Annals of Botany **XX**: 1–17, 2025
https://doi.org/10.1093/aob/mcaf012, available online at www.academic.oup.com/aob

ANNALS OF BOTANY
Founded 1887

ORIGINAL ARTICLE

Distinct patterns of genome size evolution in each bryophyte lineage are not correlated with whole genome duplication

Nikisha Patel[1,*], Jessica M. Budke[2] and Jillian D. Bainard[3]

[1]Biology Department, Trinity College, 300 Summit Street, Hartford, CT 06106, USA, [2]Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA and [3]Agassiz Research and Development Centre, Agriculture and Agri-food Canada, Agassiz, BC V0M 1A2, British Columbia, Canada
[*]For correspondence. E-mail nikisha.patel@trincoll.edu

Received: 2 January 2025 Returned for revision: 2 January 2025 Editorial decision: 20 January 2025 Accepted: 25 January 2025

• **Background and Aims** Genome size varies by orders of magnitude across land plants, and the factors driving evolutionary increases and decreases in genome size vary across lineages. Bryophytes have the smallest genomes relative to other land plants, and there is growing evidence for frequent whole genome duplication (WGD) across the lineage. However, the broad patterns of genome size, chromosome number and WGD have yet to be characterized across bryophytes in a phylogenetic context.
• **Methods** In the present study, we adopt a phylogenetic comparative approach and leverage previously published data on genome size, chromosome number and WGD to reconstruct the evolutionary history of these traits across the three major bryophyte lineages: hornworts, liverworts and mosses. We infer ancestral haploid chromosome numbers for each lineage and introduce a novel metric for assessing polyploidy using chromosome counts.
• **Key Results** Each lineage of bryophytes exhibits a distinct pattern of genome size evolution and prevalence of WGD, with mosses having the most dynamic genome sizes and highest propensity for WGD. We found that 21.3 % of mosses and 13 % of liverworts species have naturally occurring polyploids. In addition, haploid genome size (1C) is most dynamic in the mosses, which includes at least 15 transitions to larger genomes and nine reversals, largely in the orders Dicranales and Hypnales.
• **Conclusions** There is no correlation between genome size and WGD or between genome size and chromosome number, potentially suggesting rapid genome downsizing following WGD. Given that bryophytes are poikilohydric (desiccation-tolerant) plants, having large genomes might be physiologically prohibitive given the cost to growth and metabolism associated with them. These findings emphasize the unique evolution of the bryophytes broadly and of the hornworts, liverworts and mosses individually, and should therefore serve as impetus for more in-depth experimental studies of genome size evolution and WGD in bryophytes.

Key words: Polyploidy, genome size, chromosome number, bryophytes, mosses, hornworts, liverworts.

INTRODUCTION

The evolution in genome size and chromosome number across eukaryotes has long been of interest to biologists and is a crucial area of study given their implications for both speciation and diversification (Leitch and Leitch, 2012; Van de Peer *et al.*, 2017). Across plants, genome size and chromosome number vary substantially (e.g. 1C = 0.07–152.23 pg and $2n$ = 4–1440; Uhl, 1978; Bennett *et al.*, 1986; Khandelwal, 1990; Pellicer *et al.*, 2010; Leitch and Leitch, 2012; Fleischmann *et al.*, 2014). These two traits, haploid genome size (1C) and chromosome number ($1n$), are not necessarily correlated, and the nature of their relationship is highly lineage dependent. For example, the pteridophytes include several species with both exceptionally large chromosome numbers and large genomes (Clark *et al.*, 2016; de Abreu *et al.*, 2024; Fernández *et al.*, 2024), and broadly across ferns, genome size and chromosome number are positively correlated (Nakazato *et al.*, 2008; Patel *et al.*, 2018; Liu *et al.*, 2019; Barrington *et al.*, 2020). In contrast, in the flowering plants, positive correlations between genome size and chromosome number are observed in some lineages (Pandit *et al.*, 2014), but not others (Fleischmann *et al.*, 2014).

The mechanisms underlying evolutionary shifts in genome size and chromosome number can impact one or both of these traits and the nature of their relationship (Bennetzen *et al.*, 2005; Wang *et al.*, 2021). In plants, genome size increase is driven primarily by repeat elements and whole genome duplication (WGD) (Carta *et al.*, 2020). The accumulation of

repetitive elements, largely long terminal repeat retrotransposons, in centromeric regions is a major contributor to the increase in genome size in species with both small and large genomes, including *Arabidopsis* (1C = 0.13 pg; Arabidopsis Genome Initiative, 2000; Zhang and Wessler, 2004) and maize (1C = 2.56 pg; Sanmiguel and Bennetzen, 1998; Anderson *et al.*, 2019). In addition, increases in genome size in gymnosperms are largely attributable to the accumulation of repetitive elements (Liu *et al.*, 2022*b*). Although repetitive elements impact total DNA content rather than chromosome number and can accumulate rapidly or gradually, WGD increases both genome size and chromosome number through genome doubling, which occurs instantaneously (Soltis *et al.*, 2009; Barker *et al.*, 2016; Szövényi *et al.*, 2021). Rapid shifts in genome size and composition associated with this process have been tied to speciation and diversification (Jiao *et al.*, 2012; Landis *et al.*, 2018). Although neopolyploids may establish as new species, the nature of polyploidy as a driver of speciation in the long term remains contentious (Mayrose *et al.*, 2015; Soltis *et al.*, 2015) The propensity of plant genomes to accumulate repetitive elements and experience WGD can lead to 'genomic obesity', barring the subsequent loss of DNA through unequal recombination (Wicker *et al.*, 2003). WGD can instead be followed by gradual diploidization, or genic loss through successive chromosomal rearrangements (del Pozo and Ramirez-Parra, 2015; Li *et al.*, 2021). Crucially, this can result in lineages with haploid chromosome numbers distinct from progenitors and close relatives (Mandáková and Lysak, 2018), potentially resulting in speciation. Although diploidization reduces chromosome number and genome size in some land plant lineages (Soltis *et al.*, 2015; Mandáková *et al.*, 2018), others are known to retain most or all chromosomes following WGD over evolutionary time. For instance, it has been proposed that ferns become functionally diploid via gene silencing rather than gene loss (Haufler, 1987; Liu *et al.*, 2019; Huang *et al.*, 2022). In either instance, WGD is a major mechanism by which chromosome number and genome size can change both immediately and over evolutionary time, potentially influencing speciation and diversification.

The genome size of a given lineage is also influenced by selection (Levin, 2002). Some environmental conditions might favour smaller genome sizes because lower nuclear DNA content is associated with faster growth, shorter generation time and a higher metabolic rate (e.g. Bennett, 1972; Gregory, 2001); thus smaller genomes are often correlated with invasiveness (Pandit *et al.*, 2014; Suda *et al.*, 2015). Indeed, studies of genome size across angiosperms in a phylogenetic framework posit selection for smaller genomes as a means of constraining cell size and optimizing stomatal density, thereby directly influencing primary productivity (Knight *et al.*, 2005; Simonin and Roddy, 2018; Escudero and Wendel, 2020). In contrast, in spite of the metabolic cost of retaining large genomes, they might allow for more rapid evolution of novel genes and adaptive traits (Francis *et al.*, 2008; Pellicer *et al.*, 2018; Nieto Feliner *et al.*, 2020).

Bryophytes, which are sister to the rest of land plants, remain understudied in terms of broad evolutionary trends in genome size and chromosomal evolution (Rensing *et al.*, 2012; Duckett, 2020; Linde *et al.*, 2021; Bechteler *et al.*, 2023). Relative to other major plant lineages, most bryophyte genome sizes are small, with an average haploid genome size of 0.92 pg (N = 334, Kew C-values database; Leitch *et al.*, 2019) and a range from 1C = 0.73 pg in the hornwort *Nothoceros endiviifolius* (Bainard and Villarreal, 2013) to 1C = 20.46 pg in the liverwort *Phyllothallia fuegiana* (Bainard *et al.*, 2013). The predominantly small genomes might be, in part, a product of selection, given that these plants are also desiccation tolerant and smaller genomes might confer metabolic benefits (Bainard, 2011), and genomic studies of the bryophytes demonstrate historical gene loss in the ancestors of bryophytes (Harris *et al.*, 2022).

Although the haploid nuclear genome size (1C) is generally small in bryophytes, haploid chromosome number (1n) varies significantly both within and across species (Patel *et al.*, 2021), suggesting that WGD plays a significant role in the evolution of this lineage. Although historically, WGD, either conspecific genome doubling (autopolyploidy) or genome doubling following hybridization (allopolyploidy), was thought to play a small role in the evolution, speciation and diversification of bryophytes (Vitt, 1971), more recent investigations underscore the importance of WGD and hybridization to the evolution of the group (Jesson *et al.*, 2011; Rensing *et al.*, 2012; Perley and Jesson, 2015; Bechteler *et al.*, 2023; Shen *et al.*, 2024). Numerous instances of sterile hybrids, fertile hybrids and allopolyploid lineages have been observed in bryophytes (Natcheva and Cronberg, 2004). In addition, Patel *et al.* (2021) investigated intraspecific variation in moss chromosome data to identify cryptic autopolyploid species, which might constitute 17 % of total moss diversity, representing a major source of undescribed diversity among mosses.

In light of extant bryophytes having relatively small haploid genomes, in addition to orders of magnitude of interspecific and intraspecific variation in chromosome number suggesting WGD, several important questions arise. First, what are the patterns of bryophyte genome size in relationship to WGD across their evolution? If WGD is a major driver of genome size variation in the bryophytes, then we might expect bryophyte lineages with a higher frequency of WGD also to have larger haploid genome sizes. Second, how do these patterns vary among the major lineages of bryophytes, the Anthocerophyta (hornworts), Marchantiophyta (liverworts) and Bryophyta (mosses)? Although these three lineages are often lumped together in broad studies of plant evolution, the hornworts diverged from liverworts and mosses ~400 Mya, and liverworts and mosses diverged from each other ~300 Mya (Bechteler *et al.*, 2023). In addition, although the bryophytes are often considered 'ancient' given that they are sister to vascular plants, the majority of extant bryophyte genera evolved much more recently, and many are considered part of a recent diversification event (Laenen *et al.*, 2014). It is therefore crucial to consider patterns of evolution in the hornworts, liverworts and mosses individually (Lang *et al.*, 2018; Li *et al.*, 2020; Zhang *et al.*, 2020; Linde *et al.*, 2023).

In the present study, we investigated the relationship between haploid chromosome number and haploid genome size, in addition to the relationship of WGD to these metrics in and across the three lineages of bryophytes. Understanding what controls plant genome size is a crucial and pressing line of inquiry in plant evolution (Armstrong *et al.*, 2023), for which we laid the framework in bryophytes by analysing previously published

chromosome counts and genome size estimates in a phylogenetic framework.

MATERIALS AND METHODS

Phylogenetic data collection (Pyphlawd)

A DNA matrix of 3890 taxa and ten concatenated loci (5.8S, 26S, *atpb*, *nad5*, *psbA-trnH*, *psbA*, *rbcL*, *rps4-trnS*, *trnA* and *trnL-F*) was assembled using PyPhlawd (Smith and Walker, 2019). Concatenation was chosen to maximize resolution given the smaller number of loci sampled relative to phylogenomic studies. PyPhlawd was run using 'Embryophyte' as the taxonomic group, then narrowed to all bryophyte species (hornworts, liverworts and mosses) using a taxon list including all listed species available on GenBank (https://www.ncbi.nlm.nih.gov/). The clustering function in PyPhlawd produced 311 clusters, of which the ten with the largest taxonomic representation were selected for further analysis.

Phylogenetic analysis

Matrices for each locus were aligned using MAFFT v.7.419 with default settings (Katoh and Standley, 2013). The loci were concatenated, and each locus was assigned a separate partition. Phylogenetic inference was implemented in IQ-TREE v.2 (Minh *et al.*, 2020) using ModelFinder Plus (Kalyaanamoorthy *et al.*, 2017) to find best-fitting models for each partition. IQ-TREE was implemented using 1000 bootstrap replicates. The tree was rooted with the hornworts clade. The concatenated matrix, associated GenBank accession numbers and phylogenetic tree are available in the Dryad digital repository (https://doi.org/10.5061/dryad.31zcrjdwm).

Genome size and chromosome data collection

Genome size data. Bryophyte haploid genome size estimates were compiled from multiple sources. Original publications included the following: Bainard and Villarreal (2013) ($N = 29$); Bainard *et al.* (2013) ($N = 82$); Bainard *et al.* (2020) ($N = 60$); Greilhuber *et al.* (2003) ($N = 11$); Li *et al.* (2023) ($N = 209$); Melosik *et al.* (2005) ($N = 5$); Orzechowska *et al.* (2010) ($N = 3$); Orzechowska *et al.* (2018) ($N = 4$); Pustahija *et al.* (2013) ($N = 1$); Ricca *et al.* (2008) ($N = 58$); Schween *et al.* (2003) ($N = 1$); Temsch *et al.* (1998) ($N = 66$); Temsch *et al.* (2010) ($N = 157$); and Voglmayr (2000) ($N = 330$). In addition, 258 estimates were retrieved from the Darwin Tree of Life (https://www.darwintreeoflife.org/), which is continually generating new genome size estimates as part of an ambitious project to sequence the genomes of all complex life in Britain and Ireland. All estimates were generated by stain-based techniques, i.e. Feulgen image analysis or flow cytometry. The compiled list of taxa and estimates can be found in Supplementary Data, Table S1. These data consisted of 1274 genome size observations, representing 607 species. Any values that were not identified to the species level were removed (i.e. 'Genus sp'.).

The vast majority of the flow cytometry-based genome size estimates we compiled used an internal standard, which is the best practice for plant DNA content analysis (Sliwińska *et al.*, 2022). However, a recent paper (Li *et al.*, 2023) estimated genome size with flow cytometry using external standards for 209 samples, representing 160 bryophyte species. We were able to compare the data generated using these two standardization methods for 32 species. For the majority of these species, the externally standardized data were ±25 % in comparison to the internally standardized data, which is within the range of variation for a given species observed between different studies using internal standards in the present study. Additionally, for 128 of these species, this represents the only data available in the literature, hence these externally standardized flow cytometry-based genome size estimates were retained in the analyses.

Chromosome count data. Chromosome counts were compiled from literature, with the majority from the book *Index to Bryophyte Chromosome Counts* (Fritsch, 1991). Chromosome counts generated subsequent to the publication the book, from 1991 to 2024, were assembled from literature searches. In addition to species and unique chromosome counts, where available, the authors, number of populations, number of studies and type (mitotic or meiotic) were also recorded. The assembled chromosome counts for each species are available in Supplementary Data, Table S2.

Compiled chromosome counts were reviewed and cleaned. Chromosome counts presented as a range or as ambiguous [e.g. $1n = 7–9$, ~7, 7(?)] were removed.

Taxonomic updates and ordinal classifications

Orders were assigned using online or published classification systems for mosses (Goffinet and Buck, 2020; Bechteler *et al.*, 2023), liverworts (Söderström *et al.*, 2016) and hornworts (Villarreal and Goffinet, 2023). Families and scientific names were assigned using the Taxonomic Name Resolution Service v.5.1 (Boyle *et al.*, 2013; https://tnrs.biendata.org/; 5 December 2023, date last accessed) and were updated from the original publications for the genome size, chromosome and phylogenetic datasets. In the genome size data, when there were multiple observations for a given species, all these data were included in the calculation of the mean value for the species. For the chromosome data, when there were multiple karyotypes for a species from different publications, the data were merged into a single list for each species. In the phylogenetic tree, when synonymizing of a species resulted in duplicate tip names, redundant tips were pruned randomly. Modifications to names on the phylogenetic tree were made in R (R Core Team, 2021) using the Ape package (Paradis and Schliep, 2019).

Genome size and chromosome number metrics

Forty per cent (242 of 607 species) of species for which a haploid genome size estimate is published were represented by multiple, independent estimates. Among species with published genome size estimates represented in the molecular phylogeny (440 species), 42 % (188 of 440) were represented by multiple, independent estimates. For subsequent analysis, including ancestral character state reconstruction, phylogenetic signal and

correlation analysis, the mean value for each species was calculated, and this value is referred to hereafter as the haploid genome size.

In total, 59 % (1276 of 2138) of species with published karyotypic data were represented by karyotypes from two or more populations. Among species with published karyotypic data that are represented in the molecular phylogeny (1088 species), 70 % (759 of 1088 species) were represented by karyotypes from two or more populations. Four metrics were derived from the karyotype data: haploid chromosome number, range, polyploidy present/absent (p/a) and ploidy rank. The metric haploid chromosome number is the lowest chromosome number recorded for a species. Range is calculated as the lowest observed chromosome count subtracted from the highest for each species; for species for which only one chromosome count is available, the range is zero. The metric polyploidy p/a is calculated by dividing each chromosome number observed for a given species by the lowest chromosome number observed for the same species (i.e. $1n = 7$, 14, 21 yields the values 2 and 3). In species for which this calculation yielded a whole number greater than or equal to two, it was scored as polyploidy present. This quantification of polyploidy most probably identified neopolyploids rather paleopolypoids, though the term 'polyploidy' is used hereafter. Alternatively, in species for which only fractional numbers resulted from this calculation, it was scored as polyploidy absent. In this way, karyotypic evidence for genome doubling in the strictest sense was discerned. For ploidy rank, the metric range was divided by the metric haploid chromosome number, and therefore species for which only one haploid chromosome number was observed, the ploidy rank is zero. Values of at least one indicated the presence of polyploidy in the strictest possible sense, consistent with the metric polyploidy p/a. Values of ploidy rank scored as equal to or greater than one were recorded as one. Values from zero to one are therefore a proxy for how close to WGD the observed chromosome counts were, allowing for some inferences about aneuploidy or miscounted chromosomes. The metrics range, polyploidy p/a and ploidy rank were calculated from the chromosome dataset trimmed to species for which karyotypes were performed for two or more populations, allowing for the possibility of observing a range of values or WGD within a species. The mean value for each of these metrics for each sampled taxonomic order was calculated, as was the mean value for the hornworts, liverworts and mosses.

Phylogenetic signal and correlation analysis

For the metrics haploid genome size, haploid chromosome number, range, polyploidy p/a and ploidy rank, both phylogenetic signal and correlation among each metric were calculated. The metrics haploid genome size, haploid chromosome number, range and ploidy rank were treated as continuous characters, and phylogenetic signal was calculated as Pagel's lambda (Pagel, 1999) using 'phylosig' as part of the package Phytools v.2 (Revell, 2024) implemented in R (R Core Team, 2021). The metric polyploidy p/a was treated as a discrete character, and phylogenetic signal was calculated as a *D*-statistic (Felsenstein's *D*) for phylogenetic signal for discrete binary characters using the function 'phylo.d' as part of the package Caper (R Development Core Team, 2011) implemented in R. Correlation among each metric was also assessed using the package Caper with the function 'Corr.ic'. Each calculation of phylogenetic signal and correlation analysis was implemented with a phylogenetic tree pruned to exclude tips with missing data. Each calculation of phylogenetic signal for a metric in addition to correlation among metrics was performed on four datasets: the total dataset, mosses only, liverworts only and hornworts only, with the exception of correlation for metrics within hornworts, because the sampling when pruned was prohibitively small, which here was considered fewer than ten taxa.

Ancestral character state reconstruction

For the continuous characters haploid genome size and ploidy rank, a continuous ancestral character state reconstruction was estimated according to Brownian motion using the function 'contmap' as part of the package Phytools v.2 (Revell, 2024) implemented in R (R Core Team, 2021). Note that for the metric ploidy rank, ancestral character state reconstruction was performed using only species represented by chromosome counts from two or more populations. The resulting estimates for haploid genome size and ploidy rank were plotted onto phylogenies pruned to exclude tips with missing data for genome size estimates and chromosome number, respectively, using the colour palette *viridis* as part of the 'viridisLite' package and the function 'plot' as part of the package 'ggtree' (Yu, 2020).

For the discrete metric polyploidy p/a, an ancestral character state reconstruction was estimated using the function 'ace' as part of the package Ape (Paradis and Schliep, 2019) implemented in R (R Core Team, 2021). The model *equal rates* ('ER') was implemented, because evidence for transition asymmetry in this character was lacking. This model assumes an equal rate of transitions among discrete character states and offers a reasonable baseline model of evolution. Tip states and ancestral character likelihoods were plotted on a phylogeny pruned to exclude tips with missing data using the package 'ggtree' (Yu, 2020).

Ancestral base chromosome number estimation

The metric haploid chromosome number was used as input for reconstructing the ancestral haploid chromosome number across the tree using ChromEvol v.2.2 (Glick and Mayrose, 2014), which uses a maximum likelihood (ML) approach for modelling chromosome evolution. ChromEvol offers ten models that combine the following parameters: loss, gain, duplication, demi-duplication, linear loss, linear gain, base chromosome number, base chromosome number transition rate and the root base chromosome number optimization. The chromosome number of the root was not constrained and was therefore a free parameter, and the ancestral chromosome number was estimated. The remaining models combining the free parameters were fitted to the data, and the best-fitting model [the lowest Akaike information criterion (AIC) value; Burnham and Anderson, 2004] was used to estimate the likelihood of ancestral haploid chromosome numbers across the tree pruned to exclude tips with missing

data. In ChromEvol, the minimum chromosome number was set to one, and the maximum chromosome number was set to one greater than the maximum haploid chromosome number in our dataset.

RESULTS

Phylogenetic analysis summary

The concatenated matrix was 25 050 bp long and included 3890 taxa. After taxonomic updates, 3619 taxa remained in the phylogenetic tree. The resulting phylogenetic tree was inspected for singleton species disrupting the monophyly of taxonomic orders, which were presumed erroneous on the basis of misidentification or poor sequence quality, resulting in 3561 taxa. In total, 55 orders were represented in the phylogeny, including all five hornwort orders, 12 of 23 liverwort orders, and 38 of 46 moss orders (Brinda and Atwood, 2024). The majority of represented orders were monophyletic, with notable exceptions among the mosses, including three orders: the Hypnales, which formed a grade and within which the monophyletic orders Hypopterigiales and Hookeriales were nested, and the polyphyletic Dicranales and Grimmiales. The topology presented here was largely consistent with recent studies of broad-scale phylogenetic relationships within the bryophytes (e.g. Bainard *et al.*, 2020; Bechteler *et al.*, 2023) and offers a robust phylogenetic hypothesis upon which to map genome size and chromosomal data.

Haploid genome size evolution

Total dataset. The complete genome size dataset included 1274 accessions representing 607 species. The highest mean haploid genome size (1C) was in the liverworts (1C = 1.35 pg, $N = 338$), followed by the mosses (1C = 0.56 pg, $N = 905$) and the hornworts (1C = 0.25 pg, $N = 31$). Among the liverworts and mosses, the orders with the highest mean haploid genome sizes represented by at least nine species were the Metzgeriales (1C = 2.34 pg, $N = 9$) and the Bryales (1C = 0.96 pg, $N = 37$), respectively. The orders with the lowest mean haploid genome size among the liverworts and mosses were Orthotrichales (1C = 0.44 pg, $N = 16$) and the Marchantiales (1C = 0.65 pg, $N = 18$), respectively. Hornwort sampling was limited, and haploid genome size ranged from Dendrocerotales (1C = 0.36 pg, $N = 7$) to the Leiosporocerotales (1C = 0.18, $N = 2$) (Supplementary Data, Table S3).

Phylogenetic dataset. The phylogenetic haploid genome size (1C) dataset (trimmed to exclude species not represented in the phylogeny) consisted of 1008 accessions representing 440 species. Across the genome size data represented in the phylogenetic tree, mean haploid genome size values ranged from 0.16 to 20.25 pg, although 89 % of genome size estimates are <1 pg, and only ten species had a genome size greater than 1C = 4 pg, all of which are liverworts.

The highest average haploid genome size (1C) was in the liverworts (1C = 1.52 pg), followed by the mosses (1C = 0.53 pg) and the hornworts (1C = 0.25 pg) (Table 1). Among the mosses, for orders represented by at least ten species, the order Bryales had the highest genome size (1C = 0.92 pg) and the Orthotrichales had the lowest (1C = 0.40 pg). Among the liverworts, the order Metzgeriales had the highest mean genome size (1C = 3.01 pg). The order with the highest genome sizes in the hornworts was the Dendrocerotales (1C = 0.36 pg) (Supplementary Data, Table S4).

Haploid genome size ancestral character state reconstruction

The ancestral genome size was highest for liverworts, lowest for the hornworts, and intermediate for the mosses. Here, evolutionary transitions are defined as changes in the value of the haploid genome size to above or below the mean for each lineage. Within the hornworts, the genome sizes were uniformly small, with one transition to a genome size higher than the mean for hornworts (1C = 0.23 pg) in the clade including genera *Nothoceros*, *Phymatoceros*, *Megaceros* and *Dendroceros* (Fig. 1). In contrast, in the liverworts there were at least five transitions to haploid genome sizes above the mean (1C = 1.52 pg) and at least four subsequent reversals to smaller genomes. Transitions to larger haploid genome sizes were found throughout the class Jungermanniopsida (Fig. 1). In the mosses, there were ≥15 transitions to haploid genome sizes above the mean (1C = 0.53 pg) (Fig. 1). The majority of transitions to haploid genome sizes above the mean occurred in the orders Dicranales, Bryales and Hypnales.

Chromosome number evolution

Data. The initial dataset resulting from chromosome counts compiled within the book by Fritsch (1991), in combination with subsequent published counts (19 publications), consisted of 14 758 individual chromosome counts representing 2270 species. After taxonomic corrections, including synonymizing,

Table 1. *Mean values of haploid genome size (1C), haploid chromosome number (1n), ploidy rank, range and polyploidy p/a in the hornworts, liverworts and mosses for all species represented in the molecular phylogenetic dataset. Standard deviations are given for each value. The metric ploidy rank is calculated for each species with karyotypic data by dividing the range of chromosome counts by the lowest chromosome count, here considered the 'haploid' chromosome number. The metric polyploidy p/a is scored as present or absent on the basis of whole number multiples of haploid chromosome number for a given species.*

	Haploid genome size (1C)	Haploid chromosome number (1*n*)	Ploidy rank	Chromosome range	Polyploidy p/a
Mosses	0.53 ± 0.26	10.7 ± 4.78	0.33 ± 0.41	4.12 ± 6.97	0.23 ± 0.42
Liverworts	1.52 ± 2.59	9.3 ± 2.93	0.21 ± 0.37	2.34 ± 5.20	0.16 ± 0.37
Hornworts	0.25 ± 0.12	4.83 ± 0.75	0.4 ± 0.75	1.67 ± 1.89	0.16 ± 0.41

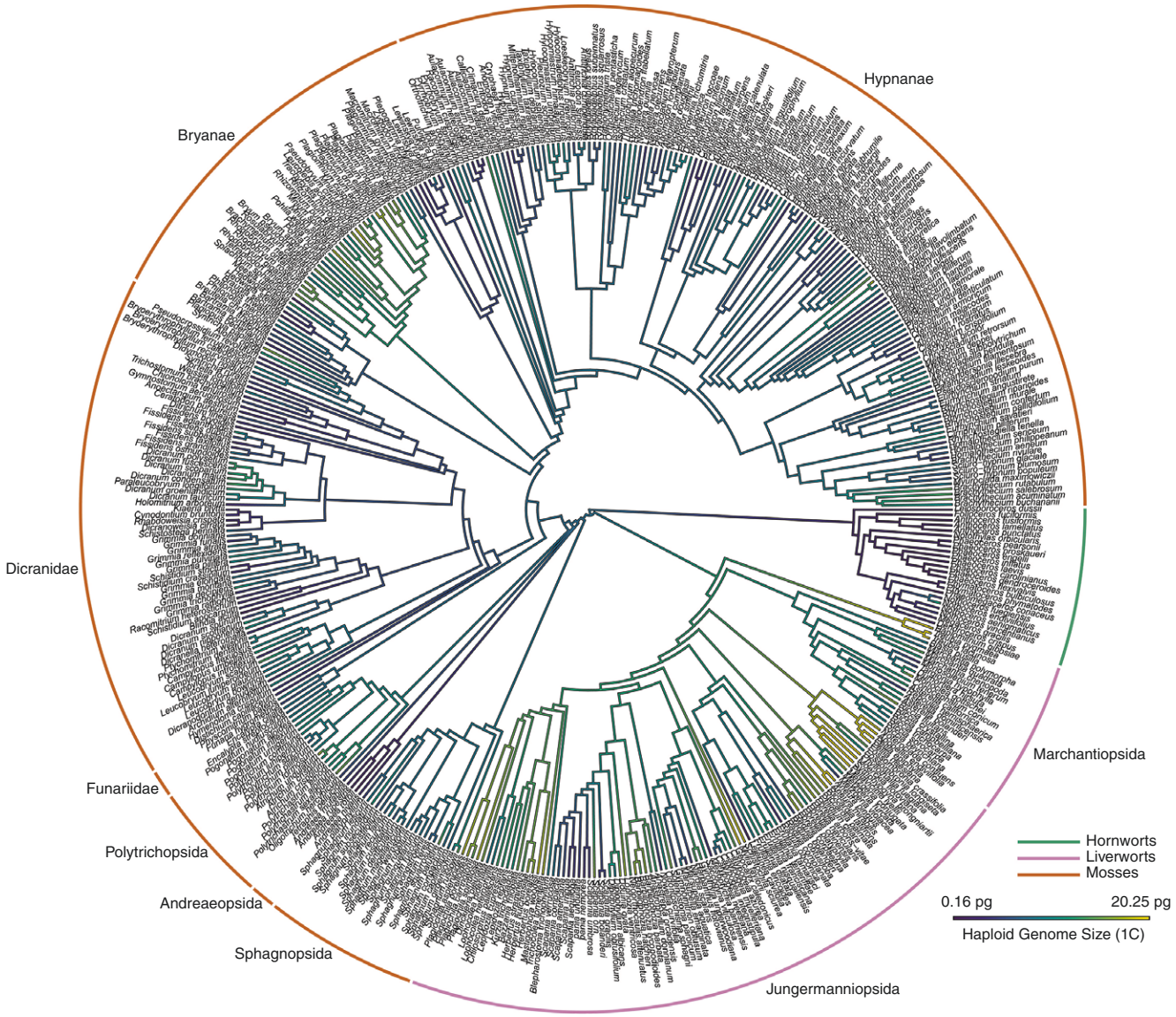

FIG. 1. Continuous ancestral character state mapping of haploid genome size (1C) values reconstructed using the 'contmap' function with default settings from the Phytools package in R on a maximum likelihood phylogenetic tree, pruned to exclude species with missing data, resulting in a tree including 440 species.

the corrected dataset consisted of 14 702 individual chromosome counts representing 2139 species. Among these, 1089 species were represented in the phylogenetic dataset. After the removal of species represented by a chromosome count from only one population, 759 remained, and this final dataset was used in subsequent ancestral character state reconstruction, phylogenetic signal and phylogenetically independent correlation analyses (Supplementary Data, Table S5).

Chromosome metrics

Total dataset. In the total dataset including karyotypic data both represented and excluded from phylogenetic reconstruction, mosses had the highest average haploid chromosome number ($1n$ = 12.12), the highest average range of chromosome counts within species (2.57) and the highest average ploidy rank (0.19). Liverworts were intermediate between mosses and hornworts for the metrics haploid chromosome number ($1n$ = 9.60) and average range of chromosome counts (0.96). In contrast, liverworts had values slightly lower than hornworts in metrics pertaining to polyploidy, ploidy rank (0.11) and number of polyploids per species (polyploidy p/a = 0.12). Hornworts had an average haploid chromosome number of $1n$ = 5.27, an average range of 0.55, ploidy rank of 0.13, and average number of polyploids per species of 0.22 (Supplementary Data, Table S3). However, it should be noted that the chromosome counts were sampled for only nine hornwort species, of which two, *Phaeoceros laevis* and *Anthoceros punctatus*, exhibited polyploidy (polyploidy p/a).

Phylogenetic dataset. In the dataset including only chromosome count data represented in the phylogeny, mosses had the highest average haploid chromosome number ($1n$ = 10.7), the

highest average range of chromosome counts within species (4.12), the highest average ploidy rank (0.47) and the highest average occurrence of polyploidy (polyploidy p/a = 0.23) (Table 1). In total, at least one instance of polyploidy (as quantified using the metric polyploidy p/a) was found in 21 % of all moss species sampled in the phylogenetic dataset. Liverworts were intermediate between mosses and hornworts in haploid chromosome number ($1n = 9.3$) and average range of chromosome counts (2.34), and for the metrics ploidy rank (0.21) and number of polyploids per species (0.16), they had values exactly the same or slightly lower than hornworts. Hornworts had an average haploid chromosome number of $1n = 4.83$, an average range of chromosome counts of 1.67, ploidy rank of 0.4, and number of polyploids per species of 0.16. As in the total dataset, however, it should be noted that species sampling in hornworts was limited.

The metric ploidy rank was derived from the metric range and haploid chromosome number. Within the mosses, the taxonomic order with the highest average ploidy rank value across all sampled bryophytes was the Funariales (0.58). Among orders represented by ten or more species, the order with the lowest ploidy rank was the Sphagnales (0.21). Among the liverwort orders represented by at least ten species, the highest average ploidy rank was in the Marchantiales (0.36). The hornworts were represented by three orders in the chromosomal dataset, the Dendrocerotales, Anthocerotales and Notothyladales, which had average ploidy rank values of 0, 0.6 and 0.4, respectively (Supplementary Data, Table S4).

Ploidy rank ancestral character state reconstruction

Reconstructed ploidy rank values ranged from zero to one. Here, major evolutionary transitions were considered changes in ploidy rank values to above or below 0.5. In the hornworts, there were two transitions to a ploidy rank value above 0.5. The liverworts included 14 transitions to a ploidy rank value above 0.5. In the mosses, there were 73 transitions to ploidy rank values above 0.5. Within the liverworts, a larger proportion of species exhibited a ploidy rank higher than 0.5 in the Marchantiopsida than the Jungermanniopsida, largely as a result of several species with high ploidy rank in the genus *Riccia*. In the Jungermanniopsida, the highest density of species with a ploidy rank of ~0.5 was observed in the genus *Riccardia*. Among the mosses, there was a higher density of species with ploidy rank higher than 0.5 in the superorder Hypnanae than in the Bryanae. In the Hypnanae, the highest density of species with high ploidy rank was in the genera *Sciuro–Hypnum* and *Brachythecium* (Fig. 2).

Ancestral base chromosome number estimation

Model optimization in ChromEvol, for haploid chromosome number associated with each species in the chromosome dataset, indicated the best fit for models 'CONST_RATE' (AIC = 239.6), 'CONST_RATE_DEMI' (AIC = 239.7) and 'CONST_RATE_NO_DUPL' (AIC = 239.8) (Supplementary Data, Table S6). Although differences in fit among these models were negligible, data were optimized according to the model with the lowest AIC, 'CONST_RATE', which includes a constant rate of chromosomal gains and losses and allows for 'half-duplications', which can result from WGD followed by chromosomal loss. Across the full phylogeny, ChromEvol reconstructed 94.78 duplications, 187.5 gains, 757.869 losses and 35.43 demi-duplications. For reconstructed chromosome numbers at all nodes, see Supplementary Data, Fig. S1. The estimated ancestral haploid chromosome number was $1n = 12$ for all bryophytes (likelihood = 0.56), $1n = 6$ for hornworts (likelihood = 0.67), $1n = 10$ for liverworts (likelihood = 0.73) and $1n = 12$ for mosses (likelihood = 0.55) (Supplementary Data, Fig. S1). Within the mosses, the most likely ancestral chromosome number for the superorders Hypnanae and Bryidae was $1n = 12$, and for Dicranidae $1n = 13$. The order Hypnales was non-monophyletic and included a large grade, in addition to a clade including genera *Plagiothecium*, *Herzogiella*, *Acrocladium* and *Pseudotaxiphyllum*, for which the most likely ancestral chromosome number was $1n = 11$. The Bryanae included a transition to $1n = 9$ for the clade consisting of orders Hedwigiales, Splachnales and Bartramiales. Within the Dicranidae, there were two transitions to $1n = 13$ and one transition to $1n = 12$. Among the mosses, the lowest ancestral haploid chromosome number was $1n = 7$ associated with the clade consisting of the orders Disceliales, Encalyptales and Funariales. Within the liverworts, the most likely ancestral chromosome number for superorders Jungermanniopsida and Marchantiopsida was $1n = 10$. Within the Jungermanniopsida, there is one transition to $1n = 9$.

Phylogenetic signal and correlation

Phylogenetic signal. In the total dataset including hornworts, liverworts and mosses, the strongest phylogenetic signal among continuous characters (haploid genome size, haploid chromosome number, range and ploidy rank) was associated with haploid genome size ($\lambda = 0.93$) and the lowest phylogenetic signal was associated with ploidy rank ($\lambda = 0.19$). Polyploidy p/a was treated as a discrete character, and *D*-statistic values close to zero indicated strong phylogenetic signal. The *D*-statistic for this metric was 0.86, suggesting low phylogenetic signal (Supplementary Data, Fig. S2). Among the mosses, the strongest phylogenetic signal among continuous characters was associated with haploid chromosome number ($\lambda = 0.72$), and polyploidy p/a had a *D*-statistic of 0.89. Among liverworts, the strongest phylogenetic signal among continuous characters was associated with haploid genome size ($\lambda = 0.96$), and polyploidy p/a had a *D*-statistic of (0.72). Among the hornworts, the strongest phylogenetic signal among continuous characters was associated with haploid genome size ($\lambda = 0.40$), and polyploidy p/a had a *D*-statistic of 15.5 (Table 2).

Correlation. There were no statistically significant phylogenetically independent correlations between haploid genome size and any chromosome metric (Supplementary Data, Table S5).

DISCUSSION

Haploid genome size evolution

The haploid nuclear genome size (1C) in bryophytes is known to be small relative to other land plants (Temsch *et al.*, 1998;

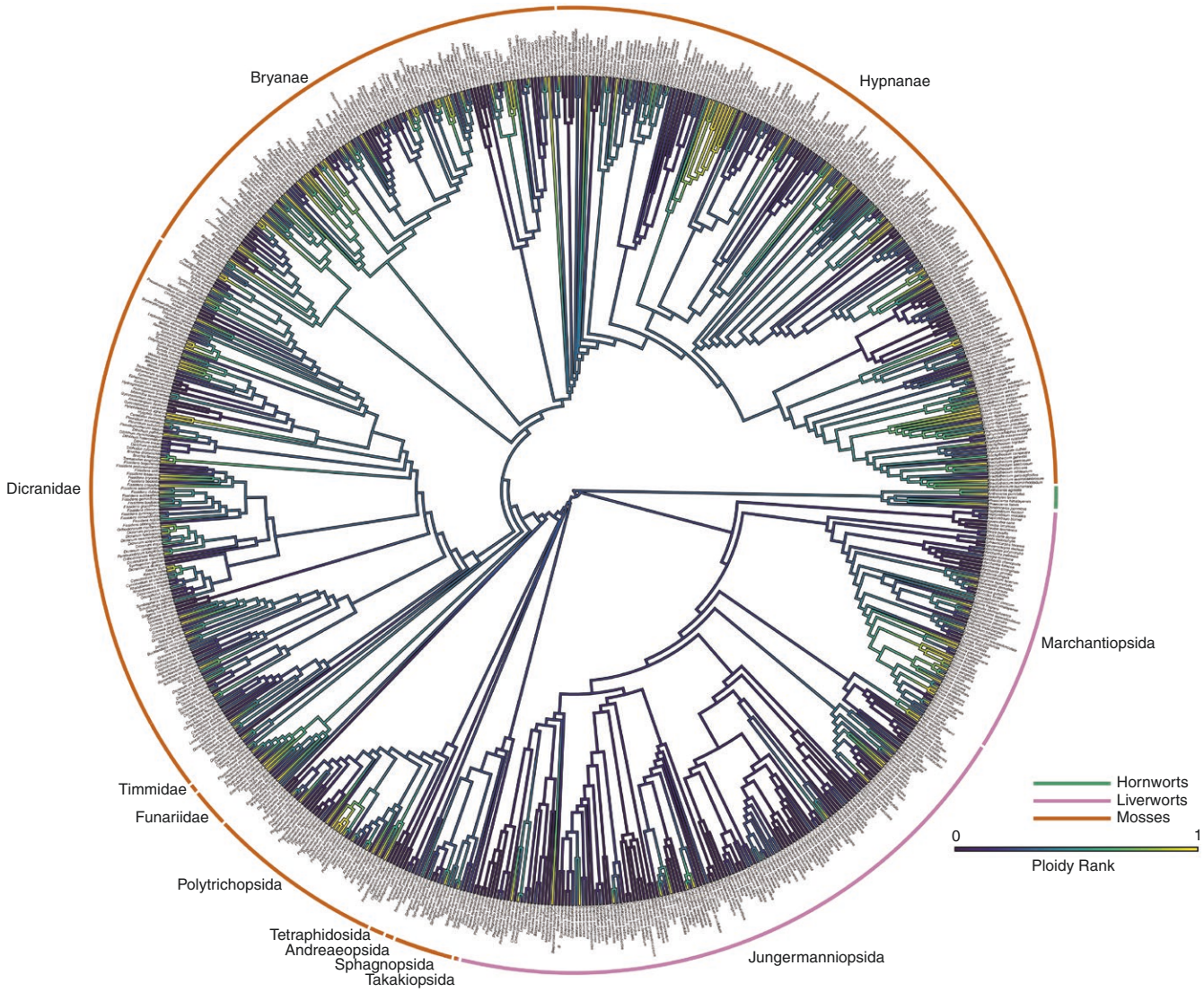

FIG. 2. Continuous ancestral character state mapping reconstruction of ploidy rank reconstructed using the 'contmap' function with default settings from the Phytools package in R on a maximum likelihood phylogenetic tree (Supplementary Data, Fig. S1) that is pruned to exclude species with missing data, resulting in a tree including 759 species. The metric ploidy rank is calculated for each species with karyotypic data by dividing the range of chromosome counts by the lowest chromosome count, here considered the 'haploid' chromosome number. For a given species, all values greater than one indicate the presence of at least one whole number multiple of the haploid chromosome number, indicating whole genome duplication (WGD). Ploidy rank values greater than one are all collapsed and mapped onto the tree as equal to one.

Voglmayr, 2000; Bainard and Villarreal, 2013; Bainard *et al.*, 2013). Here we found a mean haploid genome size of 0.72 pg for bryophytes in the total dataset (Supplementary Data, Table S3), whereas the mean haploid genome sizes in angiosperms, gymnosperms and monilophytes are 5.13, 18.35 and 12.11 pg, respectively (Leitch *et al.*, 2019). Distinct genome sizes across land plant lineages are a product of unique evolutionary histories and selection pressures (Bennett and Leitch, 2005). Although genome size is uniformly small in the bryophytes, haploid genome size varies among the hornworts, liverworts and mosses, each of which diverged ≥300 Mya (Villarreal and Renner, 2012). Haploid genome size indeed exhibited a strong phylogenetic signal (Table 2; $\lambda = 0.93$), suggesting lineage-specific factors shaping genome size evolution across the bryophytes.

The hornworts had the smallest mean haploid genome size (1C = 0.25 pg) and little variation in haploid genome size across the sampled diversity, ranging from 1C = 0.18 to 0.75 pg (Table 1; Supplementary Data, Table S3). Bainard and Villarreal (2013) observed the highest genome sizes among more recently evolved species, consistent with our finding that the most recently diverged genus, *Nothoceros*, had the largest genome size among hornworts (Fig. 1). This might suggest a transition to a larger genome in the ancestor of *Nothoceros*. However, overall, genome size in hornworts exhibited little phylogenetic pattern, and the phylogenetic signal for genome size was moderate (Table 2; $\lambda = 0.4$). Among the bryophytes, hornworts most often establish symbioses with cyanobacteria relative to other bryophytes (Villarreal and Renzaglia, 2006; Li *et al.*, 2020), are prone to horizontal gene transfer (Villarreal

TABLE 2. *Phylogenetic signal, given as Pagel's lambda (λ), for the chromosomal metrics treated as continuous traits, haploid genome size, haploid chromosome number, ploidy rank and chromosome range. The metric ploidy rank is calculated for each species with karyotypic data by dividing the range of chromosome counts by the lowest chromosome count, here considered the 'haploid' chromosome number. The* D*-statistic (Felsenstein's* D*, phylogenetic signal for binary discrete characters) is given for the discrete character polyploidy p/a (polyploidy scored as present or absent on the basis of whole number multiples of haploid chromosome number for a given species), which indicates the presence or absence of at least one polyploid cytotype for a given species.*

Parameter	All bryophytes	Mosses	Liverworts	Hornworts
Haploid genome size (1C)	0.93460	0.34	0.96454	0.401377
Haploid chromosome number (1*n*)	0.67182	0.72059	0.016545	5.71×10^{-5}
Ploidy rank	0.13539	0.18533	0.09632	5.71×10^{-5}
Chromosome range	0.25191	0.25901	0.09687	5.71×10^{-5}
Polyploidy p/a	0.86059	0.89796	0.72745	15.52363

and Renzaglia, 2015) and exhibit little genetic redundancy. A recent phylogenomic study found evidence of an ancient large-scale duplication in the ancestor of the Anthocerotopsida; however, genomic studies of some hornwort species found a lack of either ancient or recent signatures of WGD (Li *et al.*, 2020; Zhang *et al.*, 2020). Their small genomes might therefore reflect an evolutionary strategy consisting of a highly streamlined genome with genetic novelty frequently resulting from horizontal gene transfer with bacteria and fungi (Zhang *et al.*, 2020).

Liverworts had the largest average bryophyte genomes, with a mean haploid genome size of 1.35 pg in the total dataset and a range of 0.28–20.25 pg (Table 1; Supplementary Data, Table S3). Similar to the hornworts, genomic studies of liverworts thus far have not found genomic signatures of WGD and little genetic redundancy (Diop *et al.*, 2020; Dong *et al.*, 2022; Linde *et al.*, 2023), although Shen *et al.* (2024) found evidence of an ancient large-scale duplication in the ancestor of the Jungermanniopsida. Increases in genome size in the liverworts appear to be driven by the accumulation of transposable elements (TEs), as noted by Linde *et al.* (2023), who compared TE content of *Marchantia polymorpha* (1C = 0.228 pg) with that of *Lunularia cruciata*, a species with approximately double the haploid genome size (1C = 0.593 pg). Linde *et al.* (2023) point out that the bursts of TEs that lead to increases in genome size appear rare among the liverworts. Notably, however, we observed substantial variability in haploid genome size across the liverworts, including five major transitions, some of which occurred in several species with no evidence of WGD (e.g. *Treubia lacunosa*; Fig. 1). These transitions might be driven by shifts in TEs. In addition, the average haploid genome size of the two classes Marchantiopsida and Jungermanniopsida differed substantially (1C = 2.55 and 1.23 pg in the phylogenetic dataset, respectively). Genomic studies of species spanning major genome size transitions, such as these, are crucial for understanding the genome dynamics of lineages such as the liverworts, which have a relatively austere morphology and low rates of morphological evolution, but potentially dynamic systems of genetic control.

The mosses are by far the most species-rich lineage of bryophytes, with >13 000 described species (Patel *et al.*, 2021) and featured a mean genome size (1C) of 0.56 pg in the total dataset and a range of 0.16–2.21 pg (Table 1; Supplementary Data, Table S3). Consistent with Bainard *et al.* (2020), we found the majority of transitions to larger genome sizes in relatively recently diverged clades, such as the Dicranales *sensu stricto*, Bryales and Hypnales (Fig. 1). In some instances, genome size can support the definition of taxonomic groups. For instance, Bechteler *et al.* (2023) supported the recognition of several orders in the polyphyletic order Dicranales that were formerly recognized at the family level. The mean haploid genome sizes for two of these orders, Ditrichales and Distichiales, were 0.38 and 0.36 pg, respectively, which are smaller than the average genome size of Dicranales *sensu stricto* (0.5 pg; Supplementary Data, Table S4).

Multiple factors might influence the dynamic genome size evolution in the mosses relative to the hornworts and liverworts. Unlike the liverworts and hornworts, there is evidence for genomic signatures of WGDs in several moss species, including *Ceratodon purpureus*, *Funaria hygrometrica*, *Physcomitrium patens* and *Syntrichia caninervis* (Carey *et al.*, 2021; Silva *et al.*, 2021; Kirbis *et al.*, 2022), in addition to phylogenomic evidence for several WGD events across the evolutionary tree of mosses (Gao *et al.*, 2022; Patel *et al.*, 2023). In addition, genomic analysis of the moss model system *P. patens*, in addition to the closely related *F. hygrometrica*, reveals that transposable elements make up 50–60 % of their genomes and that differences in the genome size between these two species are attributable, in part, to the accumulation of TEs (Rensing *et al.*, 2008; Lang *et al.*, 2018; Kirbis *et al.*, 2022). Mosses have mechanisms for rapid genome downsizing through deletion and through chromosomal fission and fusion, in particular following WGD (Lang *et al.*, 2018; Liu *et al.*, 2022*a*). For instance, *P. patens* lacks helitron transposons, which are otherwise abundant in land plants, suggesting that they have been rapidly purged (Rensing *et al*, 2008; Leitch and Leitch, 2012). In addition, the relatively low proportion of protein-coding genes in tandem array in *P. patens* (1 %) compared with the angiosperms (Rensing *et al.*, 2008) suggests a low frequency of local gene duplication (Tuskan *et al.*, 2006).

Chromosome evolution

Haploid chromosome number. The ancestral chromosome numbers inferred in this study were consistent with previously published values. For the hornworts, we reconstructed an ancestral chromosome number of 1*n* = 6, and for the liverworts, 1*n* = 10. This aligns with the 'basic chromosome numbers' estimated to be 1*n* = 5 or 6 for hornworts and 1*n* = 9 or 10 for the liverworts (Steere, 1954, 1972; Crum, 2001). We inferred that the ancestral chromosome number for mosses is 1*n* = 12, a value that has historically been more difficult to estimate given the breadth of inter- and intraspecific variation in chromosome counts across the species (Figs 3 and 4). This is reflected in the wide breadth and even distribution of chromosome numbers

among nodes and extant species for the mosses (Fig. 4) and in the ambiguity of the ancestral chromosome number estimate of $1n = 12$ (likelihood = 0.55).

Modelling the evolution of haploid chromosome numbers across the bryophytes with ChromEvol revealed constant and frequent gains and losses of chromosomes, regardless of the involvement of WGD (Supplementary Data, Fig. S1). Two models fitted the data comparably well (Supplementary Data, Table S6), suggesting that both single gains and losses of chromosomes ('CONST_RATE'), in addition to WGD followed by some degree of genetic loss towards a more haploid-like condition ('CONST_RATE_DEMI'), are equally likely explanations for the patterns of haploid chromosome evolution observed across bryophytes. For instance, the only haploid chromosome numbers observed in the Funariales in this dataset were $1n = 13$, 19, 26 and 27, but the most likely ancestral haploid chromosome number was seven. This indicates the potential for at least one WGD event in the Funariales followed by haploidization. In contrast, in the liverworts and hornworts, WGD was decidedly less frequent (Table 1), and the haploid chromosome number remained relatively consistent (Fig. 3).

These findings provided phylogenetic and evolutionary context for the distinct factors influencing genome size evolution in each lineage of bryophytes. Among mosses, highly dynamic haploid chromosome numbers suggest that WGD is a crucial process for genome evolution in this lineage. In the hornworts and liverworts, highly uniform and evolutionarily stable haploid chromosome numbers suggest that genome size evolution might be driven more often by other genomic elements, such as TEs (Fig. 3; Supplementary Data, Fig. S1).

Whole genome duplication

Among the bryophytes, both allopolyploidy and autopolyploidy are most prevalent in the mosses (Natcheva and Cronberg, 2004; Patel *et al.*, 2021, 2023). Consistent with these observations, we found that mosses have the highest average number of whole number multiple polyploids per species (0.23; the metric 'polyploid p/a'), in addition to the highest average ploidy rank (0.33), although neither exhibited strong phylogenetic signal, and there were numerous transitions from lower to higher ploidy rank in addition to reversals across bryophytes (Tables 1 and 2). This suggests that polyploidy evolved repeatedly and independently across mosses (Fig. 2; Wyatt *et al.*, 1988; Crawford *et al.*, 2009; Patel *et al.*, 2021). Metrics representing polyploidy in the present study (ploidy rank and polyloidy p/a) were based on potential genome doubling within a morphological species, which therefore represents autopolyploidy. Although the possibility of genome doubling following hybridization (allopolyploidy) was not examined directly in this study, the families and orders in which allopolyploidy was often observed (e.g. Sphagnales and Funariales) also exhibited a high frequency of WGD, as quantified by ploidy rank and polyploidy p/a. This suggests that some reported cytotypic variation might represent allopolyploidy and, more generally, that cytotypic variation within morphological species could represent auto- or allopolyploids. For instance, consistent with numerous studies of the Funariales finding frequent polyploidy (Beike *et al.*, 2014; Medina *et al.*, 2018, 2019; Patel *et al.*, 2023), we found that this moss order has the highest ploidy rank (0.58). Funariaceae have been highlighted repeatedly as a natural study system for polyploidy, given their propensity for both allo- and autopolyploidy (Leitch and Leitch, 2012). Likewise, the genus *Sphagnum* is well studied in the context of polyploidy and hybridization (e.g. Ricca and Shaw, 2010), and we found multiple transitions to a ploidy rank of >0.5 in the genus, although its mean ploidy rank is low. In both the Sphagnales and Funariales, we also found some species with no published records of polyploidy, but relatively high ploidy rank values that might be worthy of further study (e.g. *Sphagnum squarrosum*).

Although ploidy rank and polyploid p/a were lower on average in the liverworts than in the mosses, there was a notable pattern to the distribution of species with high ploidy rank. As noted above, the liverwort class Jungermanniopsida had multiple transitions to haploid genome sizes above the average for the liverworts, whereas the Marchantiopsida included species uniformly smaller than the average liverwort genome size (Fig. 1). Conversely, ploidy rank was higher on average in the Marchantiopsida than in the Jungermanniopsida, suggesting that WGD might figure more prominently in the evolution of Marchantiopsida (Fig. 2). In contrast to these findings, Shen *et al.* (2024) found phylogenomic evidence of an ancient large-scale duplication event in the ancestor of the Jungamanniopsida and not the Marchantiopsida, and Linde *et al.* (2023) also did not find signatures of WGD in either *Marchantia polymorpha* or *Lunularia cruciata*. We also found that both these species had a relatively low ploidy rank (0.125). However, further genomic studies might find distinct patterns of WGD in other members of the Marchantiopsida that have karyotypic evidence of WGD (e.g. *Dumortiera*, *Riccia*).

Although WGD is generally considered either rare or absent in the hornworts (Schuster, 1966; Kuta and Przywara, 2000), it has been documented in two hornwort species, *Anthoceros punctatus* and *Phaeoceros laevis* (Fig. 2; Supplementary Data, Table S2). Despite the published karyotypes documenting polyploidy in these species, genomic analyses of *A. punctatus* have not found evidence of WGD (Li *et al.*, 2020). To reconcile these findings, it should be noted that in our dataset *A. punctatus* and *P. laevis* were karyotyped for by far the largest number of populations among the hornworts (27 and 26 populations, from 10 and 14 studies, respectively). Patel *et al.* (2021) note a strong positive correlation between the number of populations karyotyped for a given species and the probability of recovering at least one polyploid cytotype. Overall, the hornworts are sparsely sampled for both genome size estimation (10 % of species) and karyotyping (8 % of species; Leitch and Leitch, 2012), and more intensive sampling within species might reveal evidence of additional polyploid cytotypes.

Polyploidy and haploid genome size

The metric ploidy rank was independent from haploid genome size, because these values were not phylogenetically correlated (Supplementary Data, Table S5). The lack of correlation between these two variables, particularly on a higher taxonomic level, further underscores their evolutionary independence in the bryophytes. For instance, the order Funariales

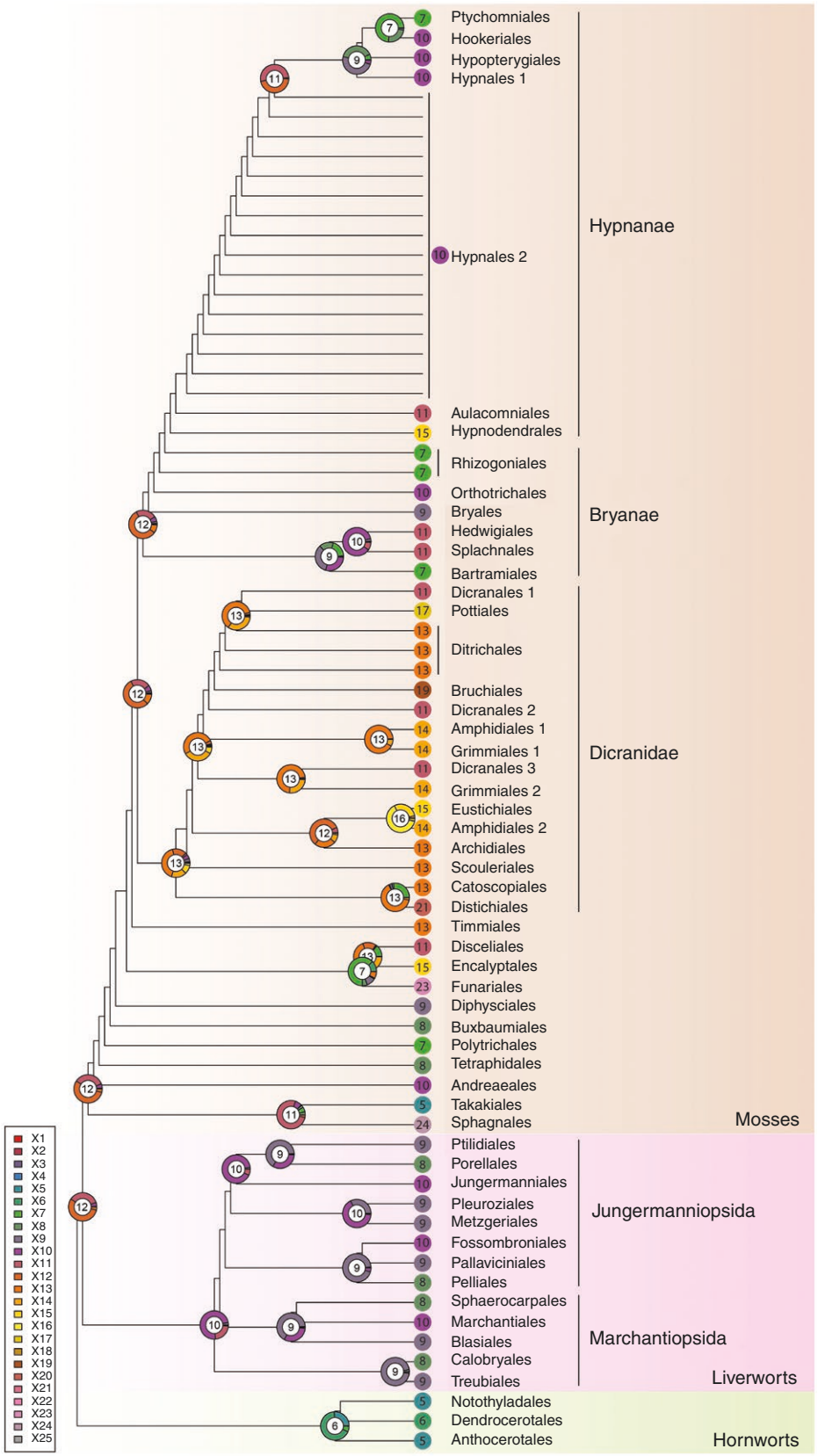

FIG. 3. Reconstruction of the ancestral haploid chromosome numbers ($1n$) inferred using ChromEvol after model fitting. The total species-level phylogeny, which includes 1088 species (Supplementary Data, Fig. S1), was collapsed to show the ordinal-level relationships in this tree. Note that pies depicting the likelihood of the inferred ancestral chromosome numbers are from the total species-level phylogeny, and colours and numbers at the tips represent the mean haploid chromosome number for each order. Given that the likelihood of the inferred ancestral chromosome number being above $1n = 25$ is never $>5.12 \times 10^{-23}$, only likelihoods of $1n$ values <25 are plotted, as indicated in the key.

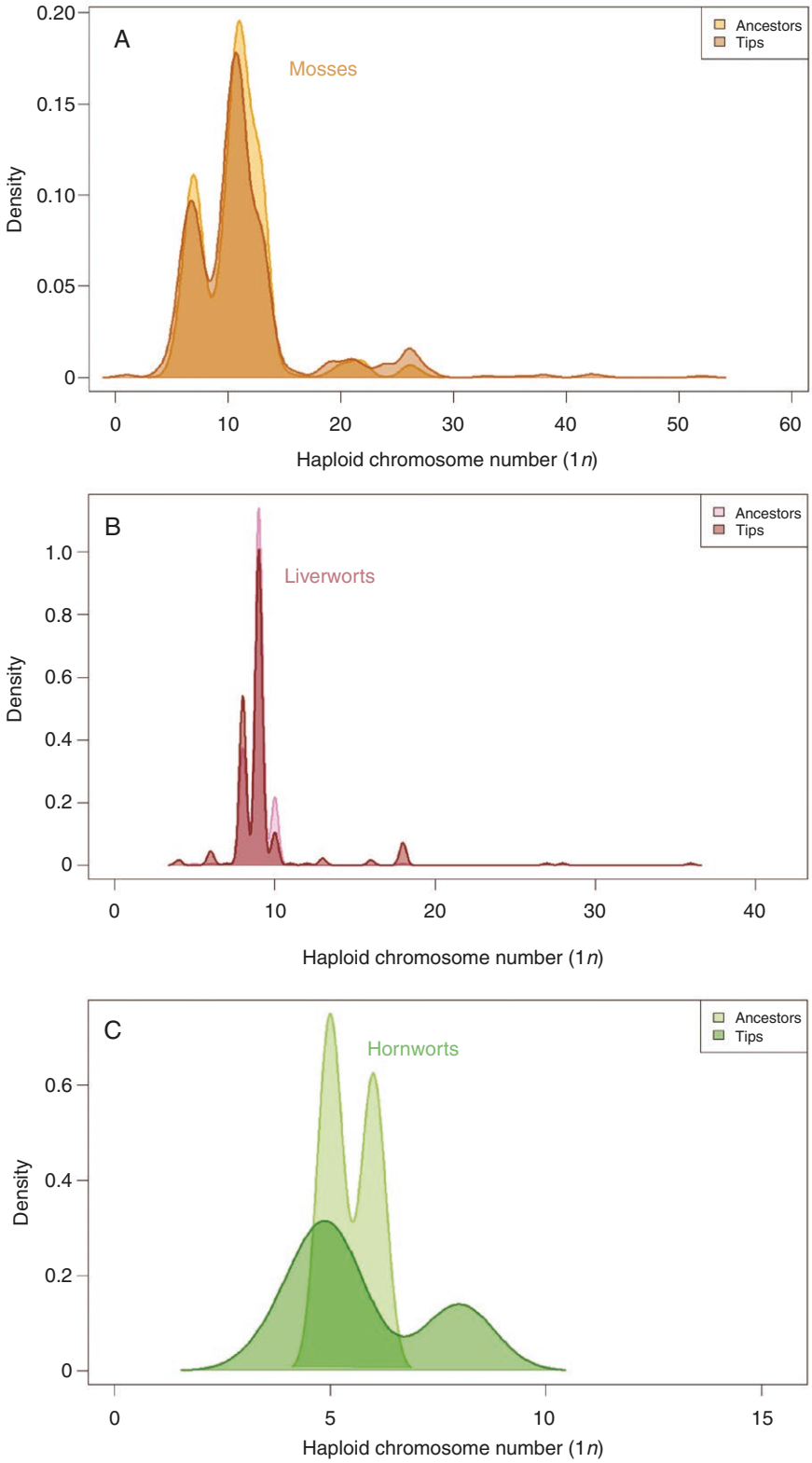

FIG. 4. Density plots depicting extant (tip state) haploid chromosome numbers (1*n*) across all karyotyped bryophyte species in comparison to the most likely ancestral haploid chromosome numbers at all internal nodes reconstructed using Chromevol, for each bryophyte lineage: mosses (A), liverworts (B) and hornworts (C).

had the highest mean ploidy rank (0.58) but a relatively small genome size among the mosses (1C = 0.46; Supplementary Data, Table S4). The lack of correlation between haploid genome size and ploidy rank, in addition to haploid genome size and polyploidy p/a, might reflect a strong evolutionary constraint on genome size across the bryophytes and particularly in the mosses. In the angiosperms, WGD can be followed by 'diploidization', which appears to result in genome downsizing or deletion of genomic content (Leitch and Bennett, 2004; Soltis *et al.*, 2015). In contrast, in the ferns, diploidization is not necessarily linked to genome downsizing (Barker, 2009; Barker and Wolf, 2010; Dauphin *et al.*, 2016), and within some fern lineages the chromosome number is positively correlated with genome size (Barker, 2009; Henry *et al.*, 2014; Clark *et al.*, 2016). Several studies have postulated selection for smaller genomes in angiosperms as a product of nucleotypic control, which refers to the phenotypic impacts of genome size without regard for genomic content or composition, particularly on cell size, growth rate and metabolism (Doyle and Coate, 2019). For instance, angiosperms have smaller genomes on average than gymnosperms, and it has been suggested that the smaller genome size values of angiosperms might have contributed to their greater diversity and abundance relative to gymnosperms (Carta and Peruzzi, 2016; Simonin and Roddy, 2018). Smaller genomes are also associated with higher reproductive rates and faster generation times in angiosperms (Bennett, 1972, 1987; Midgley and Bond, 1991). The patterns of WGD and haploid genome size in the bryophytes observed here suggest genome size constraints similar to the angiosperms, in which WGD is followed by rapid haploidization.

Nucleotypic effects are particularly important in the bryophytes given their poikilohydry and desiccation tolerance abilities (Proctor *et al.*, 1998). Many species endure periods of desiccation followed by metabolic reactivation when water becomes available (Duckett *et al.*, 2024). The capacity of organisms with smaller genomes to undergo rapid cell division when water is available might be advantageous to poikilohydric species (Bewley, 1979; Baniaga *et al.*, 2016; VanBuren *et al.*, 2018). In addition to nucleotypic controls impacting vegetative growth and metabolism, genome size might also be constrained by gametic cells. Bryophytes have flagellated sperm, which swim through thin films of water to achieve fertilization. Smaller haploid genome sizes result in streamlined cells, which might facilitate the movement of these biflagellate sperm (Renzaglia *et al.*, 1995), whereas flagellate plants with larger haploid genomes, such as the ferns, might necessitate multiflagellate sperm to enhance the dispersal of their larger sperm (Renzaglia and Garbary, 2001).

Desiccation tolerance has not been assessed in a sufficiently wide breadth of taxa to draw conclusions about patterns of evolution in this trait, and the relationship between haploid genome size and desiccation tolerance in bryophytes discussed here has not been tested. However, some genomic, physiological and ecological studies, in conjunction with our findings here, suggest that this relationship should be explored further. For example, Bainard (2011) found a negative correlation between haploid genome size and desiccation tolerance across 85 moss species. In contrast, some studies of vascular plants find that large genomes and concomitantly larger cells might be beneficial in water relationships (Scholes and Paige, 2015). In mosses, Zumel *et al.* (2023) considered the relationship between drought tolerance and polyploidy (including endopolyploidy, which is WGD in a subset of cells) across *Ceratodon* species. They found that polyploids and species with a higher degree of endopolyploidy occupy drier niches, suggesting a potential adaptive benefit of larger genomes and therefore larger cells. Perhaps mosses derive the benefits of larger genomes through a plastic increase via endopolyploidy (which is prevalent across mosses), thereby avoiding the drawbacks to growth and metabolism inherent to generative polyploidy. Building our understanding of the relationship between desiccation tolerance and genome size will require physiological experiments paired with genome size assessments analysed in a comparative framework for multiple groups of closely related species. These findings should be the impetus for further study of the complex relationship among generative polyploidy, endopolyploidy and desiccation tolerance in mosses.

CONCLUSION

In recent years, there has been an increasing focus on the patterns of genome size and chromosome number across the major lineages of land plants. By analysing published genome size estimates and karyotypes, we found distinct patterns of genome size and chromosome number evolution in each bryophyte lineage. Crucially, our comparative phylogenetic analyses revealed that WGD is not correlated with genome size in bryophytes. Although these findings should encourage further study of the haploidization process and how nucleotypic controls might constrain genome size in bryophytes, they also underscore the need for more karyotypic and flow cytometry data. These are key biological characteristics of eukaryotes that are insufficiently recorded in the literature for bryophytes, with only 10 % of species represented by at least one karyotype and only 5 % by at least one genome size estimate. Although karyotypic data, which are typically produced by chromosome squashing, are currently published with much less frequency than flow cytometry-based genome size estimates, we emphasize the need for the continued generation of both types of data. Studying karyotypic data, we not only found a polyploid frequency of 21 % among the mosses, but also numerous species that might include polyploid cytotypes on the basis of ploidy rank. Given the possibility of rapid haploidization following WGD, flow cytometry-based estimates might not consistently enable the recognition of polyploids. Overall, comparative phylogenetic studies strengthen our understanding of both past genome evolution and selection pressures on genome size in land plants (Liu *et al.*, 2019; Carta *et al.*, 2020), and our expansion of these approaches here, and in future studies, will undoubtedly underscore the unfolding complexity and dynamism of bryophyte nuclear genomes.

SUPPLEMENTARY DATA

Supplementary data are available at *Annals of Botany* online and consist of the following.

Figure S1: reconstruction of the ancestral haploid chromosome number (1*n*) inferred using ChromEvol after model fitting. Figure S2: ancestral character state reconstruction using

the 'ace' function from the package Phytools implemented in R, for the binary metric polyploidy p/a (present versus absent). Table S1: genome size estimates recovered from literature and database searches listed by species with the reference indicated. Table S2: chromosome counts recovered from literature searches listed by species with the reference indicated. Table S3: average haploid genome size, range, ploidy rank, polyploidy p/a and haploid chromosome number for each taxonomic order for all species recovered from literature and database searches in the total dataset, which includes species that do not have sequence data and thus could not be included in the phylogenetic analyses. Table S4: average haploid genome size, range, ploidy rank, polyploidy p/a and haploid chromosome number for each taxonomic order including only species represented in the molecular phylogenetic dataset. Table S5: R^2 and *P*-values for the regressions between each chromosome metric and haploid genome size. Table S6: model fit summary for the eight models fitted to the haploid chromosome number data using ChromEvol.

FUNDING

J.M.B. was supported by a grant from the United States National Science Foundation (DEB-2046467). The genome size data generated for the Darwin Tree of Life by Sahr Main was supported by the Wellcome Sanger Institute through a Darwin Tree of Life Discretionary Award (218328).

ACKNOWLEDGEMENTS

We are grateful to Sahr Mian for the genome size data estimated as part of the Darwin Tree of Life project. We also thank Ilia Leitch for facilitating our access to the data. We also appreciate the assistance of Lily Cooper in assembling the chromosome number dataset.
Conflict of interests: None declared.

AUTHOR CONTRIBUTIONS

All authors contributed to conceiving of the research and collecting the data. N. Patel performed analyses and wrote the first draft of the manuscript. All authors contributed to the improvement of analysis and manuscript drafts.

DATA AVAILABILITY

All phylogenetic data (matrices, trees and GenBank accession numbers) are available on Dryad https://doi.org/10.5061/dryad.31zcrjdwm. All raw chromosomal and genome size data are available in the Supplementary Data.

LITERATURE CITED

Anderson SN, **Stitzer MC**, **Brohammer AB**, ***et al***. **2019**. Transposable elements contribute to dynamic genome content in maize. *The Plant Journal: for Cell and Molecular Biology* **100**: 1052–1065.

Arabidopsis Genome Initiative. **2000**. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796–815.

Armstrong EM, **Larson ER**, **Harper H**, ***et al***. **2023**. One hundred important questions facing plant science: an international perspective. *New Phytologist* **238**: 470–481.

Bainard JD. 2011. *Patterns and biological implications of DNA content variation in land plants.* PhD thesis, University of Guelph, Guelph, ON, Canada.

Bainard JD, **Villarreal JC. 2013**. Genome size increases in recently diverged hornwort clades. *Genome* **56**: 431–435.

Bainard JD, **Forrest LL**, **Goffinet B**, **Newmaster SG. 2013**. Nuclear DNA content variation and evolution in liverworts. *Molecular Phylogenetics and Evolution* **68**: 619–627.

Bainard JD, **Newmaster SG**, **Budke JM. 2020**. Genome size and endopolyploidy evolution across the moss phylogeny. *Annals of Botany* **125**: 543–555.

Baniaga AE, **Arrigo N**, **Barker MS. 2016**. The small nuclear genomes of *Selaginella* are associated with a low rate of genome size evolution. *Genome Biology and Evolution* **8**: 1516–1525.

Barker MS. 2009. Evolutionary genomic analyses of ferns reveal that high chromosome numbers are a product of high retention and fewer rounds of polyploidy relative to angiosperms. *American Fern Journal* **99**: 136–141.

Barker MS, **Wolf PG. 2010**. Unfurling fern biology in the genomics age. *BioScience* **60**: 177–185.

Barker MS, **Arrigo N**, **Baniaga AE**, **Li Z**, **Levin DA. 2016**. On the relative abundance of autopolyploids and allopolyploids. *The New Phytologist* **210**: 391–398.

Barrington DS, **Patel NR**, **Southgate MW. 2020**. Inferring the impacts of evolutionary history and ecological constraints on spore size and shape in the ferns. *Applications in Plant Sciences* **8**: e11339.

Bechteler J, **Peñaloza-Bojacá G**, **Bell D**, ***et al***. **2023**. Comprehensive phylogenomic time tree of bryophytes reveals deep relationships and uncovers gene incongruences in the last 500 million years of diversification. *American Journal of Botany* **110**: e16249.

Beike AK, **von Stackelberg M**, **Schallenberg-Rüdinger M**, ***et al***. **2014**. Molecular evidence for convergent evolution and allopolyploid speciation within the *Physcomitrium-Physcomitrella* species complex. *BMC Evolutionary Biology* **14**: 158. https://doi.org/10.1186/1471-2148-14-158

Bennett MD. 1972. Nuclear DNA content and minimum generation time in herbaceous plants. *Proceedings of the Royal Society of London. Series B, Biological Sciences* **181**: 109–135.

Bennett MD. 1987. Variation in genomic form in plants and its ecological implications. *New Phytologist* **106**: 177–200.

Bennett MD, **Leitch IJ. 2005**. Genome size evolution in plants. In: **Gregory TR**. ed. *The evolution of the genome*. Cambridge, UK: Academic Press, 89–162.

Bennett MD, **Smith JB**, **Seal AG. 1986**. The karyotype of the grass *Zingeria biebersteiniana* (2*n* = 4) by light and electron microscopy. *Canadian Journal of Genetics and Cytology* **28**: 554–562.

Bennetzen JL, **Ma J**, **Devos KM. 2005**. Mechanisms of recent genome size variation in flowering plants. *Annals of Botany* **95**: 127–132.

Bewley JD. 1979. Physiological aspects of desiccation tolerance. *Annual Review of Plant Physiology* **30**: 195–238.

Boyle B, **Hopkins N**, **Lu Z**, ***et al***. **2013**. The taxonomic name resolution service: an online tool for automated standardization of plant names. *BMC Bioinformatics* **14**: 16. https://doi.org/10.1186/1471-2105-14-16

Brinda JC, **Atwood JJ. 2024**. *A classification of the Bryophyta. The bryophyte nomenclator*. https://www.bryonames.org/ (1 December 2024, date last accessed).

Burnham KP, **Anderson DR. 2004**. Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research* **33**: 261–304.

Carey SB, **Jenkins J**, **Lovell JT**, ***et al***. **2021**. Gene-rich UV sex chromosomes harbor conserved regulators of sexual development. *Science Advances* **7**: eabh2488.

Carta A, **Peruzzi L. 2016**. Testing the large genome constraint hypothesis: plant traits, habitat and climate seasonality in Liliaceae. *The New Phytologist* **210**: 709–716.

Carta A, **Bedini G**, **Peruzzi L. 2020**. A deep dive into the ancestral chromosome number and genome size of flowering plants. *The New Phytologist* **228**: 1097–1106.

Clark J, **Hidalgo O**, **Pellicer J**, ***et al***. **2016**. Genome evolution of ferns: evidence for relative stasis of genome size across the fern phylogeny. *The New Phytologist* **210**: 1072–1082.

Crawford M, **Jesson LK**, **Garnock-Jones PJ. 2009**. Correlated evolution of sexual system and life-history traits in mosses. *Evolution* **63**: 1129–1142.
Crum HA. 2001. *Structural diversity of bryophytes*. Ann Arbor, MI, USA: The University of Michigan Herbarium.
Dauphin B, **Grant J**, **Mráz P. 2016**. Ploidy level and genome size variation in the homosporous ferns *Botrychium s.l.* (Ophioglossaceae). *Plant Systematics and Evolution* **302**: 575–584.
de Abreu FPD, **Gonzatti F**, **Bahima IA**, **Souza-Chies TTD**, **Kaltchuk-Santos E. 2024**. A global review of chromosome number and genome size for the filmy ferns family (Hymenophyllaceae, Polypodiopsida). *Rodriguésia* **75**: e00552023.
del Pozo JC, **Ramirez-Parra E. 2015**. Whole genome duplications in plants: an overview from *Arabidopsis*. *Journal of Experimental Botany* **66**: 6991–7003.
Diop SI, **Subotic O**, **Giraldo-Fonseca A**, ***et al***. **2020**. A pseudomolecule-scale genome assembly of the liverwort *Marchantia polymorpha*. *The Plant Journal* **101**: 1378–1396.
Dong S, **Yu J**, **Zhang L**, **Goffinet B**, **Liu Y. 2022**. Phylotranscriptomics of liverworts: revisiting the backbone phylogeny and ancestral gene duplications. *Annals of Botany* **130**: 951–964.
Doyle JJ, **Coate JE. 2019**. Polyploidy, the nucleotype, and novelty: the impact of genome doubling on the biology of the cell. *International Journal of Plant Sciences* **180**: 1–52.
Duckett J. 2020. Towards completing understanding of genome size characters in plants. A commentary on: 'Genome size and endopolyploidy evolution across the moss phylogeny'. *Annals of Botany* **125**: iv.
Duckett JG, **Renzaglia KS**, **Pressel S. 2024**. Reflections on the absence of stomata on the gametophyte generation of extant land plants: a focus on poikilohydry. *Botanical Journal of the Linnean Society*: boae055. https://academic.oup.com/botlinnean/advance-article-abstract/doi/10.1093/botlinnean/boae055/7753969
Escudero M, **Wendel JF. 2020**. The grand sweep of chromosomal evolution in angiosperms. *The New Phytologist* **228**: 805–808.
Fernández P, **Amice R**, **Bruy D**, ***et al***. **2024**. A 160 Gbp fork fern genome shatters size record for eukaryotes. *iScience* **27**: 109889.
Fleischmann A, **Michael TP**, **Rivadavia F**, ***et al***. **2014**. Evolution of genome size and chromosome number in the carnivorous plant genus *Genlisea* (Lentibulariaceae), with a new estimate of the minimum genome size in angiosperms. *Annals of Botany* **114**: 1651–1663.
Francis D, **Davies MS**, **Barlow PW. 2008**. A strong nucleotypic effect on the cell cycle regardless of ploidy level. *Annals of Botany* **101**: 747–757.
Fritsch R. 1991. *Index to bryophyte chromosome counts*. Berlin, Germany: J. Cramer.
Gao B, **Chen MX**, **Li XS**, ***et al***. **2022**. Ancestral gene duplications in mosses characterized by integrated phylogenomic analyses. *Journal of Systematics and Evolution* **60**: 144–159.
Glick L, **Mayrose I. 2014**. ChromEvol: assessing the pattern of chromosome number evolution and the inference of polyploidy along a phylogeny. *Molecular Biology and Evolution* **31**: 1914–1922.
Goffinet B, **Buck WR. 2020**. *Classification of the Bryophyta*. http://bryology.uconn.edu/classification/ (25 January 2023, date last accessed).
Gregory TR. 2001. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biological Reviews of the Cambridge Philosophical Society* **76**: 65–101.
Greilhuber J, **Såstad SM**, **Flatberg KI. 2003**. Ploidy determination in *Sphagnum* samples from Svalbard, Arctic Norway, by DNA image cytometry. *Journal of Bryology* **25**: 235–239.
Harris BJ, **Clark JW**, **Schrempf D**, ***et al***. **2022**. Divergent evolutionary trajectories of bryophytes and tracheophytes from a complex common ancestor of land plants. *Nature Ecology & Evolution* **6**: 1634–1643.
Haufler CH. 1987. Electrophoresis is modifying our concepts of evolution in homosporous pteridophytes. *American Journal of Botany* **74**: 953–966.
Henry TA, **Bainard JD**, **Newmaster SG. 2014**. Genome size evolution in Ontario ferns (Polypodiidae): evolutionary correlations with cell size, spore size, and habitat type and an absence of genome downsizing. *Genome* **57**: 555–566.
Huang X, **Wang W**, **Gong T**, ***et al***. **2022**. The flying spider-monkey tree fern genome provides insights into fern evolution and arborescence. *Nature Plants* **8**: 500–512.
Jesson LK, **Cavanagh AP**, **Perley DS. 2011**. Polyploidy influences sexual system and mating patterns in the moss *Atrichum undulatum sensu lato*. *Annals of Botany* **107**: 135–143.
Jiao Y, **Leebens-Mack J**, **Ayyampalayam S**, ***et al***. **2012**. A genome triplication associated with early diversification of the core eudicots. *Genome Biology* **13**: R3. https://doi.org/10.1186/gb-2012-13-1-r3
Kalyaanamoorthy S, **Minh BQ**, **Wong TK**, ***et al***. **2017**. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**: 587–589.
Katoh K, **Standley DM. 2013**. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.
Khandelwal S. 1990. Chromosome evolution in the genus *Ophioglossum* L. *Botanical Journal of the Linnean Society* **102**: 205–217.
Kirbis A, **Rahmatpour N**, **Dong S**, ***et al***. **2022**. Genome dynamics in mosses: extensive synteny coexists with a highly dynamic gene space. *bioRxiv*. doi:10.1101/2022.05.17.492078 (preprint: not peer reviewed).
Knight CA, **Molinari NA**, **Petrov DA. 2005**. The large genome constraint hypothesis: evolution, ecology and phenotype. *Annals of Botany* **95**: 177–190.
Kuta E, **Przywara L. 2000**. Cytotaxonomy of bryophytes in the Department of Plant Cytology and Embryology, Institute of Botany, Jagiellonian University, Kraców, Poland. *IOPB Newsletter* **32**: 3–8.
Laenen B, **Shaw B**, **Schneider H**, ***et al***. **2014**. Extant diversity of bryophytes emerged from successive post-Mesozoic diversification bursts. *Nature Communications* **5**: 5134.
Landis JB, **Soltis DE**, **Li Z**, ***et al***. **2018**. Impact of whole-genome duplication events on diversification rates in angiosperms. *American Journal of Botany* **105**: 348–363.
Lang D, **Ullrich KK**, **Murat F**, ***et al***. **2018**. The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *The Plant Journal: for Cell and Molecular Biology* **93**: 515–533.
Leitch IJ, **Bennett MD. 2004**. Genome downsizing in polyploid plants. *Biological Journal of the Linnean Society* **82**: 651–663.
Leitch IJ, **Leitch AR. 2012**. Genome size diversity and evolution in land plants. In: **Greilhuber J**, **Dolezel J**, **Wendel JF**. eds. *Plant genome diversity, Vol. 2, Physical structure, behaviour and evolution of plant genomes*. Vienna, Austria: Springer Vienna, 307–322.
Leitch IJ, **Johnston E**, **Pellicer J**, ***et al***. **2019**. *Plant DNA C-values database (release 7.1)*. https://cvalues.science.kew.org/ (24 June 2024, date last accessed).
Levin DA. 2002. *The role of chromosomal change in plant evolution*. Oxford, UK: Oxford University Press.
Li D, **Luo G**, **Guo S**, ***et al***. **2023**. Nuclear DNA amounts in Chinese bryophytes estimated by flow cytometry: variation patterns and biological significances. *Plants (Basel, Switzerland)* **12**: 1564.
Li FW, **Nishiyama T**, **Waller M**, ***et al***. **2020**. *Anthoceros* genomes illuminate the origin of land plants and the unique biology of hornworts. *Nature Plants* **6**: 259–272.
Li Z, **McKibben MT**, **Finch GS**, **Blischak PD**, **Sutherland BL**, **Barker MS. 2021**. Patterns and processes of diploidization in land plants. *Annual Review of Plant Biology* **72**: 387–410.
Linde AM, **Eklund DM**, **Cronberg N**, **Bowman JL**, **Lagercrantz U. 2021**. Rates and patterns of molecular evolution in bryophyte genomes, with focus on complex thalloid liverworts, Marchantiopsida. *Molecular Phylogenetics and Evolution* **165**: 107295.
Linde AM, **Singh S**, **Bowman JL**, **Eklund M**, **Cronberg N**, **Lagercrantz U. 2023**. Genome evolution in plants: complex thalloid liverworts (Marchantiopsida). *Genome Biology and Evolution* **15**: evad014.
Liu HM, **Ekrt L**, **Koutecky P**, ***et al***. **2019**. Polyploidy does not control all: lineage-specific average chromosome length constrains genome size evolution in ferns. *Journal of Systematics and Evolution* **57**: 418–430.
Liu H, **Dai Z**, **Zhou J**, ***et al***. **2022***a*. The genome of the viviparous moss *Sphagnum palustre*: a resource for studying desiccation tolerance and polyploidy. *Plant Communications* **3**: 100351.
Liu H, **Zhao W**, **Zhang R**, **Mao J**, **Wang X. 2022***b*. Repetitive elements, sequence turnover and cyto-nuclear gene transfer in gymnosperm mitogenomes. *Frontiers in Genetics* **13**: 867736.
Mandáková T, **Lysak MA. 2018**. Post-polyploid diploidization and diversification through dysploid changes. *Current Opinion in Plant Biology* **42**: 55–65.
Mandáková T, **Guo X**, **Özüdoğru B**, **Mummenhoff K**, **Lysak MA. 2018**. Hybridization-facilitated genome merger and repeated chromosome fusion after 8 million years. *The Plant Journal* **96**: 748–760.
Mayrose I, **Zhan SH**, **Rothfels CJ**, ***et al***. **2015**. Methods for studying polyploid diversification and the dead end hypothesis: a reply to Soltis *et al*. (2014). *The New Phytologist* **206**: 27–35.

Medina R, **Johnson M**, **Liu Y**, ***et al***. **2018**. Evolutionary dynamism in bryophytes: phylogenomic inferences confirm rapid radiation in the moss family Funariaceae. *Molecular Phylogenetics and Evolution* **120**: 240–247.

Medina R, **Johnson MG**, **Liu Y**, **Wickett NJ**, **Shaw AJ**, **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 *Aphanorrhegma*. *Journal of Systematics and Evolution* **57**: 404–417.

Melosik I, **Ordzykoski II**, **Sliwinska E. 2005**. Delimitation of taxa of *Sphagnum subsecundum s.l.* (Musci, Sphagnaceae) based on multienzyme phenotype and cytological characters. *Nova Hedwigia* **80**: 397–412.

Midgley JJ, **Bond WJ. 1991**. Ecological aspects of the rise of angiosperms: a challenge to the reproductive superiority hypotheses. *Biological Journal of the Linnean Society* **44**: 81–92.

Minh BQ, **Schmidt HA**, **Chernomor O**, ***et al***. **2020**. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**: 1530–1534.

Nakazato T, **Barker MS**, **Rieseberg LH**, **Gastony GJ. 2008**. Evolution of the nuclear genome of ferns and lycophytes. In: **Ranker TA**, **Haufler CH**. eds. *Biology and evolution of ferns and lycophytes*. Cambridge, UK: Cambridge University Press, 175–198.

Natcheva R, **Cronberg N. 2004**. What do we know about hybridization among bryophytes in nature? *Canadian Journal of Botany* **82**: 1687–1704.

Nieto Feliner G, **Casacuberta J**, **Wendel JF. 2020**. Genomics of evolutionary novelty in hybrids and polyploids. *Frontiers in Genetics* **11**: 792. https://doi.org/10.3389/fgene.2020.00792

Orzechowska M, **Siwinska D**, **Maluszynska J. 2010**. Molecular cytogenetic analyses of haploid and allopolyploid *Pellia* species. *Journal of Bryology* **32**: 113–121.

Orzechowska M, **Figura K**, **Siwinska D. 2018**. Chromosomal distribution of rRNA genes in the karyotypes of two dioicous liverwort species from the genus *Pellia* Raddi. *Journal of Bryology* **40**: 384–392.

Pagel M. 1999. Inferring the historical patterns of biological evolution. *Nature* **401**: 877–884.

Pandit MK, **White SM**, **Pocock MJ. 2014**. The contrasting effects of genome size, chromosome number and ploidy level on plant invasiveness: a global analysis. *New Phytologist* **203**: 697–703.

Paradis E, **Schliep K. 2019**. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**: 526–528.

Patel N, **Li CX**, **Zhang LB**, **Barrington DS. 2018**. Biodiversity and apomixis: insights from the East-Asian holly ferns in *Polystichum* section *Xiphopolystichum*. *Molecular Phylogenetics and Evolution* **127**: 345–355.

Patel N, **Medina R**, **Johnson M**, **Goffinet B. 2021**. Karyotypic diversity and cryptic speciation: have we vastly underestimated moss species diversity? *Bryophyte Diversity and Evolution* **43**: 150–163.

Patel N, **Medina R**, **Williams LD**, **Lemieux O**, **Goffinet B**, **Johnson MG. 2023**. Frequent allopolyploidy with distant progenitors in the moss genera *Physcomitrium* and *Entosthodon* (Funariaceae) identified via subgenome phasing of targeted nuclear genes. *Evolution* **77**: 2561–2575.

Pellicer J, **Fay MF**, **Leitch IJ. 2010**. The largest eukaryotic genome of them all? *Botanical Journal of the Linnean Society* **164**: 10–15.

Pellicer J, **Hidalgo O**, **Dodsworth S**, **Leitch IJ. 2018**. Genome size diversity and its impact on the evolution of land plants. *Genes* **9**: 88.

Perley DS, **Jesson LK. 2015**. Hybridization is associated with changes in sexual system in the bryophyte genus *Atrichum*. *American Journal of Botany* **102**: 555–565.

Proctor MCF, **Nagy Z**, **Csintalan Z**, **Takács Z. 1998**. Water-content components in bryophytes: analysis of pressure–volume relationships. *Journal of Experimental Botany* **49**: 1845–1854.

Pustahija F, **Brown SC**, **Bogunić F**, ***et al***. **2013**. Small genomes dominate in plants growing on serpentine soils in West Balkans, an exhaustive study of 8 habitats covering 308 taxa. *Plant and Soil* **373**: 427–453.

R Core Team. **2021**. *R: a language and environment for statistical computing, v.4.1.0*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

R Development Core Team. 2011. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Rensing SA, **Lang D**, **Zimmer AD**, ***et al***. **2008**. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* **319**: 64–69.

Rensing SA, **Beike AK**, **Lang D. 2012**. Evolutionary importance of generative polyploidy for genome evolution of haploid-dominant land plants. In: **Greilhuber J**, **Dolezel J**, **Wendel JF**. eds. *Plant genome diversity, Vol. 2, Physical structure, behaviour and evolution of plant genomes*. Vienna, Austria: Springer Vienna, 295–305.

Renzaglia KS, **Garbary DJ. 2001**. Motile gametes of land plants: diversity, development, and evolution. *Critical Reviews in Plant Sciences* **20**: 107–213.

Renzaglia KS, **Rasch EM**, **Pike LM. 1995**. Estimates of nuclear DNA content in bryophyte sperm cells: phylogenetic considerations. *American Journal of Botany* **82**: 18–25.

Revell L. 2024. phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ* **12**: e16505.

Ricca M, **Shaw AJ. 2010**. Allopolyploidy and homoploid hybridization in the *Sphagnum subsecundum* complex (Sphagnaceae: Bryophyta). *Biological Journal of the Linnean Society* **99**: 135–151.

Ricca M, **Beecher FW**, **Boles SB**, ***et al***. **2008**. Cytotype variation and allopolyploidy in North American species of the *Sphagnum subsecundum* complex (Sphagnaceae). *American Journal of Botany* **95**: 1606–1620.

Sanmiguel P, **Bennetzen JL. 1998**. Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Annals of Botany* **82**: 37–44.

Scholes DR, **Paige KN. 2015**. Plasticity in ploidy: a generalized response to stress. *Trends in Plant Science* **20**: 165–175.

Schuster RM. 1966. *The Hepaticae and Anthocerotae of North America, east of the hundredth meridian*, Vol. **I**. New York, NY, USA: Columbia University Press.

Schween G, **Hohe A**, **Koprivova A**, **Reski R. 2003**. Effects of nutrients, cell density and culture techniques on protoplast regeneration and early protonema development in a moss, *Physcomitrella patens*. *Journal of Plant Physiology* **160**: 209–212.

Shen C, **Li H**, **Shu L**, **Huang WZ**, **Zhu RL. 2024**. Ancient large-scale gene duplications and diversification in bryophytes illuminate the plant terrestrialization. *New Phytologist,* 1–17.

Silva AT, **Gao B**, **Fisher KM**, ***et al***. **2021**. To dry perchance to live: insights from the genome of the desiccation-tolerant biocrust moss *Syntrichia caninervis*. *The Plant Journal: for Cell and Molecular Biology* **105**: 1339–1356.

Simonin KA, **Roddy AB. 2018**. Genome downsizing, physiological novelty, and the global dominance of flowering plants. *PLoS Biology* **16**: e2003706.

Sliwińska E, **Loureiro J**, **Leitch IJ**, ***et al***. **2022**. Application-based guidelines for best practices in plant flow cytometry. *Cytometry. Part A : The Journal of the International Society for Analytical Cytology* **101**: 749–781.

Smith SA, **Walker JF. 2019**. Py PHLAWD: A python tool for phylogenetic dataset construction. *Methods in Ecology and Evolution* **10**: 104–108.

Söderström L, **Hagborg A**, **Von Konrat M**, ***et al***. **2016**. World checklist of hornworts and liverworts. *PhytoKeys* **59**: 1–828.

Soltis DE, **Albert VA**, **Leebens-Mack J**, ***et al***. **2009**. Polyploidy and angiosperm diversification. *American Journal of Botany* **96**: 336–348.

Soltis PS, **Marchant DB**, **Van de Peer Y**, **Soltis DE. 2015**. Polyploidy and genome evolution in plants. *Current Opinion in Genetics & Development* **35**: 119–125.

Steere WC. 1954. Chromosome number and behavior in arctic mosses. *Botanical Gazette* **116**: 93–133.

Steere WC. 1972. Chromosome numbers in bryophytes. *The Journal of the Hattori Botanical Laboratory* **35**: 99–125.

Suda J, **Meyerson LA**, **Leitch IJ**, **Pyšek P. 2015**. The hidden side of plant invasions: the role of genome size. *The New Phytologist* **205**: 994–1007.

Szövényi P, **Gunadi A**, **Li FW. 2021**. Charting the genomic landscape of seed-free plants. *Nature Plants* **7**: 554–565.

Temsch EM, **Greilhuber J**, **Krisai R. 1998**. Genome size in *Sphagnum* (peat moss). *Botanica Acta* **111**: 325–330.

Temsch EM, **Greilhuber J**, **Krisai R. 2010**. Genome size in liverworts. *Preslia* **82**: 63–80.

Tuskan GA, **Difazio S**, **Jansson S**, ***et al***. **2006**. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596–1604.

Uhl CH. 1978. Chromosomes of Mexican *Sedum* II. Section *Pachysedum*. *Rhodora* **80**: 491–512.

VanBuren R, **Man WC**, **Pardo J**, ***et al***. **2018**. Desiccation tolerance evolved through gene duplication and network rewiring in *Lindernia*. *The Plant Cell* **30**: 2943–2958.

Van de Peer Y, **Mizrachi E**, **Marchal K. 2017**. The evolutionary significance of polyploidy. *Nature Reviews Genetics* **18**: 411–424.

Villarreal JC, **Goffinet B. 2023**. *Classification of the Phylum Anthocerotophyta Stotl. & Crand.-Stotl.* http://bryology.uconn.edu/classification-hornworts/ (25 January 2023, date last accessed).

Villarreal JC, **Renner SS. 2012**. Hornwort pyrenoids, carbon-concentrating structures, evolved and were lost at least five times during the last 100 million years. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 18873–18878.

Villarreal AJC, **Renzaglia KS. 2006**. Structure and development of *Nostoc* strands in *Leiosporoceros dussii* (Anthocerotophyta): a novel symbiosis in land plants. *American Journal of Botany* **93**: 693–705.

Villarreal JC, **Renzaglia KS. 2015**. The hornworts: important advancements in early land plant evolution. *Journal of Bryology* **37**: 157–170.

Vitt DH. 1971. The infrageneric evolution, phylogeny, and taxonomy of the genus *Orthotrichum* (Musci) in North America. *Nova Hedwigia* **21**: 683–711.

Voglmayr H. 2000. Nuclear DNA amounts in mosses (Musci). *Annals of Botany* **85**: 531–546.

Wang D, **Zheng Z**, **Li Y**, ***et al***. **2021**. Which factors contribute most to genome size variation within angiosperms? *Ecology and Evolution* **11**: 2660–2668.

Wicker T, **Yahiaoui N**, **Guyot R**, ***et al***. **2003**. Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genomes of wheat. *The Plant Cell* **15**: 1186–1197.

Wyatt R, **Odrzykoski IJ**, **Stoneburner A**, **Bass HW**, **Galau GA. 1988**. Allopolyploidy in bryophytes: multiple origins of *Plagiomnium medium*. *Proceedings of the National Academy of Sciences of the United States of America* **85**: 5601–5604.

Yu G. 2020. Using ggtree to visualize data on tree-like structures. *Current Protocols in Bioinformatics* **69**: e96.

Zhang J, **Fu XX**, **Li RQ**, ***et al***. **2020**. The hornwort genome and early land plant evolution. *Nature Plants* **6**: 107–118.

Zhang X, **Wessler SR. 2004**. Genome-wide comparative analysis of the transposable elements in the related species *Arabidopsis thaliana* and *Brassica oleracea*. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 5589–5594.

Zumel D, **Diéguez X**, **Werner O**, **Moreno-Ortiz MC**, **Muñoz J**, **Ros RM. 2023**. High endoreduplication after drought-related conditions in haploid but not diploid mosses. *Annals of Botany* **132**: 1249–1258.