Reconstructing trait evolution in plant evo–devo studies

Pierre-Marc Delaux, Alexander Hetherington, Yoan Coudert, Charles Delwiche, Christophe Dunand, Sven Gould, Paul Kenrick, Fay-Wei Li, Herve Philippe, Stefan Rensing, et al.

Reconstructing trait evolution: A guideline for plant evo-devo studies (and beyond)

Pierre-Marc Delaux[1]*, Alexander J. Hetherington[2], Yoan Coudert[3], Charles Delwiche[4], Christophe Dunand[1], Sven Gould[5], Paul Kenrick[6], Fay-Wei Li[7,8], Hervé Philippe[9,10], Stefan A. Rensing[11,12,13], Mélanie Rich[1], Christine Strullu-Derrien[6,14], Jan de Vries[15,16,17]

[1]Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, CNRS, UPS, Castanet-Tolosan, France.

[2]Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, UK.

[3]Laboratoire Reproduction et Développement des Plantes, Ecole Normale Supérieure de Lyon, CNRS, INRA, Université Claude Bernard Lyon 1, INRIA, 46 Allée d'Italie, Lyon, 69007, France.

[4]CBMG, University of Maryland, College Park, MD 20742, USA

[5]Institute for Molecular Evolution, Heinrich Heine University, 40225 Düsseldorf, Germany

[6]Department of Earth Sciences, The Natural History Museum, Cromwell Road, London, SW7 5BD, UK.

[7]Boyce Thompson Institute, Ithaca, NY, USA.

[8]Plant Biology Section, Cornell University, Ithaca, NY, USA.

[9]Centre de Théorisation et de Modélisation de la Biodiversité, Station d'Écologie Théorique et Expérimentale, UMR CNRS 5321, Moulis, France.

[10]Département de Biochimie, Université de Montréal, Montréal, Québec, Canada.

[11]Plant Cell Biology, Faculty of Biology, University of Marburg, 35043 Marburg, Germany

[12]BIOSS Centre for Biological Signalling Studies, University Freiburg, Germany

[13]SYNMIKRO Research Center, University of Marburg, 35043 Marburg, Germany

[14]Institut de Systématique, Évolution, Biodiversité, UMR 7205, Muséum National d'Histoire Naturelle, Paris, France.

[15]Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada

[16]Institute of Microbiology, Technische Universitaet Braunschweig, 38106 Braunschweig, Germany

[17]Institute for Microbiology and Genetics, Bioinformatics, University of Göttingen, Goldschmidtstr. 1, 37077 Göttingen, Germany

* Correspondence: pierre-marc.delaux@lrsv.ups-tlse.fr

Abstract

Our planet is teeming with an astounding diversity of plants. In a mere single group of closely related species tremendous diversity can be observed in their form and function: the colour of petals in flowering plants, the shape of the fronds in ferns, and the branching pattern of the gametophyte in mosses. Diversity can equally be found in subtler traits such as the resistance to pathogens or the ability to recruit symbiotic microbes from the environment. Plant traits can also be extremely conserved: at the cellular and metabolic levels, entire biosynthetic pathways are present in all plant groups, and morphological characteristics such as vascular tissues have been conserved for hundreds of million so years. The research community that seeks to understand these traits –both the diverse and the conserved– by taking an evolutionary point of view on plant biology is growing. Here, we summarize a subset of the different aspects of plant evolutionary biology, provide a guide for structuring comparative biology approaches and discuss the pitfalls that (plant) researchers should avoid when embarking on such studies.

Plants are extremely diverse, whether this be in the range of petal colours in angiosperms [1], the shape of the fronds in ferns [2], the branching pattern of the gametophyte in mosses [3] or their interactions with microbes and the environment. To understand this diversity, it is essential to explore the genetic framework underlying any of these traits in light of evolution. Researchers studying the evolution of traits aim at determining "*which genes and what kinds of changes in their sequences are responsible for the evolution of [morphological] diversity*" [4]. This way of approaching diversity was initiated by developmental biologists leading to the emergence of evolutionary developmental biology (evo-devo) that was later expanded to all aspects of plant biology such as the interactions between plants and microbes [5,6] (evo-MPMI, for evolutionary molecular plant-microbe interactions) or the study of cellular biology [7] (evo-cell biology). The terms are different but these fields of research rely on the same comparative approaches to characterise trait evolution; they, hence, face similar challenges.

The research community taking an evolutionary point of view on plant biology is growing. There are a number of reasons for this. A foremost reason is the availability of whole plant genomes and transcriptomes covering the breadth of the plant phylogeny. This goes hand in hand with the development of model systems beyond the flowering plant *Arabidopsis thaliana* (L.) Heynh. Such prospering models include a huge diversity of angiosperm species, from trees such as *Populus trichocarpa*, to a range of species to study flower development (e.g. *Antirrhinum majus* and *Aquilegia caerulea*), together with enormous progress with gymnosperms including the first genome assemblies and studies of gene expression and conservation. Major advances have been also made with lycophytes or monilophytes such as the model fern *Ceratopteris richardii*, and finally the development of liverwort, moss as well as algal model systems. The liverwort *Marchantia polymorpha* is establishing itself as a major model in plant science, mainly because it has a short life-cycle and a relatively small genome with fewer *paralogs* (see Glossary box) than most other land plant species [8]. Similar to the model moss *Physcomitrella patens*—which has been extensively used throughout the last two decades [9]—*M. polymorpha* is genetically tractable [10]. The diversification of model species offers unprecedented opportunities to explore fundamental biological processes. Furthermore, comparative studies with other plant clades have the potential to unravel evolutionary events that shaped the diversity of extant plant species. The access to many plant genomes, transcriptomes and new model species therefore makes it an exciting time for studying plant evolution. However, as the number of model species diversifies it is important to pay close attention to the method used to draw evolutionary comparisons between species. This becomes particular important when drawing conclusions with the benchmark of *A. thaliana,* when the species last shared a common ancestor with this Brassicaceae hundreds of millions of years ago.

The aim of this article is to provide a guide for structuring the rationale that is used when employing comparative biology approaches, starting first by highlighting the importance of drawing conclusions based on precisely reconstructed species phylogenies before subsequently outlining a 5-step guide for structuring evolutionary studies, illustrated with examples from the literature.

Species phylogenies and tree thinking

Evolutionary relationships between organisms are best expressed through phylogenetic trees. The use of DNA sequences to reconstruct phylogenies, particularly multi-loci phylogenomics, has improved the resolution of the plant tree of life (Figure 1). It is now widely accepted that the closest *extant* relatives of land plants are the streptophyte algae in the Zygnematophyceae class [11–13]. Within the land plants, the relationship among the major lineages is relatively well supported (Figure 1). Living land plants constitute two main groups, vascular and non-vascular plants. Vascular plants (tracheophytes) form a *monophyletic* group encompassing the lycophytes, ferns and seed plants. Uncertainty still remains regarding the branching order of the non-vascular plants, the bryophytes (hornworts, liverworts and mosses), with three main hypotheses equally well supported [11,12,14] (Figure 1a-c).

One requirement for the study of evolution is the ability to navigate phylogenies. Rooted phylogenetic trees contain deep and shallow branches. In addition, for practical reasons trees often include more species that are closely related to the focal organisms – which is referred to as selective, or biased, taxon sampling– often resulting in depictions that can resemble ladders with certain clades of organisms of interest (*e.g.* humans, angiosperms) at the “top” of the trees. However, phylogenetic trees can be rotated at any node without changing their evolutionary meaning (Figure 1a, d, and e). Consequently, to put trait evolution into an evolutionary context, proper tree thinking —which is not intuitive— is required [15].

While different extant organisms might share a more recent or distant common ancestor, they have all been subject to evolutionary changes. In the case of land plant evolution, for example, this means that the angiosperms are not “higher plants” and, in turn, bryophytes are not “lower”, “basal”, or “primitive” plants. The best practice, hence, is to refer to any given organism by the name of the group to which it belongs. *M. polymorpha*, *P. patens*, and any other extant bryophyte are as “evolved” as any other plant that is living today. Extant bryophytes and angiosperms are equally divergent from the most recent common ancestor (MRCA) of all land plants [12]. Similarly, any trait present if the MRCA of Zygnematophyceae and land plants has experienced an equal opportunity for divergence in each lineage since the time they derived from their MRCA. The fact that every organism is composed of ancestral and derived traits — reflected by independent gains and losses of genomic parts, expansions,

diversifications and genome rearrangements — has long been discussed as ‘mosaic evolution’ or *heterobathmy* [16].

Plant evolutionary biology studies therefore first involve the investigation of the species of interest and where they fall on a phylogenetic tree, keeping in mind that all extant lineages have evolved from an ancient MRCA.

A guideline for plant evolutionary biology studies

Step 1: Inferring trait evolution

Any study on the evolution of form and function (traits) should start by precisely defining the trait of interest. Given that the same term can be used by different authors or by different fields to mean different things it is essential to be explicit about the definition of the trait of interest. With all the diversity of the species investigated and their bouquet of traits brought to the table it might be necessary to clearly define the trait of interest—e.g., what is meant by broad terms such as ‘multicellular’ or ‘complex body plan’? It further is essential to bear in mind that over broad evolutionary timescales, ancestral traits derive independently in different lineages and, while being homologous, may result in completely different forms, thus obscuring homology at a first glance. For instance, the 3D-leaves of succulent plants are homologous to the leaves of all euphyllophytes but display a completely different, tube-like, structure at the macroscopic level [17]. Reversely, convergent evolution or *homoplasy* may lead to similar but yet not homologous structures such as the megaphyllous leaves in euphyllophytes and microphyllous leaves in lycophytes [18].

Once defined, one can map the trait of interest onto a species tree that captures a range of organisms salient to the question. Mapping means scoring all (or as many as possible) species in the tree for the state of the trait (typically presence or absence). The range of organisms is defined by starting with a focal species harboring the trait. This is key to applying the comparative method: given a rooted phylogeny, in an iterative process related species have to be investigated —from closely to distantly related. Eventually, a deep-enough node in the phylogeny will be reached that includes species lacking the trait. For instance, to study the evolution of flower development, one may start with any model angiosperms, such as Arabidopsis or Antirrhinum and expand the sampling to the entire angiosperm clade. To be informative, comparative analyses must include species that fall beyond the clade of interest (the *ingroup*), forming the *outgroup*. For example to define the origin of the flower, a *synapomorphy* of angiosperms (the ingroup), one can choose to compare with the gymnosperms (the outgroup) that lack flowers. By mapping a character of interest onto a species tree of both the in and outgroup it is possible to define the evolution of a trait based on extant species.

A trait mapped onto a phylogeny only tells us about the presence of the trait in the living species. However, it does not reveal how the trait evolved in the past. To do this we must infer ancestral character states. An ancestral character states is the definition of a character at a node within the tree rather than in living species (which sit on the tips, *i.e.* "leaves", of the tree). Inferring ancestral character states hinges on the usage of the appropriate statistical methods (Figure 2) [19,20]. For instance, combining three different methods (maximum parsimony, maximum likelihood and Bayesian approaches) on a database listing flower morphologies of 792 extant flowering plant species Sauquet *et al.* recently inferred a morphology for the common ancestor of all extant flowers: bisexual and radially symmetric [21]. In a different context, Werner *et al.* [22] inferred losses and gains of the arbuscular mycorrhizal (AM) symbiosis formed by most land plants with members of the Glomeromycotina fungi, a mutualistic symbiosis that improves plant nutrition [23]. Using a database of 3,736 seed plant species they predicted that AM symbiosis was present in the MRCA of seed plants and was lost multiple times during seed plant evolution [22]. Importantly, they compared their inferences of AM symbiosis losses with the evolution of alternative nutrient-uptake strategies and discovered a strong correlation [22]. Such inferences are the basis for comparative studies and in most cases are directly followed by step 2 (see below). However, an ancestral state reconstruction is a prediction that can only be tested by looking for direct or indirect evidence of the character of interest in the past. When available, the fossil record is therefore essential for testing predictions of ancestral character states and fossils can be integrated with well-supported phylogenies to improve such inferences [24]. DNA data are never available from ancient fossils, but various forms of information about the fossils can be integrated to understand trait evolution. Although fossils are not always straightforward to interpret and may represent derived traits of extinct lineages, integration of fossil data may lead to strong reinterpretation of predicted ancestral states, such as in the case of flowers [25]. Here, we will provide two examples: first from the evolution of plant-fungal symbiosis and second from the anatomical evolution of rooting structures.

Most extant land plants (85 %) form AM symbiosis which likely evolved in the MRCA of land plants, as proposed by phylogenetic inferences [22,26]. However, members of two bryophytes lineages (*i.e.,* liverworts and hornworts) as well as some vascular plants (*i.e.,* lycopods and ferns) [27] can develop endosymbiotic associations with not just the Glomeromycotina but also the Mucoromycotina, sometimes simultaneously [28]. Based on this observation, it can be proposed that dual colonization might have been a trait present in the earliest land plants. This prediction can be tested by examining fossil plants. Fossil sites such as the 407 million year old Rhynie chert, approaching the predicted age of the first land plants that originated c. 515-475 million years ago, provides a window into the early evolution of land plants [29]. Two exceptionally preserved fossils of early land plant described from the

Rhynie chert, *Aglaophyton majus* (a non-vascular plant [30]) and *Horneophyton lignieri* (a vascular plant [31]) show hallmarks of symbioses with both Glomeromycotina and Mucoromycotina, thus validating the ancestral state inference [32–34]. These fossils, when examined within a phylogenetic framework, allow us to confirm that early non-vascular and vascular species developed endosymbioses with one or both groups of fungi [35,36].

Almost all extant tracheophytes develop specialized rooting organs termed 'true' roots that develop from a root meristem with a root cap [37]. Hence, the prediction based on living species is that the common ancestor of vascular plants possessed a true root. To test this hypothesis, we can again call upon the Rhynie chert plants, whose exceptional preservation provides numerous anatomical characters that allow for confident placement of these species on a phylogeny of land plants [30,38]. It was found that species in the Rhynie chert spanned the origin of the vascular lineage and the majority of species developed rhizoid (filamentous outgrowth) based rooting systems [39–41]—a finding at odds with the prediction that the common ancestor of vascular plants developed a true root. This suggests that roots of extant vascular plants do not originate from a shared common ancestor, *i.e.* are not homologous but in fact evolved independently at least twice [40,42,43]—making their similarities in anatomy the product of convergent evolution. This theory was cemented with the examination of a rooting structure from another Rhynie chert fossil, *Asteroxylon mackiei*. The meristem of the rooting axes of this lycophyte lacks root caps and therefore displays a transitional suite of characters with some but not all characters of extant plant roots [44]. This extinct transitional stage sheds light on gradual character evolution leading to the roots of extant plants.

The aim of the first step is to infer the trait evolution, mostly relying on extant species. Fossils are not essential but, when available, strongly improve these inferences.

Step 2: Reconstructing the evolution of genes associated with the trait of interest.

More than two decades of genetics in model angiosperms such as *A. thaliana* have established causal links between traits and genes, thus allowing insights to be gained on the genetic and biochemical levels. Studying the evolution of these well-described genes by phylogenetic inference opens the door towards a better understanding of the evolution of the traits and the formulation of working hypotheses.

In step 1 we have outlined how to work with a given species phylogeny. Through a mapping approach, traits were projected onto the species phylogeny and ancestral character states of the MRCAs of the species in that phylogeny were reconstructed. Here, we will hone in on gene phylogenies to understand the evolution of the genetics that underpin these traits. While the evolutionary history of gene families is tied to vertical evolution of the species, additional genetic processes (gene duplication, loss, and—occasionally— horizontal gain) add an additional layer of complexity to this. For instance, the scattered distribution of a trait may

be the result of convergent gain (*homoplasy*) or convergent losses such as complex leaves in the Brassicaceae [45]. Resolving the phylogeny of *REDUCED COMPLEXITY* (*RCO*), a gene known to regulate complex leaf morphology in *Cardamine hirsute* [46], identified multiple independent losses in species with simpler leaves, indicative of convergent losses of the trait [45]. Another example is the well characterized resistance to phytophagous insects in ferns. Phylogeny of the candidate gene, *Tma12,* indicated its likely horizontal gene transfer from bacteria, providing a putative scenario for the origin of this trait in ferns [47].

To be fully informative, phylogenetic tree reconstruction must be conducted on an appropriately curated dataset covering the entire species-space of interest defined in Step 1. Indeed, restricting the analysis to a too narrow set of species may yield misleading results. For instance, using the advent of more genomes and transcriptome sequence of Zygnematophyceae it has recently been demonstrated that many of the genes once considered to be typical of or even unique to land plants are in fact present in the algal *sister group* of land plants [26,48–51].

A number of DNA sequence databases are now available for plant biologists. These include Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) that hosts a select set of well-curated genomes, and the 1KP database (https://db.cngb.org/onekp/) that includes RNA-Seq data from a range of land plant and algal sequences many of which still lack genomes. A database covering both species diversity and sequencing depth (*i.e.* genome and not transcriptomes) should result from the recently launched 10KP initiative that aims at sequencing all plant genera [52].

After identifying homologs of the target genes by similarity searches (such as BLAST), orthology is determined using phylogenetic analyses [53]. Two genes are orthologs if they have been vertically inherited from a common ancestor only through speciation, in other words if the gene phylogeny matches the species phylogeny. Orthology is often complicated by *paralogy* (duplicated genes), but phylogenetic inference helps to define ancestral states of gene family and sub-family evolution [54].

The aim of step 2 is, once a gene of interest has been identified, to reconstruct the evolutionary history of the gene based on phylogenetic analysis –making use of the extensive sequence databases. Ultimately, comparisons between species phylogenies and gene phylogenies allows for the correlation of genotype and phenotype and the formulation of hypotheses. It is prudent to note that testing such hypotheses that hinge on correlative predictions require functional validation, which are being explored in the next two steps.

Step 3: Determining the evolution of biochemical properties.

Orthology is a statement about linear descent that is inferred from phylogenies. That means that it does not have to coincide with (completely) conserved biochemical properties. This assumption needs to be experimentally tested.

Many genes have been functionally characterized in the flowering plant *A. thaliana*. If phylogenetic inference (Step 2) shows that there are orthologous sequences of a gene in other plants, one can derive and test the hypothesis that such genes have a conserved molecular function –a concept sometimes referred to as functional orthologs. Molecular function encompasses all the features that are important for the protein action such as protein-protein interactions, enzymatic activities, DNA-binding abilities or regulation sites (such as phosphorylation). In the hypothetical case where all these functions are known, *in vitro* assays can be conducted. However, a more comprehensive approach to test all these features at once is an inter-species complementation assay [55]. If the gene of one species can complement the loss-of-function mutation of an ortholog in another species it indicates conservation of its molecular function. There are many examples of deep conservation, between the flowering plant *A. thaliana*, the bryophytes *P. patens* or *M. polymorpha* [51,56] and streptophyte and chlorophyte algae [26,51]. For instance, complementation assays of the *Medicago truncatula dmi3* mutant (which is unable to form symbiosis with Glomeromycotina fungi) has been conducted with *DMI3* orthologs from liverworts, hornworts [57], Zygnematophyceae, Chlorokybophyceae and chlorophytes [26]. While liverwort, hornwort, Zygnematophyceae and Chlorokybophyceae orthologs were able to complement the *M. truncatula dmi3* mutant, the chlorophyte orthologs failed to do so [26,57]. This indicates that the molecular functions of DMI3 are not completely conserved between the chlorophyte and the streptophyte orthologs, suggesting the gain of a new function in the MRCA of the streptophytes [26]. Another example is the floral regulator *LEAFY*, which evolved novel functions apparent in seed plants that are not shared with *e.g.* mosses or the streptophyte alga *Klebsormidium nitens* [58]. Note that functional data do not impact orthology of genes/proteins. Orthologs are orthologous to one another, no matter whether shown to be functionally conserved or divergent.

In step 3, we suggest to study the evolution of biochemical properties by conducting inter-species complementation assays of knock out mutants with orthologs from multiple plant clades.

Step 4: Determining the evolution of a gene's biological role.

Conserved biochemical properties of a protein does not mean that its biological role itself is conserved. Testing the conservation of a gene's biological role is typically performed by generating mutants that are devoid of the function of the respective gene (often known as reverse genetics) in multiple species. If knock out mutants of an orthologous gene in two

species display similar defects, it can be hypothesized that the biological role of that gene has been conserved since the divergence of these two species, *i.e.* since their MRCA lived. As an example, the function of ABI3 as a key transcription factor controlling the response to dehydration has been found in both *A. thaliana* and *P. patens* suggesting its conservation at least since the MRCA of land plants [59]. Another example, where such analyses have been extended to a broader range of species, is the conserved function of a clade of bHLH transcription factors in the formation of cells with rooting function in dicots, monocots, mosses and liverworts, a clear example of *deep homology* [60–64].

When inferring a gene's biological role using comparisons between multiple species it is important to keep in mind that the power of the comparative analysis is limited by the number of species sampled. For instance, if a biological role is not conserved between two species it may either be that the role was not present in their MRCA or alternatively that the biological role was present in the MRCA but has subsequently been lost in one or the other species. To distinguish between these two scenarios it is of crucial importance to add more lineages, represented by emerging model species, to the equation [65]. For instance, the transcription factor LEAFY has been studied in a dozen eudicots, two monocots, *P. patens* and more recently the fern *C. richardii* [66]. A conserved defect in *leafy* loss of function mutants in both *P. patens* and *C. richardii* is the absence of division in the sporophyte zygote—a defect not found in angiosperm mutants [66]. Because this role is found in mosses and ferns, it can be proposed as the ancestral state (since one loss in seed plants is more parsimonious and likely than two independent gains in mosses and ferns). Although previously proposed [67], experimentally determining whether that role is ancestral (lost or reduced in angiosperms) or derived (gain in mosses) with only *P. patens* and angiosperm mutants was almost impossible. Obviously, such inferences will become stronger as model systems become available from other major clades, such as the Zygnematophyceae [65,68].

Besides the limited number of model species, a common limitation of reverse genetic approaches is redundancy, often caused by a close homolog or a paralog. Because gene duplications are very common in plant genomes, it remains difficult to reject the presence of recent, species-specific, paralogs. Thus, absence of phenotypes must be taken with caution. In addition, following duplication, subfunctionalization may occur differently between the paralogs in different species. For instance, in Arabidopsis *AGAMOUS* regulates the specification of reproductive organs while this function is taken over by its paralog, *PLENA*, in *A. majus* [69].

In step 4, the biological roles of orthologous genes are explored in phylogenetically diverse model plants using reverse genetics.

Step 5: Synthesizing the molecular evolution of the trait of interest.

Projecting gene phylogeny, conservation of the biochemical properties, and conservation of the biological role onto the species phylogenetic tree provides a means of unravelling the evolutionary mechanisms that shaped the trait of interest. An elegant example is the evolution of the jasmonate (JA) receptor COI1.

Jasmonates are a class of plant hormones that play multiple essential roles in plants [70]. In tracheophyte COI1 perceives the conjugated form of JA (JA-Ile), while the bryophyte version does not [71]. So, what is COI1's story? The trait here is the perception of JA-Ile, which is found in tracheophytes but not in any of the bryophytes that the author tested, including *P. patens*, *M. polymorpha*, and the hornworts *Anthoceros agrestis* (Step 1). Phylogenetic analyses identified a clear *COI1* ortholog in all investigated land plants (Step 2). However, the protein from bryophytes does not allow JA-Ile perception (Step 3), indicating non-conserved biochemical properties. A single mutation discriminating COI1 from tracheophytes from those in bryophytes was identified [71]. Importantly, the authors included lycophytes in the sequence comparison, pinpointing the amino-acid switch at the origin of the tracheophytes. While the metabolic precursor of JA-Ile, OPDA, inhibits the growth of *M. polymorpha* and other bryophytes, the *M. polymorpha coi1* mutant is insensitive to that molecule. Using a biochemical screening, the authors [71] showed that the ligand of the *M. polymorpha* COI1 receptor is dinor-OPDA, another molecule derived from OPDA. Phenotypically, the *coi1* mutant was more sensitive to a generalist herbivore, suggesting a biological role reminiscent to the one in angiosperms despite differences in the biochemical function (Step 4). Altogether, these data suggest that COI1 functioned in the perception of OPDA-derived molecules in the MRCA of land plants. In tracheophytes, possibly a single point mutation led to the capacity of COI1 to perceive JA-Ile. The signaling pathways downstream of COI1 have likely been conserved since the land plant MRCA. To test this hypothesis, the authors [71] mimicked evolution by replacing in the *M. polymorpha COI1* sequence the codon discriminating bryophytes and tracheophytes with the tracheophytes version. This was sufficient to make *M. polymorpha* sensitive to JA-Ile, thus validating the proposed scenario [71].

All the previous steps are combined in step 5, hereby producing a holistic understanding of molecular evolution of the trait of interest. Such inference is valid repending on available data. Experimental validations in additional species and the sequencing of more genomes may lead to refined likely scenarios. Importantly, 'likely scenario' is the appropriate wording here—it conveys the due humbleness. Since we can only investigate extant species at the genetic level, we will always capture a mere snapshot of the phenotypes that evolution has given rise to.

Concluding remarks

Plant science experiences a diversification of species accessible for genetic work. Comparisons of functional data gathered in model angiosperms, mosses, liverworts as well as multiple representatives of the major plant clades will allow the plant community to paint a more comprehensive picture of the evolutionary mechanisms that have led to the diversity observed in extant species. Reporting such findings requires appropriate methodology and common terminology that we describe in this commentary.

Acknowledgements

The authors would like to thank Prof. Liam Dolan (University of Oxford) for his critical reading and suggestions on earlier drafts of the manuscript. P-MD, MR and CD are part of the LRSV laboratory, which belongs to the TULIP Laboratoire d'Excellence (LABEX) (ANR-10-LABX-41). P-MD and MR are funded by the ANR grant EVOLSYM (ANR-17-CE20-0006-01). JdV gratefully acknowledges funding through a Research Fellowship (VR132/1-1) and Return Grant (VR132/3-1) awarded by the German Research Foundation (DFG). YC acknowledges funding from the French National Center for Scientific Research (ATIP-AVENIR research fellowship) and the University of Lyon (IMPULSION grant). AJH was funded by the George Grosvenor Freeman Fellowship by Examination in Sciences, Magdalen College (Oxford). CFD is supported in part by the Maryland Agricultural Experiment Station (MAES). F-WL is supported by the National Science Foundation grant DEB-1831428.

References

1. Moyroud, E., and Glover, B.J. (2017). The Evolution of Diverse Floral Morphologies. Curr. Biol. *27*, R941–R951.
2. Vasco, A., Moran, R.C., and Ambrose, B.A. (2013). The evolution, morphology, and development of fern leaves. Front. Plant Sci. *4*, 345.
3. Coudert, Y., Bell, N.E., Edelin, C., and Harrison, C.J. (2017). Multiple innovations underpinned branching form diversification in mosses. New Phytol. *215*, 840–850.
4. Carroll, S.B. (2008). Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. Cell *134*, 25–36.
5. Rich, M., and Delaux, P.-M. (2018). Taking the step: from Evo-Devo to plant-microbe interaction evolution with the liverwort Marchantia. New Phytol. *218*,882-884.
6. Upson, J.L., Zess, E.K., Białas, A., Wu, C. hang, and Kamoun, S. (2018). The coming of age of EvoMPMI: evolutionary molecular plant–microbe interactions across multiple timescales. Curr. Opin. Plant Biol. *44*,108-116
7. Lynch, M., Field, M.C., Goodson, H. V., Malik, H.S., Pereira-Leal, J.B., Roos, D.S., Turkewitz, A.P., and Sazer, S. (2014). Evolutionary cell biology: Two origins, one

objective. Proc. Natl. Acad. Sci. *111*, 16990-16994.

8. Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., *et al.* (2017). Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. Cell *171*, 287-304.e15.
9. Lang, D., Ullrich, K.K., Murat, F., Fuchs, J., Jenkins, J., Haas, F.B., Piednoel, M., Gundlach, H., Van Bel, M., Meyberg, R., *et al.* (2018). The Physcomitrella patens chromosome-scale assembly reveals moss genome structure and evolution. Plant J. *93*, 515-533.
10. Ishizaki, K., Nishihama, R., Yamato, K.T., and Kohchi, T. (2016). Molecular genetic tools and techniques for marchantia polymorpha research. Plant Cell Physiol. *57*, 262–270.
11. Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., Ayyampalayam, S., Barker, M.S., Burleigh, J.G., Gitzendanner, M.A., *et al.* (2014). Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc. Natl. Acad. Sci. *111*, E4859–E4868.
12. Puttick, M.N., Morris, J.L., Williams, T.A., Cox, C.J., Edwards, D., Kenrick, P., Pressel, S., Wellman, C.H., Schneider, H., Pisani, D., *et al.* (2018). The Interrelationships of Land Plants and the Nature of the Ancestral Embryophyte. Curr. Biol. *28*, 733-745.e2.
13. Timme, R.E., Bachvaroff, T.R., and Delwiche, C.F. (2012). Broad phylogenomic sampling and the sister lineage of land plants. PLoS One *7*, e29696.
14. de Sousa, F., Foster, P.G., Donoghue, P.C.J., Schneider, H., and Cox, C.J. (2018). Nuclear protein phylogenies support the monophyly of the three bryophyte groups (Bryophyta Schimp.). New Phytol. *222*, 565-575.
15. Baum, D.A., Smith, S.D., and Donovan, S.S.S. (2005). Evolution. The tree-thinking challenge. Science. *310*, 979–980.
16. Hennig, W. (1950). Grundzüge einer Theorie der phylogenetischen Systematik (Deutscher Zentralverlag).
17. Ogburn, R.M., and Edwards, E.J. (2013). Repeated origin of three-dimensional leaf venation releases constraints on the evolution of succulence in plants. Curr. Biol. *23*, 722–726.
18. Harrison, C.J., Coriey, S.B., Moylan, E.C., Alexander, D.L., Scotland, R.W., and Langdale, J.A. (2005). Independent recruitment of a conserved developmental mechanism during leaf evolution. Nature. *434*, 509-514.
19. Ronquist, F. (2004). Bayesian inference of character evolution. Trends Ecol. Evol. *19*, 475-81.
20. Nunn, C.L. (2011). The comparative approach in evolutionary anthropology and biology (University of Chicago Press)

21. Sauquet, H., von Balthazar, M., Magallón, S., Doyle, J.A., Endress, P.K., Bailes, E.J., Barroso de Morais, E., Bull-Hereñu, K., Carrive, L., Chartier, M., *et al.* (2017). The ancestral flower of angiosperms and its early diversification. Nat. Commun. *8*, 16047.
22. Werner, G.D.A., Cornelissen, J.H.C., Cornwell, W.K., Soudzilovskaia, N.A., Kattge, J., West, S.A., and Kiers, E.T. (2018). Symbiont switching and alternative resource acquisition strategies drive mutualism breakdown. Proc. Natl. Acad. Sci. *115*, 5229–5234.
23. Vigneron, N., Radhakrishnan, G. V, and Delaux, P.-M. (2018). What have we learnt from studying the evolution of the arbuscular mycorrhizal symbiosis? Curr. Opin. Plant Biol. *44*, 49–56.
24. Betts, H.C., Puttick, M.N., Clark, J.W., Williams, T.A., Donoghue, P.C.J., and Pisani, D. (2018). Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. Nat. Ecol. Evol. *2*, 1556–1562.
25. Rümpler, F., and Theißen, G. (2019). Reconstructing the ancestral flower of extant angiosperms: the "war of the whorls" is heating up. J. Exp. Bot. *70*, 2615-2622.
26. Delaux, P.-M., Radhakrishnan, G. V., Jayaraman, D., Cheema, J., Malbreil, M., Volkening, J.D., Sekimoto, H., Nishiyama, T., Melkonian, M., Pokorny, L., *et al.* (2015). Algal ancestor of land plants was preadapted for symbiosis. Proc. Natl. Acad. Sci. *112*, 13390–13395.
27. Rimington, W.R., Pressel, S., Duckett, J.G., and Bidartondo, M.I. (2015). Fungal associations of basal vascular plants: Reopening a closed book? New Phytol. *205*, 1394-1398.
28. Strullu-Derrien, C., Selosse, M.-A., Kenrick, P., and Martin, F.M. (2018). The origin and evolution of mycorrhizal symbioses: from palaeomycology to phylogenomics. New Phytol. *220*, 1012-1030.
29. Edwards, D., Kenrick, P., and Dolan, L. (2018). History and contemporary significance of the Rhynie cherts-our earliest preserved terrestrial ecosystem. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *373*, 20160489.
30. Kenrick, P., and Crane, P.R. (1997). The origin and early evolution of plants on land. Nature *389*, 33–39.
31. Cascales-Miñana, B., Steemans, P., Servais, T., Lepot, K., and Gerrienne, P. (2019). An alternative model for the earliest evolution of vascular plants. Lethaia. *52*, 445-453.
32. Remy, W., Taylor, T.N., Hass, H., and Kerp, H. (1994). Four hundred-million-year-old vesicular arbuscular mycorrhizae. Proc. Natl. Acad. Sci. *91*, 11841–11843.
33. Taylor, T.N., Remy, W., Hass, H., and Kerp, H. (1995). Fossil Arbuscular Mycorrhizae from the Early Devonian. Mycologia *87*, 560–573.
34. Strullu-Derrien, C., Kenrick, P., Pressel, S., Duckett, J.G., Rioult, J.P., and Strullu,

D.G. (2014). Fungal associations in Horneophyton ligneri from the Rhynie Chert (c. 407 million year old) closely resemble those in extant lower land plants: Novel insights into ancestral plant-fungus symbioses. New Phytol. *203*, 964–979.

35. Bidartondo, M.I., Read, D.J., Trappe, J.M., Merckx, V., Ligrone, R., and Duckett, J.G. (2011). The dawn of symbiosis between plants and fungi. Biol. Lett. *7*, 574–577.
36. Feijen, F.A.A., Vos, R.A., Nuytinck, J., and Merckx, V.S.F.T. (2018). Evolutionary dynamics of mycorrhizal symbiosis in land plant diversification. Sci. Rep. *8*, 10698.
37. Foster, A.S. (Adriance S., and Gifford, E.M. (1974). Comparative morphology of vascular plants (W.H. Freeman).
38. Kenrick, P.R., and Crane, P.R. (1997). The origin and early diversification of land plants: a cladistic study. Nature *389*, 33–39.
39. Edwards, D. (2003). Embryophytic sporophytes in the Rhynie and Windyfield cherts. Trans. R. Soc. Edinb. Earth Sci. *94*, 397–410.
40. Kenrick, P., and Strullu-Derrien, C. (2014). The Origin and Early Evolution of Roots. Plant Physiol. *166*, 570–580.
41. Hetherington, A.J., and Dolan, L. (2018). Bilaterally symmetric axes with rhizoids composed the rooting structure of the common ancestor of vascular plants. Philos. Trans. R. Soc. B Biol. Sci. *373*.
42. Raven, J.A., and Edwards, D. (2001). Roots: evolutionary origins and biogeochemical significance. J. Exp. Bot. *52*, 381–401.
43. Boyce, C.K. (2010). The evolution of plant development in a paleontological context. Curr. Opin. Plant Biol. *13*, 102–107.
44. Hetherington, A.J., and Dolan, L. (2018). Stepwise and independent origins of roots among land plants. Nature *561*, 235–238.
45. Kiefer, C., Willing, E.-M., Jiao, W.-B., Sun, H., Piednoël, M., Hümann, U., Hartwig, B., Koch, M.A., and Schneeberger, K. (2019). Interspecies association mapping links reduced CG to TG substitution rates to the loss of gene-body methylation. Nat. Plants. 5, 846-855.
46. Vlad, D., Kierzkowski, D., Rast, M.I., Vuolo, F., Dello Ioio, R., Galinha, C., Gan, X., Hajheidari, M., Hay, A., Smith, R.S., *et al.* (2014). Leaf shape evolution through duplication, regulatory diversification and loss of a homeobox gene. Science. *343*, 780–783.
47. Li, F.W., Brouwer, P., Carretero-Paulet, L., Cheng, S., De Vries, J., Delaux, P.M., Eily, A., Koppers, N., Kuo, L.Y., Li, Z., *et al.* (2018). Fern genomes elucidate land plant evolution and cyanobacterial symbioses. Nat. Plants. *4*, 460-472
48. Nishiyama, T., Sakayama, H., de Vries, J., Buschmann, H., Saint-Marcoux, D., Ullrich, K.K., Haas, F.B., Vanderstraeten, L., Becker, D., Lang, D., *et al.* (2018). The Chara

Genome: Secondary Complexity and Implications for Plant Terrestrialization. Cell *174*, 448-464.e24.

49. Wilhelmsson, P.K.I., Mühlich, C., Ullrich, K.K., and Rensing, S.A. (2017). Comprehensive Genome-Wide Classification Reveals That Many Plant-Specific Transcription Factors Evolved in Streptophyte Algae. Genome Biol. Evol. *9*, 3384–3397.
50. de Vries, J., Curtis, B.A., Gould, S.B., and Archibald, J.M. (2018). Embryophyte stress signaling evolved in the algal progenitors of land plants. Proc. Natl. Acad. Sci. *115* E3471-E3480.
51. Ju, C., Van De Poel, B., Cooper, E.D., Thierer, J.H., Gibbons, T.R., Delwiche, C.F., and Chang, C. (2015). Conservation of ethylene as a plant hormone over 450 million years of evolution. Nat. Plants *1*, 14004.
52. Cheng, S., Melkonian, M., Smith, S.A., Brockington, S., Archibald, J.M., Delaux, P.-M., Li, F., Melkonian, B., Mavrodiev, E. V, Sun, W., *et al.* (2018). 10KP: A Phylodiverse Genome Sequencing Plan. Gigascience. 7, 1-9.
53. Mindell, D.P., and Meyer, A. (2001). Homology evolving. Trends Ecol. Evol. *16*, 434–440.
54. Altenhoff, A.M., and Dessimoz, C. (2012). Inferring Orthology and Paralogy. In Methods in molecular biology (Clifton, N.J.), pp. 259–279.
55. Kramer, E.M. (2015). A stranger in a strange land: The utility and interpretation of heterologous expression. Front. Plant Sci. *6*, 734.
56. Rensing, S.A. (2018). Great moments in evolution: the conquest of land by plants. Curr. Opin. Plant Biol. *42*, 49–54.
57. Wang, B., Yeun, L.H., Xue, J.Y., Liu, Y., Ané, J.M., and Qiu, Y.L. (2010). Presence of three mycorrhizal genes in the common ancestor of land plants suggests a key role of mycorrhizas in the colonization of land by plants. New Phytol. *186*, 514–525.
58. Sayou, C., Monniaux, M., Nanao, M.H., Moyroud, E., Brockington, S.F., Thévenon, E., Chahtane, H., Warthmann, N., Melkonian, M., Zhang, Y., *et al.* (2014). A promiscuous intermediate underlies the evolution of LEAFY DNA binding specificity. Science. *343*, 645–648.
59. Khandelwal, A., Cho, S.H., Marella, H., Sakata, Y., Perroud, P.-F., Pan, A., and Quatrano, R.S. (2010). Role of ABA and ABI3 in Desiccation Tolerance. Science. *327*, 546–546.
60. Menand, B., Yi, K., Jouannic, S., Hoffmann, L., Ryan, E., Linstead, P., Schaefer, D.G., and Dolan, L. (2007). An ancient mechanism controls the development of cells with a rooting function in land plants. Science. *316*, 1477–1480.
61. Proust, H., Honkanen, S., Jones, V.A.S., Morieri, G., Prescott, H., Kelly, S., Ishizaki,

K., Kohchi, T., and Dolan, L. (2016). RSL Class i Genes Controlled the Development of Epidermal Structures in the Common Ancestor of Land Plants. Curr. Biol. *26*, 93–99.

62. Kim, C.M., and Dolan, L. (2016). ROOT HAIR DEFECTIVE SIX-LIKE Class I Genes Promote Root Hair Development in the Grass Brachypodium distachyon. PLoS Genet. *12*, e1006211.
63. Kim, C.M., Han, C.D., and Dolan, L. (2017). RSL class I genes positively regulate root hair development in Oryza sativa. New Phytol. *213*, 314-323.
64. Shubin, N., Tabin, C., and Carroll, S. (2009). Deep homology and the origins of evolutionary novelty. Nature *457*, 818–23.
65. Rensing, S.A. (2017). Why we need more non-seed plant models. New Phytol. *216*, 355–360.
66. Plackett, A.R., Conway, S.J., Hewett Hazelton, K.D., Rabbinowitsch, E.H., Langdale, J.A., and Di Stilio, V.S. (2018). LEAFY maintains apical stem cell activity during shoot development in the fern Ceratopteris richardii. Elife *7*.
67. Moyroud, E., Kusters, E., Monniaux, M., Koes, R., and Parcy, F. (2010). LEAFY blossoms. Trends Plant Sci. *15*, 346-52.
68. Chang, C., Bowman, J.L., and Meyerowitz, E.M. (2016). Field Guide to Plant Model Systems. Cell *167*, 325–339.
69. Causier, B., Castillo, R., Zhou, J., Ingram, R., Xue, Y., Schwarz-Sommer, Z., and Davies, B. (2005). Evolution in action: Following function in duplicated floral homeotic genes. Curr. Biol. *15*, 1508-12.
70. Wasternack, C., and Hause, B. (2013). Jasmonates: Biosynthesis, perception, signal transduction and action in plant stress response, growth and development. Ann. Bot. *111*, 1021-58.
71. Monte, I., Ishida, S., Zamarreño, A.M., Hamberg, M., Franco-Zorrilla, J.M., García-Casado, G., Gouhier-Darimont, C., Reymond, P., Takahashi, K., García-Mina, J.M., *et al.* (2018). Ligand-receptor co-evolution shaped the jasmonate pathway in land plants. Nat. Chem. Biol. *14*, 480–488.

Figure 1. Navigating the plant phylogeny: one tree, different views. (a-c) The dendrograms depict the three most highly supported branching orders for algal and land plant groups forming the green lineage (after Puttick *et al.*[12]). (a) Bryophyte monophyletic; (b) Mosses and liverworts monophyletic, hornworts sister to tracheophytes; (c) Mosses and liverworts monophyletic with hornworts sister to all land plants. (d-e) Alternative views of (a) showing that nodes can be rotated in a tree without changing the topology or relationships between sister groups. (f) Alternative view of cladograms (a-c) depicting the paraphyly of streptophyte algae, the position of the Zygnematophyceae as sister lineage to land plants, and the position of the chlorophyte sister lineage to the streptophytes. (g) Condensed view of (f). (h) Cladogram depicting the relationships between the major groups of land plants as in (a, d-e). (i) Condensed view of (h) highlighting bryophyte and tracheophyte monophyly.

Figure 2. Reconstructing trait evolution and ancestral states from phylogenetic trees. Top (a-c) and bottom (d-f) panels show two plausible topologies of the land plant phylogeny, where bryophyte (green) and tracheophyte (yellow) lineages are monophyletic (a-c), or mosses and liverworts are monophyletic, and hornworts form the sister group to tracheophytes (d-f). White and red-filled circles represent the absence and presence of a given hypothetical character in extant groups, respectively. In the first scenario (a, d), the tree topology has no impact on ancestral character state reconstruction. Reconstructions suggest that the character was present in the land plant common ancestor and that it was lost after the divergence between lycophytes and other tracheophytes, because under a parsimonious model of evolution a single loss (in the MRCA of ferns and seed plants) is more probably than two independent gains (in the MRCA of bryophytes and in lycophytes). In the second scenario (b, e), the character is present only in bryophytes and the tree topology impacts on ancestral character state reconstruction. In (b), reconstruction suggests that the character was present in the bryophyte common ancestor and absent in the tracheophyte common ancestor, and the reconstruction for the land plant common ancestor is uncertain. In (e), reconstruction suggests that the character was present in the land plant common ancestor and subsequently lost in the tracheophyte common ancestor. The number of required changes (losses/gains) of the characters remains the same in each scenario. In the third scenario (c, f), the character is present only in mosses and liverworts. In (c), reconstruction suggests that the character was absent in the land plant common ancestor and was gained in the moss/liverwort common ancestor. In (f), reconstruction suggests that the character was gained in the moss/liverwort common ancestor but the situation in the land plant common ancestor is uncertain. Again, the required number of changes to explain the evolution of the trait remains the same.