Utilising CRISPR/Cas9 Mutagenesis to Understand Stress-Induced Flavonoid Biosynthesis and Regulation in *Marchantia polymorpha*

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science in Biological Sciences

at Massey University, Palmerston North, New Zealand

Rebecca Maree Yorker

2023

Abstract

Flavonoids are secondary metabolites that evolved during land plant colonisation to provide protection against the plethora of abiotic and biotic stresses present in a terrestrial environment. The bryophytes lineage is thought have arisen at the first branching event during evolution of the extant land plants, making them the sister group to all other land plants. Moreover, extant bryophytes retain features in common with early land plant ancestors. Together, these aspects make bryophytes particularly important for understanding land plant evolution. While the biosynthesis and regulation of flavonoids in angiosperms (flowering plants) is well understood, this knowledge is lacking for bryophytes. Thus, there is growing interest in using *Marchantia polymorpha* (hereafter, Marchantia), a model species for bryophytes, in comparative studies with angiosperms to investigate the conservation of the flavonoid biosynthetic pathway among land plants.

Marchantia produces two main flavonoids under stress conditions: colourless flavones that function in UV-B light protection and a cell wall-bound red pigment called auronidin. The Marchantia *Flavone synthase* (Mp*FNSI*) and *BTB-TAZ Domain 2* (Mp*BT2*) candidate genes were hypothesised to be involved in flavone biosynthesis and auronidin regulation, respectively. Gene function was assessed by using CRISPR/Cas9 mutagenesis and *Agrobacterium tumefaciens*-mediated Marchantia spore transformation to generate *fnsI* and *bt2* loss-of-function mutants and characterising their ability to produce flavonoids under abiotic stress conditions. The *fnsI* mutants did not produce flavones, confirming that MpFNSI is required for flavone biosynthesis. Flavonoid analysis of *bt2* mutants indicated that MpBT2 does not regulate auronidin production, as no significant differences in auronidin content were observed compared to wild-type control plants.

Functional characterisation of MpFNSI and MpBT2 contributed to understanding of the conservation of the flavonoid biosynthetic pathway among land plants. Flavones can be produced by two types of FNS enzymes in flowering plants, so identifying that Marchantia uses FNSI established a conserved function of 2-oxoglutarate-dependent dioxygenase-type enzymes in flavone biosynthesis among land plants. The retention of FNS throughout land plant evolution also emphasises the importance of flavones in protecting plants against damaging UV-B light in a terrestrial environment. Identifying that MpBT2 does not regulate auronidin production in Marchantia established differences in the regulation of red pigmentation among land plants, because the production of anthocyanins is regulated by BT2 in flowering plants. This proposes that the regulation of flavonoids by BT2 proteins was acquired with the evolution of anthocyanins after the divergence from bryophytes.

Acknowledgements

I firstly thank my supervisors Dr Nick Albert, Dr Kevin Davies and Associate Professor Jennifer Tate. This research would not have been possible without your continuous support and encouragement, and I cannot express how grateful I am for this. Special thanks go to Nick. I am incredibly thankful for your patience, help with troubleshooting, and tireless efforts and enthusiasm to impart your knowledge. Kevin, your expertise and high-level discussions have been invaluable. The guidance both of you have provided has greatly contributed to the success of this project. Jen, you kept me on track and provided support and advice when I needed it, and for this I am grateful.

I would also like to thank Plant and Food Research for allowing this research to be undertaken and the 'Metabolite Traits in Plants' team for providing sound advice and encouragement. In particular, I thank Dr Yanfei Zhou who helped with the original development of this project and continued to provide useful insight. I have made good friends at Plant and Food Research. To Jasmine, Chethi, Henry and Nathanael, I value the time spent with you both at and outside of work, and I am truly grateful for the immense support you have given me throughout this thesis.

I thank Sarah Cordiner, Dr Tony McGhie, Dr John van Klink, Dr Nigel Perry and Dr Christelle Andre for performing chemical analysis, aiding in compound identification and evocating helpful discussions.

I am thankful to have had my fees funded by a James Cook Research Fellowship (JCF-PAF2001) awarded to Kevin Davies (Plant and Food Research). I also thank Massey University and associated sponsors for providing me with the "Sir Alan Stewart Postgraduate Scholarship", "Murray Richards Bursary" and "Hudson and McGowan Scholarship" scholarships.

Lastly, I thank my parents Nadine and Simon, my siblings Kristen and Nick, and other family and friends for supporting me on this journey. You have been there for me during the highs and the lows, and I would not have been able to complete this thesis without you.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2-OGD	2-oxoglutarate-dependent dioxygenase/hydroxylase enzyme superfamily
4CL	4-coumarate:CoA ligase
A	absorbance
ANS	anthocyanidin synthase
AS	acetosyringone
bHLH	basic helix-loop-helix
bp	base pairs
BT2	bric-a-brac, tramtrack and broad (BTB) – transcriptional adapter zinc finger (TAZ) domain 2
CaMV35S	35S promoter from the cauliflower mosaic virus (CaMV)
C4H	cinnamate 4-hydroxylase
CDS	coding sequence
CHI	chalcone isomerase
CHIL	chalcone isomerase-like
CHS	chalcone synthase
Cm	centimetres
crRNA	CRISPR RNA
CRISPR/Cas9	clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9
CTAB	cetyltrimethylammonium bromide
CYP450	cytochrome P450 monooxygenase enzyme superfamily
°C	degrees in Celsius

d	days
DNA	deoxyribonucleic acid
dNTP	deoxy-nucleotide-triphosphate
DW	dry weight
EDTA	ethylenediaminetetraacetic acid
EF1α	elongation factor 1-alpha
F2H	flavanone 2-hydroxylase
F3'H	flavonoid 3'-hydroxylase
F3H	flavanone 3-hydroxylase
FLS	flavonol synthase
FNS	flavone synthase
FW	fresh weight
g	grams
GFP	green fluorescence protein
GLK	golden 2-like
h	hours
HPLC	high performance liquid chromatography
HS	heat-shock treatment
kb	kilobases
LB	lysogeny broth; bacterial growth media
ng	nanogram
nm	nanometre
M	molar; moles per litre
m/z	mass-to-charge ratio
min	minute
µL	microlitre
µM	micromolar
mg	milligram
mL	millilitre
mm	millimetre
mM	millimolar
NOS	nopaline synthase
OCS	octopine synthase
PAL	phenylalanine ammonia-lyase

PAM	protospacer motif
PCR	polymerase chain reaction
PDS	phytoene desaturase
PPO	polyphenol oxidase
RNA	ribonucleic acid
ROS	reactive oxygen species
rpm	revolutions per minute
s	seconds
sgRNA	single guide RNA
T-DNA	transfer DNA; DNA transferred from *Agrobacterium tumefaciens* into the genome of host plant cells
TBE	tris-borate EDTA buffer
tracrRNA	trans-activating CRISPR RNA
Tris	*tris*(hydroxymethyl)aminomethane
tRNA	transfer RNA
U	enzyme units
UV	ultra violet
v/v	volume/volume
w/v	weight/volume
WT	wild-type

- Gene names are italicised; full names are lower case (e.g., *Flavone Synthase I*) and abbreviations are capitalised (e.g., *FNSI*)

- Gene mutants are italicised and all lower case (e.g., *fnsI*)

- Proteins are capitalised and not italicised (e.g., FLAVONE SYNTHASE I)

Chapter 1

Introduction

1.1 Land plant evolution

Land plants, or embryophytes, are a group of plants whose ancestors colonised land 470-500 million years ago, transitioning from living in water to the terrestrial environment (Morris et al., 2018; Wodniok et al., 2011). Land plant colonisation shaped the earth's environment and ecosystems in many major aspects, including contributing to global cooling by reducing atmospheric carbon dioxide (CO_2; a greenhouse gas) and increasing atmospheric oxygen (Pires & Dolan, 2012; Quirk et al., 2015). Additionally, plant productivity paved the way for the evolution of ecosystems on land and thereby the evolution of terrestrial animal life over 400 million years ago, which eventually lead to the evolution of humans (Garwood & Edgecombe, 2011; Pires & Dolan, 2012).

Land plants comprise two main lineages, the non-vascular (non-tracheophyte) bryophytes: liverworts, hornworts and mosses, and the vascular plants (tracheophytes): lycophytes, ferns, gymnosperms (conifers) and angiosperms (flowering plants), as shown in Figure 1.1 (Bowman et al., 2017; Davies et al., 2020). These two lineages separated early during land plant evolution, so bryophytes are the sister lineage to all other land plants. The land plant lineage is thought to have evolved from an ancestor of the extant charophyte green algae, which are mainly aquatic but include some terrestrial species (Karsten & Holzinger, 2014). Together, the land plants and charophytes form a monophyletic group (a group of plants that are all of the descendants of a single ancestor) called streptophytes that is sister to the chlorophytes, another ancestral group of freshwater green algae (Becker & Marin, 2009).

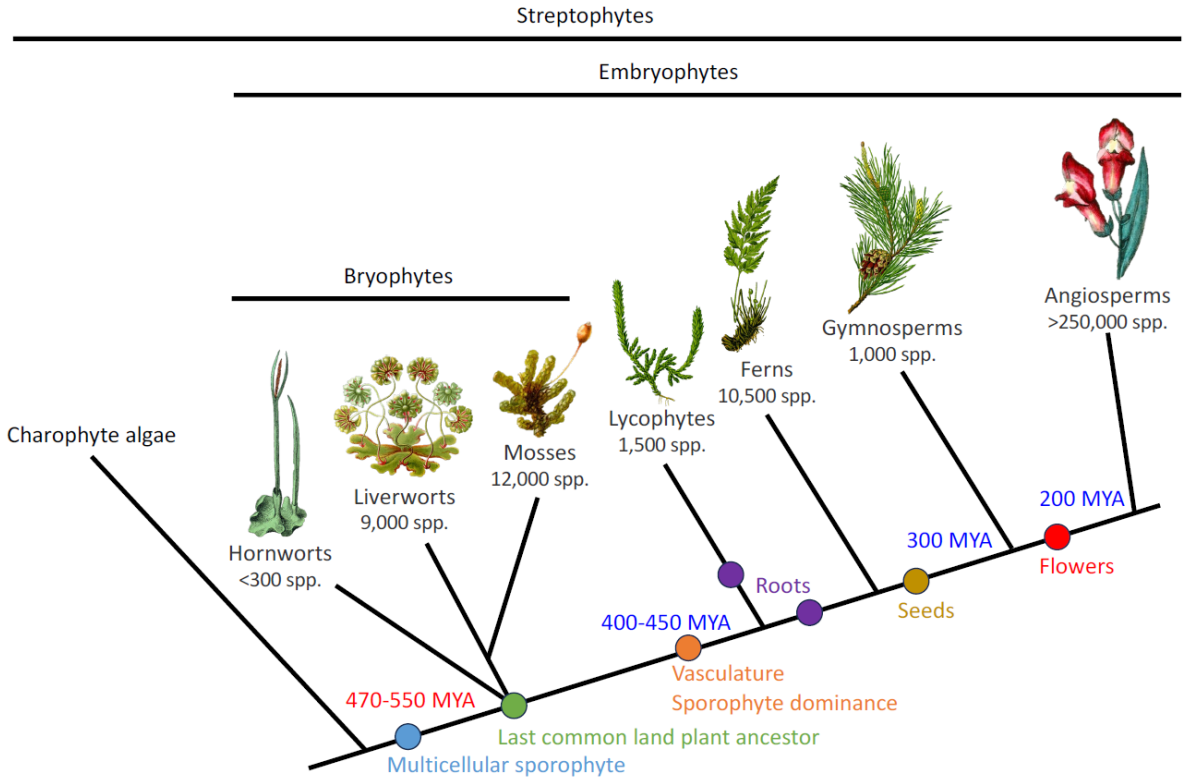


Figure 1.1. Phylogeny illustrating the evolution of land plants. Land plants (embryophytes) are thought to have evolved from an ancestor of the extant charophytes (green algae), which together with embryophytes form the streptophytes. Bryophytes form a monophyletic group that is the sister lineage to all other land plants, and thus the last common land plant ancestor (indicated by a green circle) is at the branch point of bryophytes and the lineage leading to angiosperms (flowering plants). (Figure provided by Kevin Davies.)

During land plant evolution, there was extensive plant diversification and generation of novel plant characteristics that allowed plants to survive in a range of different terrestrial environments. This diversification was driven by several historic evolutionary processes. These included adaptive radiation (the rapid emergence of distinct plant groups from ancestral species in response to environmental changes) (Jønsson et al., 2012) and the occurrence of whole genome duplication (WGD) events followed by subsequent genome reorganisation, gene loss and neofunctionalization (one gene copy generates a new function) (Crow & Wagner, 2005; Mühlhausen & Kollmar, 2013). All of these events contributed to the increase in morphological and genetic diversification and complexity that occurred in the lineage that gave rise to vascular plants (Hagolani et al., 2021). As a result, angiosperms have considerable species diversification, comprising approximately 90% of the world's extant flora (Condamine et al., 2020; Tank et al., 2015). Comparatively, the extant bryophytes appear to have undergone less major diversification as they have relatively little species diversity and seemingly simple plant structures, with less duplicated genomes and smaller total gene

numbers (Kenrick & Crane, 1997). Albeit this may also be a result of bryophytes having a dominant haploid genome life phase (Ishizaki, 2017).

Transition to a terrestrial environment entailed abiotic and biotic stresses that were not present in the aquatic environment of the land plant ancestors. This drove the evolution of morphological and physiological characteristics that enabled plants to recognise and respond to such stressors to facilitate their survival and reproduction. The current proposal is that the major innovation underpinning the early success of land plants was the evolution of an "alternation of generations", where haploid gametophytes form diploid sporophytes that produce offspring by sexual reproduction and thereby regenerate the gametophyte (Bennici, 2008). This strategy, a hallmark of embryophytes, is advantageous because the zygote is protected within the archegonia of the female sporophyte. Additionally, it allows the formation of desiccation resistant propagules, such as spores or seeds, that can travel large distances and survive seasonal changes (Bowman et al., 2017; Niklas & Kutschera, 2010). Bryophytes have a dominant gametophyte life stage, whereas vascular plants have a dominant sporophyte stage (Bowman, 2022). The alternation of generations is one of a complex of characters that evolved and facilitated land colonisation. In extant vascular plants, roots are used for structural support and water and nutrient uptake; stomata can regulate gas exchange and control water loss; lignified vascular tissue facilitates the long-distance transport of water, nutrients and sugars within the plant; an extracellular hydrophobic cuticle protects plants from desiccation and UV light; and a great diversity of specialised small metabolites protects plants against abiotic and biotic stressors (Arteaga-Vazquez, 2016; Kong et al., 2020; Kulshrestha et al., 2022; Ligrone et al., 2012). However, bryophytes lack some of these characteristics. They do not produce lignified vascular tissue, rather absorbing water and nutrients on their thalloid surfaces and diffusing it throughout the plant (Turetsky, 2003). Stomata are lacking on all bryophyte gametophytic structures, instead using permanently open air pores on their thalloid surfaces to conduct gas exchange. Some hornworts and mosses have pseudostomata on sporophyte structures, but the function of these is debated, and they are absent in liverworts (Harris et al., 2020).

Determining what features the first land plants may have gained during land colonisation is crucial because it can improve our understanding of how they survived the various stresses present in a terrestrial environment. This knowledge is also invaluable in the present day, because climate change is forcing plants to respond to an uncertain environment. Thus, knowing the characteristics that evolved during land colonisation to facilitate plant tolerance to stress conditions may provide insight into how we can ensure plants survive the changing climate in the future. Angiosperms are often the subject of studies trying to understand abiotic stress tolerance, since they are abundant,

diverse and include commercially important crops (Soltis et al., 2007; Yang et al., 2020). Whereas, other land plant lineages are comparatively understudied, despite the fact they are useful for understanding land plant colonisation and contain characteristics absent from angiosperms.

1.2 *Marchantia polymorpha* as a model species

Marchantia is the only genus in the family Marchantiaceae, which is part of the order Marchantiales that comprises the "complex" thalloid liverworts. *Marchantia* is widely distributed world-wide, but the main species found in New Zealand are *M. polymorpha*, *M. foliacea*, *M. berteroana* and *M. macropora* (Campbell et al., 1979). *M. polymorpha* is the most commonly used species for genetic studies. There are three *M. polymorpha* subspecies: *polymorpha*, *ruderalis* and *montivagens* (Linde et al., 2020). *M. polymorpha* naturally grows on moist, shaded soils or rocks, and has a branched, flat, non-leafy thalloid plant body. It reproduces asexually through clonal gemmae produced in gemmae cups and sexually through spores contained within the sporangia generated on the archegoniophores (female sporophyte), following fertilisation of the egg from sperm contained on the antheridiophores (male sporophyte). Once sporangia are formed, they burst open to release the spores (Figure 1.2). The plants are dioicous, meaning gametangia (gamete-producing structures) form on separate male and female plants (Bowman et al., 2022; Shimamura, 2015).

The *Marchantia* species used in this study is *M. polymorpha* subsp. *ruderalis* (hereafter, Marchantia). Marchantia has recently become a model species for the bryophyte lineage (Bowman et al., 2022), partly because of its dominant haploid gametophytic phase which has one copy of each chromosome. This makes targeted gene mutagenesis simpler because mutations do not have to be made in multiple alleles of a gene (Sugano et al., 2018). Comparatively, other model bryophyte species such as *Physcomitrium* (*Physcomitrella*) *patens* (a moss) are not as useful for mutagenesis because their genomes have undergone a relatively recent whole genome duplication, resulting in families of very similar (duplicated) genes (Lang et al., 2016). Additionally, Marchantia is a good model system because the plants are small, fast growing with a short life cycle, well-suited to tissue culture and easily reproduced as large numbers of clonal plants of the same developmental stage. A fully annotated reference genome has also been generated (Bowman et al., 2022; Bowman et al., 2017; Shimamura, 2015).

Figure 1.2. *Marchantia polymorpha* sporangia formation. The archegoniophore (female gametangium) contains archegonia that house eggs and the antheridiophore (male gametangium) produces sperm on top of the antheridia. Fertilisation of eggs by the sperm generates sporangia (appears yellow) on the archegoniophore that rupture to release spores (right). A gemmae cup on the thallus (gametophyte), containing gemmae, is also indicated. (Image on the right provided by Kevin Davies.)

Based on fossil records, bryophytes have retained morphological features in common with early land plant ancestors (Bowman et al., 2022; Kenrick & Crane, 1997), so Marchantia is useful in comparative studies with other land plant species to determine features of ancestral land plants that could have been important during land plant colonisation. The utilisation of Marchantia as a model species has already led to important discoveries in land plant biology, including regarding their stress responses, evolution, and growth and development (Bowman et al., 2022). In recent years, significant advances have been made in understanding flavonoid biosynthesis and regulation in Marchantia. However, very little is known compared to research on angiosperms and there are still important unanswered questions.

1.3 Flavonoid production

Flavonoids are an incredibly diverse group of plant secondary metabolites that have various biological functions (Figure 1.3). They are hypothesised to have evolved during land plant colonisation as a mechanism of coping with abiotic and biotic stresses present in the terrestrial environment, which plants living a submerged lifestyle are not exposed to. Thus, the pathway is unique to land plants (Davies et al., 2020; Li et al., 2020a). Flavonoids are broadly classified based on their chemical structure into groups such as anthocyanins, proanthocyanins, flavonols, flavones, isoflavones, aurones and the recently identified auronidins. These core compounds can undergo secondary chemical modifications, such as glycosylation, methylation and acylation, creating huge diversity within each class (Alseekh et al., 2020; Stracke et al., 2007).

1.3.1 Function and evolution of flavonoids

By far the most well-known flavonoids are the anthocyanins, which confer colours ranging from orange to blue and are predominantly reported from the vascular plants gymnosperms and angiosperms (Campanella et al., 2014). The colour resulting from the presence of anthocyanins in tissues depends on a combination of anthocyanin structure, including the number of hydroxyls and substituted groups, and the cellular environment (Hurtado et al., 2009; Kähkönen & Heinonen, 2003; Landi et al., 2015). Anthocyanins provide colour to many flowers and fruit, thereby functioning in pollinator attraction and seed dispersal (Albert et al., 2022; Koes et al., 2005; Winkel-Shirley, 2002). Additionally, they are involved in leaf senescence and provide tolerance to various abiotic stresses, including light, nutrient deprivation, cold, drought and metal ions (An et al., 2020b; Landi et al., 2015; Liang & He, 2018; Zhang et al., 2012). Anthocyanins function in photoprotection by absorbing excess white light that would otherwise cause damage to the chloroplasts and photosystems by the generation of reactive oxygen species (ROS), inhibiting photosynthesis. However, they can also act as plant antioxidants to reduce cellular stress under excess white light conditions by scavenging these ROS (Agati et al., 2012; Gould et al., 2002; Landi et al., 2015; Pietrini et al., 2002). Which of these two roles is the primary function of anthocyanins has been a topic of much controversy over the last several decades.

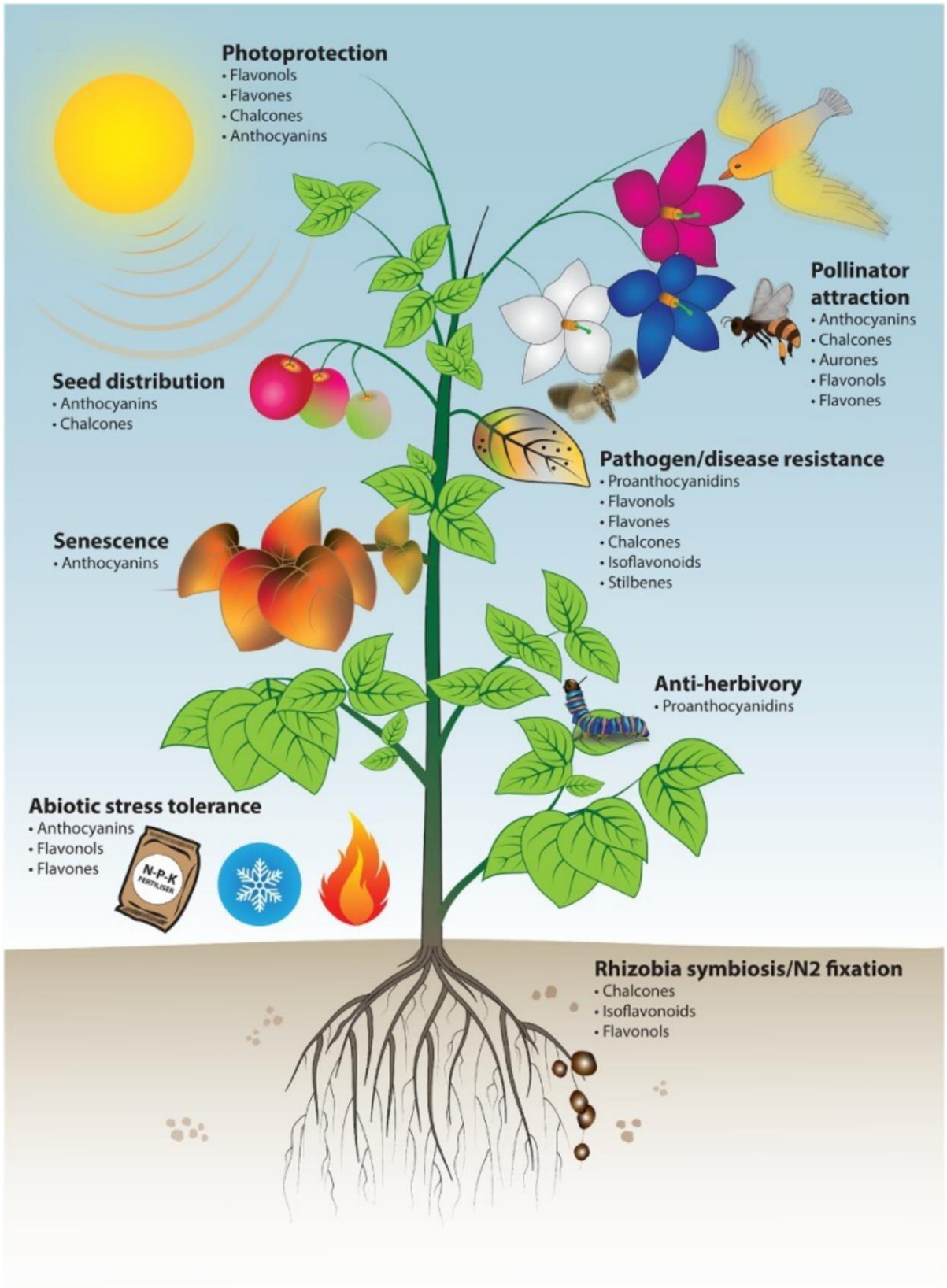
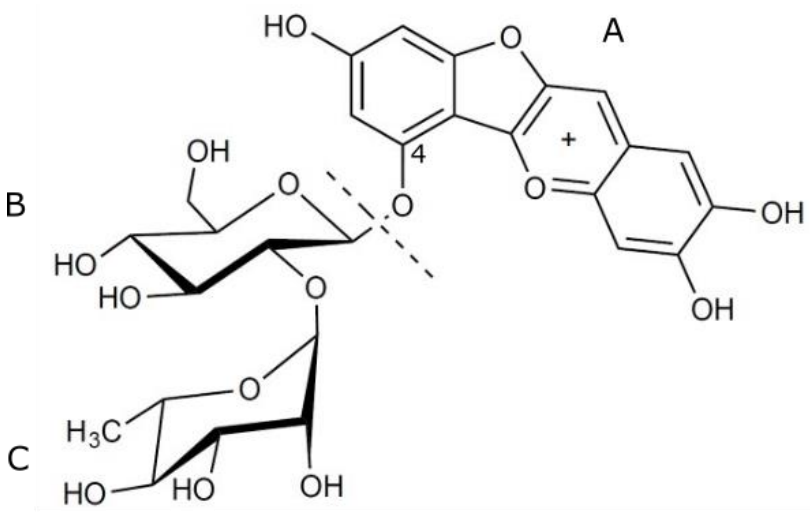


Figure 1.3. Functions of different flavonoid classes in plants. (From Albert et al. (2022).)

Other red pigments are present in non-vascular plants. Liverworts (including Marchantia) produce auronidins (originally identified as a pigment called riccionidin A) and mosses sphagnorubin. These compounds were initially considered to be unusual anthocyanidins because they are cell wall-bound instead of being in the vacuole and have a different chemical structure to 'normal' anthocyanins (Kunz et al., 1993; Rudolph & Jöhnk, 1982). However, Berland et al. (2019) later discovered that riccionidin A was not an anthocyanidin, instead being a distinct class of flavonoids that was previously uncharacterisied. Thus, the pigment was given the name "auronidins", due to its structural similarity to both aurones and anthocyanidins. The two auronidin compounds identified in Marchantia are the auronidin aglycone (no glycosylation) and auronidin 4-neohesperidoside (chemical structure shown in Figure 1.4). It is unclear whether sphagnorubins are part of a distinct class of flavonoids as the biosynthetic route to their production has not been resolved. However, they may be derived from an anthocyanin precursor, 3-deoxyanthocyanins (Davies et al., 2022; Piatkowski et al., 2020). The presence of auronidins and sphagnorubins in bryophytes and anthocyanins in gymnosperms and angiosperms suggests that red pigmentation has evolved several times among land plants and begs the question of what the driving force for this was. Auronidins and sphagnorubins, like anthocyanins, function in photoprotection by screening excess white light. Thus, it is possible that red pigmentation was conserved for the constant need of photoprotection, but the specific pigment used was adapted for additional functions such as for attracting pollinators, senescence and biotic stress tolerance (Davies et al., 2022). However, more work is necessary to gain a better understanding of the evolution of these compounds.



Auronidin 4-neohesperidoside

Figure 1.4. Chemical structure of the auronidin 4-neohesperidoside compound present in *Marchantia polymorpha*. A is the auronidin aglycone backbone, which is attached to a rhamnose sugar (B) at position 4, which is attached to a glucose (C).

Flavones and flavonols are two closely related groups of flavonoids in structure, biosynthesis and (apparently) function. Although they are less visible flavonoids, because they are typically colourless, they provide essential UV-B photoprotective functions by absorbing more UV-B light and detoxifying more ROS than other flavonoids (Agati & Tattini, 2010; Clayton et al., 2018). The presence of either flavones or flavonols is nearly ubiquitous among land plants, being reported in bryophytes, ferns, gymnosperms and angiosperms (Davies et al., 2020; Jarial et al., 2018; Wang et al., 2021; Winkel-Shirley, 2002; Yonekura-Sakakibara et al., 2019). However, liverworts and hornworts lack flavonols and flavones are of sporadic occurrence, suggesting that either their production has evolved several times or that production was lost in some plant lineages (Wen et al., 2020).

1.3.2 Conservation of the flavonoid biosynthetic pathway

Flavonoid biosynthesis is part of the larger phenylpropanoid pathway. In land plants, the phenylpropanoid pathway begins at phenylalanine, and the first committed step of flavonoid biosynthesis is catalysed by CHALCONE SYNTHASE (CHS) that converts *p*-coumaroyl CoA into naringenin chalcone (Figure 1.5) (Davies et al., 2020). While there can be variations in the enzymes used in different plants to make slightly different compounds, the flavonoid biosynthetic pathway shown in Figure 1.5 (summarised to include flavonoids produced by all land plants) has two main pathways branching from chalcones. The predominant pathway includes the production of flavones, flavonols, then anthocyanins; the other pathway is comparatively less extensive and poorly characterised and includes the production of aurones then auronidins (Davies et al., 2020). While auronidin is unique to liverworts, aurones are produced in bryophytes and angiosperms (Berland et al., 2019; Boucherle et al., 2017).

The major enzymes involved in the phenylpropanoid pathway, including PHENYLALANINE AMMONIA-LYASE (PAL), CINNAMATE 4-HYDROXYLASE (C4H) and 4-COUMARATE:COA LIGASE (4CL), are thought to be conserved among land plants. CHS and CHALCONE ISOMERASE-LIKE (CHIL), which are proposed to work together to form naringenin chalcone, are also conserved (Tohge et al., 2013; Wohl & Petersen, 2020; Yonekura-Sakakibara et al., 2019). This is supported by functional studies in Marchantia that show *chil* mutants having a reduced flavone and auronidin content (Berland et al., 2019). CHALCONE ISOMERASE (CHI) catalyses the first step of the pathway leading to flavone biosynthesis, converting naringenin chalcone to flavanones, and is also conserved among land plants (Tohge et al., 2013; Yonekura-Sakakibara et al., 2019). Berland et al. (2019) and Clayton et al. (2018) additionally showed that Marchantia *chi* mutants have no detectable flavones but have normal auronidin production, so CHI is not involved in the biosynthesis of auronidin.

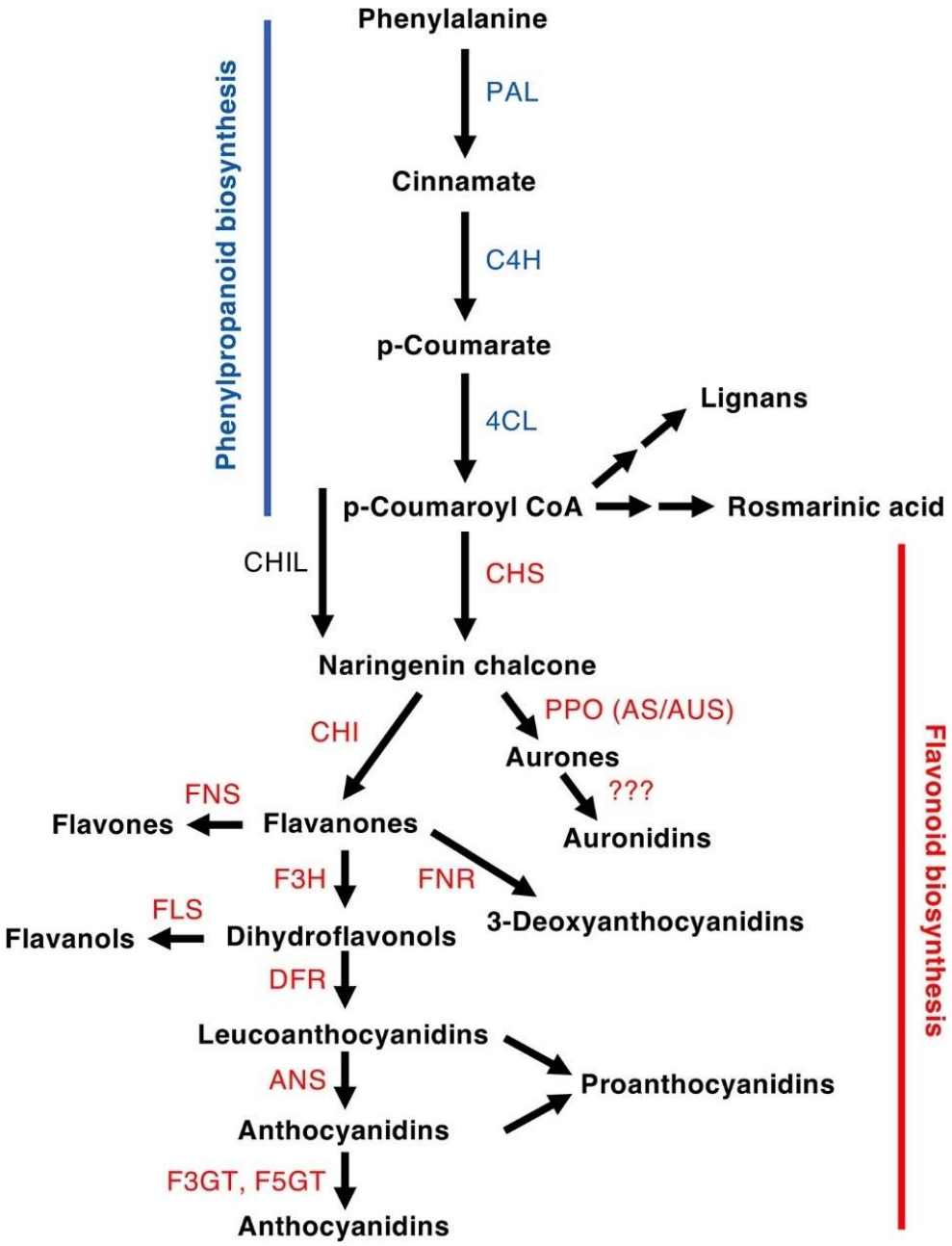
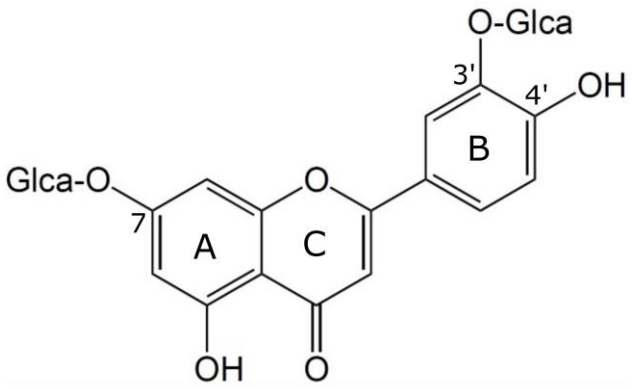


Figure 1.5. Phenylpropanoid biosynthetic pathway leading to the production of the major compounds produced in land plants. Phenylpropanoid enzymes are shown in blue text and flavonoid biosynthetic enzymes in red text. Abbreviations mentioned are: PAL, PHENYLALANINE AMMONIA-LYASE; C4H, CINNAMATE 4-HYDROXYLASE; 4CL, 4-COUMARATE:COA LIGASE; CHS, CHALCONE SYNTHASE; CHIL, CHALCONE ISOMERASE-LIKE; CHI, CHALCONE ISOMERASE; FNS, FLAVONE SYNTHASE (I = 2-OGD, II = CYP450); F3H, FLAVANONE 3-HYDROXYLASE; FLS, FLAVONOL SYNTHASE; ANS, ANTHOCYANIDIN SYNTHASE; PPO, POLYPHENOL OXIDASE; AS, AURONE SYNTHASE; AUS, AUREUSIDIN SYNTHASE. (From Davies et al. (2020).)

Two superfamilies of hydroxylase enzymes exist for flavonoid biosynthesis. The cytochrome P450 (CYP450) monooxygenases include the CYP93 enzymes FLAVONE SYNTHASE II (FNSII) and FLAVANONE 2-HYDROXYLASE (F2H), the CYP73A CINNAMATE 4-HYDROXYLASE (C4H) and CYP75 FLAVONOID 3'-HYDROXYLASE (F3'H); 2-oxoglutarate-dependent dioxygenases (2-OGD) include the DOXC28 clade enzymes FLAVONE SYNTHASE I (FNSI) and FLAVANONE 3-HYDROXYLASE (F3H) and the DOXC47 enzymes FLAVONOL SYNTHASE (FLS) and ANTHOCYANIDIN SYNTHASE (ANS) (Yonekura-Sakakibara et al., 2019). Since Marchantia does not produce flavonols or anthocyanins, it is expected that F3H and ANS enzymes are absent. The characterised luteolin flavones of Marchantia have 3' hydroxylation of the B-ring (shown in Figure 1.6) (Markham & Porter, 1974), so F3'H or an alternative hydroxylation enzyme must be present but has not been identified.

The evolution of flavone biosynthesis is a controversial topic, mainly owing to the difficulty of determining gene function purely based on sequence similarity. Generally, flavone *O*-glycosides are synthesised by FNSI or FNSII enzymes, which convert flavanones into the flavone backbone that then gets glycosylated. Whereas, flavone *C*-glycosides are synthesised by F2H converting flavanones into 2-hydroxyflavanones, which are glycosylated and dehydrated into flavone glycosides (Jiang et al., 2016). Most of the flavones detected in Marchantia to date are flavone *O*-glycosides (example chemical structure shown in Figure 1.6) (Markham et al., 1998). The evolution of FNSI, FNSII and F2H in land plants is complicated. Phylogenetic analysis suggests that FNSI enzymes may be present in bryophytes and some angiosperms, whereas FNSII is only present in angiosperms (Du et al., 2016; Gebhardt et al., 2005; Li et al., 2020a; Pucker & Iorizzo, 2023; Wang et al., 2020). Additionally, F2H may have evolved from an ancestral CYP93 gene following a loss of dehydratase activity that subsequently prevented the conversion of 2-hydroxyflavanones into flavones (Jiang et al., 2016). These ideas raise important questions around how enzymes required for flavone biosynthesis evolved throughout land plant evolution and whether there was a common ancestral gene in early land plants that the enzymes evolved from. Because current phylogenetic proposals are based solely on sequence similarity, more accurate determinations need to be made by assessing the function of genes of interest *in planta*. A Marchantia gene has been identified that is a good Mp*FNSI* candidate (unpublished data of Drs Nick Albert and Kevin Davies, Plant and Food Research, New Zealand). However, functional characterisation of the gene is necessary to verify this.



Luteolin-7,3'-di-*O*-glucuronide

Figure 1.6. Chemical structure of the flavone *O*-glycoside luteolin-7,3'-di-*O*-glucuronide present in *Marchantia polymorpha*. The flavone backbone has A, B and C rings (indicated). The hydroxyl group at positions 7 (ring A) and 3' (ring B) are glycosylated with glucuronic acid (Glca).

Branching off the main flavonoid biosynthetic pathway is the production of aurones from naringenin chalcone, which are thought to be the substrates for auronidin biosynthesis (Figure 1.5). While aurone biosynthesis is well understood in both bryophytes and angiosperms (Boucherle et al., 2017), it is currently unknown what enzymes are required for the subsequent formation of auronidin. The characterised genes for aurone biosynthesis encode POLYPHENOL OXIDASE (PPO) enzymes that have AURONE SYNTHASE (AS) or AUREUSIDIN SYNTHASE (AUS) activity (Berland et al., 2019; Furudate et al., 2023; Nakayama et al., 2000). However, based on *in vitro* assays, peroxidases have also been proposed to catalyse the formation of aurones in angiosperms (Farag et al., 2009; Nakayama, 2022). Aurone biosynthesis is not the most common function of either of these enzyme classes; in many plant species, PPOs function as pathogen and herbivore defence enzymes by producing reactive metabolites that result in the brown tissue colouration commonly observed in response to cell damage (Araji et al., 2014), and peroxidases function in oxidative stress-induced plant defense by promoting the metabolism of ROS such as hydrogen peroxide (H_2O_2) (Almagro et al., 2008). However, there are now multiple well characterised examples of PPOs being required for aurone biosynthesis. These PPOs catalyse the oxidation of chalcones by hydroxylase and/or dehydration activity (depending on whether the substrate has one or two hydroxyl groups) that is required to eventually form 3',4'-dihydroxyaurones (Nakayama et al., 2001). Whereas the peroxidases are proposed to produce 4'-monohydroxyaurones using the reduction of H_2O_2 (Farag et al., 2009; Nakayama, 2022). The PPO and peroxidase enzymes that function in aurone biosynthesis are found sporadically within land plants which, together with the fact that aurones are also produced sporadically, indicates that the biosynthesis of aurones has evolved independently several times (Davies et al., 2022; Nakayama, 2022).

The size of *PPO* gene families varies greatly in plants. In angiosperms it varies from zero (e.g., for *Arabidopsis thaliana*, hereafter *Arabidopsis*) to 11 in *Glycine max* (soybean), while the moss *P. patens* has 13 *PPO* genes (Tran et al., 2012). Notably, 64 *PPO* genes were identified in Marchantia, at least 46 of which are actively transcribed (Davies et al., 2020). However, as yet, only one of these has a characterised function – Mp*AS1* that encodes an aurone synthase required for auronidin biosynthesis (Berland et al., 2019; Davies et al., 2020; Furudate et al., 2023). This poses an interesting question of why the *PPO* gene family is so extensive in Marchantia. Given the role of PPOs in biotic stress-induced plant defence, it is possible that they were accumulated to help these non-vascular land plants tolerate the biotic stresses present in a terrestrial environment and that other characteristics (not present in Marchantia) were adopted later in land plant evolution for this purpose (e.g., a waxy cuticle and trichomes).

1.3.3 Abiotic stress-induction and regulation of flavonoids

A major function of flavonoids is to provide plants with protection against abiotic stress, which was prevalent during land colonisation when flavonoids are thought to have evolved. Therefore, flavonoid production is induced in response to several types of abiotic stress, including cold, excess white or UV-B light, nutrient deprivation (particularly nitrogen and phosphorus), drought and salinity (Agati & Tattini, 2010; Davies et al., 2018; Tattini et al., 2004). This stress-responsive function of flavonoids requires their production to be under tight spatial and temporal regulation.

In angiosperms, the key transcription factors regulating the developmental, spatial and temporal patterns of flavonoid production are MYBs, in particular R2R3-MYBs. The MYBs are one of the largest families of transcription factors in plants, being involved in a range of cellular processes including cell division, development, signalling pathways, secondary metabolism and plant defence (Katiyar et al., 2012; Mehrtens et al., 2005). MYB transcription factors are characterised by the presence of a MYB DNA-binding domain. R2R3-MYBs are those that share the R2 and R3 repeats in the domain, and are divided into subgroups based on the conservation of amino acid motifs in the protein sequence (Stracke et al., 2001). Subgroups regulating phenylpropanoid biosynthesis include at least SG4, 5, 6 and 7, and these tend to be associated with different functions.

R2R3-MYBs that activate anthocyanin and proanthocyanin biosynthesis work with transcription factor partners. Studies using angiosperms have proposed that the MYBs form a complex with basic helix-loop-helix (bHLH) transcription factors and WD40-repeat (WDR) proteins, now referred to as MYB-bHLH-WDR (MBW) complexes (Lloyd et al., 2017). The formation of these complexes was confirmed in a study of *Arabidopsis* by Baudry et al. (2004), which identified the physical interaction

between TTG1 (WDR), TT2 (SG5 MYB) and TT8 (bHLH). Prior to this study, these proteins were known to be necessary for the expression of the *BAN* gene, which encodes an enzyme required for proanthocyanin biosynthesis in *Arabidopsis*, but not that they interacted. However, there has been some debate about how MBW complexes regulate flavonoid biosynthesis; Borevitz et al. (2000) showed that MYB/bHLH/TTG1 regulates all stages of the phenylpropanoid pathway, whereas, Gonzalez et al. (2008) demonstrated that the complex only regulates the later genes of the flavonoid biosynthetic pathway and not the earlier genes. This controversy emphasises the fact that different proteins of the MBW complex can regulate different branches of the flavonoid biosynthetic pathway, that this regulation may be different in among plants, and that redundancy of regulation is present for some steps of the pathway (Gonzalez et al., 2008).

Angiosperms have complex systems of R2R3-MYB-mediated regulation of flavonoid production in response to abiotic stress. Such systems have been characterised in many species, but particularly in *Arabidopsis*. The production of anthocyanins in nitrogen and phosphate limiting conditions is directly activated by MYB75/PAP1 and MYB90/PAP2 (SG6 R2R3-MYBs) (Borevitz et al., 2000; Morcuende et al., 2007; Scheible et al., 2004), and predominantly by MYB75/PAP1 under excess white light and UV-B light (Cominelli et al., 2008; Vanderauwera et al., 2005). Flavonols and flavones are similarly induced by excess white light and UV-B light, and this response is regulated by MYB12 (a SG7 R2R3-MYB) (Mehrtens et al., 2005; Stracke et al., 2010; Stracke et al., 2007). While SG5 and 6 R2R3-MYBs usually function in an MBW complex, SG7 MYBs tend to work cooperatively with bZIP proteins. For example, MYB12 works alongside the bZIP protein ELONGATED HYPOCOTYL 5 (HY5) (Stracke et al., 2010). Flavonols and flavones are induced by nitrogen and phosphorus. This response could be regulated by MYB12, but further characterisation in other plant species is required to confirm this (Song et al., 2020; Stewart et al., 2001).

Abiotic stress-induction of flavonoid (or phenylpropanoid) production by R2R3-MYBs is somewhat conserved in Marchantia. However, the R2R3-MYB gene family is considerably less extensive than in angiosperms; Bowman et al. (2017) found only 21 R2R3-MYB genes in the Marchantia genome sequence, whereas angiosperm species commonly have 100-200 genes (Liu et al., 2015; Stracke et al., 2001). While none of these Marchantia genes are orthologous to the angiosperm phenylpropanoid-related genes of SG4, 5, 6 and 7, two genes (Mp*Myb14* and Mp*Myb02*) are basal to a large clade that contains all of these genes from *Arabidopsis* (Bowman et al., 2017). Functional analysis demonstrated that MpMYB14 and MpMYB02 are required for phenylpropanoid regulation in Marchantia; Mp*Myb14* expression is induced by nitrogen and phosphorus deficiency, excess white and UV-B light and high salinity to positively regulate auronidin biosynthesis and MpMYB02 activates

the production of marchantins, bibenzyls common in Marchantia that are thought to be plant defence compounds (Albert et al., 2018; Kubo et al., 2018; Rico-Reséndiz et al., 2020). Thus, the use of R2R3-MYBs for stress-induced phenylpropanoid regulation may be a conserved characteristic among land plants that arose early during land colonisation to facilitate plant survival in a terrestrial environment. Furthermore, the conservation of a phenylpropanoid-related clade of R2R3-MYBs in Marchantia and *Arabidopsis* suggests that the last common land plant ancestor may have contained an R2R3-MYB ancestral gene that was retained throughout liverwort evolution, while the gene may have duplicated and undergone neofunctionalization during the diversification of plant species and secondary metabolite production in angiosperms.

An interesting point of comparison between Marchantia and angiosperms is that the equivalent red pigments auronidins and anthocyanins are both induced by nutrient deprivation, and this is regulated by R2R3-MYBs. Anthocyanins are also regulated by nitrogen-responsive proteins called BRIC-A-BRAC, TRAMTRACK AND BROAD (BTB) – TRANSCRIPTIONAL ADAPTER ZINC FINGER (TAZ) DOMAIN 2 (BT2). In apple (*Malus domestica*), MdBT2 negatively regulates anthocyanin biosynthesis in the presence of nitrogen by promoting the ubiquitination and degradation of MdMYB1 (a SG6 R2R3-MYB, synonymously known as MdMYB10), which is required for anthocyanin biosynthesis (Wang et al., 2018). A gene has been identified in Marchantia that is a strong candidate for Mp*BT2* (unpublished data of Dr Yanfei Zhou, Plant and Food Research, New Zealand). Thus, functional analysis should be performed on the gene to determine if it acts on MpMYB14 to regulate auronidin biosynthesis, and therefore whether the regulatory action of BT2 proteins is conserved among land plants.

Regulation of flavone production by R2R3-MYBs seems to be lacking in Marchantia. In angiosperms, SG7 R2R3-MYBs regulate the production of flavones in response to UV-B light. Marchantia produces flavones in response to UV-B light (Clayton et al., 2018), but no SG7 R2R3-MYBs exist. Mp*Myb14* is thought to be the closest related sequence to the SG7 R2R3-MYBS. However, a normal amount of flavones are produced in *myb14* mutants, and no candidate genes for FNSI or F2H (that synthesise flavones) were upregulated in plants over-expressing Mp*Myb14* (Albert et al., 2018). This suggests that MpMYB14 does not activate flavone biosynthesis. Clayton et al. (2018) demonstrated that UV-B light-induced flavone production in Marchantia is dependent on activity of the *UV Resistance Locus 8* (*UVR8*) and *HY5* genes. This is the same as in *Arabidopsis*, whereby, through its interaction with CONSTITUTIVE PHOTOMORPHIC 1 (COP1), UVR8 activates the UV-B stress response by reducing the protein degradation of the HY5 transcription factor, which activates MYB12 and the genes involved in phenylpropanoid biosynthesis (Binkert et al., 2014). The role of REPRESSOR OF UV-B

PHOTOMORPHOGENESIS 1 (RUP1) was also shown to be conserved with *Arabidopsis*, where it modulates the activity of UVR8 by displacing COP1 from HY5 and redimerising UVR8 (Gruber et al., 2010). Given the dependence on HY5 for flavone biosynthesis in Marchantia, and the absence of candidate R2R3-MYB factors, it would be interesting to determine if HY5 is able to directly activate the flavone biosynthetic genes.

A current question of considerable interest is whether bryophytes are able to form MBW complexes for regulating flavonoids. The gymnosperm *Picea abies* (Norway spruce) contains an MBW complex and the liverwort *Plagiochasma appendiculatum* has a *bHLH* sequence that phylogenetically clusters with angiosperm *bHLH* genes that form part of an MBW complex (Nemesio-Gorriz et al., 2017; Wu et al., 2018). Thus, there is evidence of MBW complexes originating prior to the divergence of vascular plants. Marchantia contains a Mp*bHLH12* gene that has sequence similarity to the *P*. *appendiculatum bHLH* and could be involved in flavonoid regulation based on the transcriptomic profile of plants over-expressing the gene (Arai et al., 2019). Additionally, the Marchantia genome does contain conserved WDR proteins (Bowman et al., 2017). However, Albert et al. (2018) discovered that MpMYB14 and MpMYB02 only have partial conservation of the amino acid motif required for their interaction with a bHLH. Given the conflicting evidence, further investigation into the possible presence of MBW complexes in Marchantia is required.

Thus, there is a lot more to understand about the stress-induced biosynthesis and reglation of flavonoids in Marchantia. Genetic modification techniques have been shown to be useful for studying this, including generating transenic plants that over-express flavonoid-related genes of interest and creating gene-specific loss-of-funtion mutants.

1.4 Using genetic modification to characterise gene function

Genetic modification is the process whereby foreign DNA is introduced into the genome of an organism. An incredibly useful application of genetic modification is being able to functionally characterise genes of interest by generating plants that over-express them. Part of genetic modification that has become of particular interest in recent years is gene editing, where permanent changes are made to the host organism's DNA. CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9) is a widely used gene editing technique that allows you to make mutations in a specific location of targeted genes and is therefore an invaluable tool for determining the function of particular genes (Hsu et al., 2014).

Generating transgenic plants requires that DNA can be transformed into the plant cells. This is significantly harder than transforming cells of many other eukaryotes because plants have rigid cell

walls, making it hard to directly deliver DNA into the cell (Demirer et al., 2018). Several plant transformation techniques have been developed to overcome the challenges posed by the cell wall. The first is by transforming protoplasts, where the degradation of the cell wall enzymatically makes it easier for the cells to uptake DNA. Transgenic plants can then be regenerated from the protoplast culture (Takebe et al., 1968), although both production of suitable protoplasts and regeneration of them into plants can be challenging. Particle bombardment (or biolistics) is another transformation method that involves binding DNA to gold (or tungsten) particles and using a gene gun to physically shoot them into plant cells. The DNA can then be transiently expressed in cells that the particles penetrate and may, in a few cells, also get incorporated into the genome (Klein et al., 1987).

Probably the most widely used technique is *Agrobacterium tumefaciens*-mediated transformation. *A*. *tumefaciens* is a plant pathogen that naturally infects plants through wound sites, forming crown gall tumours that are used to introduce and integrate T-DNA from its tumour-inducing (Ti) plasmid into the host genome (Ashby et al., 1988; Larebeke et al., 1974). The infection process works by a process called chemotaxis, where plant wound sites secrete phenolic compounds (e.g., acetosyringone) that induce the expression of *virulence* genes on the Ti plasmid that are required for T-DNA transfer. This works whereby the VirA protein receptor recognises and binds to the phenolic compound, which causes autophosphorylation of VirG. Activation of VirG induces the expression of the other *vir* genes (Jin et al., 1990; Palmer & Shaw, 1992). The VirD2 endonuclease creates a single-stranded nick at the T-DNA left and right borders, creating a T-strand that gets transferred to the plant cell. VirD2 is also thought to direct the transport of the T-strand by its nuclear localisation signal (Stachel et al., 1986; Tinland et al., 1992). VirD4 and VirB proteins form the type IV secretion system that allows the VirD2/T-strand complex to be transported by forming protein channels in the cell membrane. VirB proteins also generate an extracellular T-pilus that attaches itself to the host plant cell, facilitating T-strand transfer (Lai et al., 2000). Once the T-strand is in the plant cell, it is coated by VirE2 to protect it from degradation before it gets incorporated into the plant genome (Lapham et al., 2021).

The natural infection response of *A*. *tumefaciens* can be artificially simulated for genetic modification. However, introducing foreign genes into the T-DNA of Ti plasmids is difficult both because the Ti plasmid is relatively large and because the T-DNA region does not have unique restriction endonuclease sites, required so that only the region the genes are cloned into is digested (Gelvin, 2003). To circumvent these challenges, Hoekema et al. (1983) developed a “binary-vector system”. The binary vector is a plasmid containing the foreign DNA that is flanked by T-DNA border sequences to ensure that the natural T-DNA transfer mechanism is used. Once cloned, it gets

transformed into a disarmed *A*. *tumefaciens* strain (e.g., GV3101) that contains the Ti plasmid necessary for T-DNA transfer but that has the T-DNA region removed from it to prevent tumour formation. Since both plasmids would then be present in the in the same *A*. *tumefacien* cells, plant cells can be transformed (Gelvin, 2003). Plants regenerated using this method are expected to be transgenic.

Generating CRISPR/Cas9 mutant plants often requires the use of a plant transformation system, and usually *A*. *tumefaciens* is used. The standard procedure for *A*. *tumefaciens*-mediated CRISPR/Cas9 transformation in plants is illustrated in Figure 1.7, using Marchantia as an example. The binary vector transformed into *A*. *tumefaciens* contains a gene encoding Cas9 and single guide RNAs (sgRNAs) that specify where on a target gene the double-stranded break occurs (Zhang et al., 2019). A sgRNA has two components: a spacer (CRISPR (cr) RNA) and a scaffold sequence (trans-activating CRISPR (tracr) RNA). The crRNA is typically a 20 nucleotide-long sequence that can be deisgned to bind the target DNA anywhere that is adjascent to a protospacer motif (PAM), an NGG sequence. The binding of the crRNA guides the Cas9 enzyme to the target DNA, where it recognises the PAM and makes a double-stranded cut 3 base pairs (bp) upstream of it (Jinek et al., 2012). The tracrRNA binds to pre-crRNA (the untranscribed precursor of crRNA) and promotes its processing into the crRNA active form that allows Cas9-editing to occur (Deltcheva et al., 2011). The cell tries to repair the double-stranded breaks by non-homologous end-joing (NHEJ), the process of ligating the ends of the break together without using a DNA template, but insertion or deletion mutations often occur at the cleavage site because the lack of a template makes NHEJ error-prone. Mutations created usually result in loss of gene function. Thus, gene-specific knock-out mutants are generated (Guo et al., 2018).

To functionally characterise Marchantia genes of interest, an *A*. *tumefaciens*-mediated transformation protocol needed to be developed. Protocols exist for the transformation of Marchantia thallus, gemmae and spores (Kubota et al., 2013; Tsuboyama & Kodama, 2013; Tsuboyama & Kodama, 2018a, 2018b; Tsuboyama et al., 2018). The method used is dependent on the purpose of the transformation, as some protocols are more complicated than others, have different transformation efficiencies and can regenerate plants with different genetic backgrounds. However, there is not currently a spore transformation protocol that is easy, practical and highly efficient, and attempts to reproduce existing protocols (by Dr Nick Albert) have not achieved a workable regeneration. Thus, considerable work needed to be invested into improving the efficiency of spore transformation protocols. Additionally, using spore transformation for generating CRISPR/Cas9 mutants requires a high Cas9 editing efficiency so that many mutants are created from

few spores and so that fewer chimeras (plants with both wild-type and mutant sectors) are produced, which are common in Marchatia and are impractical to work with.

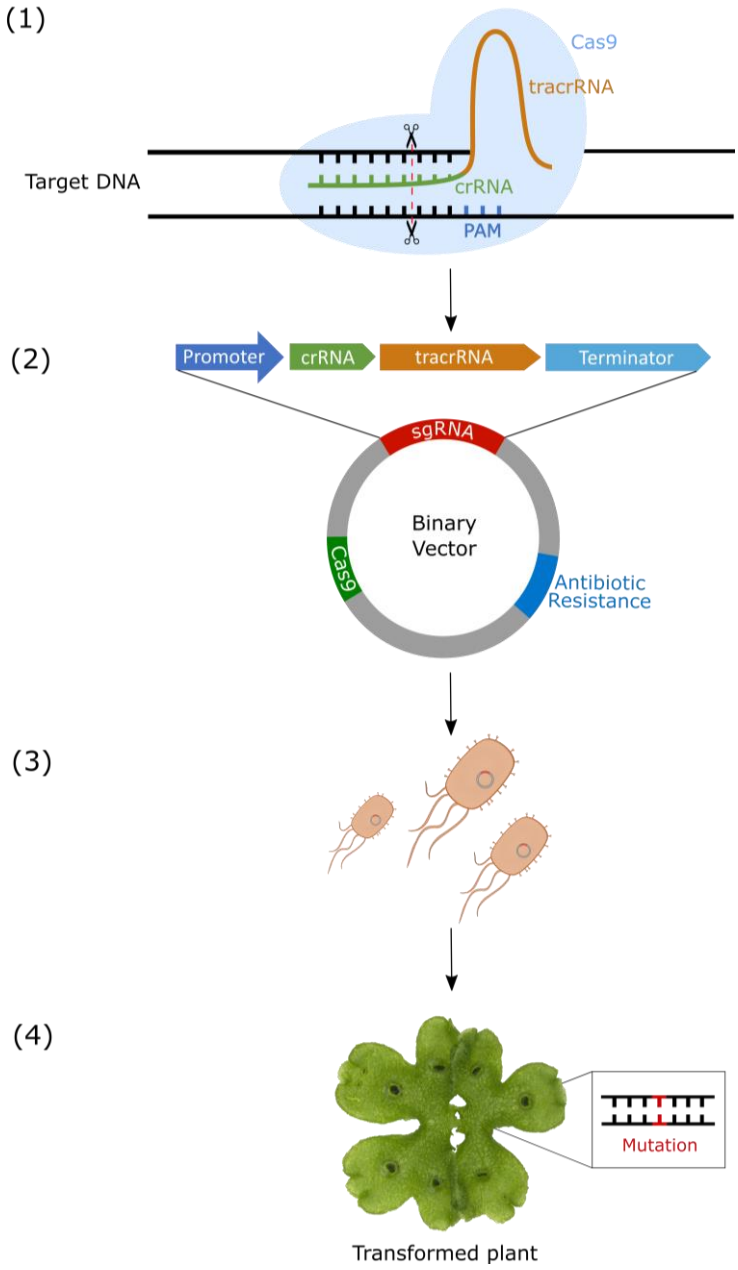


Figure 1.7. Standard procedure for *Agrobacterium tumefaciens*-mediated CRISPR/Cas9 plant transformation, using *Marchantia polymorpha* as an example. (1) Single guide RNAs (sgRNAs), comprising of CRISPR RNAs and trans-activating CRISPR RNAs, are specifically designed to bind to a target DNA sequence of interest adjascent to a protospacer motif (PAM). This guides the Cas9 enzyme to make a double-stranded break three base pairs upstream of the PAM site. (2) The sgRNAs, designed to be driven by a plant promoter and terminator sequence, are cloned into an *A*. *tumefaciens* binary vector that contains a selectable marker gene along with a gene encoding Cas9. (3) The plasmid is transformed into *A*. *tumefaciens*, which is used to transform the *M*. *polymorpha* plant tissue. Transgenic plants are selected for using the marker gene. (4) Many plants regenerated using this method will contain a mutation(s) in the targeted gene.

1.5 Project hypotheses and aims

In recent years, there has been considerable interest in using Marchantia (as a model species of the bryophyte lineage) to investigate the conservation of the flavonoid biosynthetic pathway among land plants. As a result, the main flavonoids produced by Marchantia have been identified, as well as many of the enzymes required for their production. However, there is still a lot about the pathway that is not understood. Thus, further investigation will help to elucidate the extent to which the flavonoid biosynthetic pathway of Marchantia is conserved with that known for angiosperms. Given Marchantia is thought to share many features in common with the ancestors of the land plant lineage, and is a representative of the bryophyte sister lineage to all other land plants, determining this could also provide insight into what flavonoid characteristics the ancestors may have acquired during land colonisation to provide tolerance against the abiotic stresses present in a terrestrial environment. The particular focus of this thesis was on investigating the genetic requirements of stress-induced flavone biosynthesis and auronidin regulation in Marchantia.

This research question was addressed by proposing the following hypotheses:

- MpFNSI is required for flavone biosynthesis in Marchantia.

- MpBT2 negatively regulates auronidin biosynthesis in Marchantia by promoting the degradation of MpMYB14.

To investigate these hypotheses, the overall aim of the thesis was to use CRISPR/Cas9 and Marchantia spore transformation to characterise the function of the Mp*FNSI* and Mp*BT2* candidate genes in stress-induced flavone biosynthesis and auronidin regulation, respectively. This included a plan to increase the efficiencies of the key transformation technologies being used. Thus, the project had the following specific aims:

- To optimise the Marchantia spore transformation protocol by improving the transformation efficiency and Cas9 editing efficiency (Chapter 3).

- To determine the function of the Mp*FNSI* candidate gene in Marchantia by using CRISPR/Cas9 to generate *fnsI* mutants and characterising the ability of the mutant lines to produce flavones under excess white light conditions (Chapter 4).

- To determine the function of the Mp*BT2* candidate gene in Marchantia by generating plants that over-express *BT2* and using CRISPR/Cas9 to create *bt2* mutants *and* characterising the ability of these lines to produce auronidin under nutrient deprivation conditions (Chapter 5).

Chapter 2

General materials and methods

Transgenic Marchantia plants containing CRISPR/Cas9 gene-specific mutations or over-expressing genes of interest were generated by *A. tumefaciens*-mediated spore transformation. This chapter outlines the materials and methods required to generate these plants, including the designing of CRISPR/Cas9 and over-expression constructs (section 2.1), cloning of the constructs into binary vectors (section 2.2), Marchantia plant material and growth conditions used (section 2.3), the *A. tumefaciens*-mediated spore transformation protocol (section 2.4), how Marchantia plants were imaged (section 2.5) and the statistical analysis used (section 2.6). Specific methods, as well as methods for experiments unrelated to spore transformation, are described in the results chapters where they are discussed.

2.1 CRISPR/Cas9 and over-expression construct design

Marchantia plants over-expressing a gene of interest were created by designing DNA constructs that flank the gene coding sequence (CDS; extracted from the Marchantia genome sequence) with attL sites. To generate CRISPR/Cas9 mutants, multiplexing was used to design DNA constructs containing multiple sgRNA molecules (crRNA and tracrRNA fusions) targeting specific gene regions of interest (extracted from the Marchantia genome sequence) (Figure 2.1). Each sgRNA was flanked by an *Arabidopsis* glycine transfer RNA (tRNA) molecule, which improves the processing of individual sgRNAs because tRNAs are spliced during RNA processing (Xie et al., 2015). A restriction enzyme recognition site (in this case, *Bsa*I) was designed on each end of the sgRNA cassette for cloning purposes.

The “CRISPR sites” function of Geneious version 2022.0 (Biomatters, Limited) was used to design the crRNAs adjacent to a PAM site, typically within the first few exons of the gene of interest. The tracrRNA (sequence from Dang et al. (2015)) was then copied onto the 3' end of each crRNA. sgRNA fusions created were quality checked to make sure no off-target binding could occur in the CDS and the RNA structure would fold correctly at 37°C according to the “Turner” free-energy model (Mathews et al., 2004); forming a hairpin, ensuring the crRNA was not bound to itself, and the rest of the sgRNA had a strong binding probability (see Appendix 1 for examples of high quality and bad quality sgRNAs). Once four high quality sgRNAs were selected, the CRISPR/Cas9 construct could be created.

Figure 2.1. Design of the multiplex single guide RNA (sgRNA) cassette used to generate CRISPR/Cas9 mutants. The cassette consists of sgRNAs that were flanked by *Arabidopsis thaliana* transfer RNA (tRNA) molecules and *Bsa*I restriction enzyme recognition sites that were designed on each end. The expression of the cassette was driven by a *U6* promoter and terminated by *NOS* (*Nopaline Synthase*).

CRISPR/Cas9 and over-expression constructs, once designed, were synthesised and cloned into a pUC57 plasmid as a commercial service by GenScript (USA). Once the synthesised DNA constructs were received, they were prepared for plasmid cloning (section 2.2) by resuspending the dried DNA in 50 µL sterile water.

2.2 Construct preparation by DNA cloning

DNA cloning was used to remove the constructs (designed in section 2.1) from pUC57 so that they could be, through a series of steps, recombined into a binary vector containing the necessary machinery to be transformed into *A. tumefaciens* for Marchantia spore transformation (section 2.4).

The first step of DNA cloning was to heat-shock transform the synthesised pUC57 plasmid, containing the DNA construct, into *Escherichia coli* (section 2.2.3) to both stabilise and replicate the plasmid. The cells were spread onto lysogeny broth (LB)-agar plates with 100 mg L^{-1} ampicillin selection (see Appendix 2.3 for preparation of all antibiotics), because pUC57 has an ampicillin resistance gene, and grown overnight (section 2.2.1). The plasmid was extracted from the liquid cultures made from these cells (section 2.2.4) and used for further cloning.

The remainder of DNA cloning varied for the CRISPR/Cas9 and over-expression constructs. CRISPR/Cas9 constructs were firstly excised from the extracted pUC57 plasmid using *Bsa*I in a restriction endonuclease enzyme digest (section 2.2.6). Following separation of the digested fragments by gel electrophoresis (section 2.2.7), the smaller DNA band corresponding to the construct was excised and purified from the gel (section 2.2.8). Because *Bsa*I makes non-complementary sticky ends by cutting outside of its restriction site, the DNA construct could be directionally ligated into a p$U6_{pro}$/pENTR entry vector (section 2.2.9; Figure 2.2, step 1) (see Appendix 3.1 for plasmid map). The p$U6_{pro}$/pENTR plasmid, created by Dr Nick Albert, contains a Marchantia-specific *U6* promoter (a type III RNA polymerase III promoter) that drives the expression of the introduced sgRNA cassette (Figure 2.1). The ligation product was heat-shock transformed into

E. coli (section 2.2.3) and cells spread onto LB-agar plates containing 50 mg L^{-1} kanamycin selection (selection 2.2.1), corresponding to the kanamycin resistance gene on p*U6*$_{pro}$/pENTR. Correct ligation was confirmed by extracting the plasmid from liquid cultures of the cells (section 2.2.4), digesting it with the *Pvu*II restriction enzyme (that has two recognition sites on the plasmid; see Appendix 3.1) (section 2.2.6) and separating the two fragments by gel electrophoresis to confirm they are of the expected size (section 2.2.7). The plasmid was also checked by DNA sequencing (section 2.2.11) using the NA428 forward primer (Appendices 3.1 and 4).

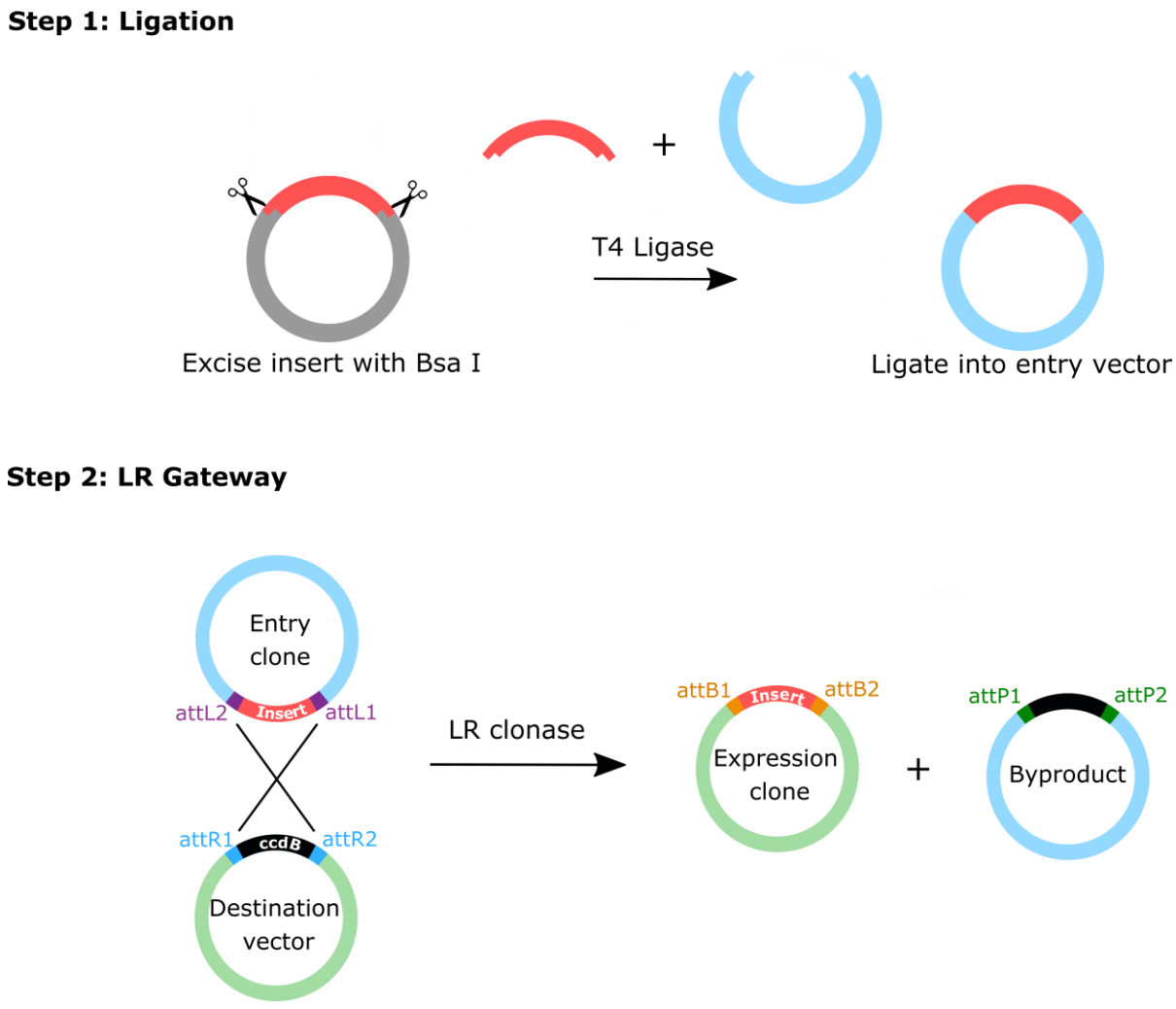


Figure 2.2. The two main DNA cloning steps for creating a CRISPR/Cas9 construct. Step 1, excision of DNA construct (red) synthesised into pUC57 (grey) and ligation into an entry vector (blue) catalysed by T4 ligase, creating an entry clone. Step 2, LR gateway reaction between the entry clone attL sites and destination vector (green) attR sites to recombine the DNA construct into the destination vector, creating an expression clone.

p*U6*$_{pro}$/pENTR also contains attL sites, making the ligation product a substrate for the next step – an LR gateway reaction (section 2.2.10). The presence of the Invitrogen™ Gateway™ LR Clonase™ II Enzyme Mix catalysed the recombination between the entry clone attL sites and the pMpGE010 destination vector (sourced from Sugano et al. (2018)) attR sites, creating an expression clone with the sgRNA cassette/*U6* promoter inserted into the destination vector (Figure 2.2; step 2) (e.g., pRMY07; see Appendix 3.2 for plasmid map). This ensured the expression of the sgRNA cassette was terminated by *NOS* (from the *A. tumefaciens Nopaline Synthase* gene) (Figure 2.1). The vector also contains machinery required for CRISPR/Cas9 and *A. tumefaciens*-mediated spore transformation, including the gene encoding AtcoCas9-NLS (*Arabidopsis*-codon-optimized Cas9, tagged with a nuclear localisation signal) driven by a Marchantia-specific *EF1*α (*Elongation Factor 1-alpha*) promoter and terminated by *pea 3A(T)*. Additionally, plant resistance to hygromycin was conferred by the *hptII* gene (*Streptomyces* origin), flanked by a *CaMV35S* (*35S*) promoter and terminator.

DNA cloning required for over-expression constructs was simpler, only involving an LR gateway reaction (step 2; Figure 2.2). This is because the synthesised pUC57 was designed to contain attL sites flanking the gene CDS. Thus, recombination between the pUC57 attL sites and the pNWA101 destination vector (created by Dr Nick Albert) attR sites created an expression clone that flanked the CDS with a *35S* promoter and *OCS* terminator (from the *A. tumefaciens Octopine Synthase* gene) (e.g., pRMY27; see Appendix 3.3 for plasmid map). The plasmid also contained a bacterial origin of replication gene and a *nptII* gene (*E. coli* origin), flanked by a *NOS* promoter and *OCS* terminator, that conferred plant resistance to geneticin (G-418) and kanamycin.

The binary vectors containing CRISPR/Cas9 and over-expression constructs were heat-shock transformed into *E. coli* (section 2.2.3) to stabilise and replicate the plasmid, and cells were spread onto LB-agar plates with 100 mg L^{-1} spectinomycin selection (corresponding to the spectinomycin resistance genes on pMpGE010 and pNWA101) (section 2.2.1). Plasmids extracted from the liquid cultures (section 2.2.4) were used as a template for DNA sequencing (section 2.2.11) to confirm correct recombination. CRISPR/Cas9 constructs were screened using the NA428 forward primer (Appendices 3.1 and 4) alone, and over-expression constructs using NAg176 (forward), NAg177 (reverse), NA354 (forward) and NA355 (reverse) primers (Appendix 4). Once sequenced, the plasmids were then electroporated into *A. tumefaciens* (section 2.2.12). Transformed cells were spread onto LB-agar plates containing 100 mg L^{-1} spectinomycin and 25 mg L^{-1} gentamicin selection and grown for 48 h (section 2.2.1). Gentamicin was required to maintain the *A. tumefaciens* Ti plasmid. Liquid cultures generated from the cells were stored as glycerol bead stocks (section 2.2.1). Cells from these glycerol stocks were grown on LB-agar plates during *A. tumefaciens*-mediated Marchantia spore transformation.

The remainder of section 2.2 outlines the materials and methods required for the DNA cloning process described above.

2.2.1 Bacterial growth conditions

The XL1-blue *E. coli* strain used was grown on LB-agar plates or as liquid cultures in LB-broth (Appendix 2.4), with the appropriate selection. The plates and cultures were incubated overnight (~17 h) at 37°C, in a standing incubator or Infors HT Multitron shaking incubator at 250 rpm (revolutions per minute), respectively.

The GV3101 *A. tumefaciens* strain used was grown on LB-agar plates or as liquid cultures in LB-broth, with the appropriate selection. Plates and cultures were incubated at 28°C in a standing incubator for 48 hours, or Infors HT Multitron shaking incubator at 250 rpm overnight (~17 h), respectively.

E. coli and *A. tumefaciens* liquid cultures were placed in Microbank™ glycerol and beads at -80°C for long term storage.

2.2.2 Generation of competent cells

E. coli and *A. tumefaciens* competent cells were prepared for transformation of DNA plasmids.

2.2.2.1 *E. coli* chemically competent cells

Competent XL1-blue *E. coli* cells were prepared as described in Inoue et al. (1990), using an antibiotic selection of 12.5 mg L^{-1} tetracycline.

2.2.2.2 *A. tumefaciens* competent cells

Competent GV3101 *A. tumefaciens* cells were prepared as described in McCormac et al. (1998), with the following additions. Antibiotic selection was 25 mg L^{-1} rifampicin for the transformed plasmid and 25 mg L^{-1} gentamicin to maintain the Ti helper plasmid. Cells were grown until the absorbance at 600 nm (A_{600}) reached 0.6, before being pelleted by centrifugation at 5,000 rpm (Eppendorf Centrifuge 5424 R). Collected cells were washed three times – first with 1 volume, then 0.5 volumes, of sterile, ice-cold water, then 0.02 volumes of sterile, ice-cold 10% (v/v) glycerol. Washed, pelleted cells were re-suspended in 0.01 volumes of sterile, ice-cold 10% glycerol, before aliquoting 25 µL of cells and snap freezing them in liquid nitrogen. The cells were stored at -80°C until they were used in cloning.

2.2.3 *E. coli* heat-shock transformation

Heat-shock transformation of chemically competent XL1-blue *E. coli* cells with plasmid DNA was routinely used to stabilise and replicate plasmids. A 50 µL aliquot of competent cells were thawed on ice. A 1 µL aliquot of GenScript prepared plasmid DNA (described in section 2.1), 5 µL of a ligation reaction (section 2.2.9), or 2 µL of a LR gateway reaction (section 2.2.10) were incubated with the competent cells on ice for 15 min, then given a heat-shock of 42°C for 45 s before being put immediately back on ice. Cells were recovered from the heat-shock by adding 350 µL of LB-broth + glucose and putting them in a 37°C shaking incubator (250 rpm) for 1 h. Transformed cells were spread onto LB-agar plates with the appropriate selection, and the resulting colonies made into liquid cultures.

2.2.4 *E. coli* plasmid extraction

Plasmid DNA was extracted from *E. coli* liquid cultures for DNA cloning purposes such as restriction endonuclease digests and setting up ligations, LR gateway and DNA sequencing reactions. Plasmid from overnight *E. coli* liquid cultures was extracted using QIAprep® Mini-prep Kit, following the manufacturer's instructions but using P1, P2, N3, PB and PE solutions prepared in-house (Appendix 2.5). Extracted DNA was quantified spectrophotometrically (section 2.2.5.1).

2.2.5 DNA quantification

Plasmid DNA, extracted from *E. coli* liquid cultures (section 2.2.4) and gels (section 2.2.8), and genomic DNA were routinely quantified to check the quality and quantity of the DNA, and to inform volume requirements for downstream applications. Quantification methods used were Nanodrop Spectrophotometry (section 2.2.5.1) and Qubit fluorometry (section 2.2.5.2).

2.2.5.1 NanoDrop Spectrophotometry

A 1.5 µL aliquot of the DNA sample was loaded onto the pedestal of the NanoDrop® spectrophotometer ND-1000. The absorbance at 260 nm (A_{260}) was used to determine the concentration of the DNA, blanked with the elution substance (usually sterile water). Absorbance ratios determined the purity of the DNA. Generally, a 260/280 ratio below 1.8 can indicate protein contamination and 260/230 below 1.8 can indicate carbohydrate contamination.

2.2.5.2 Qubit fluorometry

DNA was quantified using an Invitrogen™ Qubit™ Fluorometer and the Invitrogen Qubit dsDNA Broad Range Assay Kit, following the manufacturer's instructions.

2.2.6 Restriction endonuclease digests

Several different restriction endonuclease enzyme digests were used during DNA cloning to extract required sequences or check correct ligation products were obtained. In a standard 15 µL reaction, 1.5 µL of 10× buffer suitable to the enzyme used was added with 500 ng DNA and 10 U (1 µL) of enzyme (Roche enzyme and SuRE/Cut Buffers). Reactions were incubated at 37°C for 1–2 h before digested DNA fragments were separated by agarose gel electrophoresis (section 2.2.7). When digests were used to extract DNA constructs from pUC57, DNA bands matching the required extract were excised from the gel and the DNA purified (section 2.2.8).

2.2.7 Agarose gel electrophoresis

Agarose gel electrophoresis was routinely used for diagnostic purposes involving the separation of Polymerase Chain Reaction (PCR) products and restriction endonuclease-digested DNA. The required amount of agarose to make a 1% (w/v) gel (or other appropriate percentage) was heat-dissolved in 1× Tris-borate EDTA (TBE) buffer and, once cooled to 56°C, 0.5 µg mL^{-1} Invitrogen Ethidium Bromide Solution was added to the gel solution. The gel was poured into a casting tray containing an appropriately sized comb to form wells and left to set. Once set, the casting tray was submerged in a gel tank containing 1× TBE buffer. Gel loading dye (10×) was added to the DNA samples at a final concentration of 1× before the samples, alongside a 1 kb (kilobase) Plus ladder for comparative size estimates, were loaded into the wells of the gel. The Bromophenol Blue dye (runs at ~200 bp) was usually used. However, Xylene Cyanol dye (runs at ~ 2 kb) was used when bands at ~200 bp were expected. Small gels were run at 120 V and large gels at 150 V, until desired DNA separation was achieved. The addition of fluorescent ethidium bromide allowed DNA bands to be visualised using a UV transilluminator (BioRad Gel Doc XR) and photographed using Image Lab software. The preparation of all solutions and dyes used for gel electrophoresis are described in Appendix 2.6.

2.2.8 DNA gel extraction

DNA gel extraction was regularly used following restriction endonuclease enzyme digestion (section 2.2.6) and agarose gel electrophoresis (section 2.2.7), to extract specific DNA fragments for use in future cloning steps. Before the gel was imaged, it was placed onto a UV table, where a razor blade

was used to cut the desired bands out of the gel. Care was taken to reduce the exposure of the desired DNA band to UV light, to prevent damage that would inhibit downstream applications. The excised bands were transferred to a 1.5 mL micro-centrifuge tube. The rest of the gel was then visualised (section 2.2.7) to check that the target band had been removed. DNA was extracted and purified from the gel slice using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research), following the manufacturer's instructions.

2.2.9 Ligation reactions

Before ligation reactions could be performed, the p*U6*$_{pro}$/pENTR entry vector was digested (section 2.2.6) with the same restriction endonuclease enzyme (e.g., *Bsa*I) as the insert. The vector and insert were quantified spectrophotometrically (section 2.2.5.1), and the following formula was used to determine the amount of insert and vector to use, where the ratio of insert:vector was 3:1.

$$\text{ng insert required} = \frac{\text{\_kb insert} \times 50 \text{ ng vector}}{\text{\_kb vector}} \times 3$$

In a standard 20 µL ligation reaction, the appropriate volume of insert and digested vector were added to 5 U (1 µL) KAPA T4 DNA Ligase (Kapa Biosystems, Inc) and 2 µL of matching T4 DNA Ligase buffer. The reaction was incubated at 16°C all day (~6 h) before 5 µL was heat-shock transformed into *E. coli* (section 2.2.3). Products of the reaction were screened for correct ligation by a restriction endonuclease digest (section 2.2.6).

2.2.10 LR Gateway recombination reactions

LR gateway recombination reactions were set up by quantifying the entry clone and destination vector spectrophotometrically (section 2.2.5.1) and adding 500 ng and 200 ng of each, respectively, to an 8 µL reaction with 1× Tris EDTA (TE) buffer. A 1 µL aliquot of Invitrogen Gateway LR Clonase II Enzyme Mix was added, before leaving the reaction to incubate at room temperature (~25°C) all day (~6 h). A 1 µL aliquot of Proteinase K solution (2 µg µL^{-1}) was added to the reaction, incubated at 37°C for 10 min, then 2 µL was heat-shock transformed into *E. coli* (section 2.2.3). Products of the reaction were screened for correct recombination by DNA sequencing (section 2.2.11).

2.2.11 DNA sequencing

DNA sequencing was routinely performed on plasmid DNA or PCR products for diagnostic purposes. Sequencing was performed in the forward direction, or forward and reverse directions, depending on the DNA product size. In a 20 µL sequencing reaction, 300 ng of plasmid DNA was used for

smaller plasmids and 600 ng for larger plasmids. For ExoSAP-treated PCR-products (section 2.4.5.3), ~25 ng was used. Primers specific to the region of interest were added to a final concentration of 4 pmol and the remaining volume was made up with sterile water. Reactions were incubated at 98°C for 5 min to denature the DNA (by inducing single strand breaks), which can improve sequencing quality from supercoiled templates. Sanger sequencing was performed using Applied Biosystems Inc (ABI) instrumentation and the BigDye Terminator V3.1 Cycle Sequencing Kit as a commercial service by the Massey Genome Service (Massey University, Palmerston North). Geneious version 2022.0 was used to analyse sequencing reads, which involved aligning reads to a reference sequence and manually correcting ambiguities introduced during sequencing.

2.2.12 Electroporation transformation of *A. tumefaciens*

Electroporation of competent GV3101 *A. tumefaciens* cells was used to transform completed CRISPR/Cas9 and over-expression DNA constructs into *A. tumefaciens* for use in Marchantia spore transformation. A 0.5 µL aliquot of plasmid DNA was added to 25 µL GV3101 cells thawed on ice, mixed gently, then put between the electrodes of an ice-cold BioRad Gene Pulser/MicroPulser Electroporation Cuvette (0.1 cm gap). Cells were electroporated in a BioRad Gene Pulser Xcell™ at 2400 V (voltage), 25 µF (capacitance) and 200-ohm (resistance), then LB-broth + glucose was used to flush out the cells into a new 1.5 mL micro-centrifuge tube. Cells were left to recover for 2 h in a 28°C shaking incubator (250 rpm), before being plated onto LB-agar plates with 100 mg L^{-1} spectinomycin and 25 mg L^{-1} gentamicin and grown in a 28°C standing incubator for 48 h.

2.3 Plant material and growth conditions

Marchantia accessions used were Aud-2 (female) and Sey-1 (male). Male and female sexual structures (antheridiophore and archegoniophore, respectively) were induced by growing Aud-2 and Sey-1 on Jiffy-7 sphagnum peat pellets (Egmont Seeds, New Zealand) under excess white light conditions, then with additional supplemented far-red light, and were crossed to produce sporangia (containing spores for transformation) as described in Chiyoda et al. (2008). Sporangia were originally generated and archegoniophore heads collected (in 1.5 mL micro-centrifuge tubes) by Dr Nick Albert.

Unless otherwise stated, plants were grown in controlled environment conditions, at 25°C, with a 16 h photoperiod of 30 µmol m^{-2} s^{-1} light intensity provided by fluorescent tubes. Plants were grown in tubs on 0.5× B5 media (Appendix 2.7) either with or without sterile Munktell 70 mm filter paper discs or Membrane Solutions MCE membrane filters. Plants were maintained, or propagated for new

experiments, by plating gemmae (somatic propagules) onto 0.5× B5 media, with the addition of 200 µM acetosyringone (Appendix 2.7) when required.

All plants were grown within a PC2 containment Facility at Plant and Food Research, Palmerston North, New Zealand.

2.4 *M. polymorpha* spore transformation and regeneration

Transgenic plants were routinely generated by transforming Marchantia spores with *A. tumefaciens* harbouring DNA constructs created through DNA cloning (sections 2.1 and 2.2). The transformation protocol is shown in Figure 2.3 and outlined in the sections following. The basic protocol was developed by Dr Nick Albert as an improvement of the S-AgarTrap method (Tsuboyama & Kodama, 2013; Tsuboyama & Kodama, 2018a), establishing a more simple, repeatable and efficient method. However, the protocol described also includes the optimisations discovered in Chapter 3 since these were required to perform future experiments.

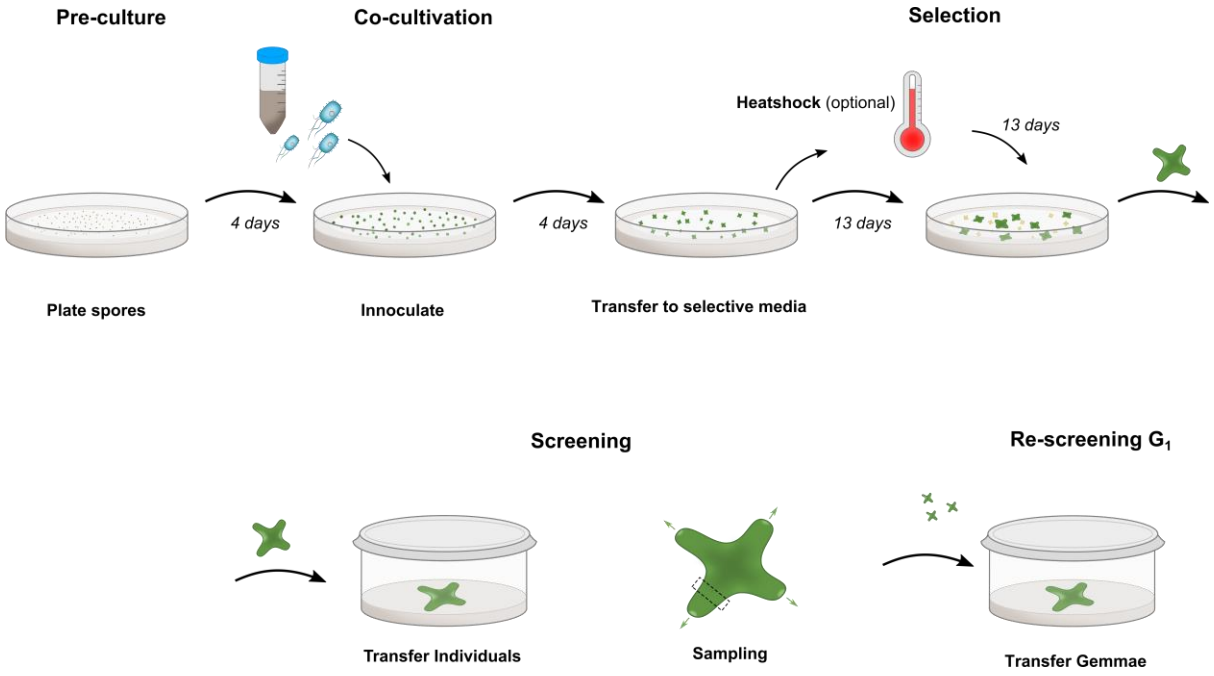


Figure 2.3. Standard *Agrobacterium tumefaciens*-mediated *Marchantia polymorpha* spore transformation protocol. Sterilised spores were plated, grown for four days, co-cultivated with *A. tumefaciens* harbouring a construct of interest for four days, then transferred to a selective media for two weeks (with the optional addition of a heat-shock). Regenerated sporelings were transferred to individual tubs and grown for two weeks before plants were sampled for screening and the positive transgenic plants propagated through gemmae (G_1 stage) and re-screened. (Figure provided by Dr Nick Albert.)

2.4.1 Spore sterilisation and plating

Marchantia spores required surface-sterilisation before plating to prevent contamination. Spores were collected from sporangia by adding a 0.5 g L^{-1} chlorine solution (sodium dichloroisocyanurate dissolved in sterile water, made up from Milton Anti-bacterial tablets) to Marchantia archegonia and crushing open the sporangia using a plastic pestle until the solution turned yellow. This spore/chlorine solution was transferred to a new 1.5 mL micro-centrifuge tube and incubated for 15 min. Spores were pelleted by centrifugation (12,000 rpm), then washed twice with sterile water before being diluted in 80 mL sterile water. For transformation, a 1 mL aliquot of spores was plated onto sterile filter paper on 0.5× B5 media containing 200 µM acetosyringone and swirled around the tub for even distribution. Spores were grown for four d.

2.4.2 Growth of *A. tumefaciens*

A. tumefaciens harbouring the DNA constructs of interest were grown up from glycerol stocks. Frozen glycerol bead stocks were thawed on ice, then a single bead containing the *A. tumefaciens* was streaked across an LB-agar plate containing 100 mg L^{-1} spectinomycin and 25 mg L^{-1} gentamicin selection. Plates were incubated at 28°C for 48 h.

2.4.3 Co-cultivation of spores with *A. tumefaciens*

A. tumefaciens cells were scraped into 10 mM $MgCl_2$, MES (pH = 5.6; Appendix 2.8) and 200 µM acetosyringone and left on the bench for 2–4 h until the A_{600} was ~1.0. The *A. tumefaciens* solution was diluted 10-fold in 10 mM $MgCl_2$, 10 mM MES, making the A_{600} ~0.1. Once the spores had grown for four d, a 1 mL aliquot of the diluted *A. tumefaciens* solution was added to the tubs of spores and swirled to distribute evenly. Spores and *A. tumefaciens* were left to co-cultivate for four d.

2.4.4 Spore regeneration

After four d co-cultivation, 3 mL of 500 mg L^{-1} ticarcillin was pipetted onto each tub to loosen any *A. tumefaciens* overgrowth, before transferring the filter paper containing the spores onto selective 0.5× B5 media containing 500 mg L^{-1} ticarcillin (to kill *A. tumefaciens*) and an antibiotic specific to the antibiotic resistance gene the DNA construct contained. This ensured only transformed spores survived. For generating CRISPR/Cas9 mutants, antibiotic selection was 8 mg L^{-1} hygromycin, corresponding to the hygromycin resistance gene (*hptII*) of pMpGE010. This concentration was chosen based on the results of a hygromycin kill curve performed (Appendix 5). Selection for generating over-expression plants was 8 mg L^{-1} geneticin (G-418), corresponding to the *nptII*

resistance gene on the pNWA101 plasmid. This selection was chosen based on Albert et al. (2018) using a similar geneticin concentration of 7.5 mg L^{-1}. Spores were grown for 13 d (based on knowledge from the kill curve) until a tight kill was achieved, which always delivered many putative transgenic sporelings. For the next 2–3 weeks, as sporelings started to develop a thallus, ~30 putatively transgenic plants were lifted onto individual tubs of 500 mg L^{-1} ticarcillin 0.5× B5 media to prevent *A. tumefaciens* growth. Plants were grown for a further ~two weeks, until they were large enough to be sampled for genomic DNA extraction.

2.4.5 Screening regenerated spores

Sections of regenerated plants were screened for presence of the transgene by extracting genomic DNA from collected tissue (section 2.4.5.1), PCR-amplifying the target region (section 2.4.5.2), treating PCR products with Exonuclease I and Shrimp Alkaline Phosphatase (ExoSAP) to remove excess primers and nucleotides, before sending samples for DNA sequencing to confirm that plants were transgenic (section 2.4.5.3).

2.4.5.1 Genomic DNA extraction

Once plants were 4–5 cm in diameter, an individual thallus branch was sampled behind the meristem and stored at -20°C. Only the meristem was grown on (the rest of the plant discarded) to minimise problems from chimeric tissue. This sampling method ensured that if mutant tissue was identified it could continue to be propagated. Genomic DNA was extracted from each sample by using the Bead Ruptor Elite Bead Mill Homogenizer (Omni International) to grind the tissue in 500 µL of CTAB extraction buffer (Appendix 2.9) and 6 µL β-2-mercaptoethanol, using two ceramic beads. Cells were lysed by incubating samples at 60°C for 30 min, then DNA was separated from proteins and other cell debris using an equal volume of chloroform:IAA (24:1; v:v) followed by centrifugation (12,000 rpm). DNA in the collected supernatant was precipitated using 500 µL of cold isopropanol, followed by a 5 min incubation on ice, was pelleted by centrifugation (15,000 rpm) then washed with 70% (w/v) ethanol to remove salts. Once the ethanol was removed and the pellet was dry, 25 µL of 10 mM Tris HCL pH 8.0 and 1 µL RNase A was added to the DNA. DNA was left to rehydrate in the fridge overnight before quantification (section 2.2.5.1).

2.4.5.2 Touchdown Polymerase Chain Reaction (PCR)

'Touchdown PCR' was routinely used for genomic DNA amplification, to determine if putative CRISPR/Cas9 mutants contained gene-specific mutations or if putative over-expression plants contained specific transgenic regions. It was used to prevent non-specific DNA amplification (which

often occurred in standard PCR conditions), by starting with a high annealing temperature and reducing it by 1°C per cycle for five cycles, until the optimal annealing temperature was reached. This enriched the reaction for more specific products in the early PCR rounds, which acted as a template for later cycles.

Each PCR contained 4 nmol of dNTPs (deoxy-nucleotide-triphosphates), 4 pmol of each primer, 4 µL of KAPA2G 5× buffer A, 2U KAPA2G Robust HotStart *Taq*-like polymerase enzyme, and 1 µL of DNA, made up to 20 µL with sterile water. PCR primers were designed using Geneious version 2022.0 and ordered from Integrated DNA Technologies™. Primer stocks were prepared at 100 µM with sterile water, stored at -20°C and were diluted to 10 µM for use in PCRs. Primers used are listed in Appendix 4. Annealing temperature and length of extension varied depending on the primers used and the product size, respectively, but the general touchdown PCR conditions were the following: initial denaturation at 94°C for 2 min; 5 cycles of 95°C for 30 s (denaturation), 60°C (-1°C per cycle) for 30 s (annealing), 72°C for 1 min 30 s (extension); 31 cycles of 95°C for 30 s (denaturation), 55°C for 30 s (annealing), 72°C for 30 s (extension); final extension at 72°C 1 min; hold at 12°C. PCR products were visualised by gel electrophoresis on a 1.5% (w/v) agarose gel, which is a higher % than standard, to allow better separation of the relatively small PCR products (section 2.2.7).

2.4.5.3 ExoSAP treatment and DNA sequencing of PCR products

Following results from gel electrophoresis, the first 10 putatively transgenic lines were treated with ExoSAP to remove excess primers and dNTPs, before being submitted for DNA sequencing (section 2.2.11).

ExoSAP treatment involved adding 0.15 µL Exonuclease I (20 U μL^{-1}; New England BioSciences) and 0.9 µL Shrimp Alkaline Phosphatase (rSAP 1 U μL^{-1}; New England BioSciences), made up to 3 µL with sterile water, to 7 µL of PCR sample. The reaction was incubated at 37°C for 30 min, then 95°C for 5 min.

Confirmed CRISPR/Cas9 mutant plants were kept and propagated through gemmae (G_1 generation) to ensure non-chimeric plants were obtained, because gemmae are derived from a single somatic cell. G_1 plants were re-screened by PCR and sequencing. Confirmed transgenic over-expression plants were not re-screened before being used in further experiments as CRISPR/Cas9 was not used and therefore chimeric plants were not likely.

2.5 Imaging *M. polymorpha* plants

For microscope images, the Leica M205 FA Microscope was used with a Leica DFC550 Camera and Leica CLS 150 X Light source, alongside the LAS V4.8 imaging software.

For whole tub images, a Nikon D5600 Digital camera and Godox LED170 side lights was used.

2.6 Statistical analysis

The statistical significance of data generated from manipulation of transgenic plants was analysed by performing a one-way Analysis of Variance (ANOVA) of the means, using a Fisher's protected least significant differences (LSD) test and specifying an error rate of 0.05 (95% confidence). All means within an experiment were compared to each other (called "multiple comparisons"), rather than analysing the difference of one mean to another. Different letters were assigned to significantly different means. Analysis was performed using the Genstat version 22 (VSN International) software.

Chapter 3

Optimising transformation and Cas9 editing efficiencies

3.1 Introduction

Marchantia spore transformation is an important technique for generating transgenic plants, including plants with CRISPR/Cas9 mutations. However, prior attempts (by Dr Nick Albert) to implement existing protocols did not achieve a workable regeneration efficiency, so improvements to the protocol were necessary. This component of my project focussed on determining ways to optimise the transformation of Marchantia spores and to increase the Cas9 mutation rate, to produce a robust and efficient protocol for generating transgenic/mutant lines.

Recent transformation protocols use *A*. *tumefaciens* to introduce foreign DNA to plant cells. Acetosyringone (AS) is a phenolic compound that is a potent inducer of *A*. *tumefaciens vir* genes (Stachel et al., 1985). For this reason, inclusion of AS in the *A. tumefaciens* co-cultivation medium has increased transformation efficiency in many plant species, including *Arabidopsis* (Sheikholeslam & Weeks, 1987), durum wheat (*Triticum turgidum* L. var. *durum*) (He et al., 2010) and *Citrus* (Dutt & Grosser, 2009). This is standard practice in Marchantia spore transformation (section 2.4), but was not sufficient to achieve a high transformation efficiency. Guivarc'h et al. (1993) and Aggarwal et al. (2011) reported increases in transformation efficiency by including AS in the pre-culture media, rather than only during *A. tumefaciens* inoculation. This has not been attempted in Marchantia before, so I examined whether adding AS to the pre-culture media during Marchantia spore transformation could improve the transformation efficiency.

An ideal CRISPR/Cas9 system would generate non-chimeric plants. Chimeric plants are common in Marchantia spore transformation and generate when mutations occur in some spore cells but not others. This results in a single plant containing wild type and mutant sectors, which adds complexity to screening and handling of candidate mutant plants. The number of chimeric plants could be reduced if Cas9 gene editing occurred early in spore development with high efficiency. This may also encourage larger mutations, which usually result in complete loss of gene function, because mutations are more likely to simultaneously occur at several sgRNA target sites. Studies have reported that providing plants with a heat-shock (HS) of 37°C improved Cas9 mutagenesis (Blomme et al., 2022; LeBlanc et al., 2018).

The effect of HS treatments on Cas9 editing efficiency has not been investigated in Marchantia but may be a way to improve recovery of mutant lines. Thus, I examined whether providing spores with a HS during the selection stage of spore transformation could improve the Cas9 editing efficiency. This required CRISPR/Cas9 mutants to be generated that had a simple, visible phenotype that could be scored for editing efficiency based on the size of mutant sectors. Yelina et al. (2023) discovered that Marchantia mutants in the *Golden 2-like* (*GLK*) gene have a pale green phenotype owing to reduced chloroplast development, so CRISPR/Cas9 mutants of the Mp*GLK* gene (*Mapoly0108s0060*) were generated for this study.

3.2 Materials and methods

3.2.1 CRISPR-Cas9 and over-expression constructs

A CRISPR/Cas9 construct targeting the Marchantia Mp*GLK* gene was designed and cloned into a binary vector as described in section 2.1 ad 2.2 (pRMY32; Appendix 3.2 for plasmid map). Two high quality sgRNAs (as described in section 2.1) were designed in exon four (sgRNA 277) and exon six (sgRNA 192) of the Mp*GLK* (*Mapoly0108s0060*) coding sequence (Figure 3.1), purposely avoiding the first three exons because the gene models in the Marchantia reference genome suggest alternative transcription start sites exist. Marchantia gene IDs used in this thesis are based on an older, but still high quality, genome assembly. Conversions of all gene IDs to the newer MpTak_v6.1 genome version are given in Appendix 6.

The *35S*:*Myb14* and *35S*:*GFP* constructs were previously designed and cloned by Dr Nick Albert, as described in section 2.1 and 2.2.

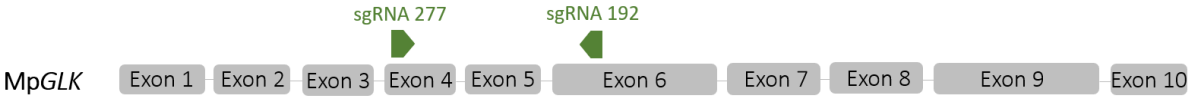


Figure 3.1. Position of the two single guide RNA (sgRNA) molecules designed to target exons four and six of the *Golden 2-like* (Mp*GLK*; *Mapoly0108s0060*) gene in *Marchantia polymorpha*.

3.2.2 Stable transgenic lines and growth conditions

Mutants of Mp*GLK* and *35S*:*Myb14* and *35S*:*GFP* transgenic plants were generated by *A. tumefaciens*-mediated Marchantia spore transformation (section 2.4). During transformation with the Mp*GLK* CRISPR/Cas9 construct, spores were diluted 5-fold (in 400 mL sterile water instead of 80 mL) before being plated, to ensure better visualisation of phenotypes when spores were regenerating. Antibiotic selection was 8 mg L^{-1} hygromycin for the Mp*GLK* construct and 8 mg L^{-1} G-418 for generating *35S*:*Myb14* and *35S*:*GFP* plants. Screening the *glk* and over-expression plants generated based on visual phenotypes (*glk*: pale tissue; *Myb14*: red/bronze thallus; *GFP*: green fluorescence), and not by PCR, was sufficient.

The effects of AS were assessed by adding different concentrations of it (0, 50, 200 and 1000 µM) to 0.5× B5 pre-culture media during the spore transformation of *35S*:*Myb14* and *35S*:*GFP* plants. This range of concentrations was chosen as 200 µM is used during co-cultivation and previous studies on other plants used 100 µM or 200 µM in pre-culture media (Aggarwal et al., 2011; Boase et al., 1998; Guivarc'h et al., 1993). Once transgenic sporelings started to regenerate, they were imaged (section 2.5) and used to calculate the transformation efficiency (section 3.2.3). Three biological replicates, consisting of separate tubs, were used per transgenic line and treatment.

The HS conditions applied to spores transformed with the Mp*GLK* CRISPR/Cas9 construct were: 28°C/16 h, 28°C/24 h, 28°C/48 h, 37°C/16 h or 37°C/24 h. Treatments were applied on the same day as they were transferred to the selective media, using non-shaking bacterial incubators lacking illumination. A no-HS control at 25°C (standard culture conditions) was also included. The 37°C HS was chosen because the *Cas9* gene originated from the bacterium *Streptococcus pyogenes* (Gao et al., 2019; Le Rhun et al., 2019), which has an optimal growing temperature of 37°C (Gera & McIver, 2013), and the 28°C HS because of availability of incubators at this temperature and as it is lower than 37°C but higher than the 25°C regular growing temperature. Treatment times were chosen based on prior knowledge gained from a pilot experiment conducted by Dr Nick Albert (unpublished). Regenerated sporelings were imaged (section 2.5) 13 d after transfer to selective media to calculate transformation efficiency (section 3.2.3). For the next two to six weeks, regenerated sporelings were picked onto new tubs of 0.5× B5 media containing 500 mg L^{-1} ticarcillin until there were at least 30 plants per treatment. Plants were imaged (section 2.5) and phenotypically scored for calculating editing efficiency (section 3.2.4).

3.2.3 Calculating transformation efficiency

Spore transformation efficiency was determined by taking four representative microscopic images (section 2.5) of the tubs of spores, one in each quadrant of the tub (Figure 3.2), keeping the same layout for each tub/biological replicate both within and between treatments. The same images were taken with a GREEN FLUORESCENCE PROTEIN (GFP) filter as this provided clearer images of the dead spores. Transformation efficiency was calculated by counting the number of transgenic (alive) spores and the number of dead spores in each image and using the following equation. The average transformation efficiency per treatment was the average between the four images of a tub, and then between all tubs/biological replicates.

$$\text{Transformation efficiency (\%)} = \frac{\text{Number of transgenic spores}}{\text{Total number of spores}} \times 100$$

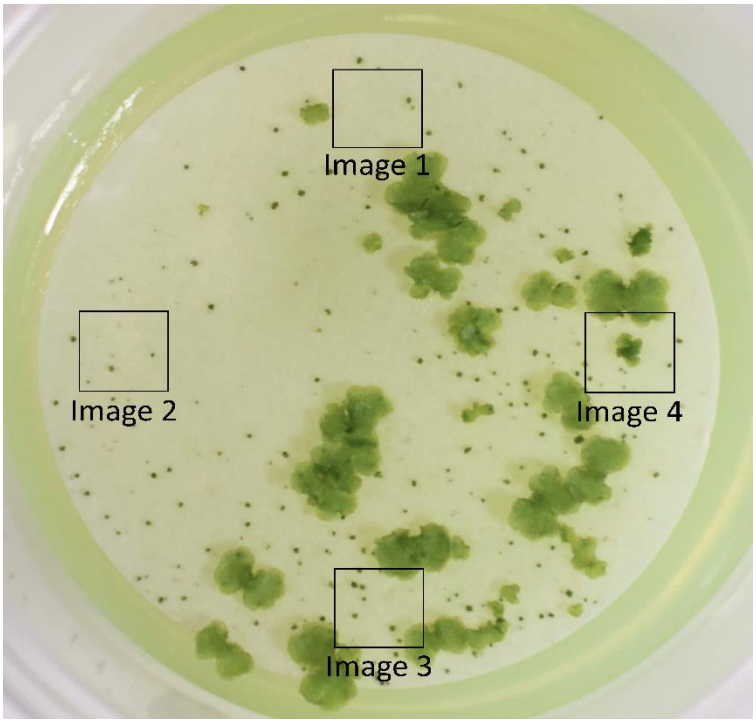


Figure 3.2. Example layout of the four microscopic images taken for each tub of *Marchantia polymorpha* spores to calculate the transformation efficiency.

3.2.4 Phenotypic scoring of Cas9 editing efficiency

Cas9 editing efficiency was determined by scoring plants based on the physical size of the mutant sectors they contained. The size of mutant sectors depends on when the editing occurred; mutations occurring during early stages of tissue regeneration/development will give rise to larger sectors. The scoring involved assigning plants to five categories: 1) plants appeared wild-type, 2) plants were mostly wild-type with at least one mutant section less than 2 mm, 3) plants were mostly wild-type with at least one mutant section of at least 2 mm, 4) more than half of the plant had the mutant phenotype, and 5) plants were entirely mutant (Figure 3.3). Following scoring, data were quantified by determining the percentage of total plants in each phenotypic category, for each treatment.

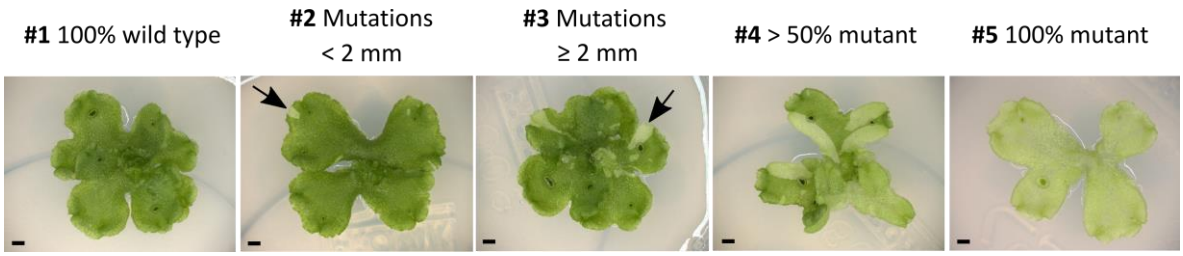


Figure 3.3. Examples of *Marchantia polymorpha* plants assigned to each phenotypic category used to score Cas9 editing efficiency. Category 1: plants are wild-type; Category 2: plants are mostly wild-type with at least one mutant section less than 2 mm (indicated by an arrow); Category 3: plants are mostly wild-type with at least one mutant section of at least 2 mm (sector indicated by an arrow); Category 4: more than half of the plant is mutant; Category 5: plants are entirely mutant. Scale bars represent 2 mm.

3.3 Results

3.3.1 Pre-culture acetosyringone improves transformation efficiency

Various concentrations of AS were added to the pre-culture media during spore transformation to determine whether it improves transformation efficiency. Pre-culture is a period where the spores can germinate into sporelings and begin to grow prior to exposure to *A. tumefaciens*.

The addition of AS to the media altered the germination and growth of spores into sporelings. The control (0 µM) had a dense lawn of sporelings. However, with increasing concentrations of AS, the density was reduced and spores were smaller, reaching the point where there was no spore growth using 1000 µM AS (Figure 3.4).

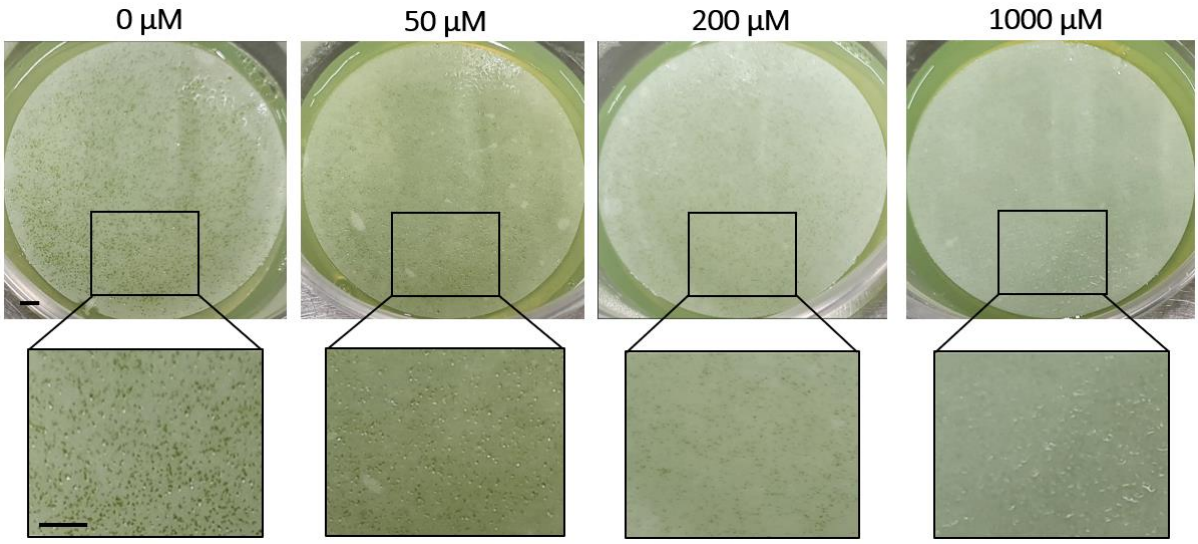


Figure 3.4. Representative density of *35S*:*Myb14 Marchantia polymorpha* spores following pre-culture with different acetosyringone concentrations. Images taken after co-cultivation, but prior to selection. Scale bar represents 500 µm.

Once the transformed spores had gone through selection, the sporeling regeneration was assessed. In both the *35S*:*Myb14* and *35S*:*GFP* transformations, regeneration did not occur when AS was absent from the media (0 µM), as shown by the tubs containing only dead sporelings (Figure 3.5). The addition of 50 µM AS resulted in a visually small amount of regeneration, and 200 µM resulted in a substantial number of regenerated sporelings. No regeneration was observed when 1000 µM AS was used, owing to the lack of spore growth (Figure 3.4).

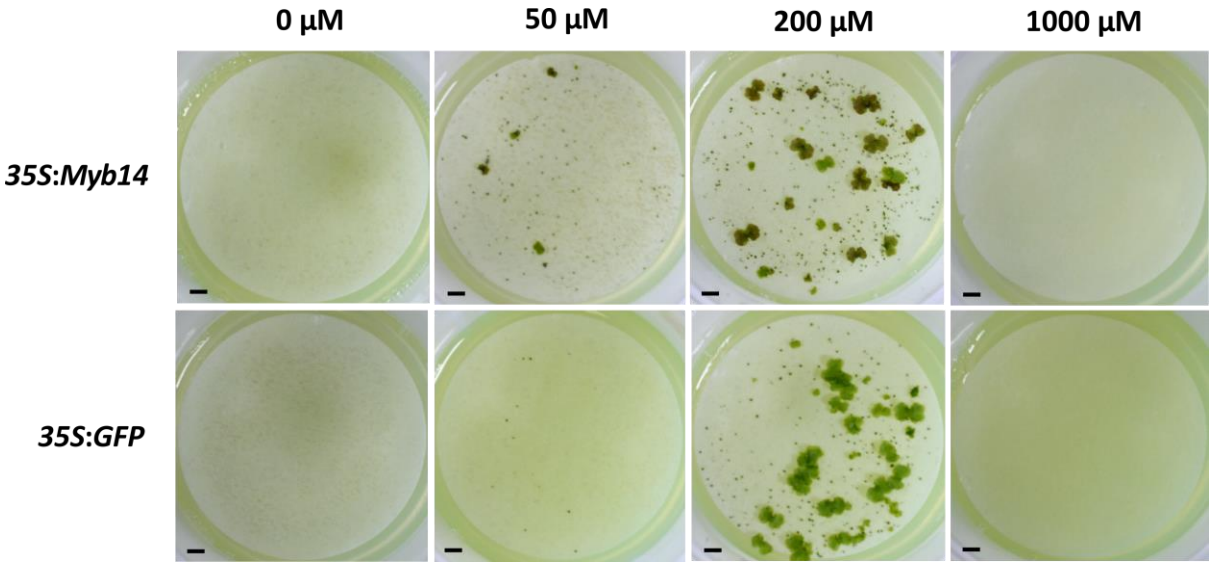


Figure 3.5. Representative density of regenerated *35S*:*Myb14* and *35S*:*GFP Marchantia polymorpha* sporelings following pre-culture with different acetosyringone concentrations. Images were taken 13 days after transfer to selection. *35S*:*Myb14* sporelings are darker-coloured due to auronidin production. Scale bars represent 500 µm.

The effect of pre-culture AS on spore regeneration was quantified by calculating the spore transformation efficiency for at least three tubs (biological replicates) of each AS treatment and transgenic plant line (*35S*:*Myb14* and *35S*:*GFP*). Similarly to the result shown in Figure 3.5, the absence of AS in the pre-culture medium (0 µM) resulted in a transformation efficiency of 0% with the *35S*:*GFP* construct and 0.1% for *35S*:*Myb14* (Figure 3.6). Addition of 50 µM AS to the pre-culture media resulted in a higher transformation efficiency, being 0.6% in the *35S*:*GFP* line and 4% in the *35S*:*Myb14* line. The 200 µM concentration produced the highest spore regeneration, having a transformation efficiency of 6% for the *35S*:*GFP* line and 11% for the *35S*:*Myb14* line. The 1000 µM AS concentration had a transformation efficiency 0% for both transgenic lines.

Another interesting observation, both visually (Figure 3.5) and quantifiably (Figure 3.6), was that there was a noticeable difference in transformation efficiency between the *35S*:*Myb14* and *35S*:*GFP*; the latter having a lower transformation efficiency than the former, for all AS treatments.

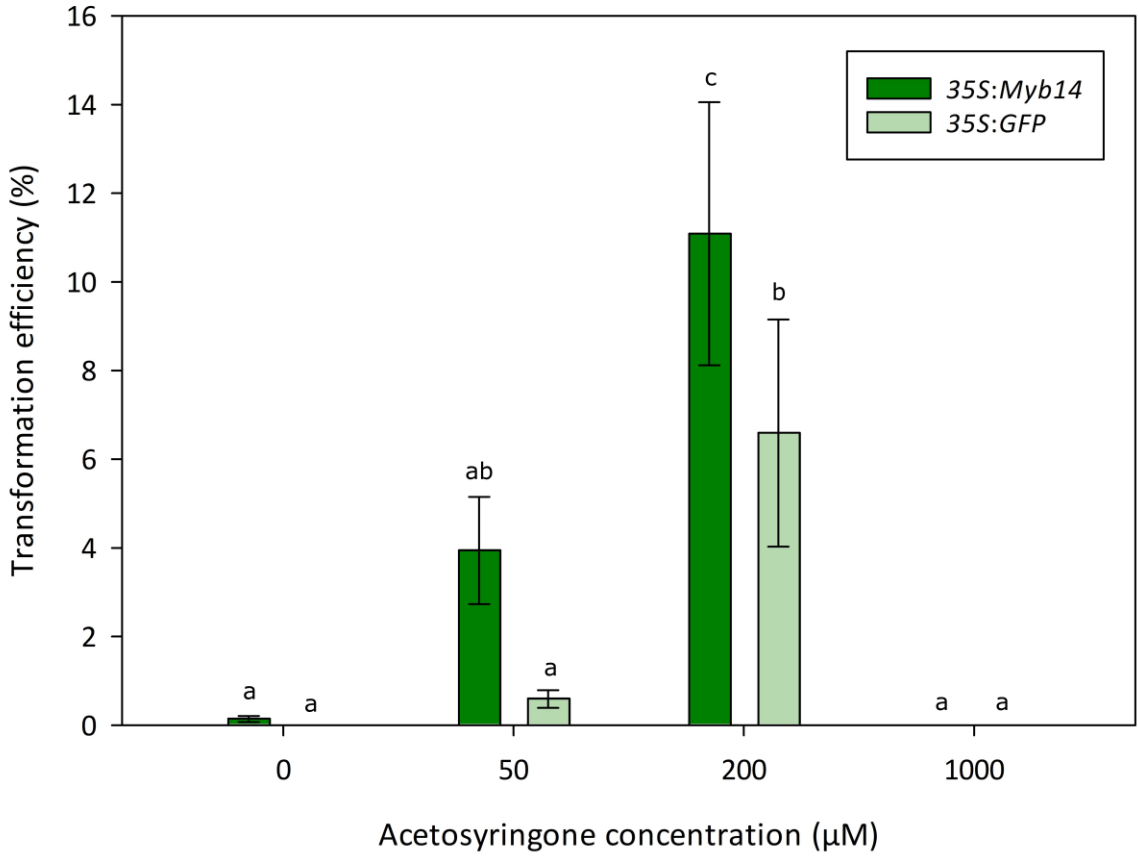


Figure 3.6. Transformation efficiency of 35S:*Myb14* and 35S:*GFP Marchantia polymorpha* spores following pre-culture with different acetosyringone concentrations. Error bars are the standard error of the means, n = 3. Significantly different means are represented by different letters (a - c).

3.3.2 Heat-shock improves Cas9 editing efficiency

The effect HS treatments had on Cas9 editing efficiency was examined using CRISPR/Cas9 and *A. tumefaciens*-mediated spore transformation to generate *glk* mutants. Regenerated sporelings were scored for editing efficiency by assignment into five categories based on the phenotypic size of the pale green coloured mutations they contained (Figure 3.3).

To quantify Cas9 editing efficiency, the proportion of total plants allocated to each category (hereafter, 'editing frequency') was calculated for all HS treatments. As shown in Figure 3.7, the 28°C/48 h, 37°C/16 h and 37°C/24 h treatments had an equally high total editing frequency (accounting for categories #2 - #5) of ~50%. Whereas the 28°C/16 h and no-HS treatments had a lower editing frequency of ~30% and the 28°C/24 h treatment had the lowest frequency of ~10%. However, a high Cas9 editing efficiency is more accurately represented by category #4 (mutations across more than half of the plant) and #5 (100% mutant) plants, because Cas9 must have been more active earlier in spore development (once the HS treatment was applied) for phenotypically larger mutant sectors to exist. Thus, considering the number of plants allocated to individual categories, the 28°C/48 h treatment had the highest Cas9 editing efficiency because it had the highest proportion of category #4 and #5 mutants (10% category #4 and 20% category #5). Moreover, the 37°C/24 h treatment had a higher editing efficiency than the 37°C/16 h treatment, because it had a higher proportion of category #5 mutants with no category #4 mutants and a higher proportion of category #3 mutants with no category #2 mutants. Additionally, the editing efficiency of the 28°C/16 h treatment was higher than the no-HS control because it produced considerably more category #5 mutants and less category #2 and #3 mutants. Surprisingly, the 28°C/24 h treatment had a lower editing efficiency than these treatments.

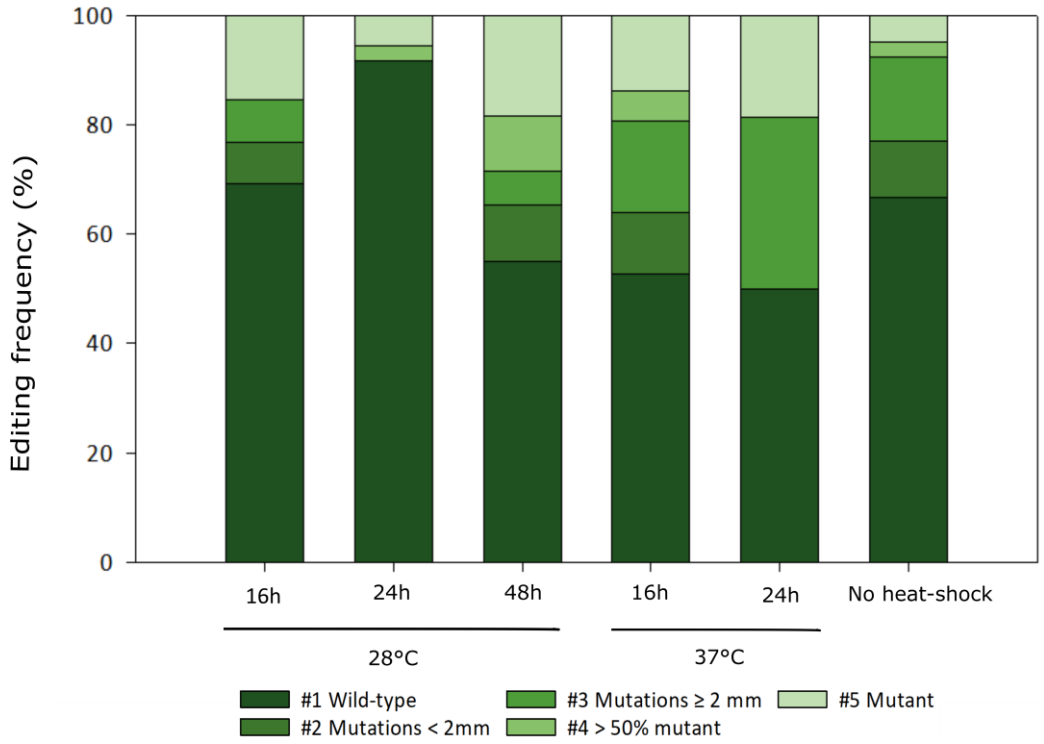


Figure 3.7. Cas9 editing efficiency of *Marchantia polymorpha* spores treated with various heat-shock conditions. Editing efficiency was determined by the number of plants assigned to each phenotypic category (#1 - #5) based on the size of mutations they contained. Error bars could not be calculated for percentages.

Generally during spore transformation, there is an initial abundance of plants regenerated following the first two weeks that sporelings are on selective media. However, for the next four weeks, new generations of transformed sporelings emerge. An aspect of interest in this study was whether there was any difference in mutation rate between the first regeneration of sporelings and the later regenerations. To assess this, the sporelings scored for mutations over the six-week period (Figure 3.7) were grouped into two-weeklong regeneration periods (Figure 3.8). Plants produced following 37°C HS treatments had the highest total editing frequency (accounting for categories #2 - #5) within the first two weeks of regeneration, having mutations in ~70% and 80% of plants in the 16 h and 24 h treatments, respectively. This reduced to 25% and 0% in the third and fourth weeks and increased slightly in the fifth and sixth weeks to 35% and 40%, respectively. Comparatively, the no-HS control had a fairly even editing frequency across each two-week period, having mutations in 50% of plants for the first two weeks of regeneration, and ~30% over the next four weeks. The 28°C treatments had a different result again: the 16 h and 24 h treatments had a similar editing frequency in the first and second and fifth and six weeks, being between 10% and 30%, with no mutants regenerated in the third and fourth weeks. The editing frequency of the 48 h

treatment appeared to increase with time, being ~30% in the first two weeks of regeneration and increasing to 60% by the last two weeks.

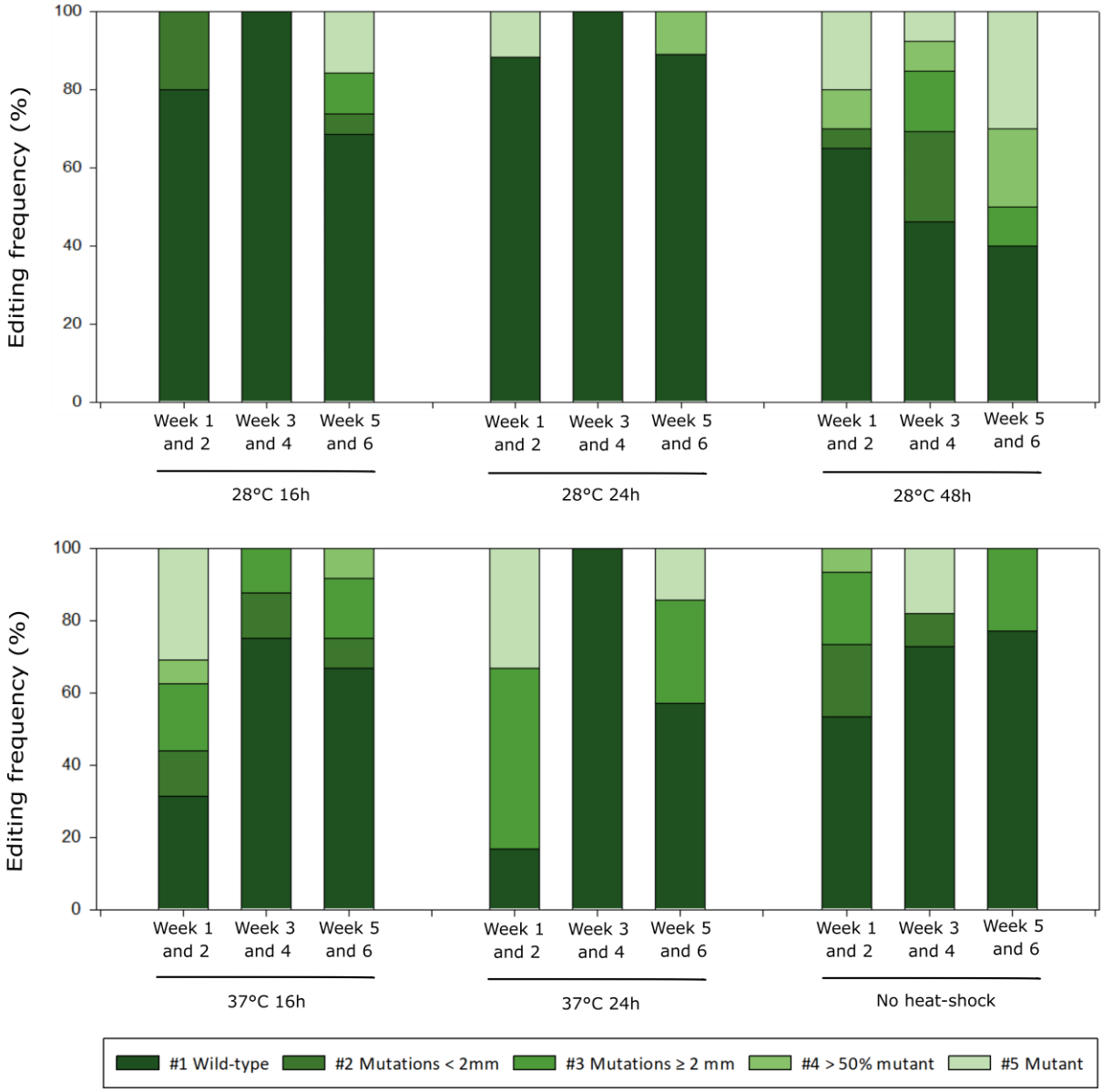


Figure 3.8. Effect of *Marchantia polymorpha* spore regeneration time on Cas9 editing efficiency of spores treated with various heat-shock conditions. Editing efficiency was scored by assigning sporelings from each heat-shock treatment into five phenotypic categories (#1 - #5) as they regenerated for a period of six weeks. The resulting editing frequencies calculated were grouped based on fortnightly regenerations. Error bars could not be calculated for percentages.

Given the HS treatments were above the normal growing temperature (25°C), their effect on spore death was determined by calculating the spore transformation efficiency (Figure 3.9). Although there were differences between HS treatments, no clear pattern was apparent for the effect of temperature or duration of treatment. This is because the 37°C/16 h HS treatment had a similar transformation efficiency to the no-HS control (~12%) and unexpectedly, the 28°C/16 h and 28°C/24 h treatments had a transformation efficiency of 8% and 6%, respectively, that were also not significantly different. Predictably, the 37°C/24 h treatment caused extensive spore death, having a transformation efficiency of ~3% that was significantly lower than the 37°C/16 h and no-HS treatments. However, surprisingly, the 28°C/48 h treatment had a transformation efficiency of ~17%, which is higher than the no-HS control and significantly higher than the 28°C/16 h and 28°C/24 h treatments.

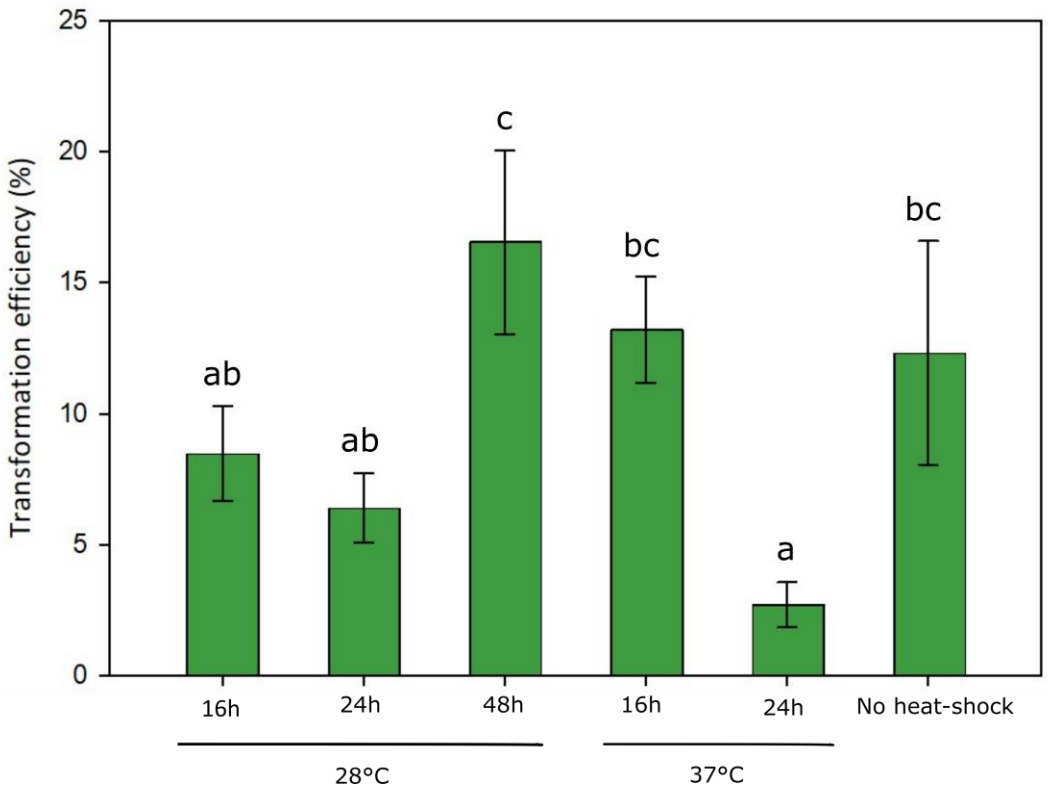


Figure 3.9. Transformation efficiency of *Marchantia polymorpha* spores provided with various heat-shock conditions. Error bars are the standard error of the means, n = 3. Significantly different means are represented by different letters (a - c).

3.4 Discussion

3.4.1 Pre-culture acetosyringone improves transformation efficiency

Adding AS to the pre-culture media during Marchantia spore transformation improved the transformation efficiency. This was observed both visually (Figure 3.5) and quantifiably (Figure 3.6). The 200 µM treatment had a transformation efficiency of 11% for the *35S*:*Myb14* line and 7% for the *35S*:*GFP* line. This was significantly higher than that of the no-AS control, which was 0.1% for *35S*:*Myb14* and 0% for *35S*:*GFP*. In subsequent experiments using the optimal 200 µM treatment, transformation efficiencies of up to ~20% have been observed (Appendix 7). It is difficult to compare the improvements of using AS in the pre-culture media obtained here to other published results, as the transformation protocol used differs markedly to published methods.

The spore transformation protocol used in this study was developed to be efficient but easier to conduct than published Marchantia spore transformation protocols. Several problems exist in published methods. For example, some protocols use liquid culture (Ishizaki et al., 2008; Nasu et al., 1997), where rhizoids of sporelings can get tangled and transformed plants are harder to distinguish. Other protocols plate spores directly onto solid media (Tsuboyama & Kodama, 2013; Tsuboyama & Kodama, 2018a), which are prone to being washed away upon *A*. *tumefaciens* co-cultivation. Comparatively, many factors make the method in the present study robust, considerably easier and highly efficient. These include plating spores onto filter paper on a solid medium, which makes it far easier to manage regenerated sporelings and the spores imbed into the filter paper so are not washed off. It also has the advantage that the filter paper can be transferred to a new medium for selection, rather than applying a concentrated antibiotic directly to spores, which other methods do (Tsuboyama & Kodama, 2018a).

Other Marchantia transformation protocols have been published reporting higher efficiencies than found here. Thallus transformation (Kubota et al., 2013; Tsuboyama & Kodama, 2018a) has a reported transformation efficiency of 30-60% and gemmae transformation (Tsuboyama & Kodama, 2018c; Tsuboyama et al., 2018) almost 100%. However, because these methods are different to the method used in this study, the transformation efficiencies are not comparable. These protocols require more extensive tissue culture procedures and start with far smaller numbers of target 'explants' (thallus pieces or gemmae). Spore-based transformation protocols are simpler and start with thousands of spores, so that lower transformation efficiencies still deliver many more transgenic events. Although more work is required to produce a similar number of transgenics,

gemmae and thallus protocols allow production of transgenic plants in a uniform genetic background, while spores (being sexually produced) generate unique transgenic plant backgrounds.

There are several possible explanations for why pre-culture with AS improves transformation efficiency. Since AS is included in the *A*. *tumefaciens* inoculation medium before being cultivated with the spores, induction of *vir* gene expression should already occur regardless of whether AS was included in the Marchantia pre-culture medium. Therefore, the basis of the improved transformation efficiency probably corresponds to changes in the Marchantia tissue. AS slowed pre-culture spore growth (Figure 3.4). The reasons for this are unknown, and it has not been previously reported. Nevertheless, it is possible that this retardation of growth maintains spores in a developmental state that is more receptive to *A*. *tumefaciens* infection. However, a more probable explanation is that AS primes the spores to attract *A*. *tumefaciens*. In many eudicot plant transformation methods, plants are initially wounded during explant preparation, which causes secretion of phenolic compounds such as AS that induce the chemotaxis response of *A. tumefaciens* and thereby facilitate transformation (Ashby et al., 1987; Palmer & Shaw, 1992). Marchantia may not be a natural host for *A. tumefaciens*, and therefore it may not produce phenolic compounds that attract *A. tumefaciens*. Wounding is not part of spore transformation and, assuming Marchantia is able to take up AS from the media, it is plausible that including AS in pre-culture creates a concentration gradient that attracts *A. tumefaciens* and replaces the need for wounding (Escudero & Hohn, 1997; Kuta & Tripathi, 2005). Therefore, it seems likely that pre-culture AS improves transformation efficiency because it attracts *A*. *tumefaciens* to Marchantia sporelings, making it easy to find and infect them.

Reported angiosperm transformation efficiencies indicate that different plant species may require different concentrations of AS to have the same effect on transformation efficiency. Aggarwal et al. (2011) and Guivarc'h et al. (1993) reported increased transformation efficiency in *Eucalyptus tereticornis* and carrot with pre-treatment of 100 µM and 25 µM AS, respectively. Additionally, given the optimal concentration for Marchantia (of those tested) was 200 µM, it is advisable that a range of concentrations are initially tested for each individual species being transformed. It is likely that higher AS concentrations will give a higher transformation efficiency but based on the detrimental effects of 1000 µM AS on Marchantia spore growth (Figure 3.4), it should be expected that at a certain point, high AS concentrations will result in a lower transformation efficiency.

In the present study, the *35S*:*Myb14* gene construct consistently gave a higher transformation efficiency than the *35S*:*GFP* construct (Figure 3.6). The cause of this difference is not clear because the transformation efficiencies calculated were variable, as shown by the large error bars, and there

are many variables to consider in spore transformation. Thus, the experiment would have to be repeated to determine which variables may have caused the observed differences in transformation efficiency. However, a possible explanation is that the exact *A*. *tumefaciens* strain used in each transformation differed. *A*. *tumefaciens* used in co-cultivation originates from a single colony, and colonies could have different infection abilities resulting from variation in genetic or epigenetic aspects arising from the long-term maintenance of strains in the laboratory (termed 'lab attenuation') (Gelvin, 2003). Other possible explanations for a difference in transformation efficiency could be related to the amount of T-DNA that is integrated into the plant genome after transformation, and the expression of the transgene following integration (Gelvin, 2003). In this study, the antibiotic resistance gene is designed at the left border on the binary vector. Because T-DNA integration works from the right border to the left, incomplete T-DNA transfer could result in transformed spores not being regenerated because the antibiotic resistance gene is not active and therefore cause a lower transformation efficiency.

3.4.2 Heat-shock improves Cas9 editing efficiency

Applying a HS during Marchantia spore transformation improved Cas9 editing efficiency. The 28°C/48 h treatment had the highest editing efficiency because it produced more category #4 and #5 mutants than the 37°C/16 h and 37°C/24 h treatments (Figure 3.7). Because these categories have the largest mutant sectors, this suggests that the 28°C/48 h treatment resulted in fewer chimeric plants. However, the editing efficiency of the 28°C/48 h treatment was achieved over the entire six-week regeneration period (Figure 3.8). Whereas the 37°C/16 h and 37°C/24 h treatments produced the majority of their mutants, and importantly most of the category #5 mutants, within the first two weeks of regeneration. Thus, because enough sporelings are typically produced for downstream applications within the first two weeks of regeneration, and the 37°C/24 h treatment did not produce sufficient plants for downstream applications (Figure 3.9), a 37°C/16 h HS is the preferrable treatment to achieve a high Cas9 editing efficiency and thereby produce few chimeric plants.

The 28°C/16 h and no-HS treatments had a similar editing efficiency, as expected, but surprisingly the 28°C/24 h treatment had a considerably lower efficiency (Figure 3.7). Because the 28°C/24 h treatment also had the lowest transformation efficiency (Figure 3.9), this result could be explained by the potential loss of mutants due to contamination. Another unforeseen result was that the 28°C/48 h treatment had the highest transformation efficiency and the 37°C/16 h treatment had a similar transformation efficiency to the no-HS control (although these differences are not significant) (Figure 4.9). This suggests that these heat treatments did not affect spore growth, even though the

28°C/16 h and 28°C/24 h treatments had a reduced transformation efficiency. Currently, there is no logical explanation for these results, but it would be interesting to determine whether these results are the same with repeated transformations.

Studies using angiosperm species have shown that a HS at 37°C increases Cas9 mutation rate, but the length of treatment required to achieve the same effect varies between plant species. In *Arabidopsis*, studies have shown that three or four cycles of 37°C for 30 h, alternated with a recovery of 22°C for 42 h, can increase mutagenesis (Blomme et al., 2022; LeBlanc et al., 2018). In *Citrus*, seven cycles of 37°C for 24 h with 24°C/24 h recoveries in between was optimal (LeBlanc et al., 2018). These studies suggest that when transforming whole plants, 37°C is the optimal temperature for inducing Cas9 mutations, which is in agreement with my study. However, a single cycle of 37°C/16 h was sufficient to improve editing efficiency in Marchantia. This short treatment time may be because sporelings were used, which probably cannot tolerate as high temperatures as whole plants.

Applying a 37°C HS during Marchantia transformation probably improves gene editing because Cas9 has optimal activity at this temperature, as a result of the gene having a bacterial origin. LeBlanc et al. (2018) and Kurokawa et al. (2021) demonstrated that the rate Cas9 cleaved a PCR amplified target site was higher at 37°C than at lower temperatures, under *in vitro* conditions. This high Cas9 activity may explain why in my study the 37°C HS treatments produced a majority of the mutants within the first two weeks of spore regeneration, whereas the lower Cas9 activity following the 28°C treatments resulted in fewer mutants being produced initially (Figure 3.8). Another possible explanation for why a 37°C HS improves gene editing is because DNA repair mechanisms may fail to work correctly at this temperature. Mutations result when edits made by Cas9 are not repaired, so if cellular repair mechanisms are not functional, mutations are more likely to occur (Oei et al., 2015).

Malzahn et al. (2019) showed that growing rice (*Oryza sativa*) protoplasts at 22°C and 28°C for 48 h had no significant effect on Cas9 mutagenesis. However, in my study, although the 28°C/16 h and 28°C/24 h HS treatments had minimal effects on Cas9 editing efficiency, the 28°C/48 h treatment substantially improved editing. These experiments should be repeated to confirm the initial findings, but it is possible that Marchantia spores may be more susceptible to heat than angiosperm vegetative explants. This poses the question of whether culturing Marchantia plants at 28°C for a prolonged period of time could improve Cas9 editing even further. Malzahn et al. (2019) additionally showed that growing *Arabidopsis* plants at 29°C for two weeks significantly improved the mutation frequency of AsCas12a, an enzyme that originates from *Acidominococcus* species that grow optimally at 37°C, similar to SpCas9. Furthermore, Milner et al. (2020) found that growing wheat

(*Triticum spp*.) callus cultures at 28°C for a 28 d selection period resulted in a significantly increased Cas9 mutation rate compared to growth at 25°C. These studies and the improved mutagenesis of the 28°C/48 h treatment in my study both suggest that growing transformed Marchantia spores at 28°C, at least during the two-week selection, could improve editing efficiency (assuming the increased length of exposure to 28°C would not kill the spores).

Many studies target the *Phytoene Desaturase* (*PDS*) gene when establishing or optimising CRISPR/Cas9 mutagenesis, because mutations result in a visually distinct phenotype that is easily scored. PDS functions in carotenoid biosynthesis, which are required for chloroplast photoprotection, thus, *pds* mutants have white (albino) leaves due to photo bleaching (Odipio et al., 2017; Watira et al., 2020; Wilson et al., 2019). Mp*PDS* (*Mapoly0108s0060*) was initially targeted in this study but, surprisingly, the phenotype observed was wild-type plants with white speckles (see Appendix 8 for example plants). This phenotype could result if mutations in the Mp*PDS* gene were lethal in Marchantia, such that only small sectors could survive if supported by surrounding wild-type tissue. However, in other plant species, *pds* mutants survive and grow if placed on a medium supplemented with sugar (Hooghvorst et al., 2019; Qin et al., 2007), which was also provided in the present study. Therefore, as no *pds* mutants have been reported or studied for a bryophyte before, the reason for the phenotypic difference with Marchantia is not known and would be an interesting area for further research. Notably, carotenoid cleavage-derived phytohormones (related to strigolactones) were recently found to regulate vegetative growth in Marchantia for adaptation to the environment (Komatsu et al., 2023), which may be related to why the *pds* phenotype in my study has abundant wild-type tissue.

Since the *pds* phenotype was not distinctive enough to use in this study, the *Golden 2-like* (Mp*GLK*) gene was targeted instead. The *glk* mutants had a pale green thallus due to reduced chloroplast production. Thus is consistent with the phenotype obtained by Yelina et al. (2023), the only other report of Marchantia *glk* mutants. Interestingly, several different strengths of mutant phenotype were observed in my study, ranging from almost white plants ('strong'), to a green colour barely paler than the wild type phenotype ('weak'). Yelina et al. (2023) found similar observations and correlated 'weak' phenotypes to having a single mutation in exon three and 'strong' phenotypes to having frameshifts and large deletions in downstream exons. This suggests that the alternative start site, suggested to be present by the Marchantia genome gene models, may be downstream of exon three and is therefore allowing a mostly functional GLK protein to be translated, resulting in the relatively high chlorophyll content in 'weak' phenotypes. In my study, sgRNAs were designed to target exons four and six (Figure 3.1), so it is possible that 'weak' mutants arose from mutations in

exon four, upstream of the alternative start site, and 'strong' mutants were generated from mutations in exon 6. A logical next step in my study would be to sequence *glk* mutants of different phenotypic strengths and measure their chlorophyll content, to determine if the size and position of mutations correlates to the severity of the mutant. Re-transforming using additional sgRNAs could also help distinguish the cause of the variable phenotypes.

In summary, the data suggest that *A*. *tumefaciens*-mediated Marchantia spore transformation is optimised by including 200 µM AS in the pre-culture media and Cas9 editing efficiency increased by providing a 37°C/16 h heat-shock. Implementation of these resulted in an efficient transformation system that produces high frequencies of substantially edited non-chimeric plants that are suitable for downstream analysis. This efficient protocol has been applied to characterising the function of numerous candidate Marchantia flavonoid-related genes, and the results of some of these studies are presented in Chapters 4 and 5.

Chapter 4

Functional characterisation of Mp*FNSI* in *M. polymorpha*

4.1 Introduction

Flavones and flavonols are two major classes of flavonoids that are commonly induced by abiotic stresses, particularly excess white or UV-B light (Agati & Tattini, 2010; Clayton et al., 2018; Wang et al., 2021; Winkel-Shirley, 2002). Providing similar functions across different plant groups, a particular plant species typically will only produce one of them: in angiosperms, flavonols are common while in bryophytes such as Marchantia, flavones are produced (Li et al., 2020a). Because flavonoids are accumulated in response to particular signals, their production is tightly controlled. In angiosperms, this is predominantly by R2R3-MYB transcription factors (Stracke et al., 2001). For example, AtMYB12, and orthologous genes from other species, can regulate the production of several flavonoids, including flavones and flavonols, by inducing the expression of key genes of the flavonoid biosynthetic pathway (Wang et al., 2016). However, no regulators of flavone biosynthesis have been identified in Marchantia so far, with the only Marchantia R2R3-MYBs that group with MYB12 in a phylogenetic analysis regulating auronidin or bibenzyl production (Albert et al., 2018; Kubo et al., 2018). Thus, flavones may be regulated independently of MYBs.

While flavone biosynthesis is poorly understood in Marchantia, it is widely known that angiosperms contain an enzyme called FLAVONE SYNTHASE (FNS) that performs this function. Two types of FNS enzyme have been identified in angiosperms: FNSI, part of the 2-OGD superfamily and FNSII, part of the CYP450 superfamily (Yonekura-Sakakibara et al., 2019). These enzymes are not genetically related, but functional characterisation has determined that they provide the same function in flavone biosynthesis. Thus, few plants have been identified to contain both enzymes (Jiang et al., 2016).

Marchantia and angiosperms are thought to have many flavonoid biosynthetic enzymes in common (Davies et al., 2020), and *in vitro* enzymatic assays demonstrate that FNSI may function in flavone biosynthesis in several liverworts (*M. polymorpha* was not examined) (Li et al., 2020a). A strong BLAST hit with 67% protein identity to the liverwort *Plagiochasma appendiculatum FNSI* sequence has been identified in Marchantia (*Mapoly0002s0224*) that belongs to the 2-OGD sub-group DOXC28 (unpublished data of Drs Nick Albert and Kevin Davies). The expression of the gene was also shown

to be upregulated in Marchantia plants exposed to UV-B light (Appendix 9.1), when flavone biosynthesis was induced (Clayton et al., 2018). Thus, it is a strong candidate for a *Flavone Synthase I* (Mp*FNSI*) gene and may encode a functional enzyme for flavone biosynthesis, but this requires verification.

The function of the candidate Mp*FNSI* gene was characterised by transiently expressing the gene in *Nicotiana tabacum* (tobacco) alongside At*Myb12* to induce flavonoid production, for a gain-of-function assay, and by using CRISPR/Cas9 and *A. tumefaciens*-mediated spore transformation (optimised in Chapter 3) to generate loss-of-function mutants. In both cases, plants were examined for their ability to produce flavones.

4.2 Materials and methods

4.2.1 Mp*FNSI* CRSIPR-Cas9 and over-expression constructs

MpFNSI CRISPR/Cas9 and over-expression constructs were designed and cloned into a binary vector as described in section 2.1 and 2.2 (pRMY07 and pRMY24; Appendix 3.2 and 3.3 for plasmid map, respectively). For the CRISPR/Cas9 construct, four high quality sgRNAs were designed to target exon one of *Mapoly0002s0224*, the Mp*FNSI* candidate gene (Figure 4.1).

Figure 4.1. Position of the four single guide RNA (sgRNA) molecules designed to target exon one of *Mapoly0002s0224,* a candidate gene for *Flavone Synthase I* (Mp*FNSI*) in *Marchantia polymorpha*.

4.2.2 *Venus-N7* over-expression construct

A construct was created that over-expressed the translational fusion of Venus and the N7 nuclear localisation domain (*35S*:*Venus-N7*). The *Venus-N7* coding sequence from an existing plasmid (pVenusDJL; problematic because of an impractical antibiotic selection gene on the plasmid) was first amplified by PCR using the iProof™ High-Fidelity DNA polymerase (BioRad), following the manufacturer's instructions. A high fidelity (proof-reading) polymerase was used to reduce the chance of introducing errors into the construct and to ensure the PCR product had blunt ends for

downstream cloning. PCR cycling conditions were: 98°C/20 s initial denaturation, 35 cycles of 98°C/5 s, 68°C/20 s, 72°C/30 s, a final extension of 72°C for 1 min and held at 12°C (primers were NA717 and NA718; Appendix 4). The *Venus-N7* PCR product was separated by gel electrophoresis (section 2.2.7) to confirm correct amplification, before being quantified using the Qubit (section 2.2.5.2). This provides a more accurate reading of PCR product concentration than the NanoDrop. The blunt-ended *Venus-N7* PCR product was directionally cloned into the Invitrogen pENTR™/D-TOPO® vector using TOPO® Cloning, following the manufacturer's instructions. The product was heat-shock transformed into *E. coli* (section 2.2.3) and cells were spread onto 50 mg L^{-1} kanamycin LB-agar plates (section 2.2.1). Cells were screened for the presence of the plasmid using colony PCR. Reaction conditions were the same as in section 2.4.5.2, except that the DNA template was provided by 1 µL of a cells/LB-broth solution and the annealing temperature was 55°C for 35 cycles instead of reducing the annealing temperature for five cycles. Primers used were NA585 and NA586 (Appendix 4). PCR products were separated by gel electrophoresis (section 2.2.7). Correct cloning was confirmed by using plasmid DNA extracted from the liquid cultures of positive colonies (section 2.2.4) as a template for DNA sequencing (2.2.11), using the same primers. A clone containing the correct plasmid was used as template for LR gateway recombination with the pNWA101 destination vector, as described for over-expression constructs in section 2.2, to create a binary vector containing the *35S*:*Venus-N7* construct that was then transformed into *A. tumefaciens* (pRMY21; Appendix 3.3 for plasmid map).

4.2.3 Plant material and growth conditions

Marchantia plants were exposed to excess white light conditions by growing them under a constant 120 µmol m^{-2} s^{-1} light intensity, provided by LED lights, for three weeks. Tobacco plants were grown from seed in ambient conditions, in a PC2 greenhouse that was heated at 15°C and vented at 25°C.

4.2.4 Stable *fnsI* mutant plants

Marchantia *fnsI* mutants were generated by *A. tumefaciens*-mediated spore transformation (section 2.4) using *A. tumefaciens* harbouring the Mp*FNSI* CRISPR/Cas9 construct. Regenerated sporelings were screened (section 2.4.5) for mutations in the Mp*FNSI* gene using primers NA713 and NA714 (Appendix 4). Confirmed mutant plants were retained and propagated through gemmae (G_1 generation) and re-screened to obtain pure mutant lines.

Selected *fnsI* plants, along with *chi* mutant (low or no flavones) and wild-type Sey-1 control plants, were grown under excess white light conditions (section 4.2.3) for three weeks to induce flavonoid

production, before being harvested for flavonoid analysis (section 4.2.6). Gemmae were plated onto MCE membrane filters on 0.5× B5 media. Four gemmae per tub and at least four biological replicates, consisting of separate tubs, were used per plant line. A randomised block design was used to position the tubs under the light source.

4.2.5 Mp*FNSI* transient expression in *N. tabacum* leaves

Tobacco plants were grown for seven weeks until the leaves were ~20 cm in diameter, when they were infiltrated with *A. tumefaciens* cells harbouring gene constructs of interest – *35S*:*Venus-N7*, *35S*:*FNSI* and *35S*:At*Myb12*. *A. tumefaciens* cells were grown up from glycerol stocks and scraped into a 10 mM $MgCl_2$, 10 mM MES (pH 5.6) and 200 µM acetosyringone solution and left at room temperature for 2−4 h to induce virulence. The following combinations of these *A. tumefaciens* preparations were mixed together to infiltrate the leaves: *35S*:*Venus-N7* only; *35S*:*Venus-N7/35S*:At*Myb12*; and *35S*:*Venus-N7/35S*:At*Myb12/35S*:*FNSI*. *Venus-N7* was used as an infiltration control to detect transformed cells because Venus is a YELLOW FLUORESCENT PROTEIN (YFP) that emits easily detectable fluorescence. Additionally, it was used to make up an equal total volume of each infiltration mix, so that the same volume of *35S*:*FNSI* and *35S*:At*Myb12* *A. tumefaciens* could be added to each combination.

The *A. tumefaciens* preparations were infiltrated into the tobacco leaves using 1 mL syringes pressed against the underside of the leaves, where the solution was pushed through the stomata. A 3−4 mL aliquot of solution was used to infiltrate half of a leaf, and three leaves/biological replicates were infiltrated for each *A. tumefaciens* mix. Leaves were left for five d post-infiltration before being harvested for flavonoid analysis (section 4.2.6).

4.2.6 Flavonoid analysis

Flavonoid analysis was performed on Marchantia stable *fnsI* mutants and tobacco *35S*:*FNSI* transient expression plants. Plant material was firstly frozen in liquid N_2, stored at −80°C, then finely ground using a mortar and pestle. Ground tobacco tissue was weighed into ~200 mg fresh weight (FW) samples, while Marchantia tissue was freeze-dried then ~11 mg dry weight (DW) was weighed out. Total flavonoids were extracted in 1 mL methanol:water:formic acid (80:19:1; v:v:v) then left at 4°C overnight, inverting regularly, before clearing the extract by centrifugation and using the supernatant for chemical analysis.

4.2.6.1 High Performance Liquid Chromatography

The *fnsI* mutants and tobacco samples were analysed by high performance liquid chromatography (HPLC), performed by Sarah Cordiner (Plant and Food Research, New Zealand), to separate and quantify metabolites. The HPLC set up, solvent composition and analysis was as described in Albert et al. (2018), with the only variation being that the binary pump used in the LC system was HPG-3400RS instead of HPR-3400RS.

Flavones and auronidins were quantified by integrating areas under chromatograms peaks with peak maxima at 340 nm and 484 nm, respectively, and comparison against calibration curves generated from luteolin-7-*O*-glucoside and pelargonidin-3-*O*-glucoside standards (Extrasynthese, Genay, France). Pelargonidin-3-*O*-glucoside is not an auronidin compound (it is an anthocyanin) but has a lambda max of 499 nm that is close to that of auronidin, which ranges between 484 nm and 493 nm. It is also stable and easily sourced, making it a useful equivalent standard. By contrast, auronidin standards are not commercially available, and those prepared by chemical synthesis proved to be somewhat insoluble at higher concentrations, making them an unreliable quantification standard. Final concentrations accounted for the extraction solvent volume and the weight of the ground tissue used for extraction. Auronidin compounds were identified by UV-VIS absorption spectra, information from Berland et al. (2019) and comparison of elution times to Albert et al. (2018). Flavones were identified by UV-VIS absorption spectra, previous mass-spectrometry analysis data (unpublished data of Dr Tony McGhie, Plant and Food Research, New Zealand) and comparison of elution times to Markham et al. (1998) and Albert et al. (2018).

4.3 Results

4.3.1 Stable *fnsI* mutants

The Marchantia Mp*FNSI* candidate gene was characterised by using CRISPR/Cas9 and *A. tumefaciens*-mediated spore transformation to generate *fnsI* mutants. The first 10 lines (G_1 stage) were screened for Mp*FNSI* mutations by PCR-amplification of a targeted region (700 bp), followed by gel electrophoresis and DNA sequencing. Figure 4.2 shows the PCR products of several plants transformed with the CRISPR/Cas9 construct. Some of these had size differences to the wild-type plant product, suggesting that deletions had occurred. The wild-type plant had an amplicon band of 700 bp, whereas *fnsI*/5 had a visible size difference of ~200 bp, *fnsI*/6 a difference of ~100 bp and *fnsI*/1 a small, but noticeable, size difference. All other *fnsI* lines shown did not appear to have size differences, but separation by gel electrophoresis does not distinguish small base pair differences (at

least when using a 1.5% agarose gel). The negative PCR control (a no-DNA sample) had a few bands on the gel, which could suggest DNA contamination. However, the bands were considered more likely to be due to non-specific primer binding, as they were faint and did not correspond to the size of the *fnsI* or wild-type amplicons. The PCR products of all 10 transformed plant lines were sequenced regardless of whether size differences were observed because even 1 bp mutations can be sufficient for loss of gene function.

Sequencing results revealed that mutations were present in 7 of the 10 CRISPR/Cas9 plants screened, and the majority of these were deletions (Table 4.1). Thus, *fnsI* mutants were successfully generated. The mutations present in the three *fnsI* mutants selected for further study are summarised in Figure 4.3. In comparison to the wild-type sequence, *fnsI*/5 contains a single 246 bp deletion spanning from sgRNA #94 to sgRNA #109 (see Figure 4.1 for sgRNA positions in the Mp*FNSI* coding sequence); *fnsI*/6 contains a 121 bp deletion spanning from sgRNA #94 to sgRNA #5 and another 1 bp deletion at sgRNA #109, and *fnsI*/8 contains two 1 bp deletions at sgRNAs #5 and #109, respectively. These 1 bp deletions were not detected following gel electrophoresis, which emphasises the need for DNA sequencing.

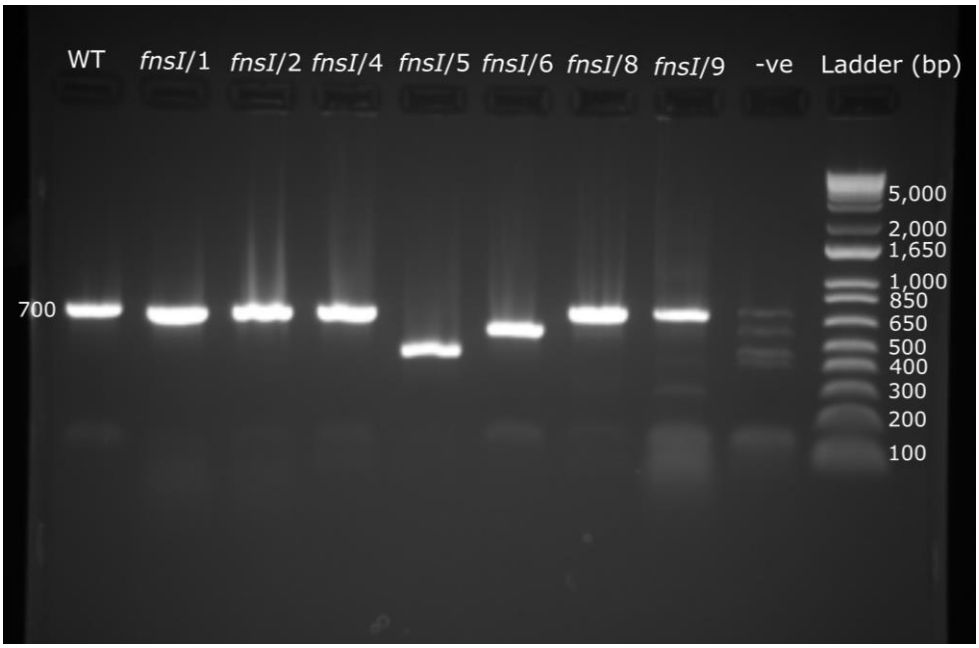


Figure 4.2. Screening *Marchantia polymorpha* plants generated using CRISPR/Cas9 for size differences in the *Flavone Synthase I* (*FNSI*) gene. PCR-amplification of a Mp*FNSI* target region (700 bp) in Mp*FNSI* CRISPR/Cas9 plants compared to a wild-type (WT) plant, showing size differences. A negative PCR control (-ve) was included.

Table 4.1. Summary of identified *Marchantia polymorpha fnsI* mutants generated by CRISPR/Cas9. DNA sequencing results of the PCR amplicons of all 10 Mp*FNSI* CRISPR/Cas9 plants, showing the types of mutations present and the single guide RNAs (sgRNAs) where they are located.

Plant line ID	Mutant?	Mutation description	sgRNA #
fnsI/1	Y	1 bp insertion, 9 bp deletion	109, 11
fnsI/2	Y	3 bp deletion, 1 bp deletion	94, 109
fnsI/3	N	-	-
fnsI/4	Y	3 bp deletion, 1 bp deletion, 1 bp deletion	94, 5, 109
fnsI/5	Y	246 bp deletion	94 and 109
fnsI/6	Y	121 bp deletion, 1 bp deletion	94 and 5, 109
fnsI/7	N	-	-
fnsI/8	Y	1 bp deletion, 1 bp deletion	5, 109
fnsI/9	Y	1 bp deletion	109
fnsI/10	N	-	-

Figure 4.3. Examples of mutations in the *Flavone Synthase I* (*FNSI*) gene of *Marchantia polymorpha* plants generated using CRISPR/Cas9. DNA sequencing results of three *fnsI* mutant lines (*fnsI*/5, *fnsI*/6 and *fnsI*/8) are shown compared to a wild-type sequence. Two single guide RNAs (sgRNAs; #5 and #109) are shown in green text, followed by their adjacent PAM (protospacer adjacent motif) site in blue text. Mutations (deletions) are shown in red text as dashes.

4.3.2 Flavonoid analysis of *fnsI* mutant plants

Flavone content was initially determined in two confirmed *fnsI* mutants (*fnsI*/5 and *fnsI*/6), along with the *chi* mutant and wild-type Sey-1 plants, all grown in standard culturing conditions. The wild-type produced two main flavone compounds: peak 1 was identified as luteolin-7,3'-di-*O*-glucuronide (lambda max 267, 339 nm) and peak 2 as luteolin-7-*O*-glucuronide (lambda max 251, 347 nm) (Figure 4.4). Flavones were not detected in the *fnsI* mutants, as evidenced by there not being any peaks with the same retention time as the flavones detected in the wild-type plants. Comparatively, the *chi* mutant produced trace amounts of luteolin-7,3'-di-*O*-glucuronide (peak 1).

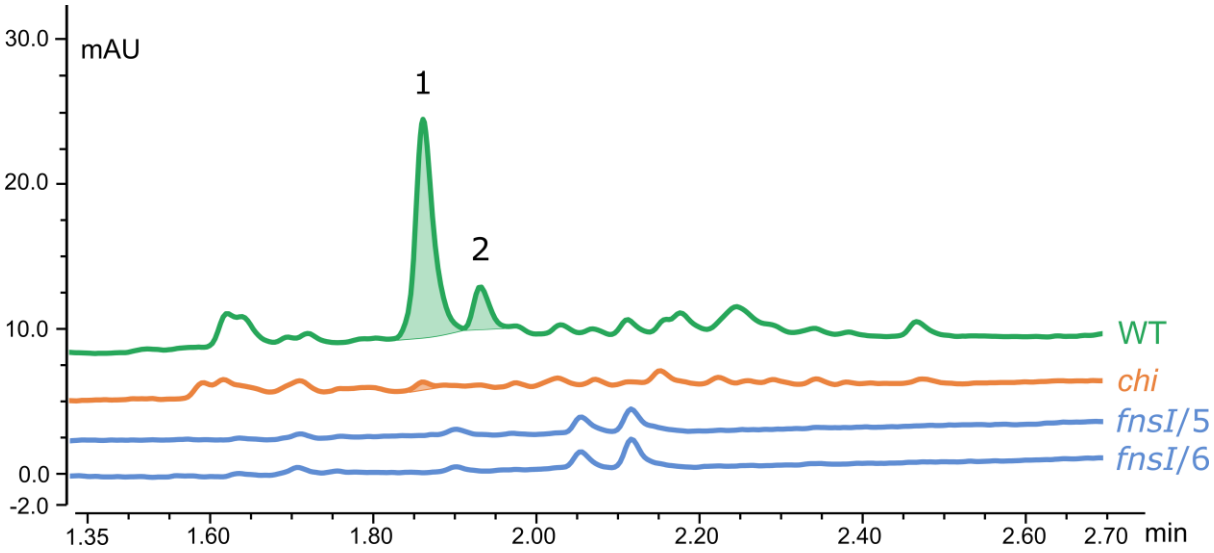


Figure 4.4. Flavonoid analysis of *Marchantia polymorpha fnsI* mutants. High performance liquid chromatography (HPLC) chromatogram of flavones in *fnsI* mutants (*fnsI*/5 and *fnsI*/6; blue), a *chi* mutant (orange) and a wild-type (WT) plant (green). Flavones were monitored at an absorbance of 340 nm. Peaks were identified based on spectral properties and comparison to Markham et al. (1998). Peaks are (1) luteolin-7,3'-di-*O*-glucuronide and (2) luteolin-7-*O*-glucuronide. mAU: absorbance units in nm; min: retention time.

Following the preliminary results, *fnsI* mutants (*fnsI*/5, *fnsI*/6 and *fnsI*/8), *chi* mutant and wild-type plants were grown under excess white light conditions to induce higher flavonoid production. The phenotypes of these plants are shown in Figure 4.5. HPLC analysis shows that the wild-type Sey-1 plants produced six main flavone compounds under excess white light conditions: luteolin-7,4'-di-*O*-glucuronide (1; lambda max 268, 337 nm), luteolin-7,3'-di-*O*-glucuronide (2; lambda max 267, 339 nm), luteolin-7-*O*-glucuronide (3; lambda max 265, 346 nm), a mixture of luteolin-3'-*O*-glucuronide and apigenin-mono-*O*-glucuronide (4; lambda max 266, 331 nm), apigenin-7-*O*-glucuronide (5; lambda max 266, 336 nm) and luteolin-4'-*O*-glucuronide (6; lambda max 267, 339 nm) (Figure 4.6A). The retention time of the flavone peaks cannot be compared to those in Figure 4.4 as the analysis was performed separately. These compounds were not detected in any of the *fnsI* mutants or the *chi* mutant. However, additional metabolites (likely to be phenolic compounds) with relatively similar retention times were detected, but their spectral properties are not consistent with flavones (an example flavone spectrum is shown on the figure). In *chi*, these had a lambda max of 397 nm (A) and in *fnsI* a lambda max of 283 nm (B) (spectra in Appendix 10.1). The wild-type Sey-1, *fnsI* and *chi* plants all produced the auronidin aglycone peaks (1 and 2; Figure 4.6B). Because all of the *fnsI* mutant lines displayed the same results, only one line was included in each of the chromatograms in Figure 4.6.

Quantification of the excess white light HPLC results showed the wild-type Sey-1 plants contained an average flavone content of ~4 mg g^{-1} DW, while flavones were not detected in the *fnsI* and *chi* mutants (Figure 4.7A). Auronidin content in the wild-type Sey-1 and *fnsI*/5 plants was ~1 mg g^{-1} DW, the *fnsI*/8 and *chi* plants had ~0.7 mg g^{-1} DW and ~0.8 mg g^{-1} DW, respectively, and *fnsI*/6 had a lower content of 0.5 mg g^{-1} DW (Figure 4.7B). However, the differences in auronidin content were not statistically significant.

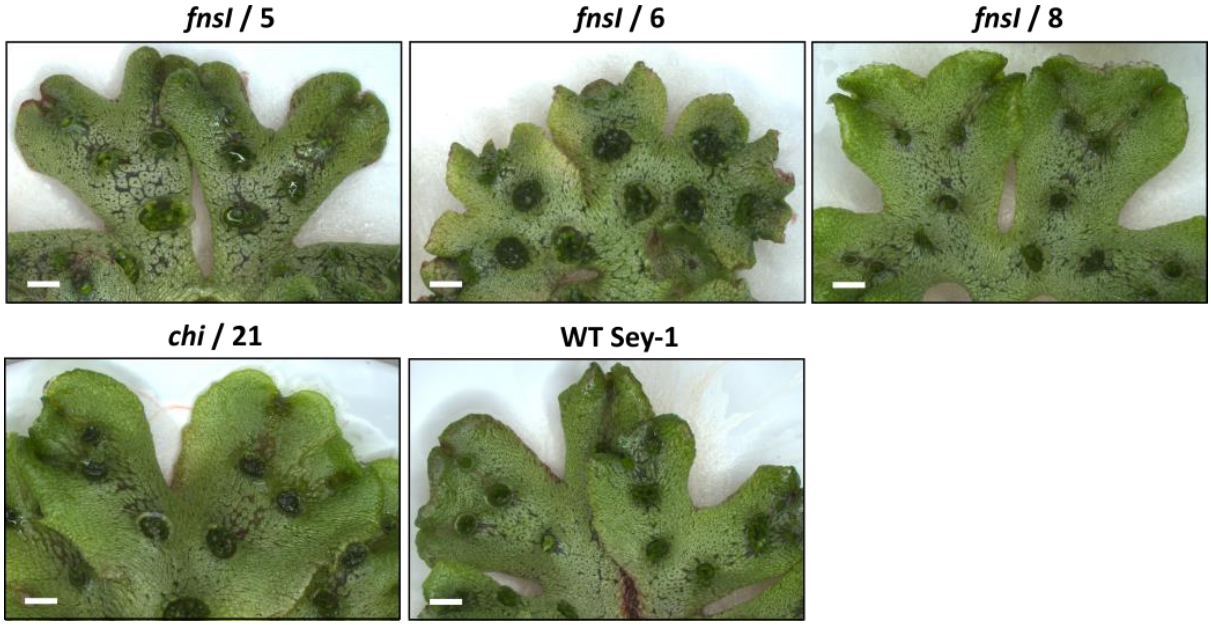


Figure 4.5. Phenotype of *fnsI*, *chi* and wild-type (WT) Sey-1 *Marchantia polymorpha* plants grown under excess white light conditions for three weeks. Scale bars represent 2 mm.

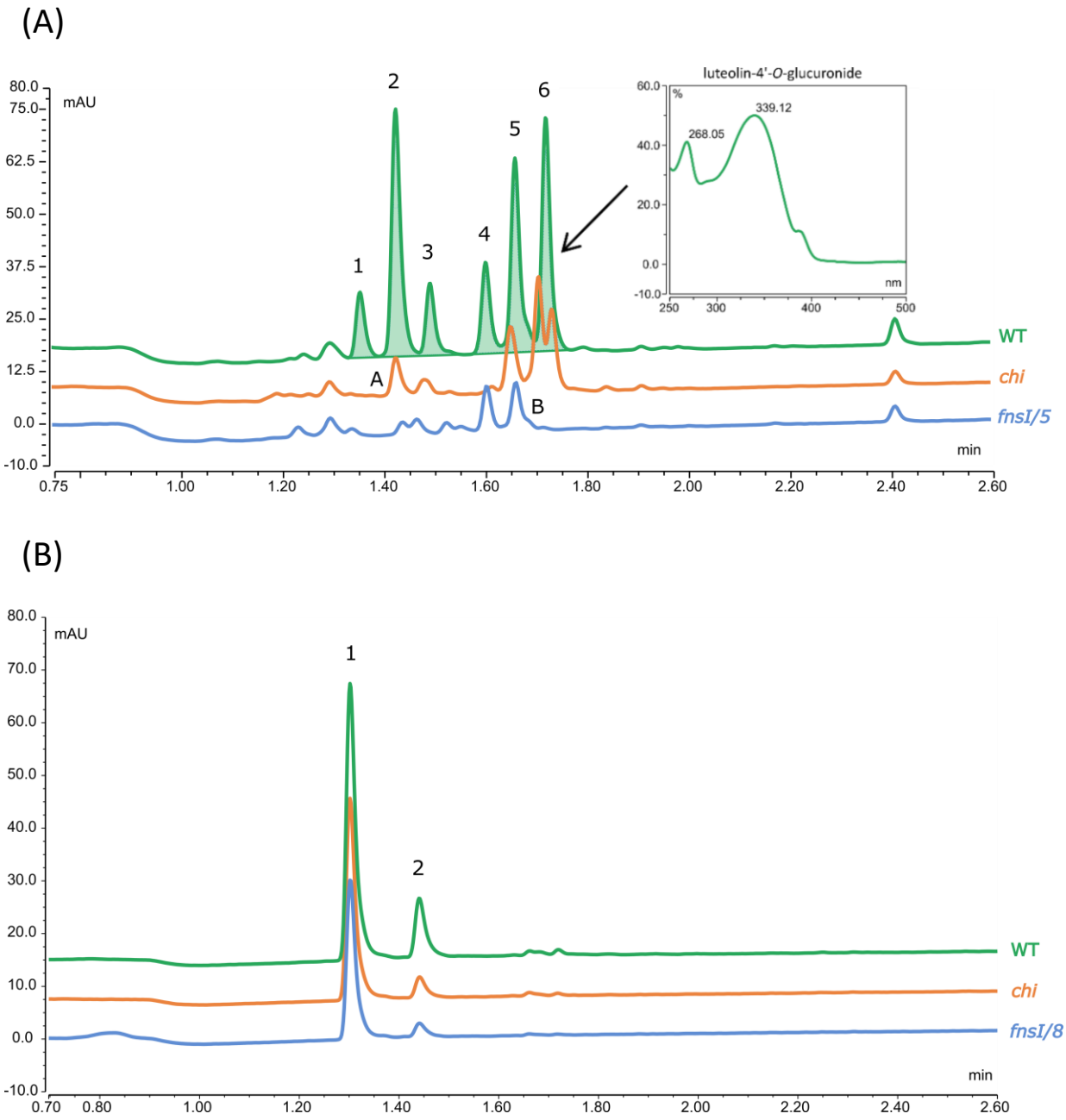


Figure 4.6. Flavonoid analysis of *Marchantia polymorpha fnsI* mutants grown under excess white light conditions. HPLC chromatograms of a *fnsI* mutant (blue), a *chi* mutant (orange) and a wild-type (WT) plant (green). mAU: absorbance units in nm; min: retention time. (A) Detection of flavone compounds at 340 nm. Peaks were identified by a previous mass-spectrometry analysis (unpublished data of Dr Tony McGhie, Plant and Food Research, New Zealand) and comparison to Markham et al. (1998) and Albert et al. (2018). (1) luteolin-7,4'-di-*O*-glucuronide; (2) luteolin-7,3'-di-*O*-glucuronide; (3) luteolin-7-*O*-glucuronide; (4) a mixture of luteolin-3'-*O*-glucuronide and apigenin-mono-*O*-glucuronide; (5) apigenin-7-*O*-glucuronide and (6) luteolin-4'-*O*-glucuronide. Production of additional metabolites in *chi* (A) and *fnsI* (B) mutants are also indicated. (B) Detection of auronidin compounds at 484 nm. Peaks were identified by comparison to Albert et al. (2018) and Berland et al. (2019). Peaks (1) and (2) are different conformations of the auronidin aglycone.

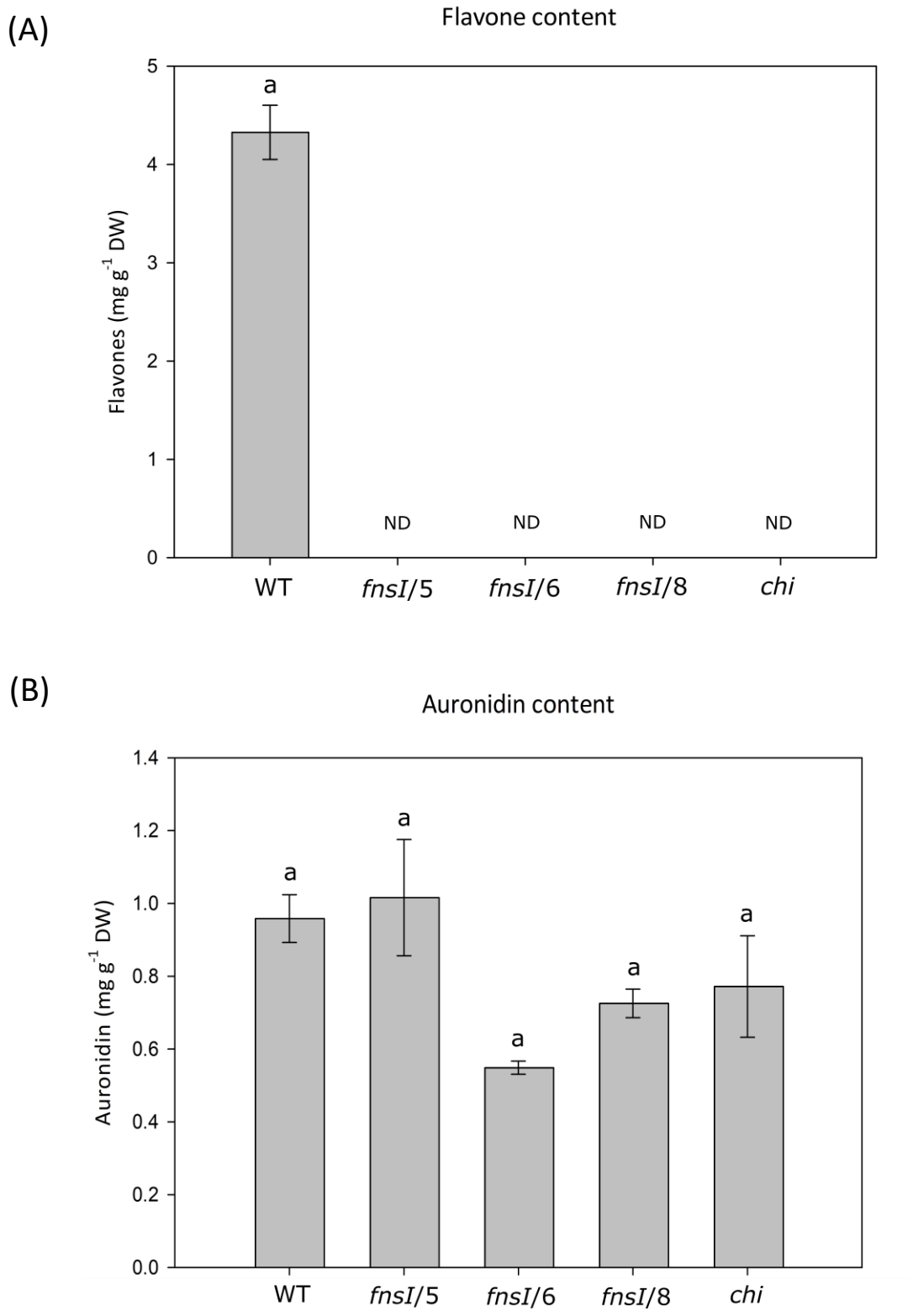


Figure 4.7. Flavonoid content in *Marchantia polymorpha fnsI* mutants grown under excess white light conditions. High performance liquid chromatography (HPLC) quantification of flavones (A) and auronidin (B) in wild-type (WT) plants, *fnsI* mutants and a *chi* mutant. Error bars are the standard error of the means, n = 3. ND: not detected. Significantly different means are represented by different letters.

4.3.3 Transient Mp*FNSI* expression in *N. tabacum*

Tobacco produces flavonols as its primary flavonoid class in leaves and lacks flavones. Thus, *A. tumefaciens*-mediated infiltration of tobacco leaves and subsequent flavonoid analysis was used to determine whether transiently expressing Mp*FNSI* could confer flavone production.

Before infiltration was performed, stable Marchantia *35S*:*Venus-N7* plants were generated to confirm that fluorescence could be easily detected (Appendix 11). Infiltration was then tested by transiently expressing the *35S*:*Venus-N7* construct into tobacco leaves. An example of an infiltrated leaf is shown in Figure 4.8A and the fluorescence emission from *35S*:*Venus-N7* in Figure 4.8B. Strong fluorescence was observed from the infiltrated (right) side of the leaf , but not the un-infiltrated (left) side. Therefore, infiltration was successfully established.

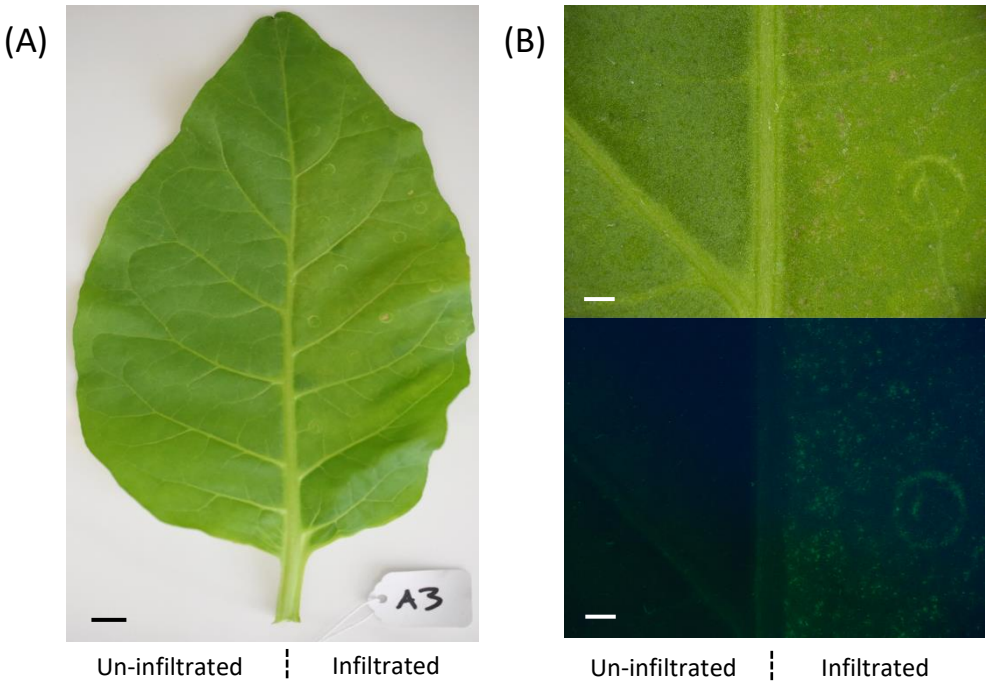


Figure 4.8. Transient expression of *35S*:*Venus-N7* in *Nicotiana tabacum* (tobacco). (A) Example of an infiltrated tobacco leaf, showing no infiltration on the left side of the leaf and infiltration on the right. Scale bar represents 1 cm. (B) 20× magnification of the same leaf with no filter (top panel) and a YELLOW FLUORESCENCE PROTEIN (YFP) microscope filter (bottom panel), showing fluorescence on the infiltrated (right) side of the leaf and no fluorescence on the un-infiltrated (left) side. Scale bars represent 200 µm.

Tobacco leaves were then infiltrated with *35S*:At*Myb12* and *35S*:*FNSI* constructs. At*Myb12* was infiltrated to upregulate the flavonoid biosynthetic pathway, making it easier to detect any flavones. Transient expression of At*Myb12* alone resulted in the production of two peaks (1 and 2; Figure 4.9). Two additional peaks (3 and 4; Figure 4.9) were identified in extracts from the *35S*:At*Myb12/35S*:*FNSI* sample. Only one HPLC trace, which best represents the described results, is shown for each expression line. Based on spectral properties, peaks 1 and 2 matched closely to the flavonol standard and peaks 3 and 4 matched closely to the flavone standard (Figure 4.10).

Subsequent analysis by LC-HRAM-MS/MS, performed by Dr John van Klink (University of Otago, New Zealand), confirmed the peak identities as flavonols and flavones, respectively (Appendix 10.2). Flavonol peak 1 was tentatively identified as quercetin-3-*O*-rutinoside (*m/z* 609.14) and peak 2 as kaempferol-3-*O*-rutinoside (*m/z* 593.15, lambda max 265, 348 nm). Peaks 3 and 4 were identified as flavones; peak 3 as apigenin-7-*O*-glucoside (*m/z* 431.0987, lambda max 267, 335 nm) and peak 4 (*m/z* 461) closely matched chrysoeriol-7-*O*-glucoside and luteolin-7-*O*-glucuronide but could not be conclusively identified.

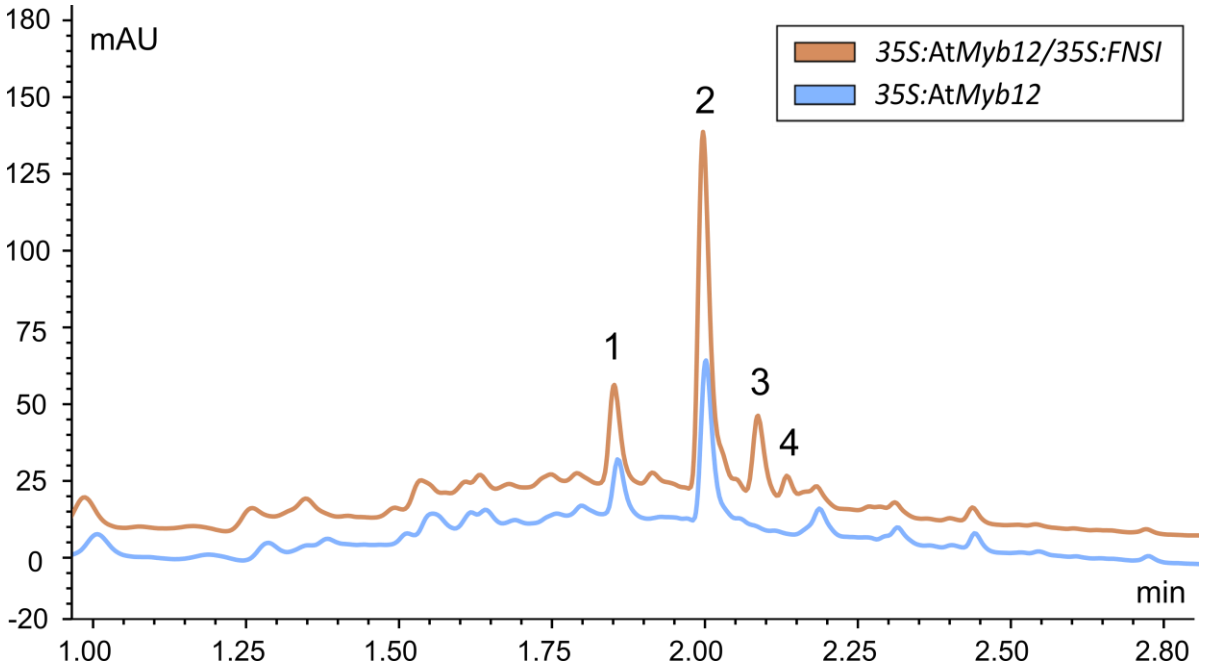


Figure 4.9. Detection of flavonoids in *Nicotiana tabacum* (tobacco) plants expressing *Flavone Synthase I* (Mp*FNSI*) and At*Myb12*. High Performance Liquid Chromatography (HPLC) chromatogram of flavonol and flavone compounds detected at 340 nm in representative *35S*:*FNSI/35S*:At*Myb12* (orange) and *35S*:At*Myb12* control (blue) tobacco plants. Based on spectral properties, peaks 1 and 2 are flavonols and peaks 3 and 4 are flavones. mAU: absorbance units in nm; min: retention time.

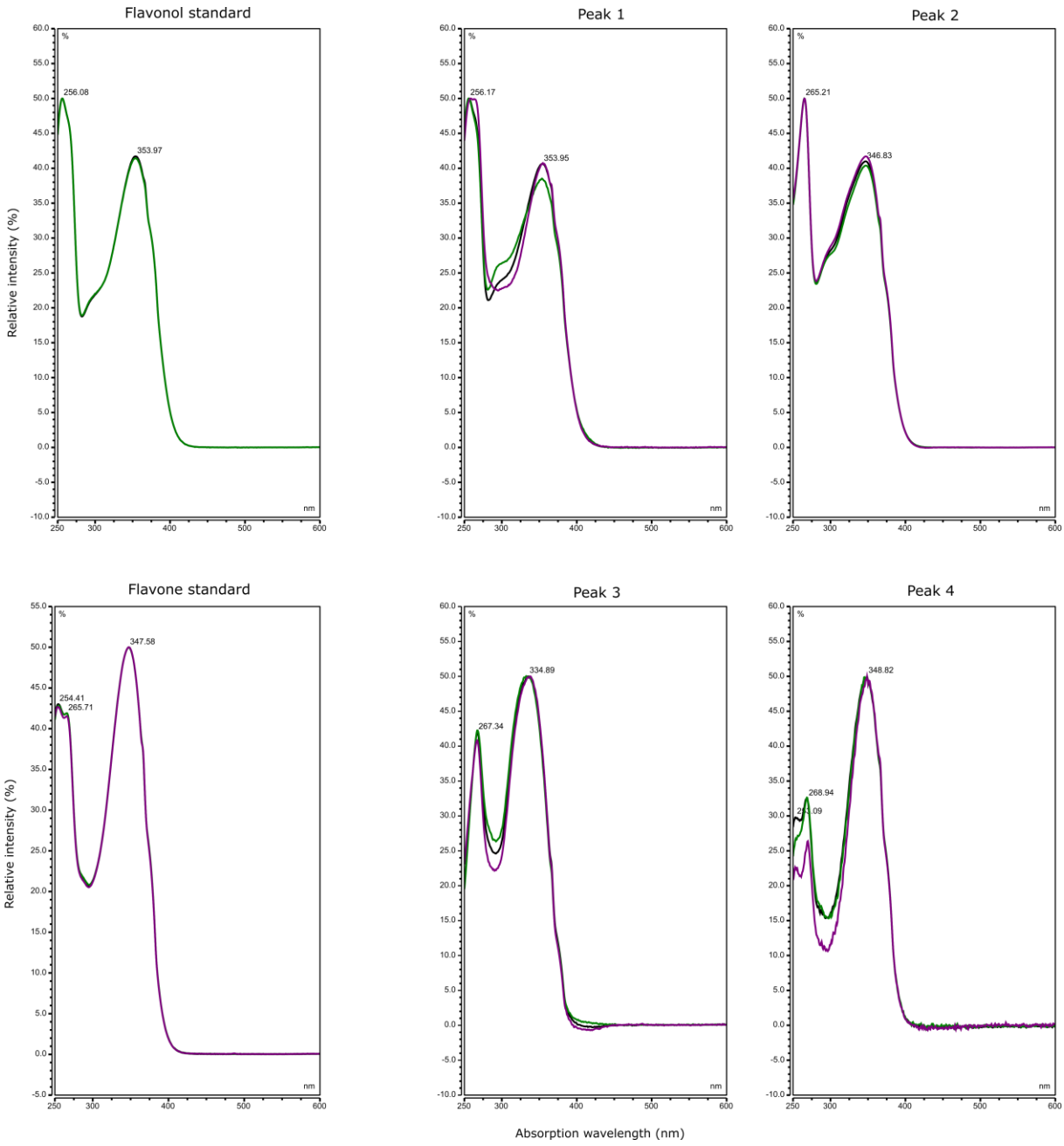


Figure 4.10. Spectral properties of flavonol and flavone peaks identified in *Nicotiana tabacum* (tobacco) plants expressing *Flavone Synthase I* (Mp*FNSI*) and At*Myb12* compared to the standards.
Peaks 1 - 4 refer to the flavonol and flavone compounds detected at 340 nm in Figure 4.9.
nm: absorption wavelength; "%": relative intensity.

4.4 Discussion

The function of the Mp*FNSI* candidate gene in Marchantia was characterised by using CRISPR/Cas9 and *A*. *tumefaciens*-mediated spore transformation to generate *fnsI* mutants. Once plants containing the Mp*FNSI* CRISPR/Cas9 construct were generated, they were screened for Mp*FNSI* mutations. Both PCR-amplification (Figure 4.2) and DNA sequencing (Table 4.1 and Figure 4.3) results showed that *fnsI* mutants were successfully created. All mutants shown produced deletions; ranging from 1 to 246 bp. Single base pair deletions originate from incorrect DNA repair following Cas9 editing at a single sgRNA (Jiang et al., 2013). These are highly effective for generating non-functional alleles because they cause a frame-shift, changing the coding of all downstream amino acids, creating a completely different protein and potentially introducing a premature stop codon that truncates the protein (Harfe & Jinks-Robertson, 1999). Large deletions arise when in a multiplex system more than one sgRNA is edited by Cas9 at a similar time and the DNA within the two double-stranded breaks 'drops out' during repair. This can cause many crucial amino acids to be lost or induce a frame-shift, usually forming a non-functional protein (Cong et al., 2013; Pagel et al., 2019). In this study, *fnsI*/5 contained a 246 bp deletion due to edits being made in sgRNAs #94 and #109 which are 246 bp apart, and *fnsI*/6 contained a 121 bp deletion resulting from editing events at sgRNAs #94 and #5 which are 121 bp apart. Interestingly, *fnsI*/6 also contained a 1 bp deletion at sgRNA #109. Because this edit did not result in a larger deletion, it is likely that the editing event did not occur at the same time as those at the other two sgRNAs.

HPLC analysis showed that *fnsI* mutants do not produce flavones. This was initially observed when mutants were grown in standard culture conditions (Figure 4.4). Wild-type plants produced two main flavone compounds (luteolin-7,3'-di-*O*-glucuronide and luteolin-7-*O*-glucuronide), but these were not detected in the *fnsI* mutants. Therefore, *Mapoly0002s0224/MpFNSI* encodes a functional FNSI enzyme that is required for flavone biosynthesis in Marchantia, at least in the tissues and growth conditions tested. This result indicates only a single *FNS* gene is present in the species, matching the previous phylogenetic analysis that found *Mapoly0002s0224* as the only 2-OGD gene that clustered in the DOXC28 sub-group. Surprisingly, the *chi* mutant produced trace amounts of luteolin-7,3'-di-*O*-glucuronide, but this was only detected in the absence of excess white light stress (Figure 4.4). CHI functions in converting naringenin chalcone to flavanones, the substrate of flavones, and previous studies have reported *chi* mutants having no detectable flavones in Marchantia (Berland et al., 2019; Clayton et al., 2018). Thus, detecting trace amounts of flavones in this study poses an interesting theory: that the reaction catalysed by CHI could partially occur

spontaneously. This idea is supported by small amounts of flavonoids being accumulated in *chi* mutants of *Arabidopsis* (*tt5*) and rice (*Oryza sativa*; *gh1*) (Hong et al., 2012; Jiang et al., 2015).

Excess white light conditions resulted in the accumulation of six flavones in wild-type plants (Figure 4.6A). Quantification showed the total flavone content was ~4 mg g^{-1} DW (Figure 4.7A), which is higher than previous reports of ~2 mg g^{-1} DW flavones in Marchantia plants exposed to 200 µmol m^{-2} s^{-1} high intensity white light (Albert et al., 2018). The abundance of flavones in this study may be a result of the 24-h light exposure, which provides additional light stress compared to a typical 16-h exposure photoperiod. Flavones remained undetected in the *fnsI* mutants under excess white light stress (Figure 4.6 A and Figure 4.7A), which reinforces the function of MpFNSI in flavone biosynthesis. Conversely, the trace amounts of flavones detected in the *chi* mutant under standard culturing conditions (Figure 4.4) were not detected in plants grown under excess white light conditions. This result aligns with previous studies (Berland et al., 2019; Clayton et al., 2018) and suggests that the small flavone peak in *chi* may be masked by the increase in flavones and other phenolic compounds under these conditions. Mass-spectrometry may be useful to determine if these flavones in the *chi* mutant can be detected.

A lack of flavone accumulation may result in the production of other phenolic compounds as a result of accumulation of substrate in the pathway and subsequent redistribution into other pathway branches. This was evident in the present study, as the *fnsI* and *chi* mutants produced additional stress-induced phenolic compounds (currently unidentified) that were not detected in wild-type plants (Figure 4.6A). The fact that these phenolic compounds were different in *fnsI* and *chi* mutants (lambda max of 283 and 397 nm, respectively) is an interesting finding that could be explored further. However, given the specific function of flavones in UV-B protection, *fnsI* and *chi* mutants would still be expected to be sensitive to UV-B light. Clayton et al. (2018) investigated this hypothesis by exposing *chi* mutants, generated through CRISPR/Cas9, to high- and low-fluence UV-B light. Exposure to both UV-B treatments resulted in the *chi* plants showing extensive signs of stress that were not evident in wild-type plants, including reduced growth, epidermal browning and silver 'glazing', and an increased ROS signal indicative of reduced ROS scavenging and/or higher generation of ROS. I attempted to replicate the conditions of Clayton et al. (2018), which were performed at Lincoln University (New Zealand), in a similar experiment investigating whether *fnsI* mutants were sensitive to UV-B light. However, the stress phenotype was much more intense than that reported in Clayton et al. (2018), with the majority of the tissue being dead (Appendix 12). Because of the severity of this phenotype, differences between the wild-type and *fnsI* mutants could not be

meaningfully interpreted. Significant attempts were made to replicate the conditions of Clayton et al. (2018), but the issues could not be resolved within the timeframe of this thesis.

Exposure of *fnsI* mutants to excess white light stress also induced auronidin production. This was mainly visible on the ventral (bottom) surface of the plants prior to them being harvested for HPLC (data not shown), but two auronidin peaks (different conformations of the same compound) were identified in the wild-type, *fnsI* and *chi* plants (Figure 4.6B). Quantification revealed slight differences in auronidin production between plant lines (Figure 4.7B). However, these were not statistically significant, and may reflect the varied genetic backgrounds of lines produced by spore transformation. This may also explain the slight differences in growth morphology observed (Figure 4.5).

Tobacco leaves were infiltrated with the Mp*FNSI* candidate gene to examine whether flavone production was conferred. When leaves were infiltrated with just *35S*:At*Myb12*, only flavonols were produced. Whereas, the co-expression of *35S*:At*Myb12* and *35S*:*FNSI* resulted in the accumulation of two flavone peaks, showing Mp*FNSI* was able to introduce flavone production into plants that do not naturally produce them. Similar results were observed when *FNSI* from mulberry (*Morus notabilis*), *Chrysanthemum morifolium, Lonicera japonica* or *Lonicera macranthoides* were transiently expressed in tobacco (Li et al., 2020b; Wang et al., 2021; Wu et al., 2016). This further established that *Mapoly0002s0224* is the Mp*FNSI* gene for flavone biosynthesis in Marchantia.

Identifying and determining the function of the Mp*FNSI* gene in Marchantia is a significant finding both with regard to understanding Marchantia biology and the evolution of flavone biosynthesis in land plants. FNS is also required for flavone biosynthesis in many angiosperm plants, which are significantly genetically diverged from Marchantia, so the function of FNS may be conserved among land plants. However, angiosperms have evolved to use either FNSI (2-OGD superfamily) or FNSII (CYP450 superfamily) enzymes for flavone biosynthesis (Yonekura-Sakakibara et al., 2019), whereas FNSII has not been identified in bryophytes. There has been some controversy over the evolution of FNS enzymes because it is hard to deduce enzyme function purely based on sequence; functional studies are required. Currently, the most logical explanation is that an ancestral gene in early land plants, that may or may not have been required for flavone biosynthesis, evolved into FNSI in bryophytes and F3H in the lineage that eventually gave rise to angiosperms (Li et al., 2020a). This is supported by Li et al. (2020a) and Wang et al. (2020) identifying that some liverworts (such as the *Marchantia emarginata* and *Marchantia paleacea*) have FNSI enzymes, other liverworts have enzymes with both FNSI and F2H activity (*Plagiochasma appendiculatum* and *Conocephalum conicum* and *Conocephalum japonicum*), and the moss *P. patens* and lycophyte *S. moellendorffii*

contain enzymes with FNSI and F3H activity. While this is based on *in vitro* enzymatic assays, functionally characterising MpFNSI in *M. polymorpha in planta* in my study demonstrates the presence of FNSI enzymes in bryophytes. Additionally, the Apiaceae family (and a few other related angiosperms) have FNSI enzymes, but these are not similar in sequence to those in bryophytes and phylogenetic analysis indicates that they evolved from F3H through tandem gene duplication and subsequent neofunctionalization. Thus, current evidence suggests that FNSI has evolved multiple times (Gebhardt et al., 2005; Li et al., 2020a; Martens et al., 2003; Pucker & Iorizzo, 2023). Comparatively, FNSII is phylogenetically much more widespread than FNSI. Phylogenetic analyses suggest that FNSII evolved from the ancestral CYP93 gene following the divergence of monocots and eudicots, as FNSII identified in these plants cluster within different CYP450 sub-groups: CYP93G (monocots) and CYP93B (eudicots) (Du et al., 2016; Jiang et al., 2016; Yonekura-Sakakibara et al., 2019).

The proposed evolution of FNS enzymes is further complicated by the evolution of F3H, allowing for the production of flavonols and anthocyanins in angiosperms (Li et al., 2020a). Given FNSII is theorised to have evolved after F3H, flavones may have only been produced in angiosperms following the evolution of flavonols and anthocyanins. This idea is supported by some (probably basal) angiosperms not producing flavones, instead accumulating flavonols that perform similar functions in protection against UV-B light (Agati & Tattini, 2010; Li et al., 2020a; Winkel-Shirley, 2002). Furthermore, the evolution of FNSI in Apiaceae (a eudicot) poses the question of whether these plants gained FNSI for a different purpose, or whether FNSI was lost in other eudicots once similar functions could be performed by FNSII. Nevertheless, the fact that most angiosperms have either FNSI or FNSII suggests that the UV-B protection function of flavones must have remained important in angiosperms (Martens & Mithöfer, 2005).

The Mp*FNSI* candidate gene was functionally characterised in *M. polymorpha* by generation of *fnsI* mutants and transient expression of Mp*FNSI* in tobacco. The results showed that the gene encodes a MpFNSI enzyme necessary for flavone biosynthesis in this liverwort. The similarities to known or candidate FNS enzymes in other bryophytes and angiosperms provides evidence that the origin and function of FNSI is conserved among land plants. This implies that the last common ancestor of these lineages, which may have existed relatively soon after plants first colonised land, may have also contained a similar FNSI gene. Therefore, the UV-B protective function of flavones may have been an important characteristic that these early land plants gained during land colonisation.

Chapter 5

Functional characterisation of Mp*BT2* in *M. polymorpha*

5.1 Introduction

Auronidin is a red/orange-coloured flavonoid unique to liverworts that provides tolerance against several abiotic stresses, such as excess light and nutrient deprivation (Albert et al., 2018; Berland et al., 2019). Auronidin is predominantly accumulated on the ventral surface of the thallus and has two forms: the predominant form is the aglycone, which is cell wall-bound and water-insoluble; the other is auronidin 4-neohesperidoside, which is the auronidin aglycone with a rhamnose sugar attached that in turn has a glucose attached. This form is found in the vacuole and is water-soluble, but has only been detected in transgenic plants that have increased auronidin production (Berland et al., 2019).

Because auronidin is a recently identified compound, the pathway to its biosynthesis is poorly understood. It is known that the precursors of auronidin, aurones, are synthesised from chalcones by POLYPHENOL OXIDASES (PPOs) such as AUREUSIDIN SYNTHASE (AUS) and AURONE SYNTHASE (AS) (Berland et al., 2019; Furudate et al., 2023). However, the enzyme(s) that converts aurones into auronidin is currently unknown. The only current information on the regulatory pathway for auronidins is that a R2R3-MYB transcription factor, MYB14, is required for auronidin biosynthesis - *myb14* mutants lack auronidin and *Myb14* over-expression plants have increased auronidin production (Albert et al., 2018; Kubo et al., 2018).

The auronidins are analogous to anthocyanins produced by many angiosperm plants, having similar red pigmentation, being induced by abiotic stress (Berland et al., 2019) and regulated by similar classes of transcription factors (Albert et al 2018). In apple, a protein called BTB-TAZ DOMAIN 2 (BT2) negatively regulates anthocyanin biosynthesis by promoting the ubiquitination and degradation of MdMYB1, a transcription factor required for anthocyanin biosynthesis (Wang et al., 2018). BT proteins have also been identified in other land plants, such as in the lycophyte *Selaginella tamariscina* (Wang et al., 2018), suggesting their presence in land plants other than angiosperms. Currently, BT proteins have not been identified in bryophytes. However, because nitrogen deprivation induces auronidin production in Marchantia and this is regulated by an R2R3-MYB, similarly to anthocyanins (Albert et al., 2018; Davies et al., 2020), it would not be surprising if there

was a BT2 protein in Marchantia. Recently, the protein Mapoly0010s0173 was identified in the Marchantia genome that is a strong BLAST match to the apple and *Arabidopsis* BT2 proteins, with 44% and 39% protein identity, respectively (Appendix 13). Furthermore, the next highest BLAST hit only had 20% protein identity with both the apple and *Arabidopsis* BT2 proteins, and is truncated so may lack certain domains. Thus, Mapoly0010s0173 seems to be the only homolog of angiosperm BT2 proteins. It is also a convincing candidate for a gene encoding a BTB-TAZ DOMAIN 2 (MpBT2) orthologue because it has all three domains usually present in BT2 proteins (BTB/POZ, BACK and TAZ) plus the nuclear localisation signal. This suggests MpBT2 could have a role comparable with that of MdBT2, negatively regulating auronidin biosynthesis by inducing the degradation of MYB14.

The Mp*BT2* candidate gene was therefore functionally characterised by using CRISPR/Cas9 and *A. tumefaciens*-mediated spore transformation (optimised in Chapter 3) to generate loss-of-function mutants. Additionally, the Mp*BT2* gene was over-expressed in wild-type and *35S:Myb14* genetic backgrounds. In all cases, the transgenic plants were assessed for their ability to produce auronidin.

5.2 Materials and methods

5.2.1 CRISPR-Cas9 and over-expression constructs

The Mp*BT2* CRISPR/Cas9 construct and over-expression construct (used in spore transformation; section 5.5.2) were designed and cloned into a binary vector as described in section 2.1 and 2.2 (pRMY09 and pRMY27, respectively; Appendix 3.2 and 3.3 for plasmid map). The Mp*BT2* over-expression construct used in thallus transformation (section 5.2.3) (pRMY30; Appendix 3.3 for plasmid map) was designed and cloned similarly, except that a pNWA99 destination vector (created by Dr Nick Albert) was used for LR gateway recombination to provide plant resistance to hygromycin instead of geneticin. For the CRISPR/Cas9 construct, three high quality sgRNAs were designed to target exon one and one to target exon two of the Mp*BT2* candidate *Mapoly0010s0173* because exon 1 is an extended N terminal region that may not be conserved, but exon 2 is conserved. The position of these sgRNAs is shown in Figure 5.1 below.

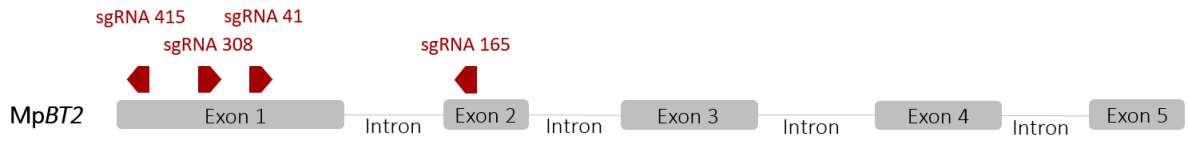


Figure 5.1. Position of the four single guide RNA (sgRNA) molecules designed to target exons one and two of *Mapoly0010s0173,* the *BTB-TAZ Domain 2* (Mp*BT2*) candidate gene in *Marchantia polymorpha*.

5.2.2 Stable transgenic lines – spore transformation

Marchantia plants transformed with the Mp*BT2* CRISPR/Cas9 construct and plants over-expressing Mp*BT2* were created using spore transformation (section 2.4) with *A. tumefaciens* harbouring the constructs described in section 5.2.1. Regenerated CRISPR/Cas9 plants were screened for Mp*BT2* mutations by PCR (section 2.4.5.2) and DNA sequencing (section 2.2.11), using the gene-specific primers NA711 and NA712 (Appendix 4) and were propagated through gemmae (G_1 stage) and re-screened, ensuring the generation of pure mutant lines. Regenerated Mp*BT2* over-expression plants were screened for the presence of the transgene by PCR-amplification of the region between the *35S* promoter on the pNWA101 destination vector and the Mp*BT2* gene, using NA354 and NA712 primers (Appendix 4). PCR products were not sequenced because the presence or absence of the transgene was sufficient to confirm plants were transgenic.

5.2.3 Stable transgenic lines – thallus transformation

MpBT2 over-expression plants were additionally generated by *A. tumefaciens*-mediated transformation of *35S*:*Myb14* thallus tissue. The *35S*:*Myb14* background was used to increase auronidin production, as MYB14 upregulates auronidin biosynthesis. This protocol was developed by Dr Yanfei Zhou and is based on similar thallus transformation methods (Kubota et al., 2013; Tsuboyama & Kodama, 2018a).

Gemmae were grown on 0.5× B5 media for 13 d until the thallus was 1.5 - 2 cm in diameter. Meristems were removed with a scalpel and the remaining thallus was cut into ~5 mm pieces. The thallus pieces were grown on 0.5× B5 media with 200 µM acetosyringone for three d. *A. tumefaciens* harbouring the Mp*BT2* over-expression construct (pRMY30) was grown and the co-cultivation medium prepared as described in section 2.4. The final A_{600} was adjusted to 0.2. Tween20 was added at a final concentration of 0.02% (v/v). A 1 mL aliquot of the suspension was pipetted onto the thallus pieces, before being left to co-cultivate in dark conditions for four d. Following co-cultivation, the pieces were washed in sterile water three times to remove *A. tumefaciens* then transferred to filter paper on selective 0.5× B5 media containing 10 mg L^{-1} hygromycin and 100 mg L^{-1} cefalexin, ensuring that only transgenic tissue regenerated. Hygromycin was used because the *35S*:*Myb14* transgenic line was generated using geneticin (G-418) selection, so is sensitive to hygromycin. After three weeks, regenerated plants were transferred to 0.5× B5 media with 500 mg L^{-1} ticarcillin as they appeared. Plants were screened for Mp*BT2* transgenic regions as described for the Mp*BT2* over-expression plants in section 5.2.2.

5.2.4 Plant material and growth conditions

Nutrient stress was used to induce flavonoid production in *bt2* mutants and *35S*:*BT2* plants. These conditions were provided by nutrient deprivation media, containing 1% (w/v) sucrose and 1% (w/v) agar. Control plants were *chiL* and Aud-2 female and Sey-1 male wild-type plants. A no-stress treatment was also included for all plant lines. Gemmae were grown on MCE membrane filters placed on 0.5× B5 media for three weeks. Plants were then transferred to nutrient deprivation media (nutrient stress treatment), or onto fresh 0.5× B5 media (no treatment) and grown for a further 13 d. Four gemmae per tub and at least four biological replicates, consisting of separate tubs, were grown per plant line and treatment.

The *35S*:*Myb14*/*35S*:*BT2* plants were not exposed to nutrient deprivation stress. Instead, gemmae (four per tub and five tubs/biological replicates per plant line) were grown on MCE membrane filters placed on 0.5× B5 media for four weeks. Four of the biological replicates were used for chemical analysis. The fifth was grown for an extra four weeks to show auronidin saturation.

Plants were imaged (section 2.5), then harvested and prepared for flavonoid analysis by HPLC as described for *fnsI* mutants in section 4.2.6. Auronidin and flavones were quantified as described. A peak for the bibenzyl precursor lunularic acid was also identified, based on a previous mass-spectrometry analysis (unpublished data of Dr Christelle Andre, Plant and Food Research, New Zealand), and quantified. However, as a lunularic acid standard was not available, quantification was based on the integrated peak area at 310 nm (absorbance units multiplied by retention time; mAU.min), the extraction solvent volume and the weight of the ground tissue used for extraction.

5.3 Results

5.3.1 Stable *bt2* mutant and Mp*BT2* over-expression plants

The Marchantia Mp*BT2* candidate gene was characterised by using CRISPR/Cas9 and *A. tumefaciens*-mediated spore transformation to generate *bt2* mutants and both spore transformation and thallus transformation to generate plants over-expressing Mp*BT2*.

Once Mp*BT2* CRISPR/Cas9 plants were generated, the first 10 lines (G_1 stage) were selected and screened for Mp*BT2* mutations using PCR-amplification of a target region (1,330 bp) followed by gel electrophoresis and DNA sequencing. PCR products of several transformed plant lines showed size differences to the wild-type product, suggesting that deletions may have occurred (Figure 5.2). The wild-type plant had a faint amplicon band of 1,330 bp, as well as a low molecular weight non-specific band. In comparison, bands with a size difference of ~800 bp, ~50-100 bp and ~10 bp were present in the samples from the *bt2*/8, *bt2*/9 and *bt2*/4 plant lines, respectively. PCR products from all other lines did not have noticeable size differences. The negative PCR control had one faint DNA band, but this did not correspond to the expected size of the amplicon. The PCR products of all 10 transformed plant lines were sequenced regardless of whether size differences were observed because even 1 bp mutations can be sufficient for loss of gene function.

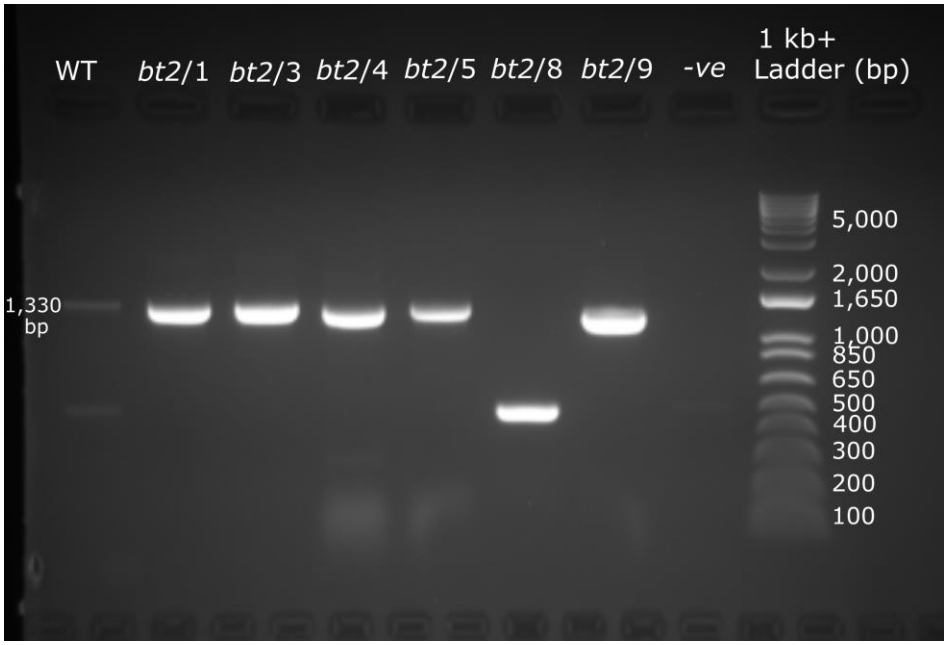


Figure 5.2. **Screening *Marchantia polymorpha* plants generated using CRISPR/Cas9 for size differences in the *BTB-TAZ Domain 2* (*BT2*) candidate gene.** PCR-amplification of a Mp*BT2* target region (1,330 bp) in *BT2* CRISPR/Cas9 plants compared to a wild-type (WT) plant, showing size differences. A negative PCR control (-ve) was also included.

Sequencing results revealed that mutations were present in five of the 10 CRISPR/Cas9 plants screened, most of which were deletions (Table 5.1). Thus, *bt2* mutants were successfully generated. The mutations present in three *bt2* mutants selected for further study are summarised in Figure 5.3. In comparison to the wild-type sequence, *bt2*/5 contains a 1 bp deletion at sgRNA #41 (which was not evident in the PCR product), *bt2*/9 contains an 87 bp deletion spanning sgRNAs #308 and #41 and a 1 bp deletion at sgRNA #165, and *bt2*/8 contains an 822 bp deletion from sgRNAs #41 to #165 (see Figure 5.1 for sgRNA positions in the Mp*BT2* coding sequence). Interestingly, *bt2*/8 also contains an 11 bp insertion at sgRNA #165, which is actually a duplication of a nearby sequence (underlined) rather than a random insertion of bases pairs. Another curious observation is that the deletions described in Figure 5.3 are further than 3 bp upstream of the PAM site as expected with the CRISPR/Cas9 system.

Table 5.1. Summary of identified *Marchantia polymorpha bt2* mutants generated using CRISPR/Cas9. DNA sequencing results of all 10 Mp*BT2* CRISPR/Cas9 plants, showing the types of mutations present and the single guide RNAs (sgRNAs) where they are located.

Plant line ID	Mutant? (Y/N)	Mutation description	sgRNA #
bt2/1	Y	1 bp deletion	41
bt2/2	N	-	-
bt2/3	N	-	-
bt2/4	Y	1 bp deletion, 8 bp deletion	41, 165
bt2/5	Y	1 bp deletion	41
bt2/6	N	-	-
bt2/7	N	-	-
bt2/8	Y	822 bp deletion, 11 bp insertion	41 and 165, 165
bt2/9	Y	87 bp deletion, 1 bp deletion	308 and 41, 165
bt2/10	N	-	-

Figure 5.3. Examples of mutations in the *BTB-TAZ Domain 2* (*BT2*) candidate gene of *Marchantia polymorpha* plants generated using CRISPR/Cas9. DNA sequencing results of three *bt2* mutants (*bt2*/5, *bt2*/8 and *bt2*/9) compared to a wild-type sequence. Two single guide RNAs (sgRNAs; #41 and #165) are shown in green text, followed by their adjacent PAM (protospacer adjacent motif) site in blue text. Mutations are shown in red text. The "//" symbol represents an introduced break in the sequence.

Plants over-expressing Mp*BT2* were generated by transforming Marchantia spores with a *35S*:*BT2* construct (hereafter, *35S*:*BT2* plants) and by transforming Marchantia *35S*:*Myb14* thallus with a *35S*:*BT2* construct (hereafter, *35S*:*Myb14*/*35S*:*BT2* plants). The first 30 regenerated plants were screened for the presence of the transformed transgene by PCR-amplification of a 1,750 bp target region, which should be absent from wild-type DNA samples because the primers used for PCR-amplification were specific to the transgene. Figure 5.4 shows that several *35S*:*BT2* and *35S*:*Myb14*/*35S*:*BT2* plants had amplicons of the expected size, albeit some were faint, confirming the presence of the transgene. The wild-type and negative (no-DNA sample) PCR controls did not contain any DNA bands. Thus, transgenic plants were successfully generated. Three lines of each of the *35S*:*BT2* and *35S*:*Myb14*/*35S*:*BT2* plants were selected for further studies.

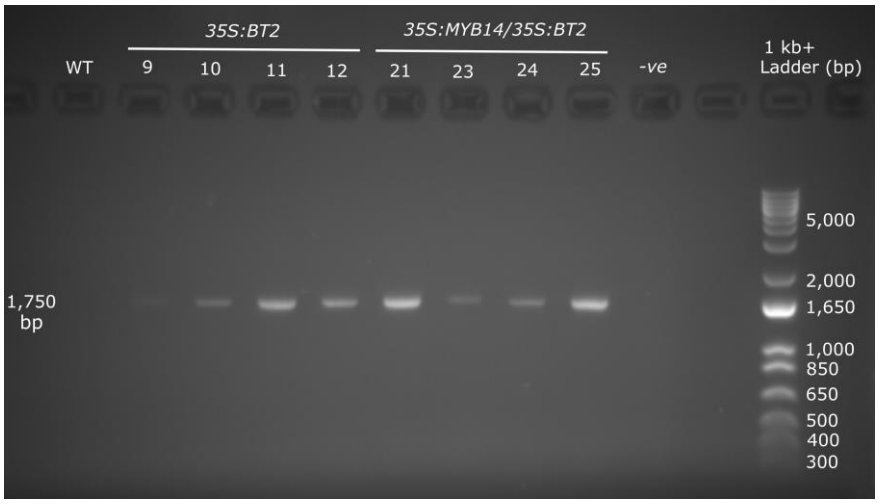


Figure 5.4. Screening *Marchantia polymorpha 35S*:*BT2* and *35S*:*Myb14*/*35S*:*BT2* plants for the presence of the *BTB-TAZ Domain 2* (*BT2*) candidate gene. PCR-amplification of a Mp*BT2* target region (1,750 bp) in transformed plants compared to a wild-type (WT) plant, showing the presence or absence of transgenic DNA. A negative PCR control (-ve) was also included.

5.3.2 Flavonoid analysis of *bt2* mutant and *35S*:*BT2* plants

To determine the function of the Mp*BT2* candidate gene in Marchantia, confirmed *bt2* mutants and *35S*:*BT2* plants were exposed to nutrient stress to induce flavonoid production, then characterised by HPLC for their ability to produce auronidin (and other phenolic compounds).

Phenotypes of plants exposed to both nutrient stress and no-stress conditions are illustrated in Figure 5.5. A key observation is that the plants not exposed to nutrient stress had a normal appearance, whereas nutrient deprivation caused a visual increase in the production of auronidin (red pigment). Expectedly, the *chiL* mutant (known to have reduced flavone and auronidin production (Berland et al., 2019)), had no visible auronidin but had some brown colouration indicative of stress. This was particularly noticeable on the ventral side of the thallus. The no-stress treatment plants were not included in the HPLC analysis because auronidin cannot be reliably detected without stress-induction (Albert et al., 2018; Berland et al., 2019).

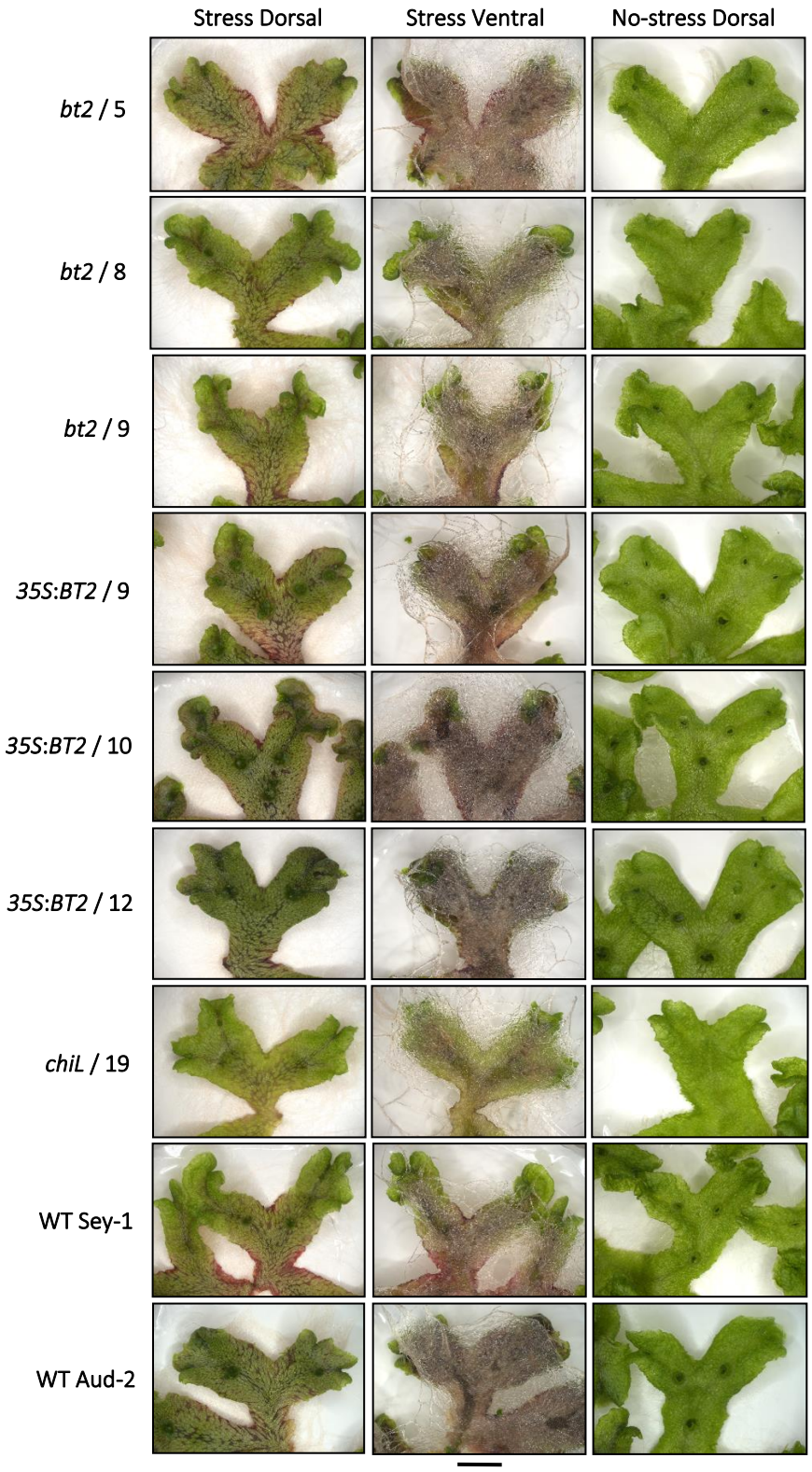


Figure 5.5. Representative phenotypes of *bt2*, *35S*:*BT2*, *chiL* and wild-type (WT) *Marchantia polymorpha* plants under nutrient stress conditions. Phenotypes are shown for the dorsal (top) and ventral (bottom) sides of stressed plants (left and middle columns, respectively) and the dorsal side of non-stressed plants (right column). Plants were grown on 0.5× B5 media for three weeks, then stressed plants were transferred to a nutrient deprivation media and non-stress plants to fresh 0.5× B5 for two weeks. Scale bar represents 5 mm.

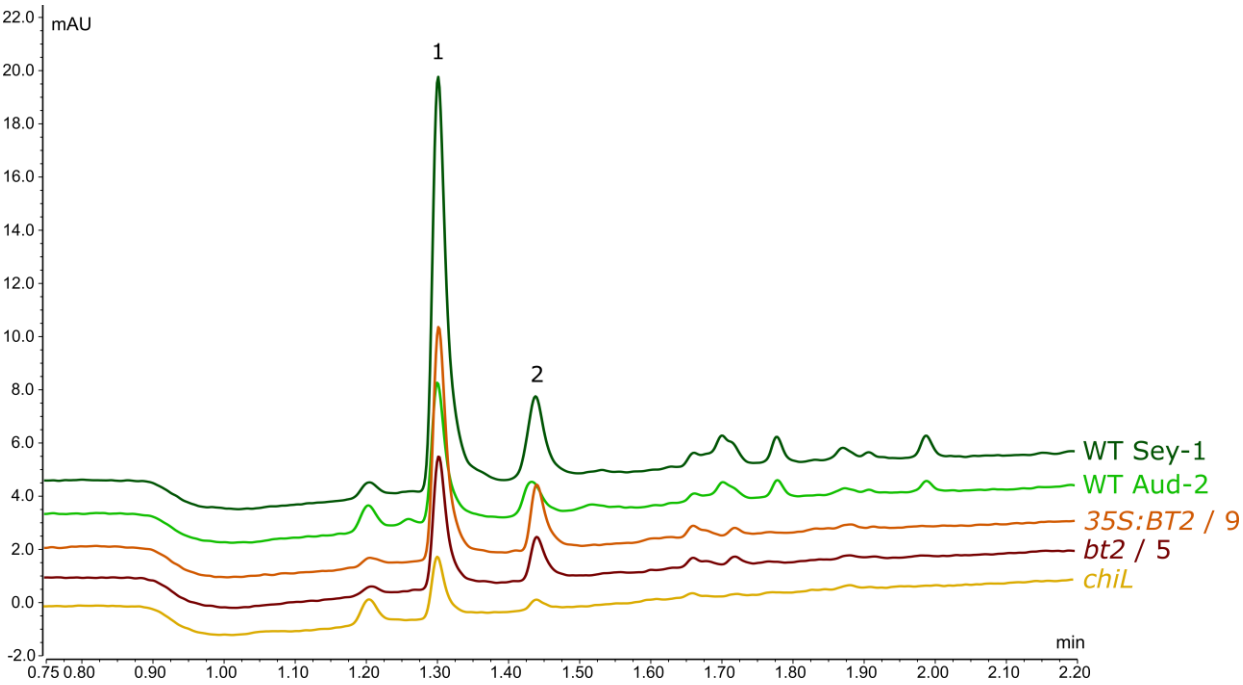
HPLC analysis revealed that all plants exposed to nutrient stress produced two auronidin aglycone peaks (1 and 2; Figure 5.6A), six flavone peaks, which were identified as luteolin-7,4'-di-*O*-glucuronide (1), luteolin-7,3'-di-*O*-glucuronide (2), luteolin-7-*O*-glucuronide (3), a mixture of luteolin-3'-*O*-glucuronide and apigenin-mono-*O*-glucuronide (4), apigenin-7-*O*-glucuronide (5) and luteolin-4'-*O*-glucuronide (6) (Figure 5.7A), and a lunularic acid peak (Figure 5.8A). Only qualitative differences in the abundance of these compounds were evident based on the HPLC traces. For practical reasons, the HPLC trace from only one line of *bt2* mutants and *35S*:*BT2* plants, that is representative of the overall result, is shown.

HPLC quantification showed that the parental lines used to generate spores for transformation differ in auronidin content; male Sey-1 wild-type plants had a significantly higher auronidin content than the female Aud-2 wild-type, of ~0.35 and 0.12 mg g^{-1} DW, respectively (Figure 5.6B). The *bt2* mutants and *35S*:*BT2* plants had an auronidin content either similar to or in between that of the wild-type plants. The *35S*:*BT2*/9 and *35S*:*BT2*/10 plant lines had an auronidin content of ~0.22 mg g^{-1} DW, which is significantly higher than the *35S*:*BT2*/12, *bt2*/5 and *bt2*/8 plants that had a similar auronidin content to the female wild-type. Additionally, the *bt2*/9 plants had a significantly higher content than the other *bt2* and *35S*:*BT2* plants, of ~0.3 mg g^{-1} DW. Comparatively, the *chiL* mutant had a significantly lower auronidin content than all other plant lines (~0.8 mg g^{-1} DW).

Similarly, flavone quantification revealed that the male wild-type plants produced significantly more flavones than the female wild-type, ~1.5 compared to 1.0 mg g^{-1} DW (Figure 5.7B). The *bt2* mutants had a similar flavone content to the wild-type plants. However, interestingly, all of the *35S*:*BT2* plant lines had a significantly higher flavone content of ~2.0 – 2.5 mg g^{-1} DW. Within the *35S*:*BT2* lines, *35S*:*BT2*/*9* produced significantly more flavones than the *35S*:*BT2*/10 and *35S*:*BT2*/12 plants. The *chiL* mutant again had a significantly lower flavone content than all other plants assayed (~0.5 mg g^{-1} DW).

Another compound identified in HPLC analysis was lunularic acid, which is the precursor to the cyclic *bis*-bibenzyl marchantin A. The *chiL* mutant and *35S*:*BT2*/*10* and *35S*:*BT2*/*12* plants had a lunularic acid content of ~250 mAU.min g^{-1} DW (310 nm), which was significantly higher than all other plants assayed (Figure 5.8B). The *35S*:*BT2*/*9* plant line had a lunularic acid content more comparable to the female wild-type and *bt2* mutants, of ~150 – 200 mAU.min g^{-1} DW. The male wild-type plants had the lowest lunularic acid content (~120 mAU.min g^{-1} DW).

(A)



(B)

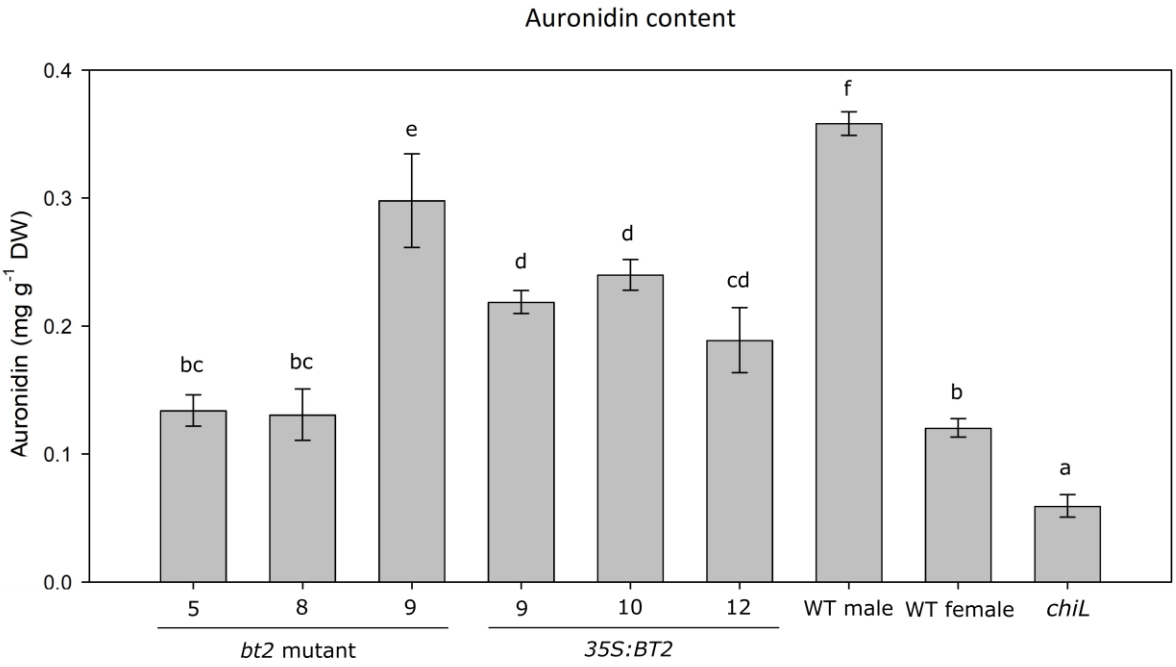


Figure 5.6. Auronidin content of *bt2* and *35S*:*BT2* *Marchantia polymorpha* plants under nutrient stress conditions. (A) High performance liquid chromatography (HPLC) chromatograms of representative *bt2* (red), *35S*:*BT2* (orange), *chiL* (yellow) and wild-type (WT; green) plants. Auronidin was detected at 484 nm. Peaks were identified by comparison to Albert et al. (2018) and Berland et al. (2019). Peaks 1 and 2 are different conformations of the auronidin aglycone. mAU: absorbance units in nm; min: retention time. (B) Quantification of auronidin content in *bt2*, *35S*:*BT2*, *chiL* and wild-type Sey-1 (WT male) and Aud-2 (WT female) plants. Error bars are the standard error of the means, n = 4. Significantly different means are represented by different letters (a – f).

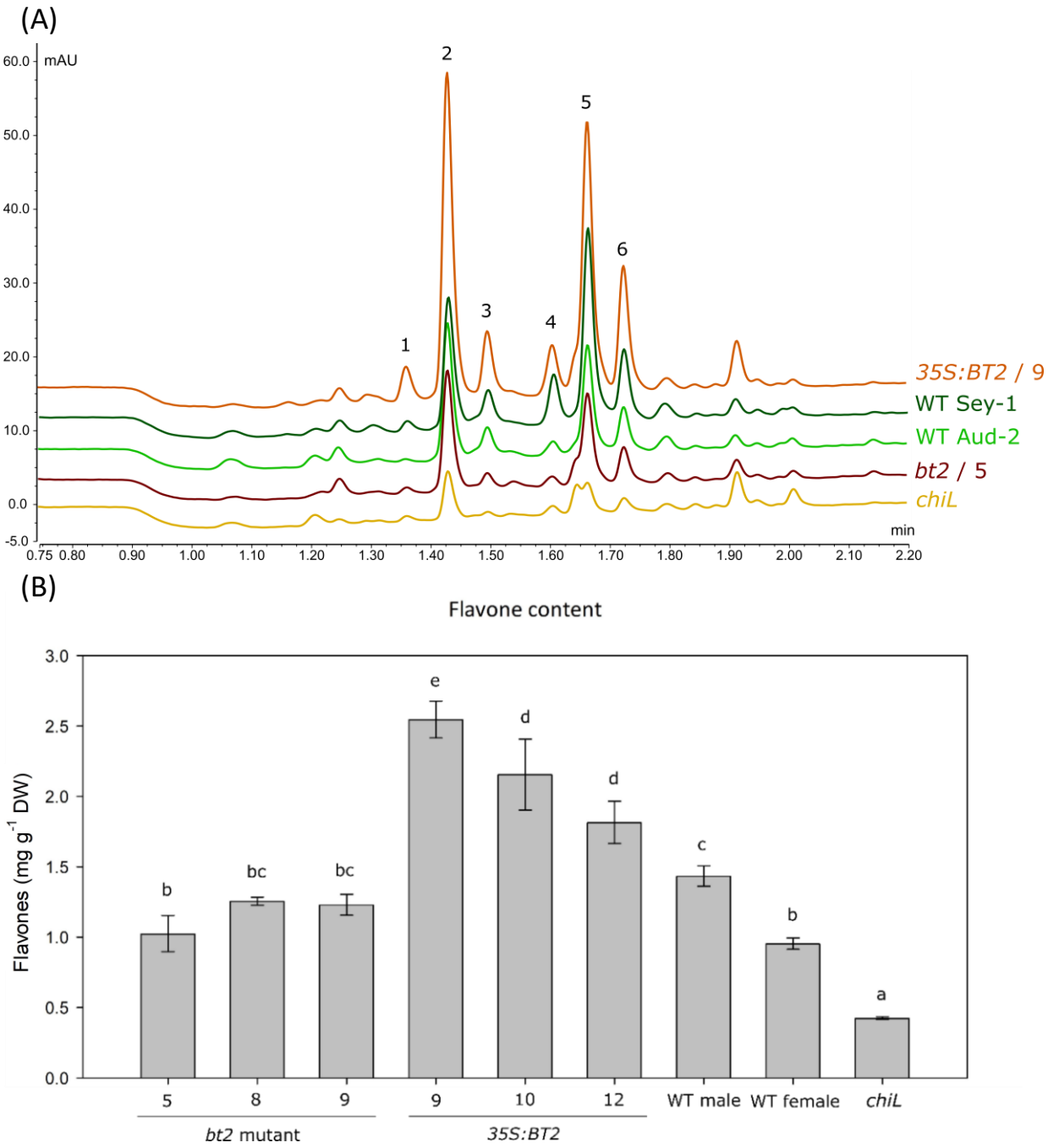


Figure 5.7 Flavone content in *bt2* and *35S*:*BT2 Marchantia polymorpha* plants under nutrient stress conditions. (A) High performance liquid chromatography (HPLC) chromatogram of representative *bt2* (red), *35S*:*BT2* (orange), *chiL* (yellow) and wild-type (WT; green) plants. Flavones were detected at 340 nm and peaks were identified by a previous mass-spectrometry analysis (unpublished data of Dr Tony McGhie, Plant and Food Research, New Zealand) and comparison to Markham et al. (1998) and Albert et al. (2018). (1) luteolin-7,4'-di-*O*-glucuronide; (2) luteolin-7,3'-di-*O*-glucuronide; (3) luteolin-7-*O*-glucuronide; (4) a mixture of luteolin-3'-*O*-glucuronide and apigenin-mono-*O*-glucuronide; (5) apigenin-7-*O*-glucuronide and (6) luteolin-4'-*O*-glucuronide. mAU: absorbance units in nm; min: retention time. (B) Quantification of flavone content in *bt2*, *35S*:*BT2*, *chiL* and wild-type Sey-1 (WT male) and Aud-2 (WT female) plants. Error bars are the standard error of the means, n = 4. Significantly different means are represented by different letters (a – e).

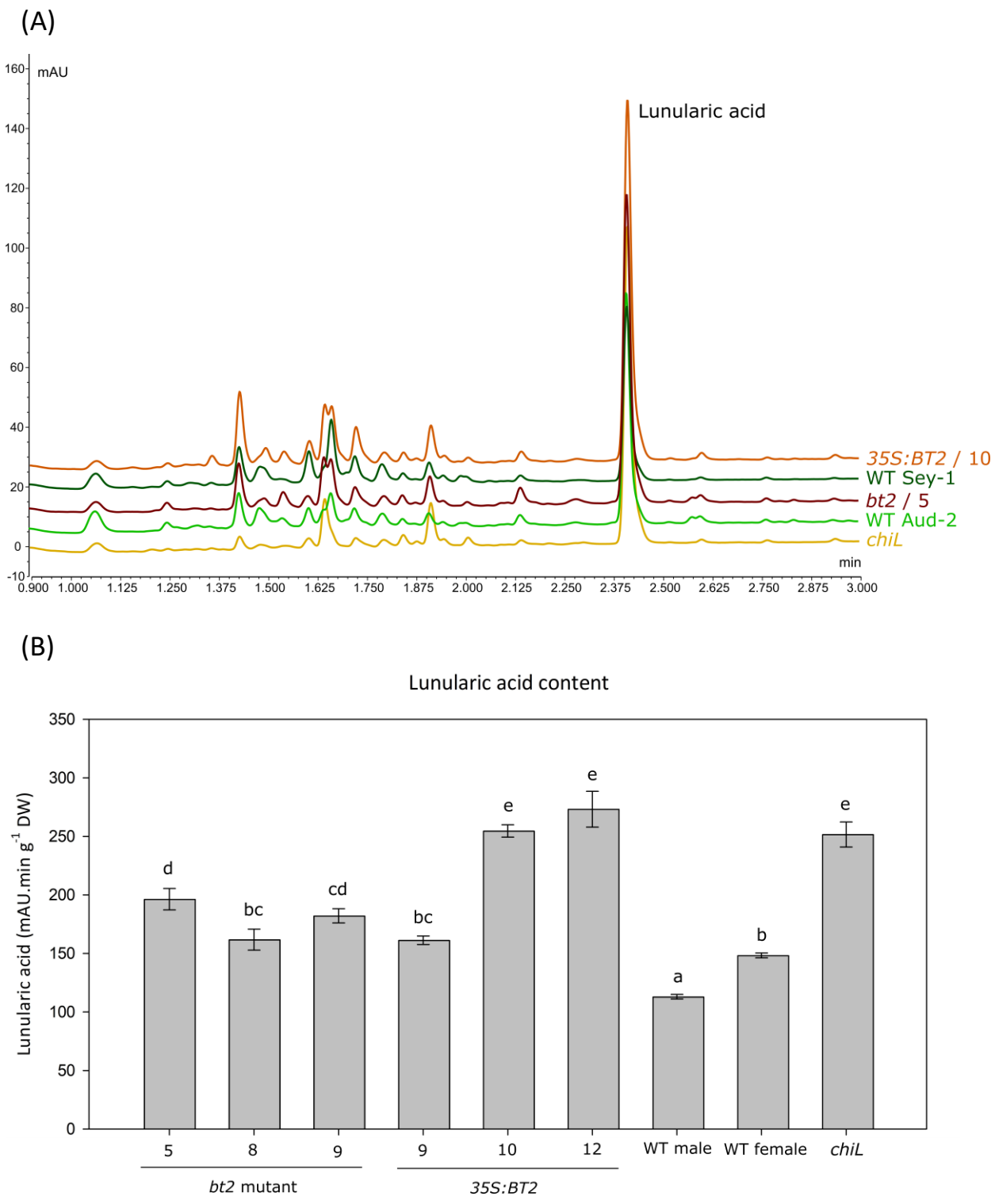


Figure 5.8 Lunularic acid content in *bt2* and *35S*:*BT2 Marchantia polymorpha* plants under nutrient stress conditions. (A) High performance liquid chromatography (HPLC) chromatogram of representative *bt2* (red), *35S*:*BT2* (orange), *chiL* (yellow) and wild-type (WT; green) plants. The lunularic acid peak detected at 310 nm was identified in a previous mass-spectrometry analysis (unpublished data of Dr Christelle Andre, Plant and Food Research, New Zealand). mAU: absorbance units in nm; min: retention time. (B) Quantification of lunularic acid content in *bt2*, *35S*:*BT2*, *chiL* and wild-type Sey-1 (WT male) and Aud-2 (WT female) plants. Error bars are the standard error of the means, n = 4. Significantly different means are represented by different letters (a – e).

5.3.3 Flavonoid analysis of *35S*:*Myb14*/*35S*:*BT2* plants

The function of the Mp*BT2* gene was also characterised by using thallus transformation to over-express Mp*BT2* in a single *35S*:*Myb14* plant line (*35S*:*Myb14*/*35S*:*BT2* plants). This was performed so that regenerated plants had a uniform genetic background, unlike those generated from spore transformation. Exposing plants to nutrient stress was not required to induce flavonoid production because *35S*:*Myb14* plants constitutively produce increased auronidin and flavone amounts.

Representative phenotypes of *35S*:*Myb14*/*35S*:*BT2* and *35S*:*Myb14* control plants are shown in Figure 5.9A. The plants had abundant auronidin pigmentation (red/orange colour) on the dorsal and ventral surfaces. To illustrate why plants were sampled at this early developmental stage for HPLC analysis, the same plant lines were grown for an additional four weeks and phenotypes imaged again (Figure 5.9B). The auronidin content was visibly saturated, as the entire thallus was a deep red/orange colour and all of the plants appeared to have the same intensity of pigmentation.

Flavonoid analysis revealed that representative *35S*:*Myb14* control and *35S*:*Myb14*/*35S*:*BT2* plants produced auronidin 4-neohesperidoside (1) and the two auronidin aglycone peaks (2 and 3; Figure 5.10A). Four flavone compounds were detected, and identified as luteolin-7,3'-di-*O*-glucuronide (1), luteolin-7-*O*-glucuronide (2), apigenin-7-*O*-glucuronide (3) and luteolin-4'-*O*-glucuronide (4) (Figure 5.11A). A single lunularic acid peak was also identified (Figure 5.12A).

Quantification showed no significant difference in auronidin content between the plant lines (Figure 5.10B). There were few differences in flavone content, except that the *35S*:*Myb14*/*35S*:*BT2*/23 plants produced significantly more flavones than all other lines (Figure 5.11B). Interestingly, all of the *35S*:*Myb14*/*35S*:*BT2* plant lines had a significantly lower lunularic acid content than the *35S*:*Myb14* control (Figure 5.12B).

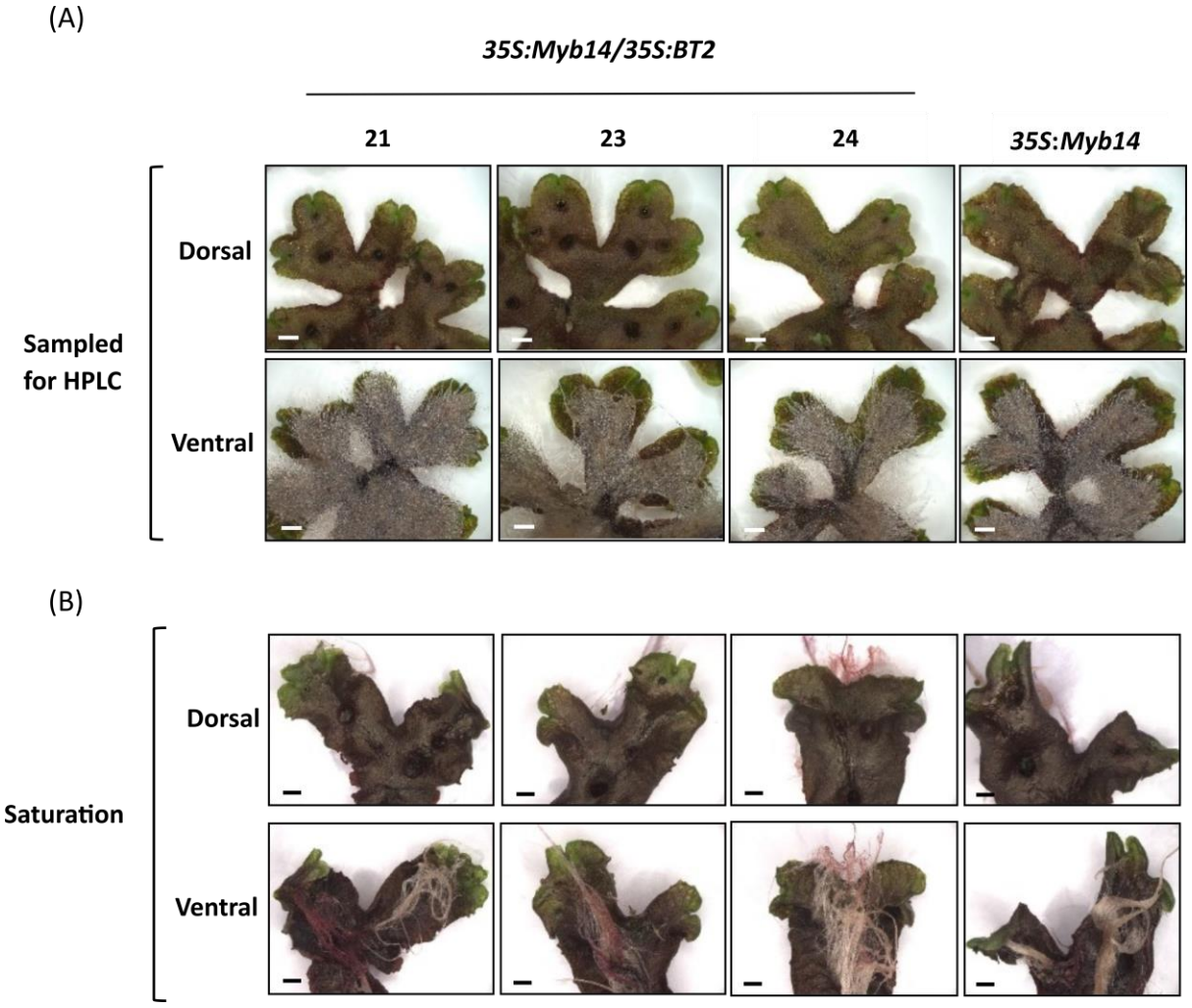


Figure 5.9. Phenotypes of *35S*:*Myb14/35S*:*BT2* and *35S*:*Myb14* control *Marchantia polymorpha* plants. The dorsal (top) side of the plants is shown in the top panels and ventral (bottom) side in the bottom panels. The *35S*:*Myb14/35S*:*BT2* lines examined were 21, 23 and 24. Scale bars represent 2 mm. (A) Phenotypes of plants harvested for HPLC after four weeks of growth. (B) Phenotypes of plants grown for eight weeks, showing auronidin saturation.

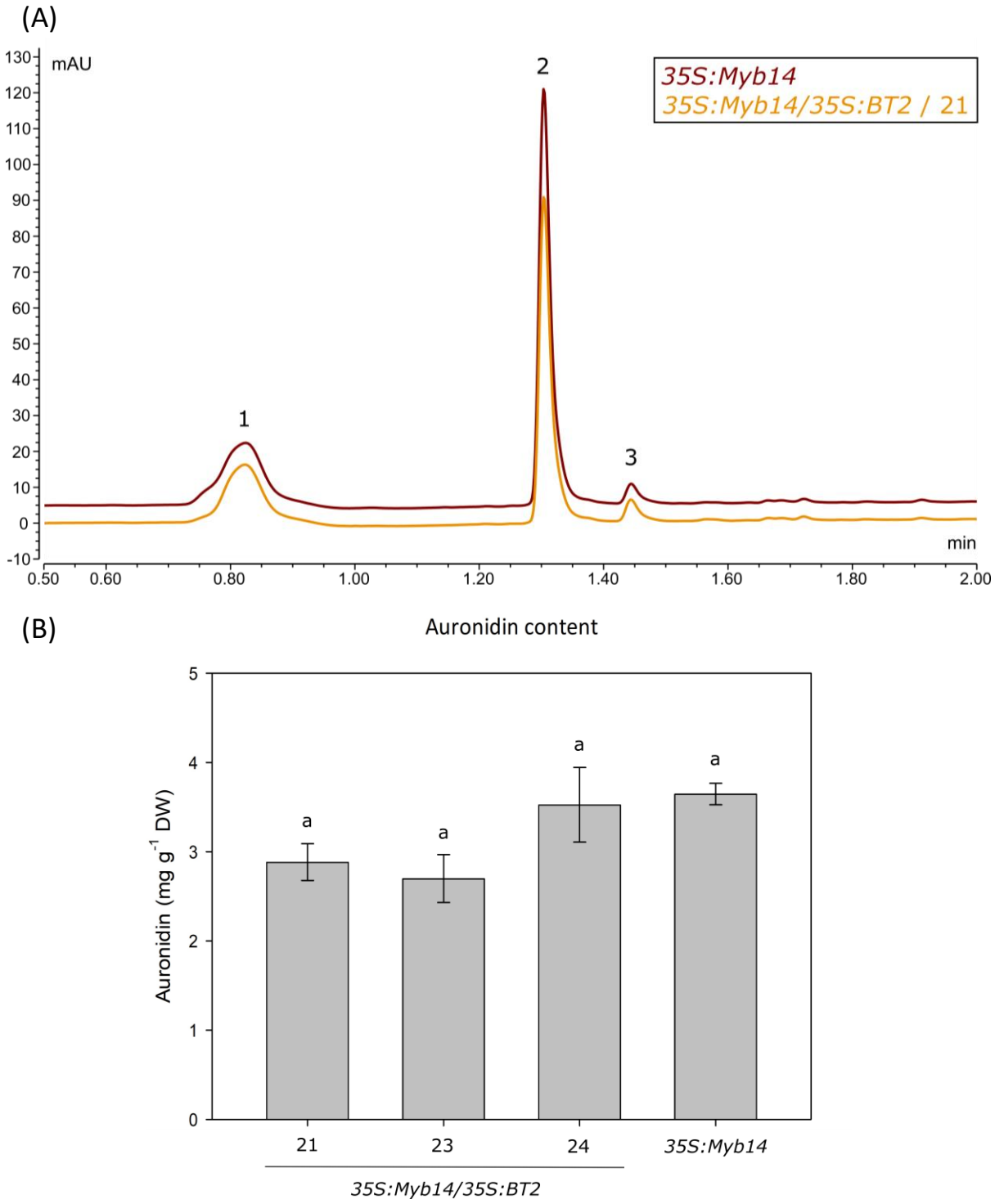


Figure 5.10. Auronidin content in *35S*:*Myb14*/*35S*:*BT2* *Marchantia polymorpha* plants. (A) High performance liquid chromatography (HPLC) chromatograms of representative *35S*:*Myb14*/*35S*:*BT2* (orange) and *35S*:*Myb14* control (red) plants. Auronidin was detected at 484 nm. Peaks were identified by comparison to Albert et al. (2018) and Berland et al. (2019). Peak 1 is auronidin 4-neohesperidoside and peaks 2 and 3 are different conformations of the auronidin aglycone. mAU: absorbance units in nm; min: retention time. (B) Quantification of auronidin content in *35S*:*Myb14*/*35S*:BT2 (lines examined were 21, 23 and 24) and *35S*:*Myb14* control plants. Error bars are the standard error of the means, n = 4. Significantly different means are represented by different letters.

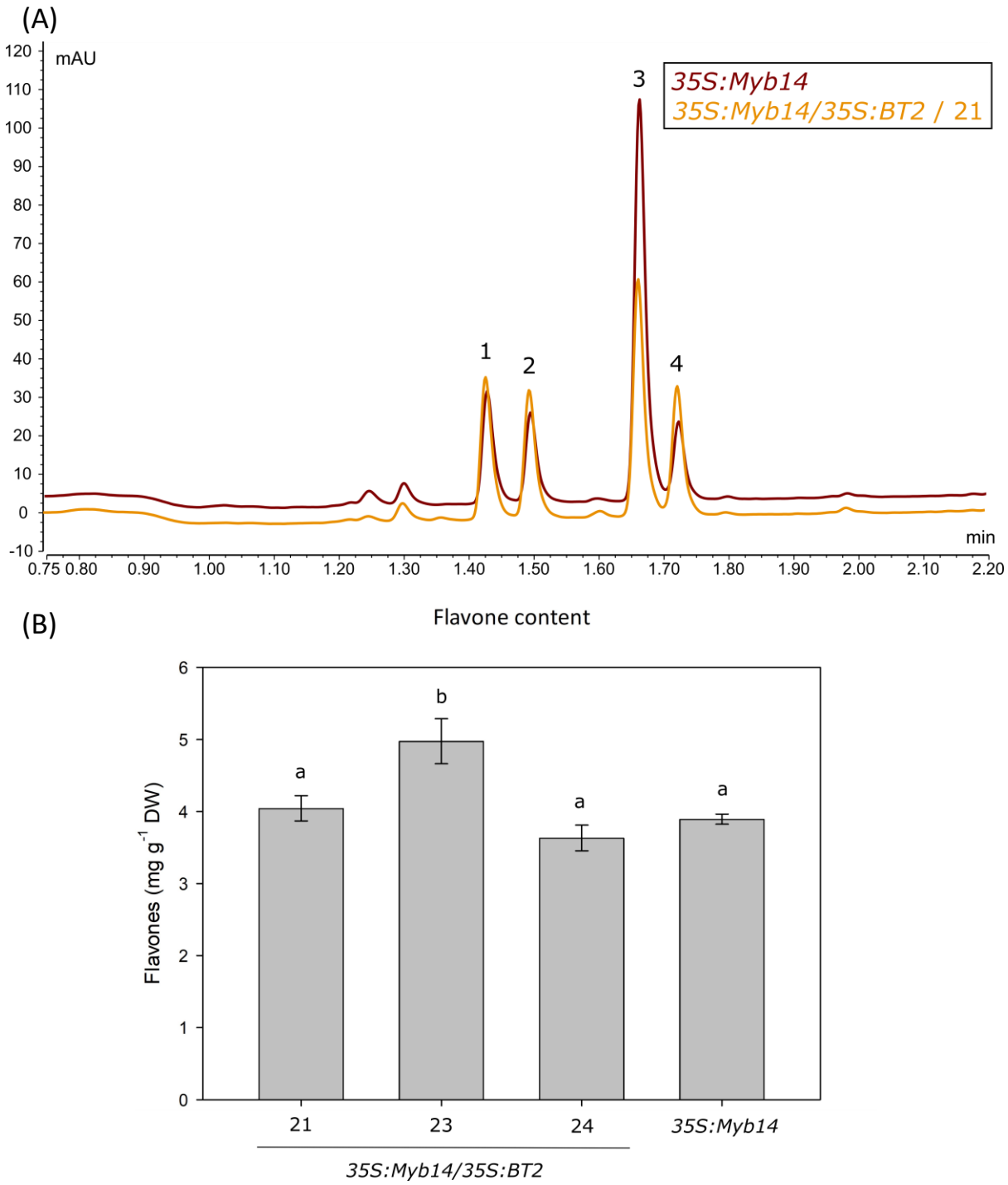


Figure 5.11 Flavone content in *35S*:*Myb14*/*35S*:*BT2* *Marchantia polymorpha* plants. (A) High performance liquid chromatography (HPLC) chromatogram of representative *35S*:*Myb14*/*35S*:*BT2* (orange) and *35S*:*Myb14* control (red) plants. Flavones were detected at 340 nm and peaks were identified by a previous mass-spectrometry analysis (unpublished data of Dr Tony McGhie, Plant and Food Research, New Zealand) and comparison to Markham et al. (1998) and Albert et al. (2018). (1) luteolin-7,3'-di-*O*-glucuronide; (2) luteolin-7-*O*-glucuronide; (3) apigenin-7-*O*-glucuronide and (4) luteolin-4'-*O*-glucuronide. mAU: absorbance units in nm; min: retention time. (B) Quantification of flavone content in *35S*:*Myb14*/*35S*:*BT2* (lines examined were 21, 23 and 24) and *35S*:*Myb14* control plants. Error bars are the standard error of the means, n = 4. Significantly different means are represented by different letters (a and b).

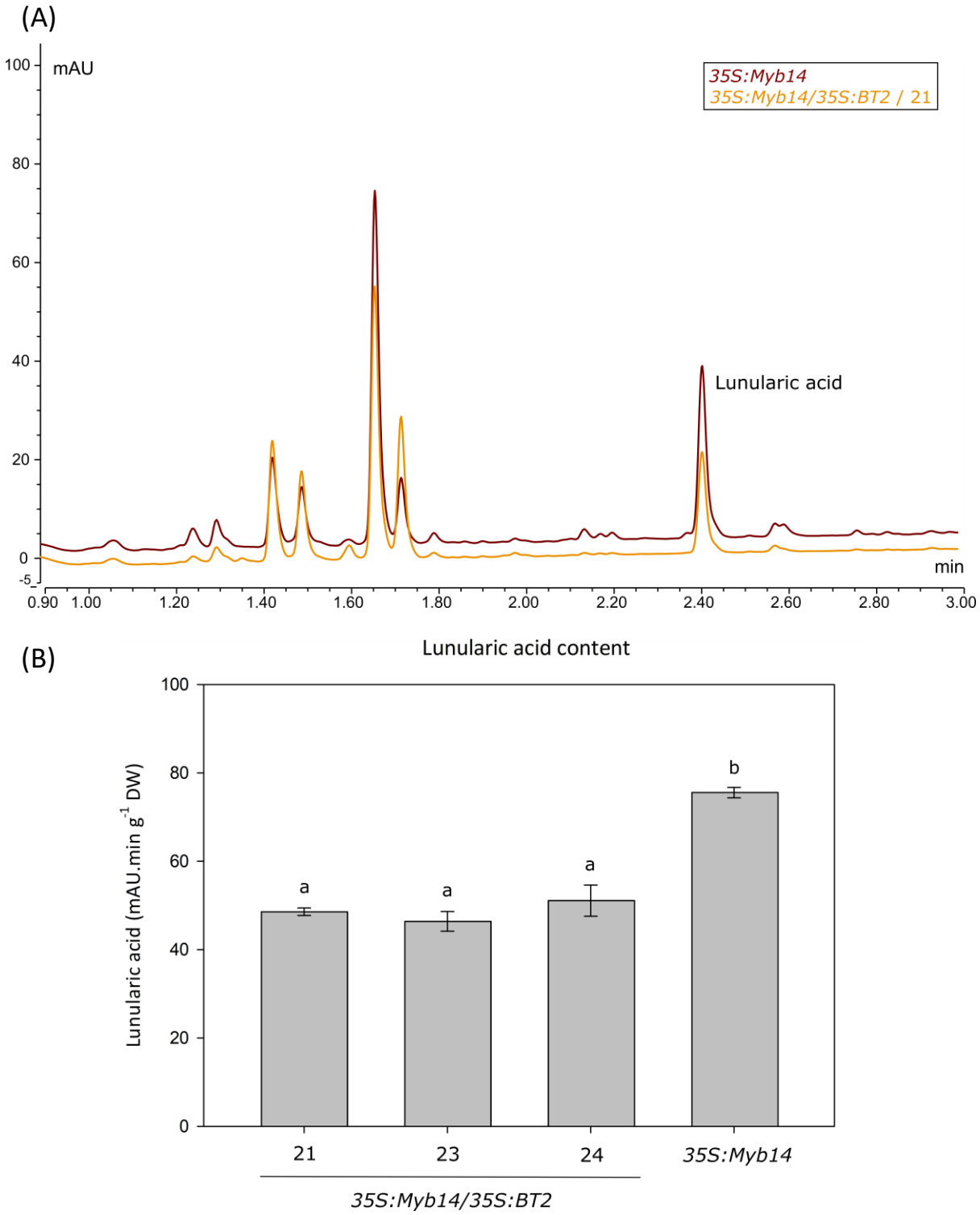


Figure 5.12 Lunularic acid content in *35S*:*Myb14*/*35S*:*BT2* *Marchantia polymorpha* plants. (A) High performance liquid chromatography (HPLC) chromatogram of representative *35S*:*Myb14*/*35S*:*BT2* (orange) and *35S*:*Myb14* control (red) plants. The lunularic acid peak detected at 310 nm was identified in a previous mass-spectrometry analysis (unpublished data of Dr Christelle Andre, Plant and Food Research, New Zealand). mAU: absorbance units in nm; min: retention time.
(B) Quantification of lunularic acid content in *35S*:*Myb14*/*35S*:*BT2* (lines examined were 21, 23 and 24) and *35S*:*Myb14* control plants. Error bars are the standard error of the means, n = 4. Significantly different means are represented by different letters (a and b).

5.4 Discussion

The function of the Mp*BT2* candidate gene in Marchantia was initially examined by transforming spores with the Mp*BT2* CRISPR/Cas9 construct. PCR-amplification of the target region (Figure 5.2) and DNA sequencing (Table 5.1 and Figure 5.3) revealed that *bt2* mutants were successfully generated. Plants over-expressing Mp*BT2* were also created using spore transformation and thallus transformation, as evident by PCR-amplification of the target sequence (Figure 5.4).

Based on the results of my study, MpBT2 is not a key regulator of auronidin biosynthesis in Marchantia. Flavonoid analysis of *bt2* mutants and *35S*:*BT2* plants grown under nutrient stress conditions showed them having an auronidin content within the range found for the wild-type male and female plants (Figure 5.6 B). The female wild-type plant contains a segregating quantitative trait locus (QTL) that results in reduced auronidin production (unpublished data of Dr Rubina Jibran, Plant and Food Research, New Zealand). Because spores for transformation were sexually produced from this female parent, the resultant plants segregate for the locus. However, the locus is not sex-linked; Marchantia plants have either a female (U) or male (V) sex chromosome, and it does not map to either of these. Currently, the only way to determine if plants have the low auronidin trait is by using gene mapping, which was not performed in this study. Therefore, it is not known whether the auronidin production of the *bt2* mutants and *35S*:*BT2* plants should be compared to the male or female wild-type plants, so they were compared to both. Thus, plants would have to produce a significantly lower (e.g., the *chiL* mutant) or higher auronidin content than both wild-type plants to conclude that MpBT2 regulates auronidin in the study. In this context, it is reasonable to conclude that MpBT2 does not have a major effect on auronidin content, but a more subtle action could not be excluded.

In functional studies, gene mutants are the most useful for determining gene function because they show the effect of a gene being non-functional. Thus, several lines of loss-of-function mutants targeting the same gene are expected to functional similarly. However, a surprising observation in this study was that the *bt2*/9 mutant had a significantly higher auronidin content than the *bt2*/5 and *bt2*/8 mutants (Figure 5.6 B). Sequencing results clearly showed that knock-out mutants were created (Table 5.1 and Figure 5.3). Thus, the best explanation for this result is that the *bt2*/5 and *bt2*/8 mutants had the low auronidin QTL, so the auronidin content was comparable to the female wild-type, and the *bt2*/9 mutant did not so was comparable to the male wild-type plant.

Because the segregating QTL confounded the analysis, thallus transformation of a single genotype (*35S*:*Myb14*) was also performed so that regenerated plants would not contain the low auronidin

QTL. Flavonoid analysis of *35S*:*Myb14*/*35S*:*BT2* plants showed that Mp*BT2* over-expression did not cause any significant differences in auronidin production (Figure 5.10 B). This reiterates that MpBT2 does not regulate auronidin biosynthesis.

Several studies on apple used a similar approach to my project to determine the function of MdBT2 (An et al., 2020a; Kang et al., 2022; Wang et al., 2018); stable *MdBT2-OX* (over-expression) and *MdBT2-anti* (antisense) lines were generated and characterised for their ability to produce anthocyanins. Plants over-expressing Md*BT2* were consistently found to produce significantly less anthocyanins than the wild-type control, and antisense lines significantly more anthocyanins. Similar effects on auronidin production would have been expected for Mp*BT2* over-expression plants and *bt2* mutants in my study if MpBT2 regulated auronidin biosynthesis, but this was not the case.

Rather than an effect on auronidin production, my results suggest that MpBT2 may indirectly regulate the production of bibenzyls (as indicated by changes in the content of the bibenzyl precursor lunularic acid) and flavones. While the *bt2* mutants had a similar flavone content to the wild-type plants, all three *35S*:*BT2* lines contained a significantly higher flavone content (Figure 5.7 B). However, there were no consistent significant differences in flavone production between the *35S*:*Myb14*/*35S*:*BT2* and *35S*:*Myb14* control lines (Figure 5.11 B). Moreover, all *35S*:*Myb14*/*35S*:*BT2* lines had a significantly lower lunularic acid content than the *35S*:*Myb14* control (Figure 5.12 B), but there were no consistent significant differences in lunularic acid production in the *bt2* mutants or *35S*:*BT2* plants (Figure 5.8 B). These opposing results suggest MpBT2 can influence flavone and bibenzyl content, but it is unclear exactly what the regulatory action of MpBT2 is.

Until now, MpBT2 had not been identified in bryophytes. However, extensive work has investigated the role of BT2 in angiosperms, particularly in apple. MdBT2 is a central regulator in apple that supresses stress-response pathways in the absence of stress by promoting the ubiquitination and degradation of the key regulators involved. A major pathway MdBT2 functions in is anthocyanin biosynthesis. In response to nitrogen deprivation, UV-B light, ABA, wounding and drought stresses, the transcription factors MYB1, bZIP44, WRKY40 and ERF38, respectively, are upregulated to promote anthocyanin accumulation. The function of MdBT2 is inhibited in these conditions. However, in the absence of these stresses, MdBT2 is activated and negatively regulates the transcription factors to prevent anthocyanin production (An et al., 2020b; An et al., 2019a; An et al., 2018; An et al., 2019b; Linde et al., 2020; Wang et al., 2018). MdBT2 also functions in dark/low light conditions to prevent the self-ubiquitination of MdCOP1, an E3 ligase that ubiquitinates MdMYB1 and presumably also MdTCP46 (a transcription factor that upregulates MdMYB1 expression in these

conditions), thereby suppressing anthocyanin biosynthesis (An et al., 2020a; Kang et al., 2022; Li et al., 2012). However, in high light conditions, MdBT2 (and MdCOP1)-mediated ubiquitination and degradation of MdMYB1 and MdTCP46 does not occur, so anthocyanin is produced (An et al., 2020a). These findings are supported by Wang et al. (2018) and Mandadi et al. (2009) also demonstrating that MdBT2 is more highly expressed in the dark than the light.

Functional characterisation of the apple BT2 protein suggests that it has a crucial function in regulating anthocyanin production in response to stress. Whereas, my study indicates that MpBT2 does not regulate the production of auronidin, the equivalent red pigment in Marchantia. This suggests that the function of BT2 proteins in regulating flavonoids is not conserved in Marchantia. Although the production of auronidin is induced by stress and regulated by R2R3-MYBs similarly to anthocyanins, auronidin not being regulated by MpBT2 could be because its different chemical structure places it in a distinct class of flavonoids from anthocyanins that may be regulated differently. Therefore, the regulation of flavonoids by BT2 proteins could have been acquired with the evolution of anthocyanins in land plants, after the divergence of bryophytes. Despite MpBT2 not regulating auronidin production, it could still be regulating stress response. RT-qPCR analysis has demonstrated that the apple Md*BT2* is expressed in the presence of nitrogen but not under nitrogen deprivation conditions (Wang et al., 2018). Comparatively, unpublished (Appendix 9.1; Dr Yanfei Zhou) and published (Marchantia eFP browser, University of Toronto, Canada; Appendix 9.2) RNA-sequencing differential analysis has shown that the expression of Mp*BT2* is constitutive in Marchantia, meaning that it is not affected by nitrogen status. However, it is possible that MpBT2 is regulated by these conditions post-translationally. Thus, more work is required to determine whether MpBT2 is regulating plant response to nitrogen or other stress conditions.

In this study, the *M. polymorpha* Md*BT2* candidate gene was functionally characterised by generating *bt2* mutants and over-expressing Md*BT2* in different genetic backgrounds. Flavonoid analysis indicated that MpBT2 does not regulate auronidin biosynthesis, which suggests that some functions of BT2 proteins may not be conserved among land plants.

Chapter 6

General discussion and future directions

This project aimed to improve our understanding of how conserved the flavonoid biosynthetic pathway of Marchantia is with the well-characterised pathway of flowering plants. The particular focus was on determining the genetic requirements of the stress-induced biosynthesis of flavones and regulation of auronidin, the two key flavonoids produced by Marchantia. This research question was addressed by generating transgenic plants, including using CRISPR/Cas9-based technologies to create gene-edited mutant lines. However, a practical tissue culture regeneration efficiency could not be achieved by replicating published Marchantia spore transformation protocols and many chimeric mutant plants resulted, which added complexity to screening and handling of candidate transgenic plants. Thus, the first specific aim of the project was to optimise the efficiency of the CRISPR/Cas9 spore transformation system of Marchantia.

My research achieved significant improvements to the existing methods for Marchantia spore transformation and gene editing. Results presented in this thesis suggest that the addition of 200 µM acetosyringone to the pre-culture media during spore transformation enables the production of sufficient transgenic plants for downstream applications. Although the resulting system was adequate for my purposes, if desired, a future study could test a narrower range of acetosyringone concentrations (e.g., 100, 200, 300 and 400 µM) to determine if the transformation efficiency could be further optimised. Additionally, providing spores with a 37°C/16 h heat-shock at the beginning of the selection stage of transformation increased the Cas9 editing efficiency. This generated more mutant plants, fewer of which were chimeric, within the first two weeks of regeneration when transformed sporelings are usually taken through further analysis. Thus, this component of my study contributed to establishing a Marchantia CRISPR/Cas9 spore transformation protocol that is not only highly efficient but is also accessible because it uses the commonly available GV3101 *A*. *tumefaciens* strain.

Because chimeras are common in Marchantia, scoring plants based on the size of mutant sectors they contained (which is determined by the amount of Cas9 activity) was necessary to accurately represent Cas9 editing efficiency, rather than simply noting which plants had a mutant phenotype. However, using this method made it hard to conclude which treatments produced a higher Cas9 editing efficiency because reporting a single quantifiable percentage did not take into account the variable size of mutant sectors. Additionally, although the heat-shock improved the experimental

outcomes, a larger replicated experiment using several spore transformations would have been preferable to conclude the effects of heat-shock. However, this would be a large undertaking (and was impractical within the confines of my study). Thus, to enable generation of a larger, conclusive dataset, a screening method more suited to high throughput analysis should be developed. For example, because the mutant phenotype (pale green tissue) is distinctive, more plants could be grown per tub and an image analysis software that is able to distinguish and score different colours (e.g., ZEISS ZEN Microscopy Software) could be used to calculate the percentage of wild-type and mutant tissue per tub. This method would be less onerous and provide a quantifiable result while accounting for the size of mutant sectors.

The second aim of this project was to characterise the function of the Mp*FNSI* candidate gene hypothesised to be required for flavone biosynthesis. Results presented in this thesis clearly demonstrate that the hypothesis was correct; all three CRISPR/Cas9 *fnsI* mutants analysed lacked flavones and expressing Mp*FNSI* in *N. tabacum* conferred flavone accumulation, providing both the loss- and gain-of-function evidence that is generally accepted as proof of gene function. Flavones provide protection against UV-B light (Agati & Tattini, 2010; Clayton et al., 2018; Winkel-Shirley, 2002). Thus, future work could investigate whether *fnsI* mutants are susceptible to UV-B light damage. I tried to replicate the daily low fluence (DLF) UV-B conditions used for a similar purpose in Clayton et al. (2018). However, plants became more stressed than expected, so no differences in phenotype were observed between the *fnsI* mutant and wild-type plants. The reason for this severe phenotype is unknown, as the UV-B intensity used (1.24 W m^{-2}) was similar to that of other Marchantia UV-B studies (~1.3 W m^{-2}) (Kondou et al., 2019; Soriano et al., 2019). Thus, repeated experiments should ensure that the UV-B light spectral readings are accurate, and the intensity reduced if necessary.

M. polymorpha requires MpFNSI for flavone biosynthesis. Wang et al. (2020) and Li et al. (2020a) used *in vitro* enzymatic assays to demonstrate that enzymes in other liverworts (*Marchantia paleacea*, *Marchantia emarginata*, *Plagiochasma appendiculatum* and *Conocephalum japonicum*), a moss (*P. patens*) and a lycophyte (*S. moellendorffii*) have at least partial FNSI function, being able to synthesise flavones from various substrates. This suggests that the function of FNSI is conserved within the bryophyte plant lineage. It would be useful to confirm the function of these genes by CRISPR/Cas9 characterisation. However, only some of these species have published transformation and CRISPR/Cas9-mediated mutation systems, and to develop those for additional species would be a large undertaking.

Given both bryophytes and angiosperms produce flavones, it is likely that using flavones to protect plants against UV-B light is a conserved characteristic that was probably gained early during land colonisation to facilitate plant survival. The importance of this function for liverworts is supported by Clayton et al. (2018) demonstrating that Marchantia plants lacking flavones are more susceptible to damage from UV-B light. While identifying that Marchantia has an FNSI enzyme required for flavone biosynthesis is a significant finding, the origin and evolution of flavones and FNS enzymes in land plants remains complicated because angiosperms can use either FNSI or FNSII for flavone biosynthesis. FNSI itself may have evolved on more than one occasion, and some plants instead produce flavonols for a similar function (Gebhardt et al., 2005; Pucker & Iorizzo, 2023; Yonekura-Sakakibara et al., 2019). In general, there has been insufficient sampling of different lineages to present an overall theory for the evolution of flavone biosynthesis in land plants. However, the bryophyte and angiosperm FNSI enzymes are found in the same 2-OGD phylogenetic clade (DOXC28) (Bowman et al., 2017; Li et al., 2020a), so it is possible that they evolved from the same ancestral gene. Further data are required to support this proposal, such as comparing the exon/intron structure of the gene models and, when lineages are sufficiently closely related, determining the microsynteny (the conservation of gene order) around the FNS genes. Based on current knowledge, the occurrence of FNSI in both bryophytes and angiosperms is best explained by the common ancestor of these lineages containing a progenitor 2-OGD gene that evolved into *FNSI* multiple, independent times throughout land plant evolution. Alternatively, the ancestral gene could have been *FNSI* (if flavones were present in these plants) but given how few extant plants contain an *FNSI* gene, extensive loss of *FNSI* would have had to occur for this to be the case. How the origin of FNSII fits into this evolutionary model, as well as how it relates to the evolution of F3H for flavonol production, is still an open question. However, the absence of FNSII in bryophytes suggests that it evolved later than the bryophyte FNSI enzymes and it is likely to have an independent origin to FNSI as it is part of a different enzyme family (CYP450s) (Du et al., 2016; Yonekura-Sakakibara et al., 2019). Nevertheless, the fact that FNS enzymes have evolved several times throughout land plant evolution signifies that the function of flavones in providing UV-B protection remained important during evolution towards angiosperms, not only in early diverging land plants.

The last aim of the project was to characterise the function of the Mp*BT2* candidate gene in Marchantia. It was hypothesised that MpBT2 negatively regulates auronidin production. However, results presented in this thesis suggest that this is not the case as, under nutrient stress conditions, the CRISPR/Cas9 *bt2* mutants and plants over-expressing Mp*BT2* in a wild-type background contained a similar auronidin content to the wild-type control plants. This aspect of the study was complicated by the fact that the spores used for transformation segregate for auronidin content (a

low auronidin QTL originated in the female wild-type parent). To eradicate this limitation, Mp*BT2* was over-expressed in a single *35S*:*Myb14* background, which accumulates high amounts of auronidins, using thallus transformation. These *35S*:*Myb14/35S*:*BT2* plants did not contain significantly different auronidin content to the *35S*:*Myb14* control plants, which again indicates MpBT2 does not regulate auronidin production. This thesis did not determine whether Mp*BT2* over-expression produced active MpBT2 protein. While this could explain the lack of differences in auronidin content observed, there is no reason to suspect that MpBT2 protein was not produced as the *35S* promoter is effective in Marchantia and the expression system consistently delivers active transgenes. However, future work could confirm that the Mp*BT2* over-expression lines have increased mRNA transcript and protein abundance by using qPCR and western blots, respectively.

The Mp*BT2* candidate gene was investigated in this study to determine if it had a similar function to angiosperm BT2 proteins in regulating the production of stress-induced flavonoids. In apple, MdBT2 negatively regulates anthocyanin production in the presence of nitrogen by interacting with and promoting the degradation of MdMYB1, which is required for anthocyanin production (Wang et al., 2018). Whereas during nitrogen deprivation, MdGRF11 (a 14-3-3 phosphopeptide-binding protein) destabilises MdBT2, thereby promoting its self-ubiquitination and degradation that allows MdMYB1 to induce anthocyanin biosynthesis (Ren et al., 2021a). Although MpBT2 does not regulate auronidin biosynthesis, auronidin is an important flavonoid in Marchantia that provides tolerance against several abiotic stresses, such as excess white light and nutrient deprivation (Albert et al., 2018; Berland et al., 2019). Therefore, some abiotic stress signalling pathways are still expected to be conserved across land plants. In particular, auronidin biosynthesis is regulated by an R2R3-MYB (MYB14), so is controlled similarly to flavonoids in angiosperms that are predominantly regulated by different subgroups of R2R3-MYBs (e.g., the SG6 protein MdMYB1) (Cheynier et al., 2013; Feller et al., 2011). Albeit there are only two R2R3-MYBs known to regulate the phenylpropanoid pathway in Marchantia, the other being MYB02 (Albert et al., 2018; Berland et al., 2019; Kubo et al., 2018). Nevertheless, the use of R2R3-MYBs for regulating stress-induced flavonoid production is conserved among land plants. However, being unable to elucidate the function of MpBT2 means that none of the regulatory factors upstream of these MYBs have been identified.

Angiosperm BT2 proteins also have nitrogen-responsive functions unrelated to regulating anthocyanin accumulation. In apple, MdBT2 reduces nitrogen use efficiency under low nitrogen conditions by promoting the ubiquitination and degradation of the MYB88 and MYB124 transcription factors that positively regulate nitrogen transporters (NRTs) to increase nitrogen usage (Zhang et al., 2021). Conversely, MdBT2 promotes the ubiquitination of MdRGL3a (a DELLA protein)

under high nitrogen conditions to increase plant growth (Ren et al., 2021b), probably by increasing nitrogen use efficiency. The *Arabidopsis* BT2 (AtBT2) functions similarly in response to nitrogen status (Araus et al., 2016; Sato et al., 2017). However, Mandadi et al. (2009) additionally found that At*BT2* expression was reduced by high sugar (glucose), opposite to the effect of nitrate, and At*BT2* over-expression increased germination rate by supressing the growth inhibition response to high sugar. This indicates that AtBT2 may regulate both nitrate and sugar signalling, probably by sensing the endogenous C:N ratio and controlling plant growth accordingly (Mandadi et al., 2009). Again, this is analogous to the induction of auronidin biosynthesis by a high C:N ratio (high carbon, low nitrogen), further supporting the initial rationale for the study. Similarly to angiosperm BT2 proteins, MpBT2 could be regulating nitrogen-responsive processes such as nitrogen uptake and plant growth or responding to changes in the C:N ratio in the plant. Such regulation may re-allocate carbon in the plant, and could have caused the significant changes in auronidin, flavone and lunularic acid accumulation observed in my study that could not be explained by direct regulation from MpBT2.

Since MpBT2 does not appear to regulate auronidin biosynthesis, future work should investigate what its function may be. RNA-seq differential analysis suggests that Mp*BT2* gene expression is not regulated by nitrogen status. However, it is possible that the gene is only regulated post-translationally under these conditions, so a function of MpBT2 in nitrogen signalling should not be ruled out. This could be investigated by using a western blot to quantify the abundance of the MpBT2 protein under different nutrient status conditions. Future work should also examine whether MpBT2 regulates any other abiotic or biotic stress responses, because angiosperm BT2 proteins also regulate stress conditions such as excess white and UV-B light, drought, wounding and the presence of ABA (An et al., 2020c). Additionally, there are four other classes of BT proteins in angiosperms (BT1, BT3, BT4 and BT5) that have stress-related functions, including in responding to drought, salinity and hydrogen peroxide, regulating the plant hormones ethylene, salicylic acid and jasmonic acid that are important in stress signalling, and providing resistance to plant diseases and pathogens (Du & Poovaiah, 2004; Fujita et al., 2007; Zheng et al., 2019; Zhou et al., 2022; Zhou et al., 2020). Given MpBT2 does not have the conserved function of angiosperm BT2 proteins in regulating flavonoid production, and was the only homolog of angiosperm BT proteins, it is possible that an angiosperm BT protein other than BT2 could be a better match for the function of MpBT2. This emphasises the importance of determining the actual function of MpBT2. Several methods could be used to investigate this. Performing RNA-sequencing analysis on the *bt2* mutants and Mp*BT2* over-expression plants may identify the pathways affected by changes in Mp*BT2* expression. Additionally, bindings assays such as the Yeast Two-hybrid (Y2H) and pulldown could be performed to investigate

which proteins MpBT2 may interact with, as it has a BTB domain that is known to mediate protein-protein binding (Figueroa et al., 2005).

The results presented in this thesis have improved our understanding of the conservation of the flavonoid biosynthetic pathway between Marchantia and flowering plants. Notably, it has added to the growing evidence for conservation of the pathway branch leading to production of flavones, but has informed that there are distinctions in the production and regulation of red flavonoid pigment pathways. The identification and functional characterisation of MpFNSI confirms that the same pathway operates from the first phenylpropanoid enzyme, PAL, to at least the flavone aglycone in land plants (Berland et al., 2019; Clayton et al., 2018; Davies et al., 2020; Yonekura-Sakakibara et al., 2019). The activation of the flavone pathway through UVR8/HY5 is also conserved (Clayton et al., 2018). However, with the exception of MYB14 (an R2R3-MYB) as the direct activator of auronidin biosynthesis, the fact that MpBT2 does not regulate auronidin biosynthesis means that differences in the regulatory pathways for auronidins and anthocyanins are now established.

This thesis helped to establish a highly efficient Marchantia CRISPR/Cas9 spore transformation protocol that was utilised to characterise the function of candidate flavonoid-related genes in Marchantia. MpFNSI was identified to have a conserved function in flavone biosynthesis and it was established that MpBT2 does not regulate auronidin production. Thus, gene editing was demonstrated to be a powerful tool to improve our understanding of the conservation of flavonoid biosynthesis and regulation in Marchantia. This has provided invaluable insight into what characteristics plants may have gained to facilitate land colonisation, but also into how plants may respond to a changing climate in the future (Agati et al., 2012).

References

Agati, G., Azzarello, E., Pollastri, S., & Tattini, M. (2012). Flavonoids as antioxidants in plants: Location and functional significance. *Plant Science*, *196*, 67-76. https://doi.org/10.1016/j.plantsci.2012.07.014

Agati, G., & Tattini, M. (2010). Multiple functional roles of flavonoids in photoprotection. *New Phytologist*, *186*(4), 786-793. https://doi.org/10.1111/j.1469-8137.2010.03269.x

Aggarwal, D., Kumar, A., & Sudhakara Reddy, M. (2011). *Agrobacterium tumefaciens* mediated genetic transformation of selected elite clone(s) of *Eucalyptus tereticornis*. *Acta Physiologiae Plantarum*, *33*(5), 1603-1611. https://doi.org/10.1007/s11738-010-0695-3

Albert, N. W., Lafferty, D. J., Moss, S. M. A., & Davies, K. M. (2022). Flavonoids – flowers, fruit, forage and the future. *Journal of the Royal Society of New Zealand*, *53*(3), 304-331. https://doi.org/10.1080/03036758.2022.2034654

Albert, N. W., Thrimawithana, A. H., McGhie, T. K., Clayton, W. A., Deroles, S. C., Schwinn, K. E., . . . Davies, K. M. (2018). Genetic analysis of the liverwort *Marchantia polymorpha* reveals that R2R3MYB activation of flavonoid production in response to abiotic stress is an ancient character in land plants. *New Phytologist*, *218*(2), 554-566. https://doi.org/10.1111/nph.15002

Almagro, L., Gómez Ros, L. V., Belchi-Navarro, S., Bru, R., Ros Barceló, A., & Pedreño, M. A. (2008). Class III peroxidases in plant defence reactions. *Journal of Experimental Botany*, *60*(2), 377-390. https://doi.org/10.1093/jxb/ern277

Alseekh, S., Perez de Souza, L., Benina, M., & Fernie, A. R. (2020). The style and substance of plant flavonoid decoration; towards defining both structure and function. *Phytochemistry*, *174*, 112347. https://doi.org/10.1016/j.phytochem.2020.112347

An, J.-P., Liu, Y.-J., Zhang, X.-W., Bi, S.-Q., Wang, X.-F., You, C.-X., & Hao, Y.-J. (2020a). Dynamic regulation of anthocyanin biosynthesis at different light intensities by the BT2-TCP46-MYB1 module in apple. *Journal of Experimental Botany*, *71*(10), 3094-3109. https://doi.org/10.1093/jxb/eraa056

An, J.-P., Zhang, X.-W., Bi, S.-Q., You, C.-X., Wang, X.-F., & Hao, Y.-J. (2020b). The ERF transcription factor MdERF38 promotes drought stress-induced anthocyanin biosynthesis in apple. *The Plant Journal*, *101*(3), 573-589. https://doi.org/10.1111/tpj.14555

An, J. P., Wang, X. F., & Hao, Y. J. (2020c). BTB/TAZ protein MdBT2 integrates multiple hormonal and environmental signals to regulate anthocyanin biosynthesis in apple. *Journal of Integrative Plant Biology*, *62*(11), 1643-1646. https://doi.org/10.1111/jipb.12940

An, J. P., Wang, X. F., Zhang, X. W., Bi, S. Q., You, C. X., & Hao, Y. J. (2019a). MdBBX22 regulates UV-B-induced anthocyanin biosynthesis through regulating the function of MdHY5 and is targeted by MdBT2 for 26S proteasome-mediated degradation. *Plant Biotechnology Journal*, *17*(12), 2231-2233. https://doi.org/10.1111/pbi.13196

An, J. P., Yao, J. F., Xu, R. R., You, C. X., Wang, X. F., & Hao, Y. J. (2018). Apple bZIP transcription factor MdbZIP44 regulates abscisic acid-promoted anthocyanin accumulation. *Plant, Cell and Environment*, *41*(11), 2678-2692. https://doi.org/10.1111/pce.13393

An, J. P., Zhang, X. W., You, C. X., Bi, S. Q., Wang, X. F., & Hao, Y. J. (2019b). MdWRKY40 promotes wounding-induced anthocyanin biosynthesis in association with MdMYB1 and undergoes MdBT2-mediated degradation. *New Phytologist*, *224*(1), 380-395. https://doi.org/10.1111/nph.16008

Arai, H., Yanagiura, K., Toyama, Y., & Morohashi, K. (2019). Genome-wide analysis of MpBHLH12, a IIIf basic helix-loop-helix transcription factor of *Marchantia polymorpha*. *Journal of Plant Research*, *132*(2), 197-209. https://doi.org/10.1007/s10265-019-01095-w

Araji, S., Grammer, T. A., Gertzen, R., Anderson, S. D., Mikulic-Petkovsek, M., Veberic, R., . . . Escobar, M. A. (2014). Novel roles for the polyphenol oxidase enzyme in secondary metabolism and the regulation of cell death in walnut. *Plant Physiology*, *164*(3), 1191-1203. https://doi.org/10.1104/pp.113.228593

Araus, V., Vidal, E. A., Puelma, T., Alamos, S., Mieulet, D., Guiderdoni, E., & Gutiérrez, R. A. (2016). Members of BTB gene family of scaffold proteins suppress nitrate uptake and nitrogen use efficiency. *Plant Physiology*, *171*(2), 1523-1532. https://doi.org/10.1104/pp.15.01731

Arteaga-Vazquez, M. A. (2016). Land plant evolution: Listen to your elders. *Current Biology*, *26*(1), R26-R29. https://doi.org/10.1016/j.cub.2015.12.001

Ashby, A. M., Watson, M. D., Loake, G. J., & Shaw, C. H. (1988). Ti plasmid-specified chemotaxis of *Agrobacterium tumefaciens* C58C1 toward *vir*-inducing phenolic compounds and soluble factors from monocotyledonous and dicotyledonous plants. *Journal of Bacteriology*, *170*(9), 4181-4187. https://doi.org/10.1128/jb.170.9.4181-4187.1988

Ashby, A. M., Watson, M. D., & Shaw, C. H. (1987). A Ti-plasmid determined function is responsible for chemotaxis of *Agrobacterium tumefaciens* towards the plant wound product acetosyringone. *FEMS Microbiology Letters*, *41*(2), 189-192. https://doi.org/10.1111/j.1574-6968.1987.tb02194.x

Baudry, A., Heim, M. A., Dubreucq, B., Caboche, M., Weisshaar, B., & Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of *Banyuls* and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *The Plant Journal*, *39*(3), 366-380. https://doi.org/10.1111/j.1365-313X.2004.02138.x

Becker, B., & Marin, B. (2009). Streptophyte algae and the origin of embryophytes. *Annals of Botany*, *103*(7), 999-1004. https://doi.org/10.1093/aob/mcp044

Bennici, A. (2008). Origin and early evolution of land plants: Problems and considerations. *Communicative & Integrative Biology*, *1*(2), 212-218. https://doi.org/10.4161/cib.1.2.6987

Berland, H., Albert, N. W., Stavland, A., Jordheim, M., McGhie, T. K., Zhou, Y. F., . . . Andersen, O. M. (2019). Auronidins are a previously unreported class of flavonoid pigments that challenges when anthocyanin biosynthesis evolved in plants. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(40), 20232-20239. https://doi.org/10.1073/pnas.1912741116

Binkert, M., Kozma-Bognár, L., Terecskei, K., De Veylder, L., Nagy, F., & Ulm, R. (2014). UV-B-responsive association of the *Arabidopsis* bZIP transcription factor ELONGATED HYPOCOTYL5 with target genes, including its own promoter. *The Plant Cell*, *26*(10), 4200-4213. https://doi.org/10.1105/tpc.114.130716

Blomme, J., Develtere, W., Köse, A., Arraiza Ribera, J., Brugmans, C., Jaraba-Wallace, J., . . . Jacobs, T. (2022). The heat is on: a simple method to increase genome editing efficiency in plants. *BMC Plant Biology*, *22*(1), 142. https://doi.org/10.1186/s12870-022-03519-7

Boase, M. R., Bradley, J. M., & Borst, N. K. (1998). An improved method for transformation of regal pelargonium (*Pelargonium X domesticum* Dubonnet) by *Agrobacterium tumefaciens*. *Plant Science*, *139*(1), 59-69. https://doi.org/10.1016/S0168-9452(98)00177-0

Borevitz, J. O., Xia, Y., Blount, J., Dixon, R. A., & Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *The Plant Cell*, *12*(12), 2383-2393. https://doi.org/10.1105/tpc.12.12.2383

Boucherle, B., Peuchmaur, M., Boumendjel, A., & Haudecoeur, R. (2017). Occurrences, biosynthesis and properties of aurones as high-end evolutionary products. *Phytochemistry*, *142*, 92-111. https://doi.org/10.1016/j.phytochem.2017.06.017

Bowman, J. L. (2022). The origin of a land flora. *Nature Plants*, *8*(12), 1352-1369. https://doi.org/10.1038/s41477-022-01283-y

Bowman, J. L., Arteaga-Vazquez, M., Berger, F., Briginshaw, L. N., Carella, P., Aguilar-Cruz, A., . . . Zachgo, S. (2022). The renaissance and enlightenment of Marchantia as a model system. *The Plant Cell*, *34*(10), 3512-3542. https://doi.org/10.1093/plcell/koac219

Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., . . . Schmutz, J. (2017). Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell*, *171*(2), 287-304. https://doi.org/10.1016/j.cell.2017.09.030

Campanella, J. J., Smalley, J. V., & Dempsey, M. E. (2014). A phylogenetic examination of the primary anthocyanin production pathway of the Plantae. *Botanical Studies*, *55*(1), 10. https://doi.org/10.1186/1999-3110-55-10

Campbell, E. O., Markham, K. R., Moore, N. A., Porter, L. J., & Wallace, J. W. (1979). Taxonomic and phylogenetic implications of comparative flavonoid chemistry of species in the family Marchantiaceae. *The Journal of the Hattori Botanical Laboratory*, *45*, 185-199. https://doi.org/10.18968/jhbl.45.0_185

Cheynier, V., Comte, G., Davies, K. M., Lattanzio, V., & Martens, S. (2013). Plant phenolics: recent advances on their biosynthesis, genetics, and ecophysiology. *Plant Physiology and Biochemistry*, *72*, 1-20. https://doi.org/10.1016/j.plaphy.2013.05.009

Chiyoda, S., Ishizaki, K., Kataoka, H., Yamato, K. T., & Kohchi, T. (2008). Direct transformation of the liverwort *Marchantia polymorpha* L. by particle bombardment using immature thalli developing from spores. *Plant Cell Reports*, *27*(9), 1467-1473. https://doi.org/10.1007/s00299-008-0570-5

Clayton, W. A., Albert, N. W., Thrimawithana, A. H., McGhie, T. K., Deroles, S. C., Schwinn, K. E., . . . Davies, K. M. (2018). UVR8-mediated induction of flavonoid biosynthesis for UVB tolerance is conserved between the liverwort *Marchantia polymorpha* and flowering plants. *Plant Journal*, *96*(3), 503-517. https://doi.org/10.1111/tpj.14044

Cominelli, E., Gusmaroli, G., Allegra, D., Galbiati, M., Wade, H. K., Jenkins, G. I., & Tonelli, C. (2008). Expression analysis of anthocyanin regulatory genes in response to different light qualities in *Arabidopsis thaliana*. *Journal of Plant Physiology*, *165*(8), 886-894. https://doi.org/10.1016/j.jplph.2007.06.010

Condamine, F. L., Silvestro, D., Koppelhus, E. B., & Antonelli, A. (2020). The rise of angiosperms pushed conifers to decline during global cooling. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(46), 28867-28875. https://doi.org/10.1073/pnas.2005571117

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., . . . Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*(6121), 819-823. https://doi.org/10.1126/science.1231143

Crow, K. D., & Wagner, G. P. (2005). What is the role of genome duplication in the evolution of complexity and diversity? *Molecular Biology and Evolution*, *23*(5), 887-892. https://doi.org/10.1093/molbev/msj083

Dang, Y., Jia, G., Choi, J., Ma, H., Anaya, E., Ye, C., . . . Wu, H. (2015). Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biology*, *16*(1), 280. https://doi.org/10.1186/s13059-015-0846-3

Davies, K., Jibran, R., Zhou, Y., Albert, N., Brummell, D., Jordan, B., . . . Schwinn, K. (2020). The evolution of flavonoid biosynthesis: A bryophyte perspective. *Frontiers in Plant Science*, *11*, 7. https://doi.org/10.3389/fpls.2020.00007

Davies, K. M., Albert, N. W., Zhou, Y., & Schwinn, K. E. (2018). Functions of flavonoid and betalain pigments in abiotic stress tolerance in plants. In J. A. Roberts (Ed.), *Annual Plant Reviews online* (pp. 21-62). https://doi.org/10.1002/9781119312994.apr0604

Davies, K. M., Landi, M., van Klink, J. W., Schwinn, K. E., Brummell, D. A., Albert, N. W., . . . Bowman, J. L. (2022). Evolution and function of red pigmentation in land plants. *Annals of Botany*, *130*(5), 613-636. https://doi.org/10.1093/aob/mcac109

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., . . . Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602-607. https://doi.org/10.1038/nature09886

Demirer, G. S., Zhang, H., Matos, J. L., Goh, N., Cunningham, F., Sung, Y., . . . Landry, M. P. (2018). High aspect ratio nanomaterials enable delivery of functional genetic material without DNA integration in mature plants. *bioRxiv*, 179549. https://doi.org/10.1101/179549

Du, H., Ran, F., Dong, H.-L., Wen, J., Li, J.-N., & Liang, Z. (2016). Genome-wide analysis, classification, evolution, and expression analysis of the cytochrome P450 93 family in land plants. *PLOS ONE*, *11*(10), e0165020. https://doi.org/10.1371/journal.pone.0165020

Du, L., & Poovaiah, B. W. (2004). A novel family of Ca2+/calmodulin-binding proteins involved in transcriptional regulation: Interaction with fsh/Ring3 class transcription activators. *Plant Molecular Biology*, *54*(4), 549-569. https://doi.org/10.1023/B:PLAN.0000038269.98972.bb

Dutt, M., & Grosser, J. W. (2009). Evaluation of parameters affecting *Agrobacterium*-mediated transformation of citrus. *Plant Cell, Tissue, and Organ Culture*, *98*(3), 331-340. https://doi.org/10.1007/s11240-009-9567-1

Escudero, J., & Hohn, B. (1997). Transfer and integration of T-DNA without cell injury in the host plant. *Plant Cell*, *9*(12), 2135-2142. https://doi.org/10.1105/tpc.9.12.2135

Farag, M. A., Deavours, B. E., de Fatima, A., Naoumkina, M., Dixon, R. A., & Sumner, L. W. (2009). Integrated metabolite and transcript profiling identify a biosynthetic mechanism for hispidol in *Medicago truncatula* cell cultures. *Plant Physiology*, *151*(3), 1096-1113. https://doi.org/10.1104/pp.109.141481

Feller, A., Machemer, K., Braun, E. L., & Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant Journal*, *66*(1), 94-116. https://doi.org/10.1111/j.1365-313X.2010.04459.x

Figueroa, P., Gusmaroli, G., Serino, G., Habashi, J., Ma, L., Shen, Y., . . . Deng, X. W. (2005). *Arabidopsis* has two redundant Cullin3 proteins that are essential for embryo development and that interact with RBX1 and BTB proteins to form multisubunit E3 ubiquitin ligase complexes in vivo. *Plant Cell*, *17*(4), 1180-1195. https://doi.org/10.1105/tpc.105.031989

Fujita, M., Mizukado, S., Fujita, Y., Ichikawa, T., Nakazawa, M., Seki, M., . . . Shinozaki, K. (2007). Identification of stress-tolerance-related transcription-factor genes via mini-scale Full-length cDNA Over-eXpressor (FOX) gene hunting system. *Biochemical and Biophysical Research Communications*, *364*(2), 250-257. https://doi.org/10.1016/j.bbrc.2007.09.124

Furudate, H., Manabe, M., Oshikiri, H., Matsushita, A., Watanabe, B., Waki, T., . . . Takanashi, K. (2023). A polyphenol oxidase catalyzes aurone synthesis in *Marchantia polymorpha*. *Plant and Cell Physiology*, *64*(6), 637-645. https://doi.org/10.1093/pcp/pcad024

Gao, N. J., Al-Bassam, M. M., Poudel, S., Wozniak, J. M., Gonzalez, D. J., Olson, J., . . . Valderrama, J. A. (2019). Functional and proteomic analysis of *Streptococcus pyogenes* virulence upon loss of its native Cas9 nuclease. *Frontiers in Microbiology*, *10*(1967). https://doi.org/10.3389/fmicb.2019.01967

Garwood, R. J., & Edgecombe, G. D. (2011). Early terrestrial animals, evolution, and uncertainty. *Evolution: Education and Outreach*, *4*(3), 489-501. https://doi.org/10.1007/s12052-011-0357-y

Gebhardt, Y., Witte, S., Forkmann, G., Lukačin, R., Matern, U., & Martens, S. (2005). Molecular evolution of flavonoid dioxygenases in the family Apiaceae. *Phytochemistry*, *66*(11), 1273-1284. https://doi.org/10.1016/j.phytochem.2005.03.030

Gelvin, S. B. (2003). *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiology and Molecular Biology Reviews*, *67*(1), 16-37. https://doi.org/10.1128/mmbr.67.1.16-37.2003

Gera, K., & McIver, K. S. (2013). Laboratory growth and maintenance of *Streptococcus pyogenes* (the Group A *Streptococcus*, GAS). *Current Protocols in Microbiology*, *30*, 9d.2.1-9d.2.13. https://doi.org/10.1002/9780471729259.mc09d02s30

Gonzalez, A., Zhao, M., Leavitt, J. M., & Lloyd, A. M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis*

seedlings. *The Plant Journal*, *53*(5), 814-827. https://doi.org/10.1111/j.1365-313X.2007.03373.x

Gould, K. S., McKelvie, J., & Markham, K. R. (2002). Do anthocyanins function as antioxidants in leaves? Imaging of H_2O_2 in red and green leaves after mechanical injury. *Plant, Cell & Environment*, *25*(10), 1261-1269. https://doi.org/10.1046/j.1365-3040.2002.00905.x

Gruber, H., Heijde, M., Heller, W., Albert, A., Seidlitz, H. K., & Ulm, R. (2010). Negative feedback regulation of UV-B-induced photomorphogenesis and stress acclimation in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(46), 20132-20137. https://doi.org/10.1073/pnas.0914532107

Guivarc'h, A., Caissard, J. C., Brown, S., Marie, D., Dewitte, W., Van Onckelen, H., & Chriqui, D. (1993). Localization of target cells and improvement of *Agrobacterium*-mediated transformation efficiency by direct acetosyringone pretreatment of carrot root discs. *Protoplasma*, *174*(1), 10-18. https://doi.org/10.1007/BF01404037

Guo, T., Feng, Y.-L., Xiao, J.-J., Liu, Q., Sun, X.-N., Xiang, J.-F., . . . Xie, A.-Y. (2018). Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. *Genome Biology*, *19*(1), 170. https://doi.org/10.1186/s13059-018-1518-x

Hagolani, P. F., Zimm, R., Vroomans, R., & Salazar-Ciudad, I. (2021). On the evolution and development of morphological complexity: A view from gene regulatory networks. *PLOS Computational Biology*, *17*(2), e1008570. https://doi.org/10.1371/journal.pcbi.1008570

Harfe, B. D., & Jinks-Robertson, S. (1999). Removal of frameshift intermediates by mismatch repair proteins in *Saccharomyces cerevisiae*. *Molecular and Cellullar Biology*, *19*(7), 4766-4773. https://doi.org/10.1128/mcb.19.7.4766

Harris, B. J., Harrison, C. J., Hetherington, A. M., & Williams, T. A. (2020). Phylogenomic evidence for the monophyly of bryophytes and the reductive evolution of stomata. *Current Biology*, *30*(11), 2001-2012.e2002. https://doi.org/10.1016/j.cub.2020.03.048

He, Y., Jones, H. D., Chen, S., Chen, X. M., Wang, D. W., Li, K. X., . . . Xia, L. Q. (2010). *Agrobacterium*-mediated transformation of durum wheat (*Triticum turgidum* L. var. *durum* cv Stewart) with improved efficiency. *Journal of Experimental Botany*, *61*(6), 1567-1581. https://doi.org/10.1093/jxb/erq035

Hoekema, A., Hirsch, P. R., Hooykaas, P. J. J., & Schilperoort, R. A. (1983). A binary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature*, *303*(5913), 179-180. https://doi.org/10.1038/303179a0

Hong, L., Qian, Q., Tang, D., Wang, K., Li, M., & Cheng, Z. (2012). A mutation in the rice *Chalcone Isomerase* gene causes the *golden hull and internode 1 phenotype*. *Planta*, *236*(1), 141-151. https://doi.org/10.1007/s00425-012-1598-x

Hooghvorst, I., López-Cristoffanini, C., & Nogués, S. (2019). Efficient knockout of *Phytoene Desaturase* gene using CRISPR/Cas9 in melon. *Scientific Reports*, *9*(1), 17077. https://doi.org/10.1038/s41598-019-53710-4

Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, *157*(6), 1262-1278. https://doi.org/10.1016/j.cell.2014.05.010

Hurtado, N. H., Morales, A. L., González-Miret, M. L., Escudero-Gilete, M. L., & Heredia, F. J. (2009). Colour, pH stability and antioxidant activity of anthocyanin rutinosides isolated from tamarillo fruit (*Solanum betaceum* Cav.). *Food Chemistry*, *117*(1), 88-93. https://doi.org/10.1016/j.foodchem.2009.03.081

Inoue, H., Nojima, H., & Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene*, *96*(1), 23-28. https://doi.org/10.1016/0378-1119(90)90336-p

Ishizaki, K. (2017). Evolution of land plants: insights from molecular studies on basal lineages. *Bioscience, Biotechnology, and Biochemistry*, *81*(1), 73-80. https://doi.org/10.1080/09168451.2016.1224641

Ishizaki, K., Chiyoda, S., Yamato, K. T., & Kohchi, T. (2008). *Agrobacterium*-mediated transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for plant biology. *Plant and Cell Physiology*, *49*(7), 1084-1091. https://doi.org/10.1093/pcp/pcn085

Jarial, R., Shard, A., Thakur, S., Sakinah, M., Zularisam, A. W., Rezania, S., . . . Singh, L. (2018). Characterization of flavonoids from fern *Cheilanthes tenuifolia* and evaluation of antioxidant, antimicrobial and anticancer activities. *Journal of King Saud University - Science*, *30*(4), 425-432. https://doi.org/10.1016/j.jksus.2017.04.007

Jiang, N., Doseff, A. I., & Grotewold, E. (2016). Flavones: From biosynthesis to health benefits. *Plants*, *5*(2), 27. https://www.mdpi.com/2223-7747/5/2/27

Jiang, W., Yin, Q., Wu, R., Zheng, G., Liu, J., Dixon, R. A., & Pang, Y. (2015). Role of a chalcone isomerase-like protein in flavonoid biosynthesis in *Arabidopsis thaliana*. *Journal of Experimental Botany*, *66*(22), 7165-7179. https://doi.org/10.1093/jxb/erv413

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., & Weeks, D. P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Research*, *41*(20), e188-e188. https://doi.org/10.1093/nar/gkt780

Jin, S. G., Prusti, R. K., Roitsch, T., Ankenbauer, R. G., & Nester, E. W. (1990). Phosphorylation of the VirG protein of *Agrobacterium tumefaciens* by the autophosphorylated VirA protein: essential role in biological activity of VirG. *Journal of Bacteriology*, *172*(9), 4945-4950. https://doi.org/10.1128/jb.172.9.4945-4950.1990

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816-821. https://doi.org/10.1126/science.1225829

Jønsson, K. A., Fabre, P.-H., Fritz, S. A., Etienne, R. S., Ricklefs, R. E., Jørgensen, T. B., . . . Irestedt, M. (2012). Ecological and evolutionary determinants for the adaptive radiation of the Madagascan vangas. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(17), 6620-6625. https://doi.org/doi:10.1073/pnas.1115835109

Kähkönen, M. P., & Heinonen, M. (2003). Antioxidant activity of anthocyanins and their aglycons. *Journal of Agricultural and Food Chemistry*, *51*(3), 628-633. https://doi.org/10.1021/jf025551i

Kang, H., Zhang, T.-T., Li, Y.-Y., Lin-Wang, K., Espley, R., Du, Y.-P., . . . Wang, X.-F. (2022). The apple BTB protein MdBT2 positively regulates MdCOP1 abundance to repress anthocyanin biosynthesis. *Plant Physiology*, *190*(1), 305-318. https://doi.org/10.1093/plphys/kiac279

Karsten, U., & Holzinger, A. (2014). Green algae in alpine biological soil crust communities: acclimation strategies against ultraviolet radiation and dehydration. *Biodiversity and Conservation*, *23*(7), 1845-1858. https://doi.org/10.1007/s10531-014-0653-2

Katiyar, A., Smita, S., Lenka, S. K., Rajwanshi, R., Chinnusamy, V., & Bansal, K. C. (2012). Genome-wide classification and expression analysis of *MYB* transcription factor families in rice and *Arabidopsis*. *BMC Genomics*, *13*(1), 544. https://doi.org/10.1186/1471-2164-13-544

Kenrick, P., & Crane, P. R. (1997). The origin and early evolution of plants on land. *Nature*, *389*(6646), 33-39. https://doi.org/10.1038/37918

Klein, T. M., Wolf, E. D., Wu, R., & Sanford, J. C. (1987). High-velocity microprojectiles for delivering nucleic acids into living cells. *Nature*, *327*(6117), 70-73. https://doi.org/10.1038/327070a0

Koes, R., Verweij, W., & Quattrocchio, F. (2005). Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends in Plant Science*, *10*(5), 236-242. https://doi.org/10.1016/j.tplants.2005.03.002

Komatsu, A., Kodama, K., Mizuno, Y., Fujibayashi, M., Naramoto, S., & Kyozuka, J. (2023). Control of vegetative reproduction in *Marchantia polymorpha* by the KAI2-ligand signaling pathway. *Current Biology*, *33*(7), 1196-1210.e1194. https://doi.org/10.1016/j.cub.2023.02.022

Kondou, Y., Miyagi, Y., Morito, T., Fujihira, K., Miyauchi, W., Moriyama, A., . . . Kohchi, T. (2019). Physiological function of photoreceptor UVR8 in UV-B tolerance in the liverwort *Marchantia polymorpha*. *Planta*, *249*(5), 1349-1364. https://doi.org/10.1007/s00425-019-03090-w

Kong, L., Liu, Y., Zhi, P., Wang, X., Xu, B., Gong, Z., & Chang, C. (2020). Origins and evolution of cuticle biosynthetic machinery in land plants. *Plant Physiology*, *184*(4), 1998-2010. https://doi.org/10.1104/pp.20.00913

Kubo, H., Nozawa, S., Hiwatashi, T., Kondou, Y., Nakabayashi, R., Mori, T., . . . Ishizaki, K. (2018). Biosynthesis of riccionidins and marchantins is regulated by R2R3-MYB transcription factors in *Marchantia polymorpha*. *Journal of Plant Research*, *131*(5), 849-864. https://doi.org/10.1007/s10265-018-1044-7

Kubota, A., Ishizaki, K., Hosaka, M., & Kohchi, T. (2013). Efficient *Agrobacterium*-mediated transformation of the liverwort *Marchantia polymorpha* using regenerating thalli. *Bioscience, Biotechnology, and Biochemistry*, *77*(1), 167-172. https://doi.org/10.1271/bbb.120700

Kulshrestha, S., Jibran, R., van Klink, J. W., Zhou, Y., Brummell, D. A., Albert, N. W., . . . Davies, K. M. (2022). Stress, senescence, and specialized metabolites in bryophytes. *Journal of Experimental Botany*, *73*(13), 4396-4411. https://doi.org/10.1093/jxb/erac085

Kunz, S., Burkhardt, G., & Becker, H. (1993). Riccionidins a and b, anthocyanidins from the cell walls of the liverwort *Ricciocarpos natans*. *Phytochemistry*, *35*(1), 233-235. https://doi.org/10.1016/S0031-9422(00)90540-5

Kurokawa, S., Rahman, H., Yamanaka, N., Ishizaki, C., Islam, S., Aiso, T., . . . Kaya, H. (2021). A simple heat treatment increases SpCas9-mediated mutation efficiency in *Arabidopsis*. *Plant and Cell Physiology*, *62*(11), 1676-1686. https://doi.org/10.1093/pcp/pcab123

Kuta, D., & Tripathi, L. (2005). *Agrobacterium*-induced hypersensitive necrotic reaction in plant cells: A resistance response against *Agrobacterium*-mediated DNA transfer. *African Journal of Biotechnology*, *4*, 752-757. https://doi.org/10.4314/AJB.V4I8.15199

Lai, E.-M., Chesnokova, O., Banta, L. M., & Kado, C. I. (2000). Genetic and environmental factors affecting T-pilin export and T-pilus biogenesis in relation to flagellation of *Agrobacterium tumefaciens*. *Journal of Bacteriology*, *182*(13), 3705-3716. https://doi.org/10.1128/jb.182.13.3705-3716.2000

Landi, M., Tattini, M., & Gould, K. S. (2015). Multiple functional roles of anthocyanins in plant-environment interactions. *Environmental and Experimental Botany*, *119*, 4-17. https://doi.org/10.1016/j.envexpbot.2015.05.012

Lang, D., van Gessel, N., Ullrich, K. K., & Reski, R. (2016). The genome of the model moss *Physcomitrella patens*. In S. A. Rensing (Ed.), *Advances in Botanical Research* (Vol. 78, pp. 97-140). Academic Press. https://doi.org/10.1016/bs.abr.2016.01.004

Lapham, R. A., Lee, L. Y., Xhako, E., Gómez, E. G., Nivya, V. M., & Gelvin, S. B. (2021). *Agrobacterium* VirE2 protein modulates plant gene expression and mediates transformation from its location outside the nucleus. *Frontiers in Plant Science*, *12*, 684192. https://doi.org/10.3389/fpls.2021.684192

Larebeke, N. V., Engler, G., Holsters, M., Den Elsacker, S. V., Zaenen, I., Schilperoort, R. A., & Schell, J. (1974). Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. *Nature*, *252*(5479), 169-170. https://doi.org/10.1038/252169a0

Le Rhun, A., Escalera-Maurer, A., Bratovič, M., & Charpentier, E. (2019). CRISPR-Cas in *Streptococcus pyogenes*. *RNA Biology*, *16*(4), 380-389. https://doi.org/10.1080/15476286.2019.1582974

LeBlanc, C., Zhang, F., Mendez, J., Lozano, Y., Chatpar, K., Irish, V. F., & Jacob, Y. (2018). Increased efficiency of targeted mutagenesis by CRISPR/Cas9 in plants using heat stress. *The Plant Journal*, *93*(2), 377-386. https://doi.org/10.1111/tpj.13782

Li, D. D., Ni, R., Wang, P. P., Zhang, X. S., Wang, P. Y., Zhu, T. T., . . . Cheng, A. X. (2020a). Molecular basis for chemical evolution of flavones to flavonols and anthocyanins in land plants. *Plant Physiology*, *184*(4), 1731-1743. https://doi.org/10.1104/pp.20.01185

Li, H., Li, D., Yang, Z., Zeng, Q., Luo, Y., & He, N. (2020b). Flavones produced by mulberry flavone synthase type I constitute a defense line against the ultraviolet-B stress. *Plants*, *9*(2), 215. https://www.mdpi.com/2223-7747/9/2/215

Li, Y. Y., Mao, K., Zhao, C., Zhao, X. Y., Zhang, H. L., Shu, H. R., & Hao, Y. J. (2012). MdCOP1 ubiquitin E3 ligases interact with MdMYB1 to regulate light-induced anthocyanin biosynthesis and red fruit coloration in apple. *Plant Physiology*, *160*(2), 1011-1022. https://doi.org/10.1104/pp.112.199703

Liang, J., & He, J. (2018). Protective role of anthocyanins in plants under low nitrogen stress. *Biochemical and Biophysical Research Communications*, *498*(4), 946-953. https://doi.org/10.1016/j.bbrc.2018.03.087

Ligrone, R., Duckett, J. G., & Renzaglia, K. S. (2012). Major transitions in the evolution of early land plants: a bryological perspective. *Annals of Botany*, *109*(5), 851-871. https://doi.org/10.1093/aob/mcs017

Linde, A.-M., Sawangproh, W., Cronberg, N., Szövényi, P., & Lagercrantz, U. (2020). Evolutionary history of the *Marchantia polymorpha* complex. *Frontiers in Plant Science*, *11*. https://doi.org/10.3389/fpls.2020.00829

Liu, J., Osbourn, A., & Ma, P. (2015). MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Molecular plant*, *8*(5), 689-708. https://doi.org/10.1016/j.molp.2015.03.012

Lloyd, A., Brockman, A., Aguirre, L., Campbell, A., Bean, A., Cantero, A., & Gonzalez, A. (2017). Advances in the MYB-bHLH-WD repeat (MBW) pigment regulatory model: Addition of a WRKY factor and co-option of an anthocyanin MYB for betalain regulation. *Plant and Cell Physiology*, *58*(9), 1431-1441. https://doi.org/10.1093/pcp/pcx075

Malzahn, A. A., Tang, X., Lee, K., Ren, Q., Sretenovic, S., Zhang, Y., . . . Qi, Y. (2019). Application of CRISPR-Cas12a temperature sensitivity for improved genome editing in rice, maize, and *Arabidopsis*. *BMC Biology*, *17*(1), 9. https://doi.org/10.1186/s12915-019-0629-5

Mandadi, K. K., Misra, A., Ren, S., & McKnight, T. D. (2009). BT2, a BTB protein, mediates multiple responses to nutrients, stresses, and hormones in *Arabidopsis*. *Plant Physiology*, *150*(4), 1930-1939. https://doi.org/10.1104/pp.109.139220

Markham, K. R., & Porter, I. J. (1974). Flavonoids of the liverwort *Marchantia polymorpha*. *Phytochemistry*, *13*(9), 1937-1942. https://doi.org/10.1016/0031-9422(74)85120-4

Markham, K. R., Ryan, K. G., Bloor, S. J., & Mitchell, K. A. (1998). An increase in the luteolin : apigenin ratio in *Marchantia polymorpha* on UV-B enhancement. *Phytochemistry*, *48*(5), 791-794. https://doi.org/10.1016/S0031-9422(97)00875-3

Martens, S., Forkmann, G., Britsch, L., Wellmann, F., Matern, U., & Lukačin, R. (2003). Divergent evolution of flavonoid 2-oxoglutarate-dependent dioxygenases in parsley. *FEBS Letters*, *544*(1), 93-98. https://doi.org/10.1016/S0014-5793(03)00479-4

Martens, S., & Mithöfer, A. (2005). Flavones and flavone synthases. *Phytochemistry*, *66*(20), 2399-2407. https://doi.org/10.1016/j.phytochem.2005.07.013

Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., & Turner, D. H. (2004). Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(19), 7287-7292. https://doi.org/10.1073/pnas.0401799101

McCormac, A., Elliott, M., & Chen, D.-f. (1998). A simple method for the production of highly competent cells of *Agrobacterium* for transformation via electroporation. *Molecular biotechnology*, *9*, 155-159. https://doi.org/10.1007/BF02760816

Mehrtens, F., Kranz, H., Bednarek, P., & Weisshaar, B. (2005). The *Arabidopsis* transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiology*, *138*(2), 1083-1096. https://doi.org/10.1104/pp.104.058032

Milner, M. J., Craze, M., Hope, M. S., & Wallington, E. J. (2020). Turning up the temperature on CRISPR: Increased temperature can improve the editing efficiency of wheat using CRISPR/Cas9. *Frontiers in Plant Science*, *11*(583374). https://doi.org/10.3389/fpls.2020.583374

Morcuende, R., Bari, R., Gibon, Y., Zheng, W. M., Pant, B. D., Blasing, O., . . . Scheible, W. R. (2007). Genome-wide reprogramming of metabolism and regulatory networks of *Arabidopsis* in response to phosphorus. *Plant Cell and Environment*, *30*(1), 85-112. https://doi.org/10.1111/j.1365-3040.2006.01608.x

Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., . . . Donoghue, P. C. J. (2018). The timescale of early land plant evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(10), E2274-E2283. https://doi.org/10.1073/pnas.1719588115

Mühlhausen, S., & Kollmar, M. (2013). Whole genome duplication events in plant evolution reconstructed and predicted using myosin motor proteins. *BMC Evolutionary Biology*, *13*(1), 202. https://doi.org/10.1186/1471-2148-13-202

Nakayama, T. (2022). Biochemistry and regulation of aurone biosynthesis. *Bioscience, Biotechnology, and Biochemistry*, *86*(5), 557-573. https://doi.org/10.1093/bbb/zbac034

Nakayama, T., Sato, T., Fukui, Y., Yonekura-Sakakibara, K., Hayashi, H., Tanaka, Y., . . . Nishino, T. (2001). Specificity analysis and mechanism of aurone synthesis catalyzed by aureusidin synthase, a polyphenol oxidase homolog responsible for flower coloration. *FEBS Letters*, *499*(1), 107-111. https://doi.org/10.1016/S0014-5793(01)02529-7

Nakayama, T., Yonekura-Sakakibara, K., Sato, T., Kikuchi, S., Fukui, Y., Fukuchi-Mizutani, M., . . . Nishino, T. (2000). Aureusidin synthase: A polyphenol oxidase homolog responsible for flower coloration. *Science*, *290*(5494), 1163-1166. https://doi.org/10.1126/science.290.5494.1163

Nasu, M., Tani, K., Hattori, C., Honda, M., Shimaoka, T., Yamaguchi, N., & Katoh, K. (1997). Efficient transformation of *Marchantia polymorpha* that is haploid and has very small genome DNA. *Journal of Fermentation and Bioengineering*, *84*(6), 519-523. https://doi.org/10.1016/S0922-338X(97)81904-6

Nemesio-Gorriz, M., Blair, P. B., Dalman, K., Hammerbacher, A., Arnerup, J., Stenlid, J., . . . Elfstrand, M. (2017). Identification of norway spruce MYB-bHLH-WDR transcription factor complex members linked to regulation of the flavonoid pathway. *Frontiers in Plant Science*, *8*(305). https://doi.org/10.3389/fpls.2017.00305

Niklas, K. J., & Kutschera, U. (2010). The evolution of the land plant life cycle. *New Phytologist*, *185*(1), 27-41. https://doi.org/10.1111/j.1469-8137.2009.03054.x

Odipio, J., Alicai, T., Ingelbrecht, I., Nusinow, D., Bart, R., & Taylor, N. (2017). Efficient CRISPR/Cas9 genome editing of *Phytoene Desaturase* in cassava. *Frontiers in Plant Science*, *8*, 1780. https://doi.org/10.3389/fpls.2017.01780

Oei, A. L., Vriend, L. E., Crezee, J., Franken, N. A., & Krawczyk, P. M. (2015). Effects of hyperthermia on DNA repair pathways: one treatment to inhibit them all. *Radiation Oncology*, *10*, 165. https://doi.org/10.1186/s13014-015-0462-0

Pagel, K. A., Antaki, D., Lian, A., Mort, M., Cooper, D. N., Sebat, J., . . . Radivojac, P. (2019). Pathogenicity and functional impact of non-frameshifting insertion/deletion variation in the human genome. *PLOS Computational Biology*, *15*(6), e1007112. https://doi.org/10.1371/journal.pcbi.1007112

Palmer, A. C. V., & Shaw, C. H. (1992). The role of VirA and VirG phosphorylation in chemotaxis towards acetosyringone by *Agrobacterium tumefaciens*. *Microbiology*, *138*(12), 2509-2514. https://doi.org/10.1099/00221287-138-12-2509

Piatkowski, B. T., Imwattana, K., Tripp, E. A., Weston, D. J., Healey, A., Schmutz, J., & Shaw, A. J. (2020). Phylogenomics reveals convergent evolution of red-violet coloration in land plants and the origins of the anthocyanin biosynthetic pathway. *Molecular Phylogenetics and Evolution*, *151*, 106904. https://doi.org/10.1016/j.ympev.2020.106904

Pietrini, F., Iannelli, M. A., & Massacci, A. (2002). Anthocyanin accumulation in the illuminated surface of maize leaves enhances protection from photo-inhibitory risks at low temperature,

without further limitation to photosynthesis. *Plant, Cell & Environment*, *25*(10), 1251-1259. https://doi.org/10.1046/j.1365-3040.2002.00917.x

Pires, N. D., & Dolan, L. (2012). Morphological evolution in land plants: new designs with old genes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1588), 508-518. https://doi.org/10.1098/rstb.2011.0252

Pucker, B., & Iorizzo, M. (2023). Apiaceae FNS I originated from F3H through tandem gene duplication. *PLOS ONE*, *18*(1), e0280155. https://doi.org/10.1371/journal.pone.0280155

Qin, G., Gu, H., Ma, L., Peng, Y., Deng, X. W., Chen, Z., & Qu, L.-J. (2007). Disruption of *Phytoene desaturase* gene results in albino and dwarf phenotypes in *Arabidopsis* by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. *Cell Research*, *17*(5), 471-482. https://doi.org/10.1038/cr.2007.40

Quirk, J., Leake, J. R., Johnson, D. A., Taylor, L. L., Saccone, L., & Beerling, D. J. (2015). Constraining the role of early land plants in Palaeozoic weathering and global cooling. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1813), 20151115. https://doi.org/doi:10.1098/rspb.2015.1115

Ren, Y.-R., Zhao, Q., Yang, Y.-Y., Zhang, T.-E., Wang, X.-F., You, C.-X., & Hao, Y.-J. (2021a). The apple 14-3-3 protein MdGRF11 interacts with the BTB protein MdBT2 to regulate nitrate deficiency-induced anthocyanin accumulation. *Horticulture Research*, *8*(1), 22. https://doi.org/10.1038/s41438-020-00457-z

Ren, Y. R., Zhao, Q., Yang, Y. Y., Zhang, R., Wang, X. F., Zhang, T. E., . . . Hao, Y. J. (2021b). Interaction of BTB-TAZ protein MdBT2 and DELLA protein MdRGL3a regulates nitrate-mediated plant growth. *Plant Physiology*, *186*(1), 750-766. https://doi.org/10.1093/plphys/kiab065

Rico-Reséndiz, F., Cervantes-Pérez, S. A., Espinal-Centeno, A., Dipp-Álvarez, M., Oropeza-Aburto, A., Hurtado-Bautista, E., . . . Cruz-Ramírez, A. (2020). Transcriptional and morpho-physiological responses of *Marchantia polymorpha* upon phosphate starvation. *International Journal of Molecular Sciences*, *21*(21), 8354. https://www.mdpi.com/1422-0067/21/21/8354

Rudolph, H., & Jöhnk, A. (1982). Physiological aspects of phenolic compounds in the cell walls of *Sphagna*. *Journal of Hattori Botanical Laboratory*, *53*, 195-203. https://doi.org/10.18968/jhbl.53.0_195

Sato, T., Maekawa, S., Konishi, M., Yoshioka, N., Sasaki, Y., Maeda, H., . . . Yanagisawa, S. (2017). Direct transcriptional activation of BT genes by NLP transcription factors is a key component of the nitrate response in *Arabidopsis*. *Biochemical and Biophysical Research Communications*, *483*(1), 380-386. https://doi.org/https://doi.org/10.1016/j.bbrc.2016.12.135

Scheible, W.-R. d., Morcuende, R., Czechowski, T., Fritz, C., Osuna, D., Palacios-Rojas, N., . . . Stitt, M. (2004). Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of *Arabidopsis* in response to nitrogen. *Plant Physiology*, *136*(1), 2483-2499. https://doi.org/10.1104/pp.104.047019

Sheikholeslam, S. N., & Weeks, D. P. (1987). Acetosyringone promotes high efficiency transformation of *Arabidopsis thaliana* explants by *Agrobacterium tumefaciens*. *Plant Molecular Biology*, *8*(4), 291-298. https://doi.org/10.1007/BF00021308

Shimamura, M. (2015). *Marchantia polymorpha* : Taxonomy, phylogeny and morphology of a model system. *Plant and Cell Physiology*, *57*(2), 230-256. https://doi.org/10.1093/pcp/pcv192

Soltis, D. E., Chanderbali, A. S., Kim, S., Buzgo, M., & Soltis, P. S. (2007). The ABC model and its applicability to basal angiosperms. *Annals of Botany*, *100*(2), 155-163. https://doi.org/10.1093/aob/mcm117

Song, Z., Luo, Y., Wang, W., Fan, N., Wang, D., Yang, C., & Jia, H. (2020). NtMYB12 positively regulates flavonol biosynthesis and enhances tolerance to low Pi stress in *Nicotiana tabacum*. *Frontiers in Plant Science*, *10*(1683). https://doi.org/10.3389/fpls.2019.01683

Soriano, G., Del-Castillo-Alonso, M.-Á., Monforte, L., Tomás-Las-Heras, R., Martínez-Abaigar, J., & Núñez-Olivera, E. (2019). Photosynthetically-active radiation, UV-A and UV-B, causes both common and specific damage and photoprotective responses in the model liverwort *Marchantia polymorpha* subsp. ruderalis. *Photochemical & Photobiological Sciences*, *18*, 400-412. https://doi.org/10.1039/c8pp00421h

Stachel, S. E., Messens, E., Van Montagu, M., & Zambryski, P. (1985). Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. *Nature*, *318*(6047), 624-629. https://doi.org/10.1038/318624a0

Stachel, S. E., Timmerman, B., & Zambryski, P. (1986). Generation of single-stranded T-DNA molecules during the initial stages of T-DNA transfer from *Agrobacterium tumefaciens* to plant cells. *Nature*, *322*(6081), 706-712. https://doi.org/10.1038/322706a0

Stewart, A. J., Chapman, W., Jenkins, G. I., Graham, I., Martin, T., & Crozier, A. (2001). The effect of nitrogen and phosphorus deficiency on flavonol accumulation in plant tissues. *Plant, Cell & Environment*, *24*(11), 1189-1197. https://doi.org/10.1046/j.1365-3040.2001.00768.x

Stracke, R., Favory, J. J., Gruber, H., Bartelniewoehner, L., Bartels, S., Binkert, M., . . . Ulm, R. (2010). The *Arabidopsis* bZIP transcription factor HY5 regulates expression of the *PFG1/MYB12* gene in response to light and ultraviolet-B radiation. *Plant Cell and Environment*, *33*(1), 88-103. https://doi.org/10.1111/j.1365-3040.2009.02061.x

Stracke, R., Ishihara, H., Barsch, G. H. A., Mehrtens, F., Niehaus, K., & Weisshaar, B. (2007). Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant Journal*, *50*(4), 660-677. https://doi.org/10.1111/j.1365-313X.2007.03078.x

Stracke, R., Werber, M., & Weisshaar, B. (2001). The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology*, *4*(5), 447-456. https://doi.org/10.1016/S1369-5266(00)00199-0

Sugano, S. S., Nishihama, R., Shirakawa, M., Takagi, J., Matsuda, Y., Ishida, S., . . . Kohchi, T. (2018). Efficient CRISPR/Cas9-based genome editing and its application to conditional genetic analysis in *Marchantia polymorpha*. *PLOS ONE*, *13*(10), e0205117. https://doi.org/10.1371/journal.pone.0205117

Takebe, I., Otsuki, Y., & Aoki, S. (1968). Isolation of tobacco mesophyll cells in intact and active state. *Plant and Cell Physiology*, *9*(1), 115-124. https://doi.org/10.1093/oxfordjournals.pcp.a079318

Tank, D. C., Eastman, J. M., Pennell, M. W., Soltis, P. S., Soltis, D. E., Hinchliff, C. E., . . . Harmon, L. J. (2015). Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist*, *207*(2), 454-467. https://doi.org/10.1111/nph.13491

Tattini, M., Galardi, C., Pinelli, P., Massai, R., Remorini, D., & Agati, G. (2004). Differential accumulation of flavonoids and hydroxycinnamates in leaves of *Ligustrum vulgare* under excess light and drought stress. *New Phytologist*, *163*(3), 547-561. https://doi.org/10.1111/j.1469-8137.2004.01126.x

Tinland, B., Koukolíková-Nicola, Z., Hall, M. N., & Hohn, B. (1992). The T-DNA-linked VirD2 protein contains two distinct functional nuclear localization signals. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(16), 7442-7446. https://doi.org/10.1073/pnas.89.16.7442

Tohge, T., Watanabe, M., Hoefgen, R., & Fernie, A. R. (2013). The evolution of phenylpropanoid metabolism in the green lineage. *Critical Reviews in Biochemistry and Molecular Biology*, *48*(2), 123-152. https://doi.org/10.3109/10409238.2012.758083

Tran, L. T., Taylor, J. S., & Constabel, C. P. (2012). The polyphenol oxidase gene family in land plants: Lineage-specific duplication and expansion. *BMC Genomics*, *13*(1), 395. https://doi.org/10.1186/1471-2164-13-395

Tsuboyama, S., & Kodama, Y. (2013). AgarTrap: A simplified *Agrobacterium*-mediated transformation method for sporelings of the liverwort *Marchantia polymorpha* L. *Plant and Cell Physiology*, *55*(1), 229-236. https://doi.org/10.1093/pcp/pct168

Tsuboyama, S., & Kodama, Y. (2018a). AgarTrap protocols on your benchtop: Simple methods for *Agrobacterium*-mediated genetic transformation of the liverwort *Marchantia polymorpha*. *Plant Biotechnology*, *35*(2), 93-99. https://doi.org/10.5511/plantbiotechnology.18.0312b

Tsuboyama, S., & Kodama, Y. (2018b). Highly efficient G-AgarTrap-mediated transformation of the *Marchantia polymorpha* model strains Tak-1 and Tak-2. *Plant Biotechnology*, *35*(4), 399-403. https://doi.org/10.5511/plantbiotechnology.18.0917a

Tsuboyama, S., & Kodama, Y. (2018c). Highly efficient G-AgarTrap-mediated transformation of the *Marchantia polymorpha* model strains Tak-1 and Tak-2. *Plant Biotechnol (Tokyo)*, *35*(4), 399-403. https://doi.org/10.5511/plantbiotechnology.18.0917a

Tsuboyama, S., Nonaka, S., Ezura, H., & Kodama, Y. (2018). Improved G-AgarTrap: A highly efficient transformation method for intact gemmalings of the liverwort *Marchantia polymorpha*. *bioRxiv*, 329839. https://doi.org/10.1101/329839

Turetsky, M. R. (2003). The role of bryophytes in carbon and nitrogen cycling. *The Bryologist*, *106*(3), 395-409. https://doi.org/10.1639/05

Vanderauwera, S., Zimmermann, P., Rombauts, S., Vandenabeele, S., Langebartels, C., Gruissem, W., . . . Van Breusegem, F. (2005). Genome-wide analysis of hydrogen peroxide-regulated gene expression in *Arabidopsis* reveals a high light-induced transcriptional cluster involved in anthocyanin biosynthesis. *Plant Physiology*, *139*(2), 806-821. https://doi.org/10.1104/pp.105.065896

Wang, F., Kong, W., Wong, G., Fu, L., Peng, R., Li, Z., & Yao, Q. (2016). AtMYB12 regulates flavonoids accumulation and abiotic stress tolerance in transgenic *Arabidopsis thaliana*. *Molecular Genetics and Genomics*, *291*(4), 1545-1559. https://doi.org/10.1007/s00438-016-1203-2

Wang, H., Liu, S., Wang, T., Liu, H., Xu, X., Chen, K., & Zhang, P. (2020). The moss flavone synthase I positively regulates the tolerance of plants to drought stress and UV-B radiation. *Plant Science*, *298*, 110591. https://doi.org/10.1016/j.plantsci.2020.110591

Wang, X.-F., An, J.-P., Liu, X., Su, L., You, C.-X., & Hao, Y.-J. (2018). The nitrate-responsive protein MdBT2 regulates anthocyanin biosynthesis by interacting with the MdMYB1 transcription factor. *Plant Physiology*, *178*(2), 890-906. https://doi.org/10.1104/pp.18.00244

Wang, Y., Zhou, L.-J., Wang, Y., Liu, S., Geng, Z., Song, A., . . . Chen, F. (2021). Functional identification of a *Flavone Synthase* and a *Flavonol Synthase* genes affecting flower color formation in *Chrysanthemum morifolium*. *Plant Physiology and Biochemistry*, *166*, 1109-1120. https://doi.org/10.1016/j.plaphy.2021.07.019

Watira, T., Park, J. J., Chen, M., Cong, L., Ge, Y., Jiang, Q., . . . Wang, Z. (2020). Improving the genome editing efficiency of CRISPR/Cas9 in *Arabidopsis* and *Medicago truncatula*. *Planta*, *252*. https://doi.org/10.1007/s00425-020-03415-0

Wen, W., Alseekh, S., & Fernie, A. R. (2020). Conservation and diversification of flavonoid metabolism in the plant kingdom. *Current Opinion in Plant Biology*, *55*, 100-108. https://doi.org/10.1016/j.pbi.2020.04.004

Wilson, F. M., Harrison, K., Armitage, A. D., Simkin, A. J., & Harrison, R. J. (2019). CRISPR/Cas9-mediated mutagenesis of *Phytoene Desaturase* in diploid and octoploid strawberry. *Plant Methods*, *15*(1), 45. https://doi.org/10.1186/s13007-019-0428-6

Winkel-Shirley, B. (2002). Biosynthesis of flavonoids and effects of stress. *Current Opinion in Plant Biology*, *5*(3), 218-223. https://doi.org/10.1016/S1369-5266(02)00256-X

Wodniok, S., Brinkmann, H., Glöckner, G., Heidel, A. J., Philippe, H., Melkonian, M., & Becker, B. (2011). Origin of land plants: Do conjugating green algae hold the key? *BMC Evolutionary Biology*, *11*(1), 104. https://doi.org/10.1186/1471-2148-11-104

Wohl, J., & Petersen, M. (2020). Functional expression and characterization of cinnamic acid 4-hydroxylase from the hornwort *Anthoceros agrestis* in *Physcomitrella patens*. *Plant Cell Reports*, *39*(5), 597-607. https://doi.org/10.1007/s00299-020-02517-z

Wu, J., Wang, X.-C., Liu, Y., Du, H., Shu, Q.-Y., Su, S., . . . Wang, L.-S. (2016). Flavone synthases from *Lonicera japonica* and *L. macranthoides* reveal differential flavone accumulation. *Scientific Reports*, *6*(1), 19245. https://doi.org/10.1038/srep19245

Wu, Y.-F., Zhao, Y., Liu, X.-Y., Gao, S., Cheng, A.-X., & Lou, H.-X. (2018). A bHLH transcription factor regulates bisbibenzyl biosynthesis in the liverwort *Plagiochasma appendiculatum*. *Plant and Cell Physiology*, *59*(6), 1187-1199. https://doi.org/10.1093/pcp/pcy053

Xie, K., Minkenberg, B., & Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(11), 3570-3575. https://doi.org/10.1073/pnas.1420294112

Yang, L., Su, D., Chang, X., Foster, C. S. P., Sun, L., Huang, C. H., . . . Zhong, B. (2020). Phylogenomic insights into deep phylogeny of angiosperms based on broad nuclear gene sampling. *Plant Communications*, *1*(2), 100027. https://doi.org/10.1016/j.xplc.2020.100027

Yelina, N. E., Frangedakis, E., Schreier, T. B., Rever, J., Tomaselli, M., Haseloff, J., & Hibberd, J. M. (2023). Streamlined regulation of chloroplast development in the liverwort *Marchantia polymorpha*. *bioRxiv*, 2023.2001.2023.525199. https://doi.org/10.1101/2023.01.23.525199

Yonekura-Sakakibara, K., Higashi, Y., & Nakabayashi, R. (2019). The origin and evolution of plant flavonoid metabolism. *Frontiers in Plant Science*, *10*(943). https://doi.org/10.3389/fpls.2019.00943

Zhang, B., Hu, Z., Zhang, Y., Li, Y., Zhou, S., & Chen, G. (2012). A putative functional MYB transcription factor induced by low temperature regulates anthocyanin biosynthesis in purple kale (*Brassica Oleracea* var. *acephala* f. tricolor). *Plant Cell Reports*, *31*(2), 281-289. https://doi.org/10.1007/s00299-011-1162-3

Zhang, D., Yang, K., Kan, Z., Dang, H., Feng, S., Yang, Y., . . . Guan, Q. (2021). The regulatory module MdBT2–MdMYB88/MdMYB124–MdNRTs regulates nitrogen usage in apple. *Plant Physiology*, *185*(4), 1924-1942. https://doi.org/10.1093/plphys/kiaa118

Zhang, Z., Hua, L., Gupta, A., Tricoli, D., Edwards, K. J., Yang, B., & Li, W. (2019). Development of an *Agrobacterium*-delivered CRISPR/Cas9 system for wheat genome editing. *Plant Biotechnology Journal*, *17*(8), 1623-1635. https://doi.org/10.1111/pbi.13088

Zheng, X., Xing, J., Zhang, K., Pang, X., Zhao, Y., Wang, G., . . . Dong, J. (2019). Ethylene response factor ERF11 activates BT4 transcription to regulate immunity to *Pseudomonas syringae*. *Plant Physiology*, *180*(2), 1132-1151. https://doi.org/10.1104/pp.18.01209

Zhou, F., Zhang, K., Zheng, X., Wang, G., Cao, H., Xing, J., & Dong, J. (2022). BTB and TAZ domain protein BT4 positively regulates the resistance to *Botrytis cinerea* in *Arabidopsis*. *Plant Signaling & Behavior*, *17*(1), 2104003. https://doi.org/10.1080/15592324.2022.2104003

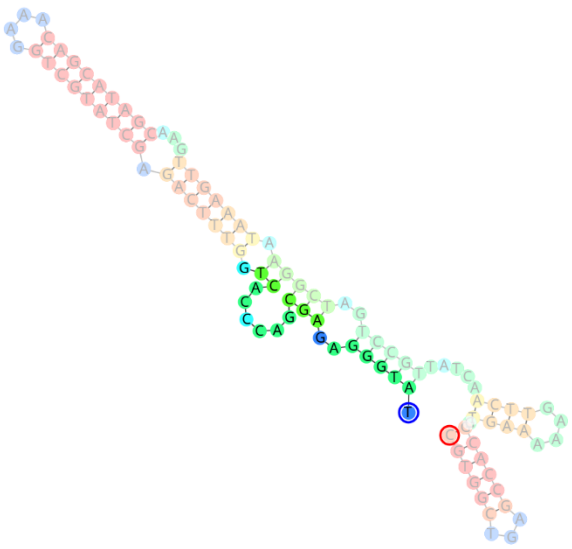
Zhou, Y., Zhai, H., He, S., Zhu, H., Gao, S., Xing, S., . . . Liu, Q. (2020). The sweetpotato BTB-TAZ protein gene, Ib*BT4*, enhances drought tolerance in transgenic *Arabidopsis*. *Frontiers in Plant Science*, *11*(877). https://doi.org/10.3389/fpls.2020.00877

Appendices

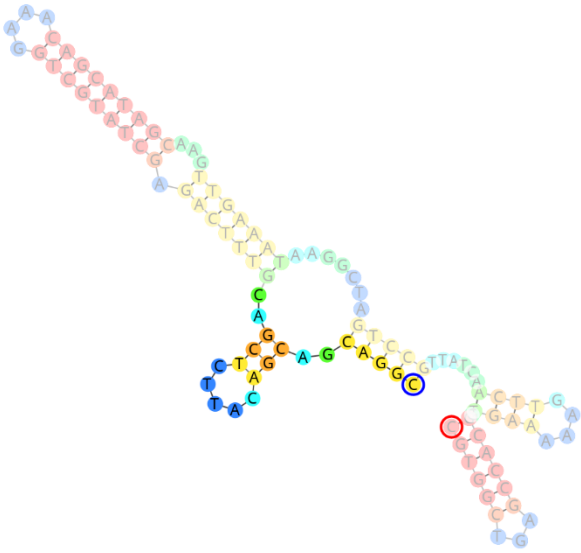
Appendix 1: sgRNA quality

sgRNAs used to generate CRISPR/Cas9 mutants need to be high quality. Examples of good (Mp*FNSI* sgRNA 5) and bad (Mp*FNSI* sgRNA 11) quality sgRNAs are shown below. Red-coloured nucleotides represent a high binding probability and green a lower binding probability.

Mp*FNSI* sgRNA 5



Mp*FNSI* sgRNA 11



Mp*FNSI* sgRNA 5 is high quality because the 20-nucleotide crRNA (highlighted) is mostly unbound, so can easily bind to the target DNA.

Mp*FNSI* sgRNA 11 is bad quality because the 20-nucleotide crRNA (highlighted) is bound to itself, forming a hairpin, and is therefore not accessible.

Appendix 2: Chemicals, media and solutions

2.1 Chemicals

All chemicals were purchased from Roche, New England Biolabs (NEB), or Invitrogen, unless otherwise stated.

2.2 Autoclaving

All media and solutions that required sterilising were autoclaved at 120°C and 15 psi, for 20 min.

Solutions that did not require sterilising were prepared with autoclaved milliQ water.

2.3 Antibiotic stocks

Gentamicin stock solution (25 mg mL^{-1}) was made by dissolving 0.25 g of gentamicin sulfate salt in a final volume of 10 mL of milliQ water. The gentamycin solution was filter sterilised (0.2 µm) and stored at 4°C.

Spectinomycin stock solution (100 mg mL^{-1}) was made by dissolving 1 g of spectinomycin dihydrochloride pentahydrate in a final volume of 10 mL of milliQ water. The spectinomycin solution was filter sterilised (0.2 µm) and stored at 4°C.

Ampicillin stock solution (100 mg mL^{-1}) was made by dissolving 1 g of ampicillin in a final concentration of 10 mL of milliQ water. The ampicillin solution was filter sterilised (0.2 µm) and stored at -20°C.

Kanamycin stock solution (50 mg mL^{-1}) was made by dissolving 0.5 g of kanamycin sulfate in a final volume of 10 mL of milliQ water. The ampicillin solution was filter sterilised (0.2 µm) and stored at 4°C.

Geneticin stock solution (8 mg mL^{-1}) was made by dissolving 0.08 g of G-418 sulfate in a final volume of 10 mL of milliQ water. The geneticin solution was filter sterilised (0.2 µm) and stored at -20°C.

Hygromycin stock solution (8 mg mL^{-1}) was made by dissolving 0.08 g of hygromycin B in a final volume of 10 mL of milliQ water. The hygromycin solution was filter sterilised (0.2 µm) and stored at -20°C.

Ticarcillin stock solution (100 mg mL^{-1}) was made by dissolving 1 g of ticarcillin 2NA & clavulanate K in a final volume of 10 mL milliQ water. The ticarcillin solution was filter sterilised (0.2 µm) and stored at -20°C.

Cefalexin stock solution (100 mg mL^{-1}) was made by dissolving 1 g of cefalexin in a final volume of 10 mL of milliQ water. The cefalexin solution was filter sterilised (0.2 µm) and stored at 4°C.

2.4 Lysogeny broth (LB) and agar

LB contains 10 g L^{-1} tryptone, 10 g L^{-1} sodium chloride and 5 g L^{-1} yeast extract dissolved in 900 mL sterile water, made up to 1 L with sterile water, and autoclaved.

LB-agar is LB with 15 g L^{-1} agar added and autoclaved.

All LB-agar was cooled to 50°C and LB-broth to room temperature before the addition of required antibiotics.

2.5 Plasmid DNA Mini-prep extraction buffers

Solution P1 contains 50 mM of Tris-Cl (pH 8.0), 10 mM of EDTA (pH 8.0), and 0.5 mg mL^{-1} RNase A, made up to 500 mL with milliQ water. Stored at 4°C.

Solution P2 contains 0.2 M NaOH and 10 g L^{-1} SDS, made up to 250 mL with milliQ water, heated to dissolve. Stored at room temperature.

Solution N3 contains 4 M Guanidine HCl and 0.5 M K-acetate dissolved in 250 mL milliQ water, made up to 500 mL with milli Q water, then pH adjusted to 4.2 with Glacial Acetic Acid. Stored at room temperature.

Solution PB contains 5 M Guanidine HCl dissolved in 50 mL milliQ water, with 20 mM Tris-Cl pH 6.6 and 66 mL 2-propanol added, made up to 200 mL with milliQ water, and pH adjusted to 6.6. Stored at room temperature.

Solution PE contains 20 mM NaCl, 20 mM Tris-Cl pH 7.5, and 400 mL 100% (v/v) ethanol, made up to 500 mL with milliQ water. Stored at room temperature.

2.6 Gel electrophoresis solutions

Standard **1% (w/v) Agarose gel** was made by heat-dissolving 2 g Agarose in 200 mL 10× TBE buffer. The ratio of Agarose to 10× TBE changed depending on the required gel percentage (w/v).

10× TBE contains 0.89 M Tris Base, 0.89 M Boric Acid and 25 mM NaEDTA.$_2H_20$ dissolved in 800 mL sterile water, made up to 1 L with sterile water, and autoclaved.

10× Bromophenol Blue dye stock was made by adding 50 mg Bromophenol Blue sodium salt (Sigma Aldrich) to 2.5 g Ficoll that has been dissolved in a few mL milliQ water. The final volume was brought up to 10 mL with milliQ water.

10× Xylene Cyanol dye stock was made by adding 50 mg Xylene Cyanol F.F. (AJAX Chemicals) to 2.5 g Ficoll that has been dissolved in a few mL milliQ water. The final volume was brought up to 10 mL with milliQ water.

1 kb Plus ladder contains 15 µg of 50 ng μL^{-1} Invitrogen 1 kb Plus DNA Ladder, 35 µL of each of the 10× Bromophenol Blue and Xylene Cyanol dyes, and 20 mM NaCl, made up to 300 µL with sterile water.

2.7 0.5x B5 spore transformation media

0.5× B5 media was made by dissolving 1% (w/v) sucrose and 0.79% (w/v) Duchefa Gamborg B5 medium including vitamins in milliQ water, bringing it up to 1 L, pH adjusting to 6.0 with NaOH and adding to 1% (w/v) Agar. Autoclaved before use.

Acetosyringone and required antibiotics were added to the media once it had cooled down to 50°C.

Acetosyringone solution (10 mg mL^{-1}) was made by dissolving 0.1 g of 3',5'-dimethoxy-4'-hydroxy-acetophenone powder in a final volume of 10 mL of milliQ water. The solution was filter sterilised (0.2 µm) and stored at 4°C.

2.8 $MgCl_2$ MES

$MgCl_2$ MES was made by dissolving 100 mM of both $MgCl_2$ and MES in milliQ water, bringing it up to 200 mL, and pH adjusting to 5.6 with NaOH. Autoclaved before use.

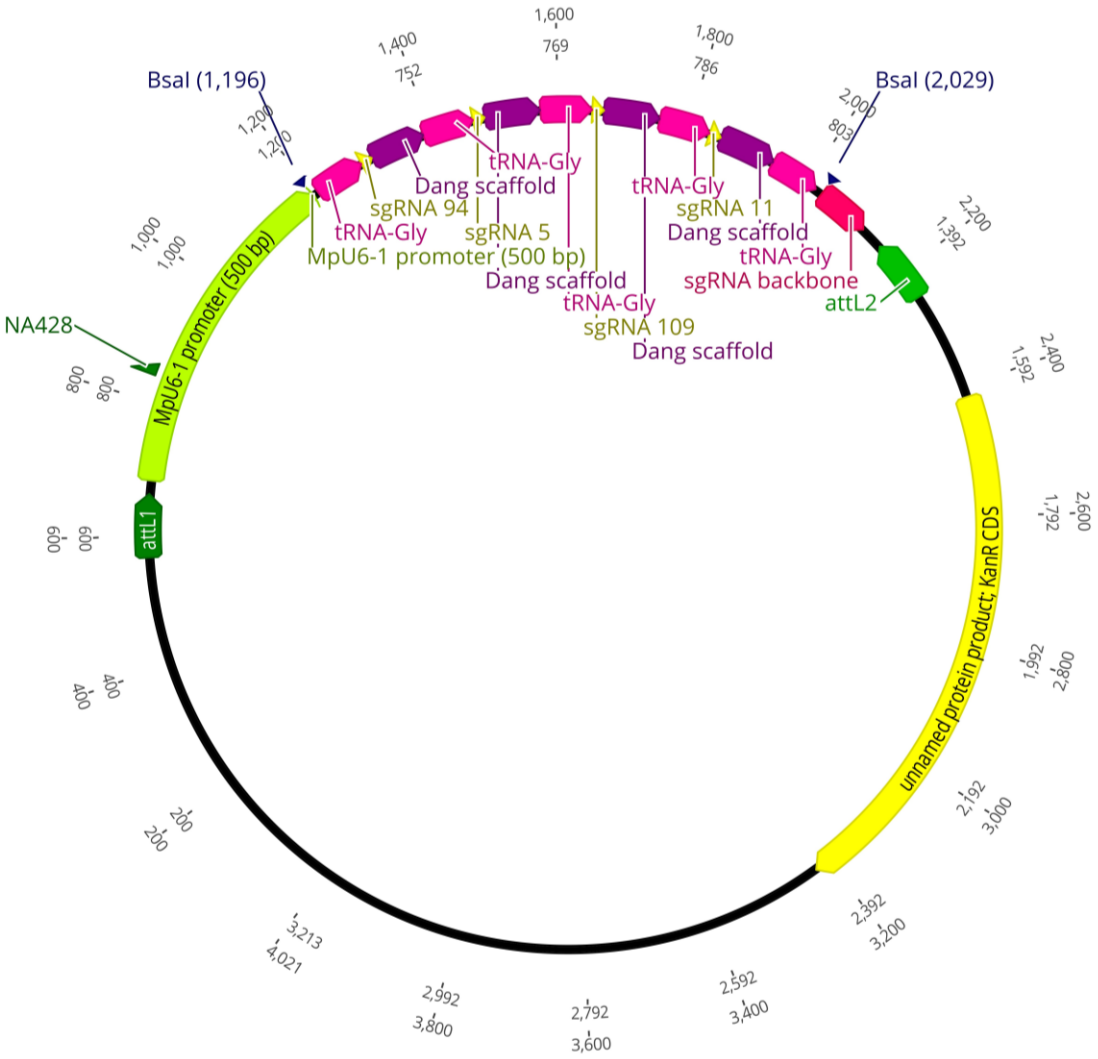
2.9 Genomic DNA extraction buffers

CTAB was made by dissolving 2% (w/v) Tris-HCl in 800 mL sterile water and adjusting the pH to 8.0 with NaOH (100 mM), before the addition of the following components: 20 mM EDTA (pH 4.0), 2% (w/v) CTAB, 1.42 M NaCl, 2% (w/v) PVP-40 and 5 mM ascorbic acid. The solution was made up to

1 L with sterile water and stored at room temperature. Just before use, 0.6% (w/v) β-2-mercaptoethanol was added to the CTAB buffer.

Appendix 3: Plasmid Maps

3.1 Plasmid map for sgRNA cassette and p*U6*$_{pro}$/pENTR ligation

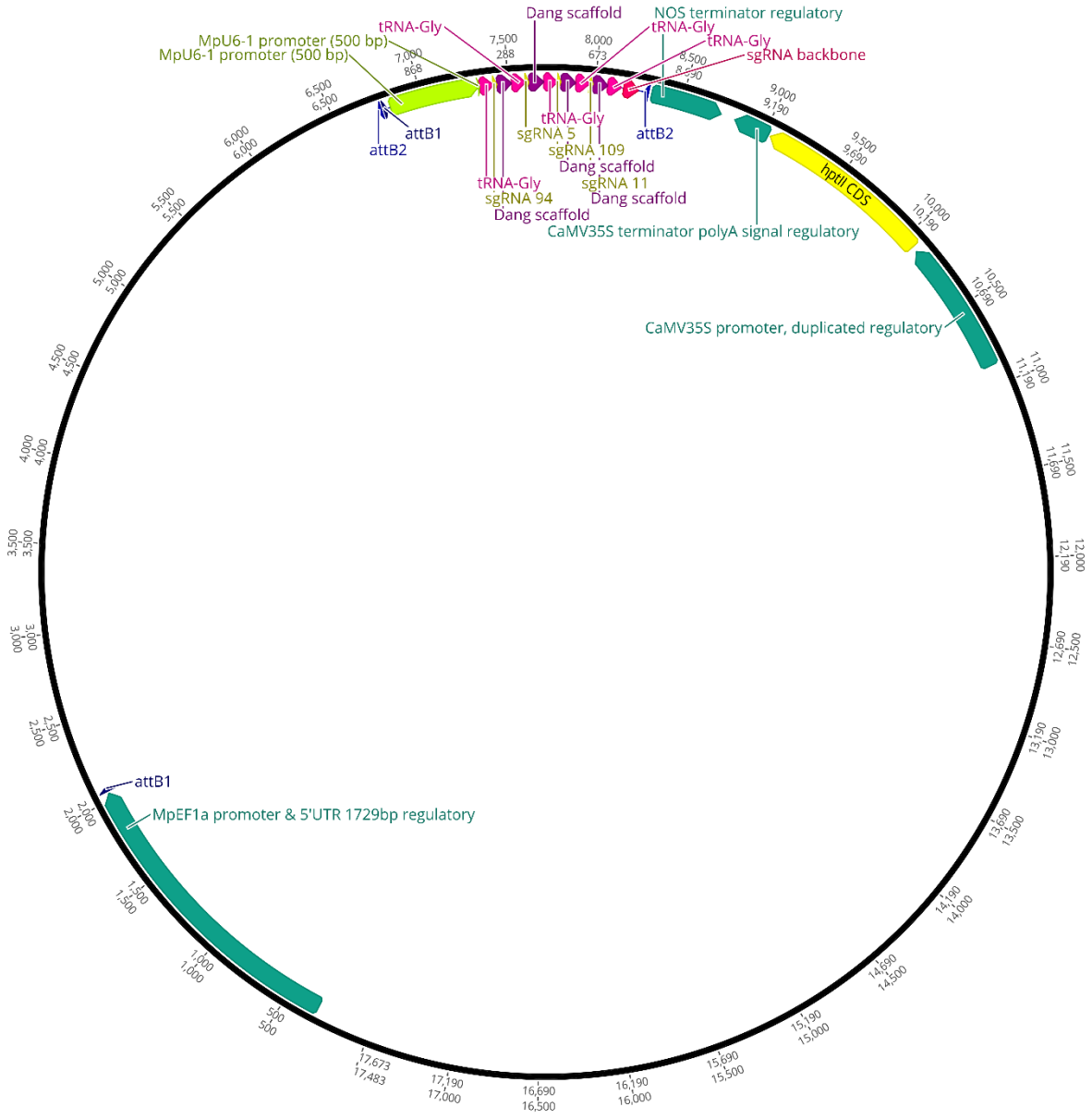


From Geneious version 2022.0, created by Biomatters.

Map is from cloning of the Mp*FNSI* CRISPR/Cas9 construct. The same plasmid design was used for creating Mp*BT2*, Mp*GLK* and Mp*PDS* constructs, except the sgRNAs were re-designed to target each particular gene.

sgRNA cassette includes four sgRNAs (including the 20-nucleotide crRNA designed to target a gene of interest ("sgRNA xxx"; yellow) and tracrRNA scaffold ("Dang scaffold"; purple)), flanked by "tRNA-Gly" molecules (pink).

3.2 Plasmid map for pRMY07

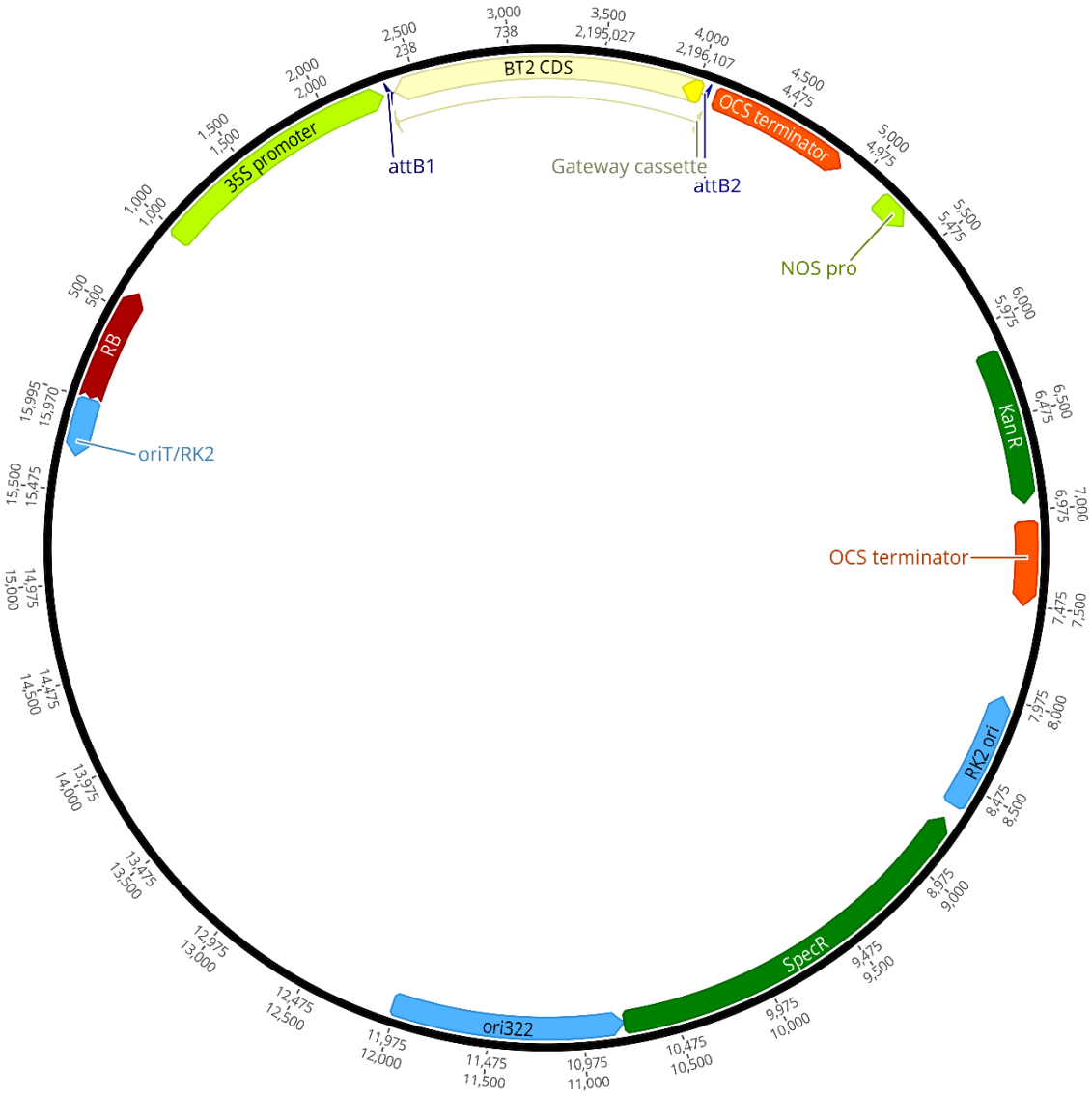


From Geneious version 2022.0, created by Biomatters.

Map is from cloning of the Mp*FNSI* CRISPR/Cas9 construct (pRMY07). The same plasmid design was used for Mp*BT2* (pRMY09), Mp*GLK* (pRMY32) and Mp*PDS* (pRMY13) constructs.

The attB1 and attB2 sites are a result of LR recombination and define the DNA transferred from p$U6_{pro}$/pENTR (the sgRNA cassette and *U6* promoter).

3.3 Plasmid map for pRMY27



From Geneious version 2022.0, created by Biomatters.

Map is from cloning of the Mp*BT2* over-expression construct for spore transformation (pRMY27). The same plasmid design was used for the Mp*FNSI* (pRMY24) and *Venus-N7* over-expression (pRMY21) constructs. A similar design was used for the Mp*BT2* over-expression construct for thallus transformation (pRMY30), except that it instead contained a plant hygromycin resistance gene (*hptII*).

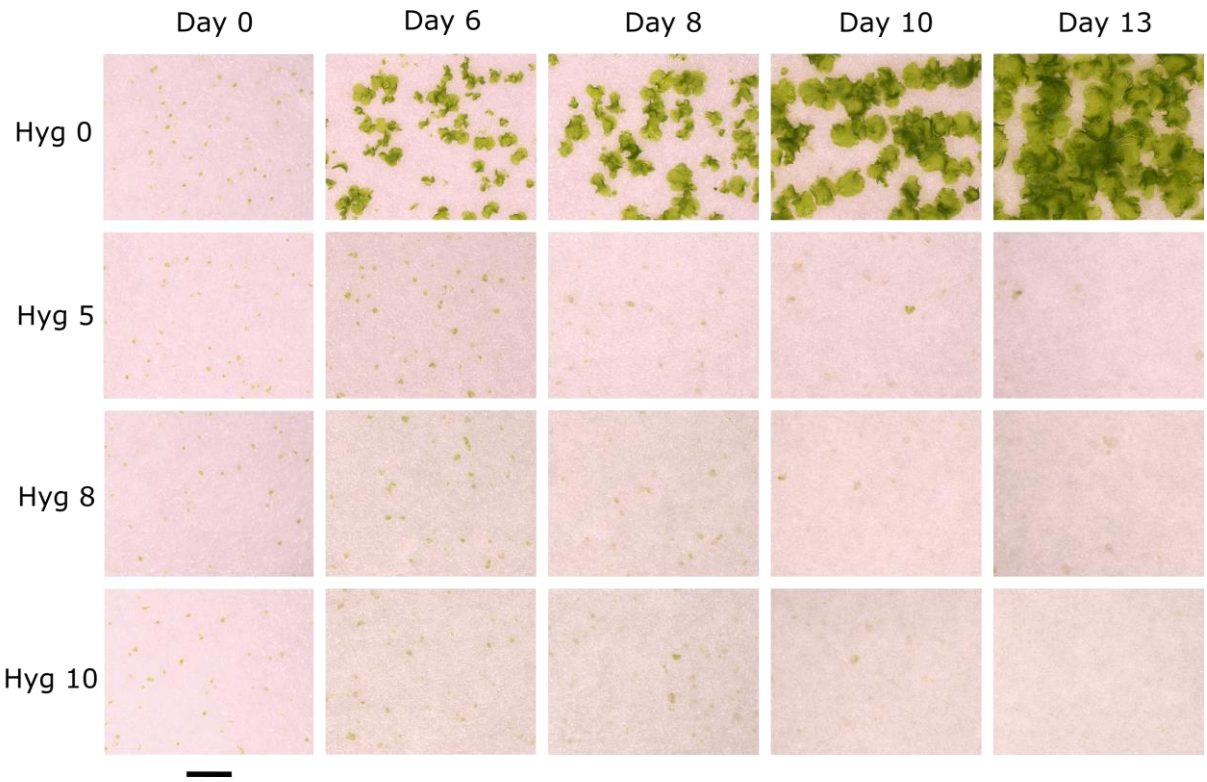
The attB1 and attB2 sites are a result of LR recombination and define the DNA transferred from pUC57 (the Mp*BT2* coding sequence).

Appendix 4: List of primers

Primer name	Target gene	Purpose	Primer sequence (5'-3')
NAg176	Between *35S* promoter and attB1	Sequencing LR gateway products	GCTCGAGGAATTCGGTACC
NAg177	Between *OCS* terminator and attB2	Sequencing LR gateway products	TCTAGAGGATCCAAGCTTATCGA
NA354	*CaMV35S* promoter	Screening LR gateway products and *BT2* over-expression lines	CAGAACTCGCCGTGAAGACT
NA355	*OCS* terminator	Sequencing LR gateway products	CGGACGGCCAATACTCAACT
NA428	*U6* promoter	Sequencing ligation and LR gateway products for CRISPR/Cas9	GCTCAGAGTGACGCAACTG
NA585	p$U6_{pro}$/pENTR	*35S*:*Venus-N7* cloning	ATGTTTTCCCAGTCACGACG
NA586	p$U6_{pro}$/pENTR	*35S*:*Venus-N7* cloning	CAGAGCTGCCAGGAAACAG
NA711	*Mapoly0010s0173*	Screening *bt2* mutants	GTTAGTGGGTTCGAGTTCGTG
NA712	*Mapoly0010s0173*	Screening *bt2* mutants and *BT2* over-expression lines	ACAACGTACCGTGAAGAGTAC
NA713	*Mapoly0002s0224*	Screening *fnsI* mutants	CGAGCTGAGTGTAGTGAGAG
NA714	*Mapoly0002s0224*	Screening *fnsI* mutants	GGAAGTTGTACCTGTATGTGTCG
NA717	*Venus-N7* CDS	Amplification from existing plasmid	CACCATGGTGAGCAAGGGCGAG
NA718	*Venus-N7* CDS	Amplification from existing plasmid	GAGGATCCTTACTCTTCTTCTTGATCAGCT

Appendix 5: Hygromycin kill curve

A hygromycin kill curve was performed to determine the optimal hygromycin concentration that should be used as a selection for *M*. *polymorpha* plants transformed with CRISPR/Cas9 constructs. Images of spore survival following treatment with 0, 5, 8 and 10 mg L^{-1} hygromycin were taken on day 0, 6, 8, 10 and 13 after treatment application (shown below).



Scale bar represents 500 µm.

The 5 mg L^{-1} hygromycin spores did not cause sufficient spore death as green spores were still present after 13 days. Whereas, the 10 mg L^{-1} hygromycin treatment caused complete spore death by day 13, meaning that some transformed spores that contain the hygromycin resistance gene could die. The 8 mg L^{-1} was chosen as the optimal hygromycin concentration because majority, but not all, of the spores died following the treatment. Thus, although some non-transformed escapes could be generated, none of the transformed spores will die.

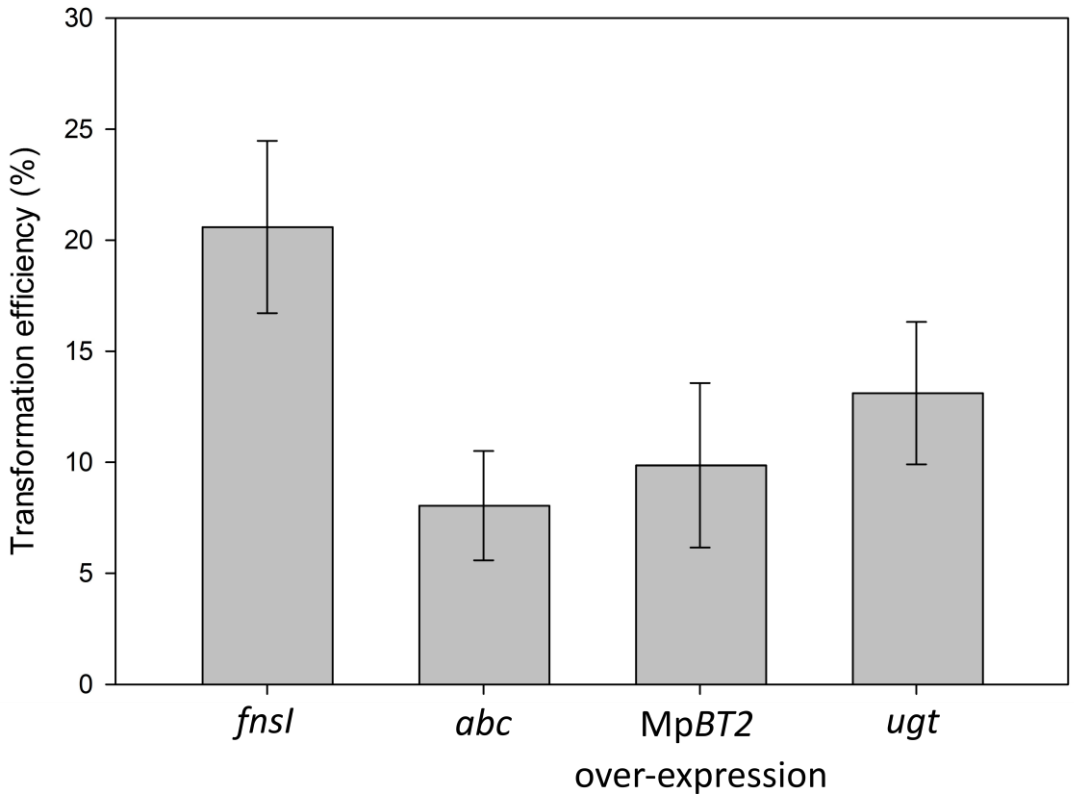
Appendix 6: Conversion of M. *polymorpha* gene names

The ID numbers of *Marchantia polymorpha* genes of interest used in this thesis were based on an earlier (but still good quality) version of the genome assembly (JGI3.1). The most recent version is MpTak_v6.1. Conversion of the gene names from the JGI3.1 genome version to the MpTak_v6.1 version is shown below.

Gene name	Gene ID JGI3.1 (old version)	Gene ID MpTak_v6.1 (new version)
Mp*FNSI*	*Mapoly0002s0224*	*Mp1g26540*
Mp*BT2*	*Mapoly0010s0173*	*Mp5g22830*
Mp*GLK*	*Mapoly0156s0007*	*Mp7g09740*
Mp*PDS*	*Mapoly0108s0060*	*Mp8g14330*

Appendix 7: Extra transformation efficiency data

Addition of 200 µM acetosyringone to the pre-culture media during *M*. *polymorpha* spore transformation achieved a transformation efficiency of 11% and 6% for *35S*:*Myb14* and *35S*:*GFP* lines, respectively. The transformation efficiency was additionally calculated for other spore transformations performed using 200 µM, including in the generation of *fnsI* mutants, Mp*BT2* over-expression plants, and mutants in candidate genes for MpABC and MpUGT proteins (not discussed in this thesis). The *fnsI* and *ugt* mutants had a higher transformation efficiency than those reported for *35S*:*Myb14* and *35S*:*GFP* lines, at 20% and 14%, respectively.



Error bars are the standard error of the means, n = 3.

Appendix 8: Phenotype of *M. polymorpha pds* mutants

CRISPR/Cas9 mutants were made in the *M. polymorpha Phytoene Desaturase* (*PDS*) gene (*Mapoly0108s0060*) (using the pRMY13 plasmid; see Appendix 3.2 for plasmid map) to try to produce a visually distinct phenotype. However, the resulting plants (shown below) had mostly wild-type (green) tissue, with white/translucent speckles present throughout the thallus that were assumed to be mutant tissue.

Scale bars represent 500 µm.

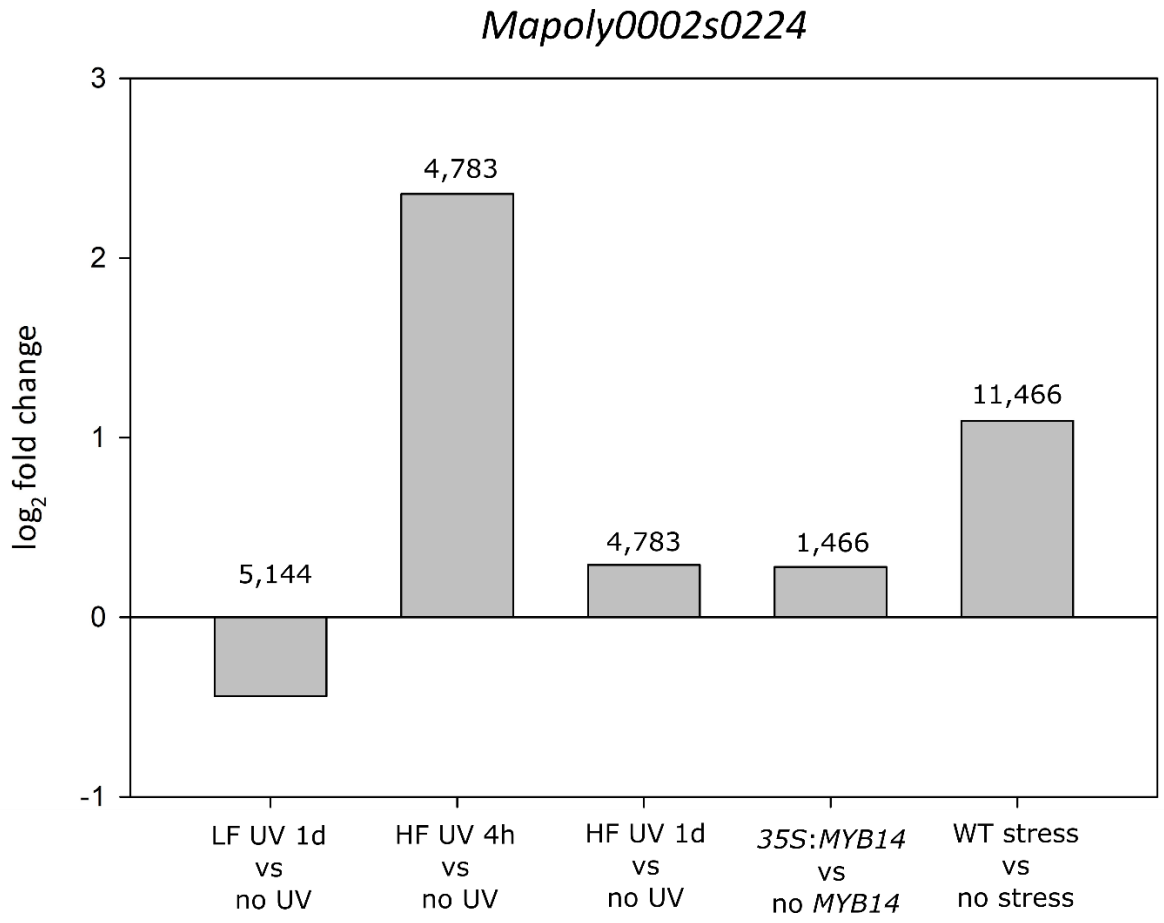
Appendix 9: RNA-sequencing data

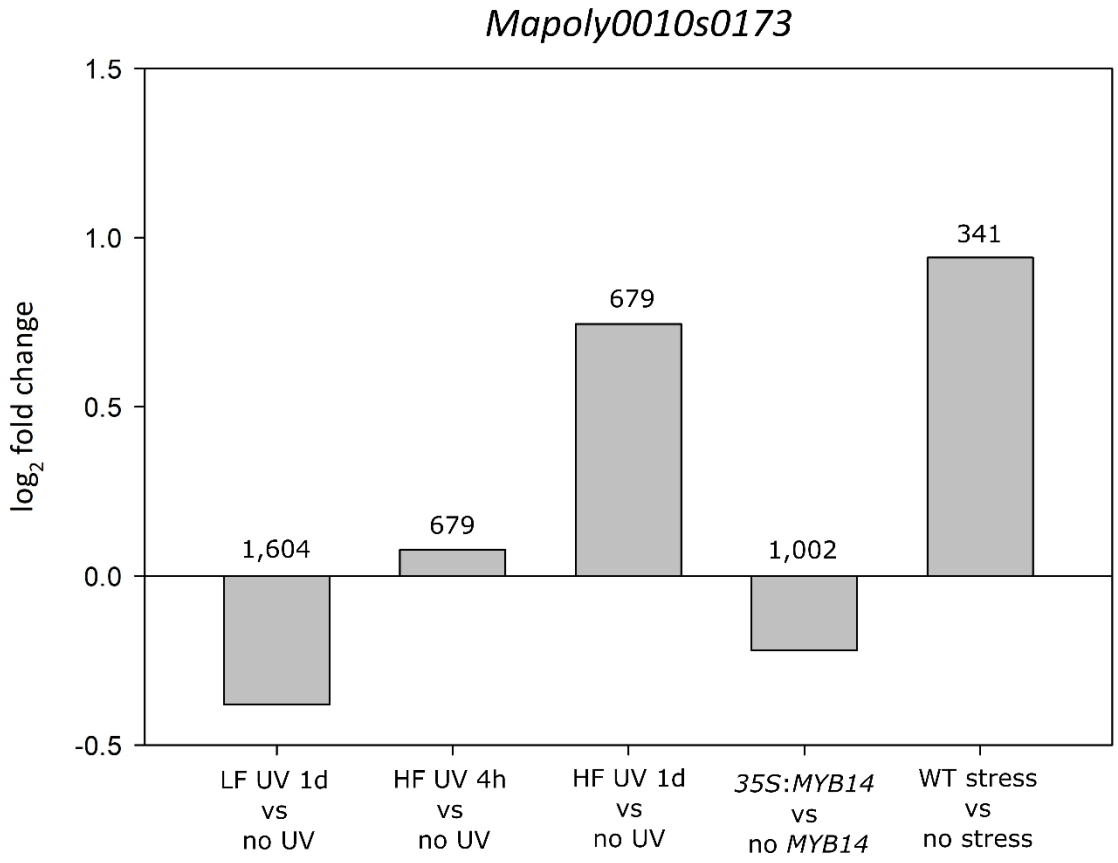
9.1 Mp*FNSI* and Mp*BT2* RNA-seq

The Mp*FNSI* candidate gene (*Mapoly0002s0224*) was of interest in this thesis because RNA-sequencing data showed the expression of the gene being upregulated under UV-B light conditions (~2-fold difference in expression between the UV and no UV treatments) that also induced the prodcution of flavones.

The Mp*BT2* candidate gene (*Mapoly0010s0173*) showed no significant differential expression under nutrient deprivation stress conditions, as the difference in expression between the stress and no stress treatments was only ~1-fold.

The RNA-seq data (unpublished data of Dr Yanfei Zhou) for these genes are shown below.





LF UV = exposure to low fluence UV-B light.

HF UV = exposure to high fluence UV-B light.

35S:*MYB14* = over-expression of the MpMYB14 gene required for auronidin biosynthesis.

WT stress = wild-type plants exposed to nutrient deprivation stress.

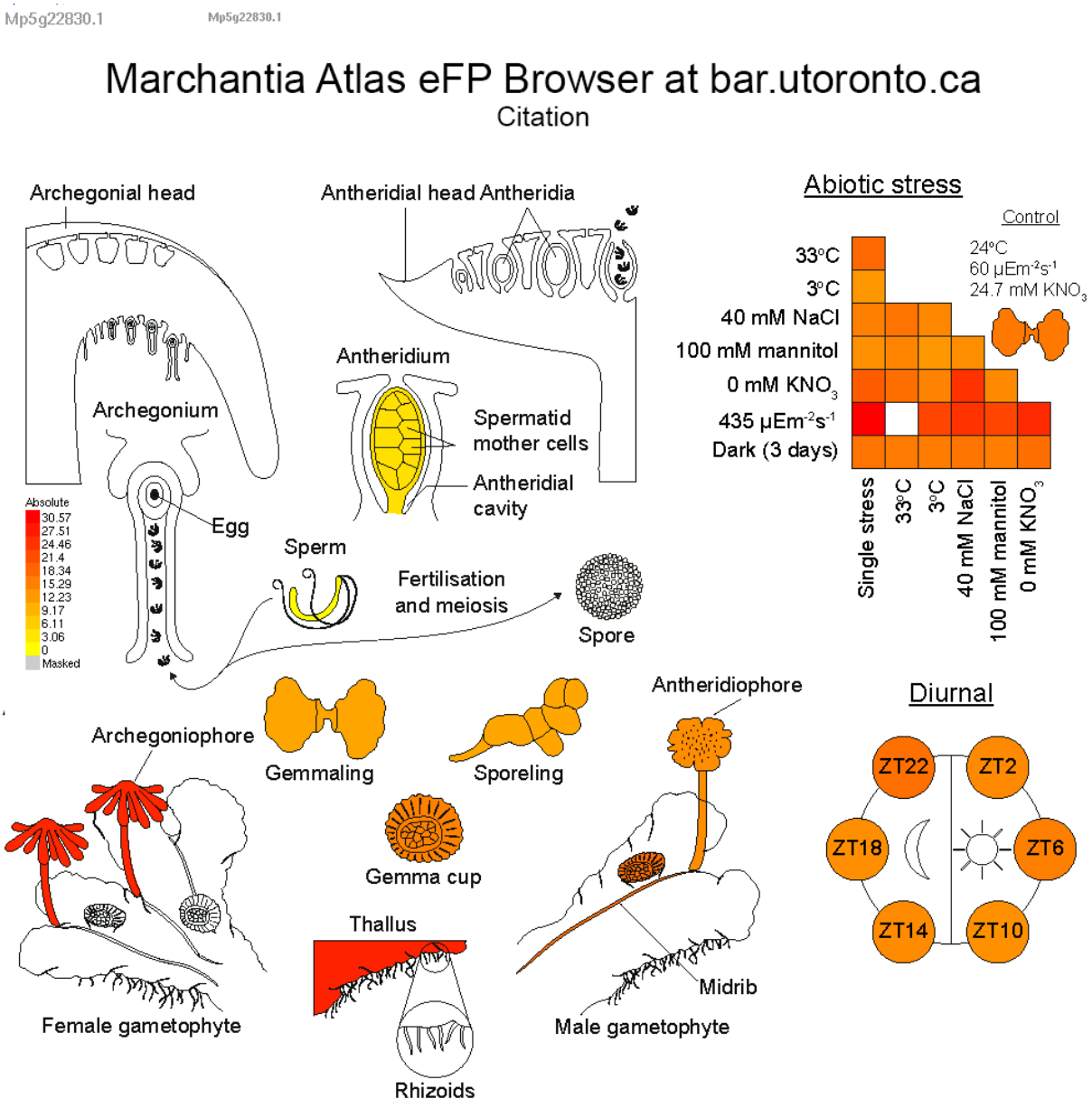
The log_2 fold change indicates how much the gene expression changed between the presence and absence of the treatment.

The base mean across the two paired treatments (in RPKM; Reads Per Kilobase per Million mapped reads) is shown above each bar, indicating the level of gene expression under the different conditions. The base means of the HF UV and no UV comparisons (4 h and 1 d) are the same because this is the mean between the 4 h and 1 d HF UV time points and the no UV.

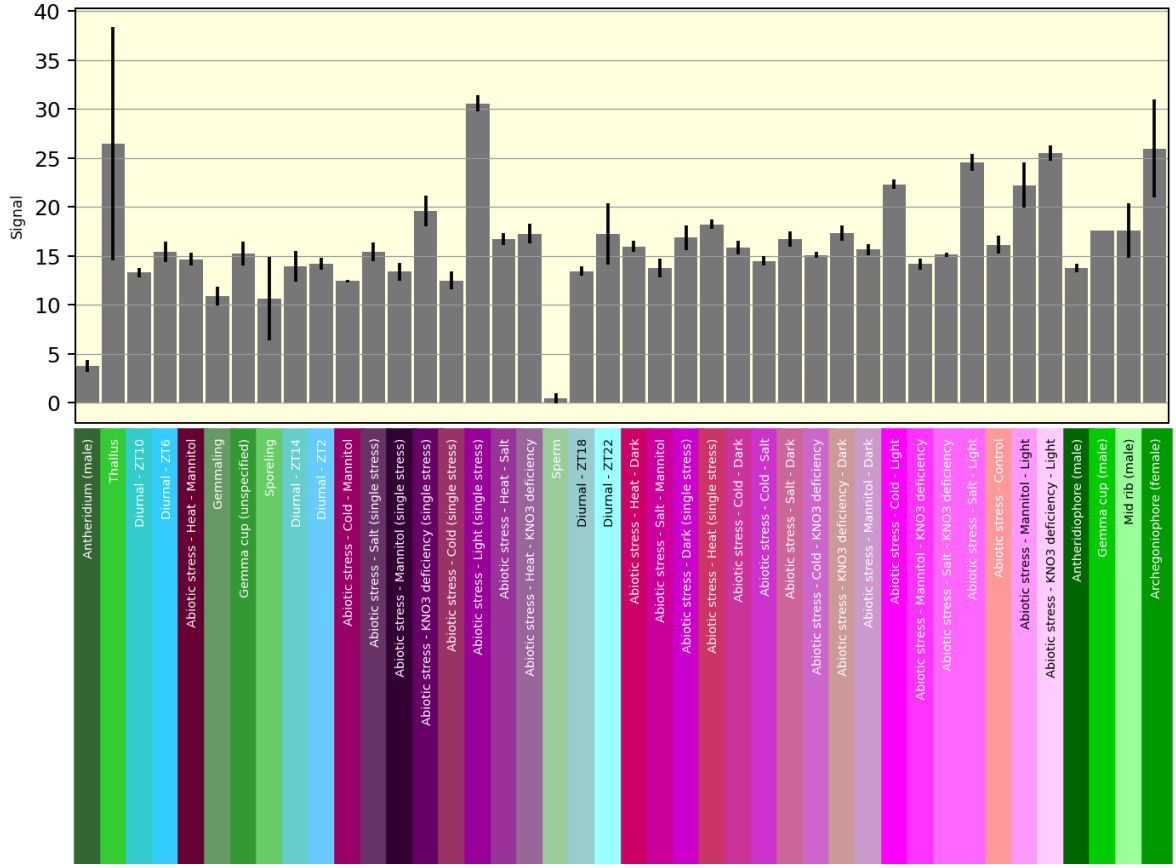
9.2 Additional Mp*BT2* RNA-seq

The Marchantia eFP Browser also shows RNA-sequencing data for the Mp*BT2* gene.

The below figure and following graphical results show the transcript abundance for Mp*BT2* in many tissue types and under different abiotic stress conditions. The expression is highly constitutive under most of the abiotic stresses tested, including under nutrient deprivation, as the expression is similar between the treatments and the control.



From Marchantia eFP Browser, University of Toronoto, Canada.



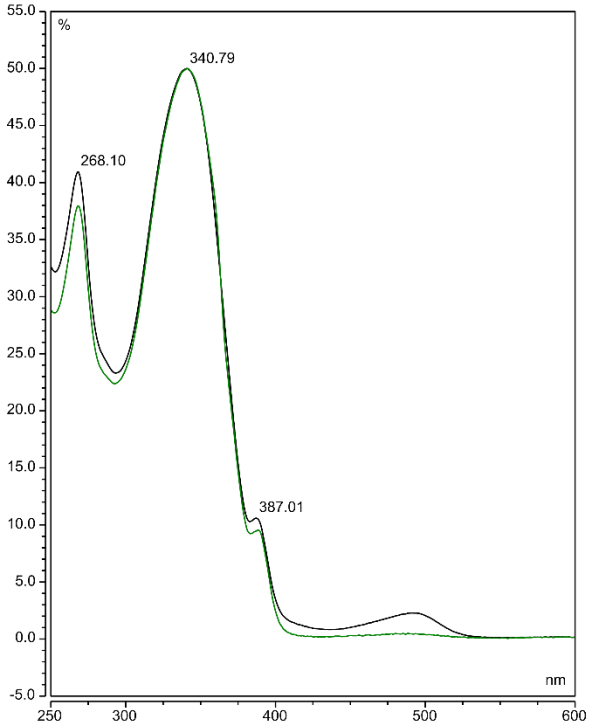
From Marchantia eFP Browser, University of Toronoto, Canada.

Appendix 10: Additional chemical data

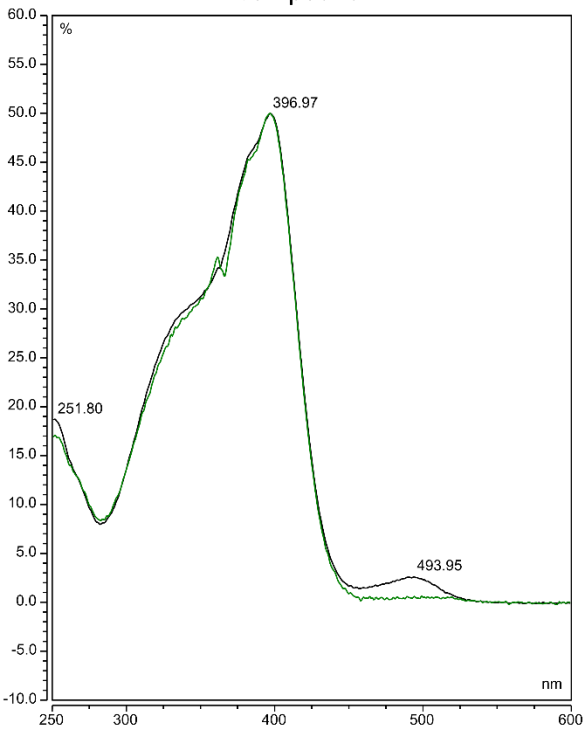
10.1 Stress-induced metabolites in *fnsI* and *chi* mutants

HPLC analysis of *M*. *polymorpha fnsI* and *chi* mutants grown under excess white light conditions determined that several metaabolites detected in these plants (A in *chi* and B in *fnsI*) have similar retention times to certain flavone compounds (Figure 4.6A). Compound A matches closely to luteolin-7,3'-di-*O*-glucuronide and Compound B to apigenin-7-*O*-glucuronide. However, the spectral properties of these compounds (shown below) indicate that they are not flavones.

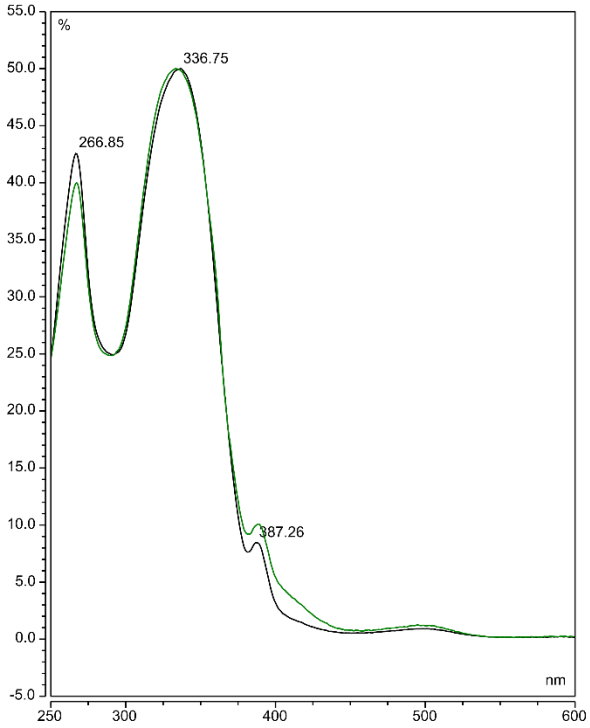
luteolin-7,3'-di-*O*-glucuronide



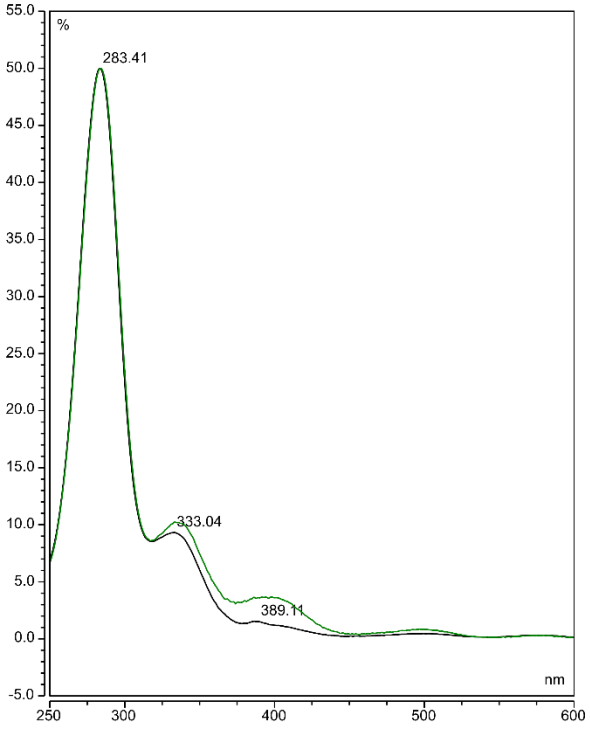
Compound A



apigenin-7-*O*-glucuronide

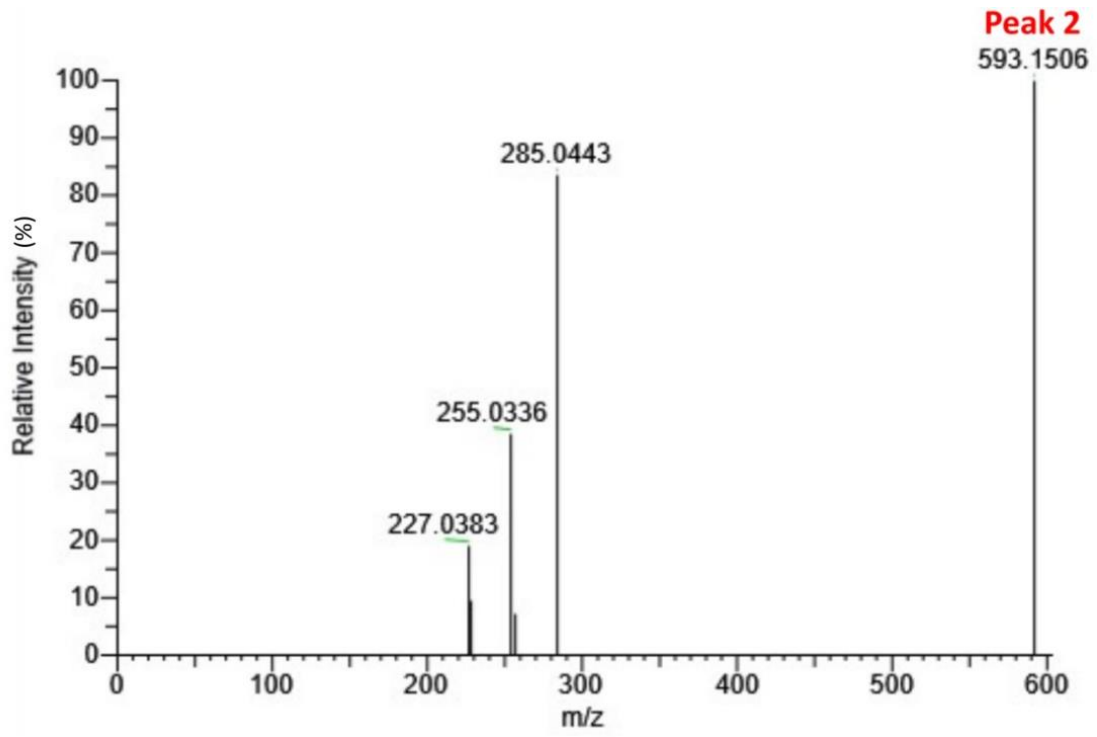


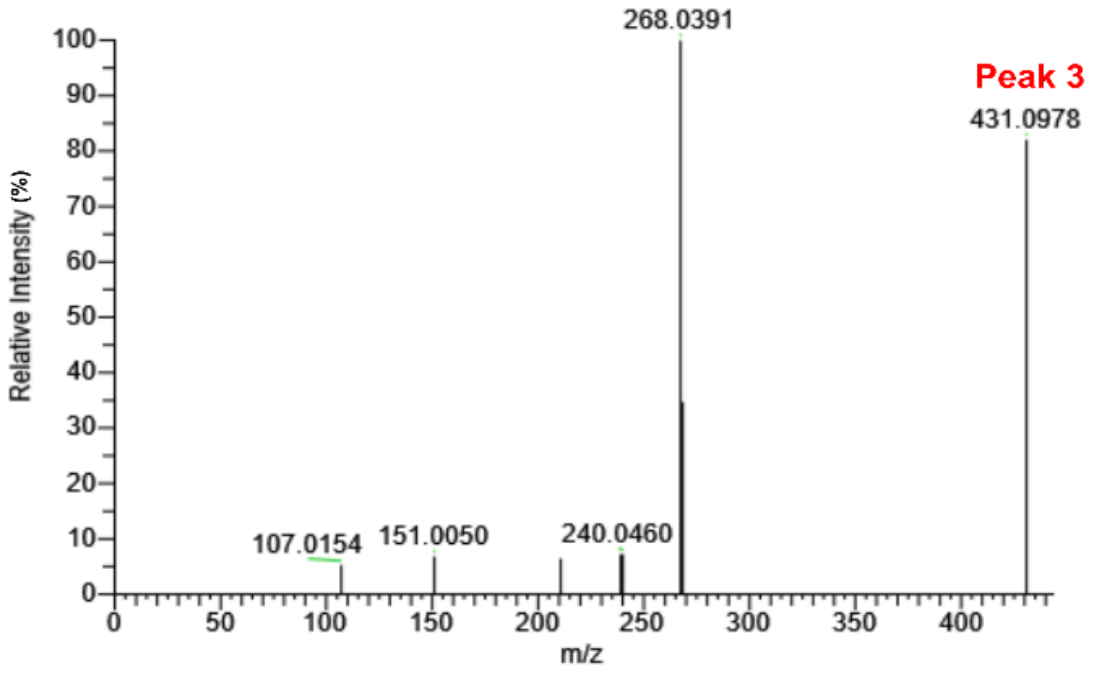
Compound B

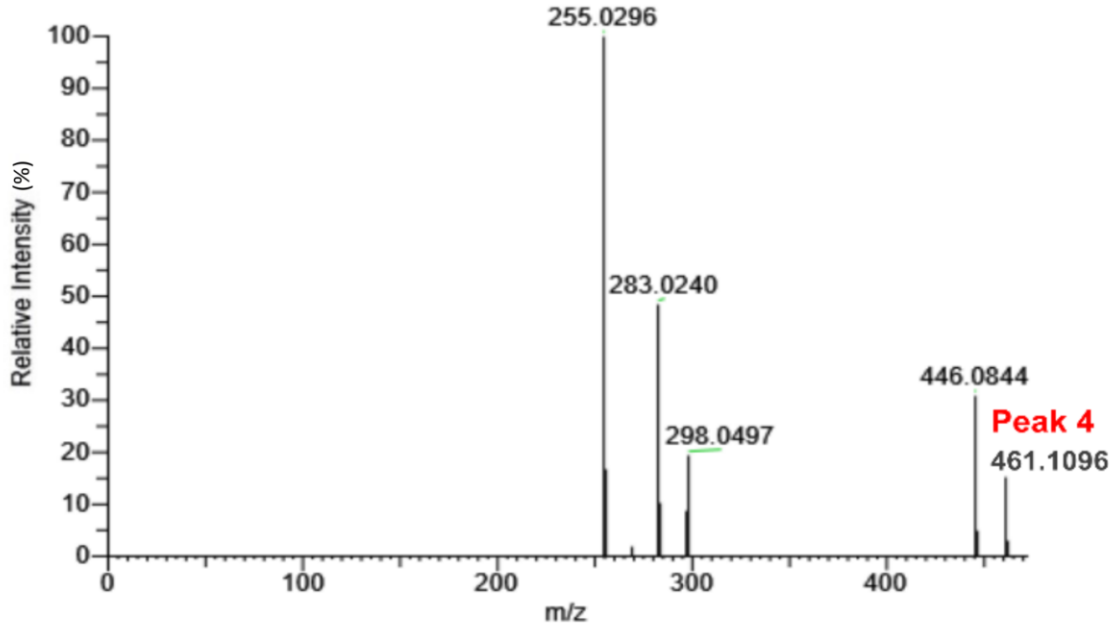


10.2 LC-HRAM-MS/MS identification of *N. tabacum* flavonols and flavones

Flavonols and flavones identified in *Nicotiana tabacum* (tobacco) plants over-expressing *Flavone synthase I* (*FNSI*) and At*MYB12* by liquid chromatography-mass spectrometry (LC-MS) are shown below.







m/z: mass-to-charge ratio.

Peak 1 was tentatively identified as quercetin-3-*O*-rutinoside (*m/z* 609.14). Data not shown because it could not be conclusively identified.

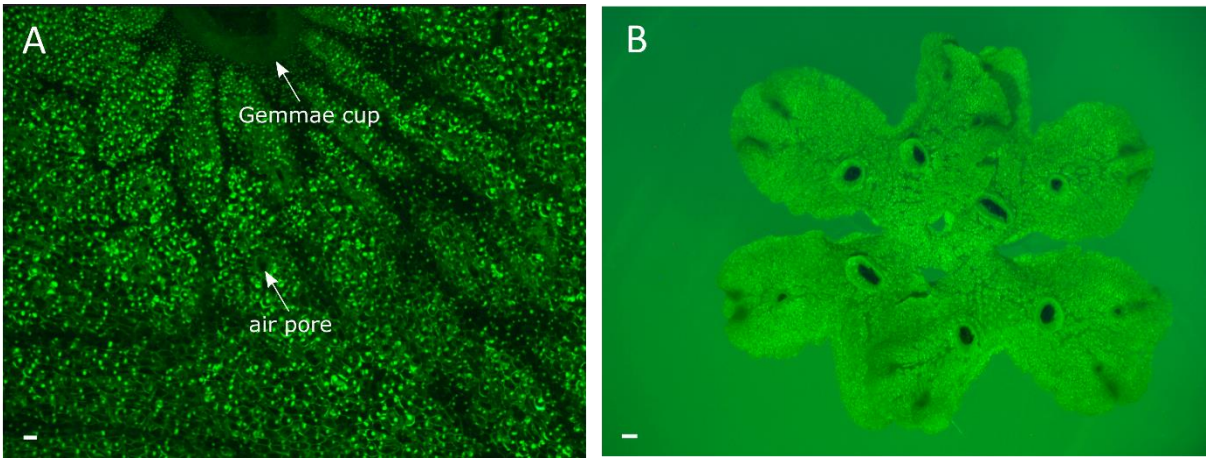
Peak 2 is kaempferol-3-*O*-rutinoside (*m/z* 593.15).

Peak 3 is apigenin-7-*O*-glucoside (*m/z* 431.0987)

Peak 4 is chrysoeriol-7-*O*-glucoside or luteolin-7-glucuronide (*m/z* 461.1096).

Appendix 11: Confirmation of fluorescence in *M. polymorpha 35S*:*VenusN7* plants

Fluorescence emission from the *35S*:Venus-N7 construct (was initially confirmed by transforming it into *M*. *polymorpha*. Detection of fluorescence under a microscope, using a GFP filter, is shown below.

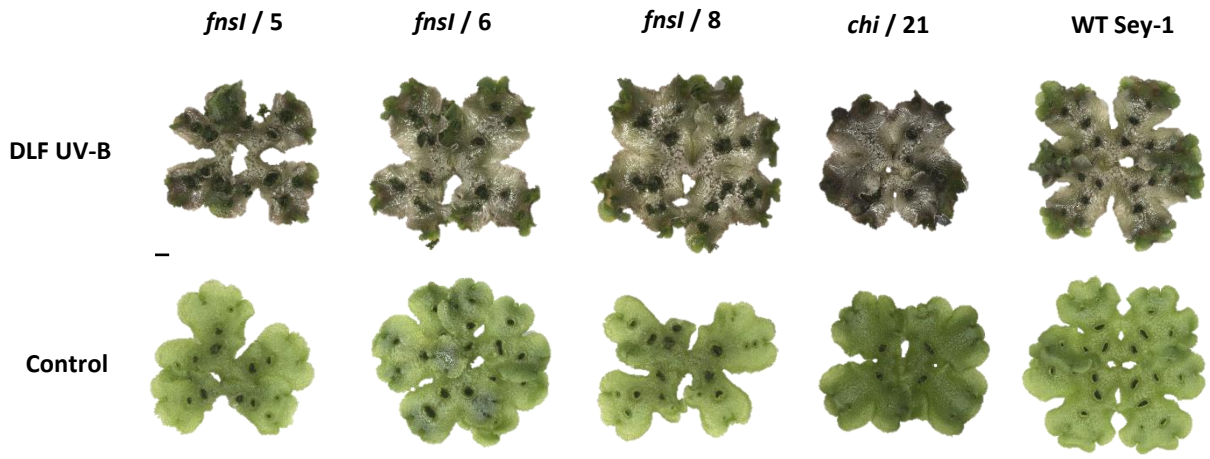


(A) Fluorescence localised to the nucleus of individual cells. Each spot of fluorescence indicates a single cell. The large circle at the top of the image is a gemmae cup and smaller circles are air pores (both indicated by arrows). Scale bar represents 100 µm.

(B) Fluorescence of the entire plant. Scale bar represents 1 mm.

Appendix 12: UV-B light-induced stress phenotype of *fnsI* mutants

To determine if a lack of flavones makes *Marchantia polymorpha* plants more sensitive to UV-B light, *fnsI* mutants (and *chi* mutant and wild-type (WT) Sey-1 controls) were exposed to daily low-fluence UV-B light (DLF UV-B) (as described in section 4.4). The resulting plant phenotypes, compared to control (no UV-B) conditions, are shown below.



Scale bar represents 2 mm.

Appendix 13: MpBT2 BLAST results

A BLAST search was performed of the *Arabidopsis* and apple BT2 proteins (AtBT2 and MdBT2, respectively) against the Marchantia genome, to determine if Marchantia contained any BT2 candidates. The best BLAST match was Mapoly0010s0173, which had a protein identity of 39% and 44% to the *Arabidopsis* and apple BT2 proteins. The second-best BLAST match had a significantly lower protein identity of 20% to both proteins, so is unlikely to be a BT2 homolog. A pairwise alignment of the protein identities described is shown below.

	AtBT2_NP_...	BT4_ARATH	Mapoly001...	Mapoly002...	MdBT2_XP_...	MdBT4_XP_...
AtBT2_NP_566902_AT3...		39%	39%	20%	57%	40%
BT4_ARATH	39%		49%	22%	41%	68%
Mapoly0010s0173_BT2	39%	49%		19%	44%	52%
Mapoly0021s0088.1_BTB	20%	22%	19%		20%	22%
MdBT2_XP_008374350	57%	41%	44%	20%		42%
MdBT4_XP_008341058	40%	68%	52%	22%	42%	

From Geneious version 2022.0, created by Biomatters.

MpBT2 = Mapoly0010s0173

AtBT2 = AtBT2_NP_56902

MdBT2 = MdBT2_XP_008374350

Other MpBT candidate = Mapoly0021s0088.1