REVIEW ARTICLE
https://doi.org/10.1038/s41477-021-00888-z

nature plants

Charting the genomic landscape of seed-free plants

Péter Szövényi[1]✉, Andika Gunadi[2] and Fay-Wei Li[2,3]

During the past few years several high-quality genomes has been published from Charophyte algae, bryophytes, lycophytes and ferns. These genomes have not only elucidated the origin and evolution of early land plants, but have also provided important insights into the biology of the seed-free lineages. However, critical gaps across the phylogeny remain and many new questions have been raised through comparing seed-free and seed plant genomes. Here, we review the reference genomes available and identify those that are missing in the seed-free lineages. We compare patterns of various levels of genome and epigenomic organization found in seed-free plants to those of seed plants. Some genomic features appear to be fundamentally different. For instance, hornworts, *Selaginella* and most liverworts are devoid of whole-genome duplication, in stark contrast to other land plants. In addition, the distribution of genes and repeats appear to be less structured in seed-free genomes than in other plants, and the levels of gene body methylation appear to be much lower. Finally, we highlight the currently available (or needed) model systems, which are crucial to further our understanding about how changes in genes translate into evolutionary novelties.

For over 150–200 million years after colonization of the land, the terrestrial vegetation was dominated by seed-free plants. Modern-day seed-free plants are a paraphyletic assemblage represented by bryophytes (mosses, liverworts and hornworts), lycophytes and ferns (Fig. 1). From the evolutionary perspective, seed-free plants hold the key to retracing the major transitions in land plant evolution; from the applied perspective, they are the vital outgroup to better understand the biology of agronomically important traits such as seeds, fruits and flowers.

The phylogenetic relationships of seed-free lineages have been widely debated, especially the relationships among the bryophytes. Almost all the possible combinations of branching orders between mosses, liverworts, hornworts and vascular plants have been proposed on the basis of morphological, ribosomal and/or organellar DNA evidence (reviewed in refs. [1–3]). Only recently have phylogenomic studies with transcriptomic and genomic datasets started to provide more definitive answers.

Wickett et al.[1] were the first to apply a large number of nuclear genes to infer the phylogeny of Viridiplantae. In their study, a sister relationship between mosses and liverworts was consistently recovered with strong support, whereas the position of hornworts varied depending on the data types (nucleotide versus amino acids), subsets (codon position or filtering threshold) and inference methods (concatenation versus species-tree method or maximum likelihood versus Bayesian)[1]. Subsequently, Puttick et al.[2] and de Sousa et al.[3] reanalysed the Wickett et al.[1] dataset with methods that better modelled rate and compositional heterogeneities. Both studies confirmed that mosses and liverworts comprised a single clade, and de Sousa et al.[3] further resolved bryophytes as monophyletic with high confidence. However, it should be stressed that the Wickett et al.[1] dataset has a very limited hornwort representation, with transcriptomes from only two closely related *Nothoceros* species. A more balanced sampling came in 2019 with the full release of the One Thousand Plants (1KP) transcriptomes[4]. The analyses by 1KP[4] and by Harris et al.[5] both supported the placement of hornworts as sister to mosses and liverworts. The monophyly of all bryophytes was further bolstered by the recent analyses of hornwort genomes[6,7]. Mounting evidence suggests that extant land plants are essentially composed of two monophyletic groups: bryophytes and vascular plants (Fig. 1). This phylogenetic framework is facilitating new research directions to revisit the major transitions of land plants as well as the underlying genetic changes.

There has been a recent renaissance of studies into seed-free plants. This has been driven partly by the development of efficient gene-editing methodologies in mosses and liverworts, as well as an increasing number of high-quality genomes from across seed-free plants and the close algal relatives of land plants (Charophyte algae) (Fig. 1). In this Review, we first lay out the current genomic landscape across seed-free plants and point out the critical gaps that need to be filled. We then highlight some of the unique features that have emerged from genomic studies of seed-free plants, and discuss their significance in comparison to seed plant genomes. Finally, we outline future research directions to advance our understanding of genome evolution in seed-free and seed-bearing plants.

The current gaps

Plant genomic research has historically focused on crop plants, which are phylogenetically restricted to a small number of angiosperm clades. This imbalance is reflected in the fact that the moss *Physcomitrium patens* and lycophyte *Selaginella moellendorfii* had been the only available seed-free genomes for more than six years[8,9]. Availability of genomes began to accelerate in 2017, thanks largely to the advent of long-read sequencing, and in the subsequent 3 years a total of 28 genomes have been published or made available for seed-free plants[6–8,10–22] as well as the Charophyte algae[23–28] (Fig. 1 and Table 1). High-quality genomes are now available for almost every major lineage of plants (Fig. 1). The noteworthy exceptions are discussed below.

Bryophytes. It is important to emphasize that the three lineages of bryophytes (mosses, liverworts and hornworts) are a result of more than 400 million years of independent evolution, going back to the Cambrian–Ordovician period[29]. Although multiple genome assemblies are available for all three lineages of bryophytes, the phylogenetic diversity of each group is poorly covered. All four published hornwort genomes are from *Anthoceros*[6,7], and all the liverwort genomes are from *Marchantia*[10–14] (Fig. 1). For mosses, high-quality

[1]Department of Systematic and Evolutionary Botany, University of Zurich and Zurich-Basel Plant Science Center, Zurich, Switzerland. [2]Boyce Thompson Institute, Ithaca, NY, USA. [3]Plant Biology Section, Cornell University, Ithaca, NY, USA. ✉e-mail: peter.szoevenyi@uzh.ch

genomes based on long reads are available only for *P. patens*[15], *Ceratodon purpureus*[22] and *Sphagnum* spp. (http://phytozome.jgi.doe.gov/), none of which belong to the hyperdiverse pleurocarpous lineage. Future work is needed to strategically sequence genomes that can better cover the bryophyte phylogeny.

Isoetales. Although members of Isoetales had once dominated the Earth's terrestrial ecosystem, only a single extant genus, *Isoetes* (quillworts), remains. The closest living relative of *Isoetes* is *Selaginella* in Selaginellales (Fig. 1), which diverged around 380 million years ago (Ma)[30]. Many morphological and developmental features of *Isoetes* are unique among land plants; for example, *Isoetes* roots are developed from a specialized rhizomorph structure and are isotomously branched. *Isoetes* species are mostly aquatic, and when underwater, they carry out photosynthesis via the crassulacean acid metabolism (CAM) pathway, which is typically a water-conserving mechanism in xeric-adapted plants. CAM in *Isoetes* is believed to be an adaptation to limited CO_2 availability in aquatic environments[31]. Comparing *Isoetes* genomic features and diel gene expression profiles with other CAM plants will provide new insights into the convergent evolution of CAM. For genome sequencing, because many *Isoetes* species are recent polyploids (up to dodecaploid), it is imperative to locate diploid species. The 1C value of a diploid species is around 1.7 Gbp[32].

Lycopodiales. This lineage is sister to *Isoetes* and *Selaginella*, and diverged around 400 Ma[29]. Lycopodiales are homosporous (that is producing one type of spores), which is distinct from the heterosporous *Isoetes* and *Selaginella* (having separate mega- and microspores). The transition between homospory and heterospory has been hypothesized to have large impacts on genome structure. In particular, homosporous individuals have the potential for intragametophytic selfing, which would result in completely homozygous offspring. To compensate for such sudden loss of heterozygosity, polyploidization would be favoured as a means to inject genetic diversity[33]. To test this hypothesis, lycophytes and ferns are the two prime groups, given they each have experienced independent homo–heterospory transitions. In both groups, homosporous lineages tend to have much larger genomes than the heterosporous relatives, consistent with the hypothesis[34]. The life cycle of Lycopodiales is also unique in that the gametophytes are mycoheterotrophic, relying on fungal partners to provide carbon sources. Very little is known about this mycoheterotrophic interaction. Lycopodiales genome sizes range from 2.4–5.6 Gbp[32].

Homosporous ferns. Both of the published fern genomes (*Azolla filiculoides* and *Salvinia cucullata*) belong to the heterosporous Salviniales[6], whose genome sizes are on average about an order of magnitude smaller than those of homosporous ferns[34]. Because Salviniales are nested within ferns, their genomes were probably secondarily reduced and unlikely to represent a general fern genome. The model fern *Ceratopteris richardii* is homosporous and has a large genome of 14 Gbp, typical of most fern lineages[34]. Recently a fragmented, short-read assembly of *C. richardii* was published[35], which unfortunately covered only a third of the estimated genome size and highlighted the difficulty in assembling a homosporous fern genome.

Whole-genome duplications

Whole-genome duplication (WGD) and the generated redundancy is thought to be one of the key drivers in plant evolution[36]. For flowering plants, various studies have suggested that WGDs preceded and are correlated with the evolution of key diversification events[4,37], and have decreased extinction risk over the Cretaceous–Paleogene mass-extinction event[38]. In contrast to the multiple rounds of WGDs in all major lineages of flowering plants[4], the picture appears to be more variable in seed-free plants. No WGD was detected in the hornwort[6,7], liverwort[10] and *Selaginella*[9] genomes sequenced to date. Analyses based on a broader sampling of transcriptomes also found no or only a few confident WGD events in these lineages[4,15]. By contrast most mosses, ferns and homosporous lycophytes have experienced at least one well-supported WGD in their history[4,6,7,15,20,35,38–40].

It is unclear what genomic or historical factors made some of the seed-free plant lineages more prone to WGDs than others. It has been proposed that the paucity of WGD in liverworts may be a consequence of the early evolution of dimorphic sex chromosomes, which, when duplicated, could cause difficulties during meiotic segregation[10]. By contrast, sex chromosomes are assumed to be relatively young in mosses and angiosperms, enabling WGDs to have accumulated. However, recent evidence contradicts this hypothesis: the genome of the moss *C. purpureus* was revealed to have experienced at least one WGD but also harbours an ancient sex chromosome system[22]. Similarly, phylogenetic evidence suggests relatively recent origins of sex chromosomes in hornworts, which is counter to the hypothesis that the age of the sex chromosomes would restrict WGDs[41].

How WGDs have contributed to genome evolution and innovations in seed-free plants is even less clear. There is some evidence that WGDs probably promoted the diversification of peat mosses and the pleurocarp mosses, at species, physiological and morphological levels[42,43]. For ferns, the limited number of studies have produced somewhat contradictory results[40,44,45]. For instance, reanalysis of *Equisetum* transcriptomes by Clark et al.[46] did not support WGD conferring a reduced extinction risk at the Cretaceous–Palaeogene boundary as previously suggested[45]. Huang et al.[40] recently identified 19 WGDs in ferns on the basis of transcriptome data, and found a positive correlation between WGD occurrence and shifts in diversification rates. However, aside from the uncertainty in inferring rate shifts (especially with a small sample size), the majority of the WGD events did not actually coincide with major diversification events. In other words, the extent to which WGDs fuel fern diversification remains to be resolved.

The limited understanding summarized above is in contrast to the vast amount of information available for seed plants about the impact of WGDs on genomic instability, genome downsizing and reshuffling, epigenetic changes, speciation, phenotypic diversification, adaptation and extinction resistance[47–52]. Many of these aspects have yet to be investigated in seed-free plants, and we believe that seed-free plants could actually provide an ideal system to complement the studies on seed plants. The highly variable frequency of WGDs among seed-free plant lineages would enable multiple comparisons between WGD-poor and WGD-rich lineages to better understand the contribution of WGD to genome evolution as well as overall diversification rate. Similar comparisons could also be made at a shallower phylogenetic scale by comparing groups of

Fig. 1 | Phylogeny of streptophytes (charophyte algae and land plants) and the available genomic resources. The topology was largely based on refs. [4,172,173]. Genome sizes, scaffold and contig N50 lengths were calculated by the summarizeAssembly.py function of PBSuite[174]. Genomes with scaffold N50 > 100 kbp are shown in bold, and asterisks indicates genomes assembled into pseudochromosomes. The draft assembly of *Ceratopteris richardii*[35] was not included here, as only a third of the genome was assembled. The recently published genome of *Synthrichia caninervis* was a short-read-only assembly but scaffolded onto pseudochromosomes[135]. WGD events inferred by 1KP[4] are marked by inverse triangles and associated identifiers; only WGDs that were supported by a combination of Ks (number of synonymous substitutions per synonymous sites) plots, orthologue divergence and multi-taxon paleopolyploidy search algorithm (MAPS) analyses[4] are plotted here.

closely related species with contrasting WGD histories. In addition, bryophytes and some ferns are attractive models in which to investigate the molecular underpinnings of a wide range of processes associated with WGDs (including genomic shock, paralogue expression, functional diversification, epigenetic changes and adaptation) due to their moderate genome sizes and reverse-genetic tractability. Finally, the haploid-dominant life cycle of bryophytes and the free-living gametophyte phase of most ferns and lycophytes

Table 1 | Summary of seed-free plant model systems

	In vitro propagation	Genetic transformation	Gene silencing	Genome editing
Ferns				
Adiantum capillus-veneris	Full life cycle[145] and somatic embryogenesis[146]	Particle bombardment[147]	DNAi[148]	–
*Azolla filiculoides**	Full life cycle[149]	–	–	–
Ceratopteris spp.	Full life cycle[150]	*Agrobacterium*-mediated[126,151] and particle bombardment[125]	RNAi[152], DNAi[153]	–
Marsilea vestita	Gametophyte[154] and sporophyte[155]	–	RNAi[154]	–
Pteris vittata	Full life cycle[156]	*Agrobacterium*-mediated[151]	RNAi[157]	–
Salvinia spp.*	Full life cycle[158]	–	–	–
Lycophytes				
Huperzia selago	Gametophyte[131] and sporophyte[159]	–	–	–
Isoetes echinospora	Full life cycle[160]	–	–	–
Lycopodium spp.	Full life cycle[161]	–	–	–
Selaginella spp.*	Full life cycle[132,162]	–	–	–
Mosses				
*Ceratodon purpureus**	Full life cycle[163]	Protoplast transfection[77,164]	–	Gene targeting[77]
*Funaria hygrometrica**	Full life cycle[163]	–	–	–
*P. patens**	Full life cycle[133,163]	*Agrobacterium*-mediated, particle bombardment and protoplast transfection[133]	RNAi[133]	CRISPR-targeted mutagenesis and gene targeting[133]
Sphagnum spp.*	Gametophyte[137,165]	–	–	–
*Syntrichia caninervis**	Gametophyte[136]	–	–	–
Syntrichia ruralis	Full life cycle[163]	–	–	–
Liverworts				
*M. polymorpha**	Full life cycle[166]	*Agrobacterium*-mediated and particle bombardment[166,167]	RNAi[168]	CRISPR-targeted mutagenesis[169] and gene targeting[138]
Riccia spp.	Gametophyte[141]	*Agrobacterium*-mediated[141]	–	–
Hornworts				
Anthoceros spp.*	Full life cycle[170], gametophyte and cyanobacteria symbiosis[71,171]	*Agrobacterium*-mediated[142] and particle bombardment (A.G., unpublished results)	–	–

Model systems with assembled genomes (see in Fig. 1) are marked with an asterisk. CRISPR, clustered regularly interspaced short palindromic repeats.

provides a unique opportunity to study how extensive exposure in the haploid phase affects the frequency of WGDs and subsequent genomic evolution. This question cannot be properly addressed in seed plants because of the highly reduced haploid phase. Altogether, we are confident that future studies on seed-free plants will provide a much more holistic view of the biological significance of WGDs.

Genome size variation and its potential drivers

Genome sizes of land plants are highly variable, with a greater variation in angiosperms than in seed-free plants[53]. Among seed-free plants, ferns and lycophytes exhibit larger and more variable genome sizes overall than bryophytes[54,55]. In addition, genome size distribution is less skewed towards smaller sizes in seed-free plants than in seed plants, especially in lycophytes and ferns. These observations suggest that the trajectory of genome size evolution might differ between seed-free and seed-bearing lineages in various aspects. The evidence suggests contrasting patterns between ferns and lycophytes and bryophytes; we therefore discuss them separately below.

Ferns and lycophytes. Genome sizes and chromosome numbers are significantly correlated in ferns and lycophytes, but not in any other plant lineages[56]. This correlation is probably the result of at least two processes. First, ferns and lycophytes may differ from angiosperms in their post-polyploidization genomic processes. The predominantly small (yet historically polyploid) genomes of angiosperms imply that post-polyploidization genome fractionation and downsizing is effective and frequent. By contrast, the diploidization process in ferns might be slower[56], although the extent to which it differs from that of angiosperms is unclear. Second, it has been hypothesized that DNA content per chromosome is constrained in ferns and lycophytes, resulting in more chromosomes being needed to sustain a larger genome[55]. Available data suggest that chromosomes of ferns and lycophytes are smaller and more uniform in size compared with those of angiosperms[57–59]. In particular, chromosome size variation is 3,100-fold in angiosperms but only 31-fold in ferns[55]. However, the underlying mechanisms constraining chromosome sizes in ferns and lycophytes remain unknown. Observations in angiosperms suggest that above a certain chromosome arm/

spindle length ratio, mitotic divisions will fail[60]. It is tempting to suggest that this threshold might be lower overall in most ferns and lycophytes, but further studies are needed.

Besides polyploidy, activity of transposable elements (TEs), especially long terminal repeat (LTR) transposons are thought to contribute to the majority of genome size variation in angiosperms and gymnosperms[61]. Until recently, it has been assumed that the role of LTRs in genome size evolution may be less prominent in ferns and lycophytes because chromosome numbers are well correlated with genome sizes. Indeed, a few studies have suggested that there is either no correlation between LTR abundance and genome size or that LTR activity may have been too recent to have an effect[9,19,62]. By contrast, Baniaga & Barker[63] recently found that, similar to seed plants, the timing of median LTR activity is also positively correlated with genome size in fern and lycophyte taxa. It is therefore possible that LTR accumulation, together with a slower post-WGD diploidization rate, may have jointly contributed to larger genome sizes in lycophytes and ferns.

Altogether, genomic processes driving genome size evolution in ferns and lycophytes are still poorly understood and several hypotheses remain to be tested. More information is needed regarding the activity of LTR elements in ferns and lycophytes. Current analyses of LTRs have been largely restricted to comparisons between homosporous versus heterosporous groups, whose genome size difference may be confounded by their contrasting reproductive strategies[63]. Therefore, estimating LTR abundance and activity among closely related species with similar life history traits should provide much-needed information on the contribution of TEs to genome size evolution. Studies at the infraspecific level would also help to clarify the dynamics of TEs over a shorter time scale. In particular, apomixis—a form of asexual reproduction in which fertilization is bypassed during phase transition—is prevalent in many ferns, and its frequency can vary greatly within a species or a species complex[64]. Therefore, it would be very interesting to test whether genome size varies between populations of different reproductive modes (sexual or apomictic), and how much of that variation is caused by TEs. Because genome reduction with TE purging may happen within a few generations in angiosperms[65], such comparative analyses could be complemented by experimental evolution studies varying the selfing rates in the fast-cycling fern model *C. richardii* to further interrogate the relationship between genome size and TEs. Techniques that have been successfully applied to *Arabidopsis thaliana* could also be adapted for ferns and lycophytes to track TE activity in real time[66].

Bryophytes. In contrast to ferns and lycophytes, there seems to be no correlation between genome size and chromosome number in mosses and potentially across all bryophytes[67]. Whereas genome sizes vary between 122–719 Mbp (mode, 176 Mbp; median, 205 Mbp)[6,68] and 206 Mbp–20 Gbp (mode, 740 Mbp; median, 751 Mbp)[67], respectively for hornworts and liverworts, both lineages have relatively constant chromosome numbers: $n = 4$–6 in hornworts and $n = 8$–9 in liverworts[69]. Contrastingly, mosses exhibit a much larger variation in chromosome numbers ($n = 6$–38)[70], but their genome sizes are relatively stable (minimum, 170 Mbp; maximum, 2 Gbp; mode, 442 Mbp; median, 433 Mbp)[71]. Phylogenetic analyses in liverworts and mosses suggest that genome size evolution is not a one-way process and that genome size increase and decrease both occurred along the phylogeny[54,67]. Analysis of hornworts suggests a different pattern, with a gradually increasing genome size across the phylogeny[68].

It is unclear how repeat elements (especially LTRs) contributed to genome size differences in bryophytes, because very few reliable estimates of repeat content are available. WGDs are frequent in mosses but rare in liverworts and absent in investigated hornwort genomes[4,6,7,10,15,39]. It is therefore possible that increase in genome size is driven mainly by repeat expansions in hornworts and liverworts, but less so in mosses. Overall repeat content varies considerably among moss genomes but it does not seem to correlate with genome size differences[15–17,22]. By contrast, in liverworts, some early cytological studies revealed massive tracks of heterochromatic regions, which may contribute to genome size expansions[72]. In hornworts, LTR content can partially explain the genome size difference, at least among three closely related strains that were analysed[6]. We can speculate that hornwort and liverwort genomes may follow a unique evolutionary trajectory in which WGD is either rare or absent and genome size variation is driven primarily by LTRs. In contrast, WGDs and post-WGD fractionations may have a larger role in genome size evolution in moss.

This overview shows that current data are insufficient to disentangle the dominant mechanisms contributing to genome size variation in bryophytes. We have identified several research areas that should be pursued. First, while it is often stated that genome sizes and chromosome numbers are uncorrelated in bryophytes[73], the evidence supporting this statement is not very strong. This is because genome sizes and chromosome numbers can substantially vary among geographically distant accessions of the very same species[54,67]. Studies are needed to measure genome size and chromosome number in the same individuals to carefully test their correlation in all three groups of bryophytes. The measurements are further complicated by the need to account for the frequent occurrence of endopolyploidy in mosses[54]. Second, to better understand the role of WGDs in bryophyte genome size evolution, phylogenetic comparative analysis of a properly assembled dataset incorporating chromosome counts, genome sizes and WGD events is necessary. Similarly, more information is needed on the contribution of TEs, which should be investigated in groups with highly divergent genome sizes but relatively stable chromosome numbers. Genome-skimming approaches[62] could be a relatively inexpensive way to gather repeat data from across a wide range of species. To this end, the pleurocarpous mosses, several groups of thalloid and leafy liverworts and hornworts would provide appropriate systems.

Finally, the processes that have kept moss and most bryophyte genomes relatively small remain unclear. Recent evidence indicates that the constraints imposed by sperm size and/or high frequency of homologous recombination[74,75] are insufficient to explain the relatively small genome sizes in mosses and potentially in most bryophytes[76,77]. Conversely, both theory and experimental evidence suggest that asexual reproduction and selfing can lead to decreased genome sizes and rapid loss of TEs[65,78,79]. Indeed, frequent selfing and/or asexual reproduction do occur in all three groups of bryophytes and may thus prevent runaway genome expansions. Evidence is accumulating that some TEs may be active throughout the life cycle of the model moss *P. patens*[15,80]. Therefore, the interplay among TEs, breeding system and genome size could be tested in experimental evolutionary studies, which, together with a larger collection of high-quality genomes, should shed light on the factors contributing to bryophyte genome size evolution.

Overall chromosome structure

Organization of angiosperm chromosomes appears to be conserved at a large scale (Fig. 2). Typically, metacentric chromosomes are characterized by a gene-poor centromeric, pericentromeric and telomeric regions, with most genes located between the pericentromeric regions and the telomeres. Centromeric regions of angiosperms investigated so far are occupied by tandem satellite repeats with relatively long repeat units that are interspersed with TEs[81]. Telomeres are composed of shorter tandem repeats whose actual sequence may vary across taxa[82]. Finally, pericentromeric regions are usually enriched for TEs, especially retrotransposons[81] (Fig. 2).

Recent results suggest that some chromosomal properties of seed-free plants deviate from this pattern (Fig. 2). Analysis of

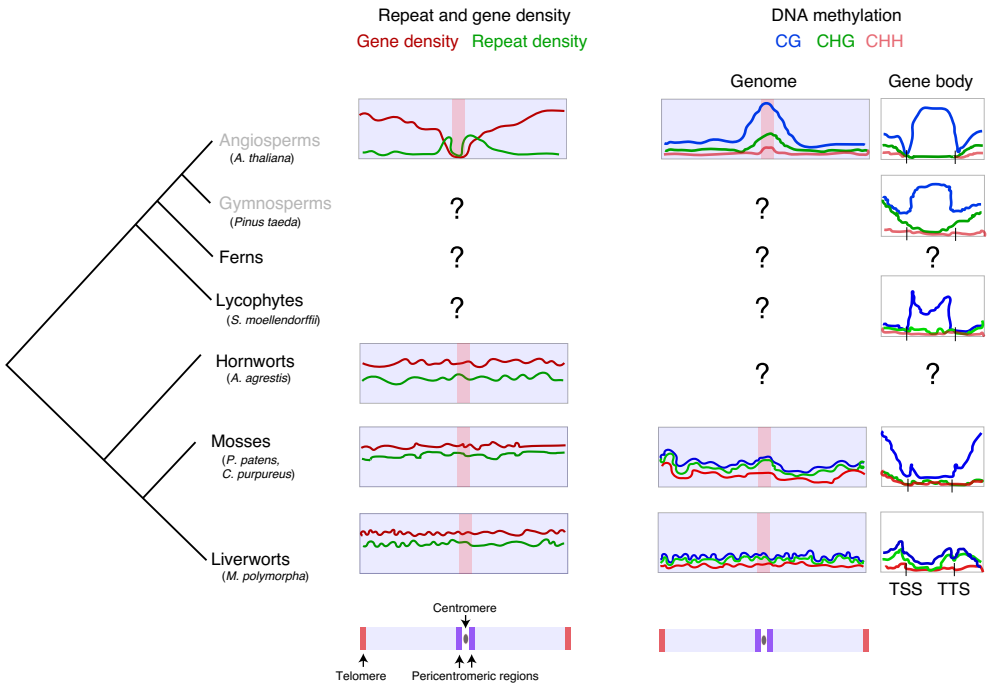

Fig. 2 | Distribution of gene/transposon density and DNA methylation. Distribution of gene/transposon density and DNA methylation (in the genome and in gene bodies) in seed-free (black) and seed plant (grey) genomes. Boxes represent idealized metacentric chromosomes, red rectangles refer to the centromeres (*x* axis, physical position; *y* axis, density or percent methylation). Idealized chromosomes are shown at the bottom. gbM is shown between the transcription start site (TSS) and transcription stop site (TTS) and along their 1-kbp flanking regions. Idealized plots are redrawn from refs. [12,100]. DNA methylation is presented in all three sequence contexts (CG, CHG and CHH). Latin names in brackets refer to the species for which data are available. Question marks indicate missing or ambiguous data points.

P. patens, *M. polymorpha* and *A. agrestis* genomes show that pericentromeric regions are not enriched for TEs[6,10,12,13,15]. Furthermore, in contrast to many angiosperm genomes, the TEs and genes are relatively evenly dispersed along the chromosomes, resulting in a fine-grained landscape of alternating gene- and repeat-rich islands. This lack of TE clustering might be an ancestral state of the Viridiplantae, because repeat elements of *Mesostigma* are also similarly distributed[28]. Alternatively, it could also be a lineage-specific trait that is shared by all bryophytes.

Very little information is available on large-scale chromosome structure in ferns and lycophytes[83]. Telomere structure has been investigated only for the lycophyte *Selaginella*, which were found to have similar telomere structure to *Arabidopsis*[84]. Putative centromeric regions of the water fern genomes were inferred, but the genomes were not assembled at the chromosomal level[20]. Consequently, whether pericentromeric regions with an elevated TE density are present in ferns and lycophytes is unclear. Nevertheless, data on the *Selaginella* genome suggest that genes and repeats seem to be evenly dispersed along the chromosomes, similar to the investigated bryophyte genomes[9,19]. Therefore, some evidence supports similar spatial arrangements of repeats and genes in the fern, lycophyte and bryophyte genomes (Fig. 2). If this is the case, it would indicate a radical difference in genome structure between seed plants and seed-free plants.

With the limited data currently available, we could only speculate about the biological processes behind such contrasting patterns. Because many genomic features (epigenetic state, nucleotide composition, recombination rate and higher-level three-dimensional (3D) organization) are significantly correlated with the density of both TEs and genes, their actual contributions are difficult to dissect. Frequent selfing has been proposed to have potentially influenced the distribution of transposons[15], but the outcrossing mating system of *M. polymorpha* makes this explanation unlikely[12,13,15]. Alternatively, it is possible that the unique distribution of TEs and genes is a consequence of idiosyncratic TE dynamics. For instance, in *Caenorhabditis inopinata*, a relative of the model nematode *Caenorhabditis elegans*, a similarly even distribution of genomic features was observed and can be attributed to the activities of a few specific TE classes[85]. Therefore, it would be necessary to learn more about the 'community ecology' of TEs in seed-free plants in a similar way as was described in Stitzer et al.[86] using available genome data. Such studies could be further complemented by investigating TEs in population genomic datasets and in mutant lines with compromised epigenetic machinery.

Another critical piece of information that is missing for seed-free plants is the process by which centromeric and pericentromeric regions evolve. In seed plants, centromere evolution begins by epigenetic reprogramming of the DNA and continues with the accumulation of retrotransposons, which contribute to the formation of tandem satellite repeats and stability of the centromere[81,87]. It is possible that centromere evolution in seed-free plants proceeds in a similar manner, although no experimental evidence is available. Comparative analyses of centromeres in closely related species and investigation of de novo evolved centromeres (neocentromeres) have provided insights into centromere biology in seed plants[88,89]. Because neocentromeres in seed-free plants are not well characterized, experiments in which chromosome fragments are created and the rapid evolution of neocentromeres is tracked may provide a viable strategy to investigate centromere evolution[90]. Furthermore, with the new sequencing technologies it will become possible to investigate the composition and dynamics of centromeres in a larger group of seed-free plants by selective capture and/or sequencing of centromeres[91–93]. Owing to their small genomes, the model bryophytes (*M. polymorpha*, *A. agrestis*

and *P. patens*) and closely related species would be ideal systems for such investigations.

Collinearity across seed-free plant genomes

Although plant genomes are highly dynamic, collinearity (the conserved order of genes on corresponding chromosomes) is detectable and can be used to reconstruct the ancestral genome structure and gene content[94]. Collinearity is substantial within core eudicots and grasses, but more limited between the two groups[95]. Collinearity is expected to decrease over time, leading to less genomic collinearity between more deeply diverged genomes.

Very little is known about collinearities among seed-free plants and between seed-free and seed plants. This is due in part to the sparse availability of high-quality genome assemblies for seed-free plants. The few available observations are somewhat contradictory. A study on the moss genome, *P. patens*, reported the presence of several hundred collinear blocks with some angiosperm genomes, which may have been conserved since the most recent common ancestor of land plants[15]. Genes in such collinear blocks tend to be co-expressed and preferentially contain genes related to stress and essential biological processes. Nevertheless, reanalysis of the available genomes from each lineage of bryophytes—the mosses, liverworts and hornworts—revealed that very few (if any) collinear blocks can be found that are shared by all three bryophyte clades and by most vascular plant lineages[6,12]. Therefore, the collinear blocks inherited from the common ancestor of land plants must have been largely broken up by rearrangements and/or fractionation following WGDs. Still, collinearity between each of the bryophyte lineages and vascular plants are present, suggesting that different and unique collinear blocks may have been retained between mosses and vascular plants, between liverworts and vascular plants, and between hornworts and vascular plants. It is currently unknown whether the limited collinearity among bryophyte lineages reflects functional significance, or is simply an artefact of the small number of genomes investigated. Limited collinearity may be the result of streamlined small genomes of the model species sequenced to date or a consequence of accelerated genome dynamics in bryophytes. Further high-quality genome assemblies are needed to resolve these issues and better characterize the dynamics of bryophyte genomes.

Information on collinearity among fern and lycophyte genomes is even more limited, mostly because none of the published genomes are resolved at the chromosomal scale[9,19,20,35]. Because fern chromosomes are small and their size is less variable than in angiosperms, it has been hypothesized that they are less dynamic[56,83] and might have retained more collinear blocks. The upcoming chromosome-scale assemblies of *C. richardii* and others will make it possible to learn more about collinearity in fern and lycophyte genomes.

DNA methylation

Chromosomes are also decorated with various modifications and proteins involved in regulating their transcription. The most frequently investigated epigenomic features are DNA methylation and histone modifications. Both have wide-ranging effects on the activity of genic and intergenic regions and on overall genome stability and dynamics. Methylation of cytosine in the fifth position (5mC) is involved in the silencing of TEs, condensation of DNA into heterochromatin and regulation of gene expression. DNA methylation is an ancient feature and has been found in green algae and land plants[96].

Methylated cytosines show a well-conserved distribution across angiosperm genomes[96,97] (Fig. 2). At the chromosomal scale, pericentromeric and centromeric regions are highly methylated, close to tenfold the level of the less methylated chromosome arms and telomeres[97]. Cytosines can be methylated in three major sequence contexts: CG, CHG and CHH (in which H corresponds to A, T or C). Repeat and TE sequences are highly methylated in all three sequence contexts, and the levels of methylation correlate well with TE activity. In genic regions, CG methylation is the lowest in the vicinity of the TSS[96,97]. TSS methylation is usually inversely correlated with the expression level of genes. In most angiosperm and gymnosperm taxa, gene bodies are also methylated mainly in the CG context, and modestly expressed genes are more methylated than those showing more extreme (very high or very low) expression levels[98,99] (Fig. 2).

Interestingly, methylome profiling in seed-free plants revealed a different pattern from seed plants, at least in some bryophytes and in the lycophyte *Selaginella*[12,96,100] (Fig. 2). Overall, cytosine methylation is at least fivefold lower in the moss, liverwort and lycophyte genomes compared with *Arabidopsis*[12,96,100]. Furthermore, TSSs are rarely methylated, and if they are, no significant correlation between expression level and extent of methylation is observed[12,13,15,101]. In contrast to the overall lower level of DNA methylation, gene body methylation (gbM) appears to be more variable in seed-free plants. In *Selaginella*, *M. polymorpha* and *P. patens*, most of the genes exhibit no sign of gbM[12,15,96,100,101], whereas gbM levels in some ferns and non-*Selaginella* lycophytes are similar to those in gymnosperms and angiosperms[100,102]. The presence of very few gbM genes in bryophytes was previously explained by the lack of *CMT3* clade genes, which is associated with the loss of gbM in flowering plants[100]. Although gbM genes are rare in the bryophyte genomes, they reportedly exist in mosses and liverworts. Gene Ontolgy analysis of gbM genes in *P. patens* and *M. polymorpha* suggests that gbM genes may be of functional importance, and some are preferentially methylated during sexual reproduction in *M. polymorpha* and *P. patens*[15,101,103]. Intriguingly, gbM genes appear to have contrasting characteristics in *M. polymorpha* and *P. patens*. In *M. polymorpha*, gbM genes are longer, contain more exons, and are more broadly expressed than non-methylated genes, a pattern that is somewhat similar to the one in flowering plants[12]. By contrast, gbM genes in *P. patens* show more tissue-specific expression, have lower GC content and are expressed at lower levels than non-methylated genes[15]. Very little is known about the function of gbM genes in ferns and other lycophytes. Nevertheless, orthologues of angiosperm gbM genes tend to also be methylated in ferns with similar structural features, suggesting conserved function among ferns, lycophytes and seed plants[102]. Without further information on the dynamic changes of gbM in various seed-free plant genomes, its functional significance is difficult to evaluate[101,103]. Furthermore, it is currently not well known how gbM is established in plants without *CMT3* clade genes[104].

The limited information outlined above suggests that in most seed-free plants, the overall level of DNA methylation appears to be lower and gbM appears to be less common than in seed plants. Nevertheless, there are exceptions to this observation and patterns of DNA methylation in seed-free plants may be more diverse. Therefore, it is imperative to collect more data on DNA methylome variation in seed-free plants to be able to make generalizations. Although whole-genome methylome sequencing requires a reference genome, information on genic methylation could also be obtained using transcriptomes as the reference[102]. With this latter approach, information on gbM could be gained in a more cost-effective manner for a wide range of taxa. Another even more poorly understood topic is whether the effect of DNA methylation on genome evolution differs between seed and seed-free plants. In seed plants, DNA methylation has important roles in silencing mobile elements, increasing mutation rates (mutagenic factor), affecting the distribution of recombination hot spots, influencing the retention of gene duplicates, contributing to phenotypic plasticity and epigenetic inheritance, and gene expression[105–111]. However, very little is known about these processes in seed-free plants, and this should be a priority for future research[112].

3D genome structure

During the past decade, important advancements have been made to characterize the 3D genome structure of flowering plants[113,114]. In flowering plants, genomes are spatially organized at various levels starting from the chromatin territories, in which telomeres and centromeres occupy different parts of the nucleus[113]. At the megabase scale, chromatin is organized into A and B compartments that preferentially interact with each other; these compartments are characterized by well-defined epigenetic states, and often correspond to heterochromatic, pericentromeric regions and euchromatic chromosome arms. Topologically associating domains (TADs) are 3D structural entities along the linear chromosome representing further genomic organization at a finer scale[115]. Chromosomal regions within a TAD exhibit higher contact frequencies than with regions outside of TADs, and are delineated by their transcriptional and epigenetic state[113]. Finally, chromosomal loops represent the smallest unit that has well-proven functional significance. Various studies have found that 3D interaction frequency is a good indicator of transcriptional activity and the epigenetic state of genes, and can be heavily modulated by abiotic factors[116].

The 3D genome structure and its functional significance in seed-free plants are not well known. Because spatial distribution of genes and TEs in seed-free plants differ from those of seed plants, it is of particular importance to know how the genome structures compare between the two groups. Information on the 3D structure of seed-free genomes is available only for the liverwort *M. polymorpha*, which has both shared and distinct features compared with those of flowering plants[13]. *M. polymorpha* telomeres have been described as clustering together at interphase, similar to the bouquet structure detected in some flowering plants[113]. Furthermore, borders of TAD-like domains in *M. polymorpha* have been associated with active gene expression and have served as units of expression regulation, features also shared with other flowering plant genomes[117]. Conversely, the *M. polymorpha* genome contains a distinct type of TAD that has not been observed in other plants or animals[117]. These TADs are enriched with TCP1 transcription factors, and the complex of TCP1 and TAD collectively repressed gene expression. It is unknown whether such transcription-factor-enriched TADs exist in other land plants or whether they are unique to *M. polymorpha*.

Another unique feature of the *M. polymorpha* genome is the presence of strong intra- and interchromosomal interactions. Regions with such interactions are depleted in heterochromatic (H3K27me1) and euchromatic (H3K4me3 and H3K36me3) histone marks, but enriched in DNA methylation similar to KNOT regions of *Arabidopsis*[118]. The *Arabidopsis* KNOT is a 3D nuclear structure that is involved in defence against invasive DNA elements independent of methylation and epigenetic silencing via small RNAs. In addition, H3K27me3 is strongly associated with heterochromatic domains in *M. polymorpha*; this is in contrast to flowering plants, in which such domains are marked by H3K9me1 and H3K27me1 (ref. [13]). Therefore, H3K27me3 is likely to be important in forming heterochromatic domains and repressing TEs in *M. polymorpha*, a feature shared with some ciliates but not with flowering plants.

While 3D genome structure is rarely considered as a direct factor affecting genome evolution, it is clear that structures such as TADs are not only spatial features, but are also regulatory units that demarcate the range of enhancer activity, control enhancer–gene interactions, synchronize replication timing and coordinate correlated gene expression[119,120]. Therefore, changes in TADs and probably other 3D structures of the genome will lead to regulatory evolution and ultimately new phenotypes[121,122]. Furthermore, proper 3D structure of the genome is required for genomic stability and can potentially impose constraints on genome evolution[123]. Nevertheless, it is not well known how and to what extent 3D genome structure contributes to genome evolution[124]. We believe that comparative analyses of 3D genome structures in seed and seed-free plants will prove particularly useful for gaining deeper insights into this question, especially given their seemingly contrasting distributions of TEs and genes.

Linking genes to phenotypes

Next, having evaluated the emerging questions about the genomes of seed-free plants, we turn to the progress in the development of functional genetic tools. Although comparative genomics is a powerful tool for formulating hypotheses concerning the evolution and diversification of seed-free plants, linking genotypes to phenotypes requires targeted investigation of gene function. While reverse-genetic tools are available for an increasing number of seed plant model systems, seed-free plants remain underdeveloped in this respect. Below, we provide a brief overview of the currently available and aspiring seed-free model systems, all of which are amenable to in vitro propagation and can provide biological data (that is, biochemical, physiological and genetics) of interest. We also highlight the problems and challenges faced in using these model systems.

Ferns. While several model species are available for ferns, few are amenable to transformation and gene silencing (Table 1). The homosporous *C. richardii* is the most popular model fern, and has a short life cycle in the laboratory, well-established reverse-genetic tools[125,126] and a representative genome size of 14 Gbp[35]. *C. richardii* has been used extensively to elucidate sporophyte and gametophyte development, cell division and sex determination in ferns[127]. Besides *C. richardii*, the homosporous *A. capillus-veneris* has been studied extensively for its unique photobiology[128], *P. vittata* for its phytoremediation potential[129], *M. vestita* for its intron retention and splicing during spermatogenesis[130], and *Azolla* for its nitrogen-fixing symbiosis[20]. The application of genome-editing tools in these model ferns have not been reported yet, but their use is highly anticipated.

Lycophytes. The absence of a reliable genetic transformation method poses a major challenge towards gene functional analysis in lycophytes (Table 1). Notably, in vitro propagation methods are available for several Lycopodiaceae species, although the length of time needed to complete their life cycle (spanning months to years for spore germination and sporophyte development) presents a major hurdle[131]. Among Selaginellales, *S. apoda* is a promising model species due to its short life cycle[132], which can be completed within 85 days.

Mosses. *P. patens* and *C. purpureus* are the two most developed model moss species (Table 1), and *P. patens* is also the most prominent bryophyte and seed-free plant model system overall, owing to the existence of effective methods for in vitro propagation, genetic transformation and high gene-targeting efficiency that is on par with that of the budding yeast[133]. Other model mosses under development (Table 1) include *F. hygrometrica*[21], the desiccation-tolerant *S. ruralis*[134] and *S. caninervis*[135,136], as well as the agriculturally and economically important *Sphagnum* spp.[137].

Liverworts. *M. polymorpha* is the only well-established model species for liverworts (Table 1). Although not as efficient as in *P. patens*, gene targeting in *M. polymorpha* has been found to be more efficient (approximately 2%) than in most land plants[138]. *M. polymorpha* and *P. patens* have recently begun to revolutionize our understanding of plant evolutionary development[139]. *M. paleacea*, a sister species to *M. polymorpha*, is also being developed as a model for studying arbuscular mycorrhizal symbiosis in bryophytes[14]. In vitro gametophyte propagation of several other liverwort species besides *Marchantia* spp. have been documented[140]. *Riccia* spp. is a promising model liverwort, and is the focus of several studies on

the effects of environmental cues on gametophyte development[141]. Altogether, diversified development of genetic transformation and genome-editing tools for species other than *M. polymorpha* are still needed to better represent the diversity of liverworts.

Hornworts. Functional genetic tools are under development for hornworts in the genus *Anthoceros*, while other hornwort clades lag behind (Table 1). A reliable *Agrobacterium*-mediated genetic transformation method was recently reported for *A. agrestis*[142]. *A. punctatus* has been especially useful as a model for the characterization of plant symbiosis with nitrogen-fixing cyanobacteria[6]. While almost all known hornwort species are able to support such symbiosis, the presence of pyrenoids (for their carbon-concentrating mechanism), sex chromosomes, stomata and arbuscular mycorrhizal symbiosis varies between species[143]. Efforts are underway to develop in vitro gametophyte cultures of representative species from all five hornwort families, paving the way towards elucidating the genetic basis of the unique traits in hornworts.

Conclusions and perspectives

During the past ten years, the availability of new sequencing technologies have contributed to an increasing number of plant genomes. While the majority of these genomes are from seed plants, the gaps in seed-free plants are being filled. Despite these efforts, genomic information for seed-free plants is far from satisfactory. In particular, genome sequences for some major clades are missing (for example, Isoetales and Lycopodiales), and most large and diverse clades are represented by a single reference genome, few of which have been assembled at the chromosomal level (Fig. 1). To gain a better understanding of genome evolution across the plant tree of life, more high-quality and phylodiverse reference genomes are needed. Future work should also consider pan-genomic approaches to capture the genomic diversity at shallower phylogenetic scales[144].

Studies to date have suggested that seed-free plant and seed plant genomes may differ in various aspects. Nevertheless, given the gaps in the genome information available for seed-free plants, the generality of these findings remains to be tested. Key questions including the evolution of collinearity, genome size, gene content, WGD, overall chromosome structure, 3D genome conformation, epigenetics and diverse aspects of gene regulation in seed-free plants need to be readdressed when more data are available.

While comparative genomic information is necessary to put forward evolutionary hypotheses, functional verification can only be achieved when amenable model systems are available. Much remains to be done in this respect for seed-free plants. Model systems with a proper reverse-genetic toolbox need to be developed for various groups of seed-free plants together with genomic resources (Table 1). We are confident that further investigations into seed-free plants will not only help to address classical evolutionary questions, but also lead to new discoveries, some of them may be of applied importance.

Received: 10 August 2020; Accepted: 25 February 2021;
Published online: 5 April 2021

References

1. Wickett, N. J. et al. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc. Natl Acad. Sci. USA* **111**, E4859–E4868 (2014).
2. Puttick, M. N. et al. The interrelationships of land plants and the nature of the ancestral embryophyte. *Curr. Biol.* **28**, 733–745 (2018).
3. de Sousa, F., Foster, P. G., Donoghue, P. C. J., Schneider, H. & Cox, C. J. Nuclear protein phylogenies support the monophyly of the three bryophyte groups (Bryophyta Schimp.). *New Phytol.* **222**, 565–575 (2019).
4. One Thousand Plant Transcriptomes Initiative. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**, 679–685 (2019).
5. Harris, B. J., Harrison, C. J., Hetherington, A. M. & Williams, T. A. Phylogenomic evidence for the monophyly of bryophytes and the reductive evolution of stomata. *Curr. Biol.* **30**, 2001–2012 (2020).
6. Li, F.-W. et al. *Anthoceros* genomes illuminate the origin of land plants and the unique biology of hornworts. *Nat. Plants* **6**, 259–272 (2020).
7. Zhang, J. et al. The hornwort genome and early land plant evolution. *Nat. Plants* **6**, 107–118 (2020).
8. Rensing, S. A. et al. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* **319**, 64–69 (2008).
9. Banks, J. A. et al. The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* **332**, 960–963 (2011).
10. Bowman, J. L. et al. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* **171**, 287–304 (2017).
11. Marks, R. A., Smith, J. J., Cronk, Q., Grassa, C. J. & McLetchie, D. N. Genome of the tropical plant *Marchantia inflexa*: implications for sex chromosome evolution and dehydration tolerance. *Sci. Rep.* **9**, 8722 (2019).
12. Diop, S. I. et al. A pseudomolecule-scale genome assembly of the liverwort *Marchantia polymorpha*. *Plant J.* **101**, 1378–1396 (2020).
13. Montgomery, S. A. et al. Chromatin organization in early land plants reveals an ancestral association between H3K27me3, transposons, and constitutive heterochromatin. *Curr. Biol.* **30**, 573–588 (2020).
14. Radhakrishnan, G. V. et al. An ancestral signalling pathway is conserved in intracellular symbioses-forming plant lineages. *Nat. Plants* **6**, 280–289 (2020).
15. Lang, D. et al. The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J.* **93**, 515–533 (2018).
16. Pederson, E. R. A., Warshan, D. & Rasmussen, U. Genome sequencing of *Pleurozium schreberi*: the assembled and annotated draft genome of a pleurocarpous feather moss. *G3* **9**, 2791–2797 (2019).
17. Yu, J. et al. Draft genome of the aquatic moss *Fontinalis antipyretica* (Fontinalaceae, Bryophyta). *Gigabyte* **1**, https://doi.org/10.46471/gigabyte.8 (2020).
18. Xu, Z. et al. Genome analysis of the ancient tracheophyte *Selaginella tamariscina* reveals evolutionary features relevant to the acquisition of desiccation tolerance. *Mol. Plant* **11**, 983–994 (2018).
19. VanBuren, R. et al. Extreme haplotype variation in the desiccation-tolerant clubmoss *Selaginella lepidophylla*. *Nat. Commun.* **9**, 13 (2018).
20. Li, F.-W. et al. Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nat. Plants* **4**, 460–472 (2018).
21. Kirbis, A. et al. Transcriptional landscapes of divergent sporophyte development in two mosses, *Physcomitrium* (*Physcomitrella*) *patens* and *Funaria hygrometrica*. *Front. Plant Sci.* **11**, 747 (2020).
22. Carey, S. B. et al. Chromosome fusions shape an ancient UV sex chromosome system. Preprint at *bioRxiv* https://doi.org/10.1101/2020.07.03.163634 (2020).
23. Hori, K. et al. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nat. Commun.* **5**, 3978 (2014).
24. Nishiyama, T. et al. The *Chara* genome: secondary complexity and implications for plant terrestrialization. *Cell* **174**, 448–464 (2018).
25. Cheng, S. et al. Genomes of subaerial Zygnematophyceae provide insights into land plant evolution. *Cell* **179**, 1057–1067 (2019).
26. Jiao, C. et al. The *Penium margaritaceum* genome: hallmarks of the origins of land plants. *Cell* **181**, 1097–1111 (2020).
27. Wang, S. et al. Genomes of early-diverging streptophyte algae shed light on plant terrestrialization. *Nat. Plants* **6**, 95–106 (2020).
28. Liang, Z. et al. *Mesostigma viride* genome and transcriptome provide insights into the origin and evolution of Streptophyta. *Adv. Sci.* **7**, 1901850 (2020).
29. Morris, J. L. et al. The timescale of early land plant evolution. *Proc. Natl Acad. Sci. USA* **115**, E2274–E2283 (2018).
30. Larsén, E. & Rydin, C. Disentangling the phylogeny of *Isoetes* (Isoetales), using nuclear and plastid data. *Int. J. Plant Sci.* **177**, 157–174 (2016).
31. Keeley, J. E. Aquatic CAM photosynthesis: a brief history of its discovery. *Aquat. Bot.* **118**, 38–44 (2014).
32. Pellicer, J. & Leitch, I. J. The plant DNA C-values database (release 7.1): an updated online repository of plant genome size data for comparative studies. *New Phytol.* **226**, 301–305 (2020).
33. Klekowski, E. J. & Baker, H. G. Evolutionary significance of polyploidy in the Pteridophyta. *Science* **153**, 305–307 (1966).
34. Kuo, L.-Y. & Li, F.-W. A roadmap for fern genome sequencing. *Am. Fern. J.* **109**, 212–223 (2019).
35. Marchant, D. B. et al. The C-Fern (*Ceratopteris richardii*) genome: insights into plant genome evolution with the first partial homosporous fern genome assembly. *Sci. Rep.* **9**, 18181 (2019).
36. Landis, J. B. et al. Impact of whole-genome duplication events on diversification rates in angiosperms. *Am. J. Bot.* **105**, 348–363 (2018).
37. Soltis, P. S. & Soltis, D. E. Ancient WGD events as drivers of key innovations in angiosperms. *Curr. Opin. Plant Biol.* **30**, 159–165 (2016).

38. Vanneste, K., Maere, S. & Van de Peer, Y. Tangled up in two: a burst of genome duplications at the end of the Cretaceous and the consequences for plant evolution. *Philos. Trans. R. Soc. Lond. B* **369**, 20130353 (2014).
39. Gao, B. et al. Integrated phylogenomic analyses reveal recurrent ancestral large-scale duplication events in mosses. Preprint at *bioRvix* https://doi.org/10.1101/603191 (2019).
40. Huang, C.-H., Qi, X., Chen, D., Qi, J. & Ma, H. Recurrent genome duplication events likely contributed to both the ancient and recent rise of ferns. *J. Integr. Plant Biol.* **62**, 433–455 (2020).
41. Villarreal, J. C. & Renner, S. S. Hornwort pyrenoids, carbon-concentrating structures, evolved and were lost at least five times during the last 100 million years. *Proc. Natl Acad. Sci. USA* **109**, 18873–18878 (2012).
42. Devos, N. et al. Analyses of transcriptome sequences reveal multiple ancient large-scale duplication events in the ancestor of Sphagnopsida (Bryophyta). *New Phytol.* **211**, 300–318 (2016).
43. Johnson, M. G., Malley, C., Goffinet, B., Shaw, A. J. & Wickett, N. J. A phylotranscriptomic analysis of gene family expansion and evolution in the largest order of pleurocarpous mosses (Hypnales, Bryophyta). *Mol. Phylogenet. Evol.* **98**, 29–40 (2016).
44. Zhang, R. et al. Dating whole genome duplication in *Ceratopteris thalictroides* and potential adaptive values of retained gene duplicates. *Int. J. Mol. Sci.* **20**, 1926 (2019).
45. Vanneste, K., Sterck, L., Myburg, A. A., Van de Peer, Y. & Mizrachi, E. Horsetails are ancient polyploids: evidence from *Equisetum giganteum*. *Plant Cell* **27**, 1567–1578 (2015).
46. Clark, J. W., Puttick, M. N. & Donoghue, P. C. J. Origin of horsetails and the role of whole-genome duplication in plant macroevolution. *Proc. R. Soc. B* **286**, 20191662 (2019).
47. Ren, R. et al. Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. *Mol. Plant* **11**, 414–428 (2018).
48. Comai, L. The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* **6**, 836–846 (2005).
49. Edger, P. P. et al. The butterfly plant arms-race escalated by gene and genome duplications. *Proc. Natl Acad. Sci. USA* **112**, 8362–8366 (2015).
50. Carretero-Paulet, L. & Fares, M. A. Evolutionary dynamics and functional specialization of plant paralogs formed by whole and small-scale genome duplications. *Mol. Biol. Evol.* **29**, 3541–3551 (2012).
51. Carretero-Paulet, L. & Van de Peer, Y. The evolutionary conundrum of whole-genome duplication. *Am. J. Bot.* **107**, 1101–1105 (2020).
52. Walden, N. et al. Nested whole-genome duplications coincide with diversification and high morphological disparity in Brassicaceae. *Nat. Commun.* **11**, 3795 (2020).
53. Pellicer, J., Hidalgo, O., Dodsworth, S. & Leitch, I. Genome size diversity and its impact on the evolution of land plants. *Genes* **9**, 88 (2018).
54. Bainard, J. D., Newmaster, S. G. & Budke, J. M. Genome size and endopolyploidy evolution across the moss phylogeny. *Ann. Bot.* **125**, 543–555 (2020).
55. Clark, J. et al. Genome evolution of ferns: evidence for relative stasis of genome size across the fern phylogeny. *New Phytol.* **210**, 1072–1082 (2016).
56. Barker, M. S. & Wolf, P. G. Unfurling fern biology in the genomics age. *Bioscience* **60**, 177–185 (2010).
57. Wagner, W. H. Jr & Wagner, F. S. Polyploidy in pteridophytes. *Basic Life Sci.* **13**, 199–214 (1979).
58. Nakazato, T., Barker, M. S., Rieseberg, L. H. & Gastony, G. J. in *Biology and Evolution of Ferns and Lycophytes* (eds Ranker, T. A. & Haufler, C. H.) 175–198 (Cambridge Univ. Press, 2008).
59. Liu, H. et al. Polyploidy does not control all: lineage-specific average chromosome length constrains genome size evolution in ferns. *J. Syst. Evol.* **57**, 418–430 (2019).
60. Schubert, I. & Oud, J. L. There is an upper limit of chromosome size for normal development of an organism. *Cell* **88**, 515–520 (1997).
61. Wendel, J. F., Jackson, S. A., Meyers, B. C. & Wing, R. A. Evolution of plant genome architecture. *Genome Biol.* **17**, 37 (2016).
62. Wolf, P. G. et al. An exploration into fern genome space. *Genome Biol. Evol.* **7**, 2533–2544 (2015).
63. Baniaga, A. E. & Barker, M. S. Nuclear genome size is positively correlated with median LTR-RT insertion time in fern and Lycophyte genomes. *Am. Fern J.* **109**, 248 (2019).
64. Grusz, A. L. A current perspective on apomixis in ferns. *J. Syst. Evol.* **54**, 656–665 (2016).
65. Roessler, K. et al. The genome-wide dynamics of purging during selfing in maize. *Nat. Plants* **5**, 980–990 (2019).
66. Lanciano, S. et al. Sequencing the extrachromosomal circular mobilome reveals retrotransposon activity in plants. *PLoS Genet.* **13**, e1006630 (2017).
67. Bainard, J. D., Forrest, L. L., Goffinet, B. & Newmaster, S. G. Nuclear DNA content variation and evolution in liverworts. *Mol. Phylogenet. Evol.* **68**, 619–627 (2013).
68. Bainard, J. D. & Villarreal, J. C. Genome size increases in recently diverged hornwort clades. *Genome* **56**, 431–435 (2013).
69. Berrie, G. K. Cytology and phylogeny of liverworts. *Evolution* **17**, 347–357 (1963).
70. Wyatt, R. & Stoneburner, A. in *Plant Biosystematics* (ed. Grant, W. F.) 519–542 (Academic Press, 1984).
71. Voglmayr, H. Nuclear DNA amounts in mosses (Musci). *Ann. Bot.* **85**, 531–546 (2000).
72. Smith, A. J. E. in *Advances in Botanical Research* Vol. 6 (ed. Woolhouse, H. W.) 195–276 (Academic Press, 1979).
73. Cove, D. in *Bryophyte Biology* (eds. Shaw, A. J. & Goffinet, B.) 182–198 (Cambridge Univ. Press, 2000).
74. Reski, R. Development, genetics and molecular biology of mosses. *Bot. Acta* **111**, 1–15 (1998).
75. Schween, G., Gorr, G., Hohe, A. & Reski, R. Unique tissue-specific cell cycle in *Physcomitrella*. *Plant Biol.* **5**, 50–58 (2003).
76. Wiedemann, G. et al. RecQ helicases function in development, DNA repair, and gene targeting in *Physcomitrella patens*. *Plant Cell* **30**, 717–736 (2018).
77. Trouiller, B., Charlot, F., Choinard, S., Schaefer, D. G. & Nogué, F. Comparison of gene targeting efficiencies in two mosses suggests that it is a conserved feature of bryophyte transformation. *Biotechnol. Lett.* **29**, 1591–1598 (2007).
78. Bast, J., Jaron, K. S., Schuseil, D., Roze, D. & Schwander, T. Asexual reproduction reduces transposable element load in experimental yeast populations. *eLife* **8**, e48548 (2019).
79. Bast, J. et al. Consequences of asexuality in natural populations: Insights from stick insects. *Mol. Biol. Evol.* **35**, 1668–1677 (2018).
80. Vendrell-Mir, P., López-Obando, M., Nogué, F. & Casacuberta, J. M. Different families of retrotransposons and DNA transposons are actively transcribed and may have transposed recently in *Physcomitrium* (*Physcomitrella*) *patens*. *Front. Plant Sci.* **11**, 1274 (2020).
81. Oliveira, L. C. & Torres, G. A. Plant centromeres: genetics, epigenetics and evolution. *Mol. Biol. Rep.* **45**, 1491–1497 (2018).
82. Peska, V. & Garcia, S. Origin, diversity, and evolution of telomere sequences in plants. *Front. Plant Sci.* **11**, 117 (2020).
83. Sessa, E. B. et al. Between two fern genomes. *Gigascience* **3**, 15 (2014).
84. Shakirov, E. V. & Shippen, D. E. *Selaginella moellendorffii* telomeres: conserved and unique features in an ancient land plant lineage. *Front. Plant Sci.* **3**, 161 (2012).
85. Woodruff, G. C. & Teterina, A. A. Degradation of the repetitive genomic landscape in a close relative of *Caenorhabditis elegans*. *Mol. Biol. Evol.* **37**, 2549–2567 (2020).
86. Stitzer, M. C., Anderson, S. N., Springer, N. M. & Ross-Ibarra, J. The genomic ecosystem of transposable elements in maize. Preprint at *bioRxiv* https://doi.org/10.1101/559922 (2019).
87. Ma, J., Wing, R. A., Bennetzen, J. L. & Jackson, S. A. Plant centromere organization: A dynamic structure with conserved functions. *Trends Genet.* **23**, 134–139 (2007).
88. Talbert, P. B. & Henikoff, S. What makes a centromere? *Exp. Cell Res.* **389**, 111895 (2020).
89. Zhao, H. et al. Recurrent establishment of de novo centromeres in the pericentromeric region of maize chromosome 3. *Chromosome Res.* **25**, 299–311 (2017).
90. Liu, Y. et al. Rapid birth or death of centromeres on fragmented chromosomes in maize. *Plant Cell* **32**, 3113–3123 (2020).
91. Miga, K. H. et al. Telomere-to-telomere assembly of a complete human X chromosome. *Nature* **585**, 79–84 (2020).
92. Payne, A. et al. Readfish enables targeted nanopore sequencing of gigabase-sized genomes. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-020-00746-x (2020).
93. Kovaka, S., Fan, Y., Ni, B., Timp, W. & Schatz, M. C. Targeted nanopore sequencing by real-time mapping of raw electrical signal with UNCALLED. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-020-0731-9 (2020).
94. Pont, C. et al. Paleogenomics: reconstruction of plant evolutionary trajectories from modern and ancient DNA. *Genome Biol.* **20**, 29 (2019).
95. Zhao, T. & Schranz, M. E. Network-based microsynteny analysis identifies major differences and genomic outliers in mammalian and angiosperm genomes. *Proc. Natl Acad. Sci. USA* **116**, 2165–2174 (2019).
96. Zemach, A., McDaniel, I. E., Silva, P. & Zilberman, D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**, 916–919 (2010).
97. Schmitz, R. J., Lewis, Z. A. & Goll, M. G. DNA methylation: Shared and divergent features across eukaryotes. *Trends Genet.* **35**, 818–827 (2019).
98. Bewick, A. J. et al. On the origin and evolutionary consequences of gene body DNA methylation. *Proc. Natl Acad. Sci. USA* **113**, 9111–9116 (2016).
99. Takuno, S. & Gaut, B. S. Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *Proc. Natl Acad. Sci. USA* **110**, 1797–1802 (2013).

100. Bewick, A. J. et al. The evolution of CHROMOMETHYLASES and gene body DNA methylation in plants. *Genome Biol.* **18**, 65 (2017).
101. Schmid, M. W. et al. Extensive epigenetic reprogramming during the life cycle of *Marchantia polymorpha*. *Genome Biol.* **19**, 9 (2018).
102. Takuno, S., Ran, J.-H. & Gaut, B. S. Evolutionary patterns of genic DNA methylation vary across land plants. *Nat. Plants* **2**, 15222 (2016).
103. Meyberg, R., Perroud, P. F., Haas, F. B. & Schneider, L. Characterization of evolutionarily conserved key players affecting eukaryotic flagellar motility and fertility using a moss model. *New Phytol.* **227**, 440–454 (2020).
104. Yaari, R. et al. RdDM-independent de novo and heterochromatin DNA methylation by plant CMT and DNMT3 orthologs. *Nat. Commun.* **10**, 1613 (2019).
105. Choi, J., Lyons, D. B., Kim, M. Y., Moore, J. D. & Zilberman, D. DNA methylation and histone H1 jointly repress transposable elements and aberrant intragenic transcripts. *Mol. Cell* **77**, 310–323 (2020).
106. Shi, T. et al. Distinct expression and methylation patterns for genes with different fates following a single whole-genome duplication in flowering plants. *Mol. Biol. Evol.* **37**, 2394–2413 (2020).
107. Hämälä, T. & Tiffin, P. Biased gene conversion constrains adaptation in *Arabidopsis thaliana*. *Genetics* **215**, 831–846 (2020).
108. Eriksson, M. C., Szukala, A., Tian, B. & Paun, O. Current research frontiers in plant epigenetics: an introduction to a virtual issue. *New Phytol.* **226**, 285–288 (2020).
109. Alonso, C., Ramos-Cruz, D. & Becker, C. The role of plant epigenetics in biotic interactions. *New Phytol.* **221**, 731–737 (2019).
110. Mirouze, M. et al. Loss of DNA methylation affects the recombination landscape in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **109**, 5880–5885 (2012).
111. Li, N. et al. DNA methylation repatterning accompanying hybridization, whole genome doubling and homoeolog exchange in nascent segmental rice allotetraploids. *New Phytol.* **223**, 979–992 (2019).
112. Domb, K. et al. DNA methylation mutants in *Physcomitrella patens* elucidate individual roles of CG and non-CG methylation in genome regulation. *Proc. Natl Acad. Sci. USA* **117**, 33700–33710 (2020).
113. Grob, S. Three-dimensional chromosome organization in flowering plants. *Brief. Funct. Genomics* **19**, 83–91 (2020).
114. Sotelo-Silveira, M., Chávez Montes, R. A., Sotelo-Silveira, J. R., Marsch-Martínez, N. & de Folter, S. Entering the next dimension: plant genomes in 3D. *Trends Plant Sci.* **23**, 598–612 (2018).
115. Bonev, B. & Cavalli, G. Organization and function of the 3D genome. *Nat. Rev. Genet.* **17**, 661–678 (2016).
116. Doğan, E. S. & Liu, C. Three-dimensional chromatin packing and positioning of plant genomes. *Nat. Plants* **4**, 521–529 (2018).
117. Karaaslan, E. S. et al. *Marchantia* TCP transcription factor activity correlates with three-dimensional chromatin structure. *Nat. Plants* **6**, 1250–1261 (2020).
118. Grob, S., Schmid, M. W. & Grossniklaus, U. Hi-C analysis in *Arabidopsis* identifies the KNOT, a structure with similarities to the flamenco locus of *Drosophila*. *Mol. Cell* **55**, 678–693 (2014).
119. Nützmann, H.-W. et al. Active and repressed biosynthetic gene clusters have spatially distinct chromosome states. *Proc. Natl Acad. Sci. USA* **117**, 13800–13809 (2020).
120. Xie, T. et al. Biased gene retention during diploidization in *Brassica* linked to three-dimensional genome organization. *Nat. Plants* **5**, 822–832 (2019).
121. Picart-Picolo, A., Grob, S., Picault, N. & Franek, M. Large tandem duplications affect gene expression, 3D organization, and plant–pathogen response. *Genome Res.* **30**, 1583–1592 (2020).
122. Choudhary, M. N. K. et al. Co-opted transposons help perpetuate conserved higher-order chromosomal structures. *Genome Biol.* **21**, 16 (2020).
123. Lee, Y. C. G. et al. Pericentromeric heterochromatin is hierarchically organized and spatially contacts H3K9me2 islands in euchromatin. *PLoS Genet.* **16**, e1008673 (2020).
124. Eres, I. E. & Gilad, Y. A TAD skeptic: is 3D genome topology conserved? *Trends Genet.* **37**, 216–223 (2021).
125. Plackett, A. R. G., Huang, L., Sanders, H. L. & Langdale, J. A. High-efficiency stable transformation of the model fern species *Ceratopteris richardii* via microparticle bombardment. *Plant Physiol.* **165**, 3–14 (2014).
126. Bui, L. T., Cordle, A. R., Irish, E. E. & Cheng, C.-L. Transient and stable transformation of *Ceratopteris richardii* gametophytes. *BMC Res. Notes* **8**, 214 (2015).
127. Petlewski, A. R. & Li, F.-W. Ferns: the final frond-tier in plant model systems. *Am. Fern J.* **109**, 192–211 (2019).
128. Kanegae, T. & Kimura, I. A phytochrome/phototropin chimeric photoreceptor of fern functions as a blue/far-red light-dependent photoreceptor for phototropism in *Arabidopsis*. *Plant J.* **83**, 480–488 (2015).
129. Cai, C. et al. Three genes define a bacterial-like arsenic tolerance mechanism in the arsenic hyperaccumulating fern *Pteris vittata*. *Curr. Biol.* **29**, 1625–1633 (2019).
130. Boothby, T. C., Zipper, R. S., van der Weele, C. M. & Wolniak, S. M. Removal of retained introns regulates translation in the rapidly developing gametophyte of *Marsilea vestita*. *Dev. Cell* **24**, 517–529 (2013).
131. Whittier, D. P. & Storchova, H. The gametophyte of *Huperzia selago* in culture. *Am. Fern J.* **97**, 149–154 (2007).
132. Schulz, C. et al. An overview of the morphology, anatomy, and life cycle of a new model species: the Lycophyte *Selaginella apoda* (L.) Spring. *Int. J. Plant Sci.* **171**, 693–712 (2010).
133. Rensing, S. A., Goffinet, B., Meyberg, R., Wu, S.-Z. & Bezanilla, M. The moss *Physcomitrium* (*Physcomitrella*) *patens*: a model organism for non-seed plants. *Plant Cell* **32**, 1361–1376 (2020).
134. Oliver, M. J. et al. Chloroplast genome sequence of the moss *Tortula ruralis*: gene content, polymorphism, and structural arrangement relative to other green plant chloroplast genomes. *BMC Genomics* **11**, 143 (2010).
135. Silva, A. T. et al. To dry perchance to live: insights from the genome of the desiccation-tolerant biocrust moss Syntrichia caninervis. *Plant J.* https://doi.org/10.1111/tpj.15116 (2020).
136. Zhou, X. et al. Practices of biological soil crust rehabilitation in China: experiences and challenges. *Restor. Ecol.* **28**, S45–S55 (2020).
137. Heck, M. A. et al. Axenic in-vitro cultivation of nineteen peat-moss (*Sphagnum* L.) species as a resource for basic biology, biotechnology and paludiculture. *New Phytol.* **229**, 861–876 (2020).
138. Ishizaki, K., Johzuka-Hisatomi, Y., Ishida, S., Iida, S. & Kohchi, T. Homologous recombination-mediated gene targeting in the liverwort *Marchantia polymorpha* L. *Sci. Rep.* **3**, 1532 (2013).
139. Delaux, P.-M. et al. Reconstructing trait evolution in plant evo–devo studies. *Curr. Biol.* **29**, R1110–R1118 (2019).
140. Wurzel, G. & Becker, H. Growth and terpenoid production of an axenic culture from the liverwort *Ricciocarpos natans*. *Z. Naturforsch.* **45c**, 13–18 (1989).
141. Althoff, F. & Zachgo, S. Transformation of *Riccia fluitans*, an amphibious liverwort dynamically responding to environmental changes. *Int. J. Mol. Sci.* **21**, E5410 (2020).
142. Frangedakis, E. et al. A simple *Agrobacterium*-mediated stable transformation technique for the hornwort model *Anthoceros agrestis*. Preprint at *bioRxiv* https://doi.org/10.1101/2021.01.07.425778 (2021).
143. Frangedakis, E. et al. The hornworts: morphology, evolution and development. *New Phytol.* **229**, 735–754 (2020).
144. Bayer, P. E., Golicz, A. A., Scheben, A., Batley, J. & Edwards, D. Plant pan-genomes are the new reference. *Nat. Plants* **6**, 914–920 (2020).
145. Maridass, M., Mahesh, R., Raju, G. & Muthuchelian, K. Clonal propagation of *Adiantum capillus-veneris*. *Int. J. Biol. Technol.* **1**, 33–37 (2010).
146. Li, X., Han, J.-D., Fang, Y.-H., Bai, S.-N. & Rao, G.-Y. Expression analyses of embryogenesis-associated genes during somatic embryogenesis of *Adiantum capillus-veneris* L. in vitro: new insights into the evolution of reproductive organs in land plants. *Front. Plant Sci.* **8**, 658 (2017).
147. Tsuboi, H., Nakamura, S., Schäfer, E. & Wada, M. Red light-induced phytochrome relocation into the nucleus in *Adiantum capillus-veneris*. *Mol. Plant* **5**, 611–618 (2012).
148. Kawai-Toyooka, H. et al. DNA interference: a simple and efficient gene-silencing system for high-throughput functional analysis in the fern *Adiantum*. *Plant Cell Physiol.* **45**, 1648–1657 (2004).
149. Brouwer, P. et al. *Azolla* domestication towards a biobased economy? *New Phytol.* **202**, 1069–1082 (2014).
150. Banks, J. A. Gametophyte development in ferns. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 163–186 (1999).
151. Muthukumar, B., Joyce, B. L., Elless, M. P. & Neal Stewart, C. Stable transformation of ferns using spores as targets: *Pteris vittata* and *Ceratopteris thalictroides*. *Plant Physiol.* **163**, 648–658 (2013).
152. Stout, S. C., Clark, G. B., Archer-Evans, S. & Roux, S. J. Rapid and efficient suppression of gene expression in a single-cell model system, *Ceratopteris richardii*. *Plant Physiol.* **131**, 1165–1168 (2003).
153. Rutherford, G., Tanurdzic, M., Hasebe, M. & Banks, J. A. A systemic gene silencing method suitable for high throughput, reverse genetic analyses of gene function in fern gametophytes. *BMC Plant Biol.* **4**, 6 (2004).
154. Wolniak, S. M., van der Weele, C. M., Deeb, F., Boothby, T. & Klink, V. P. Extremes in rapid cellular morphogenesis: post-transcriptional regulation of spermatogenesis in *Marsilea vestita*. *Protoplasma* **248**, 457–473 (2011).
155. White, R. A. Experimental and developmental studies of the fern sporophyte. *Bot. Rev.* **37**, 509–540 (1971).
156. Trotta, A. et al. Arbuscular mycorrhizae increase the arsenic translocation factor in the As hyperaccumulating fern *Pteris vittata* L. *Chemosphere* **65**, 74–81 (2006).
157. Indriolo, E., Na, G., Ellis, D., Salt, D. E. & Banks, J. A. A vacuolar arsenite transporter necessary for arsenic tolerance in the arsenic hyperaccumulating fern *Pteris vittata* is missing in flowering plants. *Plant Cell* **22**, 2045–2057 (2010).

158. Nakamura, M. & Maeda, M. Isolation and culture of protoplasts from young sporophytes of *Salvinia natans* aseptically obtained by co-culture of female and male gametophytes. *Plant Cell Tissue Organ Cult.* **36**, 237–242 (1994).
159. Szypula, W. J., Mistrzak, P. & Olszowska, O. A new and fast method to obtain in vitro cultures of *Huperzia selago* (Huperziaceae) sporophytes, a club moss which is a source of huperzine A. *Acta Soc. Bot. Pol.* **82**, 313–320 (2013).
160. Hetherington, A. J., Emms, D. M., Kelly, S. & Dolan, L. Gene expression data support the hypothesis that *Isoetes* rootlets are true roots and not modified leaves. *Sci. Rep.* **19**, 21547 (2019).
161. Benca, J. P. Cultivation techniques for terrestrial clubmosses (Lycopodiaceae): conservation, research, and horticultural opportunities for an early-diverging plant lineage. *Am. Fern J.* **104**, 25–48 (2014).
162. Jha, T. B., Mukherjee, S., Basak, A. & Adhikari, J. In vitro morphogenesis in *Selaginella microphylla* (Kunth.) Spring. *Plant Biotechnol. Rep.* **7**, 239–245 (2013).
163. Hohe, A. & Reski, R. From axenic spore germination to molecular farming. One century of bryophyte in vitro culture. *Plant Cell Rep.* **23**, 513–521 (2005).
164. Finiuk, N. S. et al. Genetic transformation of moss *Ceratodon purpureus* by means of polycationic carriers of DNA. *Cytol. Genet.* **48**, 345–351 (2014).
165. Beike, A. K. et al. Clonal in vitro propagation of peat mosses (*Sphagnum* L.) as novel green resources for basic and applied research. *Plant Cell Tissue Organ Cult.* **120**, 1037–1049 (2015).
166. Ishizaki, K., Nishihama, R., Yamato, K. T. & Kohchi, T. Molecular genetic tools and techniques for *Marchantia polymorpha* research. *Plant Cell Physiol.* **57**, 262–270 (2016).
167. Alam, A. & Pandey, S. *Marchantia polymorpha* L.: an emerging model plant system to study contemporary plant biology – a review. *Plant Sci. Today* **3**, 88–99 (2016).
168. Flores-Sandoval, E., Dierschke, T., Fisher, T. J. & Bowman, J. L. Efficient and inducible use of artificial microRNAs in *Marchantia polymorpha*. *Plant Cell Physiol.* **57**, 281–290 (2016).
169. Sugano, S. S. et al. Efficient CRISPR/Cas9-based genome editing and its application to conditional genetic analysis in *Marchantia polymorpha*. *PLoS ONE* **13**, e0205117 (2018).
170. Szövényi, P. et al. Establishment of *Anthoceros agrestis* as a model species for studying the biology of hornworts. *BMC Plant Biol.* **15**, 98 (2015).
171. Enderlin, C. S. & Meeks, J. C. Pure culture and reconstitution of the *Anthoceros-Nostoc* symbiotic association. *Planta* **158**, 157–165 (1983).
172. Cole, T. C. H., Hilger, H. H. & Goffinet, B. Bryophyte phylogeny poster (BPP). Preprint at *PeerJ Preprints* https://doi.org/10.7287/peerj.preprints.27571v3 (2019).
173. The Pteridophyte Phylogeny Group. A community-derived classification for extant lycophytes and ferns. *J. Syst. Evol.* **54**, 563–603 (2016).
174. English, A. C. et al. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS ONE* **7**, e47768 (2012).

Acknowledgements

This project was supported by National Science Foundation grants IOS-1923011 and DEB-1831428 to F.-W.L., Swiss National Science Foundation grant nos. 160004, 131726 and 184826, the European Union's Horizon 2020 Research and Innovation Programme (PlantHUB-no. 722338), the Deutsche Forschungsgemeinschaft (German Research Foundation) under the Priority Programme 'MAdLand—Molecular Adaptation to Land: Plant Evolution to Change' (SPP 2237, 440370263), the Georges and Antoine Claraz Foundation, and the Forschungskredit and the University Research Priority Program 'Evolution in Action' of the University of Zurich to P.S. We thank K. Pryer and J. Banks for coining the term 'seed-free plants' and championing research in this area.

Author contributions

P.S., A.G. and F-W.L. wrote and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to P.S.

Peer review information *Nature Plants* thanks John Bowman, Yves van de Peer and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.