsive endogenous peptide regulates root development in Arabidopsis', *Journal of Integrative Plant Biology*. Wiley Online Library, 56(7), pp. 635–647.
- Yang, X. *et al.* (2009) 'Genome-wide identification of lineage-specific genes in Arabidopsis, Oryza and Populus', *Genomics*. Elsevier, 93(5), pp. 473–480.
- Yang, Y. *et al.* (2006) 'High-affinity auxin transport by the AUX1 influx carrier protein', *Current Biology*. Elsevier, 16(11), pp. 1123–1127.
- Yu, Y. *et al.* (2017) 'Ss-Rhs1, a secretory Rhs repeat-containing protein, is required for the virulence of *Sclerotinia sclerotiorum*', *Molecular plant pathology*. Wiley Online Library, 18(8), pp. 1052–1061.
- Yuasa, K. *et al.* (2005) 'Membrane-anchored prolyl hydroxylase with an export signal from the endoplasmic reticulum', *The Plant Journal*. Wiley Online Library, 41(1), pp. 81–94.
- Zazimalova, E. *et al.* (2010) 'Auxin transporters—why so many? Cold Spring Harb Perspect Biol 2: a001552'.
- Zechmeister, H. G., Grodzińska, K. and Szarek-Łukaszewska, G. (2003) 'Bryophytes', in *Trace Metals and other Contaminants in the Environment*. Elsevier, pp. 329–375.
- Zhang, H. *et al.* (2018) 'A plant phytosulfokine peptide initiates auxin-dependent immunity through cytosolic Ca²⁺ signaling in tomato', *The Plant Cell*. American Society of Plant Biologists, 30(3), pp. 652–667.
- Zhang, H. and Yin, T. (2015) 'Advances in lineage-specific genes', *Yi Chuan= Hereditas*, 37(6), pp. 544–553.
- Zhang, L. *et al.* (2009) 'Identification of an apoplastic protein involved in the initial phase of salt stress response in rice root by two-dimensional electrophoresis', *Plant Physiology*. American Society of Plant Biologists, 149(2), pp. 916–928.
- Zhang, Y. *et al.* (2019) 'Evolution of fast root gravitropism in seed plants', *Nature*

communications. Nature Publishing Group UK London, 10(1), p. 3480.

Zhao, Z. and Ma, D. (2021) 'Genome-wide identification, characterization and function analysis of lineage-specific genes in the tea plant *Camellia sinensis*', *Frontiers in Genetics*. Frontiers Media SA, 12.

Zimmer, A. D. *et al.* (2013) 'Reannotation and extended community resources for the genome of the non-seed plant *Physcomitrella patens* provide insights into the evolution of plant gene structures and functions', *BMC genomics*. Springer, 14(1), pp. 1–20.

Annexure

7. Annexures

Table 7.1. List of peptides detected upon MS/MS analyses of the bands observed in anti-GFP western blot

Line name	Band Size	Sequence	TDR number /GFP	Peptide number
ΔN SHLF-mEGFP: <i>shlf</i>	GFP (27 kDa) and ~ 101.6 SHLF+GFP (without N-terminal)	IQLAFKQITTASDNLR	1st TDR	P1
		QITTASDNLR	ALL TDR	P2
		IISAIAEYTIK	ALL TDR	P3
		IVLAFKQITTASDNLR	2,3,4	P5
		LTLKFICTTGK	GFP	PG
		FICTTGK	GFP	PG
		SAMPEGYVQER	GFP	PG
		SAMPEGYVQER	GFP	PG
		SAMPEGYVQER	GFP	PG
		SAMPEGYVQERTIFFK	GFP	PG
		TIFFKDDGNYK	GFP	PG
		DDGNYKTRAEVK	GFP	PG
		TRAEVKFEGDTLVNR	GFP	PG
		AEVKFEGDTLVNR	GFP	PG
		AEVKFEGDTLVNR	GFP	PG
		AEVKFEGDTLVNR	GFP	PG
		FEGDTLVNR	GFP	PG
		FEGDTLVNRIELK	GFP	PG
		IELKGIDFKEDGNILGHK	GFP	PG
		GIDFKEDGNILGHK	GFP	PG
		EDGNILGHK	GFP	PG
		EDGNILGHK	GFP	PG
		QKNGIKVNFK	GFP	PG
		NGIKVNFK	GFP	PG
		NGIKVNFKIR	GFP	PG
		LSKDPNK	GFP	PG
LSKDPNEKRDHVM	GFP	PG		

		INEGGVKNDNSPLPDAD AK	1st TDR	P5
		LGETMKEIAISA	C	P6
		FSVSGEGEGDATYGK	GFP	PG
		EIAISAMVSK	C+GFP	PG
ΔC SHLF- mEGFP:shlf		No GFP specific peptide was detected.		
ΔNΔC SHLF- mEGFP:shlf	27 kDa GFP	FSVSGEGEGDATYGK	GFP	PG
N-TDR1-C- eGFP:shlf	27 kDa GFP and 43 kDa (+GFP)	EIAISAMVSK	C+GFP	PG
		IQLAFK	1st TDR	P7
		QITTASDNLRKETQL	1st TDR	P8
		INEGGVKNDNSPLPDAD AK	1st TDR	P9
		FSVSGEGEGDATYGK	GFP	PG
		FICTTGK	GFP	PG
		QHDFFKSAMPEGYVQER	GFP	PG
		SAMPEGYVQER	GFP	PG
		SAMPEGYVQER	GFP	PG
		TIFFKDDGNYK	GFP	PG
		TIFFKDDGNYK	GFP	PG
		IFFKDDGNYK	GFP	PG
		DDGNYKTRAEVK	GFP	PG
		AEVKFEGDTLVNR	GFP	PG
		FEGDTLVNR	GFP	PG
		FEGDTLVNRIELK	GFP	PG
		GIDFKEDGNILGHK	GFP	PG
		DFKEDGNILGHK	GFP	PG
		LEYNYNSHNVY	GFP	PG
		QKNGIKVNFK	GFP	PG
NGIKVNFKIR	GFP	PG		
IRHNIEDGSVQ	GFP	PG		
IISAIAEYTIK	1st TDR	P10		
N-TDR1-2- eGFP lower band	GFP (27 kDa) and ~ 44 SHLF+GFP (2nd R+GFP)	QITTASDNLR	ALL TDR	P11
		IAEGFAKIISAIAEYTIK	ALL TDR	P12
		HGLLTLIPFFPIR	ALL TDR	P13

	AFALIAEIPTQK	ALL TDR	P14
	SPLPDADAQIVVQSL	2,3,4 TDR	P16
	VAITTYSSPILESNNLQE MNTT	2,3,4 TDR	P15
	IVLAFKQITTASDNLK	2,3,4 TDR	P17
	KEVQLINIINAPLQGFK	2,3,4 TDR	P18
	VSKGEELFTG	GFP	PG
	VSKGEELFTGVVPILVEL DGDVNGHK	GFP	PG
	GEELFTGVVPILVELDGD VNGHK	GFP	PG
	LDGDVNGHK	GFP	PG
	FSVSGEGEGDATYGK	GFP	PG
	FSVSGEGEGDATYGKLT LK	GFP	PG
	LPVPWPTLVTTLT	GFP	PG
	LPVPWPTLVTTLTYG VQCF	GFP	PG
	VTTLTYGVQCF	GFP	PG
	SRYPDHMKQHDFFK	GFP	PG
	YPDHMKQHDFFK	GFP	PG
	YPDHMKQHDFFK	GFP	PG
	YPDHMKQHDFFKSAMPE GYVQER	GFP	PG
	QHDFFKSAMPEGYVQER	GFP	PG
	SAMPEGYVQER	GFP	PG
	SAMPEGYVQERTIFFK	GFP	PG
	MPEGYVQERTIFFK	GFP	PG
	PEGYVQER	GFP	PG
	TIFFKDDGNYK	GFP	PG
	TIFFKDDGNYKTR	GFP	PG
	AEVKFEGDTLVNR	GFP	PG
	FEGDTLVNR	GFP	PG
	FEGDTLVNRIELK	GFP	PG
	IELKGIDFKEDGNILGHK	GFP	PG
	IELKGIDFKEDGNILGHK	GFP	PG
	GIDFKEDGNILGHK	GFP	PG

	DFKEDGNILGHK	GFP	PG
	EDGNILGHK	GFP	PG
	LEYNYNSHNVY	GFP	PG
	LEYNYNSHNVYIMADK	GFP	PG
	LEYNYNSHNVYIMADKQ K	GFP	PG
	YIMADKQKNGIK	GFP	PG
	YIMADKQKNGIK	GFP	PG
	NGIKVNFKIR	GFP	PG
	NGIKVNFKIR	GFP	PG
	VNFKIR	GFP	PG
	IRHNIEDGGSVQ	GFP	PG
	IRHNIEDGGSVQLADHYQ QN	GFP	PG
	IRHNIEDGGSVQLADHYQ QNTPIGDGPVLLPDNHYL STQSK	GFP	PG
	GPVLLPDNHYLSTQSK	GFP	PG
	PVLLPDNHYLSTQSK	GFP	PG
	GPVLLPDNHYLSTQSK	GFP	PG
	DNHYLSTQSK	GFP	PG
	QSKLSKDPNEKR	GFP	PG
	LSKDPNEK	GFP	PG
	LSKDPNEKR	GFP	PG
	LSKDPNEKRDHMVL	GFP	PG
	RDHMLLEFVTAAGITL GMDELYK	GFP	PG
	VTAAGITLGMDEL	GFP	PG
	SLALCHSVNAGMSNVEK	N+1 TDR	P19
	IQLAFKQITTASDNLRK	1st TDR	P20
	KETQLINIINAPLQGSK	1st TDR	P21
	INEGGVKNDNSPLPDAD AK	1st TDR	P22
	IVLAFKQITTASDNLR	2,3,4 TDR	P23
	IISAIAEYTIK	ALL TDR	P24
	VVQSLTTFVQVHQ	ALL TDR	P25

		HGLLTLPFFPIR	ALL TDR	P26
		ALIAEIPTQKPAA	ALL TDR	P27
		QKPAADVQFGSL	ALL TDR	P28

Table 7.2 List of secretory proteins identified from WT secretome

Proteins identified from WT secretome					
Accession IDs					
Pp3c14_22870V3.4	Pp3c18_8100V3.3	Pp3c16_22330V3.8	Pp3c2_6650V3.5	Pp3c2_6690V3.2	Pp3c27_4150V3.1
Pp3c14_22870V3.3	Pp3c18_8100V3.2	Pp3c16_22330V3.7	Pp3c2_6650V3.4	Pp3c2_6690V3.1	Pp3c22_17810V3.2
Pp3c14_22870V3.1	Pp3c18_8100V3.1	Pp3c16_22330V3.6	Pp3c2_6650V3.3	Pp3c1_23670V3.3	Pp3c22_17810V3.1
Pp3c14_22870V3.1	Pp3c20_23270V3.3	Pp3c16_22330V3.5	Pp3c2_6650V3.2	Pp3c1_23670V3.2	Pp3c18_5520V3.3
Pp3c14_22870V3.2	Pp3c20_23270V3.2	Pp3c16_22330V3.4	Pp3c2_6650V3.1	Pp3c1_23670V3.1	Pp3c10_23470V3.3
Pp3c6_12510V3.2	Pp3c20_23270V3.1	Pp3c16_22330V3.3	Pp3c2_6160V3.2	Pp3c20_17920V3.3	Pp3c5_190V3.3
Pp3c6_12510V3.1	Pp3c20_23240V3.3	Pp3c16_22330V3.2	Pp3c2_6160V3.1	Pp3c20_17920V3.1	Pp3c5_190V3.2
Pp3c4_17240V3.2	Pp3c20_23240V3.2	Pp3c16_22330V3.1	Pp3c2_10310V3.6	Pp3c10_8570V3.7	Pp3c5_190V3.1
Pp3c4_17240V3.1	Pp3c20_23240V3.1	Pp3c16_22330V3.1	Pp3c2_10310V3.5	Pp3c10_8570V3.6	Pp3c3_3140V3.2
Pp3c2_35930V3.2	Pp3c13_6570V3.2	Pp3c2_6770V3.3	Pp3c2_10310V3.4	Pp3c10_8570V3.5	Pp3c15_710V3.6
Pp3c2_35930V3.1	Pp3c13_6570V3.1	Pp3c2_6770V3.2	Pp3c2_10310V3.3	Pp3c10_8570V3.4	Pp3c15_710V3.5
Pp3c22_5610V3.2	Pp3c1_32200V3.3	Pp3c2_6770V3.1	Pp3c2_10310V3.2	Pp3c10_8570V3.3	Pp3c15_710V3.4
Pp3c22_5610V3.1	Pp3c1_32200V3.2	Pp3c2_6690V3.5	Pp3c2_10310V3.1	Pp3c10_8570V3.2	Pp3c15_710V3.3
Pp3c6_15360V3.1	Pp3c1_32200V3.1	Pp3c2_6690V3.4	Pp3c1_23750V3.1	Pp3c10_8570V3.1	Pp3c15_710V3.2

Pp3c5_17290V3.2	Pp3c6_15360V3.2	Pp3c2_6690V3.3	Pp3c1_23750V3.2	Pp3c10_8620V3.2	Pp3c15_710V3.1
Pp3c5_17290V3.1	Pp3c6_15360V3.1	Pp3c23_15870V3.2	Pp3c2_26540V3.4	Pp3c5_4200V3.2	Pp3c9_25950V3.3
Pp3c3_2620V3.2	Pp3c5_17290V3.2	Pp3c23_15870V3.1	Pp3c2_26540V3.3	Pp3c5_4200V3.1	Pp3c9_25950V3.2
Pp3c3_2620V3.1	Pp3c5_17290V3.1	Pp3c18_21210V3.3	Pp3c2_26540V3.2	Pp3c22_2920V3.2	Pp3c9_25950V3.1
Pp3c1_2220V3.2	Pp3c3_2620V3.2	Pp3c18_21210V3.2	Pp3c2_26540V3.1	Pp3c22_2920V3.1	Pp3c9_25700V3.2
Pp3c1_2220V3.1	Pp3c3_2620V3.1	Pp3c18_21210V3.1	Pp3c8_24120V3.4	Pp3c7_5160V3.5	Pp3c9_25700V3.1
Pp3c1_2210V3.1	Pp3c1_2220V3.2	Pp3c16_23780V3.3	Pp3c8_24120V3.3	Pp3c7_5160V3.4	Pp3c9_26020V3.2
Pp3c1_2170V3.3	Pp3c1_2220V3.1	Pp3c16_23780V3.2	Pp3c8_24120V3.2	Pp3c7_5160V3.3	Pp3c10_23470V3.3
Pp3c1_2170V3.2	Pp3c1_2210V3.1	Pp3c16_23780V3.1	Pp3c8_24120V3.1	Pp3c7_5160V3.2	Pp3c14_9020V3.1
Pp3c2_5450V3.2	Pp3c1_2170V3.3	Pp3c25_760V3.4	Pp3c24_20010V3.4	Pp3c7_5160V3.1	Pp3c20_19140V3.4
Pp3c2_5450V3.1	Pp3c1_2170V3.2	Pp3c25_760V3.3	Pp3c24_20010V3.3	Pp3c4_5690V3.2	Pp3c20_19140V3.3
Pp3c24_18690V3.2	Pp3c2_5450V3.2	Pp3c25_760V3.2	Pp3c24_20010V3.2	Pp3c12_25460V3.1	Pp3c20_19140V3.2
Pp3c24_18690V3.1	Pp3c2_5450V3.1	Pp3c25_760V3.1	Pp3c3_26540V3.1	Pp3c12_25430V3.2	Pp3c20_19140V3.1
Pp3c8_13980V3.2	Pp3c24_18690V3.2	Pp3c16_22330V3.9	Pp3c27_2310V3.2	Pp3c12_25430V3.1	Pp3c1_19060V3.2
Pp3c10_8620V3.1	Pp3c3_25110V3.1	Pp3c24_20010V3.1	Pp3c27_2310V3.1	Pp3c2_28200V3.3	Pp3c1_19060V3.1
Pp3c2_26540V3.5	Pp3c3_26540V3.2	Pp3c3_25110V3.2	Pp3c16_5670V3.2	Pp3c2_28200V3.2	Pp3c19_5170V3.4
Pp3c4_5680V3.2	Pp3c4_5680V3.1	Pp3c12_25460V3.2	Pp3c16_5670V3.1	Pp3c2_28200V3.1	Pp3c16_16040V3.3
Pp3c26_6990V3.2	Pp3c26_6990V3.2	Pp3c26_6990V3.2	Pp3c26_6990V3.2	Pp3c26_6990V3.2	Pp3c16_16040V3.2
Pp3c12_22990V3.2	Pp3c26_6810V3.1	Pp3c20_2050V3.3	Pp3c16_12390V3.1	Pp3c27_660V3.1	Pp3c16_16040V3.1
Pp3c12_22990V31	Pp3c26_3070V3.2	Pp3c20_2050V3.2	Pp3c6_20300V3.2	Pp3c1_7510V3.2	Pp3c1_34240V3.2
Pp3c12_3460V3.5	Pp3c26_3070V3.1	Pp3c20_2050V3.1	Pp3c6_20300V3.1	Pp3c1_7510V3.1	Pp3c1_34240V3.1

Pp3c12_3460V3.4	Pp3c25_4480V3.3	Pp3c14_11470V3.2	Pp3c5_10730V3.2	Pp3c16_2860V3.2	Pp3c1_34210V3.2
Pp3c12_3460V3.3	Pp3c25_4480V3.2	Pp3c14_11470V3.1	Pp3c5_10730V3.1	Pp3c16_2860V3.1	Pp3c1_34210V3.1
Pp3c12_3460V3.2	Pp3c25_4480V3.1	Pp3c10_25415V3.1	Pp3c25_3840V3.3	Pp3c15_14410V3.2	Pp3c17_5070V3.3
Pp3c12_3460V3.1	Pp3c20_2100V3.9	Pp3c10_25410V3.1	Pp3c25_3840V3.2	Pp3c15_14410V3.1	Pp3c17_5070V3.2
Pp3c15_13040V3.2	Pp3c20_2100V3.8	Pp3c10_25371V3.4	Pp3c25_3840V3.1	Pp3c21_6170V3.4	Pp3c17_5070V3.1
Pp3c15_13040V3.1	Pp3c20_2100V3.7	Pp3c10_25371V3.3	Pp3c9_15380V3.2	Pp3c21_6170V3.3	Pp3c1_2210V3.4
Pp3c15_12980V3.2	Pp3c20_2100V3.6	Pp3c10_25371V3.2	Pp3c9_15380V3.1	Pp3c21_6170V3.2	Pp3c1_2210V3.1
Pp3c27_7020V3.1	Pp3c20_2100V3.5	Pp3c10_25371V3.1	Pp3c6_19290V3.2	Pp3c21_6170V3.1	Pp3c10_23470V3.3
Pp3c20_20260V3.3	Pp3c20_2100V3.4	Pp3c10_25360V3.1	Pp3c6_19290V3.1	Pp3c6_22090V3.3	Pp3c14_9020V3.1
Pp3c20_20260V3.2	Pp3c20_2100V3.3	Pp3c10_18650V3.2	Pp3c27_820V3.2	Pp3c6_22090V3.2	Pp3c4_29270V3.2
Pp3c4_5690V3.1	Pp3c20_2100V3.2	Pp3c10_18650V3.1	Pp3c27_820V3.1	Pp3c6_22090V3.1	Pp3c4_29270V3.1
Pp3c18_5520V3.2	Pp3c20_2100V3.1	Pp3c16_12390V3.2	Pp3c27_660V3.2	Pp3c27_7020V3.2	Pp3c4_29150V3.3
Pp3c18_5520V3.1	Pp3c10_12250V3.1	Pp3c10_12250V3.1	Pp3c10_12250V3.1	Pp3c10_12250V3.1	Pp3c4_29150V3.2
Pp3c27_7010V3.2	Pp3c1_2210V3.2	Pp3c9_26020V3.1	Pp3c5_7940V3.2	Pp3c3_10800V3.2	Pp3c4_29150V3.1
Pp3c27_7010V3.1	Pp3c9_450V3.2	Pp3c8_1130V3.3	Pp3c5_7940V3.1	Pp3c3_10800V3.1	Pp3c3_12510V3.3
Pp3c16_16040V3.4	Pp3c9_450V3.1	Pp3c8_1130V3.2	Pp3c4_3440V3.3	Pp3c2_34620V3.2	Pp3c3_12510V3.2
Pp3c8_1130V3.1	Pp3c4_3440V3.2	Pp3c10_16420V3.2	Pp3c14_21480V3.1	Pp3c2_34620V3.1	Pp3c3_22150V3.1
Pp3c17_19310V3.2	Pp3c4_3440V3.1	Pp3c10_16420V3.1	Pp3c10_23470V3.3	Pp3c2_34580V3.2	Pp3c14_23460V3.1
Pp3c17_19310V3.1	Pp3c4_20940V3.2	Pp3c10_23470V3.3	Pp3c14_9020V3.1	Pp3c2_34580V3.1	Pp3c3_12510V3.1
Pp3c14_24450V3.2	Pp3c4_20940V3.1	Pp3c24_20960V3.2	Pp3c20_20260V3.1	Pp3c2_23590V3.6	Pp3c19_5170V3.3
Pp3c14_24450V3.1	Pp3c3_35240V3.2	Pp3c23_21090V3.2	Pp3c3_3160V3.4	Pp3c27_4150V3.5	

Pp3c14_23460V3.5	Pp3c3_35240V3.1	Pp3c13_14980V3.1	Pp3c3_3160V3.3	Pp3c27_4150V3.4	
Pp3c14_23460V3.4	Pp3c3_22150V3.4	Pp3c3_29970V3.2	Pp3c3_3160V3.2	Pp3c27_4150V3.3	
Pp3c14_23460V3.3	Pp3c3_22150V3.3	Pp3c3_29970V3.1	Pp3c3_3160V3.1	Pp3c27_4150V3.2	
Pp3c14_23460V3.2	Pp3c3_22150V3.2	Pp3c14_21480V3.2	Pp3c14_9020V3.2	Pp3c17_6520V3.1	

Table 7.3 List of secretory proteins identified from *shlf* secretome

Proteins identified from <i>shlf</i> secretome					
Accession IDs					
Pp3c2_26540V3.2	Pp3c2_26540V3.2	Pp3c2_26540V3.2	Pp3c2_26540V3.2	Pp3c2_26540V3.2	Pp3c2_6770V3.2
Pp3c2_26540V3.1	Pp3c2_26540V3.1	Pp3c2_26540V3.1	Pp3c2_26540V3.1	Pp3c2_26540V3.1	Pp3c2_6770V3.1
Pp3c9_5310V3.3	Pp3c9_5310V3.3	Pp3c9_5310V3.3	Pp3c9_5310V3.3	Pp3c9_5310V3.3	Pp3c2_6690V3.5
Pp3c9_5310V3.2	Pp3c9_5310V3.2	Pp3c9_5310V3.2	Pp3c9_5310V3.2	Pp3c9_5310V3.2	Pp3c2_6690V3.4
Pp3c9_5310V3.1	Pp3c9_5310V3.1	Pp3c9_5310V3.1	Pp3c9_5310V3.1	Pp3c9_5310V3.1	Pp3c2_6690V3.3
Pp3c9_5310V3.4	Pp3c9_5310V3.4	Pp3c9_5310V3.4	Pp3c9_5310V3.4	Pp3c9_5310V3.4	Pp3c2_6690V3.2
Pp3c9_5280V3.2	Pp3c9_5280V3.2	Pp3c9_5280V3.2	Pp3c9_5280V3.2	Pp3c9_5280V3.2	Pp3c2_6690V3.1
Pp3c9_5280V3.1	Pp3c9_5280V3.1	Pp3c9_5280V3.1	Pp3c9_5280V3.1	Pp3c9_5280V3.1	Pp3c2_6650V3.5
Pp3c15_12410V3.3	Pp3c15_12410V3.3	Pp3c15_12410V3.3	Pp3c15_12410V3.3	Pp3c15_12410V3.3	Pp3c3_4140V3.2
Pp3c15_12410V3.2	Pp3c15_12410V3.2	Pp3c15_12410V3.2	Pp3c15_12410V3.2	Pp3c15_12410V3.2	Pp3c3_4140V3.1
Pp3c15_12410V3.1	Pp3c15_12410V3.1	Pp3c15_12410V3.1	Pp3c15_12410V3.1	Pp3c15_12410V3.1	Pp3c2_36570V3.3
Pp3c13_7000V3.2	Pp3c13_7000V3.2	Pp3c13_7000V3.2	Pp3c13_7000V3.2	Pp3c13_7000V3.2	Pp3c2_36570V3.2
Pp3c13_7000V3.1	Pp3c13_7000V3.1	Pp3c13_7000V3.1	Pp3c13_7000V3.1	Pp3c13_7000V3.1	Pp3c2_36570V3.1

Pp3c3_510V3.3	Pp3c3_510V3.3	Pp3c3_510V3.3	Pp3c3_510V3.3	Pp3c3_510V3.3	Pp3c2_36500V3.1
Pp3c3_510V3.2	Pp3c3_510V3.2	Pp3c3_510V3.2	Pp3c3_510V3.2	Pp3c3_510V3.2	Pp3c2_36480V3.1
Pp3c3_510V3.1	Pp3c3_510V3.1	Pp3c3_510V3.1	Pp3c3_510V3.1	Pp3c3_510V3.1	Pp3c2_36380V3.2
Pp3c25_760V3.4	Pp3c25_760V3.4	Pp3c25_760V3.4	Pp3c25_760V3.4	Pp3c25_760V3.4	Pp3c2_36380V3.1
Pp3c25_760V3.3	Pp3c25_760V3.3	Pp3c25_760V3.3	Pp3c25_760V3.3	Pp3c25_760V3.3	Pp3c2_36220V3.2
Pp3c25_760V3.2	Pp3c25_760V3.2	Pp3c25_760V3.2	Pp3c25_760V3.2	Pp3c25_760V3.2	Pp3c2_36220V3.1
Pp3c25_760V3.1	Pp3c25_760V3.1	Pp3c25_760V3.1	Pp3c25_760V3.1	Pp3c25_760V3.1	Pp3c2_36190V3.2
Pp3c20_21960V3.9	Pp3c20_21960V3.9	Pp3c20_21960V3.9	Pp3c20_21960V3.9	Pp3c20_21960V3.9	Pp3c2_36190V3.1
Pp3c20_21960V3.8	Pp3c20_21960V3.8	Pp3c20_21960V3.8	Pp3c20_21960V3.8	Pp3c20_21960V3.8	Pp3c26_6740V3.2
Pp3c20_21960V3.7	Pp3c20_21960V3.7	Pp3c20_21960V3.7	Pp3c20_21960V3.7	Pp3c20_21960V3.7	Pp3c26_6740V3.1
Pp3c20_21960V3.6	Pp3c20_21960V3.6	Pp3c20_21960V3.6	Pp3c20_21960V3.6	Pp3c20_21960V3.6	Pp3c21_3950V3.2
Pp3c20_21960V3.5	Pp3c20_21960V3.5	Pp3c20_21960V3.5	Pp3c20_21960V3.5	Pp3c20_21960V3.5	Pp3c21_3950V3.1
Pp3c20_21960V3.4	Pp3c20_21960V3.4	Pp3c20_21960V3.4	Pp3c20_21960V3.4	Pp3c20_21960V3.4	Pp3c13_5930V3.2
Pp3c20_21960V3.3	Pp3c20_21960V3.3	Pp3c20_21960V3.3	Pp3c20_21960V3.3	Pp3c20_21960V3.3	Pp3c13_5930V3.1
Pp3c20_21960V3.2	Pp3c20_21960V3.2	Pp3c20_21960V3.2	Pp3c20_21960V3.2	Pp3c20_21960V3.2	Pp3c10_3020V3.3
Pp3c20_21960V3.1	Pp3c20_21960V3.1	Pp3c20_21960V3.1	Pp3c20_21960V3.1	Pp3c20_21960V3.1	Pp3c10_3020V3.2
Pp3c3_2620V3.2	Pp3c3_2620V3.2	Pp3c3_2620V3.2	Pp3c3_2620V3.2	Pp3c3_2620V3.2	Pp3c3_4140V3.2
Pp3c3_2620V3.1	Pp3c3_2620V3.1	Pp3c3_2620V3.1	Pp3c3_2620V3.1	Pp3c3_2620V3.1	Pp3c3_4140V3.1
Pp3c6_12510V3.2	Pp3c6_12510V3.2	Pp3c6_12510V3.2	Pp3c6_12510V3.2	Pp3c6_12510V3.2	Pp3c2_36570V3.3
Pp3c6_12510V3.1	Pp3c6_12510V3.1	Pp3c6_12510V3.1	Pp3c6_12510V3.1	Pp3c6_12510V3.1	Pp3c2_36570V3.2
Pp3c4_17240V3.2	Pp3c4_17240V3.2	Pp3c4_17240V3.2	Pp3c4_17240V3.2	Pp3c4_17240V3.2	Pp3c2_36570V3.1

Pp3c4_17240V3.1	Pp3c4_17240V3.1	Pp3c4_17240V3.1	Pp3c4_17240V3.1	Pp3c4_17240V3.1	Pp3c2_36500V3.1
Pp3c2_35930V3.2	Pp3c2_35930V3.2	Pp3c2_35930V3.2	Pp3c2_35930V3.2	Pp3c2_35930V3.2	Pp3c2_36480V3.1
Pp3c2_35930V3.1	Pp3c2_35930V3.1	Pp3c2_35930V3.1	Pp3c2_35930V3.1	Pp3c2_35930V3.1	Pp3c2_36380V3.2
Pp3c22_5610V3.2	Pp3c22_5610V3.2	Pp3c22_5610V3.2	Pp3c22_5610V3.2	Pp3c22_5610V3.2	Pp3c2_36380V3.1
Pp3c22_5610V3.1	Pp3c22_5610V3.1	Pp3c22_5610V3.1	Pp3c22_5610V3.1	Pp3c22_5610V3.1	Pp3c2_36220V3.2
Pp3c18_8100V3.3	Pp3c18_8100V3.3	Pp3c18_8100V3.3	Pp3c18_8100V3.3	Pp3c18_8100V3.3	Pp3c2_36220V3.1
Pp3c18_8100V3.2	Pp3c18_8100V3.2	Pp3c18_8100V3.2	Pp3c18_8100V3.2	Pp3c18_8100V3.2	Pp3c2_36190V3.2
Pp3c18_8100V3.1	Pp3c18_8100V3.1	Pp3c18_8100V3.1	Pp3c18_8100V3.1	Pp3c18_8100V3.1	Pp3c2_36190V3.1
Pp3c3_4100V3.2	Pp3c3_4100V3.2	Pp3c3_4100V3.2	Pp3c3_4100V3.2	Pp3c3_4100V3.2	Pp3c26_6740V3.2
Pp3c3_4100V3.1	Pp3c3_4100V3.1	Pp3c3_4100V3.1	Pp3c3_4100V3.1	Pp3c3_4100V3.1	Pp3c26_6740V3.1
Pp3c5_7180V3.2	Pp3c5_7180V3.2	Pp3c5_7180V3.2	Pp3c5_7180V3.2	Pp3c5_7180V3.2	Pp3c21_3950V3.2
Pp3c5_7180V3.1	Pp3c5_7180V3.1	Pp3c5_7180V3.1	Pp3c5_7180V3.1	Pp3c5_7180V3.1	Pp3c21_3950V3.1
Pp3c5_7150V3.2	Pp3c5_7150V3.2	Pp3c5_7150V3.2	Pp3c5_7150V3.2	Pp3c5_7150V3.2	Pp3c13_5930V3.2
Pp3c5_7150V3.1	Pp3c5_7150V3.1	Pp3c5_7150V3.1	Pp3c5_7150V3.1	Pp3c5_7150V3.1	Pp3c13_5930V3.1
Pp3c5_7110V3.2	Pp3c5_7110V3.2	Pp3c5_7110V3.2	Pp3c5_7110V3.2	Pp3c5_7110V3.2	Pp3c10_3020V3.3
Pp3c5_7110V3.1	Pp3c5_7110V3.1	Pp3c5_7110V3.1	Pp3c5_7110V3.1	Pp3c5_7110V3.1	Pp3c10_3020V3.2
Pp3c3_4140V3.2	Pp3c3_4140V3.2	Pp3c3_4140V3.2	Pp3c10_3020V3.1	Pp3c10_3020V3.1	Pp3c10_3020V3.1
Pp3c3_4140V3.1	Pp3c3_4140V3.1	Pp3c3_4140V3.1	Pp3c2_6770V3.3	Pp3c2_6770V3.3	Pp3c2_6770V3.3
Pp3c2_36570V3.3	Pp3c2_36570V3.3	Pp3c2_36570V3.3	Pp3c2_6770V3.2	Pp3c2_6770V3.2	Pp3c13_5930V3.2
Pp3c2_36570V3.2	Pp3c2_36570V3.2	Pp3c2_36570V3.2	Pp3c2_6770V3.1	Pp3c2_6770V3.1	Pp3c13_5930V3.1
Pp3c2_36570V3.1	Pp3c2_36570V3.1	Pp3c2_36570V3.1	Pp3c2_6690V3.5	Pp3c2_6690V3.5	Pp3c10_3020V3.3

Pp3c2_36500V3.1	Pp3c2_36500V3.1	Pp3c2_36500V3.1	Pp3c2_6690V3.4	Pp3c2_6690V3.4	Pp3c10_3020V3.2
Pp3c2_36480V3.1	Pp3c2_36480V3.1	Pp3c2_36480V3.1	Pp3c2_6690V3.3	Pp3c2_6690V3.3	Pp3c13_5930V3.2
Pp3c2_36380V3.2	Pp3c2_36380V3.2	Pp3c2_36380V3.2	Pp3c2_6690V3.2	Pp3c2_6690V3.2	Pp3c13_5930V3.1
Pp3c2_36380V3.1	Pp3c2_36380V3.1	Pp3c2_36380V3.1	Pp3c2_6690V3.1	Pp3c2_6690V3.1	Pp3c10_3020V3.3
Pp3c2_36220V3.2	Pp3c2_36220V3.2	Pp3c2_36220V3.2	Pp3c2_6650V3.5	Pp3c2_6650V3.5	Pp3c10_3020V3.2
Pp3c2_36220V3.1	Pp3c2_36220V3.1	Pp3c2_36220V3.1	Pp3c10_3020V3.1	Pp3c10_3020V3.1	Pp3c13_5930V3.2
Pp3c2_36190V3.2	Pp3c2_36190V3.2	Pp3c2_36190V3.2	Pp3c2_6770V3.3	Pp3c2_6770V3.3	Pp3c13_5930V3.1
Pp3c2_36190V3.1	Pp3c2_36190V3.1	Pp3c2_36190V3.1	Pp3c2_6770V3.2	Pp3c2_6770V3.2	Pp3c10_3020V3.3
Pp3c26_6740V3.2	Pp3c26_6740V3.2	Pp3c26_6740V3.2	Pp3c2_6770V3.1	Pp3c2_6770V3.1	Pp3c10_3020V3.2
Pp3c26_6740V3.1	Pp3c26_6740V3.1	Pp3c26_6740V3.1	Pp3c2_6690V3.5	Pp3c2_6690V3.5	Pp3c2_6690V3.2
Pp3c21_3950V3.2	Pp3c21_3950V3.2	Pp3c21_3950V3.2	Pp3c2_6690V3.4	Pp3c2_6690V3.4	Pp3c2_6690V3.1
Pp3c21_3950V3.1	Pp3c21_3950V3.1	Pp3c21_3950V3.1	Pp3c2_6690V3.3	Pp3c2_6690V3.3	Pp3c2_6650V3.5

Table 7.4 List of secretory SHLF peptides identified from moss lines used in this study

Secretory SHLF peptides detected in: WT, SHLF-eGFP and N-TDR1-2-eGFP (miniSHLF)

TDR	Sequence	Peptide
N and 1st TDR	LCHSVNAGMSNVEK	P50
N and 1st TDR	HSVNAGMSNVEK	P51
N and 1st TDR	NAGMSNVEKI	P52
1st TDR	FIFSLALCHSVNAGMSNVEK (3)	P53
1st TDR	SNVEKIQL	P54
1st TDR	IQLAFK (2)	P55
1st TDR	IQLAFKQITTA	P56
1st TDR	IQLAFKQITTASDNLR (3)	P57
1st TDR	ETQLINIINA	P58

1st TDR	ETQLINIINAPLQGSK	P59
1st TDR	INAPLQGSKIAEGFAK	P60
1st TDR	IISAIAEYTIK	P61
1st TDR	INEGGVKNDNSPLPDADAK	P62
1st TDR	SPLPDADAKLVVQ	P63
1st TDR	SPLPDADAQIVVQ	P64
1st TDR	PLQGSKIAEGFAK	P65
1st TDR	ETQLINIINAPLQGSK (5)	P66
1st TDR	ETQLINIINAPLQGSKIAEGFAKI	P67
1st TDR	INIINAPLQGSKIAEGFA	P68
1st TDR	ISIAIAEYTIKINEGGVKNDNS	P69
1st TDR	INEGGVKNDNSPL	P70
1st TDR	INEGGVKNDNSPLPDADAK (4)	P71
1st TDR	NEGGVKNDNSPL	P72
1st TDR	NDNSPLPDADAK	P73
1st TDR	DNSPLPDADAK	P74
1st TDR	SITLQVAITTYSS	P75
1st TDR	ISIAIAEYTIKINEGGVKNDNS	P76
2nd - 4th TDR	SAIAEYTIKINEGGVKNYNSP	P77
2nd - 4th TDR	ALLNVVIGKH	P78
2nd - 4th TDR	ALLNVVIGKHGLL	P79
2nd - 4th TDR	EVQLINIINAPLQGFK	P80
2nd - 4th TDR	IINAPLQGFKI	P81
2nd - 4th TDR	APLQGFKIAEGFAK	P82
2nd - 4th TDR	EGGVKNYSPLPDADAQI	P83
2nd - 4th TDR	ADVQFGSLSVTL	P84
2nd - 4th TDR	ADVQFGSLSVTLQV	P85
2nd - 4th TDR	SSPILESNNLQEMNTTEK	P86
2nd - 4th TDR	LESNNLQEM	P87
2nd - 4th TDR	IVLAFK	P88
2nd - 4th TDR	VLAFKQITTASDNLR	P89
2nd - 4th TDR	QITTASDNLRKEVQLINIINAPL	P90
2nd - 4th TDR	TASDNLRKEVQ	P91

2nd - 4th TDR	KEVQLINIINAPLQGFK (9)	P92
2nd - 4th TDR	EVQLINIINAPLQGFKIAEGFAKI	P93
2nd - 4th TDR	INEGGVKNYNSPLPDAD (2)	P94
2nd - 4th TDR	EGGVKNYNSPLPDADAQI	P95
All TDRs	KQITTASDNLR (2)	P96
All TDRs	QITTASDNLR (6)	P97
All TDRs	QITTASDNLRK (13)	P98
All TDRs	TFVQVHQALLNVVIGK	P99
All TDRs	IVLAFKQITTASDNLR (2)	P100
All TDRs	IAEGFAK	P101
All TDRs	IISAIAEYTIK (10)	P102
All TDRs	IISAIAEYTIKI	P103
All TDRs	ISAIAEYT	P104
All TDRs	TIKINEGGVK	P105
All TDRs	SLTTFVQVHQALLN	P106
All TDRs	IGKHGLLTLIPFFE	P107
All TDRs	HGLLTLIPFFEPIR	P108
All TDRs	LLTLIPFFEPI	P109
All TDRs	ALIAEIPTQKPAA	P110

* No secretory SHLF peptides were detected in:

1. *shlf*
2. Δ N SHLF-mEGFP
3. Δ C SHLF-mEGFP
4. Δ C SHLF-mEGFP
5. Δ N Δ C SHLF-mEGFP
6. N-TDR1-C-eGFP
7. TDR1

Table 7.5 List of most up-regulated genes between WT and *shlf* at p value <0.01

Gene IDs	Gene name	log2Fold Change
Pp3c6_5150	HEAVY METAL TRANSPORT/DETOXIFICATION DOMAIN-CONTAINING PROTEIN	8.53
Pp3c18_18530	SH3 Domains-Co-expressed with genes in gametophores specific co-expression subnetwork	8.12
Pp3c17_13620	AP2 domain containing TF	8.08
Pp3c17_17450	SMI1/KNR4-like	7.69
Pp3c10_1530	Unnamed family	7.53
Pp3c6_1720	Unknown family	7.49
Pp3c24_9670	EARLY LIGHT-INDUCED PROTEIN 1, CHLOROPLASTIC-RELATED	7.32
Pp3c2_6110	Rubredoxin-like domain	7.00
Pp3c3_21910	FLOTILLIN-RELATED	6.80
Pp3c8_17610	Unnamed family	6.78
Pp3c14_1598	Unnamed family	6.75
Pp3c15_21480	CHLORIDE INTRACELLULAR CHANNEL, ISOFORM B	6.41
Pp3c16_23550	Translation factors	6.17
Pp3c4_3600	CHITINASE-RELATED	6.01
Pp3c23_22700	Cytolysin/lectin	5.99
Pp3c4_18020	LATE EMBRYOGENESIS ABUNDANT DOMAIN-CONTAINING PROTEIN-RELATED	5.96
Pp3c11_1420	Chitin recognition protein (Chitin_bind_1) // Rare lipoprotein A (RlpA)-like double-psi beta-barrel (DPBB_1)	5.85
Pp3c3_14410	Vaccinia Virus protein VP39	5.77
Pp3c12_11870	Unnamed family	5.70

Pp3c3_17090	Seed maturation protein (SMP)	5.60
Pp3c16_20740	Small hydrophilic plant seed protein (LEA_5)	5.59
Pp3c2_31380	EARLY LIGHT-INDUCED PROTEIN 1, CHLOROPLASTIC-RELATED	5.57
Pp3c14_1600	COPPER TRANSPORT PROTEIN ATOX1-RELATED	5.49
Pp3c7_10780	AP2 domain (AP2)	5.44
Pp3c1_37850	PROTEIN LURP-ONE-RELATED 1-RELATED	5.42
Pp3c17_4070	Unnamed family	5.32
Pp3c20_18310	Unnamed family	5.31
Pp3c11_7280	EARLY LIGHT-INDUCED PROTEIN 1, CHLOROPLASTIC-RELATED	5.16
Pp3c23_830	Unnamed family	5.16
Pp3c8_4130	PROTEIN PLANT CADMIUM RESISTANCE 2-RELATED	5.13
Pp3c10_13570	Unnamed family	5.12
Pp3c14_21880	Apolipoprotein (APOA1 / APOA4 / APOE FAMILY)	4.99
Pp3c5_2060	Unnamed family	4.98
Pp3c17_8560	LATE EMBRYOGENESIS ABUNDANT PLANTS LEA-RELATED	4.96
Pp3c8_19800	Unnamed family	4.95
Pp3c9_3690	DEHYDROGENASE RELATED	4.88
Pp3c1_14230	AP2 domain (AP2)	4.88
Pp3c11_4920	Unnamed family	4.77
Pp3c22_8470	THIOREDOXIN PEROXIDASE	4.72
Pp3c26_12920	GLUCOSE AND RIBITOL DEHYDROGENASE HOMOLOG 1-RELATED	4.71
Pp3c21_10570	Unnamed family	4.57
Pp3c14_10490	Rubredoxin domain	4.39
Pp3c9_24460	Cytochrome c oxidase subunit VIII/photosystem I reaction centre subunit IX	4.35

Pp3c12_22330	Late embryogenesis abundant protein (LEA_4)	4.25
Pp3c11_14310	Unnamed family	4.24
Pp3c20_2930	Peptidase Do / Protease Do	4.16
Pp3c12_26160	Phosphatidylinositol diacylglycerol-lyase / Phosphatidylinositol phospholipase C	4.15
Pp3c11_3620	CGI-141-RELATED/LIPASE CONTAINING PROTEIN	4.12
Pp3c16_23630	Aldehyde Oxidoreductase; domain 4	4.08
Pp3c19_6700	Unnamed family	4.05

Table 7.6 List of most down-regulated genes between WT and *shlf* at p value <0.01

Gene ID	Gene name	log2FoldChange
Pp3c2_14400	Unnamed family	-1.88
Pp3c7_13340	HIGH AFFINITY NITRATE TRANSPORTER 2.5	-1.89
Pp3c15_19480	Unnamed family	-1.90
Pp3c13_18810	AQUAPORIN TRANSPORTER	-1.91
Pp3c12_6700	FAMILY NOT NAMED	-1.91
Pp3c7_1040	PHOSPHATIDYLGLYCEROPHOSPHATASE AND PROTEIN-TYROSINE PHOSPHATASE 1	-1.95
Pp3c9_2960	UNCHARACTERIZED NODULIN-LIKE PROTEIN	-1.95
Pp3c16_22860	PRICHEXTENSN	-1.96
Pp3c18_17700	Unnamed family	-2.00
Pp3c6_10310	Non-reducing end alpha-L-arabinofuranosidase / Arabinosidase	-2.07
Pp3c15_19571	Unnamed family	-2.07
Pp3c6_4610	PRICHEXTENSN	-2.10
Pp3c1_2040	ZINC FINGER FYVE DOMAIN CONTAINING PROTEIN	-2.14

Pp3c26_1140	Gamma-glutamyl hydrolase / Pteroyl-poly-alpha-glutamate hydrolase	-2.16
Pp3c10_19460	PECTINESTERASE 8-RELATED	-2.16
Pp3c26_11940	FASCICLIN-LIKE ARABINOGALACTAN PROTEIN 11	-2.17
Pp3c19_15950	Oxysterol-binding protein	-2.17
Pp3c14_25670	METHYLESTERASE 12, CHLOROPLASTIC-RELATED	-2.18
Pp3c16_25520	Protein of unknown function (DUF3445) (DUF3445)	-2.20
Pp3c22_22380	ALPHA-AMYLASE	-2.24
Pp3c6_25650	PRICHEXTENSIN	-2.28
Pp3c8_19540	Xyloglucan:xyloglucosyl transferase / Xyloglucan endotransglycosylase	-2.34
Pp3c7_19020	Phosphoglycerate dehydrogenase / Phosphoglyceric acid dehydrogenase	-2.36
Pp3c23_21120	Unnamed family	-2.37
Pp3c11_5230	Tetratricopeptide repeat-containing domain	-2.39
Pp3c16_10460	Unnamed family	-2.50
Pp3c25_14630	PTHR11527:SF110 - 17.4 KDA CLASS I HEAT SHOCK PROTEIN-RELATED	-2.53
Pp3c20_11960	3-phytase / Phytate 6-phosphatase	-2.54
Pp3c14_12730	SERINE-RICH PROTEIN-LIKE PROTEIN	-2.81
Pp3c22_4140	Chloramphenicol acetyltransferase-like domain	-2.82
Pp3c17_14500	COPPER CHAPERONE FOR SUPEROXIDE DISMUTASE	-2.84
Pp3c15_10680	PLASMA MEMBRANE IRON PERMEASE	-2.94
Pp3c6_5720	PTHR10209:SF155 - 2-OXOGLUTARATE (2OG) AND FE(II)-DEPENDENT OXYGENASE SUPERFAMILY PROTEIN-RELATED	-2.96
Pp3c14_10900	PTHR12461:SF46 - HYPOXIA-INDUCIBLE FACTOR 1 ALPHA INHIBITOR-RELATED	-3.11
Pp3c11_19760	PATHOGENESIS-RELATED PROTEIN	-3.11
Pp3c17_14510	SUPEROXIDE DISMUTASE [FE] 2, CHLOROPLASTIC	-3.11

Pp3c13_16030	BURP domain (BURP)	-3.16
Pp3c11_14210	Unnamed family	-3.17
Pp3c13_3830	SHN SHINE , DNA BINDING / TRANSCRIPTION FACTOR	-3.19
Pp3c3_14700	NADH-ubiquinone reductase complex 1 MLRQ subunit (B12D)	-3.30
Pp3c20_2890	Unnamed family	-3.40
Pp3c19_6200	RNA-directed DNA polymerase / Revertase	-3.42
Pp3c5_3169	PROKAR_LIPOPROTEIN	-3.85
Pp3c14_22870	RNA binding K homology domain type I	-4.49
Pp3c6_3710	17.6 KDA CLASS I HEAT SHOCK PROTEIN 2	-4.82
Pp3c24_290	Very-long-chain 3-oxoacyl-CoA synthase / Very-long-chain beta-ketoacyl-CoA synthase	-4.86
Pp3c27_1590	DnaJ domain (DnaJ) // 4Fe-4S single cluster domain of Ferredoxin I (Fer4_13)	-5.05
Pp3c5_22920	CHLOROPHYLL A/B BINDING PROTEIN	-5.17
Pp3c11_18420	Unnamed family	-6.94
Pp3c9_21600	Unnamed family	-10.91

Table 7.7 List of all metabolites differentially regulated between WT and *shlf*

Metabolites M.wt/retention time (s)	Fold Change	log2(FC)	Metabolites M.wt/retention time (s)	Fold Change	log2(FC)
333.1916__643.58	354.23	8.4685	347.2072__664.15	7.3805	2.8837
265.1476__835.49	0.03644	-4.7783	494.2471__615.69	7.1764	2.8433
537.2336__608.29	18.884	4.2391	519.2664__408.64	0.14663	-2.7698
587.3326__657.17	18.091	4.1772	348.1791__667.66	6.7257	2.7497
369.1251__827.4	16.941	4.0824	379.187__638.15	6.4684	2.6934

362.9699__597.64	0.063022	-3.988	186.0535__711.93	0.15611	-2.6793
538.2368__610.44	15.603	3.9638	175.0464__123.25	0.15939	-2.6494
372.1016__466.91	15.067	3.9133	430.8354__126.29	0.1604	-2.6402
493.2437__609.01	14.328	3.8407	479.177__583.4	6.2232	2.6377
123.0451__466.72	14.051	3.8126	545.0637__467.21	6.2075	2.634
546.0658__465.61	11.452	3.5175	467.0863__466.61	5.7448	2.5223
327.1085__466.91	11.253	3.4923	391.1652__663.54	5.7006	2.5111
475.1564__635.94	10.976	3.4563	317.0906__711.93	0.18029	-2.4716
539.2357__626.79	10.854	3.4402	265.0716__567.14	5.493	2.4576
551.2701__574.5	0.09672	-3.37	308.1344__684.02	0.18463	-2.4373
328.1118__467	10.233	3.3552	572.7897__140.09	0.19429	-2.3637
163.0947__596.22	0.11028	-3.1807	413.1587__636.8	5.1421	2.3623
480.1802__582.45	8.8435	3.1446	289.0619__580.81	4.9802	2.3162
371.0984__462.26	8.4984	3.0872	161.0243__508.88	4.9785	2.3157
495.2499__627.33	8.3054	3.054	292.1846__638.15	4.9453	2.3061
499.1302__158.32	8.2506	3.0445	286.0437__570.23	4.8563	2.2799
286.0436__607.2	4.8165	2.268	339.1996__689.34	4.0695	2.0248
373.1108__459.28	4.7887	2.2596	278.1198__517.81	4.0115	2.0041
331.176__623.11	4.7648	2.2524	291.1705__637.35	3.9993	1.9998
335.1971__637.07	4.735	2.2434	351.1294__423.81	3.968	1.9884
597.1062__149.1	4.6816	2.227	324.0578__91.43	0.25213	-1.9878
357.1036__93.63	4.5482	2.1853	438.1514__429.23	0.25314	-1.982
395.1552__430.63	4.4886	2.1663	589.0974__586.8	3.9421	1.979
400.2342__657.86	4.4842	2.1649	273.0379__158.39	0.25621	-1.9646
378.185__637.25	4.4417	2.1511	353.8581__125.35	0.25859	-1.9513
377.1815__637.03	4.3987	2.1371	261.0917__94.19	0.26028	-1.9419

387.1143__118.42	4.3879	2.1335	411.1508__413.93	3.8302	1.9374
474.8126__139.72	0.22813	-2.1321	271.0996__596.53	3.8166	1.9323
467.0529__360.47	4.3784	2.1304	371.8684__126.29	0.2627	-1.9285
334.195__638.17	4.2988	2.1039	569.0722__691.99	3.7818	1.9191
151.0036__580.83	4.2931	2.102	474.173__447.61	3.7767	1.9171
388.0962__482.31	4.2833	2.0987	147.0451__567.78	3.7657	1.9129
224.0677__92.64	0.23568	-2.0851	242.0786__92.11	0.26578	-1.9117
512.2429__527.54	4.2399	2.084	433.8333__36.48	3.7082	1.8907
391.028__834.23	4.2367	2.0829	293.891__588.81	0.27301	-1.873
591.1396__588.48	4.2245	2.0788	277.1657__610.32	3.6521	1.8687
590.1012__586.94	4.2188	2.0768	265.1476__702.27	0.27439	-1.8657
586.8536__139.13	0.23904	-2.0647	306.1639__590.55	3.6365	1.8626
349.1857__596.92	4.164	2.058	230.1397__591.12	3.6164	1.8546
223.2064__642.82	0.27863	-1.8436	274.8808__437.87	3.16	1.6599
319.1761__610.73	3.5769	1.8387	260.0889__92.92	0.3202	-1.6429
537.1307__610.4	3.5464	1.8263	512.8228__139.72	0.32291	-1.6308
320.1793__610.34	3.5359	1.8221	429.1386__544.09	0.32365	-1.6275
387.1144__714.81	3.5246	1.8175	321.1189__517.81	3.0782	1.6221
370.8687__124.92	0.28407	-1.8157	277.1291__517.45	3.0717	1.619
459.1812__541.28	0.28582	-1.8068	343.1301__496.02	3.0707	1.6186
287.056__580.61	3.4792	1.7988	281.0667__496.64	3.0611	1.6141
473.1696__637.2	3.4697	1.7948	367.0669__555.55	3.0533	1.6104
288.0595__581	3.4676	1.7939	597.1358__476.78	3.024	1.5965
500.1337__124.95	3.4611	1.7912	274.8807__671.42	3.0182	1.5937
211.1338__837.98	0.28898	-1.791	557.2335__528.81	2.9991	1.5845
258.0386__130.62	0.29027	-1.7846	450.2061__582.64	2.9955	1.5828

377.2175__593.97	3.4211	1.7745	293.8911__437.67	2.9875	1.5789
571.0875__674.85	3.3668	1.7514	431.8353__30.79	0.33504	-1.5776
305.1604__589.35	3.3569	1.7471	414.8461__140.26	0.3364	-1.5717
296.8694__143.91	0.29961	-1.7388	373.06__403.81	2.967	1.569
394.8459__141.78	0.30398	-1.718	511.2395__527.18	2.9602	1.5657
506.0906__481.8	0.30423	-1.7168	397.2052__825.77	0.33977	-1.5573
552.1483__637.31	3.2517	1.7012	499.1531__80.36	0.34039	-1.5547
428.0917__513.26	3.2079	1.6816	503.212__573.98	2.9117	1.5418
495.0839__517.31	3.197	1.6767	352.8581__294.12	0.34392	-1.5399
499.1306__125.62	3.1839	1.6708	395.1193__423.8	2.8985	1.5353
552.2739__578.66	0.34556	-1.533	227.1033__528.37	2.6562	1.4094
408.1955__575.32	2.8919	1.532	180.0668__124.52	0.37656	-1.409
212.1372__612.55	2.8889	1.5305	186.1135__514	2.6491	1.4055
211.1337__611.76	2.8761	1.5241	174.9557__54.06	2.6396	1.4003
161.0092__143.18	2.8472	1.5095	455.1403__413.93	2.6238	1.3917
389.0878__573.35	2.8365	1.5041	432.8336__29.76	0.38169	-1.3895
412.8251__36.88	2.8275	1.4995	534.0346__136.65	2.6149	1.3867
551.1463__637.2	2.8205	1.4959	511.1679__662.02	2.6131	1.3858
443.1537__570.02	0.35747	-1.4841	260.1138__515.83	0.38294	-1.3848
353.1817__512.64	2.7911	1.4809	261.135__409.75	0.38475	-1.378
553.1475__637.25	2.791	1.4808	410.1569__500.43	0.38563	-1.3747
186.9734__832.15	0.35996	-1.4741	254.1398__602.56	2.5924	1.3743
275.1038__455.69	0.36057	-1.4717	565.1625__635.94	2.5897	1.3728
289.1293__531.22	2.7692	1.4694	376.2784__816.8	2.5851	1.3702
568.143__548.62	0.36155	-1.4678	489.1658__83.93	0.3878	-1.3666
359.0982__439.31	2.7502	1.4595	234.9219__124.88	0.38824	-1.365

556.2325__527.79	2.7195	1.4434	294.8892__126.29	0.38827	-1.3649
335.8798__140.5	0.36892	-1.4386	421.0566__570.73	2.5678	1.3605
596.1334__476.71	2.7011	1.4335	435.0535__134.61	2.5643	1.3586
312.8998__124.52	0.37098	-1.4306	310.9019__441.16	2.5604	1.3563
432.856__139.41	0.3712	-1.4297	370.8687__439.53	2.56	1.3561
569.1426__548.97	0.3743	-1.4177	169.9627__775.86	0.3919	-1.3514
431.0831__482.87	2.6613	1.4121	575.146__176.6	2.542	1.346
557.1085__619.95	2.5269	1.3374	389.1755__467.71	0.41041	-1.2849
310.9022__124.58	0.39619	-1.3357	595.1303__476.58	2.4363	1.2847
210.1216__605.07	2.52	1.3334	587.0825__595.57	2.4352	1.284
274.1557__602.2	2.5186	1.3326	291.1449__561.94	2.4304	1.2812
352.8898__136.69	0.39718	-1.3321	401.1086__482.22	2.429	1.2804
431.202__670.38	2.5151	1.3306	384.1015__512.23	2.4225	1.2765
341.3058__833.76	0.39866	-1.3268	246.0984__453.93	0.41323	-1.275
494.2753__647.23	0.39864	-1.3268	302.1012__66.21	0.41382	-1.2729
233.9244__126.29	0.39918	-1.3249	202.1085__503.53	2.4155	1.2723
273.1536__602.2	2.4955	1.3193	384.2029__591.93	2.415	1.2721
555.2295__527.88	2.494	1.3184	420.0674__451.62	2.4149	1.272
173.038__617.16	2.4921	1.3174	433.095__410.22	0.41425	-1.2714
272.231__812.06	0.40184	-1.3153	500.2778__769.52	2.3912	1.2577
311.9017__124.39	0.40235	-1.3135	208.0617__467.37	2.3906	1.2574
294.1184__468.4	0.40267	-1.3123	236.9027__143.18	0.41896	-1.2551
362.9696__833.46	0.40268	-1.3123	361.0904__283.13	2.3711	1.2455
449.8459__26.31	2.4832	1.3122	136.9367__584.29	0.42316	-1.2407
310.9018__618.18	2.4815	1.3112	270.9256__30.94	2.3621	1.2401
291.0694__100.71	0.40333	-1.31	272.1503__602.56	2.356	1.2363

375.2751__816.01	2.4791	1.3098	383.0981__513.03	2.3513	1.2335
273.1341__834.2	0.40394	-1.3078	383.083__601.88	2.3472	1.231
449.2029__611.12	2.4722	1.3058	574.1055__607.53	2.3466	1.2306
334.8795__141.13	0.41028	-1.2853	289.1658__554.17	2.3426	1.2281
302.2205__826.42	0.42841	-1.2229	485.2821__836.73	0.44445	-1.1699
586.4396__748.99	2.3342	1.2229	269.0545__126.75	0.44518	-1.1675
307.0796__112.72	0.42873	-1.2219	356.0917__98.25	2.2421	1.1649
365.1444__508.85	2.3312	1.2211	403.1472__108.57	0.44663	-1.1628
516.2011__538.56	2.3259	1.2178	505.0728__168.18	2.2381	1.1623
277.0562__130.84	2.3228	1.2159	567.1416__552.66	0.44886	-1.1557
250.1014__360.45	2.3214	1.215	583.1362__481.78	0.45022	-1.1513
314.88__144.29	0.4315	-1.2126	190.0509__499.24	2.2189	1.1499
570.0761__583.72	2.3174	1.2125	436.1214__94.13	0.45117	-1.1483
585.4371__749.27	2.3141	1.2105	345.0826__427.04	2.2146	1.1471
293.1239__455.75	0.43243	-1.2095	317.212__771.95	2.2063	1.1416
258.0981__741.73	0.43429	-1.2033	270.1346__593.38	2.2056	1.1412
354.8561__31.62	2.2996	1.2014	364.2055__589	2.2013	1.1383
387.0929__483.18	2.2911	1.1961	533.0314__135.06	2.1971	1.1356
342.1093__489.34	2.2894	1.195	309.1165__318.2	2.1862	1.1284
466.8855__138.7	0.43722	-1.1936	294.9275__139	0.45875	-1.1242
351.0807__98.4	0.43847	-1.1894	405.2125__611.07	2.1778	1.1229
505.1269__545.51	2.2785	1.1881	440.2295__608.95	0.46134	-1.1161
293.8913__124.92	0.44031	-1.1834	491.1499__554.51	2.1606	1.1114
192.0664__531.65	0.44048	-1.1828	415.088__443.36	2.1559	1.1083
570.0761__635.8	2.2693	1.1822	507.1313__642.34	0.46427	-1.107
510.8109__33.74	2.2644	1.1791	386.0987__489.28	2.1536	1.1067

271.2276__812.06	0.44371	-1.1723	292.1482__562.3	2.1495	1.104
294.0912__359.96	2.1323	1.0924	518.1447__85.55	2.0695	1.0493
571.0866__632.5	2.1288	1.0901	436.0577__134.64	2.0649	1.0461
338.0811__131.92	2.1282	1.0896	172.9579__36.79	2.0644	1.0457
318.0938__711.79	0.47148	-1.0847	270.9325__138.35	0.48449	-1.0455
276.1095__92.7	0.47296	-1.0802	393.1066__711.79	0.48475	-1.0447
319.0528__162.74	2.1141	1.0801	288.9093__30.08	0.48574	-1.0417
273.0149__103.01	2.1061	1.0746	377.0646__451.69	2.0576	1.041
488.8771__138.73	0.47523	-1.0733	307.1761__555.57	2.0515	1.0367
570.076__683.51	2.0949	1.0669	234.9137__155.5	0.48857	-1.0334
301.217__826.42	0.47745	-1.0666	329.0878__453.14	2.0461	1.0329
292.8914__531.22	0.47808	-1.0647	488.1626__81.25	0.49011	-1.0288
276.8786__31.26	2.0896	1.0632	258.1165__564.99	2.0353	1.0252
237.1104__652.92	0.47862	-1.063	332.8907__145.14	0.49264	-1.0214
256.119__565.2	2.0882	1.0623	307.1553__639.13	0.4945	-1.016
175.0251__99.77	0.47903	-1.0618	219.0511__122.17	0.49494	-1.0147
388.1725__468.4	0.47909	-1.0616	392.9043__139.75	0.49525	-1.0138
218.9633__667.17	2.0839	1.0593	397.1135__529.12	0.49534	-1.0135
509.0998__563.65	2.0802	1.0567	416.9661__154.99	0.49745	-1.0074
383.1868__397.58	2.0773	1.0547			
473.2821__833.51	0.48199	-1.0529			
589.3336__767.58	2.0728	1.0516			
571.0867__607.53	2.072	1.051			
233.9242__711.15	0.48306	-1.0497			

Curriculum Vitae – SHIRSA PALIT

A. Personal data

Name Shirsa Palit
Date and place of birth 20th September, 1994, India
Sex Female
Marital Status Single
Nationality Indian



Address at work Department of Biology, Indian Institute of Science Education and Research (IISER) Pune, Dr.Homi Bhabha Road, Pashan, Pune – 411008, Maharashtra, India

Phone: +91-9628367595

Contact

Email: palit.shirsa@students.iiserpune.ac.in

Web: <http://sites.iiserpune.ac.in/~akb/phd.html>

B. Education

PhD in Biological Science, Indian Institute of Science Education and Research (IISER), Pune, Maharashtra, India

Aug. 2017-Current

Supervisor: Prof. Anjan K Banerjee, Biology, IISER Pune, Maharashtra, India

Thesis: Investigating the molecular role of bryophyte-specific protein SHORT-LEAF (SHLF) in gametophore development of moss (*Physcomitrium patens*)

M.Sc. in Biotechnology awarded **cum laude**, Allahabad University, Allahabad, Uttar Pradesh, India

Aug. 2015-July 2017

Supervisor: Dr. Awadh Bihari Yadav, Allahabad University, Uttar Pradesh, India

Thesis: Antioxidant properties of mushroom extract loaded chitosan nanoparticles

Summer dissertation project, CSIR-IGIB, Delhi, India

April-July, 2016

Supervisor: Dr. Ashok K Singh, Institute of Genomics and Integrative Biology (CSIR-IGIB), Delhi, India

Thesis: Recent advances in cardiac troponin I based sensors for detection of human heart attack

Aug. 2012-July 2015

B.Sc in Biotechnology, Biochemistry and Genetics awarded cum laude from Bangalore University, Bangalore, Karnataka, India

C. Work experience:

Aug. 2017-Current: PhD Research Scholar in Molecular Plant Biology lab at IISER Pune, Maharashtra, India

D. Research Interests:

- Comparative regulation of seed and non-seed plant innate immunity in response to biotic and abiotic stress.
- Evolution and functional conservation of immune signaling components in plants
- Role of secretory peptides and small molecules in plant development and defense response
- Translational and synthetic biology using moss as a platform for heterologous protein expression

E. Professional memberships:

- American Society of Plant Biologists (ASPB, USA)
- ASPB Early Career Plant Scientists section (ASPB-ECPS, USA)

F. Publications:

- **Palit S.**, Bhide A J., Mohanasundaram B., Pala M., & Banerjee, A. K. (2023) Secretory peptides from conserved tandem repeats of SHORT-LEAF regulate gametophore development in moss. (Under revision in **Plant Physiology**)
- Mohanasundaram, B., **Palit, S.**, Bhide, A. J., Girigosavi P & Banerjee, A. K (2023). SCARECROW regulated leaf blade and lamina development in the moss (*Physcomitrium patens*) (Under revision in **Plant molecular biology**)
- Mohanasundaram, B., Bhide, A. J., **Palit, S.**, Chaturvedi, G., Lingwan, M., Masakapalli, S. K., & Banerjee, A. K. (2021). The unique bryophyte-specific repeat-containing protein SHORT-LEAF regulates gametophore development in moss. **Plant Physiology**, 187(1), 203-217.
- Bhatnagar, D., **Palit, S.**, Singh, M. P., Kaur, I., & Kumar, A. (2016). Recent advances in cardiac troponin I based sensors for detection of human heart attack. **Cellular and Molecular Biology**, 62(3), 1000142.

G. Manuscripts under preparation:

- **Palit S.**, Bhide A J., Banerjee A.K (2023) Small secretory SHLF peptides mediate innate auto-immune responses in the moss *P. patens*
- **Palit S.**, Agarwal P, Banerjee A.K (2023) Bryophyte specific SHLF peptides induce autoimmune response in flowering plants.

H. Awards and Honor:

- First prize for outstanding oral presentation at the International Conference on “Current Trends and Future Prospects of Plant Biology (CTFPPB-2023), University of Hyderabad, Telangana, India
- Best Poster presentation award at EMBO 2022- An Integrated View of Early land Plant Evolution, NISER Bhubaneshwar
- ASPB travel award to attend ASPB Plant Biology 2022, Portland, Oregon
- Infosys travel award to attend ASPB Plant Biology 2022, Portland, Oregon
- Recipient of Gold medal for securing first position in MSc Biotechnology batch of 2017, Allahabad University, Allahabad, Uttar Pradesh, India
- Graduate Aptitude Test in Engineering- GATE (Feb. 2017)
- Junior Research Fellowship- CSIR NET (June 2016 and December 2016)
- Recipient of Gold medal for securing first position in BSc Biotechnology, Biochemistry and Genetics batch of 2015, Bangalore University, Bangalore, Karnataka, India

I. Conferences attended:

- Oral talk presenter at International Conference on “Current Trends and Future Prospects of Plant Biology (CTFPPB-2023), 23rd – 25th February 2023, University of Hyderabad, Telangana, India
- Poster presenter at EMBO 2022- An Integrated View of Early land Plant Evolution, November 8th -11th, 2022, Bhubaneshwar, INDIA, “*Secretory peptides of the bryophyte-specific protein SHORT-LEAF regulate gametophore development in moss P. patens*”
- Poster presenter at ASPB Plant Biology 2022, Portland, Oregon, USA 9th -13th July 2022, “*Characterization of SHORT-LEAF (SHLF) domains indicates that SHLF acts through small secretory cryptic peptides to regulate moss gametophore development*”
- International Conference on Plant Developmental Biology & 3rd National Arabidopsis Meeting, December 12–16, 2017, NISER, Bhubaneshwar, India.

J. Teaching Experience:

Teaching Assistant

- Undergraduate Courses on Basic Cell biology, and plant developmental biology at IISER Pune.
- Experienced in scientific and laboratory training of undergraduate and junior graduate students.

K. Research Skills:

- **Molecular Biology and biochemistry:** Specialized PCR techniques, RT- qPCR analysis, conventional and gateway cloning, immunoblotting, DNA and RNA isolation, genome editing, recombinant protein expression in heterologous systems, co-immunoprecipitation, protein expression and purification, In-house antibody generation in mammalian systems (rabbits), synthetic peptide supplementation assays, studying cis-regulatory elements using ChIP-Seq, RNA-Seq analyses, Whole genome sequencing, sample and library preparation for omics characterization.
- **Stress Biology:** Studying the molecular and physiological effect of biotic and abiotic stressors on plant growth and development.
- **MS/MS and LC/MS-based** detailed targeted and non-targeted proteomics, peptidomics and metabolomics from total plant tissues and secretomes.
- **Bioinformatics and statistics:** NGS techniques including Whole Genome Sequencing, RNA Seq analysis (ILLUMINA), AlphaFold2-based protein structure prediction, Constructing and inferring phylogenetic trees, Custom designing NGS pipelines for data mining from hundreds of genomes and transcriptomes using Python. Advanced statistical analysis using R and Python. Experience with PC and UNIX operating systems. Experience with databases of plant genomes (Phytozome), and methods of analysis of biological sequence data including R software suite.
- **Genetics:** Mutagenesis, mutant population establishment, maintenance and cryopreservation of

moss. Microdissection of moss leaf for phenotyping. Forward and reverse genetics-based gene-discovery in moss.

- **Microscopy:** Advanced confocal microscopy and TEM-based and apotome-based, analysis to study subcellular protein localization, intercellular protein trafficking, cellular organelles and intercellular structures.
- **Plant physiology and tissue culture:** Culturing, maintenance, protoplast and agrobacterium mediated transformation and phenotypic characterization of transgenic *Physcomitrium*, *Marchantia*, *Nicotiana* and *Arabidopsis*.
- **Grant proposal writing:** I am experienced in writing national research grant proposals. Recently, the Council for Scientific and Industrial Research (CSIR-Govt. Of India) accepted our translational research grant for heterologous protein expression using moss-based photobioreactors.

L. Soft skills:

I am a proactive and persistent critical thinker with good communication, networking, organization and interpersonal skills. I am highly motivated and passionate about science and am courageous enough to comfortably deal with challenging projects. I am willing to take risks and learn from all my experiences. I am resilient in the face of adversity and approach problems with a solution-based approach. I am comfortable both in assuming leadership in a project and/or working in a collaborative manner to attain set goals in a timely manner.

M. Hobbies and Co-curricular activities:

- Arts and Literary: 5-year piano course certified by the London School of Music.
- Fine Arts: 5-year Post Graduate Diploma in Fine Arts
- Avid bibliophile and aspiring poet
- Amateur nature photographer
- Design and maintenance of terrariums with indigenous moss species

