Forward and reverse genetic approaches to investigate cellulose biosynthesis in *Physcomitrium patens*

by

Sara Behnami

A thesis submitted to the
School of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Applied Bioscience

Faculty of Science

University of Ontario Institute of Technology (Ontario Tech University)

Oshawa, Ontario, Canada

April 2021

THESIS EXAMINATION INFORMATION

Submitted by: Sara Behnami

Doctor of Philosophy in Applied Bioscience

Thesis title: Forward and reverse genetic approaches to investigate cellulose biosynthesis in *Physcomitrium patens*

An oral defense of this thesis took place on March 8th, 2021, in front of the following examining committee:

Examining Committee:

Chair of Examining Committee	Dario Bonetta, PhD
Research Supervisor	Dario Bonetta, PhD
Examining Committee Member	Andrea Kirkwood, PhD
Examining Committee Member	Sean Forrester, PhD
University Examiner	Julia Green-Johnson, PhD
External Examiner	Keiko Yoshioka, PhD, University of Toronto

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

Abstract

Cellulose biosynthesis is a common feature of land plants and involves multimeric complexes composed of cellulose synthase A (CESA) proteins and other structural proteins. The exact stoichiometry of CESA proteins and interactions among proteins within cellulose synthase complex (CSC) is not well understood. Therefore, cellulose biosynthesis inhibitors (CBIs) are useful tools in decoding fundamental aspects of cellulose biosynthesis. Here, I characterize the CBI indaziflam, with a unique mode of action for resistance management, which prevents plant growth by inhibiting cellulose biosynthesis.

Arabidopsis thaliana indaziflam resistant mutants were identified through forward genetic screening. Since indaziflam is also active in moss *Physcomitrium patens*, a forward genetic approach to screen indaziflam resistance was applied on *P. patens,* and positional cloning combined with next-generation sequencing revealed two point mutations in CULLIN1 (CUL1) and AUXIN/INDOLE-3-ACETIC ACID INDUCED (Aux/IAA) which both are involved in auxin signaling pathways. The mutants were also cross-resistant to synthetic auxin 2,4-D. It is predicted that indaziflam affects plant growth and development and impacted the production and remodeling of plant cell wall directly or indirectly.

Moreover, to gain insight into the nature of the protein composition of CSCs, I employed a strategy called Biotin identification (BioID), aimed at identifying proximate and vicinal proteins *in vivo* associated with CESAs in *P. patens*. I generated multiple BioID-CESA translational fusions by homologous recombination to identify biotinylated proximate proteins. Due to limitations, including but not limited to different behaviors of fused proteins tagged at the C- or N-terminus, decreased expression level, longer incubation time with biotin, higher incubation temperature, and large size BirA* tag, I was not able to identify any interacting proteins.

Another finding of this thesis is that *P. patens* can be used to produce known natural products that are difficult to obtain by chemical synthesis. An *in vivo* combinatorial biosynthesis approach was pursued in *P. patens* to obtain rare cannabinoids with beneficial biological activity. The outcome was to produce rare cannabinoids and some pathway

intermediates. This idea's other significant result is designing different drug-candidate-producing moss strains, especially within the chemical class of cannabinoids.

Keywords: *Physcomitrium patens*; Cellulose biosynthesis inhibitor (CBI); BioID; Auxin; combinatorial biosynthesis

AUTHOR'S DECLARATION

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology (Ontario Tech University) to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Ontario Institute of Technology (Ontario Tech University) to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

_____ Sara Behnami

STATEMENT OF CONTRIBUTIONS

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication. I have used standard referencing practices to acknowledge ideas, research techniques, or other materials that belong to others. Furthermore, I hereby certify that I am the sole source of the creative works and/or inventive knowledge described in this thesis.

This thesis is dedicated to my parents,
Tahereh Ghamati and Ghodrat Behnami.

Thank you for your constant love and support every single day.
I am eternally grateful for all of the sacrifices you have
made so that I could succeed.

ACKNOWLEDGEMENTS

I would like to express my profound appreciation and the highest level of admiration for my principal investigator, Dr. Dario Bonetta. You truly inspire me to do my best work each day. Thank you for your constant support, guidance, patience and for giving me the freedom to learn throughout this degree.

A special thanks to my committee members: Dr. Andrea Kirkwood and Dr. Sean Forrester. Your support and advice have been invaluable to me.

To all the Bonetta lab members, Dr. Elysabeth Reavell-Roy, Zachary Kileeg, Luis Salgado, and Dr. Patricia Stronghill, thank you for your help. Thank you for setting aside time in your schedules to discuss and troubleshoot problems. I am highly grateful for your friendship each and every day in the lab. Work would not have been near as enjoyable without you.

Thank you to Dr. Julian Northey and his company Frontier Agri-Science.

My sincere gratitude also goes to Dr. Maryam Daneshpour at Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences. You were the first scientist to trust me in a research setting and give me the opportunity to learn. I found my passion for research in your lab, and I am forever grateful for the opportunity to work with you.

I would like to thank my family and friends, to whom I owe a great deal. To my late dad, Ghodrat Behnami, thank you for showing me that the key to life is enjoyment. To my brother, Kasra Behnami, thank you for your support, and constantly encouraging me and reminding me to take it one step at a time. And finally, the one person who has made this all possible has been my mom Tahereh Ghamati. She has been a constant source of support and encouragement and has made an untold number of sacrifices for the entire family, and specifically for me to continue my life journey. She is a great inspiration to me. Hence, great appreciation and enormous thanks are due to her, for without her understanding, I am sure this thesis would never have been completed.

I thank you all.

Table of Contents

List of figures

List of tables

List of abbreviations and symbols

®	Registered Sign
°C	Degrees Celsius
2,4-D	2,4-Dichlorophenoxyacetic Acid
3-IAA	3-Indole Acetic Acid
AAE1	Acyl-Activating Enzyme 1
ABA	Abscisic Acid
ABP1	Auxin Binding Protein1
acid	Tetrahydrocannabinolic Acid
AFB	Auxin Signaling F-Box
AFLP	Amplified Fragment Length Polymorphism
AP-MS	Affinity Purification Coupled With Mass Spectrometry
ARF	Auxin Response Factor
ARPC3	Actin Related Protein 2/3 Complex Subunit 3
ASK1	Skp1-Like
Aux/IAA	Auxin/INDOLE-3-ACETIC ACID
AUX1/LAX	Auxin1/Like Auxin Permease
AXR	Auxin-Resistant
AXR1	Auxin Resistant 1
BSA	Bovine Serum Albumin
CaMV	Cauliflower Mosaic Virus
CBD	Cannabidiol
CBDA	Cannabidiolic Acid
CBDAS	Cannabidiolic Acid Synthase
CBG	Cannabigerol
CBI	Cellulose Biosynthesis Inhibitor
CDS	Coding Sequence
CESA	Cellulose Synthase A
CH	Cullin Homology Domain
CK	Cytokinin
CMC	4-Chloro-2(Methylthio)Pyrimidine-5-Carbontrile
CMF	Cellulose Micro-And Microfibrils
CMT	Cortical Microtubules
co-IP	Co-Immunoprecipitation
CR	Conserved Region
CR1	Cullin Repeat 1
cryo-EM	Cryo-Electron Microscopy
CSC	Cellulose Synthase Complex
CSL	Cellulose Synthase-Like
cSNP	Coding SNP
CSR	Class-Specific Region
Ct	Threshold Cycle
CTAB	Cetyltrimethylammonium Bromide
CUL1	Cullin1

DCB	2,6-Dichlorobenzonitrile
DPC	4,6-Dichloro Pyrimidine-5-Carbonitrile
EDTA	Ethylenediaminetetraacetic Acid
eFP	Electronic Fluorescent Pictographs
EPA	Environmental Protection Agency
EST	Expressed Sequence Tag
FLP	Flippase
FPX	Flupoxam
FRT	Flippase Recognition Target
GC	Gas Chromatography
GC-FID	Gas Chromatography-Flame Ionization Detector
GH3	Gretchen Hagen 3
Gly	Glycine
Gr	Gransden
GT	Glycosyltransferase
HC	Hemicellulose
HRP	Horseradish Peroxidase
IBA	Indole-3-Butyric Acid
IBR5	Iba Response5
ISX	Isoxaben
izr	Indaziflam Resistant
kDa	Kilodalton
KO	Knock Out
KOR1	KORRIGAN Gene
LB	Lysogeny Broth
LG	Linkage Group
LRL	Roothairless Like
MAP20	Microtubule Association Protein 20
MW	Molecular Weight
MYA	Million Years Ago
NAA	1-Naphthalene Acetic Acid
NEDD8	Neural Precursor Cell-Expressed, Developmentally Down-Regulated Protein8
NPA	1-N-Naphthylphthalamic Acid
NPTII	Neomycin Phosphotransferase
NRT	No-Reverse Transcriptase
NTD	N-Terminal Domains
NTP	No-Template
OAC	Olivetolic Acid Cyclase
OLS	Olivetol Synthase
ORF	Open Reading Frame
PCR	Plant Conserved Region
PCR	Polymerase Chain Reaction
PCW	Primary Cell Wall
PEG	Polyethylene Glycol

Phe	Phenylalanine
PIN	Pin-Formed
PM	Plasma Membrane
PT4	TPA-Exp-Prenyltransferase 4
qPCR	Quantitative PCR
qRT-PCR	Quatritative Reverse Transcriptase PCR
RBX1	RING-Box1
RNAi	RNA Interference
ROC2	Regulator Of Cullins-1
ROS	Reactive Oxygen Species
RSL	Root Hair Defective-Six Like
SAUR	Small Auxin Upregulated Rna
SCF	Skp, Cullin, F-Box Complex
SCW	Secondary Cell Wall
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SKP	S-Phase Kinase Associated Protein 1
SNPs	Single-Nucleotide Polymorphisms
SOC	Super Optimal Broth With Catabolites Repression
TAE	Tris-Acetate-Edta
TBC	3-(Trifluoromethyl)Benzofuran-5-Carbonitrile
TBS	Tris-Buffered Saline
TCAG	The Centre For Applied Genomics
TEMED	Tetramethylethylenediamine
THCA	Tetrahydrocannabinolic Acid Synthase
TIR1/AFB	Transport Inhibitor Response 1/Auxin Signaling F-Box
TM	Transmembrane
Tween	Tris-Buffered Saline With 0.1% Tween
™	Trademark Symbol
UDP	Uridine Diphosphate
UTR	Untranslated Region
UV	Ultraviolet
VR	Variable Region
Vx	Villersexel K3
w/v	Weight/Volume
WGS	Whole Genome Sequencing
WT	Wild Type
YFP	Yellow Fluorescent Protein
Zn	Zinc-Finger Domain

Chapter one: Introduction and literature review

1.1 Plant cell wall biology

Finding renewable forms of energy that can compete with fossil fuels has been a significant scientific challenge. One push has been towards using non-food, biomass crops that can be used to produce energy from cellulose and other cell wall polysaccharides rather than seed starch [1, 2]. Plant biotechnology can enhance the energy density of biomass species and improve biomass conversion to biofuels. Biomass conversion to biofuels involves harvesting the biomass, drying it, and transporting it to a cellulosic ethanol refinery where it turns the cellulose, hemicelluloses (HCs), and pectin into fermentable sugars. The term saccharification denotes the process of enzymatic bioconversion of cellulose into fermentable sugars. Cellulose represents a significant source of fermentable sugars in lignocellulosic biomass. Biotechnological applications focused on reducing the recalcitrance of cellulose to enzymatic digestion have received major attention. The extensive hydrogen bonding between glucan chains in native cellulose is a major recalcitrance source [3]. Unfortunately, the combined costs of hydrolytic enzymes and biomass pretreatments have caused cellulosic feedstocks to be far more costly than starch for biofuel production [4]. With these challenges in mind, a long-term goal has been to manipulate the biosynthetic processes via genetic engineering strategies to understand the cell wall biosynthetic process and alter cellulose structure to reduce its recalcitrance to chemical and enzymatic hydrolysis.

1.2 The plant cell wall

The natural structure and chemistry of plant cell walls fulfill a vast range of functions, including the ability of cells to expand, become rigid, provide structural support and mechanical stability, transport water over great distances, protect plants from biotic and abiotic stresses, provide cell-cell communication and support growth and cell shape changes [5, 6]. Without cell walls, plants would be pliant piles of cells, more like molds than the trees and other greenery that grace our planet [7]. The plant cell wall's physical features are evident when a cell adopts its characteristic shape and size. For example,

during cell elongation, cell walls must be strong in the lateral areas but flexible in apical areas to enable cell expansion [8]. In general, the cell wall can be compared to the building block of the plant organism. Indeed, it is impossible to conceive of a viable plant cell without its cell wall. Furthermore, cell walls are the central storage of the energy from photosynthesis and most fixed carbon sources on earth [8].

In general, plant cell walls belong to one of two classes: primary and secondary cell walls. Primary cell walls (PCW) differ from secondary walls (SCW) in both composition and function [9]. Every plant cell makes a PCW early in its life during cell division and throughout cell growth [10]. All PCWs contain cellulose and a matrix of hemicelluloses and pectins, with some structural proteins [11]. In addition to PCWs, all plants deposit a thick and rigid SCW after cell growth has been completed. The SCW contains cellulose, hemicellulose, and lignin. The SCW is also the main component of wood cells found in trees. The hemicellulose in the PCW of dicots is mostly xyloglucans, while xylans are the major hemicellulose in SCW [12], and PCW contains considerably more pectins and proteins compared to SCW [7].

Cellulose is a homopolymer of glucose that is synthesized by plants and some bacteria. Beta-1,4 -glycosidic bonds join these monomers to form very long glucan chains (made up of thousands of glucose monomers). Extensive hydrogen bonding between glucan chains results in a water-insoluble, rigid, semi-crystalline polysaccharide [13]. Glucan chains are arranged in fibrillar units, which are further assembled as microfibril bundles [14, 15]. The key enzyme involved in the process of cellulose biosynthesis is called cellulose synthase A (CESA), which is a uridine diphosphate-α-D-sugar (UDP-sugar) dependent β-glycosyltransferase [16], and it belongs to the processive glycosyltransferase (GT) II family, which also includes chitin and hyaluronan synthases. GT II family enzyme can 'invert' the stereochemistry of the donor UDP-sugar's anomeric (C-1). Thus cellulose synthase generates β-linked UDP-glucose from the α-linked UDP-glucose (Figure 1.1). A catalytic aspartate (D) acting as a base, a DxD motif required to complex metal ions to bind UDP-glucose, a QxxRW motif (Q369, R370, V371, R372, and W373) in the glucan binding domain, and an additional D are the characteristic features of all CESAs [17]. Finally, it is

well accepted that elongation of the glucan chains occurs at the non-reducing end of the chain, supported by silver labeling experiments [18] and x-ray crystallography data [19].

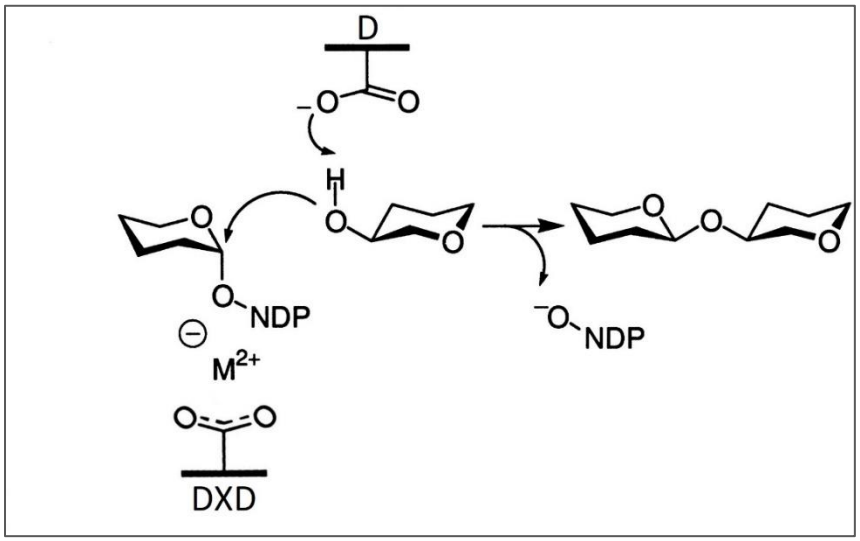


Figure 1.1: The putative mechanism for glycan synthesis. This mechanism needs two catalytic carboxylates. One to activate the acceptor species (D) and one to a divalent metal ion associated with the sugar (DXD) (Adapted from [16]).

CESA proteins function together as cellulose synthase complexes (CSCs) in the plasma membrane (Figure 1.2A) with a total of 25-30 nm diameter. The CSC asymmetrical rosette structure contains six globular protein complexes (Figure 1.2B) [20, 21]. Each lobe of the CSC rosette has a similar structure to CESA subunits [22, 23]. Recent studies have shown that 18, 24, or 36 CESA proteins per CSC may mediate microfibrils' synthesis with a corresponding maximum number of cellulose chains [24]. Seed plants contain six functional and phylogenetic classes of CESA proteins, three for PCW synthesis and three for the lignified SCW synthesis of tracheary elements and fibers [25]. Any mutations of secondary *CESA* genes cause a distinctive irregular xylem phenotype and collapsed xylem tracheary elements, and weak stems [26].

In the model plant, *Arabidopsis thaliana*, there are ten genes (*AtCESA1-10*) encoding CESA proteins [27]. Combined biochemical and genetic studies suggest three *CESA* genes, *CESA4*, *CESA7*, and *CESA8*, are necessary for cellulose deposition during SCW formation and interact directly [25]. Related evidence also suggests that *CESA1*, *CESA3*, and *CESA6* are

essential CSC members of PCW [28]. *CESA2*, *CESA5*, and *CESA9* have partial redundancy to *CESA6* [28, 29]. By comparison, there are 12 *CESA* genes in maize [30], 18 genes in poplar [31], and eight genes in barley [32].

Three CESA isoforms in both primary and secondary wall rosette CSCs are equally consistent with 18 CESAs (three per lobe) in a 1:1:1 ratio (Figure 1.2B) [33]. As cellulose is synthesized, CSCs slide within the plasma membrane (PM) underlying cortical microtubules (CMTs), which can be detected by live-cell imaging of fluorescent protein-tagged CSCs [34]. Understanding CSC dynamics during cellulose synthesis, trafficking of CSCs to and from the plasma membrane (PM), and uncovering CSC proteins' roles and some other proteins that impact cellulose synthesis are some of the most critical aspects of cell wall biology.

Plants are expressing different CESA isoforms, which often are not functionally equivalent [35]. The isoforms interact, possibly by forming heterotrimeric complexes [35]. Conserved residues mapping of the three isoforms reveals that the plant conserved region (PCR) and transmembrane (TM) region's complex interfaces are nearly identical [35]. Reconstitutions of single CESA isoforms from *Physcomitrium patens* and *Populus tremula x tremuloides* revealed that single isoforms with no additional plant-derived factors could be sufficient to form cellulose fibers resembling authentic wall micro-and microfibrils (CMFs) [36, 37]. Plant CESA contains three distinct regions that mediate oligomerization: the N-terminal domains (NTDs), PCR, and class-specific regions (CSR). NTDs form extended cytosolic structures with a RING-like domain with the hypervariable regions. PCR is inserted into the catalytic glycosyltransferase domain, and CSR follows a conserved substrate-binding motif near the active site [8, 33]. A cryo-electron microscopy (cryo-EM) structure revealed a CESA trimeric complex in the formation of CMF in SCW during cellulose biosynthesis [35, 38]. The data shows CESA trimer complexes produce cellulose chains containing three glucan chains to align with CMFs synthesized by a six CesA trimers rosette and assemble into microfibrils containing 18 cellulose chains [35].

The length of individual glucan chains differ in cell types and plant species, but in general, it can be from 1000-8000 (PCW) to as much as 16000 (SCW) glucose molecules in length [39]. Researchers speculated an analogous pore structure present in plant CESA, which

might be responsible for positioning the glucan chain's correct hydrogen bonds to emerge the CESA into the cell wall [40]. Based on this assumption, mutations within the transmembrane helices of plant CESAs could affect glucan chain threading and microfibril crystallization. Both glucan chain length (degree of polymerization) and hydrogen bonding (degree of crystallization) are two fundamental properties of cellulose that affect its enzymatic hydrolysis [41].

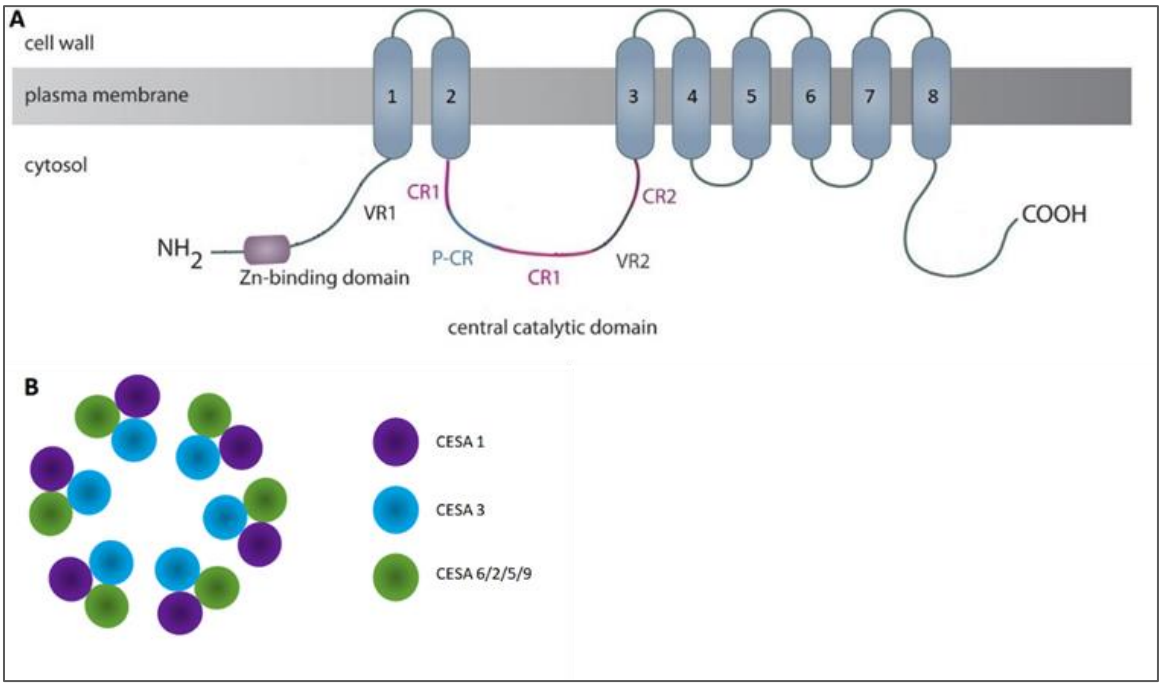


Figure 1.2: Schematic Structure of a CESA Protein and a CSC in *Arabidopsis*. (A) Domain structure of a CESA. The intracellular N-terminal domain contains a Zn binding domain and a variable region (VR) followed by two transmembrane domains. Two conserved regions (CR1 and CR2) flank on both sides by large cytoplasmic central catalytic. The cytoplasmic C-terminal domain follows the six subsequent transmembrane domains. (B) A schematic of a CSC consisting of 18 individual CESA proteins. CSC is composed of six lobes that contain three CESA isoforms (Adapted from [42]).

1.3 Moss as a model for understanding cell wall biology and as a vehicle for biotechnology

Bryophytes are ancient land plants and considered the missing link between algae and vascular plants, and their morphologies and life cycles are different from flowering plants significantly [43].

Physcomitrium patens live in open habitats and colonize damp soil or dry areas; these environments can be exposed to varying light conditions (Figure 1.3A-E)[44]. Opposite of flowering plants, in moss, the gametophyte is the haploid generation and dominant over the diploid sporophyte generation. In three months of its life cycle, mosses propagate through spores from filamentous cells called protonemata, which exhibit apical growth (Figure 1.3D). They germinate, grow, differentiate, and produce callus-like buds that develop into gametophores, an upright haploid leafy shoot. The moss life cycle's vegetative stage avoids any genetic instabilities and allows stable production [45].

Figure 1.3: Habitus and laboratory culture of *Physcomitrium patens*. (A) Natural *Physcomitrium patens* population covering rocks near Toronto, Canada. (B) Close-up of the sporophyte, comprising the short seta that bears the spore capsule with its calyptra. (C) Laboratory culture of gametophytes is growing on a solid medium in a Petri dish. (D) Filamentous protonemata are growing on a solid medium in a Petri dish. (E) Gametophytes are growing on peat pellets in a glass jar.

Leafy shoots differentiate in the haploid generation, and the gametophytes, i.e., pollen tubes and embryo sacs, are part of the haploid plant bodies in which fertilization occurs. The tissues of the gametophytic leafy shoots of mosses are much simpler than the sporophytic shoots of flowering plants. The gametophore entails a photosynthetic stem with the leaves distributed in a helical pattern around it. Along with aging gametophore, rhizoids grow from the stem base. The monoecious moss carries both sex organs on one single gametophore and self-fertilize. Archegonia and antheridia differentiate at the tips of the leafy shoots to form eggs and sperm, respectively. After gamete production and fertilization, the zygote divides to produce a matrotrophic sporophyte consisting of a foot, setum, and capsule. Spores are made within the capsule (Figure 1.3B), which eventually breaks open irregularly to release the spores [44, 46, 47].

Physcomitrium is a small plant that can grow in controlled laboratory conditions on simple media (Figure 1.3C)[48]. The dominant haploid gametophytic generation makes *Physcomitrium* an excellent model organism for genetic studies since mutations' phenotypic effects can be observed directly in the progeny. On the other hand, *Physcomitrium* has the highest reported ratio of homologous recombination to non-homologous recombination of all land plants. The favorable result of *Physcomitrium* to mutagenesis makes the moss an attractive organism for reverse genetics [49, 50].

Many moss genes have been isolated and disrupted, and sequenced genes (ESTs) projects are ongoing, and *P. patens* is the most suitable moss for genetic comparisons with the flowering plants [46]. Many gene families present in *A. thaliana* are also present in *P. patens*, hence most likely in their common ancestor. This allows for direct comparisons of gene sequences and comparative gene function [51, 52].

In contrast to vascular plants with both primary and secondary cell walls, mosses have only a primary cell wall. Among bryophytes, mosses have very well specialized cells, such as stereids, hydroids, and leptoids, similar to the higher plants' vascular tissue. Stereids have thickened cell walls specializing for support, hydroids are water-conducting cells, and

leptoids are sugar-conducting cells. Unlike the cells in the vascular tissue of higher plants, these cells lack lignin [53].

Mosses have diverged from the lineage leading to vascular plants nearly 500 MYA [54] and, thus, hold a phylogenetically informative position. Although mosses are simpler than vascular plants, many moss features are similar to those of tracheophytes, and development processes are analogous to vascular plants, and the majority of genes are homologous to vascular plant genes, too [55].

Physcomitrium patens can be genetically engineered to knock out the genes responsible for undesirable post-translational modifications [45]. Additionally, it is feasible to target any gene of interest to a well-known naturally highly expressed locus. For all these reasons, *P. patens* provides an ideal system of choice for basic molecular, cytological, developmental questions and the heterologous expression of proteins and metabolites [45].

1.4 The evolution of land plants and preadaptation

Using a bryophyte model, *Physcomitrium patens* allows comparative genomics to understand evolutionary processes that allowed land plants to thrive and become successful. Bryophytes are the earliest branching land plant lineages, representing the closest existing plants to all land plants' ancestors [56]. It is crucial to understand the evolutionary processes that allowed the first land plants' genesis to understand how the infertile land became colonized and transformed fully. It is commonly accepted that land plants evolved from aquatic green algae (Figure 1.4)[57]. Some suggested liverworts are the earliest branching lineage and sister to all other land plants [58].

In contrast, others suggested that all three bryophyte lineages, mosses, liverworts, and hornworts, were monophyletic [59]. These conflicts between these views are due to the different methods and taxa used. The early land plants lacked anatomical complexity, so the land's downfall was dependent on molecular adaptations. As a result, they would have required changes and developments in stress signaling to tolerate the variable environment on lands, such as exposure to dehydration, increased UV radiation, and the

higher and more rapid temperature fluctuations. So, to make a move onto land, they indeed required some preadoption to face the stresses.

Plant hormones regulate plant growth and development. Also, they are managing responses to external stimuli, including stresses. It has been hypothesized that some hormone biosynthetic pathway components were established in charophytes before conquesting the land [56]. The evidence for auxin biosynthetic origins in the charophytes is based on relatively low protein homology scores and questionable phylogenetics [60]. This finding proved that while charophytes contain auxin, the ability to regulate growth and development through auxin conjugation was a land plant-specific adaptation [61].

Other than hormone-mediated stress signaling, other biochemical changes were needed to help protect cells from high UV and the lack of buoyancy. Important compounds include phenolics that are found in charophytes [62]. For example, lignin-like polymers present in bryophytes as a critical evolutionary change allowing lignin synthesis a necessary adaptation in vascular plants [63]. Charophytes have also been known for having resistant cell walls, tolerating the high temperature, and the extra structural support required on land [64]. A study by Mikkelsen et al. proposed that cell wall polysaccharide complexity significantly increased in charophytes before conquering land [65].

Our knowledge of cellulose production in land plants is from the model angiosperm *Arabidopsis* [38], and most of the cellulose-related components in *Arabidopsis* have orthologs in other plant species [66]. Although these data were derived from seed plants, the 'core' components of cellulose synthesis are also present in earlier diverging lineages, including mosses (i.e., *Physcomitrium*), indicating coordinated expression of the genes is essential to arrange the synthesis activities.

Seed plants have larger *CESA* gene families in comparison to seedless plants. The hypothesis can explain this difference: rosettes in seed plants are composed of CESA subunits encoded by two or more different genes [67]. Another reason for these differences may be the cell-type-specific expression, which reflects functional or regulatory specialization. For example, in *Arabidopsis thaliana,* the genes *AtCESA4*, *AtCESA7*, and *AtCESA8* are expressed during secondary cell wall deposition in vascular tissue. *AtCESA1*,

AtCESA3, and *AtCESA6* are expressed in expanding cells during primary cell wall formation [25, 68-72]. Based on studies of evolution and functional specialization of *CESAs* and cellulose synthase-like genes (Csls), a variation of *CESAs* can be required for vascular tissues' evolution [67].

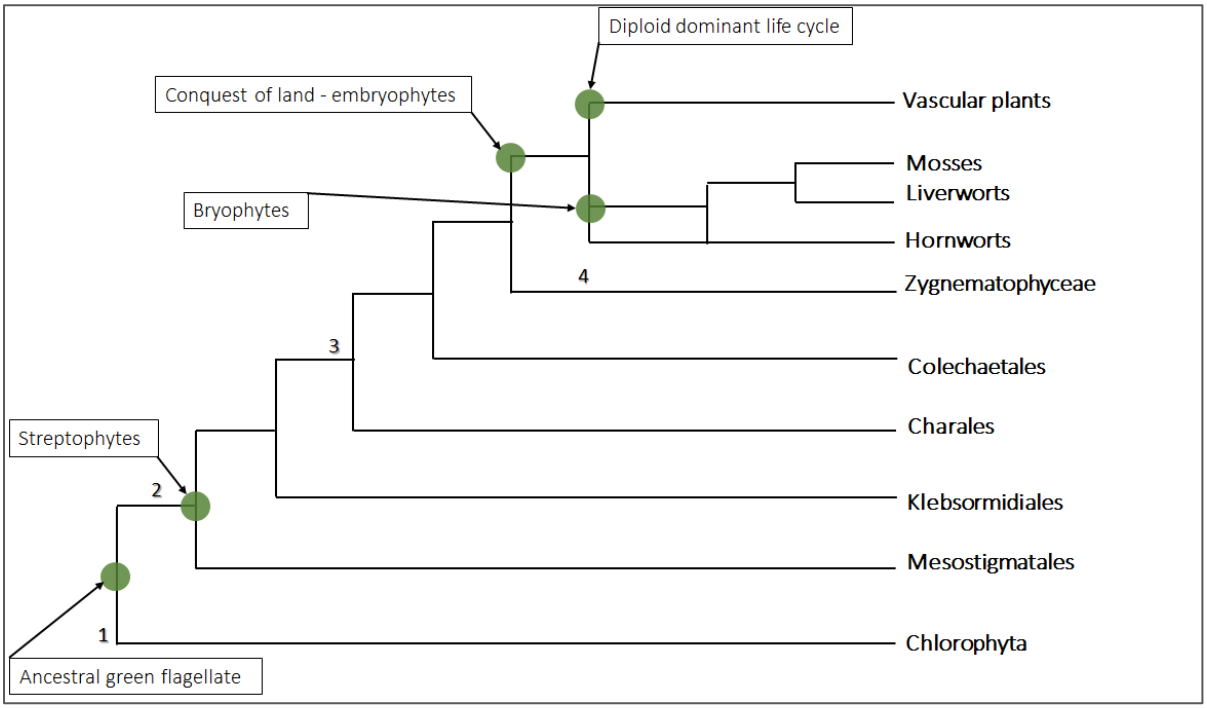


Figure 1.4: Phylogenetic relationship between land plants and green algae. Not all charophyte lineages are shown, but zygnematophyceae are recovered as sister to land plants. Bryophytes are the earliest branching land plants anatomically comparable to the earliest land plants. Still, hornworts' position is the least confident, possibly representing a distinct land plant lineage. The ancestral green flagellate shows the origin of key groupings. Streptophytes include charophyte green algae and land plants. The development of vascularization and the anatomical complexity is associated with the switch to a diploid dominant life cycle. (1) Evidence for the presence of all five major plant hormones (ABA, ethylene, auxin, gibberellin, and cytokinin) in microalgal. (2) Origins of charophytes accompanied by some traits shared with land plants. (3) Evolution of morphological complexity. (4) The subsequent loss of morphological complexity in these single cells or filamentous algae and loss of motile gametes resulted in conjugation. (Adapted from [73])

The moss *Physcomitrium patens* (Hedw.) lacks lignified secondary cell walls, which are characteristic of vascular plants but have support cells with thick walls called stereids [74]. However, their unlignified cell walls contain cellulose, and they also have cells called hydroids that conduct water with a thin cell wall and undergo programmed cell death like tracheary elements. Unlike in *Arabidopsis*, the CESA isoforms' specific functions in *P.*

patens are unknown [75]. Additionally, seed plant CSCs contain three *CESA* isoforms [76], whereas CSCs in *P. patens* might be homo- or hetero-oligomeric [77]. However, these features have yet to be determined experimentally.

Moss has seven *CESA*s genes that do not cluster with the six *CESA* clades in seed plants based on the phylogenetic analysis [75]. Three pairs show high similarity (*PpCESA6* and *PpCESA7*, *PpCESA3* and *PpCESA4*, *PpCESA8,* and *PpCESA10*) and three pseudogenes (*PpCESA1*, *PpCESA2*, and *PpCESA11*) [75, 77]. ESTs' relative abundances show expression of *PpCESA6*, *PpCESA7*, and *PpCESA8* are induced by auxin [75] and *PpCESA6* for tip-growing protonemata, rhizoids, and axillary hairs induction [53].

Moss *CESA* genes have isoforms and diversified from seed plants *CESA*s, but their functions are still unknown [75, 77]. Knowing about the functional differences between *PpCESA*s and seed plants can provide us an insight into CSC evolution and the role of each CESA isoform in the complex assembly [78]. Analysis of mutations has shown only one of the seven *P. patens CESA* is functional. *PpCESA5* is necessary for gametophore formation, and the gametophore buds of mutants do not accumulate crystalline cellulose in their cell walls and fail to develop and form deformed cell clumps [77]. Also, double *PpCESA6/7* knockout shows a shorter gametophore [53]. Another double *PpCESA3/8* knockout produced fewer gametophores, leading to the assumption that they contribute to structural support but are not required for morphogenesis [21]. The genes identified as *CESA* superfamily in moss *Physcomitrium patens* were assigned to *CESA* and *CSL* families. Phylogenetic analysis confirmed this hypothesis that the CESA and CSL families' divergence predated the separation of mosses and vascular plants. There is a similarity in SCW microfibril texture as well as convergent evolution between mosses and vascular plants. *P. patens* midribs include hydroid cells to transport water, and stereid cells provide support for aerial organs and have thick and cellulose-rich cell walls [75]. Based on previous studies, the genes that encode secondary *PpCESA3* and *PpCESA8,* and primary *PpCESA5* are subfunctionalized and specialized, so they encode interchangeable proteins. The diversification and duplication of *CESA*, followed by changes in the regulatory elements of primary and secondary *CESA*s in mosses and vascular plants, occurred independently based on phylogenetic studies [75].

1.5 The role of phytohormone signaling in plant cell wall formation

The size and shape of a plant cell are controlled by cell expansion and localized cell division during plant growth. The cell wall expansion is the only construction process to shape the plant organs by loosening the cell wall matrix and deposition of new wall components. These expansions happen either isotropically, in all directions, or anisotropically which means a specific directional growth [79]. How cell expansion occurs depends on growth effectors such as phytohormones and light. Some of the particular hormones control developmental responses. In terms of plant growth and development, these hormones and cell wall organizations' importance suggests that synthesis and modification of cell wall components and hormone production should be connected directly or indirectly [80]. Over the last century, many molecular and physiological studies have shown the link between auxin-promoted cell division and elongation to cell wall remodeling [80, 81].

Phytohormones play a crucial role in plant growth, development, and stress responses [82]. One class of phytohormones, cytokinins (CKs), has a critical role in cell division and proliferation during plant development [83], chloroplast division, senescence delay [84], and abiotic stress responses [85]. Some evidence suggests Cks are involved in stress responses [86]. *P. patens* is a new diverging land plant encoding proteins involved in cytokinin biosynthesis, metabolism, and signaling [87]. It is the only plant that encodes both classical and newly identified cytokinin receptors in its genome [87]. There are 20 different endogenous cytokinins detected in *P. patens*, and cytokinin-deficient plants showed the importance of extracellular cytokinins for bud formation [88]. *P. patens*' development is controlled mainly by evolutionary phytohormones such as auxin, Ck, and abscisic acid (ABA) [89, 90]. Cks, in particular, are essential for the evolution of moss' life cycle as they induce the production of buds (three-faced apical cells) and the transition from protonemic growth to the formation of leafy structure, the gametophore [91]. An exogenous Ck application results in an overproduction of buds [92].

In 2012, Goss et al. showed that *CESA5* in moss is upregulated by cytokinin, which also induced gametophore development [77]. While cytokinesis is unaffected in *Arabidopsis cesa* mutants, in moss *cesa5KO*, cell expansion and cytokinesis are disrupted [77, 93]. Yet,

cytokinesis defects have been noted in *Arabidopsis* KOR1 mutants [94], but the role of KOR1 in cellulose biosynthesis is still unclear. While CESAs in seed plants are specific for primary and secondary cell wall biosynthesis [25, 28, 29], *P. patens* CESAs seem to be specified for primary cell wall biosynthesis in different cell types [75]. The gametophore-specific phenotype is consistent with the improvement of *PpCESA5* expression by cytokinin, promoting gametophore initiation. In contrast to *PpCESA5*, *PpCESA6* is expressed in tip-growing protonemal filaments, axillary hairs, and rhizoids [53], and its ESTs are overrepresented in libraries derived from cultures treated with auxin, which promotes caulonema and rhizoid development [75]. These differences in the patterns of functional specialization of *P. patens* and *A. thaliana* CESAs are consistent with phylogenetic analysis showing independent diversification of CESAs in mosses and seed plants [77].

1.6 From auxin biosynthesis to signaling and transport

The phytohormone auxin is synthesized in a wide range of streptophyte plants (the most basal land plants), seed plants, and some algae species, thus suggesting that some components of the auxin signaling pathway might be conserved. There is evidence of presenting auxin signaling machinery in *Physcomitrium patens* identical to flowering plants [52, 95].

Auxin is required for plant growth through cell elongation, cell division, and cell differentiation, and specifically in moss development of protonemal, gametophore, and sporophyte tissue requires the action of auxin. [91, 96]. Auxin stimulates the gradual transition from chloronemal to caulonemal cells at the tip of protonemal filaments [97]. Various studies showed an increase in auxin levels in the tip cells of protonema filaments [98]. It has also been reported that the transition of chloronema to caulonema is affected by light quality through cryptochrome-mediated changes in auxin sensitivity. Furthermore, auxin can influence the development of gametophore shoots and their different organs [99].

Moss EST databases reveal the presence of three genes encoding Auxin/INDOLE-3-ACETIC ACID (Aux/IAA) proteins [100], which bind to AUXIN RESPONSE FACTOR (ARF) transcription

factors and block them from the upregulation of auxin-responsive genes. The molecular function of these components is conserved in land plants [101]. When auxin is present, TRANSPORT INHIBITOR RESPONSE 1/AUXIN SIGNALING F-BOX (TIR1/AFB) acts as a co-receptor with Aux/IAA for auxin. The study of auxin antagonists in angiosperms and *P. patens* suggests that auxin response mediated by TIR1, Aux/IAA, and ARF proteins is an ancient mechanism [102]. A recent study has shown that the moss Aux/IAA proteins interact with *PpAFB (Arabidopsis* TIR1 moss homologs) and reduce *PpAFB* levels, and mutations in *Aux/IAA* genes lead to an auxin-resistant phenotype [100]. On the other hand, the genomic analysis uncovered the presence of genes encoding (AUXIN BINDING PROTEIN 1 (ABP1) and IBA RESPONSES 5 (IBR5) proteins in bryophytes [52], suggesting that the three known auxin signaling pathways were present in the last common ancestor of extant land plants.

When auxin binds, Aux/IAA is degraded by a 26S protease through ubiquitination, and ARF is then free to act upon auxin-responsive genes. In ABP1 signaling, ABP1 is anchored on the external surface of the plasma membrane [103]. The *P. patens* genome contains *TIR1/AFB*, *Aux/IAA*, and *ABP1* homologs [52], and auxin-resistant mutants in moss have shown the same behavior of the TIR1/AFB signaling pathway in *Arabidopsis* [100].

The sensitivity of plants to auxin depends on the function of a protein ubiquitin ligase complex of the Skp, Cullin, F-box containing (SCF) type. SCF complexes consist of an F-box protein, a SKP1 homolog, a Cullin/CDC53 homolog, and a RBX1/ROC1 (RING-box1/ regulator of cullins-1) homolog. F-box proteins consist of various protein-protein interaction domains, the domain responsible for recruiting specific substrates to the SCF complex to target and ubiquitinate protein degradation by the 26S proteasome [104].

Two distinct pathways play a role in auxin transport: an active cell-to-cell polar transport and a passive distribution through vascular tissue [101]. The polar auxin transport is essential for auxin distribution over both short and long distances, which occurs in a cell-to-cell manner and depends on the specific influx and efflux carrier proteins that enable the uptake and release of auxin from/to the apoplast [101]. Some auxin carriers are characterized: the PIN-FORMED (PIN) proteins [105] and several ABCB and ABCG

transporter protein families [106-108], which are involved in auxin efflux from the cell and the AUX1/LIKE AUXIN PERMEASE (AUX1/LAX) proteins are involved in auxin influx [109, 110]. Four homologs of the auxin influx carriers *AUX1/LAX* have been detected in the *Physcomitrium* genomes [52]. The influx and efflux carriers are sensitive to several inhibitors, such as 1-N-naphthylphthalamic acid (NPA). Treatments with the inhibitors also cause changes in the distribution of auxin in the sporophyte and result in abnormal embryo development but not in the shoot [111].

1.7 Auxin and crosstalk with other phytohormones

Auxin interacts with all the other phytohormone pathways and impacts morphogenesis [112]. Different hormone pathways affect auxin homeostasis by modifying auxin transport, biosynthesis, or signaling components [113]. For example, cytokinins induce the expression of *IAA3*, which has a negative effect on the auxin-signaling pathway at the root's transition zone, thus causing auxin-induced cell proliferation instead of differentiation [114, 115]. Ethylene also controls auxin levels by changing auxin transporters' expression, such as AUX and PIN families and biosynthetic enzymes [116-119].

Recently, cytokinins were shown to control PIN protein endocytosis and induce their degradation in lytic vacuoles [120]. High concentrations of jasmonate also induce PIN2 endocytosis and degradation [121]. On the contrary, gibberellin signaling and the secretory peptides from the GOLVEN family limit PIN trafficking to lytic vacuoles [122, 123]. Consequently, other than auxin, many other hormones control auxin efflux and the intracellular trafficking of PIN. Since auxin is involved in many aspects of plant growth and development, it is not unexpected that auxin regulation turns out to be so complex relative to other plant hormones.

1.8 Synthetic auxin herbicides

Many synthetic small molecules can induce an auxin response. Synthetic auxin herbicides are similar to natural auxin IAA. There are five classes of auxinic herbicides: phenoxy-carboxylic acids (e.g., 2,4-dichlorophenoxyacetic acid); benzoic acids (e.g., dicamba); pyridineacids (e.g., picloram, clopyralid); quinolinecarboxylic acids (e.g., quinclorac); and

pyrimidine carboxylic acid (e.g., aminocyclopyrachlor). Binding to receptor proteins depends on structural variation in each herbicide molecule [124]. These herbicides are more stable and can persist for a longer time, in contrast to IAA in plants that degrade rapidly. These molecules can yield growth and development in horticulture and agriculture at lower concentrations, but they are also lethal at higher concentrations as an herbicide for weed control [125].

Due to several reasons, resistance to auxinic herbicides is challenging to evolve because genetic alteration in the complex signaling network governed by auxin decreases survival range in severe environmental conditions [126, 127]. Also, auxins affect the metabolism dynamics in the cells at various levels of the organization, such as endocytosis, cell polarity, and cell cycle control. Finally, auxins act in regulatory processes and direct embryogenesis, patterning tissues, and new organ formation [128].

The auxin mimicking herbicides 2,4-dichlorophenoxyacetic acid (2,4-D) and 1-naphthalene acetic acid (NAA) were used in this study.

1.9 Cellulose biosynthesis inhibitors (CBI) as a molecular toolbox to study cellulose biosynthesis

In previous studies, cellulose synthesis was shown to be important in stabilizing tip growth in plants [76]. Both pollen tubes and root hairs extend through tip growth, similar to *P. patens* protonemal tissue [129]. Inhibiting cellulose biosynthesis in the cell wall causes loss of directional cellular expansion and radially swollen tissues [130]. These combined characteristics, leading to swollen tissues and stunted growth, provide a tractable phenotype for genetic screens [130].

Cellulose biosynthesis inhibitors (CBIs) are a chemically diverse group of herbicides that reduce cellulose production specifically. Although these molecules interrupt the cellulose biosynthetic pathway, the exact mode of action remains vague [131]. Genetic screening of mutagenized populations with CBIs has led to the identification of mutations leading to CBI resistance. Due to the specificity of CBIs to the cellulose biosynthesis pathway, mutations leading to CBI resistance are found primarily in genes involved in cellulose biosynthesis [131-133]. Continued development of new and current CBIs is expected to be driven by

their utility in cellulose biosynthesis research and weed control agents [131]. Therefore the precise associations between CBIs and CESA are correlative. The responses to chemical inhibition have been studied via live-cell imaging thus far. One group of CBIs promotes the clearing of CESA from the plasma membrane such as isoxaben, another group stops the movement of CSC in the plasma membrane which of 2,6-dichlorobenzonitrile (DCB) is categorized in this group, and the third group disrupts the mechanism of the CSC which is co-disturbance of both CESA and microtubules [130] (Figure 1.5). This classification is based on short-term [^{14}C] radioisotope tracer studies and the incorporation of glucose into cellulose [134] and time-lapse confocal microscopy [135]. Interestingly, however, in the moss *Funaria hygrometrica*, the number of rosette structures decreased under treatment with the CBI, DCB, as visualized with freeze-fracture electron microscopy [130].

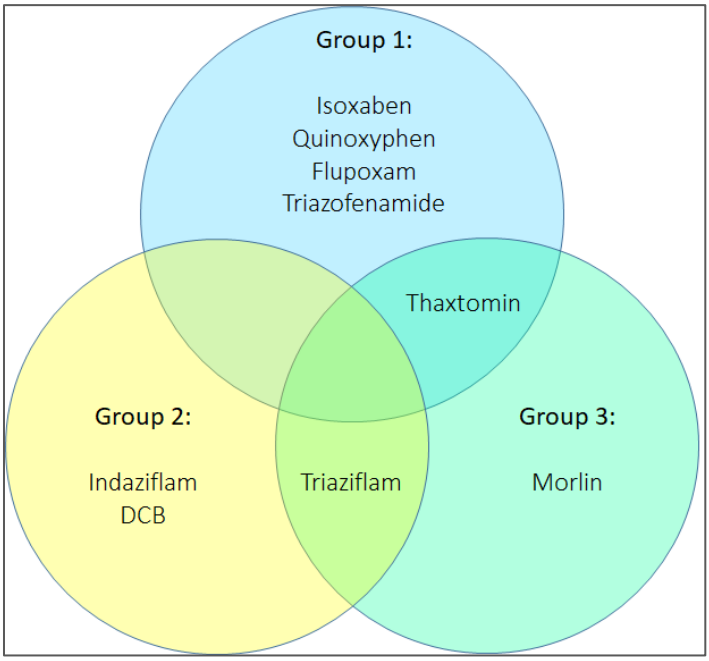


Figure 1.5: A Venn diagram showing the three classes of CBIs. The overlapping regions represent how a CBI can show a range of modes of action that can pose a challenge to their classification.

The CBI used in this study is indaziflam. The herbicide was introduced by Bayer Crop Science and had a photosystem II inhibition value of 9.4 [136]. Indaziflam is an alkylazine herbicide used to control annual grasses by eliminating the invasive grasses, increase the competitiveness of continuing co-occurring species, and ultimately increase biomass [137].

The mechanism indaziflam controls weeds are different from other CBIs, such as isoxaben. Indaziflam has a different molecular target than isoxaben that eventually causes differential effects on CESA particle velocity, clearing the plasma membrane in indaziflam treated plants [136]. In a dose-dependent manner in *Arabidopsis*, cellulose production is inhibited and a primary mode of action for indaziflam, consisting of inhibition of cellulose biosynthesis [136]. Their results do not support that indaziflam has the same action mechanism as isoxaben and other CBI like quinoxyphen [136]. Based on these findings, indaziflam causes radial swelling and ectopic lignification in *Arabidopsis* and *Poa annua* treated seedlings. These findings suggest that indaziflam affects a different component of the complex cellulose biosynthetic process than other CBIs [136]. Since no published alleles confer resistance to indaziflam, its mode of action is still not completely understood.

1.10 The use of *Physcomitrium patens* as a vehicle for expressing novel compounds

P. patens has been used to study plant metabolic engineering for a very long time [138]. The high efficiency of *P. patens* homologous recombination among plants has established *P. patens* as a model organism to study gene functions [49, 52, 139]. *P. patens* are also used as a heterologous expression host for recombinant proteins and natural products of therapeutic and commercial value [138, 140-143]. *P. patens*, as a photosynthetic organism, can play a role as a biotechnological host and an expression system for plant biosynthetic pathways [144, 145]. Stable and heterologous expression of genes encoding various biosynthetic enzymes in planta is an alternative to microbial host expression for biotechnological production, a valuable tool for functional characterization, and an alternative industrial-scale production of biopharmaceuticals. The moss *P. patens* has already been recognized as viable for reconstructing heterologous pathways as a biotechnological host [146].

Polyethylene glycol (PEG) mediated transformation of *P. patens* protoplasts with a linearized plasmid is being used as a standard method to introduce heterologous DNA into *P. patens* cells [147]. Multiple insertion sites targeting one single transformation event are also another important feature of *P. patens* [148]. The homologous recombination

machinery of *P. patens* can be applied to assemble multi-fragment constructs with 5′ and 3′ flanking sequences and integrate them into the genome. This is a similar methodology using *S. cerevisiae,* although the development of a similar method for *P. patens* can further increase its utility as a plant biotechnological framework [149]. Improvement in targeted techniques of large-scale genome manipulation in plant systems is valuable and could help understand and improve genetic structure and function at the chromosome and genome levels [149]. The application and optimization of *in vivo* DNA assembly methodologies can significantly improve the potential of *P. patens* as a biotechnological host species and as a photosynthetic chassis for synthetic biology.

1.11 Thesis objectives and hypothesis

The work outlined in this thesis has focused on using *P. patens* as a tool to explore basic biological questions and exploit the organism for biotechnology applications. I have used *Physcomitrium patens* for a line of inquiry into cellulose biosynthesis using reverse and forward genetics techniques. I have also tried to determine how amenable *P.patens* is for reconstituting a metabolic pathway, the cannabinoid biosynthetic pathway from *Cannabis sativa*.

Objective 1 – Forward genetics to identify possible CBI targets.

Cellulose synthesis has been investigated through the use of cellulose synthesis inhibitors (CBIs) [130]. CESAs and the accessory proteins involved in the CSC are targets for the CBIs, which can influence cellulose biosynthesis and may induce a CBI-like mode of action [150]. In this thesis, I will characterize the effects of CBIs on the deposition of cellulose, explore the mechanism by which CESA responds to the CBIs, and identify CBI resistant mutants in *Physcomitrium patens* by using the forward genetic approach. In *Arabidopsis thaliana* and *Triticum aestivum,* CBI resistance was caused by point mutations in different *CESA* genes (Bonetta, unpublished). I hypothesized that forward genetic screening would identify new CBI resistant mutants in moss, similar to *A. thaliana* and *T. aestivum* point mutations in primary cell wall CESAs, which would alter CBI binding, resulting in CBI resistance. If this hypothesis is correct, it will be evident that mutations affecting cellulose production in *P.*

patens are transferable among land plants, and abiotic stress signaling is subject to cellulose synthesis, which likely played a vital role in the conquest of land by plants. Also, identifying numerous genes and pathways connected to CBI sensitivity will allow for a better understanding of CBI's mode of action. Functional conservation of the cellulose biosynthesis composition has been established across 500 million years of land plant evolution. This has been a successful approach that has helped build our understanding of the cellulose synthesis pathways likely to be crucial in anatomically simple early land plants. The results of this objective will be described in Chapter 2.

Objective 2 – Identification of CSC components

Since there is no reliable *in vitro* assay for cellulose synthase activity from plant material, the second objective of this research is to characterize CSC's composition in moss that would help uncover new components that haven't been identified yet by using a method called biotin identification (BioID). Forward genetic screens have implicated several proteins other than CESAs that are part of the CSC. However, this reverse genetics approach, BioID, is necessarily limited to testing our hypotheses based on a previous study, which has been done on bacteria, yeast, plants, and mammalian cells. Knowing the CSC composition should lead to defining a better mechanistic role of non-CESA proteins in cellulose synthesis. For example, there is currently no known activator of the CSC. This work is discussed in chapter three.

Objective 3 – The use of moss as a vehicle for cannabinoid biosynthesis.

Researchers worldwide have chosen moss *Physcomitrium patens* as the candidate organism to produce drug candidates derived from plants, and it has been used as a vehicle for the production of pharmaceuticals [113]. Some drugs like cannabinoids are difficult to obtain by chemical synthesis, and they are not available in sufficient amounts from their original plant. To overcome the public demand, we require a different, efficient production system. Combinatorial biosynthesis has already been successfully done for the production of compounds such as antibiotics in microorganisms [151]. An *in vivo* combinatorial

biosynthesis approach will be pursued in *Physcomitrium*, primarily to obtain known natural products with useful biological activity such as cannabinoids. A critical aspect of this work is to produce not just the most common cannabinoids, also rare cannabinoids and pathway intermediates.

An important additional outcome of this idea is that knowledge gained during the establishment of a *Physcomitrium* strain producing cannabinoids will enable the design of several other drug-candidate-producing moss strains, especially within the chemical class of cannabinoids. The possibilities of using *Physcomitrium* as a new production platform for plant-derived natural products are novel and exciting. The results of this objective will be described in Appendix 1.

Chapter two: Forward genetics in *Physcomitrium patens* identifies indaziflam resistant mutants

2.1 Introduction

Physcomitrium is established as a plant model system for study plant gene functions using forward and reverse genetics. It is also used as an experimental system to answer molecular, developmental, and evolutionary questions that mostly rely on the excellent gene targeting capabilities mediated by homologous recombinations [49]. As a bryophyte, the position of *Physcomitrium* at the base of the land plant lineage can help understand the evolutionary processes for comparative genomics, which led to the origins and creation of land plants, although there are features unique to bryophytes that showed extra to the requirements of subsequent land plants [55]. Likewise, the bryophyte's lineage-specific features indicate the numerous evolutionary routes can take, and that can prove compelling and unexpected facts [55]. For example, the finding of neochrome in fern *Adiantum capillus-veneris* helped with new gene formation in evolutionary processes. The neochrome is a chimeric protein in which the N-terminal chromophore binding domains from the far-red/red sensing phytochromes are fused with the blue-light sensing phototropin [152]. This neochrome arose in the hornworts and is thought to have been passed to ferns by horizontal gene transfer [153]. Similar neochrome has also been discovered in the green alga *Mougeotia scalaris,* in which they are thought to have arisen independently [154]. Nonetheless, of their fascinating origins, this neochrome has been proposed to confer a selective advantage to ferns in low-light habitats [153]. The discovery of neochrome, initially through the analysis of phototropic mutants in Adiantum, validates the benefit of forward genetics in discovering the unexpected and novel.

Forward genetics is based on the available genetic resources in the *Physcomitrium* gene discovery. This approach is based on whole-genome sequencing of bulked segregant pools to identify mutants and the screen's related phenotypes, which encourages the pursuit of the mutants' analysis, particularly in *Physcomitrium*, for novel gene discoveries given the

novel mechanisms involved in bryophyte signaling pathways and understand some early events in land plant evolution.

This thesis's forward genetic approach is based on single nucleotide polymorphisms (SNPs) created by a mutagen. An SNP site where the nucleotide substitution leads to an amino acid substitution or otherwise affects a protein is a coding SNP (cSNP). SNPs can occur anywhere in genomes, including in coding regions of genes where they can be associated with natural variation in gene functions and often a defeat in organisms. In the context of mapping SNPs in crosses where two distinct ecotypes are used, such as in *Physcomitrium patens* the Gransden (Gr) and Villersexel K3 (Vx) ecotypes, the SNP genotyping allows identifying the parental origin of regions in the genome. In *P. patens*, the forward genetic screening has taken advantage of identifying thousands of SNPs between the Gr and Vx ecotypes.

The details of the approach used in this thesis and the basic principles and theory behind SNP-based mapping will be described in the following sections. This approach deals with bi-allelic SNPs (A↔B) in which allele A occurs in genotype A (Gransden ecotype), and allele B occurs in genotype B (Villersexel ecotype). The crossing of these genetically distinct ecotypes results in a diploid sporophyte genome containing SNPs from both parents, which become shuffled following recombination during meiosis. In mosses, the gametophytes produce the sperm (antheridium) and egg (archegonium), creating organs, and the gametes are produced by mitosis. Following fertilization, the diploid sporophytes form and grow from the archegonium. These form diploid sporangia at their tips, which then have spores through meiosis. The spore-derived progeny are haploids and are equivalent to recombinant inbred lines in diploid species (Figure 2.1). This process relies on the basic principle that SNPs closer together are less likely to be separated by recombination events between them, which approaches 0% when adjacent. In contrast, SNPs on different chromosomes (linkage groups) or sufficiently separated on the same chromosomes (i.e., at opposite ends) will show no association (50% assortment in progeny) [89].

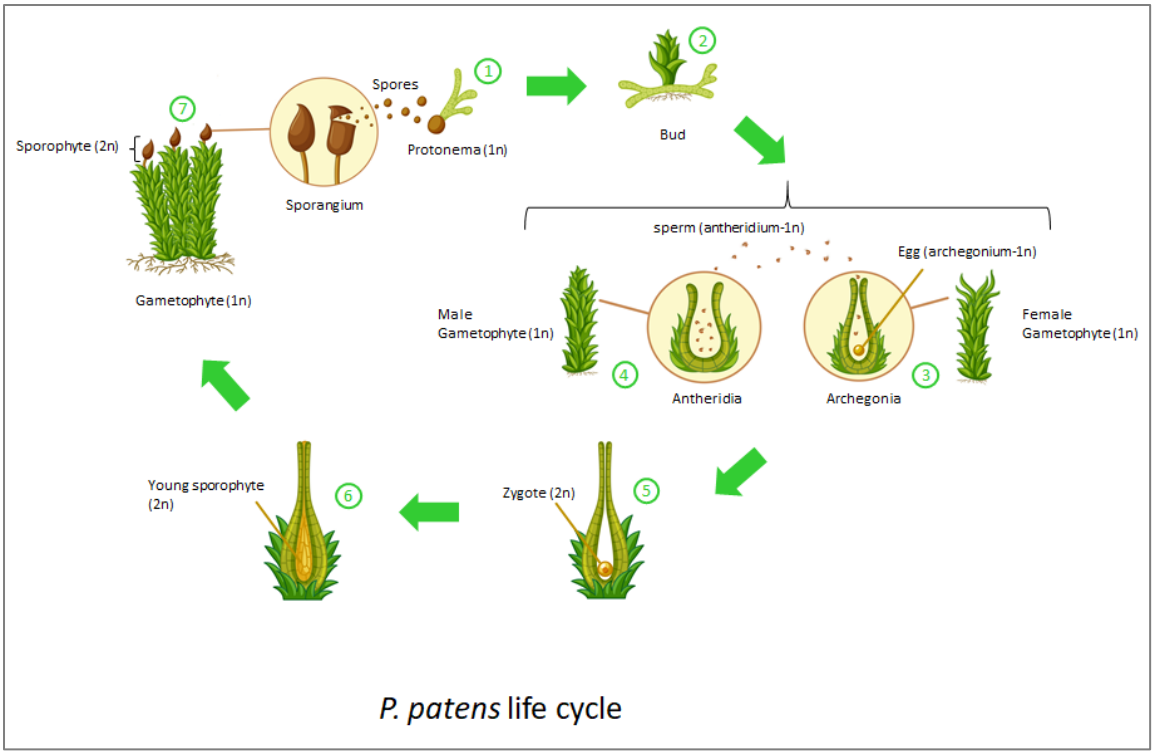


Figure 2.1: *Physcomitrium patens* life cycle. 1: Spore dispersed and germination, 2: Protonema development, 3: New female gametophyte, 4: New male gametophyte, 5: Egg growth, 6: sporophyte development. 7: Mature sporangium (Adapted from GraphicsRF.com).

By implementing this positional mapping process on indaziflam (*izr*) mutants in *Physcomitrium patens*, we aimed to identify novel genes involved in the cellulose biosynthesis pathway. The typical wild-type response to indaziflam is characterized by slow growth since it blocks cell elongation, which eventually causes cell death. Any mutant in which this response no longer occurs was scored as resistant, which progressed through normal growth and development without the inhibitory effects of indaziflam being manifest. Previous findings suggested that indaziflam affects a different component of the complex cellulose biosynthetic process, causes radial swelling and ectopic lignification in *Arabidopsis thaliana* and *Poa annua* treated seedlings [136].

Indaziflam is known as a CBI. Investigation of CESA behavior while the plant is treated with CBIs can provide us a platform to answer questions regarding the mechanism of delivery, activation, movement, and array organization during cellulose biosynthesis. By identifying mutations from forward genetic studies of CBI-treated plants that confer resistance to

these chemicals, it is possible to learn about the different mechanisms involved in cellulose biosynthesis. Chemicals like DCB [155], flupoxam (FPX) [156], indaziflam [136], isoxaben (ISX) [157] and, quinoxyfen [158] are widely used as herbicides and contain diverse chemical structures (Figure 2.2). The herbicides group includes DCB, and indaziflam cause changed motility of CSCs across the plasma membrane. DCB and indaziflam act to completely inhibit the CSC movement across the plasma membrane [136, 159].

Indaziflam, N-[(1R,2S)-2,3-dihydro-2,6-dimethyl-1H-inden-1-yl]-6-[(1RS)-1fluoroethyl]-1,3,5-triazine-2,4-diamine, is a strong herbicide produced by Bayer Crop Science (U.S. Environmental Protection Agency [EPA], 2010), and is used to control broadleaf weeds and annual grasses and is effective at low concentrations, and became a desirable candidate for widespread use in agriculture. However, to be most effective, indaziflam would need to be paired with indaziflam resistant crops. To develop indaziflam resistance in crops, firstly, we need to understand its herbicidal mode of action. Brabham in 2014 has shown that indaziflam can cause a decrease in CSC velocity and prevent plant growth [136]. The evidence for indaziflam's mode of action currently rests with the *ixr 1-1* allele, which has been shown to confer a low level of resistance against indaziflam [136]. Indaziflam was also shown to inhibit the production of cellulose in a dose-dependent manner [136]. Previous studies have demonstrated that CBI-resistant *A. thaliana* strains can be identified through forward genetic screening [131, 132, 158, 160][Bonetta unpublished]. However, they were not successful in finding a strong resistance to indaziflam, which suggested indaziflam may have multiple targets in plants [131] [Bonetta unpublished].

To characterize the effects of CBIs and the mechanism with the forward genetic approach, I will identify CBI-resistant mutants in *P. patens.* These mutations may alter CBI interaction and result in CBI resistance, and I will determine whether the mutations are causing the relative changes in *P. patens* genes and pathways are connected to CBI sensitivity to provide a better understanding of CBI's mode of action.

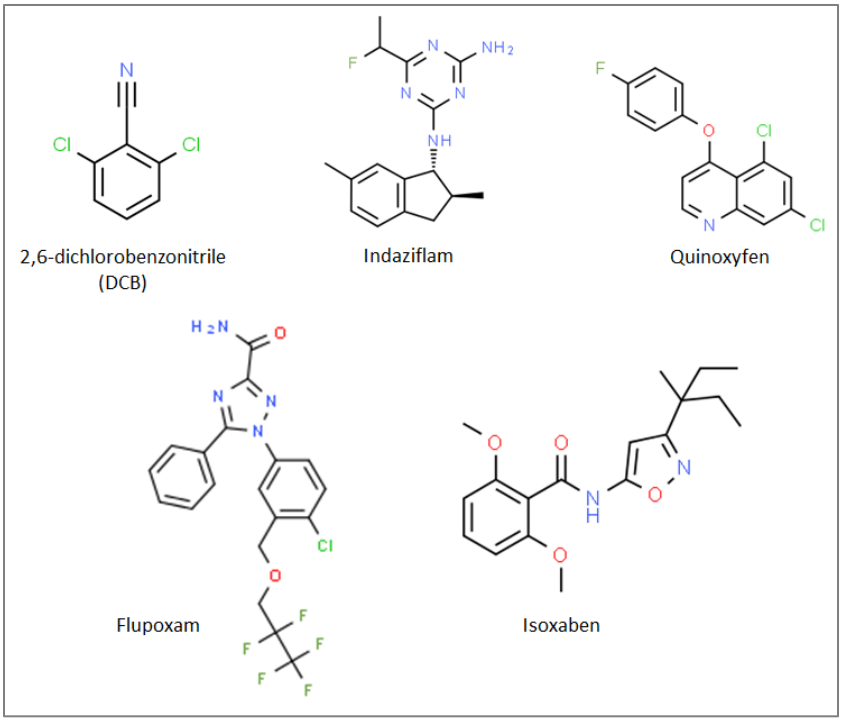


Figure 2.2: Cellulose biosynthesis inhibitor structures. Chemical structures are from www.chemspider.com.

2.2 Materials and Methods

2.2.1 *Physcomitrium* strains and growth conditions

Physcomitrium parens wild-type ecotypes Gransden and Villersexel (V3) were used for all experiments and observations. All samples of *P. patens* were grown on the media called BCD and BCDAT (BCD medium supplemented with 5mM ammonium tartrate), described in table 2.1. The media overlayed with sterile cellophane disks (RPI Research Products International, ULTRA CLEAR CELLOPHANE) cut into circles and autoclaved in glass Petri plates and incubated approximately at 21-25°C under continuous illumination at 50–80 mmol/m^2/s supplied by white fluorescent tubes. The plants were subcultured by homogenizing tissue in 4-6 ml sterile water with a PowerGen 125 grinder equipped with Omni International hard tissue Omni disposable tip probes (USA Scientific, Ocala, FL, USA) and plating them onto BCDAT medium overlaid with cellophane disks and grown under the same conditions for 5 to 7 days for protoplast isolation.

For sporophyte induction, gametophytes were incubated at 16°C [161] in 8 hours light and 16 hours dark at 50–80 mmol/m^2/s (Adapted from [162]) on sterile peat pellets (Jiffy-7, 42 mm).

Table 2.1: Media for culture and transformation of *P. patens* [163].

Solution	BCD (per L)	BCDTA (per L)	PRMB (per L)	PRMT (per L)	PRML (per L)
$MgSO_4$ $7H_2O$	0.25 g	0.25 g	0.25 g	0.25 g	0.25 g
KH_2PO_4 (pH6.5)	0.25 g	0.25 g	0.25 g	0.25 g	0.25 g
KNO_3	1.0 g	1.0 g	1.0 g	1.0 g	1.0 g
$FeSO_4$ $7H_2O$	12.5 mg	12.5 mg	12.5 mg	12.5 mg	12.5 mg
Ammonium Tartrate	-	0.92 g	0.92 g	0.92 g	0.92 g
Alternative TES (Trace elements)*	1.0 mL	1.0 mL	1.0 mL	1.0 mL	1.0 mL
Mannitol	-	-	60 g	80 g	80 g
Agar	7.0 g	7.0 g	8.0 g	5.0 g	-
$CaCl_2$ solution 55.5 g/L (sterile) Add after autoclaving	1.0 mL	1.0 mL	10 mL	10 mL	10 mL

***** 55 mg/L cupric sulfate pentahydrate; 55 mg/L zinc sufate heptahydrate; 614 mg/L boric acid; 389 mg/L manganous chloride tetrahydrate; 55 mg/L cobalt chloride hexahydrate; 28 mg/L potassium iodide; 25 mg/L sodium molybdate dehydrate.

2.2.2 Protoplast isolation and CBI screen

To study the effect of CBIs on moss, protoplasts were isolated according to Roberts et al. [162] using the wild-type 5-7 days old chloronemal tissue was digested using 2% Driselase™ (Sigma-Aldrich) solution, which was dissolved in 50 mL 8.5% D-mannitol. Protoplasts were isolated by treating 5 to 7 days old protonemata with 0.5% (w/v) Driselase (Sigma Chemical Co., Dorset, UK) in 8.5% (w/v) mannitol solution for 30-60 minutes. Protoplasts were washed twice by centrifuging at 200 rpm for 3 minutes in 8.5% mannitol. PRMT medium is melted and equilibrated to 45°C in the water bath; protoplast regeneration plates were prepared with PRMB medium and overlain with sterile cellophane disks. Protoplasts were resuspended in 5 mL PRMT held at 45°C, and 1.5 mL of suspension was spread on each PRMB plate. Plates were incubated for five days at 25°C with continuous light at 50–80 mmol/m^2/s [162].

Cellophane disks were then transferred to BCDAT media supplemented with 20 µM of different CBIs listed in table 2.2. This represents an in-house CBI collection compiled by screening Arabidopsis plants on a 20000 compound collection from Cambridge Corp. for cellulose deficient phenotypes and validation by cell wall cellulose content (Bonetta, unpublished). The CBI screening concentration was determined based on the concentration that has growth effects on *Arabidopsis thaliana* (Bonetta et al. unpublished) and wild-type sensitivity. The explants were removed from the CBI supplemented media and transferred to BCD plates after four weeks of growth.

2.2.3 Dose-response curve

Growth curves must be completed on wild-type plants using CBIs to establish the screening conditions and test whether CBIs affect protonemal tip growth and cell wall regeneration, affecting phenotype and influenced growth in *Physcomitrium patens*. Therefore the growth curves were completed on different concentrations; 0 µM (control), one µM, 5 µM, 10 µM, 20 µM, and 30 µM for wild-type plant. Samples with a zero herbicide concentration contained a solvent equal to the highest herbicide concentration.

It was found that there was incomplete or no growth at 20 µM, which is shown in Figure 2.5. Therefore 20 µM was chosen to screen the mutagenized protoplasts. CBI treatment of *P. patens* was carried out by transferring 5-day-old previously described regenerated protoplasts to solid BCDAT media supplemented with 20 µM of 32 CBIs listed in table 2.2 and incubating for two weeks, and digital photographs were taken [164].

2.2.4 Creation and screening of *izr* mutants by UV mutagenesis and indaziflam screen

Protoplasts were isolated as described previously and embedded in a 3ml PRMT agar medium. Approximately 50,000 protoplasts were suspended in PRMT for each 9cm Petri dish, then spread on PRMB plates overlaying with cellophane disks. The protoplasts were exposed to 25,000µJ.cm^{-2} ultraviolet radiation (280 nm) in a UV Stratalinker 2400 and incubated in darkness 24 hours. Since photoreactivation is used to repair UV damage in DNA, incubation in the dark helps avoid the repair mechanism. Then the plates were

incubated at 25°C, under continuous illumination, and allowed to regenerate. This treatment resulted in 10-20% of the irradiated protoplasts surviving and regenerating as protonemata. After five days to enable cell wall regeneration, the cellophane bearing the embedded regenerants was transferred to plates containing standard BCDAT-agar medium (1mM $CaCl_2$, no mannitol) supplemented with 20 µM indaziflam for four weeks. By that time, *izr* mutants were distinguishable and could be routinely subcultured.

Table 2.2: List of CBIs tested in *Physcomitrium patens*

CBI#	Chemical Name	Molecular Weight (g/mole)
3	2-(2-bromo-4-methylphenoxy)-N-(2,4-dimethoxyphenyl)acetamide	380
7	N-{amino[(4-methyl-2-quinazolinyl)amino]methylene}-4-methylbenzenesulfonamide	355
8	N-[2-(2-isopropyl-5-methylphenoxy)ethyl]-3-phenylpropanamide	325
9	3,5-diethoxy-N-(4-methyl-2-pyridinyl)benzamide	300
14	2-{3-[(4-fluorophenoxy)methyl]-4-methoxyphenyl}-3-propyl-2,3-dihydro-4(1H)-quinazolinone	420
15	3-(2-furylmethyl)-2-(4-methoxyphenyl)-2,3-dihydro-4(1H)-quinazolinone	334
17	7-ethoxy-4-ethyl-2H-chromen-2-one	218
21	2-(1,3-benzodioxol-5-yl)-3-benzyl-2,3-dihydro-4(1H)-quinazolinone	358
23	2-(1,3-benzodioxol-5-yl)-3-propyl-2,3-dihydro-4(1H)-quinazolinone	310
26	2-(4-bromophenyl)-N-(4-methyl-2-pyridinyl)acetamide	305
27	N-(2,4-dimethoxyphenyl)-4-iodobenzamide	383
30	N-[1-(anilinocarbonyl)-2-(1,3-benzodioxol-5-yl)vinyl]-5-bromo-2-furamide	455
31	N-(2-methoxybenzyl)-5-nitro-2-pyridinamine	259
32	2-hydroxybenzaldehyde (4,6-dimethyl-2-pyrimidinyl)hydrazone	242
33	2-{[4-(phenylsulfonyl)phenyl]thio}ethanol	294
34	2-[(4-chlorophenyl)thio]-N-(2-methoxybenzyl)acetamide	322
36	N-{4-[(4-chlorobenzyl)amino]phenyl}acetamide	275
37	(4-methoxyphenyl)(1-propyl-1H-benzimidazol-2-yl)methanol	296
38	1,3-benzodioxol-5-yl(1-propyl-1H-benzimidazol-2-yl)methanol	310
40	3-methyl-2-phenyl-5,6,7,8-tetrahydro[1]benzothieno[2,3-b]pyridin-4-amine	294
41	1-methyl-4-(7-nitro-1,2,3,4-tetrahydro-4aH-xanthen-4a-yl)piperazine	329
42	N-(5-chloro-2-methylphenyl)-3-(1,3-dioxo-1,3-dihydro-2H-isoindol-2-yl)propanamide	343
43	5-(2,3-dichlorobenzyl)-1,3-thiazol-2-amine	259
44	4-(2-bromo-4,5-dimethoxyphenyl)-3,4-dihydrobenzo[h]quinolin-2(1H)-one	412
45	isonicotinaldehyde (4-methyl-2-quinolinyl)hydrazone	262
Indaziflam	2-N-[(1R,2S)-2,6-dimethyl-2,3-dihydro-1H-inden-1-yl]-6-(1-fluoroethyl)-1,3,5-triazine-2,4-diamine	301.37
CMC	4-chloro-2(methylthio)pyrimidine-5-carbontrile	185.63
ACC	4-amino-6-chloro pyrimidine-5-carbonitrile	154.56
DPC	4,6-dichloro pyrimidine-5-carbonitrile	173.99
APC	2-aminopyrimidine-5-carbonitrile	120.11
DCB	2,6-dichlorobenzonitrile	172
Isoxaben	2,6-dimethoxy-N-[3-(3-methylpentan-3-yl)-1,2-oxazol-5-yl]benzamide	332.4
Flupoxam	1-[4-chloro-3-(2,2,3,3,3-pentafluoropropoxymethyl)phenyl]-5-phenyl-1,2,4-triazole-3-carboxamide	460.79

2.2.5 DNA preparation and whole-genome sequencing (WGS)

A DNA purification kit for Solexa was used to prepare genomic DNA for whole-genome sequencing. Genomic DNA was collected from stable *izr* mutants on 20 µM indaziflam. Gametophyte tissues were frozen with liquid nitrogen and ground to produce a fine powder. Ground plant tissue was moved to a 1.5 ml tube, and 600 µl of Cell Lysis Solution

(Qiagen®, Cat#158906) was added. The solution was mixed and incubated at 65°C for an hour. After one hour, 1.2 µl of 10 mg/ml RNase A solution was added to the tube, mixed by inversion, and allowed to incubate at 37°C for an hour. After the sample was settled and cooled at room temperature, 200 µl of Protein Precipitation Solution (Qiagen®, Cat#158910) was added, and the tube was mixed by inversion for approximately 2 minutes. The sample was incubated on ice for 1 hour. After incubation, the sample was centrifuged at 13,000 rpm for 10 minutes at 4°C. After centrifugation, the supernatant was transferred to a new tube and incubated on ice for 5 minutes. The 10 minutes centrifugation, transfer to a new tube, and 5 minutes incubation were repeated three more times to ensure DNA purity. The purified supernatant was transferred to a final 1.5 ml tube containing 600 µl of cold 100% isopropanol and mixed by inversion 50 times. The tube was centrifuged at 13,000 rpm at 4°C to collect the genomic DNA pellet. The supernatant was discarded, and the pellet was washed with 300 µl of 70% ethanol. After final centrifugation at 13,000 rpm at 4°C for 5 minutes, the 70% ethanol supernatant was discarded, and the DNA pellet was air-dried for approximately 30 minutes. The genomic DNA pellet was rehydrated into 25 µl of 10 mM Tris buffer (pH 8.0).

The solution was gently tapped to bring DNA into the solution and rehydrated at 65°C for 1 hour. DNA concentration and quality were measures using a BioDrop Duo (80-3006-61). Whole Genome Sequencing (WGS) was completed by BGI Genomics, using the Illumina HiSeq 2500 platform. Sequencing data was compared to the *Physcomitrium patens* Gransden ecotype reference genome to identify SNPs' location across the sequenced genome. The identified SNPs were filtered to find SNPs caused by UV mutagenesis (G to A, or C to T transitions) located in the DNA coding regions. Data were analyzed using genomics tools. Dr. Hossein Lanjanian from Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, analyzed the WGS data. We have chosen five different genomics tools: bowtie2, samtools, bcftools, annovar, and Vep. These are the most cited based on literature and accept standard alignment input formats. These programs have varied approaches to variant identification and produce different variant calls [165, 166].

2.2.6 Somatic hybridization and creation of mapping population

Physcomitrium patens is a monoicous moss, meaning the plant bears both sperm and egg on the same gametophyte and can self-pollinate predominantly in the wild. Crossing techniques usually have been used between the Gransden (Gr) ecotype and the genetically divergent Villersexel (Vx) ecotype for genetic mapping. Vx is the most divergent amongst *P. patens* and can produce sporophytes upon crossing [167]. Vx displays a Single Nucleotide Polymorphism (SNP) rate five to ten times higher than Gr [168]. Also, Vx is the ecotype other than Gransden, commonly used in forward genetics studies. Somatic hybridization between indaziflam mutants and the Vx wild-type was performed using the PEG-induced protoplast fusion method, with some modifications [169]. Protoplasts were isolated from *izr* mutants, and Vx used the same protoplast isolation procedure as described previously and combined the two strains for hybridization. Cells were resuspended in 6 ml 8% mannitol at a final cell density of 1×10^6 cells ml^{-1}. One milliliter of each cell suspension (1×10^6 cells total) was combined, mixed gently, and then centrifuged at 200g for 4 minutes at room temperature with no braking. The supernatants are discarded, and protoplasts were resuspended in 250 µl of the wash solution containing 8.5% mannitol and 10 mM $CaCl_2$. Somatic hybridization was started by adding 750 µl fusion solution (Polyethylene glycol (PEG) MW 6000 50% (w/v), $CaCl_2$ 5mM) and mixing gently. After 40 minutes, 1.5 ml of wash solution was added. After 50 minutes, another 10 ml of wash solution was added. A further 10 ml of wash solution was added after 60 minutes. Following adding the wash solution, protoplasts were mixed gently and incubated at room temperature. Finally, After 70 minutes, protoplasts were centrifuged at 200g for four minutes at room temperature with no braking, and the pellet was resuspended in PRMT held at 45°C, and 1.5 mL of suspension was spread on each PRMB plates. Plates were incubated for five days at 25°C with continuous light at 50–80 $mmol/m^2/s$ [162, 170]. After 5-7 days, cellophane discs were removed and transferred to BCDAT medium supplemented with 20 µM indaziflam and incubated under the same conditions for several weeks to select for stable hybrids insensitive to indaziflam. Control plates on each component strain were carried out in parallel.

The second technique, sporulation, and crossing experiment were performed between *izr* mutants (with Gr background) and Vx ecotype. Protonema or gametophore tissues were inoculated from BCD plates to plant pots on autoclaved peat pellets for five weeks at 25°C with continuous light at 50–80 mmol/m^2/s. Next, the pots were transferred to 16°C with a short daylight cycle of 8 hours of light and 16 hours of the dark regime and a reduced fluence rate of 20 µmol/m^2/s white light for the rest of the experiment [161]. After two weeks at 16°C (upon gametangia formation), the culture is submerged with water for 24 hours, after which the excess water is drained, leaving a moist culture. The same operation is repeated one week later to assure complete fertilization. The first visible green sporophytes could be observed three weeks after the first watering. The *izr* mutant plants, which had a Gransden strain background, exhibit low levels of male fertility, and accordingly, the appearance of developing sporophytes on this strain was generally indicative of cross-fertilization by the Vx parent. Spore capsules on the Gransden parent were collected when capsules were becoming uniformly dark brown and mature, usually 2 to 4 weeks later [171]. Sporophytes were air-dried in sterile tubes for one week before storage at room temperature and performing germination assays. Mature sporophytes were surface-sterilized with 30% bleach for five minutes, then washed with sterilized water 4-5 times. Sporangia were broken open in sterile water to make a spore solution. Next, the spore solution was spread on BCDAT plates supplemented with 10mM calcium chloride, overlaid with cellophane disks at 25°C under continuous light to induce the development of protonema. Spore germination was recorded seven days after inoculation. After 7-10 days, the germinated spores were transferred to the BCDAT medium supplemented with 20 µM indaziflam and incubated under the same conditions to select for indaziflam insensitive plants (Figure 2.3).

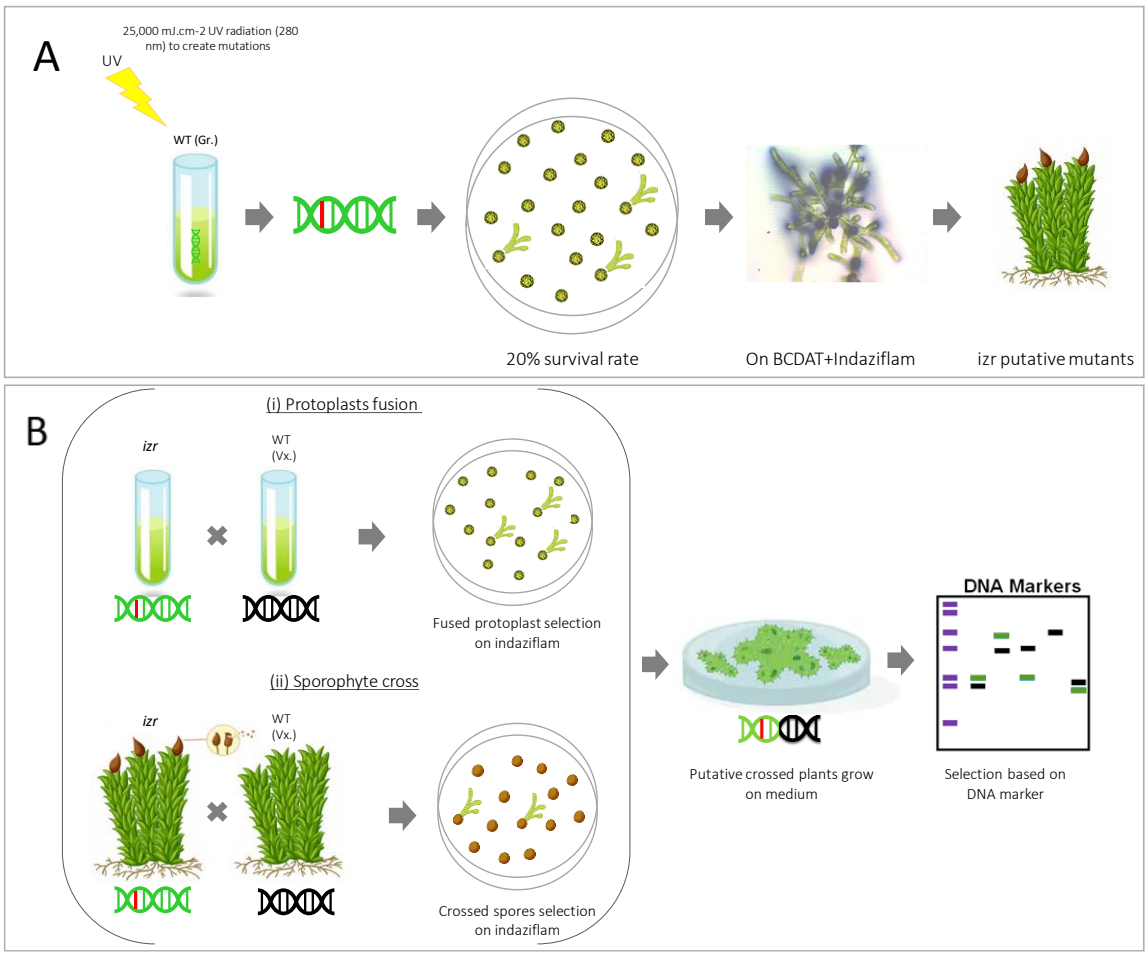


Figure 2.3: (A) Diagram outlining forward genetic screening by UV mutagenesis. (B) Two bulk segregant mapping techniques (i) protoplast fusion and (ii) sporophyte cross in *Physcomitrium patens*.

2.2.7 AFLP analysis of *Physcomitrium patens* ecotypes

The genetic analysis of the progeny of a cross between a mutant insensitive to indaziflam (*izr*) generated in the Gr background with Vx allowed the identification of the causal gene for the detected phenotype, and only ones demonstrating the phenotype of interest were retained to remove mutations that were not responsible for the phenotype. UV-induced mutations and the mutations found in multiple mutants were subtracted to leave only a handful of variations to examine in a relatively small portion of the genome. From there, we would be able to identify the mutation causing the *izr* phenotype [172].

An amplified fragment length polymorphism (AFLP) marker has been used to detect crossed sporophytes and discriminate between the Villersexel and Gransden genotypes. In

the Gransden-specific allele, the intercalation in linkage group (LG) 11, a marker appears to be inverted and then causes a misidentification by SCALPHUNTER [173]. This intercalation in the Gransdon genome causes a mutation A113G/G113V, eliminating the MseI restriction enzyme site in the Villersexel genotype and the 2-bp difference in fragment length between the Villersexel and Gransden alleles [173].

Primers were used to amplify the allelic fragments from both ecotypes genomic DNA (MseI-Forward: AATTCACAATTCGTGGATGCA and MseI-Reverse: AGAGGAGGAAATGGTATAGCCATC). About 500 bp DNA amplicon was generated from each parental genotype and digested with MseI restriction enzyme (NEB), which showed an additional MseI site in the fragment amplified from the Gransden genotype, the existence of this restriction enzyme site generates the Gransden-specific AFLP.

2.2.8 Identification of the causal mutation

Identification of missense mutations in the *izr* mutants' gene was confirmed by polymerase chain reaction (PCR) amplification and restriction enzyme digestion of a fragment. The genomic DNA was extracted based on the genomic DNA extraction cetyltrimethylammonium bromide (CTAB) method with minor modifications [174]. Gametophore tissues were frozen with liquid nitrogen to yield fine powder and transferred to a tube. A 500 µl volume of CTAB buffer containing 2% w/v CTAB, 1.4 M NaCl, 100nM Tris HCl pH 8.0 buffer, and 20 mM EDTA was added to the ground plant tissue. The solution was incubated at 55°C for an hour. After incubation, 500µl of chloroform was added to the cooled solution. The solution was mixed by inversion, followed by centrifugation for 7 minutes at 16000 g. The aqueous phase was transferred to a new tube, and 0.08 volumes of cold 7.5M ammonium acetate and 0.54 volumes of cold isopropyl alcohol were added. The solution was mixed by inversion, followed by incubation on ice for 30 minutes. Following incubation, genomic DNA was collected by centrifugation for 5 minutes at 16000g. The supernatant was discarded, and the DNA pellet was washed with 700 µl 70% ethanol, followed by an additional wash with 700 µl 95% ethanol. The DNA pellet was allowed to dry and then rehydrated in sterilized water [174].

The candidate gene's amplification was carried out by PCR using the custom primers from mutant, wild-type, and wild-type segregants using *Taq* DNA polymerase (Takara). The PCR fragments contain the mutation site in *izr* mutant and wild-type segregants. Genomic DNA samples were taken from three biological replicates to ensure that the identified mutation was present in multiple resistant plants. The *izr* mutation may create a restriction enzyme site. In such a case, the PCR product was cut with the enzyme, and fragments were run on a 3% agarose gel to separate DNA fragments and determine the genotype.

2.2.9 Growth assays

The *izr* response was assayed by measuring the growth of explants, small tweezer-picked fragments of protonemal tissue inoculated on BCDAT, and 20 µM indaziflam supplemented media. Two replica plates were set up in which each was split into three equal sections, within which ten explants from each line were inoculated with even spacing. These lines included the wild-type Gransden strain, *izr6*, and *izr10*. Digital photographs were taken every 3-4 days and analyzed using ImageJ software to calculate the mean colony size over time [164, 175]. The software converts the digital images to binary format and determines the colony area based on counting the number of pixels corresponding to each colony [164]. Colony area determinations based on different photographs were normalized for each colony using the plate's estimated area. Data processing included removing irregular explants such that the minimum sample size analyzed per line was n=8.

3.2.10 Phenotypic characterization of the *P. patens izr* mutant lines

Fresh protonemal tissue was inoculated on BCDAT and BCD medium supplemented with various synthetic phytohormones (Figure 2.4) and grown into well-developed colonies for one month to observe phenotypic characterization. Photographs were taken for each colony, and typical representations are shown in the following sections.

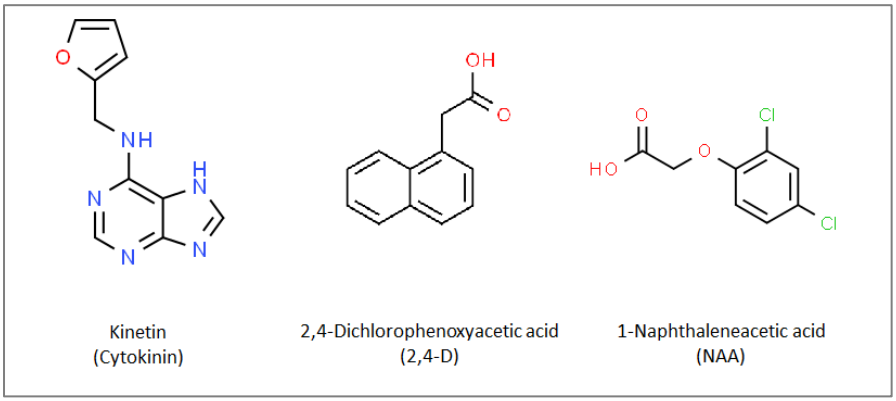


Figure 2.4: Synthetic plant hormones. Chemical structures are from www.chemspider.com.

2.2.11 Quantitative reverse-transcription polymerase chain reaction (qRT-PCR) analysis

The expression of the upregulated and downregulated genes in response to auxin and cytokinin was measured using qRT-PCR to confirm the reliability of the results for the dataset. The 15 top genes from auxin and cytokinin pathways were selected, and gene-specific primers were generated for each (Table 2.3).

Total RNA from 2-week-old protonema tissue of *P. patens* grown on BCDAT was extracted with TRIzol® based on the protocol [176], and 5µM IAA was used for auxin treatments. One µg of total RNA was used to synthesize double-stranded cDNA with a random primer mix using M-MuLV Reverse Transcriptase (New England BioLab®). Quantitative RT-PCR was performed using a CFX Connect Real-Time System (BIO-RAD) and SsoFast™ EvaGreen® Supermix (BIO-RAD), with three independent biological replicates. *PpEF1α* (elongation factor 1-alpha, Phypa_439314) was used as a reference gene to calculate the relative expression. qRT-PCR standard curves were prepared for each primer set using a series of 10X dilution, including no-template (NTP) and no-reverse transcriptase (NRT) controls. Five µl of the 1/10 diluted cDNA was used, and qRT-PCR reactions were performed for 30 cycles. Relative fold change in gene expression was estimated following the threshold cycle (Ct) values. On the one hand, differences in the ΔCt were detected among analyzed (treated) and reference (untreated) *P. patens* tissues (ΔΔCt = ΔCttreated – ΔCtuntreated).

Table 2.3: List of primers used for qRT-PCR.

Primer name	Sequence	Primer name	Sequence
PpEF1α-F’	ACGCGTTGTTGGCTTTCACTTTG	PpGH3-1-F'	AGAGATCGTGGAGGACATCA
PpEF1α-R'	GTGGTTGCGTCCATCTTGTTGC	PpGH3-1-R'	CTACTTGCAAGGTCTGGGTTAG
PpIAA1A-F’	Atccgggagtccgagcttc	PpGH3-2-F'	CGACAAGTCCTGTTGACATAGT
PpIAA1A-R’	ggttctgcgcaggaggtg	PpGH3-2-R'	CAGACATCTGCTCCACTTCTT
PpIAA1B-F’	cggtggtcagaatgggtca	PpRSL2-F'	AGCAGGCTAGAAAGCATCAG
PpIAA1B-R’	Cccacagtctggttctgcg	PpRSL2-R'	AGGCTTGAGGTTGGTGTTC
PpIAA2-F’	tgccttgggactggttcatc	PpRSL1-F'	GGACCAAACTGGATCGATGAATA
PpIAA2-R’	cacagcaccttgggctttca	PpRSL1-R'	TGGTGAGCTAAAGCACTGATAC
PpAUX1-F'	TGGAGGTCACGGAATCACCATC	PpSOUR-F'	CTCTAGCGCTACCTCCAATTC
PpAUX1-R'	ACACTGTCACTGCTATGCACAC	PpSOUR-R'	TGCGAACATCAGTCCTACTTC
PpARFa8-F'	CAGATGAGTCAGGAGCAACTTGAG	PpCUL1-06-F'	GAGGCGATCGAGGATACAC
PpARFa8-R'	ATTCGACGTGTTCACAGGAGACG	PpCUL1-06-R'	GCACTTTTGTCAAAGAGCAGC
PpARFa6-F'	CAGATGAGTCAGGAGCAACTTGAG	PpCUL1-05-F'	CAAACAAGCTGAAGATGCAGTC
PpARFa6-R'	ATTCGACGTGTTCACAGGAGACG	PpCUL1-05-R'	CAATCACCTTTCGCACAAAGAC
PpRSL4-F'	TCAAACGGCCGAAACATTCTACG	PpCUL1-16-F’	GTGAAGCAAGCTGAAGATGCT
PpRSL4-R'	CAGCTCCGCTCCTTTCAGAATATG	PpCUL1-16-R'	GACCTTCCGCACAAAGGC

2.3 Results

2.3.1 Morphological analysis of CBI screen

P. patens germinates from either a haploid spore or a protoplast. There are two different cell types in *P. patens*, chloronemal and caulonemal cells form the protonemata. They have the characteristic tubular shape, which is growing at the tip while connecting at one end to the next filament cell. These filaments grow by two main mechanisms: unequal cell division and cell elongation at the tip [177, 178]. Protonemal filaments of *P. patens* extend by apical cell division and tip growth, branching to form colonies [97]. These filaments are characterized by rapid wall synthesis and a polarised structure from the first regenerated cells [179].

The result for testing the in-house CBI collection at 20 µM concentration on protoplast regeneration in wild-type moss is summarized in table 2.4, shown in figure 2.5. The concentration has been chosen based on the growth effects on *Arabidopsis thaliana* (Bonetta, unpublished). This experiment helped us to determine possible compounds for

the CBI screens. The regeneration response to different CBIs was varied. Some CBIs did not affect growth and development compared to control, and for some of the CBIs, moss was very sensitive and died. On the other hand, some CBIs had effects on the chloronemal tissue development and caused a bids-like phenotype. Based on previous studies on protoplast tip growth and the cell wall formation [177-179], and based on the fact that CBI is targeting cell wall synthesis components [129, 131], I believe that in the affected plants by the CBIs, the polarized tip growth was disrupted and cells are unable to regenerate the cell wall.

Since among all CBIs, indaziflam can inhibit cellulose production [136], and *A. thaliana* mutants showed a decreased sensitivity to the auxin mimicking herbicide, 2,4-D [180], it was predicted that a cell wall-related protein would cause indaziflam resistance. In this screen for CBI resistance, indaziflam was chosen to study the mode of action associated with this potent herbicide.

Table 2.4: Effect of CBI collection of *P. patens* growth

Growth effect of *P. patens*	CBI Name
No effect (normal growth)	CBI5, CBI8, CBI9, CBI14, CBI15, CBI19, CBI26
Abnormal growth	CBI3, CBI17, CBI21, CBI23, CBI31, CBI34, CBI36, CBI37, CBI41, CBI42, CBI45, indaziflam, CMC, DCB, DPC, Quinoxyfen, TBC
No growth or death	CBI18, CBI27, CBI32, CBI33, CBI38, CBI43

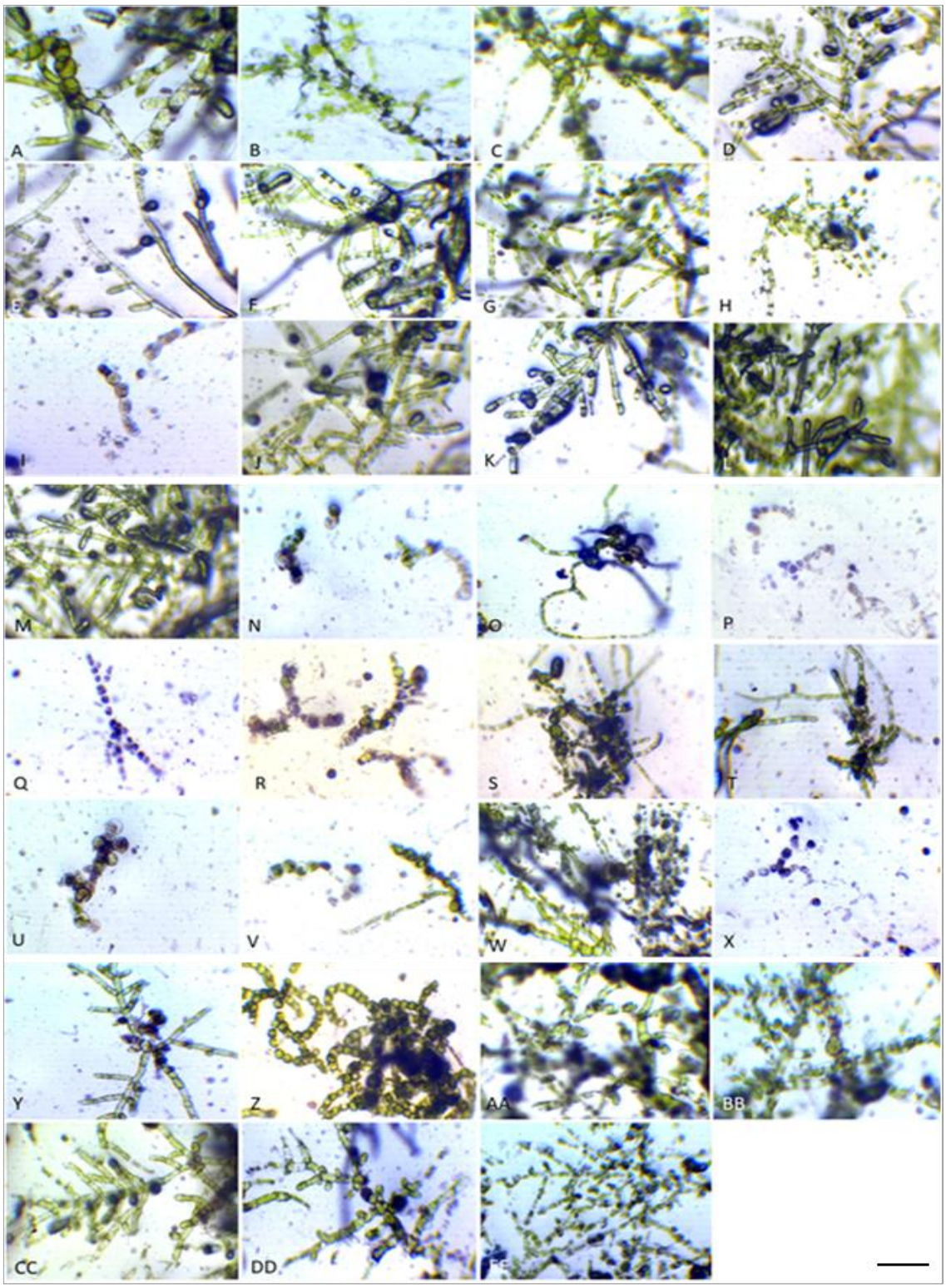


Figure 2.5: CBIs collection testing on WT plant: A: No CBI, B:CBI3, C:CBI5, D:CBI8, E:CBI9, F:CBI14, G:CBI15, H:CBI17, I:CBI18, J:CBI19, K:CBI21, L:CBI23, M:CBI26, N:CBI27, O:CBI31, P:CBI32, Q:CBI33, R:CBI34, S:CBI36, T:CBI37, U:CBI38, V:CBI41 W:CBI42, X:CBI43,Y:CBI45, Z:indaziflam, AA:CMC BB:DCB, CC:DPC DD:Quinoxyfen EE:TBC (Scale bars: 100µm)

2.3.2 Indaziflam non-responsive mutants displaying strong indaziflam resistance

Sunlight contains three UV categories based on the respective wavelength: UV-A (320 to 400 nm), UV-B (290 to 320 nm), and UV-C (<290 nm). UV-B and UV-A can effectively reach the earth's surface while UV-C is filtered out in the atmosphere [181]. UV-A radiation has less DNA damaging effect since native DNA cannot absorb it. However, UV-A and visible light energy (up to 670 - 700 nm) can damage DNA via indirect photosensitizing reactions-mediated reactive oxygen species (ROS) generation, especially singlet oxygen ($^1O_2$) [182]. UV-C has the maximum absorption by DNA and is the most potent mutagen [183]. UV produces DNA damage via a direct or an indirect mechanism. The direct mechanism affects two neighboring pyrimidines (thymines or cytosines). The indirect mechanism leads to oxidative mutations through ROS production, which attacks cellular nucleotide pools, producing oxidized nucleotides such as 8-hydroxydeoxyguanosine-triphosphate (8OH-dGTP), and causes oxidative base damage [183, 184].

UV-C was used to induce mutation in the Gransden background protoplasts. Mutations affecting herbicide growth responses were identified by the germination of the protoplasts on indaziflam-containing media. If wild-type protoplasts or protonemal explants are regenerated in the presence of indaziflam, the cells show a characteristic change in growth pattern. Changes in growth patterns were characterized by cellular differentiation producing small, round, thick-walled cells in which cell expansion is suppressed (Figure 2.6), the development of chloronemal cells disrupted, and the protonemal tissue was not developed which leads to cell death. Some plants show a highly reduced growth rate with a ruptured tip and dwarf and twisted protonemal morphology in case of survival.

In such a way, the indaziflam non-responsive mutants have a typical growth rate despite the presence of indaziflam, which contrasts with the wild-type response. Some of these mutants exhibited premature aerial gametophore tissue, and some of them were not able to form aerial gametophore tissue. Several independent mutants showing the indaziflam non-responsive (*izr*) phenotypes were isolated.

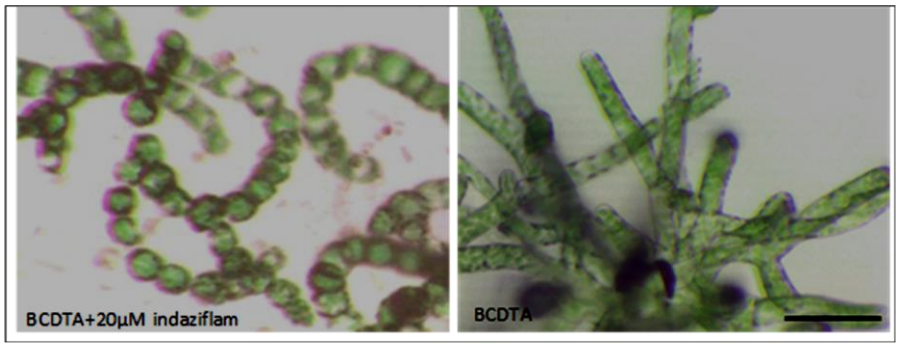


Figure 2.6: Wild-type response of developing protonemal tissue to 20 µM indaziflam (Scale bar: 100 µm)

2.3.3 Indaziflam resistance screen yielded 46 *izr* mutant lines

Approximately 5,000,000 protoplasts from 11 pools were screened on 20µM indaziflam. All protoplasts with at least a minor decrease in sensitivity were selected as potential *izr* mutants. In total, 112 potential *izr* explants were selected and moved to BCD medium after 4-6 weeks of growth on indaziflam supplemented media to grow to maturity. Of the 112 potential *izr* explants, 46 lines grew to maturity. Of the 46 possible *izr* lines, 16 lines were not able or delayed to produce aerial gametophore tissue (Figure 2.7). Failing in developmental transition in protonema and leafy gametophore formation can indicate a defect in the phytohormone signaling pathway [185]. The phenotype of some of the *izr* lines which were not able or delayed to produce gametophore tissue was quite similar to auxin (*aux/iaaΔ*) [100] and cytokinin [186] mutants.

Auxin and cytokinin play an essential role and antagonistic activity in the root and shoot development of *Physcomitrium patens* [100]. Wild-type chloroplast-rich chloronemal cells in moss divide and form elongated caulonemal cells, producing buds and developing into leafy gametophores with brown-pigmented rhizoid filaments [96]. Auxin promotes the differentiation of chloronemata, which is a chloroplast-rich filament into elongated filaments with fewer chloroplasts called caulonemata, at the protonemal stage, while cytokinin is involved in the regulation of the *Aux/IAA* gene [100]. In contrast, at the stage of leafy gametophore formation, auxin stimulates stem elongation and rhizoid development [187]. Previous studies by Estelle et al. have shown the auxin-resistant (*aux*) mutants cannot transit from the primary chloronemal to caulonema stage, or this process

happens with a delay compared to wild-type and exogenous auxin increases this transition and converts gametophore tissues to rhizoid cells. Some mutants also fail to form leafy gametophores unless they are treated with exogenous cytokinin [188, 189].

Figure 2.8 shows two resistant lines on a range of indaziflam concentrations (zero as a control to 50 µM). Selected indaziflam non-responsive putative mutants were compared to wild-type plants on 1µM, 5, and 10µM 2, 4-D, NAA, and cytokinin.

Of those *izr* mutants with growth defects, only *izr10* was resistant to all three phytohormones and chosen for further analysis. Of the *izr* mutants exhibiting normal phenotype compared to wild-type, *izr6* showed strong resistance to NAA, 2,4-D, and cytokinin.

Prioritized mutants were prepared for whole-genome sequencing and moved to the soil for crossing in preparation for positional cloning. Further genotypic and phenotypic characterization was carried out to understand the cause of indaziflam resistance.

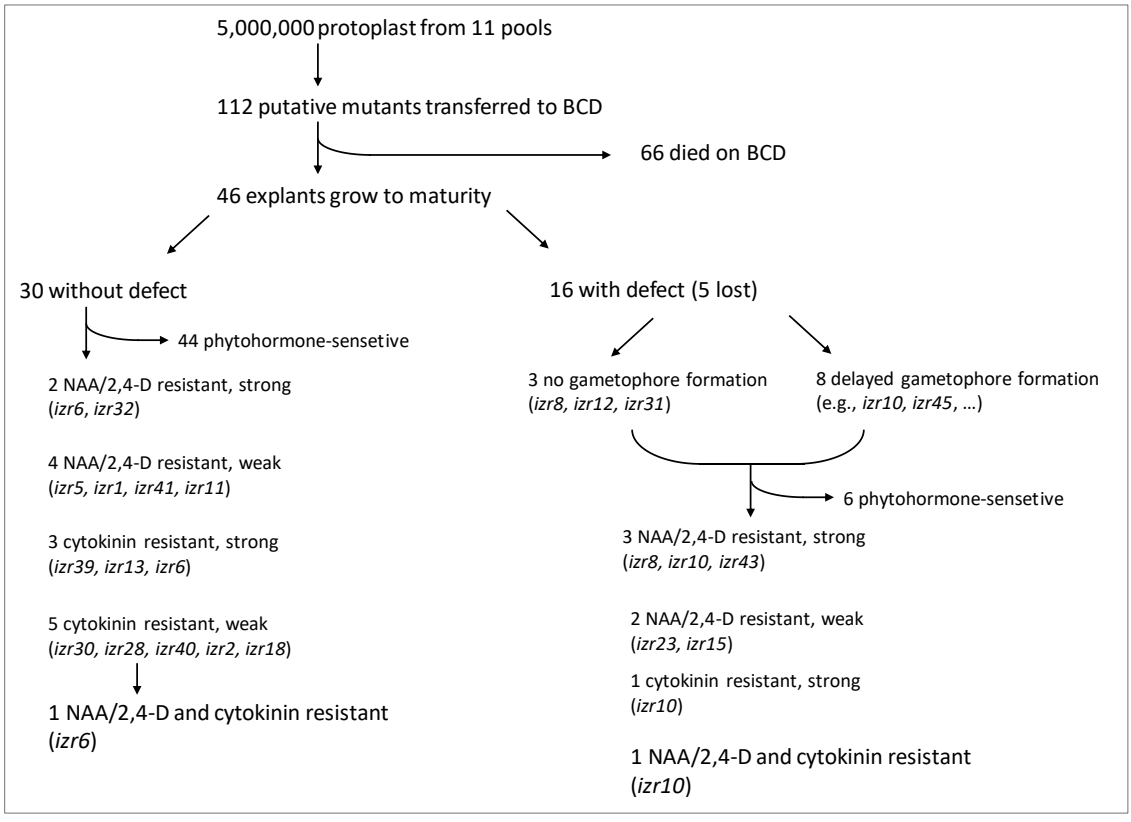


Figure 2.7: Flow chart of the "*izr* mutant screen." Five million UV-mutagenized wild-type protoplasts from 11 pools were screened on the indaziflam.

2.3.4 Mapping of causal loci

Genomic DNA from pooled *izr10* explants was used to obtain whole-genome sequencing analyzed by Dr. Hossein Lanjaninan, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences. Next-Generation Mapping identifies the causative mutation location by determining areas in the sequencing sample with low SNPs variation compared to the appropriate ecotype's reference genome [190]. Since *izr* is in the Gransden ecotype background, the sequencing results were compared to the Gransden reference genome. The regions, referred to as SNP, can be used to determine the location of a causative mutation [190], but a causative mutation could not be determined by next-generation mapping alone. The presence of SNPs can quantify the variation between genomic DNA from different ecotypes. Due to crossover frequency and haploid spore-derived progeny, DNA surrounding the causative mutation will also consist of homozygous DNA from the background ecotype. This process relies on the basic principle that SNPs closer together are less likely to be separated by recombination events between them, which approaches 0% when adjacent [190].

The independent indaziflam non-responsive mutants, *izr6* and *izr10*, were cross-fertilized by co-culture with Vx strain and protoplast hybridized crossing, respectively, to generate segregating populations [52]. In cross-fertilization, hybrid diploid sporogonia recovered from the mutant with Gr maternal background plants, sporelings were identified by spore germination and growth on indaziflam-supplemented medium. Hybrid sporogonia yielded spores that segregated 1:1 for the *izr* and wild-type phenotypes. In *Physcomitrium*, the gametophyte produces gametes by mitosis that fuse to generate the diploid sporophyte and generates haploid spores by meiosis, so the spores obtained from hybrid sporogonia represent a segregating population containing recombinant chromosomes, the equivalent of recombinant inbred lines in diploid species [52]. Ten segregants from the *izr6* mutants were selected on 20µM indaziflam and pooled.

PCR-based genotyping using the selected SNP-specific primers to distinguish Gr from Vx confirmed the mapped region as segregating with the phenotype [52] in the *izr10* x Vx cross (i10V) segregants. Hybrids from protoplast fusion were obtained using the SNP in the

Gransdon genome, which causes a mutation and eliminates the MseI restriction enzyme site in the Villersexel genotype [173]. Strains that contain Villersexel SNP and resistant to indaziflam were characterized as a segregation population for *izr10* mutants. All *izr* segregants (20 explants) displayed the Gr genotype, while all wild-type segregants (54 explants) displayed the Vx genotype as expected.

SNPs on different chromosomes (linkage groups) or sufficiently separated on the same chromosomes (i.e., at opposite ends) will show no association (50% assortment in progeny) [73]. To confirm the causative mutation, candidate genes are required sequencing within the mapped region. Twenty segregants from the (i10V) were selected on 20µM indaziflam and pooled. Non-crossed *izr10* explants were also selected on 20µM indaziflam. The pooled i10v segregants' genome would consist of an equal mixture of DNA derived from Vx parental strain with neutral marker allele as a result of random crossing-over during meiosis and since allowed the identification of the causal genes for the *izr* phenotype and eliminated mutations that were not responsible for the phenotype. The SNP and the location of a causative mutation are determined by analyzing the pooled i10V genomic DNA sample due to the subtle resistance phenotype of the *izr* explant on indaziflam.

2.3.5 Identification of the *izr* locus in *Physcomitrium patens* genes

Along with the resistance to indaziflam, the *izr6* and *izr10* explants demonstrate a clear resistance to the synthetic auxin NAA, 2,4-D, and cytokinin (Figures 2.8-2.10). 2,4-D and cytokinin allowed for a specific selection of explants for the causative *izr* alleles.

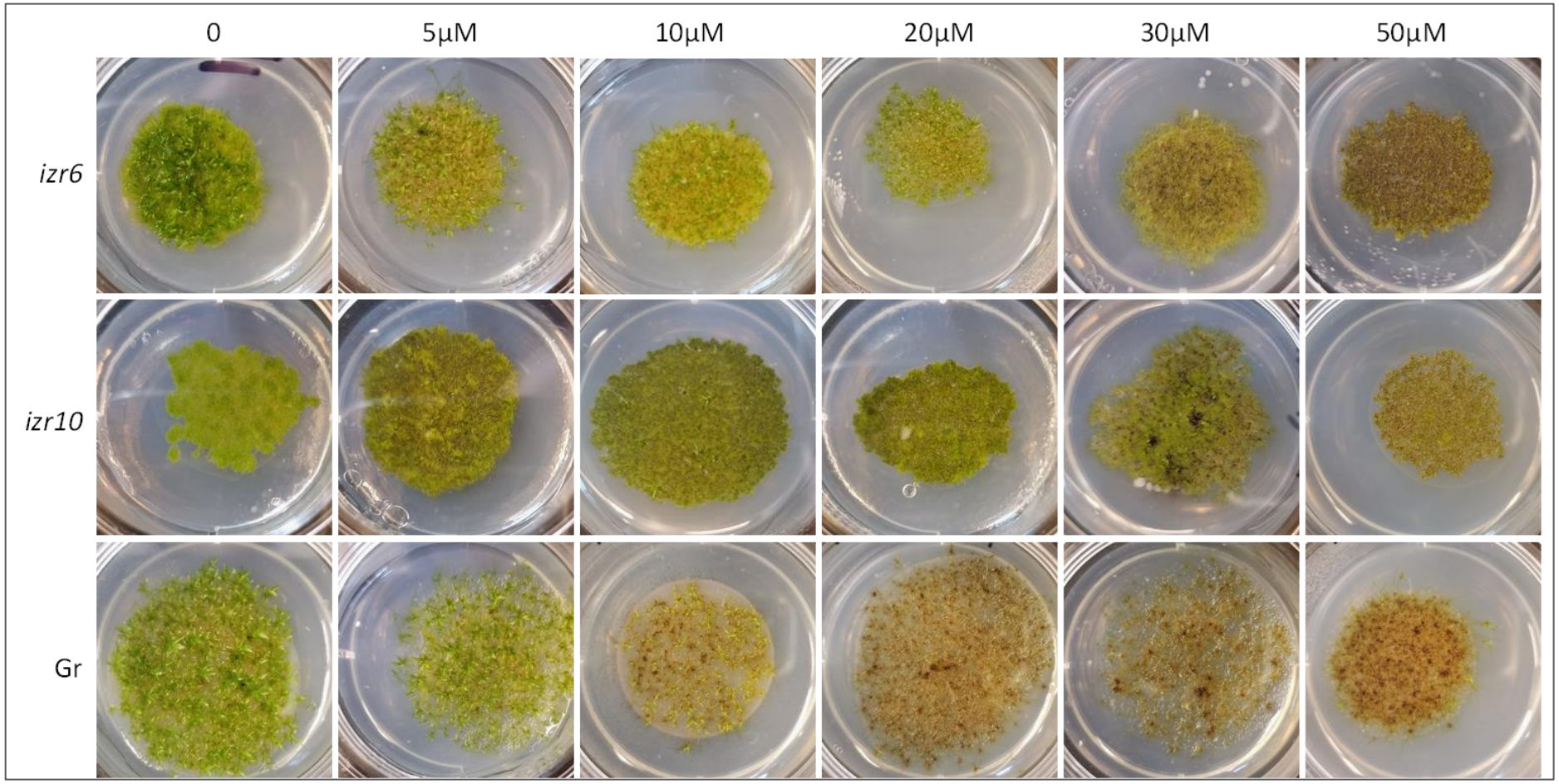


Figure 2.8: *izr6* and *izr10* mutants are grown BCD media contain indaziflam (0-50µM) and compared to Wild-type (WT). The panel's top row shows the *izr6* mutant growth, the photographs in the middle show *izr10* mutant growth, and the pictures in the bottom row show WT. Photos are taken after one-month growth on BCD media supplemented with zero (control) to 50 µM indaziflam.

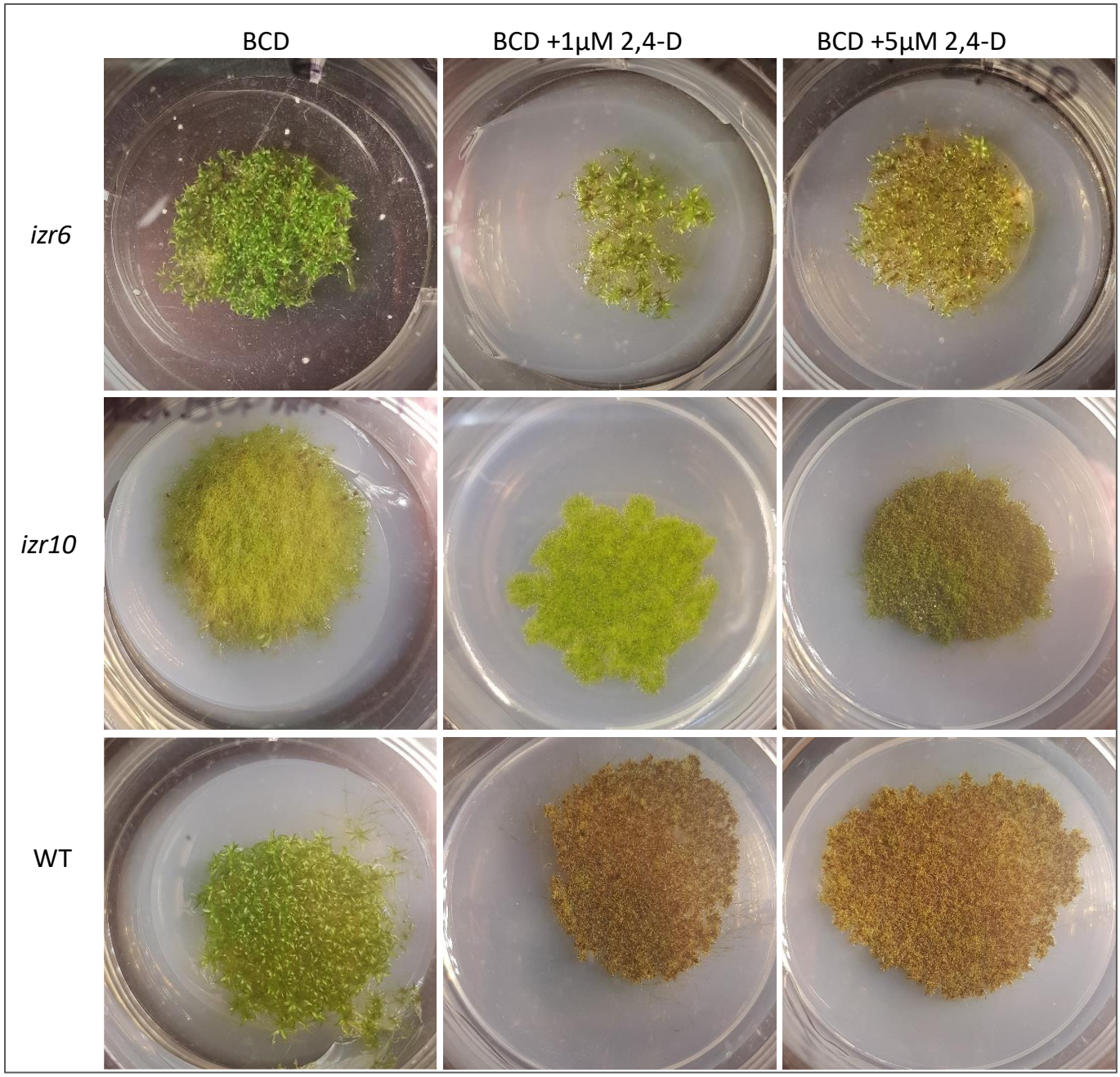


Figure 2.9: *izr6* and *izr10* mutants are resistant to 1µM and 5µM 2.4-D compared to Wild-type (WT). The left micrograph in each panel shows the plant growth on BCD media after one month of subculture. The photographs in the middle of each panel show WT and mutants' phenotypes after one month on BCD media supplemented with one µM 2,4-D. The pictures on each panel's right show WT and mutants' phenotypes after one month on BCD media increased with five µM 2,4-D.

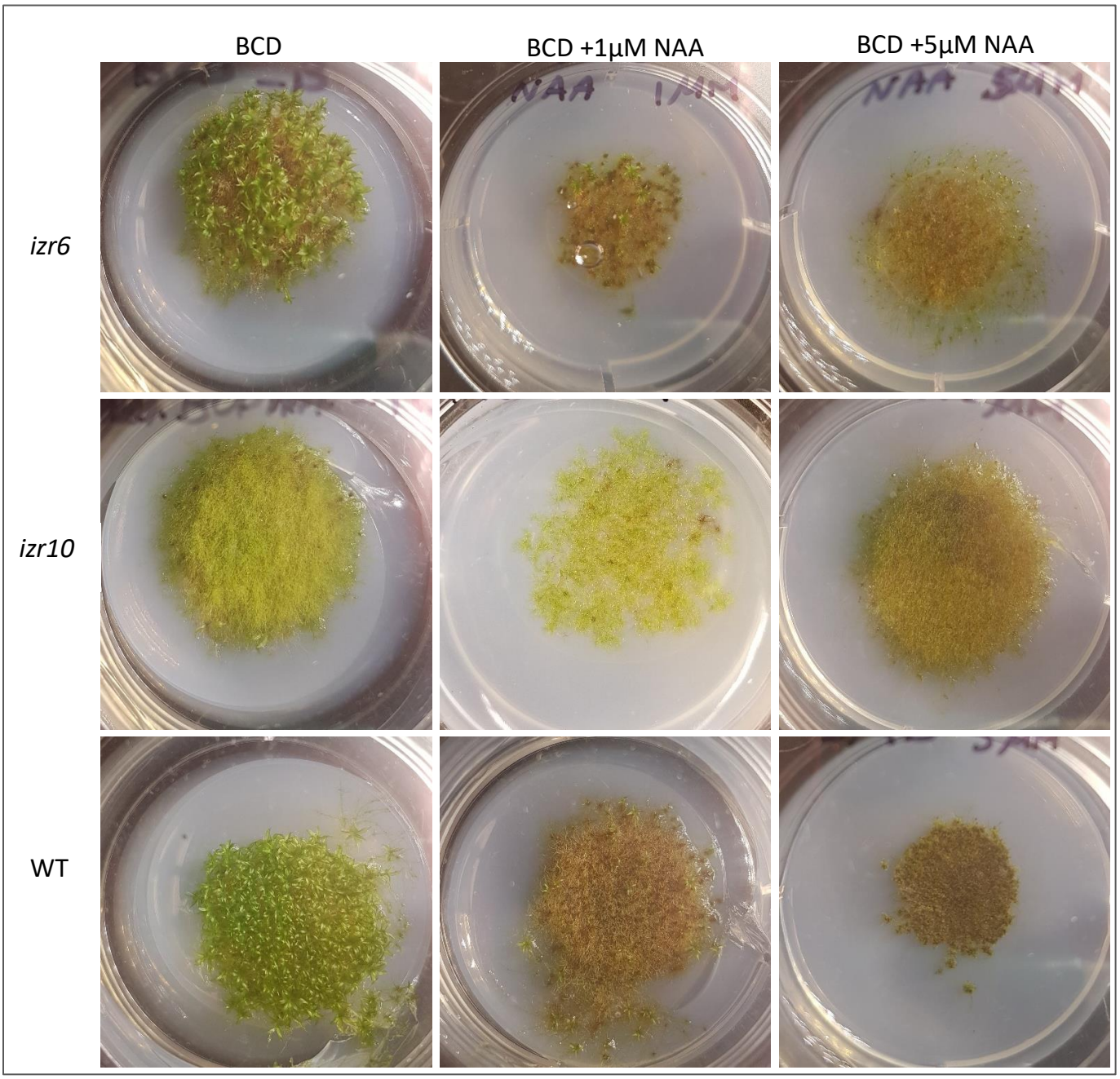


Figure 2.10: *izr6* and *izr10* mutants are resistant to 1µM and 5µM NAA compared to wild-type (WT). The left micrograph in each panel shows the plant growth on BCD media after one month of subculture. The photographs in the middle of each panel show WT and mutants' phenotypes after one month on BCD media supplemented with one µM NAA. The pictures on each panel's right show WT and mutants' phenotypes after one month on BCD media increased with five µM NAA.

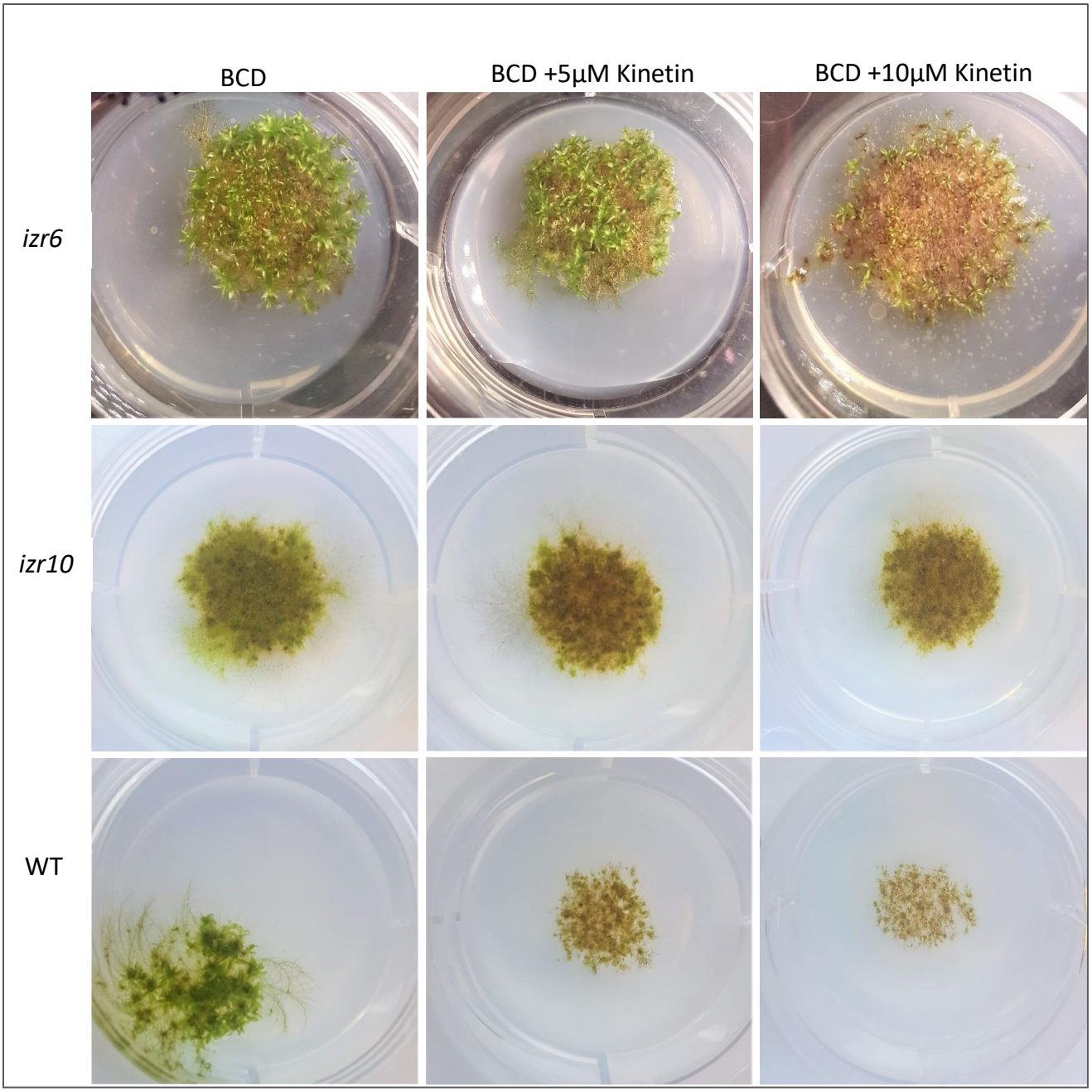


Figure 2.11: *izr6* and *izr10* mutants are resistant to 5µM and 10µM kinetin compared to wild-type (WT). The left micrograph in each panel shows the plant growth on BCD media after one month of subculture. The photographs in the middle of each panel show WT and mutants' phenotypes after one month on BCD media supplemented with five µM kinetin. The pictures on each panel's right show WT and mutants' phenotypes after one month on BCD media increased with ten µM NAA.

The whole-genome sequencing (WGS) of the *izr10* mutant demonstrated two causative mutations located in chromosome 8 and chromosome 16 (Figure 2.12A-B). Sequencing of SNP's coding region from WGS reveals a point mutation predicted amino acid substitution in chromosome 8, which caused missense mutation GGT>TTT in the open reading frame (ORF) (GGT>TTT: Gly207Phe). This mutation is a nonconservative substitution causing the change to the amino acid, which has different biochemical properties [191]. Glycine has a unique structure that contains hydrogen as its side chain (rather than carbon, such as all other amino acids). As a result, there is much more conformational flexibility in glycine than other amino acids with tight turns in structures. The rareness of glycine also means that it can play a distinct functional role, such as using its backbone (without a side chain) to bind to phosphates [192]. This means changing to any other amino acid could have a drastic impact on function [193]. So, glycine (Gly) substitution to phenylalanine (Phe) is unfavored to the plant.

This mutations is in the *PpIAA1B* (also known as Pp1s184_21V6; Phpat.008G051100; Pp3c8_14720 in the V3.1 assemblies). The presence of the GG to TT transition has been confirmed by amplification of the gene using polymerase chain reaction (PCR) using gene-specific primers PpIAA1B-F1 GAGAAGTGTTGCAGAGAAGA and PpIAA1B-F2 CACATGCAACAAGAGTCTCT and sequenced, using Sanger sequencing, at The Centre for Applied Genomics (TCAG) at Sick Kids Hospital (Figure 2.12). All selected plants contained the GG to TT transition in *IAA1B*.

I disregarded insertions and deletions and retained single nucleotide polymorphisms located in splice sites, introns, or the lesions that resulted in synonymous codon changes. Although mutations consistent with UV induces are specific types of mutation, single-base substitutions of cytosine to thymine (C→T) at dipyrimidine sites and CC→TT tandem base substitutions, while the last one rarely occurs [194]. Many purine residues (Adenine and Guanine) can also be subjected to UV-induced mutations in DNA from UV photoproducts at moderate UV dosages [195].

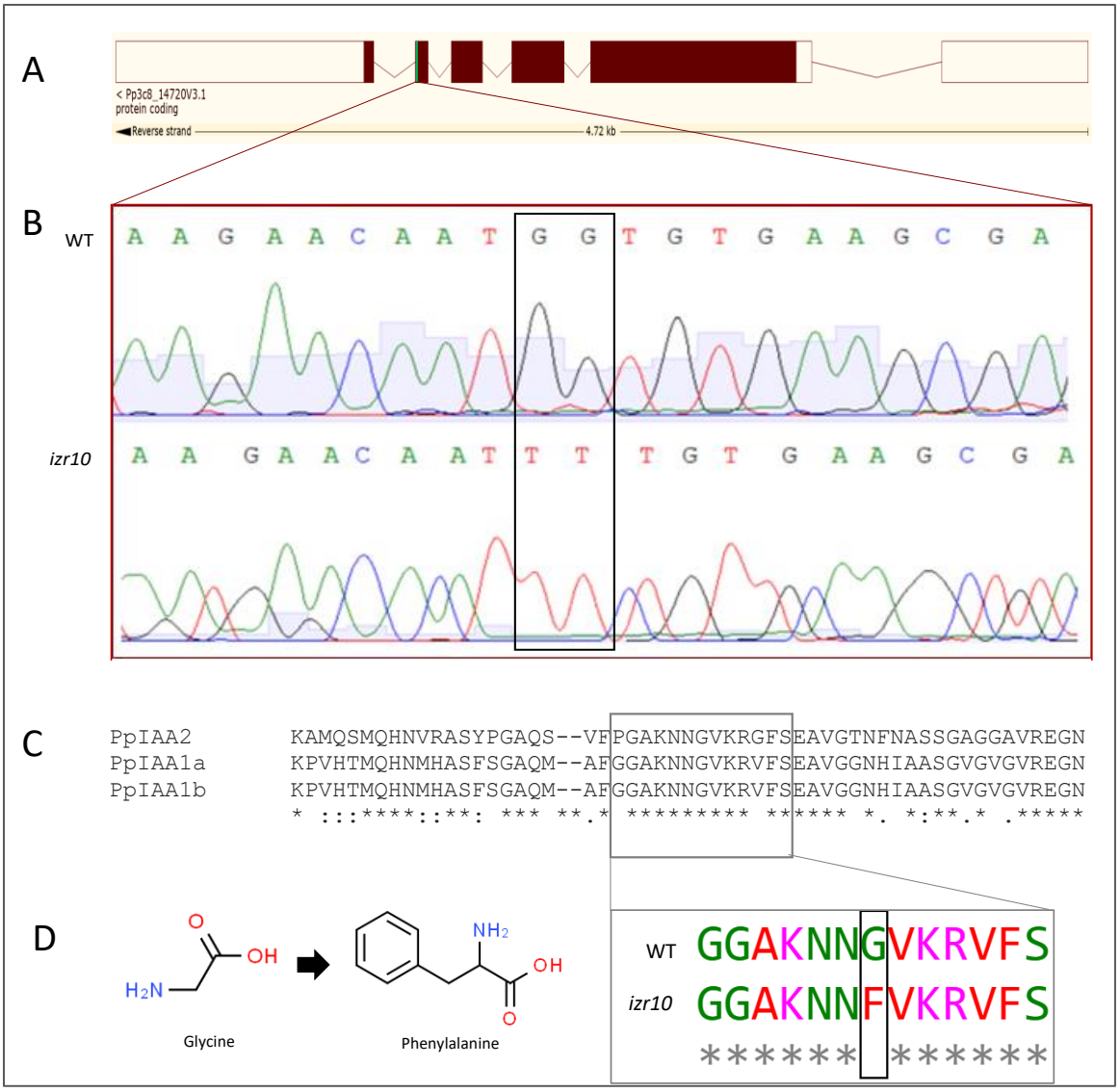


Figure 2.12: Genotyping revealed *PpIAA1B* mutation in *izr10* mutants. (A) Schematic representation of the *PpIAA1B* gene confirmation and mapped regions for *izr* mutations. (B) DNA sequencing electropherogram of the *PpIAA1B* transcript showing GG's transition to TT mutation in *izr10* mutants, black box shows the mutations' locations. (C) Partial protein sequence alignment of *PpIAA* genes. The alignment was made using Clustal Omega. An (*) indicates a conserved residue, a (:) indicates conservation between amino acids with strongly similar properties, (.) indicates semi-conservative replacement of amino acid. The protein sequence region containing the mutation is magnified, and the black box shows the mutation location. (D) Glycine is substituted for phenylalanine. Chemical structures are from www.chemspider.com.

The other missense mutation (CGT>CCT: Arg548Pro) is located on the north end of chromosome 16. The possible allele in the *izr10* mutant contained guanine to cytosine transition is one of the *CULLIN1* (Cullin/CDC53) genes in *Physcomotrium patens* referred to as *PpCUL1* (also known as Pp1s144_158V6; Phpat.016G006900; Pp3c16_1640 in the V3.1 assemblies). The presence of the G to C transition creates a BfaI restriction site in *CULLIN1*. To confirm the change in *izr10* mutant, genomic DNA was extracted from *izr10* explants selected on 20µM indaziflam. DNA fragments were amplified using *CULLIN1* gene-specific primers CUL1-16-F CGATAAAGATCTGTTTGCAGAGTTT and CUL1-16-R CATCTCTGCTGGTAAAGCCAG, and the fragments were digested with BfaI. All selected plants contained the G to C transition in *CULLIN1*, and this mutation was confirmed by testing this region in three different *izr10* mutants and 20 i10V segregants in which G to C substitution was found exclusively in the mutants and segregants (Figure 2.13E).

This allele results incorporation of proline in place of arginine at position 458 in the coding region of the CULLIN1 protein. Previous studies on *A. thaliana* showed mutations in the n-terminal of *AtCUL1*, which caused increasing tolerance to auxin and 2,4-D [180, 196, 197]. Other reports also have described that *CUL1* mutations (*cul1-6)* result in cytokinin resistance, which is an indirect effect of auxin resistance since the response to auxin or cytokinin is positively related [196]. CULLIN1 is a hydrophobic protein that acts as a scaffold for ubiquitin ligase (E3) and targets ubiquitin-mediated destruction proteins. There are also NEDD8-modified proteins (NEDD8 (NEURAL PRECURSOR CELL-EXPRESSED, DEVELOPMENTALLY DOWN-REGULATED PROTEIN8)) with neddylation activity [198]. The ubiquitin ligase activity of the SCF complex requires the neddylation states of CUL1 [199]; therefore, lack of NEDD8-activating enzyme subunit AUXIN RESISTANT 1 (AXR1) results in a reduction of response to several phytohormones, including auxin [200]. It was expected any alteration in the CUL1 proteins causes the defective phenotype, such as the one observed in *izr10* and auxin hyposensitivity.

The mutation in *izr10* identified in *PpCUL1* found in the homology domain (CH) in regions conserved between *AtCULLIN1* and all three *PpCULLIN1* homoeologs (Figure 2.13B). The complete alignment s included in Appendix 2.

The cullin family members have three Cullin repeat (CR) domains that anchor similar adaptor proteins [201]. The cullin homology domain (CH) at the c-terminal is critical for binding the RING-finger protein, which plays a role as a docking site for ubiquitin-conjugating enzymes (E2s), and the neddylation site that activates the E3 ligase by promoting substrate polyubiquitination [201]. The E2 delivers the bound ubiquitin protein to the target protein, using an E3 ligase as the recognition element in the ubiquitin 26S proteasome pathway [202].

I didn't characterize any mutation related to the cellulose synthase gene family and auxin signaling pathway in the *izr6* mutant. All the WGS and annotation data are available to analyze further and investigate the CBIs' effects.

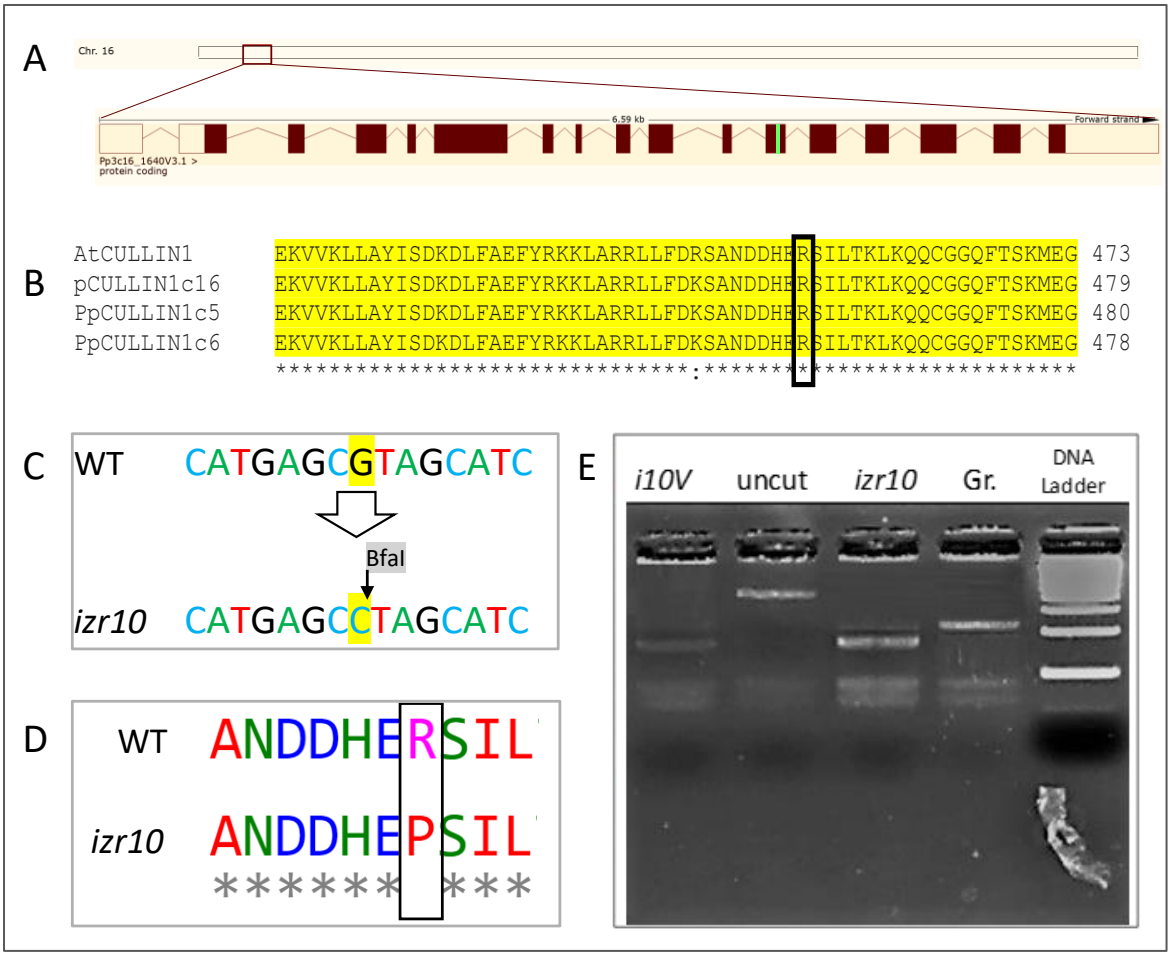


Figure 2.13: Genotyping revealed the *PpCULLIN1* mutation in *izr10* mutants. (A) Schematic representation of the *PpCUL1* gene confirmation and visualization of mapped regions for *izr* mutation. (B) Alignment of the partial CH domain of AtCULLIN1 with three *PpCULLIN1*s. The black box shows the locations of the *PpCULLIN1* mutation. (C) The DNA sequence of the PpCUL1 transcript showing the transition of G to C mutation in *izr10* mutants and the creation of BfaI restriction enzyme site. (D) Protein sequence alignment of wild-type (WT) and *izr10* mutant shows the amino acid change (Arg548Pro). Arginine is substituted for proline. Black boxes show the amino acid position at 548. The alignment was made using Clustal Omega. An (*) indicates a conserved residue, a (:) indicates conservation between amino acids with strongly similar properties, (.) indicates semi-conservative replacement of amino acid, and a ( ) indicates a non-conservation replacement. The complete alignment is included in Appendix 2. (E) CUL1 DNA fragments were digested with BfaI to confirm the presence of G to C transition in the CUL1 of *izr10* explants. i10V sample includes pooled genomic DNA from 20 i10V explants selected on 20µM indaziflam. The BfaI restriction site was present in all *izr10* and i10V samples and not present in wild-type DNA. The DNA ladder used is a 100bp DNA Ladder.

2.3.6 Phenotypic characterization of *izr* mutants

2.3.6.1 *izr10* mutants show slower, less vigorous growth

The distinct plant architecture phenotype of the *izr10* mutants, including less aerial growth and smaller gametophores, was observed from sub-culturing and growth assays. First, plants were allowed to grow over eight weeks on the BCD agar medium (without adding indaziflam). The architecture of *izr10* mutants was then compared to wild-type. The *izr10* mutants grew more slowly and to a lower density and less consistency (Figure 2.14). The phenotype is also comparable by parallel growing the mutants and wild-type in a growth medium supplemented with 2,4-D and kinetin for four weeks (Figure 2.9 and 2.11).

After several weeks, *izr10* mutants still displayed slower growth accompanied by a reduced plant bulk than the wild-type plant (Figure 2.14). While the wild-type plants develop into thick pin cushion-like structures, *izr10* gametophytes remain smaller and less developed. This is reflected in the length of the gametophores shorter in *izr10*, while the leaves are smaller (Figure 2.14). *izr10* also doesn't form sporophyte, and subsequently, no spore is produced (Figure 2.22D).

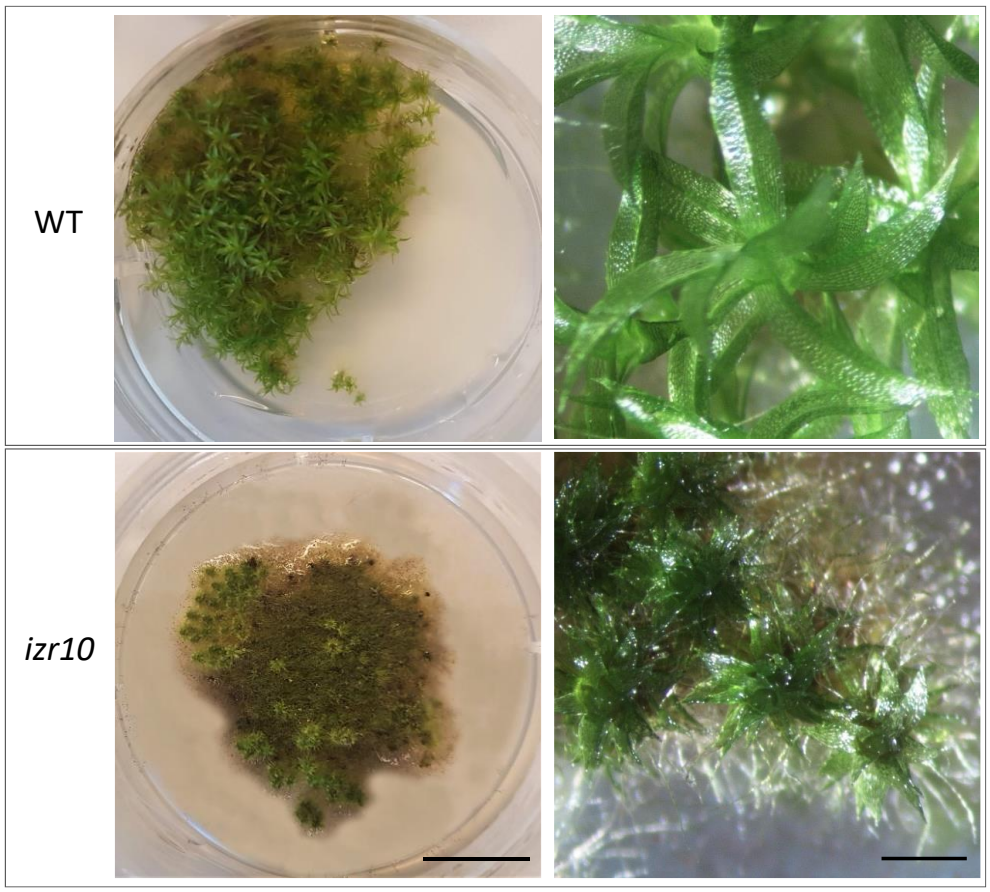


Figure 2.14: *izr10* plants show distinct and reduced growth on well-provisioned. Plant structure of *izr10* after eight weeks growth on BCD medium. This is dramatically different from the wild-type plant structure, which has denser foliage coverage and more vigorous growth. The decrease in individual gametophores can partly explain the reduction in plant size and foliage coverage (Scale bars: 0.5 cm and 0.1 mm).

2.3.6.2 *izr10* protonemal cells are shorter than wild-type

After gametophore development in *izr10* mutants, which appears to have a smaller gametophore size, including reduced leaves, further analysis was carried out on the cellular level. Further analysis of the cell morphology was analyzed by measuring *izr10* and wild-type chloronema cell lengths, assimilatory filamentous protonemal cells (identified as chloroplast-rich protonemal cells with perpendicular cross-walls). This revealed that the *izr10* mutant had cells significantly (2-tailed t-test P-value<0.005) shorter (mean=100.1µm, SD=14.1µm, n=109) than wild-type (mean=92.1µm, SD=9.8µm, n=105) (Figure 2.15). This analysis showed that the slower-growing, less dense colonies were due to shorter cells. Given that there is a mutation in the auxin signaling gene, the observed phenotype

represents an auxin-resistant phenotype, while the plant is insensitive to indaziflam. Plant hormones' interplay in regulating proper growth and development interacts with regulating the cell cycle, division, cell expansion, and gene alteration, affecting this process by failing to execute auxin-mediated cell elongation [203].

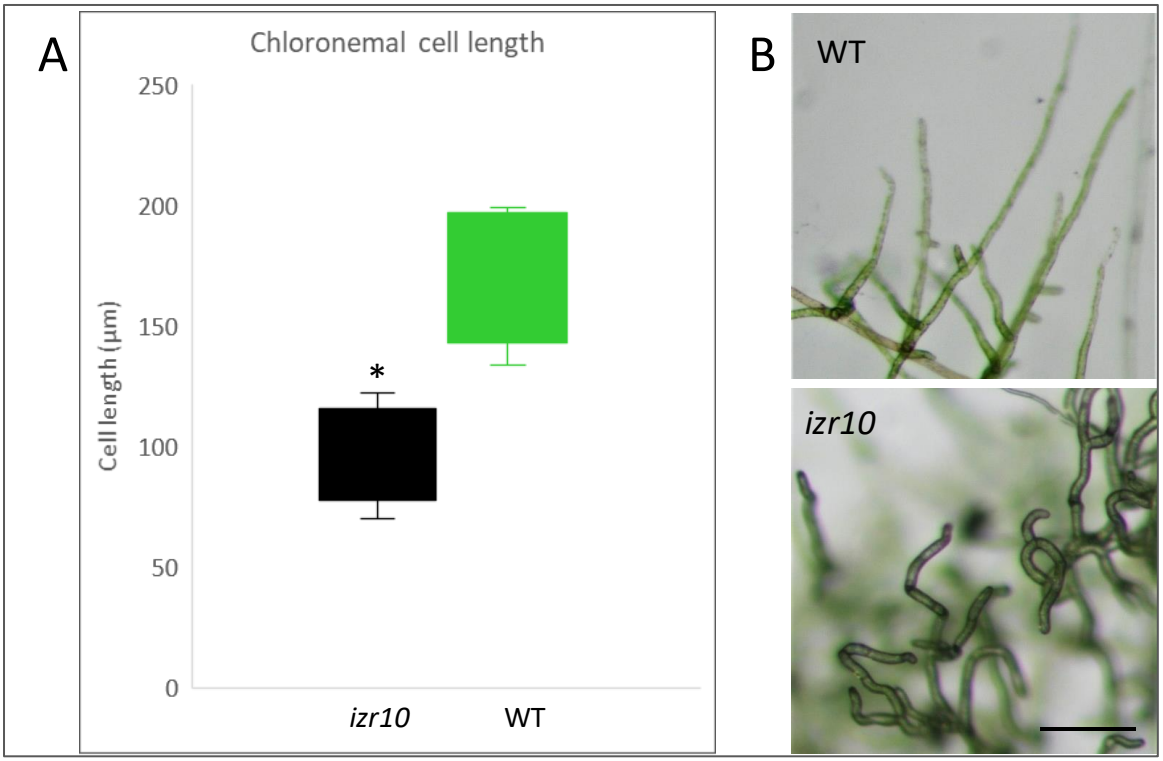


Figure 2.15: *izr10* chloronemal cells are significantly shorter than wild-type. (A) Distribution of chloronemal cell lengths in wild-type and *izr10*. *izr10* has significantly (Error bars represent s.e.m. *p<0.005 (t-test), n=30) shorter cells than wild-type. (B) Microscopic enlargement of wild-type (WT) (top) and *izr10* (bottom) chloronemal grown on BCDAT medium (scale bar 200 µm).

2.3.6.3 *izr10* plants display phenotypic similarities to auxin-related mutants

The *izr10* mutant line was compared to a documented *iaa2* mutant line [185]. The highly auxin-resistant IAA2 degran mutant is *iaa2-P328S*. The *iaa2-P328S* auxin-resistant mutant responds to exogenous auxin by developing short caulonemal filaments, and differentiation has happened at the primary chloronemal stage [185]. Like *iaa2-P328S* plants, *izr10* explants have shorter chloronemal filaments than wild-type explants (Figure 2.15). However, *izr10* produces a small and undeveloped gametophore than an *iaa2-P328S* mutant. When grown with 10µM 2,4-D, *izr10* had longer filaments compared to wild-type

explants. The *izr10* plants also display fewer and shorter rhizoids (Figure 2.16A). This observation suggests that genes that promote rhizoid development such as *PpLRL1* and *PpLRL2* (ROOTHAIRLESS LIKE) can be affected since these genes are required to initiate and develop rhizoids from buds and gametophores [204]. Based on previous studies, *Pprsl1 Pprsl2* double mutants have grown a few very short basal rhizoids and no mid-stem rhizoids. In contrast, *Pprsl1 Pprsl2* double mutant gametophores were identical to wild-type despite the apparent defects in rhizoid development [204], which indicates that for the development of both basal and mid-stem rhizoids, *PpRSL1* and *PpRSL2 (*ROOT HAIR DEFECTIVE-SIX LIKE) are required but not for the development of other cell types in the gametophore [204]. The auxin-regulatory network was present in early land plants and the last common ancestor of the mosses. *LRL* and *RSL* genes are part of this ancient gene-regulatory network that diverged since *P. patens* and *A. thaliana* shared a common ancestor. The auxin-regulatory network controls the development of tip-growing filamentous cells, rhizoids, caulonema, and root hairs, which have rooting functions at the interface between land plants and the soil (Figure 2.16B, C)[204].

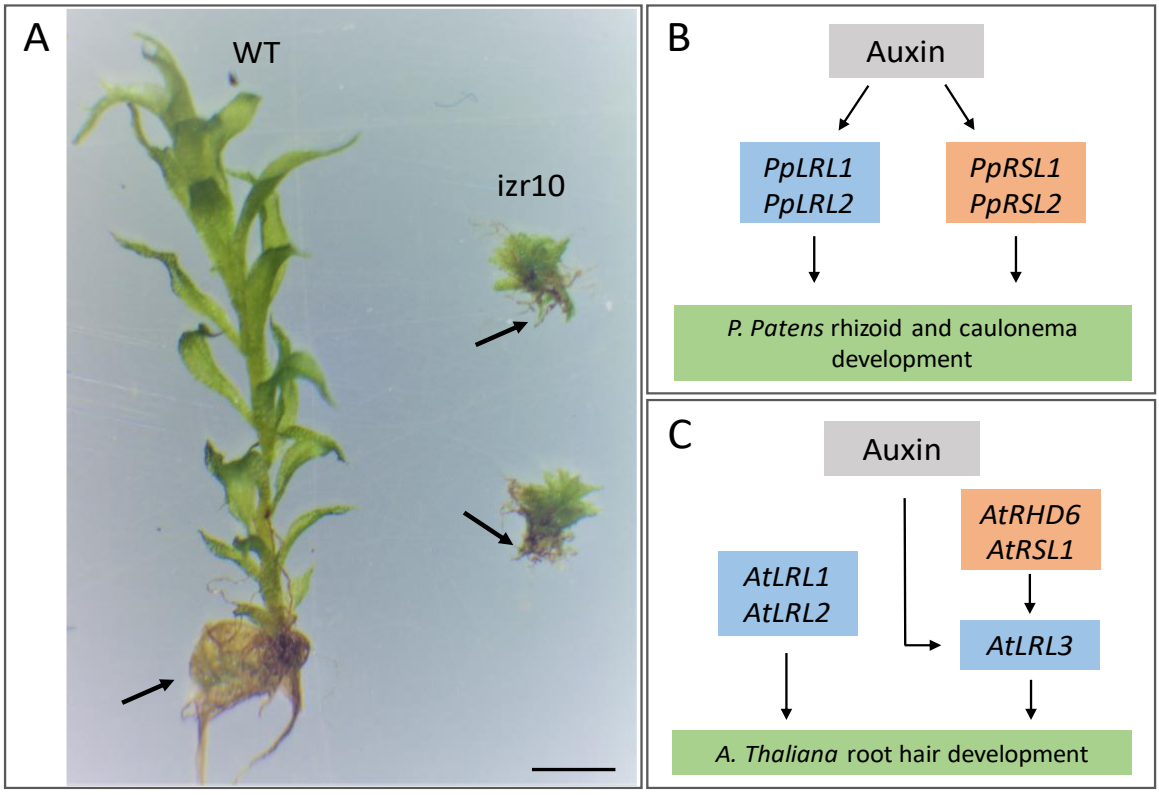


Figure 2.16: Rhizoid development in association with *RSL* genes *izr* mutant. (A) No rhizoids developed at the base of the *izr10* mutant gametophore than the wild-type (WT), shown with arrows. (Scale bars: 500µm). (B) Models indicate the transcriptional regulation between auxin, *LRL*, and class I *RSL* genes in rhizoid and caulonema development in *P. patens* compared to (B) root hair development in *A. thaliana* (Adapted from [204]).

2.3.7 Gene disruption alters auxin-regulated genes

To further characterize the effects of *PpCUL1* and *PpIAA1B* mutations and auxin on the development of gametophore and rhizoid in *izr10* mutants, achieve some insight into possible changes in gene expression of the auxin signaling transcripts, and study the influence of indaziflam on auxin signaling, I performed gene expression analysis. The genes involved in the auxin pathway were compared between wild-type, *izr6,* and *izr10* mutants with and without IAA treatment by qRT-PCR. The selected genes are *PpIAA1A, PpIAA1B, PpIAA2, PpAUX1, PpARFa6, PpARFa8, PpRSL1, PpRSL2, PpRSL4, PpSOUR, PpGH3-1, PpGH3-2, and PpCUL1s.*

2.3.7.1 Auxin transport pathways, regulation, and gene expression

Auxin is a highly mobile signaling molecule, unusual among other phytohormones since it is specifically and actively transported. Polar auxin transport proved to be essential for generating local auxin maxima and gradients. Polar auxin transport is mediated by plasma membrane-based influx and efflux carriers with asymmetrical cellular localization and functional significance [205]. Genes that act to regulate auxin homeostasis are mostly hormone-responsive. The most crucial gene groups regulated by endogenous auxin are AUXIN/INDOLE-3-ACETIC ACID INDUCED (*Aux/IAA*), GRETCHEN HAGEN 3 (*GH3*) gene families and SMALL AUXIN UPREGULATED RNA (*SAUR*). These genes are considered early or primary-response genes, considering their transcription is rapidly activated by auxin [206]. Seed plants have large families of *Aux/IAA* genes (29 in *Arabidopsis* [207]). Due to gene redundancy and lack of a plant without the *Aux/IAA* family, not many loss-of-function mutants have been studied in seed plants. The genome of *P. patens* encodes three *Aux/IAA* genes: *IAA1A*, *IAA1B*, and *IAA2* [185]. Studies have recently been done in auxin response mechanisms and plant development by generating a triple mutant lacking all three *IAA* genes in *P.patens* [185]. As a result of the *aux/iaa* triple knockout, the mutant presented a phenotype similar to wild-type plants grown on high levels of auxin, without leaves and brown-pigmented rhizoids [185], similar to what I observed in the *izr10* mutant and the segregate population since they had small gametophores with ectopic rhizoids, with insensitivity to synthetic auxin (Figure 2.9 and 2.11).

The plants were grown for four weeks on BCDAT (BCD medium supplemented with ammonium tartrate) for protonemata differentiation. Wild-type plants under these conditions responded to the high concentration of exogenous auxin by arrested at the primary chloronemal stage. In contrast, the *izr10* mutant showed the auxin-resistant phenotype by developing longer caulonemal filaments (Figure 2.17). Since the *izr10* mutant displayed the same phenotype as that of auxin-treated wild-type plants, I tested its response to adding exogenous auxin to the growth medium (5-10 µM). In wild-type plants, chloronema to caulonemal filaments differentiation and gametophore development was

slower in the presence of synthetic auxin (2,4-D), whereas the *izr10* mutant produced relatively longer chloronema filaments in response to 2,4-D (Figure 2.17).

Since I know that the *Aux/IAAs* genes have a profound role in the mechanism underlying the auxin-related genes' suppression [208], I evaluated the expression of the *IAA1A*, *IAA1B*, *IAA2*, and *AUX* genes. I analyzed the wild-type and mutants auxin-responsive gene expression by 5 µM IAA treatment for five hours.

The genes, *SAUR*, and two *GH3* homologs with a conserved function in IAA-conjugate synthesis [209] were tested too. The ARF genes that have been shown to respond to auxin in *Arabidopsis* [208, 210] were also analyzed. Our analysis identified genes that were/weren't affected by auxin treatment, as represented in the following sections.

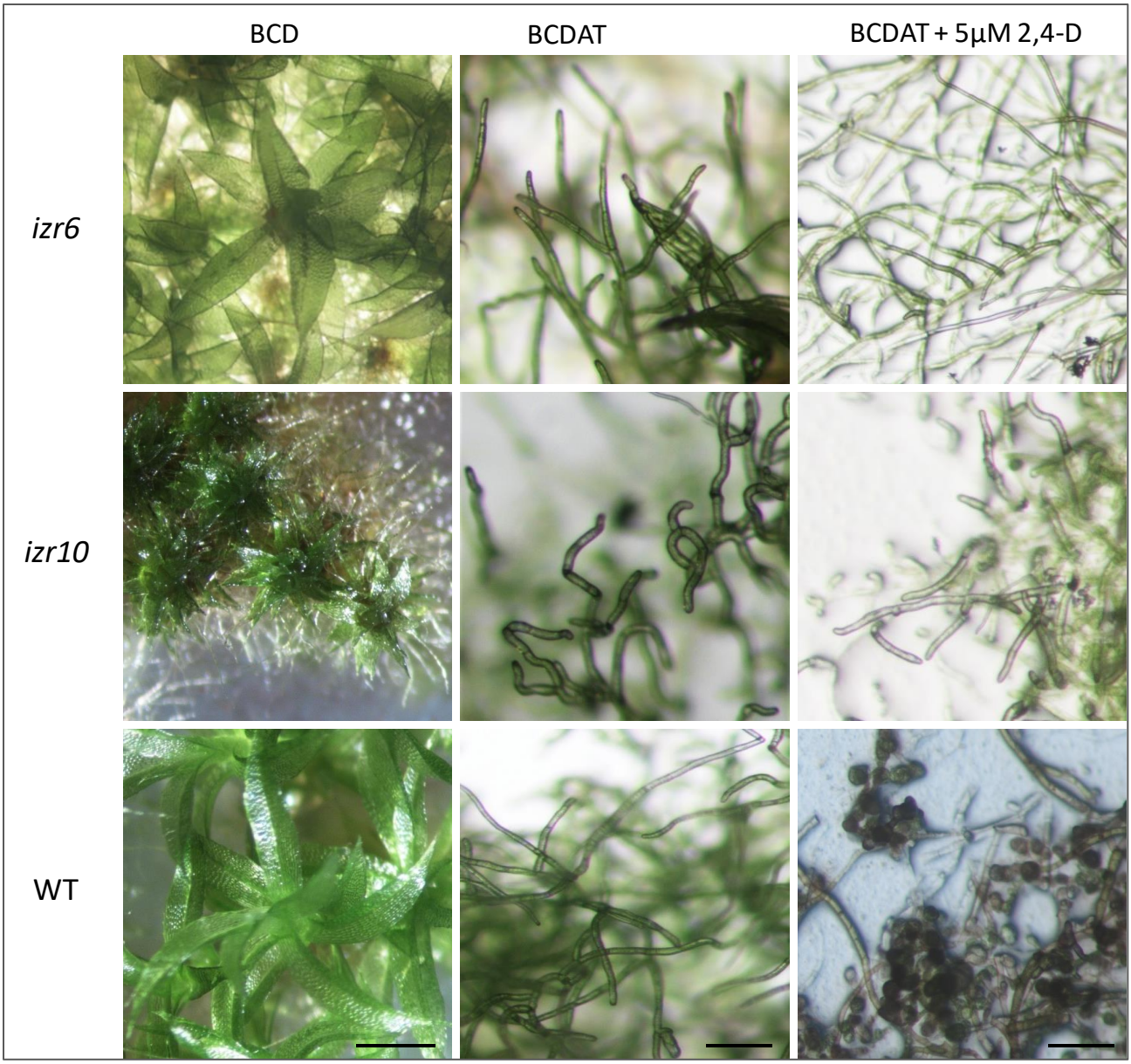


Figure 2.17: Close-up of the *izr6*, *izr10* mutants, and WT gametophores grown on BCD medium for eight weeks (left panel) (Scale bar: 5 mm). Chloronema filaments were compared after the addition of five µM 2,4-D (middle and right panel). Auxin treatment induces cell elongation, and the *izr* mutants are insensitive to synthetic auxin. (Scale bars: 500 µm).

2.3.7.1.1 *PpIAA*s and *PpARF* gene expression are disrupted in *izr10* mutant

The initial gene expression analysis revealed that *IAA* genes were downregulated in *izr6* and *izr10* (Figure 2.18). Our results show that the *izr10* and *izr6* mutants are both entirely insensitive to auxin. Strikingly *IAA* gene expressions only increased upon auxin treatment in *izr6* under our experimental conditions and statistical threshold (Figure 2.18) and displayed

a fold change further validated this finding (statistical threshold of padj <0.05). These results suggest that any mutation that caused the auxin-insensitivity in *izr6* mutant changes in gene expression must be related to the *Aux/IAAs*, emphasizing the role of these proteins in auxin signaling.

A study by Lavy et al. in 2016 showed that *ARF* genes are upregulated in the *aux/iaa* mutant compared to wild-type, which indicates the presence of an ARF-dependent feedback loop [185]. The ARF proteins can act both as activators or repressors of transcription [185]. We know that repressing ARFs requires the Aux/IAAs for their function, but it is unknown if targeting downregulated genes is caused by repressing ARFs. Auxin stimulates the degradation of Aux/IAA proteins, which repress auxin-responsive genes (ARGs) expression through interaction with activator ARFs. The network contains two negative feedback, with both Aux/IAAs and ARFs responding to auxin [211]. *P. patens* showed, higher levels of putative repressing ARFs resulted in decreased auxin response [211].

I selected *PpARFa6* and *PpARFa8* to assess the expression level in both *izr* mutants. The *izr10* mutant shows downregulation of *ARFa6*; it can be due to the negative feedback loops associated with *IAA* genes and caused the constitutive auxin phenotype of the *izr10* mutant (Figure 2.18). This result demonstrates that repressing ARFs act as a repressor in plants, and this activity requires the *IAA* genes.

The *IAA1A* and *IAA1B* genes in *izr* mutants had dramatically lower expression levels than the wild-type, indicating that activating and repressing ARFs can affect, either directly or indirectly, the same target genes. Our results specify that auxin cannot stimulate ARFa6 and ARFa8 responses in the *izr6* mutant after the treatment. But, in *izr10*, there were 2 and 3 fold changes in the expression level of *ARFa8* and *ARFa6* respectively, it is possible that auxin can promote protonemal development through other mechanisms or extended treatment, which needs further studies. Although these genes' expression in both untreated and auxin-treated *izr10* plants was reduced, the fold change was either similar or higher than the wild-type (Figure 2.18). In comparison, phenotypic analysis of the *izr10* revealed developmental defects, including shorter filaments and smaller leafy gametophores consistent with reduced auxin response (Figure 2.18). These findings further

support our hypothesis that hyposensitivity to auxin can relate to a gene mutation in the auxin signaling pathway. However, they show an unexpected trend in two different *izr* mutants.

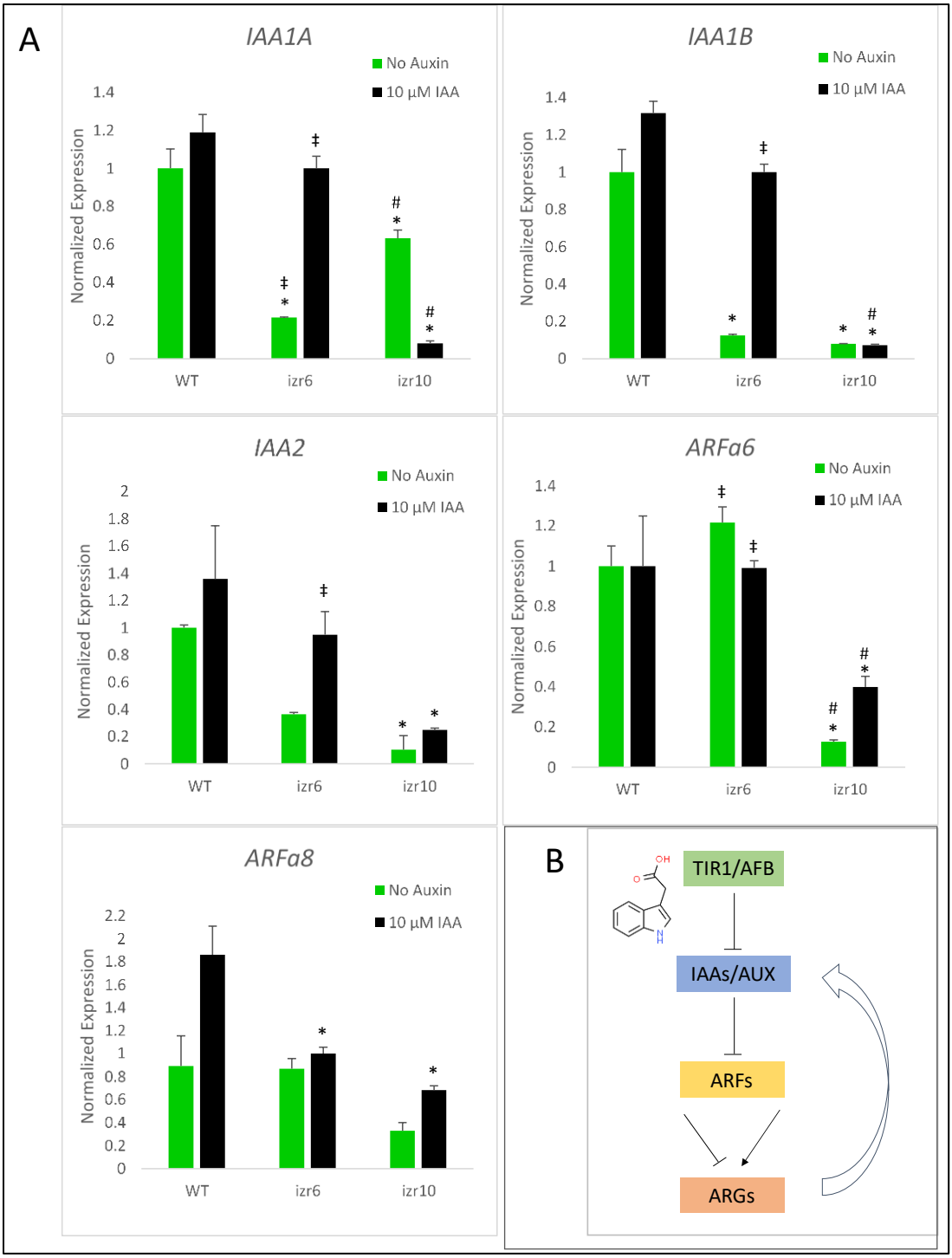


Figure 2.18: qRT-PCR showing the expression levels of auxin-responsive genes in wild-type (WT), *izr6*, and *izr10* mock- or 10 mM IAA-treated for five hours. (A) *PpIAA1A*, *PpIAA1B*, *PpIAA2*, *PpARFa6*, and *PpARFa8* expression level in response to auxin treatment. Error bars represent S.E.M. */#/‡=P<0.05 (t-test), n=2. *=t-test comparing the lines to WT, #= t-test comparing the lines to *izr6.* ‡=t-test comparing the line to *izr10.* WT, *izr6*, and *izr10* grown for one month on BCDAT without auxin. (B) Scheme representing the auxin-signaling pathway in plants.

2.3.7.1.2 *IAA* mutation can alter the *GH3* family expression

Two *GH3*-like homeologs, *PpGH3-1,* and *PpGH3-2* were expressed in *P. patens* gametophytic tissues [212]. We believe both genes have partially overlapping functions of the two GH3 proteins [213]. The previous study reveals the role of GH3 proteins in auxin homeostasis and shows an additional mechanism for the detoxification of auxin in moss [209]. There is also evidence of PpGH3 proteins' role in developing *P. patens* from studying auxin sensitivity assays with Δ*PpGH3-1* and Δ*PpGH3-2* lines [209].

I performed a gene expression study for *PpGH3-1* and *PpGH3-2.* The *PpGH3-1* expression level in the *izr* mutants was lower than the wild-type. After the treatment, the wild-type plant has a lower expression of the *PpGH3-1*, which was opposite in the *izr6* line, and has a three-fold increase (Figure 2.19). By contrast, auxin treatment slightly (but not significantly) decreased *PpGH3-2* expression level in *izr* mutants compared to wild-type (Figure 2.19). I believe this effect was GH3-independent because there were no significant differences in *PpGH3-2* expression between wild-type and *izr* mutants before the treatment. Therefore, *P. patens* GH3 proteins possibly have an intermediate position, suggesting that the GH3 function goes with further developing organs and auxin function.

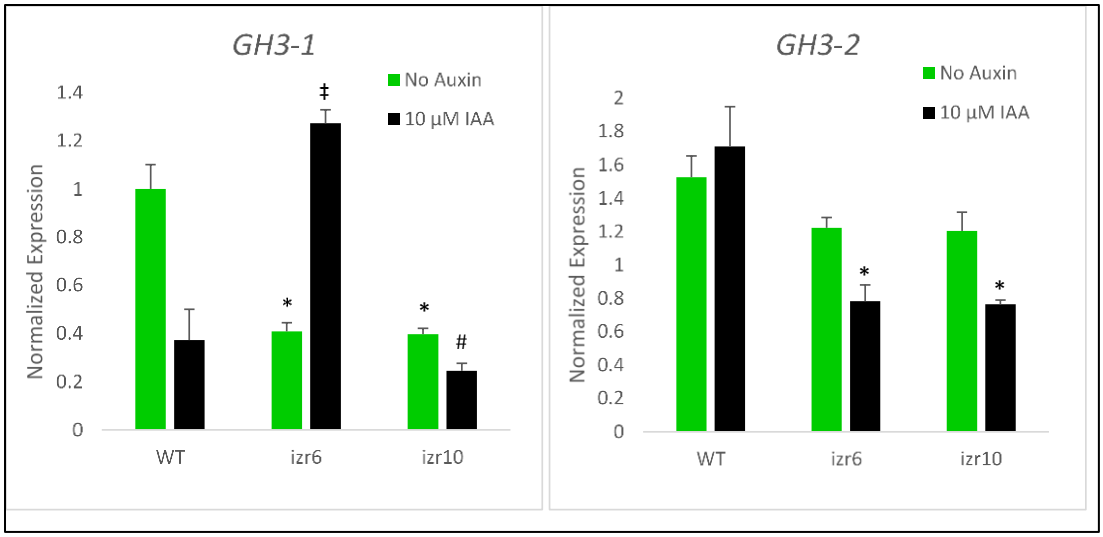


Figure 2.19: qRT-PCR showing the expression levels of auxin-responsive genes *GH3* in wild-type (WT), *izr6*, and *izr10* mock- or ten mM IAA-treated for five hours. *PpGH3-1* and *PpGH3-2* expression level in response to auxin treatment. Error bars represent S.E.M. */#/‡=P<0.05 (t-test), n=2. *=t-test comparing the lines to WT, #= t-test comparing the lines to *izr6.* ‡=t-test comparing the line to *izr10.* WT, *izr6*, and *izr10* grown for one month on BCDAT without auxin.

2.3.7.1.3 *SAUR* downregulation in *izr* mutants

The *SAUR* genes have been important for the output of the auxin response from the beginning of land plant evolution. *Physcomitrium patens* contains 18 *SAUR* genes [52]. *SAUR* overexpression affects auxin levels, auxin transport, and the expression of auxin pathway genes [214]. *The SAUR* genes can be regulated positively or negatively by auxin [215]. There is a group of *SAUR* genes, which is not responsive to auxin [215]. *AtSAUR8* in *Arabidopsis* can induce cell elongation when overexpressed and promote growth in response to stimuli other than auxin [214, 215], and it is non-responsive to auxin but is upregulated by cytokinin [215].

I analyzed the wild-type and *izr* mutant lines *PpSAUR* gene and found a significant change in gene expression between *izr* mutants and wild-type (Figure 2.20). The *PpSAUR* gene was downregulated in *izr6* and *izr10*. Also, the *PpSAUR* gene expression was slightly higher in wild-type and *izr* mutants following IAA treatment (Figure 2.20)

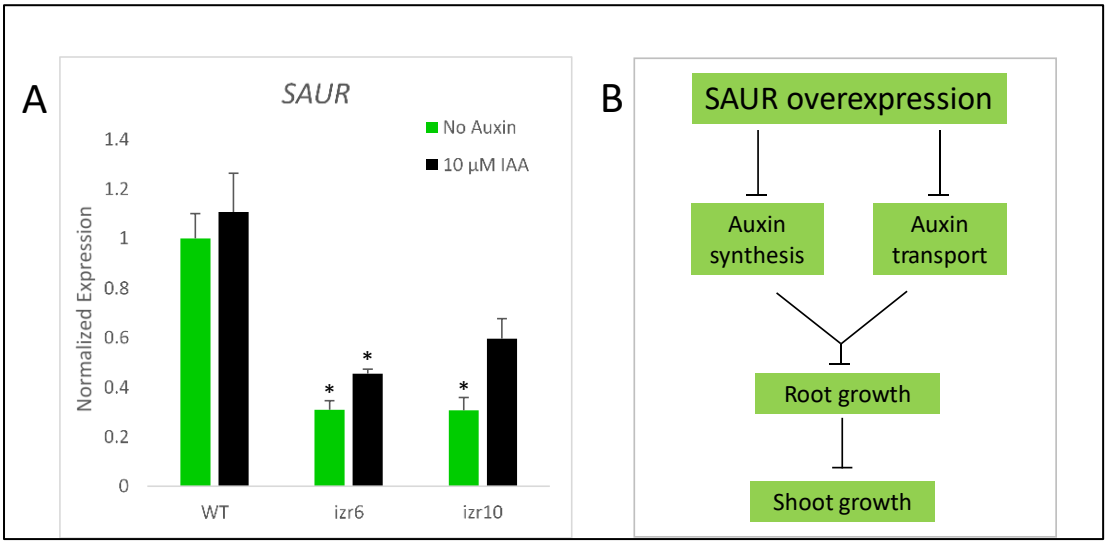


Figure 2.20: qRT-PCR showing the expression levels of auxin-responsive gene *SAUR* in wild-type (WT), *izr6*, and *izr10* mock- or ten mM IAA-treated for five hours. (A) *PpSAUR* expression level in response to auxin treatment. Error bars represent S.E.M. *=P<0.05 (t-test), n=2. *=t-test comparing the lines to WT. WT, *izr6*, and *izr10* grown for one month on BCDAT without auxin. (B) A model to explain the constitutive upregulation of *SAUR* gene function in negative regulation of auxin level in plants, resulting in reduced plant growth via reduced auxin synthesis and transport [216].

2.3.7.1.4 *CULLIN1* mutation does not affect the transcriptional expression of the genes

CULLIN1 (CUL1) is integral to various key developmental processes through its ubiquitin 26S proteasome pathway [202]. An allele of *CUL1* has been identified from a genetic screen for mutants resistant to a synthetic compound that induces the expression of auxin response genes called sirtinol. This finding indicates that *CUL1* is required for diverse signaling pathways throughout development [196].

The qRT-PCR analysis was carried out to confirm the role of *CUL1* as an auxin regulator at the molecular level in *P. patens*. I used three *PpCUL1* genes, including the one which contains the mutation in the *izr10* mutant line. As shown in Figure 2.21A, the gene expression of *PpCUL1*c5 of (*CULLIN1* located on chromosome 5) in *izr* mutants was comparable to wild-type, while *PpCUL1c6* (*CULLIN1* located on chromosome 6) and *PpCUL1c16* (*CULLIN1* located on chromosome 16) were abolished in the *izr10* mutant (Figure 2.21A).

I found that the auxin-induced expression of *PpCUL1c5* and *PpCUL1c6* genes was significantly upregulated in the *izr10* mutant compared with that in the wild-type (Figure 2.21B). Interestingly, the mutation in *PpCUL1c16* did not block the auxin-induced expression of this gene, indicating that the regulation mechanism of *PpCUL1c16* might be the same for all the *CUL1* genes in *Physcomitrium* under auxin treatment. In the *izr6* transgenic line, the auxin-induced expression recovered to a similar level to the wild-type (Figure 2.21B).

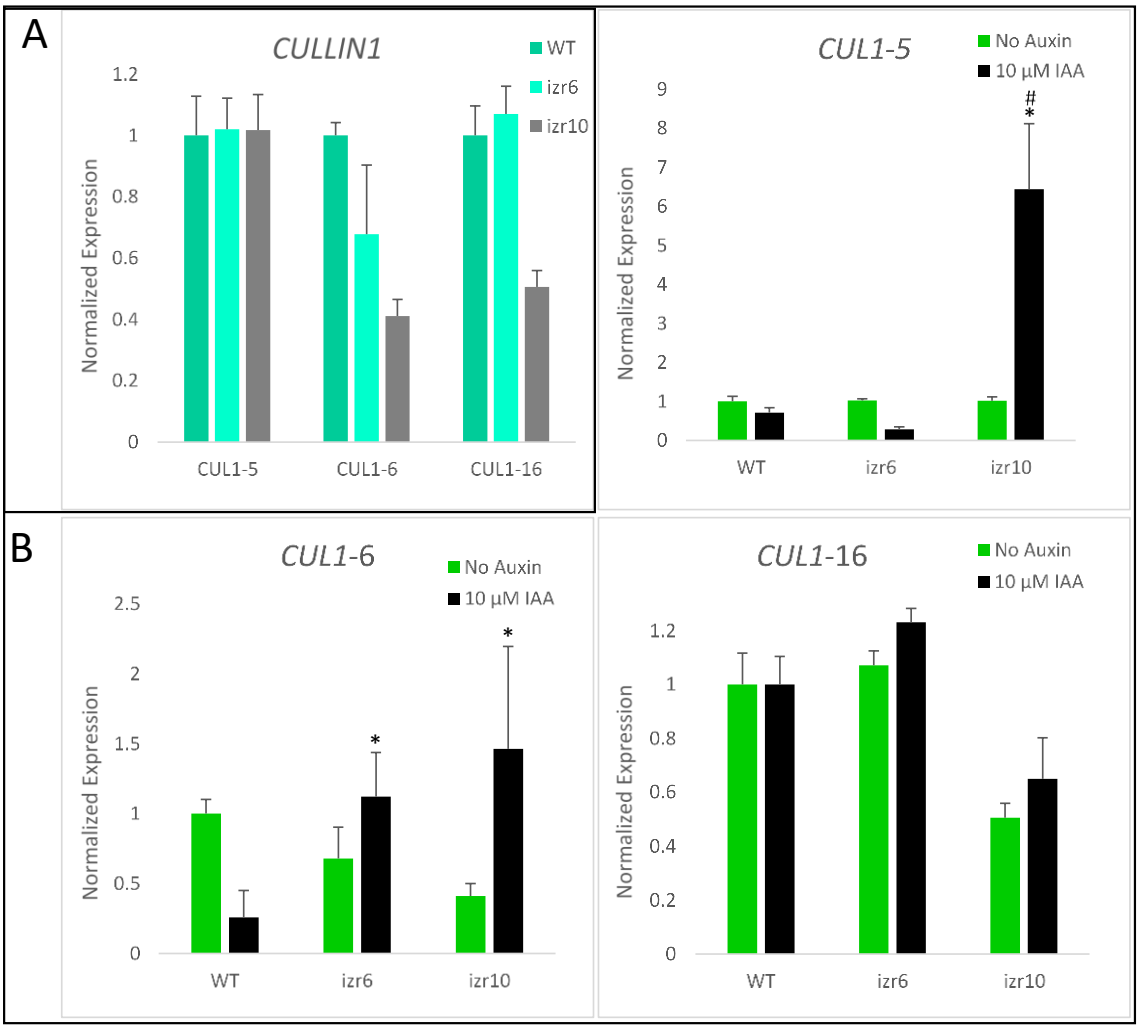


Figure 2.21: qRT-PCR showing the expression levels of *CULLIN1* genes. (A) *PpCULLIN1*s expression level in wild-type (WT) compared to *izr6* and *izr10*. (B) *PpCULLIN1*s expression level response in WT, *izr6*, and *izr10* mock- or ten mM IAA-treated for five hours. Error bars represent S.E.M. */#/‡=P<0.05 (t-test), n=2. *=t-test comparing the lines to WT, #= t-test comparing the lines to *izr6*. ‡=t-test comparing the line to *izr10*. WT, *izr6*, and *izr10* grown for one month on BCDAT without auxin.

2.3.7.1.5 RSL transcriptional expression was affected in *izr* mutant

To determine whether the expression of RSL genes was affected to inhibit rhizoid development in *izr* mutants, I tested the expression level of *PpRSL1* and *PpRSL2* in both mutants and wild-type.

There are two classes of RSL genes present in *Physcomitrium*, RSL class I (*PpRSL1* and *PpRSL2*) and RSL class II (*PpRSL3, PpRSL4, PpRSL5, PpRSL6*, and *PpRSL7*) [217]. Caulonema and rhizoids are tip-growing filamentous cells, which their differentiation is regulated by

class 1 *RSL* genes (*PpRSL1* and *PpRSL2*) and auxin [218]. The qRT-PCR analysis demonstrated that *PpRSL1* and *PpRSL2* expression levels in *izr10* were lower than wild-type and *izr6,* and the mutant line was distinguishable phenotypically from wild-type (Figure 2.22). By contrast, steady-state PpRSL1 and PpRSL2 transcript levels were elevated in the *izr6* line (Figure 2.22A, and B). We know from previous studies that *PpRSL1* and *PpRSL2* are required to develop both basal and mid-stem rhizoids [68] and are expressed in the cells that give rise to rhizoids.

In our study, *izr10* explants were distinguishable from wild-type since it has only a very few short basal rhizoids and shows the obvious defects in rhizoid development. Also, the mutant's caulonema doesn't differentiate from chloronema (Figure 2.22C). This could indicate a possible disruption in *PpRSL1* and *PpRSL2* transcription, which affected these genes' expression in the *izr10* line.

Auxin positively regulates rhizoid development [96], and because the expression of *PpRSL1* and *PpRSL2* is sufficient to promote rhizoid differentiation, I believe that auxin controlled rhizoid development by regulating the expression of *PpRSL1* and *PpRSL2* [219]. To test the effect of auxin-treatment on *PpRSL1* and *PpRSL2* expression, I checked the expression level after 5 hours of IAA treatment. As expected, auxin-treatment increased *PpRSL1* and *PpRSL2* steady-state mRNA levels in wild-type plants (Figure 2.22A and B). Auxin treatment increased the expression level of *PpRSL2* in both *izr* mutant lines, but for *PpRSL1,* the effect of auxin treatment was the opposite (Figure 2.22A and B). This shows that auxin-treatment increases *PpRSL1* and *PpRSL2* transcription and extends these genes' expression into the rhizoids. These data validate the regulatory effect of auxin in the expression of *PpRSL1* and *PpRSL2*. These data are consistent with previous studies in which *PpRSL1* is expressed in cells sensitive to auxin, and the cellular pattern of auxin sensitivity regulates the expression of *PpRSL1* and *PpRSL2*, which positively regulates rhizoid development [219]. Since *izr10* is not sensitive to auxin and shows obvious defects in rhizoid development, I expected to see *PpRSL* class I genes' downregulation. On the other hand, auxin controls rhizoid development by positively regulating the activity of *PpRSL1* and *PpRSL2* genes [219], which is consistent

with our results regarding the upregulation of *PpRSL2* upon auxin treatments in both *izr* mutants and wild-type.

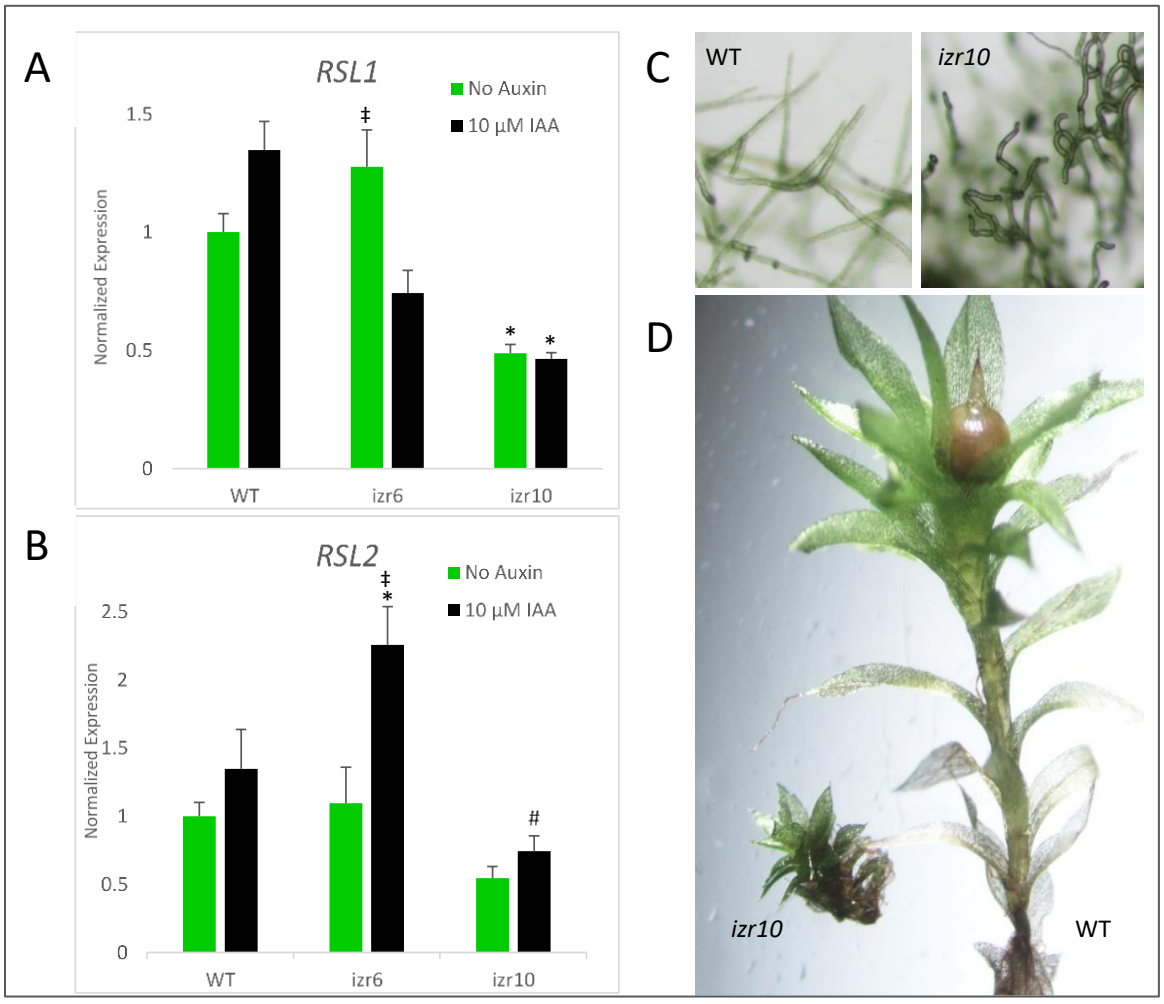


Figure 2.22: Repressing *PpRSL1* and *PpRSL2* target auxin-induced genes. (A) qRT-PCR showing the expression levels of *PpRSL1* in wild-type (WT), *izr6,* and *izr10* in the presence of 10 mM IAA-treated for five hours. (B) qRT-PCR showing the expression levels of *PpRSL2* in WT, *izr6,* and *izr10* in the presence of 10 mM IAA-treated for five hours. Error bars represent S.E.M. */#/‡=P<0.05 (t-test), n=2. *=t-test comparing the lines to WT, #= t-test comparing the lines to *izr6.* ‡=t-test comparing the line to *izr10.*WT, *izr10,* grown for two weeks on BCDAT. (C) Comparing the WT and *izr10* plants grown on BCDAT for 7 days. (D) Comparing the WT and *izr10* plants grown on soil for 12 weeks on the soil.

2.4 Discussion

2.4.1 Forward genetics as a tool in bryophyte genomics

The genome assembly of *Physcomitrium* (V3.1) [220] with chromosome-scale linkage maps makes forward genetics possible in the model bryophyte *Physcomitrium patens*. This study represents a known example of this approach, taking advantage of advances in the available genetic resources in the gene discovery of *Physcomitrium.* Also, bulk segregant analysis of separately bulked mutants and wild-type segregants can be compared and combine knowledge of SNP markers with de novo identification of likely causal mutations.

This is possible to be a widely used process as whole-genome sequencing gets cheaper and more accessible to all. However, a similar approach based on the WGS of bulked segregant pools is also possible. The reliable identification of mutants is a crucial part of any forward genetic approach, and the screening of positive phenotypes such as the faster growing *izr* mutants in the presence of indaziflam has its advantages. This suggests the analysis of other *izr* mutants should be successfully pursued in *Physcomitrium,* especially given those novel mechanisms are involved in bryophyte auxin signaling. This proof of concept should be vastly used for *Physcomitrium* for novel gene discoveries.

2.4.2 The *izr* mutants

Identifying the *izr* mutants results from detailed mapping data that identified a single region associating with the phenotype. The sequencing of potential regulatory regions in *izr10* mutant has revealed the causal mutations. One of the exciting differences between *izr6* and *izr10* is that *izr10* displays altered growth and altered leaf morphology ("wasserpest" phenotype, resembling the waterweed *Elodea*) (Figure 2.5). The mutation/s slows down the growth and increases synthetic auxin responses. While more *izr* mutants remain to be analyzed, a fascinating and unexpected story has already been started by implementing forward genetics in *Physcomitrium* and identifying such a vital and unpredictable auxin regulator.

2.4.3 Getting lucky

The mutagenesis of a different model organism, moss, and with another mutagenesis agent, UV for *izr* mutants, interestingly aided its eventual identification of the *IAA1B* and *CULLIN1* gene related to hyposensitivity to indaziflam, but just how likely was this to have happened? The previous study in Bonetta's lab resulted in identifying and characterizing a novel allele of *AtCULLIN1* in the indaziflam resistance screen.

This highlights the advantage of forward genetics and WGS technologies, which offer traditional mapping to identify novel alleles. With WGS approaches, assuming the mutation lies within DNA, coarse mapping or bulked segregant analysis may be adequate to narrow the window for mutation identification based on the segregation of novel mutations with the phenotype. And a good reference genome plays a vital role in identifying these novel mutations induced as part of the forward genetic approach.

2.4.4 Effects of *izr* mutants

Indaziflam is involved in auxin signaling, which has effects at the phenotypic and molecular levels. Growth assays confirm that *izr* mutants fail to respond to the presence of synthetic auxin 2,4-D in the medium, which in wild-type plants significantly slows growth. Both *izr6* and *izr10* show similar levels of 2,4-D non-responsiveness, but they appeared to have different phenotypes, with *izr6* having a more similar phenotype to the wild-type characterized by more gametophyte and rhizoid growth. An exciting possibility of the roles of indaziflam in the auxin signaling pathway would enable a mechanism for this to be found. For example, the recent finding of the *Arabidopsis thaliana* indaziflam resistant line (*izr1*) demonstrated reduced sensitivity to auxin mimicking herbicides, 2,4-D (Bonetta, unpublished). This lack of molecular responses to auxin mimicking herbicides showed exact mechanisms behind the short protonema cells of mutants and the inability for mutants to respond to a high concentration of synthetic auxin, 2,4-D. In anatomically simple plants such as bryophytes, molecular changes to protect the cells is crucial.

Since indaziflam as a cellulose biosynthesis inhibitor (CBI) inhibits cellulose production [136], it is predicted that a cell wall-related protein would cause indaziflam resistance as

many CBI-resistant lines to have mutations in cellulose synthases (CESAs) genes [71, 132, 158, 221]. In this screen for indaziflam resistance, two possible alleles were identified in the *IAA1B* and *CULLIN1* genes from next-generation sequencing analysis.

2.4.5 Interactions of indaziflam and auxin

The growth phenotype of *izr* mutants was also analyzed under normal conditions. Mutants appeared to grow slower than wild-type colonies as cell length is noticeably shorter in *izr10*. The consistently smaller and less dense *izr10* plants were likely due to the decreasing gametophores, which also had reduced leaves. Analysis of the plant architecture revealed that *izr* mutants grew with less uniformity and with a phenotype superficially similar to that observed in the auxin response characterized by reduced tip growth in the colony [185]. The *izr* mutant chloronemal cells were significantly shorter than wild-type, which is in strong harmony with the process of auxin signaling, which also confirms that basal land plants have conserved submergence mechanisms likely under the same hormone signaling regulation as those in higher plants.

In this study, I could find two alleles to confer the indaziflam resistance. The first one was mapped in chromosome 8, and the next-generation sequencing revealed a point mutation in the *IAA1B* gene, which caused a Gly207Phe substitution in the IAA1B protein. The Aux/IAA proteins have an essential role in auxin signaling and play roles as auxin co-receptors and transcriptional repressors. Deleting all three *Aux/IAAs* has shown the TIR-Aux/IAA pathway's effects on plant growth and development and auxin response [185]. These genes also have an extensive impact on other genes' expression levels, including those that do not respond to auxin treatment. It also emphasizes the requirement for the Aux/IAAs in auxin-responsive transcription [185]. In *P. patens*, developmental transitions and differentiation depend on auxin-dependent modulation of gene expression [185].

Our mutants analysis shows that the transition from chloronema to caulonema is regulated by auxin. When auxin response is low, such as in the *izr10* mutants, protonema cells are arrested in the chloronema stage. There is an argument that ARFs are competing with AUX/IAA for binding to the same promotors in *Arabidopsis* [222-224]. The study by Lavy et

al. in 2016 in *P. patens* supports this argument. Auxin responsive genes are targeted by ARFs and can manage the gene expression levels. Their studies hypothesize that ARFs (repressing and activating) compete for the same binding sites, form heterodimers, and induce transcription [185]. These findings can explain a wide range of gene expression levels by auxin. They also showed that *ARF* genes are upregulated in the *aux/iaa* mutant compared to wild-type, indicating an ARF-dependent feedback loop [185]. The ARF proteins can act both as activators or repressors of transcription [185]. I know that repressing ARFs requires the Aux/IAAs for their function, but it is unknown if targeting downregulated genes is caused by repressing ARFs.

The *ARF*s used in our study show expression level changes in *izr10* mutants upon auxin treatment, which supports the finding from previous studies, which means the *ARF* genes downregulated in the *izr10* mutant contains a mutation in *IAA1B* compared to wild-type. This again confirms the ARF-dependent feedback loop.

I have seen that *izr10* mutants fail to respond to exogenous auxin both at the phenotypic and molecular levels. And *IAA1B* gene expression doesn't change after auxin treatment in *izr10* mutants. This translated into a direct loss of *PpIAA1B* function and highlighted this protein's role in auxin signaling.

Auxin stimulates the degradation of Aux/IAA proteins, which repress auxin-responsive genes (*ARGs*) expression through interaction with activator ARFs [211]. *P. patens* showed higher levels of putative repressing ARFs resulting in a decreased auxin response [211]. Aux/IAAs are essential for repressing ARFs, but there is no evidence if repressing ARFs has a specific role in targeting downregulated genes. The downregulation of *ARF* in *izr10* mutant could result in negative feedback loops of ARF associated with *IAA* genes, which may contribute to auxin insensitivity.

On the other hand, GH3 proteins play a role in auxin homeostasis and contribute to auxin's detoxification in moss [209]. Since, in our mutants, there weren't any significant changes of GH3 expression in the present and absence of exogenous auxin, I suggest that the effect was GH3-independent, and GH3 proteins have an intermediate position in *P. patens.*

Another protein regulated by auxin is SAUR, which induces cell elongation and growth [214, 215]. SAUR proteins affect directly or indirectly the auxin transport since plants expressing them have higher basipetal auxin transport. Therefore *SAUR* genes induced growth [225, 226]. Our finding suggests that in *izr* mutants, the SAUR gene functionality was affected due to Aux/IAA dysfunction, leading to a lack of normal growth (Figure 2.22D).

2.4.6 Loss of CULLIN1 function and auxin sensitivity

As I will discuss in more detail below, loss of CULLIN1 function decreased sensitivity to the synthetic auxin, 2,4-D [180]. In *Arabidopsis,* the ubiquitin/26S proteasome pathway regulates the degradation of critical proteins in response to environmental or biological signals. There are three enzymes involve in the ubiquitination of proteins; ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and ubiquitin-ligating enzyme (E3) [202]. Skp, Cullin, F-box containing complex (or SCF complex), is a major class of E3s, which have been associated with hormone response and environmental factors in plants [227, 228]. The SCF is made of four protein subunits: CULLIN 1 (CUL1), RING-box 1 (RBX1), S-Phase Kinase Associated Protein 1 (SKP1), and an F-box protein. Transport inhibitor response 1 (TIR1) is an F-box protein incorporated into the SCF complex through Skp1-like (ASK1) [229]. Auxin can bind to TIR1, function as an auxin receptor, and target AUX/IAA for degradation by 26S proteasome transcriptional repressors (Figure 2.23B) [230, 231]. Identification of the *CULLIN1* alleles *axr6-1* and *axr6-2* based on their auxin insensitivity created a link between CULLIN1, AUX/IAA degradation, and auxin insensitivity [180, 197, 232, 233]. This indicates that *CUL1* is required for diverse signaling pathways throughout development [196]. CUL1 has a similar function during the formation of the SCF complex since it interacts with Skp and F-box proteins [234], and also, CUL1 is part of one of the SCF subgroups of ubiquitin ligase E3 [235] and act as an auxin receptor [236]. The CUL1 protein StCUL1 in potato *Solanum tuberosum* is localized in the cell cytoplasm [234], consistent with the subcellular localization of the F-box protein Skp. The Skp protein in *Physcomitrium patens* is also localized in the cytoplasm and cell nucleus [237], consistent with the subcellular localization predicted for StCUL1 [234]. An alteration in the CUL1 protein, based

on the CH domain's critical role, where the *izr10* mutation is located, as the recognition element in the ubiquitin 26S proteasome pathway [201, 202] will explain the insensitivity to auxin.

A recent study by Bonetta et al. (unpublished) confirmed the *izr1* allele was responsible for the weak indaziflam resistance similar to the *cul1-6* mutant in *A. thaliana* had strong 2,4-D resistance. The finding in this study also confirmed that the mutation in *PpCUL1c16* does not interfere with auxin-induced gene expression (Figure 2.21) and the regulation mechanism of *PpCUL1c16 in Physcomitrium* is the same for all *CUL1* genes under auxin treatment. However, it confirmed that amino acid substitution in CUL1 could alter indaziflam and auxin sensitivity (Figure 2.7, 2.8, 2.16). Similarly, the *ixr1-1* allele has been reported to be resistant to the CBI isoxaben [71] and indaziflam (Bonetta, unpublished). This indicates that the herbicidal mode of action of indaziflam is not limited to inhibiting cellulose synthesis; it also affects various pathways in plants with multiple protein targets. Identifying indaziflam resistant mutants can help us better understand the mode of action associated with this potent herbicide. Positional cloning is an approach to identify resistance alleles, and it is challenging to select a mapping population homozygous for the associated allele in diploid species. However, the advantage of moss as a haploid model organism makes it easier to identify the allele associated with herbicide resistance.

However, this would require further analysis as the exact mechanism of CUL1 proteins in *Physcomitrium patens* is unknown, especially when dealing with low expression, is not sufficient to make firm conclusions.

2.4.7 Complex regulation process of rhizoid development in the moss

Rhizoid and root hairs develop in gametophytes and sporophytes, respectively, in contact with soil [238] and are controlled by the interaction between several genes, including class I RSL genes, which belong to the basic helix-loop-helix transcription factor family [239]. Class I RSL (*PpRSL1* and *PpRSL2*) induce the expression of downstream transcription factors of the basic helix-loop-helix family, eventually triggering differentiation and polarized tip growth root hairs [240-242]. *RSL1* loss-of-function mutants do not develop root hairs [239]. Root

hair development is a dynamic process with complex regulation of the endogenous, environment, and phytohormone combination [243]. I believe there are particular molecular mechanisms that auxin activates root hair tip growth. Mutants, such as *aux/iaa1* (*axr5*) [244], *aux/iaa7* (*axr2*) [245, 246], *aux/iaa14* (*slr*) [247] with reduced growth in root hair cells and cell expansion have shown an impact on auxin biosynthesis, conjugation, transport, and signaling highlight the multiple and complex levels of auxin action on cell elongation (Figure 2.23A). On the other hand, gain-of-function mutants of *AUX/IAAs* [243, 248] showed root hair defect, which underlining that ARF activation is required for root hair tip growth (Figure 2.23C) [238]. The study by Velasques et al. in 2016 also suggests that to trigger cell expansion in the root hair cells, auxin has to be sensed in situ. However, the non-cell-autonomous contribution of auxin signaling can affect root hair development as well [238]. The class II *RSL* genes are also regulated by auxin in *Arabidopsis* [241, 249]. In *Physcomitrium*, auxin promotes the expression of *PpRSL1* and *PpRSL2* and not orthologs of *RSL4* [250].

In our study, *izr10* transcriptional expression of PpRSL1 and PpRSL2 significantly affected, as expected, due to the undifferentiated chloronema (Figure 2.22C) and no/short rhizoid development (Figure 2.22D). This data can be validated by analyzing the *izr6* mutant line, which shows similar *PpRSL1* and *PpRSL2* transcript levels compared to wild-type (Figure 2.22A and B).

This indicates a disruption in class I RSLs transcription in the *izr10* line, which also can happen because of mutations in upstream genes. However, auxin treatment increases the transcriptional expression levels of RSL1 and RSL2 in *Physcomitrium* [218, 250]. I did not detect the same changes in the transcriptional expression of RSL1 and RSL2 in the *izr10* mutant line, emphasizing the complex regulation process of rhizoid development, which requires further study.

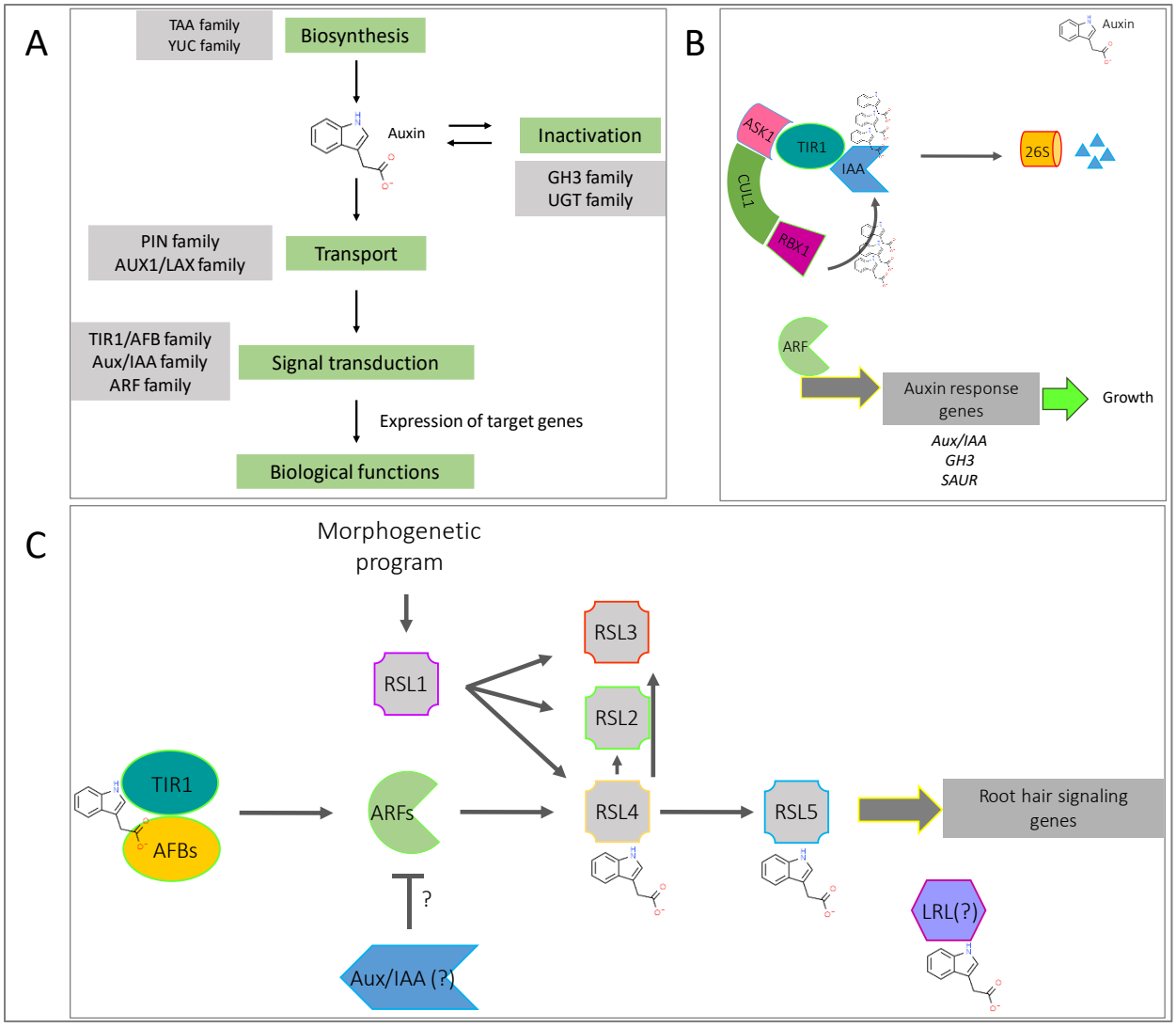


Figure 2.23: Schematic model of auxin biosynthesis, inactivation, transport, and signal transduction in plants. (A) Protein families involved in biosynthesis, transport, and signal transduction of auxin (Adapted from [251]). (B) The ARF-Aux/IAA auxin response pathway. In the presence of auxin, Aux/IAAs are ubiquitinated and degraded in a proteasome-dependent manner. Aux/IAAs mutants can no longer be ubiquitinated and affect stable repression of ARF function (Adapted from [252, 253]). (C) The main target of auxin in root hair cells. Auxin activates a transcriptional response in the nucleus viaTIR/AFB-ARFs. Auxin can directly regulate the expression of several transcription factors, such as RSL3, RSL4, RSL5, and LRL. ARFs and AUX/IAAs can be components of the auxin response in root hair cells (Adapted from [238]).

2.5 Conclusion

Forward genetic screening in plants is an effective method for identifying novel CBI-resistant mutant alleles [132] and indaziflam resistant lines in moss. Previous studies have shown that the addition of synthetic (2, 4-D) and natural (IAA) auxins to cell cultures can inhibit programmed cell death by CBIs. Also, the addition of auxin efflux inhibitors (NPA) can increase the accumulation of auxin and protect cells from CBI-induced cell death, which means auxin can stop CBI-induced cell death by making changes in the cell wall composition and compensate for reduced cellulose synthesis. However, the auxin-mediated protection against CBI-induced cell death does not depend on cell wall reinforcement. There is a possibility that auxin stimulates rapid changes in the plasma membrane to compensate for cell wall defects [254].

Moreover, studies with IAA in *A. thaliana* [255, 256] and 2,4-D in potato [112] demonstrated a negative association between CBI toxicity and auxin content within plant tissues. Recent transcriptional profiling studies in response to CBIs showed CBIs upregulated few auxin-responsive genes, but they don't act on auxin receptors [257]. Another group also showed that the root growth response in CBI mutants in *A. thaliana* (*ixr1-1*, *KOR1*, *txr1-1*) didn't change using synthetic auxins [258]. This suggests that mechanisms that cause resistance in CBI resistant mutants do not influence auxin uptake associated with root growth, so there is no direct CBI – auxin interaction [258].

Our results presented here make use of *P. patens* mutants, which provide an important resource for describing and understanding CBI resistance pathways and possible mechanisms of action. In this case, *P. patens*' auxin pathway modifications provide a link between auxin signal transduction and cellulose synthesis inhibition. Testing CBI *P. patens* mutants, as presented here, can help identify common mechanisms across herbicides and similarities in the mutants' resistance mechanisms across plant species. A better understanding of these mechanisms will help us understand how plant cells respond to toxins and key information about cellulose synthesis. Indeed, the similarity of responses expressed in this study, particularly by *cul1,* with a very recent work by Bonetta et al. (unpublished) link between indaziflam resistance and auxin transport pathway. Our finding

suggests that the indaziflam mode of action is linked to phytohormone signaling and possibly an auxin-indaziflam interaction.

On the other hand, the lack of cross-resistance of indaziflam to CESA in this study suggests that this altered target enzyme is not the only target of indaziflam. However, the mode of action and the specific target of indaziflam has not yet been identified. Further work in elucidating mechanisms is required, but identifying strong resistance to indaziflam expressed by both *izr6* and *izr10* makes these mutants a unique tool for further understanding auxin transduction processes concerning toxicity.

A previously reported *izr1* mutant (Bonneta, unpublished) demonstrated increased tolerance to the synthetic auxins, 2,4-D, and Dicamba, reduced gravitropism, apical dominance, and germination, which bring us to a conclusion that this mutation disrupts the interaction between CUL1 and ASK1 [196] and affects the auxin response. A conceivable explanation for why mutations in *CUL1* affect indaziflam tolerance is that indaziflam can bind to the auxin binding site of TIR1. This would lead to degradation of AUX/IAA repressor proteins in wild-type plants, followed by herbicide-induced auxin response. However, the *izr10* plant would have a limited response, which would result in increased indaziflam tolerance.

The phytohormone auxin contributes to many essential cellular processes, such as cellular polarization and expansion, cell wall remodeling, polar tip growth, and vesicle trafficking [238, 259-262]. The link between cell wall modifications and auxin is proven, but most of the auxin-dependent cell wall effectors are still unknown. Our study urges a stronger focus on the cell wall as an appropriate cellular system to generate such fundamental insights into the auxin-dependent cellular expansion regulation in *Physcomitrium*.

The data presented in this study predicted that indaziflam targets multiple pathways integral to early growth and development in plants. Though millions of mutagenized *P. patens* protoplasts have been screened on indaziflam, only a small number have demonstrated any indaziflam tolerance. Fully understanding the mode of action of

indaziflam is necessary to ensure that this herbicide is being used safely and will have a minimal effect on non-target species.

Future work will focus on the characterization of the allele in other *izr* mutant lines. Although this screen yielded 46 possible *izr* mutant lines, only two were selected for characterization. Testing the other lines for cross-resistance to phytohormones and herbicides may allow for better characterization of these lines. Understanding the indaziflam mode of action will help apply the knowledge to crops species and produce herbicide-resistant crops and reduce competition from weeds.

Chapter three: Determination of cellulose synthase complex composition using BioID (Biotin identification)

3.1 Introduction

The cellulose synthase complex (CSC) interactome is a complicated network that remains to be fully understood. One central issue in determining the function of the protein complexes like the CSC is to elucidate what proteins play a role in the complex's function. CESA-binding partner identification has previously been done using techniques such as yeast two-hybrid, affinity purification, and co-immunoprecipitation (co-IP) [263]. In doing so, a key consideration is whether the complex's integrity has been preserved and if weak interactions are captured.

A commonly used technique, co-IP, can be a powerful way to study protein-protein interactions. It can show direct interactions *in vitro* that need to be verified using other methods. In *Arabidopsis,* this technique was used to show the 1:1:1 molecular ratio of CESA1, CESA3, and CESA6 [264], a more recent study also using the same method proposed, other proteins may support the CSC machinery in addition to cellulose synthase proteins [265]. This technique's major limitation is tied to conditions such as the pH of the binding partners' incubation duration, leading to false positives. Furthermore, artifacts can arise because two proteins might interact as a result of physical manipulation.

Another important method for detecting protein-protein interactions is the yeast two-hybrid approach. In contrast to IP, yeast two-hybrid detects protein interactions *in vivo,* closer to the normal conditions where protein partners interact. Nevertheless, the overexpression of two proteins of interest may cause an artificial interaction because of each protein's non-physiological levels. Also, proteins might not undergo appropriate folding and post-translational modification since they are often not native to the yeast cell producing them [266]. The technique is also dependent on those proteins that can form complexes in the cell's nucleus.

Both techniques do not provide sufficient evidence for a physiologically relevant interaction, and some interactions may also be too weak to detect. In the case of the CSC, the inability to reliably assay cellulose synthase activity *in vitro* implies that the complex is

particularly labile, making the use of IP problematic. Furthermore, since the complex is membrane-bound, yeast two-hybrid assays are seriously limited. Alternative approaches for detecting CSC interacting partners should be explored. The Biotin Identification (BioID) pull-down assay is complimentary to affinity capture approaches established by Kyle Roux [267]. The approach depended on proximity-based biotinylation to characterize protein-protein interactions in living cells [263](Figure 3.1) and was designed to identify neighboring proteins with direct or indirect physical interaction with the protein of interest [263]. The system involves the fusion of a promiscuous biotin protein ligase (BirA*- R118G) from *E. coli* [268] that can add a biotin tag or, in other words, biotinylate all proteins in close proximity to the protein of interest that is fused to the BirA* enzyme and create a list of near-neighbors. Since biotinylation is an uncommon post-transcriptional modification and restricted to a few carboxylases [268], it makes the biotinylation of proximate proteins a useful tool for identifying the potential interactors. The R118G substitution in BirA* causes a reduction in activated biotin affinity (biotinoyl-5'-AMP). As a result, it can activate biotin efficiently, and the highly reactive biotin will diffuse away and react with proximal amine groups, including primary amines in lysine residues of adjacent proteins [263].

The biotinylation radius on this *in vivo* method is <20nm for practical labeling purposes [269]. The biotinylated protein may then be affinity purified using streptavidin following cell lysis and identified using mass spectrometry and creating a list of near-neighbors proteins [263].

The principle is based on DamID that uses Dam methylase that methylates DNA sequences that the protein comes into contact with when fused to a DNA binding protein and detects DNA-protein interactions [270].

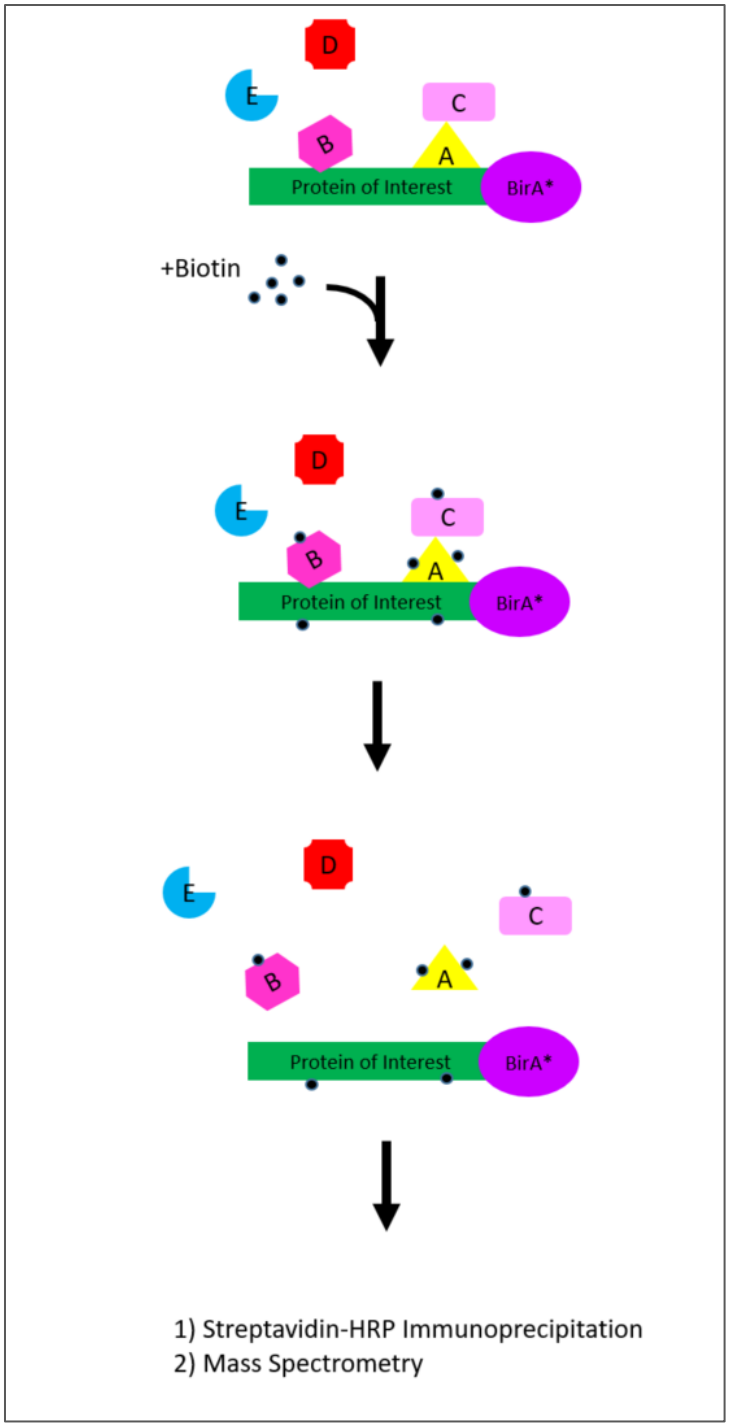


Figure 3.1: Overview of BioID method. A protein of interest is fused to the promiscuous biotin ligase BirA*. After the addition of biotin, BirA* will biotinylated proteins in a proximity-dependent fashion. BirA* will biotinylate both direct (A, B) and indirect (C) protein binding partners. It will not contain biotinylated proteins that are not within close proximity (D, E). After a 24-hour incubation with biotin, cells are lysed, and biotinylated proteins are immunoprecipitated using streptavidin-sepharose beads. Proteins are digested on-bead and subjected to mass spectrometry analysis (Adapted from [263]).

A refinement of the method came with the development of a smaller BirA*, or BioID2. BioID2 is less disruptive to the trafficking of fusion proteins to different subcellular compartments, and it is efficient. A 13 'GGGGS' repeat linker is added to either N or C terminus of BioID2 to increase the biotinylation radius without affecting the enzyme activity [271].

There are many advantages to this method. For example, there is no need for the protein of interest to be soluble and localize to any cell compartment. Since biotinylated polypeptides are purified using a binding affinity of streptavidin [272], protein interactions are recorded before cell lysis, and even weak interactions can be detected. The other advantages are that labeling occurs in a native cellular context, so *in vivo* protein processing is accurate, and the interactions do not need to be maintained throughout the processing of cell extracts leading up to affinity capture [272].

Despite these points, the BioID method does have some disadvantages. One important one is the potential of missing low abundance proteins; these proteins may not be detected in cases where it is impossible to collect sufficient starting material [263]. BirA* itself is quite large, 35 kDa, and it could alter the fusion protein's structure or function and cause its mislocalization or reduce its binding affinity for protein partners [263]. Another critical problem with using BioID is that the BirA*can have different effects depending on whether it is a C-terminal or N-terminal fusion. Studies have shown that fused proteins have very different behaviors when tagged at the carboxyl or amino terminus, and the expression level can be lower than unmodified protein [273].

Another limitation relates to solubility since the regular BioID approach needs a biotin incubation time of at least 6 hours. It is less suitable for soluble proteins because they are washed from the cells after permeabilization [274]. In cases where solubilized proteins are required, bait protein may not be compatible with stabilizing interactions with partner proteins and vice versa, and this is a major issue when it comes to weak interactions. In the case of CESA, a highly insoluble protein, this has proved to be a severe setback in the reliable identification of interacting proteins.

BioID2 seems to overcome some of these drawbacks and its influence on fusion protein localization or function. BioID2 appears to localize to the fusion protein's location more efficiently and functions best at 50°C [271], although this can be a disadvantage for some organisms.

Applying the technique to whole plant systems and proteomics application in plants have unique challenges compared to animal systems. Some of these challenges are high protease and phosphatase content and low cytoplasmic volume compared to cell wall mass, inhibiting protein detection and identification [275]. Despite all the challenges, proteomics is a crucial component of plant biology.

Here I present the BioID technique applied to the model plant *Physcomitrium patens* and use CESAs tagged BirA* and BirA* alone as proof of concept to develop BioID in *P. patens* and to compare the proximal proteome of cell-wall-associated baits.

3.2 Methods

3.2.1 *Physcomitrium patens* growth conditions

Physcomitrium patens wild-type ecotypes Gransden was used for all experiments and observations. All *P. patens* sample*s* were grown on the media called BCD and BCDAT described in the previous chapter (table 2.1). The Petri dishes containing plants on media overlayed with sterile cellophane disks were incubated at approximately 21-25°C under continuous illumination. The plants were subcultured by homogenizing tissue in sterile water and plating them onto BCDAT medium overlaid with cellophane disks and grown under the same conditions for 5 to 7 days for protoplast isolation. Culture manipulations were typically carried out under the "BIO Guard" containment hood when transfer to a new culture medium was required (Figure 3.2).

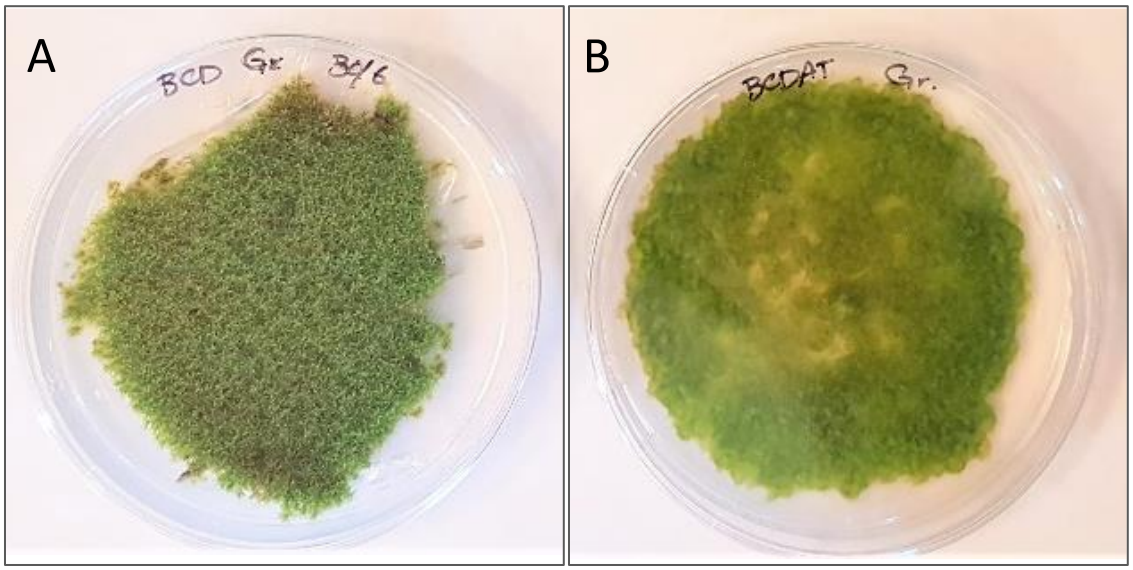


Figure 3.2: *Physcomitrium* has grown on medium plates overlaid with cellophane. (A) *P. patens* Gametophyte (Gransden ecotype) on BCD medium. (B) *P. patens* protonema (Gransden ecotype) on BCDAT medium.

3.2.2 Overview of knock-in methods

The primary methods for creating knock-ins are outlined in Roebroek et al. [276]. In essence, a construct is prepared that includes 5' and 3' intergenic sequences flanking the gene targeted for insertion. The knock-in target gene is then replaced by a targeting construct and a selection marker such as an antibiotic. Most cloning steps involved in creating the knock-in vector are accomplished using the New England BioLabs® Gibson Assembly® system. When the knock-in vector is transformed into wild-type *Physcomitrium* protoplasts, some undergo homologous recombination with the wild-type gene locus.

Knock-in mutants are selected by treating cells regenerated from the protoplasts with the antibiotic. I genotyped the mutants using PCR to ensure that homologous recombination has occurred in the appropriate location.

General information on Gibson Assembly® cloning is described in the following sections [276].

3.2.3 Searching the *Physcomitrium patens* Genome for *CESA* Genes and choosing the genes of the interest

I compared *P. patens* transcriptome data with available data from *A. thaliana* to extract *CESA* homologs genes of *P. patens*, and thus, it probably played a significant role in cellulose biosynthesis. Firstly, I have chosen *CESA3* based on the full data set of the electronic Fluorescent Pictographs (eFP) browser created by the University of Toronto (http://bar.utoronto.ca/efp_physcomitrella/cgi-bin/efpWeb.cgi). The represented data showed that *CESA3* has the highest expression level in gametophore and the lowest on protometal tissue (Figure 3.3). So it may have a specific role in gametophore cell wall development. The diagrams illustrating the various *P. patens* tissues were generated and provided an intuitive and graphical representation of the selected gene's expression level, enabling users with no bioinformatics experience to explore the data [16] quickly.

In 2012, Goss et al. showed that *CESA5* in moss plays a role in the morphogenesis of leafy gametophore and is also upregulated by cytokinin, which also induced gametophore development [77]. They showed that both cell expansion and cytokinesis are disrupted in *cesa5KO*, while cytokinesis is unaffected in *Arabidopsis CESA* mutants [77, 93]. Whereas leafy gametophores are severely affected by the absence of *PpCESA5*, protonemal filaments stay unaffected in *cesa5KO* lines [77].

To study cellulose synthase interactomes, *CESA5* was also chosen since changes in phenotype can indicate interruptions in gene expression or protein modification.

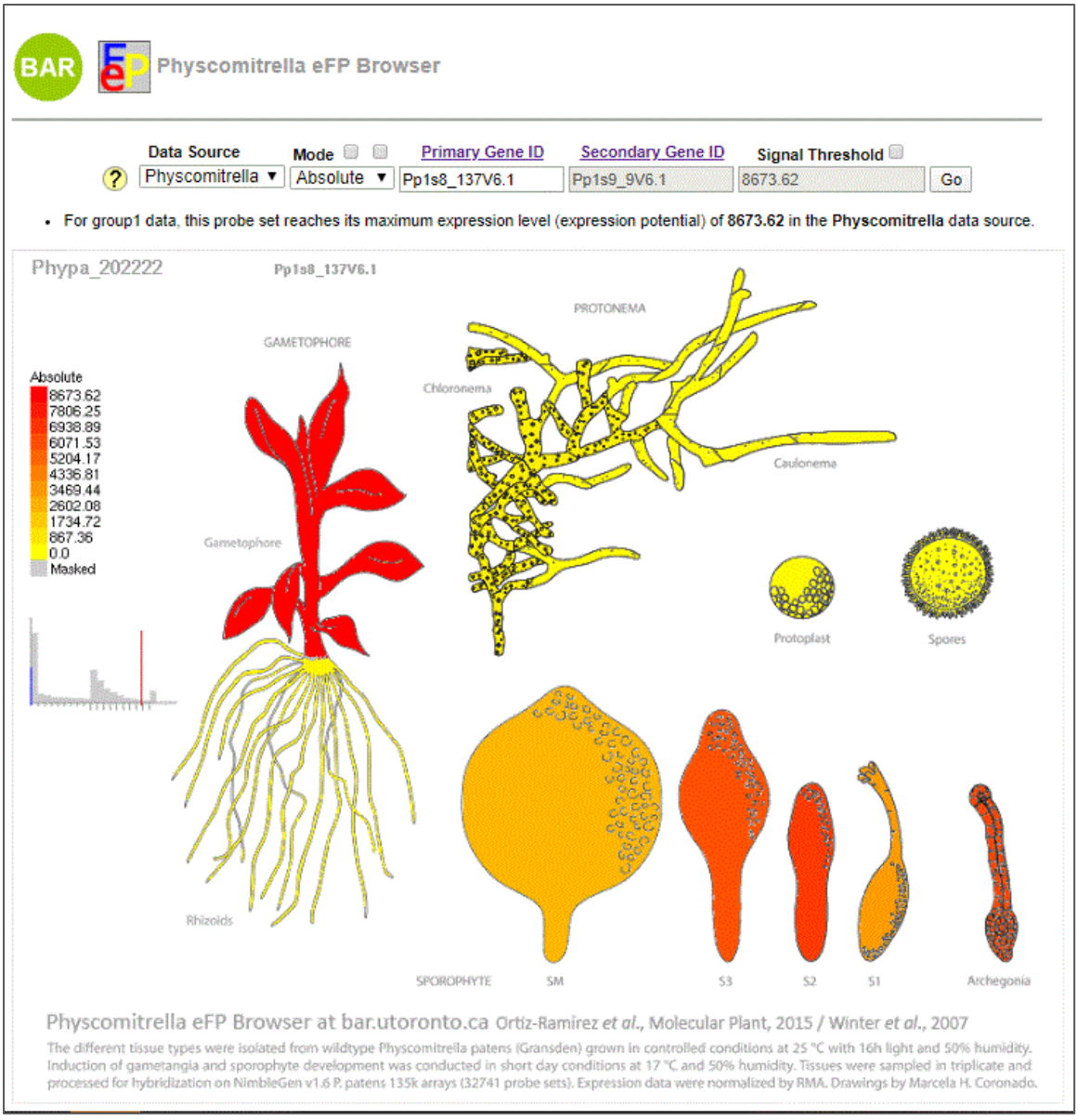


Figure 3.3: *Physcomitrium* eFP browser. Cartoon representation of *P. patens* tissues integrated into the Bio-analytical Resource for Plant Biology (BAR). Together with all expression data generated, images are available in BAR for public consultation (http://bar.utoronto.ca/efp_physcomitrella/cgi-bin/efpWeb.cgi)

3.2.4 Molecular cloning

3.2.4.1 Isolation of *Physcomitrium* genomic DNA

For genomic fragment cloning and PCR genotyping, *P. patens* DNA used for PCR amplification of intergenic regions flanking the *CESA3* and *CESA5* genes and was isolated using the method cetyltrimethylammonium bromide (CTAB) with minor modifications [277]

as described in chapter 2 (section 2.2.8). DNA samples were quantified by using a BioDrop Due (80-3006-61) using two µL samples.

3.2.4.2 Primers design, polymerase chain reaction (PCR) amplification of fragments, and cloning techniques

For molecular cloning, several sequence-independent DNA assembly techniques exist which can assemble multiple fragments efficiently. One example of these techniques includes the Gibson Assembly® to assemble PCR (polymerase chain reaction) fragments and linear plasmid [278]. Dr. Daniel Gibson and his colleagues created this technique [279, 280]. This technique is used to assemble and combine multiple DNA fragments simultaneously in one reaction fast and efficiently and create translational fusion.

The constructs' design was done in silico using ApE Plasmid Editor and Snapgene® Viewer following the instructions manual of Gibson Assembly® by New England BioLabs® (NEB). The design of primers was carried out using the online tool, Primer-BLAST from NCBI. Primers were designed to amplify fragments with appropriate 15 to 40 bp homology regions between the fragments. The Gibson Assembly® technique does not need restriction digestion and is mostly sequence-independent. The DNA fragments for assembly were obtained with PCR of the moss genome or PCR of specific plasmids, and gBlocks® synthesized fragments by Integrated DNA Technologies (IDT®), followed by PCR purification using QIAquick® PCR Purification Kit (QIAGEN). In cases where the PCR products have undesired fragments, agarose gel extraction was performed to purify further the PCR product following separation in a standard agarose electrophoretic gel buffered Tris-Acetate-EDTA (TAE) running buffer following the protocols described in Sambrook [281]. In addition to the DNA fragments, the reaction mix contains three enzymes, including 5' exonuclease, DNA polymerase, DNA ligase, and other required reagents. The 5' exonuclease chews back the 5' end of each strand so that the 3' end of neighboring strands that overlap each other can anneal and act as primers for each other in the next step; then, the DNA polymerase adds nucleotides to the 3' end; because the polymerase works faster than the exonuclease, the gaps will be filled, and the DNA ligase, seals the nick. All the components are put in a single tube and incubated at 50°C for one hour [278-280, 282].

After performing the assembly according to the manual's instructions and if the product was linear synthetic DNA, PCR was performed to verify and amplify the desired product for downstream applications such as having enough DNA to effect *Physcomitrium* transformation. If the DNA fragments were assembled into a plasmid, five µl of the reaction mix was used to transform chemically competent *E. coli* DH5α™ cells, following the NEB manual protocol. Chemically competent *E. coli* DH5α™ cells used are also purchased from NEB. If colonies appear on LB media [1 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 1 % (v/v) NaCl] plates with proper antibiotics, several of them were picked and inoculated into liquid LB media, and plasmid purification was done using Phenol-Chloroform plasmid preparation techniques as described Sambrook [281]. Transformants were verified with PCR of the purified plasmids with primers specific to the expected fragment insert.

The Gateway expression system is another way to assemble DNA fragment, developed by Thermo Fisher Scientific, enables site-specific recombination from entry vector such as pDONR/pENTR into a destination vector. The site-specific recombination is dependent on the attL1 and attL2 viral recombination sites. These recombination sites maintain the reading frame of the sequence and match it to the destination vector's reading frame. Because the recombined insert is in frame with the destination vector, the gene product can be fused N-terminally or C-terminally to additional sequences upstream or downstream of the recombination site, such as fluorescent proteins.

In each case, the insert sequence was amplified by PCR or using synthesized gblocks® with attL sites at each ends, then cloned into the pDONR plasmid using the BP clonase Gateway® cloning kit (Thermo Fisher Scientific), based on the manufacturer's specification creating an entry clone. Once I had performed the BP recombination reaction, I transformed competent *E. coli* and select for entry clones using the appropriate antibiotic (Gentamicin for pDONT207 plasmid). DH5α™ chemically competent *E. coli* was used for this purpose. The protocol to transform chemically competent is provided in the next section. A new pDONR plasmid with the new DNA fragments at the cloning sites was introduced into the destination vector pTZ-UBIgate (Figure 3.13A) by LR clonase Gateway® cloning kit (the detailed steps of the Gateway® cloning are explained in details in the following sections).

3.2.5 Plasmid construction

3.2.5.1 BioID-CESA3-pUC19

The N-terminal BirA, BioID, was amplified from the pcDNA3.1 mycBioID vector (Addgene #35700) by PCR to generate the BioID-CESA3 fusion protein in the pUC19 backbone vector using primers listed in Table 3.1. Fragments from *PpCESA3* include the 5'UTR, Exon1-Intron1, Intron1-Exon2-Intron2-Exon3-Intron3. A neomycin phosphotransferase (NPTII) gene was used as a selection marker and was amplified from the expression vector pCAMBIA2301. Vector pUC19 is also amplified with the overlap primers as the destination vector and using the Gibson Assembly® Master Mix from NEB. Following incubation at 50°C, the result is a double-stranded DNA molecule stored on ice or at –20°C for subsequent transformations to *E. coli* DH5α™. Since there are multiple fragments shown in Figure 3.4, increasing the DNA assembly's efficiency for this construct is performed in three steps. (i) Assembly one is included 5'UTR of *CESA3*, BioID, Exon1-Intron1 into pUC19 vector, and (ii) the second assembly is included the cauliflower mosaic virus (CaMV) 35S promoter, NPTII, and *CESA3* Exon2-Exon3, all of the DNA fragments were cloned into the pUC19 vector. CaMV 35S promoter is used to drive the expression of recombinant proteins in plants [283]. This study used the CaMV 35S promoter to express the NPTII gene used as a selection marker. For the third assembly (iii) 5'UTR, BioID, and Exon1-Intron1 from assembly one and CaMV 35S promoter, NPT II, and Exon2-Exon3 from assembly two were amplified from pUC19 by PCR with the overlap primers (Table 3.1). These new PCR fragments have 30 bp overlap with one another and pUC19 vector and were assembled to linearized pUC19 as the final construct.

3.2.5.2 Venus-CESA3-pUC19

The same procedure was used for the Venus-CESA3 fusion in which the Venus gene, an enhanced Yellow Fluorescent Protein (YFP) variant, replaced BioID. The Venus variant has improved maturation and brightness, as well as reduced environmental dependence [284]. In this study, I am trying to determine the subcellular localization of proteins by analyzing the fluorescently tagged fusion proteins. For this purpose, I will examine the localization of

fusion protein Venus-CESA3, and I will demonstrate its co-localization with possible interactors.

3.2.5.3 KnockOut-CESA3-pUC19

As a control, I have created another knock-in construct, in which only the NPTII gene is placed inside the intron two as it was in BioID-CESA3 and Venus-CESA3 knock-in constructs (Figure 3.4). Since the constructs contain the NPTII gene as a selection marker, if there are any interferes by NPTII to the fusion protein function, KnockOut-CESA3 will distinguish the disruption that may cause by either BioID or Venus versus NPTII.

3.2.5.4 pTZ-UBIgate-FLP-FRT

To avoid any possible lethal effects or gene disruption, I have generated a system that will utilize the Flp recombinase from yeast to excise the NPTII gene in situ [285]. To this end, I have developed a moss strain that expresses Flp recombinase fusion to the human estradiol receptor (FlpeERT2) [286]. This system makes it possible to selectively induce recombinase activity with the addition of estradiol or one of its analogs. The FLP recombinase gene that I have used was derived from pCAG-FlpeERT2 [286].

Flpe was used to excise the NPTII selection marker, which is flanked by FLP recombination targets (FRT), a 34bp sequence 5'GAAGTTCCTATTCtctagaaaGtATAGGAACTTC3' [285]. The expression vector pTZ-UBIgate (10µg) was double digested by AbsI (20units) and NotI (20units) restriction enzymes (NEB) at 37°C for 3 hours as per the manufacturer's instruction. The FLP recombinase gene was amplified from pCAG-FlpeERT2 by PCR with primers with 20-40 bp overlap with pTZ-UBIgate plasmid at restriction enzyme sites. The FLP-FRT fusion was cloned into the moss-specific vector pTZ-UBIgate (Figure 3.6) using the same Gibson Assembly® strategy as described previously. The vector is designed so that homologous recombination within the moss chromosome occurs at a redundant copy of the Actin Related Protein 2/3 Complex Subunit 3 (*ARPC3*) gene and does not affect growth or development [287].

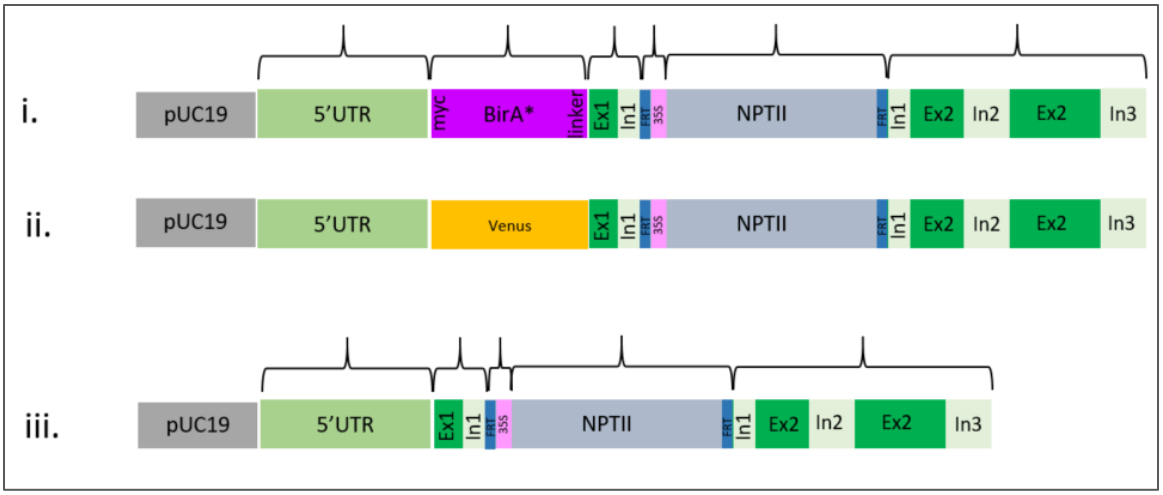


Figure 3.4: PCR fragments used for Gibson Assembly® (i) BioID-CESA3-pUC19, (ii) Venus-CESA3-pUC19, (iii) KO-CESA3-pUC19.

3.2.5.5 BioID-pTZ-UBIgate

The BioID was amplified from the pcDNA3.1 mycBioID vector (Addgene #35700) by PCR to generate the BioID-pUTZ UBIgate construct as a control. I designed primers with 20-40 bp overlap with pTZ-UBIgate plasmid at restriction enzyme sites (Table 3.1). The expression vector pTZ-UBIgate (10µg) was double digested by restriction enzymes (NEB) AbsI and NotI as per the manufacturer's instruction, and the BioID gene was cloned into the double digested vector pTZ-UBIgate using the Gibson Assembly® strategy as described previously.

3.2.5.6 BioID2-CESA5-pTZ-UBIgate

To generate the construct, the Gibson Assembly® technique has been used to assemble nine synthesized gBlocks®, including BioID2 and full-length *CESA5* cDNA (Table 3.2), adding attB1 and attB2 sequences at both N-terminus of BioID2 and C-terminus of *CESA5* cDNA into the backbone pUC19 vector that generated by PCR with the overlap primers. The Genescript company synthesizes gBlocks®. BP recombination of pDONR™ and BioID2-CESA5-pUC19 constructs were done to create the entry clones. According to the manufacturer's instructions (NEB), the respective BioID2-CESA5-pUC19 clone was linearized with the restriction enzyme HindIII to increase recombination efficiency. To perform homologous recombination, 100 ng linearized BioID2-CESA5-pUC19 clone as an entry clone

was mixed with 150 ng of pDONR™207 as the destination vector and 4µl BP Clonase reaction mix (Thermo Fisher Scientific).

The solution was brought to a total volume of 5 µl with TE buffer pH 8.0 (10 mM Tris/HCl pH 8.0, 1 mM EDTA). Following incubation at 25°C (for the highest efficiency, the incubation time increased to three hours), one µl Proteinase K solution was added (Thermo Fisher Scientific), the samples were transferred to 37°C for 10 min to terminate BP clonase activity. Afterward, one µl of the BP reaction mixture was used to transform *E. coli* stain DH5α™. For each transformation, an aliquot of competent *E. coli* was thawed on ice. 25 µl bacterial suspension was mixed with a plasmid. After incubation on ice for 30 min, the cells were heat-shocked for 30 seconds at 42°C and immediately transferred back to ice. 500 µl SOC was added. The bacteria were incubated for one hour at 37°C to enable recovery. Fifty µl bacteria suspension were spread on plates containing LB medium supplemented with appropriate antibiotics and 1 % (w/v) agar.

The kind of antibiotic depended on the resistance encoded on the plasmid used for the transformation procedure. The selection of positive clones was performed on LB media supplemented with 1 % (w/v). The destination clone pDONR™207 conferred resistance to Gentamicin (7 µg/ml). Putative positive clones were re-evaluated by performing colony PCRs. Cells from one single *E. coli* colony were transferred to a 25 µl PCR mixture. The colonies were picked from a plate containing selecting antibiotics, which was incubated overnight at 37 °C.

Following an initial denaturation for 30 min for cell lysis, the reactions were performed according to standard PCR protocol. For each colony PCR, a control sample containing no template was included to ensure the purity of the used reagents. Positive colonies for successful cloning were isolated from 5 ml LB medium cultures supplemented with antibiotics. They had been incubated overnight at 37°C under continuous shaking at 250 rpm. The extraction was performed by the plasmid Miniprep DNA purification technique. The extraction is base on the ability of silica membranes to bind plasmid DNA under high salt conditions. Before loading the column, cell debris and precipitated proteins are removed by centrifugation. RNase I digested RNA, and remaining contaminations were

discarded in several washing steps. The elution of the plasmids from membranes was performed with autoclaved deionized water. After extraction, the DNA content was checked for quality and purity using the BioDrop® Due (80-3006-61). All samples were either stored at -20°C or directly used for restriction digests and LR Clonase™ reaction cloning procedures.

For LR Clonase™ recombination of BioID2-CESA5-pTZ-UBIgate, the destination vector, pTZ-UBIgate (100 ng), and the entry vector (150 µl) were mixed with 4 µl LR Clonase® reaction mix (Thermo Fisher Scientific) and 5 µl of TE buffer pH 8.0 (10 mM Tris/HCl pH 8.0, 1 mM EDTA). Following incubation at 25 °C for 3 hours, one µl Proteinase K solution was added, and the LR clonase activity was terminated by incubation at 37 °C for 10 min. Next, one µl of the LR reaction mixture was used to transform *E. coli* stain DH5α™ as described before. The destination clones pTZ-UBIgate conferred resistance to Ampicillin (50 µg/ml). Colony PCRs, plasmid isolation, and confirmation of final construct BioID2-CESA5-pTZ-UBIgate were performed as described earlier, and all samples were purified and stored at -20°C or directly used for further procedures.

All the construct maps are generated by SnapGene® Viewer are shown in appendix 3.

Table 3.1: Primer list of *CESA3* knock-in constructs used for Gibson Assembly®

Primer name	Sequence	Use
	BioID-CESA3-pUC19 construct	
pUC19-F	CGACGTTGTAAAACGACGGCCAGTGAATTC	Gibson Assembly®
5UTR-F	CGACGTTGTAAAACGACGGCCAGTGAATTCGTCTGTTTAGTTTCGTTTGGTTTCGCTGTA	Gibson Assembly®
5UTR-Bio-R	ATCCTCTTCTGAGATGAGTTTTTGTTCCATTGCTGCAACGCCACTCCGCTCGCTCCACTG	Gibson Assembly®
Bio-F	CAGTGGAGCGAGCGGAGTGGCGTTGCAGCAATGGAACAAAAACTCATCTCAGAAGAGGAT	Gibson Assembly®
Bio-R	CATGGCCGCTGCCGCAGCGGCAGCGGCAGCCTCGAGCTTCTCTGCGCTTCTCAGGGAGAT	Gibson Assembly®
pUC19-R	CCTGCAGGTCGACTCTAGAGGATCCCCGGG	Gibson Assembly®
Exo1-F	GCGGCCGCTGCCGCTGCGGCAGCGGCCATGGAGGCTAATGCGGGCCTGGTTGCTGGATCC	Gibson Assembly®
Exo1-R	CCTGCAGGTCGACTCTAGAGGATCCCCGGGTTCCTATACTTTCTAGAGAATAGGAACTTCAAACACAGAGGAAACAGATTCATCGAGTCT	Gibson Assembly®
NPT-F	CGACGTTGTAAAACGACGGCCAGTGAATTCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCATGGAGTCAAAGATTCAAATAGAGG	Gibson Assembly®
NPT-R	TTCCTATACTTTCTAGAGAATAGGAACTTCTCAGAAGAACTCGTCAAGAAGGCGATAGAA	Gibson Assembly®
Exo2-F	GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCAGTTGAATTAGATCAACATTTTGGC	Gibson Assembly®
Exo2-R	CCTCATGCCGGTGTGGTCATGTTGATACAGCCCGGGGATCCTCTAGAGTCGACCTGCAGG	Gibson Assembly®
NPT-F2	CGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCATGGAGTCAAAGATTCAAATAGAG	Gibson Assembly®
Exo1-R2	TTCCTATACTTTCTAGAGAATAGGAACTTCAAACACAGAGGAAACAGATTCATCGAGTCT	Gibson Assembly®
	Venus-CESA3-pUC19 construct	
Exo1k-F	CAGTGGAGCGAGCGGAGTGGCGTTGCAGCAATGGAGGCTAATGCGGGCCTGGTTGCTGGA	Gibson Assembly®
Exo1k-R	GAAGTTCCTATACTTTCTAGAGAATAGGAACTTCATTCATTGCAAGCCACAAACAATTCG	Gibson Assembly®
5UTRk-R	TCCAGCAACCAGGCCCGCATTAGCCTCCATTGCTGCAACGCCACTCCGCTCGCTCCACTG	Gibson Assembly®
5UTR-F	CGACGTTGTAAAACGACGGCCAGTGAATTCGTCTGTTTAGTTTCGTTTGGTTTCGCTGTA	Gibson Assembly®
NPTk-F	GAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCATGGAGTCAAAGATTCAAATAGAGG	Gibson Assembly®
NPTk-R	GAAGTTCCTATACTTTCTAGAGAATAGGAACTTCTCAGAAGAACTCGTCAAGAAGGCGAT	Gibson Assembly®
Exo2k-F	GAAGTTCCTATTCtctagaaaGtATAGGAACTTCGTGCGTTTCCAGTATGCCGCACGTGC	Gibson Assembly®
Exo2-R	CCTCATGCCGGTGTGGTCATGTTGATACAGCCCGGGGATCCTCTAGAGTCGACCTGCAGG	Gibson Assembly®
Exo1a-R	GAAGTTCCTATACTTTCTAGAGAATAGGAACTTCAAACACAGAGGAAACAGATTCATCGA	Gibson Assembly®
	KO-CESA3-pUC19 construct	
5UTR-F	GTAAAACGACGGCCAGTGAATTCGAGCTCGGTCTGTTTAGTTTCGTTTGGTTTCGC	Gibson Assembly®
5UTR-R	TTCACCTTTAGACATTGCTGCAACGCCACTCCG	Gibson Assembly®
Venus-F	AGTGGCGTTGCAGCAATGTCTAAAGGTGAAGAATTATTC	Gibson Assembly®
Venus-R	CAGCGGCAGCGGCCGCTTTGTACAATTCATCCATACCATG	Gibson Assembly®
Exon1-F	GATGAATTGTACAAAGCGGCCGCTGCCGCTGCGGCAGCGGCCATGGAGGCTAATGCGGGC	Gibson Assembly®
Exon1-R	CATGATTACGCCAAGCTTGCATGCCTGCAGAAACACAGAGGAAACAGATTCATCG	Gibson Assembly®
	BioID-CESA3-pTZ-UBIgate/BioID-pTZ-UBIgate constructs	
BioID-CESA3-pTZ-F	GGAGTAGGGGTATCGGGCCCCCCCGTCTGTTTAGTTTCGTTTGGTTTCGCTGTA	Gibson Assembly®
BioID-CESA3-pTZ-R	TAATTCGAGCTCCACCGCGGTGGCGGCCGCCTGTATCAACATGACCACACCGGCATGAGG	Gibson Assembly®
BioID-pTZ-F	GGAGTAGGGGTATCGGGCCCCCCCATGGAACAAAAACTCATCTCAGAAGAGGATCTCGACA	Gibson Assembly®
BioID-pTZ-R	AATTCGAGCTCCACCGCGGTGGCGGCCCTCGAGCTTCTCTGCGCTTCTCAGGGAGATTTC	Gibson Assembly®
	FLP-FRT-pTZ-UBIgate	
FLP-FRT-F	GGAGTAGGGGTATCGGGCCCCCCCTCGAGGTCAAGCTGTGGCAGGGAAAC	Gibson Assembly®
FLP-FRT-R	TAATTCGAGCTCCACCGCGGTGGCGGCCGCATGAGCCAATTTGATATATT	Gibson Assembly®

Table 3.2: List of gBlocks® for assembling BioID2-CESA5 into pTZ-UBIgate backbone plasmid

gBlocks® name	Gene Fragment	Size (bp)	Use
BC51	attB1, BioID2, CESA5 cDNA	500	Gibson Assembly®
BC52	*CESA5* cDNA	500	Gibson Assembly®
BC53	*CESA5* cDNA	500	Gibson Assembly®
BC54	*CESA5* cDNA	500	Gibson Assembly®
BC55	*CESA5* cDNA	500	Gibson Assembly®
BC56	*CESA5* cDNA	500	Gibson Assembly®
BC57	*CESA5* cDNA	500	Gibson Assembly®
BC58	*CESA5* cDNA	500	Gibson Assembly®
BC59	*CESA5* cDNA, attB2	409	Gibson Assembly®

3.2.6 Protoplast isolation

Using the method of Roberts et al. [13], protoplasts were isolated using the wild-type 5-7 days old chloronemal tissue. The chloronemal tissue was digested using a 2% Driselase™ (Sigma-Aldrich) solution dissolved in 50 mL 8.5% d-mannitol. The mixture was stirred gently for 30 minutes to one hour at 21–25°C, then the protoplast suspension was passed through a nylon filter and transferred to a new conical centrifuge tube and centrifuged at speed 200g for 3 minutes, the supernatant was discarded, and protoplasts were resuspended in 10 mL 8.5% mannitol. The centrifugation of protoplast suspension was repeated three times after resuspension in 10 mL of 8.5% mannitol each time. Ten µL of suspension is loaded into a hemocytometer, the intact protoplasts were counted, and the density was estimated at approximately 2–4 × 10^5 protoplasts/mL. PRMT medium is melted and equilibrated to 45°C in the water bath; protoplast regeneration plates were prepared with PRMB medium and overlaid with sterile cellophane disks. Protoplasts were resuspended in 5 mL PRMT held at 45°C, and 1.5 mL of suspension was spread on each PRMB plate. Plates were incubated for five days at 25°C with continuous light at 50–80 mmol/m^2/s [162].

3.2.7 PEG-mediated *Physcomitrium patens* transformation

The isolated Protoplasts were washed twice by centrifuging at 200g for 3 minutes in 8.5% mannitol. Protoplasts were counted and resuspended at 1.3 x 10^6/ml in 8.5% mannitol, 15 mM $MgCl_2$, 1% MES-KOH pH5.6 (MMM solution). Thirty µl of DNA solution of each

construct (BioID-CESA3-pUC19, Venus-CESA3-pUC19, KO-CESA3-pUC19, BioID-pTZ-UBIgate, FLP-FRT-pTZ-UBIgate, and BioID2-CESA5-pTZ-UBIgate) were added to separate tubes containing protoplast suspensions and gently mixed, followed by adding 300 µl PEG/T solution (40% (w/v) PEG 6000 (Prolabo) in 9% (w/v) mannitol containing 0.1 M $Ca(NO_3)_2$, pH 8.0). The PEG (polyethylene glycol) was either autoclaved before dissolving in the sterile mannitol solution or sterilized in the final solution by filtration. The PEG solution was prepared 2-3 hours fresh before each experiment. The transformation mix was incubated for 30 minutes at room temperature. The protoplast suspension was heat-shocked at 45°C for 5 minutes, cooled to 20°C water bath for 10 minutes. The transformation mixture was diluted by adding five aliquots of 300 µl protoplast liquid medium at 3 minutes intervals and then five aliquots of 1 ml of the protoplast liquid medium at 3 minutes intervals. The diluted protoplast solution was transferred into a 6 cm Petri dish and incubated at 25˚C overnight in darkness. The following day, protoplast suspension was transferred into a 15 ml conical tube and centrifuged at 200g for three minutes at room temperature. The supernatant was discarded, and eight ml of a PRMT medium stored at 45°C was added to the protoplast pellet. Protoplasts were resuspended, and 2 ml of the protoplast suspension was poured into PRMB medium plates overlayed with cellophane disks, and the Petri dishes were sealed with surgical tape and incubated at 25˚C under continuous light with a light flux of 30 - 80 µmol/m^2/s for five days. After the 5-day incubation, the cellophane disks were transferred to medium without mannitol and containing Zeocin (50 µg/ml) or Ampicillin B (25-30 µg/ml), for pTZUBIgate vector and pUC19 vector respectively, based on their selection markers. Antibiotic-resistant plants were observed from 10 to 30 days after selection [288].

3.2.8 Selecting stable transformants

Initial transformants obtained from protoplast transformation were removed after 21 days from the BCDAT medium containing Zeocin or Ampicillin to BCD medium without antibiotics for 12-14 days at 25°C under continuous light. Transformants were then transferred back to BCDAT medium containing antibiotics for another 14 days. Plants

either died or grew at a slower rate. Stably transformed plants that transiently transformed grew at the same rate in the presence or absence of antibiotics. Stable transformants of FLP-FRT-pTZ-UBIgate were used for the second transformation procedure for BioID-CESA3-pUC19, Venus-CESA3-pUC19, KO-CESA3-pUC19 purified constructs.

3.2.9 Genotyping of knock-in moss

Putative transformants were genotyped by PCR to ensure that homologous recombination had inserted the knock-in constructs into the correct position in the genome (replacing the *CESA3* or *ARPC3* gene). Genomic DNA from individual colonies (each derived from a single protoplast) was isolated using the CTAB protocol [174]. The pellets obtained were dissolved in 200 µL of molecular grade water to serve as templates for PCR genotyping. DNA from each transformant was amplified in PCR reactions to determine the 5' and 3' insertion positions. Primer sets with a primer beyond the transformation platforms and a primer within the transformation cassette were used to determine the insertion positions. Only samples in which both corresponding PCR reactions successfully amplified a product were used for further experiments.

3.2.10 Semi-quantitative reverse transcription PCR

Total RNA was extracted from wild-type cultures (gametophore stages) and stable transformants, using the TRIzol™ extraction method [289]. 1µg of DNase I (NEB) treated total RNA was used to synthesis cDNA following the ProtoScript® II Reverse Transcriptase protocol (NEB). PCR was performed with DNA Taq polymerase (GeneDireX®) to examine knock-in constructs and related genes' expression levels.

3.2.11 Biotin labeling and processing

To perform the BioID screen, individual colonies of stable transformants were isolated and transferred to plates containing liquid BCD medium supplemented with 50 µM to 5 mM biotin (Bioshop® Life Science Products) for 24 hours at two different temperatures, room temperature and 37°C with agitation. Plants were then collected and washed twice with

ice-cold phosphate-buffered saline (PBS; 8 mM Na2HPO4, two mM NaH2PO4, 150 mM NaCl) or Tris-buffered saline (TBS; 150 mM Tris, 150 mM NaCl, pH 7.2-7.4) to remove the excess biotin and flash-frozen in liquid nitrogen for further analysis. The frozen powder was resuspended in freshly prepared ice-cold lysis buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 0.5% Triton-X 100, 10 mM sodium orthovanadate, ten mM NaF, 2µM pepstatin, 1mM PMSF, one ug/mL aprotinin and ten mM β-glycerophosphate), supplemented with a cocktail of protease inhibitors (Sigma®). Cells were lysed on ice for 20 min. Lysates were then centrifuged at 14,000 rpm for 15 min at 4°C, and the supernatant was transferred to the fresh tubes. Proteins from whole-cell lysates were separated on 10% SDS-PAGE (29:1 acrylamide/bisacrylamide; Bio Shop). Before loading samples, samples were mixed with protein loading dye (0.25 M Tris-HCl, pH 6.8, 15% β-mercaptoethanol, 30% glycerol, 7% SDS, 0.3% bromphenol blue) to a 1:1 ratio, and denatured at 95°C for five minutes, vortexed, and centrifuged at 12,000 rpm for one minute at room temperature. Proteins were separated electrophoretically on a 10% SDS-PAGE in electrophoresis buffer (25 mM Tris-base, 250 mM glycine, 0.1% SDS, 0.34% EDTA) at 100 V for at least 10 min, followed by separation at 200 V until sufficient separation. The SDS gel consisted of a running gel (375 mM Tris, pH 8.8, 12% acrylamide/bisacrylamide 37.5:1 (BioRad), 2 mM EDTA, 0.1% SDS) and a stacking gel (125 mM Tris, pH 6.8, 5% acrylamide/bisacrylamide (37.5:1), 2 mM EDTA, 0.1% SDS), details are summarized in Table 3.3. The PageRuler™ Protein Ladder (#26616, Thermo Fisher Scientific) was used to reference molecular weight estimation. Proteins were visualized using Coomassie blue stain (45% (v/v) methanol, 10% (v/v) acetic acid, 0.25% (w/v) Coomassie Brilliant Blue R-250) and de-stained in 30% (v/v) methanol with 10% (v/v) acetic acid.

SDS-PAGE gels were transferred by electroblotting onto a nitrocellulose membrane (Hybond-C extra, GE Healthcare) using the semi-dry method. I used Trans-Blot®SD Semi-Dry Electrophoteric transfer Cell by Bio-Rad for the semi-dry technique and followed the company's manual instructions. To visualize total protein levels, membranes were stained with Ponceau Red (0.2% PonceauS, 3% trichloroacetic acid) for at least 5 min, washed, and photographed. Staining was removed by washing in TBS. The membrane was blocked for

an hour at room temperature in a blocking solution (TBS, 0.5% Tween-20, and 2% BSA, bovine serum albumin). The blocked membrane was then incubated in 1:2000 dilution of streptavidin-HRP (horseradish peroxidase) (#3999, Cell Signaling Technology®) in 2% BSA and TBST (TBS; 150 mM Tris, 150 mM NaCl, pH 7.2-7.4, 0.1% Tween) overnight at 4°C. Following removing the streptavidin-HRP solution, the membrane was washed three times for 5 minutes each in TBST.

The anti-myc blot was also prepared parallel and washed three times with TBST before the addition of rabbit anti-myc (1:1000; 2278S Cell Signalling) in 2% BSA TBST overnight at 4°C, then washed three times, five minutes per wash, in TBST before the addition of anti-rabbit HRP (1:5000; 7074S Cell Signalling) in 2% BSA and TBST for 1 hour at room temperature. For detecting all blotted proteins, chemiluminescent signals were detected by adding Pierce™ ECL Western Blotting Substrate (#32106, Thermo Scientific™), and specific protein bands were detected with the C-DiGit® membrane scanner using the Image Studio Digits, the imaging software.

The whole-cell extract also was mixed was incubated with 10 µL Streptavidin (Sepharose® Bead Conjugate) (#3419, Cell Signaling Technology®) with rotation overnight at 4°C. The following day the beads were collected with the centrifuge, and the supernatant was kept as the unbound sample. The beads were then resuspended in 500 µl of 1X cell lysis buffer and rotated to wash the beads at room temperature, and then the supernatant was discarded. The wash was repeated five times, and once the supernatant was completely clear. Beads were suspended in SDS-PAGE sample buffer and heated for 5 minutes at 95°C to elute bound proteins. Eluate was analyzed by immunoblotting, same as it was described above.

Table 3.3: The mixture used for SDS-PAGE gels

Solution	10 Resolving gel	4% Stacking gel
H_2O	4 mL	10 mL
30% bis-acrylamide (37.5:1)	5.85 mL	2.5 mL
1.5 M Tris (pH)	7.35 mL	-
1 M Tris (pH)	-	1 mL
10% SDS	175 µL	140 µL
10% APS	50 µL	75 µL
TEMED	4 µL	4 µL

3.3 Results

3.3.1 Establishment of the knock-in systems

Proximity is the key to the biotinylation process. Therefore, it was vital for the promiscuous BirA enzyme to be active on the N-terminal of CESAs. For this reason, the BirA tag was cloned downstream of the N-terminus of partial-length *CESA3*. I used the pUC19 plasmid to express bait proteins fused to BirA* with a c-Myc tag. The c-Myc tag allowed for a western blotting-based assessment of protein expression levels [290]. BirA and Venus genes were fused to the N-terminus of the *CESA3* gene to produce BioID-CESA3 and Venus-CESA3 fusion proteins, targeting the plasma membrane's protein. In these constructs, NPTII, the selection marker, was placed in the *CESA3* non-coding region. We've also created a construct called KnockOut-CESA3, which only contains NPTII in the *CESA3* non-coding region (Figure 3.4).

During splicing, introns will be removed from the mRNA [291], although, in some cases, mRNAs with very long introns, only part of an intron is removed, and then the remaining intron may stay in the mRNA [292, 293]. So, to avoid this, since NPTII is a large fragment, we've created a system for auto-excision of the selectable marker gene, NPTII, via a stress-inducible FLP-FRT site-specific recombination system [294].

3.3.2 Excision of the selectable marker

FLP-FRT recombination, a site-directed recombination technology, is used to manipulate DNA under controlled conditions *in vivo*, which involves the recombination of sequences

between short flippase recognition target (FRT) sites by the recombinase flippase (FLP). The 34bp FRT site sequence has 5'GAAGTTCCTATTCtctagaaaGtATAGGAACTTC3' sequence for which flippase (FLP) binds to both 5'-GAAGTTCCTATTC-3' arms flanking the eight basepair spacer, for example, the site-specific recombination (for a crossover) in reverse orientation. FRT-mediated cleavage happens just forward from the asymmetric 8bp 5'tctagaaa3' core region on the top and bottom strands. Despite several variant FRT sites, recombination usually occurs between two identical FRTs but typically not among non-identical (heterospecific) FRTs [295].

The generated system, which utilized the Flp recombinase from yeast to excise the NPTII gene in situ [285], uses the site-specific DNA recombinase FLP [286] to selectively induce recombinase activity with the addition of estradiol or one of its analogs to excise the NPTII selection marker, which is flanked by FRT sites [285] (Figure 3.5). The estrogen receptor ETR2 is a mutated form of the human estrogen receptor (ER) ligand-binding domain (LBD), which is inducible by the synthetic anti-estrogen 4-hydroxytamoxifen but does not respond to the natural ligand estrogen [286].

In the chemical-inducible FLP DNA recombination system, removing the NPTII-selectable marker gene was confirmed by detecting PCR fragment FLP-FRT in transgenic moss. The same primers as it was used to amplify the FLP recombinase gene from pCAG-FlpeERT2 were used. Zeocin-resistant transgenic moss that had been transformed with FLP-FRT-pTZ-UBIgate was used to induce FLP-FRT.

Initially, the transformed moss with FLP-FRT-pTZ-UBIgate was selected on media supplemented with Zeocin, and stable transformants were subjected for the second round of transformation BioID-CESA3-pUC19, Venus-CESA3-pUC19, KO-CESA3-pUC19.

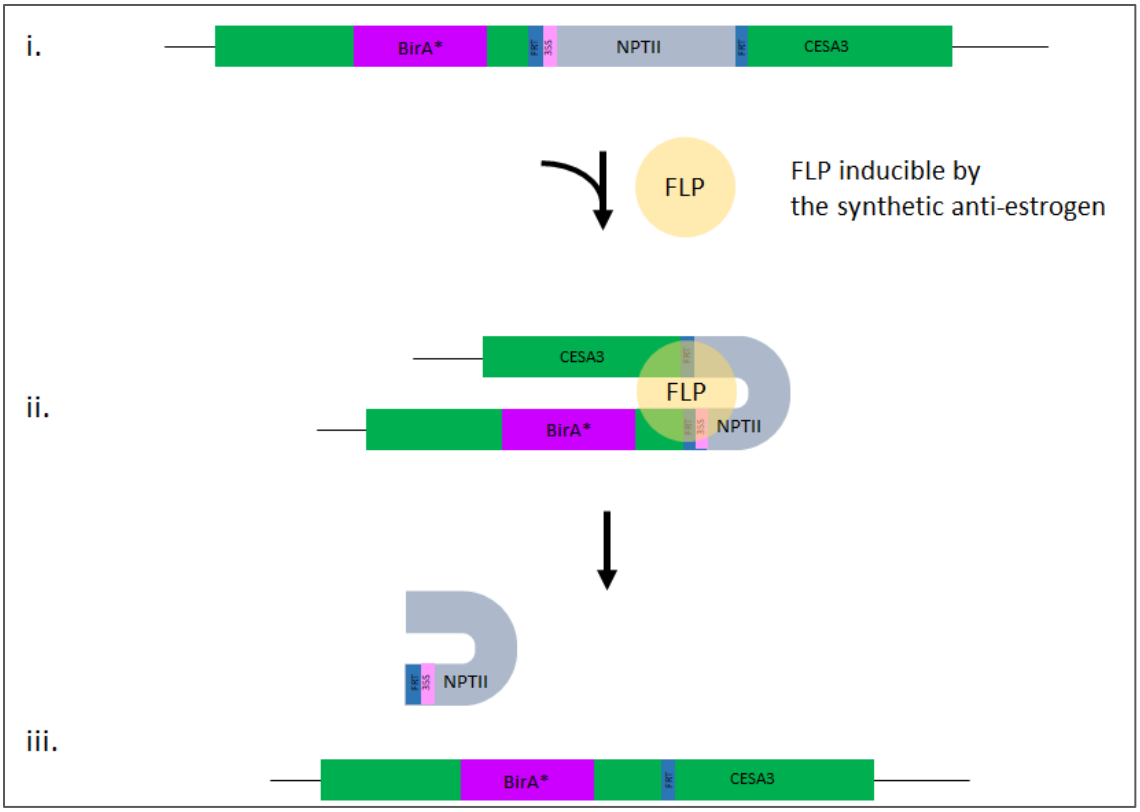


Figure 3.5: Gene excision of NPTII, using FLP recombinase.

3.3.3 Establishment of the *CESA3* knock-in systems

A total of 5-15 transgenic mosses resistant to both Zeocin and NPTII were generated. Most transgenic plants showed a normal morphological phenotype compared to the wild-type (Figure 3.8).

I carried out PCR analysis with primer pairs (Table 3.1) to detect the recombination events in both constructs on every single colony. Based on the results of the PCR analysis, ten plants in total (BioID-CESA3: 5 plants, Venus-CESA3: 2 plants, and KO-CESA3: 3 plants) were characterized, which also contain the FLP-FRT site and therefore they putatively contain two transgenic loci which one locus would have undergone excision of the NPTII fragment. To elicit Flp recombinase activity, gametophores were incubated in 1 µM 4-hydroxytamoxifen solution [286] for 16-18 hours and then cultured in BCDAT. A piece of gametophore tissue was cultured in BCDAT with antibiotics as a control to monitor for lack of growth on antibiotic supplemented media. Plants were observed to be resistant to

Zeocin and sensitive to NPTII and showed amplification of BioID-CESA3 but 1463 bp shorter than the original locus, as is shown in Figure 3.7.

Morphological analysis of the ten putative marker-free transgenic mosses showed a normal phenotype compared to the wild-type (Figure 3.8) and was then subjected to biotinylation assay following western blot analysis.

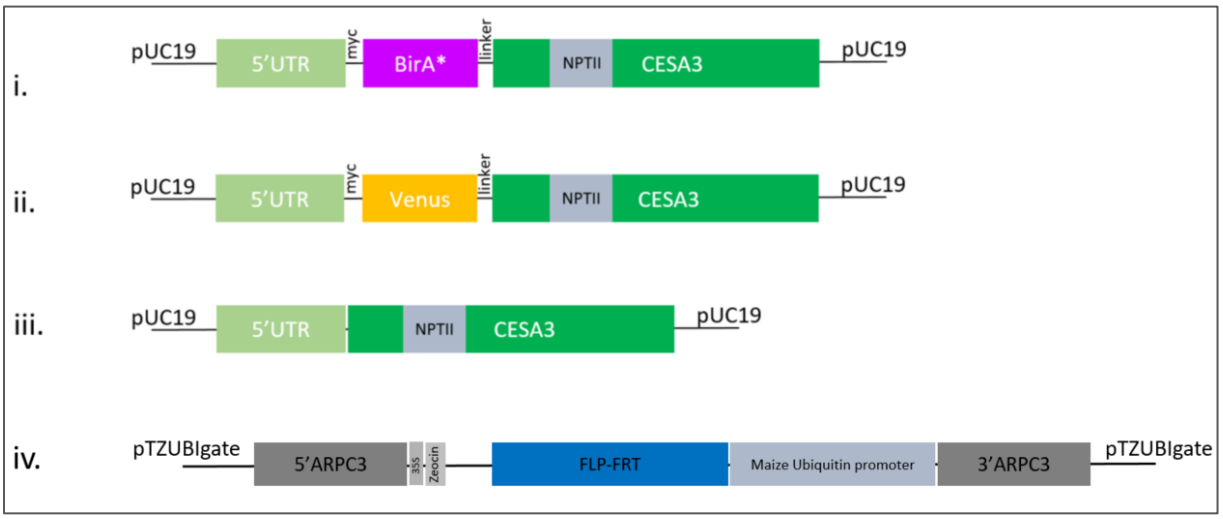


Figure 3.6: CESA3 fusion cunstructs (i) BioID-CESA3-pUC19, (ii) Venus-CESA3-pUC19, (iii) KO-CESA3-pUC19. FLP-FRT-pTZUBIgate construct.

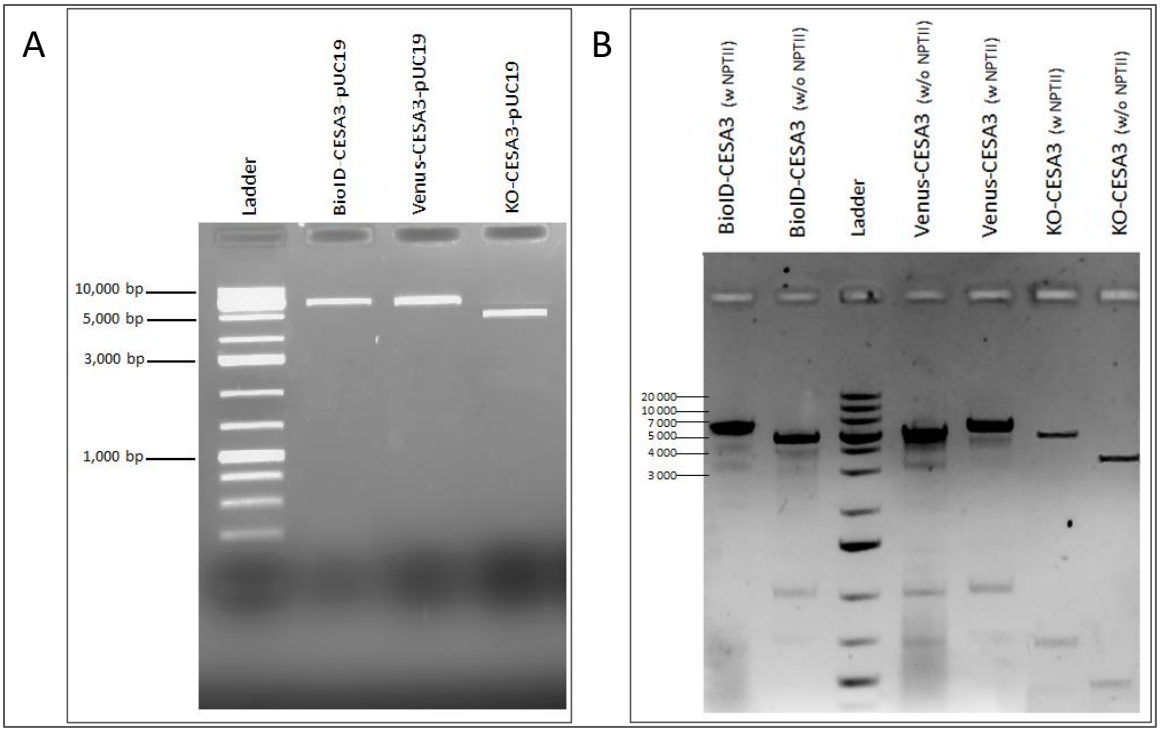


Figure 3.7: Confirmation and visualization of successful DNA assembly of CESA3 constructs. (A) CESA3 constructs cloned into pUC19 and linearized by a restriction enzyme (B) Confirmation of NPTII gene excision in CESA3 constructs.

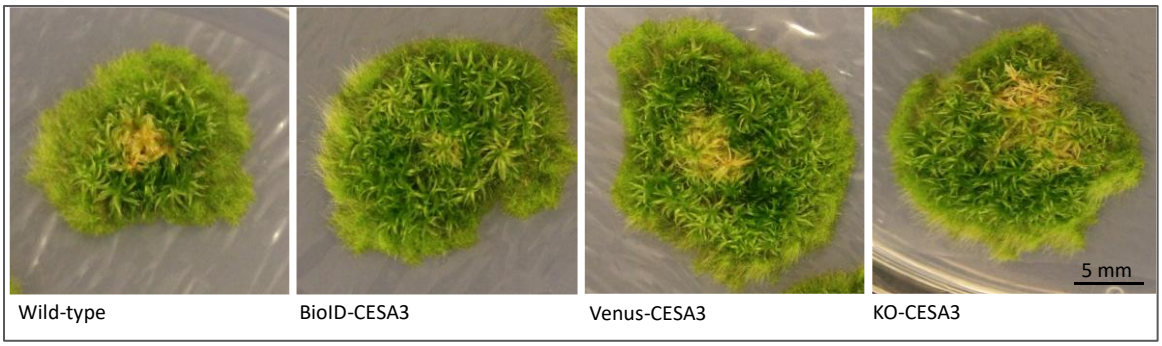


Figure 3.8: Comparison of BioID-CESA3, Venus-CESA3, and KO-CESA3 transformants phenotype with wild-type plant.

3.3.4 Biotin identification (BioID)

I generated the BioID method for labeling proteins in a proximity-dependent manner in the *Physcomitrium* cell wall. I planned to create a system based on the fusion of *PpCESA3* and *PpCESA5* proteins to an enzyme that could selectively modify vicinal proteins *in vivo*. There are two requirements for this system. The first and the most important is that the

fusion protein must be correctly targeted when expressed in cells. Next, the modification itself must help the isolation of the specifically labeled proteins.

BirA* is a 35-kD mutated DNA-binding biotin protein ligase (R118G) in *E. coli*, which is defective in self-association and DNA binding [296]. A biotin acceptor tag (BAT) is coexpressed with BirA* that is fused to a protein of interest in this system. The BAT is biotinylated and can detect with one-step high-affinity avidin/streptavidin mediated purification of the tagged protein. Since biotinylation is a rare modification in euyoukariotic cells, it is restricted to a few carboxylases [268]. In the two-step process of BirA* biotinylation, firstly, biotin and ATP form a highly reactive biotinoyl-5'-AMP [263] (bioAMP [297]). The free bioAMP reacts with primary amines [263]. BirA* biotinylates proteins promiscuously in a proximity-dependent manner [298, 299].

I generated multiple fusion proteins for expression in *Physcomitrium patens*. Plant tissue contains some natively biotinylated proteins, which can be detected by western blotting using streptavidin-HRP. Expression of BirA* alone in a plant cell must result in labeling some proteins within the cytoplasm. The expression of BirA in moss tissue produced minimal labeling without exogenous biotin substrate (Figure 3.9A). BirA* has been shown to self-biotinylate and tagged with the myc-tag; therefore, the biotinylated 35 kDa species in the BioID-CESA3 were expected to be observed on the anti-myc western blot to indicate the biotinylated fusion proteins [298]. Western blot analysis using antibodies against myc tags for BioID-CESA3 was performed to confirm the expression of the BioID in moss transformants (Figure 3.9B). A band was observed at approximately 35 kDa at the blot, highlighted in Figure 3.9B with an arrow.

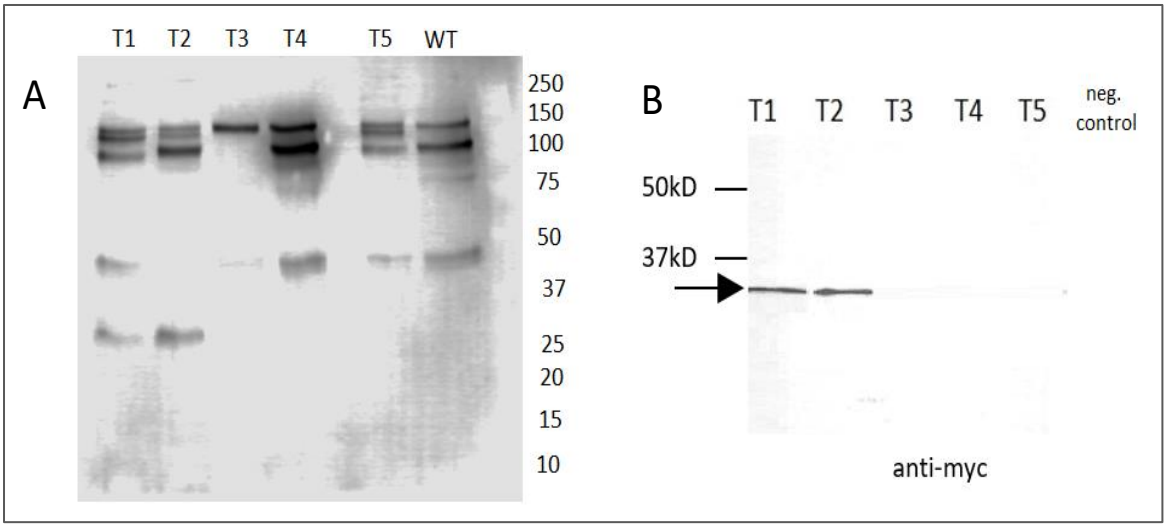


Figure 3.9: Western blot analysis of BioID transformants. (A) Plants with BioID-CESA3 fused protein show minimal biotin labeling without exogenous biotin substrate. Lane 1-5 represents five independent transformants (T1-T5) compared to the wild-type (WT). (B) anti-myc western blot shows in lanes 1-5 five independent transformants (T1-T5) compared to the WT.

3.3.5 Incorporation and metabolism of biotin

The next important step for establishing the BioID-CESA system was testing the incorporation and metabolism of biotin. Live BioID-CESA3 plants were treated and incubated with biotin according to the protocol published by Roux and colleagues [263]. Figure 3.10 demonstrates the results of this test. In the presence of exogenous biotin substrate in the growth media, the BirA fusion constructs produced significantly less labeling than expected to observe compared to wild-type (Figure 3.10). There was no differential banding pattern for the BioID-CESA3 samples as seen with wild-type (Figure 3.10); however, the bands are more predominant after incubation of BioID-CESA3 samples biotin, and only a few bands were observed in the absence of biotin (lane with – sign), a smear was seen when biotin was added (lanes with + sign). The same experiments were performed with a control cell line expressing wild-type. In these cases, the same bands were observed (Figure 3.10). I could not detect bands corresponding to each of the full-length fusion proteins' predicted molecular weights in lysates from transiently expressing knock-in mutants. The fusion protein expression was supposed to be around 100 kDa, higher than the endogenous *CESA3* expression, approximately 58 kDa. But the expression of the fusion proteins wasn't detected.

To overcome this issue, I tested various biotin concentrations, incubation temperature, and duration of incubation. A range of biotin concentrations (50 µM to 2 mM) for 18, 24, and 48 hours were used at room temperature (approximately 25°C), 37°C, and 50°, following by cell lysis and western blotting analysis using Streptavidin-HRP (Figure 3.11). Despite testing different conditions, no biotinylated protein labeling other than the minimal labeling with exogenous biotin substrate was detected in knock-in moss stains. Also, I didn't observe any differences between the various conditions mentioned above.

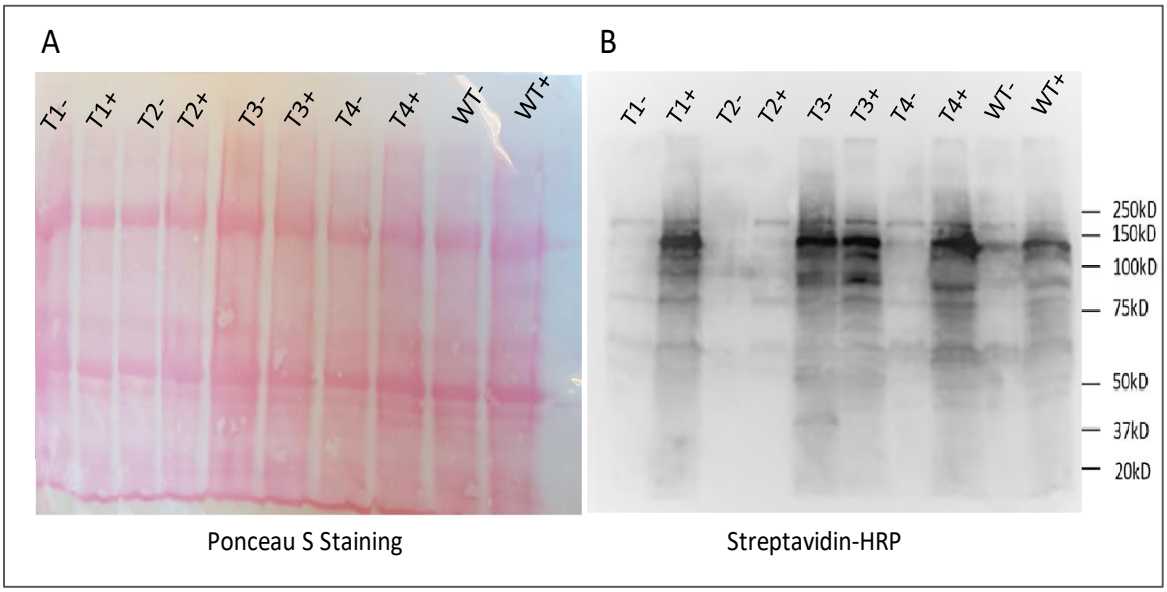


Figure 3.10: Western blot analysis for the biotin ligase activity. (A) Ponceau S Staining verified equal loading of protein. (B) Western blot result for biotinylation for *P. patens* lines derived from the transformation of BioID-CESA3 in five stable transformants (T1-T5) and wild-type (WT) as control. All stable transformants were grown in liquid BCD media with (+) or without (-) biotin final concentration of 5 mM, added 24 hours before cell lysis.

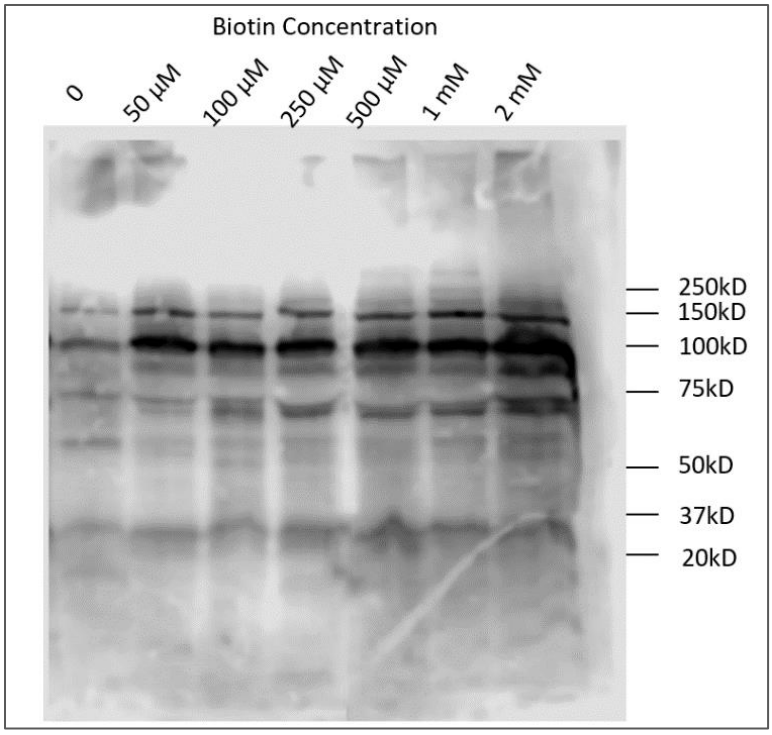


Figure 3.11: Western blot analysis of a range of 50 µM to 2 mM biotin concentrations at 37°C for 18 hrs for the transformants 1 (T1) sample.

3.3.6 Fusion proteins' expression levels measurement

I performed semi-quantitative PCR on leafy gametophores isolated from BioID-CESA3 plants to measure fusion proteins' expression levels in these tissues. The results show that BioID-CESA3 is less expressed compare to *PpCESA3* (Figure 3.12). And it appears the BioID-CESA3 fusion may significantly downregulate *CESA3* gene expression compared with the wild-type. Next, I set up the BioID2 system, using the overexpression vector pTZ-UBIgate with full-length *CESA5* to improve the biotinylation process.

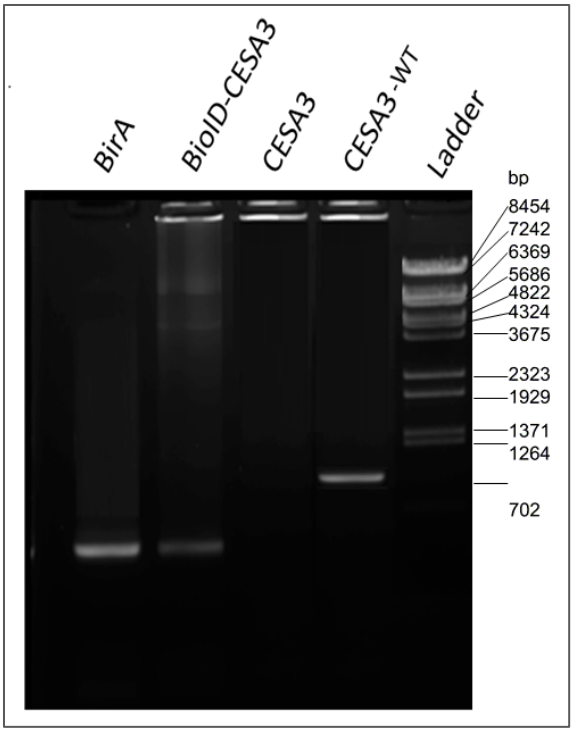


Figure 3.12: Semi-quantitative PCR result for BioID-CESA3 fusion protein. The expression level of *PpCESA3* after the BirA gene is fused to N-terminal.

3.3.7 Generation of new fusion protein BioID2-CESA5

Given the data presented in the previous section, I hypothesized that BirA* triggers a conformational change that may downregulate the expression of *CESA3*, which disrupts the interactions. To evaluate the study hypothesis and identify the CESAs interacting proteins, I used the BioID2 approach to identify PpCESA5-interacting proteins.

For this purpose, I generate the fusion protein using CESA5. As described previously, any interruption in the gene expression and/or protein modification of *PpCESA5* can alter the plant phenotype due to the role of *PpCESA5* in morphogenesis and development of leafy gametophore [77]. On the other hand, BioID2 is an improved version of the BioID platform, where a smaller promiscuous biotin ligase is fused in-frame to the protein of interest to facilitate the biotinylation of proximal proteins in live cells [271]. BioID2 uses the *Aquifex aeolicus* biotin ligase, which a mutation Arg40Gly in the catalytic domain, increases the promiscuity of biotinylation [271].

For this study, wild-type *CESA5* was cloned in the frame of the c-terminus of the BioID2 biotin ligase, and the construct, BioID2-CESA5, was cloned to overexpression vector pTZ-UBIgate (Figure 3.13A) and transformed into wild-type *Physcomitrium patens* to target the protein in the plasma membrane. I didn't observe any phenotype changes in putative transformant compared to wild-type to ensure that these fusion constructs retained native-like phenotype (Figure 3.13B).

These results from analyzing the putative transformants showed that BioID2-CESA5 fusion protein failed to incorporate into the *Physcomitrium patens* genome, so I could not detect any proximity labeling in-frame fusion CESA5.

Despite not detecting any expression of fusion proteins, I tested the activity of the biotin ligase. Firstly, proteins from transformants and wild-type were resolved by SDS–PAGE, and myc tag levels were tested with the anti-myc antibody. But no BioID2 expression was detected. Next, treating cells with growth medium supplemented with additional biotin (2 mM), the biotinylation assay for CESA5 proximal protein resulted in a weak pattern of biotinylated proteins of a low range of molecular weights as detected by probing western blots with streptavidin HRP (Figure 3.14), which correspondent to the naturally biotinylated proteins. Also, no expression of BioID2-CESA5 fusion protein was detected.

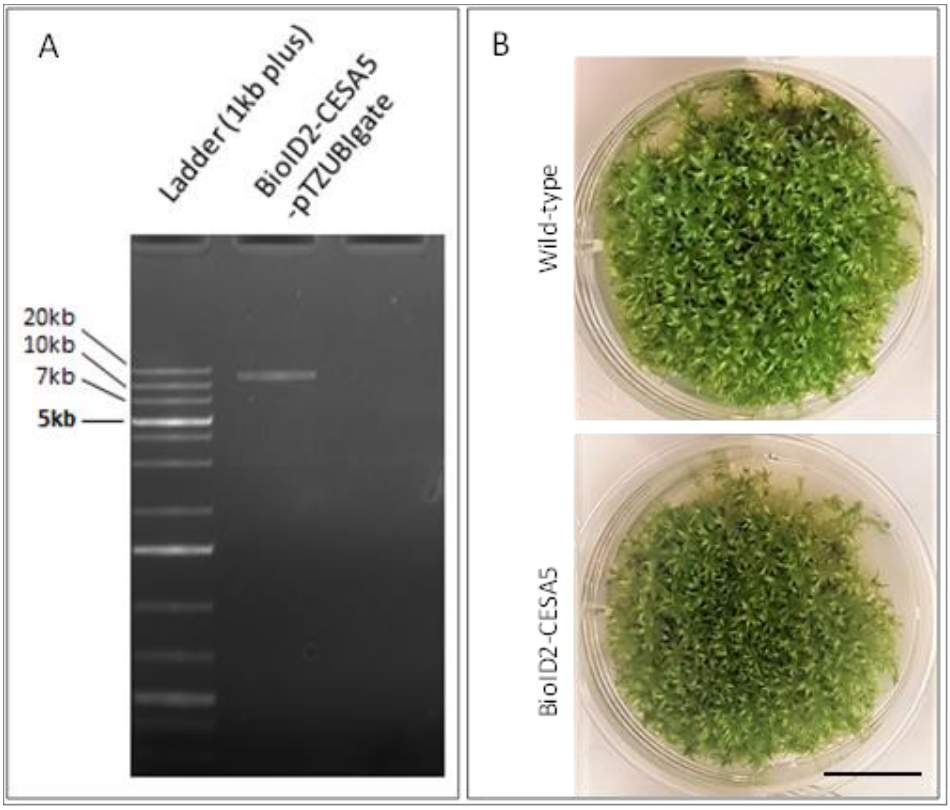


Figure 3.13: BioID2-CESA5-pTZ-UBIgate construct. (A) Confirmation of the construct assembly. (B) compare BioID2-CESA5 transformant phenotype to wild-type (Scale bar: 0.5 cm).

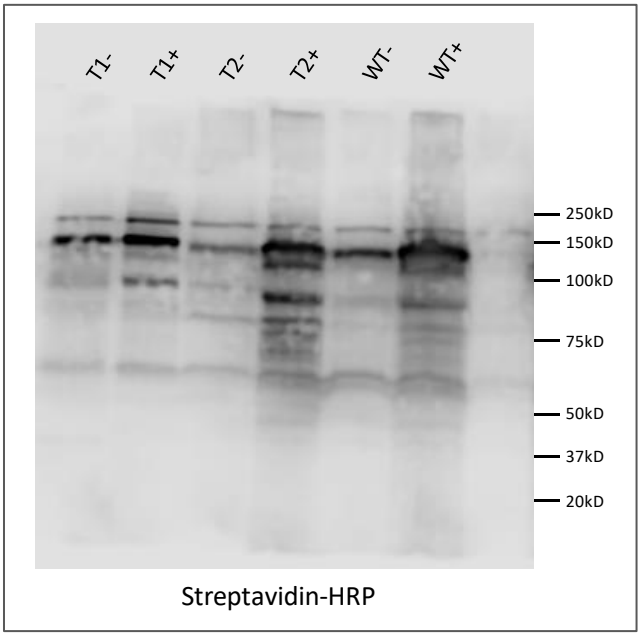


Figure 3.14: The biotin ligase activity of BioID2-CESA5 transgenic plants. An initial test for the biotinylation of two stable transformants (T1-T2) and wild-type (WT) as control. Both transformants were grown in liquid BCD media with (+) or without (-) biotin to a final concentration of 5 mM, added 24 hours before cell lysis, and incubated at 37°C for 18 hours.

3.4 Conclusion and future directions

Methods such as affinity purification coupled with mass spectrometry (AP-MS) have identified very few plant CESA interacting partners. This is because AP-MS fails to identify proteins that interact weakly or transiently with a target protein. In the past decade, enzyme-catalyzed proximity labeling approaches have been developed to overcome some of these drawbacks [263, 300, 301]. Proximity labeling methods are based on engineered ascorbate peroxidase (APEX) [302] and a mutant *E. coli* biotin ligase $BirA^{R118G}$ (BioID) [263]. The technique called proximity biotinylation helps many researchers to identify the proteins that partner together. It relies on attaching an enzyme (BioID [267] or TurboID [303]) to a protein of interest; when an interacting protein comes in close contact with this construct, it can attach biotin, and tagged proteins can then be identified, revealing the interacting molecules with the protein of interest. In contrast to AP-MS, in proximity labeling, the labeling of proteins is performed in living cells in the native cellular environment. Although, in proximity biotinylation methods, longer incubation time with biotin (16-24 h) and higher incubation temperature (37 °C) is required for efficient labeling [263], which can be toxic and lethal to some cell types. Even though a smaller biotin ligase, BioID2, from *Aquifex aeolicus* has been described, conditions required for biotinylation of interacting proteins are similar to those of BioID [271]. The necessary conditions for BioID and BioID2 are not optimal for *in vivo* proximity labeling in the whole plant, which has resulted in the limited deployment of BioID-based proximity labeling in plants [290].

Other limitations, such as the large size of BirA* tag (35 kDa) tag in the BioID system, could change the structure and function of the fused protein and cause potentially wrong location or reduced binding [263], which I believe could be the case for BioID-CESA3 fusion protein since there was no expression in the fused CESA3. Another possibility could be the problem with the protein tags that can have diverse effects depending on the c- or n-terminal ends that they are attached to. Based on the previous studies, fused proteins can have different behaviors when tagged at the carboxyl or amino terminus, and the expression level can be decreased to zero compared to the wild-type with no effect on the growth and development of plants [273].

Recent studies show that the *E. coli* biotin ligase BirA generated new promiscuous labeling variants TurboID and miniTurboID, which allow rapid nontoxic proximity labeling. TurboID is more active than BioID and can label proteins within ten minutes. However, it can be more likely to be toxic for the cell under certain conditions. It also makes mistakes and tags proteins that do not interact with the protein of interest [303].

To address these issues, recently, Kido et al. established a new technique called AirID, a new enzyme for proximity biotinylation. Experiments were confirmed that AirID could label interacting proteins in human cells. It is also less probable to mistakenly tag non-partners or to kill the cells, even over long periods, or affect the expression level compared with BioID and TurboID [304].

Chapter four: General discussion

This study aimed to investigate the cellulose biosynthesis-related processes in *Physcomitrium patens* to extend our current understanding of the proteins involved in these processes. The choice of *Physcomitrium patens* was governed by its unique experimental attributes. Over the past two decades, genetic studies on *P. patens* have expanded the available tools and resources and laid the foundation for robust forward and reverse genetic approaches. Since 1962, *P. patens* Gransden was established as a model species by H.L.K. Whitehouse. Between the 1960s to the 1980s, this moss was used to isolate and study mutants in plant morphology, hormone biology, and nutrition [96, 305-307]. Early studies revealed that the effects of phytohormones [92, 308, 309] were conserved throughout the evolutionary history of land plants [310, 311]. The efficient gene targeting via homologous recombination, which triggers the immediate generation of mutants in *P. patens*, was reported in 1997 [312, 313], following the initial transcriptomic characterization [314], established this moss as a functional genomics model [90, 315, 316]. *P. patens* was targeted for genome sequencing and designed to become a cell biology model and evolutionary developmental studies [317].

P. patens protoplasts transformation using PEG allows for the stable incorporation of DNA and transient expression assays [49, 318]. *P. patens* whole plants protoplasts can be derived from a single protoplast carrying the inserted DNA molecules into specific sites in the genome. The generation of gene deletions and insertions in *P. patens* has allowed functional gene analysis to be performed throughout development. Also, by transforming haploid protoplasts due to the predominant haploid state of *P. patens*, it is possible to isolate knock-outs and knock-ins of essential genes. Researchers have also employed gene silencing mediated by RNA interference (RNAi) [318, 319]. Transient expression of RNAi constructs is a powerful approach to studying essential genes [320] involved in development. Despite the effectiveness of gene silencing approaches, they can suffer from variable levels of silencing. Recent advances in genome editing technologies such as CRISPR-Cas9 overcome the limitations of traditional homologous recombination and gene silencing approaches [317]. Moreover, the stable integration of multiple genes occurs at

high frequency, and thus, it obtains stable transformants containing functional genes [321].

In forward genetic approaches to isolate morphological mutants in *P. patens*, ultraviolet light (UV) or chemical mutagens such as ethyl methanesulfonate treatment are used [306]. The availability of a well-annotated genome and marker lines in ecotypes facilitates the identification of crossed sporophytes [322, 323] and causal mutations [73]. Similar to the research presented in this thesis, mutagenized protoplasts with UV showed conditional loss of growth mutants [324]. Also, Moody et al. in 2018 identified UV-mutagenized mutants that were not able to form gametophores [325]. However, without gametophores to obtain a segregating population, it would not be possible to cross the mutant strain to a different ecotype. Somatic diploidization is the solution to perform a crossing between two somatic diploids [325]. The causal mutations in both groups can be successfully identified using whole-genome sequencing approaches on *P. patens* derived from their mapping population.

Physcomitrium patens with extensive annotation and short-generation time have great potential to shed light on fascinating topics, such as genomic and transcriptomic studies. These different aspects will significantly benefit from the increasing growth in data and further experiments in the field. Special attention will be paid to population genomic approaches to study individuals from single moss populations, as such studies could provide valuable insight into the effects of selfing and the mechanisms related to genome evolution in the Funariaceae.

Appendix one: Creation of an alternative biosynthetic vehicle for the production of rare cannabinoids in *Physcomitrium patens*

A1.1 Introduction

Plants can produce a variety of complex secondary metabolites in large quantities. Naturally, plants use these small organic molecules to handle many types of stress, and many time humans benefit from the biological activities of these metabolites, which make them of high commercial interest to the biotechnology industry [326]. Previous attempts have been made to produce high-value compounds more economically available by introducing their biosynthesis pathways from natural source plants into other plants, such as the moss *Physcomitrium patens* as a host production organism [327]. Another motivation for pursuing this endeavor is that some medicinally important plants are endangered species [328]. Moreover, some plants have been difficult to cultivate in high amounts, which happened to be the case with thapsigargin to treat multiple cancers [328-330]. At the time, thapsigargin was obtained from the intact fruits and roots of *Thapsia garganica* L. plants, and the increasing demand for thapsigargin could lead to plant extinction. These plants are difficult to germinate from seed and maintain in a greenhouse [328, 331], and there is no report of the successful *in vitro* regeneration of *T. garganica* [332]. Chemical synthesis is another possible method to obtain thapsigargin. Though, in 2000, it wasn't commercially feasible [333].

Besides, the natural source of desirable compounds might not be reliable or consistent. For example, terpenoids are some of the most active compounds from plants with antitumor activity, like Taxol (paclitaxel), extracted from yews [334]. The high demand for the drug has led to pressure on wild plants as a drug source. As a result, I cannot only rely on the wild population for the supply of drugs.

Therefore, there are efforts to identify the active compounds before their natural sources disappear or become compromised. *Escherichia coli* and *Saccharomyces cerevisiae* (yeast) are well established as the conventional workhorses of biotechnology for plant terpenoid production and have advantages in sustainability and similarity to plant systems [335].

Yeast is a good host in most simple pathways because of its eukaryotic membrane structure [335]; however, active molecule identification is more difficult in a complex pathway. For complex pathways, it is essential to understand the biosynthesis pathway of the target molecule fully. Although biosynthetic pathways from another kingdom also can create problems such as in isoprenoid production [336, 337]. These problems can be avoided by using host organisms like the moss, *Physcomitrium patens*. Recent studies have shown that *P. patens* is an excellent host for drug production. For example, artemisinin, the potent malaria drug from the plant *Artemisia anua*, can be produced at higher levels in *Physcomitrium* even before any production enhancements [327]. *P. patens* has also been used as a host for the biosynthesis of high-value products, such as α/β-santalene [338], patchoulol [149], sclareol [141], and recombinant therapeutic peptides [339].

Physcomitrium, as a multicellular land plant, may not be able to compete with yeast in terms of growth rate. However, efficient gene disruption in *Physcomitrium* is one of the most significant benefits in reverse genetics research in plants [340]. Also, what sets *P. patens* apart from other plant heterologous expression systems is its genome editing's unique capacity through homologous recombination matches the *in vivo* assembly of multiple heterologous DNA fragments in *P. patens* [341]. *P. patens* is a non-vascular ancient lineage of land plants, is the right choice for the stable heterologous production of terpenes. It is mostly because of its ability to undergo homologous recombination and allow direct genome editing without the requirement of CRISPR-systems or known genome sequence [341]. Besides, *Physcomitrium* can grow photo-autotrophically in sterile inorganic culture media without any need for vitamins, phytohormones, or a carbon source [305]. In most plant production systems, liquid cultures are typically composed of undifferentiated callus cells. However, *Physcomitrium* cultures can be propagated as haploid, photosynthetic protonemata [145]. The advantage is that a protonemal culture has a stable genetic background to ensure that the genetic composition of *Physcomitrium* cultures does not change over time [342]. Moreover, *Physcomitrium* is tolerant of abiotic stresses [343], which can help the non-native compound biosynthesis in moss, as mentioned earlier.

As previously stated, an important advantage of *Physcomitrium* is that it is amenable to combinatorial biosynthesis, which can be defined as the application of genetic engineering for the modification of biosynthetic pathways for the production of novel compounds [151, 344, 345].

An *in vivo* combinatorial biosynthesis approach in *Physcomitrium* to make natural products with useful biological activity has already been done [346]. In this context, the important outcome of using *Physcomitrium* is that establishing a moss strain producing cannabinoids will help create other drug-candidate-producing moss strains.

Cannabis has been used for medical purposes around the world for millennia [347]. As Canada has implemented medical and recreational cannabis policies, people using cannabis as a therapy for pain management are increasing [348]. Recent reports are conclusive evidence that cannabis is effective for chronic pain treatment in adults [349]. Cannabinoids, like some other drugs, are challenging to obtain by chemical synthesis, and *Cannabis sativa*, the original plant, cannot produce sufficient amounts for market demand [145].

In the study by Ikram et al. in 2017, the genes involved in artemisinin biosynthesis were engineered into the *P. patens* via direct *in vivo* assembly of multiple DNA fragments. The initial production was relatively high, and the study showed *P. patens* could be an efficient and sustainable industrial-scale production platform of artemisinin and supply of artemisinin with lower price and more affordable [346]. Based on the successful and efficient *in vivo* assembly of multiple linear DNA in *P. patens* for amorphadiene and artemisinin production [149, 346], I develop a concatenation assembly system for cannabinoids imitating some basic cannabinoids' natural production.

A1.2 Methods

A1.2.1 Plant material and growth conditions

The 'Villersexel' strain of *Physcomitrium patens* (Vx) was used for this research, grown, and propagated as described in previous chapters.

A1.2.2 DNA constructs and plant transformation

Targeted loci included multiple-gene insertions of five key enzymes in the cannabinoid biosynthetic pathway; cannabidiolic acid synthase (CBDAS), olivetol synthase (OLS), acyl-activating enzyme 1 (AAE1), TPA-exp-prenyltransferase 4 (PT4), olivetolic acid cyclase (OAC), tetrahydrocannabinolic acid synthase (THCA) [341] precursor using CDS clones. The cannabinoid biosynthesis pathway is shown in figure A1.1. Each synthesized CDS clone (sequences are shown in A1.5.1) was introduced to an overexpression vector pTZ-UBIgate, under the strong constitutively expressed promotor's control, maize ubiquitin promoter with Zeocin as a resistant cassette in plants. The CDSs were prepared for transformation by using Gateway® cloning kit (Thermo Fisher Scientific), based on the manufacturer's specification, as described in chapter three. The successfulness of gateway cloning was assessed by running the linearized constructs on 1% agarose gel and compared to the original empty vector (Figure A1.2A). These constructs carried 1100-1500 bp of terminal homology corresponding to the moss ARPC3 gene's redundant copy. All the construct's maps are generated by SnapGene® viewer and shown in A1.5.2.

Transformation of *Physcomitrium* followed established protocols by performing direct uptake of DNA in PEG presence (polyethylene glycol 6000), which is described in chapter three. Linear DNA was used for each transformation and obtained by MluI restriction enzyme digestion. The DNA was purified by the CTAB extraction method. Lastly, the DNA was resuspended in 30 µl ddH_2O. Following DNA transformation, protoplasts were embedded in an agar overlaying protoplast regeneration medium and incubated for seven days, by which time protoplasts had regenerated. These were then transferred to a selection medium containing Zeocin (50 mg/l) for 10–14 days. The selection protocol involved serial replica picking of regenerating colonies on to medium without Zeocin for 14 days, followed by the second period of selection for 14 days in the antibiotic's presence, to confirm the recovery of stably transformed plants.

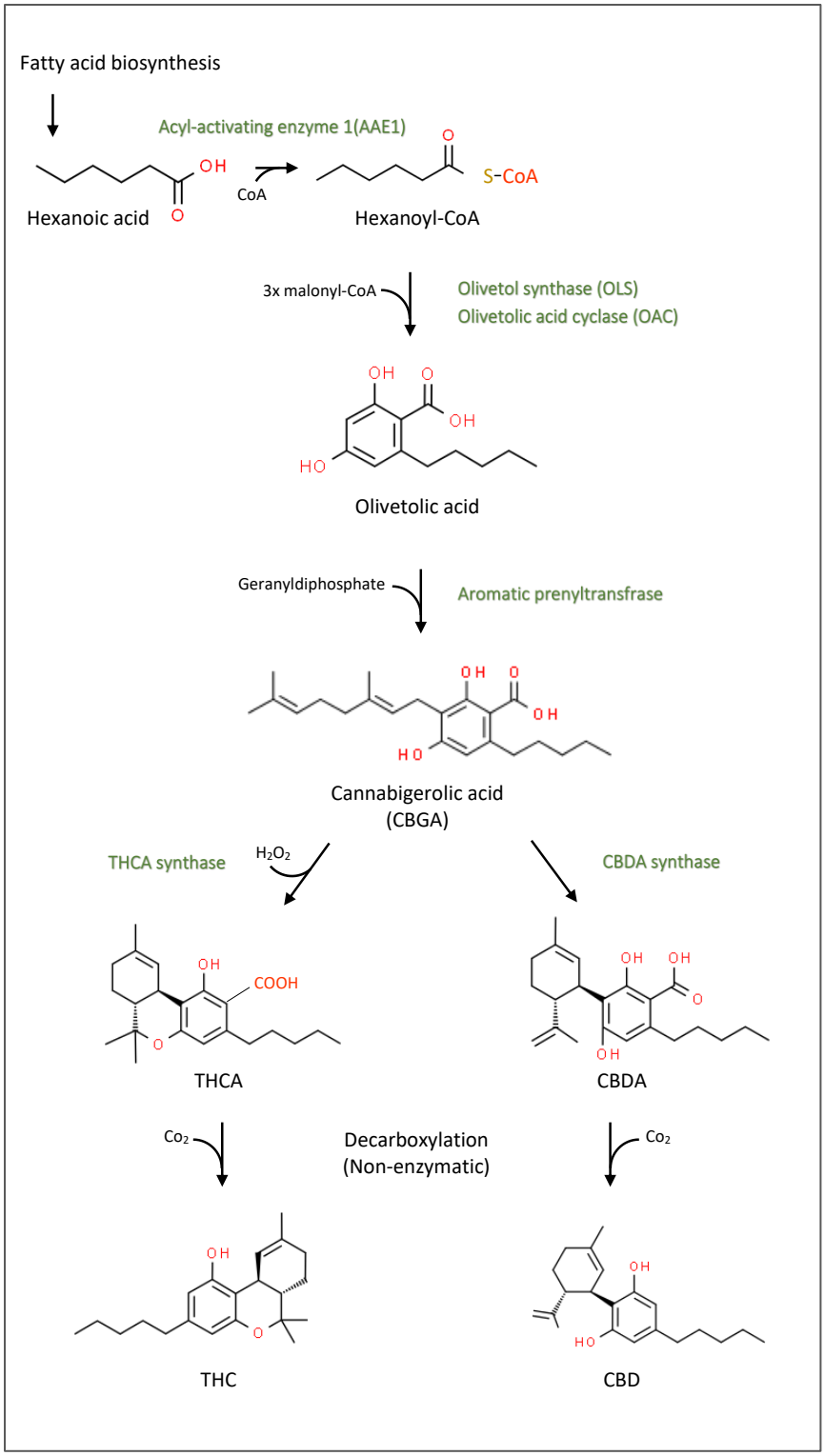


Figure A1.1: Overview of the cannabinoid biosynthetic pathway leads to the main cannabinoids. THCA and CBDA, and the non-enzymatic decarboxylation to THC and CBD (Adapted from [350]). Enzymes are shown in green. Chemical structures are from www.chemspider.com.

A1.2.3 Molecular analysis of transformed plants

PCR-based assays were used to analyze large numbers of transformed plants for the presence or absence of gene targeting events. DNA was isolated from the small protonemal inocula of transgenic mosses using the small-scale DNA isolation procedure and the CTAB DNA extraction method described in previous chapters [174]. Primers were designed, allowing amplification with gene-specific primers that annealed with the inserted genes' sequence and maize ubiquitin promoter inside the homology region incorporated in the targeting vector (Primer listed in Table A1.1).

Table A1.1: Primer list for genotyping the cannabinoid genes constructs in the pTZ-UBIgate backbone plasmid

Primer name	Sequence
OAC-F	ATGGCCGTGAAGCACTTGAT
OAC-R	GTGTAGTCGAAGATCAACAA
THCAS-F	ATCCAGGCCACCATCTTGTG
THCAS-R	TTCCGTCCACGTTCACCAAG
PT4-F	ATCTCCTCCCAGTCCAGGTC
PT4-R	TGTTGATCCTGTCGATGTCC
AAE1-F	GTTGAACGGAAACCACCACG
AAE1-R	GGTGGTGTCGTTGGAGTCC
OLS-F	ATCGGAGGACACATCAGGGA
OLS-R	GTACTTGATGGGCACGGACC
CBDAS-F	GGACAGAACGGAGCCTTCAA
CBDAS-R	TTCTTGGACACGTAGGGGGT
MaizeUBIPromotor-F	CAAATAAATAGCGTATGAAGGC
MaizeUBIPromotor-R	CTGGTGGATTTATTAATTTTGGATC

A1.2.4 Preparation of plant extracts

To test the whole plant for the production of cannabinoids, the plant was dried in a ventilated oven overnight at 40°C and then powdered by grinding. 50 mg of each sample was weighed into a centrifuge tube, and one mL methanol (MeOH) solution was added, vortexed, and incubated at room temperature for 20min. The mixture was centrifuged for 5min at 1200g. The procedure was repeated four times by combining all the supernatants, and the volume was brought to 5 mL using the solvent. 500 µL of each sample extract was taken into a separate gas chromatography (GC) vial and was injected for gas

chromatography-flame ionization detector (GC-FID) analysis [351]. All the experiments were done in triplicate.

A1.2.5 Standards and reagents

Three cannabinoids were used as reference standards in this study: CBG, CBDA, and THCA at a concentration of 1.0mg/mL. The purities of all the standards were above 99%. The solvent was MeOH of HPLC analytical grade and was purchased from Sigma-Aldrich.

A1.2.6 GC-FID analysis of samples

GC-FID analysis was performed on a Varian 3900 gas chromatograph (Varian, USA) using a Restek Rtx-5 column (10 m x 0.10 mm ID, 0.10 µm) for separation. Helium was used as a carrier gas at a flow rate of 1.0 ml/min and the FID make-up gas. The injector temperature was 240 °C; detector temperature, 270 °C; oven program, 170 °C (hold 1 min) to 250 °C at 10 °C/min (hold 3 min); and run time was 12 min with 1 µl injection volume. Data analysis was performed using the Galaxie Chromatography Data System software (Version 1.9.3.2).

A1.3 Results

A1.3.1 Multiple gene-knock-ins by a single transformation

Homologous recombination gene targeting results in exact gene replacement and targeted insertion. Gene targeting happens following batch transformation using multiple DNA constructs. It is also possible to do simultaneous targeting of multiple genes by cotransformation with batches of independent targeting constructs [148]. Then I analyzed a detailed molecular analysis of transgene integration after delivering multiple constructs [50, 352]. Effective simultaneous delivery of multiple DNA transforming has been used before to create a large population of mutant plants by transformation with a single CDS clone.

Batch transformations are used to generate mutant collections of *Physcomitrium* by targeted gene-knock-ins via homologous recombination and performed with batches of six CDSs. In these experiments, the DNA used for transformation was linearized by restriction

enzyme digestion MluI, and CDSs in the vector backbone were delivered to protoplasts. The batches were compiled by mixing equal amounts of the different CDS constructs before the transformation. Total DNA concentration was 20-100 µg per transformation, and experiments were repeated seven times. As a control, single-gene transformations were performed for OAC-pTZ-UBIgate in parallel to compare the gene-targeting frequency in single gene transformations and batch transformations. In all, 24 plants were evaluated by PCR for targeted knock-in events in single gene transformations of OAC-pTZ-UBIgate. The single-gene transformation was showed approximately 12 transformants per transformation.

Of the 135 plants derived from batch transformation with mixes of six different CDS constructs and analyzed for knock-in events by PCR, 93 independent transformants were recovered, and genotyping showed they carried targeted knock-ins integrated into the genome. PCR genotyping (Figure A1.2B) showed the integration of each construct as; OAC-pTZ-UBIgate (95.7%), THCAS-pTZ-UBIgate (44.3%), PT4-pTZ-UBIgate (21.4%), AAE1-pTZ-UBIgate (37.1%), OLS-pTZ-UBIgate (92.8%), and CBDAS-pTZ-UBIgate (82.8%). These ratios are similar to those determined for the single-gene transformation, indicating that the different CDS fragments do not interfere with each other's homologous integration. Assuming that CDSs in a batch transformation can create individually targeted knock-ins, and multiple knock-ins should occur. Indeed, 12 plants were generated by transformation with the mix of six different CDSs, showing that independent gene-targeting of the different CDSs happens, and multiple knock-ins are possible using CDS batches.

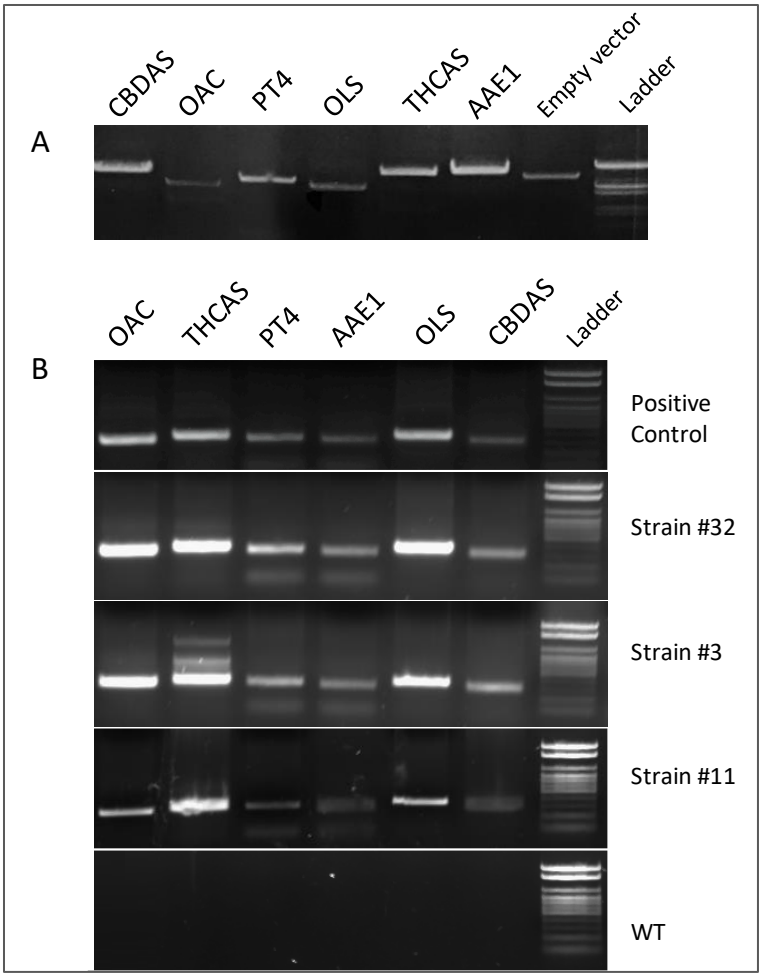


Figure A1.2: Confirmation and visualization of the successful cloning of cannabinoid constructs. (A) CDS constructs. The six CDS were cloned into the pTZUBIgate vector, and the successful insertion was compared to the vector with no insertion. Since some of the CDSs were smaller than the region of cloning sites in the vector, the construct's length appeared smaller than the empty vector. The DNA ladder used is a λDNA-BstEII digest. (B) Genotyping of new moss strains shows they have multiple gene insertions. They are compared to the positive control (fragments in positive control amplified from the constructs) and wild-type (WT) as a negative control. PCR fragments are amplified with gene-specific primers and show partial length CDS. The DNA ladder used is a 100bp DNA Ladder.

When *Physcomitrium* is transformed with targeting constructs, the coincidental integration of vector sequence was confirmed by PCR genotyping of structures that included CDS and vector sequences. The gene targeting is detected by PCR using primers specific to the promoter combined with gene-specific primers corresponding to the inserted genes (Primers listed in table A1.1). Amplification with both pairs confirmed accurate homologous recombination at each end of the construct. In some plants, amplify a fragment with both pairs was failed, which indicates transgene insertion was at an

illegitimate site. The accuracy of the homologous recombination events can be established by sequence analysis of the amplified fragments. I initially performed PCR-screens in 93 transgenic lines transformed with single to multiple insertions.

In *Physcomitrium*, the transformation was done with 20-120 µg DNA/transformation. The efficiency of transformation increased with the quantity of DNA, reaching an optimum at 100 µg, but the amount of DNA per transformation did not affect the number of integrated transgenes per plant. In contrast, in mammalian cells, the amount of DNA delivered per transformation can impact the copy number of integrated transgenes [353].

A1.3.2 Analysis of cannabinoids in *Physcomitrium patens*

Terpenoids are the largest class of chemicals among plants' many compounds. Plants employ terpenoid metabolites for various basic functions in growth and development, but most terpenoids are used to protect the abiotic and biotic environment. For decades, plant-based terpenoids have been used in the pharmaceutical and chemical industries and biofuel products [354]. Many terpenoids are detected in both bryophytes and seed plants, although some are unique to bryophytes [355]. The chemical analysis identified common monoterpenes, sesquiterpenes, and diterpenes in mosses [355]. Terpenoids in *Cannabis sativa* L. are divided into several groups. Cannabinoids are derived from the diterpene structure [356].

GC-FID is the most common technique for the analysis of cannabinoids. It is simple, fast, and sensitive for the determination of the total cannabinoids. GC-FID quantitative analysis of cannabinoids in illicit cannabis products (Marijuana, Sinsemilla, Thai sticks, Ditchweed, Hashish, and Hash Oil) has been reported previously, and samples of cannabis preparations were examined, emphasizing the levels of seven cannabinoids (D9-THC, D8-THC, CBD, CBC, CBG, CBN, and CBL) using a validated GC-FID method [357].

For this study, extracts of the transgenic *P. patens* lines containing 1-6 genes of cannabinoid biosynthesis pathway were analyzed using GC-FID. As I was expecting, few peaks appeared in GC-FID analysis for wild-type moss, present in all the transgenic mosses. Although 12 plants were carrying all the six genes, cannabinoid traces in two *P. patens*

transgenic lines were detected in extracts but not in the extracts of wild-type *P. patens* (Figure A1.3). Peaks compared to cannabidiolic acid (CBDA), tetrahydrocannabinol acid (THCA), and cannabigerol (CBG) derivatives. THCA, CBDA, and CBG traces in the transgenic *P. patens* extracts are identical to the standard controls. The analysis was performed twice within triplicates each time. The presence of cannabinoids in *P. patens* transgenic lines was confirmed by comparing a standard (Figure A1.4-A1.6).

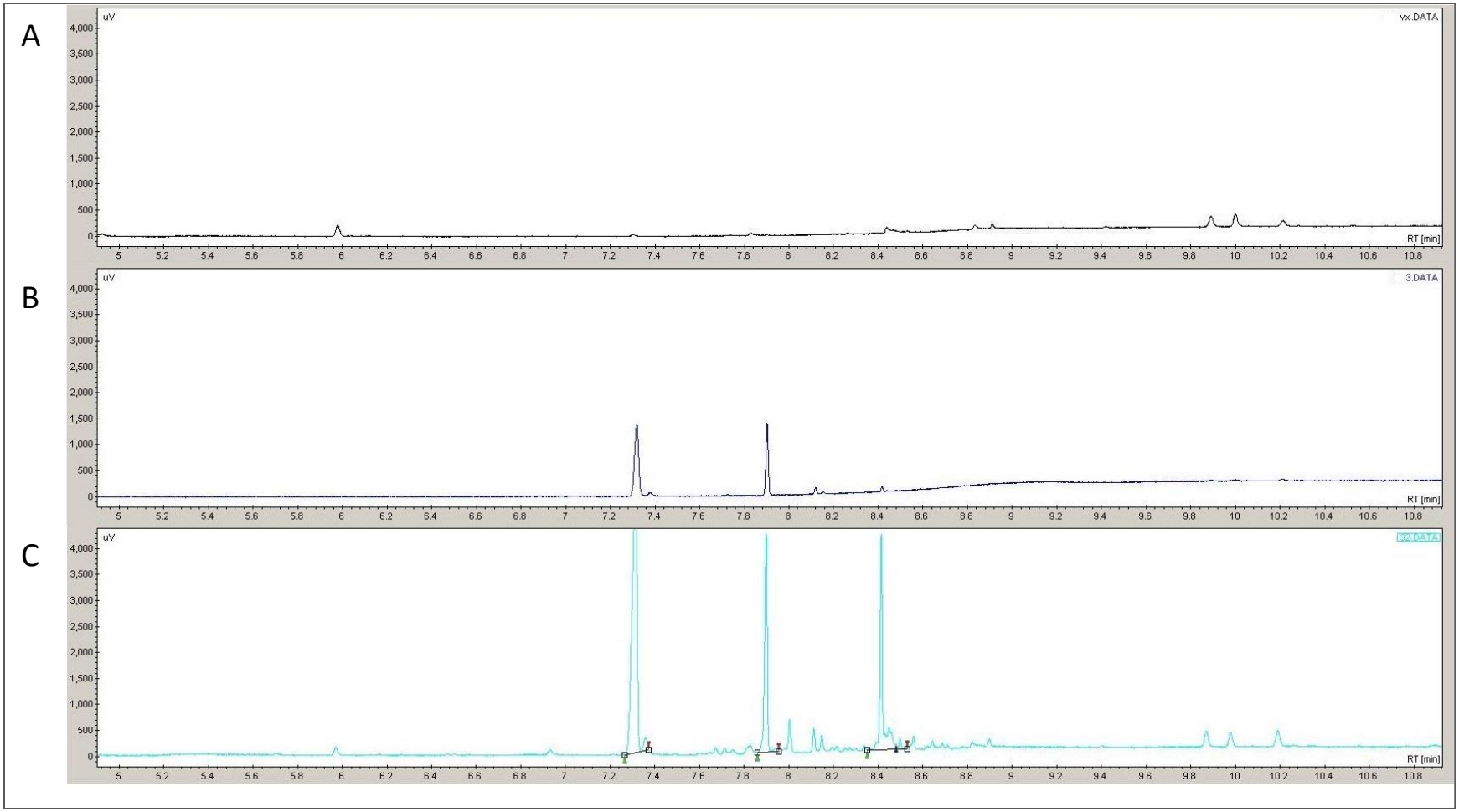


Figure A1.3: GC-FID analysis of cannabinoid produced from transgenic *Physcomitrium patens* and wild-type (WT) as control with retention time. (A) Chromatogram showing the standard peaks for terpenoids in moss. (B) Gas chromatogram of the transgenic line #3. There are two large peaks between the retention times of 7.3-8 minutes not present in wild-type. (C) Gas chromatogram of the transgenic line #32. There are multiple peaks between the retention times of 6.8-9 minutes which are not present in wild-type.

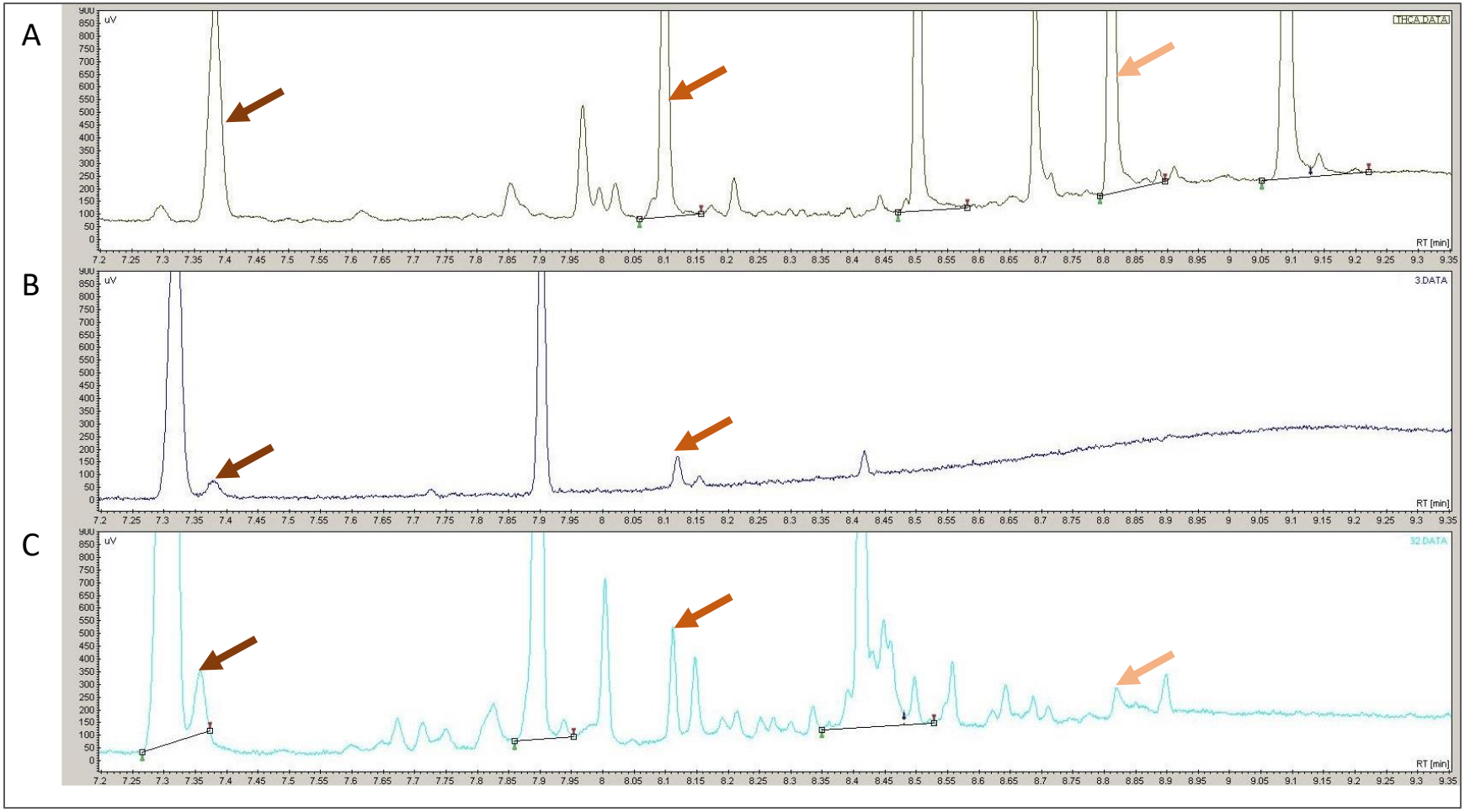


Figure A1.4: GC-FID analysis of cannabinoid produced from transgenic *Physcomitrium patens* and THCA reference sample with retention time. (A) Chromatogram showing the standard peaks for THCA derivatives in a reference sample. (B) Gas chromatogram of the transgenic line #3. There are two peaks at the retention times of 7.37 and 8.12 minutes aligned with the reference sample. (C) Gas chromatogram of the transgenic line #32. There are three peaks at the retention times of 7.37, 8.12, and 8.83 minutes and multiple small peaks aligned with the reference sample. (Arrow with the same color corresponds to the same peak in different samples)

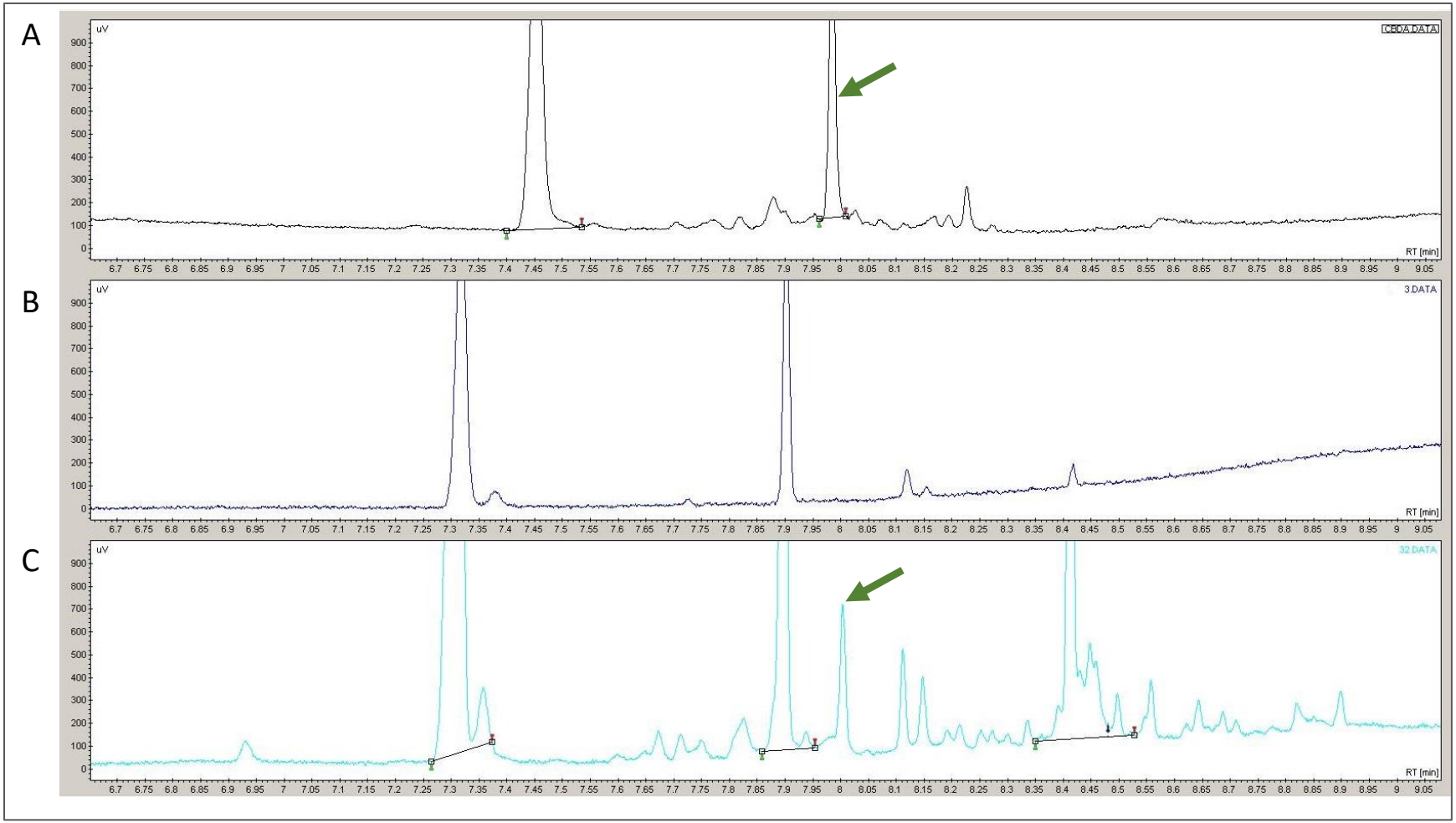


Figure A1.5: GC-FID analysis of cannabinoid produced from transgenic *Physcomitrium patens* and CBDA reference sample with retention time. (A) Chromatogram showing the standard peaks for CBDA derivatives in a reference sample. (B) Gas chromatogram of the transgenic line #3. No peak matches the reference sample. (C) Gas chromatogram of the transgenic line #32. One peak at the retention times of 8 minutes aligned with the reference sample. Arrow with the same color corresponds to the same peak in different samples)

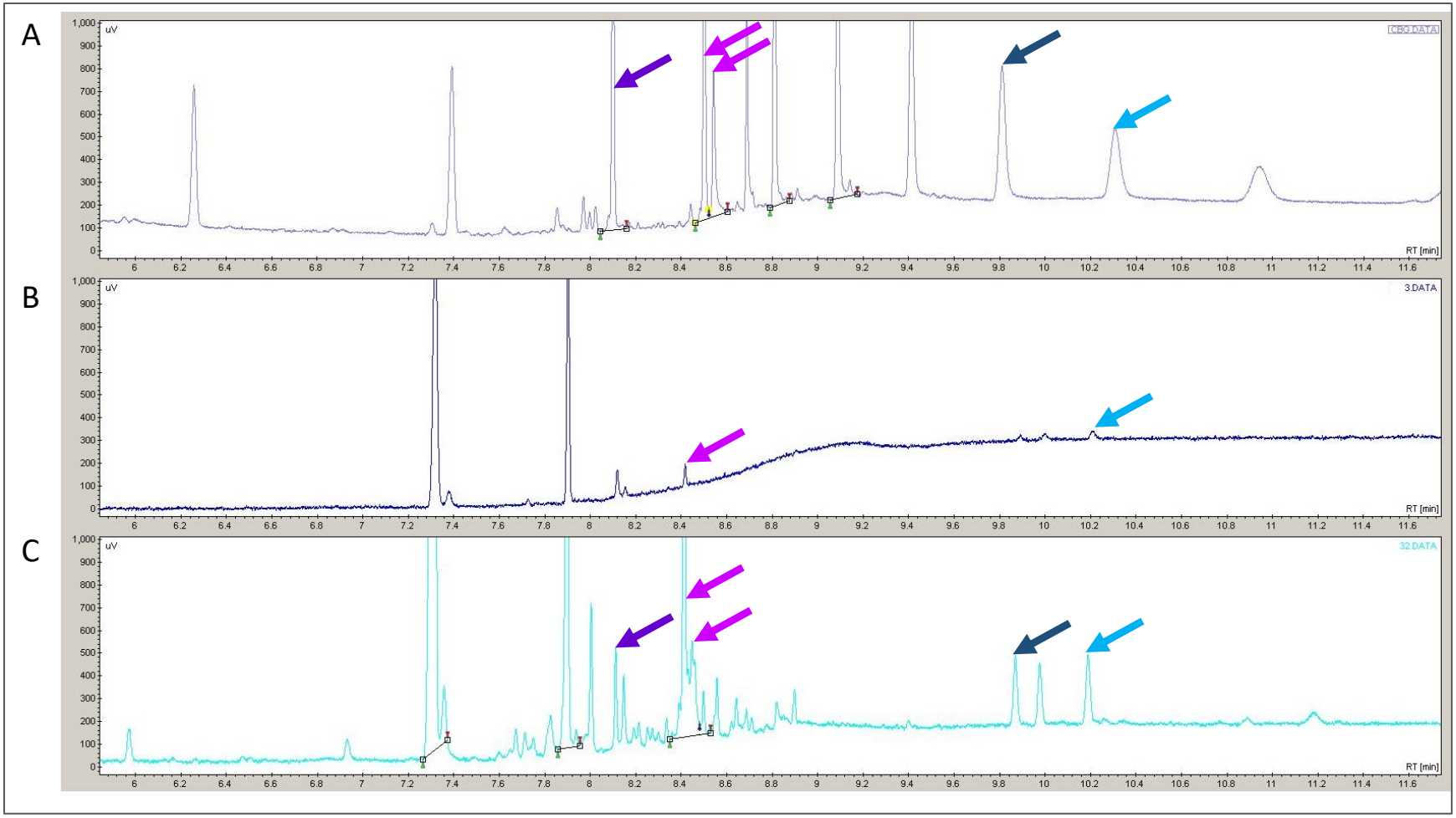


Figure A1.6: GC-FID analysis of cannabinoid produced from transgenic *Physcomitrium patens* and CBG reference sample with retention time. (A) Chromatogram showing the standard peaks for CBGA derivatives in a reference sample. (B) Gas chromatogram of the transgenic line #3. There are two small peaks at the retention times of 8.4 and 10.2 minutes, which align with the reference sample. (C) Gas chromatogram of the transgenic line #32. There are multiple peaks at the retention times of 8.1, 8.4, 9.9, and 10.2 minutes, which align with the reference sample. (Arrow with the same color corresponds to the same peak in different samples)

A1.3.3 Phenotypic analysis of *Physcomitrium patens* transgenic lines

I investigated the impact of the cannabinoid derivatives by creating a *Physcomitrium patens* transgenics where the cannabinoid biosynthesis genes are knocked in by homologous recombination. Studying the transgenic plant phenotype is one technique used to achieve this goal. Both *Physcomitrium patens* wild-type and the transgenic lines were cultivated on BCD media, and a stereomicroscope was used to characterize the difference in the phenotype between the wild-type and the selected transgenic plants (Figure A1.7).

There is a reduction in the growth rate in some of the transgenic lines (Figure A1.7). This indicates that there might be a disruption of other metabolic pathways responsible for *P. patens* growth and development. The effect of cannabinoid biosynthesis on *P. patens* growth is not much, and it might be due to the toxicity of the product and nutrients' reduction in the media, which requires further studies. This production of cannabinoids is precious for future industrial production.

From Figure A1.7, it can be observed that wild-type and transgenic plants have gametophores. However, there are some differences between strain #3 and wild-type plants. The transgenic line #3 has delayed in producing gametophyte and shows a dwarf phenotype. The first apparent difference is the size of the phyllid. The wild-type plant and strain #32 contain gametophores with large phyllids, while the strain #3 plant has gametophores with smaller phyllids. The wild-type phyllids are more extensive and broader, while the transgenic plant #3 phyllids are smaller than the wild-type (Figure A1.7). Some were turning brown quickly and presumably dying. Some transgenic lines displayed a defective phenotype, such as line #3 with smaller gametophores or no gametophore, which suggested that over accumulating some of the derivatives or one of the bioproduct could cause this, or they might have more than one copy number of the genes. Different expression profiles of the enzymes can be one other possibility of the severely stunted phenotype. On the contrary, some transgenic lines have a normal phenotype compared to wild-type with the same number of the gene inserted (Figure A1.9) compared to defective stains, which needs to be looked at in more detail in future studies.

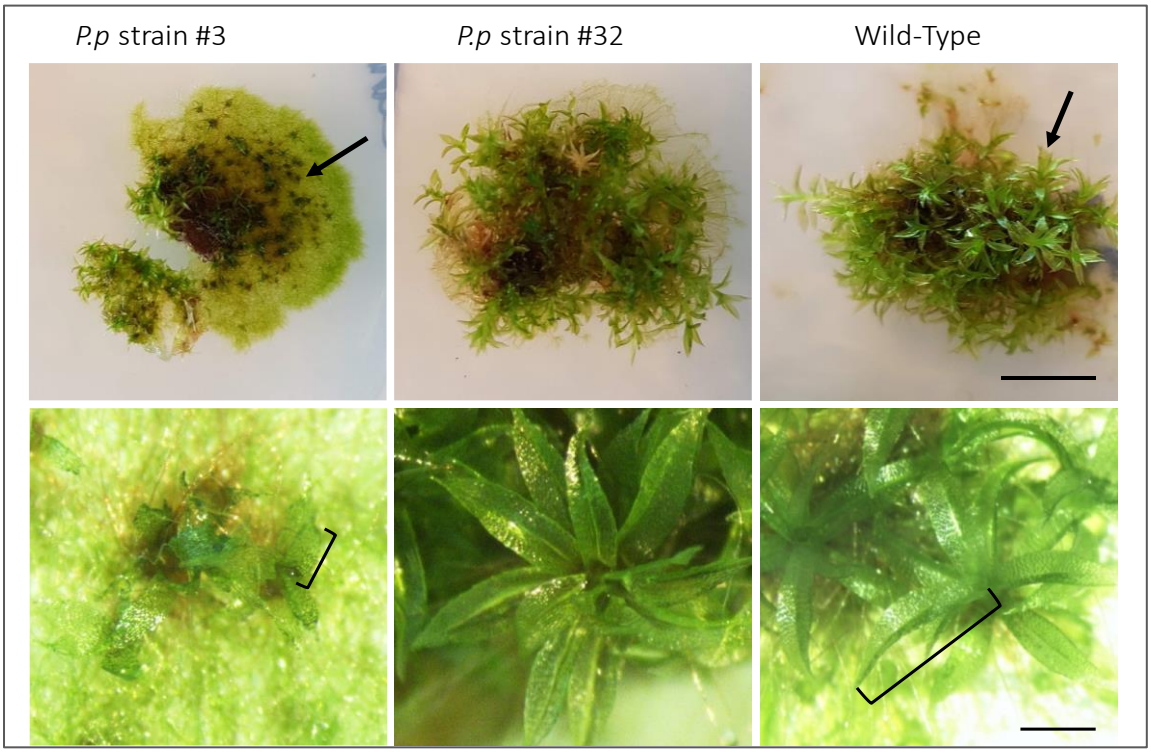


Figure A1.7: Comparison of wild-type and transgenic line #3 and #32 phenotypes. All plants were cultivated on BCD media for eight weeks. Line #3 on the left panel shows a dwarf phenotype with smaller phyllids, while line #32, shown in the middle, has more extensive and broader phyllids, similar to wild-type on the right. Phyllids are indicated (Scale bars: 0.5 cm (top) and 0.2 µm (bottom).

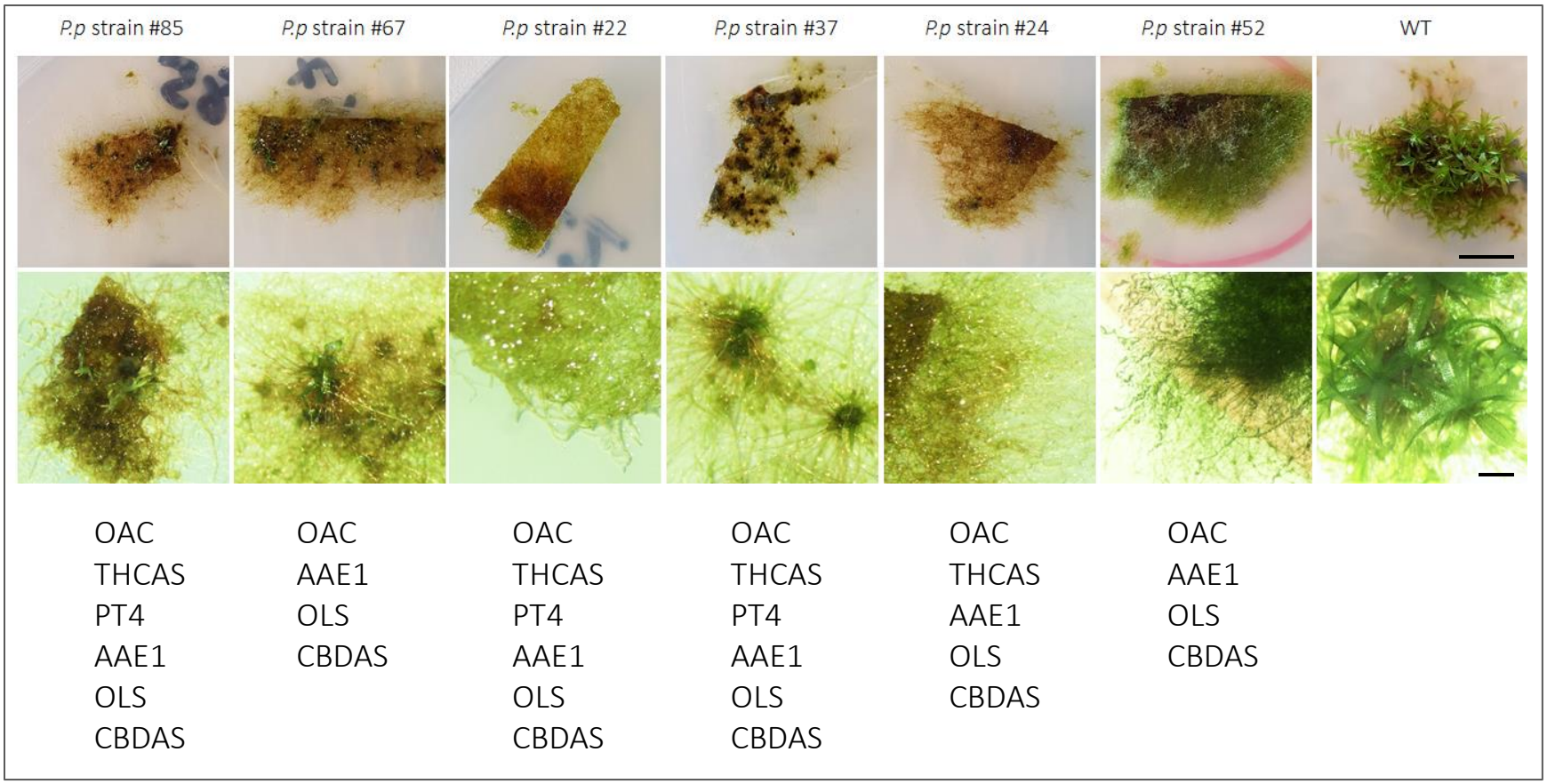


Figure A1.8: Defective phenotype of transgenic lines compared to wild-type (WT). All plants were cultivated on BCD media for eight weeks. Some transgenic lines have turned brown (#85, #67, #22, and #24) quickly, #37 has shown undifferentiated gametophore tissue, and some lines such as #52 are not producing gametophore tissue (same plants with two different magnifications are shown). The list of inserted genes is written under each picture of the transgenic line. (Scale bars: 0.5 cm (top) and 0.2 µm (bottom).

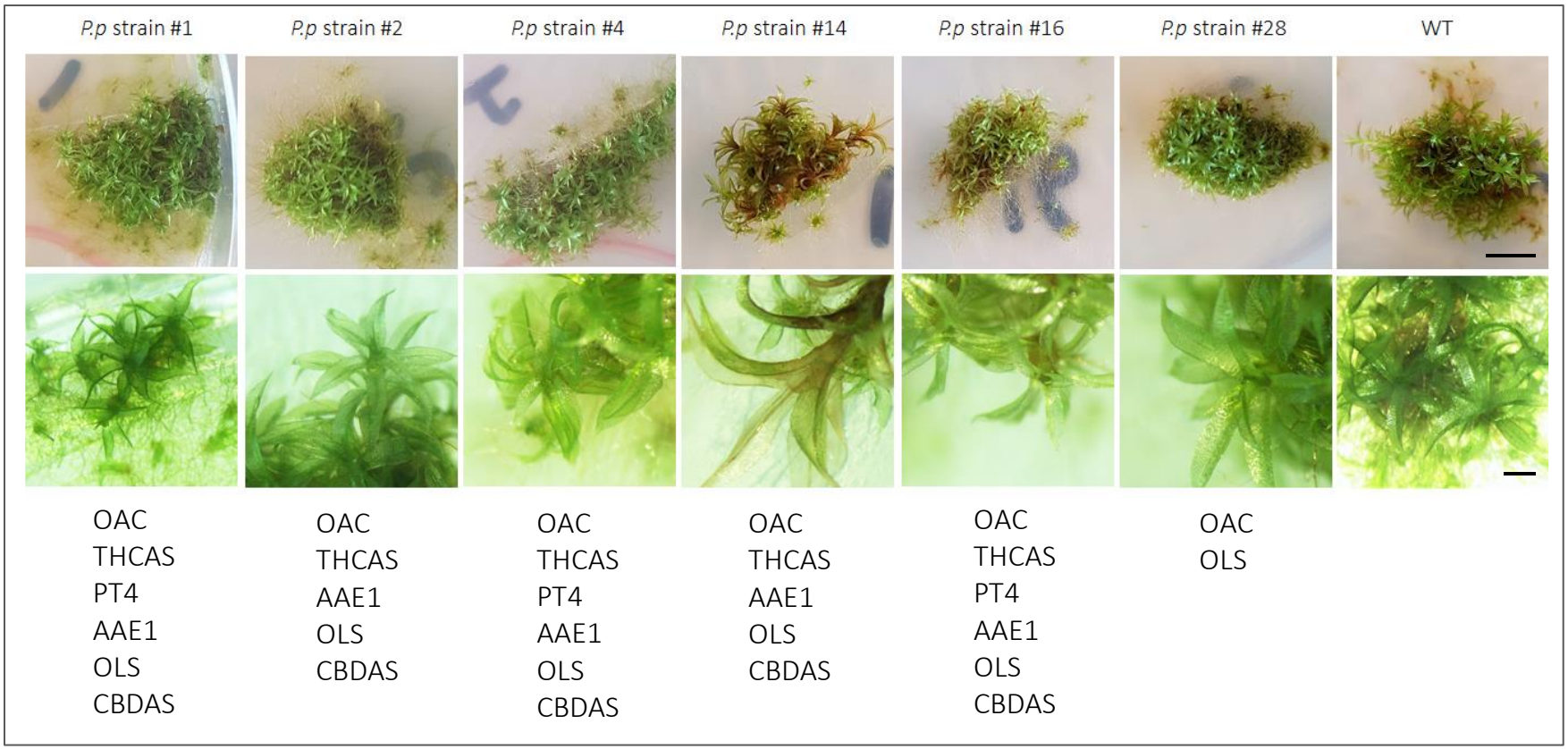


Figure A1.9: Normal phenotype of transgenic lines compared to wild-type (WT). All plants were cultivated on BCD media for eight weeks. Transgenic lines are shown with similar phenotype and developed gametophore tissue (same plants with two different magnifications are presented). The list of inserted genes is written under each picture of the transgenic line. (Scale bars: 0.5 cm (top) and 0.2 µm (bottom).

A1.4 Conclusion

Six cannabinoid biosynthetic pathway genes were engineered into the moss *Physcomitrium patens*. This work establishes the production of cannabinoids in a photosynthetic organism that can lead to large-scale industrial production. This production method will not affect environmental and ecological variables and generally have a lower environmental impact than field production, such as fertilizers, water, and petrol. Compared to yeast, it needed extensive and complex precursor pathway engineering before introducing the cannabinoid pathway genes and further optimizing the metabolic network in *P. patens* [142]. Optimization of cannabinoid yield in *P. patens* can hypothetically result in a stable, environmentally friendly, and commercially sustainable production platform. An advantage of *P. patens* as a cannabinoid production platform is that the extract only requires simple purification steps. The use of *P. patens* could lead to upscaling the production of cannabinoids in plant-based bioreactors and the production of other complexes, high-value plant-based compounds [346]. This bioengineering accomplishment expands the frontiers of synthetic biotechnology, and it is potentially more cost-effective than other carbon supplemented biotechnological platforms.

In *Physcomitrium*, I could target multiple loci independently with a single transformation experiment using a mixture of constructs. So, each locus was targeted with the same frequency as when a single construct was used. Multiple transformations were established as means by which large numbers of mutants were generated efficiently, and our analysis of multiply transformed plants confirms that the majority of targeting constructs (62.4%) delivered in a batch and combined in the generated transgenic lines. However, phenotypic analysis of all mutants requires to be observed.

Concatenation is most frequent in targeted insertion loci [321]. Significantly, it appears that it does not happen by a combination of homologous recombination and non-homologous end-joining, as it was assumed before [358], and it occurs entirely by homologous recombination integration of concatenated DNA. It is known that concatenated copies of plasmid DNA insert at one or a small number of loci when circular molecules are delivered to eukaryotic cells [321] and linear DNA delivered to plant cells

can cause the incorporation of nonspecific 'carrier' DNA in transgenic loci, following its co-delivery with a transforming plasmid [359]. Consequently, transforming DNA can undergo fragmentation and religation events after entering the cell before insertion into the genome.

In *Physcomitrium*, a typical transformation experiment produces many unstable transformants that maintain transgenes only as long as selection is maintained. These significantly decrease stable transformants when circular DNA is delivered to protoplasts. Unstable transformants contain extrachromosomal concatemers of the transforming DNA [360]. Transformation with linear DNA reduces the frequency of unstable transformants [361]. On the other hand, the numbers of maize ubiquitin promoter copies per plant have to be determined to confirm the concatenation event, and the Southern hybridization technique can estimate it.

A1.5 Supplementary data

A1.5.1 The synthesized CDS sequences of the cannabinoid biosynthesis pathway enzymes

>OAC
Atggccgtgaagcacttgatcgtgttgaagttcaaggacgagatcaccgaggcccagaaggaggagttcttcaagacctacgtgaacttggtgaacatca
tccccgccatgaaggacgtgtactggggaaaggacgtgacccagaagaacaaggaggagggatacacccacatcgtggaggtgaccttcgagtccgtg
gagaccatccaggactacatcatccaccccgcccacgtgggattcggagacgtgtacaggtccttctgggagaagttgttgatcttcgactacacccccag
gaag

>THCA
atgaactgctccgccttctccttctggttcgtgtgcaagatcatcttcttcttcttgtccttccacatccagatctccatcgccaaccccagggagaacttcttg
aagtgcttctccaagcacatccccaacaacgtggccaaccccaagttggtgtacacccagcacgaccagttgtacatgtccatcttgaactccaccatcca
gaacttgaggttcatctccgacaccacccccaagcccttggtgatcgtgacccccctccaacaactcccacatccaggccaccatcttgtgctccaagaagg
tgggattgcagatcaggaccaggtccggaggacacgacgccgagggaatgtcctacatctcccaggtgcccttcgtggtggtggacttgaggaacatgca
ctccatcaagatcgacgtgcactcccagaccgcctgggtggaggccggagccaccttgggagaggtgtactactggatcaacgagaagaacgagaactt
gtccttccccggaggatactgccccaccgtgggagtgggaggacacttctccggaggaggatacggagccttgatgaggaactacggattggccgccga
caacatcatcgacgcccacttggtgaacgtggacggaaaggtgttggacaggaagtccatgggagaggacttgttctgggccatcaggggaggaggag
gagagaacttcggaatcatcgccgcctggaagatcaagttggtggccgtgccctccaagtccaccatcttctccgtgaagaagaacatggagatccacgg
attggtgaagttgttcaacaagtggcagaacatcgcctacaagtacgacaaggacttggtgttgatgacccacttcatcaccaagaacatcaccgacaac
cacggaaagaacaagaccaccgtgcacggatacttctcctccatcttccacggaggagtggactccttggtggacttgatgaacaagtccttccccgagtt
gggaatcaagaagaccgactgcaaggagttctcctggatcgacaccaccatcttctactccggagtggtgaacttcaacaccgccaacttcaagaaggag
atcttgttggacaggtccgccggaaagaagaccgccttctccatcaagttggactacgtgaagaagcccatccccgagaccgccatggtgaagatcttgg
agaagttgtacgaggaggacgtgggagccggaatgtacgtgttgtacccctacggaggaatcatggaggagatctccgagtccgccatccccttccccca
cagggccggaatcatgtacgagttgtggtacaccgcctcctgggagaagcaggaggacaacgagaagcacatcaactgggtgaggtccgtgtacaacttt
caccacccccctacgtgtcccagaaccccaggttggcctacttgaactacagggacttggacttgggaaagaccaaccacgcctcccccaacaactacacc
caggccaggatctggggagagaagtacttccggaaagaacttcaacaggttggtgaaggtgaagaccaaggtggaccccaacaacttcttcaggaacga
gcagtccatccccccccttgccccccccaccaccac

>PT4
Tgggattgtccttggtgtgcaccttctccttccagaccaactaccacaccttgttgaaccccccacaacaagaaccccaagaactccttgttgtcctaccagc
accccaagaccccccatcatcaagtcctcctacgacaacttcccctccaagtactgcttgaccaagaacttccacttgttgggattgaactcccacaacagg
atctcctcccagtccaggtccatcagggccggatccgaccagatcgagggatcccccccaccacgagtccgacaactccatcgccaccaagatcttgaact
tcggacacacctgctggaagttgcagaggccctacgtggtgaagggaatgatctccatcgcctgcggattgttcggaagggagttgttcaacaacaggca
cttgttctcctggggattgatgtggaaggccttcttcgccttggtgcccatcttgtccttcaacttcttcgccgccatcatgaaccagatctacgacgtggaca
tcgacaggatcaacaagcccgacttgcccttggtgtccggagagatgtccatcgagaccgcctggatcttgtccatcatcgtggccttgaccggattgatcg
tgaccatcaagttgaagtccgcccccttgttcgtgttcatctacatcttcggaatcttcgccggattcgcctactccgtgcccccccatcaggtggaagcagta
ccccttcaccaacttcttgatcaccatctcctcccacgtgggattggccttcacctcctactccgccaccacctccgccttgggattgcccttcgtgtggaggc
ccgccttctccttcatcatcgccttcatgaccgtgatgggaatgaccatcgccttcgccaaggacatctccgacatcgagggagacgccaagtacggagtg
tccaccgtggccaccaagttgggagccaggaacatgaccttcgtggtgtccggagtgttgttgttgaactacttggtgtccatctccatcggaatcatctgg
ccccaggtgttcaagtccaacatcatgatcttgtcccacgccatcttggccttctgcttgatcttccagaccagggagttggccttggccaactacgcctccg
ccccctccaggcagttcttcgagttcatctggttgttgtactacgccgagtacttcgtgtacgtgttcatc

>AAE1
atgggaaagaactacaagtccttggactccgtggtggcctccgacttcatcgccttgggaatcacctccgaggtggccgagaccttgcacggaaggttgg
ccgagatcgtgtgcaactacggagccgccacccccagacctggatcaacatcgccaaccacatcttgtcccccgacttgcccttctccttgcaccagatg
ttgttctacggatgctacaaggacttcggacccgccccccccgcctggatccccgaccccgagaaggtgaagtccaccaacttgggagccttgttggaga
agaggggaaaggagttcttgggagtgaagtacaaggaccccatctcctccttctcccacttccaggagttctccgtgaggaaccccgaggtgtactggag
gaccgtgttgatggacgagatgaagatctccttctccaaggaccccgagtgcatcttgaggagggacgacatcaacaaccccggaggatccgagtggttg
cccggaggatacttgaactccgccaagaactgcttgaacgtgaactccaacaagaagttgaacgacaccatgatcgtgtggagggacgagggaaacga
cgacttgcccttgaacaagttgaccttggaccagttgaggaagagggtgtggttggtgggatacgccttggaggagatgggattggagaagggatgcgcc
atcgccatcgacatgcccatgcacgtggacgccgtggtgatctacttggccatcgtgttggccggatacgtggtggtgtccatcgccgactccttctccgcc
cccgagatctccaccaggttgaggttgtccaaggccaaggccatcttcacccaggaccacatcatcaggggaaagaagaggatccccttgtactccaggg

tggtggaggccaagtcccccatggccatcgtgatcccctgctccggatccaacatcggagccgagttgagggacggagacatctcctgggactacttcttg
gagagggccaaggagttcaagaactgcgagttcaccgccagggagcagcccgtggacgcctacaccaacatcttgttctcctccggaaccaccggagag
cccaaggccatcccctggacccaggccacccccttgaaggccgccgccgacggatggtcccacttggacatcaggaagggagacgtgatcgtgtggccc
accaacttgggatggatgatgggaccctggttggtgtacgcctccttgttgaacggagcctccatcgccttgtacaacggatccccccttggtgtccggattc
gccaagttcgtgcaggacgccaaggtgaccatgttgggagtggtgccctccatcgtgaggtcctggaagtccaccaactgcgtgtccggatacgactggt
ccaccatcaggtgcttctcctcctccggagaggcctccaacgtggacgagtacttgtggttgatgggaagggccaactacaagcccgtgatcgagatgtgc
ggaggaaccgagatcggaggagccttctccgccggatccttcttgcaggcccagtccttgtcctccttctcctcccagtgcatgggatgcaccttgtacatct
tggacaagaacggataccccatgcccaagaacaagcccggaatcggagagttggccttgggacccgtgatgttcggagcctccaagaccttgttgaacg
gaaaccaccacgacgtgtacttcaagggaatgcccaccttgaacggagaggtgttgaggaggcacggagacatcttcgagttgacctccaacggatact
accacgcccacggaagggccgacgacaccatgaacatcggaggaatcaagatctcctccatcgagatcgagagggtgtgcaacgaggtggacgacag
ggtgttcgagaccaccgccatcggagtgccccccttgggaggaggacccgagcagttggtgatcttcttcgtgttgaaggactccaacgacaccaccatc
gacttgaaccagttgaggttgtccttcaacttgggattgcagaagaagttgaaccccttgttcaaggtgaccagggtggtgcccttgtcctccttgcccagg
accgccaccaacaagatcatgaggagggtgttgaggcagcagttctcccacttcgag

>OLS
Atgaaccacttgagggccgagggacccgcctccgtgttggccatcggaaccgccaaccccgagaacatcttgttgcaggacgagttccccgactactact
tcagggtgaccaagtccgagcacatgacccagttgaaggagaagttcaggaagatctgcgacaagtccatgatcaggaagaggaactgcttcttgaacg
aggagcacttgaagcagaaccccaggttggtggagcacgagatgcagaccttggacgccaggcaggacatgttggtggtggaggtgcccaagttggga
aaggacgcctgcgccaaggccatcaaggagtggggacagcccaagtccaagatcacccacttgatcttcacctccgcctccaccaccgacatgcccgga
gccgactaccactgcgccaagttgttgggattgtccccctccgtgaagagggtgatgatgtaccagttgggatgctacggaggaggaaccgtgttgagga
tcgccaaggacatcgccgagaacaacaagggagccagggtgttggccgtgtgctgcgacatcatggcctgcttgttcaggggaccctccgagtccgactt
ggagttgttggtgggacaggccatcttcggagacggagccgccgccgtgatcgtgggagccgagcccgacgagtccgtgggagagaggcccatcttcga
gttggtgtccaccggacagaccatcttgcccaactccgagggaaccatcggaggacacatcagggaggccggattgatcttcgacttgcacaaggacgtg
cccatgttgatctccaacaacatcgagaagtgcttgatcgaggccttcacccccatcggaatctccgactggaactccatcttctggatcacccaccccgga
ggaaaggccatcttggacaaggtggaggagaagttgcacttgaagtccgacaagttcgtggactccaggcacgtgttgtccgagcacggaaacatgtcct
cctccaccgtgttgttcgtgatggacgagttgaggaagaggtccttggaggagggaaagtccaccaccggagacggattcgagtggggagtgttgttcgg
attcggacccggattgaccgtggagagggtggtggtgaggtccgtgcccatcaagtac

>CBDA
atgaagtgctccaccttctccttctggttcgtgtgcaagatcatcttcttcttcttctccttcaacatccagacctccatcgccaaccccagggagaacttcttg
aagtgcttctcccagtacatccccaacaacgccaccaacttgaagttggtgtacacccagaacaacccccttgtacatgtccgtgttgaactccaccatccac
aacttgaggttcacctccgacaccacccccaagcccttggtgatcgtgacccccctcccacgtgtcccacatccagggaaccatcttgtgctccaagaaggt
gggattgcagatcaggaccaggtccggaggacacgactccgagggaatgtcctacatctcccaggtgcccttcgtgatcgtggacttgaggaacatgagg
tccatcaagatcgacgtgcactcccagaccgcctgggtggaggccggagccaccttgggagaggtgtactactgggtgaacgagaagaacgagaacttg
tccttggccgccggatactgccccaccgtgtgcgccggaggacacttcggaggaggaggatacggacccttgatgaggaactacggattggccgccgac
aacatcatcgacgcccacttggtgaacgtgcacggaaaggtgttggacaggaagtccatgggagaggacttgttctgggccttgaggggaggaggagcc
gagtccttcggaatcatcgtggcctggaagatcaggttggtggccgtgcccaagtccaccatgttctccgtgaagaagatcatggagatccacgagttggt
gaagttggtgaacaagtggcagaacatcgcctacaagtacgacaaggacttgttgttgatgacccacttcatcaccaggaacatcaccgacaaccaggg
aaagaacaagaccgccatccacacctacttctcctccgtgttcttgggaggagtggactccttggtggacttgatgaacaagtccttccccgagttgggaat
caagaagaccgactgcaggcagttgtcctggatcgacaccatcatcttctactccggagtggtgaactacgacaccgacaacttcaacaaggagatcttg
ttggacaggtccgccggacagaacggagccttcaagatcaagttggactacgtgaagaagcccatccccgagtccgtgttcgtgcagatcttggagaagt
tgtacgaggaggacatcggagccggaatgtacgccttgtaccctacggaggaatcatggacgagatctccgagtccgccatccccttcccccacagggc
cggaatcttgtacgagttgtggtacatctgctcctgggagaagcaggaggacaacgagaagcacttgaactggatcaggaacatctacaacttcatgacc
ccctacgtgtccaagaaccccaggttggcctacttgaactacagggacttggacatcggaatcaacgaccccaagaaccccaacaactacacccaggcc
aggatctggggagagaagtacttcggaaagaacttcgacaggttggtgaaggtgaagaccttggtggaccccaacaacttcttcaggaacgagcagtcc
atccccccccttgcccaggcacaggcac

A1.5.2 The construct's maps of the cannabinoid biosynthesis pathway enzymes genes.

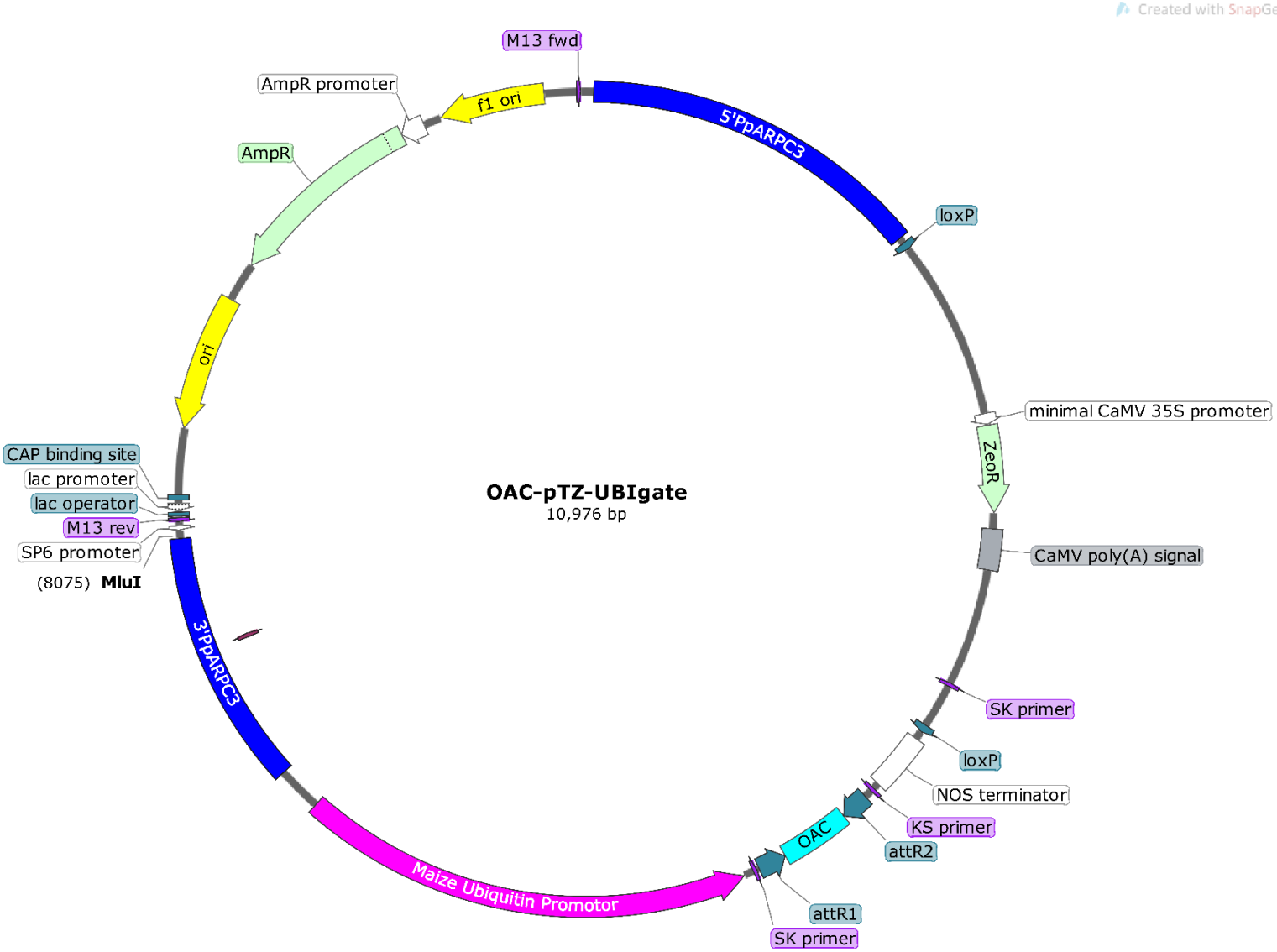


Figure A1.10: Plasmid map of the OAC-pTZ-UBIgate. Schematic of the clone used in OAC-pTZ-UBIgate. Two ~1.3kbp regions of homology flank a central Maze ubiquitin promoter and olvetolic acid cyclase (OAC) cassette with the Zeocin selection cassette enabling marker-assisted transformant selection.

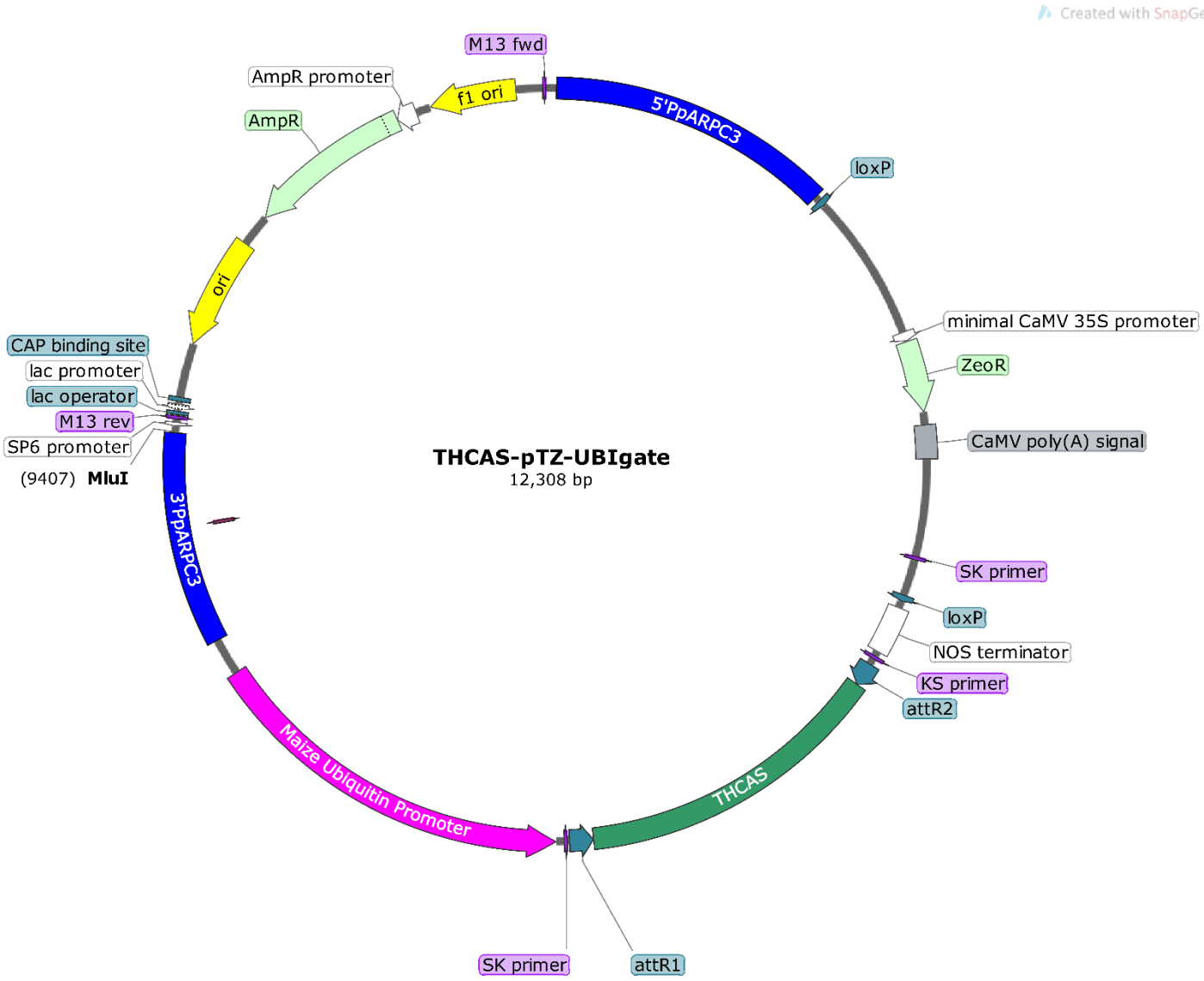


Figure A1.11: Plasmid map of the THCAS-pTZ-UBIgate. Schematic of the clone used in THCAS-pTZ-UBIgate. Two ~1.3kbp regions of homology flank a central Maze ubiquitin promoter and tetrahydro cannabinolic acid (THCAS) cassette with the Zeocin selection cassette enabling marker-assisted transformant selection.

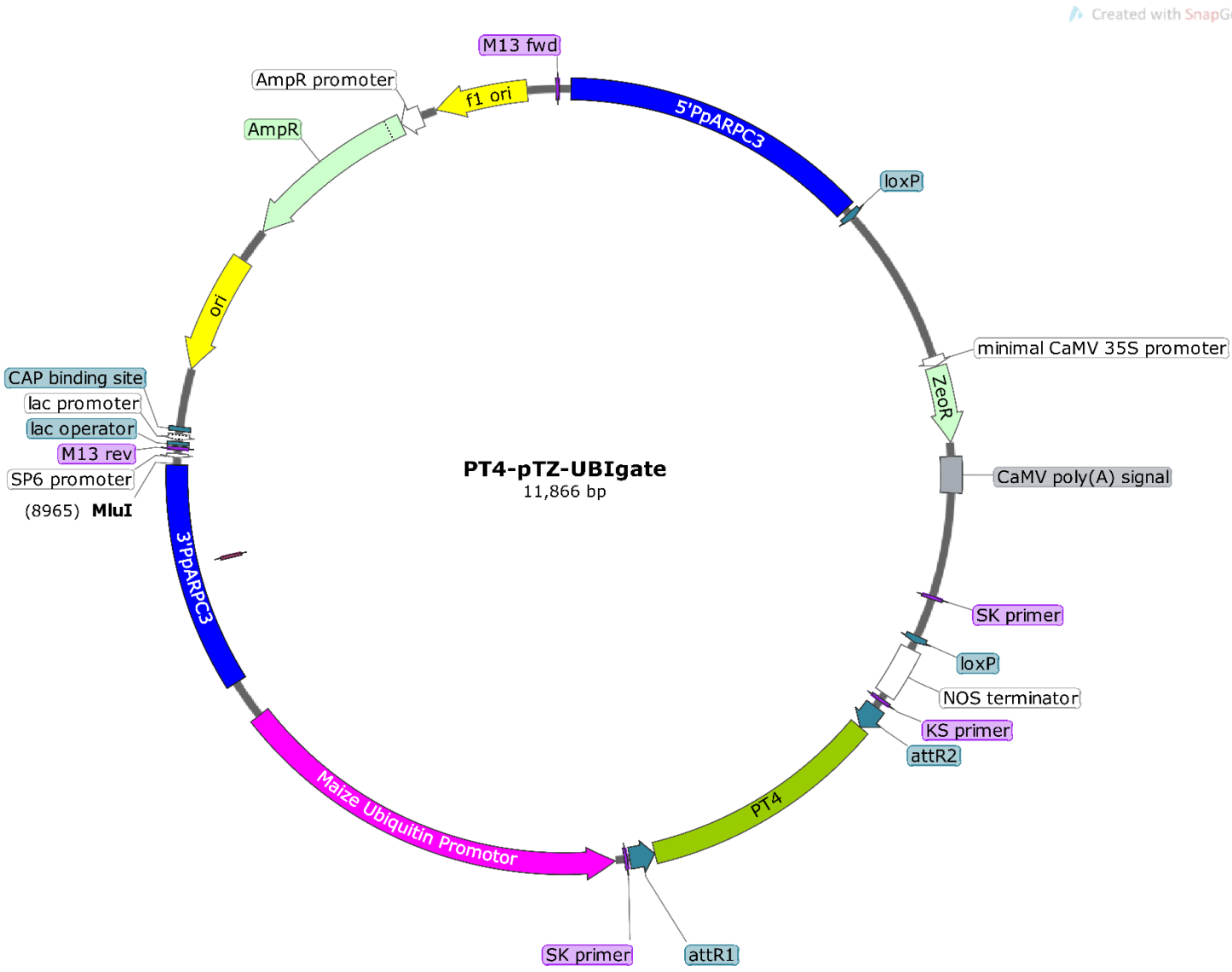


Figure A1.12: Plasmid map of the PT4-pTZ-UBIgate. Schematic of the clone used in PT4-pTZ-UBIgate. Two ~1.3kbp regions of homology flank a central Maze ubiquitin promoter and prenyltransferase 4 (PT4) cassette with the Zeocin selection cassette enabling marker-assisted transformant selection.

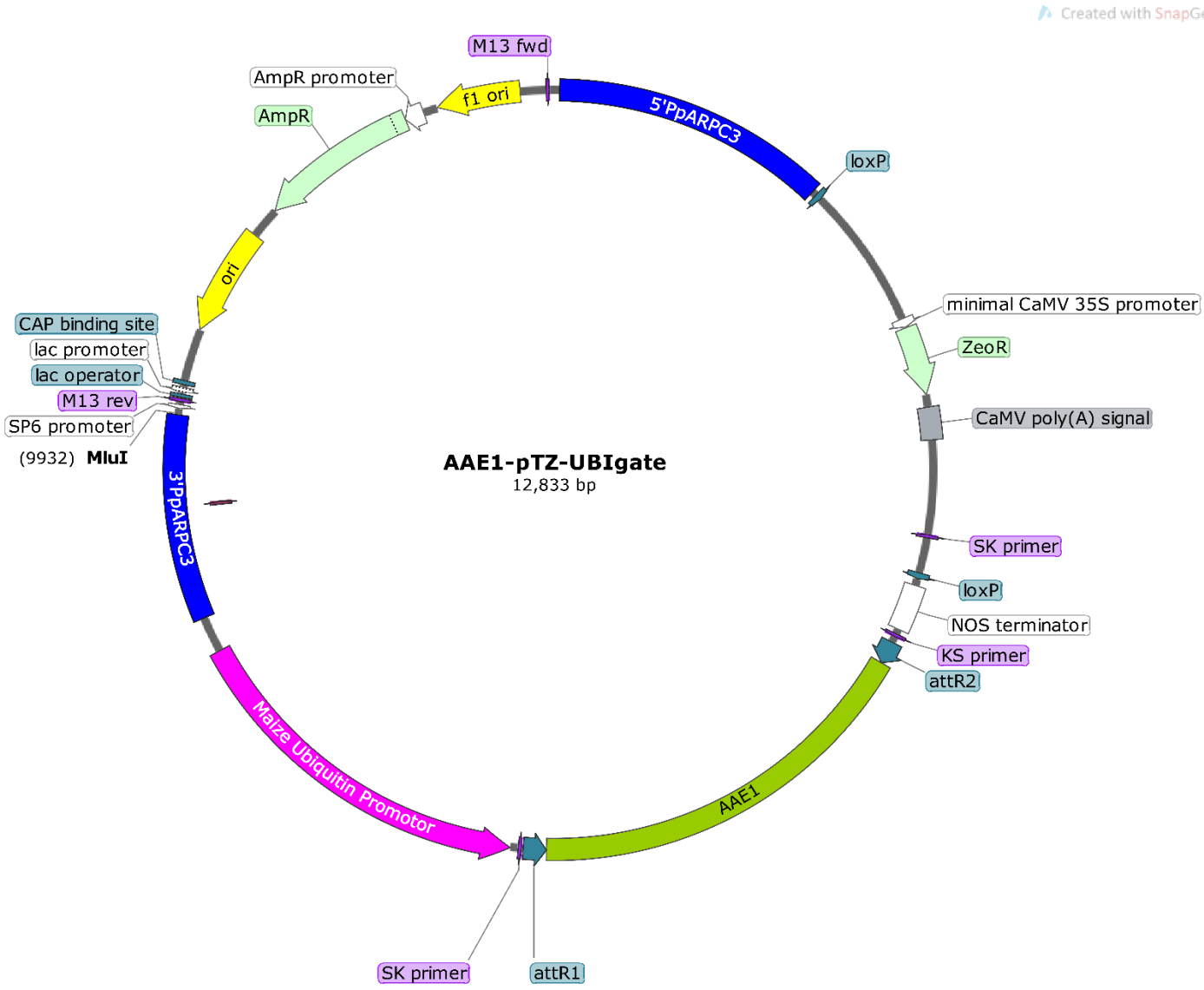


Figure A1.13: Plasmid map of the AAE1-pTZ-UBIgate. Schematic of the clone used in AAE1-pTZ-UBIgate. Two ~1.3kbp regions of homology flank a central Maze ubiquitin promoter and acyl activating enzyme 1 (AAE1) cassette with the Zeocin selection cassette enabling marker-assisted transformant selection.

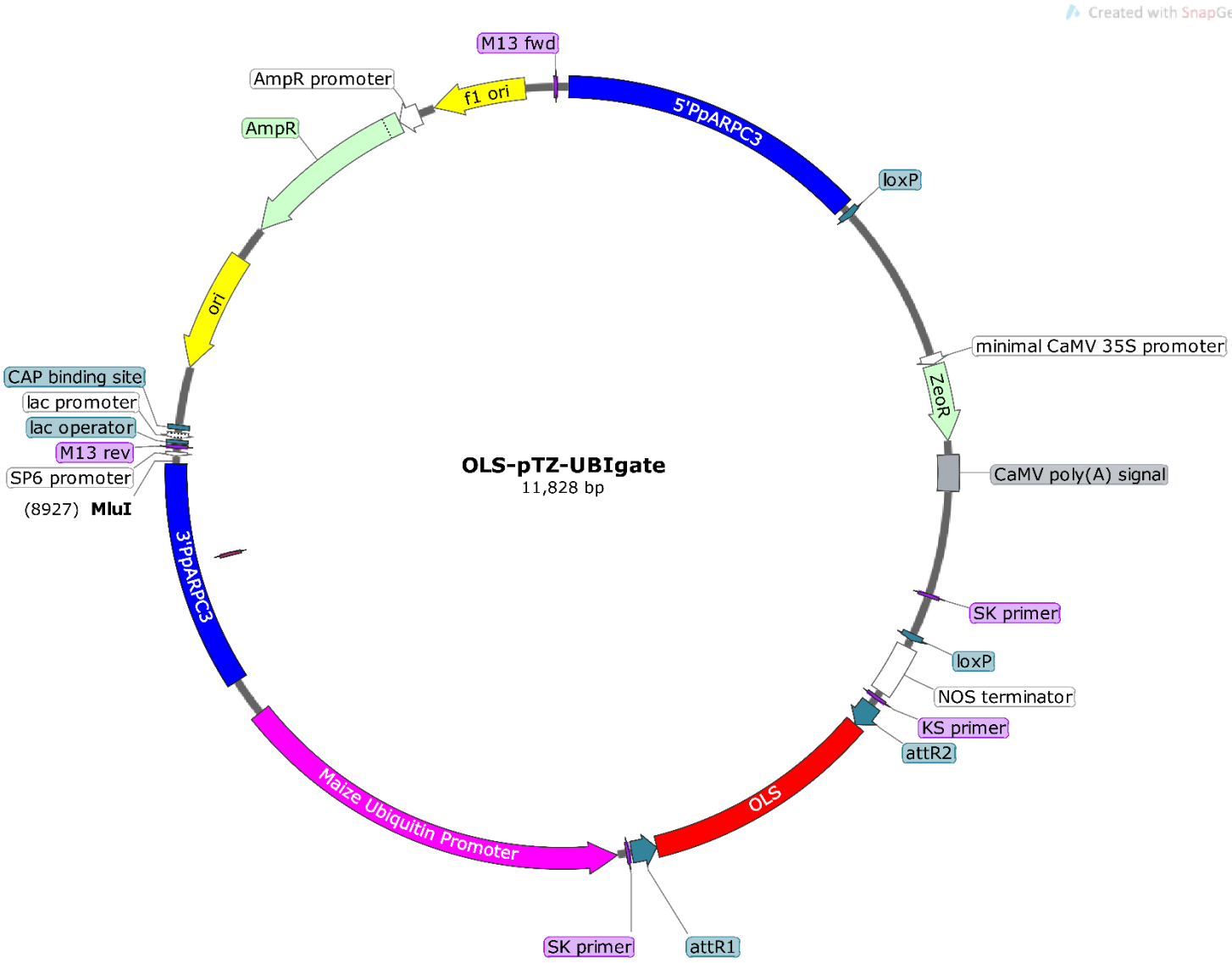


Figure A1.14: Plasmid map of the OLS-pTZ-UBIgate. Schematic of the clone used in OLS-pTZ-UBIgate. Two ~1.3kbp regions of homology flank a central Maze ubiquitin promoter and olivetol synthase (OLS) cassette with the Zeocin selection cassette enabling marker-assisted transformant selection.

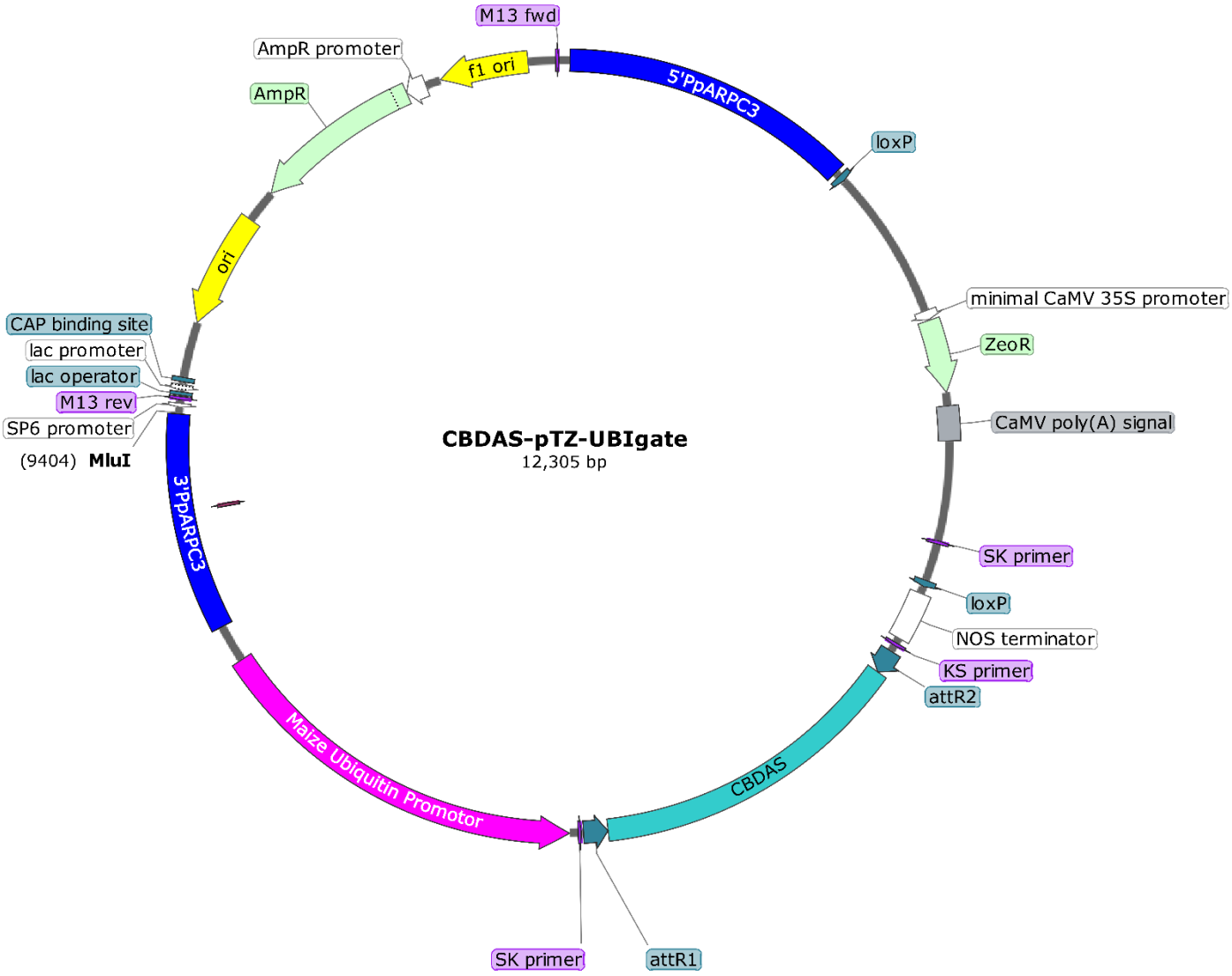
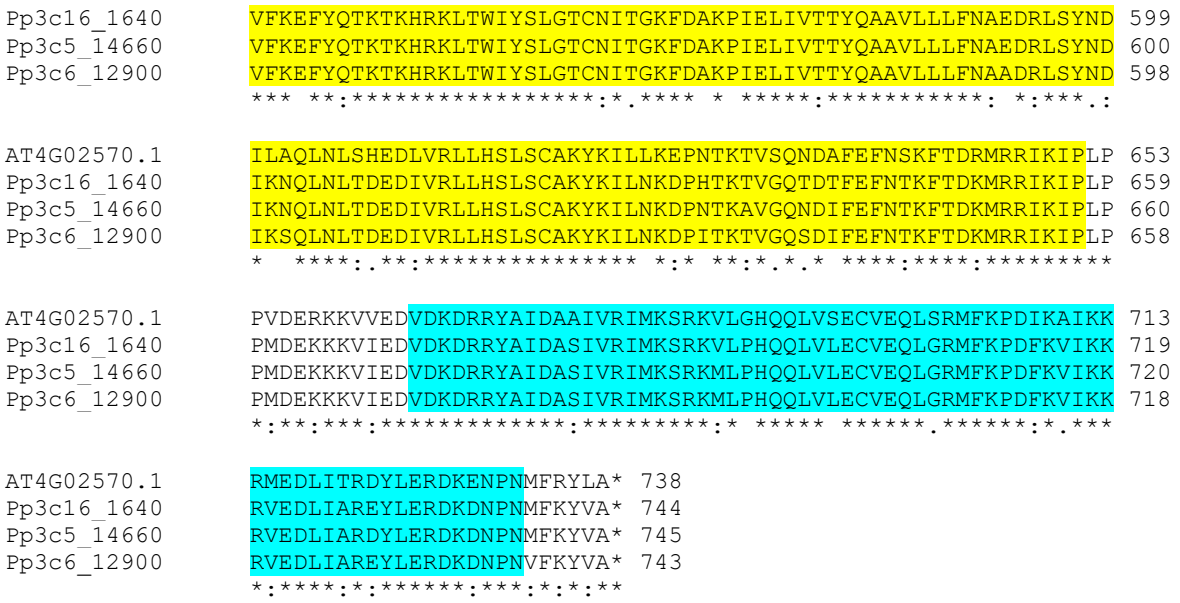


Figure A1.15: Plasmid map of the CBDAS-pTZ-UBIgate. Schematic of the clone used in CBDAS-pTZ-UBIgate. Two ~1.3kbp regions of homology flank a central Maize ubiquitin promoter and cannabidiolic acid synthase (CBDAS) cassette with the Zeocin selection cassette enabling marker-assisted transformant selection.

Appendix two: Protein Alignment of *P.patens* CUL1s and *A. thaliana* CUL1

```
CLUSTAL O(1.2.4) multiple sequence alignments

AT4G02570.1      ----MERKTIDLEQGWDYMQTGITKLKRILEGLNEPAFDSEQYMMLYTTIYNMCTQKPPH 56
Pp3c16_1640      -CMNNERKVIELEQGWSFMKKGITKLKNILEGVPEQQFSSEEYMLLYTTIYNMCTQKPPQ 59
Pp3c5_14660      MTMNNERKVIELEQGWSFMQIGITKLKNLLEGVPEQQFSSEEYMLLYTTIYNMCTQKPPQ 60
Pp3c6_12900      --MINERRVIELEQGWNFMQKGITKLKNLLEGVPEQQFSSEEYMLLYTTIYNMCTQKPPQ 58
                      **:.*:*****.:*: ******.:***: *  *.**:**:***************:

AT4G02570.1      DYSQQLYDKYREAFEEYINSTVLPALREKHDEFMLRELFKRWSNHKVMVRWLSRFFYYLD 116
Pp3c16_1640      DYSQQLYDRYRESFEDYIKSKVLPALREKHEEFMLKELVKRWDNHKIMVRWLSRFFNYLD 119
Pp3c5_14660      DYSQQLYDGYRVSFEEYINSKVLPALREKHEEFMLKELVKRWYNHKIMVRWLSRFFNYLD 120
Pp3c6_12900      DYSQQLYDRYRESFEGYINSKVLPALREKHEEFMLKELVKRWDNHKIMVRWLSRFFNYLD 118
                 ******** ** :** **:*.*********:****:**.*** ***:********* ***

AT4G02570.1      RYFIARRSLPPLNEVGLTCFRDLVYNELHSKVKQAVIALVDKEREGEQIDRALLKNVLDI 176
Pp3c16_1640      RYFIARRSLPALSEVGLMCFRDLVYAEIKINVKDAVIALIDREREGEQIDRALLKNVLGI 179
Pp3c5_14660      RYFIARRSLPALSEVGLICFRNLVYAETKINVKDAVVALIDREREGEQIDRALLKNVLGI 180
Pp3c6_12900      RYFIARRSLPALSEVGLMRFRDLVYEEMKVNVKDAVIALIDREREGEQIDRALLKNVLGI 178
                 ********** *.****  **:*** * : :**:**:**:*:****************.*

AT4G02570.1      YVEIGMGQMERYEEDFESFMLQDTSSYYSRKASSWIQEDSCPDYMLKSEECLKKERERVA 236
Pp3c16_1640      FVEIGMGNMDAYESDFETFMLQDTASYYSRKAASWIEEDSCPDYMLKAEECLKREKERVG 239
Pp3c5_14660      FVEIGMGNMDAYDTDFEQFMLEDTAAYYRRKASSWIQEDSCPDYMLKAEECLKREKERVG 240
Pp3c6_12900      FVEIGMGNMDAYETDFEAFMLEDTASYYKRKASSWIQEDSCPDYMLKAEECLKRERERVG 238
                 :******:*: *: *** ***:**::** ***:***:**********:*****:*:***.

AT4G02570.1      HYLHSSSEPKLVEKVQHELLVVFASQLLEKEHSGCRALLRDDKVDDLSRMYRLYHKILRG 296
Pp3c16_1640      HYLHASSEHKLLERVQHELLTQYEPQLLEKEHSGCHTLLRDDKVEDLSRMYRLFLRIPKG 299
Pp3c5_14660      HYLHASSEQKLLEKVQQELLTQYETQLLEKEHSGCHTLLRDDKVDDLSRMYRLFYRIPKG 300
Pp3c6_12900      HYLHASSEQKLLEKVQHELLTQYETQLLEKEHSGCHTLLRDDKVDDLSRMYRLFCRILKG 298
                 ****:*** **:*:**:***. :  **********::*******:********: :* :*

AT4G02570.1      LEPVANIFKQHVTAEGNALVQQAEDTATNQVA---NTASVQEQVLIRKVIELHDKYMVYV 353
Pp3c16_1640      LEPVAAIFKLHVTEEGTALVKQAEDAAGNKKAEKKDTVGVQEQAFVRKVIELHDKYLQYV 359
Pp3c5_14660      LEPVASIFKQHVTEEGTALVKQAEDAVSNKRAEKKDTVGVQEQVFVRKVIELHDKYLQYV 360
Pp3c6_12900      LDPVAAIFREHVTGEGTALVKQAEDAASNKKAERKDIVGVQEQAFVRKVIELHDKYLQYV 358
                 *:*** **: *** **.***:****:. *: *   : ..****.::**********: **

AT4G02570.1      TECFQNHTLFHKALKEAFEIFCNKTVAGSSSAELLATFCDNILKKGGSEKLSDEAIEDTL 413
Pp3c16_1640      SECFVNHSLFHKALKEAFEVFCNKGVAGSTSAELLATFCDNLLKKGGSEKLSDEAIEDTL 419
Pp3c5_14660      SECFANHSLFHKALKEAFEVFCNKGVAGSTSAELLATFCDNLLKKGGSEKLSDEAIEDTL 420
Pp3c6_12900      SDCFLNHSLFHKALKEAFEVFCNKGVAGSTSAELLATFCDNLLKKGGSEKLSDEAIEDTL 418
                 ::** **:***********:**** ****:***********:****************

AT4G02570.1      EKVVKLLAYISDKDLFAEFYRKKLARRLLFDRSANDDHERSILTKLKQQCGGQFTSKMEG 473
Pp3c16_1640      EKVVKLLAYISDKDLFAEFYRKKLARRLLFDKSANDDHERSILTKLKQQCGGQFTSKMEG 479
Pp3c5_14660      EKVVKLLAYISDKDLFAEFYRKKLARRLLFDKSANDDHERSILTKLKQQCGGQFTSKMEG 480
Pp3c6_12900      EKVVKLLAYISDKDLFAEFYRKKLARRLLFDKSANDDHERSILTKLKQQCGGQFTSKMEG 478
                 ******************************:****************************

AT4G02570.1      MVTDLTLARENQNSFEDYLGSNPAANPGIDLTVTVLTTGFWPSYKSFDINLPSEMIKCVE 533
Pp3c16_1640      MVTDLTLARENQTNFEEYLTENIQSSPGIDLTVTVLTTGFWPSYKSSDLALPAEMVKCVE 539
Pp3c5_14660      MVTDLTLARENQSNFEDYLSDNTKSNPGIDLTVTVLTTGFWPSYKSSDLALPAEMVKCVE 540
Pp3c6_12900      MVTDLTLARENQINFEEYLSDNTQSNPGIDLTVTVLTTGFWPSYKSSDLALPAEMVKCVE 538
                 ************ .**:** .*  :.******************** *: **:**:****

AT4G02570.1      VFKGFYETKTKHRKLTWIYSLGTCHINGKFDQKAIELIVSTYQAAVLLLFNTTDKLSYTE 593
```



Alignment of three *P. patens* CULLIN1s and *A. thaliana* CULLIN1 via Clustal Omega (European Molecular Biology Labs-European Bioinformatics Institute, Cambridge, UK, https://www.ebi.ac.uk/Tools/msa/clustalo/). The Cullin repeat 1 (CR1) domain is highlighted in light grey, and the Cullin repeat 2 (CR2) is highlighted in dark grey, the Cullin repeat 3 (CR3) is highlighted in dark yellow, the cullin homology domain (CH) is highlighted in yellow, the cullin neddylation domain is highlighted in turquoise. The location of the mutation is underlined and highlighted in red. Stars indicate a conserved residue, while colons indicate a similar structure. Periods indicate a relatively similar size, and blank spaces indicate no conservation.

Appendix three: The construct's maps of the cellulose synthase genes.

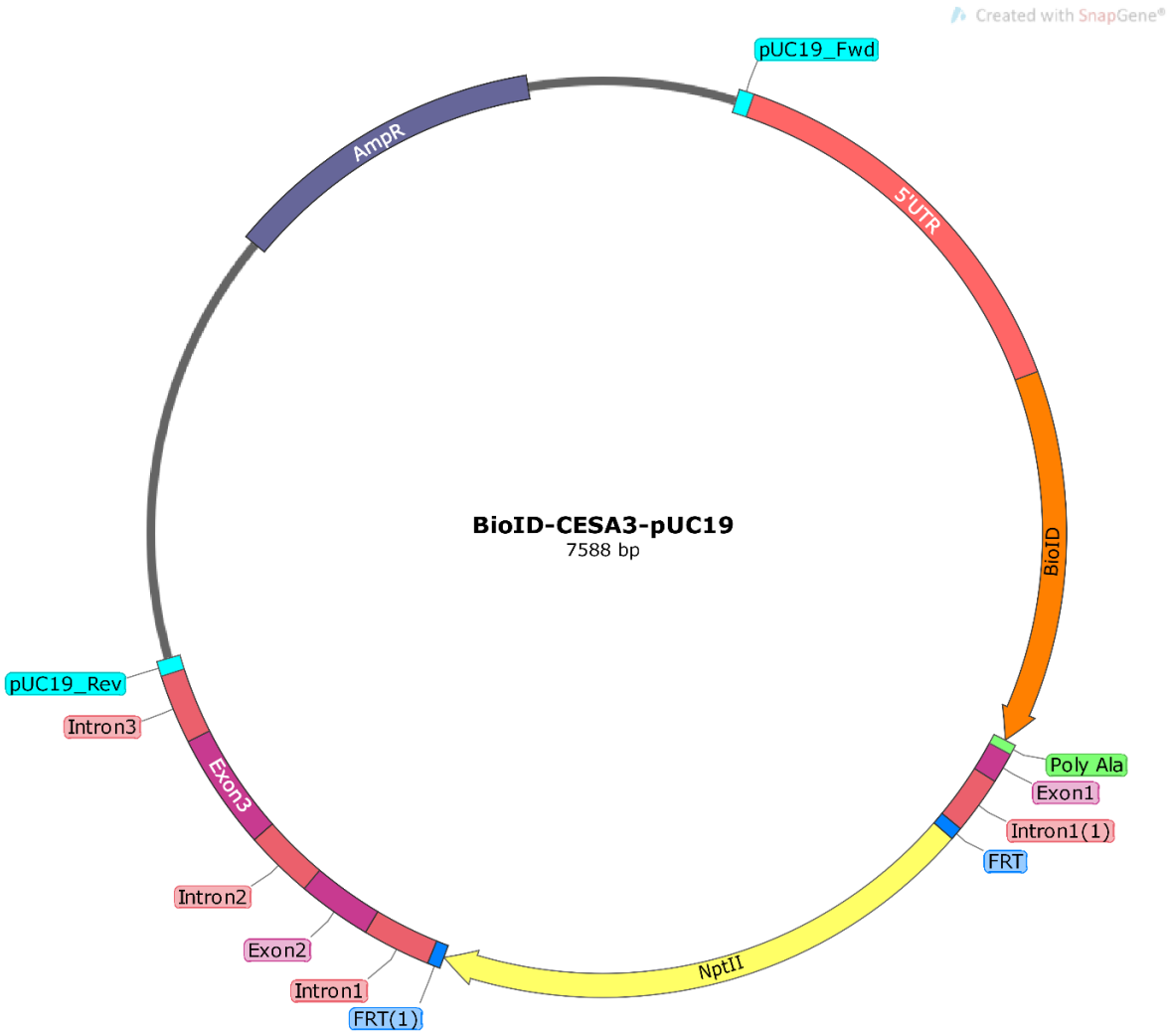


Figure A3.1: Plasmid map of the BioID-CESA3-pUC19. Schematic of the clone used in BioID-CESA3-pUC19. Multiple fragmnets cloned into pUC19 plasmid multiple cloning site by Gibson Assembly® technique.

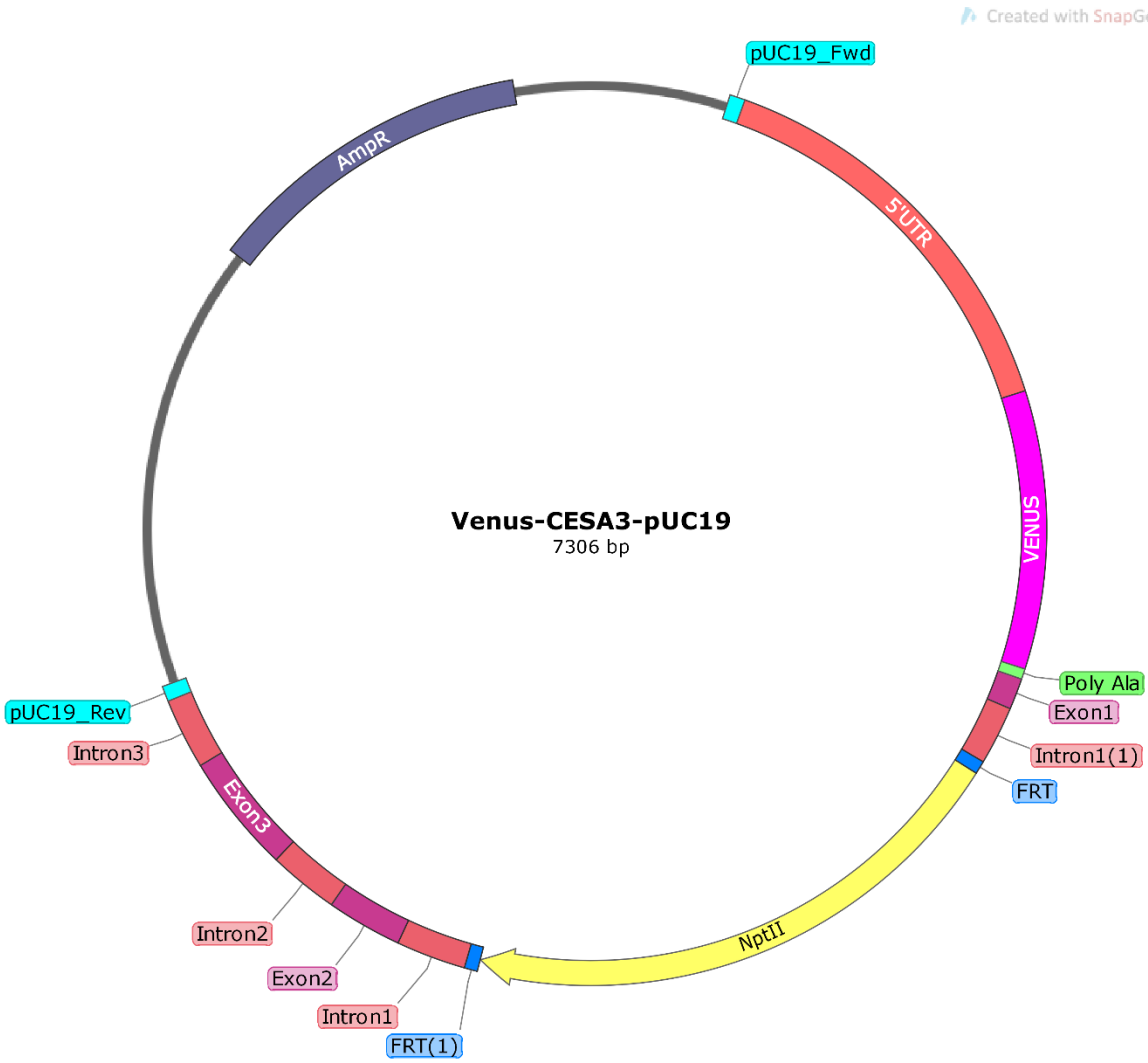


Figure A3.2: Plasmid map of the Venus-CESA3-pUC19. Schematic of the clone used in Venus-CESA3-pUC19. Multiple fragmnets cloned into pUC19 plasmid multiple cloning site by Gibson Assembly® technique.

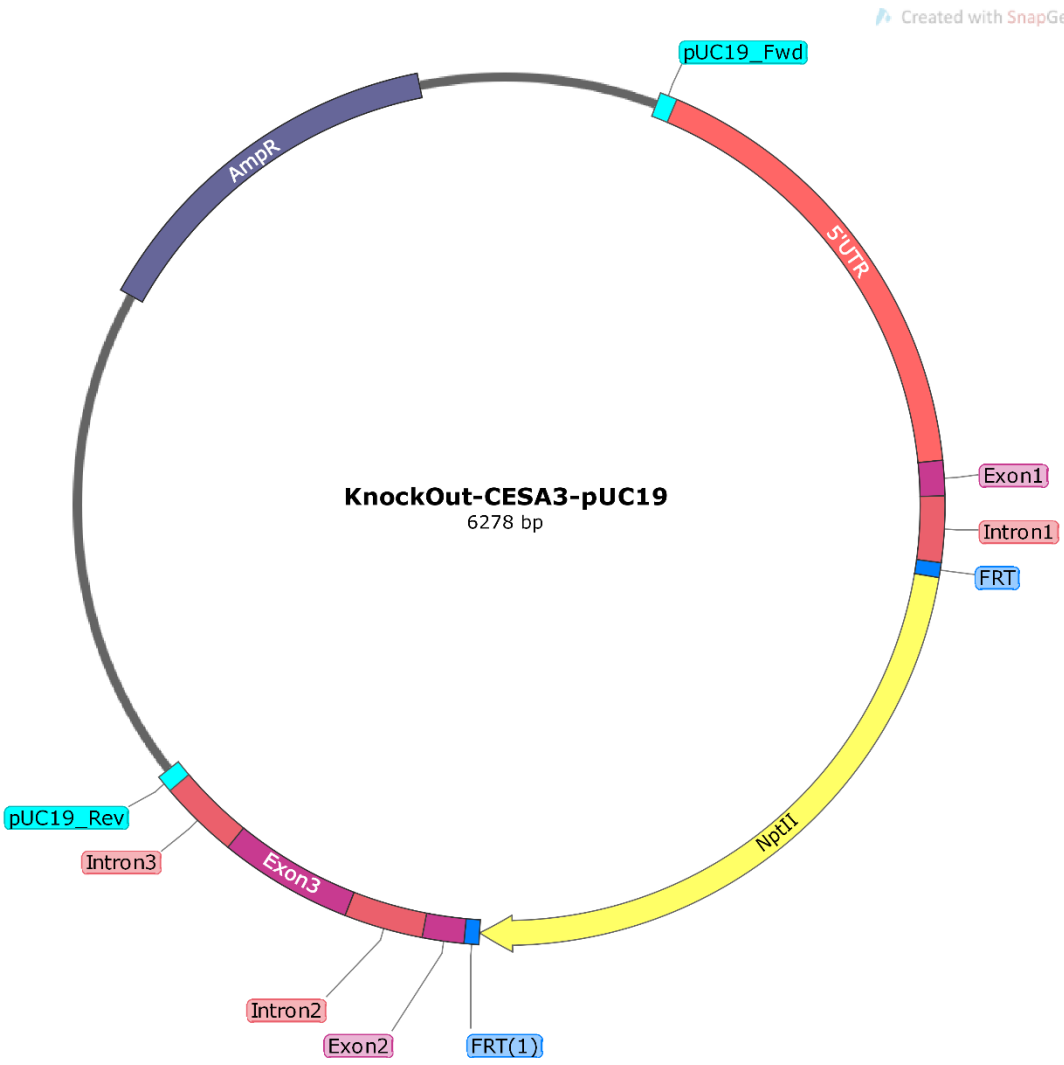


Figure A3.3: Plasmid map of the KnockOut-CESA3-pUC19. Schematic of the clone used in KnockOut-CESA3-pUC19. Multiple fragmnets cloned into pUC19 plasmid multiple cloning site by Gibson Assembly® technique.

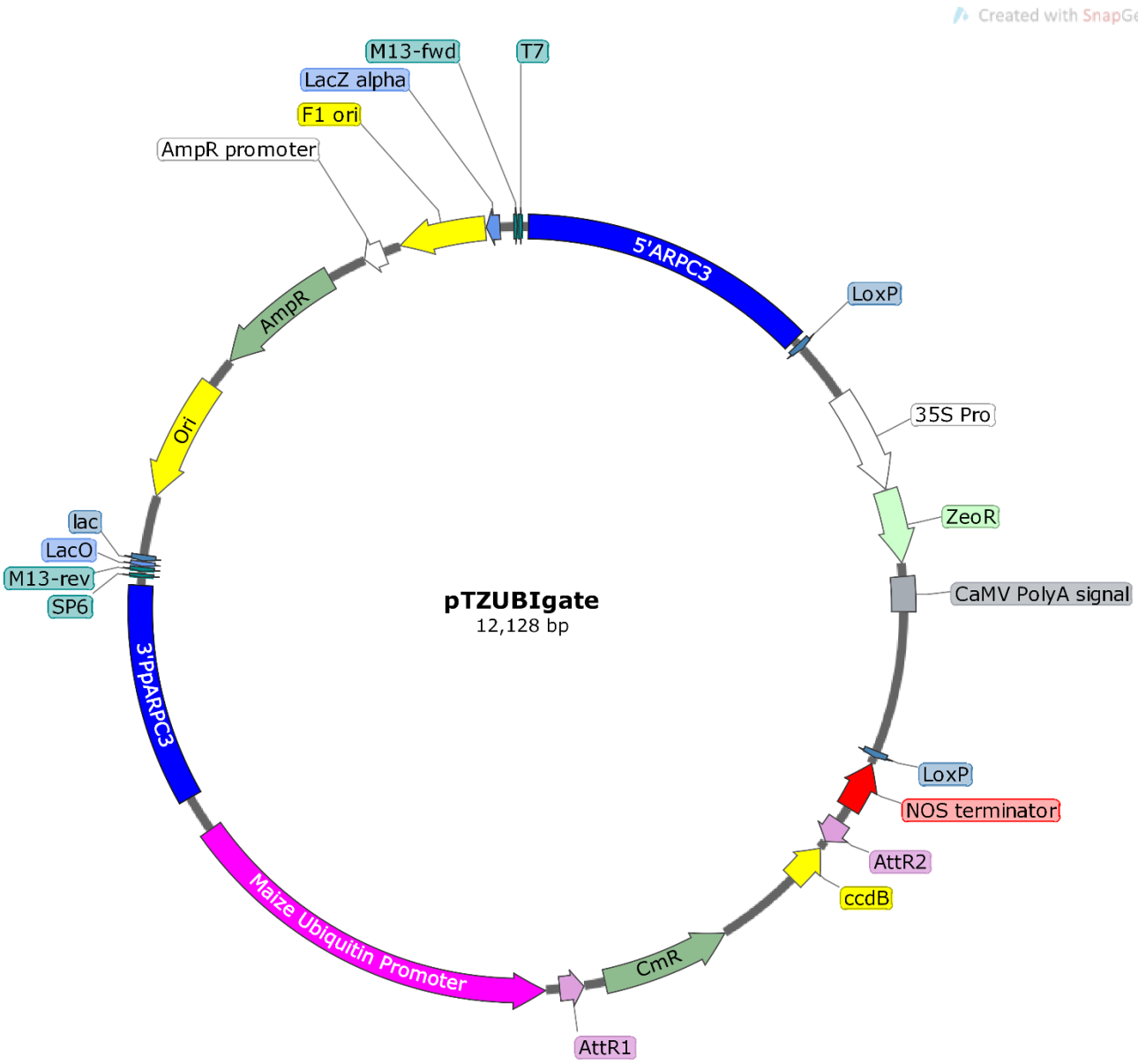


Figure A3.4: Plasmid map of the moss specific vector pTZUBIgate. Schematic of the pTZUBIgate used for gateway cloning technique. Two ~1.3kbp regions of homology flank a central Maze ubiquitin promoter cassette with the Zeocin selection cassette enabling marker-assisted transformant selection.

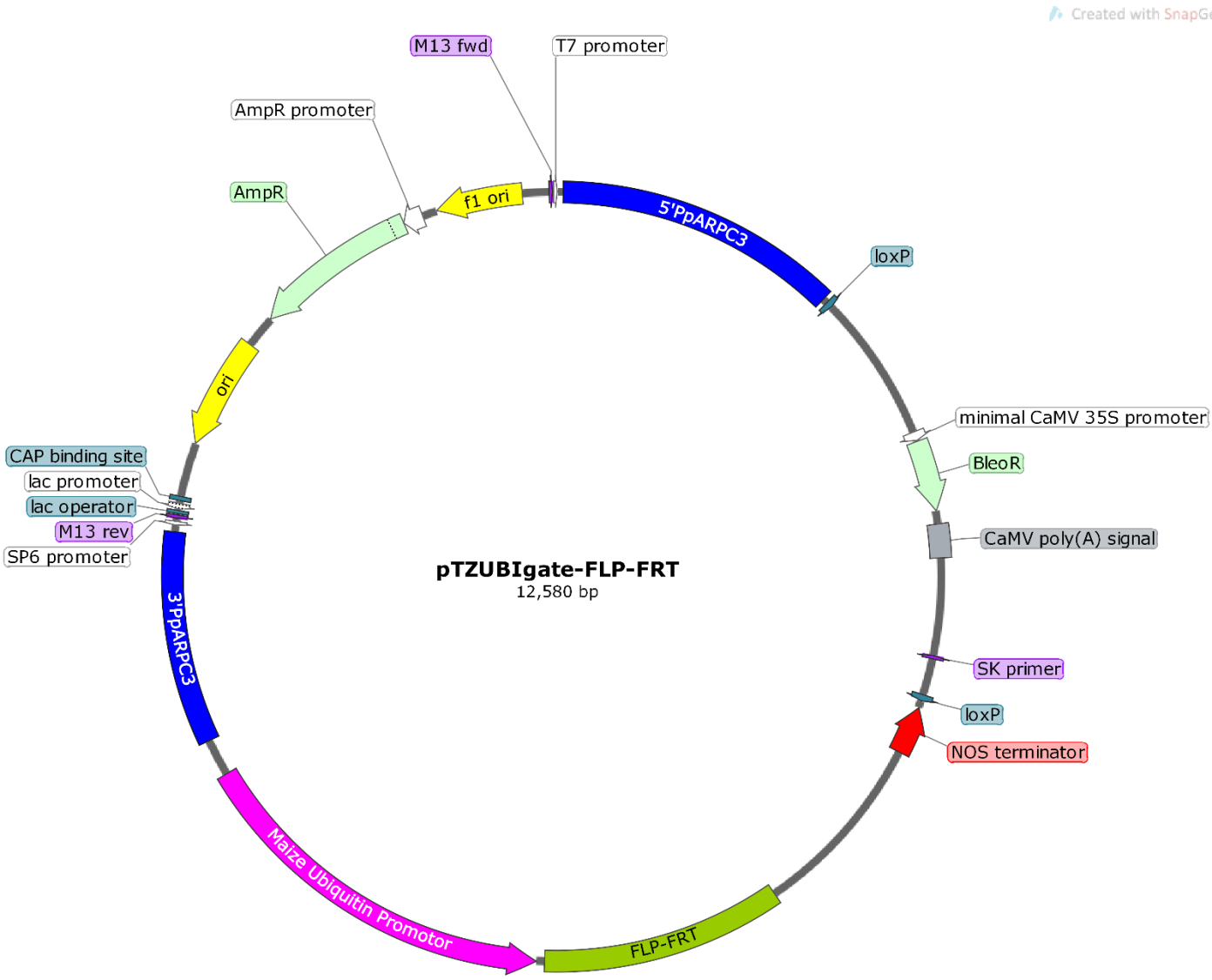


Figure A3.5: Plasmid map of pTZUBIgate-FLP-FRT construct. Schematic of the FLP-FRT fragment cloned into pTZUBIgate plasmid by gateway cloning technique. Two ~1.3kbp regions of homology flank a central Maze ubiquitin promoter and FLP-FRT with the Zeocin selection cassette enabling marker-assisted transformant selection.

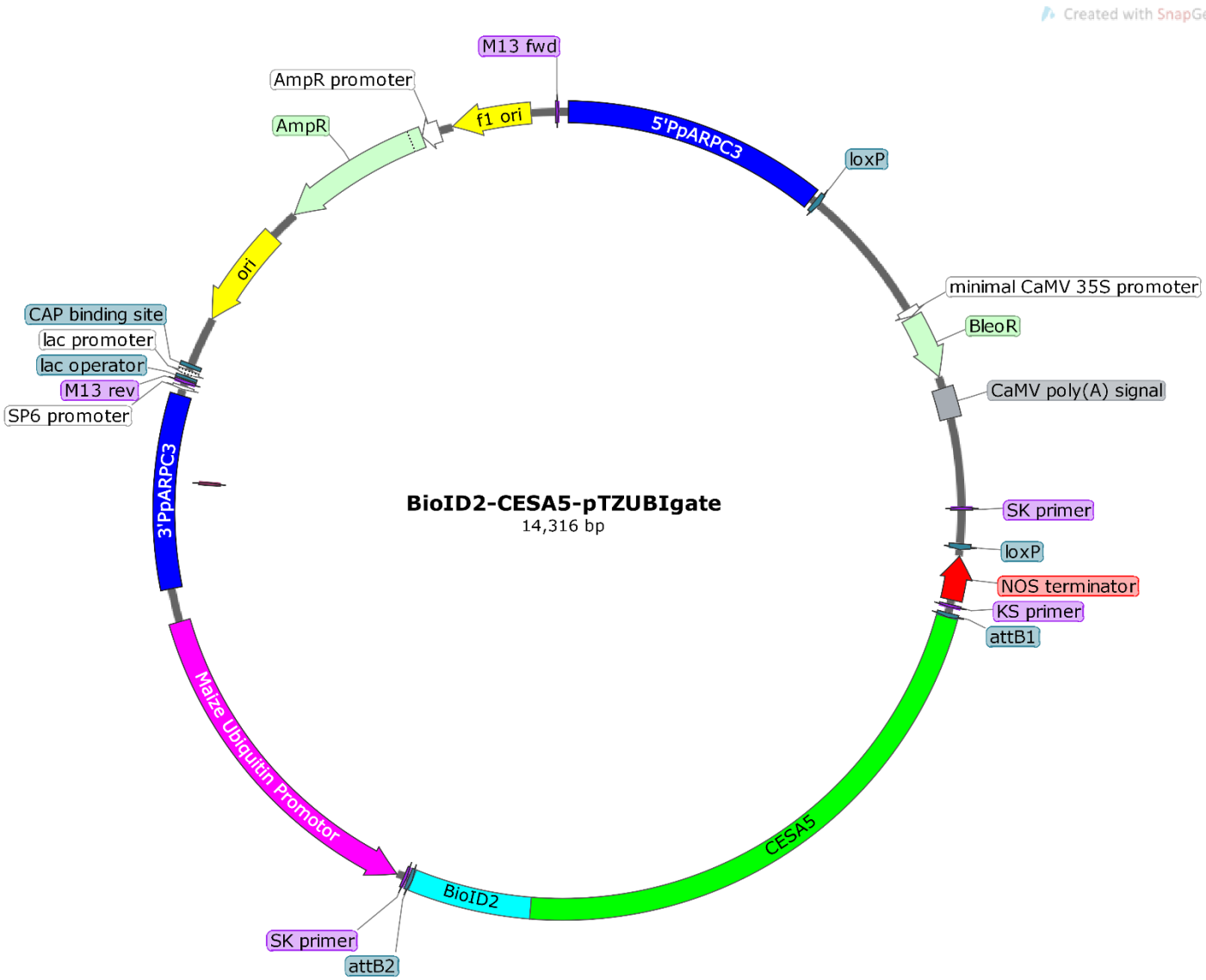


Figure A3.6: Plasmid map of the BioID2-CESA5-pTZUBIgate. Schematic of the clone used in BioID2-CESA5-pTZUBIgate. Two ~1.3kbp regions of homology flank a central Maize ubiquitin promoter and BioID2-CESA5 cassette with the Zeocin selection cassette enabling marker-assisted transformant selection.

References:

1. Campbell, M.M. and R.R. Sederoff, *Variation in Lignin Content and Composition (Mechanisms of Control and Implications for the Genetic Improvement of Plants).* Plant Physiology, 1996. **110**(1): p. 3.
2. Vitousek, P.M., D.R. Turner, and K. Kitayama, *Foliar Nutrients During Long-Term Soil Development in Hawaiian Montane Rain Forest.* Ecology, 1995. **76**(3): p. 712-720.
3. Carroll, A. and C. Somerville, *Cellulosic biofuels.* Annu Rev Plant Biol, 2009. **60**: p. 165-82.
4. Sticklen, M.B., *Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol.* Nature Reviews Genetics, 2008. **9**(6): p. 433-443.
5. Carpita, N.C., and McCann, M. , *The cell wall. In Biochemistry and Molecular Biology of Plants*, in *Plant Physiology and Development*, W.G. B. B. Buchanan, and R. L. Jones, Editor. 2000, American Society of Plant Biologists: Rockville, MD. p. 52–108.
6. Sampathkumar A., N.L., Persson S. , *The Plasma Membrane and the Cell Wall.*, in *The Plant Plasma Membrane. Plant Cell Monographs*, S.B. Murphy A., Peer W. , Editor. 2011, Springer: Berlin, Heidelberg.
7. Cosgrove, D.J., *Growth of the plant cell wall.* Nat Rev Mol Cell Biol, 2005. **6**(11): p. 850-61.
8. Somerville, C., *Cellulose synthesis in higher plants.* Annu Rev Cell Dev Biol, 2006. **22**: p. 53-78.
9. Matar, D. and A.M. Catesson, *Cell plate development and delayed formation of the pectic middle lamella in root meristems.* Protoplasma, 1988. **146**(1): p. 10-17.
10. Lodish, H., Berk, A., Zipursky, SL. et al., *The Dynamic Plant Cell Wall*, in *Molecular Cell Biology 4th edition*. 2000.
11. Carpita, N.C. and D.M. Gibeaut, *Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth.* The Plant journal : for cell and molecular biology, 1993. **3**(1): p. 1-30.
12. Mellerowicz, E.J. and B. Sundberg, *Wood cell walls: biosynthesis, developmental dynamics and their implications for wood properties.* Current opinion in plant biology, 2008. **11**(3): p. 293-300.
13. Török, B. and T. Dransfield, *Chapter 1.1 - Green Chemistry: Historical Perspectives and Basic Concepts*, in *Green Chemistry*, B. Török and T. Dransfield, Editors. 2018, Elsevier. p. 3-16.
14. Delmer, D.P. and Y. Amor, *Cellulose biosynthesis.* The Plant cell, 1995. **7**(7): p. 987-1000.
15. Saxena, I.M. and R.M. Brown, Jr., *Cellulose biosynthesis: current views and evolving concepts.* Annals of botany, 2005. **96**(1): p. 9-21.
16. Charnock, S.J., B. Henrissat, and G.J. Davies, *Three-Dimensional Structures of UDP-Sugar Glycosyltransferases Illuminate the Biosynthesis of Plant Polysaccharides.* Plant Physiology, 2001. **125**(2): p. 527.
17. Saxena, I.M., R.M. Brown, Jr., and T. Dandekar, *Structure--function characterization of cellulose synthase: relationship to other glycosyltransferases.* Phytochemistry, 2001. **57**(7): p. 1135-48.
18. Koyama, M., et al., *Parallel-up structure evidences the molecular directionality during biosynthesis of bacterial cellulose.* Proceedings of the National Academy of Sciences, 1997. **94**(17): p. 9091-9095.
19. Jing, W. and P.L. DeAngelis, *Dissection of the two transferase activities of the Pasteurella multocida hyaluronan synthase: two active sites exist in one polypeptide.* Glycobiology, 2000. **10**(9): p. 883-889.

20. Brown, R.M., *The Biosynthesis of Cellulose.* Journal of Macromolecular Science, Part A, 1996. **33**(10): p. 1345-1373.
21. Norris, J.H., et al., *Functional Specialization of Cellulose Synthase Isoforms in a Moss Shows Parallels with Seed Plants.* Plant Physiology, 2017. **175**(1): p. 210.
22. Pear, J.R., et al., *Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(22): p. 12637-12642.
23. Smith, V.C. and A.R. Ennos, *The effects of air flow and stem flexure on the mechanical and hydraulic properties of the stems of sunflowers Helianthus annuus L.* J Exp Bot, 2003. **54**(383): p. 845-9.
24. Jarvis, M.C., *Cellulose biosynthesis: counting the chains.* Plant physiology, 2013. **163**(4): p. 1485-1486.
25. Taylor, N.G., et al., *Interactions among three distinct CesA proteins essential for cellulose synthesis.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(3): p. 1450-1455.
26. Taylor, N.G., et al., *Cellulose synthesis in the Arabidopsis secondary cell wall.* Cellulose, 2004. **11**(3): p. 329-338.
27. Richmond, T., *Higher plant cellulose synthases.* Genome Biology, 2000. **1**(4): p. reviews3001.1-reviews3001.6.
28. Persson, S., et al., *Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in Arabidopsis.* Proc Natl Acad Sci U S A, 2007. **104**(39): p. 15566-71.
29. Desprez, T., et al., *Organization of cellulose synthase complexes involved in primary cell wall synthesis in Arabidopsis thaliana.* Proc Natl Acad Sci U S A, 2007. **104**(39): p. 15572-7.
30. Appenzeller, L., et al., *Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (CesA) gene family.* Cellulose, 2004. **11**(3): p. 287-299.
31. Suzuki, S., et al., *The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in Populus trichocarpa.* Plant Physiol, 2006. **142**(3): p. 1233-45.
32. Burton, R.A., et al., *The CesA gene family of barley. Quantitative analysis of transcripts reveals two groups of co-expressed genes.* Plant Physiol, 2004. **134**(1): p. 224-36.
33. Nixon, B.T., et al., *Comparative Structural and Computational Analysis Supports Eighteen Cellulose Synthases in the Plant Cellulose Synthesis Complex.* Sci Rep, 2016. **6**: p. 28696.
34. Paredez, A.R., C.R. Somerville, and D.W. Ehrhardt, *Visualization of Cellulose Synthase Demonstrates Functional Association with Microtubules.* Science, 2006. **312**(5779): p. 1491-1495.
35. Purushotham, P., R. Ho, and J. Zimmer, *Architecture of a catalytically active homotrimeric plant cellulose synthase complex.* Science, 2020. **369**(6507): p. 1089.
36. Cho, S.H., et al., *Synthesis and self-assembly of cellulose microfibrils from reconstituted cellulose synthase.* Plant physiology, 2017. **175**(1): p. 146-156.
37. Purushotham, P., et al., *A single heterologously expressed plant cellulose synthase isoform is sufficient for cellulose microfibril formation in vitro.* Proceedings of the National Academy of Sciences, 2016. **113**(40): p. 11360-11365.
38. McFarlane, H.E., A. Döring, and S. Persson, *The cell biology of cellulose synthesis.* Annu Rev Plant Biol, 2014. **65**: p. 69-94.

39. Triplett, B.A. and J.D. Timpa, *Characterization of cell-wall polymers from cotton ovule culture fiber cells by gel permeation chromatography.* In Vitro Cellular & Developmental Biology - Plant, 1995. **31**(3): p. 171-175.
40. Harris, D., J. Stork, and S. Debolt, *Genetic modification in cellulose-synthase reduces crystallinity and improves biochemical conversion to fermentable sugar.* GCB Bioenergy, 2009. **1**(1): p. 51-61.
41. Mansfield, S.D., C. Mooney, and J.N. Saddler, *Substrate and Enzyme Characteristics that Limit Cellulose Hydrolysis.* Biotechnology Progress, 1999. **15**(5): p. 804-816.
42. Polko, J.K. and J.J. Kieber, *The Regulation of Cellulose Biosynthesis in Plants.* The Plant Cell, 2019. **31**(2): p. 282.
43. Bold, H.C., Alexopoulos, C. J. & Delevoryas, T., *Morphology of Plants*

and Fungi 1987: HarperCollins, New York.
44. Schaefer, D.G. and J. Zryd, *The Moss Physcomitrella patens, Now and Then.* Plant Physiology, 2001. **127**(4): p. 1430-1438.
45. Decker, E.L. and R. Reski, *Current achievements in the production of complex biopharmaceuticals with moss bioreactors.* Bioprocess Biosyst Eng, 2008. **31**(1): p. 3-9.
46. Nishiyama, T., et al., *Comparative genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution.* Proc Natl Acad Sci U S A, 2003. **100**(13): p. 8007-12.
47. Cove, D.J. and C.D. Knight, *The Moss Physcomitrella patens, a Model System with Potential for the Study of Plant Reproduction.* The Plant Cell, 1993. **5**(10): p. 1483.
48. Cove, D., et al., *Mosses as model systems for the study of metabolism and development.* Annu Rev Plant Biol, 2006. **57**: p. 497-520.
49. Schaefer, D.G. and J.P. Zryd, *Efficient gene targeting in the moss Physcomitrella patens.* Plant J, 1997. **11**(6): p. 1195-206.
50. Egener, T., et al., *High frequency of phenotypic deviations in Physcomitrella patens plants transformed with a gene-disruption library.* BMC Plant Biology, 2002. **2**(1): p. 6.
51. Lang, D., et al., *Representation and high-quality annotation of the Physcomitrella patens transcriptome demonstrates a high proportion of proteins involved in metabolism in mosses.* Plant Biol (Stuttg), 2005. **7**(3): p. 238-50.
52. Rensing, S.A., et al., *The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants.* Science, 2008. **319**(5859): p. 64-9.
53. Wise, H.Z., I.M. Saxena, and R.M. Brown, *Isolation and characterization of the cellulose synthase genes PpCesA6 and PpCesA7 in Physcomitrella patens.* Cellulose, 2011. **18**(2): p. 371-384.
54. Zimmer, A., et al., *Dating the early evolution of plants: detection and molecular clock analyses of orthologs.* Mol Genet Genomics, 2007. **278**(4): p. 393-402.
55. Shaw, J. and K. Renzaglia, *Phylogeny and diversification of bryophytes.* Am J Bot, 2004. **91**(10): p. 1557-81.
56. Kenrick, P. and P.R. Crane, *The origin and early evolution of plants on land.* Nature, 1997. **389**(6646): p. 33-39.
57. Leliaert, F., Smith, D R., Moreau, H., Herron, M D., Verbruggen, H., Delwiche, Ch F., Clerck, O De., Critical Reviews in Plant Sciences, *Phylogeny and Molecular Evolution of the Green Algae.* Critical Reviews in Plant Sciences, 2012. **31**: p. 1 - 46.
58. Qiu, Y.L., et al., *The gain of three mitochondrial introns identifies liverworts as the earliest land plants.* Nature, 1998. **394**(6694): p. 671-4.

59. Nishiyama, T., et al., *Chloroplast Phylogeny Indicates that Bryophytes Are Monophyletic.* Molecular Biology and Evolution, 2004. **21**(10): p. 1813-1819.
60. Wang, C., et al., *Origin of plant auxin biosynthesis in charophyte algae.* Trends Plant Sci, 2014. **19**(12): p. 741-3.
61. Cooke, T.J., et al., *Evolutionary patterns in auxin action.* Plant Molecular Biology, 2002. **49**(3): p. 319-338.
62. Aigner, S., et al., *Unusual phenolic compounds contribute to ecophysiological performance in the purple-colored green alga Zygogonium ericetorum (Zygnematophyceae, Streptophyta) from a high-alpine habitat.* Journal of Phycology, 2013. **49**(4): p. 648-660.
63. Ligrone, R., et al., *Immunocytochemical detection of lignin-related epitopes in cell walls in bryophytes and the charalean alga Nitella.* Plant Systematics and Evolution, 2008. **270**(3): p. 257-272.
64. Kroken, S.B., L.E. Graham, and M.E. Cook, *Occurrence and evolutionary significance of RESISTANT CELL WALLS IN CHAROPHYTES AND BRYOPHYTES.* American Journal of Botany, 1996. **83**(10): p. 1241-1254.
65. Mikkelsen, M.D., et al., *Evidence for land plant cell wall biosynthetic mechanisms in charophyte green algae.* Ann Bot, 2014. **114**(6): p. 1217-36.
66. Little, A., et al., *Revised Phylogeny of the Cellulose Synthase Gene Superfamily: Insights into Cell Wall Evolution.* Plant Physiology, 2018. **177**(3): p. 1124.
67. Roberts, A.W. and E. Roberts, *Cellulose synthase (CesA) genes in algae and seedless plants.* Cellulose, 2004. **11**(3): p. 419-435.
68. Holland, N., et al., *A Comparative Analysis of the Plant Cellulose Synthase (CesA) Gene Family.* Plant Physiology, 2000. **123**(4): p. 1313.
69. Taylor, N.G., S. Laurie, and S.R. Turner, *Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in Arabidopsis.* The Plant cell, 2000. **12**(12): p. 2529-2540.
70. Fagard, M., et al., *PROCUSTE1 encodes a cellulose synthase required for normal cell elongation specifically in roots and dark-grown hypocotyls of Arabidopsis.* Plant Cell, 2000. **12**(12): p. 2409-2424.
71. Scheible, W.-R., et al., *Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in Arabidopsis Ixr1 mutants.* Proceedings of the National Academy of Sciences, 2001. **98**(18): p. 10079.
72. Burn, J.E., et al., *Functional analysis of the cellulose synthase genes CesA1, CesA2, and CesA3 in Arabidopsis.* Plant Physiol, 2002. **129**(2): p. 797-807.
73. R., S.S., *Forward genetics analysis in Physcomitrella patens identifies a novel ABA regulator.*, in *Centre for Plant Sciences*. 2015, University of Leeds. p. 237.
74. Roberts, A., E. Roberts, and C. Haigler, *Moss cell walls: structure and biosynthesis.* Frontiers in Plant Science, 2012. **3**(166).
75. Roberts, A.W. and J.T. Bushoven, *The cellulose synthase (CESA) gene superfamily of the moss Physcomitrella patens.* Plant Mol Biol, 2007. **63**(2): p. 207-19.
76. Taylor, N.G., *Cellulose biosynthesis and deposition in higher plants.* New Phytologist, 2008. **178**(2): p. 239-252.
77. Goss, C.A., et al., *A CELLULOSE SYNTHASE (CESA) gene essential for gametophore morphogenesis in the moss Physcomitrella patens.* Planta, 2012. **235**(6): p. 1355-67.
78. Tran, M.L., *Cellulose synthesis in Physcomitrella patens: Gene expression and mutation analysis*. 2015, University of Rhode Island: DigitalCommons@URL.
79. Baskin, T.I., *Anisotropic expansion of the plant cell wall.* Annu Rev Cell Dev Biol, 2005. **21**: p. 203-22.

80. Sánchez-Rodríguez, C., et al., *Phytohormones and the cell wall in Arabidopsis during seedling growth.* Trends Plant Sci, 2010. **15**(5): p. 291-301.
81. Hobbie, L. and M. Estelle, *Genetic approaches to auxin action.* Plant, Cell & Environment, 1994. **17**(5): p. 525-540.
82. Hyoung, S., et al., *Cytokinin oxidase PpCKX1 plays regulatory roles in development and enhances dehydration and salt tolerance in Physcomitrella patens.* Plant Cell Rep, 2020. **39**(3): p. 419-430.
83. Nishiyama, R., et al., *Analysis of cytokinin mutants and regulation of cytokinin metabolic genes reveals important regulatory roles of cytokinins in drought, salt and abscisic acid responses, and abscisic acid biosynthesis.* Plant Cell, 2011. **23**(6): p. 2169-83.
84. Heintz, D., et al., *Rapid alteration of the phosphoproteome in the moss Physcomitrella patens after cytokinin treatment.* J Proteome Res, 2006. **5**(9): p. 2283-93.
85. Zwack, P.J. and A.M. Rashotte, *Interactions between cytokinin signalling and abiotic stress responses.* J Exp Bot, 2015. **66**(16): p. 4863-71.
86. Tran, L.-S.P., et al., *Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in Arabidopsis.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(51): p. 20623-20628.
87. von Schwartzenberg, K., et al., *CHASE domain-containing receptors play an essential role in the cytokinin response of the moss Physcomitrella patens.* Journal of experimental botany, 2016. **67**(3): p. 667-679.
88. von Schwartzenberg, K., et al., *Cytokinins in the bryophyte Physcomitrella patens: analyses of activity, distribution, and cytokinin oxidase/dehydrogenase overexpression reveal the role of extracellular cytokinins.* Plant Physiol, 2007. **145**(3): p. 786-800.
89. Decker, E.L., et al., *Moss systems biology en route: phytohormones in Physcomitrella development.* Plant Biol (Stuttg), 2006. **8**(3): p. 397-405.
90. Cove, D., *The moss Physcomitrella patens.* Annu. Rev. Genet., 2005. **39**: p. 339-358.
91. Cove, D.J. and N.W. Ashton, *Growth regulation and development in Physcomitrella patens: an insight into growth regulation and development of bryophytes.* Botanical Journal of the Linnean Society, 1988. **98**(3): p. 247-252.
92. Reski, R. and W.O. Abel, *Induction of budding on chloronemata and caulonemata of the moss, Physcomitrella patens, using isopentenyladenine.* Planta, 1985. **165**(3): p. 354-8.
93. Robert, S., G. Mouille, and H. Höfte, *The mechanism and regulation of cellulose synthesis in primary walls: lessons from cellulose-deficient Arabidopsis mutants.* Cellulose, 2004. **11**(3): p. 351-364.
94. Nicol, F., et al., *A plasma membrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in Arabidopsis.* The EMBO journal, 1998. **17**(19): p. 5563-5576.
95. Banks, J.A., et al., *The Selaginella genome identifies genetic changes associated with the evolution of vascular plants.* science, 2011. **332**(6032): p. 960-963.
96. Ashton, N.W., N.H. Grimsley, and D.J. Cove, *Analysis of gametophytic development in the moss, Physcomitrella patens, using auxin and cytokinin resistant mutants.* Planta, 1979. **144**(5): p. 427-35.
97. Cove, D., *The moss Physcomitrella patens.* Annu Rev Genet, 2005. **39**: p. 339-58.
98. Bierfreund, N.M., R. Reski, and E.L. Decker, *Use of an inducible reporter gene system for the analysis of auxin distribution in the moss Physcomitrella patens.* Plant Cell Reports, 2003. **21**(12): p. 1143-1152.

99. Imaizumi, T., et al., *Cryptochrome light signals control development to suppress auxin sensitivity in the moss Physcomitrella patens.* Plant Cell, 2002. **14**(2): p. 373-86.
100. Prigge, M.J., et al., *Physcomitrella patens auxin-resistant mutants affect conserved elements of an auxin-signaling pathway.* Curr Biol, 2010. **20**(21): p. 1907-12.
101. Finet, C. and Y. Jaillais, *AUXOLOGY: When auxin meets plant evo-devo.* Developmental Biology, 2012. **369**(1): p. 19-31.
102. Hayashi, K.-i., et al., *Small-molecule agonists and antagonists of F-box protein–substrate interactions in auxin perception and signaling.* Proceedings of the national academy of sciences, 2008. **105**(14): p. 5632-5637.
103. Enders, T.A. and L.C. Strader, *Auxin activity: Past, present, and future.* American journal of botany, 2015. **102**(2): p. 180-196.
104. Leyser, O., *Auxin signalling: the beginning, the middle and the end.* Curr Opin Plant Biol, 2001. **4**(5): p. 382-6.
105. Gälweiler, L., et al., *Regulation of polar auxin transport by AtPIN1 in Arabidopsis vascular tissue.* Science, 1998. **282**(5397): p. 2226-2230.
106. Cho, M., S.H. Lee, and H.-T. Cho, *P-Glycoprotein4 displays auxin efflux transporter–like action in Arabidopsis root hair cells and tobacco cells.* The Plant Cell, 2007. **19**(12): p. 3930-3943.
107. Geisler, M., et al., *Cellular efflux of auxin catalyzed by the Arabidopsis MDR/PGP transporter AtPGP1.* The Plant Journal, 2005. **44**(2): p. 179-194.
108. Ruzicka, K., et al., *Łangowski Ł, Nejedlá E, Fujita H, Ito H, Syono J, et al (2010) Arabidopsis PIS1 encodes the ABCG37 transporter of auxinic compounds including the auxin precursor indole-3-butyric acid.* Proc Natl Acad Sci USA. **107**: p. 10749-10753.
109. Bennett, M.J., et al., *Arabidopsis AUX1 gene: a permease-like regulator of root gravitropism.* Science, 1996. **273**(5277): p. 948-950.
110. Swarup, R., et al., *Localization of the auxin permease AUX1 suggests two functionally distinct hormone transport pathways operate in the Arabidopsis root apex.* Genes & development, 2001. **15**(20): p. 2648-2653.
111. Fujita, T., et al., *Convergent evolution of shoots in land plants: lack of auxin polar transport in moss shoots.* Evolution & development, 2008. **10**(2): p. 176-186.
112. Jaillais, Y. and J. Chory, *Unraveling the paradoxes of plant hormone signaling integration.* Nature structural & molecular biology, 2010. **17**(6): p. 642-645.
113. Vert, G. and J. Chory, *Crosstalk in cellular signaling: background noise or the real thing?* Developmental cell, 2011. **21**(6): p. 985-991.
114. Ioio, R.D., et al., *A genetic framework for the control of cell division and differentiation in the root meristem.* Science, 2008. **322**(5906): p. 1380-1384.
115. Moubayidin, L., et al., *The rate of cell differentiation controls the Arabidopsis root meristem growth phase.* Current Biology, 2010. **20**(12): p. 1138-1143.
116. Růžička, K., et al., *Ethylene regulates root growth through effects on auxin biosynthesis and transport-dependent auxin distribution.* The Plant Cell, 2007. **19**(7): p. 2197-2212.
117. Stepanova, A.N., et al., *TAA1-mediated auxin biosynthesis is essential for hormone crosstalk and plant development.* Cell, 2008. **133**(1): p. 177-191.
118. Stepanova, A.N., et al., *Multilevel interactions between ethylene and auxin in Arabidopsis roots.* The Plant Cell, 2007. **19**(7): p. 2169-2185.
119. Swarup, R., et al., *Ethylene upregulates auxin biosynthesis in Arabidopsis seedlings to enhance inhibition of root cell elongation.* The Plant Cell, 2007. **19**(7): p. 2186-2196.
120. Marhavý, P., et al., *Cytokinin modulates endocytic trafficking of PIN1 auxin efflux carrier to control plant organogenesis.* Developmental cell, 2011. **21**(4): p. 796-804.

121. Sun, J., et al., *Jasmonate modulates endocytosis and plasma membrane accumulation of the Arabidopsis PIN2 protein.* New Phytologist, 2011. **191**(2): p. 360-375.
122. Whitford, R., et al., *GOLVEN secretory peptides regulate auxin carrier turnover during plant gravitropic responses.* Developmental cell, 2012. **22**(3): p. 678-685.
123. Willige, B.C., et al., *Gibberellin regulates PIN-FORMED abundance and is required for auxin transport–dependent growth and development in Arabidopsis thaliana.* The Plant Cell, 2011. **23**(6): p. 2184-2195.
124. Plant Biol (Stuttg).
125. Grossmann, K., *Auxin herbicides: current status of mechanism and mode of action.* Pest Manag Sci, 2010. **66**(2): p. 113-20.
126. Jasieniuk, M. and B.D. Maxwell, *Populations genetics and the evolution of herbicide resistance in weeds.* Phytoprotection, 1994. **75**(4): p. 25-35.
127. Mortensen, D.A., et al., *Navigating a Critical Juncture for Sustainable Weed Management.* BioScience, 2012. **62**(1): p. 75-84.
128. Woodward, A.W. and B. Bartel, *Auxin: regulation, action, and interaction.* Ann Bot, 2005. **95**(5): p. 707-35.
129. Tran, M.L., et al., *Direct observation of the effects of cellulose synthesis inhibitors using live cell imaging of Cellulose Synthase (CESA) in Physcomitrella patens.* Scientific reports, 2018. **8**(1): p. 735-735.
130. Brabham, C. and S. DeBolt, *Chemical genetics to examine cellulose biosynthesis.* Frontiers in Plant Science, 2012. **3**: p. 309.
131. Tateno, M., C. Brabham, and S. DeBolt, *Cellulose biosynthesis inhibitors - a multifunctional toolbox.* J Exp Bot, 2016. **67**(2): p. 533-42.
132. Shim, I., et al., *Alleles Causing Resistance to Isoxaben and Flupoxam Highlight the Significance of Transmembrane Domains for CESA Protein Function.* Front Plant Sci, 2018. **9**: p. 1152.
133. Brabham, C. and S. Debolt, *Chemical genetics to examine cellulose biosynthesis.* Frontiers in plant science, 2013. **3**: p. 309-309.
134. Heim, D.R., et al., *Isoxaben Inhibits the Synthesis of Acid Insoluble Cell Wall Materials In Arabidopsis thaliana.* Plant Physiology, 1990. **93**(2): p. 695-700.
135. Paredez, A.R., C.R. Somerville, and D.W. Ehrhardt, *Visualization of cellulose synthase demonstrates functional association with microtubules.* Science, 2006. **312**(5779): p. 1491-5.
136. Brabham, C., et al., *Indaziflam Herbicidal Action: A Potent Cellulose Biosynthesis Inhibitor.* Plant Physiology, 2014. **166**(3): p. 1177.
137. Sebastian, D.J., et al., *Indaziflam: a new cellulose-biosynthesis-inhibiting herbicide provides long-term control of invasive winter annual grasses.* Pest Manag Sci, 2017. **73**(10): p. 2149-2162.
138. Reski, R., J. Parsons, and E.L. Decker, *Moss-made pharmaceuticals: from bench to bedside.* Plant Biotechnol J, 2015. **13**(8): p. 1191-8.
139. Vesty, E.F., et al., *The decision to germinate is regulated by divergent molecular networks in spores and seeds.* New Phytologist, 2016. **211**(3): p. 952-966.
140. Decker, E.L., J. Parsons, and R. Reski, *Glyco-engineering for biopharmaceutical production in moss bioreactors.* Frontiers in plant science, 2014. **5**: p. 346-346.
141. Pan, X.-W., et al., *Sclareol production in the moss Physcomitrella patens and observations on growth and terpenoid biosynthesis.* Plant Biotechnology Reports, 2015. **9**(3): p. 149-159.

142. Zhan, X., et al., *Metabolic engineering of the moss Physcomitrella patens to produce the sesquiterpenoids patchoulol and α/β-santalene.* Frontiers in plant science, 2014. **5**: p. 636-636.
143. Anterola, A., et al., *Production of taxa-4(5),11(12)-diene by transgenic Physcomitrella patens.* Transgenic Res, 2009. **18**(4): p. 655-60.
144. Ikram, N.K.B.K., et al., *Stable heterologous expression of biologically active terpenoids in green plant cells.* Frontiers in plant science, 2015. **6**: p. 129-129.
145. Simonsen, H.T., D.P. Drew, and C. Lunde, *Perspectives on using physcomitrella patens as an alternative production platform for thapsigargin and other terpenoid drug candidates.* Perspectives in medicinal chemistry, 2009. **3**: p. 1-6.
146. Bach, S.S., et al., *Heterologous stable expression of terpenoid biosynthetic genes using the moss Physcomitrella patens.* Methods Mol Biol, 2014. **1153**: p. 257-71.
147. Cove, D.J., et al., *Transformation of the moss Physcomitrella patens using direct DNA uptake by protoplasts.* Cold Spring Harb Protoc, 2009. **2009**(2): p. pdb.prot5143.
148. Hohe, A., et al., *An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, Physcomitrella patens.* Curr Genet, 2004. **44**(6): p. 339-47.
149. King, B.C., et al., *In vivo assembly of DNA-fragments in the moss, Physcomitrella patens.* Scientific Reports, 2016. **6**(1): p. 25030.
150. Chen, S., D.W. Ehrhardt, and C.R. Somerville, *Mutations of cellulose synthase (CESA1) phosphorylation sites modulate anisotropic cell expansion and bidirectional mobility of cellulose synthase.* Proc Natl Acad Sci U S A, 2010. **107**(40): p. 17188-93.
151. Weissman, K.J. and P.F. Leadlay, *Combinatorial biosynthesis of reduced polyketides.* Nature reviews. Microbiology, 2005. **3**(12): p. 925-936.
152. Kawai, H., et al., *Responses of ferns to red light are mediated by an unconventional photoreceptor.* Nature, 2003. **421**(6920): p. 287-290.
153. Lei, L., et al., *Activation of MKK9-MPK3/MPK6 enhances phosphate acquisition in Arabidopsis thaliana.* New Phytol, 2014. **203**(4): p. 1146-60.
154. Suetsugu, N., et al., *A chimeric photoreceptor gene, NEOCHROME, has arisen twice during plant evolution.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(38): p. 13705.
155. Sabba, R.P. and K.C. Vaughn, *Herbicides That Inhibit Cellulose Biosynthesis.* Weed Science, 1999. **47**(6): p. 757-763.
156. Hoffman, J.C. and K.C. Vaughn, *Flupoxam Induces Classic Club Root Morphology but Is Not a Mitotic Disrupter Herbicide.* Pesticide Biochemistry and Physiology, 1996. **55**(1): p. 49-53.
157. Heim, D.R., et al., *Isoxaben Inhibits the Synthesis of Acid Insoluble Cell Wall Materials In <em>Arabidopsis thaliana</em&gt.* Plant Physiology, 1990. **93**(2): p. 695.
158. Harris, D.M., et al., *Cellulose microfibril crystallinity is reduced by mutating C-terminal transmembrane region residues CESA1A903V and CESA3T942I of cellulose synthase.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(11): p. 4098-4103.
159. DeBolt, S., et al., *Nonmotile cellulose synthase subunits repeatedly accumulate within localized regions at the plasma membrane in Arabidopsis hypocotyl cells following 2,6-dichlorobenzonitrile treatment.* Plant physiology, 2007. **145**(2): p. 334-338.
160. Heim, D.R., et al., *Mutation of a Locus of Arabidopsis thaliana Confers Resistance to the Herbicide Isoxaben.* Plant Physiol, 1989. **90**(1): p. 146-50.

161. Hohe, A., et al., *Day Length and Temperature Strongly Influence Sexual Reproduction and Expression of a Novel MADS-Box Gene in the Moss Physcomitrella patens.* Plant Biology, 2002. **4**(5): p. 595-602.
162. Roberts, A.W., et al., *Knocking out the wall: protocols for gene targeting in Physcomitrella patens.* Methods Mol Biol, 2011. **715**: p. 273-90.
163. Knight, C.D., Cove, D. J., Cuming, A. C., Quatrano, R. S., *Moss gene technology*, in *Molecular Plant Biology — A Practical Approach.*, P.M. Gilmartin, C. B, eds., Editor. 2002, Oxford Press: Oxford, New York. p. 285–301.
164. Kamisugi, Y., et al., *MRE11 and RAD50, but not NBS1, are essential for gene targeting in the moss Physcomitrella patens.* Nucleic acids research, 2012. **40**(8): p. 3496-3510.
165. Purcell, S., et al., *PLINK: a tool set for whole-genome association and population-based linkage analyses.* American journal of human genetics, 2007. **81**(3): p. 559-575.
166. Li, M., et al., *Robust and rapid algorithms facilitate large-scale whole genome sequencing downstream analysis in an integrative framework.* Nucleic Acids Res, 2017. **45**(9): p. e75.
167. von Stackelberg, M., S.A. Rensing, and R. Reski, *Identification of genic moss SSR markers and a comparative analysis of twenty-four algal and plant gene indices reveal species-specific rather than group-specific characteristics of microsatellites.* BMC Plant Biology, 2006. **6**(1): p. 9.
168. Yokota, Y. and A.N. Sakamoto, *The Moss Physcomitrella patens Is Hyperresistant to DNA Double-Strand Breaks Induced by γ-Irradiation.* Genes, 2018. **9**(2): p. 76.
169. Cove, D.J., et al., *Somatic hybridization in the moss Physcomitrella patens using PEG-induced protoplast fusion.* Cold Spring Harb Protoc, 2009. **2009**(2): p. pdb.prot5141.
170. Cove, D.J., et al., *The moss Physcomitrella patens: a novel model system for plant development and genomic studies.* Cold Spring Harb Protoc, 2009. **2009**(2): p. pdb.emo115.
171. Hiss, M., et al., *Sexual reproduction, sporophyte development and molecular variation in the model moss Physcomitrella patens: introducing the ecotype Reute.* Plant J, 2017. **90**(3): p. 606-620.
172. Zuryn, S., et al., *A strategy for direct mapping and identification of mutations by whole-genome sequencing.* Genetics, 2010. **186**(1): p. 427-30.
173. Kamisugi, Y., et al., *A sequence-anchored genetic linkage map for the moss, Physcomitrella patens.* The Plant journal : for cell and molecular biology, 2008. **56**(5): p. 855-866.
174. Doyle, J., *DNA Protocols for Plants*, in *Molecular Techniques in Taxonomy*, G.M. Hewitt, A.W.B. Johnston, and J.P.W. Young, Editors. 1991, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 283-293.
175. Abràmoff, M.D., P.J. Magalhães, and S.J. Ram, *Image processing with ImageJ.* Biophotonics international, 2004. **11**(7): p. 36-42.
176. Chomczynski, P., *A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples.* Biotechniques, 1993. **15**(3): p. 532-4, 536-7.
177. Burgess, J. and P.J. Linstead, *Cell-wall differentiation during growth of electrically polarised protoplasts of Physcomitrella.* Planta, 1982. **156**(3): p. 241-248.
178. *Tip Growth in the Moss Physcomitrella Patens*, in *Annual Plant Reviews online*. p. 143-166.
179. Burgess, J. and P.J. Linstead, *Studies on the growth and development of protoplasts of the moss, Physcomitrella patens, and its control by light.* Planta, 1981. **151**(4): p. 331-338.

180. Hobbie, L., et al., *The axr6 mutants of Arabidopsis thaliana define a gene involved in auxin response and early development.* Development, 2000. **127**(1): p. 23.
181. Biedermann, S., S. Mooney, and H. Hellmann, *Recognition and repair pathways of damaged DNA in higher plants.* Selected topics in DNA repair, 2011: p. 201-236.
182. Alscher, R.G., J.L. Donahue, and C.L. Cramer, *Reactive oxygen species and antioxidants: relationships in green cells.* Physiologia Plantarum, 1997. **100**(2): p. 224-233.
183. Barsoum, M., et al., *Ultraviolet Mutagenesis Coupled with Next-Generation Sequencing as a Method for Functional Interrogation of Powdery Mildew Genomes.* Molecular Plant-Microbe Interactions®, 2020. **33**(8): p. 1008-1021.
184. Sekiguchi, M. and T. Tsuzuki, *Oxidative nucleotide damage: consequences and prevention.* Oncogene, 2002. **21**(58): p. 8895-904.
185. Lavy, M., et al., *Constitutive auxin response in Physcomitrella reveals complex interactions between Aux/IAA and ARF proteins.* Elife, 2016. **5**.
186. von Schwartzenberg, K., et al., *Cytokinins in the Bryophyte Physcomitrella patens: Analyses of Activity, Distribution, and Cytokinin Oxidase/Dehydrogenase Overexpression Reveal the Role of Extracellular Cytokinins.* Plant Physiology, 2007. **145**(3): p. 786.
187. Tao, S. and M. Estelle, *Mutational studies of the Aux/IAA proteins in Physcomitrella reveal novel insights into their function.* New Phytol, 2018. **218**(4): p. 1534-1542.
188. Ashton, N.W. and D.J. Cove, *Mutants as tools for the analytical dissection of cell differentiation in Physcomitrella patens gametophytes.* Bryophyte development: physiology and biochemistry, 1990: p. 17-31.
189. Ashton, N.W., et al. *Developmental Studies of Physcomitrella patens Using Auxin and Cytokinin Sensitivity Mutants*. in *Plant Growth Substances 1988*. 1990. Berlin, Heidelberg: Springer Berlin Heidelberg.
190. Austin, R.S., et al., *Next-generation mapping of Arabidopsis genes.* Plant J, 2011. **67**(4): p. 715-25.
191. Betts, M.J., et al., *Bioinformatics for geneticists.* Wiley, Chichester, 2003: p. 289.
192. Schulze-Gahmen, U., H.L. De Bondt, and S.H. Kim, *High-resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: bound waters and natural ligand as guides for inhibitor design.* J Med Chem, 1996. **39**(23): p. 4540-6.
193. *Amino Acid Properties and Consequences of Substitutions*, in *Bioinformatics for Geneticists*. 2003. p. 289-316.
194. Ikehata, H. and T. Ono, *The mechanisms of UV mutagenesis.* J Radiat Res, 2011. **52**(2): p. 115-25.
195. Becker, M.M. and Z. Wang, *Origin of ultraviolet damage in DNA.* Journal of Molecular Biology, 1989. **210**(3): p. 429-438.
196. Moon, J., et al., *A New CULLIN 1 Mutant Has Altered Responses to Hormones and Light in Arabidopsis.* Plant Physiology, 2007. **143**(2): p. 684.
197. Quint, M., et al., *Characterization of a novel temperature-sensitive allele of the CUL1/AXR6 subunit of SCF ubiquitin-ligases.* Plant J, 2005. **43**(3): p. 371-83.
198. Hua, Z. and R.D. Vierstra, *The cullin-RING ubiquitin-protein ligases.* Annu Rev Plant Biol, 2011. **62**: p. 299-334.
199. Petroski, M.D. and R.J. Deshaies, *Function and regulation of cullin-RING ubiquitin ligases.* Nat Rev Mol Cell Biol, 2005. **6**(1): p. 9-20.
200. Vain, T., et al., *Selective auxin agonists induce specific AUX/IAA protein degradation to modulate plant development.* Proceedings of the National Academy of Sciences, 2019. **116**(13): p. 6463-6472.

201. Sarikas, A., T. Hartmann, and Z.-Q. Pan, *The cullin protein family.* Genome Biology, 2011. **12**(4): p. 220.
202. Moon, J., G. Parry, and M. Estelle, *The ubiquitin-proteasome pathway and plant development.* Plant Cell, 2004. **16**(12): p. 3181-95.
203. Perrot-Rechenmann, C., *Cellular responses to auxin: division versus expansion.* Cold Spring Harbor perspectives in biology, 2010. **2**(5): p. a001446-a001446.
204. Tam, T.H., B. Catarino, and L. Dolan, *Conserved regulatory mechanism controls the development of cells with rooting functions in land plants.* Proc Natl Acad Sci U S A, 2015. **112**(29): p. E3959-68.
205. Petrásek, J. and J. Friml, *Auxin transport routes in plant development.* Development, 2009. **136**(16): p. 2675-88.
206. Abel, S. and A. Theologis, *Early Genes and Auxin Action.* Plant Physiology, 1996. **111**(1): p. 9.
207. Luo, J., J.-J. Zhou, and J.-Z. Zhang, *Aux/IAA Gene Family in Plants: Molecular Structure, Regulation, and Function.* International journal of molecular sciences, 2018. **19**(1): p. 259.
208. Paponov, I.A., et al., *Comprehensive transcriptome analysis of auxin responses in Arabidopsis.* Mol Plant, 2008. **1**(2): p. 321-37.
209. Ludwig-Müller, J., et al., *Moss (Physcomitrella patens) GH3 proteins act in auxin homeostasis.* New Phytologist, 2009. **181**(2): p. 323-338.
210. Lau, S., G. Jürgens, and I. De Smet, *The evolving complexity of the auxin pathway.* The Plant Cell, 2008. **20**(7): p. 1738-1746.
211. Plavskin, Y., et al., *Ancient trans-Acting siRNAs Confer Robustness and Sensitivity onto the Auxin Response.* Developmental Cell, 2016. **36**(3): p. 276-289.
212. Bierfreund, N.M., et al., *Loss of GH3 function does not affect phytochrome-mediated development in a moss, Physcomitrella patens.* Journal of plant physiology, 2004. **161**(7): p. 823-835.
213. Bierfreund, N., R. Reski, and E. Decker, *Use of an inducible reporter gene system for the analysis of auxin distribution in the moss Physcomitrella patens.* Plant cell reports, 2003. **21**(12): p. 1143-1152.
214. Stortenbeker, N. and M. Bemer, *The SAUR gene family: the plant's toolbox for adaptation of growth and development.* Journal of Experimental Botany, 2019. **70**(1): p. 17-27.
215. van Mourik, H., et al., *Divergent regulation of Arabidopsis SAUR genes: a focus on the SAUR10-clade.* BMC plant biology, 2017. **17**(1): p. 245.
216. Kant, S. and S. Rothstein, *Auxin-responsive SAUR39 gene modulates auxin level in rice.* Plant signaling & behavior, 2009. **4**(12): p. 1174-1175.
217. Pires, N.D., *Evolution of bHLH transcription factors that control cell differentiation in plants*, in *School of Biological Sciences*. 2010, University of East Anglia.
218. Jang, G., *Functional analysis of class 1 RSL genes in caulonema and rhizoid differentiation of Physcomitrella patens*, in *School of Biological Sciences*. 2011, University of East Anglia.
219. Jang, G., et al., *RSL genes are sufficient for rhizoid system development in early diverging land plants.* Development, 2011. **138**(11): p. 2273-81.
220. Lang, D., et al., *The Physcomitrella patens chromosome-scale assembly reveals moss genome structure and evolution.* Plant J, 2018. **93**(3): p. 515-533.
221. Desprez, T., et al., *Resistance against herbicide isoxaben and cellulose deficiency caused by distinct mutations in same cellulose synthase isoform CESA6.* Plant Physiol, 2002. **128**(2): p. 482-90.

222. Vernoux, T., et al., *The auxin signalling network translates dynamic input into robust patterning at the shoot apex.* Molecular Systems Biology, 2011. **7**(1): p. 508.
223. Boer, D.R., et al., *Structural basis for DNA binding specificity by the auxin-dependent ARF transcription factors.* Cell, 2014. **156**(3): p. 577-89.
224. Korasick, D.A., et al., *Molecular basis for AUXIN RESPONSE FACTOR protein interaction and the control of auxin response repression.* Proceedings of the National Academy of Sciences, 2014. **111**(14): p. 5427.
225. Chae, K., et al., *Arabidopsis SMALL AUXIN UP RNA63 promotes hypocotyl and stamen filament elongation.* Plant J, 2012. **71**(4): p. 684-97.
226. Gallavotti, A., *The role of auxin in shaping shoot architecture.* J Exp Bot, 2013. **64**(9): p. 2593-608.
227. Gray, W.M., et al., *Identification of an SCF ubiquitin-ligase complex required for auxin response in Arabidopsis thaliana.* Genes Dev, 1999. **13**(13): p. 1678-91.
228. Smalle, J. and R.D. Vierstra, *The ubiquitin 26S proteasome proteolytic pathway.* Annu Rev Plant Biol, 2004. **55**: p. 555-90.
229. Gray, W.M., et al., *Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins.* Nature, 2001. **414**(6861): p. 271-6.
230. Dharmasiri, N., et al., *Plant development is regulated by a family of auxin receptor F box proteins.* Dev Cell, 2005. **9**(1): p. 109-19.
231. Winkler, M., et al., *Variation in auxin sensing guides AUX/IAA transcriptional repressor ubiquitylation and destruction.* Nature Communications, 2017. **8**(1): p. 15706.
232. Esteve-Bruna, D., et al., *incurvata13, a Novel Allele of AUXIN RESISTANT6, Reveals a Specific Role for Auxin and the SCF Complex in Arabidopsis Embryogenesis, Vascular Specification, and Leaf Flatness.* Plant Physiology, 2013. **161**(3): p. 1303.
233. Shen, W.-H., et al., *Null Mutation of AtCUL1 Causes Arrest in Early Embryogenesis in Arabidopsis.* Molecular Biology of the Cell, 2002. **13**(6): p. 1916-1928.
234. Pang, P.-X., et al., *Cloning and expression analysis of the StCUL1 gene in potato.* Journal of Plant Biochemistry and Biotechnology, 2019. **28**(4): p. 460-469.
235. Bornstein, G., D. Ganoth, and A. Hershko, *Regulation of neddylation and deneddylation of cullin1 in SCFSkp2 ubiquitin ligase by F-box protein and substrate.* Proc Natl Acad Sci U S A, 2006. **103**(31): p. 11515-20.
236. Weijers, D. and D. Wagner, *Transcriptional Responses to the Auxin Hormone.* Annual Review of Plant Biology, 2016. **67**(1): p. 539-574.
237. Elzanati, O., et al., *Genome-wide analysis, classification, expression and interaction of Physcomitrella patens SKP1-like (PpSKP) and F-box (FBX) genes.* Plant Gene, 2017. **12**: p. 13-22.
238. Velasquez, S.M., et al., *Auxin and Cellular Elongation.* Plant Physiology, 2016. **170**(3): p. 1206.
239. Menand, B., et al., *An ancient mechanism controls the development of cells with a rooting function in land plants.* Science, 2007. **316**(5830): p. 1477-1480.
240. Karas, B., et al., *Conservation of lotus and Arabidopsis basic helix-loop-helix proteins reveals new players in root hair development.* Plant physiology, 2009. **151**(3): p. 1175-1185.
241. Yi, K., et al., *A basic helix-loop-helix transcription factor controls cell growth and size in root hairs.* Nature genetics, 2010. **42**(3): p. 264.
242. Bruex, A., et al., *A gene regulatory network for root epidermis cell differentiation in Arabidopsis.* PLoS Genet, 2012. **8**(1): p. e1002446.

243. Cho, H.-T. and R.D.W. Lee, *Auxin, the organizer of the hormonal/environmental signals for root hair growth.* Frontiers in Plant Science, 2013. **4**: p. 448.
244. Park, J.Y., H.J. Kim, and J. Kim, *Mutation in domain II of IAA1 confers diverse auxin-related phenotypes and represses auxin-activated expression of Aux/IAA genes in steroid regulator-inducible system.* The Plant Journal, 2002. **32**(5): p. 669-683.
245. Wilson, A.K., et al., *A dominant mutation inArabidopsis confers resistance to auxin, ethylene and abscisic acid.* Molecular and General Genetics MGG, 1990. **222**(2-3): p. 377-383.
246. Timpte, C., A.K. Wilson, and M. Estelle, *The axr2-1 mutation of Arabidopsis thaliana is a gain-of-function mutation that disrupts an early step in auxin response.* Genetics, 1994. **138**(4): p. 1239-1249.
247. Fukaki, H., et al., *Lateral root formation is blocked by a gain-of-function mutation in the SOLITARY-ROOT/IAA14 gene of Arabidopsis.* The Plant Journal, 2002. **29**(2): p. 153-168.
248. Knox, K., C.S. Grierson, and O. Leyser, *AXR3 and SHY2 interact to regulate root hair development.* Development, 2003. **130**(23): p. 5769-5777.
249. Pires, N.D., et al., *Recruitment and remodeling of an ancient gene regulatory network during land plant evolution.* Proceedings of the National Academy of Sciences, 2013. **110**(23): p. 9571-9576.
250. Jang, G. and L. Dolan, *Auxin promotes the transition from chloronema to caulonema in moss protonema by positively regulating PpRSL1and PpRSL2 in Physcomitrella patens.* New Phytologist, 2011. **192**(2): p. 319-327.
251. Kasahara, H., *Current aspects of auxin biosynthesis in plants.* Biosci Biotechnol Biochem, 2016. **80**(1): p. 34-42.
252. Bargmann, B.O. and K.D. Birnbaum, *Positive fluorescent selection permits precise, rapid, and in-depth overexpression analysis in plant protoplasts.* Plant Physiol, 2009. **149**(3): p. 1231-9.
253. Vierstra, R.D., *The ubiquitin–26S proteasome system at the nexus of plant biology.* Nature Reviews Molecular Cell Biology, 2009. **10**(6): p. 385-397.
254. Awwad, F., et al., *Auxin protects Arabidopsis thaliana cell suspension cultures from programmed cell death induced by the cellulose biosynthesis inhibitors thaxtomin A and isoxaben.* BMC Plant Biology, 2019. **19**(1): p. 512.
255. Tegg, R.S., et al., *Plant cell growth and ion flux responses to the streptomycete phytotoxin thaxtomin A: calcium and hydrogen flux patterns revealed by the non-invasive MIFE technique.* Plant and Cell Physiology, 2005. **46**(4): p. 638-648.
256. Tegg, R.S., et al., *Auxin-induced resistance to common scab disease of potato linked to inhibition of thaxtomin A toxicity.* Plant Disease, 2008. **92**(9): p. 1321-1328.
257. Duval, I. and N. Beaudoin, *Transcriptional profiling in response to inhibition of cellulose synthesis by thaxtomin A and isoxaben in Arabidopsis thaliana suspension cells.* Plant cell reports, 2009. **28**(5): p. 811-830.
258. Tegg, R.S., et al., *Enhanced resistance to the cellulose biosynthetic inhibitors, thaxtomin A and isoxaben in Arabidopsis thaliana mutants, also provides specific co-resistance to the auxin transport inhibitor, 1-NPA.* BMC plant biology, 2013. **13**: p. 76-76.
259. Ischebeck, T., et al., *Phosphatidylinositol 4, 5-bisphosphate influences PIN polarization by controlling clathrin-mediated membrane trafficking in Arabidopsis.* The Plant Cell, 2013. **25**(12): p. 4894-4911.
260. Ichikawa, M., et al., *Syntaxin of plant proteins SYP123 and SYP132 mediate root hair tip growth in Arabidopsis thaliana.* Plant and Cell Physiology, 2014. **55**(4): p. 790-800.

261. Löfke, C., et al., *Asymmetric gibberellin signaling regulates vacuolar trafficking of PIN auxin transporters during root gravitropism.* Proceedings of the National Academy of Sciences, 2013. **110**(9): p. 3627-3632.
262. Kiefer, C.S., et al., *Arabidopsis AIP1-2 restricted by WER-mediated patterning modulates planar polarity.* Development, 2015. **142**(1): p. 151-161.
263. Roux, K.J., et al., *A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells.* J Cell Biol, 2012. **196**(6): p. 801-10.
264. Gonneau, M., et al., *Catalytic subunit stoichiometry within the cellulose synthase complex.* Plant Physiol, 2014. **166**(4): p. 1709-12.
265. Song, D., J. Shen, and L. Li, *Characterization of cellulose synthase complexes in Populus xylem differentiation.* New Phytol, 2010. **187**(3): p. 777-90.
266. Shoemaker, B.A. and A.R. Panchenko, *Deciphering Protein–Protein Interactions. Part I. Experimental Techniques and Databases.* PLOS Computational Biology, 2007. **3**(3): p. e42.
267. Roux, K.J., D.I. Kim, and B. Burke, *BioID: A Screen for Protein-Protein Interactions*, in *Current Protocols in Protein Science*. 2001, John Wiley & Sons, Inc.
268. Chapman-Smith, A. and J.E. Cronan Jr, *Molecular Biology of Biotin Attachment to Proteins.* The Journal of Nutrition, 1999. **129**(2): p. 477S-484S.
269. Kim, D.I., et al., *Probing nuclear pore complex architecture with proximity-dependent biotinylation.* Proc Natl Acad Sci U S A, 2014. **111**(24): p. E2453-61.
270. van Steensel, B. and S. Henikoff, *Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase.* Nat Biotechnol, 2000. **18**(4): p. 424-8.
271. Kim, D.I., et al., *An improved smaller biotin ligase for BioID proximity labeling.* Mol Biol Cell, 2016. **27**(8): p. 1188-96.
272. Roux, K.J., *Marked by association: techniques for proximity-dependent labeling of proteins in eukaryotic cells.* Cell Mol Life Sci, 2013. **70**(19): p. 3657-64.
273. Mattiuzzo, M., et al., *Abnormal kinetochore-generated pulling forces from expressing a N-terminally modified Hec1.* PloS one, 2011. **6**(1): p. e16307-e16307.
274. Ummethum, H. and S. Hamperl, *Proximity Labeling Techniques to Study Chromatin.* Frontiers in Genetics, 2020. **11**(450).
275. Hurley, B., et al., *Proteomics of effector-triggered immunity (ETI) in plants.* Virulence, 2014. **5**(7): p. 752-760.
276. Roebroek, A.J.M., P.L.S.M. Gordts, and S. Reekmans, *Knock-In Approaches*, in *Transgenic Mouse Methods and Protocols*, M.H. Hofker and J. van Deursen, Editors. 2011, Humana Press: Totowa, NJ. p. 257-275.
277. Lukowitz, W., C.S. Gillmor, and W.-R. Scheible, *Positional Cloning in Arabidopsis. Why It Feels Good to Have a Genome Initiative Working for You.* Plant Physiology, 2000. **123**(3): p. 795.
278. Gibson, D.G., et al., *Enzymatic assembly of DNA molecules up to several hundred kilobases.* Nature Methods, 2009. **6**(5): p. 343-345.
279. Gibson, D.G., et al., *Creation of a bacterial cell controlled by a chemically synthesized genome.* Science, 2010. **329**(5987): p. 52-6.
280. Merryman, C. and D.G. Gibson, *Methods and applications for assembling large DNA constructs.* Metab Eng, 2012. **14**(3): p. 196-204.
281. Sambrook, J. and D.W. Russell, *The condensed protocols from molecular cloning: a laboratory manual*. 2006.
282. Gibson, D.G., *Enzymatic assembly of overlapping DNA fragments.* Methods Enzymol, 2011. **498**: p. 349-61.

283. Zheng, X., et al., *The cauliflower mosaic virus (CaMV) 35S promoter sequence alters the level and patterns of activity of adjacent tissue- and organ-specific gene promoters.* Plant Cell Rep, 2007. **26**(8): p. 1195-203.
284. Rekas, A., et al., *Crystal structure of venus, a yellow fluorescent protein with improved maturation and reduced environmental sensitivity.* J Biol Chem, 2002. **277**(52): p. 50573-8.
285. Akbudak, M.A. and V. Srivastava, *Improved FLP Recombinase, FLPe, Efficiently Removes Marker Gene from Transgene Locus Developed by Cre–lox Mediated Site-Specific Gene Integration in Rice.* Molecular Biotechnology, 2011. **49**(1): p. 82-89.
286. Werdien, D., G. Peiler, and G.U. Ryffel, *FLP and Cre recombinase function in Xenopus embryos.* Nucleic acids research, 2001. **29**(11): p. E53-e53.
287. Wu, S.Z. and M. Bezanilla, *Myosin VIII associates with microtubule ends and together with actin plays a role in guiding plant cell division.* Elife, 2014. **3**.
288. Schaefer, D., et al., *Stable transformation of the moss Physcomitrella patens.* Molecular and General Genetics MGG, 1991. **226**(3): p. 418-424.
289. Chomczynski, P. and K. Mackey, *Short technical reports. Modification of the TRI reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources.* Biotechniques, 1995. **19**(6): p. 942-5.
290. Conlan, B., et al., *Development of a Rapid in planta BioID System as a Probe for Plasma Membrane-Associated Immunity Proteins.* Frontiers in Plant Science, 2018. **9**(1882).
291. William Roy, S. and W. Gilbert, *The evolution of spliceosomal introns: patterns, puzzles and progress.* Nature Reviews Genetics, 2006. **7**(3): p. 211-221.
292. Sibley, C.R., et al., *Recursive splicing in long vertebrate genes.* Nature, 2015. **521**(7552): p. 371-375.
293. Duff, M.O., et al., *Genome-wide identification of zero nucleotide recursive splicing in Drosophila.* Nature, 2015. **521**(7552): p. 376-379.
294. Woo, H.-J., et al., *Auto-excision of selectable marker genes from transgenic tobacco via a stress inducible FLP/FRT site-specific recombination system.* Transgenic Research, 2009. **18**(3): p. 455-465.
295. Zhu, X.D. and P.D. Sadowski, *Cleavage-dependent ligation by the FLP recombinase. Characterization of a mutant FLP protein with an alteration in a catalytic amino acid.* J Biol Chem, 1995. **270**(39): p. 23044-54.
296. Kwon, K. and D. Beckett, *Function of a conserved sequence motif in biotin holoenzyme synthetases.* Protein science : a publication of the Protein Society, 2000. **9**(8): p. 1530-1539.
297. Lane, M.D., et al., *THE ENZYMATIC SYNTHESIS OF HOLOTRANSCARBOXYLASE FROM APOTRANSCARBOXYLASE AND (+)-BIOTIN. II. INVESTIGATION OF THE REACTION MECHANISM.* J Biol Chem, 1964. **239**: p. 2865-71.
298. Choi-Rhee, E., H. Schulman, and J.E. Cronan, *Promiscuous protein biotinylation by Escherichia coli biotin protein ligase.* Protein Science : A Publication of the Protein Society, 2004. **13**(11): p. 3043-3050.
299. Cronan, J.E., *Targeted and proximity-dependent promiscuous protein biotinylation by a mutant Escherichia coli biotin protein ligase.* J Nutr Biochem, 2005. **16**(7): p. 416-8.
300. Kim, D.I. and K.J. Roux, *Filling the Void: Proximity-Based Labeling of Proteins in Living Cells.* Trends Cell Biol, 2016. **26**(11): p. 804-817.
301. Li, P., et al., *Proximity Labeling of Interacting Proteins: Application of BioID as a Discovery Tool.* Proteomics, 2017. **17**(20).

302. Lam, S.S., et al., *Directed evolution of APEX2 for electron microscopy and proximity labeling.* Nat Methods, 2015. **12**(1): p. 51-4.
303. Branon, T.C., et al., *Efficient proximity labeling in living cells and organisms with TurboID.* Nature Biotechnology, 2018. **36**(9): p. 880-887.
304. Kido, K., et al., *AirID, a novel proximity biotinylation enzyme, for analysis of protein-protein interactions.* Elife, 2020. **9**.
305. Ashton, N.W. and D.J. Cove, *The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants of the moss, Physcomitrella patens.* Molecular and General Genetics MGG, 1977. **154**(1): p. 87-95.
306. Engel, P.P., *The induction of biochemical and morphological mutants in the moss Physcomitrella patens.* American Journal of Botany, 1968. **55**(4): p. 438-446.
307. Schuster, R.M., *New manual of bryology.* 1983.
308. Knight, C.D., et al., *Molecular responses to abscisic acid and stress are conserved between moss and cereals.* The Plant Cell, 1995. **7**(5): p. 499-506.
309. Machuka, J., et al., *Sequence analysis of expressed sequence tags from an ABA-treated cDNA library identifies stress response genes in the moss Physcomitrella patens.* Plant and Cell Physiology, 1999. **40**(4): p. 378-387.
310. Khandelwal, A., et al., *Role of ABA and ABI3 in desiccation tolerance.* Science, 2010. **327**(5965): p. 546-546.
311. Richardt, S., et al., *Microarray analysis of the moss Physcomitrella patens reveals evolutionarily conserved transcriptional regulation of salt stress and abscisic acid signalling.* Plant molecular biology, 2010. **72**(1): p. 27-45.
312. Schaefer, D.G. and J.P. Zrÿd, *Efficient gene targeting in the moss Physcomitrella patens.* The Plant Journal, 1997. **11**(6): p. 1195-1206.
313. Girke, T., et al., *Identification of a novel D6-acyl-group desaturase by targeted gene disruption in Physcomitrella patens.* The Plant Journal, 1998. **15**(1): p. 39-48.
314. Rensing, S.A., et al., *Moss transcriptome and beyond.* Trends in Plant Science, 2002. **7**(12): p. 535-538.
315. Reski, R. and D.J. Cove, *Physcomitrella patens.* Current Biology, 2004. **14**(7): p. R261-R262.
316. Frank, W., E.L. Decker, and R. Reski, *Molecular tools to study Physcomitrella patens.* Plant Biol (Stuttg), 2005. **7**(3): p. 220-7.
317. Rensing, S.A., et al., *The Moss Physcomitrium (Physcomitrella) patens: A Model Organism for Non-Seed Plants.* The Plant cell, 2020. **32**(5): p. 1361-1376.
318. Bezanilla, M., A. Pan, and R.S. Quatrano, *RNA interference in the moss Physcomitrella patens.* Plant physiology, 2003. **133**(2): p. 470-474.
319. Nakaoka, Y., et al., *An inducible RNA interference system in Physcomitrella patens reveals a dominant role of augmin in phragmoplast microtubule generation.* The Plant Cell, 2012. **24**(4): p. 1478-1493.
320. Augustine, R.C., et al., *Actin depolymerizing factor is essential for viability in plants, and its phosphoregulation is important for tip growth.* The Plant Journal, 2008. **54**(5): p. 863-875.
321. Kamisugi, Y., et al., *The mechanism of gene targeting in Physcomitrella patens: homologous recombination, concatenation and multiple integration.* Nucleic Acids Res, 2006. **34**(21): p. 6205-14.
322. Perroud, P.F., et al., *An experimental method to facilitate the identification of hybrid sporophytes in the moss Physcomitrella patens using fluorescent tagged lines.* New Phytologist, 2011. **191**(1): p. 301-306.

323. Perroud, P.F., R. Meyberg, and S. Rensing, *Physcomitrella patens Reute mC herry as a tool for efficient crossing within and between ecotypes.* Plant Biology, 2019. **21**: p. 143-149.
324. Ding, X., et al., *Conditional genetic screen in Physcomitrella patens reveals a novel microtubule depolymerizing-end-tracking protein.* PLoS genetics, 2018. **14**(5): p. e1007221.
325. Moody, L.A., et al., *Genetic Regulation of the 2D to 3D Growth Transition in the Moss Physcomitrella patens.* Curr Biol, 2018. **28**(3): p. 473-478.e5.
326. Newman, D.J. and G.M. Cragg, *Natural products as sources of new drugs over the last 25 years.* J Nat Prod, 2007. **70**(3): p. 461-77.
327. Ikram, N.K.K., et al., *Insights into Heterologous Biosynthesis of Arteannuin B and Artemisinin in Physcomitrella patens.* Molecules (Basel, Switzerland), 2019. **24**(21): p. 3822.
328. Makunga, N.P., A.K. Jäger, and J. van Staden, *Micropropagation of Thapsia garganica—a medicinal plant.* Plant Cell Reports, 2003. **21**(10): p. 967-973.
329. Denmeade, S.R., et al., *Prostate-specific antigen-activated thapsigargin prodrug as targeted therapy for prostate cancer.* J Natl Cancer Inst, 2003. **95**(13): p. 990-1000.
330. Janssen, S., et al., *Pharmacokinetics, biodistribution, and antitumor efficacy of a human glandular kallikrein 2 (hK2)-activated thapsigargin prodrug.* The Prostate, 2006. **66**(4): p. 358-368.
331. Ekiert, H., *Medicinal plant biotechnology: the Apiaceae family as the example of rapid development.* Pharmazie, 2000. **55**(8): p. 561-7.
332. Makunga, N.P., A.K. Jäger, and J. van Staden, *Improved in vitro rooting and hyperhydricity in regenerating tissues of Thapsia garganica L.* Plant Cell, Tissue and Organ Culture, 2006. **86**(1): p. 77-86.
333. Ball, M., et al., *Total synthesis of thapsigargin, a potent SERCA pump inhibitor.* Org Lett, 2007. **9**(4): p. 663-6.
334. Cragg, G.M. and D.J. Newman, *Plants as a source of anti-cancer agents.* J Ethnopharmacol, 2005. **100**(1-2): p. 72-9.
335. Paramasivan, K. and S. Mutturi, *Progress in terpene synthesis strategies through engineering of Saccharomyces cerevisiae.* Crit Rev Biotechnol, 2017. **37**(8): p. 974-989.
336. Withers, S.T. and J.D. Keasling, *Biosynthesis and engineering of isoprenoid small molecules.* Appl Microbiol Biotechnol, 2007. **73**(5): p. 980-90.
337. Kirby, J. and J.D. Keasling, *Metabolic engineering of microorganisms for isoprenoid production.* Nat Prod Rep, 2008. **25**(4): p. 656-61.
338. Zhan, X., et al., *Metabolic engineering of the moss Physcomitrella patens to produce the sesquiterpenoids patchoulol and α/β-santalene.* Frontiers in Plant Science, 2014. **5**(636).
339. Reski, R., H. Bae, and H.T. Simonsen, *Physcomitrella patens, a versatile synthetic biology chassis.* Plant Cell Reports, 2018. **37**(10): p. 1409-1417.
340. Reski, R., *Physcomitrella and Arabidopsis: the David and Goliath of reverse genetics.* Trends in Plant Science (United Kingdom), 1998.
341. Banerjee, A., et al., *Engineering modular diterpene biosynthetic pathways in Physcomitrella patens.* Planta, 2019. **249**(1): p. 221-233.
342. Decker, E.L. and R. Reski, *Moss bioreactors producing improved biopharmaceuticals.* Curr Opin Biotechnol, 2007. **18**(5): p. 393-8.
343. Frank, W., D. Ratnadewi, and R. Reski, *Physcomitrella patens is highly tolerant against drought, salt and osmotic stress.* Planta, 2005. **220**(3): p. 384-94.

344. Floss, H.G., *Combinatorial biosynthesis—Potential and problems.* Journal of Biotechnology, 2006. **124**(1): p. 242-257.
345. Piel, J., *Combinatorial biosynthesis in symbiotic bacteria.* Nature Chemical Biology, 2006. **2**(12): p. 661-662.
346. Khairul Ikram, N.K.B., et al., *Stable Production of the Antimalarial Drug Artemisinin in the Moss Physcomitrella patens.* Front Bioeng Biotechnol, 2017. **5**: p. 47.
347. Vučković, S., et al., *Cannabinoids and Pain: New Insights From Old Molecules.* Frontiers in pharmacology, 2018. **9**: p. 1259-1259.
348. Manzanares, J., M. Julian, and A. Carrascosa, *Role of the cannabinoid system in pain control and therapeutic implications for the management of acute and chronic pain episodes.* Current neuropharmacology, 2006. **4**(3): p. 239-257.
349. Häuser, W., et al., *Cannabinoids in Pain Management and Palliative Medicine.* Deutsches Arzteblatt international, 2017. **114**(38): p. 627-634.
350. Stout, J.M., et al., *The hexanoyl-CoA precursor for cannabinoid biosynthesis is formed by an acyl-activating enzyme in Cannabis sativa trichomes.* The Plant Journal, 2012. **71**(3): p. 353-365.
351. Ibrahim, E.A., et al., *Determination of Acid and Neutral Cannabinoids in Extracts of Different Strains of Cannabis sativa Using GC-FID.* Planta Med, 2018. **84**(4): p. 250-259.
352. Schween, G., et al., *Large-Scale Analysis of 73 329 Physcomitrella Plants Transformed with Different Gene Disruption Libraries: Production Parameters and Mutant Phenotypes.* Plant Biology, 2005. **7**(3): p. 228-237.
353. Folger, K.R., et al., *Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules.* Mol Cell Biol, 1982. **2**(11): p. 1372-87.
354. Tholl, D., *Biosynthesis and biological functions of terpenoids in plants.* Adv Biochem Eng Biotechnol, 2015. **148**: p. 63-106.
355. Chen, F., et al., *Terpenoid secondary metabolites in bryophytes: chemical diversity, biosynthesis and biological functions.* Critical Reviews in Plant Sciences, 2018. **37**(2-3): p. 210-231.
356. Hanuš, L.O. and Y. Hod, *Terpenes/Terpenoids in Cannabis: Are They Important?* Medical Cannabis and Cannabinoids, 2020. **3**(1): p. 25-60.
357. ElSohly, M.A., et al., *Changes in Cannabis Potency Over the Last 2 Decades (1995-2014): Analysis of Current Data in the United States.* Biological psychiatry, 2016. **79**(7): p. 613-619.
358. Britt, A.B. and G.D. May, *Re-engineering plant gene targeting.* Trends Plant Sci, 2003. **8**(2): p. 90-5.
359. Jongsma, M., et al., *Tomato protoplast DNA transformation: physical linkage and recombination of exogenous DNA sequences.* Plant Molecular Biology, 1987. **8**(5): p. 383-394.
360. Ashton, N.W., et al., *The bryophyte Physcomitrella patens replicates extrachromosomal transgenic elements.* New Phytologist, 2000. **146**(3): p. 391-402.
361. Hohe, A. and R. Reski, *Optimisation of a bioreactor culture of the moss Physcomitrella patens for mass production of protoplasts.* Plant Science, 2002. **163**(1): p. 69-74.