The evolutionary emergence of land plants

Philip C.J. Donoghue1,*, C. Jill Harrison2, Jordi Paps2, and Harald Schneider1,3
1School of Earth Sciences, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK
2School of Biological Sciences, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK
3Center of Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglin, Yunnan, China
*Correspondence: Phil.Donoghue@bristol.ac.uk
https://doi.org/10.1016/j.cub.2021.07.038

SUMMARY

There can be no doubt that early land plant evolution transformed the planet but, until recently, how and when this was achieved was unclear. Coincidence in the first appearance of land plant fossils and formative shifts in atmospheric oxygen and CO₂ are an artefact of the paucity of earlier terrestrial rocks. Disentangling the timing of land plant bodyplan assembly and its impact on global biogeochemical cycles has been precluded by uncertainty concerning the relationships of bryophytes to one another and to the tracheophytes, as well as the timescale over which these events unfolded. New genome and transcriptome sequencing projects, combined with the application of sophisticated phylogenomic modelling methods, have yielded increasing support for the Setaphyta clade of liverworts and mosses, within monophyletic bryophytes. We consider the evolution of anatomy, genes, genomes and of development within this phylogenetic context, concluding that many vascular plant (tracheophytes) novelties were already present in a comparatively complex last common ancestor of living land plants (embryophytes). Molecular clock analyses indicate that embryophytes emerged in a mid-Cambrian to early Ordovician interval, compatible with hypotheses on their role as geoengineers, precipitating early Palaeozoic glaciations.

Introduction

There can be no doubt that early land plant evolution transformed the planet, increasing the energy budget, changing atmospheric chemistry and the albedo of the continents, complexifying biogeochemical cycles through fungal symbioses, weathering and modified styles of sedimentation, and carbon fixation and storage, and creating habitats for metazoa terrestrialization. However, understanding of the timing and nature of phytoterrestrialization has been complicated by uncertainty concerning the fundamental relationships among embryos (land plants), specifically the relationships of three principal lineages of bryophytes to the tracheophytes (vascular plants) which dominate extant land plant diversity (Box 1: Embryophyte bodyplans). Hornworts, liverworts and mosses comprise the bryophytes, all of which exhibit haploid gametophyte-dominant life cycles, much like the haploid-dominant charophyte green algal relatives of land plants. After fertilization, diploid bryophyte sporophytes develop from gametophytes to which they remain attached and from which they are nourished. In contrast, vascular plants have sporophyte-dominant and (at least primitively) independent gametophyte and sporophyte life cycle stages; their monophyly is not contested. Many bryophyte similarities are plesiomorphies (see Box 2 for glossary) inherited from freshwater algal relatives, but some bryophyte lineages exhibit characteristics that are otherwise exclusive to tracheophytes, like stomata, water conducting cells, and leaves. These features have a complex taxonomic distribution that complicates both the possibility of their homology to tracheophyte counterparts, as well as the phylogenetic relationships between bryophytes and tracheophytes. Almost every possible hypothesis for bryophyte relationships to vascular plants has been proposed (Figure 1), from a monophyletic sister-group (Figure 1A), to a paraphyletic array of sister lineages (Figure 1B–I), with hornworts, liverworts and mosses alternately considered the most distant or closest relative of tracheophytes (Figure 1B–I). Each of these hypotheses implies competing patterns of character gain and loss in the evolutionary assembly of land plant bodyplans. However, the weight of evidence has recently swung towards what might have been perceived as the least likely solution: bryophytes comprise a natural group, sister to the vascular plants. Here we review the evidence for bryophyte monophyly and consider its implications for the evolution of early land plant evolution based on insights from comparative genomics, comparative developmental biology and the fossil record. Drawing this evidence together, we infer the embryophyte ancestor to have been a more complex (more tracheophyte-like) organism and a more capable geoengineer than has been perceived hitherto.

Evolutionary relationships among embryophytes

Historical heuristic analysis and cladistic analyses of morphology have generally considered bryophytes paraphyletic to vascular plants. This view was largely corroborated by phylogenetic analyses of nuclear ribosomal, plastid and nuclear protein coding genes, but the relative ordering of the relationships of hornworts, liverworts and mosses to the vascular plants has been disputed, usually with either hornworts or liverworts resolved as the earliest branching lineage of land plants, and either hornworts or mosses considered the closest relative of vascular plants. In any of these iterations, the ancestral land plant must have been of bryophytic grade because the immediate branches emerging from it are populated by bryophytes (Figure 1B–I);
Box 1. Embryophyte bodyplans.

The four principal lineages of living land plants are the hornworts, liverworts, mosses and the vascular plants. While the bryophytes, which comprise the first three, are superficially more similar to one another than to the vascular plants, this is largely because of their shared primitive characteristics. Here we outline the principal characteristics of their bodyplans. (A) Hornworts (Anthocerotophyta): Dominant gametophyte with dependent sporophyte, homosporous. Sporophyte unbranched and horn-like, lacking a differentiated seta and terminated by a single sporangium; stomata present only in sporophyte. Gametophyte usually has a thalloid growth form and lacks stomata. No conducting tissues. Simple rhizoid rooting system. Fungal associations are recruited from both Mucoromycotina and Glomeromycota. (B) Liverworts (Marchantiophyta): Dominant gametophyte with dependent sporophyte, homosporous. Sporophyte unbranched but differentiated into a setae and a single terminal sporangium capsule. Gametophyte bodyplans range from thalloid to elongate shoot with leaves organized in parallel rows, with the exception of Haplomitrium, which has spiral leaf arrangements. Water conducting tissues (where present) occur only in the gametophyte; stomata absent. Simple rhizoid rooting system. Fungal associations are recruited from both Mucoromycotina and Glomeromycota. (C) Mosses (Bryopsida): Dominant gametophyte with dependent sporophyte, homosporous. Sporophyte unbranched but differentiated into a setae and a single terminal sporangium capsule; stomata occur only in the sporophyte. Gametophyte usually differentiated into elongated shoots with leaves usually in spiral arrangements. Simple rhizoid rooting system. Where present, conducting tissues occur in both the gametophyte and sporophyte. In some species, the shoots contain differentiated water conducting cells. Fungal associations are usually absent. (D) Vascular plants (Tracheophyta): Dominant sporophyte with both generations independent (homosporous ferns, largely homosporous lycophytes) or the gametophyte dependent (heterosporous plants). Sporophyte branched with multiple sporangia formed per sporophyte, vascular tissue differentiated, stomata present in aerial parts of the sporophyte (with very few exceptions), often with bodyplans differentiated into shoot, root, and leaves, vascular tissue. Gametophyte reduced to highly reduced compared to sporophyte, free living gametophytes are autotrophic or heterotrophic with the later associated with fungal partners, free living gametophytes thalloid to filamentous. Branched rooting systems primitive. Fungal associations are found frequently in the sporophyte generation and free-living gametophytes. Fungal associations are recruited from both Mucoromycotina and Glomeromycota.

Bodyplans of living embryophytes.

Living embryophytes are divided into four fundamental lineages: (A) hornworts (Phaeoceros evanidus); (B) liverworts (Marchantia berteroana); (C) mosses (Hypnodendron menziesii); and (D) vascular plants (Cycas). Images A–C courtesy of Sylvia Pressel (NHM, London).
tracheophytes emerged from within the bryophytes. Thus, the successive bryophyte sister lineages of tracheophytes have been used experimentally to infer the nature of successive ancestors within the assembly of the vascular land plant bodyplan. Perhaps the most widely accepted phylogenetic scenario until very recently envisaged bryophytes as paraphyletic, with hornworts and/or mosses the closest relative of tracheophytes and liverworts the most distant among land plants, e.g.17 (Figures 1B,D,I). It is in this light that the fossil record has been rationalized18, models selected for studying early land plant development, genome composition and physiology17,19,20, and the evolutionary assembly of the embryophyte and tracheophyte

### Box 2. Glossary of morphological and phylogenetic terminology.

- **Capsule**: the swollen spore bearing structure of mosses and liverworts.
- **Coalescent analyses**: phylogenetic methods that evaluate the evolutionary histories of each gene individually in the inference of the species tree.
- **Compositional heterogeneity**: variation in the proportion of different nucleotides or amino acids in a molecular sequence, which in some cases may violate the assumptions of simple evolutionary models.
- **Compositional substitution models**: statistical models that profile nucleotide or amino acid changes in a molecular sequence over evolutionary time, taking into account the impact of sequence compositional heterogeneity.
- **Concatenation**: phylogenetic approach in which different genetic markers for a set of species are appended one after another in a supermatrix. The parameters of the substitution model can be calculated for each gene independently, or in sets of genes or sites (partitioning), or considering all the markers together.
- **Crown-ancestor**: the last common ancestor of all members of a group of living species — the crown group.
- **Crown-group**: clade defined by its living members, their most recent common ancestor plus all of its descendants, living and extinct (e.g. living birds comprise their crown-group which also includes extinct phylogenetic intermediates).
- **Cryptospores**: a grade of permanently fused spores that occurs early in the land plant fossil record.
- **Developmental recapitulation**: the general view that the sequence of events in ontogeny parallels their phylogenetic sequence of evolutionary origin. Originally invoked as the Biogenetic Law by Ernst Haekel, we here simply observe the ontogenetic–phylogenetic parallel.
- **Dyad spores**: a pair of spores derived from a single and complete meiotic division.
- **Heterogeneous substitution models**: statistical models that profile nucleotide or amino acid changes in a molecular sequence over evolutionary time, assuming that the rate of change varies between nucleotide positions within genes, between lineages, and/or between time intervals.
- **Homogeneous substitution models**: statistical models that profile nucleotide or amino acid changes in a molecular sequence over evolutionary time, assuming that the rate of change is the same across all the positions within genes, among different lineages, and/or over time.
- **Homosporous**: having spores of a single size.
- **Horizontal Gene Transfer (HGT)**: also known as lateral gene transfer, is the transference of genetic material between organisms that do not have a parent to offspring relationship, in contrast with ‘vertical’ transfer, which is associated with reproduction.
- **Initials**: plant stem cells.
- **Intercalary proliferation**: while proliferation to supply new cells for shoot development in most vascular plants occurs apically, intercalary meristems are active between the apex and base of a shoot axis.
- **Molecular clock**: use of molecular sequences to infer the time of origin of clades, calibrating the rate of molecular changes using geological events or the fossil record.
- **Monophyletic group or clade**: a group of organisms descended from a last common ancestor and including all its descendants. Clades are considered natural groupings that inform taxonomy and classification.
- **Paraphyletic**: a group of organisms descended from a single common ancestor but not including all of its descendants (e.g. fish, which excludes tetrapods).
- **Phylogenomics**: inference of the evolutionary relationships of organisms using large-scale gene datasets, often derived from genome or transcriptome sequencing.
- **Plesiomorphies**: characters or character states that are shared among members of a specific group of organisms but inherited from a remote ancestor and, therefore, not informative of the relationships among the specific group.
- **Reniform**: kidney shaped.
- **Rhizomes**: horizontal underground shoot systems.
- **Setaphyta**: the monophyletic group of liverworts and mosses.
- **Stem-group**: paraphyletic assemblage of extinct lineages sister to a clade defined by its living members (e.g. dinosaurs are the stem-group birds).
- **Tetrad spores**: a cluster of four spores derived from a single and complete meiotic division.
- **Total-group**: clade comprised of a stem-group and its respective crown group (e.g. dinosaurs plus birds).
and amino acid substitution rates between sites within the genome and between evolutionary lineages, which are not addressed in the homogeneous substitution models conventionally employed in molecular phylogenetics. Such biases have impacted some of the highest profile controversies in phylogenetics, including the relationships among cnidarians and sponges with respect to other animals, the fundamental relationships among mammals, and the coherence of some animal supergroups.

Cox and colleagues have shown that support for bryophyte paraphyly in legacy plastid datasets is an artefact of among-lineage compositional heterogeneity—biases in the use of related codons by different sites within the genome. Attempts to control for this using a bespoke substitution matrix favour bryophyte monophyly. Use of site-homogeneous models to analyse an initial 1KP dataset of 852 nuclear protein coding genes recovered bryophyte paraphyly, while a coalescent analysis of amino acid gene trees recovered monophyletic bryophytes. These data have been reanalysed using models that address both site- and lineage-specific compositional heterogeneity, using both concatenation (where all genes are analysed together) as well as coalescent (where genes are analysed individually) approaches, inductively seeking optimal trees and deductively exploring the fit of the data to prior competing topologies. Puttick and colleagues found unequivocal support for the setaphyte clade of liverworts plus mosses (named for the seta which supports the sporangium in these taxa) and, while the sum of their analyses strongly favoured bryophyte monophyly, they could not formally reject the possibility of hornworts as sister to setaphytes plus tracheophytes. Sousa and colleagues showed that support for bryophyte paraphyly in these
data is an artefact of synonymous substitutions, and analysis using a lineage-heterogenous compositional model supported bryophyte monophyly. Their analyses of more comprehensive mitochondrial datasets reached similar conclusions, but composition-aware modelling of plastid datasets supports bryophyte monophyly. Analysis of an expanded taxon dataset including increased sampling of hornworts recovered both seta-phyte and bryophyte monophyly through coalescent analysis of nuclear and plastid genes, as well as through analysis of a concatenated nuclear gene dataset using a site-specific model of compositional heterogeneity.

While relationships among bryophytes and between bryo-phytes and tracheophytes will remain the focus of phylogenetic investigation for some time to come, the problem has been addressed by diverse phylogenetic methods applied to datasets that have resolved the long-standing shortfall in the taxonomic and genomic sampling of bryophytes. Phylogenetic studies now commonly recover the Setaphyta clade of liverworts and mosses and bryophyte monophyly has become widely accepted as the conventional working hypothesis for elucidating the nature of early land plants and the evolutionary assembly of their bodyplans and genomes.

While our understanding of bryophyte relationships has been in a state of flux through this period, the monophyly of tracheo-phytes has remained uncontentious, with lycophytes (club-mosses, spikemosses, quillworts and scale trees) resolved as sister to the euphyllophyte clade of monilophytes (ferns and horsetails) and spermatophytes (seed plants including gymno-sperms and angiosperms). Meanwhile, hypotheses on the phylo-genetic relationships among land plant relatives have proven more contentious (Figure 2). Chara and Colechaete (Charophy-ceae and Coleochaetophyceae in Figure 2) have long featured in scenarios for the origin of land plants (e.g.) because they exhibit a number of land plant characteristics (branching, tissues, cell walls with plasmodesmata, apical meristems, asymmetric cell division and zygotes that produce sporopollenin) and many phylogenetic studies have recovered them as sister to embryophytes (Figure 2A–C). However, most recent molecular phylogenetic analyses identify the aflagellate, conjugating and largely unicellular Zygnematophyceae (though some forms are colonial, filamentous and exhibit branching) as the immediate sister-lineage to embryophytes.

**The long road of genome assembly versus phenotypic traits**

The new genomic resources delivered by a revolution in genome sequencing technologies has led not only to a revised understanding of land plant phylogeny, but also to fundamental new insights into the evolution of land plant genomes. More than anything, comparative genomics has revealed that the genes associated with many of the key anatomical and physiological traits of land plants have a much more ancient evolutionary origin. The origin of land plants is as much associated with gene co-option, exaptation and even horizontal gene transfer as it is with fundamental innovation.

Chromosome-level events (e.g., whole genome duplications, emergence of epigenetic mechanisms) have surely been influen-tial in early land plant evolution, but their study is complicated by the plasticity of plant genomes (e.g., hybridisation, polyploidy, dysploidy) and the poor contiguity of the genomes currently available for species that phylogenetically bracket major evolutionary transitions. Consequently, focus has shifted towards tracing the evolution of gene complements, including well-studied candidate genes whose roles are known to be influential in the development of key innovations, but also in the discovery of new genes using comparative genomics. Bowles and colleagues undertook a forensic comparative analysis of genomes from over 200 eukaryotes, including 158 land plant species, focussed on dissecting the evolutionary assembly of plant genomes. These analyses show unprecedented levels of gene novelty associated with two critical branches in plant phylogeny, reflecting the origin of streptophytes (the clade composed of land plants and their algal relatives, Zygnemophytaeae, Coleochaetophyceae, Charophyceae, etc.).
A similar burst of new genes occurred at the origin of animals44, sent in most taxa, and thus inferred to be biologically important. To account for the evolutionary distance between living species (whose genomes were investigated) at the tips of the phylogenetic tree, the ancient ancestors (represented by nodes in the tree) in which the genes are inferred to have originated60. However, hypotheses on the ancestral roles of genes from derived model organisms like Arabidopsis thaliana are open to testing through reverse genetics in extant lineages that bracket evolutionary transitions. For example, commonality of function in bryophytes and tracheophytes that is different to orthologous gene function in zygnematophytes and coleochaetophytes would indicate gain of function in the embryophyte stem-lineage.

In the ancestor of streptophytes, a large proportion of the genetic novelties are transcription factors, genes involved in gene regulation and usually linked with the emergence of cell types in multicellular organisms45. In the branch leading to land plants, most of the novel core genes are involved in functions typically related to terrestrial environments, such as UV light protection, root development, interactions between plants and microbes, tolerance to drought, and phytohormones involved in response to land stressors. Other recent studies expanding the sampling of charophytes have also shown that transcription factors, phytohormones, and genes involved in adaptations to life on land (e.g., drought resistance) emerged coincident with the origin of embryophytes56-68. Other remarkable land plant novelties include genes of the MYB class genes or CYP77A (cuticle biosynthesis)69, as well as CLAVATA (3D growth)59.

However, genomic novelty is not the only process involved in these transitions. Some candidate genes evolved before the phenotypic novelty with which they are causally implicated, reflecting gene co-option. For example, genes involved in multicellular development, embryonic morphogenesis, roots, trichomes, seeds, flowers, or lateral organs predate the emergence of their respective anatomical novelties69. PIN auxin transporters involved in 3D growth evolved in streptophyte algae70. Transcription factors from gene regulatory networks involved in differentiation of land plant structures have been shown to have emerged in streptophyte algae and have been retained throughout land plants71. For example, Class I and II KNOX genes, homeobox transcription factors central to land plant development, arose through duplication of a single ancestral orthologue, prior to the origin of the Phragmoplastyphtya clade54.

Perhaps surprisingly, some innovations arose through horizontal gene transfer (HGT) — the handover of genes from lineages that are distant in the Tree of Life55. The first genome of a moss revealed horizontal transfer from bacteria, fungi, and viruses of genes involved in xylem formation, plant defence, or metabolism76. A new analysis of genomes of zygnematophyceae algae shows that stress-response genes GRAS and PYL might have been incorporated by transfer from soil bacteria to the ancestor of these algae and land plants67. Similarly, Bowles et al.45 found 323 gene families potentially shared only between fungi and land plants related to gene regulation and protein modification. The evolutionary importance of the interactions between plants and microbes, such as arbuscular mycorrhizae (which supply nitrogen to plants in exchange for photosynthate), cannot be overstated77.

Gene gains by duplication or whole genome duplication, exaptation of older genes, and HGT all contributed to shaping the genome of early land plants. Some of the mechanisms underlying these processes are not well understood and require further investigation. Genome reduction, in particular, has been implicated as a major evolutionary force shaping genomes78 and shown to be influential in animal genome evolution60,61. However, though it has clearly played a role in Zygnematophyceae79,80, the role of gene loss in land plant evolution has yet to be the subject of systematic investigation.

Developmental evolution

The evolutionary implications of comparative developmental genetic studies are contingent on the phylogenetic framework within which they are considered. Since the engineering of moss83, liverwort84 and now even hornwort85 developmental models, it is no longer necessary to rely on a single species as a proxy for the embryophyte ancestor, which is reassuring since each model is a chimera of ancestral and derived characteristics. However, by marshalling insights from all of these lineages in comparison to tracheophytes and algal relatives, it is possible to leverage insights into the developmental biology of embryophyte, tracheophyte and bryophyte ancestors and the developmental evolution of their bodyplans. Here we consider fundamental insights that comparative developmental genetics has revealed about the homology and evolution of key phenotypic novelties of land plants.

Cuticle and stomata

Dehydration due to evaporative water loss and UV irradiation would have been among the most proximal of environmental challenges faced by plants on land, and the innovation of the waxy cuticle in the last shared common ancestor of land plants conferred protection56. The genetic toolkit for cutin and cuticular wax biosynthesis is shared among land plants, including transcriptional regulators belonging to the MYB class, phenolic metabolism enzymes such as CYP77A, and secretory pathway components49,67. Reverse genetic analysis in bryophytes has begun to reveal functional conservation and identified mutants with more permeable cuticles (e.g.,58,69).

2D to 3D growth

Land plants also gained capacity to rotate cell divisions through multiple planes during apical growth70, which enabled volumetric space occupation (3D growth). The algal sister lineages to land plants lack this capacity and are constrained to matlike or filamentous growth habits77,78. The evolutionary innovation of 3D growth is recapitulated by developmental transitions in bryophytes, and mosses in particular undergo a protracted filamentous 2D growth phase prior to the onset of 3D growth in the gametophyte stage of the life cycle78. Rotating cell division plane orientations during this transition are coupled to changes in cell growth and cell fate and have been proven amenable to reverse and forward genetic perturbation without problems
associated with tissue complexity or lethality\textsuperscript{73}. Reverse genetic approaches have identified roles for APB transcription factors\textsuperscript{74}, PIN auxin transporters\textsuperscript{75,76}, CLAVATA receptor-like kinases\textsuperscript{77,78} and the DEK1 calpain protease\textsuperscript{79} during the transition to 3D growth, and forward genetic analyses are starting to dissect genetic interactions\textsuperscript{79–81}.

**Gametophyte bodies**

The initials of bryophyte gametophytes either generate leafy gametophores (mosses and leafy liverworts) or thalli (thalloid liverworts and hornworts), and although these serve the ultimate purpose of gamete production, they have widely divergent morphologies and can be monoecious (hermaphrodite) or dioecious (having male and female plants). Diversity among bryophytes points to convergent evolution of indeterminate leaf and branching forms in gametophytes and sporophytes. While some of the genetic mechanisms regulating these traits have been independently recruited to regulate gametophyte and sporophyte development (e.g., PIN- and CLV- regulation of apical function\textsuperscript{80,81}, HD-Zip regulation of leaf development\textsuperscript{82}) other genetic solutions to morphological innovations are distinct (e.g., divergent auxin transport mechanisms in the regulation of gametophyte and sporophyte branching in a moss\textsuperscript{83}). The independent evolution of multicellularity in the haploid gametophyte stage of the life cycle and the diploid sporophyte stage of the life cycle of land plants offers opportunities to determine the genetic basis of convergent evolution.

**Embryos, dichotomous branching and axis extension**

While the algal relatives of land plants have diverse gametophyte morphologies, fertilisation is followed rapidly and uniformly by meiosis\textsuperscript{85}. Land plant life cycles are characterised by delayed meiosis, with a period of embryonic growth prior to sporangium development in bryophytes or vegetative development in vascular plants\textsuperscript{86–88}. While bryophyte sporophytes usually have a single growth axis, perturbation of polar auxin transport or TCP transcription factor function in mosses can induce branching, potentially by an early and equal division of the apical cell or respecification of epidermal cells as apical cells\textsuperscript{89,90}, and the innovation of bifurcation may have boosted spore production in stem-tracheophytes by amplifying the number of branches, each bearing a terminal sporangium. A nascent capacity for vegetative growth is indicated by KNOX promotion of proliferation and elongation from an intercalary region in moss sporophyte axes\textsuperscript{89,91}, and the conservation of KNOX function in promoting apical proliferation among vascular plants with indeterminate shoot growth\textsuperscript{92–96}. Intercalary proliferation in moss sporophytes occurs after embryonic axis establishment by apical and basal stem cells, serving to push the developing capsule through maternal tissues\textsuperscript{97}.

**Vegetative shoots**

In vascular plant shoot apices, stem cells and proliferative zones marked by KNOX activity are juxtaposed, and the displacement of sporangia away from the shoot tips may have enabled the innovation of indeterminate shoot growth\textsuperscript{98}. In contrast to the innovation of indeterminacy, sporophyte leaves arose many times convergently by modification of branching shoot systems in different parts of the land plant phylogeny\textsuperscript{99}. The cellular and developmental processes involved in leaf morphogenesis in different lineages are not fully clear. Lycophyte leaves are thought to have arisen by enation (epidermal outgrowth)\textsuperscript{98,100} or the reduction of lateral branches bearing sporangia (sterilisation)\textsuperscript{101} and fern frond initials arise from the apical epidermis\textsuperscript{102,103}, but pools of cells from the shoot apical meristem are recruited into seed plant leaf development\textsuperscript{104}. The extent of conservation in pathways for leaf development among vascular plants is currently unclear but expression analyses indicate that roles for HD-Zip genes in leaf initiation and leaf patterning are conserved within seed plants, but not to lycophytes\textsuperscript{105–107}, and roles for KANADI transcription factors in specifying the abaxial leaf domain are conserved to the level of eukaryophytes\textsuperscript{108}. While mutually exclusive KNOX (meristem) and ARP (leaf) expression characterises flowering plant leaf development, ferns have overlapping KNOX and ARP expression in leaves\textsuperscript{99,34,96}, and some lycophytes gained a mutually exclusive pattern independently of flowering plants\textsuperscript{92}.

**Vascular and supporting tissues**

Shoot growth necessitated the evolution of transport mechanisms, and it is generally accepted that vascular and supporting tissues evolved with tracheophytes, and tracheids with annual-helical thickenings are a synapomorphy of tracheophytes. However, mosses have long been recognised to possess water (hydroid) and food (leptoid) conducting cells, as well as thickened supporting cells (stereids)\textsuperscript{99,109,110} but their homology to vascular tissues in tracheophytes has been contentious\textsuperscript{111,112}. Xu and colleagues\textsuperscript{112} have shown that VNS transcription factors, which are functionally implicated in xylem, phloem and secondary wall thickening in Arabidopsis, are functionally implicated in moss hydroid and stereid development, influencing water conductance, hydroid cell maturation, secondary wall thickening and apoptosis. VNS transcription factors and their orthologues regulate the same genes in both mosses and vascular plants, providing support for the homology of a VNS-based gene regulatory system underpinning development of water-conducting and supporting tissues in all land plants\textsuperscript{112}. Meanwhile, Norris and colleagues\textsuperscript{113} have demonstrated that moss stereids have structurally distinct secondary walls and the deposition of primary and secondary cell walls is regulated by distinct isoforms of cellulose synthase genes, paralleling the diversification of the cellulose synthase gene family and its regulatory function in seed plants. Similarly, TM05 and LHW basic helix-loop-helix (bHLH) transcription factors, which promote periclinal and radial cell division to expand the girth of vascular tissues in Arabidopsis, arose prior to land plants and cross-species complementation experiments from Marchantia to Arabidopsis suggest conservation of TM05 function within land plants\textsuperscript{112}.

**Roots**

Root systems comprising rhizoids for anchorage arose in the algal sister lineages of land plants\textsuperscript{114} and the presence of similar rooting systems in bryophyte and monilophyte gametophytes suggests that these represent the earliest land plant anchorage system\textsuperscript{115}. The mechanisms promoting rhizoid development are conserved between bryophyte groups, and also conserved with mechanisms regulating root hair development in flowering plant sporophytes\textsuperscript{116–118}. The bHLH transcription factors involved predate the origin of land plants\textsuperscript{119} but a Chara braunii bHLHVIII sister gene to the land plant RSL gene clade functions non-equivalently to land plant genes, suggesting that neofunctionalisation was involved in the emergence of rhizoid-based rooting systems\textsuperscript{120}.
Less is known about the genetic mechanisms involved in root evolution. WOX regulators of meristem function are conserved between the shoots (WUS) and roots (WOX5) of Arabidopsis, and a T3 WOX similar to WUS and WOX5 is expressed in the root tips of Ceratopteris, a fern. CWUL can complement wus-1 mutants (shoot apex) but not wox5 mutants (root apex) when expressed heterologously in Arabidopsis. Ceratopteris expresses CrWOXB in both gametophyte and sporophyte life cycle stages, but RNAi lines showed that, while there is a sporophytic function in root and lateral root development, the gametophytic function is in promoting cell proliferation. The function of WOX genes in lycophytes remains unknown due to the lack of a genetic model, but SmiWOX genes are expressed in a broad range of tissues. Bryophytes only have WOX13-like T1 WOXes and these function in growth promotion in Physcomitrium and Marchantia. The data currently point to the emergence of ‘rooting’ WOX functions with sister genes to the T2WOX and T3WOX clade in the emergence of vascular plants. A later innovation of roots, the capacity for fast gravitropic growth responses, has been linked to seed plant innovations in PIN auxin transporter function.

**Fossil record**

The oldest possible fossil evidence for land plants occurs as late Cambrian cryptospores, but their irregular arrangements and occurrence in ‘packets’ of multiple spore-like bodies surrounded by syneocosporal walls has led to algal interpretations. The oldest unequivocal evidence for the divergence of the embryophyte lineage from algal relatives is also based on cryptospores, but these occur in regular fused geometric arrangements of dyads and tetrads of middle Ordovician (~469 Ma) age (Figure 3A–D). These are accepted as embryophytes based on similarity to permanent tetrads and dyads produced by living bryophytes, though it is unclear whether these earliest records are stem- or crown-embryophytes (see Box 3 for an explanation of crown-, stem- and total-group classification). Younger records occur within fossil sporangia on bifurcating axes, perhaps representing stem- or total-group embryophytes. The earliest unequivocal evidence for the divergence of tracheophytes and bryophytes can be constrained by a middle Silurian (~429.3 Ma) record of Cooksonia that is known from younger anatomically preserved specimens of sporophytes bearing terminal sporangia (Figure 3E,F) containing spores (Figure 3G), located on branching axes (Figure 3E) containing tracheids (Figure 3H), steme, and bearing stomata (Figure 3I)139,140. More derived stem-tracheophytes like Rhynia (Figure 3J,K) and Homeophyton (Figure 3L,M), as well as early crown-tracheophytes possessed bryophyte-like rhizoid anchoring systems, while true roots with gravitropism, a root cap and root hairs are features seen only in derived lycophytes and euphyllophytes. The earliest evidence for the tracheophyte crown-group is late Silurian (~420.7 Ma) based on zosterophylls, which preserve evidence of lateral insertion of reniform sporangia along stems, which dehisce along their distal margins into two valves, substantiating an affinity with lycophytes. Thus, while the fossil record is an incomplete archive of early land plant lineages, life cycles and their anatomy, subject to systematic biases in the rock record, it provides unique insights into the evolutionary assembly of land plant body plan characteristics. The challenge is to interpret these data phylogenetically. While the majority of early land plant fossils have been interpreted as tracheophytes, changing perspectives on bryophyte–tracheophyte relationships are leading to inference of a more complex ancestral embryophyte.

**Implications for character evolution**

The traditional land plant phylogeny, in which liverworts were considered sister to all other bryophytes and tracheophytes, led to the reconstruction of a fairly simple, gametophyte-dominant ancestral crown embryophyte, missing traits that have been considered integral to the biogeochemical impact, as well as the evolutionary and ecological success of extant land plants. These include absence of stomata, rooting systems, dichotomous branching, upright axes, leaves, dehiscent spores, etc. In large part, this is because liverworts lack many of these characters (e.g.,) and because our developmental and genome model taxa are often interpreted too literally, as model proxies for ancestral organisms. Rather, insights should be sought from comparative analysis of lineages that phylogenetically bracket evolutionary phenomena, allowing us to indirectly infer the nature of ancestral organisms, drawing on evidence, as we have, from the comparative anatomy of living and fossil species, as well as comparative developmental genetics.

Inferences of many of the characteristics of the embryophyte crown ancestor are unaffected by whether bryophytes are resolved as monophyletic or paraphyletic. The shared characteristics of living embryophytes include embryonic development, a bifasic life cycle including a 3D bodyplan — though it is not clear whether this was present in the haploid gametophyte and/or diplloid sporophyte phase. Following the same principle, a waxy cuticle would also have been present, along with rhizoids, valvate sporangia and dehiscent trilete spores — though it is not clear whether the ancestral crown embryophyte also possessed permanent non-dehiscent dyad and tetrad spores. However, without phylogenetic resolution of the relationships between bryophytes and tracheophytes, it is also unclear whether the land plants ancestrally possessed stomata, vascular tissues and branching axes, or the nature of their life cycles.

Bryophyte monophyly and, in particular, the setaphyte clade of liverworts and mosses, helps to resolve some of these uncertainties, leading to the conclusion that key embryophyte characteristics are missing through loss from liverworts. Stomata are present in mosses, hornworts and tracheophytes and, hence, following phylogenetic reasoning, they must have been lost within the lineage leading to extant liverworts after divergence from the moss lineage. It remains formally possible that stomata have evolved independently in mosses and hornworts, separately from tracheophytes, since they exhibit a patchy phylogenetic distribution in bryophyte lineages where they appear to function differently. However, Harris and colleagues have demonstrated that key components of the gene regulatory network underpinning tracheophyte stomatal development are present in mosses and hornworts, corroborating their homology and indicating that they were present in the embryophyte crown ancestor. Similarly, extant bryophytes do not exhibit dichotomous branching and so this trait has been conventionally interpreted as a tracheophyte innovation. However, as we have seen, dichotomous branching can be induced experimentally.
Figure 3. Fossil representatives of early land plants.
Fossils preserve evidence that informs the evolution of key embryophyte novelties. (A–D) sporangial mass and isolated permanent dyad and tetrad cryptospores from the middle Ordovician of Saudi Arabia; (A) sporangial cryptospore mass; (B) magnification of (A) showing individual cryptospores; (C,D) isolated cryptospores from the same sample; (E) permanent dyad cryptospore; (D) permanent tetrad cryptospore. (E) *Cooksonia pertoni* branched sporophyte with terminal sporangium (Early Devonian). (F–I) *Cooksonia* sporophyte preserved in three dimensions through charcoalification (Early Devonian); (F) sporangium and supporting axis; (G) spores in situ within the sporangium; (H) vascular tissues including tracheid from the sporophyte axis; (I) stoma from the sporophyte axis. (J–M) Rhizoids associated with stem-tracheophytes *Rhynia* and *Horneophyton* from the Early Devonian Rhynie Chert; (J) section through a *Rhynia* axis with rhizoids extending from the lower surface; (K) reconstruction of *Rhynia* from; (L) reconstruction of *Horneophyton* from; (M) section through a *Horneophyton* axis with rhizoids extending from the lower surface. Images (A–D) courtesy of Charles Wellman (Sheffield University); (E–I) courtesy of Dianne Edwards (Cardiff University); (J–M) courtesy of Alexander Hetherington (Edinburgh University).
Given the shared mechanistic basis of branching in mosses and tracheophytes, it is likely that the last common ancestor of living embryophytes was also capable of dichotomous axial branching. The fossil record demonstrates that the rhizoid-based anchoring systems of living charophycean algae and bryophytes were a feature not only of the ancestral crown-embryophyte, but of the ancestral crown-tracheophyte as well; true roots with gravitropism, a root cap and root hairs are features seen only in derived lycophytes and euphyllophytes\textsuperscript{141,142}. It is also possible that even vascular tissues are a shared primitive feature of land plants, rather than a derived characteristic of vascular plants. Scheirer\textsuperscript{109} drew comparisons between moss leptoids and phloem-like sieve elements, moss hydroids and xylem-like tracheids, arguing for their homology. Ligrone and
The originating analysis sampled few species from these lineages and the clades so-defined are not meaningful here. Associated blue distributions reflect the extant (mostly crown-clade) diversity, pinned to the median age estimate; Charophyceae and Anthoceratophyta are represented as single lineages because probability of clade age estimate, which should be interpreted in terms of its span, not the median point estimate. Icons for Mesostigmatophyceae, Klebsormidiophyceae should be anticipated in extant bryophytes, just as stoma exhibit a sporadic phylogenetic distribution among hornworts and mosses. Vestigial tracheids are not seen in extant bryophytes and so we should perhaps conclude that they are a tracheophyte innovation.

Consequently, the last common ancestor of living land plants must have been more complex and tracheophyte-like than has been perceived hitherto, since we can infer that comparatively few phenotypic characters distinguish the respective crown-ancestors of the tracheophyte and embryophyte clades (Figure 4). At the same time, identification of Zygnematophyceae as sister to Embryophyta might imply that many of the embryophyte-like characteristics of Charophyceae and Coleochaetophyceae are the consequence of convergent evolution, and that embryophytes evolved from much simpler pond scum, exemplified by their zygnematophyceae kin. Inevitably, the truth is more complicated and there is evidence for both convergence on complex bodyplans among Charophyceae and Embryophyta and simplification in Zygnematophyceae.

*Figure 4. Timescale of streptophyte phylogeny and the origin of land plant novelties.*

This summary timescale of streptophyte phylogeny is based on the dated ‘monophyletic bryophytes’ phylogeny of Morris and colleagues. The triangles reflect the extant (mostly crown-clade) diversity, pinned to the median age estimate; Charophyceae and Anthoceratophyta are represented as single lineages because probability of clade age estimate, which should be interpreted in terms of its span, not the median point estimate. Icons for Mesostigmatophyceae, Klebsormidiophyceae, Charophyceae and Zygnematophyceae are from Phylopic.org. Klebsormidiophyceae from Phylopic/Matt Crook (CC BY-SA 3.0).
adaptations to life on land have a deep evolutionary origin among green algae, such as the emergence of genes for traits generally considered distinctly land plant specific, including cell wall biosynthesis\textsuperscript{146}, hormones, signalling pathway elements, as well as drought- and light-related stress response factors\textsuperscript{77,147}. No doubt these traits evolved within freshwater green algae because they were episodically confronted by the same environmental challenges as later land plants. However, the embryophyte stem-lineage coincides with the origin of large numbers of novel genes associated with land plant traits\textsuperscript{148}, insights into which we otherwise have no recourse but to the fossil record.

**Timescale of early land plant evolution**

We should anticipate that the fossil record underestimates the true age of clades because lineage divergences are initially genetic phenomena and evidence of occurrence cannot be diagnosed in the fossil record until fossilizable features evolve within derivative lineages\textsuperscript{149}. This problem is accentuated by a dearth of early Palaeozoic terrestrial rock sequences, and when terrestrial strata become prevalent, from the middle to late Silurian, they preserve evidence of terrestrial plant and animals that are already differentiated and organised into ecosystems\textsuperscript{8,149,150}. There can be no fossil record without a rock record in which to entomb it and the paucity of early Palaeozoic terrestrial strata has the effect of telescoping the earliest fossil records of land plant lineages\textsuperscript{141}, an interpretation corroborated by Ordovician records of terrestrial spores within marine strata, tens of millions of years older than the earliest records of land plants in terrestrial sequences\textsuperscript{151}. Land plants have undoubtedly impacted on the sedimentary rock record principally as a consequence of the evolution of rooting systems that prevent erosion and promote deposition of fine sediment\textsuperscript{152,153}. This shift coincides with the earliest fossil records of embryophytes and this has been marshalled in support of a literal interpretation of the fossil record\textsuperscript{154}. However, as we have seen, rooting systems are a derived characteristic of tracheophytes, evolving convergently in lycophytes and euphyllophytes; even early crown-tracheophytes possessed simple rhizoids like those of their bryophyte and algal relatives\textsuperscript{140,156} and these are not envisaged to have had any impact on sedimentary systems. Consequently, there is a mismatch between the Devonian origin of rooting systems and Ordovician sedimentary phenomena that their evolution is envisaged to have affected.

Therefore, in the absence of a literal record of plant evolution to read, there is no recourse to methods that might help to interpret the imperfect archive of evolutionary history that the fossil record represents. While molecular clock methods have been widely employed in plant science, there have been surprisingly few attempts to date the deep divergences within land plants until recently\textsuperscript{157–160} and these differ as much in methods as in the estimates that they derive. Most of the earlier clock studies include just moss as a representative bryophyte, and so assuming bryophyte monophyly, they effectively date the ancestral embryophyte node. The application of strict clock methods and few calibrations that integrate no uncertainty resulted in late Tonian to early Ediacaran age estimates (e.g. 703 Ma ±45 Myr\textsuperscript{157}, 707 Ma ± 98 Myr\textsuperscript{158}), while methods that allow for rate variance have tended to recover much younger estimates. Sanderson’s penalized likelihood approach estimated a late Cambrian age (490 Ma) for the ancestral embryophyte\textsuperscript{160} similar to later Bayesian relaxed clock studies that have recovered ages in the range late Ediacaran to early Ordovician (513–489 Ma\textsuperscript{159}, 557–407 Ma\textsuperscript{151}, 480.35–471.35 Ma\textsuperscript{163}; 515.2–473.5 Ma\textsuperscript{163}, 559.3–459.9 Ma\textsuperscript{164}). Tracheophyta divergence has been estimated within the interval middle Ordovician to early Devonian (469–399 Ma\textsuperscript{161}; 434.34–416.28 Ma\textsuperscript{160}; 450.8–431.2 Ma\textsuperscript{163}, 457.6–438.3 Ma\textsuperscript{165}).

To be clear, these estimates are argostic to the true divergence time within their range, allowing for close to poor approximations of the calibrating fossil evidence. The differences among them reflect differences in the interpretation of the fossil record and the prior views on how the fossil evidence approximates the ages of clades\textsuperscript{165,166}. Crucial to constraining the timing of divergence of land plants is the near-indestructability of sporopollenin spore walls and their accumulation in marine sediments, transported from the land. Proterozoic and early Phanerozoic marine sequences have been extensively sampled for algal cysts used in biostratigraphy and nothing resembling a land plant spore has been recovered from sediments prior to the Middle Cambrian. The preservation of algal cysts in older strata serves as a taphonomic control, demonstrating that where land plants present, their spores would be preserved and, thus, their absence constitutes evidence that embryophytes were not present at this time\textsuperscript{163,166,167}. Given the temporal proximity of fossil constraints on the age of tracheophytes, this provides for a very informative calibration. Recent studies\textsuperscript{165,168} that have eschewed this calibration recover deep Proterozoic estimates for the origin of crown-embryophytes and even crown-tracheophytes, like those of early strict clock studies\textsuperscript{157,158}; these timescales can be readily rejected based on the available palaeontological and geological evidence. However, analyses in which these calibrations are employed estimate embryophytes to have diverged from Zygnematophyceae in a late Tonian to middle Cambrian interval, bryophyte–tracheophyte divergence (crown-Embryophyta) in the interval 514.8–473.5 Ma, hornwort–setaphyte (crown-Bryophyta) in the interval 506.4–460.3 Ma and lycophyte–euphyllophyte (crown-Tracheophyta) in the interval 450.8–431.2 Ma\textsuperscript{169}.

**Fungal associations**

The successful colonization of challenging terrestrial environments by the ancestors of land plants was made possible through partnership with fungi. Plant–fungal associations recruited fungal partners mainly from two lineages, namely the Glomeromycota — a lineage of fungi well known for their contribution to mycorrhizae in many vascular plants — and Mucoromycotina. Evidence for plant–fungal partnerships has been documented in the early land plant fossil record, especially the Rhynie Chert\textsuperscript{169}. Their widespread occurrence in most ancient plant lineages, especially hornworts, liverworts, and basal vascular plant lineages such as homosporous lycophytes and ferns, suggests the formation of this partnership dates back to the common ancestors of all land plants\textsuperscript{170}. However, not all lineages maintained these partnerships, as illustrated by the absence of fungal association in the majority of mosses\textsuperscript{170}. Whereas the fungal partnerships in the roots of vascular plants and the autotrophic gametophytes of vascular plants have been well characterized, the rapid expansion of a diversity of
fungal–plant partnerships raises core questions about the kind of interactions between these partners during the early divergence of land plants. Understanding these early partnerships will arguably be crucial to reconstruct the greening of the continents.

Implications of the greening of the continents for Earth system evolution

While it is widely appreciated that the evolution of land plants has had a transformative impact on the Earth system, the inception of this role has been unclear due to a combination of uncertainties concerning the timing of origin of both land plant lineages and the key novelties that impact the Earth system. The origin of trees and their complex and deep rooting systems in the Devonian, for example, is implicated in CO₂ draw down through silicate weathering resulting in middle Palaeozoic glaciation, but a role for plants as geoengineers prior to this time is often dismissed, for example, because early land plants are perceived to have had a negligible weathering effect and because their biomass is thought to have been limited due to desiccation intolerance. However, mosses with vascular systems, which might serve as a model for ancestral embryophytes, exhibit impressive resistance to dehydration and they have been demonstrated to enhance silicate weathering within a factor of that seen in vascular plants. Biogeochmical modelling which assumes an early Ordovician (490 Ma) origin of land plants compatible with molecular timescales assuming moss-like silicate weathering rates and 15% of current landscape vegetation can explain the otherwise paradoxical end-Ordovician glaciations. These glaciations took place after a period of elevated atmospheric CO₂ (14–22 times current levels) and high global temperatures, but the impact of the bryophyte-like physiology assumed for early land plants is sufficient to replicate the geochemical record of reducing global temperatures and atmospheric CO₂ levels sufficiently low for glaciation. Modelling has also causally implicated this early rise of land plants in the establishment of modern levels of atmospheric O₂ levels, principally through increased flux via silicate weathering of phosphorous to the oceans. This resulted in marine anoxia and organic carbon burial on geologic timescales in the form of the black shales that dominate middle and upper Ordovician strata.

The revised understanding of land plant relationships allows for a more tracheophyte-like ancestral embryophyte, as well as the contemporaneous existence of both stem-bryophytes and -tracheophytes. It would be interesting to explore the impact of these traits on the model which assumed a moss-like ancestral embryophyte; analyses using a liverwort proxy-ancestor have contested the geoengineering capabilities of the earliest land plants. In the interim, existing sensitivity tests imply that a more tracheophyte-like ancestral land plant would result in the same or greater effect as a moss-proxy, but with lower vegetation cover and/or an even later origin of crown-embryophytes. Conversely, new insights into the evolution of rooting systems imply that the impact of crown-tracheophytes on the Earth system would not have been as great as previously envisaged, at least not in terms of the origin of the clade itself. This is because the earliest crown-tracheophytes retained the same rhizoid-based rooting systems inherited from the ancestral embryophyte; branching rooting systems did not evolve until later, independently within lycophytes and euphyllophytes, and it is this phenomenon that is manifest in the fossil and geochemical records. It should not be surprising, therefore, that there is a temporal disconnect both in terms of the appearance of the first tracheophytes and their impact on the Earth system, whether that is read in terms of the fossil record or molecular timescales.

Outstanding problems

Recent advances in the availability of genome scale data from a more representative taxonomic diversity of embryophytes, as well as improved phylogenetic modelling, analyses of developmental evolution and the fossil record, have precipitated a minor revolution in our understanding of early land plant evolution. It has helped to resolve some long-standing uncertainties in our understanding of land plant bodyplan evolution, but also introduced new uncertainties, such as the question of whether the last common ancestor of living embryophytes was gametophyte- or sporophyte-dominant or, indeed, whether it was isomorphic. Here a fossil record of stem-embryophytes, -bryophytes and -tracheophytes would help but, first, more credible attempts are needed to reconcile the fossil record with molecular phylogenies of their living relatives. That is going to require a better understanding of anatomical evolution among living embryophytes and those data are now perhaps harder to assemble, due to a dearth of specialists, than are genome scale datasets for the most obscure of bryophytes. Nevertheless, a framework of anatomical evolution is integral to elucidating the evolutionary significance of the fossil early land plants exquisitely preserved to a cellular level of resolution through charcoalification and silification, and new technologies are required to fully characterise these fossils. Ultimately, however, it could be argued that our understanding of early land plant evolution effectively starts at the end of the story, with the embryophyte crown-ancestor from which bryophytes and tracheophytes emerged. Perhaps the most fundamental insights into the origin of land plants are to be leveraged from genomic and developmental analyses of their algal relatives as well as the identification of fossil stem-embryophytes – some of which may already lie in existing collections, awaiting the correct search image for their discovery.

ACKNOWLEDGEMENTS

We are very grateful to Alex Bowies who provided comments on the text and Dianne Edwards, Sandy Hetherington, Silvia Pressel and Charles Wellman for kindly providing images used in this paper. P.C.J.D. is funded by BBSRC (BB/N000919/1; BB/T012773/1), Leverhulme Trust, NERC (NE/P013678/1), and the Moore Foundation; C.J.H. is funded by the Leverhulme Trust (RPG 2011-2014); Leverhulme Trust, NERC (NE/P013678/1), and the Moore Foundation; C.J.H. is funded by the Leverhulme Trust (RPG 2011-2014); J.P. is funded by the Wellcome Trust.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES


