



Transcriptional Landscapes of Divergent Sporophyte Development in Two Mosses, *Physcomitrium (Physcomitrella) patens* and *Funaria hygrometrica*

Alexander Kirbis¹, Manuel Waller¹, Mariana Ricca², Zoe Bont³, Anna Neubauer¹, Bernard Goffinet⁴ and Péter Szövényi^{1*}

¹ Department of Systematic and Evolutionary Botany, University of Zurich, Zurich and Zurich-Basel Plant Science Center, Zurich, Switzerland, ² Department for BioMedical Research (DBMR), University of Bern, Bern, Switzerland, ³ Institute of Plant Sciences, University of Bern, Bern, Switzerland, ⁴ Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, United States

OPEN ACCESS

Edited by:

Stefan de Folter,
Centro de Investigación y Estudios
Avanzados, Instituto Politécnico
Nacional de México (CINVESTAV),
Mexico

Reviewed by:

Lydia Gramzow,
Friedrich Schiller University Jena,
Germany
Caspar Christian Cedric Chater,
The University of Sheffield,
United Kingdom

*Correspondence:

Péter Szövényi
peter.szovenyi@systbot.uzh.ch

Specialty section:

This article was submitted to
Plant Development and EvoDevo,
a section of the journal
Frontiers in Plant Science

Received: 19 February 2020

Accepted: 11 May 2020

Published: 10 June 2020

Citation:

Kirbis A, Waller M, Ricca M,
Bont Z, Neubauer A, Goffinet B and
Szövényi P (2020) Transcriptional
Landscapes of Divergent Sporophyte
Development in Two Mosses,
Physcomitrium (Physcomitrella)
patens and *Funaria hygrometrica*.
Front. Plant Sci. 11:747.
doi: 10.3389/fpls.2020.00747

Understanding the molecular basis of morphological shifts is a fundamental question of evolutionary biology. New morphologies may arise through the birth/death of genes (gene gain/loss) or by reutilizing existing gene sets. Yet, the relative contribution of these two processes to radical morphological shifts is still poorly understood. Here, we use the model system of two mosses, *Funaria hygrometrica* and *Physcomitrium (Physcomitrella) patens*, to investigate the molecular mechanisms underlying contrasting sporophyte architectures. We used comparative analysis of time-series expression data for four stages of sporophyte development in both species to address this question in detail. We found that large-scale differences in sporophytic architecture are mainly governed by orthologous (i.e., shared) genes frequently experiencing temporal gene expression shifts between the two species. While the absolute number of species-specific genes expressed during sporophyte development is somewhat smaller, we observed a significant increase of their proportion in preferentially sporophyte expressed genes, suggesting a fundamental role in the sporophyte phase. However, further functional studies are necessary to determine their contribution to diverging sporophyte morphologies. Our results add to the growing set of studies suggesting that radical changes in morphology may rely on the heterochronic expression of conserved regulators.

Keywords: transcriptomics, sporophyte development, mosses, Funariaceae, RNAseq

INTRODUCTION

The genome is constantly reshaped by diverse types of mutations providing raw material for evolution to work with. Genomic changes can ultimately lead to new phenotypes possessing a novel set of morphological characters potentially enabling adaptation to new environmental conditions (Orr, 2005). Various molecular mechanisms may underlie the origin of morphological novelties including (i) the rise of novel genes (de novo evolution and introgression of genes)

(Chen et al., 2013; Schlötterer, 2015; Li et al., 2016; Zhang et al., 2019), (ii) the utilization of existing genes for new functions (Davidson, 2010; Pires and Dolan, 2012; Das Gupta and Tsiantis, 2018; Bowman et al., 2019), (iii) and the loss of genes or gene function (Albalat and Cañestro, 2016; Xu et al., 2019). Nevertheless, in reality molecular mechanisms underlying the origin of morphological novelties are more complex and can be best described by various combinations of these three basic scenarios.

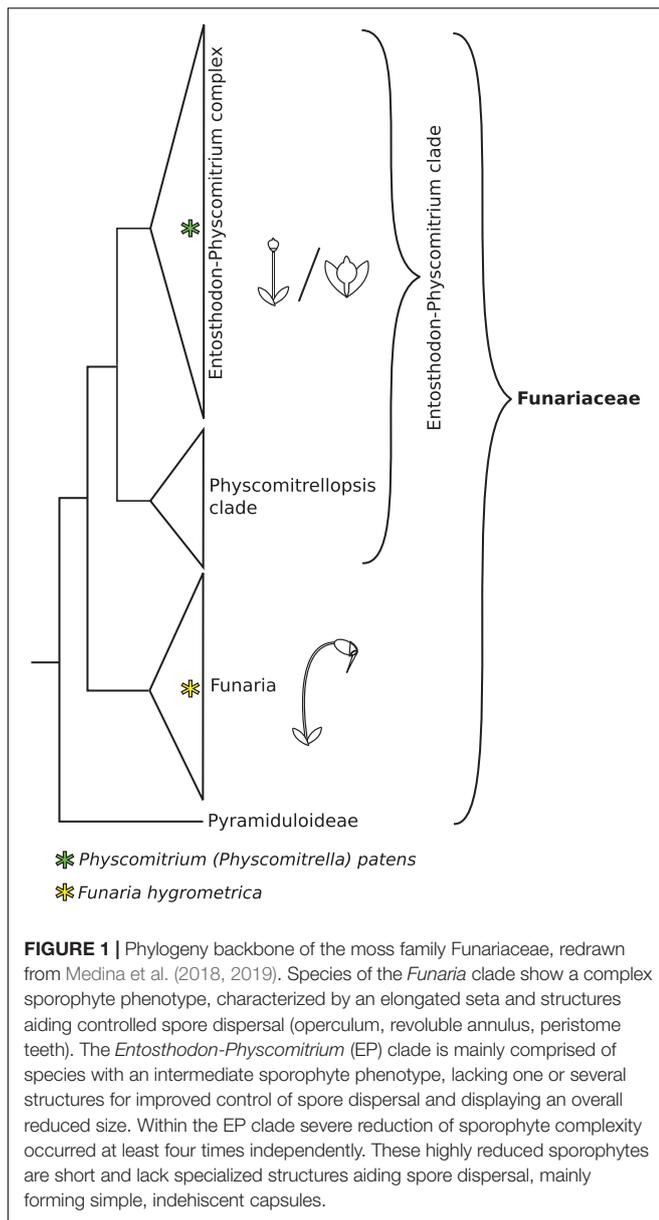
Although the primary molecular mechanisms contributing to novel phenotypes are well-documented, their relative importance is poorly understood and difficult to predict (True and Carroll, 2002; Khalturin et al., 2009; Kaessmann, 2010; Wagner and Lynch, 2010). For instance, *de novo* evolution of genes is thought to have boosted the diversification of plants by enabling the evolution of new key characters both at the molecular and macromorphological levels (Furumizu et al., 2015; Soltis and Soltis, 2016; Jill Harrison, 2017; Van de Peer et al., 2017; Clark and Donoghue, 2018; Landis et al., 2018; Whitewoods et al., 2018; One Thousand Plant Transcriptomes Initiative, 2019; Bowles et al., 2020). Similarly, new genes acquired by horizontal gene transfer or introgression/hybridization seem to have also been crucial in the evolution of key morphological features (Suarez-Gonzalez et al., 2018; Cheng et al., 2019; Wickell and Li, 2020). In contrast, the evolution of many key phenotypic characters has taken another path by co-opting existing genes or complete regulatory networks to create new morphological features (Rast-Somssich et al., 2015; Rebeiz and Tsiantis, 2017). Often, temporal shifts in the expression of conserved regulatory modules (e.g., heterochronic expression) is sufficient to give rise to new morphological innovations both in plant and animal systems (Geuten and Coenen, 2013; Buendía-Monreal and Gillmor, 2018). For instance, heterochronic expression of some key genes are major determinants of organ size and number in *Arabidopsis* (Sun et al., 2014). Finally, evidence is mounting that new phenotypes can also be acquired by loss of genes or gene functions. For instance, rapid evolution of new phenotypes seems to have proceeded by compromising gene function both in animals and plants (Olson, 1999; Nachman et al., 2003; Gujas et al., 2012; MacArthur et al., 2012; Nadeau et al., 2016; Sun et al., 2018). Therefore, the underlying molecular processes leading to the evolution of new morphological structures are diverse and it is currently unclear why evolution of particular phenotypes would follow one or the other trajectory (see i-iii above). Furthermore, whether evolutionary trajectories are canalized by various currently poorly known constraints or the prevalence of one trajectory is rather determined by random chance is highly debated (Galis et al., 2018). Revealing the molecular processes underlying phenotypic changes in a diverse set of model systems may help to discover key commonalities of the evolutionary process and assess how gene gain/loss and co-option of existing genes for new functions may contribute to morphological evolution.

Here, we use a model system of two species from a single family of mosses, *Physcomitrium* (*Physcomitrella*) *patens* (hereafter referred to as *P. patens*, see in Medina et al., 2019)

and *Funaria hygrometrica*, to begin investigating the molecular mechanisms shaping the evolution of their highly distinct sporophyte morphologies. In contrast to flowering plants, the moss life cycle possesses a dominant haploid gametophyte (1n) (consisting of the filamentous protonema and leafy shoot-like gametophores) alternating with an unbranched diploid sporophyte (2n) phase. The diploid sporophyte phase is multicellular, photosynthetic although nutritionally dependent on the maternal gametophyte to which it is permanently attached (Jonathan Shaw and Goffinet, 2000). The primary function of the sporophyte is to produce spores and control their dispersal, and therefore the sporophyte is likely under severe selection (Haig, 2013).

The moss family Funariaceae comprises about 300 species, displaying a relatively uniform gametophyte morphology, but highly variable sporophyte stature varying in size from about a millimeter to 5 cm and in complexity from a sessile sporangium lacking a differentiated mode of dehiscence to a long stalked capsule bearing highly specialized structures for controlled spore dispersal (Medina et al., 2018, 2019). Phylogenomic analyses revealed that the subfamily Funarioideae comprises the monophyletic *Funaria* and its sister lineage, the *Entosthodon-Physcomitrium* (EP) complex (**Figure 1**) with an estimated divergence time of 60 million years (95% CI: 35–70 million years ago; Medina et al., 2018). *Funaria* is characterized by an architecturally rather complex sporophyte, with an elongated seta, asymmetric capsule, dehiscing via a revoluble double annulus revealing a double peristome regulating spore dispersal (Fife, 1982; Liu et al., 2012; Medina et al., 2018). By contrast, the sporophyte of the E-P complex is more variable, spanning a gradient of architectural complexity extending to the simple sporophyte of *P. patens* whose sporophyte is composed of a short seta and a small spherical capsule lacking differentiated structures associated with spore dispersal (**Figures 1, 2**). Given the resolution of *P. patens* within a grade of *Entosthodon* and *Physcomitrium* species, all with more complex sporophytes, it is assumed that the sporophyte morphology of *P. patens* arose through reduction (Liu et al., 2012; Medina et al., 2018). Ontogenetic transformations of the sporophyte are correlated or at least followed by changes in the development of the calyptra, the protective maternal gametophytic tissue covering the apex of the developing sporophyte (Budke and Goffinet, 2016). Reduction of the diploid generation occurred multiple times in the complex, giving rise to several species with a *Physcomitrium* phenotype (Medina et al., 2019). Whether such transformations were triggered by similar genetic processes is not known, and in fact the molecular mechanisms underlying sporophyte development in mosses, in particular that of contrasting sporophyte morphologies in the Funariaceae, are poorly understood (Goffinet and Buck, 2013).

We used comparative gene expression and genome analysis to begin investigating the molecular changes that potentially contribute to the evolution of the contrasting sporophyte morphology in Funariaceae. To do so, we generated time-series gene expression data for four comparable stages of sporophyte development in two funarioid mosses, *F. hygrometrica*



(complex sporophyte) and *P. patens* (simple sporophyte), representing the two extremes of sporophyte complexity in the family. We contrasted gene expression in the two species to test whether shifts in sporophyte complexity were associated with changes in expression of shared genes or with changes in gene content, e.g., gene loss/gain. Our analyses suggest that heterochronic expression of conserved sets of developmental genes govern the development of sporophytes with contrasting complexities, although the contribution of gene gain/loss is also considerable. Our study adds to the growing set of observations that reorganization of highly conserved regulatory networks is a critical mechanism underlying major morphological shifts in evolution (Rosin and Kramer, 2009; Zhao et al., 2016; Das Gupta and Tsiantis, 2018; Bowman et al., 2019).

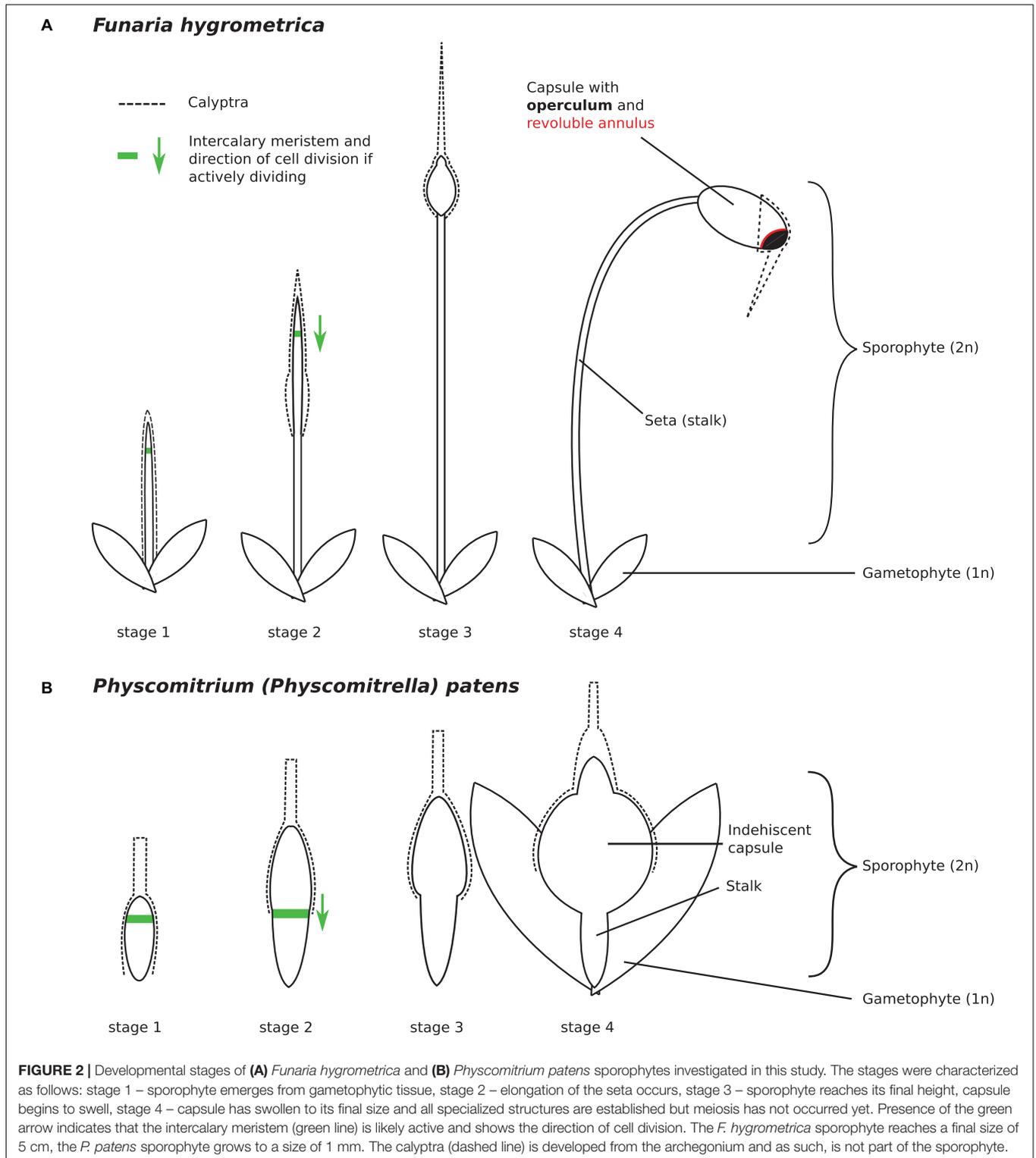
MATERIALS AND METHODS

Extraction of RNA From Sporophyte Tissue of *F. hygrometrica* and *P. patens*

The *F. hygrometrica* isolate used in this study was established from a single spore of a sporophyte collected around the city of Sankt Gallen, Switzerland. Gametophores were grown and vegetatively propagated in petri dishes on KNOP medium (Reski and Abel, 1985). To obtain sporophytes, gametophores were coarsely ground, and fragments spread on sterilized soil. Cultures were regularly sprayed with water to facilitate fertilization and sporophyte development was checked weekly. We collected sporophytes of *F. hygrometrica* corresponding to the four developmental stages described in Budke et al. (2012) and depicted in **Figure 2A**. We also sampled *P. patens* sporophytes with comparable developmental stages as shown in **Figure 2B**. For RNA-sequencing, we sampled sporophytes using forceps and scissors. We also used spores from freshly collected sporophytes to obtain plant material for various developmental stages of the gametophyte phase. We spread spores of surface sterilized sporophytes onto Knop medium overlaid with cellophane and collected germinating spores after 3 days (germinating spores 3 days) and 2 weeks (protonema 2 weeks). Young gametophores (young buds) emerging after 3 weeks were also collected. We immediately submerged collected sporophytes and gametophyte tissues into RNAlater[®] and stored them at 4°C until RNA extraction. RNA was extracted using the Spectrum[™] Plant Total RNA Kit from Sigma-Aldrich. The calypra was removed from all sporophytic tissues prior RNA extraction. Poly-A RNA of three biological replicates were sequenced with paired-end or single-end sequencing on a HiSeq 2500 or 4000 machine (see **Supplementary Table S1**).

Draft Genome Sequence of *F. hygrometrica*

We extracted DNA using a modified CTAB protocol (Porebski et al., 1997) from axenically grown gametophytes of the very same *F. hygrometrica* isolate we used for the RNA-seq experiment. We first prepared 454 libraries, which were sequenced with the titanium technology on a 454 FLX Roche machine at the Functional Genomic Center Zurich (FGCZ). We prepared short-insert DNA libraries and 3 and 5 kb jumping libraries using the Illumina DNA-seq and mate-pair library preparation kits and sequenced them at the FGCZ on HiSeq 2000 and HiSeq 2500 machines (short-insert libraries with a depth of 120x, mate-pair libraries with 26x). We quality filtered raw reads with Trimmomatic (Bolger et al., 2014) using default values for paired-end and single-end sequencing data. We assembled the genome with SOAPdenovo2 (Luo et al., 2012) using short-insert Illumina and 454 reads in the contig assembly step and 3 and 5 kb jumping libraries in the scaffolding step. We filled gaps in the final assembly using all 454 and Illumina data with Gap-closer (Luo et al., 2012). Our final assembly consisted of 13,000 scaffolds longer than 100 bp with an N50 value of ca. 100 kb and a total length of 340 Mbp. Genomic scaffolds can be found in **Supplementary Data S1**. Quality of the genome



assembly was assessed by searching translated peptide sequences of the reconstructed gene set (see in section “Read Mapping and Quantification of Expression”) against hidden Markov models of Benchmarking Universal Single-Copy Orthologs (BUSCO) (Seppy et al., 2019). For this purpose, we used the

Viridiplantae and Embryophyta datasets retrieved from OrthoDB v10.1 (Kriventseva et al., 2019). To put these results into context, publicly available gene coding sequences of *Marchantia polymorpha* (Bowman et al., 2017), *P. patens* (Lang et al., 2018), *Pleurozium schreberi* (Pederson et al., 2019), and *Sphagnum*

fallax were retrieved from Phytozome v12 (Goodstein et al., 2012) and searched against the same BUSCO sets. Results of all completeness assessments are shown in **Table 1**. Our results show that the reconstructed gene set captures the complete sequence of 80.5% BUSCOs from the Embryophyta dataset and of 91.8% from the Viridiplantae dataset. These BUSCO values are close to the figures obtained for high-quality genome assemblies generated using long-read technologies (*M. polymorpha*, *P. patens*, and *S. fallax*) and clearly exceed the values of the *P. schreberi* genome, which was likewise assembled using only short-reads. Overall, our BUSCO analysis implies that the *F. hygrometrica* draft genome is of good quality.

Read Mapping and Quantification of Expression

Mapping of the reads obtained by Illumina sequencing was performed following the HISAT2/StringTie pipeline (Pertea et al., 2016). First, adapter sequences were removed from the libraries and the reads were quality filtered using Trimmomatic v36 (Bolger et al., 2014) (ILLUMINACLIP:TruSeq3-SE.fa:2:30:10 SLIDINGWINDOW:4:5 LEADING:5 TRAILING:5 MINLEN:25). The reads were then mapped to the *P. patens* reference genome, version 3.1 (Lang et al., 2018), and the *F. hygrometrica* draft genome (**Supplementary Data S1**), respectively, using HISAT2 (Kim et al., 2015). Detailed statistics of the read mapping are available in **Supplementary Table S1**. To calculate transcript abundance for the *P. patens* data set, we used the version 3.1 gene annotation (Lang et al., 2018) in combination with quantification in StringTie. The transcriptome of *F. hygrometrica* was reconstructed using StringTie, including *de novo* assembly of transcripts and a subsequent quantification of transcript abundance (Pertea et al., 2015). To get a complete set of transcripts expressed during sporophyte development for each species, the assembled transcripts of all samples were merged using the `-merge` function of StringTie (**Supplementary Data S1**). The transcriptome reconstruction of *F. hygrometrica*

yielded 25,904 unique loci. We discarded all gene models with a summarized read count over all samples lower than 10 to remove gene models with very few, potentially misaligned, reads. This filtering step reduced the gene set to 25,460 (*F. hygrometrica*) and 22,690 (*P. patens*) predicted gene models which show detectable expression in the analyzed samples.

Preferential Expression in the Sporophyte Stage

Genes preferentially expressed in the sporophyte stage were identified using the R package DESeq2 version 1.16.1 (Love, Huber and Anders, 2014). The sporophyte expression data comprised RNAseq raw count data from three replicates of four developmental stages each (for sporophyte developmental stage 1 from *P. patens* only two replicates were used). Expression data from the gametophyte phase included raw read counts from six (*F. hygrometrica*) and eight (*P. patens*) samples from different stages of gametophyte development. After defining sporophytic and gametophytic samples, a differential expression analysis was conducted using the unmodified DESeq2 algorithm. To obtain genes highly expressed in the sporophyte stage in contrast to gametophytic stages, significant results ($p_{adj} \leq 0.05$) were filtered for gene models with a \log_2 -fold change ≥ 2 (**Supplementary Data S2, S3**). We used conventional chi-square statistics to test for enrichment of genes in the set showing preferential sporophyte expression.

Identification of Homologs and Orthologs

Homologous genes (including orthologs and paralogs) were identified by the BLASTp algorithm (Altschul et al., 1990) and the software Orthofinder2 (Emms and Kelly, 2018). We used this approach to identify *F. hygrometrica* and *P. patens* gene models that are species-specific (i.e., that have no detectable homolog in the alternate species' proteome). Besides Orthofinder, we used BLASTp to detect distant homologs that might be missed by the

TABLE 1 | Assessment of completeness of the *F. hygrometrica* draft genome using Benchmarking Universal Single-Copy Orthologs (BUSCO).

	Total	Complete		Fragmented	Missing
		Single-copy	Duplicated		
Embryophyta (n = 1614)					
<i>F. hygrometrica</i>	1300 (80.5%)	1101 (68.2%)	199 (12.3%)	43 (2.7%)	271 (16.8%)
<i>M. polymorpha</i>	1412 (87.5%)	1385 (85.8%)	27 (1.7%)	20 (1.2%)	182 (11.3%)
<i>P. patens</i>	1423 (88.1%)	1174 (72.7%)	249 (15.4%)	28 (1.7%)	163 (10.2%)
<i>P. schreberi</i>	736 (45.6%)	540 (33.5%)	196 (12.1%)	278 (17.2%)	600 (37.2%)
<i>S. fallax</i>	1447 (89.7%)	1257 (77.9%)	190 (11.8%)	22 (1.4%)	145 (8.9%)
Viridiplantae (n = 425)					
<i>F. hygrometrica</i>	390 (91.8%)	336 (79.1%)	54 (12.7%)	7 (1.6%)	28 (6.6%)
<i>M. polymorpha</i>	411 (96.7%)	409 (96.2%)	2 (0.5%)	3 (0.7%)	11 (2.6%)
<i>P. patens</i>	416 (97.8%)	361 (84.9%)	55 (12.9%)	1 (0.2%)	8 (2.0%)
<i>P. schreberi</i>	227 (53.4%)	160 (37.6%)	67 (15.8%)	107 (25.2%)	91 (21.4%)
<i>S. fallax</i>	417 (98.1%)	375 (88.2%)	42 (9.9%)	2 (0.5%)	6 (1.4%)

For comparison, BUSCO statistics for the *Marchantia polymorpha*, *Physcomitrium patens*, *Pleurozium schreberi*, and *Sphagnum fallax* genomes are shown. Data was retrieved from Phytozome v12 (Goodstein et al., 2012) and analyzed using the Viridiplantae and Embryophyta BUSCO datasets obtained from OrthoDB v10.1.

more stringent approach of Orthofinder. The BLASTp search included reference peptide sequences for each gene model in the *P. patens* and *F. hygrometrica* data sets. Hits were filtered for alignments with excess similarity by applying an *E*-value threshold of $\leq 10^{-6}$. To remove alignments that cover only small conserved domains within the query we applied a second filtering step for alignments with $\geq 80\%$ query coverage and $\geq 35\%$ sequence similarity between the query and the corresponding hit. Both filtering methods are reported to reliably detect homologous sequences even between distantly related species (Rost, 1999; Pearson, 2013). We also ran Orthofinder2 v2.3.3 (Emms and Kelly, 2018) to identify orthogroups (comprising both orthologs and paralogs) between *P. patens* and *F. hygrometrica* using default parameters (**Supplementary Data S4**). We included the following proteomes to generate orthogroups: *Citrus clementina*, *M. polymorpha*, *Theobroma cacao*, *Vitis vinifera*, *Prunus persica*, *Cucumis sativus*, *Amborella trichopoda*, *Physcomitrium patens*, *F. hygrometrica*, *Selaginella moellendorffii*, *Zea mays*, *Oryza sativa* v7_JGI, *Brassica oleracea*, *Arabidopsis thaliana*, *P. trichocarpa*, *Medicago truncatula*, and *Daucus carota*. All proteomes, except the one for *F. hygrometrica*, were retrieved from Phytozome v12 (Goodstein et al., 2012). We also let Orthofinder2 v2.3.3 automatically calculate gene trees for each orthogroup and identify one-to-one orthologs of *P. patens* and *F. hygrometrica* using phylogenetic information. We used one-to-one orthologs between *P. patens* and *F. hygrometrica* to compare gene expression in the two species.

PCA

To gain insights into the overall change of gene expression associated with sporophyte development in *F. hygrometrica* and *P. patens*, we used principal component (PCA) analyses (Hotelling, 1933) on normalized and standardized RNA-seq expression data (**Figure 3A**). Because we assumed that divergent sporophyte morphologies may be partially associated with gene gain/loss we generated PCAs separately using each species' full gene set (including both homologs and species-specific genes), the gene set with homologs in the alternate species (including orthologs and paralogs) and the species-specific gene set (non-homologs, for details see section "Identification of Homologs and Orthologs"). We argue here that the amplitude of a gene's expression variation (expression dynamics) throughout sporophyte developmental stages likely correlates with its functional relevance. Following this reasoning, we expected more pronounced differentiation across sporophyte developmental stages in gene expression using the species-specific gene set if gene gain/loss is more important in contributing to the divergent sporophyte morphologies than genes that are part of the shared gene set (homologous genes). Alternatively, expression variation across sporophyte developmental stages may be more pronounced using the shared gene set if sporophyte development is rather driven by gene expression changes in homologous genes and not the expression dynamics of species-specific genes.

To further investigate which sporophyte developmental stages showed similar or rather divergent gene expression patterns in the two species we calculated an expression divergence matrix between developmental stages, using one-to-one orthologs that

showed expression in both species in the corresponding stage. The expression divergence matrix was built by calculating pairwise Spearman's rank correlation coefficients between expression profiles of developmental stages of *F. hygrometrica* and *P. patens*. Results were plotted using the R package ggplot2 v. 3.2.1 (Wickham, 2009).

Functional Annotation

We retrieved functional annotation of *P. patens* gene models from Phytozome v11 (Goodstein et al., 2012). We translated and annotated transcript sequences of *F. hygrometrica* by running tBLASTn searches (*E*-value threshold 10^{-6}) against the plant proteomes available in PLAZA v2.0 (Proost et al., 2009). GO annotation for each transcript was obtained by transferring the GO annotation of the respective best hit protein to the *F. hygrometrica* transcript.

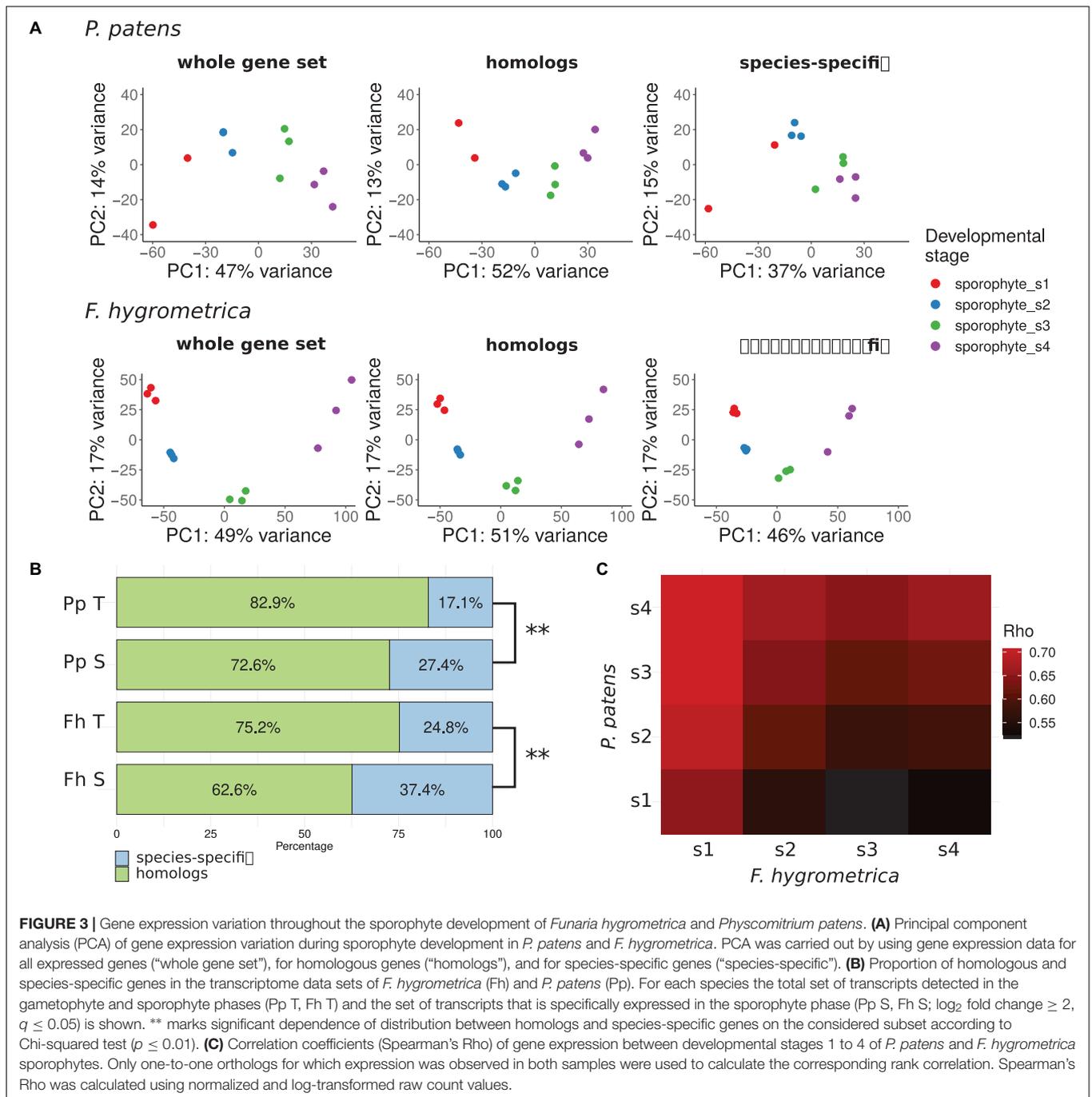
Key positions in gene regulatory networks are often occupied by transcription factors, which trigger expression of an array of downstream target genes resulting in a regulatory cascade setting off developmental transitions (Davidson and Erwin, 2006; Erwin and Davidson, 2009). Identification of gene models that encode transcription factors was done based on conserved DNA binding motifs using the PlantTFDB 4.0 resource (Jin et al., 2017). The algorithm reported 1,061 (*P. patens*) and 849 (*F. hygrometrica*) transcription factor encoding genes, which are expressed in at least one sample.

Gene Ontology Term Enrichment

To functionally characterize genes preferentially expressed in the sporophyte phase of *F. hygrometrica* and *P. patens*, we used the R package TopGO, version 2.38.1 (Kim, 2019). For the enrichment analyses we only considered GO-terms of the class "Biological Process." After defining all gene models with detectable expression in the RNA-seq data set of sporophyte and gametophyte tissue as the gene universe, and identifying all genes which are preferentially expressed in the sporophyte phase of either species as the subset of interesting genes, enriched GO-terms were computed using the Parent-Child Algorithm (Grossmann et al., 2007) and their significance tested with Fisher's exact test, both implemented in the TopGO package. Lists of enriched GO-terms (**Supplementary Data S5–S8**) were reduced using the REVIGO online resource (Supek et al., 2011) to remove redundant terms. Results of the enrichment analysis (**Figure 4**) were visualized using the R package GOsummaries version 2.22.0 (Kolde and Vilo, 2015).

Clustering by Gene Expression Profiles

The main goal of this study is to identify the genes and gene networks that may have contributed to the divergent development of the sporophyte phase in *P. patens* and *F. hygrometrica*. Since development is a dynamic, time dependent process we sought to use a method that would allow us to statistically distinguish one-to-one orthologous gene sets showing differential expression dynamics throughout sporophyte development between *P. patens* and *F. hygrometrica*. To compare expression dynamics of one-to-one orthologous genes in the two species during sporophyte development, we used



a method enabling classification of genes by their expression pattern and mapping gene expression profiles of one species against clusters of the other species. To achieve this, we carried out fuzzy c-means clustering with the R package mfuzz (Kumar and Futschik, 2007). We first clustered the *F. hygrometrica* gene expression data and then statistically assigned genes of *P. patens* to these clusters according to their expression dynamics. Because this assignment relies on the alignment of putatively functionally conserved genes, the analysis is restricted to one-to-one orthologs as detected by the

Orthofinder algorithm (see section "Materials and Methods"). To obtain the optimal cluster number, we computed the minimum centroid distance for a range between four and 40 clusters using the implemented Dmin() function. Centroid distances decreased rapidly up to a cluster number of 12. For higher cluster numbers the distance curve flattened out and changes in centroid distances were less severe. Therefore, we decided to use 12 clusters for the fuzzy c-means clustering as a compromise between describing the expression dynamics reasonably well without obtaining too

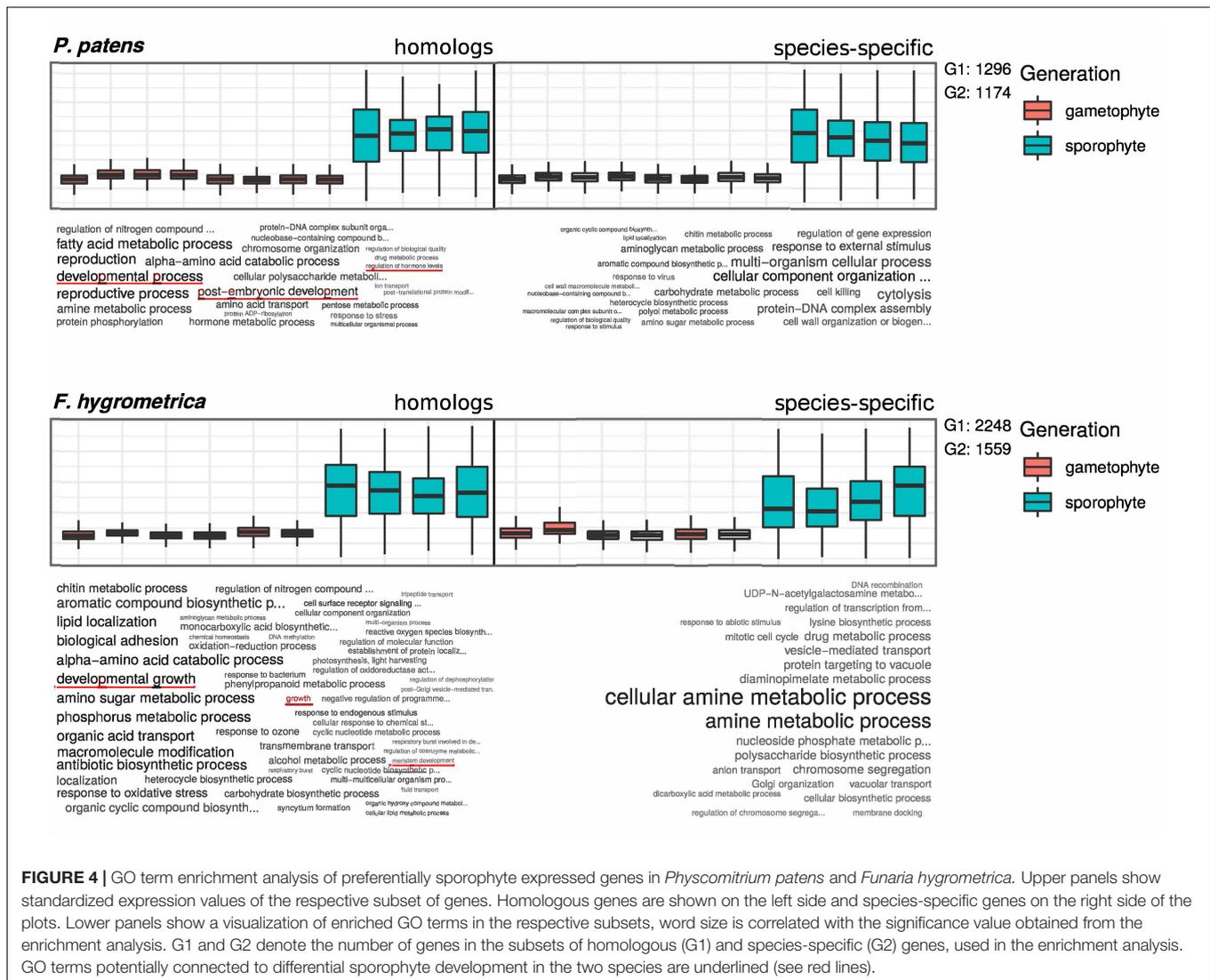


FIGURE 4 | GO term enrichment analysis of preferentially sporophyte expressed genes in *Physcomitrium patens* and *Funaria hygrometrica*. Upper panels show standardized expression values of the respective subset of genes. Homologous genes are shown on the left side and species-specific genes on the right side of the plots. Lower panels show a visualization of enriched GO terms in the respective subsets, word size is correlated with the significance value obtained from the enrichment analysis. G1 and G2 denote the number of genes in the subsets of homologous (G1) and species-specific (G2) genes, used in the enrichment analysis. GO terms potentially connected to differential sporophyte development in the two species are underlined (see red lines).

many redundant clusters. The optimal fuzzifier was computed according to the method described by Schwämmle and Jensen (2010), which is implemented in the mfuzz package (mestimate() function). Assignments of *F. hygrometrica* and *P. patens* one-to-one orthologs to the clusters can be found in **Supplementary Data S9, S10**.

RESULTS

To identify genes potentially underlying the difference in sporophyte development between *F. hygrometrica* and *P. patens*, we measured gene expression across four distinct stages of sporophyte development and several stages of gametophyte development in both species using RNA-sequencing (**Figure 2**). We assumed that most genes expressed or preferentially expressed in the sporophyte phase are important for proper sporophyte development. These data form the basis of our analyses described below.

Proportion of Preferentially Sporophyte Expressed Genes Is Slightly Higher in the Species With Greater Sporophyte Complexity

First, we performed a quantitative and qualitative assessment of the gene set expressed during sporophyte development in both species. After applying a relaxed expression threshold (see section “Materials and Methods”), 22,690 *P. patens* gene models were considered to be expressed in at least one of the sporophyte developmental stages, which is about 64% of the 35,307 currently annotated gene models (Lang et al., 2018). Similarly, of the total 25,904 gene models assembled using RNA-seq data of *F. hygrometrica* (see section “Materials and Methods”), 25,460 (98%) showed expression during sporophyte development. Of the gene set expressed during sporophyte and gametophyte development, 2,807 (12.4%) (*P. patens*) and 3,807 (15%) (*F. hygrometrica*) gene models were strongly upregulated in the sporophyte versus the gametophyte phase (\log_2 fold change

≥ 2 , false discovery rate ≤ 0.05). This suggests that the proportion of preferentially sporophyte expressed genes is slightly higher in the species with a more complex sporophyte morphology, but such genes represent only a relatively small proportion of the total gene set in both species.

Sporophytic Transcriptomes of *P. patens* and *F. hygrometrica* Are Dominated by Homologs but the Contribution of Species-Specific Genes Is Not Negligible

We then asked whether transcripts expressed during sporophyte development mainly represent genes homologous (including orthologs and paralogs, see identification of homologs in section “Identification of Homologs and Orthologs”) between the two species or, instead, species-specific genes. The Orthofinder analysis revealed that 19,667 and 19,267 genes of *P. patens* and *F. hygrometrica* fall into orthogroups containing sequences from both species. That is 55% (19,667 out of 35,307) and 74% (19,267 out of 25,904) of the *P. patens* and *F. hygrometrica* full gene set had homologs (orthologs and paralogs) in the alternate species’ proteome, respectively. Out of these homologs, 12,741 gene models were identified as one-to-one orthologs. When expressed as a proportion of genes transcribed throughout sporophyte development in *P. patens* and *F. hygrometrica*, 85.3% (19,364 of 22,690 gene models) and 75.6% (19,245 of 25,460 gene models), respectively, had homologs in the alternate species’ genome. These numbers did not change considerably when using the more relaxed BLASTp approach (see section “Identification of Homologs and Orthologs”) to include more distant homologs. To ensure that genes without homologs in either one of the species are species-specific and not missed by the annotation process, we searched protein sequences of putatively non-homologous genes against the corresponding genome sequence, using the tBLASTn algorithm (Camacho et al., 2009) with a threshold of $\geq 80\%$ query coverage and $\geq 35\%$ sequence similarity between the query and the corresponding hit. We identified an additional 2,003 gene models of *P. patens* to have putative homologs in the *F. hygrometrica* genome, which was expected due to the fact that the gene models for *F. hygrometrica* were assembled using RNAseq data and genes, which are not expressed in the sampled stages will not be detected by this method. Similarly, 980 additional gene models in the *F. hygrometrica* gene set have homologs in the *P. patens* genome. Combined with the expression data, 81% and 78% of the sporophytic transcriptome of *P. patens* and *F. hygrometrica*, respectively, thus have homologs in the alternate species’ genome. In summary, the two species express a large proportion of homologous (about 80%) genes during the development of their sporophyte. Nevertheless, about 20% of their genes expressed during sporophyte development appear species-specific, that is with no clear homologs in the other species’ genome.

F. hygrometrica Shows More Expression Variation During Sporophyte Development Than *P. patens*

Sporophyte development stages showed distinct expression patterns in *F. hygrometrica* in the shared (orthologs and

paralogs: homologs) and species-specific gene sets (Figure 3A). By contrast, differentiation in gene expression among the four sporophyte developmental stages in *P. patens* was strongly dependent on the gene set used. They were well-differentiated when we used all genes or only the set of homologs, whereas only the early and the late two developmental stages were well-distinguishable when comparing expression of species-specific genes (Figure 3A).

Species-Specific Genes Are Overrepresented Among Preferentially Sporophyte-Expressed Genes

Next, we tested how species-specificity or shared nature of genes is correlated with their putative function in sporophyte development. We assumed that preferential expression of genes in the sporophyte stage (compared to the gametophyte) can be used as a proxy for their functional role in sporophyte development. During gametophyte and sporophyte development 24,031 and 25,818 gene models were expressed in *P. patens* and *F. hygrometrica* of which 82.9% and 75.2% fell into orthogroups containing genes from both species, respectively. When using genes preferentially expressed in the sporophyte phase (\log_2 fold change ≥ 2 , false discovery rate ≤ 0.05), the proportion of genes coming from orthogroups shared by the two species decreased: only 72.6% (*P. patens*, 1,792 gene models) and 62.6% (*F. hygrometrica*, 2,385 gene models) of preferentially sporophyte expressed genes were homologous between the two species (Figure 3B). This difference in the proportion of homologous genes between preferentially sporophyte expressed and all expressed genes was highly significant in both species according to a chi-squared test (*P. patens*: $\chi^2_{df=1} = 206.16$, $p < 2.2 \times 10^{-16}$; *F. hygrometrica*: $\chi^2_{df=1} = 384.73$, $p < 2.2 \times 10^{-16}$). Altogether, this implies that species-specific genes are significantly overrepresented while homologous genes are underrepresented in the gene set preferentially expressed in the sporophyte.

Comparing gene expression at specific developmental stages between the two species (Figure 3C) using one-to-one orthologs reveals that all sporophytic stages of *P. patens* were most similar to the earliest sporophyte developmental stage of *F. hygrometrica* (stage number 1, see Figure 3C). Furthermore, overall expression similarity increased between *P. patens* and any of the *F. hygrometrica* developmental stages along a developmental chronology with the expression in the first stage of *P. patens* being most dissimilar. This suggests that the major differences in gene expression of orthologs can be found during early sporophyte development, a stage where precursors of various tissues that will form the mature sporophyte are likely established (Schwartz, 1997).

Functional Analysis of Genes Suggests a Contribution of Homologs in Establishing Contrasting Sporophyte Phenotypes

Comparing GO-terms of species-specific and homologous genes preferentially expressed in the sporophyte stage of both species (see section “Materials and Methods” for details) revealed that 103 (homologs) and 32 (species-specific) GO terms in

F. hygrometrica and 55 (homologs) and 50 (species-specific) GO terms in *P. patens* were significantly enriched in the corresponding subset (Fisher's exact test, $p \leq 0.05$, **Figure 4**). The majority of enriched GO-terms is related to various metabolic processes in both species. Nevertheless, some significantly enriched GO terms are potentially linked to the establishment of sporophyte morphology, which among others includes terms such as "regulation of growth", "developmental transitions", and "structure morphogenesis" (**Figure 4**). Importantly, only the homologous gene set was enriched for GO terms potentially linked to the morphology and development of the sporophyte phase in both species investigated (**Figure 4**). By contrast, the species-specific subset of preferentially sporophyte expressed genes were primarily enriched for general metabolic processes. In particular, homologous genes preferentially expressed in the sporophyte of *P. patens* were enriched for terms "developmental process," "post-embryonic development," and "regulation of hormone levels." Similarly, homologous genes preferentially expressed in the sporophyte in *F. hygrometrica* were enriched for GO-terms "developmental growth," "growth," and "meristem development" that are possibly linked to the establishment of the complex sporophyte morphology. By contrast, species-specific genes preferentially expressed in the sporophyte phase in both species were primarily enriched for GO terms related to general metabolic processes and not to differential development of sporophytic structures. Furthermore, while about 50% of the preferentially sporophyte expressed homologs in both species were associated with at least one GO-term, only 12% (*F. hygrometrica*) and 22% (*P. patens*) of the species-specific genes could be functionally characterized by one or more GO-terms, and consequently the contrasting pattern of homologous and species-specific genes may result from their sparse functional annotation.

Several Orthologous Developmental Regulators Show Heterochronic Expression in *P. patens* and *F. hygrometrica*

We used fuzzy c-mean clustering to identify groups of one-to-one orthologous genes experiencing differential expression dynamics throughout sporophyte development between *P. patens* and *F. hygrometrica*. We first clustered the *F. hygrometrica* gene expression data and then statistically assigned genes of *P. patens* to these clusters according to their expression dynamics. Clustering genes according to their expression profile over sporophyte development in *F. hygrometrica* resulted in 12 clusters with distinct expression patterns (referred to as minor clusters). We assigned those clusters to early, mid, and late sporophyte development (referred to as major clusters), based on the expression peak of each cluster (**Figure 5**). Following filtering (see section "Materials and Methods"), the data set contained 3,976 one-to-one orthologs assigned to one of the minor clusters. Of these, 1,807 (45.4%) were assigned to the same major cluster in both species, and 2,169 (54.6%) to different major clusters (**Supplementary Data S9**).

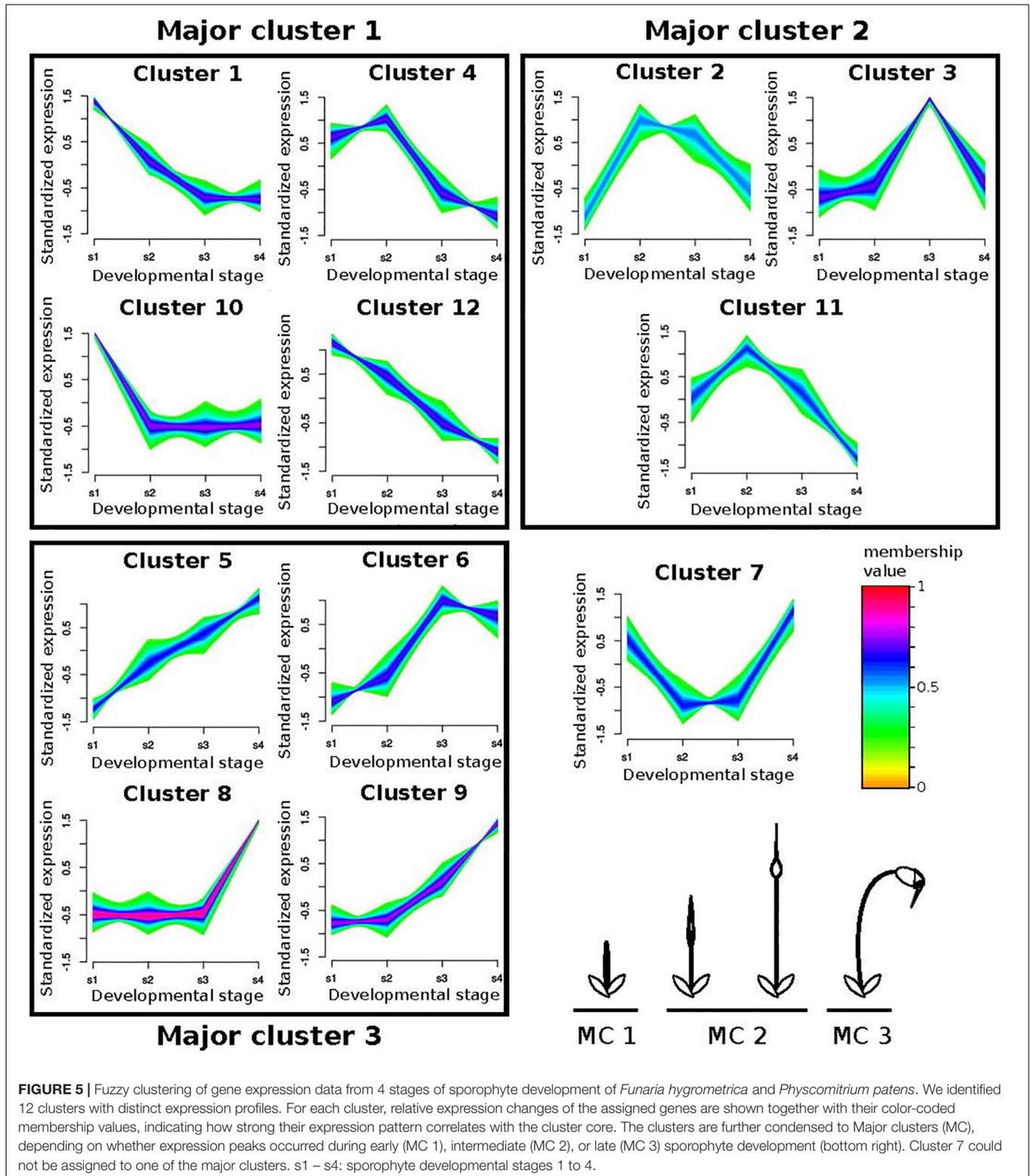
Looking at only transcription factor encoding genes did not influence our conclusion. Half of the 230 transcription factor encoding one-to-one orthologs present in the analyzed data set (i.e., 114, 51.8%) shifted between major clusters in *P. patens* and *F. hygrometrica* whereas the other half (i.e., 106, 48.2%) stayed in the same major cluster. These proportions were not significantly different from those when considering all expressed genes. Among the differentially expressed transcription factor encoding genes, many are members of gene families frequently associated with developmental control and growth in plants, including AP2/ERF, ARE, GRAS, MADS-box, TALE and WOX (see **Supplementary Data S10** for a complete table).

DISCUSSION

Morphological novelties may be associated with gene birth/death (*de novo* evolution of genes or loss of genes) or can alternatively be achieved by existing genes acquiring a new function (Davidson, 2010; Pires and Dolan, 2012; Chen et al., 2013; Schlötterer, 2015; Li et al., 2016; Das Gupta and Tsiantis, 2018; Bowman et al., 2019; Zhang et al., 2019). The relative contribution of these two major processes to the evolution of morphological novelties is still poorly known. Here, we compared the transcriptomes of two moss species, *P. patens* and *F. hygrometrica*, to investigate the relative contribution of gene gain/loss and shifts in expression dynamics of orthologous genes to the contrasting morphologies of their sporophyte phase. We found that the divergent sporophyte morphologies are primarily achieved by the heterochronic expression of a conserved set of genes, and while species-specific genes are likely to be also important, their contribution remains to be clarified. Our study contributes to the growing set of observations suggesting that shifting temporal dynamics of conserved genes represents a frequent mechanism through which radically different morphologies can be achieved (Sakamoto et al., 2009; Pires and Dolan, 2012; Lemmon et al., 2016; Olson and Nedelcu, 2016; Pham et al., 2017; Fisher et al., 2019). Moreover, it provides critical information on the molecular processes contributing to the diversification of moss sporophytes, a central topic of bryophyte taxonomy and evolutionary research. In the following paragraphs we discuss the implications of our findings for the molecular mechanisms underlying divergent sporophyte morphologies in *P. patens* and *F. hygrometrica* and for the evolution of morphological novelties in general.

Proportion of Preferentially Sporophyte-Expressed Genes Is Small

Developmental processes regulating the morphological properties of organisms are known to be composed of an interconnected set of genes organized in so-called regulatory networks (Davidson and Erwin, 2006; Davidson, 2010; Das Gupta and Tsiantis, 2018). Regulatory networks are modular, which makes them highly evolvable through the replacement of conserved modules and the rewiring of regulatory connections (Rosin and Kramer, 2009; Glassford et al., 2015; Halfon, 2017; Verd et al., 2019). Although many genes are expressed during



sporophyte development in both species, only about 12–15% indeed exhibit preferential sporophytic expression and are thus likely components of the genetic network underlying the development of the sporophyte versus the gametophyte.

Consequently, shifts from the gametophyte to the sporophyte program would only require changes in expression of a relatively small set of major genes. This is in line with multiple studies reporting how shifts in temporal or local expression of a

small set of genes can lead to radical change in morphology (Frankel et al., 2011; Ichihashi et al., 2014; Sicard et al., 2014; Di Ruocco et al., 2018; Whitewoods et al., 2020). Whether this assertion may also apply to the molecular mechanisms underlying contrasting sporophytic architectures is discussed in the paragraphs below.

Evidence for Control of Divergent Sporophyte Development by Conserved Orthologs Is Mounting, but Contribution of Species-Specific Genes Remains Unclear

Novel morphological structures can be acquired by the evolution of new genes or the differential usage of already existing gene sets, but the significance of either mechanism in shaping innovation or shifts in morphology remains ambiguous, and likely spans a broad range across lineages (Davidson, 2010; Pires and Dolan, 2012; Chen et al., 2013; Schlötterer, 2015; Li et al., 2016; Das Gupta and Tsiantis, 2018; Bowman et al., 2019; Zhang et al., 2019). Although the body plan of the moss sporophyte is fundamentally simple (Goffinet and Buck, 2013), the complexity of its mature architecture varies considerably (Vitt, 1981), and homoplastic reduction seems to be common across mosses (Schwartz, 1994; Shaw et al., 2000).

Within the Funariaceae the architecture of the sporophyte varies considerably, with *P. patens* and *F. hygrometrica* representing the two extremes of the spectrum (Liu et al., 2012; Medina et al., 2018, 2019). Their large-scale morphological difference lies in the length of the seta, the (a)symmetry of the spore bearing capsule and the presence/absence of appendages controlling spore dispersal (Fife, 1982). Sporophytes of *F. hygrometrica* consist of a 2–4.5 cm long stalk (seta) bearing an asymmetric capsule with subapical differentiated cells (annulus) enabling the release of a lid (operculum), and the then exposed mouth is lined by small hygroscopic teeth (peristome). By contrast, the sporophyte of *P. patens* is composed of an extremely short stalk, ending in a spherical capsule whose epidermal wall disintegrates when spores are mature (Figure 2).

Our study provides the first hints on the molecular mechanisms potentially underlying contrasting sporophyte architectures in *P. patens* and *F. hygrometrica*. We found that about half of the one-to-one orthologs between *P. patens* and *F. hygrometrica* show divergent expression dynamics during sporophyte development of the two species as demonstrated by our clustering analysis. Furthermore, of the transcription factor encoding orthologs that show heterochronic expression, many are members of gene families frequently associated with developmental control and growth in plants, including AP2/ERF, ARF, GRAS, MADS-box, TALE and WOX (see **Supplementary Data S10** for a complete table). Finally, our gene ontology analysis and previous functional studies on *P. patens* also provided evidence that some of these genes are likely to be key players in contributing to the divergent sporophyte phenotypes. Altogether, our data supports the notion that drastic macromorphological differences in sporophyte morphology (seta length and opening mechanisms of the capsule) are

likely achieved by the heterochronic expression of orthologous transcription factors.

Besides the overwhelming support for the contribution of orthologous transcription factors to the two contrasting sporophyte architectures, our analysis provides strong evidence that species-specific genes are also fundamental. We found that a large proportion of all (roughly 20%) and preferentially sporophyte expressed genes (between 30 and 40%) had no orthologs/homologs in the alternate species' genome and thus were truly species-specific. The proportion of species-specific genes between *P. patens* and *F. hygrometrica* appears to be high but not exceptional when considering their estimated divergence time of 60 million years (Medina et al., 2018). Comparable values can also be observed between closely related species pairs in angiosperms. For instance, *Arabidopsis thaliana* and *Arabidopsis lyrata* diverged 10 million years ago and roughly 16–18% of their gene sets are species-specific, respectively (Hu et al., 2011). Similar proportion of species-specific genes were also observed between the sister species *Zostera marina* and *Zostera muelleri*, which diverged from one another roughly 14 million years ago (Lee et al., 2016; Olsen et al., 2016). We further found that genes upregulated in the sporophyte phase were significantly enriched for species-specific genes suggesting their preferential recruitment for sporophytic functions. Unfortunately, species-specific genes are poorly annotated and thus their specific contribution in establishing divergent sporophyte architectures remains unclear and awaits clarification. In the next two paragraphs, we will focus on some major morphological features and their suggested molecular basis supported by our analyses.

Molecular Basis of Differential Seta Length in *P. patens* and *F. hygrometrica*

The seta develops through the activity of an intercalary meristem located immediately below the presumptive tissue of the capsule (Garner and Paolillo, 1973; French and Paolillo, 1976; Sano et al., 2005; Figure 2). The intercalary meristem is basipetal, such that cells are added below it and elongating, contributing to the growth of the seta. Thus, the length of the seta is likely correlated with the duration of activity of the intercalary meristem during sporophyte development. Given that the sporophyte of *P. patens* completes its development faster than *F. hygrometrica* (French and Paolillo, 1975; Sano et al., 2005; Coudert et al., 2019) as its meristem soon ceases to add new cells, it is therefore likely that expression of the underlying regulatory networks is shut down sooner or shortened compared to *F. hygrometrica*. That is, heterochronic expression of the meristem activity regulatory network can contribute to the seta length difference.

The class I KNOX gene *PpMKN2* promotes sporophyte axis extension by activating cytokinin biosynthesis pathways in *P. patens* (French and Paolillo, 1975; Coudert et al., 2019). In this species, *PpMKN2* is mainly expressed in the first sporophyte stage with expression quickly decreasing in more advanced stages of sporophyte development (Figure 5, cluster

10). By contrast, Fh_22445, the *F. hygrometrica* ortholog of *PpMKN2*, is characterized by a slower decline of expression reaching its minimum in the third and fourth phases of sporophyte development (Figure 5, cluster 1). The apparent correlation between difference in gene expression and degrees of seta length in the two species may suggest that the prolonged expression of the class I KNOX ortholog in *F. hygrometrica* accounts for the elongated seta phenotype.

A WUSCHEL-related homeobox13-like (WOX13-like) gene in *P. patens* (Pp3c15_20000) and its ortholog Fh_3609 also display differential gene expression dynamics in the two species. Fh_3609 is equally highly expressed during the first two developmental stages, but only weakly in other stages (Figure 5, cluster 4), whereas Pp3c15_20000 reaches an expression maximum only during the second stage of sporophyte development in *P. patens* (Figure 5, cluster 11). WOX13-like genes are important factors in initiating cell growth during zygote and stem cell formation by upregulating genes controlling cell wall loosening, including expansins (Sakakibara et al., 2014). This coincides with our finding that growth related genes are enriched in the set of preferentially sporophyte expressed genes in *F. hygrometrica*, among them three expansin homologs (Fh_21099, Fh_21985, Fh_4588). Furthermore, all of these expansins were assigned to the same major cluster as the WOX13-like gene Fh_3609 and two of them to the very same minor cluster. The three expansin homologs also have co-orthologous sequences in *P. patens* (Pp3c8_15200, Pp3c20_5780, Pp3c21_22300), but only Pp3c8_15200 is co-expressed with the WOX13-like gene Pp3c15_20000, whereas the other two show opposite expression patterns reaching their maximum expression in developmental stage 4 (Figure 5, cluster 5). Taken together, these results suggest that additional genes aiding in cell expansion and growth were recruited to the regulatory module initiated by expression of the WOX13-like gene Fh_3609 in *F. hygrometrica*, presumably to facilitate the rapid growth of the seta during early sporophyte development. Alternatively, regulatory connections between the WOX13-like gene of *P. patens* and some expansin homologs could have been lost in *P. patens*, contributing to the reduced growth of the seta.

Regulators of Abscission and Dehiscence Pathways Could Fulfill a Similar Function in Peristome and Annulus Development in *F. hygrometrica*

In addition to their highly contrasted seta length, the sporophytes of *F. hygrometrica* and *P. patens* differ in the presence/absence of specialized structures aiding in spore release and dispersal (Figure 2). The sporangium of *F. hygrometrica* opens via the loss of a lid (operculum), enabled by the differentiation of an annulus, a subapical ring of cells forming a predetermined breaking point between capsule body and operculum (Garner and Paolillo, 1973). The release of the operculum exposes the capsule mouth, which is lined by two rows of hygroscopic appendages, the peristome teeth, whose movement over the mouth at least partially controls the dispersal of spores.

The development of both structures, the annulus and peristome teeth, requires the establishment of tissue layers, spatially and temporally tightly regulated cell death and cell wall break down during development of the sporophyte (Budke et al., 2007; Goffinet et al., 2009; Goffinet and Buck, 2013), a process reminiscent of abscission and dehiscence found in a wide variety of plant species.

Control of abscission and dehiscence on a genetic level are well studied in vascular plants (Lenser and Theissen, 2013; Kim et al., 2019), but regulatory pathways remain unexplored in bryophytes. Central regulators of abscission zone (AZ) formation in *Arabidopsis thaliana* are the redundantly acting transcription factors *BLADE-ON-PETIOLE1* (*BOP1*) and *BOP2* (McKim et al., 2008; Wu et al., 2012; Khan et al., 2014), for which co-orthologs are present both in *F. hygrometrica* (Fh_25624, Fh_20164, Fh_1793) and *P. patens* (Pp3c17_8330, Pp3c1_35410, Pp3c14_11190). In *P. patens* the BOP co-orthologs show expression throughout sporophyte development with peaks in stages 1, 3 and 4, respectively. By contrast, expression of the *F. hygrometrica* BOP co-orthologs is more focused to intermediate developmental stages with Fh_25624 and Fh_1793 displaying peak expression in stage 2 and Fh_20164 being most strongly expressed in stage 3. The differentiation of capsule tissues begins after the seta has reached its maximum length (Garner and Paolillo, 1973), which happens around developmental stage 3. The shift of *F. hygrometrica* BOP ortholog expression to intermediate developmental stages when the capsule is not yet differentiated (stage 2) and differentiation begins (stage 3) could point to a potential role of these genes in establishing the boundary tissues that will later allow detachment of the operculum and formation of peristome teeth. However, BOP transcription factors are also involved in various other developmental and physiological processes in plants (Khan et al., 2014) and no apparent phenotypes in the sporophyte were reported in mutant analyses of PpBOP genes (Hata et al., 2019). Therefore, functional studies of BOP co-orthologs in *F. hygrometrica* are necessary to determine if these genes play a role in capsule tissue patterning that has been lost in funarioid and potentially also other species with a reduced sporophyte phenotype.

We also identified a pair of MIKCC-type MADS-box gene orthologs in *F. hygrometrica* (Fh_2640) and *P. patens* (Pp3c14_14900), which differ in their expression profile throughout sporophyte development. Fh_2640 shows an expression peak during late stages of sporophyte development (stage 3 and 4) whereas Pp3c14_14900 is primarily expressed during the elongation phase in stages 2 and 3. This shift in expression likely represents a change in the regulatory context of the orthologs, but whether this change is also accompanied by a functional diversification is currently unclear. MIKCC-type MADS-box genes are both expressed in the gametophyte and sporophyte phases in *P. patens* and are redundant negative regulators of cell division and growth in the gametophyte as well as sperm formation (Koshimizu et al., 2018). Their function in sporophyte development of mosses is poorly understood and debated. Koshimizu et al. (2018)

could not identify a phenotypic effect in knockout lines on sporophyte development in *P. patens* although all six MIKC^C MADS-box genes were strongly expressed in the sporophyte phase. By contrast, other studies reported well-recognizable sporophyte phenotypes when down regulating some MIKC^C genes in *P. patens* (Tanabe et al., 2003; Singer et al., 2007). We speculate that the MIKC^C-type MADS-box genes may also be important regulators of sporophyte development, potentially as negative regulators of cell division, growth and tissue patterning. The temporal expression shift of the *F. hygrometrica* ortholog to later stages of development may play a significant yet currently unknown role in the evolution of the more elaborate sporophyte morphology. This assumption is supported by the observations that down regulation of MIKC^C-type genes in *P. patens* led to abnormal swelling of the capsule (Singer et al., 2007) and that fluorescent reporter lines in *P. patens* showed that MIKC^C-type MADS-box genes have specific and complementary expression patterns throughout the development of the sporophyte. Altogether, it is possible that temporal expression shifts of MIKC^C-type MADS-box genes have contributed to the morphological diversification of sporophytes in Funarioid mosses, which parallels the observation that MIKC^C-type genes are key regulators in various aspects of sporophyte development including the diversification of flowers in angiosperms (Gramzow and Theissen, 2010). Nevertheless, sporophytic function of the MIKC^C-type genes in *P. patens* and *F. hygrometrica* must be determined in future studies on sporophyte development.

Implications for the Evolution of Sporophytic Characters in Mosses

The sporophyte is responsible for the production of spores, which will initiate the development of a new free-living gametophytic generation. Transformations of the sporophyte during the diversification of mosses affect a broad spectrum of traits, such as spore numbers, asymmetry of the capsule, mode of dehiscence, length of the seta, presence of stomata, responses of the capsule wall to dehydration, or architecture of the peristome controlling spore release (Crum, 2001). All these modifications may have a direct impact on fitness and are thus likely under natural selection (Vitt, 1981; Rose et al., 2016). Although some mosses have atypically amplified sporophytic traits [e.g., entomophilous Splachnaceae (Marino et al., 2009), or *Buxbaumia* (Crum, 2001)], and the macroevolutionary trend is one of increased complexity, repeated reductions through loss of complexity is rampant across the diversification of mosses, leading to capsules being immersed among vegetative leaves, lacking a complete peristome, stomata, or an operculum. Reduced sporophytic architecture characterizes many lineages distributed along a decreasing humidity gradient, and is common among epiphytic mosses, and particularly among xerophytic or short-lived annual mosses (Vitt, 1981; Rose et al., 2016). Finally, given the phylogenetic distribution of taxa with reduced morphology among congeners with more complex architectures, such as the three species traditionally referred to as *Physcomitrella* (highly reduced sporophytes) scattered among species of *Physcomitrium*

(with more elaborate sporophytes; Medina et al., 2019), reduction may result from repeated breakdown or shut-down of complex traits, and thus convergence may be readily achieved. Reduced sporophyte complexity of *P. patens* may be easily achieved by loss or reduced expression of particular *F. hygrometrica* genes involved in sporophyte development. This is in line with previous observations that temporal and/or spatial shifts in gene expression have significantly contributed to morphological diversification both in animal and plant systems (Frankel et al., 2011; Ichihashi et al., 2014; Rast-Somssich et al., 2015; Das Gupta and Tsiantis, 2018).

Our comparative study of transcriptomic profiles of developing sporophytes in *Physcomitrium patens* and *Funaria hygrometrica* provides the very first insights into the potential genetic tools shaping sporophyte morphologies. Our observations reveal that species-specific genes are preferentially expressed during sporophyte development in both species. While the actual function of these genes is poorly known, they may be involved in meiotic processes, physiological processes such as those related to endohydric, or metabolic pathways such as for components of the cuticle (Koch et al., 2009; Budke and Goffinet, 2016). Contrasts in genic expression levels suggest also that differential length of the seta and the gain/loss of a specific opening mechanism can be achieved relatively easily by heterochronic expression of major developmental regulators, and hence that sporophyte reduction can occur by temporal and/or spatial reprogramming of expression of conserved regulatory modules. Whether convergent evolution of reduced sporophyte morphologies is driven by similar molecular mechanisms across unrelated moss lineages must await similar comparative studies from independent lineages. Ultimately, once the critical genetic traits underlying reduction are identified, the hypothesis on their irreversibility/reversibility can be tested. Together these advances would be essential to further our understanding of trends in the diversification of the moss sporophyte, and their integration in systematic concepts reflecting the relationships among mosses with highly contrasted morphologies.

DATA AVAILABILITY STATEMENT

RNA-seq data generated for this study have been submitted to the European Nucleotide Archive (ENA) under study accession number PRJEB36328. The draft genome sequence of *F. hygrometrica*, assembled transcripts and their protein translations are provided on figshare (<https://doi.org/10.6084/m9.figshare.11663892.v1>).

AUTHOR CONTRIBUTIONS

AK, PS, and BG conceptualized and designed the study. MW and ZB grew *F. hygrometrica* sporophytes and gametophytes, respectively, extracted RNA and prepared libraries. MR and AN collected *P. patens* sporophytes, extracted RNA, and prepared

libraries. AK and PS analyzed the data and wrote the first version of the manuscript. All authors corrected and revised the final version of the manuscript.

FUNDING

This work was supported by the URPP Evolution in Action of the University of Zurich (PS), the Forschungskredit of the University of Zurich (MR, MW, and AK), grants of the Swiss National Science Foundation (PS 160004 and 131726), the Georges and Antoine Claraz Foundation (AK, AN, MW,

and PS), Plant Fellows Postdoctoral Fellowship (MR), the Foundation of German Business (sdw) (AN), and NSF grant DEB-1753811 (BG).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.00747/full#supplementary-material>

DATA SHEET S1 | This data can be found at the link: <https://doi.org/10.6084/m9.figshare.11663892.v1>.

REFERENCES

- Albalat, R., and Cañestro, C. (2016). Evolution by gene loss. *Nat. Rev. Genet.* 17, 379–391.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Bowles, A. M. C., Bechtold, U., and Paps, J. (2020). The origin of land plants is rooted in two bursts of genomic novelty. *Curr. Biol.* 30, 530–536.e6. doi: 10.1016/j.cub.2019.11.090
- Bowman, J. L., Briginshaw, L. N., and Florent, S. N. (2019). Evolution and co-option of developmental regulatory networks in early land plants. *Curr. Top. Dev. Biol.* 131, 35–53. doi: 10.1016/bs.ctdb.2018.10.001
- Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., et al. (2017). Insights into land plant evolution garnered from the marchantia polymorpha genome. *Cell* 171, 287–304.e6. doi: 10.1016/j.cell.2017.09.030
- Budke, J. M., and Goffinet, B. (2016). Comparative cuticle development reveals taller sporophytes are covered by thicker calyptra cuticles in mosses. *Front. Plant Sci.* 7:832. doi: 10.3389/fpls.2016.00832
- Budke, J. M., Goffinet, B., and Jones, C. S. (2012). The cuticle on the gametophyte calyptra matures before the sporophyte cuticle in the moss *Funaria hygrometrica* (Funariaceae). *Am. J. Bot.* 99, 14–22. doi: 10.3732/ajb.1100311
- Budke, J. M., Jones, C. S., and Goffinet, B. (2007). Development of the enigmatic peristome of *Timmia megapolitana* (Timmiaceae; Bryophyta). *Am. J. Bot.* 94, 460–467. doi: 10.3732/ajb.94.3.460
- Buendia-Monreal, M., and Gillmor, C. S. (2018). The times they are a-changin': heterochrony in plant development and evolution. *Front. Plant Sci.* 9:1349. doi: 10.3389/fpls.2018.01349
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform.* 10:421. doi: 10.1186/1471-2105-10-421
- Chen, S., Krinsky, B. H., and Long, M. (2013). New genes as drivers of phenotypic evolution. *Nat. Rev. Genet.* 14, 645–660. doi: 10.1038/nrg3521
- Cheng, S., Xian, W., Fu, Y., Marin, B., Keller, J., Wu, T., et al. (2019). Genomes of subaerial zygnematophyceae provide insights into land plant evolution. *Cell* 179, 1057–1067.e6. doi: 10.1016/j.cell.2019.10.019
- Clark, J. W., and Donoghue, P. C. J. (2018). Whole-genome duplication and plant macroevolution. *Trends Plant Sci.* 23, 933–945. doi: 10.1016/j.tplants.2018.07.006
- Coudert, Y., Novák, O., and Harrison, C. J. (2019). A KNOX-cytokinin regulatory module predates the origin of indeterminate vascular plants. *Curr. Biol.* 29, 2743–2750.e3. doi: 10.1016/j.cub.2019.06.083
- Crum, H. A. (2001). *Structural Diversity of Bryophytes*. Ann Arbor, MI: University of Michigan Herbarium.
- Das Gupta, M., and Tsiantis, M. (2018). Gene networks and the evolution of plant morphology. *Curr. Opin. Plant Biol.* 45, 82–87. doi: 10.1016/j.pbi.2018.05.011
- Davidson, E. H. (2010). Emerging properties of animal gene regulatory networks. *Nature* 468, 911–920. doi: 10.1038/nature09645
- Davidson, E. H., and Erwin, D. H. (2006). Gene regulatory networks and the evolution of animal body plans. *Science* 311, 796–800. doi: 10.1126/science.1113832
- Di Ruocco, G., Bertolotti, G., Pacifici, E., Polverari, L., Tsiantis, M., Sabatini, S., et al. (2018). Differential spatial distribution of miR165/6 determines variability in plant root anatomy. *Development* 145:dev153858. doi: 10.1242/dev.153858
- Emms, D. M., and Kelly, S. (2018). OrthoFinder2: fast and accurate phylogenomic orthology analysis from gene sequences. *bioRxiv*[Preprint]. doi: 10.1101/466201
- Erwin, D. H., and Davidson, E. H. (2009). The evolution of hierarchical gene regulatory networks. *Nat. Rev. Genet.* 10, 141–148. doi: 10.1038/nrg2499
- Fife, A. J. (1982). *A Generic Revision Of The Funariaceae (Bryophyta: Musci)*. Ann Arbor, MI: University Of Michigan.
- Fisher, C. R., Wegrzyn, J. L., and Jockusch, E. L. (2019). Co-option of wing-patterning genes underlies the evolution of the treehopper helmet. *Nat. Ecol. Evol.* 4, 250–260. doi: 10.1038/s41559-019-1054-4
- Frankel, N., Erezylmaz, D. F., McGregor, A. P., Wang, S., Payre, F., and Stern, D. L. (2011). Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* 474, 598–603. doi: 10.1038/nature10200
- French, J. C., and Paolillo, D. J. (1975). Effect of exogenously supplied growth regulators on intercalary meristematic activity and capsule expansion in *Funaria*. *Bryologist* 78:431. doi: 10.2307/3242165
- French, J. C., and Paolillo, D. J. (1976). Effect of the calyptra on intercalary meristematic activity in the sporophyte of *Funaria* (Musci). *Am. J. Bot.* 63, 492–498. doi: 10.1002/j.1537-2197.1976.tb11838.x
- Galis, F., Metz, J., and van Alphen, J. M. J. (2018). Development and evolutionary constraints in animals. *Annu. Rev. Ecol. Syst.* 49, 499–522. doi: 10.1146/annurev-ecolsys-110617-062339
- Furumizu, C., Alvarez, J. P., Sakakibara, K., and Bowman, J. L. (2015). Antagonistic roles for KNOX1 and KNOX2 genes in patterning the land plant body plan following an ancient gene duplication. *PLoS Genet.* 11:e1004980. doi: 10.1371/journal.pgen.1004980
- Garner, D., and Paolillo, D. J. (1973). A time-course of sporophyte development in *Funaria hygrometrica* hedw. *Bryologist* 76, 356–360.
- Geuten, K., and Coenen, H. (2013). Heterochronic genes in plant evolution and development. *Front. Plant Sci.* 4:381. doi: 10.3389/fpls.2013.00381
- Glassford, W. J., Johnson, W. C., Dall, N. R., Smith, S. J., Liu, Y., Boll, W., et al. (2015). Co-option of an ancestral hox-regulated network underlies a recently evolved morphological novelty. *Dev. Cell* 34, 520–531. doi: 10.1016/j.devcel.2015.08.005
- Goffinet, B., and Buck, W. (2013). “The evolution of body form in bryophytes,” in *Annual Plant Reviews*, Vol. 45, eds B. A. Ambrose and M. Purugganan (Chichester: John Wiley & Sons, Ltd), 51–90. doi: 10.1002/9781118305881.ch2
- Goffinet, B., Buck, W. R., and Shaw, A. J. (2009). Morphology, anatomy, and classification of the Bryophyta. *Bryophyte Biol.* 2, 55–138. doi: 10.1017/cbo9780511754807.003
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186. doi: 10.1093/nar/gkr944
- Gramzow, L., and Theissen, G. (2010). A hitchhiker's guide to the MADS world of plants. *Genome Biol.* 11:214. doi: 10.1186/gb-2010-11-6-214

- Grossmann, S., Bauer, S., Robinson, P. N., and Vingron, M. (2007). Improved detection of overrepresentation of gene-ontology annotations with parent child analysis. *Bioinformatics* 23, 3024–3031. doi: 10.1093/bioinformatics/btm440
- Gujas, B., Alonso-Blanco, C., and Hardtke, C. S. (2012). Natural *Arabidopsis* brx loss-of-function alleles confer root adaptation to acidic soil. *Curr. Biol.* 22, 1962–1968. doi: 10.1016/j.cub.2012.08.026
- Haig, D. (2013). Filial mistletoes: the functional morphology of moss sporophytes. *Ann. Bot.* 111, 337–345. doi: 10.1093/aob/mcs295
- Halfon, M. S. (2017). Perspectives on gene regulatory network evolution. *Trends Genet.* 33, 436–447. doi: 10.1016/j.tig.2017.04.005
- Hata, Y., Naramoto, S., and Kyoizuka, J. (2019). BLADE-ON-PETIOLE genes are not involved in the transition from protonema to gametophore in the moss *Physcomitrella patens*. *J. Plant Res.* 132, 617–627. doi: 10.1007/s10265-019-01132-8
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24, 417–441. doi: 10.1037/h0071325
- Hu, T. T., Pattyn, P., Bakker, E. G., Cao, J., Cheng, J.-F., Clark, R. M., et al. (2011). The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nat. Genet.* 43, 476–481. doi: 10.1038/ng.807
- Ichihashi, Y., Aguilar-Martínez, J. A., Farhi, M., Chitwood, D. H., Kumar, R., Millon, L. V., et al. (2014). Evolutionary developmental transcriptomics reveals a gene network module regulating interspecific diversity in plant leaf shape. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2616–E2621. doi: 10.1073/pnas.1402835111
- Jill Harrison, C. (2017). Development and genetics in the evolution of land plant body plans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372:20150490. doi: 10.1098/rstb.2015.0490
- Jin, J., Tian, F., Yang, D.-C., Meng, Y.-Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982
- Jonathan Shaw, A., and Goffinet, B. (2000). *Bryophyte Biology*. Cambridge: Cambridge University Press.
- Kaessmann, H. (2010). Origins, evolution, and phenotypic impact of new genes. *Genome Res.* 20, 1313–1326. doi: 10.1101/gr.101386.109
- Khalturin, K., Hemmrich, G., Fraune, S., Augustin, R., and Bosch, T. C. G. (2009). More than just orphans: are taxonomically-restricted genes important in evolution? *Trends Genet.* 25, 404–413. doi: 10.1016/j.tig.2009.07.006
- Khan, M., Xu, H., and Hepworth, S. R. (2014). BLADE-ON-PETIOLE genes: setting boundaries in development and defense. *Plant Sci.* 215–216, 157–171. doi: 10.1016/j.plantsci.2013.10.019
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Kim, J., Chun, J.-P., and Tucker, M. L. (2019). Transcriptional regulation of abscission zones. *Plants* 8:E154. doi: 10.3390/plants8060154
- Kim, J. H. (2019). *Genome Data Analysis*. Berlin: Springer.
- Koch, K., Frahm, J.-P., and Pollawat, R. (2009). The cuticle of the *Buxbaumia viridis* sporophyte. *Flora Morphol. Distrib. Funct. Ecol. Plants* 204, 34–39. doi: 10.1016/j.flora.2007.11.007
- Kolde, R., and Vilo, J. (2015). GOSummary: an R package for visual functional annotation of experimental data. *F1000Res.* 4:574. doi: 10.12688/f1000research.6925.1
- Koshimizu, S., Kofuji, R., Sasaki-Sekimoto, Y., Kikkawa, M., Shimojima, M., Ohta, H., et al. (2018). Physcomitrella MADS-box genes regulate water supply and sperm movement for fertilization. *Nat. Plants* 4, 36–45. doi: 10.1038/s41477-017-0082-9
- Kriventseva, E. V., Kuznetsov, D., Tegenfeldt, F., Manni, M., Dias, R., Simão, F. A., et al. (2019). OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Res.* 47, D807–D811. doi: 10.1093/nar/gky1053
- Kumar, L., and Futschik, M. E. (2007). Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics* 2, 5–7. doi: 10.6026/97320630002005
- Landis, J. B., Soltis, D. E., Li, Z., Marx, H. E., Barker, M. S., Tank, D. C., et al. (2018). Impact of whole-genome duplication events on diversification rates in angiosperms. *Am. J. Bot.* 105, 348–363. doi: 10.1002/ajb2.1060
- Lang, D., Ullrich, K. K., Murat, F., Fuchs, J., Jenkins, J., Haas, F. B., et al. (2018). The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J.* 93, 515–533. doi: 10.1111/tpj.13801
- Lee, H., Golicz, A. A., Bayer, P. E., Jiao, Y., Tang, H., Paterson, A. H., et al. (2016). The genome of a southern hemisphere seagrass species (*Zostera muelleri*). *Plant Physiol.* 172, 272–283. doi: 10.1104/pp.16.00868
- Lemma, Z. H., Park, S. J., Jiang, K., Van Eck, J., Schatz, M. C., and Lippman, Z. B. (2016). The evolution of inflorescence diversity in the nightshades and heterochrony during meristem maturation. *Genome Res.* 26, 1676–1686. doi: 10.1101/gr.207837.116
- Lenser, T., and Theißen, G. (2013). Conservation of fruit dehiscence pathways between *Lepidium campestre* and *Arabidopsis thaliana* sheds light on the regulation of INDEHISCENT. *Plant J.* 76, 545–556. doi: 10.1111/tpj.12321
- Li, Z.-W., Chen, X., Wu, Q., Hagmann, J., Han, T.-S., Zou, Y.-P., et al. (2016). On the origin of De Novo genes in *Arabidopsis thaliana* populations. *Genome Biol. Evol.* 8, 2190–2202. doi: 10.1093/gbe/evw164
- Liu, Y., Budke, J. M., and Goffinet, B. (2012). Phylogenetic inference rejects sporophyte based classification of the Funariaceae (Bryophyta): rapid radiation suggests rampant homoplasy in sporophyte evolution. *Mol. Phylogenet. Evol.* 62, 130–145. doi: 10.1016/j.ympev.2011.09.010
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1:18. doi: 10.1186/2047-217X-1-18
- MacArthur, D. G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., et al. (2012). A systematic survey of loss-of-function variants in human protein-coding genes. *Science* 335, 823–828. doi: 10.1126/science.1215040
- Marino, P., Raguso, R., and Goffinet, B. (2009). The ecology and evolution of fly dispersed dung mosses (Family Splachnaceae): manipulating insect behaviour through odour and visual cues. *Symbiosis* 47, 61–76. doi: 10.1007/bf03182289
- McKim, S. M., Stenvik, G.-E., Butenko, M. A., Kristiansen, W., Cho, S. K., Hepworth, S. R., et al. (2008). The BLADE-ON-PETIOLE genes are essential for abscission zone formation in *Arabidopsis*. *Development* 135, 1537–1546. doi: 10.1242/dev.012807
- Medina, R., Johnson, M., Liu, Y., Wilding, N., Hedderson, T. A., Wickett, N., et al. (2018). Evolutionary dynamism in bryophytes: phylogenomic inferences confirm rapid radiation in the moss family Funariaceae. *Mol. Phylogenet. Evol.* 120, 240–247. doi: 10.1016/j.ympev.2017.12.002
- Medina, R., Johnson, M. G., Liu, Y., Wickett, N. J., Shaw, A. J., and Goffinet, B. (2019). Phylogenomic delineation of *Physcomitrium* (Bryophyta: Funariaceae) based on targeted sequencing of nuclear exons and their flanking regions rejects the retention of *Physcomitrella*, *Physcomitridium* and *Aphanorhagma*. *J. Syst. Evol.* 57, 404–417. doi: 10.1111/jse.12516
- Nachman, M. W., Hoekstra, H. E., and D’Agostino, S. L. (2003). The genetic basis of adaptive melanism in pocket mice. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5268–5273. doi: 10.1073/pnas.0431157100
- Nadeau, N. J., Pardo-Diaz, C., Whibley, A., Supple, M. A., Saenko, S. V., Wallbank, R. W. R., et al. (2016). The gene cortex controls mimicry and crypsis in butterflies and moths. *Nature* 534, 106–110. doi: 10.1038/nature17961
- Olsen, J. L., Rouzép, P., Verhelst, B., Lin, Y.-C., Bayer, T., Collen, J., et al. (2016). The genome of the seagrass *Zostera marina* reveals angiosperm adaptation to the sea. *Nature* 530, 331–335. doi: 10.1038/nature16548
- Olson, B. J., and Nedelcu, A. M. (2016). Co-option during the evolution of multicellular and developmental complexity in the volvocine green algae. *Curr. Opin. Genet. Dev.* 39, 107–115. doi: 10.1016/j.gde.2016.06.003
- Olson, M. V. (1999). When less is more: gene loss as an engine of evolutionary change. *Am. J. Hum. Genet.* 64, 18–23. doi: 10.1086/302219
- One Thousand Plant Transcriptomes Initiative (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574, 679–685. doi: 10.1038/s41586-019-1693-2
- Orr, H. A. (2005). The genetic theory of adaptation: a brief history. *Nat. Rev. Genet.* 6, 119–127. doi: 10.1038/nrg1523
- Pearson, W. R. (2013). An introduction to sequence similarity (“homology”) searching. *Curr. Protoc. Bioinform.* Chapter 3, Unit3.1. doi: 10.1002/0471250953.bi0301s42
- Pederson, E. R. A., Warshan, D., and Rasmussen, U. (2019). Genome sequencing of *Pleurozium schreberi*: the assembled and annotated draft genome of a pleurocarpus feather moss. *G3* 9, 2791–2797. doi: 10.1534/g3.119.400279

- Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122
- Pham, T., Day, S. M., Glassford, W. J., Williams, T. M., and Rebeiz, M. (2017). The evolutionary origination of a novel expression pattern through an extreme heterochronic shift. *Evol. Dev.* 19, 43–55. doi: 10.1111/ede.12215
- Pires, N. D., and Dolan, L. (2012). Morphological evolution in land plants: new designs with old genes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 508–518. doi: 10.1098/rstb.2011.0252
- Porebski, S., Bailey, L. G., and Baum, B. R. (1997). Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol. Biol. Rep.* 15, 8–15. doi: 10.1007/bf02772108
- Proost, S., Van Bel, M., Sterck, L., Billiau, K., Van Parys, T., Van de Peer, Y., et al. (2009). PLAZA: a comparative genomics resource to study gene and genome evolution in plants. *Plant Cell* 21, 3718–3731. doi: 10.1105/tpc.109.071506
- Rast-Somssich, M. I., Broholm, S., Jenkins, H., Canales, C., Vlad, D., Kwantes, M., et al. (2015). Alternate wiring of a KNOXI genetic network underlies differences in leaf development of *A. thaliana* and *C. hirsuta*. *Genes Dev.* 29, 2391–2404. doi: 10.1101/gad.269050.115
- Rebeiz, M., and Tsiantis, M. (2017). Enhancer evolution and the origins of morphological novelty. *Curr. Opin. Genet. Dev.* 45, 115–123. doi: 10.1016/j.gde.2017.04.006
- Reski, R., and Abel, W. O. (1985). Induction of budding on chloronemata and plantonemata of the moss, *Physcomitrella patens*, using isopentenyladenine. *Planta* 165, 354–358. doi: 10.1007/BF00392232
- Rose, J. P., Kriebel, R., and Sytsma, K. J. (2016). Shape analysis of moss (Bryophyta) sporophytes: Insights into land plant evolution. *Am. J. Bot.* 103, 652–662. doi: 10.3732/ajb.1500394
- Rosin, F. M., and Kramer, E. M. (2009). Old dogs, new tricks: regulatory evolution in conserved genetic modules leads to novel morphologies in plants. *Dev. Biol.* 332, 25–35. doi: 10.1016/j.ydbio.2009.05.542
- Rost, B. (1999). Twilight zone of protein sequence alignments. *Protein Eng. Design Select.* 12, 85–94. doi: 10.1093/protein/12.2.85
- Sakakibara, K., Reisewitz, P., Aoyama, T., Friedrich, T., Ando, S., Sato, Y., 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. doi: 10.1242/dev.097444
- Sakamoto, K., Onimaru, K., Munakata, K., Suda, N., Tamura, M., Ochi, H., et al. (2009). Heterochronic shift in Hox-mediated activation of sonic hedgehog leads to morphological changes during fin development. *PLoS One* 4:e5121. doi: 10.1371/journal.pone.0005121
- Sano, R., Juárez, C. M., Hass, B., Sakakibara, K., Ito, M., Banks, J. A., et al. (2005). KNOX homeobox genes potentially have similar function in both diploid unicellular and multicellular meristems, but not in haploid meristems. *Evol. Dev.* 7, 69–78. doi: 10.1111/j.1525-142x.2005.05008.x
- Schlötterer, C. (2015). Genes from scratch—the evolutionary fate of de novo genes. *Trends Genet.* 31, 215–219. doi: 10.1016/j.tig.2015.02.007
- Schwämmle, V., and Jensen, O. N. (2010). A simple and fast method to determine the parameters for fuzzy c-means cluster analysis. *Bioinformatics* 26, 2841–2848. doi: 10.1093/bioinformatics/btq534
- Schwartz, O. M. (1994). The development of the peristome-forming layers in the Funariaceae. *Int. J. Plant Sci.* 155, 640–657. doi: 10.1086/297204
- Schwartz, O. M. (1997). Apical cell segmentation and its relationship to the peristome-forming layers in the Funariaceae. *Int. J. Plant Sci.* 158, 236–248. doi: 10.1086/297435
- Seppy, M., Manni, M., and Zdobnov, E. M. (2019). “BUSCO: assessing genome assembly and annotation completeness,” in *Gene Prediction: Methods and Protocols*, ed. M. Kollmar (New York, NY: Springer), 227–245. doi: 10.1007/978-1-4939-9173-0_14
- Shaw, A. J., Anderson, L. E., and Mishler, B. D. (2000). Paedomorphic sporophyte development in *Bruchia flexuosa* (Bruchiaceae). *Bryologist* 103, 147–155. doi: 10.1639/0007-2745(2000)103[0147:psdibf]2.0.co;2
- Sicard, A., Thamm, A., Marona, C., Lee, Y. W., Wahl, V., Stinchcombe, J. R., et al. (2014). Repeated evolutionary changes of leaf morphology caused by mutations to a homeobox gene. *Curr. Biol.* 24, 1880–1886. doi: 10.1016/j.cub.2014.06.061
- Singer, S. D., Krogan, N. T., and Ashton, N. W. (2007). Clues about the ancestral roles of plant MADS-box genes from a functional analysis of moss homologues. *Plant Cell Rep.* 26, 1155–1169. doi: 10.1007/s00299-007-0312-0
- Soltis, P. S., and Soltis, D. E. (2016). Ancient WGD events as drivers of key innovations in angiosperms. *Curr. Opin. Plant Biol.* 30, 159–165. doi: 10.1016/j.pbi.2016.03.015
- Suarez-Gonzalez, A., Lexer, C., and Cronk, Q. C. B. (2018). Adaptive introgression: a plant perspective. *Biol. Lett.* 14:20170688. doi: 10.1098/rsbl.2017.0688
- Sun, B., Looi, L.-S., Guo, S., He, Z., Gan, E.-S., Huang, J., et al. (2014). Timing mechanism dependent on cell division is invoked by Polycomb eviction in plant stem cells. *Science* 343:1248559. doi: 10.1126/science.1248559
- Sun, G., Xu, Y., Liu, H., Sun, T., Zhang, J., Hettenhausen, C., et al. (2018). Large-scale gene losses underlie the genome evolution of parasitic plant *Cuscuta australis*. *Nat. Commun.* 9:2683. doi: 10.1038/s41467-018-04721-8
- Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6:e21800. doi: 10.1371/journal.pone.0021800
- Tanabe, Y., Uchida, M., Hasebe, M., and Ito, M. (2003). Characterization of the *Selaginella remotifolia* MADS-box gene. *J. Plant Res.* 116, 71–75.
- True, J. R., and Carroll, S. B. (2002). Gene co-option in physiological and morphological evolution. *Annu. Rev. Cell Dev. Biol.* 18, 53–80. doi: 10.1146/annurev.cellbio.18.020402.140619
- Van de Peer, Y., Mizrachi, E., and Marchal, K. (2017). The evolutionary significance of polyploidy. *Nat. Rev. Genet.* 18, 411–424. doi: 10.1038/nrg.2017.26
- Verd, B., Monk, N. A., and Jaeger, J. (2019). Modularity, criticality, and evolvability of a developmental gene regulatory network. *eLife* 8:42832. doi: 10.7554/eLife.42832
- Vitt, D. H. (1981). Adaptive modes of the moss sporophyte. *Bryologist* 84, 166–186.
- Wagner, G. P., and Lynch, V. J. (2010). Evolutionary novelties. *Curr. Biol.* 20, R48–R52. doi: 10.1016/j.cub.2009.11.010
- Whitewoods, C. D., Cammarata, J., Nemeč VENZA, Z., Sang, S., Crook, A. D., Aoyama, T., et al. (2018). CLAVATA was a genetic novelty for the morphological innovation of 3D growth in land plants. *Curr. Biol.* 28, 2365–2376.e5. doi: 10.1016/j.cub.2018.05.068
- Whitewoods, C. D., Gonçalves, B., Cheng, J., Cui, M., Kennaway, R., Lee, K., et al. (2020). Evolution of carnivorous traps from planar leaves through simple shifts in gene expression. *Science* 367, 91–96. doi: 10.1126/science.aay5433
- Wickell, D. A., and Li, F.-W. (2020). On the evolutionary significance of horizontal gene transfers in plants. *New Phytol.* 225, 113–117. doi: 10.1111/nph.16022
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Berlin: Springer Science & Business Media.
- Wu, X.-M., Yu, Y., Han, L.-B., Li, C.-L., Wang, H.-Y., Zhong, N.-Q., et al. (2012). The tobacco BLADE-ON-PETIOLE2 gene mediates differentiation of the corolla abscission zone by controlling longitudinal cell expansion. *Plant Physiol.* 159, 835–850. doi: 10.1104/pp.112.193482
- Xu, Y.-C., Niu, X.-M., Li, X.-X., He, W., Chen, J.-F., Zou, Y.-P., et al. (2019). Adaptation and phenotypic diversification in *Arabidopsis* through loss-of-function mutations in protein-coding genes. *Plant Cell* 31, 1012–1025. doi: 10.1105/tpc.18.00791
- Zhang, L., Ren, Y., Yang, T., Li, G., Chen, J., Gschwend, A. R., et al. (2019). Rapid evolution of protein diversity by de novo origination in *Oryza*. *Nat. Ecol. Evol.* 3, 679–690. doi: 10.1038/s41559-019-0822-5
- Zhao, Y., Zhang, T., Broholm, S. K., Tähtiharju, S., Mouhu, K., Albert, V. A., et al. (2016). Evolutionary co-option of floral meristem identity genes for patterning of the flower-like asteraceae inflorescence. *Plant Physiol.* 172, 284–296. doi: 10.1104/pp.16.00779

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Kirbis, Waller, Ricca, Bont, Neubauer, Goffinet and Szövényi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.