

Evolutionary conservation of epigenetic control of plant sexual reproduction

Dissertation

zur Erlangung des Grades eines

Doktor der Naturwissenschaften

(Dr. rer. nat.)

des Fachbereichs Biologie der Philipps-Universität Marburg

vorgelegt von

MSc Anne Christina Genau

aus Frankfurt a.M.

Marburg, 2020

Die vorliegende Dissertation wurde von 12/2016 bis 03/2020 am Fachbereich Biologie, Philipps-Universität Marburg unter Leitung von Prof. Dr. Stefan A. Rensing angefertigt.

Tag der Einreichung: 29.03.2020 (Marburg)

Vom Fachbereich Biologie der Philipps-Universität Marburg (Hochschulkennziffer 1180) als Dissertation angenommen am 21.05.2020

Erstgutachter(in): Prof. Dr. Stefan A. Rensing

Zweitgutachter(in): Prof. Dr. Annette Becker (Justus-Liebig-Universität Gießen)

Drittgutachter(in): Prof. Dr. Uwe Maier

Viertgutachter(in): Prof. Dr. Lars M. Voll

Tag der Disputation: 18.06.2020 (Marburg)

Contents

Abbreviations	IV
Table of figures	VIII
List of tables	X
1 Summary	1
1 Zusammenfassung	2
2 Introduction	4
2.1 Haploid-dominant plants as model organisms for the water to land transition	6
2.2 Haplodiplontic life cycle of bryophytes	7
2.3 Evolution of land plant sexual reproduction	11
2.4 Transcription associated proteins (TAPs)	13
2.5 Genetic and epigenetic regulation of the alternation of generations	14
3 Aim of this study	19
4 Materials and Methods	20
4.1 Materials, buffers and media	20
4.2 Methods	23
4.2.1 Candidate gene search	23
4.2.2 Localisation and expression analyses	24
4.2.3 Culture conditions	24
4.2.4 Nucleic acid analysis	24
4.2.4.1 Genomic DNA isolation from plant material	24
4.2.4.2 Nucleic acid isolation "Quick and Dirty"	25
4.2.4.3 RNA isolation	25
4.2.4.4 Apices RNA-seq & analysis	25
4.2.4.5 cDNA synthesis	26
4.2.4.6 Polymerase chain reaction (PCR)	26
4.2.4.7 Gelectrophoresis	27
4.2.4.8 Extraction of PCR and restriction fragments	28
4.2.4.9 Plasmid isolation	28
4.2.4.9.1 Mini preparation	28
4.2.4.9.2 Midi preparation	28
4.2.4.9.3 Ethanol precipitation	29
4.2.5 Molecular cloning	29
4.2.5.1 Homologous recombination	29
4.2.5.2 Gateway cloning	30
4.2.5.3 Localisation analyses	31
4.2.5.4 CRISPR/Cas9	31

4.2.5.5	pJet cloning.....	36
4.2.5.6	Restriction.....	36
4.2.5.6.1	Restriction of DNA with compatible ends	36
4.2.5.6.2	Insert control.....	36
4.2.5.7	Ligation	37
4.2.5.8	Generation of RbCl ₂ -competent cells	37
4.2.5.9	Bacterial transformation.....	38
4.2.5.9.1	<i>E. coli</i>	38
4.2.5.9.2	<i>Agrobacterium tumefaciens</i>	38
4.2.5.10	Sequencing.....	38
4.2.6	Plant transfection	39
4.2.6.1	Transfection of <i>P. patens</i>	39
4.2.6.1.1	Linearisation of the constructs to be transformed	39
4.2.6.1.2	Liquid culture	39
4.2.6.1.3	Transfection.....	39
4.2.6.1.4	Regeneration and selection.....	40
4.2.6.2	Transfection of <i>M. polymorpha</i>	41
4.2.7	Phenotyping.....	42
4.2.7.1	Selfing and crossing analyses	42
4.2.7.2	Phenotypic analysis of gametangia and sporophytes	43
4.2.7.3	Spermatozoid analysis, DAPI and NAO staining.....	43
5	Results.....	44
5.1	Determination of conserved single copy TAPs.....	44
5.2	Candidates to study the alternation of generations	49
5.3	Localisation and expression	53
5.4	<i>P. patens</i> loss-of-function CRISPR/Cas9 mutants – loss of sporophytes.....	58
5.5	Test for male impairment.....	61
5.6	Spermatozoid analysis	64
5.7	<i>Pphag1</i> gametangia analyses	69
5.8	Summary of morphological analyses.....	73
5.9	RNA-seq analysis	73
5.10	The role of SWI3A/B and HAG1 in <i>M. polymorpha</i>	76
6	Discussion	80
6.1	TAPs as prime candidates for evolution under the Selected Single Copy Gene Hypothesis.....	80
6.2	<i>In vivo</i> confirmation of <i>in silico</i> analyses	81
6.3	Gene function during gametogenesis.....	82
6.3.1	<i>Ppswi3a/b</i>	83
6.3.2	<i>Pphag1</i>	84

6.4	Involvement of Pphag1 and Ppswi3a/b in transcriptional regulation in the alternation of generations	85
6.4.1	Ppswi3a/b	85
6.4.2	Pphag1	87
6.5	Role in all bryophyte lineages.....	88
6.6	Evolutionary conservation	90
6.7	Outlook	92
7	Conclusion	94
8	References	95
9	Supplement.....	110
	<i>Curriculum vitae</i>	153
	Publications and contributions.....	154
	Danksagung.....	155
	Eigenständigkeitserklärung	156

Abbreviations

%	percentage
°C	centigrade
µl	microlitre
µm	micrometer
µM	micromolar
<i>A. agrestis</i>	<i>Anthoceros agrestis</i>
<i>A. thaliana</i>	<i>Arabidopsis thaliana</i>
<i>A. tumefaciens</i>	<i>Agrobacterium tumefaciens</i>
AmpR	Ampicillin resistance
attL1/2	attachment sites
ATX1	Trithorax protein1
B	brown sporophyte
BNB	BONOBO
BoGa	Botanischer Garten
bom	basis of mobility
bp	base pair (s)
BP	biological process
BRG1	brahma-related gene 1
C	carbon
<i>C. reinhardtii</i>	<i>Chlamydomonas reinhardtii</i>
CaMV poly(A)	Cauliflower mosaic virus polyadenylation
Carb	Carbenicillin
CC	cellular component
CmR	Cas module-RAMP (Repeat-Associated Mysterious Proteins)
ColE1	colicin E1
CRC	Chromatin-remodelling complex
cs	crossed sporophytes
d	days
DAPI	4',6-diamidino-2-phenylindole
daw	days after watering
DE	Germany
DEG	differentially expressed gene
DNA	deoxyribonucleic acid
dNTP	deoxy nucleotide triphosphate
dpi	days post induction
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	<i>exempli gratia</i> (for example)
ES	early sporophyte
<i>et al.</i>	and colleagues
F1	first filial

FRA	France
fwd	forward
g	gram
g	gametophore
GC	generative cell
GCS1	GENERATIVE CELL-SPECIFIC 1
gDNA	genomic DNA
GEX2	GAMETE EXPRESSED 2
Gt	gigatons
h	hour
<i>H. sapiens</i>	<i>Homo sapiens</i>
H3K27	histone 3 lysine 27
HAP2	HAPLESS 2
HD	homeodomain
HF	high fidelity
hpt	Hygromycin phosphotransferase gene
HR	homologous recombination
Hyg	Hygromycin
i.e.	<i>id est</i> (that is)
KanR	Kanamycin resistance
kb	kilobase pair (s)
l	litre
LB	left border
LD	long day
LECA	last eukaryotic ancestor
M	molar
<i>M. musculus</i>	<i>Mus musculus</i>
<i>M. polymorpha</i>	<i>Marchantia polymorpha</i>
MA	mega-annum
MF	molecular function
mg	milligram
ml	millilitre
mM	millimolar
mm	millimeter
MRCA	most recent common ancestor
mRNA	messenger RNA
NAO	nonyl-acridine orange
NeoR	Neomycin resistance
NL	The Netherlands
NLS	nuclear localization sequence
nm	nanometer
NPP	net primary productivity

nw	no watering
<i>O. sativa</i>	<i>Oryza sativa</i>
OD	optical density
ori	origin of replication
Os	<i>Oryza sativa</i>
<i>P. patens</i>	<i>Physcomitrium patens</i>
PACRG	PARKIN COREGULATED GENE PROTEIN
PcG	Polycomb Group
PCR	Polymerase chain reaction
pH	pH value
PRC2	Polycomb complex 2
prom	promoter
PT	putative TAP
pVS1 oriV	Origin of replication from the plasmid pVS1
pVS1 RepA	Replication protein from the plasmid pVS1
pVS1 StaA	Stability protein from <i>Pseudomonas</i> plasmid pVS1
RB	right border
rbcS	ribulose-1,5-bisphosphate carboxylase
RCF	Relative Centrifugal Force
rev	reverse
Rif	Rifampicin
RKD	RWP-RK domain-containing
rpm	revolutions per minute
RT	room temperature
s	sporophyte
s/g	sporophyte per gametophore
SAR	Stramenopiles Alveolates Rhizaria
SC	sperm cell
SD	short day
SMC	spore mother cell
SmR	Aminoglycoside adenyltransferase; spectinomycin resistance
Spect	Spectinomycin
ssp	sub species
SWE	Sweden
TAE	TrisAcetatEDTA
TALE	three amino acid length extension
TAP	transcription associated protein
T-DNA	transfer-DNA
Term	terminator
TF	transcription factor
Tm	melting temperature
TR	transcriptional regulator

U	Unit
USA	United States of America
V	Volt
WGD	whole genome duplication
WT	wild type
x	times

Table of figures

Figure 1 Key events in the land plant evolution.....	5
Figure 2 Gradual change from the dominant generation in land plant evolution.	8
Figure 3 Cartoon of <i>P. patens</i> life cycle (sexual reproduction).	9
Figure 4 Life cycle of <i>M. polymorpha</i>	10
Figure 5 Male and female reproductive organs of angiosperms and gymnosperms.	13
Figure 6 Genetic and epigenetic regulation of land plant alternation of generations in an evolutionary context.	16
Figure 7 Control of gametophyte (n=haploid) development in <i>A. thaliana</i> and <i>M. polymorpha</i>	18
Figure 8 Knock-in vector PIG_AN.	30
Figure 9 pMpGE_En03 vector.	32
Figure 10 Detailed sequence view to visualise the principle of BsaI cut vector.	33
Figure 11 Vector cards of used destination vectors pMpGE010 and pMpGWB401.	35
Figure 12 Workflow of candidate gene search approach.	45
Figure 13 Sigma70-like phylogenetic analysis.	46
Figure 14 HAG1 phylogenetic analysis.	50
Figure 15 SWI3 phylogenetic analysis.	52
Figure 16 Expression profile of Pphag1 (Pp3c2_9560V3.1, histone acetyltransferase) and Ppswi3a/b (Pp3c2_24980V3.1, chromatin remodelling complex subunit).	53
Figure 17 Expression profile of Mphag1 (Mapoly0187s0003, histone acetyltransferase) and Mpswi3a/b (Mapoly0079s0052, chromatin remodelling complex subunit).	54
Figure 18 Expression profile of the developmental stages of Athag1 (histone acetyltransferase).	55
Figure 19 Expression profile of the developmental stages of Atswi3a and b (chromatin remodelling complex subunit).	56
Figure 20 Vector card of transiently expressing candidate gene Pphag1 with a C-terminal GFP tag under the control of the CABsh Promoter (Hiss <i>et al.</i> , 2017b).	57
Figure 21 <i>In vivo</i> localisation analysis.	58
Figure 22 Sporophytes per gametophore of loss-of-function mutants.	59
Figure 23 Phenotypic analysis of <i>swi3a/b</i> and <i>hag1</i> compared to Reute as control after watering.	60
Figure 24 Micrographs of a sporophyte of <i>swi3a/b</i> in brightfield as well as fluorescent light.	61
Figure 25 Crossing with a fluorescent male fertile strain to test for male impairment.	62
Figure 26 Crossing with a fluorescent male fertile strain to test for male impairment.	63
Figure 27 Swimming spermatozoids shortly after release.	64
Figure 28 Spermatozoid analysis of <i>swi3a/b</i> 21 dpi; mature antheridia of Reute and <i>swi3a/b</i> release spermatozoids.	65
Figure 29 Close-up of Figure 28.	66
Figure 30 Phasecontrast images of the control (left panel) and Ppswi3a/b (right panel).	66
Figure 31 Spermatozoid analysis of both candidate gene mutants via a double staining with DAPI and NAO.	67

Figure 32 Close-up of Figure 31.....	68
Figure 33 Archegonial development at 21, 22 and 28 dpi.....	69
Figure 34 Analysis of opened archegonial neck cells.	70
Figure 35 Analysis of opened antheridia.	71
Figure 36 <i>Pphag1</i> antheridia development 28 dpi.....	72
Figure 37 Summary of phenotypical analyses throughout the life cycle with emphasis on sexual reproduction.....	73
Figure 38 RNA-seq analysis.....	74
Figure 39 STRING network of AtHAG1 and AtTRAUCO.	75
Figure 40 Female plants of <i>M. polymorpha</i> fulfilled successfully the whole life cycle.....	76
Figure 41 Mutant analysis during gametangiophore induction.....	77
Figure 42 Phenotyping in petri dishes.	78
Figure 43 Phenotyping in ECO2boxes.	79
Figure 44 ARID TAP family phylogenetic analysis.	117
Figure 45 GRAS TAP family phylogenetic analysis.	118
Figure 46 CCAAT_HAP3 TAP family phylogenetic analysis.....	119
Figure 47 BES1 TAP family phylogenetic analysis.	119
Figure 48 Med6 TAP family phylogenetic analysis.	120
Figure 49 Med7 TAP family phylogenetic analysis.	120
Figure 50 HMG TAP family phylogenetic analysis.	121
Figure 51 O-FucT TAP family phylogenetic analysis.....	122
Figure 52 MADS TAP family phylogenetic analysis.....	123
Figure 53 WRKY TAP family phylogenetic analysis.....	124
Figure 54 TUB TAP family phylogenetic analysis.	125
Figure 55 tify TAP family phylogenetic analysis.	126
Figure 56 bHLH TAP family phylogenetic analysis.	127
Figure 57 C2H2 TAP family phylogenetic analysis.	128
Figure 58 C3H TAP family phylogenetic analysis.	129
Figure 59 IWS1 TAP family phylogenetic analysis.....	130
Figure 60 Jumonji_Other TAP family phylogenetic analysis.....	131
Figure 61 PcG_MSI TAP family phylogenetic analysis.....	132
Figure 62 TANGO2 TAP family phylogenetic analysis.....	132
Figure 63 MYB-related TAP family phylogenetic analysis.	133
Figure 64 GNAT TAP family phylogenetic analysis.....	134
Figure 65 FHA TAP family phylogenetic analysis.	135
Figure 66 AT5G13780 phylogenetic analysis.....	136
Figure 67 AT5G47790 phylogenetic analysis.....	137
Figure 68 AT5G06550 phylogenetic analysis.....	138
Figure 69 AT5G06110 phylogenetic analysis.....	139
Figure 70 AT5G03500 phylogenetic analysis.....	140

Figure 71 AT3G54610 phylogenetic analysis.....	141
Figure 72 AT2G47900 phylogenetic analysis.....	142
Figure 73 AT2G47620 phylogenetic analysis.....	143
Figure 74 AT2G39020 phylogenetic analysis.....	144
Figure 75 AT1G72050 phylogenetic analysis.....	145
Figure 76 AT1G64860 phylogenetic analysis.....	146
Figure 77 Sporophytes per gametophore of <i>Pphag1</i>	147
Figure 78 Sporophytes per gametophore of <i>Ppswi3a/b</i>	147
Figure 79 Crossing with a male fluorescent fertile strain to test for male impairment.....	148
Figure 80 Crossing with a male fluorescent fertile strain to test for male impairment.....	148
Figure 81 Crossing with a male infertile strain to test for male impairment.	149
Figure 82 Comparison of all RNA-seq results within both mutants and both conditions.	149
Figure 83 GO enrichment of DEGs up-regulated in <i>Ppswi3a/b</i> two days after watering compared to the control.	150
Figure 84 Vector card of pMETA.....	151
Figure 85 Vector card of pPGX8.	151
Figure 86 Schematic overview of <i>P. patens</i> complementation analyses.	152

List of tables

Table 1 Buffers, composition and their application.	20
Table 2 Media, composition and their application.	21
Table 3 Timetable for <i>A. tumefaciens</i> mediated <i>M. polymorpha</i> transfection.	41
Table 4 Candidate gene selection analysed in terms of single copy status via Bayesian inference. ...	47
Table 5 Overview of the used primers. Supplier: Sigma Aldrich (St. Louis, USA).	110
Table 6 Chemicals and supplier.	112
Table 7 Materials, devices and supplier.	115

Summary

1 Summary

During the course of evolution, land plants evolved a peculiar haplodiplontic life cycle in which both the haploid gametophyte and the diploid sporophyte are multicellular. Within the life cycle, the phase transition between these two stages is called the alternation of generations. However, within land plants the dominant generation varies. Following the alternation of generations on molecular level is difficult to analyse in flowering plants due to their diploid-dominant generation and their highly reduced gametophytic generation. Bryophytes, in contrast, are haploid-dominant plants, have more easily tractable (morphologically larger) generations and share with flowering plants the fundamental regulatory networks for switching between vegetative and reproductive growth. This switch is most often regulated by transcription associated proteins (TAPs). Combining an orthology detection tool (proteinortho), protein-family categorization (TAPscan), literature search and phylogenetic inference a set of candidate genes, representing single copy TAPs potentially involved in embryo development, was identified. In bryophytes the knowledge of sexual reproduction and evolution of plant sexual reproduction is limited as compared to seed plants. Even though, bryophyte plants defective in gametangiogenesis or embryo/sporophyte development often can be grown as gametophytic tissue and vegetatively propagated contrary to flowering plants. Two master regulators (*swi3a/b*, *hag1*) were identified. It could be shown that they are indeed crucial for sexual reproduction and thus the knowledge about the regulation of sexual reproduction in bryophytes could be amended. Deletion of the respective genes led to a loss of sporophyte development due to an impairment in the male germ line in the latest steps of spermatozoid ripening on the part of *swi3a/b*, and an impairment in the male and female germ lines (disturbed ripening) on the part of *hag1*. Hag1 is a highly conserved gene throughout all eukaryotic kingdoms, which was proven by phylogenetic analyses and underlined by mutants impaired in the male germ line in moss as well as flowering plants. In contrast, the SWI3 family diversified during land plant evolution (one SWI3 gene in algae, four SWI3 genes in *Arabidopsis thaliana*). That was underlined by the phenotypical analyses in *Physcomitrium patens* and *A. thaliana*, which showed a shift from a male exclusive role of *swi3a/b* in *P. patens* towards an involvement in male and female gametogenesis as well as embryogenesis in *A. thaliana*. These findings underline the hypothesis that on one hand TAPs can be highly conserved and can serve as single copy phylogenetic markers, whereas on the other hand TAP families expanded with plant complexity as drivers for morphogenic evolution. All in all, TAPs are important developmental regulators, that can be used for analysing the plants' potential to adapt to new conditions.

1 Zusammenfassung

Im Laufe der Evolution entwickelten Landpflanzen einen haplodiplontischen Lebenszyklus, in dem sowohl der haploide Gametophyt als auch der diploide Sporophyt mehrzellig sind. Der Phasenübergang zwischen diesen beiden Stadien des Lebenszyklus wird als Generationswechsel bezeichnet. Innerhalb der Landpflanzen variiert die dominante Generation. Die Analyse des Generationenwechsels auf molekularer Ebene ist bei Blütenpflanzen aufgrund ihrer dominanten diploiden Generation und ihrer stark reduzierten gametophytischen Generation schwierig. Bryophyten hingegen sind haploid dominante Pflanzen, haben leichter untersuchbare (morphologisch größere) Generationen und haben mit den Blütenpflanzen die grundlegenden regulatorischen Netzwerke für den Wechsel zwischen vegetativem und reproduktivem Wachstum gemeinsam. Dieser Wechsel wird häufig durch transkriptionsassoziierte Proteine (TAPs) reguliert. Durch die Kombination aus einem orthologermittlungs Programm (Proteinortho), Proteinfamilien Kategorisierung (TAPscan), Literatursuche und phylogenetischer Inferenz konnte eine Reihe von Kandidatengenen identifiziert werden. Diese repräsentieren TAPs mit nur einer Kopie, welche möglicherweise in die Embryonalentwicklung involviert sind. Bei Bryophyten ist das Wissen über sexuelle Reproduktion und die Evolution derselbigen im Vergleich zu Samenpflanzen begrenzt. Bryophyten, welche Defekte in der Gametangiogenese oder Embryo-/Sporophytenentwicklung zeigen, können als gametophytisches Gewebe angezogen werden und im Gegensatz zu Blütenpflanzen vegetativ vermehrt werden. Es wurden zwei Masterregulatoren (*swi3a/b*, *hag1*) identifiziert und für diese wurde gezeigt, dass sie essentiell für die sexuelle Reproduktion sind. Folglich wurde das Wissen über die Regulation eben jener bei Bryophyten erweitert. *Loss-of-function* Mutationen führten zu einem Verlust der Sporophytenentwicklung bedingt durch eine Beeinträchtigung der männlichen Keimbahn in den finalen Schritten der Spermatozoidreifung seitens *swi3a/b* und eine Beeinträchtigung der männlichen und weiblichen Keimbahn (gestörte Reifung) seitens *hag1*. Hag1 ist ein hochkonserviertes Gen in allen eukaryotischen Reichen, was durch phylogenetische Analysen impliziert und durch Mutanten, die in der männlichen Keimbahn sowohl in Moosen als auch in Blütenpflanzen beeinträchtigt sind, unterstrichen wurde. Im Gegensatz dazu diversifizierte die SWI3-Familie innerhalb der Evolution von Landpflanzen (ein SWI3-Gen in Algen, vier SWI3-Gene in *A. thaliana*). Dies wurde durch phänotypische Untersuchungen bei *P. patens* und *A. thaliana* verdeutlicht, die eine Verschiebung von einer ausschließlichen Kontrolle der männlichen Keimbahn von *swi3a/b* bei *P. patens* zu einer Beteiligung an der männlichen und weiblichen Gametogenese sowie der Embryogenese bei *A. thaliana* zeigten. Diese Analysen unterstreichen die Hypothese, dass einerseits TAPs hoch konserviert sind und als taxonomische Marker herangezogen werden können, andererseits TAP Familien im Zuge der

Zusammenfassung

Evolution morphologischer Komplexität expandierten. Zusammengefasst sind TAPs wichtige Entwicklungsregulatoren, die für das Verständnis des Anpassungspotentials von Pflanzen an neue Bedingungen von besonderem Interesse sind.

2 Introduction

The plant terrestrialisation 500 million years ago was one of the key events in evolutionary history. Plants had to face barren rock. It was hypothesised that filamentous freshwater algae (streptophyte algae) are sister group to all land plants (Figure 1) and performed the transition from water to land after adapting to the new conditions (Delwiche and Cooper, 2015; Wickett *et al.*, 2014). Photosynthesis carried out by cyanobacteria (3,000 Ma ago) and later eukaryotic algae (1,500 Ma ago), led to an increase of atmospheric oxygen (Lyons *et al.*, 2014). The greening of the planet resulted in tremendous changes in the atmosphere (increase in oxygen) as well as the earth (weathering) (Lenton *et al.*, 2016; Porada *et al.*, 2016). It is highly likely that mutualistic microbes facilitated the water to land transition in terms of water and nutrient uptake (Delaux *et al.*, 2015; Field *et al.*, 2015; Field *et al.*, 2016). The last common ancestor had to adapt to a completely new environment, which was associated with new biotic and abiotic stresses like drought, radiation, cell wall architecture, nutrient acquisition and sexual reproduction without constant water supply. The latter led to the evolution of the recent haplodiplontic life cycle which requires complex transcriptional regulation. The control of the gene expression related to the alternation of generations, namely the phase transition of haploid gametophyte and diploid sporophyte is carried out by TAPs, which raises with plant complexity (Levine and Tjian, 2003). The analysis of how plants dealt with these challenges is especially interesting in times of climate change, where plants have to face again tremendous changes like drought or water deprivation. The life on land and the therewith affiliated changes necessitated a highly controlled gene regulation, which was studied mainly in flowering plants. Bryophytes as sister group to streptophyte algae and streptophyte algae themselves are therefore optimal models in evo-devo approaches to analyse the origin and evolution of genes controlling the plant development (Rensing, 2017, 2018).

Introduction

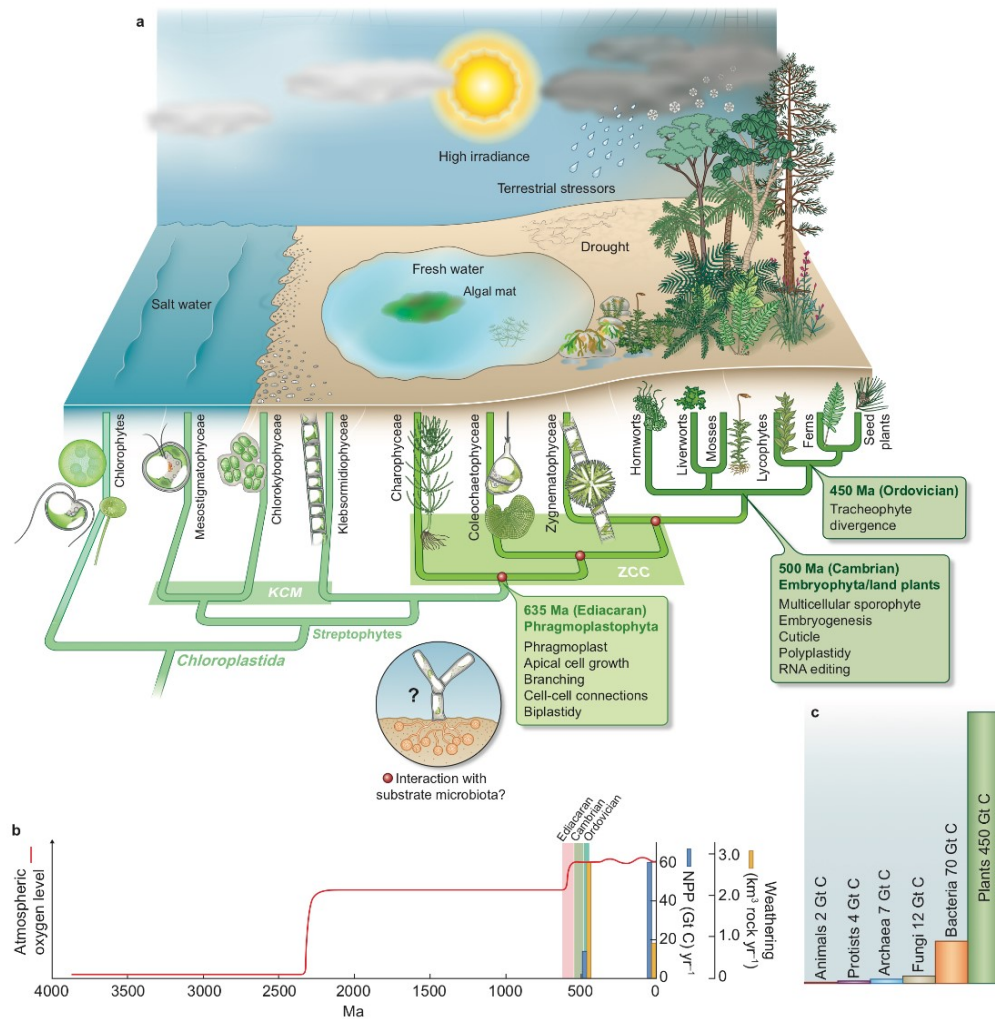


Figure 1 Key events in the land plant evolution.

During land plant evolution, the conquerors had to face a variety of new conditions. The land plant evolution is related to the rise of atmospheric oxygen and terrestrial weathering.

a) The last common ancestor of chlorophytes and streptophytes probably lived 1 billion years

ago. Streptophytes comprise the monophyletic embryophytes, the land plants, as well as the paraphyletic streptophyte algae. The streptophyte algae are classified into two grades, the KCM grade consisting of Klebsormidiophyceae, Chlorokybophyceae, and Mesostigmatophyceae as well as the ZCC grade comprising Zygnematophyceae, Coleochaetophyceae, and Charophyceae. Chlorophyte algae can be found in marine, freshwater and terrestrial habitats, whereas streptophyte algae live in freshwater and terrestrial environments. Although all chloroplastida can colonise terrestrial habitats, only the land plants dominate the terrestrial ground in the whole lineage, in terms of biomass and species richness. The most widely accepted hypothesis of the terrestrialisation of land by plants is an interaction of an organism with branching filaments or pseudoparenchymatous growth with mutual microbiota (potentially the ancestors of mycorrhizal fungi). The algal ancestor of the land plants faced terrestrial abiotic stresses like drought/desiccation, high ultraviolet and photosynthetic irradiance as well as temperature changes, so that it seems probable that the ancestor already had the genetic toolkit and therefore the appropriate physiology to assert itself. Figure modified after (de Vries and Archibald, 2018). The topology and divergence times were taken from (Morris *et al.*, 2018; Wickett *et al.*, 2014). The grene boxes indicate newly gained traits throughout Phragmoplastophyta and Embryophyta evolution. b) The increase of atmospheric oxygen concentration to the present level coincided with the arise of land plants. The blue and yellow bars show the net primary productivity (NPP) as well as the weathering of rock for Ordovician and today's flora. Figure according to data from (Lenton *et al.*, 2016; Lyons *et al.*, 2014; Porada *et al.*, 2016). c) Comparison of taxonomic groups in terms of biomass in gigatons (Gt) carbon (C), which raised in plants from 133 Gt C in the Ordovician towards 450 Gt C in these days. Data was taken from (Bar-On *et al.*, 2018).

Introduction

2.1 Haploid-dominant plants as model organisms for the water to land transition

Genome sequencing enables comparative genomics and phylogenetic inferences. Apart from that, it is also the basis for evolutionary approaches. *A. thaliana* was the first plant genome that was sequenced in 2000 (Arabidopsis Genome, 2000). The unicellular green alga *Chlamydomonas reinhardtii* was sequenced in 2007 (Merchant *et al.*, 2007). After the sequencing of the moss *P. patens* as representative of the bryophytes in 2008 (Rensing *et al.*, 2008) evo-devo studies were feasible, since the three genomes span plant evolution as well as the transition from water to land. The genetically closest land plants to algae are the bryophytes comprising mosses, liverworts and hornworts, which most likely represent best the life that evolved first on land (Kenrick, 2017). In the past years, additional plant genomes have been sequenced, which were first of all mainly angiosperms, neglecting other species-rich and less species-rich lineages, that are essential for inferring proper evolutionary analyses to avoid results coming from sample bias (Delaux *et al.*, 2019; Rensing, 2017). In the meantime, it has been recognised that for inferring proper phylogenetic analyses the sequencing of genomes spanning the whole taxonomic variety is of substantial need (Bowman *et al.*, 2017; Cheng *et al.*, 2019; Guan *et al.*, 2016; Nishiyama *et al.*, 2018; Wang *et al.*, 2020; Zhao *et al.*, 2019).

P. patens is especially of interest, due to its phylogenetic position and body plan. It serves as a model for developmental analyses (Bowman *et al.*, 2016; Sakakibara *et al.*, 2013; Sakakibara *et al.*, 2014), as both generations are easily accessible (Rensing, 2017) and evolutionary approaches as well as functional and comparative genomics are feasible (Cove *et al.*, 2006; Decker *et al.*, 2006; Quatrano *et al.*, 2007; Rensing, 2017, 2018).

P. patens can be propagated vegetatively as well as sexually (Cove, 2005; Hohe A., 2002; Strotbek *et al.*, 2013) and regenerate from a single cell (protoplast) to a whole plant (Boyd *et al.*, 1988). Apart from the genome that is fully sequenced (Rensing *et al.*, 2008), *P. patens* can be easily genetically modified by gene targeting via homologous recombination (Schaefer, 2001; Schaefer and Zryd, 1997) as well as CRISPR/Cas9 (Collonnier *et al.*, 2017), so that forward and reverse genetics can be applied (Perroud *et al.*, 2011). Apart from that, there is a broad pallet of expression data which can be accessed for gene function analysis (Fernandez-Pozo *et al.*, 2019). On top of that, histone marks were already analysed in terms of drought stress (Widiez *et al.*, 2014).

The water to land transition was an enormous challenge that plants had to cope with. Understanding how plants dealt with the adaptations in plant terrestrialisation is especially of interest in times of global warming, since both events are characterised by changing environmental conditions which require plants to adapt. Therefore, studying the past is also of

Introduction

importance for the future to learn how plants will potentially adapt to new conditions, e.g. drought (Rensing, 2017).

In addition, *P. patens* is not only interesting for finding answers to questions in life science e.g. development, architecture, biochemistry, reproduction or stress tolerance but also for biotechnological innovations. Mosses can be used as factories producing metabolites or pharmaceuticals. Moreover, moss extracts can be applied in cosmetic industry. Because of their special way to cope with habitats in terms of biotic and abiotic stresses, studying these mechanisms is important for applications in crop science. Apart from biotechnological applications, mosses can also serve as biomonitors of environmental pollution and CO₂-neutral “farming on rewetted bogs” (Decker and Reski, 2019)

2.2 Haplodiplontic life cycle of bryophytes

The plant haplodiplontic life cycle is characterised by the alternation of generations, the change of multicellular haploid and multicellular diploid phases (Hofmeister, 1851), whereas charophyte algae perform a haplontic life cycle with only the gametophyte being multicellular. The sexual generation (haploid gametophyte) develops sperm and egg cells (gametes), while the asexual generation (diploid sporophyte) undergoes meiosis, which leads to spore development (from which in turn the gametophyte develops). In bryophytes, a sister lineage to vascular plants, the haploid generation is dominant (Figure 2). In comparison, angiosperms for example are characterised by their sporophyte-dominant life cycle, having reduced their haploid generation to just a few hidden cells, pollen and embryo sac (Bowman *et al.*, 2016; Hackenberg and Twell, 2019).

Introduction

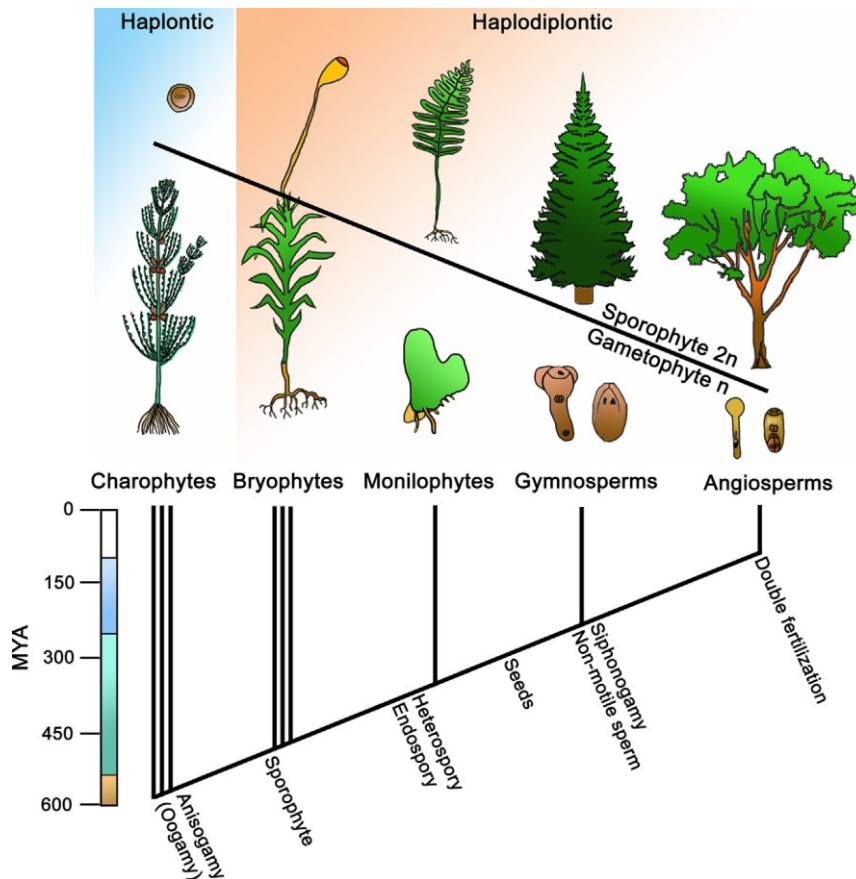


Figure 2 Gradual change from the dominant generation in land plant evolution.

The upper part shows the gradual change of the dominant generation from haploid-dominant to a diploid-dominant generation. In charophyte algae the zygote is the only diploid stage, so that charophytes are characterised by their haplontic life cycle with only the gametophyte being multicellular. The haplodiplontic life cycle which is featured by a multicellular sporophyte evolved in the land plant lineage. In angiosperms as well as gymnosperms the gametophyte is reduced to only a few hidden cells. In the lower

part a timeline of divergence of taxa and their reproductive innovations are shown starting with charophyte algae. The colour code shows geological areas: Cenozoic (white), Mesozoic (blue), Paleozoic (green), Neoproterozoic (brown). Figure was taken from (Hackenberg and Twell, 2019).

P. patens is a monoicous plant that favours self-fertilisation (Nakosteen and Hughes, 1978; Perroud *et al.*, 2011). The whole life cycle can be fulfilled in two to three months depending on the start culture (gametophore or single spore). A single spore germinates and gives rise to protonema filaments termed chloronema and caulonema. The chloronema is enriched with chloroplasts and perpendicular cell walls, whereas the caulonema is poor in chloroplasts and possesses oblique cross walls. The phase transition from 2D to 3D growth starts with the development of a bud that subsequently develops into a gametophore (Strotbek *et al.*, 2013). Sexual reproduction can be initiated by the transfer from long day conditions into cold and short day conditions, which imitates the natural life cycle in the wild where the sexual reproduction takes place in autumn (Engel, 1968; Hohe *et al.*, 2002; Nakosteen and Hughes, 1978). The gametangia are formed on the tip of the gametophore. The male sexual organ is called antheridium and develops from an antheridium initial stem cell (Kofuji *et al.*, 2018a). A mature apex develops a bundle of antheridia. Antheridia comprise a single outer cell layer that encapsulates the motile spermatozoids. Through a droplet of water, the swollen antheridia tip cell bursts and releases the gametes, which are able to swim. Shortly after initiation of

Introduction

antheridia development, archegonia development starts. Archegonia are the female gametangia, which are flask shaped. The egg develops inside the archegonial venter. During fertilisation the spermatozooids swim towards the archegonium and enter the opened archegonial neck (Hiss *et al.*, 2017a; Kofuji *et al.*, 2018b; Landberg *et al.*, 2013; Sanchez-Vera *et al.*, 2017). Upon fertilisation a zygote is formed and undergoes embryogenesis. The embryo develops into the sporophyte and performs meiosis leading to the formation of haploid spores (Figure 3). Mature spore capsules open and disperse the spores that can again germinate and develop a new plant (Hiss *et al.*, 2017a; Strotbek *et al.*, 2013).

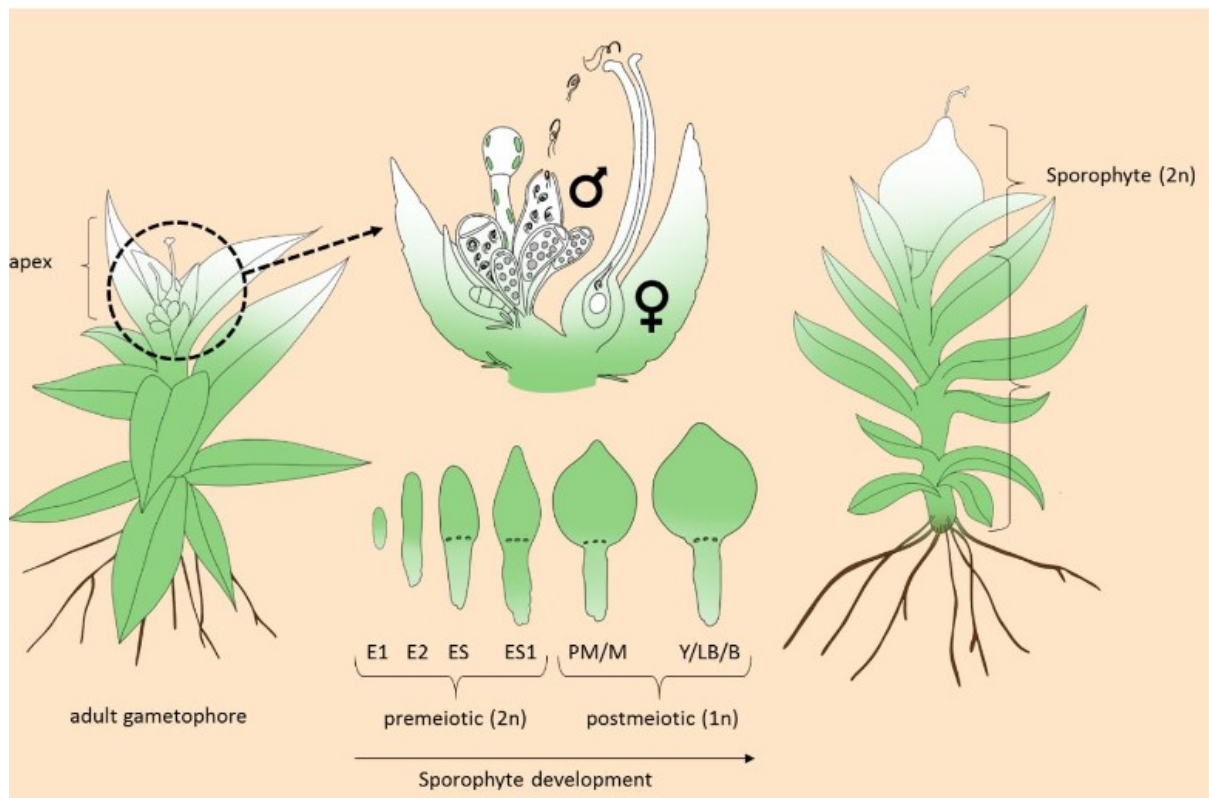


Figure 3 Cartoon of *P. patens* life cycle (sexual reproduction).

After cold day incubation, an adult gametophore develops gametangia on the gametophore (apex), which comprise archegonia (female) and antheridia (male). After ripening of the gametangia, mature spermatozooids can be released by water supply and swim towards the archegonium and enter the neck canal to reach the egg cell in the venter. Fertilization takes place and initiates embryo development. Finally, a diploid sporophyte develops. The sporophyte development can be divided into distinct stages (embryo development E1/E2 and sporophytic stages ES-B). During the sporophyte development meiosis takes place, which leads to the development of haploid spores encapsulated in the sporophyte (Hiss *et al.*, 2017a; Meyberg *et al.*, 2020). Figure was taken from (Meyberg *et al.*, 2020).

In contrast to *P. patens*, the liverwort *Marchantia polymorpha* is a dioicous plant (Bischler and Boisselier-Dubayle, 1993; Shimamura, 2016) that has two sex chromosomes (X and Y chromosome) and separate sexes, which enables the work with female as well as male plants (Bowman *et al.*, 2017). Figure 4 shows the schematic life cycle of *M. polymorpha*. As a haploid-

Introduction

dominant plant, the gametophyte generation is the dominant generation during the life cycle with the thallus as main body plan. Comparable to *P. patens*, the life cycle starts with the outgrowth of an initial short filamentous protonema with rhizoids from a unicellular spore. The body plan is thalloid. Thallus growth proceeds at the apical notches through dichotomous branching (Solly *et al.*, 2017). *M. polymorpha* can reproduce sexually as well as vegetatively via gemma cups on the dorsal part of the mature thallus, which comprise numerous multicellular gemmae. As a dioicous plant, *M. polymorpha* has antheridiophores (male gametangiophore) and archegoniophores (female gametangiophore) on separate thalli. A single egg develops inside of the venter of each archegonium. The spermatozooids develop in the antheridia. Fertilization takes place when the flagellated spermatozooids reach the egg cell. Subsequently, the zygote forms, which develops into a sporophyte (Hisanaga *et al.*, 2019b). The sporophyte is protected by a calyptra and hangs upside down beneath the archegoniophore. Due to meiotic divisions, countless haploid spores are formed in the capsule and can be released (Figure 4) (Schmid *et al.*, 2018; Shimamura, 2016).

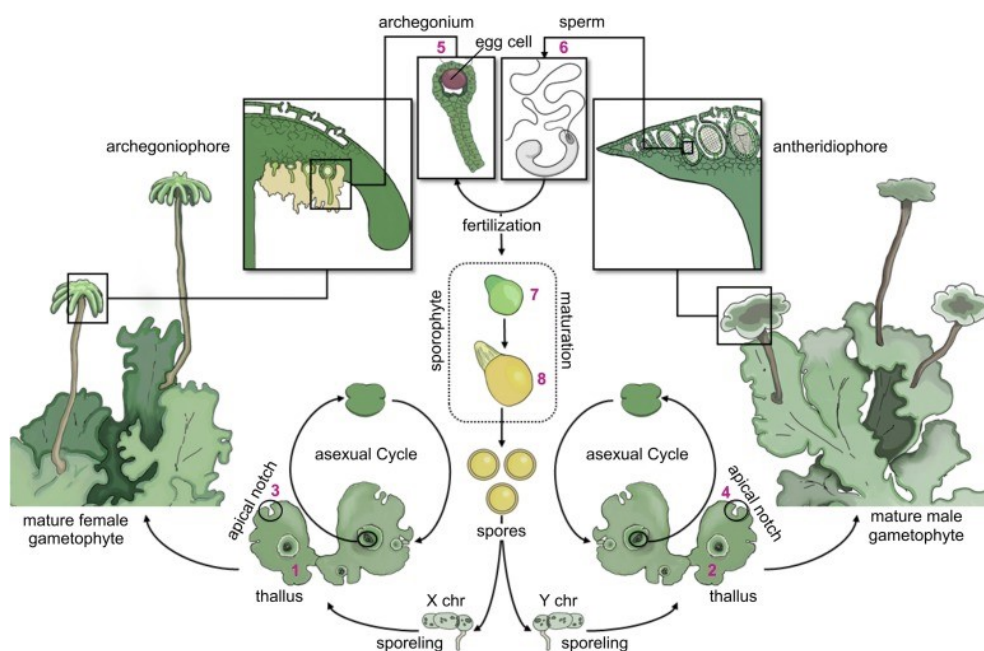


Figure 4 Life cycle of *M. polymorpha*.

M. polymorpha as dioicous plant develops antheridiophores on male gametophytes, which harbour the antheridia with the spermatozooids. The female plants develop archegoniophores which possess

archegonia that contain single egg cells. Reproduction can occur sexually and vegetatively. Within sexual reproduction, spermatozooids are released from the antheridia and swim towards the egg cells in order to inseminate them. Subsequently, a sporophyte is formed beneath the archegoniophore. The sporophyte performs meiosis whereupon spores develop and can be subsequently released. Upon suitable conditions, spores are able to germinate and develop either male or female gametophytes. The vegetative reproduction occurs through the formation of gemmae within gemma cups (Shimamura, 2016). Numbers in magenta are study specific. Figure was taken from (Schmid *et al.*, 2018).

Introduction

2.3 Evolution of land plant sexual reproduction

Sex was an evolutionary innovation that already arose in the origin of the eukaryotes about 850 million years ago featured by syngamy, nuclear fusion and meiosis. The latter was needed for ploidy reduction and DNA recombination (Cavalier-Smith, 2002). Another invention of the eukaryotic life was the vertical transmission of genetic material. The vertical lineage inheritors faced the problem of mutations that accumulated in clonal generations, called Muller's ratchet (Felsenstein, 1974; Muller, 1964). Recombination of genetic material via sexual reproduction preserved the eukaryotic lineages from extinction and helped to overcome Muller's ratchet (Crow, 2005; Felsenstein, 1974; Garg and Martin, 2016; Moran, 1996; Muller, 1964). Besides avoiding accumulation of mutations, there is evidence that meiosis was the basis for multicellular organisms to evolve from unicellular ancestors. In the case of land plant evolution of the alternation of generations, it was hypothesised that the multicellular sporophyte evolved via interposed mitotic divisions of the zygote prior to meiosis. In haplontic green algae the zygote is the only diploid stage (Bower, 1908; Lee *et al.*, 2008; Niklas and Kutschera, 2010; Rensing, 2016). Land plants are characterised by their alternation of multicellular haploid and diploid generations (haplodiplontic life cycle). The haploid multicellular generation, the gametophyte, is obtained through mitotic divisions after meiosis (Rensing, 2016).

Prior to the fusion of haploid gametes (fertilisation), it is necessary that the male gamete is able to reach the egg cell. In flowering plants, the pollen is transferred via the outgrowth of a pollen tube. A huge amount of plant species has flagellated sperm that swims towards the egg cell, which is the ancestral eukaryotic state (Mitchell, 2007; Stewart and Mattox, 1975). Flagellated gametes have been secondarily lost in some lineages e.g. Zygnematales (Transeau, 1951), which show a conjugation of aplanogametes and, as mentioned before, flowering plants (Renzaglia and Garbary, 2001; Southworth and Cresti, 1997). Flagella are not only important for gamete movement but also for the motility of the whole organism in case of *C. reinhardtii*, a unicellular alga. During reproduction, the flagella mediate a species-specific adhesion between two algae of different mating types (van den Ende *et al.*, 1990), which leads to the development of the zygote.

Streptophyte algae possess motile spermatozoids as the only motile cells, since these multicellular algae are sessile organisms. The habitat of algae is water, which favours motile sperm, because they can easily swim towards the oogonia for fertilisation (Hackenberg and Twell, 2019; McCourt *et al.*, 2004). After the water to land transition, the land plants had to face a new environment with new challenges like the lack of surrounding water. Nevertheless, the ancestor of the recent bryophytes retained flagellated spermatozoids. Certain environmental conditions like rain or moist were then needed for spermatozoids to reach the egg cell. The loss of flagellated gametes took place twice in seed plants, after the evolution of cycads and

Introduction

Ginkgo, both of them having pollen as well as flagellated sperm. It is highly likely that this loss happened in the most recent common ancestor (MRCA) of conifers and Gnetales. The second loss occurred in flowering plants (Higo *et al.*, 2018; Renzaglia *et al.*, 2000; Renzaglia and Garbary, 2001; Southworth and Cresti, 1997).

The evolutionary invention of flowers as reproductive structures probably led to the major evolutionary success of angiosperms (flowering plants), with almost 90 % of known plant species reproducing this way (Soltis *et al.*, 2008; Wozniak and Sicard, 2018). The invention and evolution of flower morphology improved the reproductive system, which led to a boost in species diversification (species radiation), starting in the early Cretaceous (Crane *et al.*, 1994; Friis *et al.*, 1994; Friis *et al.*, 2006). The enormous species radiation was followed by a diversification of flower structures as well as appearances. This species diversification was facilitated by a well orchestrated interaction with animal pollinators and their co-evolution (Crane *et al.*, 1994, 1995; Friis *et al.*, 1999; Hu *et al.*, 2008; Thien *et al.*, 2009; Wozniak and Sicard, 2018). Because of the abrupt occurrence of the flower in the fossil record (Crepet, 2000), Darwin described this phenomenon as “abominable mystery” (O'Maoileidigh *et al.*, 2014; Wozniak and Sicard, 2018).

It seems probable that flowers evolved from the male or female cone of gymnosperms (Figure 5). Gymnosperms and angiosperms are distinct in their reproductive structures. Figure 5 illustrates that the male and female reproductive units in gymnosperms are separated and the ovules are exposed, while angiosperms are typically monoicous. The flower of angiosperms is organised in whorls. *A. thaliana* is characterised by four sepals in the first whorl, four petals in the second whorl, six stamens in the third whorl, that are important for pollen development (male organ), and the female organ which comprises two fused carpels with the ovules (fourth whorl) (O'Maoileidigh *et al.*, 2014; Smyth *et al.*, 1990).

Introduction

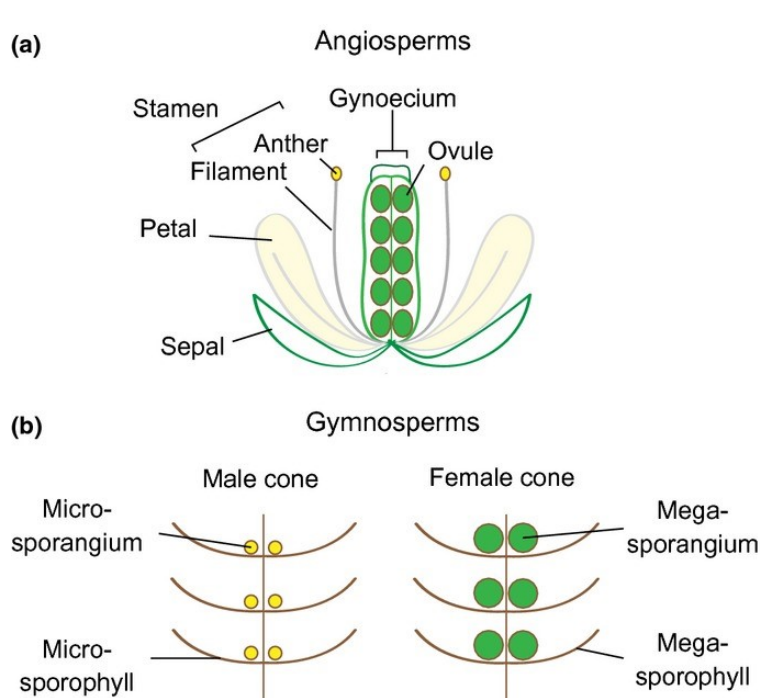


Figure 5 Male and female reproductive organs of angiosperms and gymnosperms.

a) Schematic angiosperm flower (*A. thaliana* as model flower). Angiosperm flowers are characterised by the enclosed ovules within a gynoecium. *A. thaliana* is a dioecious plant, that has male and female reproductive units in the same flower. Nevertheless, there are monoicous angiosperms as well. The composition of sepals and petals shows variations (Theissen and Melzer, 2007). On top of that, they can be replaced by tepals, which are undifferentiated perianth structures. b) The gymnosperm lineage is represented by male and

female pine (Pinaceae) cones. Microsporangia, which comprise the pollen (male), develop on top of leaf-like structures called microsporophylls. The female cones consist of uncovered ovules (megasporangia) which are located on megasporophylls in contrast to an enclosure in a gynoecium. The sporophylls are hypothesised to be gymnosperm organs that most closely resemble carpels (Melzer *et al.*, 2009; O'Maoileidigh *et al.*, 2014). Figure was taken from (O'Maoileidigh *et al.*, 2014).

2.4 Transcription associated proteins (TAPs)

A multicellular organism is dependent on a tightly regulated gene expression. Especially the regulation of transcription, the process in which the genomic DNA serves as a template for the synthesis of a messenger RNA, is of special importance. The regulation of mRNA synthesis is essential for controlling the levels of RNA and protein temporally and spatially. Transcriptional regulation is executed by TAPs, which comprise transcription factors (TFs), transcriptional regulators (TRs) and putative TAPs (PTs). TFs bind sequence-specific to *cis*-regulatory elements to activate or suppress transcription. TRs bind DNA unspecifically, act via protein-protein interactions or by chromatin modification. PTs remain mainly unexplored until today (Lang *et al.*, 2010; Richardt *et al.*, 2007; Wilhelmsson *et al.*, 2017).

The expansion of TAP gene families seems to correlate with the developmental evolution throughout all kingdoms of life (Carroll, 2005; Gutierrez *et al.*, 2004; Hsia and McGinnis, 2003; Levine and Tjian, 2003). In plants and animals, the possibility to encode TAPs resulting in a higher level of transcriptional regulation is directly linked to morphological complexity, so that raising complexity goes hand in hand with a higher number of TAPs in the genomic fraction (Levine and Tjian, 2003). On top of that, expansion of TAP families diversified the evolution of

Introduction

morphology (de Mendoza *et al.*, 2013; Lang *et al.*, 2010; Lespinet *et al.*, 2002; Levine and Tjian, 2003; Richardt *et al.*, 2007; Wilhelmsson *et al.*, 2017).

Comparative genomics between plants and animals allow to determine gains, losses and expansions of gene families, and elucidated the fact that the unicellular ancestor of animals and plants already possessed gene families that are essential for these kingdoms (Catarino *et al.*, 2016; de Mendoza *et al.*, 2013; de Mendoza *et al.*, 2015; Lang *et al.*, 2010).

In terms of land plant evolution, it was shown recently that many TAP families which were considered to be land plant-specific already originated in the streptophyte algae, which are sister to all land plants (Delaux *et al.*, 2015; Hori *et al.*, 2014; Nishiyama *et al.*, 2018; Wang *et al.*, 2015; Wilhelmsson *et al.*, 2017). TAPs are classified by their domain or domain composition and play distinct roles in plant development. Evolution of plant complexity was driven by gene duplication, which can be followed by paralog retention and a sub- or neofunctionalization, (Rensing, 2014).

Large-scale duplication events happened numerous times in eukaryotic genomes throughout evolution (Edger and Pires, 2009; Lang *et al.*, 2010; Paterson *et al.*, 2006; Van de Peer *et al.*, 2009). It was shown that plant TF paralogs were often kept after whole genome duplication events (De Bodt *et al.*, 2005; Lang *et al.*, 2010). This can be explained by gene dosis effects. Genes that are dosage-sensitive are usually retained after polyploidization events, a scenario that is called Gene Balance Hypothesis (Birchler and Veitia, 2012; Lang *et al.*, 2010; Van de Peer *et al.*, 2009). Looking at TRs the raising complexity does not necessarily go hand in hand with their expansion although they are crucial regulators as well (Wilhelmsson *et al.*, 2017). Nevertheless, there are TAPs existing as single copies in most plant genomes e.g. LEAFY which is involved in flower development and cell division, or the Polycomb complex 2 (PRC2) protein FIE (Mosquna *et al.*, 2009; Sayou *et al.*, 2014). That can be explained by the Selected Single Copy Hypothesis, completing the Gene Balance Hypothesis by postulating that paralog retention in these cases is counter-selected because of a dosage imbalance of their gene products (Duarte *et al.*, 2010; Edger and Pires, 2009).

2.5 Genetic and epigenetic regulation of the alternation of generations

The haploid and diploid life phases and the associated alternation of generations is controlled genetically and epigenetically. In unicellular, haplontic green algae TALE (three amino acid length extension) class (KNOX and BELL) homeodomain (HD) transcription factors (TFs) regulate the haploid-diploid phase transition by acting as a heterodimeric checkpoint of proper mating type fusion (Lee *et al.*, 2008). In land plants, HD-TALE (KNOX/BELL-like) genes are important for the determination of organ formation as well as the vegetative to reproductive

Introduction

phase transition (Hake *et al.*, 2004; Hay and Tsiantis, 2010). Recently, it was suggested that this checkpoint is an ancestral eukaryotic feature (Joo *et al.*, 2018).

In bryophytes, the knowledge of sexual reproduction and evolution of plant sexual reproduction is limited compared to seed plants. In the model moss *P. patens* research of sexual reproduction advanced quite recently in terms of e.g. autophagy (Sanchez-Vera *et al.*, 2017), life cycle timing (Hiss *et al.*, 2017a), spore germination (Vesty *et al.*, 2016) and flagellar assembly (Meyberg *et al.*, 2020). Studying the control of the plant life cycle on a genetic or epigenetic level is favourable in haploid-dominant plants, since both generations can be easily analysed in contrast to the reduced gametophyte in seed plants. Moreover, moss plants defective in gametangiogenesis or sporophyte development can often be grown as gametophytic tissue and vegetatively propagated (Perroud *et al.*, 2014), in contrast to embryo-lethal mutants of flowering plants.

In bryophytes, the alternation of generations is highly regulated genetically and epigenetically by certain key players. PRC2 protein FIE, which controls developmental genes via Histone 3 (H3) K27 trimethylation (Ikeuchi *et al.*, 2015; Katz *et al.*, 2004; Kawashima and Berger, 2014) (Figure 6) is important for the repression of the sporophytic body plan in the gametophyte (Mosquna *et al.*, 2009; Okano *et al.*, 2009). In comparison, HD TFs of the KNOX class 2 repress the development of the gametophytic body plan from the sporophytic embryo (Sakakibara *et al.*, 2013). Also, a BELL HD TF controls the fertilisation and/or the first zygotic division (Horst *et al.*, 2016).

Introduction

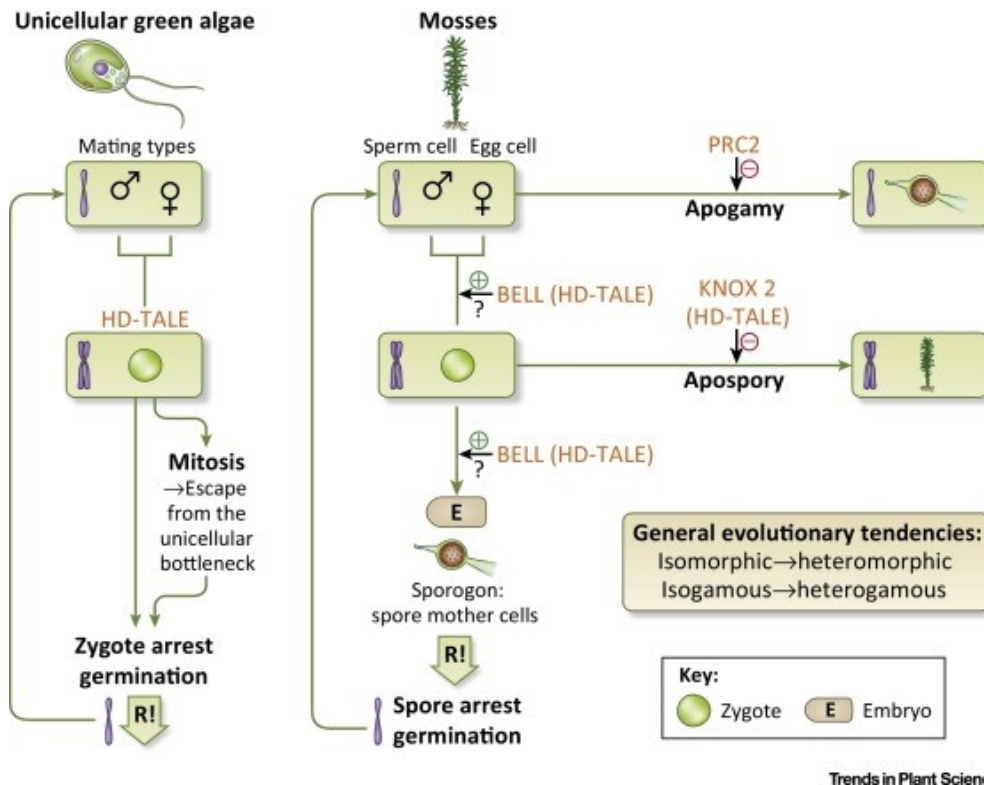


Figure 6 Genetic and epigenetic regulation of land plant alternation of generations in an evolutionary context.

Unicellular green algae e.g. *C. reinhardtii* perform a haplontic life cycle, in which the only diploid stage is the zygote (circle). Diploid stages are indicated with a

double chromosome, whereas the single chromosome marks the haploid phase. The formation of the zygote is regulated by an interplay of TALE class HD TFs, KNOX and BELL (Lee *et al.*, 2008). After undergoing meiosis the haploid gametophytic stage of two mating types of unicellular green algae is restored. It can be hypothesised that the land plant alternation of generations evolved from mitotic divisions of the zygote instead of meiotic divisions (Lee *et al.*, 2008). The moss life style is dominated by the haploid generations, the gametophyte. Contrasting the algal life cycle, the moss' zygote performs mitosis and subsequently embryogenesis. The now diploid sporophyte goes through meiosis, leading to haploid spores, which wait with germination until appropriate conditions arise and grow towards a multicellular haploid gametophyte. In mosses the gametes are heterogamous, that means egg cells and spermatozoids do not resemble each other. PRC2 prevents apogamy, which means a haploid but sporophyte-like body plan, that is involved in repressive chromatin modifications. Apospory names the contrary scenario, that means the development of a diploid but gametophyte-like body plan is repressed by KNOX2 HD TFs (TALE class). The zygote as well as the embryo formation is facilitated by a BELL HD TF (Horst *et al.*, 2016). On top of that, other (TALE) HD TFs acting as heterodimer could be involved. Figure was taken from (Rensing, 2016).

The determination of male versus female development in the dioicous liverwort is taken over by a R2R3 MYB-type transcription factor FEMALE GAMETOPHYTE MYB (MpFGMYB) (Figure 7), which is involved in the development of the female organ (Hisanaga *et al.*, 2019a), whereas MYB65/119 are involved in the female development in *A. thaliana* (Rabiger and Drews, 2013). In the liverwort *M. polymorpha*, it was shown that an RKD TF is involved in controlling the gametophyte to sporophyte transition by preventing the unfertilised egg cell from dividing (Rovekamp *et al.*, 2016). Moreover, it was shown that MpRKD is controlling the sperm differentiation, in contrast to *A. thaliana* where the RKD genes are required for embryogenesis

Introduction

and egg cell formation, but excluded from sperm development (Koi *et al.*, 2016; Rovekamp *et al.*, 2016).

Recently another germ line specification factor was identified, namely the bHLH-family transcription factor BNB. In *M. polymorpha* BNB initiates the development of male and female organs (gametangiophores which comprise male and female gametangia). In *A. thaliana* two orthologs of BNB are needed for male gametogenesis (Yamaoka *et al.*, 2018). In *M. polymorpha* BNB shows expression in gametangiophore and gametangia initial cells of male and female organs and in the germ cell lineage of antheridia and archegonia. The expression can be detected in immature egg cells and in spermatogenous cells before proliferation, but not in mature gametes. In *A. thaliana* the orthologs of BNB1/2 are male germ line specific (GC specification), in contrast to *M. polymorpha* BNB which functions in male and female gametangia development. These findings underline the hypothesis that the gene networks regulation of the germ line adapted throughout 450-500 million years of evolution, that separate bryophytes and angiosperms (Hackenberg and Twell, 2019). As mentioned before, the transcription factors RKD and BNB had a more diversified ancestral role in bryophytes compared to the sex-specific roles in angiosperms (Koi *et al.*, 2016; Rovekamp *et al.*, 2016). The evolutionary innovation of sexual reproduction required the development of new cell types. Especially the process of sperm differentiation was mainly unexplored in a lot of multicellular eukaryotic lineages. Recently, it was postulated that a single MYB TF (DUO POLLEN 1 (DUO1)) was essential for the origin of spermatogenesis in the common ancestor of Charophyceae and land plants (Higo *et al.*, 2018). DUO1 was shown to control male gamete development and differentiation in *A. thaliana* (Borg *et al.*, 2011; Brownfield *et al.*, 2009; Rotman *et al.*, 2005). The gene regulatory network of DUO1 comprises two closely related TFs, DUO1-ACTIVATED ZINC FINGER1 (DAZ1) and DAZ2 in *A. thaliana*. The DUO1/DAZ1/2 module plays a crucial role in the germ line, more specifically throughout mitosis. Apart from that it is important for the regulation of male germ line differentiation genes and is involved in fertilisation on the part of sperm fusion. In addition, DUO1 controls the expression of the sperm-specific histone variant H3.10 and the fusogenic factors GENERATIVE CELL-SPECIFIC 1 /HAPLESS 2 (GCS1/HAP2) and GAMETE EXPRESSED 2 (GEX2) (Borg *et al.*, 2014; Yamaoka *et al.*, 2018), which are essential for gamete interaction and fertilisation (Borg *et al.*, 2011).

Introduction

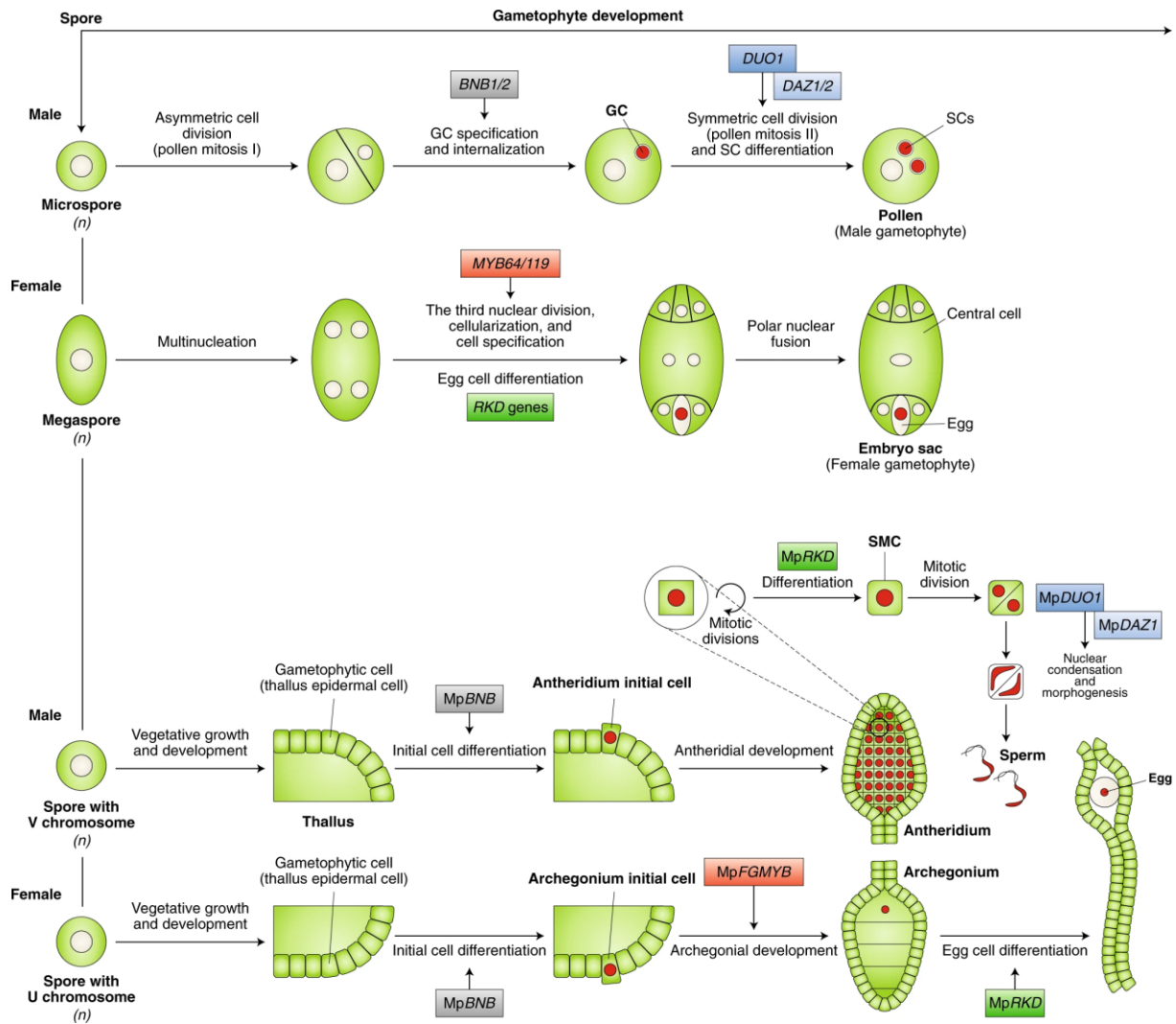


Figure 7 Control of gametophyte (n=haploid) development in *A. thaliana* and *M. polymorpha*.

In the top part, the development of the *A. thaliana* gametophyte starting from a spore is shown, whereas on the bottom part the gametophyte development in *M. polymorpha* is depicted. Germ line nuclei are shown in red, *A. thaliana* nuclei of gametophytic cells are shown in white. The genes that regulate certain steps throughout gametophytic development are shown in the boxes. The colour codes indicate homologous or orthologous genes between seed plant and liverwort. Abbreviations: generative cell (GC), sperm cells (SCs), spore mother cell (SMC). Figure was taken from (Hisanaga *et al.*, 2019b).

3 Aim of this study

The transition from water to land was a major event in the land plant evolution, that shaped the atmospheric conditions as well as the appearance of the planet itself. Therefore, the conquerors had to face various stressors such as drought, radiation, cell wall architecture, nutrient acquisition and sexual reproduction without constant water supply (Rensing, 2018). The newly gained adaptations were accompanied by the evolution of a tight gene regulation. It was shown that TAPs play a crucial role in the control of plant development. The raising plant complexity goes hand in hand with the possibility to encode TAPs (Lang *et al.*, 2010). Especially in times of global warming, it is important to study the capability of plants to adapt to new conditions, with the conquest of land by plants, a dominant example for the adaptation to terrestrial stresses (de Vries *et al.*, 2018; de Vries and Rensing, 2020; Furst-Jansen *et al.*, 2020).

Therefore, the aim of the study was the identification of single copy TAPs involved in the control of sexual reproduction via a combined approach of an orthology detection tool (proteinortho) (Lechner *et al.*, 2011), protein-family categorization (TAPscan) (Lang *et al.*, 2010; Wilhelmsson *et al.*, 2017), literature search, expression analyses and phylogenetic inference. This study aimed for analysing the evolution of sexual reproduction in more depth using bryophytes as model systems. Therefore, a set of candidate genes should be identified and analysed via forward and reverse genetics. Loss-of-function mutants of a set of candidate genes in *P. patens* as well as *M. polymorpha* and their resulting morphology in terms of sexual reproduction should give rise to the gene function in different bryophyte lineages. As described above many genes were analysed already in flowering plants as well as in mammals, which should be a prerequisite for the tackled candidates as well. Taken together, this study aimed to deepen the understanding of the evolution as well as possible conservation of the candidate's gene function. Moreover, the knowledge about the regulation of sexual reproduction should be extended in bryophytes.

Materials and Methods

4 Materials and Methods

4.1 Materials, buffers and media

All restriction enzymes used were purchased from Thermo Fisher (Dreieich, DE) or NEB (Frankfurt a.M., DE). A list of chemicals and materials can be found in supplement B. The buffers and media as well as their composition and application are listed in Table 1 and Table 2.

Table 1 Buffers, composition and their application.

buffer	composition	application
50 x TAE-buffer (1000 ml)	242 g Tris	gel electrophoresis
	57,1 ml acetic acid	
	100 ml of 0,5 M EDTA (pH 8,0)	
	autoclave	
P1-buffer	50 mM Tris-HCl, pH 8	plasmid DNA isolation (bacteria)
	10 mM EDTA	
	100 µg/ml RNase A	
P2-buffer	200 mM NaOH	plasmid DNA isolation (bacteria)
	1 % (w/v) SDS	
P3-buffer	3 M C ₂ H ₃ KO ₂ , pH 5,5 adjust with acetic acid	plasmid DNA isolation (bacteria)
10 x SDB buffer	0,3 M Tris/acetic acid pH 7,0	enzyme digestion
	0,5 M C ₂ H ₃ KO ₂	
	0,1 M Mg(C ₂ H ₃ O ₂) ₂	
	49 mM Spermidin	
	50 mM DTT	
CTAB extraction buffer	2 % (w/v)	genomic DNA isolation from plant tissue
	N-Cetyl-N,N,N-trimethylammoniumbromid (CTAB)	
	1,4 M NaCl	
	100 mM Tris/HCl pH 8,0	
	20 mM EDTA	
Shorty buffer	100 mM β-Mercaptoethanol (add directly before usage)	genomic DNA isolation from plant tissue
	0,2 M Tris-HCl, pH 9	
	0,4 M LiCl	
	25 mM EDTA	
	1 % SDS	
TE-buffer	10 mM Tris-HCl, pH 8	DNA isolation
	1 mM EDTA	
	autoclave	

Materials and Methods

HBSS	5,3 mM KCl	staining
	0,44 mM KH ₂ PO ₄	
	4,167 mM NaHCO ₃	
	137,93 mM NaCl	
	0,338 mM Na ₂ HPO ₄	
	5,56 mM D-Glucose	
RF1	100 mM RbCl ₂	RbCl ₂ -competent cells
	50 mM MnCl ₂ x 4 H ₂ O	
	30 mM C ₂ H ₃ KO ₂	
	10 mM CaCl ₂ x 2 H ₂ O	
	15 % (w/v) glycerine	
	pH 5,8 adjusted with 0,2 M acetic acid, sterile filtration	
RF2	10 mM 3-(N-Morpholino)propanesulfonic acid	RbCl ₂ -competent cells
	10 mM RbCl ₂	
	75 mM CaCl ₂ x 2 H ₂ O	
	15 % (w/v) glycerin	
	pH 5,8 adjusted with NaOH, sterile filtration	
PEG (40 %)	5 g PEG 4000	moss transfection
	7,5 g 3M-medium	
	sterile filtration	

Table 2 Media, composition and their application.

medium	composition	application
KNOP	KH ₂ PO ₄ (250 mg/l)	moss culture
	KCl (250 mg/l)	
	MgSO ₄ x 7 H ₂ O (250 mg/l)	
	Ca(NO ₃) ₂ x 4 H ₂ O (1000 mg/l)	
	FeSO ₄ x 7 H ₂ O (12,5 mg/l)	
	adjust pH to 5,8 with KOH	
	for solid medium add 9 g/l phyto agar	
	autoclave	
LB	10 g/l Trypton	bacterial culture
	5 g/l yeast extract	
	10 g/l NaCl	
	adjust pH to 7,0 with NaOH	
	for solid medium add 15 g/l bacto agar	
	autoclave	

Materials and Methods

0,5 M Mannitol (100 ml)	9,88 g Mannitol 100 ml H ₂ O (desalted water) pH adjust to 5,6 - 5,8 with HCl or KOH autoclave	moss transfection
3M-medium (100 ml)	0,3 g MgCl ₂ x 6 H ₂ O 0,13 g MES 9,44 g Mannit pH adjust to 5,6 with KOH sterile filtration	moss transfection
regeneration medium (100 ml)	5 g Glucose 4,4 g Mannit 100 ml KNOP liquid pH adjust to 5,8 with KOH sterile filtration	moss transfection
sterilisation solution	50-100 µl NaClO (12 %) 100 µl Triton-X-100 (10 %) H ₂ O millipore	liverwort transfection
liquid medium ½ Gamborg + organics for sporeling culture and as washing medium (1000 ml)	1,582 g Gamborg B5 0,3 g L-Glutamine (0,03 %) 1,0 g Casamino Acids (0,1 %) 20 g Sucrose (2 %) H ₂ O millipore	liverwort transfection
selection medium ½ Gamborg (1000 ml)	1,528 g Gamborg B5 14 g Agar Agar Kobe I autoclave cooling (approximately 55 °C) [+ selection, e.g.: Hygromycin stock solution] 1000 µl Cefotaxime (Cefo ¹⁰⁰)	liverwort transfection
solid medium ½ Gamborg (1000 ml)	1,528 g Gamborg B5 14 g Agar Agar Kobe I add 1 % Glucose for gametangiophore induction	liverwort culture
antibiotics stocksolutions	Kan ⁵⁰ : 50 mg/ml Carb ¹⁰⁰ : 100 mg/ml Rif ¹⁰⁰ : 100 mg/ml Hyg ²⁰ : 20 mg/ml Spect ⁵⁰ : 50 mg/ml Cefotaxime (Claforan): 1 g/10 ml H ₂ O millipore Acetosyringone: 1 M stock and 200 µM final concentration G418 ²⁵ (25 mg/ml)	selection

Materials and Methods

4.2 Methods

4.2.1 Candidate gene search

Proteinortho was initially used (Kristian K. Ullrich (MPI Plön, DE)) for ortholog clustering (Lechner *et al.*, 2011). The HMM-based TAPscan classification (Wilhelmsson *et al.*, 2017) was applied for assignment of TAP families. For 23 TAP families the orthology status was analysed via phylogenetic analyses using quicktree-sd (Frickenhaus and Beszteri, 2008). Details of the workflow are shown in Figure 12.

Phylogenies of the TAP families were obtained by the usage of the web interface TAPScan (<https://plantcode.online.uni-marburg.de/tapscan/index.php>). All genes of the respective TAP family and the desired species were extracted and aligned using Mafft L-INS-I v7.310 (Katoh and Standley, 2013). After manual trimming of not conserved regions, namely only one sequence in the respective part of the alignment or a minus consensus declared by Jalview (Waterhouse *et al.*, 2009), the phylogeny was calculated with the help of quicktree_sd (Frickenhaus and Beszteri, 2008).

For 11 candidate genes blastp was performed with the respective *A. thaliana* protein sequence against a database of sequenced plant and algal genomes. The obtained sequences were kept if the alignment was at least 100 positions long and at least 30 % identical. Each of the 11 protein sets was aligned using Mafft L-INS-I v7.310 (Katoh and Standley, 2013). Additionally, MUSCLE v3.8.31 was used as alignment tool (Edgar, 2004). Alignments were manually curated using Jalview 2.8 (Waterhouse *et al.*, 2009), removing identical sequences and trimming of not conserved regions, namely only one sequence in the respective part of the alignment or a minus consensus declared by Jalview. The best suited amino acid substitution model was determined using Prottest 3.4.2 (Darriba *et al.*, 2011) and turned out to be e.g. JTT+I+G (AT3G54610) and JTT+I+G (AT2G47620). Bayesian inference utilizing MrBayes 3.2.7 (Ronquist *et al.*, 2012) was carried out with two hot and cold chains until the average standard deviation of split frequencies was below 0,01 (574,350 generations) and no more trend was observable. 250 trees each were discarded as burn-in. IQ-TREE 1.6.12 (maximum likelihood) was used as alternative to MrBayes (Nguyen *et al.*, 2015). Resulting trees were visualised using FigTree 1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>). Protein domains were analysed using InterProScan 5.31-70.0 (Jones *et al.*, 2014; Mitchell *et al.*, 2019).

Materials and Methods

4.2.2 Localisation and expression analyses

The protein sequences of the selected candidate genes were analysed *in silico* with the webtool LOCALIZER (Sperschneider *et al.*, 2017). For details of *in vivo* analyses see chapter 5.3.

The candidate genes were analysed regarding their expression in the species *A. thaliana*, *M. polymorpha* and *P. patens*. Expression data of *M. polymorpha* were analysed according to (Bowman *et al.*, 2017). *A. thaliana* and *P. patens* expression data were obtained with the help of the eFP Browser (Winter *et al.*, 2007). *P. patens* was also analysed in terms of developmental stages microarray and RNA-seq data with the webtool PEATmoss (Fernandez-Pozo *et al.*, 2019).

4.2.3 Culture conditions

P. patens ecotype Reute 17 was cultivated according to (Hiss *et al.*, 2017a). For sporophyte development, and for crossing analyses the cultivation was carried out as described in (Perroud *et al.*, 2019); long day (LD) conditions ($70 \mu\text{mol m}^{-2} \text{s}^{-1}$, 16h light, 8h dark, 22 °C). Gametangia were watered in order to induce sporophyte development at 21 days after short day (SD) ($20 \mu\text{mol m}^{-2} \text{s}^{-1}$, 8h light, 16h dark, 16 °C) transfer. Liquid cultures for transformation were set up as described in (Frank *et al.*, 2005; Hohe and Reski, 2002).

M. polymorpha, ssp. *ruderalis*, Botanischer Garten (BoGa) ecotype, Osnabrück was cultivated on ½ Gamborg B5 medium. To induce gametangiophore development, the plants were transferred to ½ Gamborg B5 supplemented with 1 % glucose. After one to three weeks of cultivation in the long day chamber, the cultures were transferred to the cold for at least two weeks (6 – 8 °C, halogen/ cold white). Subsequently, the plants were put back in the long day chamber under far-red light (680 – 780 nm, peak 740 nm). Plates should remain under far-red light conditions until gametangiophore development is completed.

4.2.4 Nucleic acid analysis

4.2.4.1 Genomic DNA isolation from plant material

Isolation of whole genomic DNA was performed according to the protocol of (Doyle and Doyle, 1990). The isolated DNA comprises nuclear and organellar DNA. It can be applied to amplify specific DNA regions via polymerase chain reaction (PCR). Approximately 100 mg plant material was grinded with a pistil in a 1,5 ml reaction tube and frozen in liquid nitrogen. 600 µl extraction buffer (CTAB) were applied to the plant material, and grinded another round. The sample was vortexed for 30 seconds, followed by a 50 minutes and 70 °C incubation step. 600

Materials and Methods

µl Chloroform/Isoamylalcohol (24:1) were added followed by a 30 seconds vortex step. The samples were centrifuged for 10 minutes, 12,000 RCF and 4 °C. The aqueous phase on top was transferred to a new reaction tube and filled up with millipore water to 600 µl. The extraction step with Chloroform/Isoamylalcohol (24:1) was repeated. The aqueous phase on top was transferred to a new reaction tube. 0,7 times of the sample volume isopropanol was added to the sample. The samples were centrifuged for 20 minutes, 20,000 RCF and 4 °C. The supernatant was discarded. The nucleic acid pellet was dissolved in 600 µl 70 % ethanol and centrifuged for 10 minutes, 20,000 RCF and 4 °C. The supernatant was discarded completely. The nucleic acid pellet was dried and dissolved in 20-50 µl millipore water.

4.2.4.2 Nucleic acid isolation “Quick and Dirty”

Genomic DNA for genotyping was isolated with a fast extraction protocol, using one plant (approximately 5-10 gametophores) as described in (Cove *et al.*, 2009). Approximately 10 to 100 mg plant material was harvested in a reaction tube and directly frozen in liquid nitrogen. The frozen material was grinded with a pistil and 200 µl shorty buffer were added. The suspension was homogenised with the pistil and again 200 µl shorty buffer were added and mixed. The sample was centrifuged for five minutes and 20,000 RCF. 300 µl of the supernatant were transferred to a fresh reaction tube, which was already filled with 300 µl isopropanol. Another centrifugation step followed after inversion of the sample for 10 minutes and maximum speed. The supernatant was discarded and the pellet was washed with 250 µl 70 % ethanol. The sample was centrifuged at maximum speed for five minutes. The supernatant was discarded and the pellet was air-dried and dissolved in 200 µl millipore water overnight at 4 °C.

4.2.4.3 RNA isolation

For RNA-seq, 40 apices were harvested in 40 µl RNA later (Qiagen, Hilden, DE) after 21 days in short day incubation as well as two days after watering. RNA extraction was performed using the RNeasy plant mini Kit (Qiagen). RNA concentration and quality were measured with Agilent RNA 6000 Nano Kit (Agilent Technologies, Santa Clara, USA) using a Bioanalyzer 2100.

4.2.4.4 Apices RNA-seq & analysis

The MPI genome center Cologne (mpgc.mpipz.mpg.de) performed Truseq RNA-library preparation and RNA-seq Illumina sequencing (HiSeq3000, 18 libraries, 2 x 150 bp, paired end reads, ordersize 162 gigabases). In total, 456,185,664 fragments were sequenced. The

Materials and Methods

analyses were done according to (Perroud *et al.*, 2018). Venn diagrams were calculated with the webtool <http://bioinformatics.psb.ugent.be/webtools/Venn/> and PowerPoint.

Gene Ontology bias was analysed according to (Widiez *et al.*, 2014) based on the *P. patens* v3.3 annotation and visualised using Wordle (<http://www.wordle.net/>). The colour code indicates over-represented GO terms by green font and under-represented GO terms by red font. The significance is correlated with the size and colour of GO terms, so that the darker colours are $q < 0,0001$ and lighter colours represent $q > 0,0001$, but $< 0,05$.

4.2.4.5 cDNA synthesis

The first-strand cDNA synthesis was performed using the SuperScript™ III. In one 20 µl reaction 10 pg – 5 µg of total RNA can be used. The following components were added to a nuclease-free microcentrifuge tube: 1 µl of random primers, 10 pg – 5 µg total RNA, 1 µl 10 mM dNTP Mix (10 mM each dATP, dGTP, dCTP and dTTP at neutral pH), sterile, millipore water up to 14 µl. The mixture was heated to 65 °C for five minutes and incubated on ice for at least one minute. After a brief centrifugation step the following components were added: 4 µl 5x First-Strand Buffer, 1 µl 0,1 M DTT and 1 µl of SuperScript™ III (200 U/µl). The mixture was pipetted gently up and down. The tube was incubated at 25 °C for five minutes followed by an incubation step at 50 °C for 30 to 60 minutes. The reaction was inactivated by heating up to 70 °C for 15 minutes. The cDNA was used as a template for PCR amplification.

4.2.4.6 Polymerase chain reaction (PCR)

In general, two different polymerases were used. For genotyping as well as TA cloning OneTaq®-Polymerase was applied, whereas for the remaining cloning approaches Q5® High-Fidelity DNA Polymerase was used. For amplification, DNA of the species *A. thaliana* (Columbia), *M. polymorpha* (BC4/tak1/BoGa) and *P. patens* (Reute 17) was used.

In the following, the used PCR reactions as well as the PCR programmes are listed. To optimise the melting temperatures and annealing temperature of the used primer pairs, the NEB Tm calculator was applied (<https://tmcalculator.neb.com/#!/main>). A list of the used primers is attached in supplement A.

Materials and Methods

OneTaq®-Polymerase reaction

	stock solution	for 25 µl
DNA template	-	1 µl
Primer forward	10 µM	0,5 µl
Primer reverse	10 µM	0,5 µl
OneTaq standard reaction buffer	5x	5 µl
dNTPs	10 mM	1 µl
OneTaq®-Polymerase	5 U/µl	0,125 µl
H ₂ O millipore	-	16,875 µl

Q5®-Polymerase reaction

	stock solution	for 25 µl
DNA template	-	2 µl
Primer forward	10 µM	1,25 µl
Primer reverse	10 µM	1,25 µl
reaction buffer	5x	5 µl
dNTPs	10 mM	0,5 µl
MgCl ₂	50 mM	0,5 µl
Q5®-Polymerase	2 U/µl	0,25 µl
H ₂ O millipore	-	14,25 µl

PCR programme OneTaq®-Polymerase

step	temperature	time	
initial denaturation	94 °C	30 seconds	
denaturation	94 °C	30 seconds	cycles 30
hybridisation	primer specific	1 minute	
elongation	68 °C	1 minute/kb	
final elongation	68 °C	5 minutes	

PCR programme Q5®-Polymerase

step	temperature	time	
initial denaturation	98 °C	30 seconds	
denaturation	98 °C	10 seconds	cycles 25-35
hybridisation	primer specific	30 seconds	
elongation	72 °C	30 seconds/kb	
final elongation	72 °C	2 minutes	

4.2.4.7 Gelectrophoresis

To analyse the PCR products, gelelectrophoresis was applied with a concentration of 0,8 % agarose. 1xTAE (TrisAcetatEDTA) was used as a buffer system. For visualisation under UV light the reagent Peqgreen (4 µl/100 ml) was added. The samples were mixed with 6x loading

Materials and Methods

dye before pipetting on the gel. To correctly assign the fragment a 1 kb or a 100 bp DNA ladder was applied as well.

4.2.4.8 Extraction of PCR and restriction fragments

PCR as well as restriction fragments were cut out from the agarose gels on top of a blue light table for DNA visualisation. The samples were transferred in reaction tubes. The extraction was performed with a Hi Yield® Gel/PCR DNA Fragments Extraction Kit (SLG Süd-Laborbedarf GmbH, Gauting, DE). The kit is based on columns to which the DNA binds due to its negative charge. The protocol followed was based on information provided by the producer. 30 µl millipore water preheated to 55 °C was used to elute the DNA.

4.2.4.9 Plasmid isolation

4.2.4.9.1 Mini preparation

Plasmid isolation was carried out according to the principle of alkaline lysis using a modified protocol according to (Bimboim and Doly, 1979).

Lysis of the bacteria was caused by treatment with SDS and NaOH. 1,5 ml of bacterial cells from an overnight culture were centrifuged at 20,800 RCF and 4 °C for one minute. The supernatant was discarded. The pellet was resuspended in 200 µl P1. After addition of 200 µl P2 and careful mixing, the suspension was incubated at room temperature for a maximum of five minutes. 200 µl P3 and 30 µl chloroform were added and incubated on ice for five minutes. After mixing, centrifugation was repeated at the same parameters for 20 minutes. The supernatant was transferred to a new reaction tube and 400 µl isopropanol was added. The sample was mixed and centrifuged again at 20,800 RCF, 4 °C and 20 minutes. The supernatant was discarded. The pellet was resuspended in 400 µl 70 % ethanol and centrifuged at 20,800 RCF, 4 °C and ten minutes. After complete removal of the ethanol and air-drying, the pellet was dissolved in 30 µl millipore water.

4.2.4.9.2 Midi preparation

50 ml LB Medium equipped with 100 mg/l Ampicillin or 50 mg/l Kanamycin was inoculated with the respective bacterial cultures. The inoculated cultures were grown shaking overnight at 37 °C. Plasmid extraction was performed using NucleoBond® Xtra Midi Kit (Macherey-Nagel, Düren, DE). The plasmid DNA was eluted in 100 µl millipore water. The concentration of the DNA was determined using a NanoDrop spectrophotometer.

Materials and Methods

4.2.4.9.3 Ethanol precipitation

For precipitation of fragments after plasmid restriction, ethanol precipitation was carried out. The samples needed a minimum volume of 200 µl which were filled up accordingly with millipore water. 1/10 volume of sodium acetate (3M, pH 5,2) was added as well as 2,5 volumes of ice cold 96 % ethanol. The solution was mixed carefully and incubated at -80 °C for 30 minutes for precipitation. The DNA was recovered by centrifugation at maximum speed for 15 minutes and 4 °C. The supernatant was discarded carefully. Ice cold 70 % ethanol was added to the pellet. Another centrifugation step at maximum speed for two minutes and 4 °C followed. The supernatant was discarded. After air-drying the pellet, it was suspended in 100 µl millipore water.

4.2.5 Molecular cloning

Cloning was carried out for diverse projects: localisation analyses (*P. patens*), overexpression of *P. patens* candidate genes, complementation analyses, CRISPR/Cas9 loss-of-function mutants (*P. patens*, *M. polymorpha*).

4.2.5.1 Homologous recombination

For a stable integration of the construct in the PIG locus (Okano *et al.*, 2009) with the help of homologous recombination, the genes amplified from cDNA were ligated in an appropriate vector that is already equipped with two homologous regions e.g. vector PIG_AN, which is displayed as an example (Figure 8). The homologous regions are needed for proper integration of the construct.

Materials and Methods

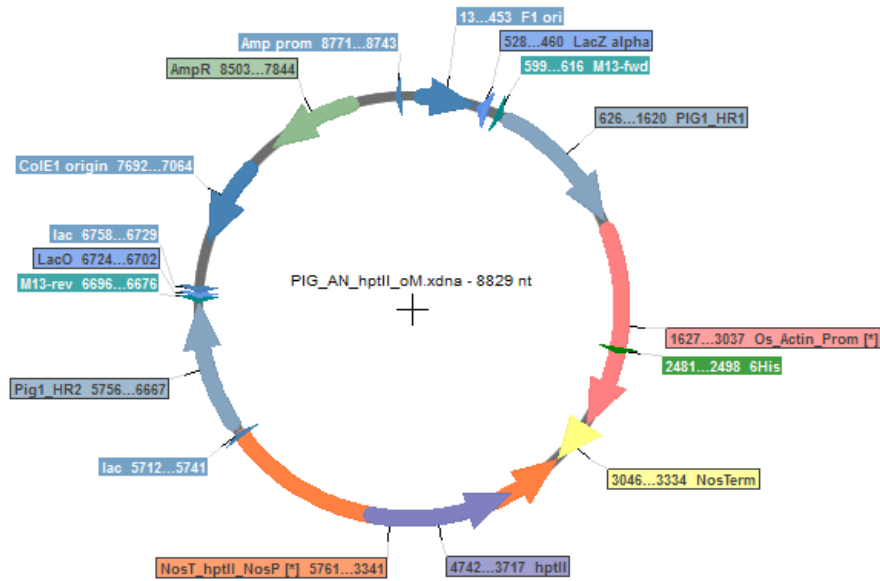


Figure 8 Knock-in vector PIG_AN.

For the complementation analyses used knock-in vector, which was applied for a stable integration of the *M. polymorpha* and *A. thaliana* homologs of the *P. patens* candidate genes. Created with the program SerialCloner 2.6.1. Features of the vector: NosTerm (terminator), hptII (Hygromycin resistance), NosT_hptII_NosP (terminator and promoter), PIG1_HR2 (homologous region), LacO (operon), ColE1 origin, M13-rev primer binding site, AmpR (Ampicillin resistance), Amp prom (promoter), F1 ori, M13-fwd primer site, PIG1_HR1, Os_Actin_Prom (promoter).

4.2.5.2 Gateway cloning

Gateway cloning was applied for obtainment of overexpressor constructs. During the Gateway reaction a gene cassette is switched between two vectors with the help of a Gateway LR Clonase™ II. The gene that should be expressed constitutively was in the first attempt ligated into the vector pMETA as entry vector via TA cloning. Therefore, the vector needed to be cut with XcmI. As destination vector pPGX8 was used (Kubo *et al.*, 2013). The destination vector was linearised using NotI-HF for better accessibility.

NotI-HF restriction approach:

4 µl	pPGX8
1 µl	10x buffer CutSmart
0,4 µl	NotI-HF (20,000U/ml)
in 10µl	H ₂ O millipore

Materials and Methods

The approach was incubated for two hours at 37 °C. The linearised vector was isolated via ethanol precipitation (see section 4.2.4.9.3). For Gateway reaction the following approach was applied:

Gateway reaction approach:

50-150 ng/μl	entry vector
150 ng/μl	pPGX8_NotI in TE
2 μl	Gateway LR Clonase™ II Enzyme Mix (5x)
<u>in 10μl</u>	TE buffer (1x)

The sample was incubated over night at 25 °C. Subsequently, the reaction was stopped by adding 1 μl of proteinase K.

4.2.5.3 Localisation analyses

To analyse the localisation of the *P. patens* candidate genes, the respective genes were tagged C-terminally in frame with GFP (Figure 20) and subsequently transformed transiently.

4.2.5.4 CRISPR/Cas9

P. patens mutants were obtained performing CRISPR/Cas9 according to (Lopez-Obando *et al.*, 2016). A small modification was made. The 500 bp fragments were cloned in pJet (<https://www.thermofisher.com/order/catalog/product/K1231>) and puc19 (<https://www.thermofisher.com/order/catalog/product/SD0061#/SD0061>). The sgRNA design was done by Fabien Nogu  (INRA, Versailles, FRA). The gBlocks gene fragments were ordered at Integrated DNA Technologies (IDT), DE.

M. polymorpha mutants were obtained performing CRISPR/Cas9 according to (Ishizaki *et al.*, 2015; Sugano *et al.*, 2018). The targets (gRNA 20 bp) were determined with the help of the webtool CRISPOR 4.97 (Concordet and Haeussler, 2018). For every loss-of-function mutant, two targets were determined. The first target should be next to the start codon while the second target should be next to the stop codon. The determined target was annealed as double strand and ligated in the vector which is already equipped with the sgRNA backbone. The targets were ordered as forward and reverse primers with BsaI overhangs.

Fwd 5' CTCG-20bp gRNA without PAM
Rev 5' AAAC-20bp gRNA in reverse complement without PAM

Materials and Methods

pMpGE_EN03 (Figure 9) ("pMpGE_En03 was a gift from Takayuki Kohchi & Sumihare Noji (Addgene plasmid # 71535 ; <http://n2t.net/addgene:71535> ; RRID:Addgene_71535)") was used as entry vector (Sugano *et al.*, 2018).

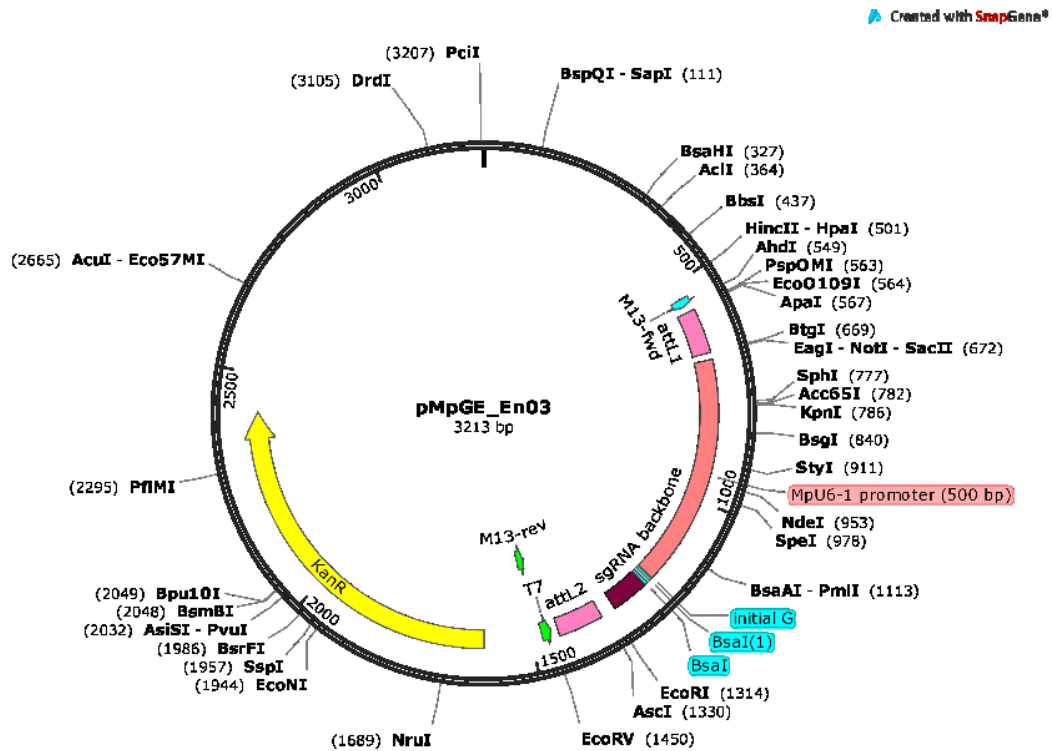


Figure 9 pMpGE_EN03 vector.

Used as entry vector which is already equipped with the sgRNA backbone. Features: MpU6-1 promoter, KanR (Kanamycin resistance), attL1, attL2 (attachment sites), sgRNA backbone, M13-fwd primer binding site, M13-rev primer binding site.

The target was ligated using the principle of cloning with restriction enzymes (BsaI) (Figure 10).

Materials and Methods

pMpGE_En03 BsaI digestion

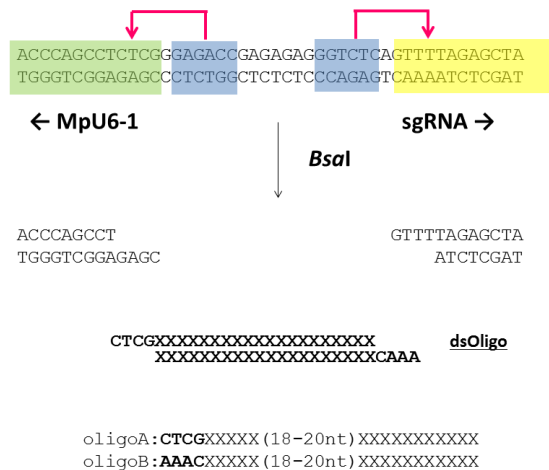


Figure 10 Detailed sequence view to visualise the principle of BsaI cut vector.

The vector can be ligated with the respective annealed targets flanked with BsaI overhangs (protocol with figures received from Shigeo Sugano, based on protocols from the Osakabe Lab).

The linearisation of the vector pMpGE_En03 was obtained with the following approach.

BsaI digestion of pMpGE_En03:

2–3 µg	pMpGE_En03
10–20 U	BsaI (10,000 U/ml)
2 µl	10x buffer CutSmart
<u>in 20 µl</u>	H ₂ O millipore

The linearisation was incubated at 37 °C for one hour. The linearised vector was isolated via gelelectrophoresis and gel extraction (see sections 4.2.4.7 and 4.2.4.8).

For annealing of the single stranded target to a double stranded target ordered as primers, the following approach was used.

Primer annealing:

sense gRNA primer 50 pmol
antisense gRNA primer 50 pmol
in 10 µl TE 1x

Programme primer annealing:

1. 95 °C, 5 minutes
2. 95–85 °C, –2 °C/s
3. 85 °C, 1 minutes
4. 85–25 °C, –0,2 °C/s
5. 25 °C hold

Materials and Methods

Ligation reaction:

x μ l	vector DNA (30–50 ng)
2 μ l	annealed primer DNA (10 pmol)
1 μ l	10x T4 ligase buffer
1 μ l	T4 ligase(400 U/ μ l)
in 10 μ l	H ₂ O millipore

The reaction was incubated at room temperature for at least one hour. Finally, the reaction was used to transform *Escherichia coli* (Kanamycin selection) (see section 4.2.5.9.1). Grown colonies were prepped via a mini preparation (see section 4.2.4.9.1). Sequencing (see section 4.2.5.10) allowed the verification of the correct insertion of the target sequence. The insertion can be checked by PCR (see section 4.2.4.6) using the M13-fwd primer and the antisense target primer. The usage of the M13-rev primer allowed to read out the sgRNA target sequence at position 400 bp.

Gateway reaction with entry vector and respective destination vector was performed as described in section 4.2.5.2. The first sgRNA next to the start codon of the respective gene (sgRNA1) was always cloned in pMpGE010 (Figure 11) (“pMpGE010 was a gift from Takayuki Kohchi (Addgene plasmid # 71536 ; <http://n2t.net/addgene:71536> ; RRID:Addgene_71536”)) equipped with the Cas9. The second sgRNA next to the stop codon of the respective gene (sgRNA2) was always cloned in pMpGWB401 (“pMpGWB401 was a gift from Takayuki Kohchi (Addgene plasmid # 68666 ; <http://n2t.net/addgene:68666> ; RRID:Addgene_68666”)) (Ishizaki *et al.*, 2015; Sugano *et al.*, 2018).

Figure 11 Vector cards of used destination vectors pMpGE010 and pMpGWB401.

Features: pVS1 StaA (stability protein), RB T-DNA repeat, M13 forward/reverse, attB1, attB2, attR1, attR2, Cas9, CmR (Cas module-RAMP (Repeat-Associated Mysterious Proteins)), ccdB (post-segregational killing), NOS terminator, CaMV poly(A) signal (Cauliflower mosaic virus polyadenylation signal), CaMV 35S promoter (enhanced), LB T-DNA repeat, SmR (Aminoglycoside adenyltransferase; spectinomycin resistance), ori, bom (basis of mobility), pVS1 oriV (Origin of replication from the plasmid pVS1), pVS1 RepA (Replication protein from the plasmid pVS1), NeoR (Neomycin)/KanR, HygR.

Materials and Methods

4.2.5.5 pJet cloning

Prior to final vector cloning, the constructs were cloned in the cloning vector pJet for amplification and verification. The following approach was used.

Ligation approach pJet:

0,5 µl	pJet
7,5 µl	fragment DNA
1,0 µl	10x Ligase buffer
1,0 µl	T4 DNA Ligase (400 U/µl)
<u>in 20µl</u>	H ₂ O millipore

The ligation approach was incubated for at least 30 minutes at room temperature.

4.2.5.6 Restriction

4.2.5.6.1 Restriction of DNA with compatible ends

The genes amplified from cDNA and ligated in the cloning vector pJet as well as the final vector were restricted with the respective enzymes (enzymes can be derived from vector maps). In the following, the restriction approaches of the insert as well as the final vector are listed.

vector:

2,0 µl vector DNA
1,0 µl 10x buffer
0,2 µl enzyme 1
0,2 µl enzyme 2
6,6 µl H₂O millipore
10 µl

insert:

4,0 µl vector DNA
1,0 µl 10x buffer
0,2 µl enzyme 1
0,2 µl enzyme 2
4,6 µl H₂O millipore
10 µl

The samples were incubated at 37 °C for at least one hour. The restriction fragments were separated and purified using an agarose gel (for purification see section 4.2.4.8).

4.2.5.6.2 Insert control

1 µl of each isolated plasmid was reviewed for correct insertion. The listed approach was applied. Enzymes already used for cloning were used for restriction as well. Buffers were chosen according to appropriate conditions (SDB/red/blue/orange/Tango/CutSmart/2.1/3.1).

Materials and Methods

restriction approach:

1 µl	plasmid DNA
0,5 µl	10x buffer
0,1 µl each	enzyme
<u>in 5µl</u>	H ₂ O millipore

Incubation was carried out at 37 °C for about half an hour to two hours. The approaches were analysed by means of gelectrophoresis. If the fragments were running as expected, these samples were sequenced.

4.2.5.7 Ligation

The following approach was used to ligate the respective constructs in the vector.

ligation approach:

2,0 µl	vector DNA, restricted
14,0 µl	fragment DNA
2,0 µl	10x ligase buffer
1,0 µl	T4 DNA Ligase (400 U/µl)
<u>in 20µl</u>	H ₂ O millipore

The sample was incubated for at least one to two hours at room temperature.

4.2.5.8 Generation of RbCl₂-competent cells

Chemically competent *E. coli* Top10 cells were used for cloning. The competent cells were generated with the help of the RbCl₂ method. A single colony of Top10 cells was grown overnight in 2 ml LB medium. From the overnight culture 100 ml LB medium were inoculated the following day. A 10 mM final concentration of MgCl₂ and MgSO₄ was adjusted in the medium. The culture was shaken at 200 rpm and 37 °C until an OD₆₀₀ of 0,5 to 0,6 was obtained. After an incubation step of the cells in 50 ml falcon tubes for 15 to 30 minutes on ice, the cells were centrifuged (3,000 x RCF, 10 minutes, 4 °C). After complete removal of the supernatant the cells were resuspended in 33 ml RF1 with a glass pipette. The cells were incubated again for 15 to 30 minutes on ice, and the cells were centrifuged again (3,000 x RCF, 10 minutes, 4 °C). The supernatant was removed and the cells were resolved in 5 ml RF2. The cells were incubated in 15 ml falcon tubes on ice for 15 to 30 minutes. Subsequently, the cells were aliquoted (100 µl aliquots). The aliquots were shock-frozen in liquid nitrogen and afterwards stored at -80 °C. The entire procedure was performed on ice and in a 4 °C cold room (Green and Rogers, 2013).

Materials and Methods

4.2.5.9 Bacterial transformation

4.2.5.9.1 *E. coli*

An aliquot of RbCl₂-competent *E. coli* cells was thawed on ice. 50 µl of these bacterial cells were mixed with the ligated vector on ice. Incubation was performed for two minutes on ice. The cells were heated at 42 °C for one minute followed by one minute on ice. The transformed bacterial cells were plated out on LB equipped with the respective antibiotic plates and incubated overnight at 37 °C. Individually grown colonies were picked the next day in 3 ml LB equipped with the respective antibiotic medium each. These liquid cultures were shaken and incubated overnight at 37 °C.

4.2.5.9.2 *Agrobacterium tumefaciens*

For *A. tumefaciens* mediated *M. polymorpha* transfection the electrocompetent *A. tumefaciens* strain C58C1 pGV2260 was used. All steps of bacterial transformation were performed at room temperature. *A. tumefaciens* was grown in 250 ml LB (Rif¹⁰⁰, Carb¹⁰⁰) for one to three days until an OD of 0,8 to 1,0 was reached. The bacterial culture was centrifuged for one minute and 11000 RCF. The supernatant was discarded. The bacteria were resuspended in 1 ml millipore water. The washing step was repeated five times. The pellet was resuspended in 2,5 ml of 10 % glycerol. The suspension was aliquoted (50 µl/reaction tube) and frozen at -80 °C. For bacteria transformation 50 µl competent cells as well as the plasmid to transform were added to a 2 mm cuvette for electroporation. The cells were shocked at 2,4 kV and 200 Ω. Afterwards, 300 µl LB was added and the cells were incubated at 28 °C for two hours. The cells were spread onto LB plates containing the respective selection antibiotics. The plates were grown for two days at 28 °C.

4.2.5.10 Sequencing

The cloned plasmids were sequenced for verification. Therefore, 5 µl of the isolated plasmids via mini preparation (1/10 dilution) were mixed with 5 µl of the respective primer (5 pmol/µl). The approaches to be sequenced were barcoded and sent to Macrogen (Amsterdam, NL). Using the program ChromasPro (version 2.1.8) the obtained sequences were aligned with the respective *in silico* set up sequences and checked for sequence identity.

Materials and Methods

4.2.6 Plant transfection

4.2.6.1 Transfection of *P. patens*

4.2.6.1.1 Linearisation of the constructs to be transformed

For stable integration of the desired constructs in the genome of *P. patens* via homologous recombination the plasmid DNA needed to be linearised. For linearisation 25-50 µg vector DNA were used. The following approach was applied.

linearisation approach:

25-50 µg	plasmid DNA
8 µl	10x buffer
2 µl	enzyme 1
2 µl	enzyme 2
<u>in 80 µl</u>	H ₂ O millipore

The restriction approach was incubated for at least one hour at 37 °C. The linearised constructs were checked for correct linearisation via gelelectrophoresis (see section 4.2.4.7) and purified via ethanol precipitation (see section 4.2.4.9.3).

4.2.6.1.2 Liquid culture

The transformation was performed in the *P. patens* Reute 17 background (Hiss *et al.*, 2017a). A filamentous grown liquid culture was used, which is weekly grinded and set up to a pH of 5,8 and 60 mg/l density. The culture was cultivated shaking under long day conditions in the long day chamber. Six to seven days prior to transfection the liquid culture was transferred to a fresh KNOP flask with a pH of 4,5.

4.2.6.1.3 Transfection

Transfection was obtained PEG mediated via protoplasts. Therefore, the cell wall had to be digested, which was done enzymatically with driselase. The plants are able to regenerate from a single cell to a whole plant (Boyd *et al.*, 1988; Cove, 2005).

0,2 g Driselase was dissolved in 5 ml 0,5 M Mannitol. The suspension was darkened with aluminium foil and rolled overhead for at least 30 minutes and centrifuged at 3500 rpm for 10 minutes. The driselase solution was filtered with a syringe and a sterile filter into a sterile falcon.

Materials and Methods

200 ml of the protonema culture was filtered with a 100 µm sieve. The moss was transferred with forceps into a fresh 9 cm petri dish and taken up into 12 ml 0,5 M mannitol. The petri dish was sealed with parafilm and shaken (30 rpm) for at least 30 minutes for equilibration. 4 ml of driselase solution were added to the equilibrated moss (final concentration 1 %). The digestion was darkened with aluminium foil and again shaken (30 rpm) for at least one hour. The following steps were performed carefully because of sensitivity of the protoplasts (wide neck pipettes/cut-off tips). The protoplast suspension was aspirated with a wide neck pipette and filtered through a 100 µm protoplast sieve. The petri dish was rinsed with 3 ml mannitol. The filtration step was repeated for a 50 µm protoplast sieve. The filtrate was distributed on two sterile glass or plastic tubes. A possible uneven distribution was compensated with mannitol. The protoplast suspension was centrifuged for ten minutes at 500 rpm (swing-out rotor). The supernatant was discarded. The protoplasts of each tube were resuspended in 10 ml of mannitol and resuspended by rolling the tubes between the hands. The centrifugation step was repeated. The protoplasts of each tube were taken up in 5 ml mannitol. Both tubes were then combined. An aliquot of 100 µl was taken and the protoplast number was calculated with the help of a Fuchs-Rosenthal counting chamber. The mean of four large squares was determined. Subsequently, another centrifugation step for ten minutes and 500 rpm followed. Calculation of the number of protoplasts: number of protoplasts per large square $\times 5000$ = number of protoplasts/ml; this number was multiplied by 10 because the protoplasts were taken up in a volume of 10 ml. The supernatant was discarded and the protoplasts were adjusted with 3M-medium to a concentration of $1,2 \times 10^6$ protoplasts/ml.

100 µl of DNA solution (5-25 µg DNA in 0,1 M $\text{Ca}(\text{NO}_3)_2$) were pipetted into sterile glass/plastic tubes. 250 µl protoplast solution was added as well as 350 µl PEG. The tubes were rolled between the hands. The PEG incubation was conducted for 30 minutes. Every five minutes the tubes were rolled between the hands. The transformation batch was then diluted every five minutes. Dilution was performed with 3M-medium starting at 1 ml, which increased by 1 ml every five minutes. The final dilution volume was 4 ml. The protoplasts were centrifuged again at 500 rpm for ten minutes. The supernatant was discarded. The protoplasts were resuspended in 1 ml (transient transfection) or 4 ml (stable transformation) regeneration medium and rolled. The transformed protoplasts were placed in the long day cabinet for regeneration.

4.2.6.1.4 Regeneration and selection

The transiently transfected protoplasts were analysed approximately after two days with the help of an upright DM6000 microscope equipped with a DFC295 camera. The stably transfected protoplasts were transferred to cellophane containing KNOP plates approximately four to five days after transfection. The plates were sealed with 3M tape and cultivated in long

Materials and Methods

day conditions. After seven days the transfected plants were transferred to KNOP plates containing selective antibiotics (e.g. G418 with a concentration of 25 µg/ml and Hygromycin 20 µg/ml). After one week another regeneration phase took place, followed by another selection round. The surviving plants were picked to fresh KNOP plates.

4.2.6.2 Transfection of *M. polymorpha*

The transfection was performed according to the protocol used in the Zachgo laboratory (University of Osnabrück). The transfection itself is a *A. tumefaciens* mediated sporeling transformation (time schedule in Table 3).

Table 3 Timetable for *A. tumefaciens* mediated *M. polymorpha* transfection.

The table shows an exemplary timetable for the different working steps that need to be done in the transfection process.

week	Mo	Tue	Wed	Thu	Fr	Sa	Sun
1.					start Mp-culture		
2.			start bacteria culture		Co-culture Mp+bacteria		
3.	selection						

The first step of the transfection process was the starting of the *M. polymorpha*, ssp. *ruderalis*, ecotype BoGa Osnabrück culture. 1 ml sterilisation solution was calculated for one spore reaction tube containing six sporangia. The tube was mixed and vortexed carefully and incubated for two minutes. The suspension was centrifuged for one minute at 6000 rpm. The supernatant was discarded. A washing step followed which was repeated three times. 1 ml millipore water was added to the sporelings in the reaction tube. The suspension was mixed carefully and incubated for two minutes. After a centrifugation step of one minute at 6000 rpm the supernatant was discarded. After the last washing step, the sporelings were resuspended in liquid medium (number of spore capsules x 100 µl). The suspension was mixed carefully again. The final step was the inoculation of liquid medium with sterilised sporelings. Therefore, 100 µl of sporelings per 25 µl liquid medium were inoculated (sterile in 100 ml flask). The suspension was mixed carefully and put in the cabinet for seven days with the appropriate conditions (22 °C, 16 hours light, 70 µmol m⁻²s⁻¹). The culture was shaken the whole time at 130 rpm.

The second step of the transfection process was the activation of the *A. tumefaciens* (C58C1 pGV2260) culture. Each construct, which should be transformed, was inoculated in 5 ml LB

Materials and Methods

medium with the appropriate selection. The suspension was incubated shaking for two days at 28 °C and 170 rpm.

The third step of the transfection process was the co-culture of *M. polymorpha* and *A. tumefaciens*. The bacteria culture was centrifuged at 20,000 RCF for 15 minutes. The bacteria were resuspended in 10 ml *M. polymorpha* liquid medium containing 2 µl Acetosyringone 1M (200 µM final concentration) and incubated for 6 hours at 28 °C and 175 rpm. 1 ml of the bacteria culture was added to 25 ml of the sporeling culture as well as 5 µl Acetosyringone 1M. The suspension was mixed carefully. The co-culture was incubated for three days shaking at 150 rpm in the cabinet (22 °C, 16 hours light 70 µmol m⁻²s⁻¹).

The final step of the transformation process was the selection. The content of the flasks was transferred in 50 ml falcons. If there was remaining culture in the flask, the flask could be washed with 10 ml Gamborg liquid medium and transferred in the falcon as well. The falcon was sealed. The cultures sedimented for 15 to 20 minutes. The supernatant was poured off until green *M. polymorpha* transformants reached the opening of the falcon. The transformants were washed with 50 ml Gamborg liquid medium. The falcon was sealed and inverted. Afterwards, the sedimentation step was repeated. The supernatant was discarded. The washing step was repeated four times. Approximately 2 ml of bacterial and *M. polymorpha* Gamborg culture should remain in the falcon. 50 to 200 µl of the remaining liquid were aspirated with a cut-off tip and distributed to four selection plates. The liquid was spread evenly with a flamed drigalski spatula. The plates were labelled accordingly and sealed with 3M tape and transferred in the cabinet (22 °C, 16 hours light, 70 µmol m⁻²s⁻¹). After one to two weeks transformants should be visible. The number of selection plates can be increased to obtain more putants.

4.2.7 Phenotyping

4.2.7.1 Selfing and crossing analyses

The number of sporophytes developed after 30 days, after watering by Reute as well as loss-of-function mutants, were counted in three independent mutant lines according to (Hiss *et al.*, 2017a) using a Leica S8Apo binocular. The crossed F1 generations were counted as described in (Hiss *et al.*, 2017a) using a fluorescence stereomicroscope SteREO LumarV12. At least 4 independent biological replicates of the selfed F1 generation were counted. At least three biological replicates per mutant line/control were counted with regard to the crossing (performed as described in (Perroud *et al.*, 2019)). The number of gametangiophores of *M. polymorpha* wild type as well as mutants were counted with the naked eye as well as with the help of a Leica S8Apo binocular. For synchronizing fertilisation, a drop of water was placed

Materials and Methods

onto an antheridiophore. Subsequently, the drop was resorbed with a pipette and pipetted onto an archegoniophore.

Microsoft Excel 2016 was used to perform statistical analyses and data visualisation. PlotsOfData was used for data visualisation (Postma and Goedhart, 2019). Statistical analyses were carried out with the webtools: <https://www.socscistatistics.com/tests/fisher/default2.aspx>, <https://www.graphpad.com/quickcalcs/contingency1/> and <http://quantpsy.org/chisq/chisq.htm>.

4.2.7.2 Phenotypic analysis of gametangia and sporophytes

A Leica S8Apo binocular was used for harvesting and counting of sporophytes and gametangia. An upright DM6000 microscope equipped with a DFC295 camera was used for taking microscopic pictures. Both devices ran under the control of the Leica Application suite version 4.4. Microsoft PowerPoint was used for processing (brightness and contrast adjustment) the images.

4.2.7.3 Spermatozoid analysis, DAPI and NAO staining

Spermatozoid analysis was performed according to (Meyberg *et al.*, 2020). Antheridia bundles were harvested 21 days after short day induction. Spermatozooids were stained with 4',6-diamidino-2-phenylindole (DAPI) staining. Antheridia bundles were harvested and opened for spermatozoid release with two ultra-fine forceps in 7 µl sterile tap water applied to an objective slide. The sample was dried at room temperature (RT) and afterwards fixed with 3:1 ethanol/acetic acid for at least five minutes. Samples were finally stained with 0,7 ng/µl DAPI in millipore water. Samples were additionally stained with a DAPI and 100 µM nonyl-acridine orange (NAO) double staining without prior fixation (Sanchez-Vera *et al.*, 2017). Samples were directly placed in the fixing solution. Sealing of samples was performed with nail polish (<https://www.marc-cain.com/index.php?cl=jaggeolandingpage&lang=0&sBackUrl=%2F>). Microscopic images were taken with an upright DM6000 microscope equipped with a DFC295 camera. Brightness and contrast of microscopy images was adjusted using Microsoft PowerPoint.

Results

5 Results

5.1 Determination of conserved single copy TAPs

TAPs are important regulatory hubs in the plant development, life cycle and biochemistry (Wilhelmsson 2017). The TAPs LEAFY (Sayou *et al.*, 2014) and PRC2 (Mosquana *et al.*, 2009; Okano *et al.*, 2009) are conserved single copy genes and serve as an example for the hypothesis that single copy genes are important key regulators in plant development that evolved under the Selected Single Copy Hypothesis. It states that copies of certain genes are not retained in most genomes, since it would lead to a stoichiometric imbalance of the gene products. The duplicates of these genes are therefore counter-selected to avoid dosage effects (Duarte *et al.*, 2010; Edger and Pires, 2009). The Selected Single Copy Hypothesis contrasts the Gene Balance Hypothesis (Birchler and Veitia, 2012) which postulates that dosage-sensitive genes show a tendency to be retained after large scale duplication events. On top of that, it was shown that plant morphogenic complexity coincides with an increasing genomic fraction of TAPs (Lang *et al.*, 2010; Van de Peer *et al.*, 2009).

The approach of identification of conserved single copy TFs/TRs started with the usage of an ortholog clustering approach, proteinortho (Lechner *et al.*, 2011). With the help of proteinortho a table with gene identifiers was created to predict orthologs in a broad range of plant species. The initial ortholog clustering approach was done by Kristian K. Ullrich (MPI Plön, DE). The table was reduced to the species *P. patens* and *M. polymorpha*, which resulted in 2,654 genes (Figure 12). The table was reduced to single copy genes (383 genes), which had to fulfil the prerequisite of being a TAP following the domain rule set described in (Wilhelmsson *et al.*, 2017) (38 genes). The 38 potential single copy TAPs were classified in 23 different TAP families (Wilhelmsson *et al.*, 2017), namely GNAT, MYB-related, SET/Sigma70-like, FHA, PcG_MSI, Jumonji_Other, IWS1, C3H, C2H2, bHLH, TANGO2, tify, TUB, WRKY, MADS, Med7, O-FucT, BES1, CCAAT_HAP3, GRAS, HMG, ARID and Med6. For these 23 families, phylogenetic analyses (Frickenhaus and Beszteri, 2008) were performed with the species set of *A. thaliana*, *Anthoceros agrestis*, *M. polymorpha*, *P. patens* as well as *Oryza sativa* to verify single copy status in the bryophytes of interest. The main focus for single copy status were the species *A. thaliana* as well as the three bryophytes *P. patens*, *M. polymorpha* and *A. agrestis*, the former because *A. thaliana* is the best characterised species in terms of mutant phenotypes, the latter three as representatives of all three lineages of bryophytes (mosses, liverworts and hornworts) (Bowman *et al.*, 2017; Wickett *et al.*, 2014). *O. sativa* was chosen as monocot representative since *A. thaliana* belongs to the dicots.

Results

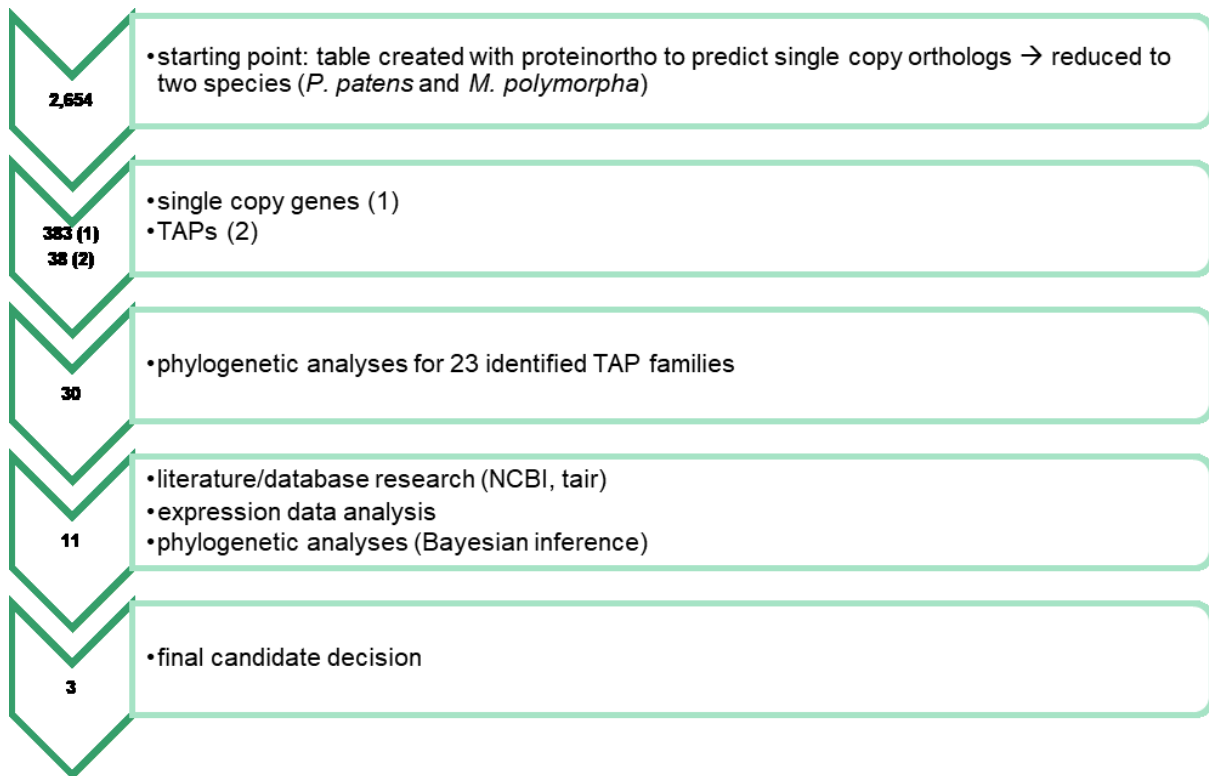


Figure 12 Workflow of candidate gene search approach.

Figure 13 displays the phylogenetic analysis of the TAP family Sigma70-like, shown as an example for the 23 TAP family phylogenies. As depicted in the flow chart (Figure 12) the phylogenetic analyses of the TAP families reduced the candidate gene number from 38 to 30. In fact, some of the phylogenies showed that the assumed single copy TAP is indeed not single copy or that there is no ortholog present in all species of interest. Finally, 14 TAP families remained in the workflow. Figure 13 shows a phylogenetic tree that confirmed the hypothesis of the gene in questions being single copy in all of the five analysed species, which is depicted by the red bars. Only one single copy ortholog cluster in all five analysed species was detected in the Sigma-70 like TAP family.

Results

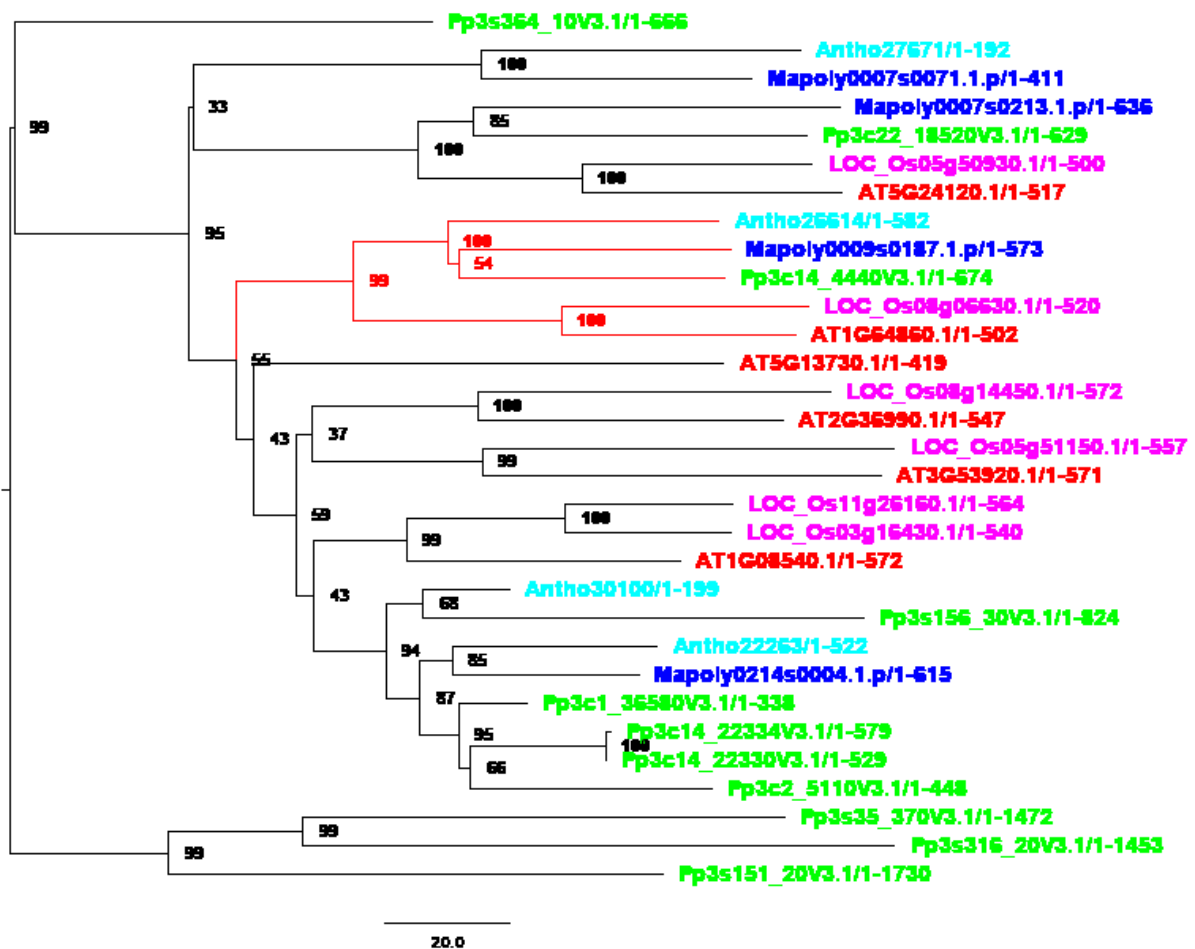


Figure 13 Sigma70-like phylogenetic analysis.

There are clear single copy orthologs of the *A. thaliana* gene (red) in the three bryophyte species of interest, *P. patens* (moss, green), *A. agrestis* (hornwort, turquoise) and *M. polymorpha* (liverwort, blue) as well as in *Oryza sativa* (pink), which is depicted with the red bars. Support values at the nodes represent bootstrap values.

The 30 confirmed single copy gene candidates were analysed via literature research as well as Tair description (*A. thaliana*), GO_annotation_BP (biological process), GO_annotation_MF (molecular function), GO_annotation_CC (cellular component) (Lang *et al.*, 2018). A table (data not shown) with all the collected data was created. Subsequently, the table was reduced to 11 candidate genes. 19 candidate genes were discarded because there was no information available about the gene in *A. thaliana*, the gene was tackled in bryophytes already in other approaches, or the literature showed no promising phenotype with a strong focus on sexual reproduction. The literature research in many cases allowed learning about protein function derived from loss-of-function mutants, usually in *A. thaliana*. Table 4 depicts the 11 candidate genes, for which Bayesian inference was applied to confirm the single copy status of the genes. The coloured arrows show whether there is one, two or no clear *P. patens* ortholog for the respective gene.

Results

Table 4 Candidate gene selection analysed in terms of single copy status via Bayesian inference.

The number of *P. patens* orthologs is depicted by the arrow colour (green=one, orange=two, red=no). The table lists information of the respective gene like *P. patens*, *A. thaliana* identifier, TAP family, Bayesian inference and Tair description/GO biological process (Lang *et al.*, 2018).

TAP family	<i>P. patens</i> Identifier	Bayesian inference	TAIR10_id	Tair description/GO biological process
C2H2	Pp3c11_480V3.1	(↑) two paralogs	AT1G72050	Encodes a transcriptional factor TFIIIA required for transcription of 5S rRNA gene. 5S rRNA is the smallest constituent of the ribosome. Work on one of the gene models AT1G72050.2 showed that it encodes a protein with nine Cys(2)-His(2)-type zinc fingers, a characteristic feature of TFIIIA proteins. AT1G72050.2 also contains a 23 amino acid spacer between fingers 1 and 2, a 66 amino acid spacer between fingers 4 and 5, and a 50 amino acid non-finger C-terminal tail. in vitro assay demonstrated that AT1g72050.2 binds to 5S rDNA and efficiently stimulates the transcription of 5S rRNA. AT1g72050.2 also binds to 5S rRNA in vitro. AT1g72050.2 is located at several nuclear foci including the nucleolus and is absent from the cytoplasm. multicellular organism development, regulation of transcription, DNA-templated, transcription, DNA-templated
FHA	Ppc12_14900V3.1	↑	AT5G47790	SMAD/FHA domain-containing protein ; CONTAINS InterPro DOMAIN/s: SMAD/FHA domain (InterPro:IPR008984), Forkhead-associated (FHA) domain (InterPro:IPR000253); BEST Arabidopsis thaliana protein match is: SMAD/FHA domain-containing protein (TAIR:AT5G38840.1)
GNAT	Pp3c5_21790V3.1	↓ no clear ortholog	AT2G39020	ornithine metabolic process; Although this locus shares considerable sequence similarity with the adjacent NATA1 gene (At2g39030), they appear to encode genes with different functions. NATA1 is involved in the production of N-delta-acetylornithine, but, overexpression of At2g39020 in tobacco does not lead to the formation of this defense compound. The mRNA is cell-to-cell mobile.
GNAT	Pp3c7_5970V3.1	(↑) two paralogs	AT5G13780	embryo development ending in seed dormancy, response to water deprivation
GNAT	Pp3c2_9560V3.1	↑	AT3G54610	Encodes a histone acetyltransferase that is plays a role in the determination of the embryonic root-shoot axis. It is also required to regulate the floral meristem activity by modulating the extent of expression of WUS and AG. In other eukaryotes, this protein is recruited to specific promoters by DNA binding transcription factors and is thought to promote transcription by acetylating the N-terminal tail of histone H3. The enzyme has indeed been shown to catalyse primarily the acetylation of H3 histone with only traces of H4 and H2A/B being acetylated. Non-acetylated H3 peptide or an H3 peptide that had been previously acetylated on K9 both serve as excellent substrates for HAG1-catalyzed acetylation. However, prior acetylation of H3 lysine 14 blocks radioactive acetylation of the peptide by HAG1. HAG1 is specific for histone H3 lysine 14. flower development, histone acetylation, positive regulation of transcription, DNA-templated, regulation of vegetative phase change, response to light stimulus, root morphogenesis, transcription, DNA-templated

Results

Jumonji_Other	Pp3c23_15970V3.1	↑	AT5G06550	Encodes a HR demethylase that acts as a positive regulator of seed germination in the PHYB-PIL5-SOM pathway. cell surface receptor signaling pathway, histone H4-R3 methylation, positive regulation of seed germination
Med7	Pp3c5_20850V3.1	(↑) two paralogs	AT5G03500	Mediator complex, subunit Med7; FUNCTIONS IN: RNA polymerase II transcription mediator activity; INVOLVED IN: regulation of transcription from RNA polymerase II promoter; LOCATED IN: mediator complex; CONTAINS InterPro DOMAIN/s: Mediator complex, subunit Med7 (InterPro:IPR009244); BEST Arabidopsis thaliana protein match is: Mediator complex, subunit Med7 (TAIR:AT5G03220.1)
MYB-related	Pp3c2_24980V3.1	↑	AT2G47620	chromatin remodelling, covalent chromatin modification, multicellular organism development, regulation of transcription, DNA-templated, transcription, DNA-templated
MYB-related	Pp3c14_5130V3.1	(↑) two paralogs	AT5G06110	DnaJ domain ;Myb-like DNA-binding domain; FUNCTIONS IN: heat shock protein binding, DNA binding; INVOLVED IN: protein folding; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Molecular chaperone, heat shock protein, Hsp40, DnaJ (InterPro:IPR015609), Heat shock protein DnaJ, N-terminal (InterPro:IPR001623), Heat shock protein DnaJ, conserved site (InterPro:IPR018253), MYB-like (InterPro:IPR017877), SANT, DNA-binding (InterPro:IPR001005), Myb, DNA-binding (InterPro:IPR014778), Homeodomain-like (InterPro:IPR009057); BEST Arabidopsis thaliana protein match is: DnaJ domain ;Myb-like DNA-binding domain (TAIR:AT3G11450.1) cell division, protein folding
SET/Sigma70-like	Pp3c14_4440V3.1	↑	AT1G64860	DNA-templated transcription, initiation, cellular response to light stimulus, cellular response to redox state, photosystem stoichiometry adjustment, regulation of RNA biosynthetic process, regulation of transcription, DNA-templated
TUB	Pp3c3_17070V3.1	↓ no clear ortholog	AT2G47900	cellular response to osmotic stress, regulation of transcription, DNA-templated, response to fungus, response to hydrogen peroxide, response to salt stress

The genes with no clear ortholog were discarded directly, whereas the genes with one or two orthologs were analysed again in terms of literature research and expression data with emphasis on the alternation of generations.

Finally, three genes were selected for further analyses because of their impact on sexual reproduction shown in the literature as well as their expression data (*A. thaliana*, *M. polymorpha* and *P. patens*) indicating an involvement in the alternation of generations. Two of the three genes are single copy genes whereas one candidate has two paralogs (AT5G13780). Only two of three genes were analysed via loss-of-function mutants, this is why only the respective two genes were elucidated in more detail. The first one is a histone acetyltransferase, orthologous to *A. thaliana* hag1, and the second gene is part of the SWITCH/SUCROSE NONFERMENTING (SWI/SNF) complex, orthologous to *A. thaliana* swi3a/b.

Results

Apart from the genes that were selected for further experiments, the approach resulted in the identification of genes known as important regulators of the plant life cycle, e.g. PRC2 (Mosquna *et al.*, 2009; Motomura *et al.*, 2016; Okano *et al.*, 2009) reinforcing the notion that the selection procedure will find important developmental regulators.

5.2 Candidates to study the alternation of generations

As mentioned above, epigenetic modifications may control cell fate (Gaspar-Maia *et al.*, 2011). Histone acetylation is a potential regulatory mechanism involved in transcriptional activation (Burk *et al.*, 2003) and therefore associated with many developmental and biological processes in eukaryotes (Verdone *et al.*, 2005). Histone acetyltransferase proteins are highly conserved in all eukaryotic lineages including plants as well as yeast, flies and mammals (Aquea *et al.*, 2017; Kim *et al.*, 2015). The *A. thaliana* HAG1 ortholog AT3G54610 (Kim *et al.*, 2015), also known as GCN5, was shown to be predominantly involved in the acetylation of H3K14 (Benhamed *et al.*, 2006; Earley *et al.*, 2007; Mahrez *et al.*, 2016). As a fundamental regulator, HAG1 is associated with a high number of developmental processes (Aquea *et al.*, 2017), e.g. timing of juvenile-to-adult phase transition (Kim *et al.*, 2015), inflorescence meristem and stamen development (Cohen *et al.*, 2009), and miRNA production (Kim *et al.*, 2009).

The phylogenetic analysis (Figure 14) of the HAG1 clade shows that there are single copy orthologs of the *A. thaliana* gene (red) in the moss *P. patens* (green), the hornwort *A. agrestis* (turquoise) and the liverwort *M. polymorpha* (blue). *Homo sapiens* is coloured in yellow. HAG1/GCN5 belongs to the GNAT TR family defined by the presence of the GNAT acetyltransferase domain (PFAM PF00583) (Wilhelmsson *et al.*, 2017). Apart from the plant kingdom, the phylogenetic analysis shows that there are orthologs in a broad range of eukaryotes, e.g. the animal kingdom, the fungi as well as SAR. In most of the eukaryotic lineages hag1 is a single copy gene, whereas in vertebrates the gene was duplicated, so that there are two copies in all the analysed vertebrates.

Results

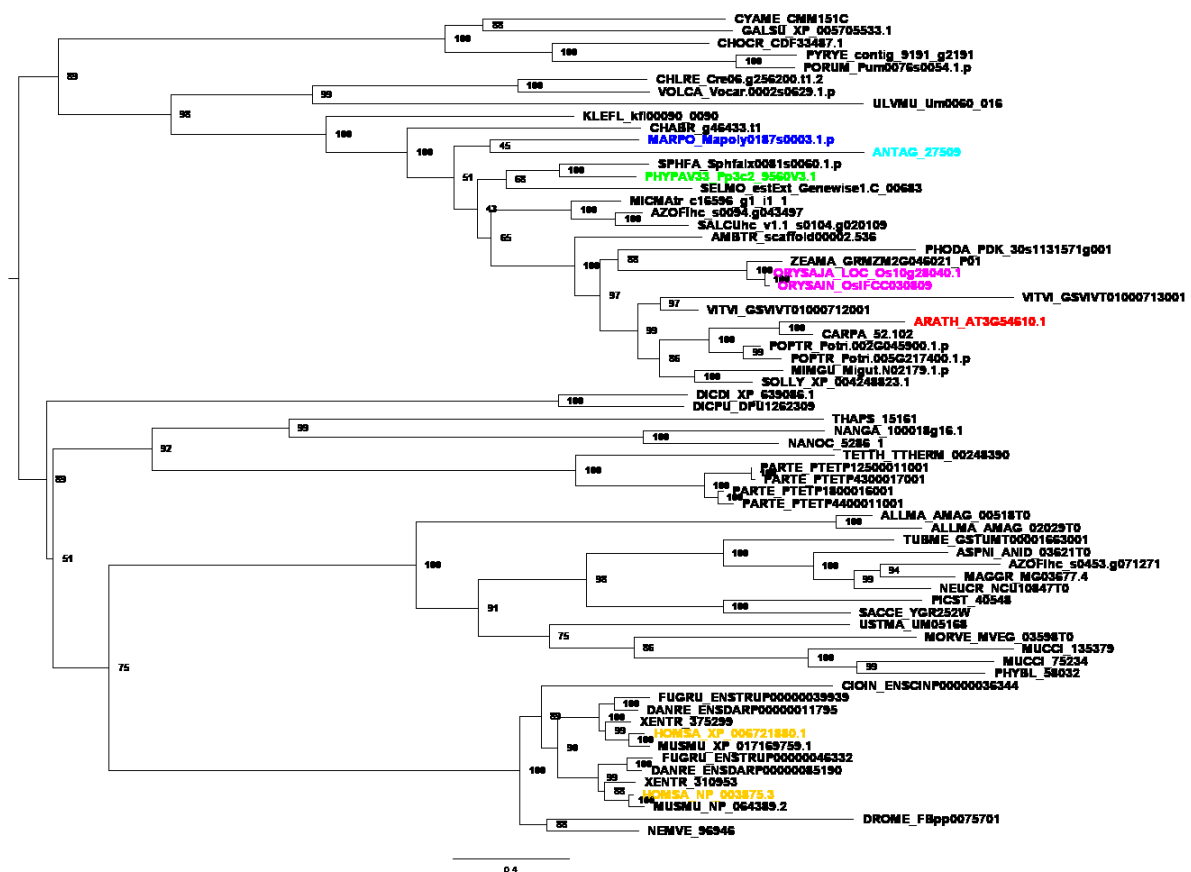


Figure 14 HAG1 phylogenetic analysis.

There are clear single copy orthologs of the *A. thaliana* gene (red) and the three bryophyte species of interest, *P. patens* (moss, green), *A. agrestis* (hornwort, turquoise) and *M. polymorpha* (liverwort, blue), *H. sapiens* (yellow). Support values in percentage at the nodes represent maximum likelihood. Species names are abbreviated in a five letter code where the first three letters represent the species and the latter two the genus e.g. *ORYza SATiva* = ORYSA [red algae: CYAME = *Cyanidioschyzon merolae*, GALSU = *Galdieria sulphuraria*, CHOCH = *Chondrus crispus*, PORUM = *Porphyra umbilicalis*, PYRYE = *Pyropia yezoensis*, chlorophyte algae: CHLRE = *Chlamydomonas reinhardtii*, ULVUM = *Ulva mutabilis*, VOLCA = *Volvox carteri*, streptophyte algae: KLEFL = *Klebsormidium nitens*, CHABR = *Chara braunii*, bryophytes: ANTAG = *Anthoceros agrestis*, MARPO = *Marchantia polymorpha*, PHYPA = *Physcomitrium patens*, SPHFA = *Sphagnum fallax*, SELMO = *Selaginella moellendorffii* (lycophyte), ferns: AZOFI = *Azolla filiculoides*, SALCU = *Salvinia cucullata*, MICMA = *Microlepia marginata*, angiosperms: ZEAMA = *Zea mays*, ORYSA = *Oryza sativa*, ARATH = *Arabidopsis thaliana*, POPTR = *Populus trichocarpa*, CARPA = *Carica papaya*, SOLLU = *Solanum lycopersicum*, AMBTR = *Amborella trichopoda*, MIMGU = *Mimulus guttatus*, VITVI = *Vitis vinifera*, PHODA = *Phoenix dactylifera*, fungi: ALLMA = *Allomyces macrogynus*, MORVE = *Mortierella verticillata*, MUCCI = *Mucor circinelloides*, PHYBL = *Phycomyces blakesleeana*, USTMA = *Ustilago maydis*, SACCE = *Saccharomyces cerevisiae*, PICST = *Scheffersomyces stipitis*, TUBME = *Tuber melanosporum*, ASPNI = *Aspergillus nidulans*, NEUCR = *Neurospora crassa*, MAGGR = *Magnaporthe grisea*, amoebozoos: DICPU = *Dictyostelium purpureum*, DICDI = *Dictyostelium discoideum*, SAR: THAPS = *Thalassiosira pseudonana*, NANOC = *Nannochloropsis oceanica* CCMP1779, NANGA = *Nannochloropsis gaditana*, TETTH = *Tetrahymena thermophila*, PARTE = *Paramecium tetraurelia*, invertebrates: CIOIN = *Ciona intestinalis*, NEMVE = *Nematostella vectensis*, DROME = *Drosophila melanogaster*, vertebrates: HOMSA = *Homo sapiens*, FUGRU = *Fugu rubripes*, XENTR = *Xenopus tropicalis*, MUSMU = *Mus musculus*, DANRE = *Danio rerio*].

Results

Chromatin-remodelling complexes (CRCs) constitute a switch point between signalling pathways and chromatin-based control of transcription. They are involved in transcriptional activation as well as repression (Roberts and Orkin, 2004; Sarnowski *et al.*, 2005). AtSWI3A/B (AT2G47620) is part of the SWITCH/SUCROSE NONFERMENTING (SWI/SNF) CRC, which is involved in ATP-dependent alteration of DNA-histone contacts (Sarnowski *et al.*, 2005). The genome of *A. thaliana* encodes four SWI3-like proteins, which are diverse in their function. Mutants of Atswi3a and Atswi3b have severe defects in embryo development, causing a lethal arrest at the early globular stage (Sarnowski *et al.*, 2005). The phylogenetic analysis (performed by Noé Fernández Pozo (University of Marburg, DE)) (Figure 15) of SWI3 shows that there are four single copy SWI3-like proteins in *A. thaliana*. SWI3A/B and SWIC/D cluster in two distinct subtrees, whereas SWIC/D seems to be the ancestral clade with the algae clustering only in the SWI3C/D subtree not in the A/B subtree. The SWI3C/D clade shows that there are no orthologs for C in the bryophytes, whereas D has single copy orthologs for *M. polymorpha* as well as *A. agrestis*. Regarding *P. patens* there are three paralogs for SWI3D. The SWI3A/B clade shows that there are single copy orthologs in the moss *P. patens* (green), the hornwort *A. agrestis* (turquoise) and the liverwort *M. polymorpha* (blue).

SWI3A/B are part of the MYB-related TF family characterised by the Myb_DNA-binding (PF00249) according to the TAPscan classification ruleset (Wilhelmsson *et al.*, 2017). Apart from this database, the PLANT TRANSCRIPTION FACTOR DATABASE, University of Potsdam (http://plntfdb.bio.uni-potsdam.de/v3.0/fam_mem.php?family_id=MYB-related&sp_id=ATH) classifies SWI3A and SWI3B as members of the MYB-related family, which is also characterised on that database by the Myb_DNA-binding domain (Martin and Paz-Ares, 1997), even annotated as DNA binding. According to the domain description of Myb_DNA-binding (PF00249), the domain family contains not only DNA binding domains but the SANT domain as well (Aasland *et al.*, 1996). They show a high structural similarity (Grune *et al.*, 2003) but the SANT domain is predicted to be incompatible with DNA binding (Boyer *et al.*, 2004). Taken together, SWI3A/B cannot clearly be classified as TF or TR.

Results

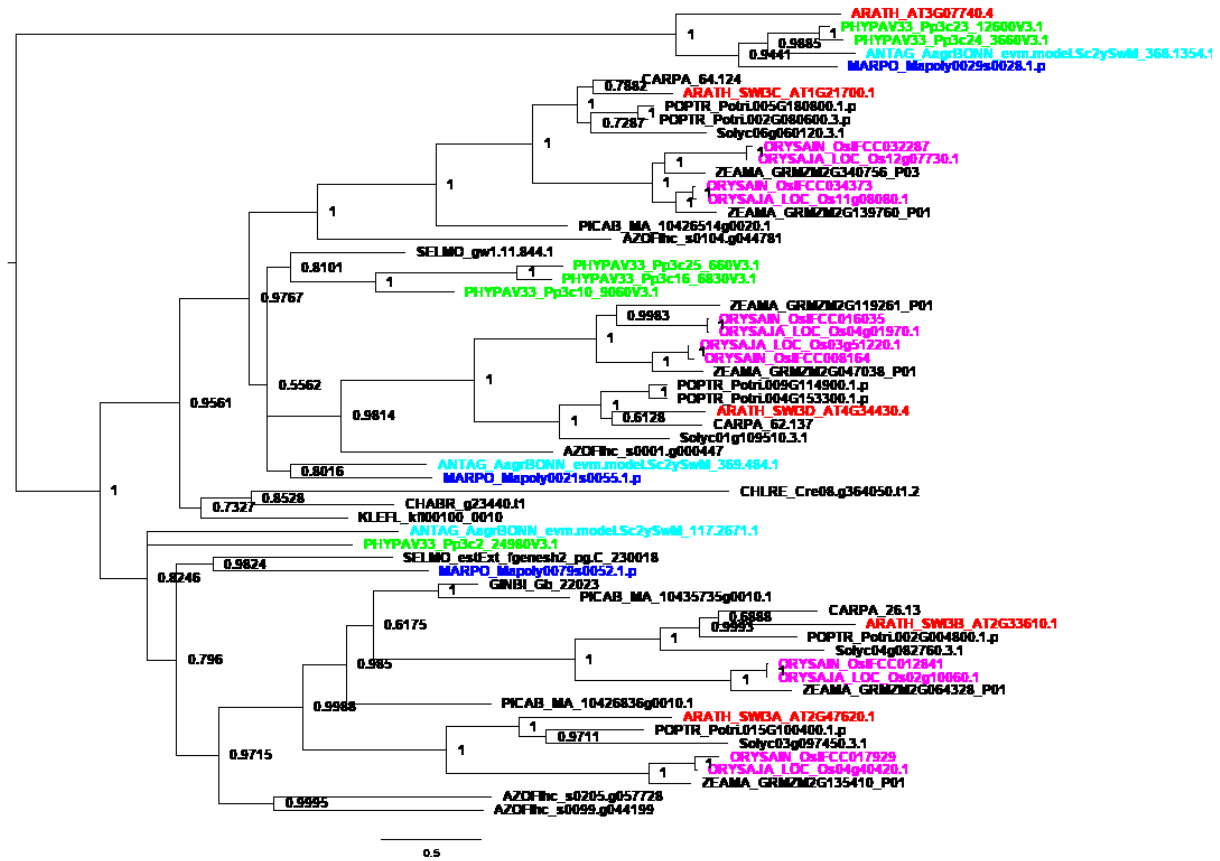


Figure 15 SWI3 phylogenetic analysis.

The phylogenetic analysis confirms single copy status for four *A. thaliana* SWI-like proteins (A,B,C,D). The SWI3A/B clade represented by the *A. thaliana* proteins (red) shows clear single copy orthologs for the three bryophyte species of interest, *P. patens* (moss, green), *A. agrestis* (hornwort, turquoise) and *M. polymorpha* (liverwort, blue). Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in a five letter code where the first three letters represent the species and the latter two the genus, e.g. ORYZa Sativa, ORYZA (rice), [angiosperms: ARATH = *A. thaliana* (thale cress), CARPA = *Carica papaya* (papaya), POPTR = *Populus trichocarpa* (poplar), SOLYC = *Solanum lycopersicum* (tomato), ZEAMA = *Zea mays* (corn), gymnosperms: PICAB = *Picea abies* (Norway spruce), GINBI = *Ginkgo biloba*, AZOFI = *Azolla filiculoides* (water fern), SELMO = *Selaginella moellendorffii* (lycophyte), bryophytes: PHYPa = *Physcomitrium patens* (moss), MARPO = *Marchantia polymorpha* (liverwort), ANTAG = *Anthoceros agrestis* (hornwort), streptophyte algae: CHABR = *Chara braunii*, KLEFL = *Klebsormidium flaccidum*, CHLRE = *Chlamydomonas reinhardtii* (chlorophyte alga)].

Results

5.3 Localisation and expression

A swi3a/b / hag1 expression profile (Figure 16) during *P. patens*' life cycle was created using RNA-seq data of major developmental stages (Fernandez-Pozo *et al.*, 2019; Meyberg *et al.*, 2020). The RNA-seq data show an expression peak in both genes in sporophytes (especially in pre-meiotic sporophytes) compared to the vegetative tissue (Hiss *et al.*, 2017a). The highest expression takes place in antheridia bundles.

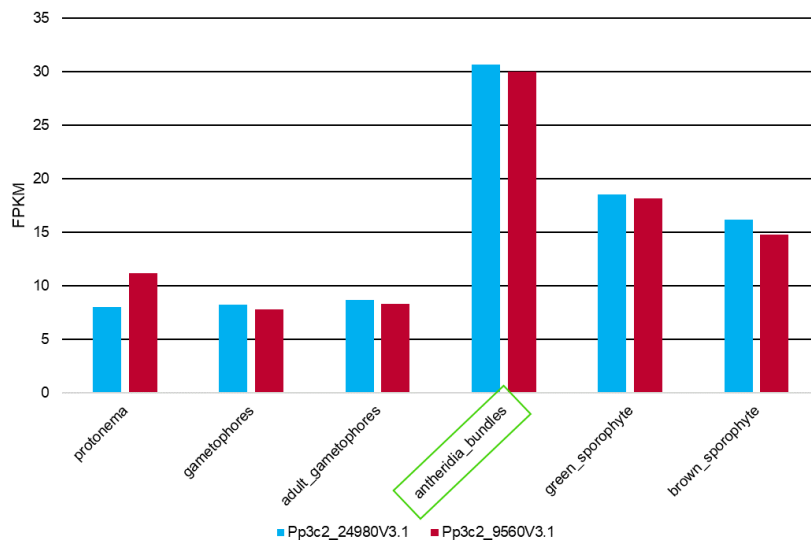


Figure 16 Expression profile of Pphag1 (Pp3c2_9560V3.1, histone acetyltransferase) and Ppswi3a/b (Pp3c2_24980V3.1, chromatin remodelling complex subunit).

The highest expression of both genes takes place in antheridia bundles (green box) (Fernandez-Pozo *et al.*, 2019). Expression data of antheridia bundles was taken from (Meyberg *et al.*, 2020).

M. polymorpha RNA-seq expression data was obtained due to the recent sequencing project of its genome (Bowman *et al.*, 2017). Figure 17 shows the expression profiles of both candidate genes throughout all developmental stages (Mphag1 Mapoly0187s0003, Mpswi3a/b Mapoly0079s0052). The green box indicates the high expression of both genes in antheridiophores as well as archegoniophores compared to the other developmental stages. Mpswi3a/b shows a high expression in the gametophore and the diploid generation (sporophyte + archegoniophore) as well. In general, the expression of Mpswi3a/b seems to be higher compared to Mphag1 in all analysed tissues.

Results

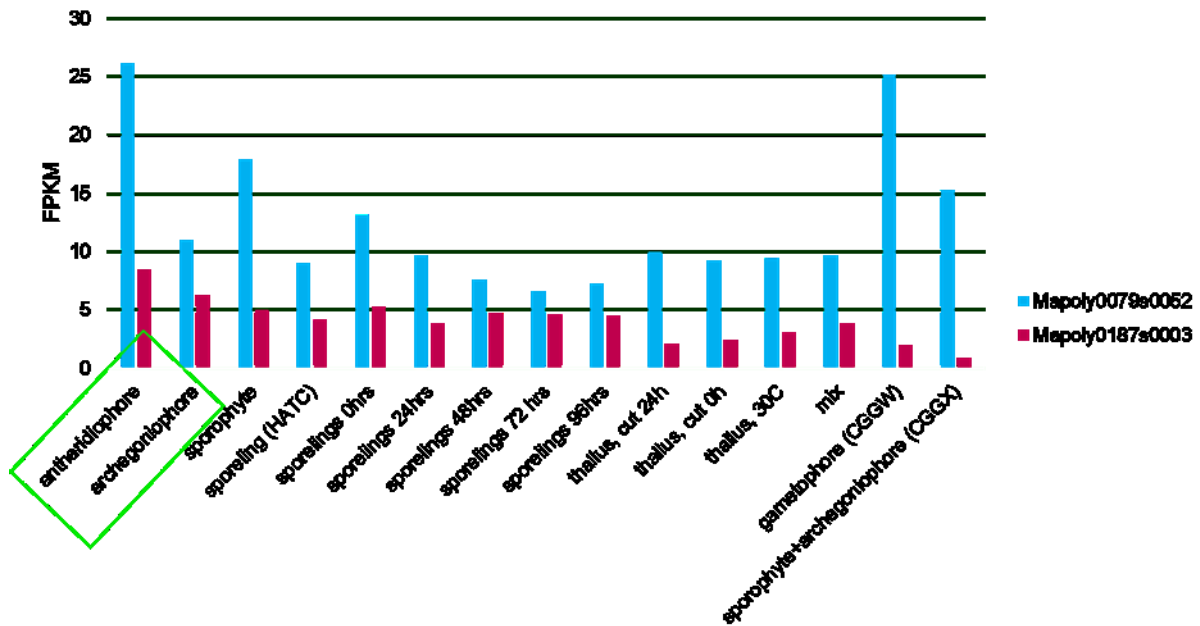


Figure 17 Expression profile of Mphag1 (Mapoly0187s0003, histone acetyltransferase) and Mpswi3a/b (Mapoly0079s0052, chromatin remodelling complex subunit).

The green box indicates the high expression of both genes in antheridiophores and archegoniophores, with the highest expression in antheridiophores, respectively (Bowman *et al.*, 2017).

A. thaliana expression analysis of developmental stages (array data) was performed with the eFP browser web tool (Winter *et al.*, 2007). Figure 18 illustrates the expression data of the developmental stages of Athag1 (AT3G54610), whereas Figure 19 represents the expression data of Atswi3a (AT2G47620) and Atswi3b (AT2G33610). The data shows high expression in the shoot apex inflorescence as well as early stages of flowering for all three genes.

Results

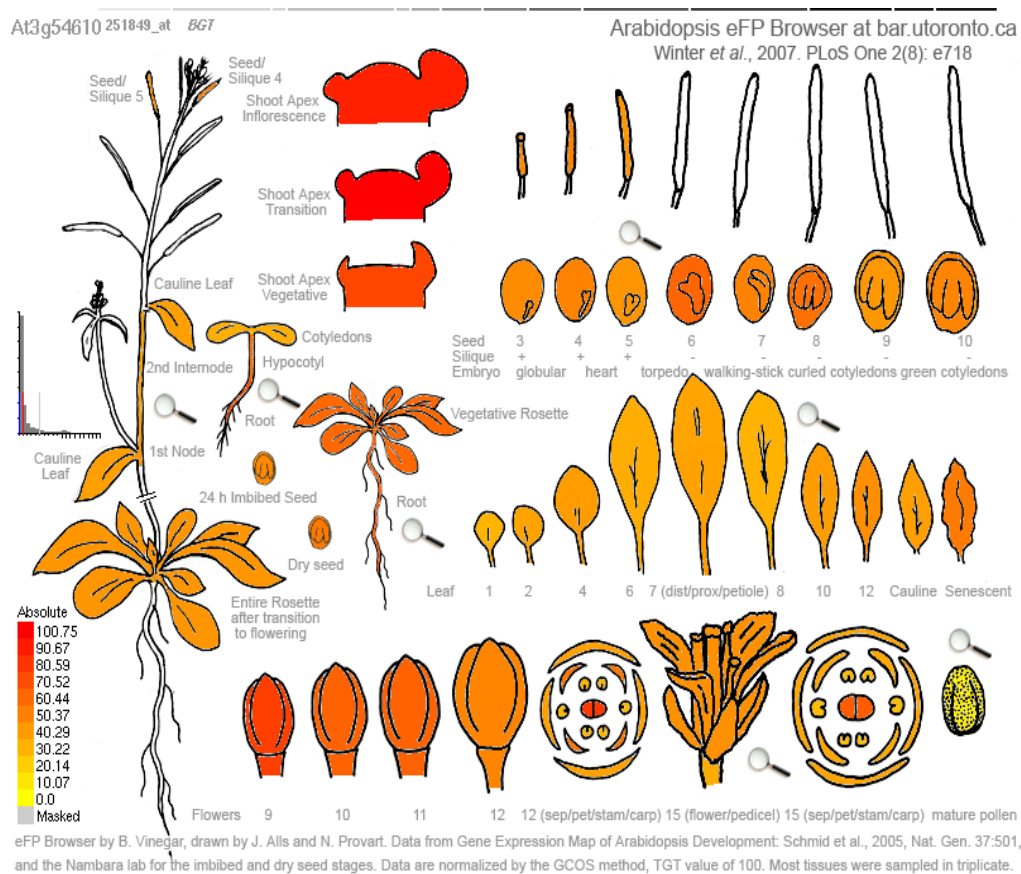


Figure 18 Expression profile of the developmental stages of *Athag1* (histone acetyltransferase).

Results

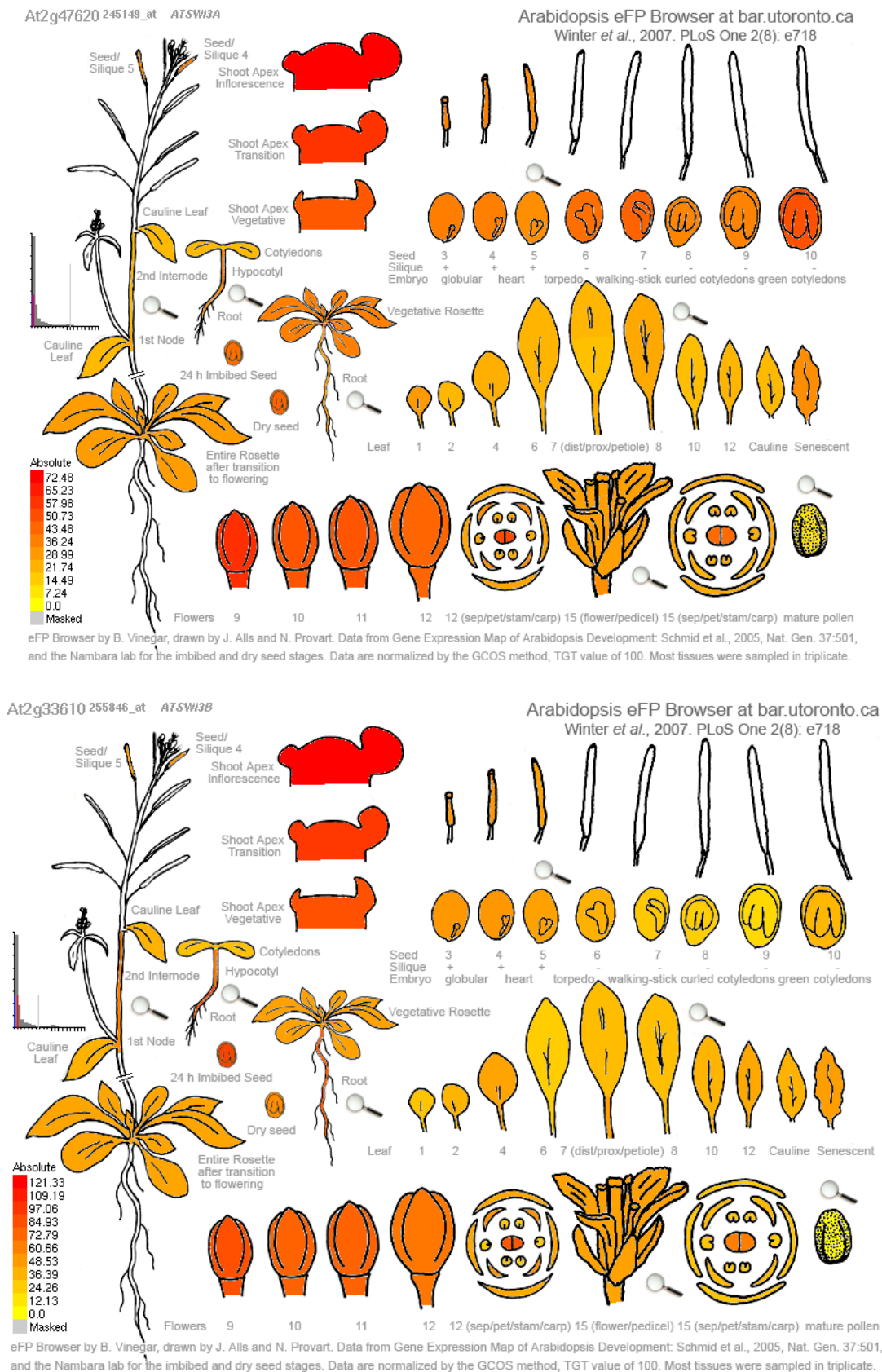


Figure 19 Expression profile of the developmental stages of Atswi3a and b (chromatin remodelling complex subunit).

Results

An *in silico* analysis predicted a nuclear localisation and a NLS (nuclear localisation sequence) of both *P. patens* gene products (Sperschneider *et al.*, 2017).

For *in vivo* localisation analysis the vector pCABsh_Loc was used (Hiss *et al.*, 2017b). The genes (Pp3c2_9560V3.1, Pp3c2_24980V3.1) were fused in frame to GFP (C-terminal GFP tag) under the control of the CABsh Promoter (Figure 20 shown as example) (Hiss *et al.*, 2017b).

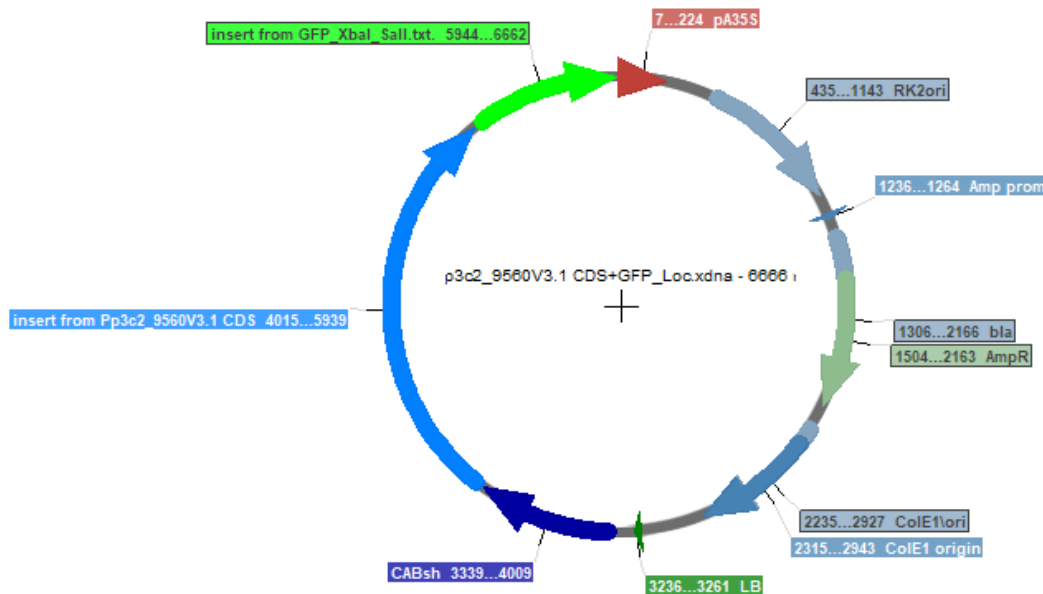


Figure 20 Vector card of transiently expressing candidate gene Pphag1 with a C-terminal GFP tag under the control of the CABsh Promoter (Hiss *et al.*, 2017b).

Pphag1 was fused in frame to GFP (C-terminal tag) for a transient expression of the gene to analyse the gene localisation. Features: terminator pA35S, RK2ori, bla/AmpR, ColE1 ori, CABsh promoter, Amp promoter.

Since a strong green fluorescent signal was observed in the nuclei of protoplasts transiently transformed with GFP-tagged genes the prediction was confirmed *in vivo* (Figure 21).

Results

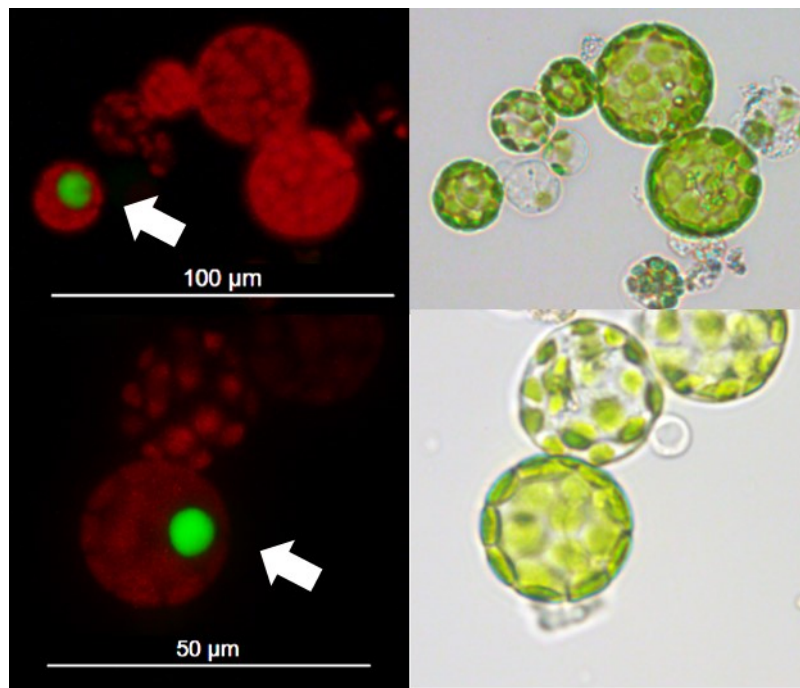


Figure 21 *In vivo* localisation analysis.

Transient expression of Ppswi3a/b and Pphag1 fused to a C-terminal GFP (upper panel Pphag1, lower panel Ppswi3a/b). The white arrows indicate the nuclear localisation.

5.4 *P. patens* loss-of-function CRISPR/Cas9 mutants – loss of sporophytes

To determine the gene function of the two candidate genes in bryophytes and to place the gene function in an evolutionary context, loss-of-function mutants were obtained with the help of CRISPR/Cas9 gene perturbation.

Mutants were identified by PCR and sequencing of amplified PCR products. Therefore, forward primers were designed upstream of the sgRNA1 target site. Reverse primers were designed 700 to 1,300 bp downstream of the sgRNA2 target site. Mutants were detected by a shortened (approximately 200 bp) PCR product. The cut out of approximately 200 bp led to either a shortened protein, if the cut out was in frame, or a frameshift resulting in a newly created stop codon. For every candidate gene at least five independent mutant lines were identified. For further analyses, namely analysis of sporophyte development as well as crossing analysis, three independent mutant lines as well as the background strain Reute as control were used. Morphological analyses (5.6, 5.7) as well as RNA-seq analysis (5.9) were carried out for one mutant line each as well as Reute as control. The used Ppswi3a/b line was chosen since that mutant line did not develop any sporophyte in all selfing experiments (Figure 22). The Pphag1 line used was chosen with the help of the crossing analysis, since there was no difference in the three analysed mutant lines with regard to the selfing experiments (Figure 22). The mutant line was chosen because of its lowest heterozygous rate of sporophyte development (Figure 26).

Results

In contrast to the control (Reute), which has a sporophyte ratio of over 98 % sporophytes per gametophore, the loss-of-function mutants of both genes showed a significant reduction of sporophytes per gametophore (Figure 22).

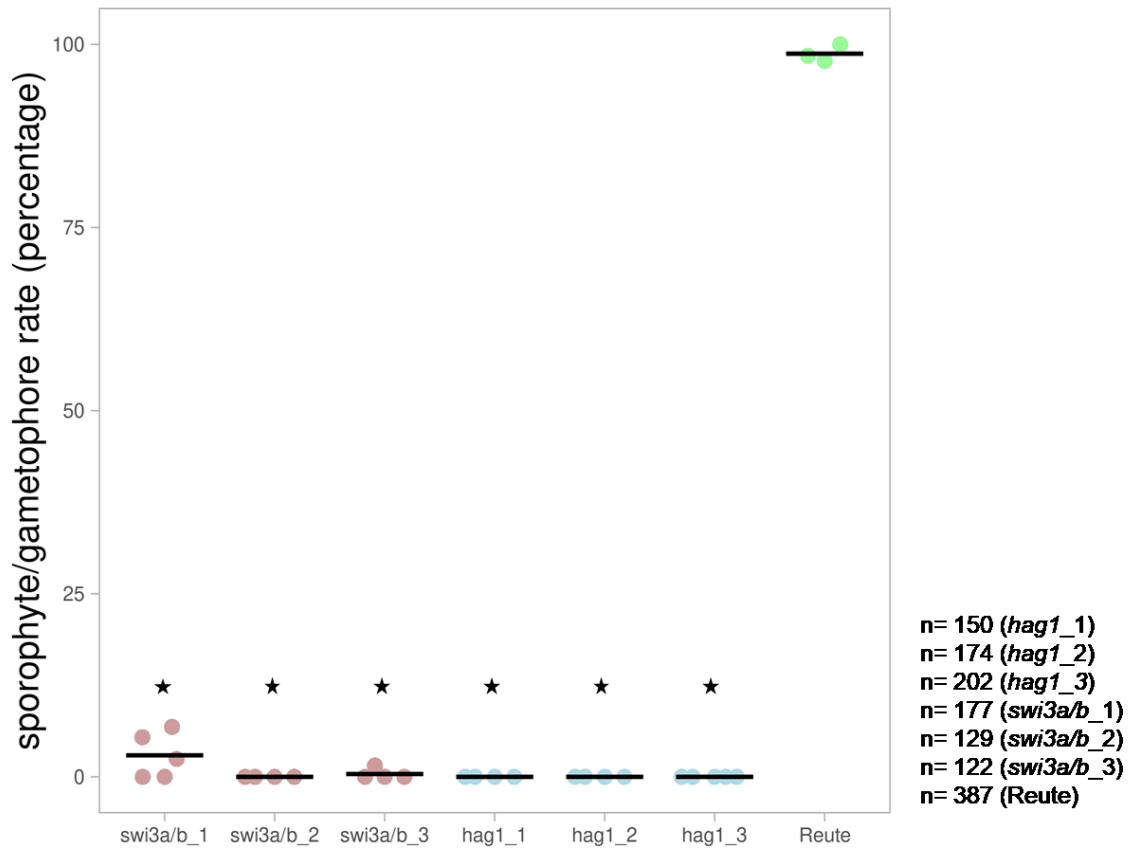


Figure 22 Sporophytes per gametophore of loss-of-function mutants.

The control (green) shows the expected selfing rate / number of sporophytes per gametophore of over 98 %. *Ppswi3a/b* (red) and *Pphag1* (blue) show a significant (asterisks) sporophyte per gametophore reduction ($p < 0.01$, Fisher's exact test applied for the conditions sporophyte yes/sporophyte no, groups mutant and control). At least three independent replicates (dots) were calculated for the mutant lines as well as the control. The total number of gametophores analysed per mutant/control is shown to the right (detailed numbers of analysed gametophores Figure 77 and Figure 78). In case of the mutants, three independent mutant lines were analysed and numbered consecutively, compared to the Reute control. The number of sporophyte per gametophore was calculated in percentage relative to the total number of gametophores. Averages of replicates are shown as horizontal lines.

Phenotypical analyses demonstrate that the control (Reute) undergoes normal development leading to an early sporophyte (ES) stage and brown mature sporophytes (B) after two and 30 days after fertilisation, respectively. Fertilization was synchronized via watering 21 days after short day transfer/gametangia induction (dpi). The mutants develop gametangia but significantly less sporophytes per gametophore (Figure 22). *Ppswi3a/b* developed cauliflower-like bundles of gametangia 30 days after fertilisation as described for a fertilisation-impaired

Results

mutant in (Meyberg *et al.*, 2020). In case of no fertilisation, *P. patens* is developing constitutively gametangia (Landberg *et al.*, 2013; Sanchez-Vera *et al.*, 2017). This observation was not as pronounced in *hag1* as in *swi3a/b*. Two days after watering (daw) the mutants show archegonia with a brown colouration of the canal cells and the egg cell, which was not detected regularly (Figure 23).

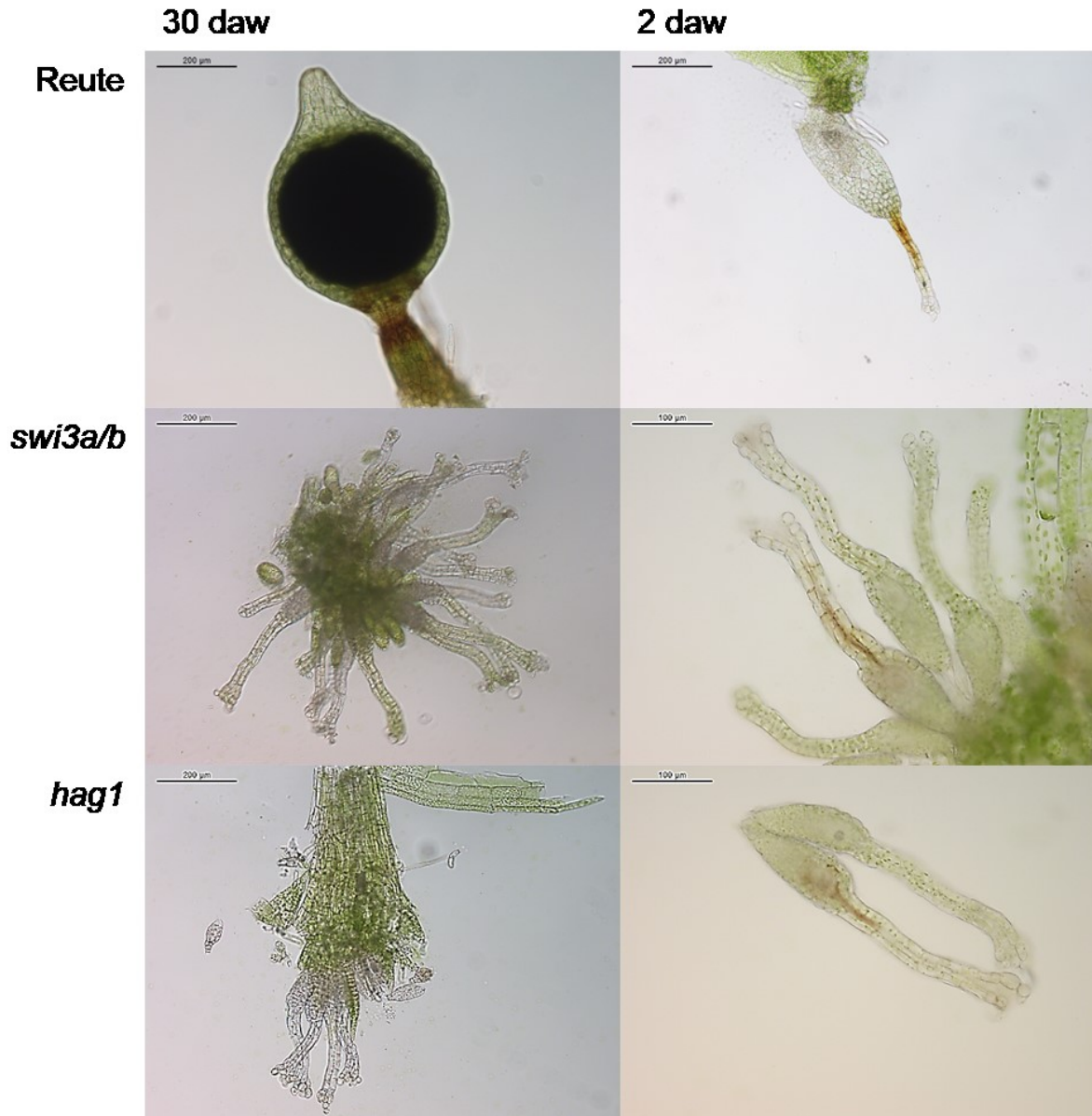


Figure 23 Phenotypic analysis of *swi3a/b* and *hag1* compared to Reute as control after watering.

Reute developed mature brown (B) sporophytes 30 days after watering. *Swi3a/b* and *hag1* developed gametangia on top of the apices. Two days after watering, Reute developed early stage (ES) sporophytes while the mutants were arrested at the archegonial stage.

Results

5.5 Test for male impairment

The lower rate of sporophytes might be due to a male or female impairment. In order to test for male impairment, the mutants were crossed with a male-fertile fluorescent strain (Perroud *et al.*, 2019). The strain Reute m-cherry stably expresses the red fluorescent protein m-cherry in all plant organs. In case of an exclusive male impairment, the mutants of the candidate genes would not be able to self-fertilise but could be fertilised in a crossing experiment. In case of crossing only the resulting sporophytic generation would show red fluorescence. Hence, successful crossing of a non-fluorescent strain with a fluorescent strain can be detected by means of glowing sporophytes, while in the paternal generation the gametophore will be non-fluorescent (Figure 24).

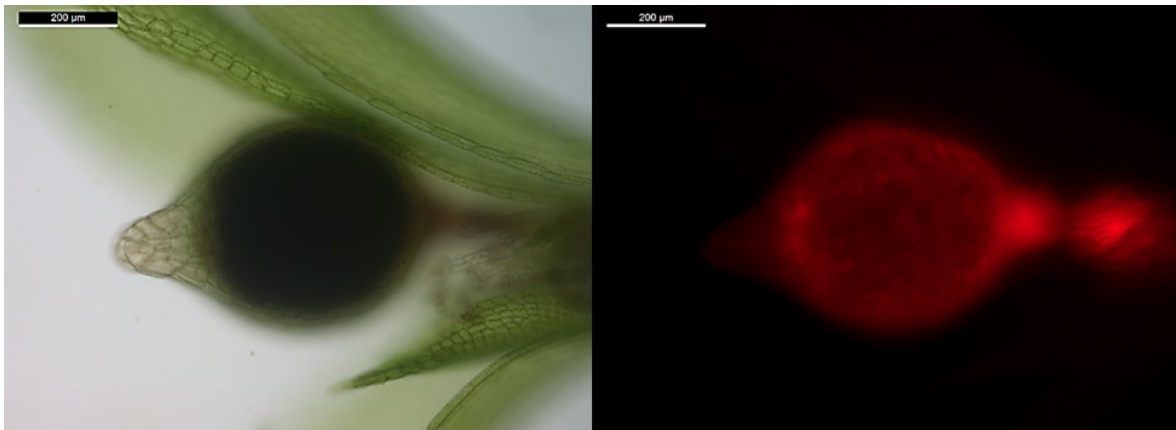


Figure 24 Micrographs of a sporophyte of *swi3a/b* in brightfield as well as fluorescent light.

Swi3a/b was crossed and subsequently fertilised by the strain Reute m-cherry. The developed sporophyte shows a red fluorescence whereas the surrounding organs (paternal generation) are non-fluorescent.

Figure 25 and Figure 26 depict the results of at least three replicates each of crossing analyses. The candidate gene mutants were crossed with Reute m-cherry to test whether the mutants are male impaired. Reute as control is highly self-fertile with a crossing ratio (heterozygous) of three percent and a sporophyte per gametophore development of 97 %. In earlier experiments, almost 2 % of outcrossing rate was described for Reute ecotypes (Perroud *et al.*, 2019). *Ppswi3a/b* shows a mean of 92 % of developed sporophytes per gametophore, whereas the crossing ratio is 98 %.

Results

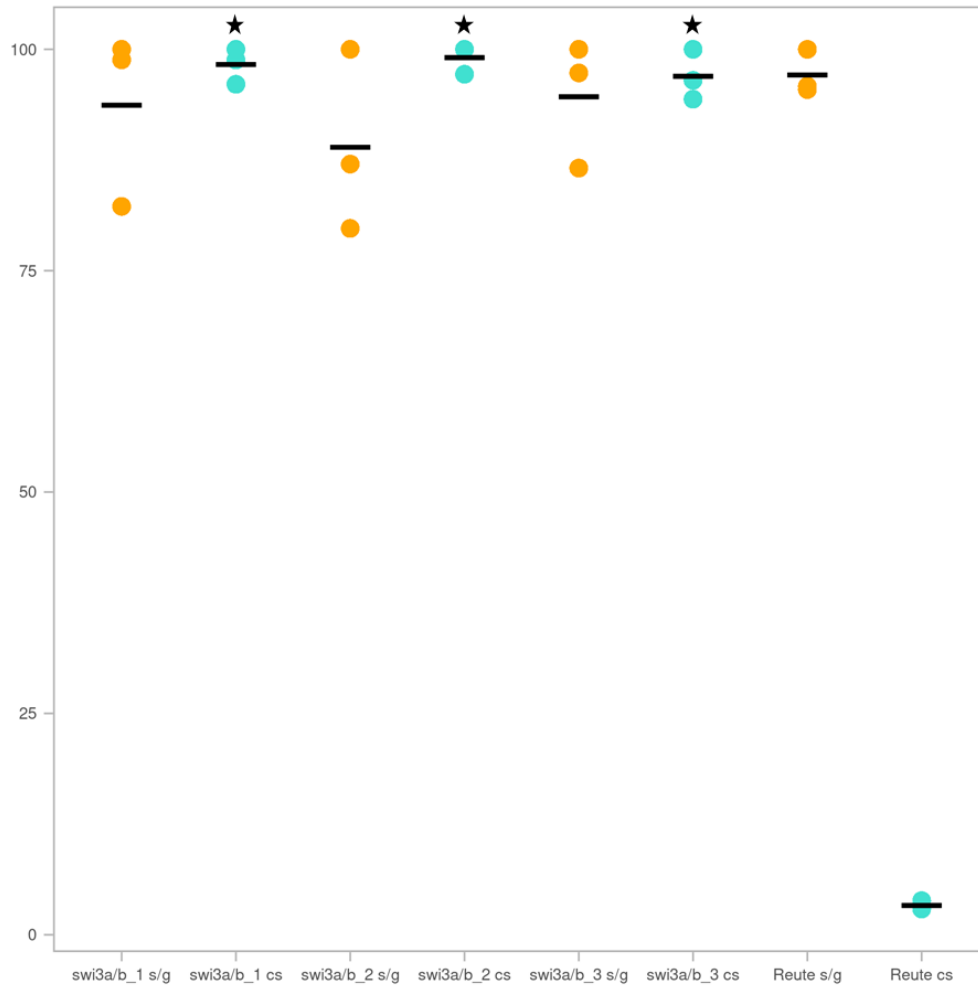


Figure 25 Crossing with a fluorescent male fertile strain to test for male impairment.

Ppswi3a/b was crossed with Re-mcherry according to (Perroud *et al.*, 2019). While most of the sporophytes in the control derive from selfing (homozygous), the significant crossing (heterozygous) ratio is almost 100 % in the mutant (asterisks), indicating a male impairment ($p < 0,01$, Fisher's exact test applied for the conditions fluorescent sporophyte/non-fluorescent sporophyte, groups mutant and control). The cross with a male fertile strain could largely restore the phenotype in swi3a/b (sporophyte ratio at least 80 %). Three independent replicates (dots) were calculated for the mutant lines as well as the control (overlay of dots possible). The total number of gametophores analysed for the mutant was 621 and 190 for the control (detailed number of analysed gametophores Figure 80). In case of the mutant, three independent mutant lines were analysed and numbered consecutively, compared to the Reute control. The turquoise dots indicate the rate of crossed sporophytes (cs) in percentage relative to the total number of sporophytes, whereas the orange dots indicate the number of sporophytes per gametophore (s/g) in percentage relative to the total number of gametophores. Averages of replicates are shown as horizontal lines.

Results

In case of *Pphag1*, 33 % of gametophores develop sporophytes, whereas 98 % of developed sporophytes are derived from crosses.

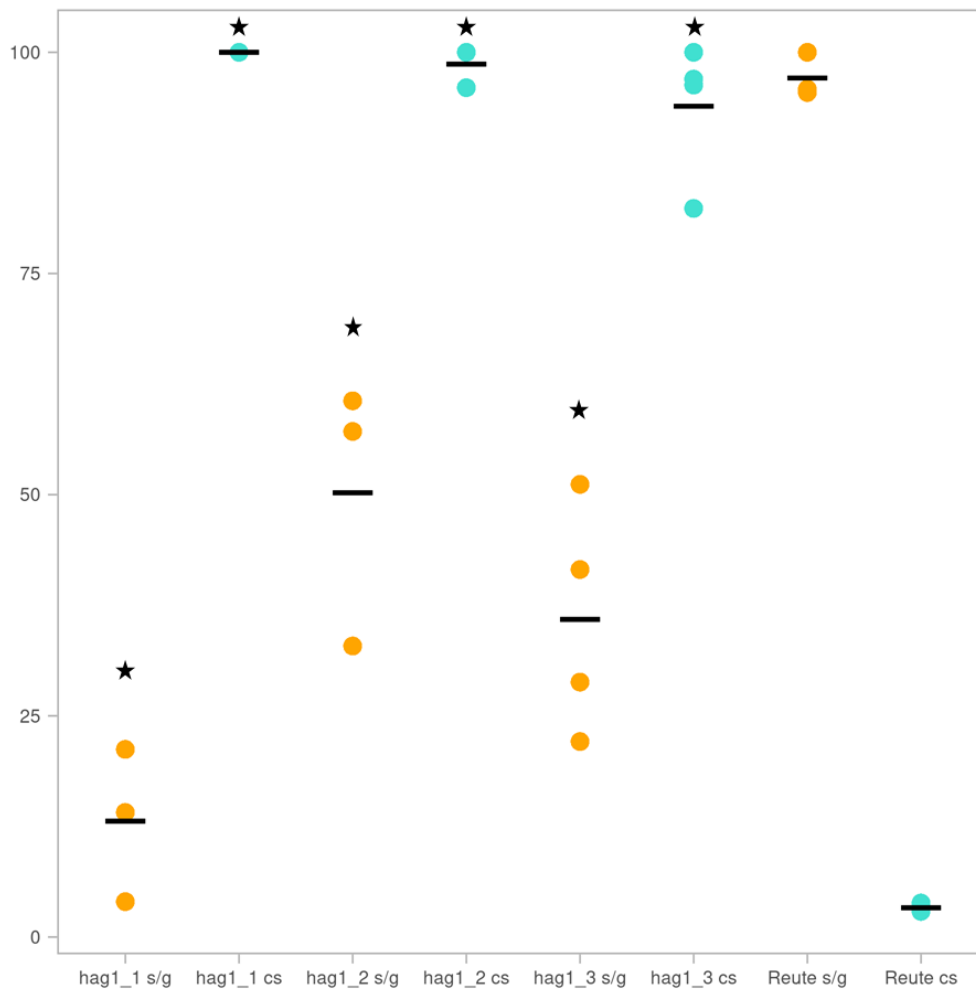


Figure 26 Crossing with a fluorescent male fertile strain to test for male impairment.

Pphag1 was crossed with Re-mcherry according to (Perroud *et al.*, 2019). While most of the sporophytes in the control derive from selfing (homozygous), the significant crossing (heterozygous) ratio is almost 100 % in the mutant (asterisks), indicating a male impairment ($p < 0,01$, Fisher's exact test applied for the conditions fluorescent sporophyte/non-fluorescent sporophyte, groups mutant and control). *Pphag1* shows a significantly lower sporophyte ratio (asterisks) compared to the control ($p < 0,01$, Fisher's exact test applied for the conditions sporophyte yes/no, groups mutant and control). At least three independent replicates (dots) were calculated for the mutant lines as well as the control (overlay of dots possible). The total number of gametophores analysed for the mutant was 770 and 190 for the control (detailed numbers of analysed gametophores Figure 79). In case of the mutant, three independent mutant lines were analysed and numbered consecutively, compared to the Reute control. The turquoise dots indicate the rate of crossed sporophytes (cs) in percentage relative to the total number of sporophytes, whereas the orange dots indicate the number of sporophytes per gametophore (s/g) in percentage relative to the total number of gametophores. Averages of replicates are shown as horizontal lines.

Results

Hence, the phenotype of *swi3a/b* can almost fully be restored by crossing with a male fertile partner. This is not the case for *hag1*, which still shows a significant reduction of sporophyte development even when crossed with a fertile male partner.

Both candidate gene mutants as well as Reute as control were crossed with a male infertile strain (Meyberg *et al.*, 2020). While the phenotype of the male infertile strain could be restored by crossing with the fertile Reute strain (sporophyte development of male infertile strain after crossing with Reute), crossing with the candidate mutants led to no sporophyte development at all (in candidate mutants as well as male infertile strain, Figure 81). Detailed crossing analysis of the male infertile strain with Reute m-cherry can be found in (Meyberg *et al.*, 2020).

5.6 Spermatozoid analysis

Spermatozoid analysis of Ppswi3a/b 21 d after induction of gametangiogenesis showed that the spermatozooids are not able to swim (Figure 27). Morphologically, they show an enclosure in caviar-like structures. Mature antheridia bundles (21 dpi) were analysed in terms of spermatozoid movement. A strong movement (swimming) of the control spermatozooids was observed shortly after release. The rate of swimming spermatozooids for the control was 100 %. With regard to Ppswi3a/b, no spermatozoid was observed that was able to swim. 28 % of spermatozooids were gently shaking after release, while the remaining 72 % of spermatozooids did not move after release.

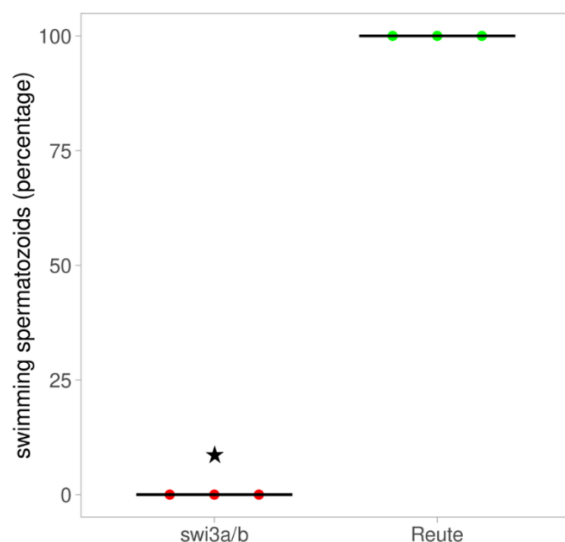


Figure 27 Swimming spermatozooids shortly after release.

Ripe antheridia (21 dpi) released their spermatozooids, which were analysed in terms of movement in three independent replicates in one mutant line and the control each. 75 (n= 29 (first replicate), 32 (second replicate), 14 (third replicate)) antheridia (Ppswi3a/b) were analysed that opened spontaneously. For the control, 35 (n= 10 (first replicate), 17 (second replicate), 8 (third replicate)) antheridia were analysed that opened spontaneously. The directly released spermatozooids of each antheridium were analysed in terms of movement. The movement was classified as swimming and no swimming. The control showed 100 % swimming spermatozooids, whereas no

swimming was observed in Ppswi3a/b (significant reduction, $p < 0,01$, Fisher's exact test applied for the conditions swimming yes/no, groups Ppswi3a/b and Reute). For Ppswi3a/b, the released spermatozooids were further classified as no movement (72 %) or gentle shaking of the whole spermatozoid cloud (28 %). Averages of replicates are shown as horizontal lines.

Results

In terms of morphology Ppswi3a/b spermatozooids develop flagella, which can move and the nuclei are able to condensate, albeit with slight structural abnormalities (Figure 28).

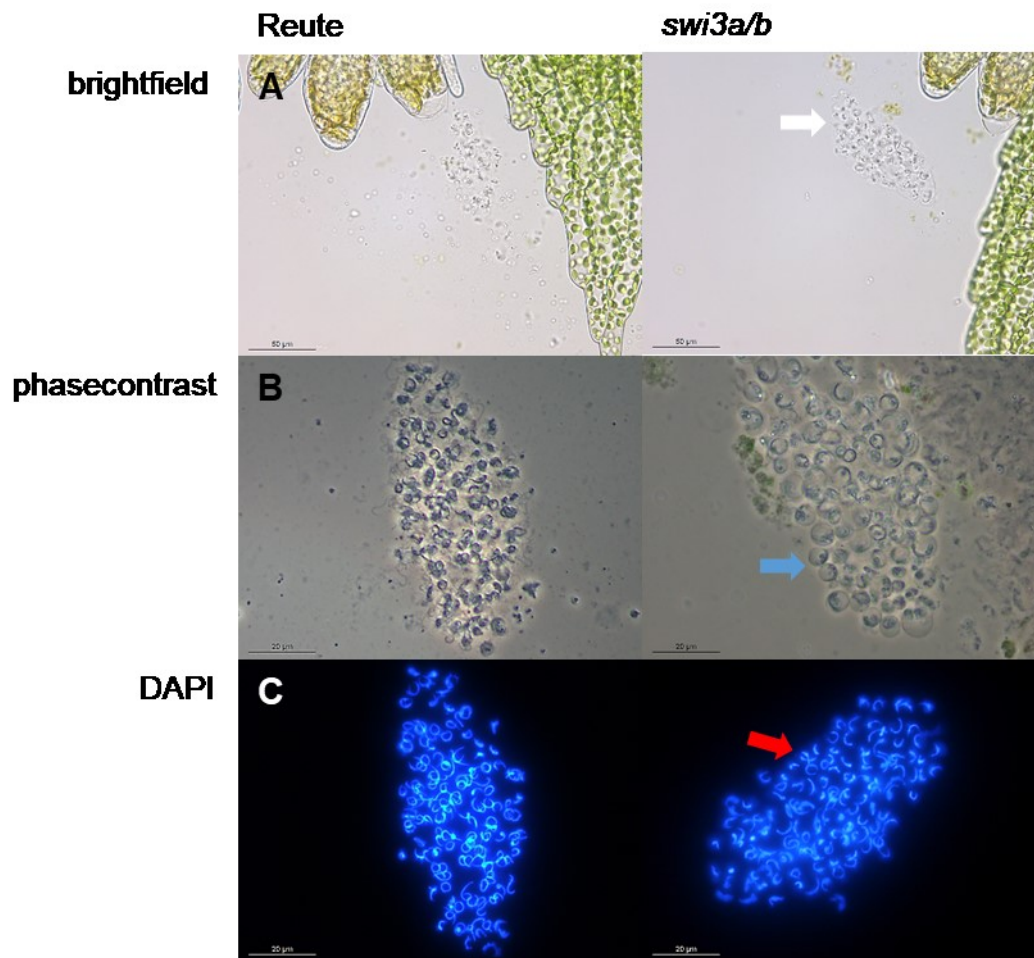


Figure 28 Spermatozoid analysis of *swi3a/b* 21 dpi; mature antheridia of Reute and *swi3a/b* release spermatozooids.

A) While spermatozooids of Reute as control start to move and swim shortly after release, *swi3a/b* spermatozooids are unable to swim and stick together without moving away from the antheridial bundle (indicated by white arrow). B) Phasecontrast images of Reute and Ppswi3a/b. The spermatozooids of *swi3a/b* show a caviar-like structure (indicated by blue arrow). C) The DAPI staining shows fully condensated nuclei in Reute, whereas Ppswi3a/b nuclei are condensated but show slight structural abnormalities (indicated by red arrow, close-up as following figure).

DAPI staining of Reute (Figure 29) demonstrated that this strain develops fully condensated nuclei, in contrast to Ppswi3a/b spermatozooids, that show slight abnormalities in the slender shape, namely a more uneven shape.

Results

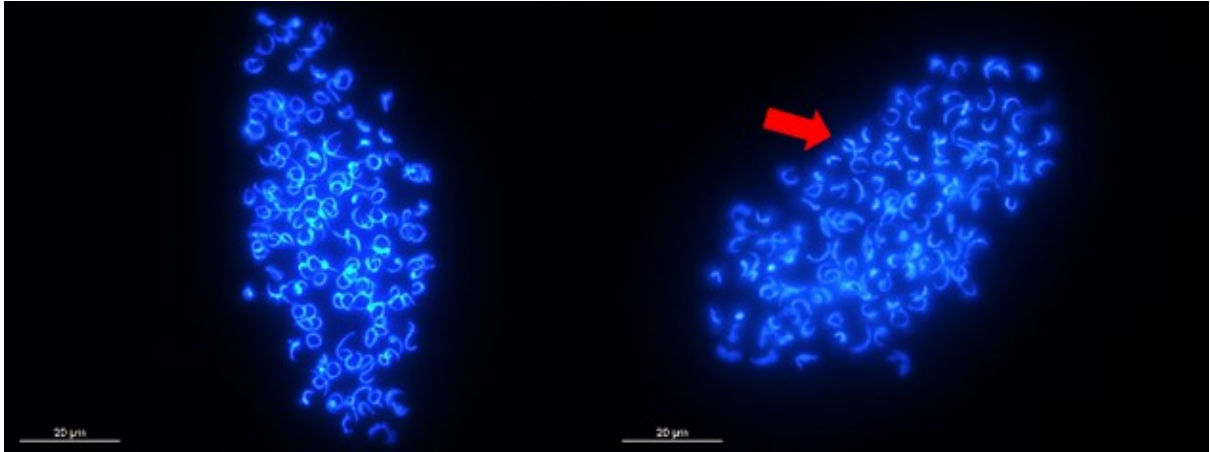


Figure 29 Close-up of Figure 28.

The red arrow indicates the slightly more uneven shape of *Ppswi3a/b* spermatozooids compared to the control on the left side. Mature wild type spermatozooids show fully condensed nuclei visualised by a DAPI staining.

As described before, *P. patens* spermatozooids are equipped with flagella to swim towards the egg cell (Meyberg *et al.*, 2020). *Ppswi3a/b* spermatozooids develop flagella which is shown in Figure 30. Qualitatively observed, no difference in flagellar assembly could be detected.

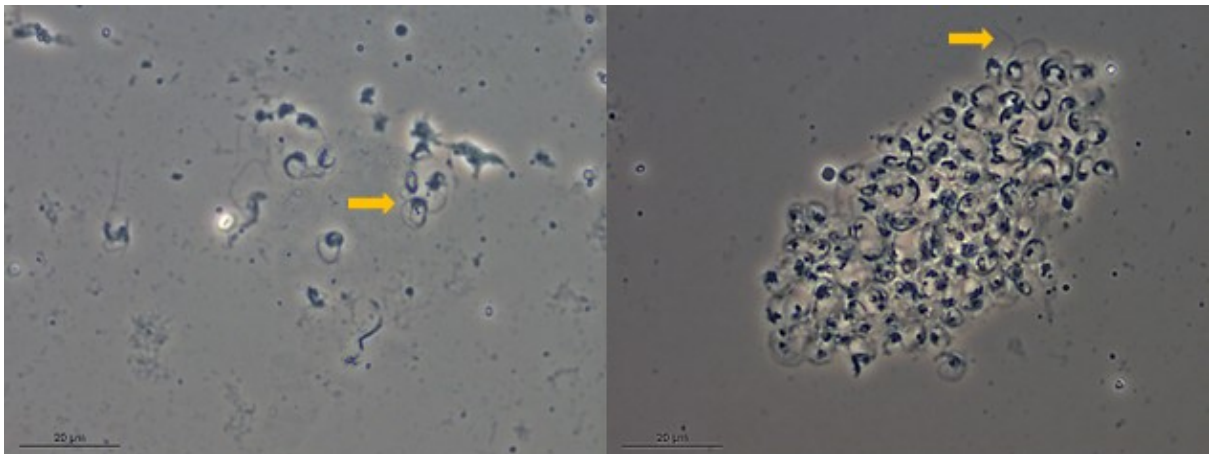


Figure 30 Phasecontrast images of the control (left panel) and *Ppswi3a/b* (right panel).

The yellow arrows indicate wild type flagella of the control (left panel) and wild type-like flagella of *Ppswi3a/b* (right panel).

The spermatozooids of both candidate gene mutants were analysed via a double staining of DAPI and NAO (Figure 31). DAPI stains the DNA, whereas NAO stains cytoplasm in *P. patens* (Sanchez-Vera *et al.*, 2017). The spermatozooids of the control are characterised by a slender shape with fully reduced cytoplasm towards the nucleus. *Hag1* antheridia do not release their spermatozooids, they remain as round-shaped structure inside the antheridia. *Swi3a/b*

Results

spermatozooids are released by the antheridia. The NAO staining reveals an impairment in cytoplasmic reduction.

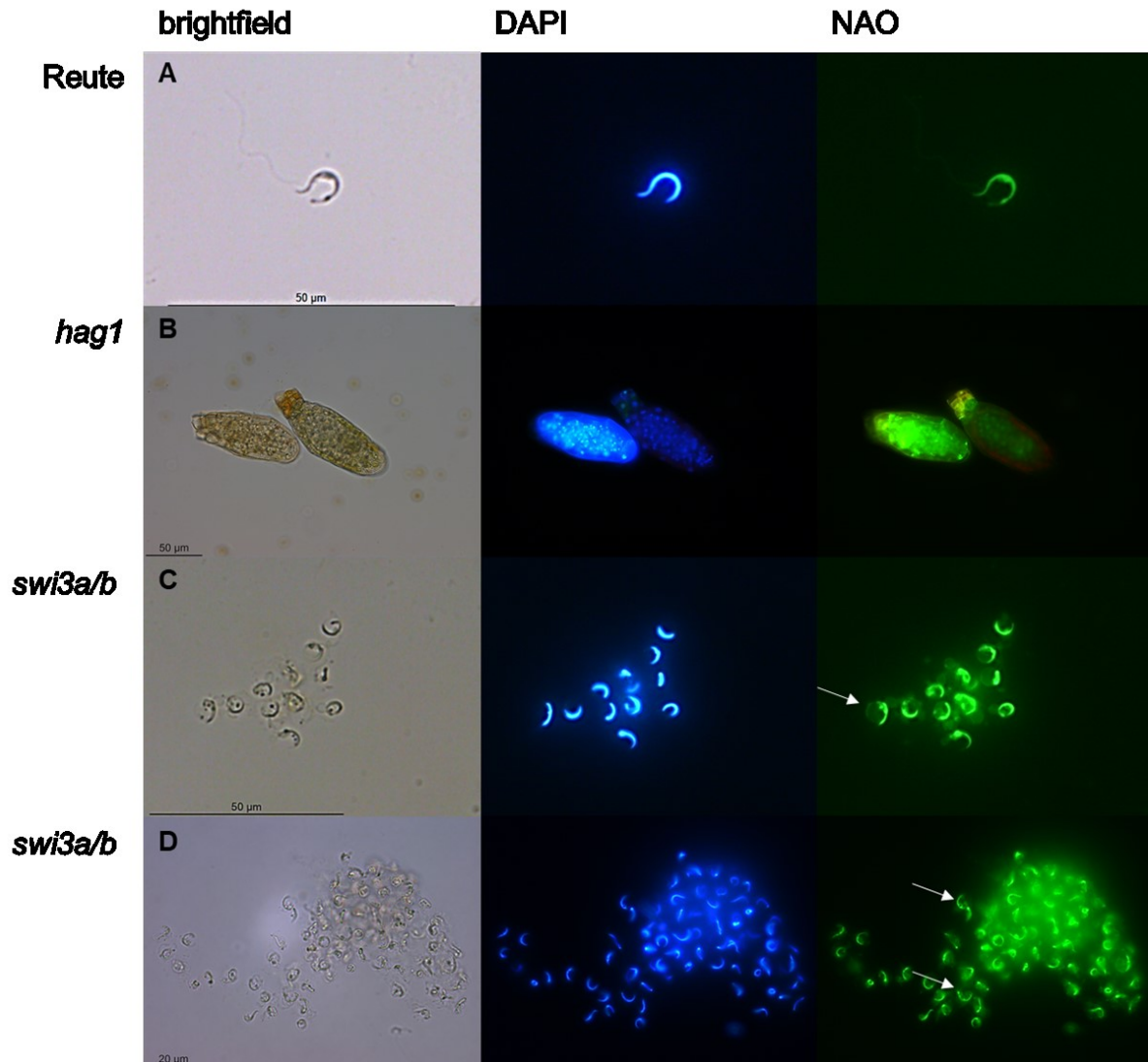


Figure 31 Spermatoid analysis of both candidate gene mutants via a double staining with DAPI and NAO. A) Reute as control shows a spermatozoid with its slender shape und fully reduced cytoplasm. B) *Hag1* does not release its spermatozooids but they remain in a round shape inside the antheridia. *Swi3a/b* antheridia release their spermatozooids, which show no (C) or incomplete cytoplasmic reduction (D). The white arrows indicate the impairment in cytoplasmic reduction (close-up as following figure).

Figure 32 shows a close-up of Figure 31 for better discrimination of potential plasma membrane and cytoplasmic content of Ppswi3a/b spermatozooids. The caviar-like structure could be seen in 37,8 % of the analysed pictures, whereas the incomplete cytoplasmic reduction (bubbly structures) was observed in 48,8 % of analysed pictures. An overlap of

Results

caviar-like structures and incomplete cytoplasmic reduction was observed in 13,5 % of analysed pictures (37 pictures in total).

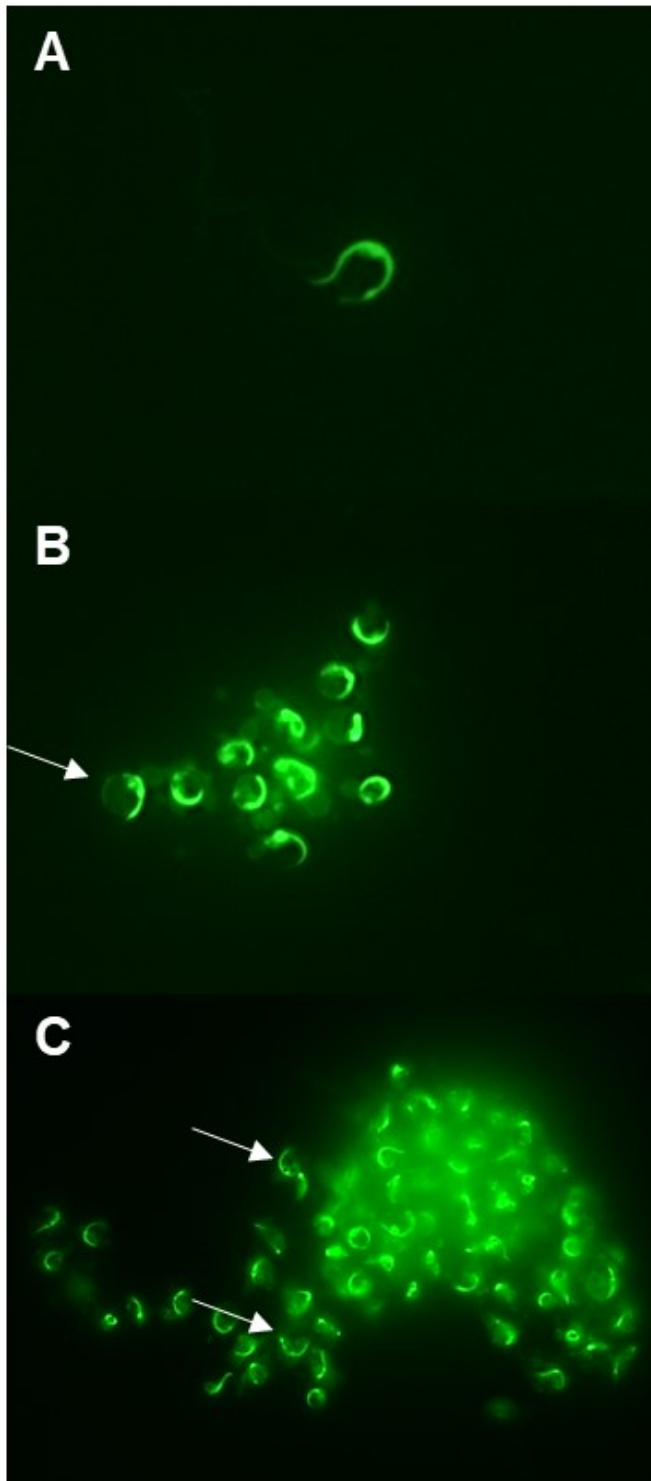


Figure 32 Close-up of Figure 31.

The white arrows indicate the impairment in cytoplasmic reduction in the *swi3a/b* spermatozoa (B, C) compared to the control (A). Mature spermatozoa of the control show a complete cytoplasmic reduction towards the nucleus (Figure 31). The caviar-structure was detected in 37,8 %, whereas the incomplete cytoplasmic reduction was detected in 48,8 % of analysed pictures (both structures in one picture 13,5 %).

Results

5.7 *Pphag1* gametangia analyses

Figure 33 displays archegonia development 21, 22 and 28 dpi. The archegonia of the control (Reute) are open 21 days after SD incubation and develop sporophytes after self-fertilisation 22 and 28 dpi even without external water supply. *Pphag1* shows archegonia which are still closed at the described timepoints. The archegonia at 22 dpi exhibit a brownish colouration and the archegonia at 28 dpi are discoloured.

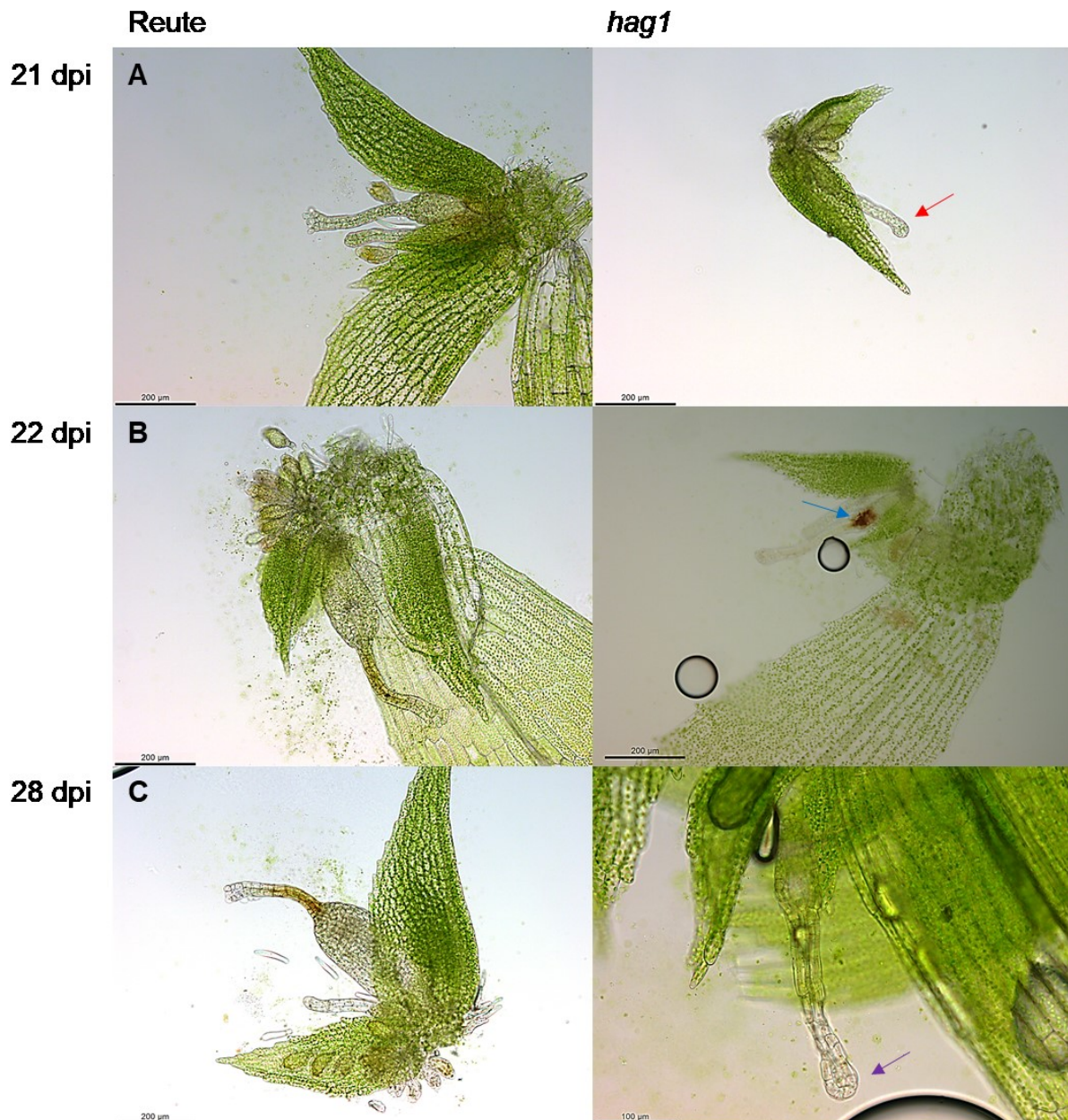


Figure 33 Archegonial development at 21, 22 and 28 dpi.

Reute archegonia are open at 21 dpi (A) and fertilised at day 22 (B) (embryo development) and 28 post induction (C), a pre-meiotic early sporophyte has developed (Hiss *et al.*, 2017a), whereas *hag1* archegonia do not open (red arrow) but develop a brownish colouration (blue arrow) and discolouration (purple arrow) 22 and 28 dpi.

Results

The number of opened archegonia was counted in the control as well as the mutant at the three described timepoints (Figure 34). In the control 88 % of archegonia are opened at 21 dpi and 100 % of archegonia are opened or already fertilised at 22 and 28 dpi. *Hag1* shows a ratio of 8 % of opened archegonia 21 dpi, 17 % at 22 dpi and 23 % at 28 dpi.

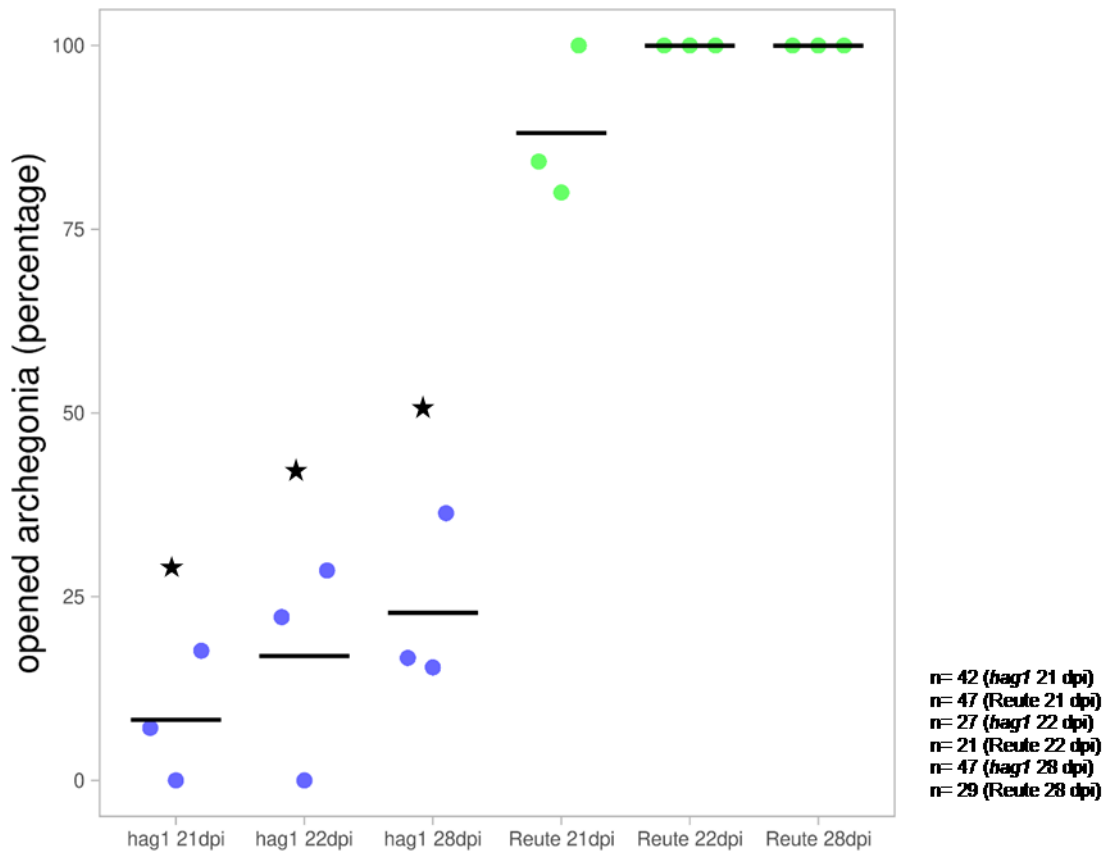


Figure 34 Analysis of opened archegonial neck cells.

The archegonia of the control as well as one mutant line of *Pphag1* were analysed in terms of archegonial opening 21, 22 and 28 dpi in three independent replicates. The blue dots visualise the three independent replicates of the mutant, whereas the green dots indicate the three independent Reute replicates. *Pphag1* shows a significantly lower (asterisks) opened archegonia rate at every analysed time point compared to the control ($p < 0,05$, Fisher's exact test was applied for the conditions opened archegonia yes/no, groups mutant and control). The total number of analysed archegonia per mutant line/control at each time point is shown to the right. The number of opened archegonia was calculated in percentage relative to the total number of counted archegonia. Averages of replicates are shown as horizontal lines.

Results

The antheridia were analysed in terms of opening as well (Figure 35). The antheridia of the control were opened in slightly lower than 100 % at 21 dpi (94 %) and opened in 100 % at 22 and 28 dpi. Antheridia of *hag1* were not opened at 21 and 22 dpi but 5 % of antheridia were opened at 28 dpi.

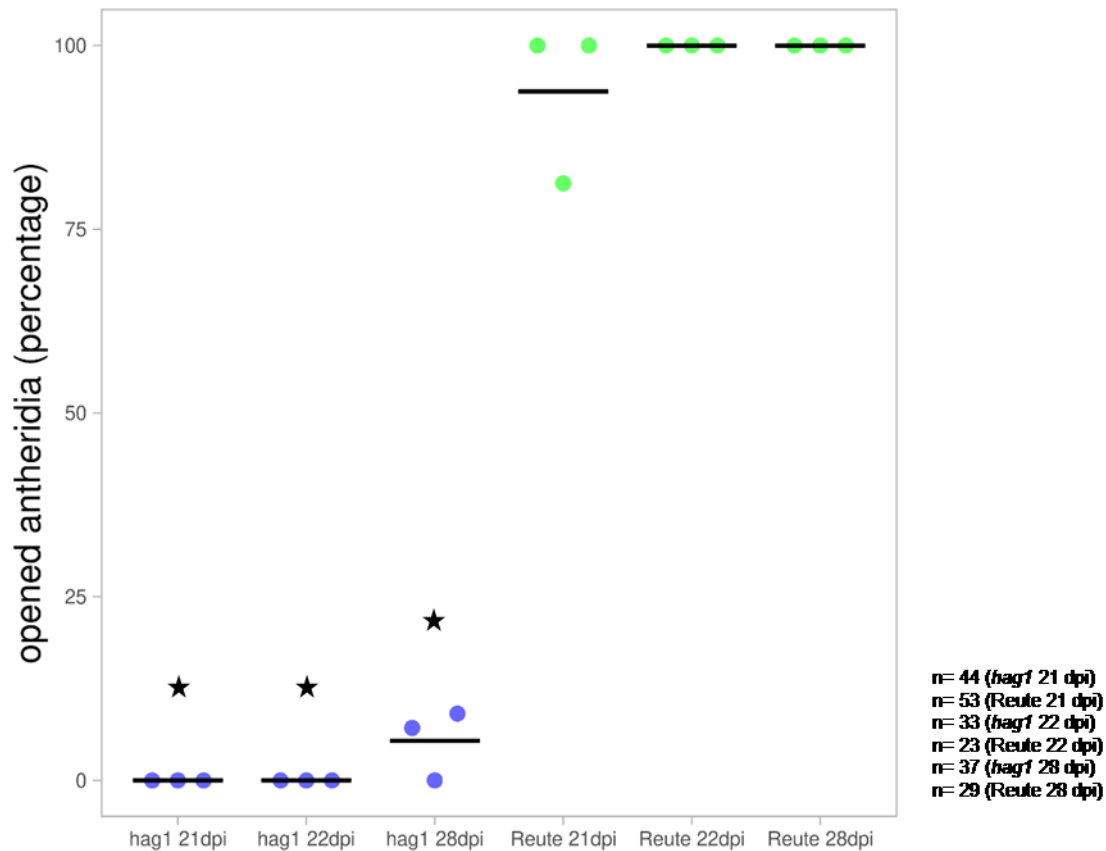


Figure 35 Analysis of opened antheridia.

The antheridia of the control as well as one mutant line of *Pphag1* were analysed 21, 22 and 28 dpi in terms of opening in three independent replicates. The blue dots visualise the three independent replicates of the mutant, whereas the green dots indicate the three independent control replicates. *Pphag1* shows a significantly lower (asterisks) opened antheridia rate at every analysed time point compared to the control ($p < 0,01$, Fisher's exact test was applied for the conditions opened antheridia yes/no, groups mutant and control). The total number of analysed antheridia per mutant line/control at each time point is shown to the right. The number of opened antheridia was calculated in percentage relative to the total number of counted antheridia. Averages of replicates are shown as horizontal lines.

Results

Figure 36 depicts antheridia bundles 28 dpi. Picture A displays closed antheridia that turned brown and discoloured. B and C show opened antheridia, which released round, shapeless spermatozooids.

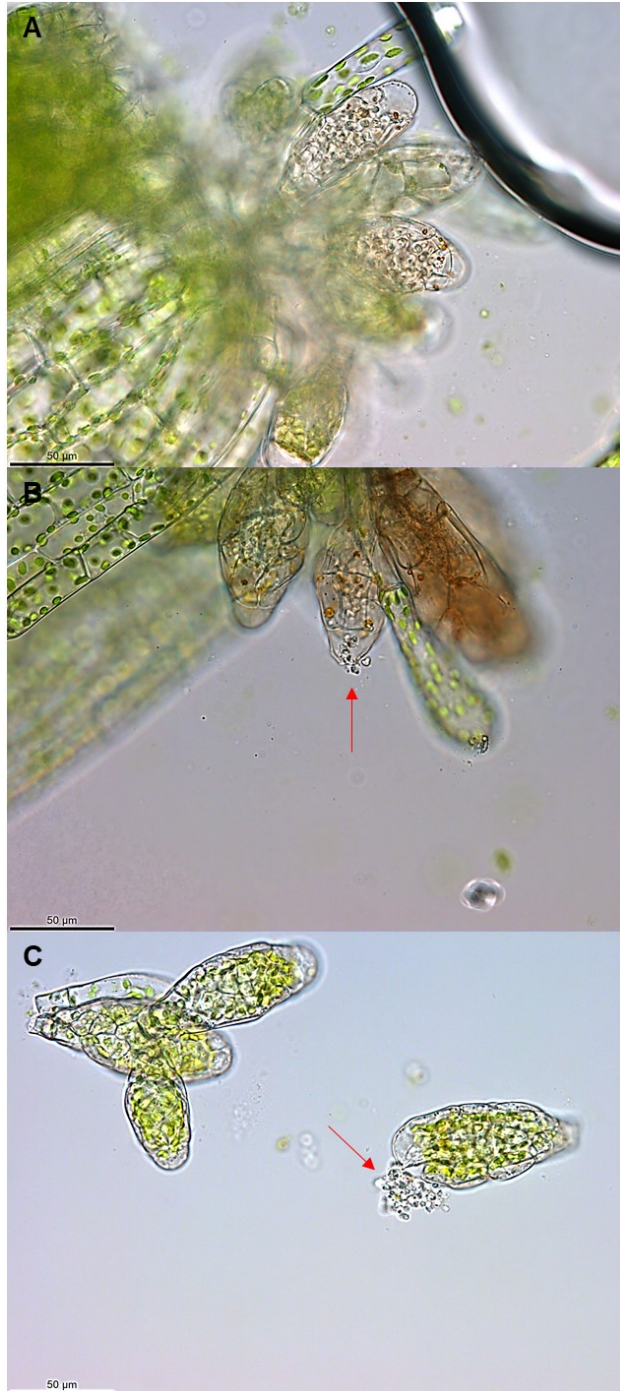


Figure 36 *Pphag1* antheridia development 28 dpi.

A) Closed antheridia turned brownish and discolored. B), C) Opened antheridia released round, bulky spermatozoid agglomerates (indicated by red arrows).

Results

5.8 Summary of morphological analyses

During the course of this study, the TAPs SWI3A/B and HAG1 were proven to be involved in the control of sexual reproduction in *P. patens*. A loss-of-function mutation (of both TAPs) led to a significant reduction of sporophyte development. A detailed morphological analysis revealed that SWI3A/B controls late steps in the male germ line, whereas HAG1 is involved in earlier steps in the male germ line. Apart from that, HAG1 regulates the female germ line as well. A summary of the phenotypical analyses is depicted as cartoon (Figure 37).

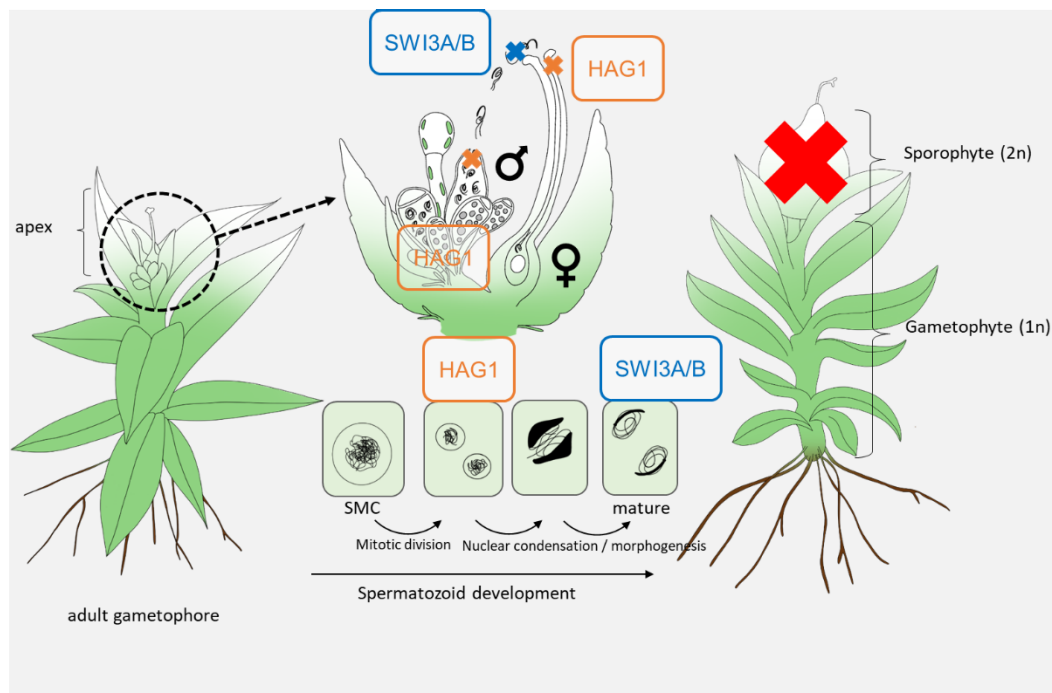


Figure 37 Summary of phenotypical analyses throughout the life cycle with emphasis on sexual reproduction.

SWI3A/B (blue) as well as HAG1 (orange) were shown to be involved in the regulation of plant sexual reproduction. Loss-of-function mutations of the respective genes resulted in a significant reduction of sporophyte development (red cross). A detailed analysis revealed that SWI3A/B controls the latest steps of spermatozoid ripening, whereas HAG1 is involved in archegonia ripening as well as earlier steps in the male germ line compared to SWI3A/B. The coloured crosses indicate the impairment throughout the respective life stages. The image was modified after (Meyberg *et al.*, 2020).

5.9 RNA-seq analysis

RNA-seq data were derived from RNA harvested from adult apices (comprising the shoot tip, the gametangia and a few phyllids). The bioinformatic analyses were carried out by Fabian Haas (University of Marburg, DE) using a previously published approach (Perroud *et al.*, 2018).

Results

Two conditions were analysed (21 days after gametangia induction (nw) as well as two days after watering (2daw/23dpi), i.e. after synchronized fertilisation). Figure 38 illustrates the number of differentially expressed genes (DEGs) in both mutants (*Ppswi3a/b* and *Pphag1*) related to the control. The Venn diagrams display the number of uniquely differentially expressed genes (DEGs) between mutant and control for the two analysed conditions. Up-regulated DEGs in the respective mutant compared to the control are coloured in green, whereas down-regulated DEGs in the respective mutant in comparison to the control are coloured in red. In case of *Pphag1* a number of 294 up-regulated unique DEGs was found two days after watering in contrast to 30 DEGs without watering. Concerning down-regulated DEGs in *Pphag1* in relation to the control, there are 158 DEGs two days after watering, whereas without watering 61 DEGs were detected.

771 up-regulated unique DEGs were observed in *Ppswi3a/b* two days after watering and 14 without watering (i.e., in the absence of fertilisation). With regard to the down-regulated DEGs in *Ppswi3a/b* related to the control, there were 89 DEGs detected two days after watering, while 22 DEGs are down-regulated compared to the control without watering.

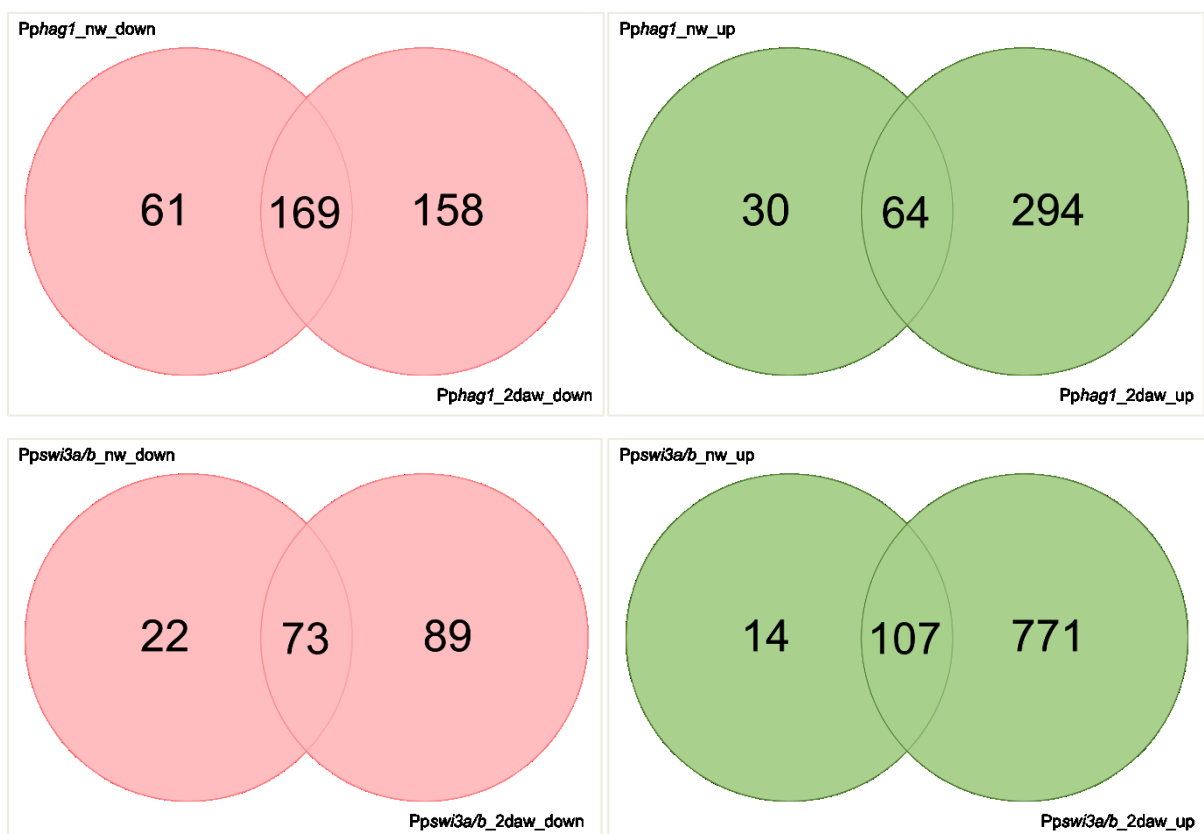


Figure 38 RNA-seq analysis.

Number of DEGs 21 days after induction of gametangiogenesis (nw/21 dpi) or two days after watering (2daw/23dpi). Each Venn diagram compares the DEGs between the two conditions nw and 2daw. The colour code indicates up-regulated DEGs (green) in the respective mutant (*Pphag1*/*Ppswi3a/b*) compared to the control as well as down-regulated DEGs (red) in the respective mutant (*Pphag1*/*Ppswi3a/b*) compared to the control.

Results

The RNA-seq analysis revealed DEGs being involved in pollen/anther, embryo and flower development, e.g. in *Pphag1* Pp3c23_9320V3.1, a homolog of *A. thaliana* *trauco*, is less abundant than in the control. In *A. thaliana* the transcript levels peak in pollen and seeds. Mutations in heterozygous plants result in an embryo-lethal phenotype, since approximately 25 % of seeds abort. It can be hypothesised that there is a paternal defect which cannot be proven, since homozygous embryos abort prior to a potential analysis. Figure 39 visualises the protein network of AtHAG1 and AtTRAUCO generated with STRING (Szklarczyk *et al.*, 2019). HAG1 is directly linked to TRAUCO via SET27 and WDR5a.

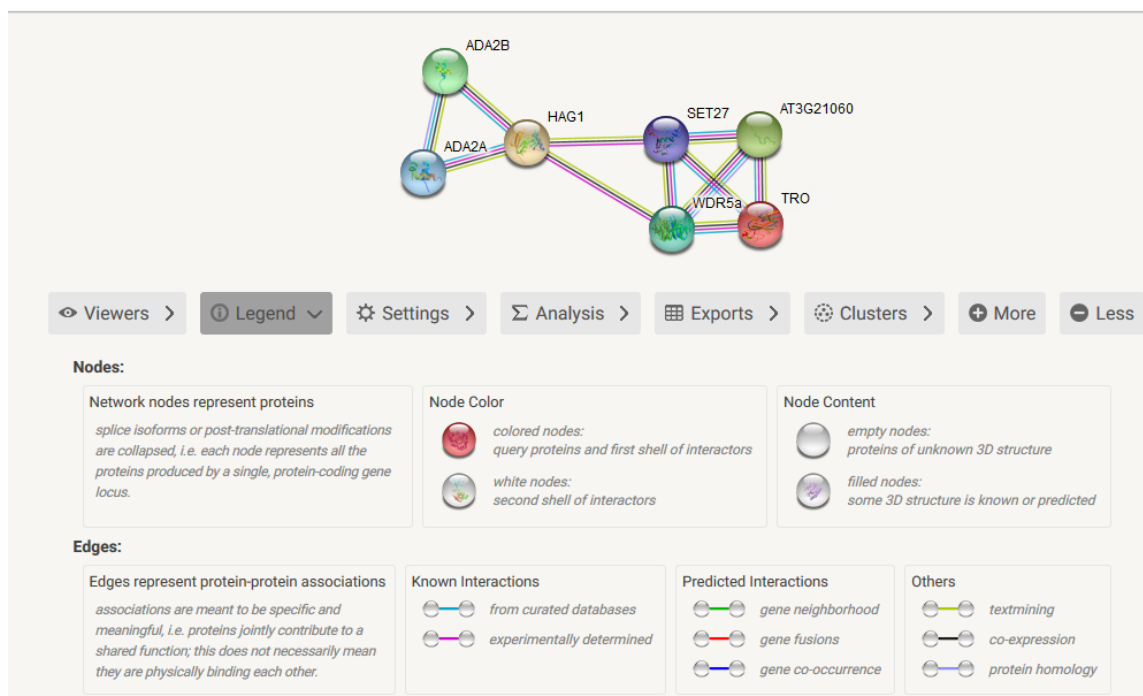


Figure 39 STRING network of AtHAG1 and AtTRAUCO.

HAG1 shows a predicted protein interaction with TRAUCO via SET27 (Histone-lysine N-methyltransferase ATX1) and WDR5a (COMPASS-like H3K4 histone methylase component WDR5a). TRAUCO and HAG1 were used as input parameters.

In *Ppswi3a/b* the *A. thaliana* *ems1* ortholog Pp3c17_21540V3.1 has a lower abundance than in the control. *Ems1* mutants in *A. thaliana* are male sterile due to a failure in microsporogenesis (Zhao *et al.*, 2002).

Apart from that, in *swi3a/b* two days after watering a PARKIN COREGULATED GENE PROTEIN (PACRG) (Pp3c9_23450V3.1) is up-regulated compared to the control and is involved in spermatid development according to its GO terms.

Results

5.10 The role of SWI3A/B and HAG1 in *M. polymorpha*

To extend the knowledge about the function of the two candidate genes not only in mosses but liverworts as well, loss-of-function mutants in *M. polymorpha* were obtained employing CRISPR/Cas9. Mutants were identified via PCR. Forward primers were designed upstream of the start codon (sgRNA1) of the respective gene. Reverse primer one was designed inside the gene, whereas primer two was designed downstream of the stop codon (sgRNA2). In case of a wild type the forward primer and reverse primer one will amplify a fragment, while in case of a deletion of the gene, there would be no amplification. As a mutant confirmation a PCR with the forward and reverse primer two would lead to a shortened fragment because of the gene deletion. In the first round of genotyping four mutants for *hag1* were identified, two male and female each. In the second round of genotyping five male mutants were identified by Zhanghai Li (University of Marburg, DE). Only one female *Mpswi3a/b* deletion mutant was identified. It has been reported that CRISPR/Cas9 efficiency can be influenced by the sgRNA design (Nomura *et al.*, 2016; Sugano *et al.*, 2018).

To analyse a gene function in terms of alternation of generations, it is a prerequisite that the organism of interest fulfils the whole life cycle in the laboratory. Therefore, fertilisation in the wild type was performed successfully, which is shown in Figure 40. The arrows indicate the yellow sporophytes developed beneath the archegoniophores, which is the diploid stage in bryophytes.



Figure 40 Female plants of *M. polymorpha* fulfilled successfully the whole life cycle.

Gametangiophores (archegoniophores) developed on top of the thallus after cold and far-red light induction. Archegoniophores were fertilised by pipetting water equipped with spermatozoids. The pink arrows indicate the developed sporophytes. The petri dish has a diameter of 92 mm.

Results

Phenotypical analyses of the developmental life stages were conducted for the mutants, which were identified in the first genotyping round. A phenotypic aberration was detectable during gametangiophore development. The first rounds of gametangiophore induction revealed a reduction in gametangiophore development in all analysed mutant lines. Only one female mutant line developed no gametangiophores at all compared to the control. Apart from that, instead of gametangiophore development, the mutants grew extra thalloid tissue (Figure 41).



Figure 41 Mutant analysis during gametangiophore induction.

The mutants show the development of extra thalloid tissue instead of gametangiophore development (indicated by a red circle).

Figure 42 shows the above described qualitatively observed reduction in gametangiophore development in female as well as male plants compared to the wild types. If developed, the gametangiophores looked like the wild type gametangiophores, but until now they were not analysed microscopically in depth. In addition, the developed mutant gametangiophores were not examined in terms of fertilisation. Phenotyping in petri dishes revealed that except for one mutant line all analysed mutant lines were able to develop gametangiophores. Gametangiophore induction in ECO2 boxes (Figure 43) contrasted these findings since either the mutant lines exhibited no gametangiophores or one mutant line developed an even higher number of gametangiophores compared to the control, indicating that the culture vessel influences gametangiophore development.

Results

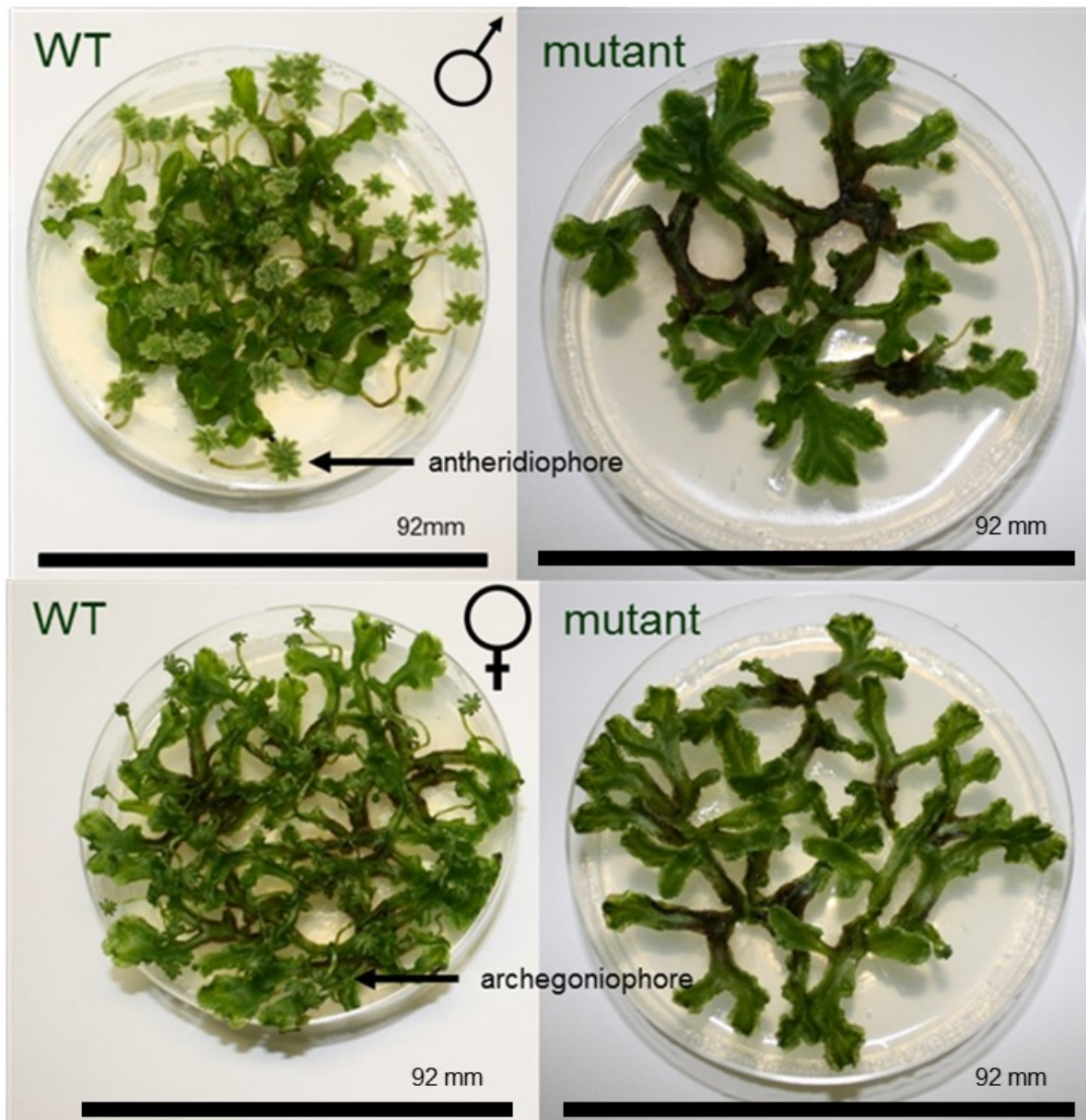


Figure 42 Phenotyping in petri dishes.

Qualitatively observed reduction in gametangiophore development in female as well as male plants compared to the female and male wild type plants (WT). First phenotyping rounds were carried out in petri dishes.

The first phenotyping rounds of gametangiophore induction were carried out in petri dishes, whereas the latter three phenotyping rounds were conducted in ECO2boxes. In each ECO2box three plants of the respective wild type or mutant were placed. Figure 43 visualises the results of the phenotyping in ECO2boxes. For every mutant line as well as the control three replicates were analysed. One replicate comprised two plants within one ECO2 box. The female mutant plants of *hag1* as well as *swi3a/b* exhibited no gametangiophore development at all, whereas one male mutant line of *hag1* developed no gametangiophore but the other one did. All wild type plants except for one replicate showed gametangiophore development.

Results

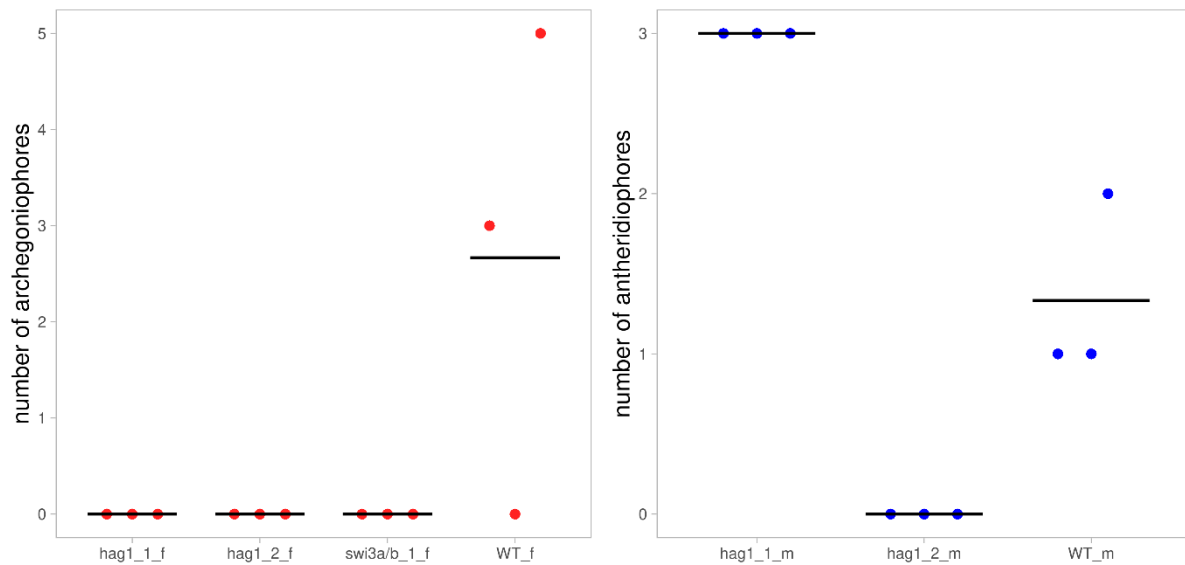


Figure 43 Phenotyping in ECO2boxes.

Female mutant plants showed no archegoniophore development (*Mpswi3a/b*, *Mphag1*) (on the left side). The female control developed no, three and five archegoniophores per ECO2box (containing two plants). *Mphag1* showed antheridiophore development in one mutant line as well as in the control (on the right side). The second male mutant line developed no antheridiophores. Three independent replicates (dots) (three ECO2 boxes with two plants respectively) were analysed with regard to the mutant lines as well as the control (WT). The mutant lines were numbered consecutively. The total number of analysed plants for every mutant line and the control thus were six in total. Blue dots represent male mutant or wild type lines, whereas red dots represent female mutant or wild type lines (m=male, f=female). Averages of replicates are shown as horizontal lines.

6 Discussion

6.1 TAPs as prime candidates for evolution under the Selected Single Copy Gene Hypothesis

The Gene Balance Hypothesis (Birchler and Veitia, 2012) describes the tendency of retainment of dosage-sensitive genes after polyploidization events. Gene duplication followed by sub- and/or neofunctionalization is a driver of the evolution of plant morphogenic complexity (Rensing, 2014). In particular, TFs show the tendency of elevated numbers with rising plant complexity, and the tendency to be retained after whole genome duplications (WGD) (Lang *et al.*, 2010; Van de Peer *et al.*, 2009; Wilhelmsson *et al.*, 2017). On top of that, during land plant evolution the regulatory hubs can be modified and co-opted (Pires *et al.*, 2013). TRs in contrast were not detected to show an equally high trend to expand with complexity as described for TFs. Nevertheless, they are essential regulators, e.g. Polycomb Group (PcG) proteins, which control the alternation of generations (Bouyer *et al.*, 2011; Mosquna *et al.*, 2009; Okano *et al.*, 2009). Beyond that, there are indeed TFs/TRs that are single copy genes in most of the analysed lineages e.g. LEAFY (Sayou *et al.*, 2014) or PRC2 protein FIE (Mosquna *et al.*, 2009). Phylogenetic analyses showed that FIE is a single copy ortholog that was already established in the last eukaryotic ancestor (LECA), but subsequently lost in some lineages. On top of that, the FIE gene tree mirrors the species tree. The high conservation of the protein with regard to its sequence throughout evolution, indicates its importance as fundamental regulator (Mosquna *et al.*, 2009). The scenario of paralog loss is called the Selected Single Copy Gene Hypothesis, which is an amendment to the Gene Balance Hypothesis, that postulates that duplicate retention of certain genes is counter-selected because of dosage effects (Duarte *et al.*, 2010; Edger and Pires, 2009). Dosage-sensitive genes which are usually retained as single copy genes are for example nuclear-encoded organellar (plastid and mitochondrial) genes (Duarte *et al.*, 2010). This can be explained due to their function in organellar signalling networks or macromolecular complexes, that need a certain stoichiometric balance with their interaction partners, which are encoded in the organellar genome. Nevertheless, it is still not researched in depth how the gene balance between nuclear and organellar genomes is regulated. The Selected Single Copy Hypothesis states that all gene duplications, no matter whether WGD or small scale duplication of these genes, will lead to a dosage imbalance of their gene products, which results in a counter-selection of the duplicate retention (Duarte *et al.*, 2010). One example for dosage-sensitive genes are ribosomal complexes: the duplicates of nuclear ribosomal protein genes are usually retained after a WGD (Aury *et al.*, 2006; Papp *et al.*, 2003). In contrast, following the Selected Single Copy Hypothesis the duplication of the

Discussion

ribosomal protein rps13, which encodes a subunit of the mitochondrial organellar ribosomal complex, might lead to dosage effects, so that there is counter-selection against retention of the duplicate (Edger and Pires, 2009).

Shared single copy genes in a broad range of lineages can serve as phylogenetic markers (Edger and Pires, 2009), i.e. gene and species tree coincide. This is also true for the phylogenetic analyses of GCN5/HAG1 throughout all eukaryotic kingdoms (Figure 14), which reflects the taxonomic relationships of the species involved. The presence of orthologs of the gene within all the described lineages shows that the gene was already present in the LECA. Taking into account the previously described examples for TAPs involved in plant development, TFs/TRs were considered prime candidates to have evolved under the Selected Single Copy Gene Hypothesis, “indeed single copy ortholog clusters contain twice as many TF/TR genes as the genome-wide fraction, which is a significant enrichment ($p = 4E-07$, Chi square test)” (Stefan Rensing, personal communication). Therefore, studying TAPs in terms of their biological function is an important tool to explore plant development and evolution of morphology and complexity. TAPs are in general classified in families by their domains, so that it is essential that these classification rules are up-to-date and easily accessible, which is given by TAPScan V2 (Wilhelmsson *et al.*, 2017).

In this study two genes were selected for further experiments. Apart from that, the candidate gene approach resulted in genes known to regulate important steps in the plant life cycle (Higo *et al.*, 2018; Mosquna *et al.*, 2009; Motomura *et al.*, 2016; Okano *et al.*, 2009), confirming the outlined workflow for candidate gene selection. Next to the identification of well-known regulatory hubs, the outlined approach could be confirmed in the laboratory with the help of loss-of-function mutants of two candidate genes, which showed an impairment in sexual reproduction. Hence, the bioinformatic approach was confirmed not only *in silico* but also *in vivo*, since the loss-of-function mutants confirmed the hypothesis that the two selected genes could be important regulators of plant development, which is indeed the case.

6.2 *In vivo* confirmation of *in silico* analyses

During the process of candidate gene selection, literature and expression profile analyses were carried out. The expression levels of the candidate genes in *P. patens* (Figure 16) and *M. polymorpha* (Figure 17) indicate an essential role in the alternation of generations, since the genes in *P. patens* have a high expression throughout sporophyte development and an expression peak in antheridia bundles, whereas *M. polymorpha* shows a high expression of swi3a/b throughout the stages of sexual reproduction, while the highest expression takes place in antheridiophores. Mphag1 is expressed highest in antheridiophores, but the second highest expression can be detected in archegoniophores. The expression levels of hag1 and swi3a/b

Discussion

in *M. polymorpha* indicate a role throughout sexual reproduction especially in the gametangiophore development most prominent in antheridiophore development (male germ line).

With the help of loss-of-function mutants, the function of the respective genes throughout plant development resulting in certain phenotypes should be determined. The loss-of-function mutants in the moss led to a significantly lower rate of sporophyte development which confirms the hypothesis of an involvement in the alternation of generations as deduced from the expression levels of the protein.

The so far performed phenotypical analyses in *M. polymorpha* loss-of-function mutants of both candidate genes partly confirmed the *in silico* analyses (see section 5.10). The highest expression of *hag1* was observed in antheridiophores and second highest in archegoniophores. One of two male mutants developed no antheridiophores at all, whereas the second was not affected in gametangiophore emergence (Figure 43). Therefore, only one of the two mutant strains confirmed the hypothesis that antheridiophore development could be affected. Thus, a higher number of male mutants are needed for clear phenotype confirmation. On top of that, RT-PCR should be performed in future work to analyse whether the confirmed mutants are still expressing *hag1* or a truncated version. The female mutants showed a qualitatively observed reduction of gametangiophore emergence cultured in petri dishes and no antheridiophore development at all cultured in ECO2 boxes (Figure 43), which confirmed the *in silico* predictions. Apart from the emergence of gametangiophores the fertility of developed gametangiophores was not yet analysed, which needs to be done in future. For *swi3a/b* only one mutant could be detected until now, which is female. The highest expression of the gene predicts the most striking phenotype associated with antheridiophore development and/or fecundity. This could not be tested because of a lack of male mutants. The second highest expression predicts a phenotype related to archegoniophore development and/or fertility. This prediction was proven by the lack of archegoniophore development in *Mpswi3a/b*. Nevertheless, it is necessary that more mutants will be identified to reliably analyse the phenotype.

6.3 Gene function during gametogenesis

Mutations of *hag1* and *swi3a/b* lead to a loss of sporophyte development, which means that the haplodiplontic life cycle could not be fully performed, since the transition from haploid to diploid generation does not occur, but they can be grown as gametophyte. The loss of the diploid generation results from a missing self-fertilisation, due to an impairment of the male germ line (Figure 22, Figure 37).

Discussion

To study the gene functions of *swi3a/b* and *hag1* in more depth, overexpression of the respective genes is to be carried out in future work. For this, the use of an inducible promoter system is a helpful tool. The system of choice is “an estrogen receptor mediated inducible gene expression system” (Kubo *et al.*, 2013). *P. patens* overexpressor mutants were already confirmed via genotyping by Zhanghai Li (four mutants for Pphag1) and should be analysed in future work with special emphasis on the transition from haploid to diploid generation. A well-known proof of principle for an overexpression approach is BELL1, a master regulation in terms of phase transition and embryonic development. The overexpression of BELL1 triggers embryo formation followed by sporophyte development without fertilisation (Horst *et al.*, 2016).

6.3.1 Ppswi3a/b

Ppswi3a/b exclusively shows a defect in spermatozoid development (see section 5.6.), which was shown on one hand by means of analysing the male germ line (Figure 27, Figure 28) and on the other hand by crossing analyses with a fertile fluorescent strain (Figure 25) (Perroud *et al.*, 2019) as well as with a male impaired strain (see section 5.5) (Meyberg *et al.*, 2020). The crossing analyses with a fertile fluorescent strain restored the phenotype almost to the level of Reute sporophyte development (Figure 25). The development of sporophytes after fertilisation by a fluorescent strain indicated an impairment in the male germ line but no defect in the female germ line. The crossing analysis with a male infertile strain underlined the hypothesis of an impairment in the male germ line. No sporophyte development of *Ppswi3a/b* as well as the male infertile strain was observed, which confirmed the hypothesis that *Ppswi3a/b* is unable to self-fertilise and cross-fertilise. The morphological analyses extended the knowledge of the defect. *Ppswi3a/b* developed 30 days after watering (synchronized fertilisation) bundles of archegonia, looking like a cauliflower at the apices (Figure 23), which usually occurs if archegonia are not fertilised (Meyberg *et al.*, 2020).

Antheridia are ripening similar to the control and release their spermatozooids (Landberg *et al.*, 2013). In contrast to the spermatozooids of the control, which start swimming shortly after release, the mutant spermatozooids are unable to swim and stay motionless or show a gentle shaking (Figure 27). The impairment in movement can be underlined by morphological analyses, namely DAPI and NAO double staining (Figure 31). The DAPI staining shows that the nuclei are able to condensate but qualitatively observed they show slight structure abnormalities compared to the control, which is characterised by the slender shape. The most striking phenotype is the incomplete cytoplasmic reduction. In mature spermatozooids of the control the cytoplasm is fully reduced towards the nucleus, which results in the narrow shape. The NAO staining exposes that the mutant spermatozooids show a bubbly caviar-like structure, which could be the plasma membrane, and incomplete cytoplasmic reduction, which indicates

Discussion

that the latest steps of spermatozoid ripening are impaired. Recently, it was shown that autophagy is essential for gamete differentiation in terms of cytoplasmic reduction (Sanchez-Vera *et al.*, 2017). To test whether autophagy is disturbed in *Ppswi3a/b*, the crossing of the mutant with an autophagy reporter line was already set up (Sanchez-Vera *et al.*, 2017) to introduce the marker into the mutant background. The marker is the autophagy-related gene *PpATG8e* tagged with GFP which shows a high expression during gametangia development. The reporter line in the mutant background should be used in future work to uncover potential differences in the autophagy flux during gametogenesis compared to the control.

Already performed analyses of archegonia related a brownish colouration of the canal cells to an entered spermatozoid (Tanahashi *et al.*, 2005). Since self-fertilisation does not happen in *Ppswi3a/b*, the brownish colouration would suggest a premature senescence that was shown to be related to autophagy (Sanchez-Vera *et al.*, 2017).

Since the spermatozooids are unable to swim, they cannot actively reach the archegonia and the egg cell within, so that fertilisation is not possible. Moss *in vitro* fertilisation as potential method to improve the understanding of the mutant impairment, would reveal whether spermatozooids cannot fertilise due to immobility or due to disturbance of the actual fertilisation process as well. Until now, this method is not established.

6.3.2 Pphag1

The loss of sporophyte development in *Pphag1* was determined to be gametangia-dependent as well. However, the crossing analyses described above revealed that *Pphag1* is not only self-infertile due to an impairment in the male germ line, but crossing with a fertile fluorescent strain cannot fully restore the phenotype, suggesting an impairment of the female structures as well (Figure 26). The sporophyte development of the respective mutant line, that was also used for further morphological analyses and RNA-seq, showed in the crossing with a fertile fluorescent strain a sporophyte/gametophore ratio of 13 % (Figure 26). That is in the range of the number of ripe/opened archegonia 21 dpi (8,3 % of opened archegonia (Figure 33)). The impairment in female gametogenesis might be due to a disturbance in ripening, leading to an arrest in archegonial development. Concluding that this finding might cause the reduction of sporophyte development in the crossing analysis, since, a prerequisite for a subsequent sporophyte development is the synchronized ripening of antheridia as well as archegonia which is characterised by their opening (Figure 34, Figure 35). Gametangia phenotyping results 28 dpi (Figure 33) show that disruptance in female gametangia ripening is not only delayed but arrested. The archegonia turn brown and colourless, indicating that they die prior to maturity (premature senescence). Premature senescence was shown to be autophagy-related (Sanchez-Vera *et al.*, 2017). On top of that there were archegonia which were open 21

Discussion

dpi, but showed a brownish colouration around the egg cell, indicating a disturbance in development as well. As described above, a brownish colouration of the canal cells was reported to be due to an entrance of a spermatozoid (Tanahashi *et al.*, 2005). These findings cannot be related, since on one hand gametangia ripening was analysed without watering as fertilisation synchronizer, on the other hand the localisation of the brownish colouration was not the same (control=canal cells, *Pphag1*=around the egg cell).

Analyses of *Pphag1* antheridia revealed that the ripening is disturbed just as in *Ppswi3a/b* but at a much earlier time point compared to *Ppswi3a/b*. Antheridia do not open 21 dpi self-reliantly at all. The double staining unveiled that spermatozoids arrest round and bulky and do not condensate towards a slender shape. It was observed that antheridia turned completely brownish although they didn't open 28 dpi, which indicates that antheridia are disturbed in apoptosis prior to maturity. This would fit a *gcn5* phenotype, described in mice, which shows embryonic lethality and an increased apoptosis (Phan *et al.*, 2005).

6.4 Involvement of *Pphag1* and *Ppswi3a/b* in transcriptional regulation in the alternation of generations

RNA-seq analyses were performed for *P. patens* loss-of-function mutants as well as the control. Therefore, apices were harvested for RNA isolation 21 dpi as well as two days after watering (23 dpi). 21 dpi, the control develops ripe gametangia. Fertilization can be synchronized via watering. Two days after watering, fertilisation occurs, so that the gene regulatory program of embryogenesis should have been initiated, if fertilisation can be obtained (Hiss *et al.*, 2017a). RNA-seq analyses were carried out to identify the involvement of the candidate genes in the transcriptional network during gametangia ripening as well as the initiation of embryogenesis.

In both mutants (*Ppswi3a/b* and *Pphag1*) self-fertilisation is not possible, leading to an arrest in the life cycle during gametangia development. The transition to the sporophytic generation needs major changes in gene expression that are often linked to epigenetic reprogramming, carbohydrate, fatty acid or lipid metabolism (O'Donoghue *et al.*, 2013).

6.4.1 *Ppswi3a/b*

771 uniquely up-regulated DEGs in *Ppswi3a/b* compared to the control upon fertilisation were detected and only 14 unique DEGs in *Ppswi3a/b* compared to the control without watering (i.e., in the absence of fertilisation). With regard to the down-regulated DEGs, there were 89 DEGs detected 2daw and 22 DEGs without watering. The comparatively low number of DEGs without

Discussion

watering prior to fertilisation indicates that there are no major differences in the transcriptional network during gametangia ripening (Figure 38). These results reflect the mutant phenotype in the sense that *Ppswi3a/b* develops wild type-like archegonia, whereas the male germ line is impaired in the latest steps of spermatozoid ripening, so that no major differences in the transcriptional network in gametangia ripening are observable. In contrast, 2daw (synchronized fertilisation) the up-regulated DEGs reflect the maintained gametangial program in *Ppswi3a/b*, since no fertilisation occurs in contrast to the control that started 2daw already the zygotic transcriptional program. On top of that, 107 DEGs (up-regulated) and 73 DEGs (down-regulated) all together were detected between the two conditions tested, showing that the mutation leads to changes in the transcriptional program separately from fertilisation. Keeping the described phenotype in mind, in case of a repetition of the experiment, it would be advantageous to perform RNA-seq with antheridia exclusively at different points of ripening, to detect the exact developmental stage where the transcriptional program varies between mutant and control.

In *swi3a/b* the *Atems1* ortholog Pp3c17_21540V3.1 shows lower abundance than in the control. *Ems1* mutants in *A. thaliana* are male sterile due to a failure in microsporogenesis (Zhao *et al.*, 2002). The lower expression of *ems1* in *Ppswi3a/b* and *Atems1* exhibit a comparable phenotype, indicating a similar role of *ems1* orthologs in bryophytes and angiosperms.

In addition, in *swi3a/b* 2daw PACRG is up-regulated compared to the control that is involved in spermatid development according to its GO terms. In mice the ortholog is highly expressed in testes during spermatogenesis and the gene product is abundant in mature sperm. Mutations lead to malformations in the sperm head as well as defects in the flagellum (Lorenzetti *et al.*, 2004). It was shown that PACRG is involved in flagellar tubulin binding (Dawe *et al.*, 2005).

For proper pollen development (*A. thaliana* male gametophyte) callose deposition is required (Dong *et al.*, 2005). In *swi3a/b* 2daw UDP-glucose pyrophosphorylase (Pp3c23_20400V3.1) is differentially expressed (up-regulated compared to the control) indicating that, in contrast to the control where sporophytic post-fertilisation gene expression takes place, the male germ line specific gene expression still is turned on in the mutant. The UDP-glucose pyrophosphorylase is an important enzyme in the UDP-glucose metabolism, which is a precursor for the synthesis of callose. Mutations in this gene lead to vegetative and reproductive growth defects, namely male sterility, since callose deposition around the microspores does not occur (Park *et al.*, 2010). It was shown, that genes which are involved in pollen-wall development are conserved between mosses and angiosperms, e.g. sporopollenin (Morant *et al.*, 2007). These findings indicate that they could play a role in the spore-wall development in mosses (Zaveska Drabkova and Honys, 2017). Spores and seeds

Discussion

are considered analogous but spores and pollen might be homologous (Rensing, 2016; Wallace *et al.*, 2011). These findings seem contra-intuitive, since fertilisation leads to sporophyte development and subsequently results in meiosis and spore development. Nevertheless, potentially the generative gene regulation is disturbed in the mutant, so that this gene is already expressed in gametangia development. On top of that, the *P. patens* gene shows highest expression in archegonia according to (Ortiz-Ramirez *et al.*, 2016), indicating a gene function throughout gametangia development just as in *A. thaliana* (pollen development). Taken together, these findings underline the hypothesis that fertilisation does not occur in *swi3a/b*. Rather, the transcriptional regulation is probably arrested during male germ line expression through altered epigenetics/chromatin modifications, as evidenced by the high number of DEGs that can be assigned to spermatozoid/pollen development. These findings can be confirmed by the mutant phenotype, which shows a defect in the latest steps of ripening, namely the cytoplasm reduction (see section 5.6).

6.4.2 Pphag1

For *Pphag1* compared to the control a number of 294 uniquely up-regulated DEGs was detected 2daw compared to 30 up-regulated DEGs 21 dpi without watering, which indicates that the gametangial transcriptional program at 2daw still is activated, since no fertilisation occurs. With regard to the down-regulated DEGs in *Pphag1* as compared to the control, 158 DEGs were detected 2daw, whereas without watering there are only 61 DEGs. 2daw, fertilisation has already occurred in the control, so that the embryogenic transcriptional regulation is activated. 169 DEGs were detected which are commonly down-regulated in *Pphag1* within both conditions. That could indicate on one hand differences in the generative life cycle with regard to the transcriptional regulation, independent of the transition from the haploid to the diploid generation. On the other hand, fertilisation could have occurred without external water supply, so that the zygotic program was activated already in the control without watering. Analysing apices still seems feasible since the development of male and female gametangia are impaired in the mutant (see sections 5.6 and 5.7). Taken together, the higher number of up- and down-regulated DEGs 2daw is comprehensible. The total number of DEGs confirms the phenotype in that regard, that the impairment takes place at the transition from gametophyte with mature gametangia to the sporophytic generation, namely from haploid to diploid generation, since self-fertilisation is almost impossible.

The data show a lower abundance of Pp3c23_9320V3.1, a homolog of Attrauco, in *hag1*. In *A. thaliana* the transcript levels show a peak in pollen and seeds. Mutations in heterozygous plants show an embryo-lethal phenotype, since approximately 25 % of seeds abort. It can be hypothesised that there is a paternal defect which cannot be proven, since homozygous

Discussion

embryos abort. The lower abundance in *hag1* mutants in *P. patens* indicates a similar role as in *A. thaliana trauco* mutants (Aquea *et al.*, 2010). The hypothesised paternal role in *A. thaliana* can be substantiated by the *P. patens* phenotype (see section 5.6). The STRING network (Figure 39) showed that AtHAG1 and AtTRAUCO are directly linked via WDR5a and SET27. SET27, which is also known as Trithorax protein1 (ATX1), was shown to be involved in the regulation of flowering time and floral organ identity (Alvarez-Venegas *et al.*, 2003). Recently, it was postulated that a homolog of ATX1 in rice (*O. sativa*) Trithorax1 (OsTrx1) is also involved in flowering timing, since the suppression of the gene expression leads to a delay of flowering time in rice (Choi *et al.*, 2014). Phenotypical analyses of *Pphag1* archegonia and antheridia showed a delay in ripening and an arrest at unripe stages indicating a conserved role of the described gene network.

6.5 Role in all bryophyte lineages

Loss-of-function mutants were initially planned for all three bryophyte lineages (mosses, liverworts and hornworts), the closest extant relatives to streptophyte algae (Kenrick, 2017; Nishiyama *et al.*, 2018; Wickett *et al.*, 2014), to determine whether the candidate genes fulfil the same function in all three bryophyte lineages and to elucidate the phylogenetic relationship between hornworts, liverworts and mosses, which is still unsolved (Morris *et al.*, 2018).

The water to land transition was a key event in evolution which has presented the conquerors with several challenges, e.g. sexual reproduction without constant water supply. Analysing genes related to sexual reproduction is therefore needed to understand land plant evolution and to learn about how sexual reproduction has adapted.

Loss-of-function mutants were carried out for the respective candidate genes *hag1* and *swi3a/b* in *P. patens* as well as *M. polymorpha*. *P. patens* is a moss, whereas *M. polymorpha* belongs to the liverworts. *P. patens* as well as *M. polymorpha* are well-established model organisms with published genomes (Bowman *et al.*, 2017; Rensing *et al.*, 2008). The hornworts are the third bryophyte lineage, which are represented by *A. agrestis*. The hornwort is currently under progress in terms of establishment as a model organism (Szovenyi *et al.*, 2015), e.g. the establishment of a transfection protocol. On top of that the “chloroplast and mitochondrial genomes, their transcriptomes and editomes” as well as the nuclear genome assembly were recently published (Gerke *et al.*, 2020) (earlier in the study, the genome was kindly provided by Peter Szovenyi (University of Zurich)). Nevertheless, mutants for the *A. agrestis* orthologs could not be obtained throughout the study. On top of that, the life cycle is not totally established in the laboratory with regard to fulfilling the life cycle, although sporophyte development was seen occasionally (Nora Stingl (University of Marburg)). Phenotyping approaches need a fully established life cycle to deduce reliable statements. Therefore, the *A.*

Discussion

agrestis orthologs should be applied in a complementation analysis to study whether the hornwort gene can restore the *P. patens* phenotype, a scenario that would indicate the same or a comparable gene function.

Comparing *M. polymorpha* and *P. patens* via loss-of-function of the respective orthologs is especially interesting since the liverwort is dioicous whereas the moss is monoicous. This could result in a different phenotype, since the expression of the sexes takes place during gametogenesis in *P. patens*, whereas there are already expression differences in the vegetative thallus in *M. polymorpha*.

As described above, *P. patens* loss-of-function mutants show a significant reduction of sporophyte development, due to an impairment in the male germ line (*swi3a/b*) and male and female germ line (*hag1*). The initial phenotyping results in *M. polymorpha* (5.10) show a reduction as well as no gametangia development in both candidate genes. These results would indicate that the gene acts differently throughout reproductive life stages, namely earlier in the germ line, which in turn leads to the reduced emergence of gametangiophores due to separate sexes in *M. polymorpha*. Apart from that, the reduced appearance of female and male gametangiophores reflects the phenotype in *P. patens*, since *Pphag1* also shows an impairment in female and male gametogenesis, even though the defects appear later throughout gametogenesis.

Ppswi3a/b shows no impairment in the female germ line, whereas *Mpswi3a/b* shows no development of archegoniophores, indicating a shift of the phenotype in *M. polymorpha* compared to *P. patens*. Nevertheless, the expression profile (Figure 17) would predict a role in the male germ line as well. Only one female mutant and no male mutant were detected for *Mpswi3a/b* until now, therefore more mutants are needed to deduce a reliable conclusion comparing moss and liverwort. Taken together, the phenotypical analyses were not consistent until now. The phenotyping in ECO2boxes would indicate that *hag1* female plants could not develop gametangiophores at all, whereas in former phenotyping rounds in petri dishes, a reduced number of gametangiophores was detected. Hence, more mutants and phenotyping rounds as well as optimised culture conditions need to be obtained. The reliable gametangiophore development in wild type plants (male and female) is therefore indispensable. Apart from that, the fecundity of developed gametangiophores was not yet tested, so that it may very well be that the *M. polymorpha* mutants show an impairment in gametangiophore ripening as well, which would reflect the *P. patens* phenotype. All in all, due to the existing problems in gametangiophore induction in *M. polymorpha*, a proven conclusion cannot be drawn between the candidate gene function *in vivo* in moss compared to liverwort.

Discussion

6.6 Evolutionary conservation

The present work as well as recently described single copy TAPs (LEAFY, PRC2 FIE) underline the hypothesis that single copy TAPs are fundamental key players in terms of plant developmental regulation (Mosquna *et al.*, 2009; Sayou *et al.*, 2014). FIE was most probably already present in the LECA and subsequently lost in certain lineages. On top of that, the protein taxonomy reflects roughly the taxonomy of the involved species (Mosquna *et al.*, 2009). The candidate genes analysed in the present study are already described as highly conserved genes in the literature, which could be extended by the present work in bryophytes.

A. thaliana encodes four SWI3 genes, which diversified their function. Apart from that, BNB, given as an example, has two copies with redundant function in *A. thaliana*. Throughout land plant evolution, BNB evolved from a broader ancestral role to a sex-specific control (Yamaoka *et al.*, 2018). This finding underlines the hypothesis that gene regulatory networks adapted during the course of evolution.

Mutations of swi3a and b show an arrest of embryo development at the globular stage. Apart from this, mutations in swi3b lead to an impairment in macrosporogenesis (arrested ovules) as well as microsporogenesis. With regard to the failure of macrosporogenesis, the analyses showed that megaspores were developed normally in the ovules after meiosis but roughly 50 % of ovules did not form embryo sacs, so that the megaspores failed to divide subsequently, but instead stayed in an arrested central position. A mutation in swi3b caused partial male sterility. A deeper analysis of microsporogenesis revealed that pollen mother cells underwent meiosis, followed by callose-separated tetrads and divided vacuolated microspores. Nevertheless, about 50 % of the microspores did not undergo further divisions but arrested, vacuolated and were exposed to senescence, which leads to lipid deposition and organellar degradation or collapsing. Mutations in swi3c lead to a semi-dwarfed body plan, root elongation is inhibited, the leaves show a curly morphogenesis, stamen development is abnormal, and the fecundity is impaired. Mutations in swi3d cause dwarfism, abnormal flower organ development and entire male and female sterility.

Together, the four SWI3 proteins have non-complementary roles, which are involved in the plant's embryogenesis as well as vegetative and reproductive life stages (Sarnowski *et al.*, 2005). The so far published phylogenetic analysis of swi3 postulated that there are four swi3 genes in *P. patens* and that one Ppswi3 gene is an ortholog to Atswi3a, whereas the remaining three Ppswi3 genes are paralogs of Atswi3c (Gao *et al.*, 2012). Our data (Figure 15), however, shows another phylogenetic relationship, since the *P. patens* gene formerly designated as Ppswi3a is orthologous to the Atswi3a/b clade. This finding is likewise true for the other bryophyte lineages. Also, the three remaining Ppswi3 paralogs are orthologous to swi3d not c. The other two bryophyte lineages are characterised again by single copy gene orthologs to

Discussion

Atswi3d. The phylogenetic analysis (Figure 15) indicates that the Atswi3c/d clade is the ancestral clade, as only in that clade algae orthologs to Atswi3c/d can be found. This underlines that the gene already evolved prior to the water to land transition and was duplicated and diversified throughout land plant evolution.

Ppswi3a/b shows a strong male impairment reminiscent of the *A. thaliana swi3b* phenotype, which indicates a gene function in mosses exclusively in the male germ line. The *P. patens* phenotype only partially reflects the *A. thaliana* phenotype, which verifies the phylogenetic analyses in that sense that the *A. thaliana swi3* complex has four gene copies with non-redundant functions. In contrast, the bryophytes have only two distinct copies, excluding *P. patens* (three paralogs for swi3d), which is nevertheless separated in two clades. The gene, and as a corollary its function, apparently evolved throughout land plant evolution from a male germ line specific to a germ line unspecific and embryogenic role. It can be hypothesised that this is due to the before mentioned need of focusing the adaption of sexual reproduction without constant water supply during the water to land transition in terms of gene regulatory networks. In summary, *Ppswi3a/b*, *Atswi3a* and *Atswi3b* function in bryophytes as well as angiosperms as fundamental developmental regulators with regard to sexual reproduction.

The SWI/SNF chromatin-remodelling complex plays a fundamental role in the germ line even in mammals. It was shown that the ATPase subunit, brahma-related gene 1 (BRG1), is involved in spermatogenesis, i.e. repair events during meiosis. Mutations lead to a meiotic arrest, accumulation of unrepaired DNA and a failure of complete synapsis (Kim *et al.*, 2012; Menon *et al.*, 2019).

Hag1, which is also termed *gcn5*, was shown to be involved in a high number of developmental processes (Aquea *et al.*, 2017) and therefore exhibits pleiotropy like dwarfism, alterations in rosette leaves and a loss of apical dominance in *A. thaliana* mutants (Benhamed *et al.*, 2006; Bertrand *et al.*, 2003; Cohen *et al.*, 2009; Long *et al.*, 2006; Vlachonasios *et al.*, 2003). Mutations in *hag1* could also be associated with the alternation of generations, namely abnormal flower development and reduced fertility (Bertrand *et al.*, 2003; Vlachonasios *et al.*, 2003). The loss of fecundity was attributed to a reduced stamen height in early-arising flowers and an increase in stamen number in early-arising flowers in contrast to late flowers that are able to produce small siliques with a few seeds (Vlachonasios *et al.*, 2003). Recently, it was shown that *clavata1* and *hag1* double mutants exhibit elongated gynoecia with reduced valves and enlarged stigma and style, deducing the hypothesis that CLAVATA signalling and GCN5 regulation play a synergistic role throughout gynoecium development. Single mutants of *gcn5* showed a milder phenotype in terms of gynoecium development (Poulios and Vlachonasios, 2018). Mutations in *P. patens* also result in reduced fertility, primarily due to defects in the male germ line. However, mutants show a weaker impairment in the female germ line as well (see section 5.6 and 5.7).

Discussion

In *Mus musculus*, it was shown that a disruption of *hag1* leads to a lethal embryonic defect (Xu *et al.*, 2000; Yamauchi *et al.*, 2000), indicating functional conservation in different eukaryotic lineages. These *gcn5* null embryos are not viable, since they exhibit an increased apoptosis, which is restricted primarily to the mesodermal and ectodermal lineages (Phan *et al.*, 2005; Xu *et al.*, 2000). On top of that, it was shown recently that GCN5 is also essential for proper spermiogenesis. The loss of *gcn5* leads to altered chromatin structures and an increase in histone retention in sperm, which then results in malformation (abnormal nuclear development) and disrupted male fertility in mice (Luense *et al.*, 2019). Interestingly, the mouse phenotype even resembles the moss phenotype.

Our data (Figure 14) shows that *gcn5* is single copy in most of the analysed lineages and can be found throughout eukaryotic evolution, which indicates that the gene was already present in the LECA as described for PRC2 FIE (Mosquna *et al.*, 2009). The only exception is the vertebrates, where paralog retention has occurred, since the invertebrates still share single copy status. The phylogenetic analysis roughly reflects the taxonomy of involved species. The only outlier is one *Azolla filiculoides* ortholog, which clusters with the fungi. This can be explained either by a horizontal gene transfer or a contamination during sequencing. Paralog retention is observable in some lineages. The analysed gymnosperm species lost the gene. The performed phylogenetic analysis underlines the hypothesis that single copy genes can serve as phylogenetic markers, as they are highly conserved throughout the whole eukaryotic evolution (Edger and Pires, 2009), which even results in a comparable phenotype between moss and mouse (pairwise divergence time for *P. patens* and *M. musculus* estimated time: 1,496 million years ago (Hedges *et al.*, 2015) <http://timetree.org/>). Therefore, mosses can even serve as a model for studying evolutionary conserved genes that lead to human diseases, as proposed in (Meyberg *et al.*, 2020).

The analyses of *swi3a/b* (TF or TR) and *hag1* (TR) reinforce the hypothesis that TRs as well as TFs play fundamental roles in plant development (Wilhelmsson *et al.*, 2017).

6.7 Outlook

In summary, the study showed that both candidate genes are important developmental regulators in the alternation of generations. *Ppswi3a/b* was shown to be involved in the latest stages of spermatozoid ripening (Figure 37). A loss-of-function mutation of the respective gene led to an impairment in cytoplasmic reduction throughout spermatozoid ripening. The unripe spermatozoids were not able to swim towards the egg cell. Therefore, no fertilisation can be obtained in *Ppswi3a/b*. To determine whether these defects are related to autophagy, the ATG8e-reporter line should be implemented in the mutant to see whether autophagy-related GFP signal differences can be detected between *Ppswi3a/b* and *Reute* as a control. Exceeding

Discussion

RNA-seq of *P. patens* apices, namely single cell sequencing should be applied to break down the transcriptional regulation to single cells, e.g. the spermatozoids.

Pphag1 revealed an arrest in spermatozoid ripening in earlier steps compared to *Ppswi3a/b* (Figure 37). The spermatozoids in the loss-of-function mutants arrested throughout ripening in a round and bulky shape. Antheridia did not open to release spermatozoids and are impaired in ripening. Therefore, Transmission electron microscopy should give rise to morphologically detailed analyses of anatomic defects in the *P. patens* mutants (*Ppswi3a/b* and *Pphag1*) compared to the control.

To analyse epigenetic marks and subsequently the controlled genes targeted by SWI3A/B and HAG1, ChIP-seq should be performed to reveal which genes are effected by the loss-of-function mutations (*Ppswi3a/b* and *Pphag1*). It was shown that SWI/SNF acts antagonistically against the epigenetic silencing of PcG proteins (Wilson *et al.*, 2010). PRC2 regulates gene silencing through trimethylation of histone 3 lysine 27 (H3K27) (Wilson *et al.*, 2010), which is therefore an interesting histone mark to address in a ChIP-seq approach. H3K14 acetylation was shown to be in *A. thaliana* exclusively related to HAG1. Therefore, H3K14 would be the epigenetic mark of choice to analyse in a ChIP-seq approach. *A. thaliana hag1* mutants were analysed compared to the wild type in a comparable approach (ChIP array). HAG1 targets a wide range of promoters (related to 40 % of the tested promoters) of genes diverse in function based on gene ontology (Benhamed *et al.*, 2006; Benhamed *et al.*, 2008).

Mphag1 was described to have a potential defect in gametangiophore development, namely a reduction in gametangiophore emergence. Therefore, *Mphag1* is planned to cross, if fertilisation can be obtained, with a *M. polymorpha* citrine tagged BNB strain. Bonobo gene expression is induced throughout initiation of gametangiophore development, which therefore can be used as marker for initiation of gametangiophore development (Yamaoka *et al.*, 2018). In terms of transcriptomics the *M. polymorpha* mutants should be analysed via RNA-seq to explore the altered gene expression driven by the mutations already happening in the vegetative thallus compared to *P. patens*.

In the present work the TAPs were mainly described in their function via loss-of-function mutations. The loss-of-function mutants should be complemented with orthologs of certain plant lineages such as charophytes, bryophytes and angiosperms to explore in more depth the evolutionary conservation (Figure 86). Apart from that, complementation with the respective *P. patens* genes should be applied as proof that the described phenotypes are truly due to the loss-of-function mutations of the respective genes themselves. Complemented mutants with the moss genes were already identified by Zhanghai Li (two mutants *Ppswi3a/b*, seven mutants *Pphag1*). The mutants need to perform the whole life cycle in future work.

Conclusion

7 Conclusion

The present work has aimed for identifying single copy TAPs that play essential roles throughout plant development, especially in the phase transition from vegetative to reproductive growth. Indeed, two fundamental regulators (TAPs) could be characterised with regard to their impact on sexual reproduction in mosses within an evo-devo approach. Therefore, the study contributed to the understanding of sexual reproduction and its underlying gene regulatory networks in the sense of evolutionary conservation and diversification. In particular, the control of the male as well as female germ line in bryophytes was explored in greater detail. Loss-of-function mutations of the respective genes revealed that both candidate genes are essential for proper development of the male germ line.

Apart from that, the present work reinforced the hypothesis that TAPs are essential in terms of plant development. It showed that mosses can be applied for the analysis of such conserved key players and that single copy genes can be used as phylogenetic markers. Furthermore, mosses can serve as model organism to study human diseases that are caused by conserved single copy TAPs.

In addition, the analysis of swi3a/b in bryophytes showed the importance of studying a gene in different organisms to understand how genes evolved throughout land plant evolution. The knowledge of conservation and adaptation of genes is especially important in light of climate change. Plants will be challenged with increasing heat and water stress which is a tremendous change that plants need to adjust to. The water to land transition was a comparable key event in the adaption of land plants towards the novelties accompanied by the new habitat. Studying land plants in evo-devo approaches thus potentially makes it possible to derive predictions with regard to plant's capabilities in facing the stresses of global warming.

References

8 References

Electronical references (last access March 29th 2020):

<https://www.arabidopsis.org/>
<https://www.socscistatistics.com/tests/fisher/default2.aspx>
<https://www.graphpad.com/quickcalcs/contingency1/>
<http://quantpsy.org/chisq/chisq.htm>
<http://bioinformatics.psb.ugent.be/webtools/Venn/>
<https://www.marc-cain.com/index.php?cl=jaggeolandingpage&lang=0&sBackUrl= %2F>
<http://timetree.org/>
<https://tmcalculator.neb.com/#!/main>
<https://plantcode.online.uni-marburg.de/tapscan/index.php>
<http://tree.bio.ed.ac.uk/software/figtree/>
<http://www.wordle.net/>
<https://www.thermofisher.com/order/catalog/product/K1231>
<https://www.thermofisher.com/order/catalog/product/SD0061#/SD0061>
http://plntfdb.bio.uni-potsdam.de/v3.0/fam_mem.php?family_id=MYB-related&sp_id=ATH

- Aasland, R., Stewart, A.F., and Gibson, T. (1996). The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. *Trends Biochem Sci* 21, 87-88.
- Alvarez-Venegas, R., Pien, S., Sadler, M., Witmer, X., Grossniklaus, U., and Avramova, Z. (2003). ATX-1, an Arabidopsis homolog of trithorax, activates flower homeotic genes. *Curr Biol* 13, 627-637.
- Aquea, F., Johnston, A.J., Canon, P., Grossniklaus, U., and Arce-Johnson, P. (2010). TRAUCO, a Trithorax-group gene homologue, is required for early embryogenesis in Arabidopsis thaliana. *J Exp Bot* 61, 1215-1224.
- Aquea, F., Timmermann, T., and Herrera-Vasquez, A. (2017). Chemical inhibition of the histone acetyltransferase activity in Arabidopsis thaliana. *Biochem Biophys Res Commun* 483, 664-668.
- Arabidopsis Genome, I. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* 408, 796-815.
- Aury, J.M., Jaillon, O., Duret, L., Noel, B., Jubin, C., Porcel, B.M., Segurens, B., Daubin, V., Anthouard, V., Aiach, N., *et al.* (2006). Global trends of whole-genome duplications revealed by the ciliate Paramecium tetraurelia. *Nature* 444, 171-178.
- Bar-On, Y.M., Phillips, R., and Milo, R. (2018). The biomass distribution on Earth. *Proc Natl Acad Sci U S A* 115, 6506-6511.

References

- Benhamed, M., Bertrand, C., Servet, C., and Zhou, D.X. (2006). Arabidopsis GCN5, HD1, and TAF1/HAF2 interact to regulate histone acetylation required for light-responsive gene expression. *Plant Cell* 18, 2893-2903.
- Benhamed, M., Martin-Magniette, M.L., Taconnat, L., Bitton, F., Servet, C., De Clercq, R., De Meyer, B., Buysschaert, C., Rombauts, S., Villarroel, R., *et al.* (2008). Genome-scale Arabidopsis promoter array identifies targets of the histone acetyltransferase GCN5. *Plant J* 56, 493-504.
- Bertrand, C., Bergounioux, C., Domenichini, S., Delarue, M., and Zhou, D.X. (2003). Arabidopsis histone acetyltransferase AtGCN5 regulates the floral meristem activity through the WUSCHEL/AGAMOUS pathway. *J Biol Chem* 278, 28246-28251.
- Bimboim, H., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids research* 7, 1513-1523.
- Birchler, J.A., and Veitia, R.A. (2012). Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. *Proc Natl Acad Sci U S A* 109, 14746-14753.
- Bischler, H., and Boisselier-Dubayle, M.C. (1993). VARIATION IN A POLYPLOID, DIOICOUS LIVERWORT, MARCHANTIA GLOBOSA. *American Journal of Botany* 80, 953-958.
- Borg, M., Brownfield, L., Khatab, H., Sidorova, A., Lingaya, M., and Twell, D. (2011). The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in Arabidopsis. *Plant Cell* 23, 534-549.
- Borg, M., Rutley, N., Kagale, S., Hamamura, Y., Gherghinoiu, M., Kumar, S., Sari, U., Esparza-Franco, M.A., Sakamoto, W., Rozwadowski, K., *et al.* (2014). An EAR-Dependent Regulatory Module Promotes Male Germ Cell Division and Sperm Fertility in Arabidopsis. *Plant Cell* 26, 2098-2113.
- Bouyer, D., Roudier, F., Heese, M., Andersen, E.D., Gey, D., Nowack, M.K., Goodrich, J., Renou, J.P., Grini, P.E., Colot, V., *et al.* (2011). Polycomb repressive complex 2 controls the embryo-to-seedling phase transition. *PLoS Genet* 7, e1002014.
- Bower, F. (1908). *The Origin of Land Flora: A Theory Based Upon the Facts of Alternation*. MacMillan.
- Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., *et al.* (2017). Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. *Cell* 171, 287-304 e215.
- Bowman, J.L., Sakakibara, K., Furumizu, C., and Dierschke, T. (2016). Evolution in the Cycles of Life. *Annu Rev Genet* 50, 133-154.
- Boyd, P.J., Grimsley, N.H., and Cove, D.J. (1988). Somatic mutagenesis of the moss, *Physcomitrella patens*. *Molecular and General Genetics MGG* 211, 545-546.
- Boyer, L.A., Latek, R.R., and Peterson, C.L. (2004). The SANT domain: a unique histone-tail-binding module? *Nat Rev Mol Cell Biol* 5, 158-163.
- Brownfield, L., Hafidh, S., Borg, M., Sidorova, A., Mori, T., and Twell, D. (2009). A plant germline-specific integrator of sperm specification and cell cycle progression. *PLoS Genet* 5, e1000430.
- Burk, D.L., Ghuman, N., Wybenga-Groot, L.E., and Berghuis, A.M. (2003). X-ray structure of the AAC(6')-II antibiotic resistance enzyme at 1.8 Å resolution; examination of oligomeric arrangements in GNAT superfamily members. *Protein Sci* 12, 426-437.

References

- Carroll, S.B. (2005). Evolution at two levels: on genes and form. *PLoS Biol* 3, e245.
- Catarino, B., Hetherington, A.J., Emms, D.M., Kelly, S., and Dolan, L. (2016). The Stepwise Increase in the Number of Transcription Factor Families in the Precambrian Predated the Diversification of Plants On Land. *Mol Biol Evol* 33, 2815-2819.
- Cavalier-Smith, T. (2002). Origins of the machinery of recombination and sex. *Heredity (Edinb)* 88, 125-141.
- Cheng, S., Xian, W., Fu, Y., Marin, B., Keller, J., Wu, T., Sun, W., Li, X., Xu, Y., and Zhang, Y. (2019). Genomes of Subaerial Zygnematophyceae Provide Insights into Land Plant Evolution. *Cell* 179, 1057-1067. e1014.
- Choi, S.C., Lee, S., Kim, S.R., Lee, Y.S., Liu, C., Cao, X., and An, G. (2014). Trithorax group protein *Oryza sativa* Trithorax1 controls flowering time in rice via interaction with early heading date3. *Plant Physiol* 164, 1326-1337.
- Cohen, R., Schocken, J., Kaldis, A., Vlachonasios, K.E., Hark, A.T., and McCain, E.R. (2009). The histone acetyltransferase GCN5 affects the inflorescence meristem and stamen development in *Arabidopsis*. *Planta* 230, 1207-1221.
- Collonnier, C., Epert, A., Mara, K., Maclot, F., Guyon-Debast, A., Charlot, F., White, C., Schaefer, D.G., and Nogue, F. (2017). CRISPR-Cas9-mediated efficient directed mutagenesis and RAD51-dependent and RAD51-independent gene targeting in the moss *Physcomitrella patens*. *Plant Biotechnol J* 15, 122-131.
- Concordet, J.P., and Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* 46, W242-W245.
- Cove, D. (2005). The moss *Physcomitrella patens*. *Annu Rev Genet* 39, 339-358.
- Cove, D., Bezanilla, M., Harries, P., and Quatrano, R. (2006). Mosses as model systems for the study of metabolism and development. *Annu Rev Plant Biol* 57, 497-520.
- Cove, D.J., Perroud, P.F., Charron, A.J., McDaniel, S.F., Khandelwal, A., and Quatrano, R.S. (2009). The moss *Physcomitrella patens*: a novel model system for plant development and genomic studies. *Cold Spring Harb Protoc* 2009, pdb emo115.
- Crane, P.R., Friis, E.M., and Pedersen, K.R. (1994). Palaeobotanical evidence on the early radiation of magnoliid angiosperms. In *Early Evolution of Flowers* (Springer), pp. 51-72.
- Crane, P.R., Friis, E.M., and Pedersen, K.R. (1995). The origin and early diversification of angiosperms. *Nature* 374, 27.
- Crepet, W.L. (2000). Progress in understanding angiosperm history, success, and relationships: Darwin's abominably "perplexing phenomenon". *Proc Natl Acad Sci U S A* 97, 12939-12941.
- Crow, J.F. (2005). Timeline: Hermann Joseph Muller, evolutionist. *Nat Rev Genet* 6, 941-945.
- Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27, 1164-1165.
- Dawe, H.R., Farr, H., Portman, N., Shaw, M.K., and Gull, K. (2005). The Parkin co-regulated gene product, PACRG, is an evolutionarily conserved axonemal protein that functions in outer-doublet microtubule morphogenesis. *J Cell Sci* 118, 5421-5430.

References

- De Bodt, S., Maere, S., and Van de Peer, Y. (2005). Genome duplication and the origin of angiosperms. *Trends Ecol Evol* 20, 591-597.
- de Mendoza, A., Sebe-Pedros, A., Sestak, M.S., Matejcic, M., Torruella, G., Domazet-Loso, T., and Ruiz-Trillo, I. (2013). Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proc Natl Acad Sci U S A* 110, E4858-4866.
- de Mendoza, A., Suga, H., Permanyer, J., Irimia, M., and Ruiz-Trillo, I. (2015). Complex transcriptional regulation and independent evolution of fungal-like traits in a relative of animals. *Elife* 4, e08904.
- de Vries, J., and Archibald, J.M. (2018). Plant evolution: landmarks on the path to terrestrial life. *New Phytol* 217, 1428-1434.
- de Vries, J., Curtis, B.A., Gould, S.B., and Archibald, J.M. (2018). Embryophyte stress signaling evolved in the algal progenitors of land plants. *Proc Natl Acad Sci U S A* 115, E3471-E3480.
- de Vries, J., and Rensing, S.A. (2020). Gene gains paved the path to land. *Nat Plants* 6, 7-8.
- Decker, E.L., Frank, W., Sarnighausen, E., and Reski, R. (2006). Moss systems biology en route: phytohormones in *Physcomitrella* development. *Plant Biol (Stuttg)* 8, 397-405.
- Decker, E.L., and Reski, R. (2019). Mosses in biotechnology. *Curr Opin Biotechnol* 61, 21-27.
- Delaux, P.M., Hetherington, A.J., Coudert, Y., Delwiche, C., Dunand, C., Gould, S., Kenrick, P., Li, F.W., Philippe, H., Rensing, S.A., *et al.* (2019). Reconstructing trait evolution in plant evo-devo studies. *Curr Biol* 29, R1110-R1118.
- Delaux, P.M., Radhakrishnan, G.V., Jayaraman, D., Cheema, J., Malbreil, M., Volkening, J.D., Sekimoto, H., Nishiyama, T., Melkonian, M., Pokorny, L., *et al.* (2015). Algal ancestor of land plants was preadapted for symbiosis. *Proc Natl Acad Sci U S A* 112, 13390-13395.
- Delwiche, C.F., and Cooper, E.D. (2015). The Evolutionary Origin of a Terrestrial Flora. *Curr Biol* 25, R899-910.
- Dong, X., Hong, Z., Sivaramakrishnan, M., Mahfouz, M., and Verma, D.P. (2005). Callose synthase (CalS5) is required for exine formation during microgametogenesis and for pollen viability in *Arabidopsis*. *Plant J* 42, 315-328.
- Doyle, J.J., and Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. *Focus* 12, 39-40.
- Duarte, J.M., Wall, P.K., Edger, P.P., Landherr, L.L., Ma, H., Pires, J.C., Leebens-Mack, J., and dePamphilis, C.W. (2010). Identification of shared single copy nuclear genes in *Arabidopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels. *BMC Evol Biol* 10, 61.
- Earley, K.W., Shook, M.S., Brower-Toland, B., Hicks, L., and Pikaard, C.S. (2007). In vitro specificities of *Arabidopsis* co-activator histone acetyltransferases: implications for histone hyperacetylation in gene activation. *Plant J* 52, 615-626.
- Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32, 1792-1797.
- Edger, P.P., and Pires, J.C. (2009). Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes. *Chromosome Res* 17, 699-717.

References

- Engel, P.P. (1968). The induction of biochemical and morphological mutants in the moss *Physcomitrella patens*. *American Journal of Botany* 55, 438-446.
- Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics* 78, 737-756.
- Fernandez-Pozo, N., Haas, F.B., Meyberg, R., Ullrich, K.K., Hiss, M., Perroud, P.F., Hanke, S., Kratz, V., Powell, A.F., Vesty, E.F., *et al.* (2019). PEATmoss (*Physcomitrella* Expression Atlas Tool): a unified gene expression atlas for the model plant *Physcomitrella patens*. *Plant J.*
- Field, K.J., Pressel, S., Duckett, J.G., Rimington, W.R., and Bidartondo, M.I. (2015). Symbiotic options for the conquest of land. *Trends Ecol Evol* 30, 477-486.
- Field, K.J., Rimington, W.R., Bidartondo, M.I., Allinson, K.E., Beerling, D.J., Cameron, D.D., Duckett, J.G., Leake, J.R., and Pressel, S. (2016). Functional analysis of liverworts in dual symbiosis with Glomeromycota and Mucoromycotina fungi under a simulated Palaeozoic CO₂ decline. *ISME J* 10, 1514-1526.
- Frank, W., Decker, E.L., and Reski, R. (2005). Molecular tools to study *Physcomitrella patens*. *Plant Biol (Stuttg)* 7, 220-227.
- Frickenhaus, S., and Beszteri, B. (2008). Quicktree-SD, Software developed by AWI-Bioinformatics.
- Friis, E.M., Pedersen, K.R., and Crane, P.R. (1994). Angiosperm floral structures from the Early Cretaceous of Portugal. In *Early Evolution of Flowers* (Springer), pp. 31-49.
- Friis, E.M., Pedersen, K.R., and Crane, P.R. (1999). Early angiosperm diversification: the diversity of pollen associated with angiosperm reproductive structures in Early Cretaceous floras from Portugal. *Annals of the Missouri Botanical Garden*, 259-296.
- Friis, E.M., Pedersen, K.R., and Crane, P.R. (2006). Cretaceous angiosperm flowers: innovation and evolution in plant reproduction. *Palaeogeography, palaeoclimatology, palaeoecology* 232, 251-293.
- Furst-Jansen, J.M.R., de Vries, S., and Vries, J. (2020). Evo-physio: on stress responses and the earliest land plants. *J Exp Bot.*
- Gao, Y., Yang, S., Yuan, L., Cui, Y., and Wu, K. (2012). Comparative Analysis of SWIRM Domain-Containing Proteins in Plants. *Comp Funct Genomics* 2012, 310402.
- Garg, S.G., and Martin, W.F. (2016). Mitochondria, the Cell Cycle, and the Origin of Sex via a Syncytial Eukaryote Common Ancestor. *Genome Biol Evol* 8, 1950-1970.
- Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2011). Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol* 12, 36-47.
- Gerke, P., Szovenyi, P., Neubauer, A., Lenz, H., Gutmann, B., McDowell, R., Small, I., Schallenberg-Rudinger, M., and Knoop, V. (2020). Towards a plant model for enigmatic U-to-C RNA editing: the organelle genomes, transcriptomes, editomes and candidate RNA editing factors in the hornwort *Anthoceros agrestis*. *New Phytol* 225, 1974-1992.
- Green, R., and Rogers, E.J. (2013). Chemical transformation of *E. coli*. *Methods in enzymology* 529, 329.
- Grune, T., Brzeski, J., Eberharter, A., Clapier, C.R., Corona, D.F., Becker, P.B., and Muller, C.W. (2003). Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. *Mol Cell* 12, 449-460.

References

- Guan, R., Zhao, Y., Zhang, H., Fan, G., Liu, X., Zhou, W., Shi, C., Wang, J., Liu, W., Liang, X., *et al.* (2016). Draft genome of the living fossil *Ginkgo biloba*. *Gigascience* 5, 49.
- Gutierrez, R.A., Green, P.J., Keegstra, K., and Ohlrogge, J.B. (2004). Phylogenetic profiling of the *Arabidopsis thaliana* proteome: what proteins distinguish plants from other organisms? *Genome Biol* 5, R53.
- Hackenberg, D., and Twell, D. (2019). The evolution and patterning of male gametophyte development. *Curr Top Dev Biol* 131, 257-298.
- Hake, S., Smith, H.M., Holtan, H., Magnani, E., Mele, G., and Ramirez, J. (2004). The role of *knox* genes in plant development. *Annu Rev Cell Dev Biol* 20, 125-151.
- Hay, A., and Tsiantis, M. (2010). KNOX genes: versatile regulators of plant development and diversity. *Development* 137, 3153-3165.
- Hedges, S.B., Marin, J., Suleski, M., Paymer, M., and Kumar, S. (2015). Tree of life reveals clock-like speciation and diversification. *Mol Biol Evol* 32, 835-845.
- Higo, A., Kawashima, T., Borg, M., Zhao, M., Lopez-Vidriero, I., Sakayama, H., Montgomery, S.A., Sekimoto, H., Hackenberg, D., Shimamura, M., *et al.* (2018). Transcription factor DUO1 generated by neo-functionalization is associated with evolution of sperm differentiation in plants. *Nat Commun* 9, 5283.
- Hisanaga, T., Okahashi, K., Yamaoka, S., Kajiwar, T., Nishihama, R., Shimamura, M., Yamato, K.T., Bowman, J.L., Kohchi, T., and Nakajima, K. (2019a). A cis-acting bidirectional transcription switch controls sexual dimorphism in the liverwort. *The EMBO journal* 38.
- Hisanaga, T., Yamaoka, S., Kawashima, T., Higo, A., Nakajima, K., Araki, T., Kohchi, T., and Berger, F. (2019b). Building new insights in plant gametogenesis from an evolutionary perspective. *Nat Plants* 5, 663-669.
- Hiss, M., Meyberg, R., Westermann, J., Haas, F.B., Schneider, L., Schallenberg-Rudinger, M., Ullrich, K.K., and Rensing, S.A. (2017a). Sexual reproduction, sporophyte development and molecular variation in the model moss *Physcomitrella patens*: introducing the ecotype Reute. *Plant J* 90, 606-620.
- Hiss, M., Schneider, L., Grosche, C., Barth, M.A., Neu, C., Symeonidi, A., Ullrich, K.K., Perroud, P.F., Schallenberg-Rudinger, M., and Rensing, S.A. (2017b). Combination of the Endogenous *lhcsr1* Promoter and Codon Usage Optimization Boosts Protein Expression in the Moss *Physcomitrella patens*. *Front Plant Sci* 8, 1842.
- Hofmeister, W. (1851). *Vergleichende Untersuchungen der Keimung, Entfaltung und Fruchtbildung höherer Kryptogamen*. . Reprint: *Hist Nat Classica* 105 Cramer, Vaduz 1979.
- Hohe, A., Rensing, S.A., Mildner, M., Lang, D., and Reski, R. (2002). Day Length and Temperature Strongly Influence Sexual Reproduction and Expression of a Novel MADS-Box Gene in the Moss *Physcomitrella patens*. *Plant Biology* 4, 595-602.
- Hohe, A., and Reski, R. (2002). Optimisation of a bioreactor culture of the moss *Physcomitrella patens* for mass production of protoplasts. *Plant Science* 163, 69-74.

References

- Hohe A., R.S.A., Mildner M., Lang D., Reski R. (2002). Day length and temperature strongly influence sexual reproduction and expression of a novel MADS-Box gene in the moss *Physcomitrella patens*. *Plant Biol.*
- Hori, K., Maruyama, F., Fujisawa, T., Togashi, T., Yamamoto, N., Seo, M., Sato, S., Yamada, T., Mori, H., Tajima, N., *et al.* (2014). *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nat Commun* 5, 3978.
- Horst, N.A., Katz, A., Pereman, I., Decker, E.L., Ohad, N., and Reski, R. (2016). A single homeobox gene triggers phase transition, embryogenesis and asexual reproduction. *Nat Plants* 2, 15209.
- Hsia, C.C., and McGinnis, W. (2003). Evolution of transcription factor function. *Curr Opin Genet Dev* 13, 199-206.
- Hu, S., Dilcher, D.L., Jarzen, D.M., and Taylor, D.W. (2008). Early steps of angiosperm–pollinator coevolution. *Proceedings of the National Academy of Sciences* 105, 240-245.
- Ikeuchi, M., Iwase, A., Rymen, B., Harashima, H., Shibata, M., Ohnuma, M., Breuer, C., Morao, A.K., de Lucas, M., De Veylder, L., *et al.* (2015). PRC2 represses dedifferentiation of mature somatic cells in *Arabidopsis*. *Nat Plants* 1, 15089.
- Ishizaki, K., Nishihama, R., Ueda, M., Inoue, K., Ishida, S., Nishimura, Y., Shikanai, T., and Kohchi, T. (2015). Development of Gateway Binary Vector Series with Four Different Selection Markers for the Liverwort *Marchantia polymorpha*. *PLoS One* 10, e0138876.
- Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., *et al.* (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236-1240.
- Joo, S., Wang, M.H., Lui, G., Lee, J., Barnas, A., Kim, E., Sudek, S., Worden, A.Z., and Lee, J.H. (2018). Common ancestry of heterodimerizing TALE homeobox transcription factors across Metazoa and Archaeplastida. *BMC Biol* 16, 136.
- Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30, 772-780.
- Katz, A., Oliva, M., Mosquana, A., Hakim, O., and Ohad, N. (2004). FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development. *Plant J* 37, 707-719.
- Kawashima, T., and Berger, F. (2014). Epigenetic reprogramming in plant sexual reproduction. *Nat Rev Genet* 15, 613-624.
- Kenrick, P. (2017). How land plant life cycles first evolved. *Science* 358, 1538-1539.
- Kim, J.Y., Oh, J.E., Noh, Y.S., and Noh, B. (2015). Epigenetic control of juvenile-to-adult phase transition by the *Arabidopsis* SAGA-like complex. *Plant J* 83, 537-545.
- Kim, W., Benhamed, M., Servet, C., Latrasse, D., Zhang, W., Delarue, M., and Zhou, D.X. (2009). Histone acetyltransferase GCN5 interferes with the miRNA pathway in *Arabidopsis*. *Cell Res* 19, 899-909.
- Kim, Y., Fedoriw, A.M., and Magnuson, T. (2012). An essential role for a mammalian SWI/SNF chromatin-remodeling complex during male meiosis. *Development* 139, 1133-1140.

References

- Kofuji, R., Yagita, Y., Murata, T., and Hasebe, M. (2018a). Antheridial development in the moss *Physcomitrella patens*: implications for understanding stem cells in mosses. *Philos Trans R Soc Lond B Biol Sci* 373.
- Kofuji, R., Yoshimura, T., Inoue, H., Sakakibara, K., Hiwatashi, Y., Kurata, T., Aoyama, T., Ueda, K., and Hasebe, M. (2018b). Gametangia development in the moss *Physcomitrella patens*. *Annual Plant Reviews online*, 167-181.
- Koi, S., Hisanaga, T., Sato, K., Shimamura, M., Yamato, K.T., Ishizaki, K., Kohchi, T., and Nakajima, K. (2016). An Evolutionarily Conserved Plant RKD Factor Controls Germ Cell Differentiation. *Curr Biol* 26, 1775-1781.
- Kubo, M., Imai, A., Nishiyama, T., Ishikawa, M., Sato, Y., Kurata, T., Hiwatashi, Y., Reski, R., and Hasebe, M. (2013). System for stable beta-estradiol-inducible gene expression in the moss *Physcomitrella patens*. *PLoS One* 8, e77356.
- Landberg, K., Pederson, E.R., Viaene, T., Bozorg, B., Friml, J., Jonsson, H., Thelander, M., and Sundberg, E. (2013). The MOSS *Physcomitrella patens* reproductive organ development is highly organized, affected by the two SHI/STY genes and by the level of active auxin in the SHI/STY expression domain. *Plant Physiol* 162, 1406-1419.
- Lang, D., Ullrich, K.K., Murat, F., Fuchs, J., Jenkins, J., Haas, F.B., Piednoel, M., Gundlach, H., Van Bel, M., Meyberg, R., *et al.* (2018). The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J* 93, 515-533.
- Lang, D., Weiche, B., Timmerhaus, G., Richardt, S., Riano-Pachon, D.M., Correa, L.G., Reski, R., Mueller-Roeber, B., and Rensing, S.A. (2010). Genome-wide phylogenetic comparative analysis of plant transcriptional regulation: a timeline of loss, gain, expansion, and correlation with complexity. *Genome Biol Evol* 2, 488-503.
- Lechner, M., Findeiss, S., Steiner, L., Marz, M., Stadler, P.F., and Prohaska, S.J. (2011). Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* 12, 124.
- Lee, J.H., Lin, H., Joo, S., and Goodenough, U. (2008). Early sexual origins of homeoprotein heterodimerization and evolution of the plant KNOX/BELL family. *Cell* 133, 829-840.
- Lenton, T.M., Dahl, T.W., Daines, S.J., Mills, B.J., Ozaki, K., Saltzman, M.R., and Porada, P. (2016). Earliest land plants created modern levels of atmospheric oxygen. *Proc Natl Acad Sci U S A* 113, 9704-9709.
- Lespinet, O., Wolf, Y.I., Koonin, E.V., and Aravind, L. (2002). The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Res* 12, 1048-1059.
- Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. *Nature* 424, 147-151.
- Long, J.A., Ohno, C., Smith, Z.R., and Meyerowitz, E.M. (2006). TOPLESS regulates apical embryonic fate in *Arabidopsis*. *Science* 312, 1520-1523.
- Lopez-Obando, M., Hoffmann, B., Gery, C., Guyon-Debast, A., Teoule, E., Rameau, C., Bonhomme, S., and Nogue, F. (2016). Simple and Efficient Targeting of Multiple Genes Through CRISPR-Cas9 in *Physcomitrella patens*. *G3 (Bethesda)* 6, 3647-3653.
- Lorenzetti, D., Bishop, C.E., and Justice, M.J. (2004). Deletion of the Parkin coregulated gene causes male sterility in the quaking(viable) mouse mutant. *Proc Natl Acad Sci U S A* 101, 8402-8407.

References

- Luense, L.J., Donahue, G., Lin-Shiao, E., Rangel, R., Weller, A.H., Bartolomei, M.S., and Berger, S.L. (2019). Gcn5-Mediated Histone Acetylation Governs Nucleosome Dynamics in Spermiogenesis. *Dev Cell* 51, 745-758 e746.
- Lyons, T.W., Reinhard, C.T., and Planavsky, N.J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature* 506, 307-315.
- Mahrez, W., Arellano, M.S., Moreno-Romero, J., Nakamura, M., Shu, H., Nanni, P., Kohler, C., Gruissem, W., and Hennig, L. (2016). H3K36ac Is an Evolutionary Conserved Plant Histone Modification That Marks Active Genes. *Plant Physiol* 170, 1566-1577.
- Martin, C., and Paz-Ares, J. (1997). MYB transcription factors in plants. *Trends Genet* 13, 67-73.
- McCourt, R.M., Delwiche, C.F., and Karol, K.G. (2004). Charophyte algae and land plant origins. *Trends Ecol Evol* 19, 661-666.
- Melzer, R., Verelst, W., and Theissen, G. (2009). The class E floral homeotic protein SEPALLATA3 is sufficient to loop DNA in 'floral quartet'-like complexes in vitro. *Nucleic Acids Res* 37, 144-157.
- Menon, D.U., Shibata, Y., Mu, W., and Magnuson, T. (2019). Mammalian SWI/SNF collaborates with a polycomb-associated protein to regulate male germline transcription in the mouse. *Development* 146.
- Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Marechal-Drouard, L., *et al.* (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318, 245-250.
- Meyberg, R., Perroud, P.F., Haas, F.B., Schneider, L., Heimerl, T., Renzaglia, K.S., and Rensing, S.A. (2020). Characterization of evolutionarily conserved key players affecting eukaryotic flagellar motility and fertility using a moss model. *New Phytol.*
- Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.Y., El-Gebali, S., Fraser, M.I., *et al.* (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res* 47, D351-D360.
- Mitchell, D.R. (2007). The evolution of eukaryotic cilia and flagella as motile and sensory organelles. *Adv Exp Med Biol* 607, 130-140.
- Moran, N.A. (1996). Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc Natl Acad Sci U S A* 93, 2873-2878.
- Morant, M., Jorgensen, K., Schaller, H., Pinot, F., Moller, B.L., Werck-Reichhart, D., and Bak, S. (2007). CYP703 is an ancient cytochrome P450 in land plants catalyzing in-chain hydroxylation of lauric acid to provide building blocks for sporopollenin synthesis in pollen. *Plant Cell* 19, 1473-1487.
- Morris, J.L., Puttick, M.N., Clark, J.W., Edwards, D., Kenrick, P., Pressel, S., Wellman, C.H., Yang, Z., Schneider, H., and Donoghue, P.C.J. (2018). The timescale of early land plant evolution. *Proc Natl Acad Sci U S A* 115, E2274-E2283.
- Mosquna, A., Katz, A., Decker, E.L., Rensing, S.A., Reski, R., and Ohad, N. (2009). Regulation of stem cell maintenance by the Polycomb protein FIE has been conserved during land plant evolution. *Development* 136, 2433-2444.

References

- Motomura, K., Berger, F., Kawashima, T., Kinoshita, T., Higashiyama, T., and Maruyama, D. (2016). Fertilization-independent Cell-fusion between the Synergid and Central Cell in the Polycomb Mutant. *Cell Struct Funct* **41**, 121-125.
- Muller, H.J. (1964). The Relation of Recombination to Mutational Advance. *Mutat Res* **106**, 2-9.
- Nakosteen, P.C., and Hughes, K.W. (1978). Sexual Life Cycle of Three Species of Funariaceae in Culture. *The Bryologist Vol. 81, No. 2*, 307-314.
- Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution* **32**, 268-274.
- Niklas, K.J., and Kutschera, U. (2010). The evolution of the land plant life cycle. *New Phytologist* **185**, 27-41.
- Nishiyama, T., Sakayama, H., de Vries, J., Buschmann, H., Saint-Marcoux, D., Ullrich, K.K., Haas, F.B., Vanderstraeten, L., Becker, D., Lang, D., *et al.* (2018). The Chara Genome: Secondary Complexity and Implications for Plant Terrestrialization. *Cell* **174**, 448-464 e424.
- Nomura, T., Sakurai, T., Osakabe, Y., Osakabe, K., and Sakakibara, H. (2016). Efficient and Heritable Targeted Mutagenesis in Mosses Using the CRISPR/Cas9 System. *Plant Cell Physiol* **57**, 2600-2610.
- O'Donoghue, M.T., Chater, C., Wallace, S., Gray, J.E., Beerling, D.J., and Fleming, A.J. (2013). Genome-wide transcriptomic analysis of the sporophyte of the moss *Physcomitrella patens*. *J Exp Bot* **64**, 3567-3581.
- O'Maileidigh, D.S., Graciet, E., and Wellmer, F. (2014). Gene networks controlling *Arabidopsis thaliana* flower development. *New Phytol* **201**, 16-30.
- Okano, Y., Aono, N., Hiwatashi, Y., Murata, T., Nishiyama, T., Ishikawa, T., Kubo, M., and Hasebe, M. (2009). A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution. *Proc Natl Acad Sci U S A* **106**, 16321-16326.
- Ortiz-Ramirez, C., Hernandez-Coronado, M., Thamm, A., Catarino, B., Wang, M., Dolan, L., Feijo, J.A., and Becker, J.D. (2016). A Transcriptome Atlas of *Physcomitrella patens* Provides Insights into the Evolution and Development of Land Plants. *Mol Plant* **9**, 205-220.
- Papp, B., Pal, C., and Hurst, L.D. (2003). Dosage sensitivity and the evolution of gene families in yeast. *Nature* **424**, 194-197.
- Park, J.I., Ishimizu, T., Suwabe, K., Sudo, K., Masuko, H., Hakozaiki, H., Nou, I.S., Suzuki, G., and Watanabe, M. (2010). UDP-glucose pyrophosphorylase is rate limiting in vegetative and reproductive phases in *Arabidopsis thaliana*. *Plant Cell Physiol* **51**, 981-996.
- Paterson, A.H., Chapman, B.A., Kissinger, J.C., Bowers, J.E., Feltus, F.A., and Estill, J.C. (2006). Many gene and domain families have convergent fates following independent whole-genome duplication events in *Arabidopsis*, *Oryza*, *Saccharomyces* and *Tetraodon*. *Trends Genet* **22**, 597-602.
- Perroud, P.F., Cove, D.J., Quatrano, R.S., and McDaniel, S.F. (2011). An experimental method to facilitate the identification of hybrid sporophytes in the moss *Physcomitrella patens* using fluorescent tagged lines. *New Phytol* **191**, 301-306.

References

- Perroud, P.F., Demko, V., Johansen, W., Wilson, R.C., Olsen, O.A., and Quatrano, R.S. (2014). Defective Kernel 1 (DEK1) is required for three-dimensional growth in *Physcomitrella patens*. *New Phytol* 203, 794-804.
- Perroud, P.F., Haas, F.B., Hiss, M., Ullrich, K.K., Alboresi, A., Amirebrahimi, M., Barry, K., Bassi, R., Bonhomme, S., Chen, H., *et al.* (2018). The *Physcomitrella patens* gene atlas project: large-scale RNA-seq based expression data. *Plant J* 95, 168-182.
- Perroud, P.F., Meyberg, R., and Rensing, S.A. (2019). *Physcomitrella patens* Reute mCherry as a tool for efficient crossing within and between ecotypes. *Plant Biol (Stuttg)* 21 Suppl 1, 143-149.
- Phan, H.M., Xu, A.W., Coco, C., Srajer, G., Wyszomierski, S., Evrard, Y.A., Eckner, R., and Dent, S.Y. (2005). GCN5 and p300 share essential functions during early embryogenesis. *Dev Dyn* 233, 1337-1347.
- Pires, N.D., Yi, K., Breuninger, H., Catarino, B., Menand, B., and Dolan, L. (2013). Recruitment and remodeling of an ancient gene regulatory network during land plant evolution. *Proc Natl Acad Sci U S A* 110, 9571-9576.
- Porada, P., Lenton, T.M., Pohl, A., Weber, B., Mander, L., Donnadieu, Y., Beer, C., Poschl, U., and Kleidon, A. (2016). High potential for weathering and climate effects of non-vascular vegetation in the Late Ordovician. *Nat Commun* 7, 12113.
- Postma, M., and Goedhart, J. (2019). PlotsOfData-A web app for visualizing data together with their summaries. *PLoS Biol* 17, e3000202.
- Poulios, S., and Vlachonasios, K.E. (2018). Synergistic action of GCN5 and CLAVATA1 in the regulation of gynoecium development in *Arabidopsis thaliana*. *New Phytol* 220, 593-608.
- Quatrano, R.S., McDaniel, S.F., Khandelwal, A., Perroud, P.F., and Cove, D.J. (2007). *Physcomitrella patens*: mosses enter the genomic age. *Curr Opin Plant Biol* 10, 182-189.
- Rabiger, D.S., and Drews, G.N. (2013). MYB64 and MYB119 are required for cellularization and differentiation during female gametogenesis in *Arabidopsis thaliana*. *PLoS genetics* 9.
- Rensing, S.A. (2014). Gene duplication as a driver of plant morphogenetic evolution. *Curr Opin Plant Biol* 17, 43-48.
- Rensing, S.A. (2016). (Why) Does Evolution Favour Embryogenesis? *Trends Plant Sci* 21, 562-573.
- Rensing, S.A. (2017). Why we need more non-seed plant models. *New Phytol* 216, 355-360.
- Rensing, S.A. (2018). Great moments in evolution: the conquest of land by plants. *Curr Opin Plant Biol* 42, 49-54.
- Rensing, S.A., Lang, D., Zimmer, A.D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P.F., Lindquist, E.A., Kamisugi, Y., *et al.* (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319, 64-69.
- Renzaglia, K.S., Duff, R.J.T., Nickrent, D.L., and Garbary, D.J. (2000). Vegetative and reproductive innovations of early land plants: implications for a unified phylogeny. *Philos Trans R Soc Lond B Biol Sci* 355, 769-793.
- Renzaglia, K.S., and Garbary, D.J. (2001). Motile Gametes of Land Plants: Diversity, Development, and Evolution. *Critical Reviews in Plant Sciences* 20, 107-213.

References

- Richardt, S., Lang, D., Reski, R., Frank, W., and Rensing, S.A. (2007). PlanTAPDB, a phylogeny-based resource of plant transcription-associated proteins. *Plant Physiol* 143, 1452-1466.
- Roberts, C.W., and Orkin, S.H. (2004). The SWI/SNF complex--chromatin and cancer. *Nat Rev Cancer* 4, 133-142.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61, 539-542.
- Rotman, N., Durbarry, A., Wardle, A., Yang, W.C., Chaboud, A., Faure, J.E., Berger, F., and Twell, D. (2005). A novel class of MYB factors controls sperm-cell formation in plants. *Curr Biol* 15, 244-248.
- Rovekamp, M., Bowman, J.L., and Grossniklaus, U. (2016). Marchantia MpRKD Regulates the Gametophyte-Sporophyte Transition by Keeping Egg Cells Quiescent in the Absence of Fertilization. *Curr Biol* 26, 1782-1789.
- Sakakibara, K., Ando, S., Yip, H.K., Tamada, Y., Hiwatashi, Y., Murata, T., Deguchi, H., Hasebe, M., and Bowman, J.L. (2013). KNOX2 genes regulate the haploid-to-diploid morphological transition in land plants. *Science* 339, 1067-1070.
- Sakakibara, K., Reisewitz, P., Aoyama, T., Friedrich, T., Ando, S., Sato, Y., Tamada, Y., Nishiyama, T., Hiwatashi, Y., Kurata, T., *et al.* (2014). WOX13-like genes are required for reprogramming of leaf and protoplast cells into stem cells in the moss *Physcomitrella patens*. *Development* 141, 1660-1670.
- Sanchez-Vera, V., Kenchappa, C.S., Landberg, K., Bressendorff, S., Schwarzbach, S., Martin, T., Mundy, J., Petersen, M., Thelander, M., and Sundberg, E. (2017). Autophagy is required for gamete differentiation in the moss *Physcomitrella patens*. *Autophagy* 13, 1939-1951.
- Sarnowski, T.J., Rios, G., Jasik, J., Swiezewski, S., Kaczanowski, S., Li, Y., Kwiatkowska, A., Pawlikowska, K., Kozbial, M., Kozbial, P., *et al.* (2005). SWI3 subunits of putative SWI/SNF chromatin-remodeling complexes play distinct roles during Arabidopsis development. *Plant Cell* 17, 2454-2472.
- Sayou, C., Monniaux, M., Nanao, M.H., Moyroud, E., Brockington, S.F., Thevenon, E., Chahtane, H., Warthmann, N., Melkonian, M., Zhang, Y., *et al.* (2014). A promiscuous intermediate underlies the evolution of LEAFY DNA binding specificity. *Science* 343, 645-648.
- Schaefer, D.G. (2001). Gene targeting in *Physcomitrella patens*. *Curr Opin Plant Biol* 4, 143-150.
- Schaefer, D.G., and Zryd, J.P. (1997). Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J* 11, 1195-1206.
- Schmid, M.W., Giraldo-Fonseca, A., Rovekamp, M., Smetanin, D., Bowman, J.L., and Grossniklaus, U. (2018). Extensive epigenetic reprogramming during the life cycle of *Marchantia polymorpha*. *Genome Biol* 19, 9.
- Shimamura, M. (2016). *Marchantia polymorpha*: Taxonomy, Phylogeny and Morphology of a Model System. *Plant Cell Physiol* 57, 230-256.
- Smyth, D.R., Bowman, J.L., and Meyerowitz, E.M. (1990). Early flower development in *Arabidopsis*. *Plant Cell* 2, 755-767.

References

- Solly, J.E., Cunliffe, N.J., and Harrison, C.J. (2017). Regional growth rate differences specified by apical notch activities regulate liverwort thallus shape. *Current Biology* 27, 16-26.
- Soltis, D.E., Bell, C.D., Kim, S., and Soltis, P.S. (2008). Origin and early evolution of angiosperms. *Ann N Y Acad Sci* 1133, 3-25.
- Southworth, D., and Cresti, M. (1997). Comparison of Flagellated and Nonflagellated Sperm in Plants. *American Journal of Botany* 84, 1301-1311.
- Sperschneider, J., Catanzariti, A.M., DeBoer, K., Petre, B., Gardiner, D.M., Singh, K.B., Dodds, P.N., and Taylor, J.M. (2017). LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Sci Rep* 7, 44598.
- Stewart, K.D., and Mattox, K.R. (1975). Comparative cytology, evolution and classification of the green algae with some consideration of the origin of other organisms with chlorophylls A and B. *The Botanical Review* 41, 104-135.
- Strotbek, C., Krinninger, S., and Frank, W. (2013). The moss *Physcomitrella patens*: methods and tools from cultivation to targeted analysis of gene function. *Int J Dev Biol* 57, 553-564.
- Sugano, S.S., Nishihama, R., Shirakawa, M., Takagi, J., Matsuda, Y., Ishida, S., Shimada, T., Hara-Nishimura, I., Osakabe, K., and Kohchi, T. (2018). Efficient CRISPR/Cas9-based genome editing and its application to conditional genetic analysis in *Marchantia polymorpha*. *PLoS One* 13, e0205117.
- Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., *et al.* (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47, D607-D613.
- Szovenyi, P., Frangedakis, E., Ricca, M., Quandt, D., Wicke, S., and Langdale, J.A. (2015). Establishment of *Anthoceros agrestis* as a model species for studying the biology of hornworts. *BMC Plant Biol* 15, 98.
- Tanahashi, T., Sumikawa, N., Kato, M., and Hasebe, M. (2005). Diversification of gene function: homologs of the floral regulator FLO/LFY control the first zygotic cell division in the moss *Physcomitrella patens*. *Development* 132, 1727-1736.
- Theissen, G., and Melzer, R. (2007). Molecular mechanisms underlying origin and diversification of the angiosperm flower. *Ann Bot* 100, 603-619.
- Thien, L.B., Bernhardt, P., Devall, M.S., Chen, Z.d., Luo, Y.b., Fan, J.H., Yuan, L.C., and Williams, J.H. (2009). Pollination biology of basal angiosperms (ANITA grade). *American Journal of Botany* 96, 166-182.
- Transeau, E.N. (1951). *The Zygnemataceae (fresh-water Conjugate Algae) with Keys for the Identification of Genera and Species: And Seven Hundred Eighty-nine Illustrations* (Ohio State University Press).
- Van de Peer, Y., Maere, S., and Meyer, A. (2009). The evolutionary significance of ancient genome duplications. *Nat Rev Genet* 10, 725-732.
- van den Ende, H., Musgrave, A., and Klis, F.M. (1990). *The Role of Flagella in the Sexual Reproduction of Chlamydomonas Gametes*. Springer, Boston, MA.

References

- Verdone, L., Caserta, M., and Di Mauro, E. (2005). Role of histone acetylation in the control of gene expression. *Biochem Cell Biol* 83, 344-353.
- Vesty, E.F., Saidi, Y., Moody, L.A., Holloway, D., Whitbread, A., Needs, S., Choudhary, A., Burns, B., McLeod, D., Bradshaw, S.J., *et al.* (2016). The decision to germinate is regulated by divergent molecular networks in spores and seeds. *New Phytol* 211, 952-966.
- Vlachonasios, K.E., Thomashow, M.F., and Triezenberg, S.J. (2003). Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect Arabidopsis growth, development, and gene expression. *Plant Cell* 15, 626-638.
- Wallace, S., Fleming, A., Wellman, C.H., and Beerling, D.J. (2011). Evolutionary development of the plant spore and pollen wall. *AoB plants* 2011.
- Wang, C., Liu, Y., Li, S.-S., and Han, G.-Z. (2015). Insights into the Origin and Evolution of the Plant Hormone Signaling Machinery. *Plant Physiology* 167, 872-886.
- Wang, S., Li, L., Li, H., Sahu, S.K., Wang, H., Xu, Y., Xian, W., Song, B., Liang, H., Cheng, S., *et al.* (2020). Genomes of early-diverging streptophyte algae shed light on plant terrestrialization. *Nat Plants* 6, 95-106.
- Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189-1191.
- Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., Ayyampalayam, S., Barker, M.S., Burleigh, J.G., Gitzendanner, M.A., *et al.* (2014). Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc Natl Acad Sci U S A* 111, E4859-4868.
- Widiez, T., Symeonidi, A., Luo, C., Lam, E., Lawton, M., and Rensing, S.A. (2014). The chromatin landscape of the moss *Physcomitrella patens* and its dynamics during development and drought stress. *Plant J* 79, 67-81.
- Wilhelmsson, P.K.I., Muhlich, C., Ullrich, K.K., and Rensing, S.A. (2017). Comprehensive Genome-Wide Classification Reveals That Many Plant-Specific Transcription Factors Evolved in Streptophyte Algae. *Genome Biol Evol* 9, 3384-3397.
- Wilson, B.G., Wang, X., Shen, X., McKenna, E.S., Lemieux, M.E., Cho, Y.-J., Koellhoffer, E.C., Pomeroy, S.L., Orkin, S.H., and Roberts, C.W. (2010). Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation. *Cancer cell* 18, 316-328.
- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2, e718.
- Wozniak, N.J., and Sicard, A. (2018). Evolvability of flower geometry: Convergence in pollinator-driven morphological evolution of flowers. *Semin Cell Dev Biol* 79, 3-15.
- Xu, W., Edmondson, D.G., Evrard, Y.A., Wakamiya, M., Behringer, R.R., and Roth, S.Y. (2000). Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development. *Nat Genet* 26, 229-232.
- Yamaoka, S., Nishihama, R., Yoshitake, Y., Ishida, S., Inoue, K., Saito, M., Okahashi, K., Bao, H., Nishida, H., Yamaguchi, K., *et al.* (2018). Generative Cell Specification Requires Transcription Factors Evolutionarily Conserved in Land Plants. *Curr Biol* 28, 479-486 e475.

References

- Yamauchi, T., Yamauchi, J., Kuwata, T., Tamura, T., Yamashita, T., Bae, N., Westphal, H., Ozato, K., and Nakatani, Y. (2000). Distinct but overlapping roles of histone acetylase PCAF and of the closely related PCAF-B/GCN5 in mouse embryogenesis. *Proc Natl Acad Sci U S A* 97, 11303-11306.
- Zaveska Drabkova, L., and Honys, D. (2017). Evolutionary history of callose synthases in terrestrial plants with emphasis on proteins involved in male gametophyte development. *PLoS One* 12, e0187331.
- Zhao, D.Z., Wang, G.F., Speal, B., and Ma, H. (2002). The excess microsporocytes1 gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the Arabidopsis anther. *Genes Dev* 16, 2021-2031.
- Zhao, Y.P., Fan, G., Yin, P.P., Sun, S., Li, N., Hong, X., Hu, G., Zhang, H., Zhang, F.M., Han, J.D., *et al.* (2019). Resequencing 545 ginkgo genomes across the world reveals the evolutionary history of the living fossil. *Nat Commun* 10, 4201.

9 Supplement

A Primer

Table 5 Overview of the used primers. Supplier: Sigma Aldrich (St. Louis, USA).

number	name	sequence
1530	PIG_Seq	AGACACTTCAAAAGAGGATCTAATGTAG
1523	Pig_SphI_fwd	GCATGCAAGCTGATAATTggtacc
1524	Pig_PmeI_SphI_rev	GCATGCGTTTAACTATGGACCCAAATTGTGAGATCCTTC
1517	PIG_HR1_fwd_SpeI	ACTAGTACCCAGGTAGGACCCTAAGATTTC
1518	PIG_HR2_rev_SpeI	ACTAGTAGGCATGCTATGGACCCAAATTG
1495	shpCAB_fwd	gatATCACGCTTGAAGAGAATGTGGG
1496	shCAB_rev_SwaIEcoRI	GAATTCATTTAAATGGTTCTAAAGCAGTTCAGAGAAACGC
1497	Actinprom_fwd_Seq	TGAATCCCTCAGCATTGTTTCATCG
1482	Mp79s0052_fwd_AscI	GGCGCGCCATGGCGGACGACAATGAG
1483	Mp79s0052_rev_AscI	GGCGCGCCTCAGCCATGCCTTGTCAAATG
1484	Mp187s0003_fwd_AscI	GGCGCGCCATGCTTTCACAAAATTCTGAGTCACTCAC
1485	Mp187s0003_rev_AscI	GGCGCGCCCTACTGCCTTCCGCCTC
1486	Mp14s0153_fwd_AscI	GGCGCGCCATGGTGTGCATACGGCAG
1487	Mp14s0153_rev_AscI	GGCGCGCCTTACGCAACAGAAGTGGCAAC
1452	AT5G13780_fwd_AscI	GGCGCGCCATGGTTTGCATCAGGCGAG
1453	AT5G13780_rev_AscI	GGCGCGCCTCATTTGGAAACTGCTTTACCATCTACG
1454	AT3G54610_fwd_AscI	GGCGCGCCATGGACTCTCACTCTTCCCAC
1455	AT3G54610_rev_AscI	GGCGCGCCCTATTGAGATTTAGCACCAGATTGGAGAC
1456	AT2G47620_fwd_AscI	GGCGCGCCATGGAAGCCACTGATCCAAGC
1457	AT2G47620_rev_AscI	GGCGCGCCTCATTTACGTACGTATGATCCCAACG
1430	Marpo_female_fwd	CACCATGGGCCTACTTGTTTCAGTCGCTGGTGG
1431	Marpo_female_rev	TCAAAGGCTAGTGTTTCCATTACTTGGAC
1432	Marpo_male_fwd	GCAGCTGTGTTTTGTGCAGATCGTC
1433	Marpo_male_rev	ATTCTGACCTTACAAGAAATCCTCC
1434	M14s0153_geno_fwd	TGCAAGCACTAAAGGGAAGC
1435	M14s0153_geno_rev	AGACCGCCATCGACTGATC
1436	M79s0052_geno_fwd	TCGTGAATGTGGAGATTCATGC
1437	M79s0052_geno_rev	ACGACCGACATACCTTGC
1438	M79s0052_geno_rev2	AGACCCTCTTGTTTCATGAGCAG
1439	M187s0003_geno_fwd	ACGTACATGAGAGTACGAAAGCA
1440	M187s0003_geno_rev	TCACGACTGCAGTACGCTC
1441	M187s0003_geno_rev2	ATACTCCTTTGTTGCACAGATGC
1388	Pp3c2_9560_geno_rev_neu	AGGTGCTCCCCGAGATACG
1358	Mapoly0079s0052_sg1_for	CTCGCTTGGGCGATAACTAAACCG
1359	Mapoly0079s0052_sg1_rev	AAACCGGTTTAGTTATCGCCCAAG
1360	Mapoly0079s0052_sg2_for	CTCGTTGTAAACTGGTGCCCAGCA
1361	Mapoly0079s0052_sg2_rev	AAACTGCTGGGCACCAGTTTACAA

Supplement

1362	Mapoly0187s0003_sg1_for	CTCGGGGTAGAGCCACAATCCACT
1363	Mapoly0187s0003_sg1_rev	AAACAGTGGATTGTGGCTCTACCC
1364	Mapoly0187s0003_sg2_for	CTCGGACGCTGGACATGTTTGCGG
1365	Mapoly0187s0003_sg2_rev	AAACCCGCAAACATGTCCAGCGTC
1366	Mapoly0014s0153_sg1_for	CTCGTCGCGCTGAAACCTCGAGCG
1367	Mapoly0014s0153_sg1_rev	AAACCGCTCGAGGTTTCAGCGCGA
1368	Mapoly0014s0153_sg2_for	CTCGCCCGCCTACCTTGGCACGAG
1369	Mapoly0014s0153_sg2_rev	AAACCTCGTGCCAAGGTAGGCGGG
1370	Pp3c2_9560_geno_fwd	TCCAACGGTGATTCTCGCAG
1371	Pp3c2_9560_geno_rev	TCGTCAGGATCCGTTCCAATG
1372	Pp3c2_24980_geno_fwd	TGCTTCCAAGGAGGCGAAAC
1373	Pp3c2_24980_geno_rev	ACTATGCAACTCCATCACTCCAGG
1374	Pp3c7_5970_geno_fwd	TCTGCAGGTGTGAGAGATAACC
1375	Pp3c7_5970_geno_rev	ACATAGTAGAGGAGCAATGAGTGC
1376	Pp3c18_3810_geno_fwd	TGGCTTGAGCTTGAGGAAGTTATG
1377	Pp3c18_3810_geno_rev	TGTATCGTGCCAGTTCACAGC
1335	sgRNA1fwd	GGGGTCTAGATCGAGCTCG
1336	sgRNA1rev	GGGGGCATGCCTCGAG
1337	sgRNA2fwd	GGGGCTCGAGTCTAGATCG
1338	sgRNA2rev	GGGGGCATGCCGACG
1288	Pp3c2_24980V3.1_fwd	ATGGTGAACCCGCGC
1289	Pp3c2_24980V3.1_rev	TCAACTGGGCGCAGGG
1290	Pp3c2_9560V3.1_fwd	ATGCGCAAGGGGGCC
1291	Pp3c2_9560V3.1_rev	TTACAAGTGTTGTGATCTGCTGACCAAC
1292	Pp3c18_3810V3.1_fwd	ATGGTGTGCGTGCGC
1293	Pp3c18_3810V3.1_rev	TCAGCTGATCACCAACGC
1294	Pp3c7_5970V3.1_fwd	ATGGTTTGCATACGGCAAGC
1295	Pp3c7_5970V3.1_rev	CTAAGACTTTGAATCAGCCTTTAAGTCAG
1296	Pp3c2_24980V3.1_fwd_EcoRI	GAATTCATGGTGAACCCGGCG
1297	Pp3c2_24980V3.1_rev_XbaI	TCTAGATGAACTGGGCGCAGGG
1298	Pp3c2_9560V3.1_fwd_EcoRI	GAATTCATGCGCAAGGGGGCC
1299	Pp3c2_9560V3.1_rev_Sall	GTCGACTGACAAGTGTTGTGATCTGCTGACCAAC
1300	Pp3c18_3810V3.1_fwd_EcoRI	GAATTCATGGTGTGCGTGCGC
1301	Pp3c18_3810V3.1_rev_XbaI	TCTAGATGAGCTGATCACCAACGC
1302	Pp3c7_5970V3.1_fwd_EcoRI	GAATTCATGGTTTGCATACGGCAAGC
1303	Pp3c7_5970V3.1_rev_XbaI	TCTAGACGAAGACTTTGAATCAGCCTTTAAGTCAG
1304	Mapoly0079s0052.1_fwd	ATGGCGGACGACAATGAG
1305	Mapoly0079s0052.1_rev	TCAGCCATGCCTTGTCAAATG
1306	Mapoly0187s0003.1_fwd	ATGCTTTCACAAAATTCTGAGTCACTCAC
1307	Mapoly0187s0003.1_rev	CTACTGCCTTCCGCCTC
1308	Mapoly0014s0153.1_fwd	ATGGTGTGCATACGGCAG
1309	Mapoly0014s0153.1_rev	TTACGCAACAGAAGTGGCAAC
1214	seq_rev_pPGX8	TCCTGTTCACGACCATGCTTCG
1210	hpt_outfor	TCCTCTAGAGTCGACCTGCAG
1202	PIG_Locus_rev_neu	AATCCACAGCACATGGACTCATGG
710	pJet_uni_neu	CAACTGCTTTAACTTGTGCCTG

Supplement

711	pJet_rev	GCTGAGAATATTGTAGGAGATCTTCTAG
703	PpPIG1gen_for	CCCATTCTTAGGACACCCCTTTCC
410	M13for	GTAAAACGACGGCCAGT
411	M13rev	GGAAACAGCTATGACCATG
842	Act1_Fw	GTGATGGTCGGTATGGGACAGAAGG
843	Act1_Rv	ATCTTCATGCTCGATGGTGCCAATGC
1223	Hyg-Cassette-fwd	ACTGTCGGGCGTACACAAAT
1092	p35ss_rev	AGGTCCTCTCCAAATGAAATGAAC

B Chemicals and devices

Table 6 Chemicals and supplier.

chemical	supplier
Acetic acid	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Acetone	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Acetosyringone	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Agar Agar Kobe I	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Agarose	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Ampicillin	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Buffer 2.1	New England Biolabs, Frankfurt a.M., DE
Buffer 3.1	New England Biolabs, Frankfurt a.M., DE
Bacto Pepton	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Buffer blue, red, orange, Tango	Thermo Scientific, Dreieich, DE
C ₂ H ₃ KO ₂	Carl Roth GmbH + Co. KG, Karlsruhe, DE
C ₂ H ₃ NaO ₂	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Ca(NO ₃) ₂ x 4 H ₂ O	Carl Roth GmbH + Co. KG, Karlsruhe, DE
CaCl ₂ x 2 H ₂ O	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Carbenicillin	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Casamino Acids	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Casein, enzymatic hydrolase	Sigma-Aldrich Chemie GmbH, Schnelldorf, DE
Cefotaxime Sodium	Duchefa, Haarlem, NL
CH ₃ CO ₂ K	Merck KGaA, Darmstadt, DE
Chloramphenicol	Sigma-Aldrich Chemie GmbH, Schnelldorf, DE
Chloroform/ Isoamylalcohol (Roti® C/I (24:1))	Carl Roth GmbH + Co. KG, Karlsruhe, DE
CTAB (N-Cetyl-N,N,N-trimethylammoniumbromid)	Carl Roth GmbH + Co. KG, Karlsruhe, DE
CuCl ₂	Carl Roth GmbH + Co. KG, Karlsruhe, DE
CutSmart	New England Biolabs, Frankfurt a.M., DE
D (+) Saccharose	Carl Roth GmbH + Co. KG, Karlsruhe, DE
DAPI	Carl Roth GmbH + Co. KG, Karlsruhe, DE

Supplement

DNA blunting enzyme	Thermo Scientific, Dreieich, DE
DNA gel electrophoresis ladder (1kb DNA Ladder)	New England Biolabs, Frankfurt a.M., DE
DNA gel electrophoresis ladder (GeneRuler 100 bp DNA ladder)	Thermo Scientific, Dreieich, DE
DNA polymerase (Q5)	New England Biolabs, Frankfurt a.M., DE
dNTP mix, 10 mM each	New England Biolabs, Frankfurt a.M., DE
Driselase Enzyme Mix	-
DTT (1,4-Dithiothreitol)	Thermo Scientific, Dreieich, DE
EDTA	Sigma-Aldrich Chemie GmbH, Schnelldorf, DE
Ethanol, denatured	Carl Roth GmbH + Co. KG, Karlsruhe, DE
FeSO ₄ x 7 H ₂ O	Carl Roth GmbH + Co. KG, Karlsruhe, DE
First-Strand Buffer (5x)	Thermo Scientific, Dreieich, DE
G418	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Gamborg B5	Sigma-Aldrich Chemie GmbH, Schnelldorf, DE
Gateway LR Clonase™ II Enzyme Mix (5x)	Thermo Scientific, Dreieich, DE
Gateway LR Clonase™ II.	Thermo Scientific, Dreieich, DE
Gel loading dye purple 6x	New England Biolabs, Frankfurt a.M., DE
Glycerol	Carl Roth GmbH + Co. KG, Karlsruhe, DE
H ₂ O (desalted water)	University VE-system
H ₂ O (filtered, desalted water, H ₂ O millipore)	University VE-system
H ₂ O (tap water)	-
HCl	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Hygromycin	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Isopropanol	Carl Roth GmbH + Co. KG, Karlsruhe, DE
K ₂ HPO ₄	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Kanamycin	Carl Roth GmbH + Co. KG, Karlsruhe, DE
KCl	Carl Roth GmbH + Co. KG, Karlsruhe, DE
KH ₂ PO ₄	Carl Roth GmbH + Co. KG, Karlsruhe, DE
KI	Carl Roth GmbH + Co. KG, Karlsruhe, DE
KOH	Carl Roth GmbH + Co. KG, Karlsruhe, DE
L-Glutamine	Carl Roth GmbH + Co. KG, Karlsruhe, DE
LiCl	Merck KGaA, Darmstadt, DE
Liquid nitrogen	Linde AG, Linde Gas Deutschland, Mainz/Kostheim, DE
Mannitol	Carl Roth GmbH + Co. KG, Karlsruhe, DE
MES	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Methanol	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Mg(C ₂ H ₃ O ₂) ₂	Merck KGaA, Darmstadt, DE
MgCl x 6 H ₂ O	Carl Roth GmbH + Co. KG, Karlsruhe, DE
MgCl ₂	Thermo Scientific, Dreieich, DE
MgSO ₄ x 7 H ₂ O	Carl Roth GmbH + Co. KG, Karlsruhe, DE
MnCl ₂ x 4 H ₂ O	Carl Roth GmbH + Co. KG, Karlsruhe, DE

Supplement

MOPS (3-(N-Morpholino)propanesulfonic acid)	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Na ₂ CO ₃	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Na ₂ HPO ₄	Carl Roth GmbH + Co. KG, Karlsruhe, DE
NaCl	Carl Roth GmbH + Co. KG, Karlsruhe, DE
NaClO	Carl Roth GmbH + Co. KG, Karlsruhe, DE
NaHCO ₃	Merck KGaA, Darmstadt, DE
NaOH	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Nicotinic acid	Sigma-Aldrich Chemie GmbH, Schnelldorf, DE
Nonyl-acridine orange (NAO)	Thermo Scientific, Dreieich, DE
OneTaq®-DNA-Polymerase	New England Biolabs, Frankfurt a.M., DE
OneTaq®-Standard-Buffer (5x)	New England Biolabs, Frankfurt a.M., DE
PEG	Carl Roth GmbH + Co. KG, Karlsruhe, DE
peqGREEN	peqLab / VWR International DE GmbH, Darmstadt, DE
Phyto-agar	Duchefa Biochemie, Haarlem, NL
pJet	Thermo Scientific, Dreieich, DE
Proteinase K	Thermo Scientific, Dreieich, DE
Random Primer Mix	New England Biolabs, Frankfurt a.M., DE
RbCl	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Rifampicin	Carl Roth GmbH + Co. KG, Karlsruhe, DE
RNase-Free DNase Set	Qiagen, Hilden, DE
SDS	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Soy peptone	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Spectinomycin	Sigma-Aldrich Chemie GmbH, Schnelldorf, DE
Spermidin	Sigma-Aldrich Chemie GmbH, Schnelldorf, DE
SuperScript™ III	Thermo Scientific, Dreieich, DE
T4 DNA Ligase	New England Biolabs, Frankfurt a.M., DE
T4-Ligase Buffer (10x)	New England Biolabs, Frankfurt a.M., DE
Tris	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Triton-X-100	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Trypton	Carl Roth GmbH + Co. KG, Karlsruhe, DE
Yeast extract	Carl Roth GmbH + Co. KG, Karlsruhe, DE
α-D(+)-Glucosemonohydrate	Carl Roth GmbH + Co. KG, Karlsruhe, DE
β-Mercaptoethanol	Carl Roth GmbH + Co. KG, Karlsruhe, DE

Supplement

Table 7 Materials, devices and supplier.

material	supplier
AQUAline AL 12	LAUDA-Brinkmann, New Jersey, USA
Autoclave	Systec, Linden, DE
Automatic-Sarpette	Sarstedt, Nümbrecht, DE
Binocular Stereomicroscope	Leica Microsystems, Wetzlar, DE
Bioanalyzer 2100	Agilent Technologies, Santa Clara, USA
Biofuge pico	Heraeus Instruments, Mülheim a.d. Ruhr, DE
Biometra®-T-Gradient	Biometra GmbH, Göttingen, DE
Cellophan for petri dishes	Hans Schütt, Halstenbek, DE
Cover slip	Menzel-Gläser, Braunschweig, DE
Cryo boxes	Sarstedt, Nümbrecht, DE
DFC295 camera	Leica Microsystems, Wetzlar, DE
DMIL LED	Leica Microsystems, Wetzlar, DE
Dry-Block DB2D	Techne, Staffordshire, United Kingdom
ECO2 box	Duchefa Biochemie, Haarlem, NL
Electrophoresis Power Supply - EPS 301	Amersham Pharmacia Biotech
Elektroporator GenePulser / GenePulser Controller	BioRad, München, DE
Environmental Shaker Incubator ES20	Biosan Ltd., Riga, Latvia
Eppendorf Research plus 0,1-2,5 µl	Eppendorf, Hamburg, DE
Eppendorf Research plus 100-1000 µl	Eppendorf, Hamburg, DE
Eppendorf Research plus 20-200 µl	Eppendorf, Hamburg, DE
Extend ED 822-CW weighing scale	Sartorius Stedim Biotech GmbH, Göttingen, DE
Falcon tubes	Sarstedt, Nümbrecht, DE
Faster KBN	Thermo Scientific, Schwerte, DE
Filter Cube I3	Leica Microsystems, Wetzlar, DE
Filter Cube L5	Leica Microsystems, Wetzlar, DE
Filter Cube N2.1	Leica Microsystems, Wetzlar, DE
Fluorescence stereomicroscope SteREO LumarV12	Carl Zeiss, Oberkochen, DE
Forceps	Dumont, DE
Fuchs-Rosenthal counting chamber	Marienfeld Superior, Lauda-Königshofen, DE
Galaxy MiniStar	VWR, Erlangen, DE
Gene Pulser®/MicroPulser™ Electroporation Cuvettes, 0,2 cm gap	BioRad, München, DE
Laminar flow cabinet Bio CL190	EHRET Labor, Emmendingen, DE
Laminar flow cabinet H-130	EHRET Labor, Emmendingen, DE
Laminar flow cabinet V-160	EHRET Labor, Emmendingen, DE
Micropore 3M tape	WiSMa GmbH, Freiburg, DE
Microscope DM6000 CS	Leica Microsystems, Wetzlar, DE
Microscope slides	Carl Roth, Karlsruhe, DE
Microwave Fairline MW 2717	Goldhand GmbH, DE

Supplement

Mikro 200R	Hettich Zentrifugen, Tuttlingen, DE
Mikro 22R	Hettich Zentrifugen, Tuttlingen, DE
Minisart sterile filter	Sartorius Stedim Biotech GmbH, Göttingen, DE
MLR-352H	Panasonic Healthcare Co., Ltd., Ora-Gun, Gunma, Japan
MLR-352-PE	Panasonic Healthcare Co., Ltd., Ora-Gun, Gunma, Japan
NanoDrop Spectrophotometer ND-1000	Thermo Scientific, Schwerte, DE
Objectives	Leica Microsystems, Wetzlar, DE
Parafilm	VWR, Erlangen, DE
PCR Cycler PEQ Star 96 Universal Gradient	Peqlab Biotechnologie GmbH, Erlangen, DE
PCR tubes	Sarstedt, Nümbrecht, DE
Petri dishes	Sarstedt, Nümbrecht, DE
pH meter 766 Calimatic	Knick GmbH & Co. KG, Berlin, DE
Phoenix II	Schuett-Biotec, Göttingen, DE
Photometer Ultrospec 2000	Pharmacia, Uppsala, SWE
Pipette tips	Sarstedt, Nümbrecht, DE
Reaction tubes	Sarstedt, Nümbrecht, DE
Rotina 380 R	Hettich Zentrifugen, Tuttlingen, DE
S6D binocular	Leica Microsystems, Wetzlar, DE
S8Apo binocular	Leica Microsystems, Wetzlar, DE
Serological pipettes	Sarstedt, Nümbrecht, DE
Thermomixer Comfort	Eppendorf, Hamburg, DE
Transilluminator	Peqlab Biotechnologie GmbH, Erlangen, DE
Tumble shaker	Otto E. Kobe KG, Marburg, DE
Ultra-Turrax	VWR, Erlangen, DE StepOnePlus™ Real-Time PCR System
UV Illuminator	Nippon Genetics Europe GmbH, Düren, DE
Vacuum Concentrator, "Speedvac"	BACHOFER GmbH Laboratoriumsgeräte, Reutlingen, DE
Vortex Genie 2	Scientific Industries, Karlsruhe, DE

Supplement

C additional figures, that were not shown in the main part

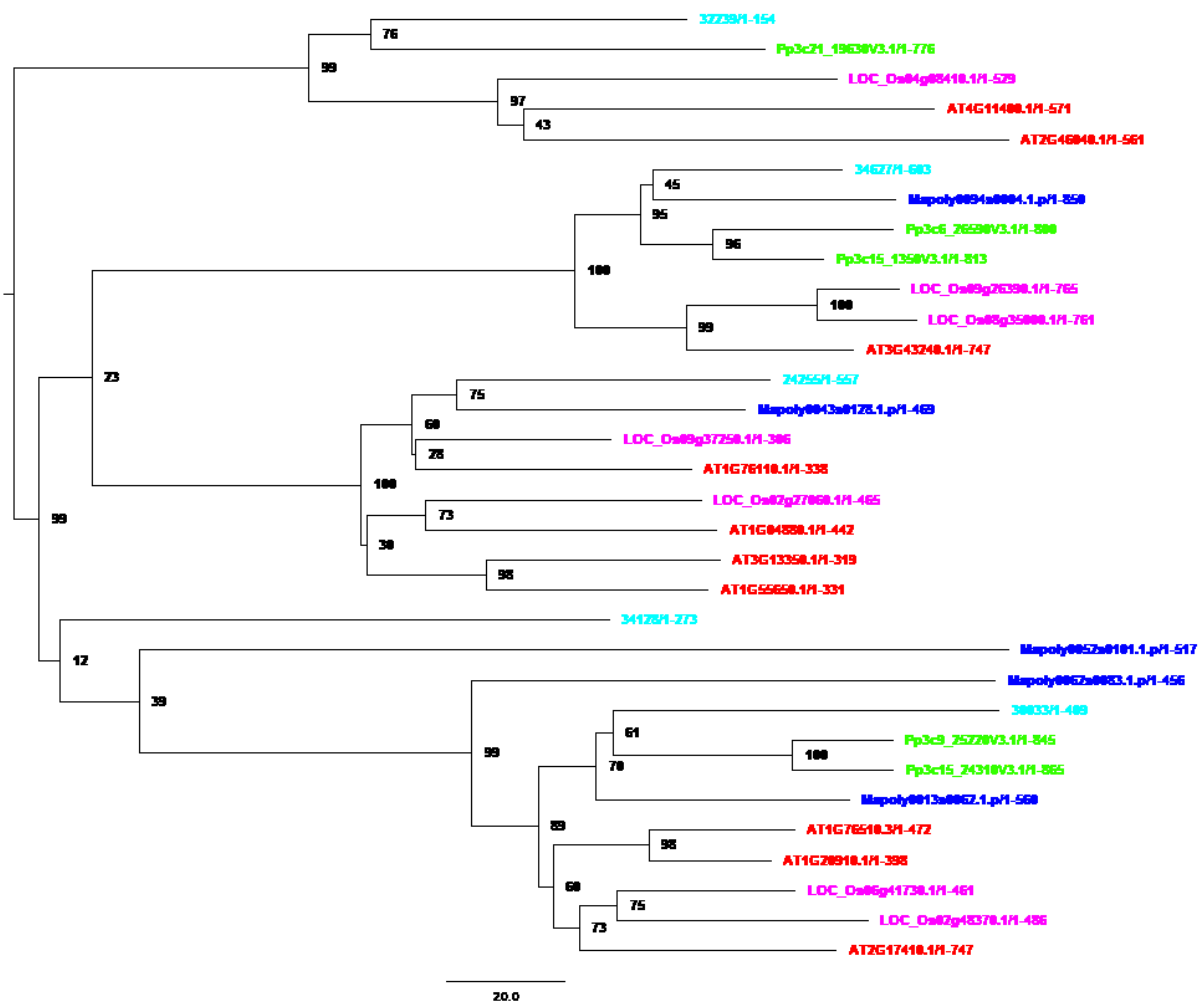


Figure 44 ARID TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.



Figure 45 GRAS TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

Supplement

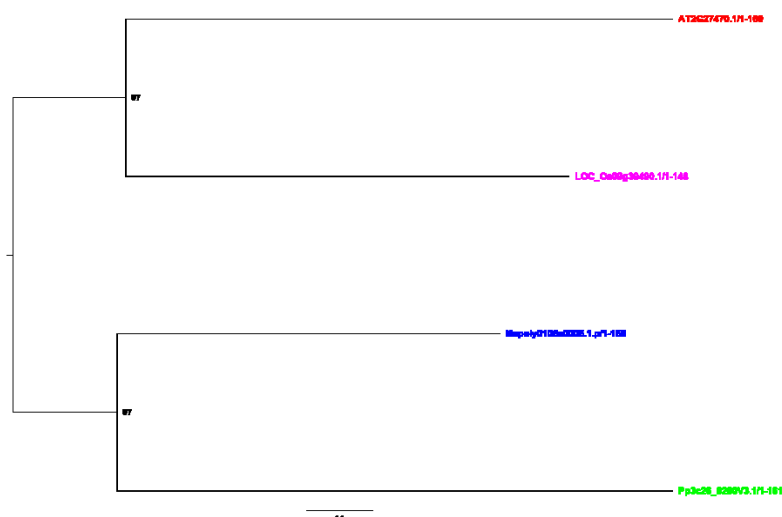


Figure 46 CCAAT_HAP3 TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*). Support values at the nodes represent bootstrap values.

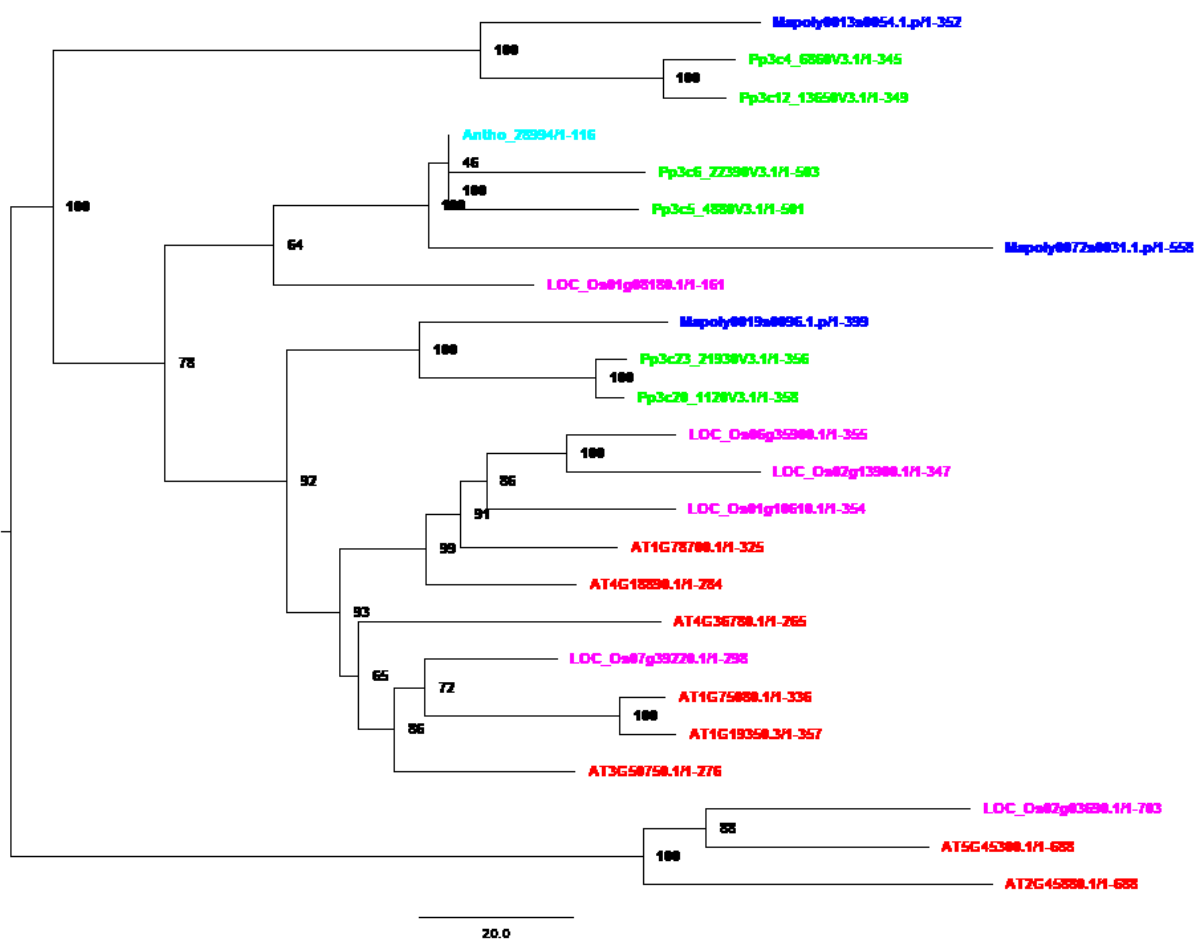


Figure 47 BES1 TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

Supplement

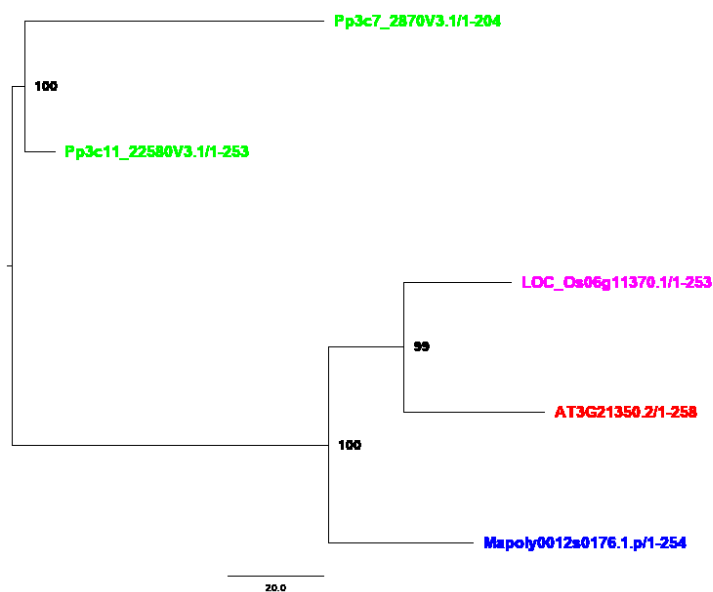


Figure 48 Med6 TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*). Support values at the nodes represent bootstrap values.

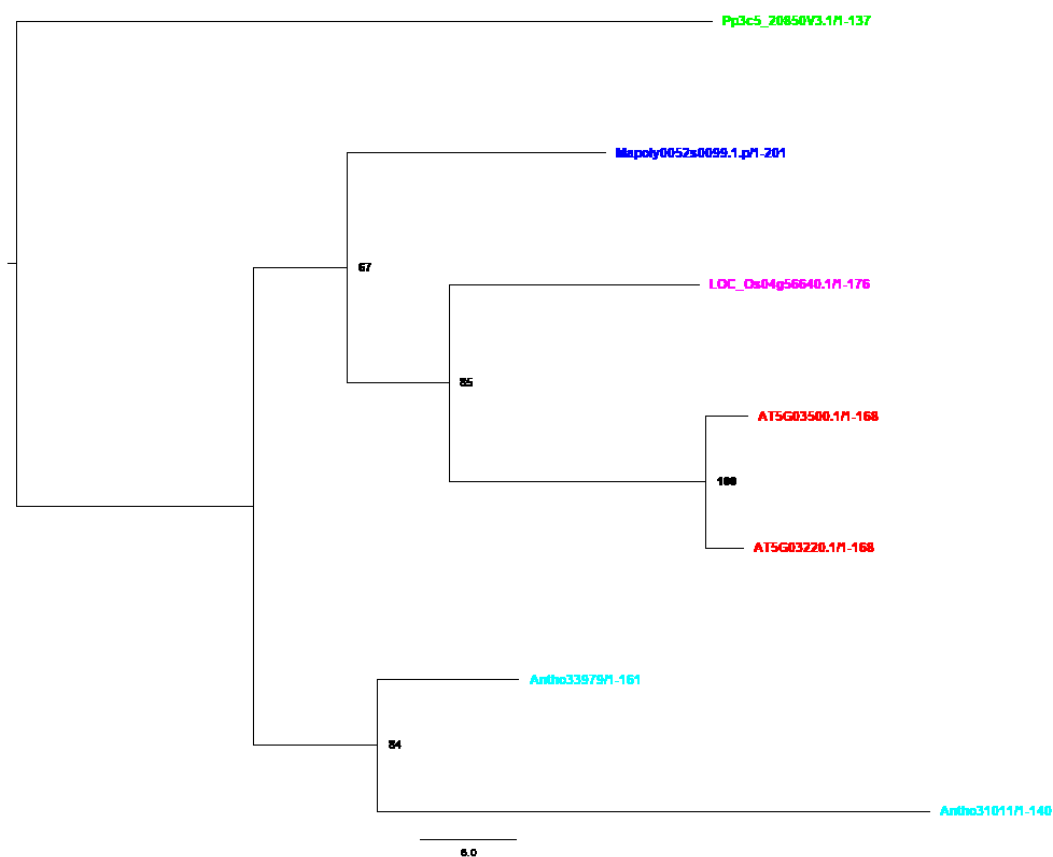


Figure 49 Med7 TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

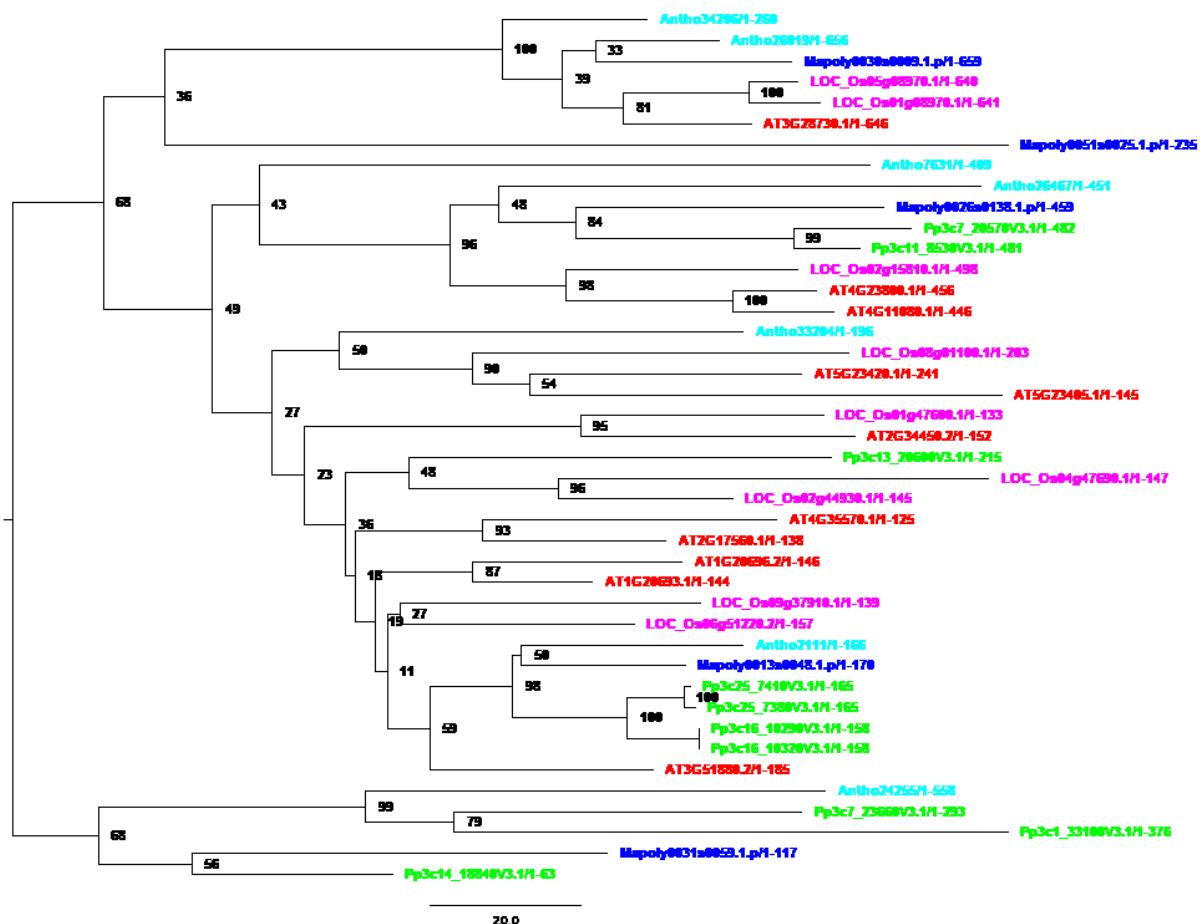


Figure 50 HMG TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

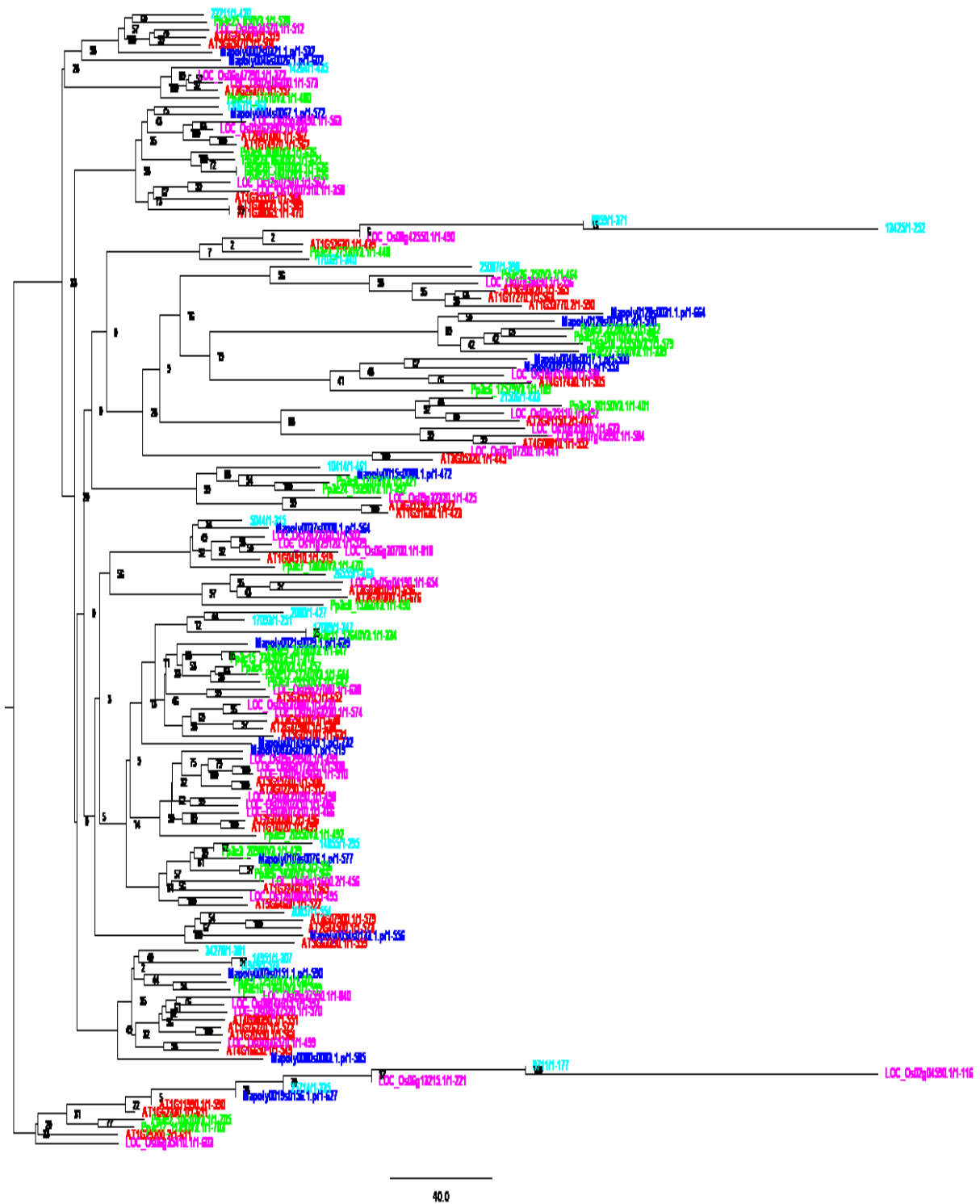


Figure 51 O-FucT TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values. Size of phylogenetic analysis limits resolution.

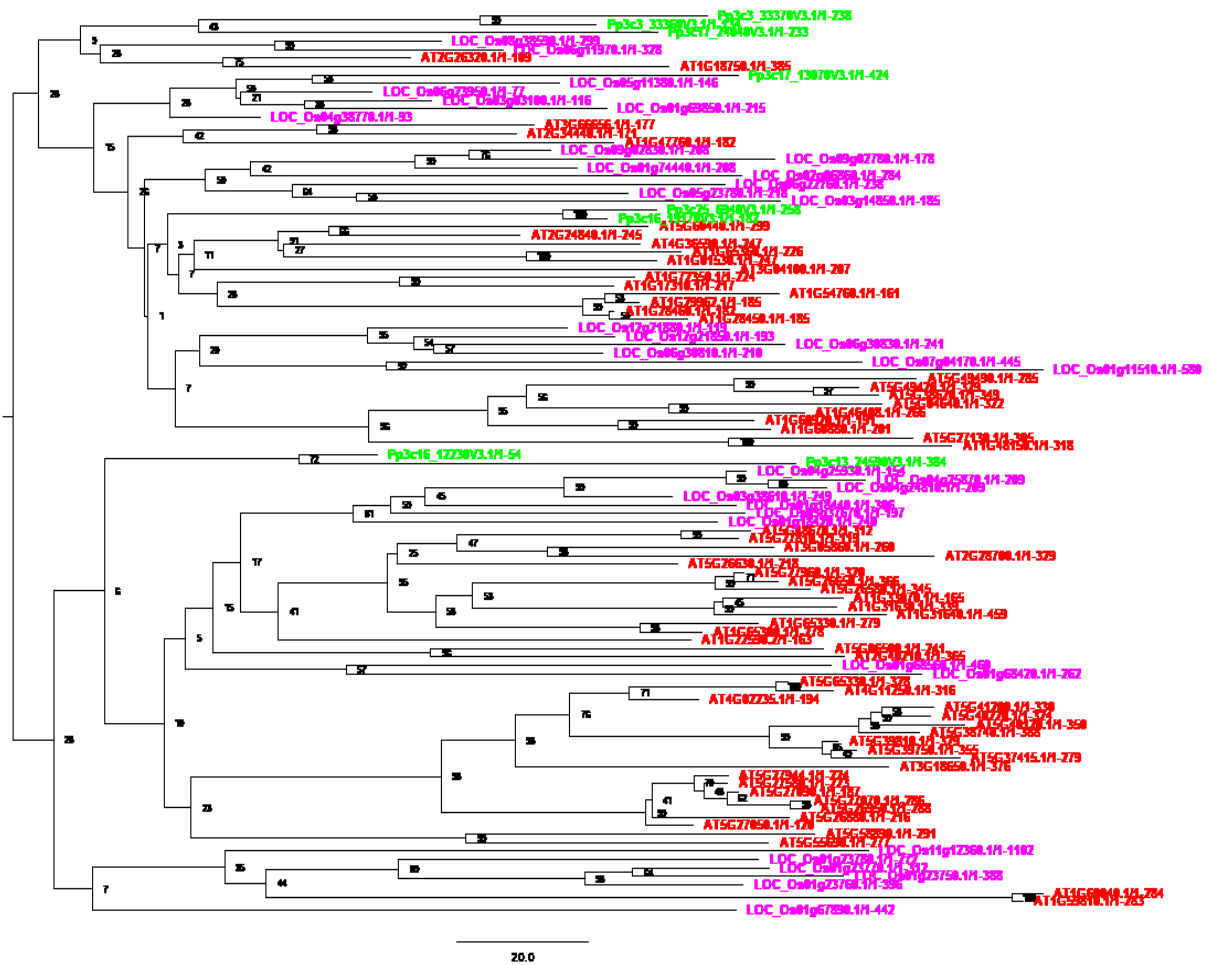


Figure 52 MADS TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *A. thaliana*, *O. sativa*, *P. patens*).

Support values at the nodes represent bootstrap values.

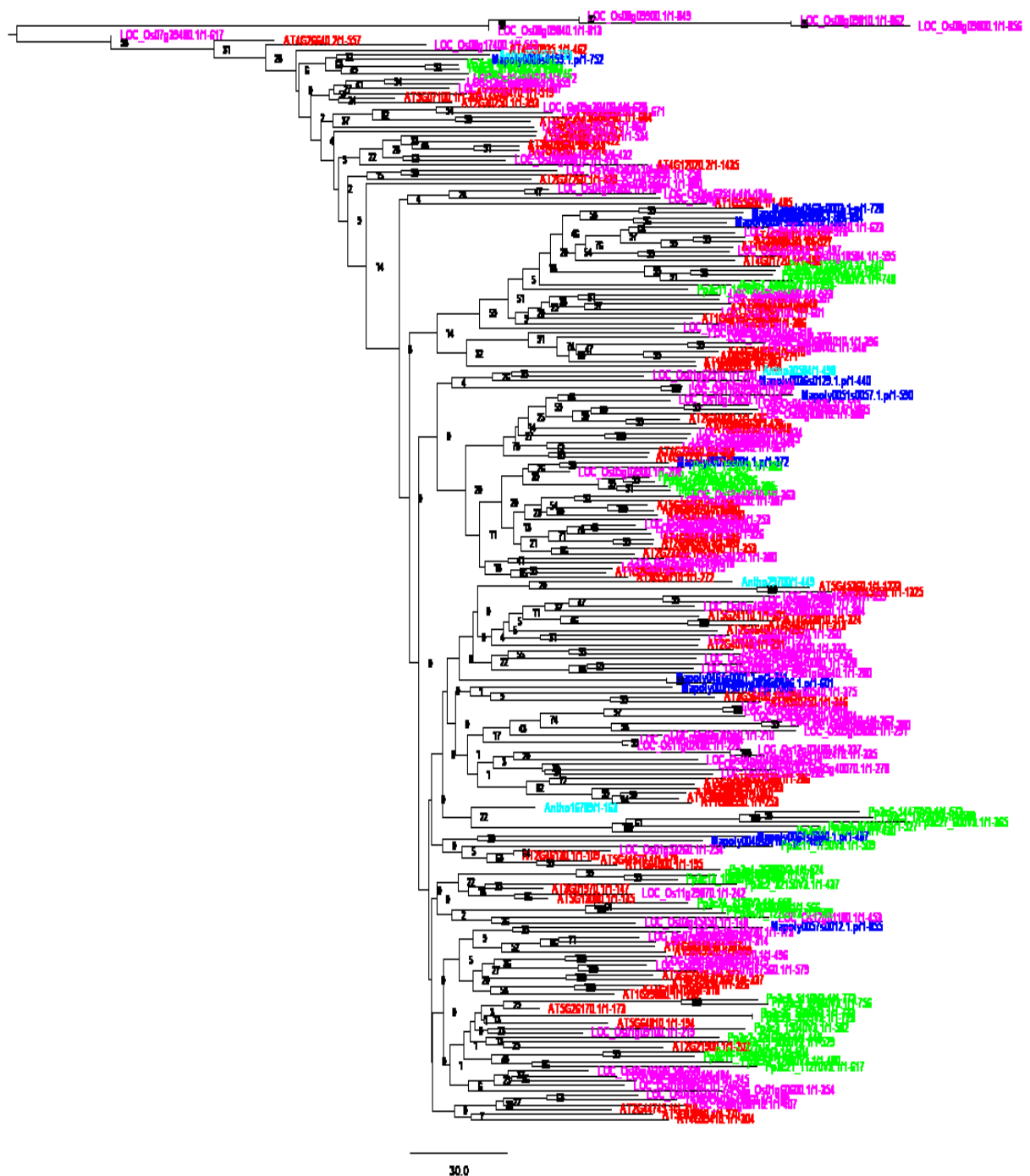


Figure 53 WRKY TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values. Size of phylogenetic analysis limits resolution.

Supplement

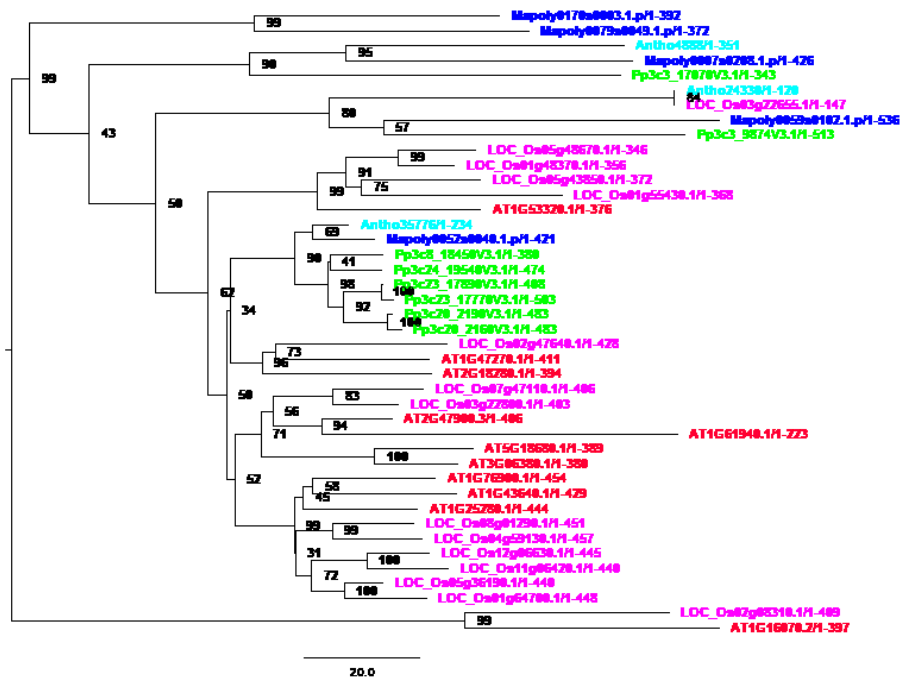


Figure 54 TUB TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

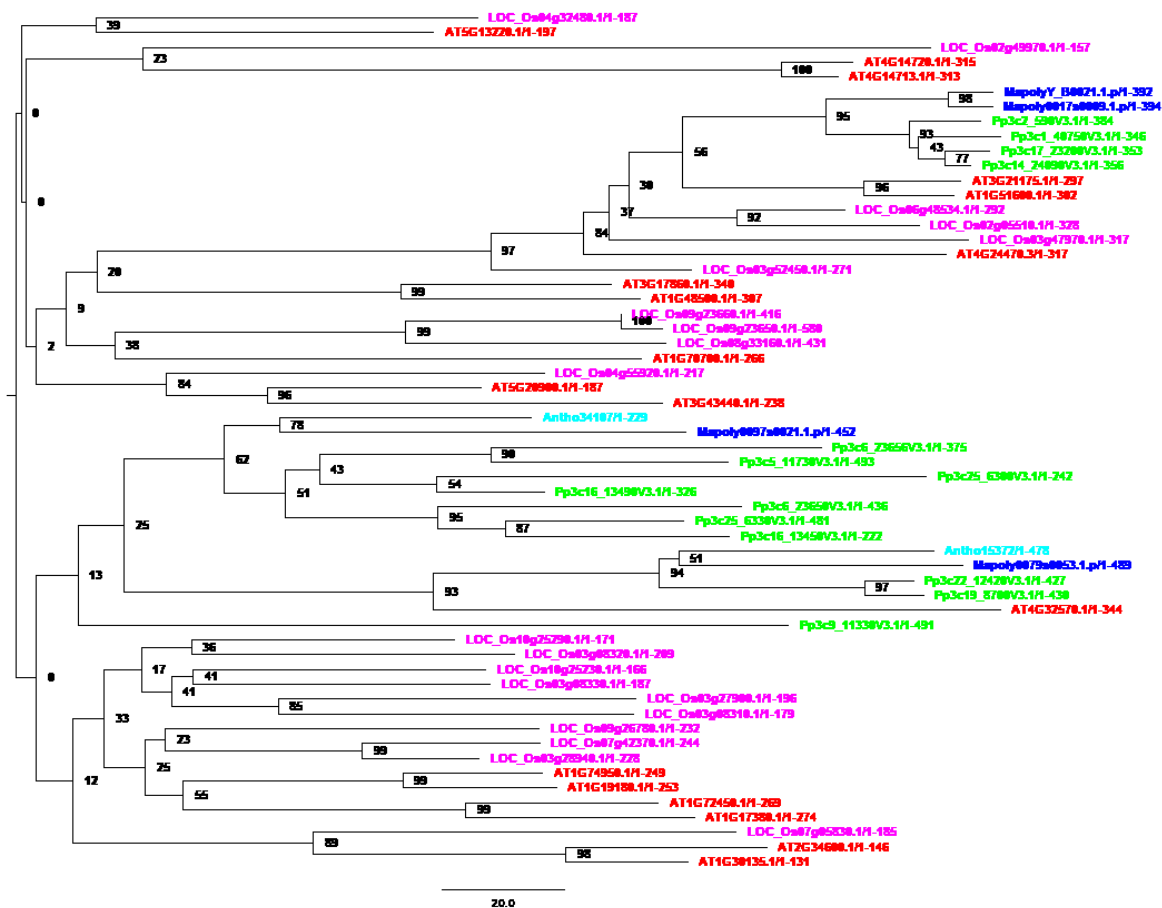


Figure 55 tify TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

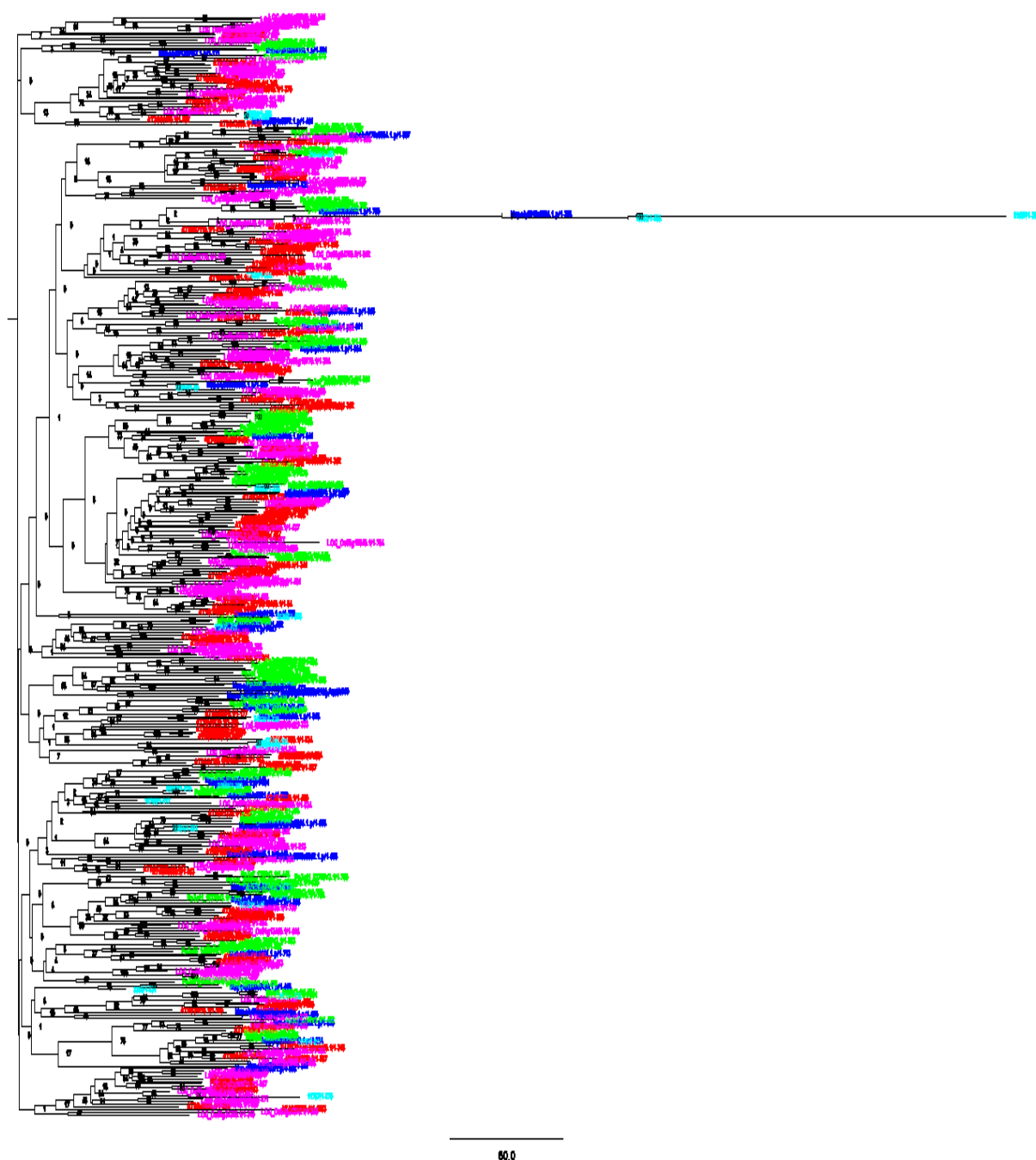


Figure 56 bHLH TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values. Size of phylogenetic analysis limits resolution.

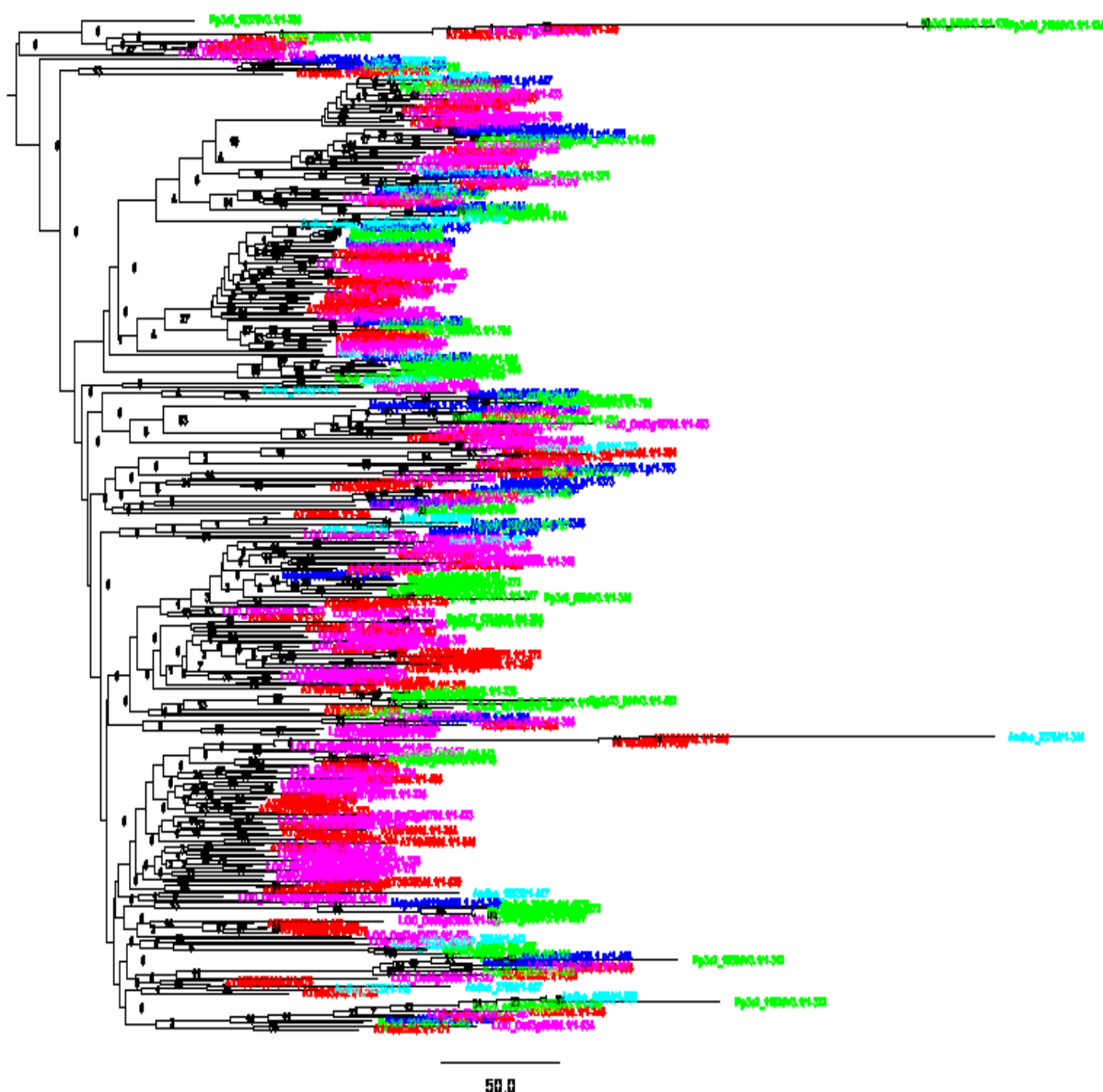


Figure 57 C2H2 TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values. Size of phylogenetic analysis limits resolution.

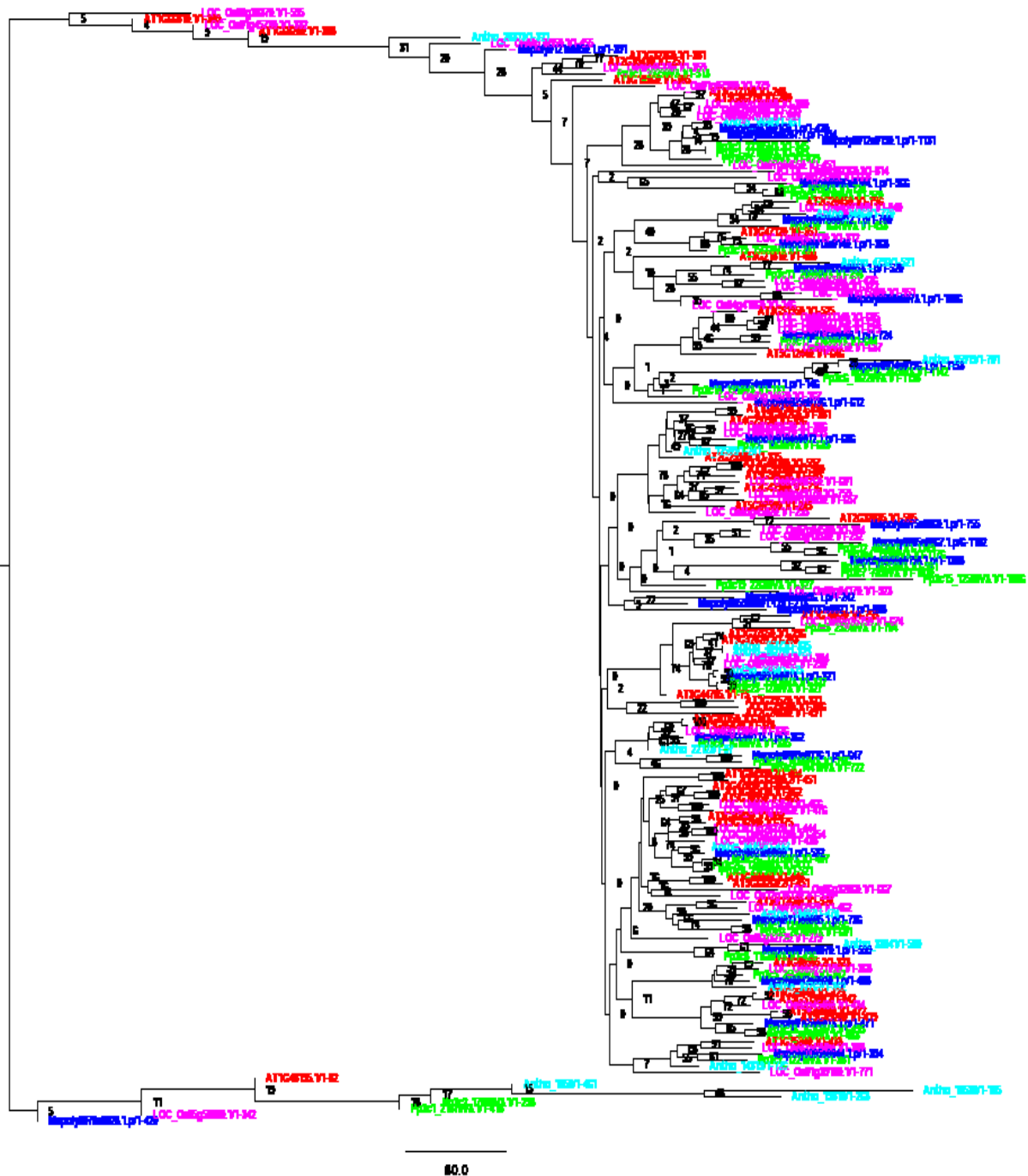


Figure 58 C3H TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values. Size of phylogenetic analysis limits resolution.

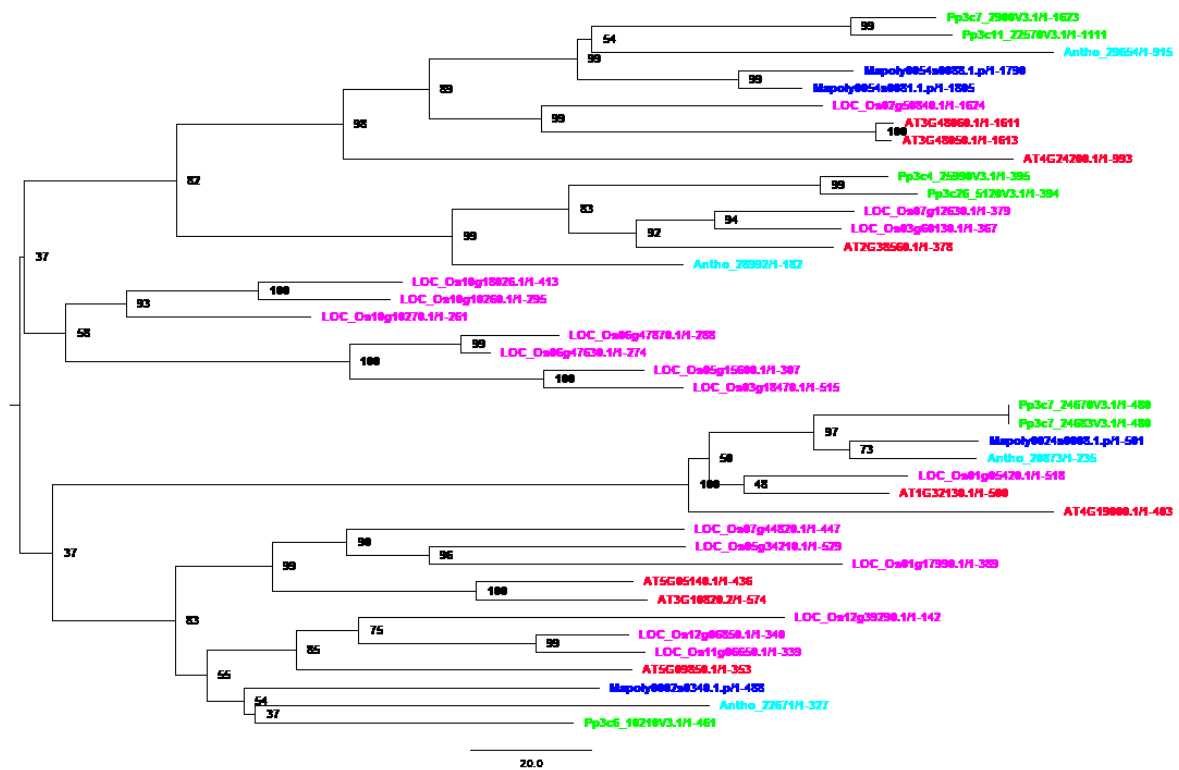


Figure 59 IWS1 TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

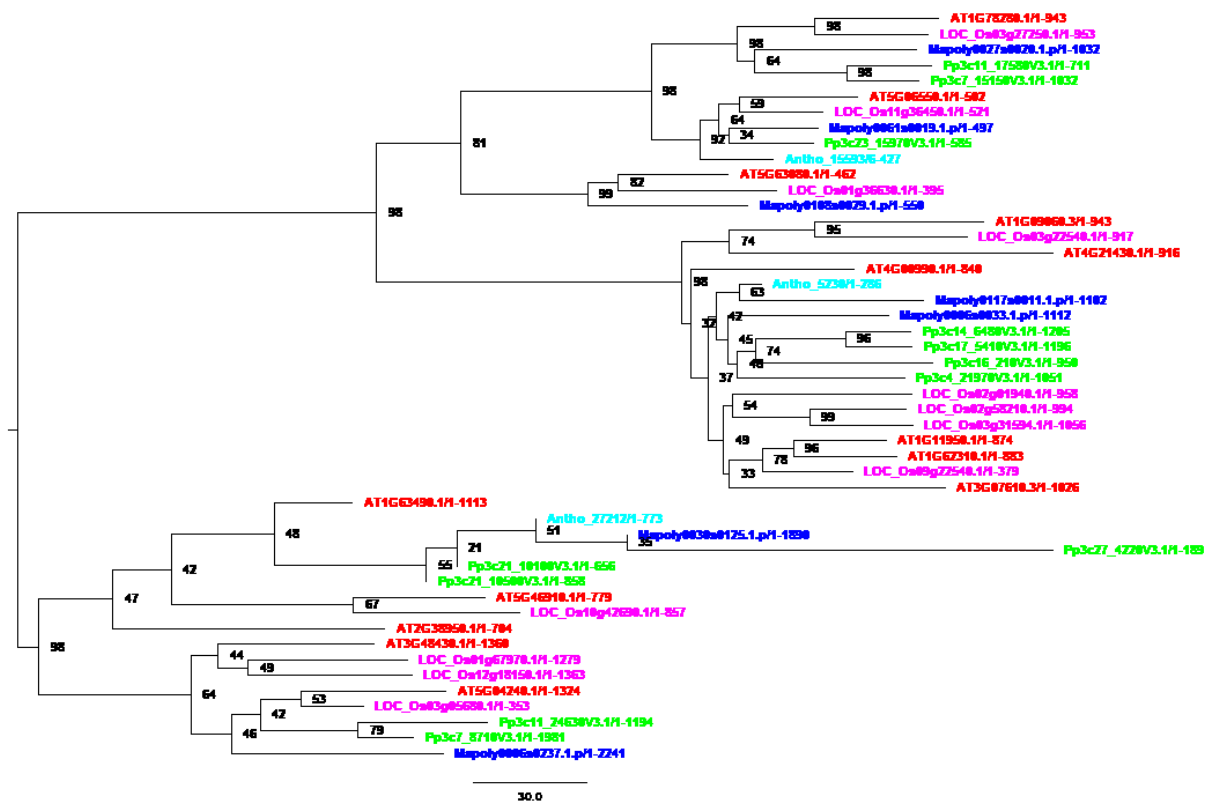


Figure 60 Jumonji_Other TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

Supplement

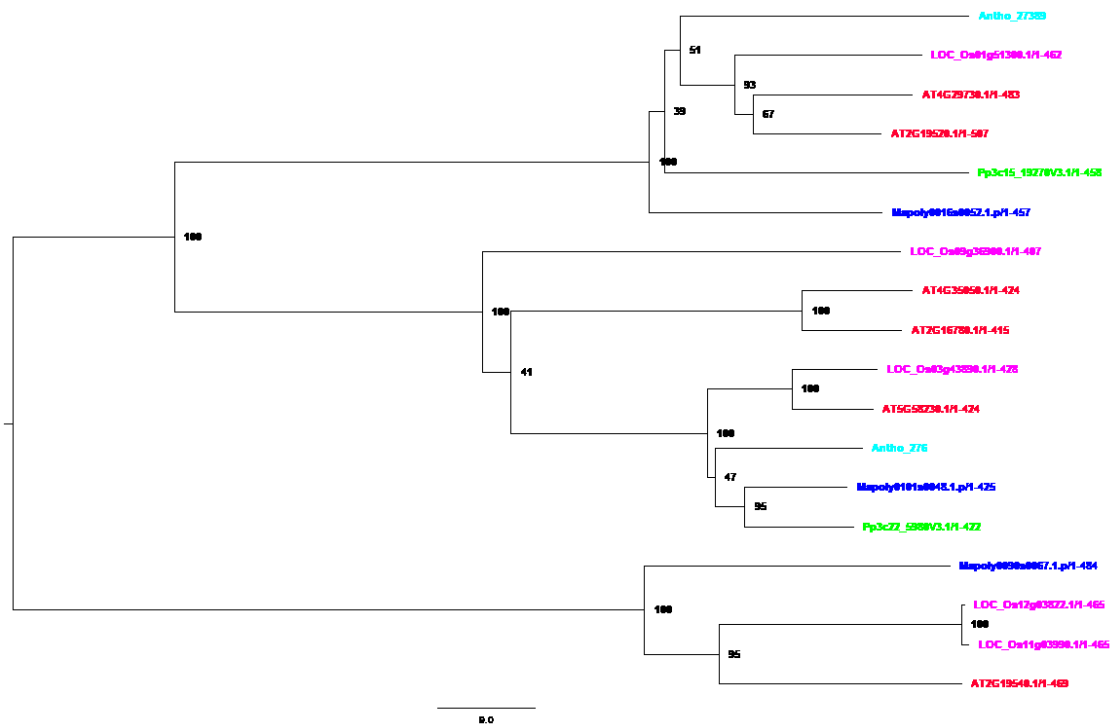


Figure 61 PcG_MSI TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

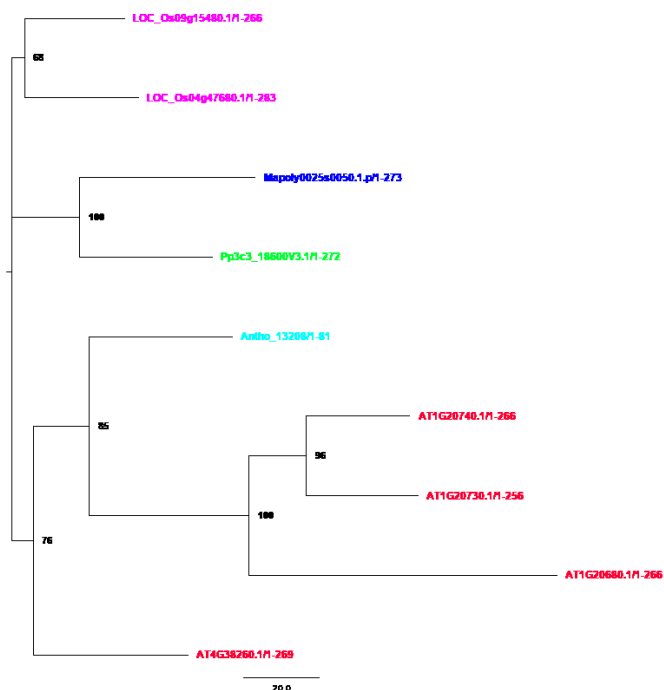


Figure 62 TANGO2 TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

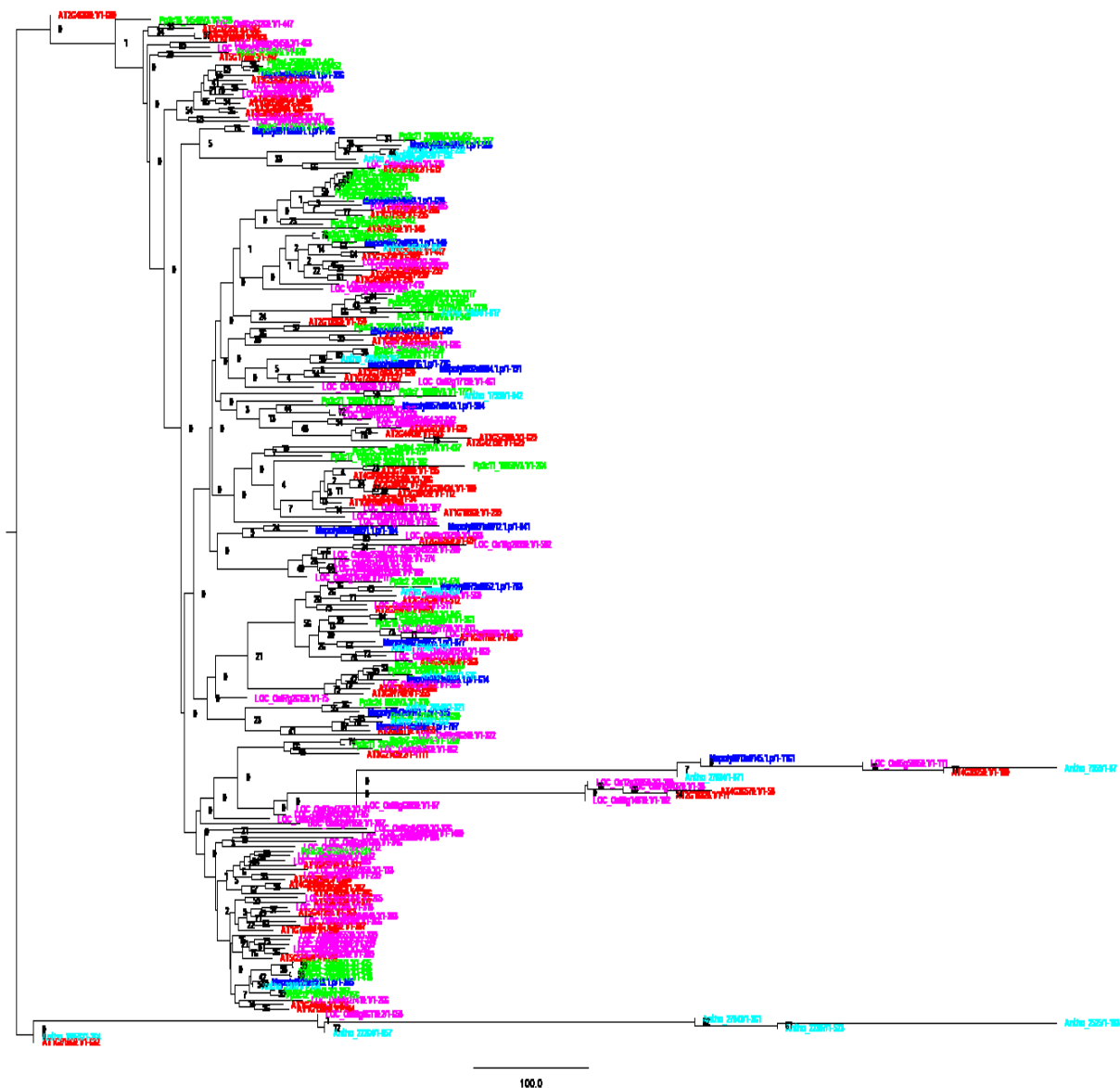


Figure 63 MYB-related TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values. Size of phylogenetic analysis limits resolution.

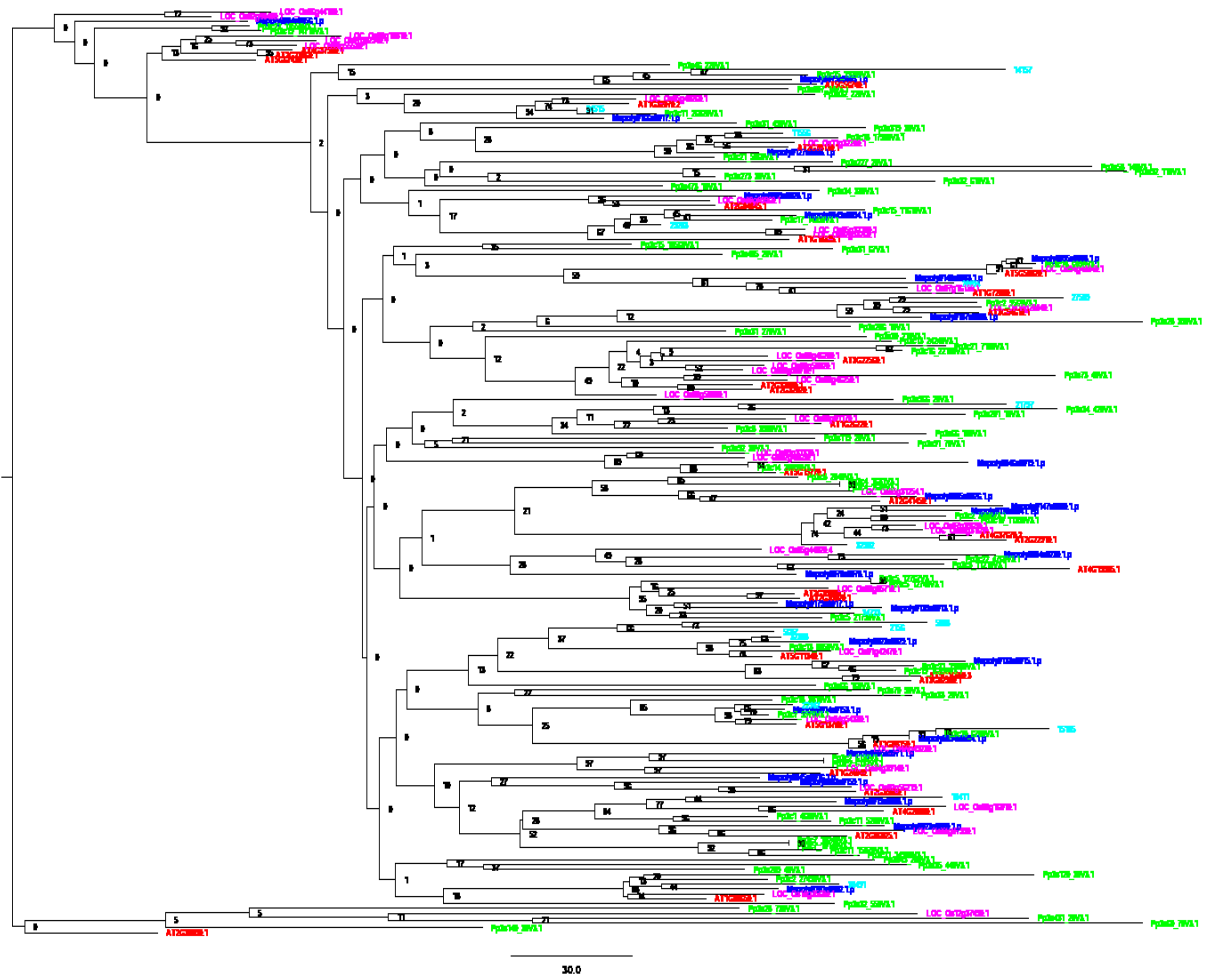


Figure 64 GNAT TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

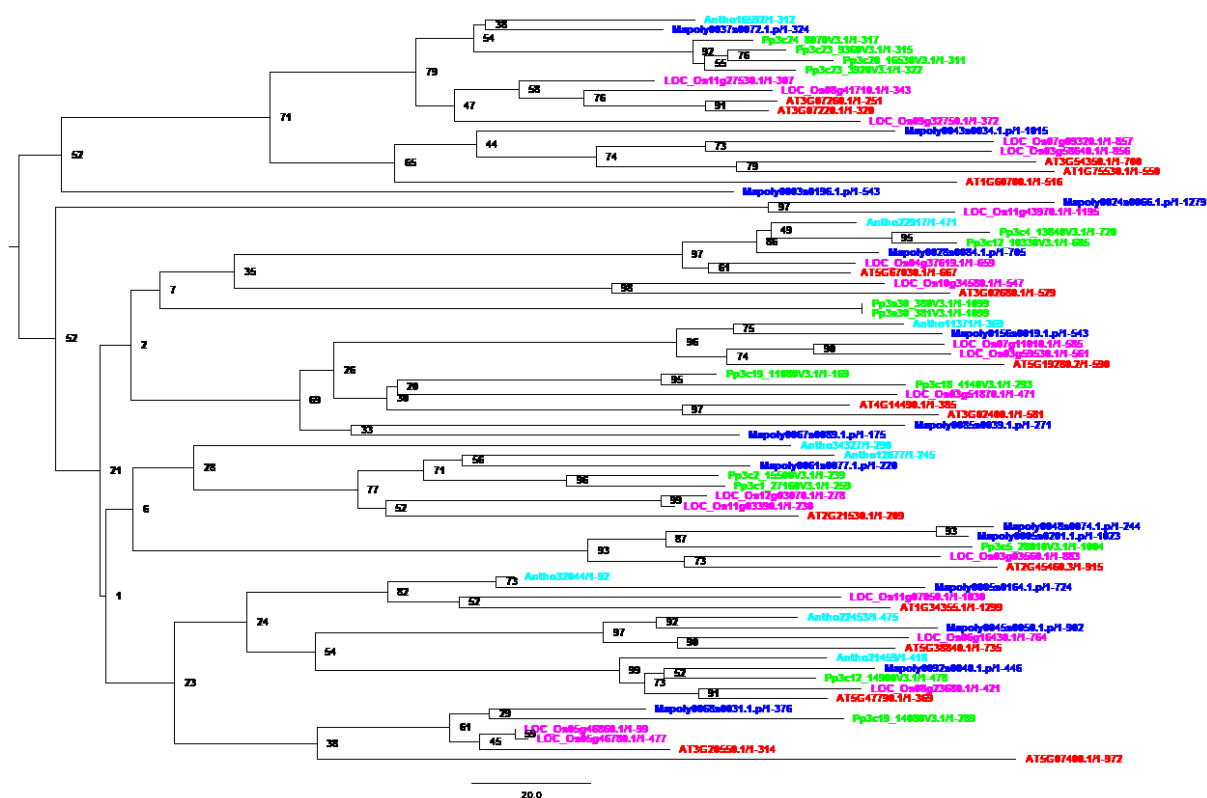


Figure 65 FHA TAP family phylogenetic analysis.

Colour code was used according to the colour code in the main part (species: *M. polymorpha*, *A. thaliana*, *O. sativa*, *P. patens*, *A. agrestis*). Support values at the nodes represent bootstrap values.

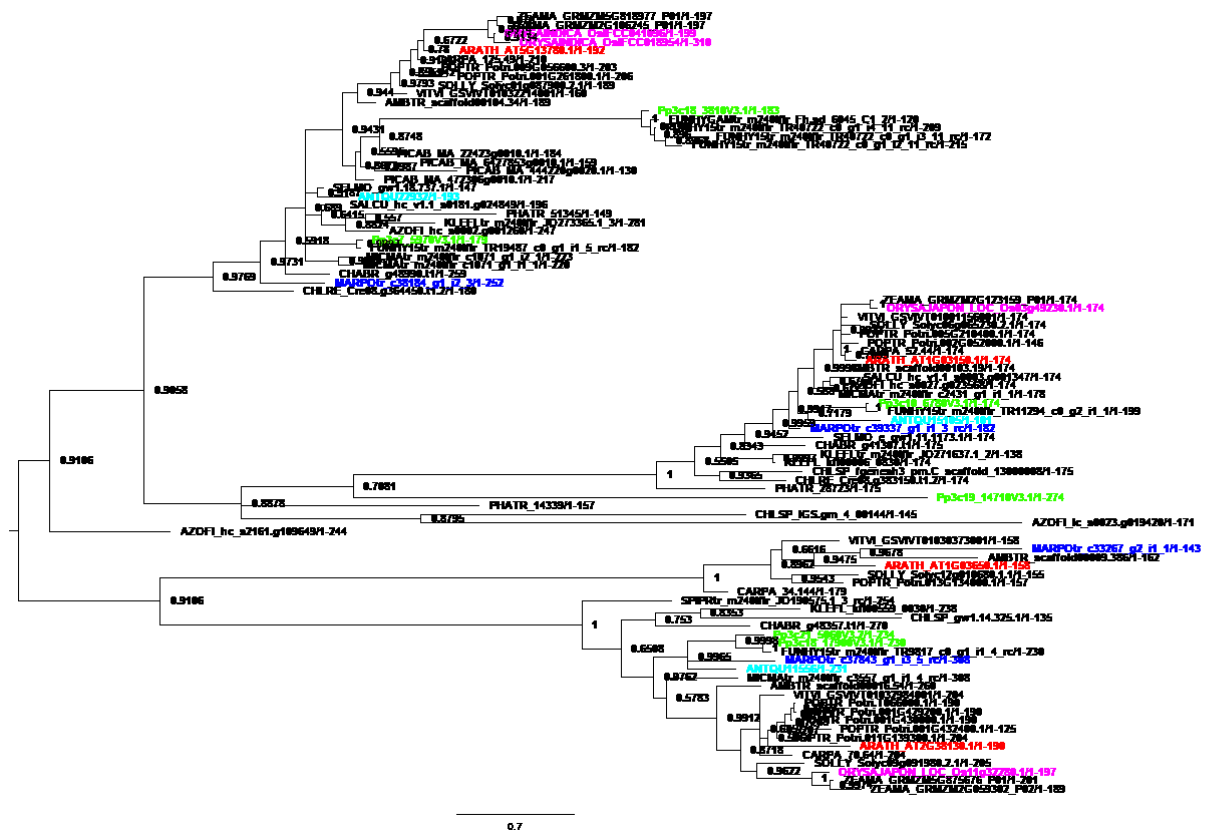


Figure 66 AT5G13780 phylogenetic analysis.

Phylogenetic analysis named after the *A. thaliana* blastp input query. Colour code was used according to the colour code in the main part. Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in the before mentioned five letter code, additional species not mentioned before: CHLSP = *Chlorella* sp. NC64A, FUNHY = *Funaria hygrometrica*, PHATR = *Phaeodactylum tricornutum*, SPIPR = *Spirogyra pratensis*.

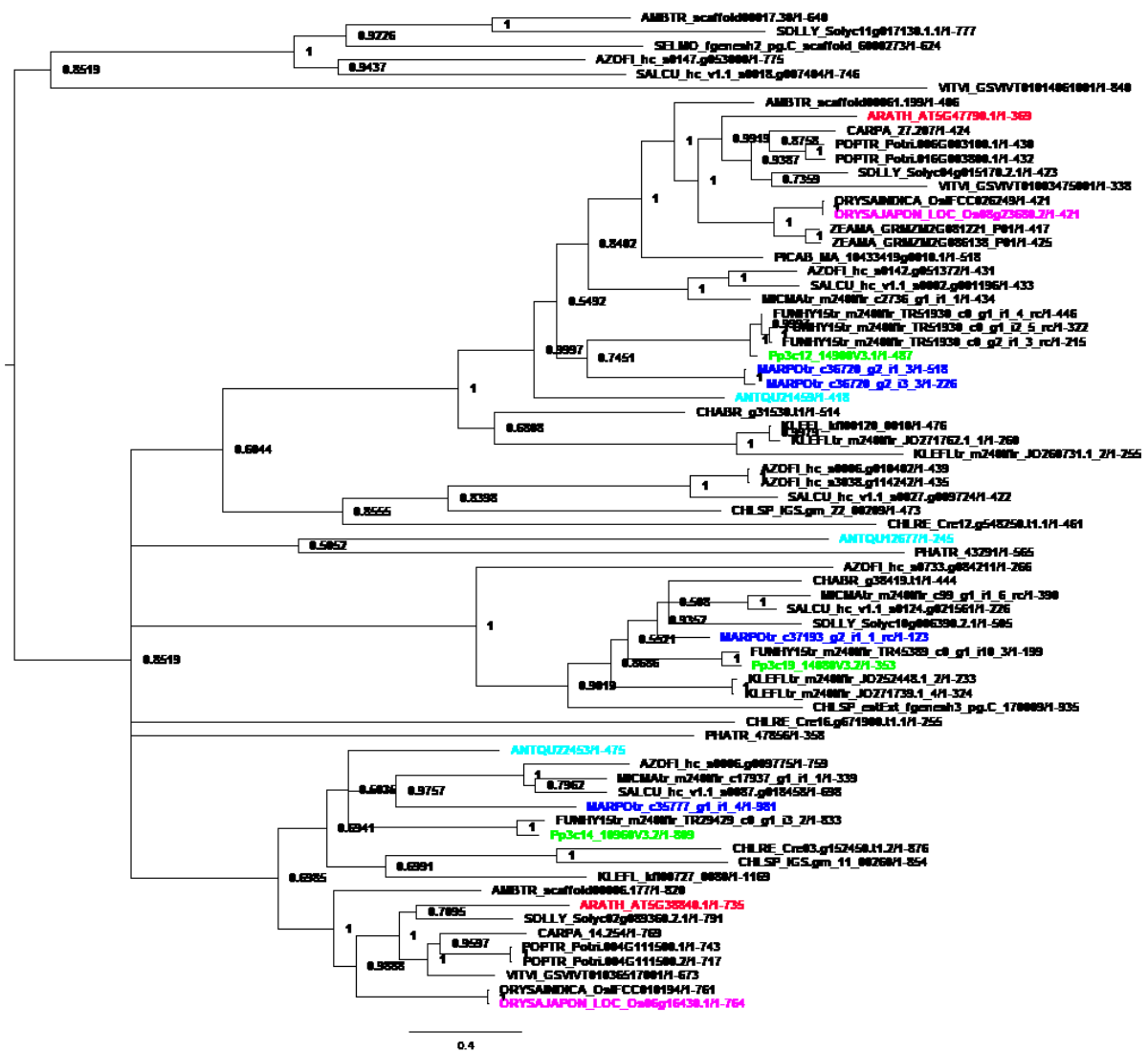


Figure 67 AT5G47790 phylogenetic analysis.

Phylogenetic analysis named after the *A. thaliana* blastp input query. Colour code was used according to the colour code in the main part. Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in the before mentioned five letter code.

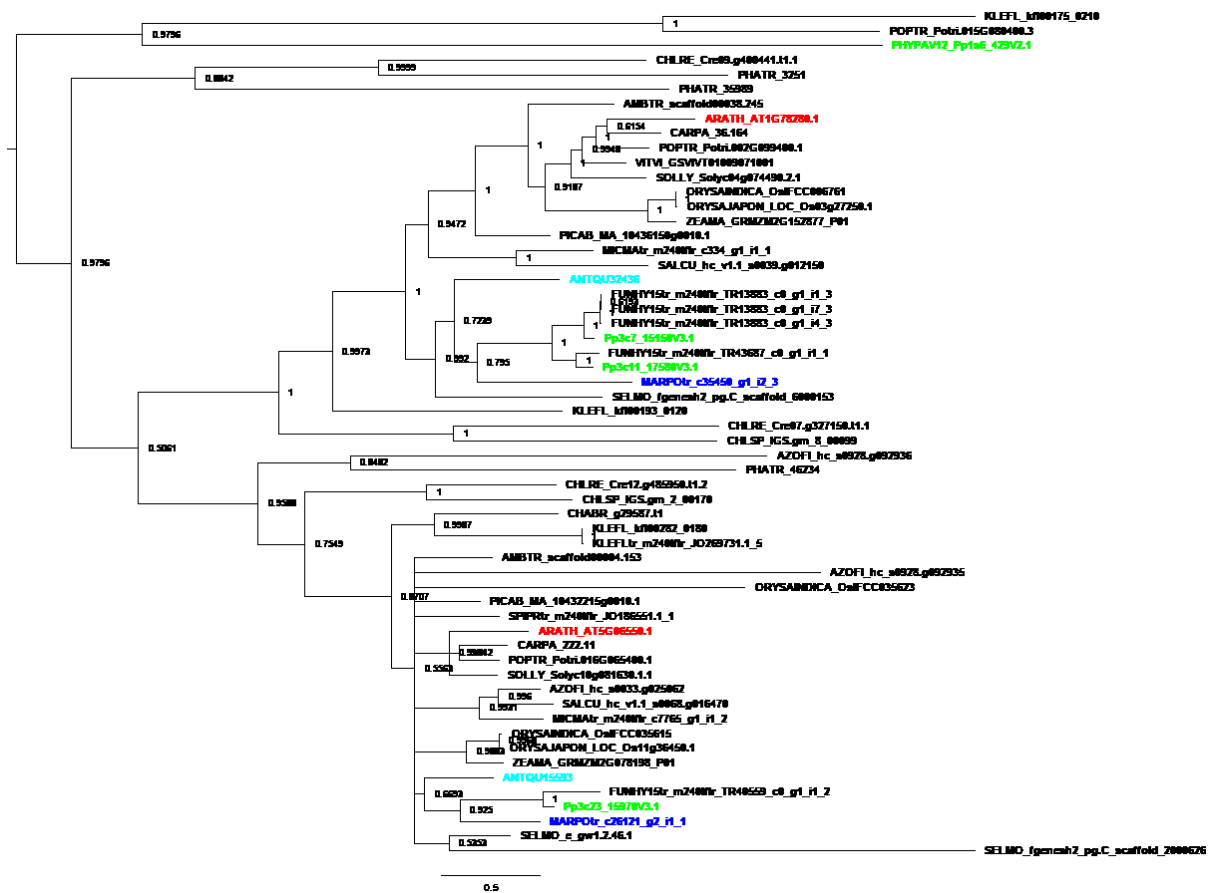


Figure 68 AT5G06550 phylogenetic analysis.

Phylogenetic analysis named after the *A. thaliana* blastp input query. Colour code was used according to the colour code in the main part. Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in the before mentioned five letter code.

Supplement

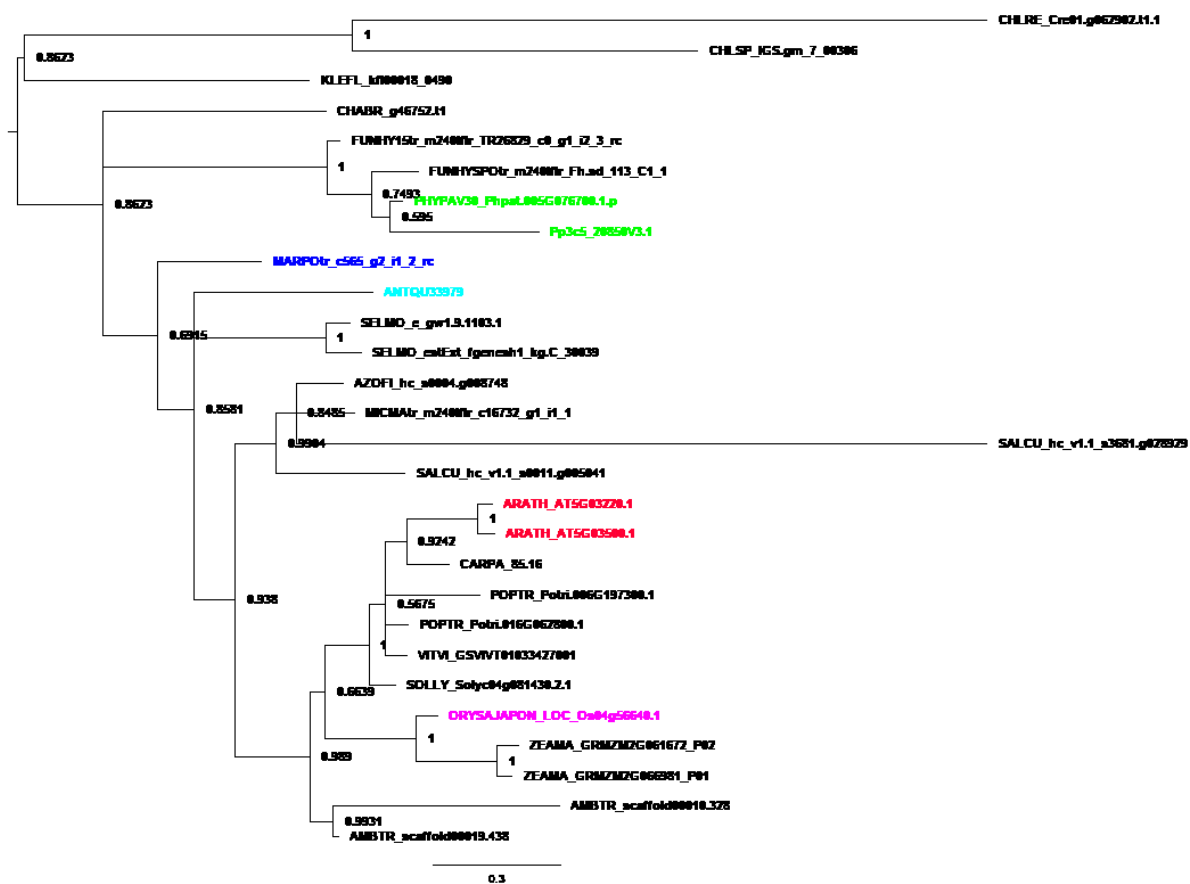


Figure 70 AT5G03500 phylogenetic analysis.

Phylogenetic analysis named after the *A. thaliana* blastp input query. Colour code was used according to the colour code in the main part. Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in the before mentioned five letter code.

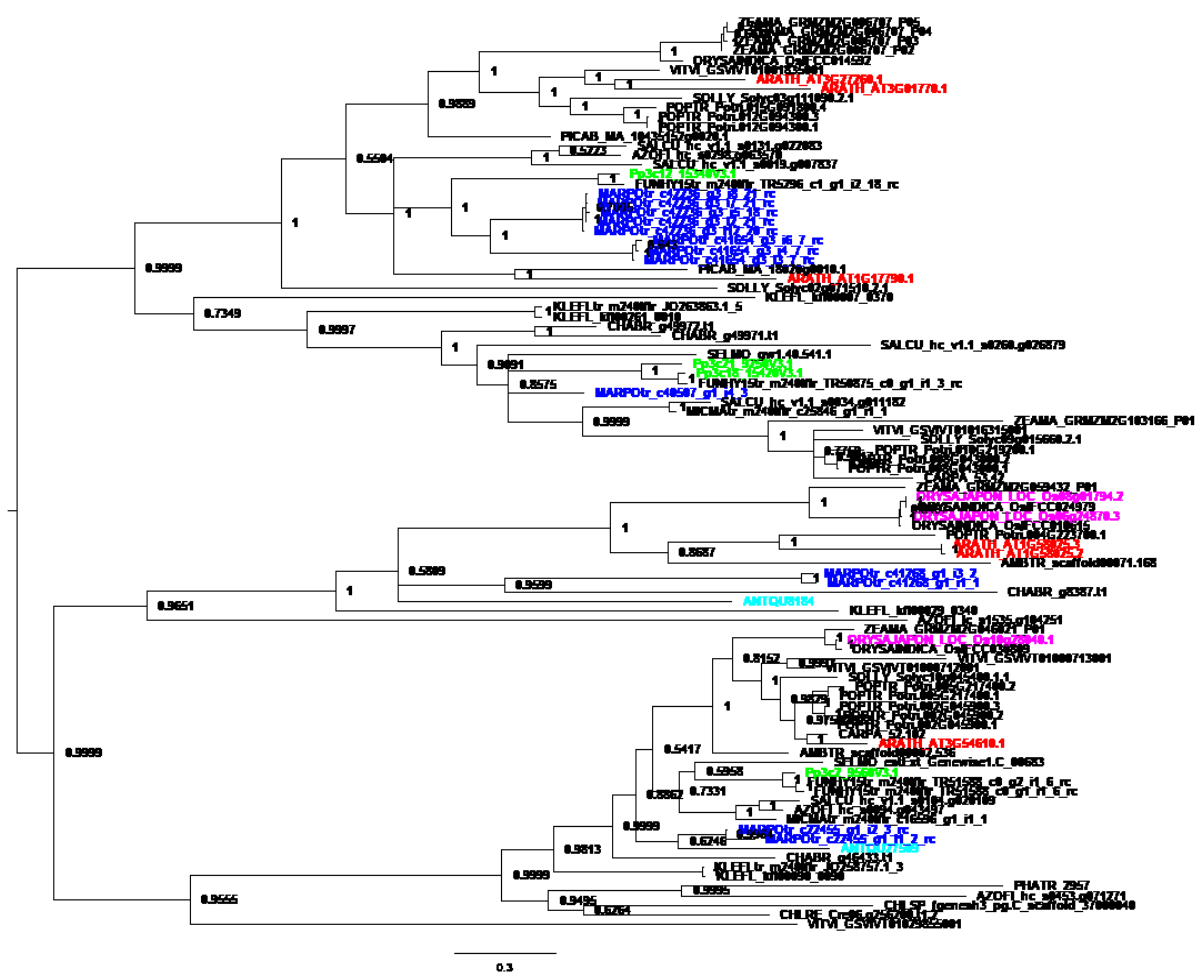


Figure 71 AT3G54610 phylogenetic analysis.

Phylogenetic analysis named after the *A. thaliana* blastp input query. Colour code was used according to the colour code in the main part. Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in the before mentioned five letter code.

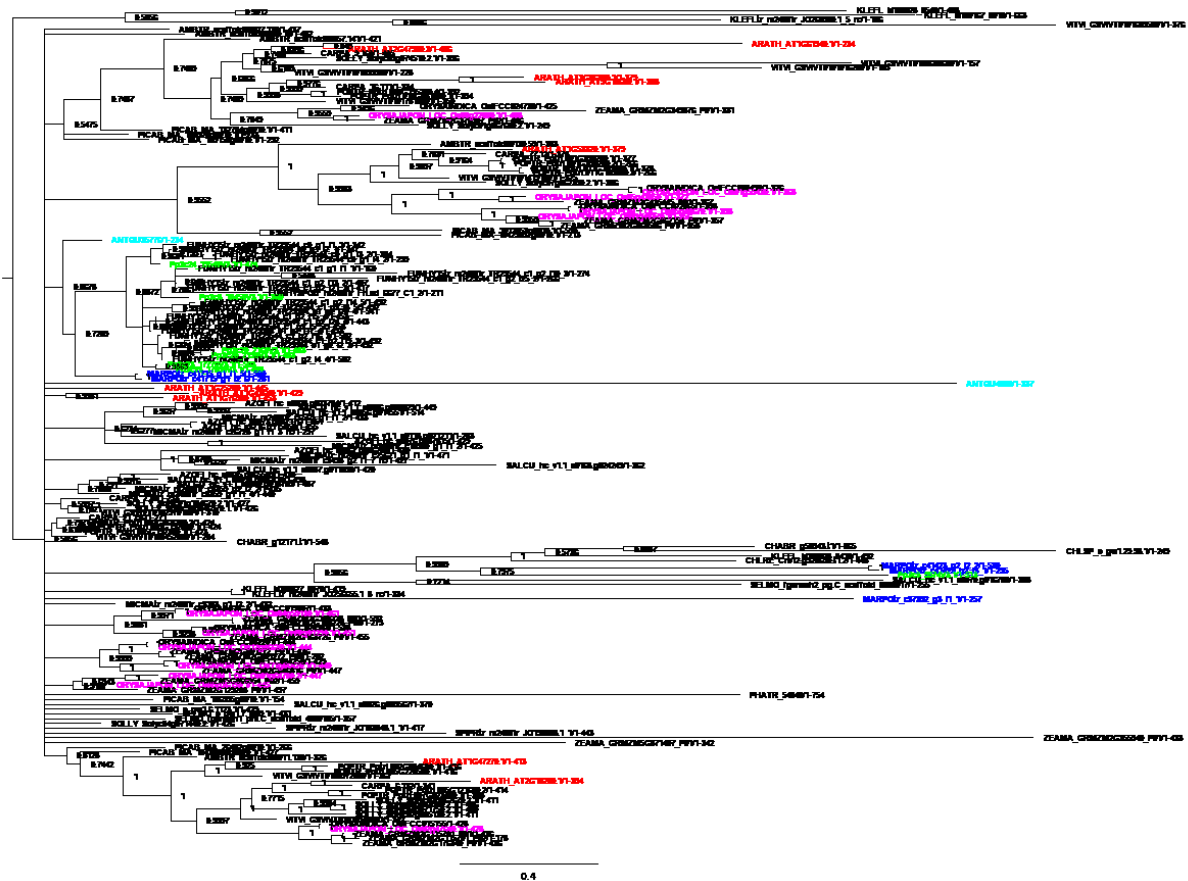


Figure 72 AT2G47900 phylogenetic analysis.

Phylogenetic analysis named after the *A. thaliana* blastp input query. Colour code was used according to the colour code in the main part. Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in the before mentioned five letter code.

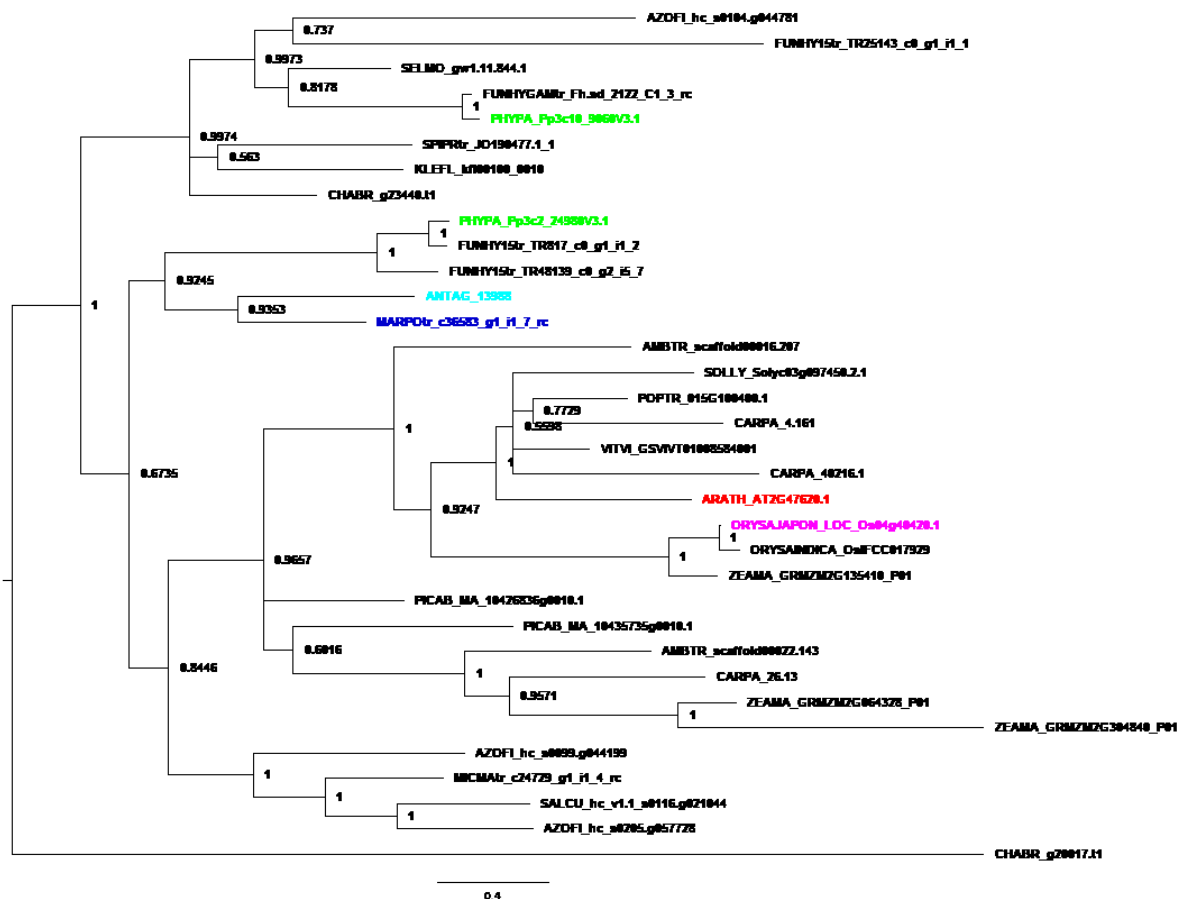


Figure 73 AT2G47620 phylogenetic analysis.

Phylogenetic analysis named after the *A. thaliana* blastp input query. Colour code was used according to the colour code in the main part. Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in the before mentioned five letter code.



Figure 74 AT2G39020 phylogenetic analysis.

Phylogenetic analysis named after the *A. thaliana* blastp input query. Colour code was used according to the colour code in the main part. Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in the before mentioned five letter code.

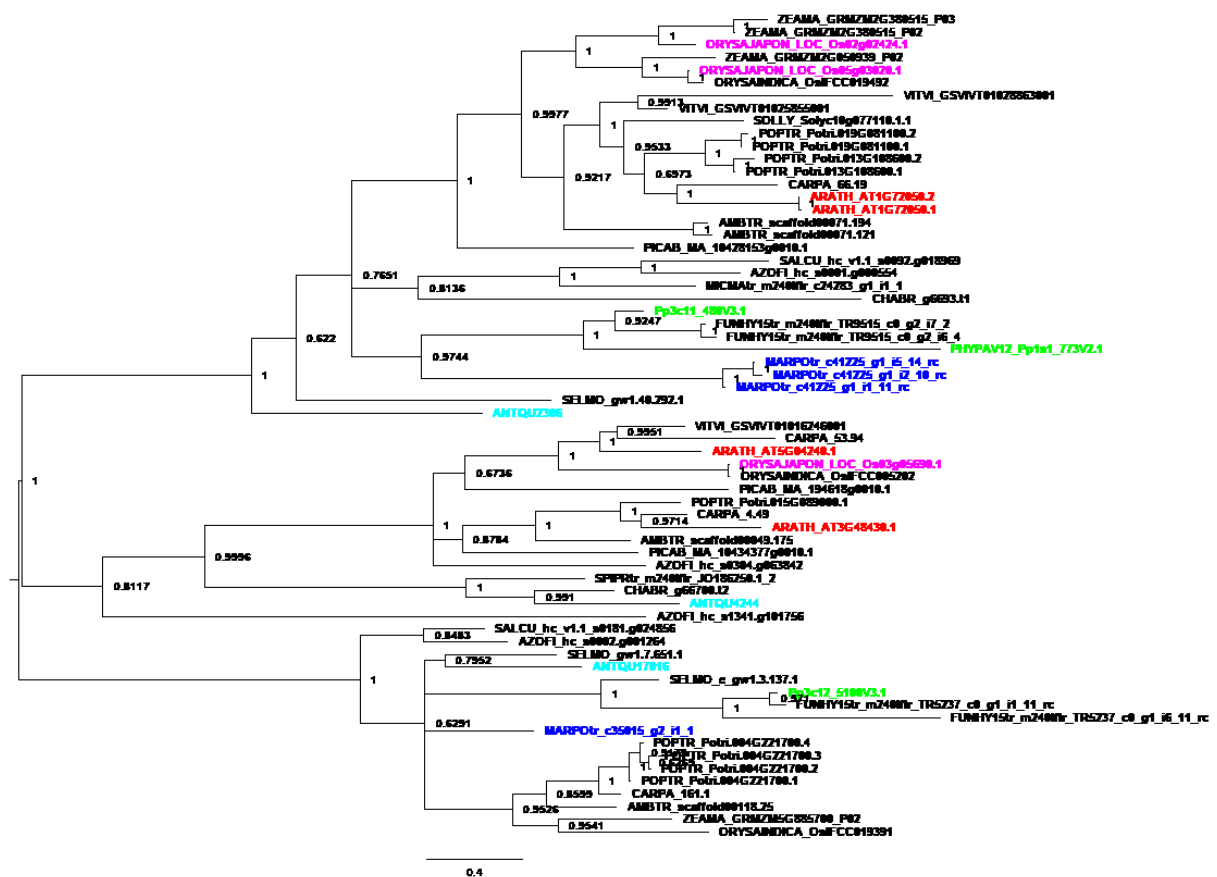


Figure 75 AT1G72050 phylogenetic analysis.

Phylogenetic analysis named after the *A. thaliana* blastp input query. Colour code was used according to the colour code in the main part. Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in the before mentioned five letter code.

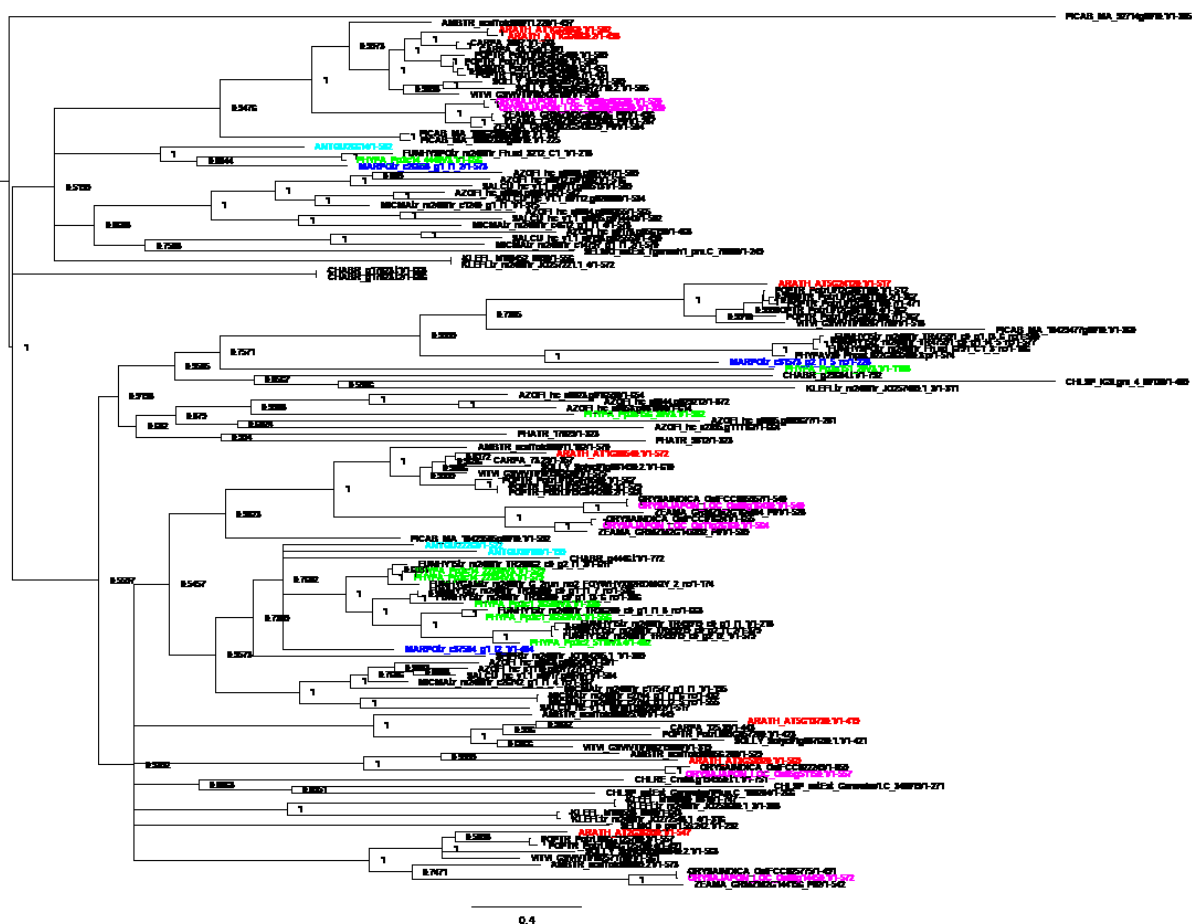


Figure 76 AT1G64860 phylogenetic analysis.

Phylogenetic analysis named after the *A. thaliana* blastp input query. Colour code was used according to the colour code in the main part. Support values at the nodes represent posterior probabilities of Bayesian inference. Species names are abbreviated in the before mentioned five letter code.

Supplement

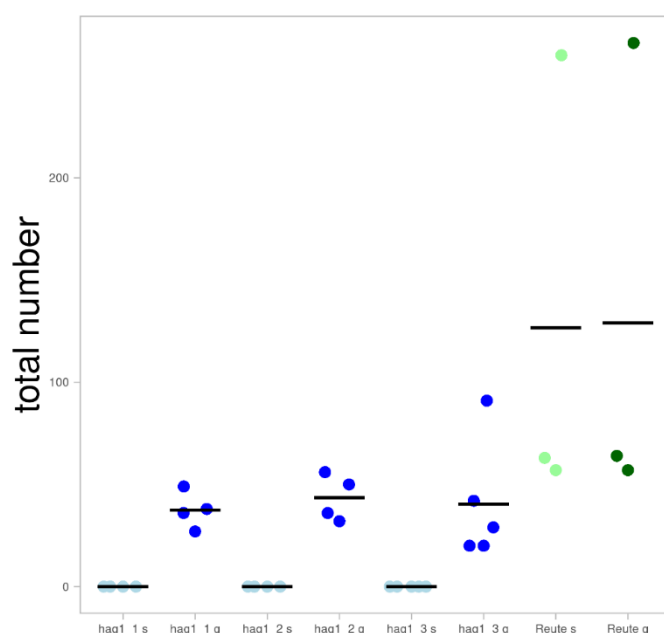


Figure 77 Sporophytes per gametophore of *Pphag1*.

The number of developed sporophytes (s) is indicated in lightgreen (control) and lightblue (*Pphag1*) and the corresponding number of gametophores (g). At least three independent replicates (dots) were calculated for the mutant lines as well as the control. The total number of gametophores analysed per mutant is shown in blue and the control is shown in darkgreen. In case of the mutants, three independent mutant lines were analysed and numbered consecutively, compared to the control. Averages of replicates are shown as horizontal lines.

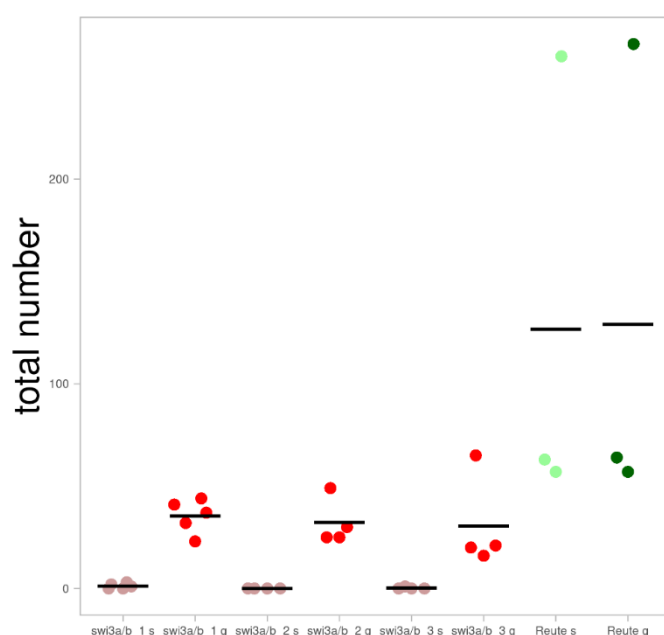


Figure 78 Sporophytes per gametophore of *Ppswi3a/b*.

The number of developed sporophytes (s) is indicated in lightgreen (control) and lightred (*Ppswi3a/b*) and the corresponding number of gametophores (g). At least three independent replicates (dots) were calculated for the mutant lines as well as the control. The total number of gametophores analysed per mutant is shown in red and the control is shown in darkgreen. In case of the mutants, three independent mutant lines were analysed and numbered consecutively, compared to the control. Averages of replicates are shown as horizontal lines.

Supplement

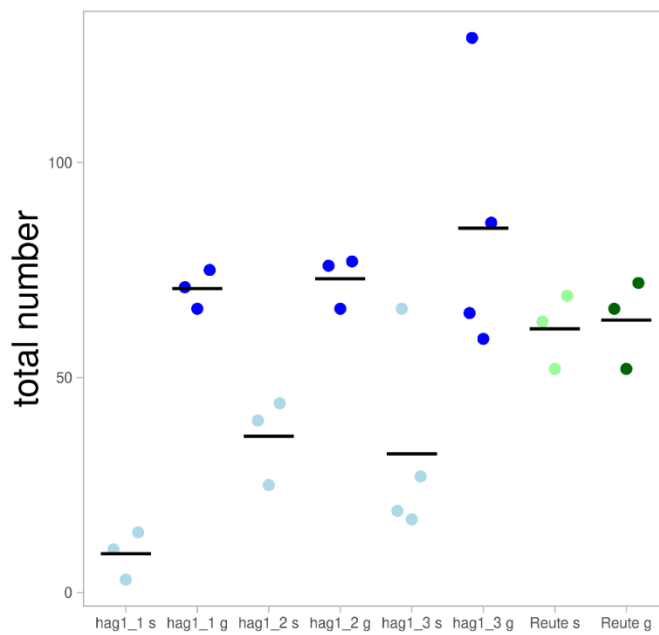


Figure 79 Crossing with a male fluorescent fertile strain to test for male impairment.

Pphag1 was crossed with Re-mcherry according to (Perroud *et al.*, 2019). *Pphag1* (lightblue) shows a lower sporophyte ratio compared to the control (lightgreen) (shown are the total numbers of developed sporophytes (s)). At least three independent replicates (dots) were calculated for the mutant lines as well as the control. The total number of gametophores (g) analysed for the mutant is shown in blue and for the control in darkgreen. In case of the mutant, three independent mutant lines were analysed and numbered consecutively, compared to the Reute control. Averages of replicates are shown as horizontal lines.

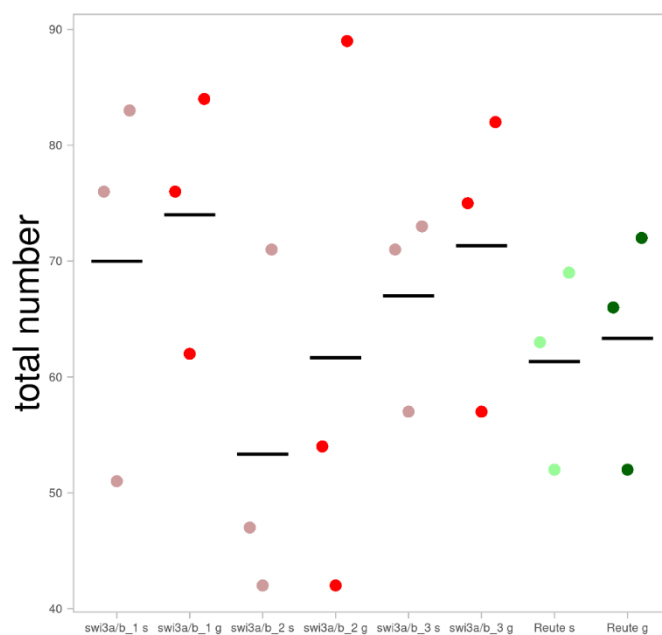


Figure 80 Crossing with a male fluorescent fertile strain to test for male impairment.

Ppswi3a/b was crossed with Re-mcherry according to (Perroud *et al.*, 2019). The cross with a male fertile strain could largely restore the phenotype in *swi3a/b* (lightred) (sporophyte ratio at least 80 %). In the control, the developed sporophytes (s) are shown in lightgreen. Three independent replicates (dots) were calculated for the mutant lines as well as the control. The total number of gametophores (g) analysed for the mutant is shown in red and for the control in darkgreen. In case of the mutant, three independent mutant lines were analysed and numbered consecutively, compared to the Reute control. Averages of replicates are shown as horizontal lines.

Supplement

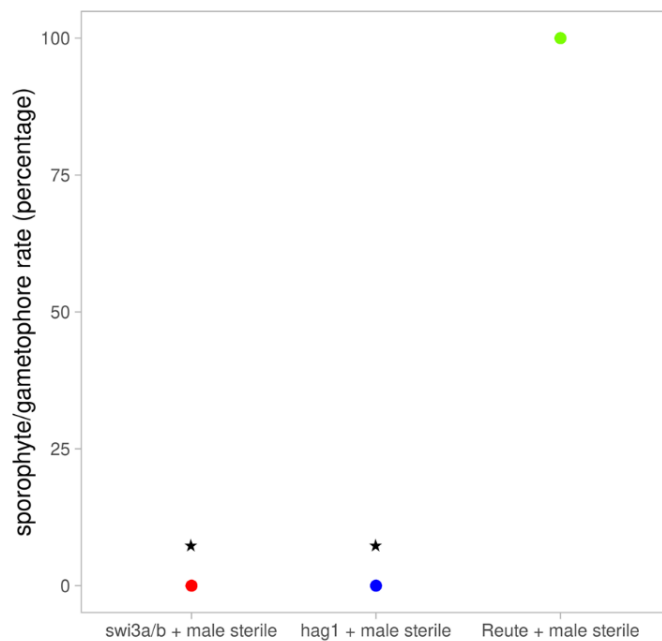


Figure 81 Crossing with a male infertile strain to test for male impairment.

Pphag1 (blue), *Ppswi3a/b* (red) and the control (Reute, green) were crossed with a male infertile strain (Meyberg *et al.*, 2020). The male sterile strain as well as the control developed 100 % of sporophytes per gametophore, whereas *Ppswi3a/b*, *Pphag1* as well as the male sterile mutant developed no sporophyte, which is a significant reduction (asterisks) ($p < 0,01$, Fisher's exact test applied for the conditions sporophyte yes/no groups mutant and control (Reute)). Only one replicate was calculated for the mutant lines and the control each (one dot). The total number of gametophores analysed per mutant/control was 73 each. In case of the mutants, one mutant line was analysed. The

dots indicate the rate of crossed sporophytes (sporophyte/gametophore rate in percentage relative to the total number of gametophores).

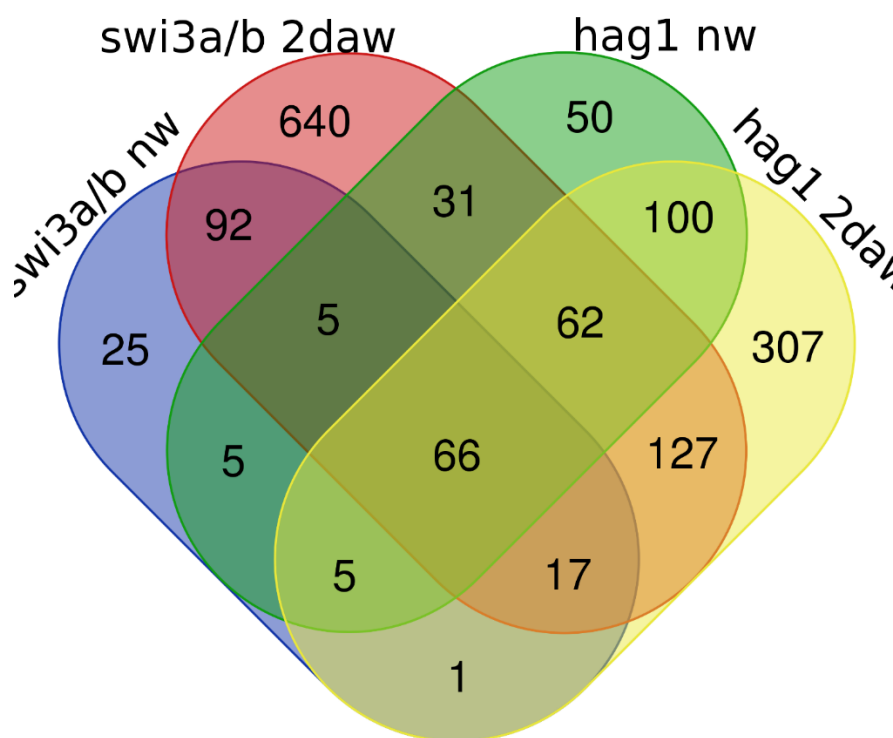


Figure 82 Comparison of all RNA-seq results within both mutants and both conditions.

Number of DEGs 21 days after induction of gametangiogenesis, with (2daw) or without (nw) watering.

Figure 83 GO enrichment of DEGs up-regulated in *Ppswi3a/b* two days after watering compared to the control.



Supplement

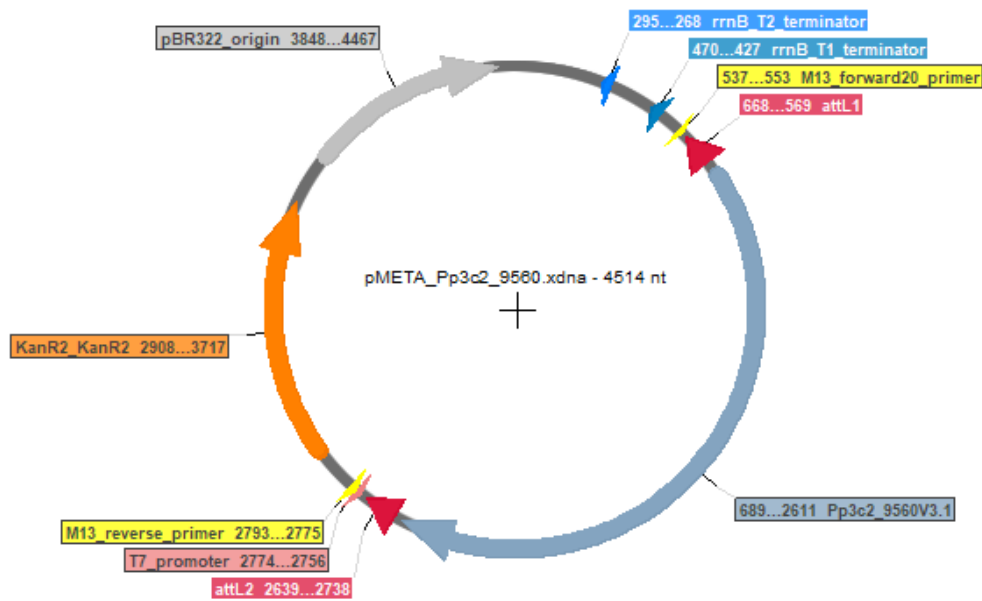


Figure 84 Vector card of pMETA.

Pphag1 was ligated in pMETA_XcmI. Features: rrnB_T2_terminator, rrnB_T1_terminator, M13_fwd, attL1, attL2, T7_promoter, M13_rev, KanR, pBR322_origin.

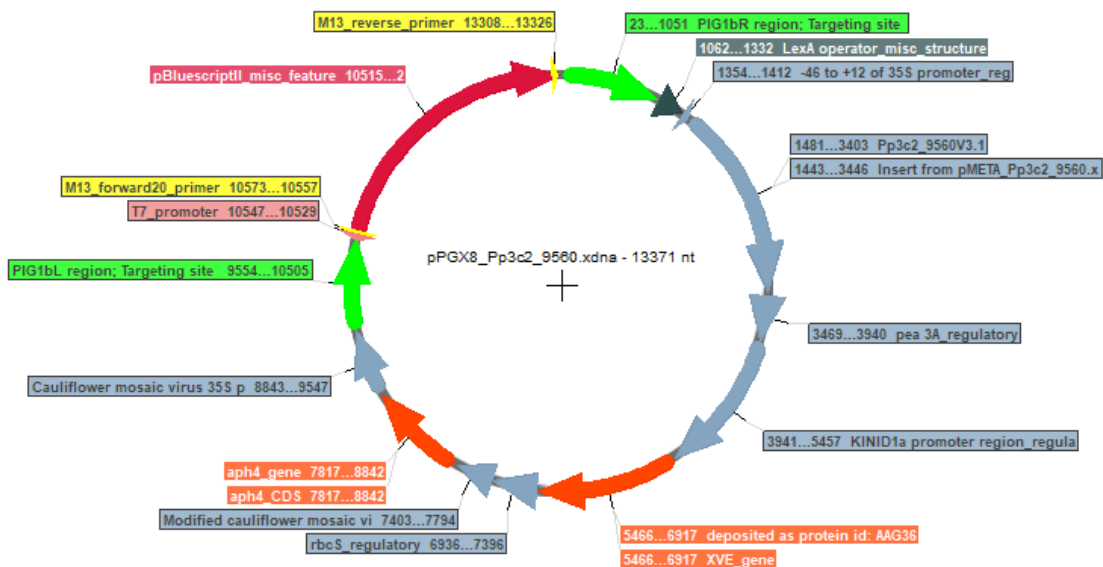


Figure 85 Vector card of pPGX8.

Pphag1 was cloned into pPGX8 via Gateway cloning. Features: PIG1bR region (HR1), LexA operator, 35S promoter, pea 3A terminator, GX8; KINID1a promoter, XVE gene estrogen chimera receptor (protein id: AAG36984), rbcS terminator, aph4 HygR, 35S polyA signal terminator, PIG1bL region (HR2), M13_fwd, M13_rev, T7_promoter.

Supplement

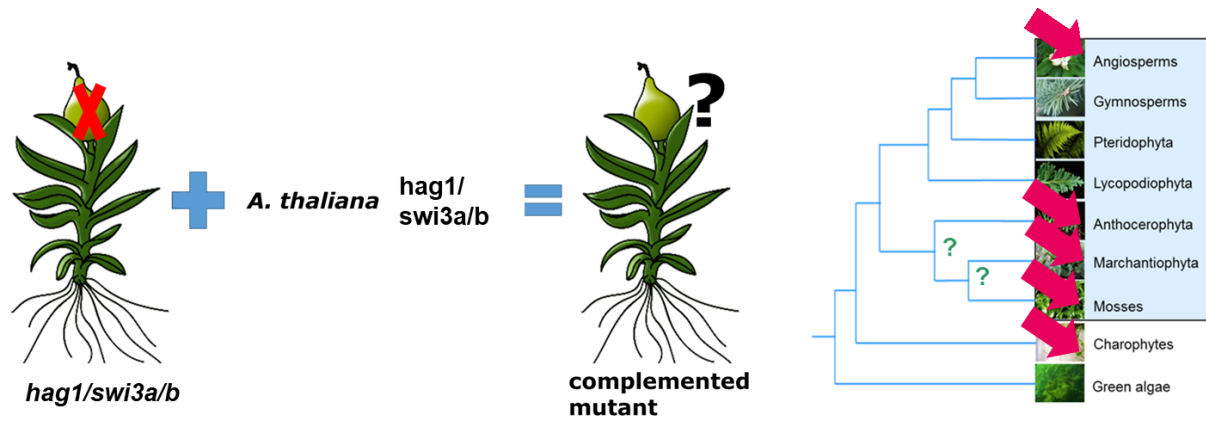


Figure 86 Schematic overview of *P. patens* complementation analyses.

The loss-of-function mutants should be complemented with the respective ortholog e.g. *A. thaliana* to see whether the phenotype can be restored by the ortholog of another species or the moss gene itself. The pink arrows indicate the species which should be used for complementation analyses.

Curriculum vitae

Curriculum vitae

Der Lebenslauf enthält persönliche Daten und ist deshalb nicht Teil der Online-Veröffentlichung.

Publications and contributions

Publications and contributions

Grosche C, **Genau AC**, Rensing SA (2018). Evolution of the Symbiosis-Specific GRAS Regulatory Network in Bryophytes, *Front. Plant Sci.* 9:1621.

Botanikertagung (2019) International Plant Science Conference

“Regulation of plant sexual reproduction – an evo-devo approach” (oral presentation)
funding: Reisestipendium der Deutschen Botanischen Gesellschaft

32rd Conference Molecular Biology of Plants of the Section Plant Physiology and Molecular Biology of the DBG (2019)

“Regulation of sexual reproduction: an evo devo approach” (oral presentation)

EMBO Workshop New shores in land plant evolution (2018) in Lisbon, Portugal

“Candidate transcriptional regulator selection to study the evolution of the alternation of generations” (poster presentation)
funding: Marburg International Doctorate Stipendium Konferenzteilnahme im Ausland

Black Forest Workshop 2018, the 1st Black Forest Flagellated Plant Workshop (2018)

“Candidate transcriptional regulator selection to study the evolution of the alternation of generations” (poster presentation)

31rd Conference Molecular Biology of Plants of the Section Plant Physiology and Molecular Biology of the DBG (2018)

“Candidate transcriptional regulator selection to study the evolution of the alternation of generations” (poster presentation)

ELIXIR workshop: Functional Plant Bioinformatics, Ghent, Belgium (2017)

Black Forest Summer School 2017, to see the (Black) Forest for the trees: next generation sequencing and phylogenetics

“Single copy ortholog identification in alternation of generation/sexual reproduction in bryophytes” (oral presentation)
funding: evoltree

Danksagung

Danksagung

Mein besonderer Dank gilt Prof. Stefan Rensing, der mir die Möglichkeit gegeben hat, dieses interessante Projekt zu bearbeiten und stets Interesse am Fortschritt des Projektes gezeigt hat. Durch hilfreiche Kommentare und Anregungen in Wort und Schrift habe ich mich in meiner gesamten Promotion von ihm unterstützt gefühlt. Ich danke ihm darüber hinaus für die Übernahme der Erstkorrektur.

Frau Prof. Annette Becker danke ich für die Übernahme der Zweitkorrektur.

Prof. Uwe Maier und Prof. Lars Voll danke ich für ihre Bereitschaft Teil des Prüfungskomitees zu sein.

Rabea, Tophier, Olivia und Heide danke ich für Korrekturen und kritische Kommentare das Manuskript betreffend.

Mein Dank gilt außerdem Noe und Fabian für Hilfestellungen in der Bioinformatik und Rabea für Diskussionen rund um das Thema sexuelle Reproduktion.

Marco und Angela danke ich sowohl für jeglichen technischen Support als auch für viele unterhaltsame Momente im Labor.

Der ganzen Arbeitsgruppe Rensing möchte ich für das gute Arbeitsklima danken.

Meinen Freunden möchte ich für Zerstreuung abseits des Laboralltages danken.

Zu guter Letzt möchte ich mich vor allem bei meinen Eltern, Geschwistern und Christian für die immerwährende Unterstützung und aufbauenden Worte in teils schwierigen Momenten bedanken.

Eigenständigkeitserklärung

Eigenständigkeitserklärung

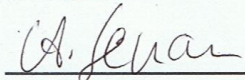
Eigenständigkeitserklärung

Eigenständigkeitserklärung

Hiermit versichere ich, Anne Christina Genau, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Die Arbeit wurde bisher weder in der jetzigen noch in einer ähnlichen Form einer Prüfungsbehörde vorgelegt.

Marburg, 2020



(Anne Christina Genau)