Review

The Molecular Phylogeny of Land Plants: Progress and Future Prospects

Guo-Qing Liu [1,2], **Lian Lian** [1] and **Wei Wang** [1,2,*]

1   State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
2   University of Chinese Academy of Sciences, Beijing 100049, China
*   Correspondence: wangwei1127@ibcas.ac.cn

Abstract: Phylogenetics has become a powerful tool in many areas of biology. Land plants are the most important primary producers of terrestrial ecosystems and have colonized various habitats on Earth. In the past two decades, tremendous progress has been made in our understanding of phylogenetic relationships at all taxonomic levels across all land plant groups by employing DNA sequence data. Here, we review the progress made in large-scale phylogenetic reconstructions of land plants and assess the current situation of phylogenetic studies of land plants. We then emphasize directions for future study. At present, the phylogenetic framework of land plants at the order and familial levels has been well built. Problematic deep-level relationships within land plants have also been well resolved by phylogenomic analyses. We pointed out five major aspects of molecular phylogenetics of land plants, which are nowadays being studied and will continue to be goals moving forward. These five aspects include: (1) constructing the genus- and species-level phylogenies for land plant groups, (2) updating the classification systems by combining morphological and molecular data, (3) integrating fossil taxa into phylogenies derived from living taxa, (4) resolving deep-level and/or rapidly divergent phylogenetic relationships using phylogenomic data, and (5) building big trees using the supermatrix method. We hope that this review paper will promote the development of plant molecular phylogenetics and other related areas.

Keywords: land plants; molecular systematics; phylogenomics; phylogeny; morphology; tree of life

1. Introduction

In his famous book, *The Origin of Species*, Charles Darwin [1] put forward the concept of the Tree of Life (TOL), which is a metaphor for presenting relationships of organisms in space and time. All organisms on Earth originated from a common ancestor and each can be found in the TOL. As a large lineage of TOL, land plants (embryophytes) are the most important primary producers of terrestrial ecosystems. Living land plants are an important source of aliments, timbers, fibers, pharmaceuticals, and other vital resources for human survival and health, and fossilized land plants become one of the sources of fossil fuels, particularly coal, with driving global economy [2]. By an array of innovations, including embryos, sperms, and eggs protected in multicellular structures, alternating generations of diploid sporophytes, and haploid gametophytes, land plants have become the most diverse group of green plants, have dominated modern terrestrial environments, and are the foundation of the vast majority of terrestrial ecosystems [3].

The rise of land plants is one of the major events in the history of life, which irreversibly changed the environments on Earth, including altering atmospheric composition by enhancing photosynthesis and influencing carbon fixation and carbon storage, and affecting evolutionary trajectories of other organisms by promoting the formation of soil and soil microbiota and the establishment of new food chains and new habitats [4,5]. The origin and evolution of land plants is a central theme in evolutionary biology and ecology [3,6]. In the

past two decades, tremendous progress has been made in reconstruction of land plant TOL by employing molecular data [4,7–9]. The order- and family-level phylogenetic framework of land plants has been well built [10–12]. In particular, problematic deep-level relationships within land plants have been largely resolved by phylogenomic analyses [4,8,9].

Here, we review the progress made in large-scale phylogenetic reconstructions of land plants and assess the current situation of phylogenetic studies of land plants. We then point out five major aspects of molecular phylogenetics of land plants, which are nowadays being studied and will continue to be goals moving forward. We hope that this review paper will promote the development of plant molecular phylogenetics and other related areas.

2. Large-Scale Phylogenetic Framework of Land Plants

With recent developments in extracting DNA, sequencing, methods of analytical technologies, and computing power, tremendous progress has been made in our understanding of phylogenetic relationships at all taxonomic levels across all land plant groups by employing DNA sequence data. At present, the large-scale phylogenetic framework of land plants has been established and the most likely sister group of land plants has also been identified. Land plants comprise five major clades: bryophytes, lycophytes, monilophytes, gymnosperms, and angiosperms (Figure 1).
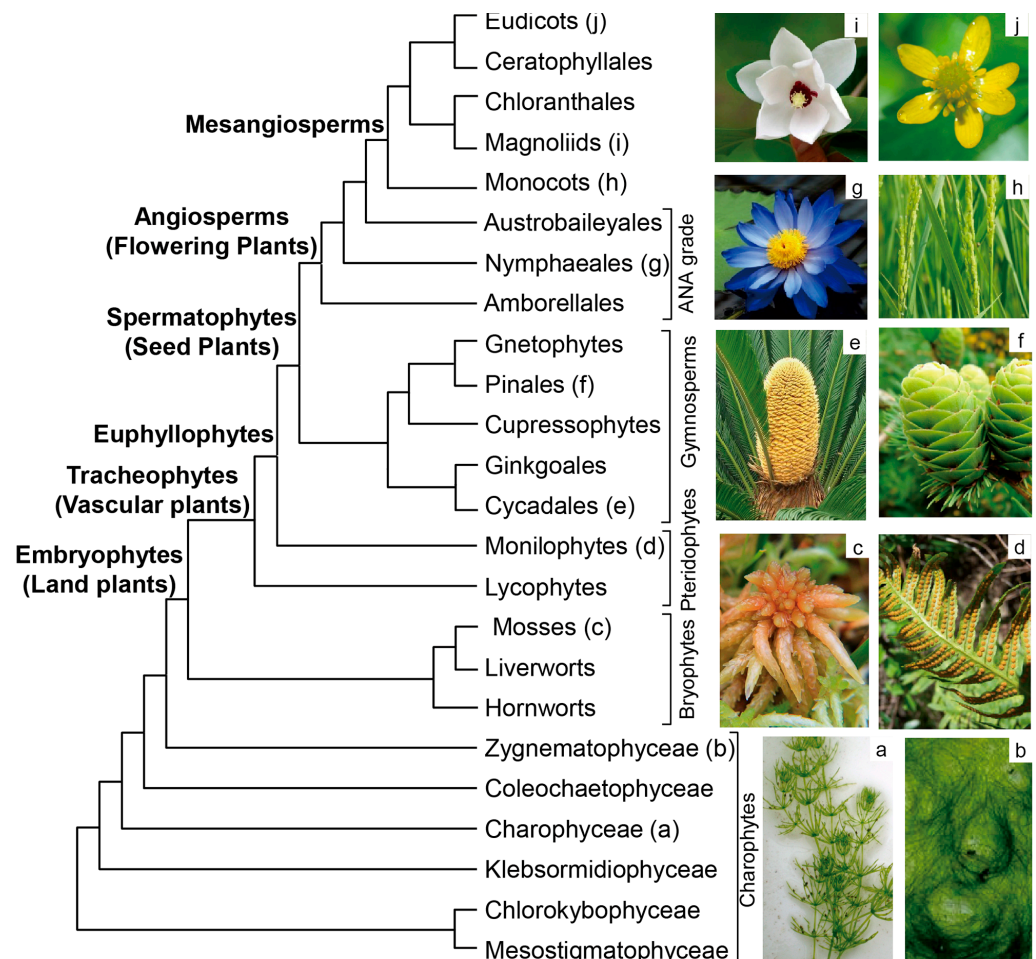

Figure 1. Summary of phylogenetic relationships among major clades of land plants. (**a**) *Chara braunii*; (**b**) *Spirogyra communis*; (**c**) *Sphagnum palustre*; (**d**) *Goniophlebium chinense*; (**e**) *Cycas revoluta*; (**f**) *Larix kaempferi*; (**g**) *Nymphaea gigantea*; (**h**) *Oryza sativa*; (**i**) *Oyama sieboldii*; (**j**) *Ranunculus sieboldii*.

2.1. Sister Group of Land Plants

Land plants originated from streptophyte algae in fresh water [13], but which green algal lineage gave rise to land plants has been a long-standing dispute. Phylogenetic studies indicate that streptophyte algae is a paraphyletic group, referred to as charophytes [14], and contains six major clades: Chlorokybophyceae, Mesostigmatophyceae, Klebsormidiophyceae, Charophyceae, Coleochaetophyceae, and Zygnematophyceae (Figure 1). Many clades of streptophytes were once suggested as sisters to land plants, such as Charophyceae [14], Zygnematophyceae [15–17], Coleochaetophyceae [18], and the clade consisting of Zygnematophyceae and Coleochaetophyceae [16,19]. Recently, phylogenomic analyses strongly support Zygnematophyceae as sister to land plants [7,8].

2.2. Bryophytes

The simplest morphological structures among land plants, bryophytes have still evolved a range of characters to adapt terrestrial habitats, such as stomata, cuticle, water-conducting cells, and embryos [20–23]. Due to a lack of seeds and a vascular system, life cycles of bryophytes are highly dependent on water [24]. Moreover, the haploid gametophyte phase of bryophytes is dominant in life history, which is distinguished from the remaining land plants. Bryophytes consist of three major monophyletic groups: hornworts, liverworts, and mosses. The relationships among these three groups and the relationships between bryophytes and vascular plants have been energetically debated, and various hypotheses have been proposed. Based on the distribution of three group II introns in the mitochondrial genes *nad1* and *cox2*, Qiu et al. [25] suggested bryophytes were paraphyletic, i.e., liverworts were sister to all other land plants and hornworts were sister to vascular plants, in agreement with the results of the multi-locus phylogenetic analyses [26,27]. Mitochondrial phylogenomic analyses by mitigating the effects of saturation, compositional heterogeneity, and codon-usage bias also supported these relationships [28], whereas plastid phylogenomic analyses by accounting for composition biases among synonymous substitutions supported bryophytes to be monophyletic [29]. Maximum likelihood analysis based on concatenated alignments of first and second codon positions for 674 nuclear genes also suggested bryophytes as a paraphyletic group, but hornworts were identified as the earliest-diverging lineages within land plants and liverworts and mosses formed a clade, sister to vascular plants, whereas the coalescent-based analysis based on 424 gene trees estimated from first and second codon position strongly recognized bryophytes monophyletic [3]. Recently, nuclear genomic phylogenetic analyses further supported bryophytes consist of a monophyletic group and hornworts are sister to liverworts and mosses [4,30,31]. These analyses highlight the negative effects of substitutional saturation and synonymous substitutions on phylogenetic references based on nuclear genomic data.

2.3. Pteridophytes

Pteridophytes consist of lycophytes and monilophytes (ferns). Traditionally, pteridophytes were considered to be monophyletic due to lycophytes and monilophytes having a similar life cycle. Phylogenetic analyses based on the 18S rDNA sequences first found that lycophytes were the earliest-diverging lineage within vascular plants and monilophytes were the sister of seed plants [32]. Subsequent multi-locus and phylogenomic analyses consistently supported these hypotheses [3,8,33,34].

2.4. Gymnosperms

Gymnosperms and angiosperms are collectively referred to as seed plants because their reproductive organ is seeds, instead of spores. The monophyly of gymnosperms has been unambiguously authenticated by molecular sequence data from nuclear and plastid genomes [35,36]. Gymnosperms consist of four groups: cycads, conifers, *Ginkgo*, and Gnetales. A phylotranscriptomic analysis with a sampling of all 13 families of gymnosperms and main lineages of angiosperms indicates that cycads plus *Ginkgo* as sister to the remaining gymnosperms and Gnetales is embedded within conifers, sister to Pinaceae [9].

Nuclear genomic data also support cycads plus *Ginkgo* as sister to the remaining gymnosperms [37,38]. Convergent molecular evolution or homoplasy is partially responsible for the phylogenetic conflicts in seed plants [9,38].

2.5. Angiosperms

As a plant group that is the most closely related to human production and life, angiosperms have spectacular morphological and species diversity, and play an irreplaceable role in global terrestrial ecosystems. Angiosperms are characterized by flowers, ovules covered by carpels, double fertilization, nutritious triploid endosperm, and vessel elements. Monophyly of angiosperms is supported by various phylogenetic analyses. Amborellales, Nymphaeales, and Austrobaileyales are successive sisters to all other angiosperms, referred to as the ANA grade [39–43]. Mesangiospermae is a monophyletic group and comprises five major lineages: monocots, magnoliids, Chloranthales, Ceratophyllales, and eudicots. The relationships among these five lineages have long been an open question because they diverged rapidly within a 3-million-year time window [44]. One Thousand Plant Transcriptomes support monocots as sister to the remaining mesangiosperms [8], and eudicots plus Ceratophyllales is sister to magnoliids plus Chloranthales (Figure 1), in agreement with the results of the nuclear genomic data [45,46]. Guo et al. [46] further found that ancient hybridization may account for the incongruent phylogenetic placements of Chloranthales + magnoliids relative to monocots and eudicots in nuclear and chloroplast trees.

3. The Potential Research Focuses on Reconstructing the TOL of Land Plants in the Long Period Future

In 1993, the article entitled "*Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene rbcL*" is a landmark work in plant molecular systematics [47]. Since then, various molecular data have been widely used to reconstruct the phylogenetic relationships of land plant groups. Development of plant molecular phylogenetics is very rapid, with a process using single locus, multi-locus, to genomic data. It is safe to say that due to the emergence of molecular phylogenetics, the achievements of plant phylogenetics made in the past twenty years have far exceeded the sum of the previous two hundred years. So, Soltis et al. [48] allege that the current period is a "golden era" in plant phylogenetics, as well as organismal phylogenetics in general. Molecular systematics is not only a subdiscipline, but has also become a powerful tool in many areas of biology, such as physiology, ecology, biogeography, paleobiology, genomics, and developmental genetics. Here, we summarized five important aspects of molecular phylogenetics of land plants, which are nowadays being studied and will continue to be goals moving forward.

3.1. Reconstructing Genus- and Species-Level Phylogenies for Land Plant Groups

With the establishment of the large-scale phylogenetic framework of angiosperms, a DNA phylogeny-based angiosperm classification system at the order and familial levels was proposed by the Angiosperm Phylogeny Group (APG) in 1998 and has been updated three times [11,49–51]. Other land plants, such as gymnosperms [12], ferns [52], and mosses [53] also have a relatively robust phylogenetic framework. The establishment of the order and familial framework is only the first step in reconstruction the land plant TOL of land plants. Altogether, 500,000 species of green plants occur on Earth, and only less than 30% of the species have reported molecular sequences so far [54,55]. In particular, taxon sampling of large-scale phylogenetic studies is sparse. For example, large-scale phylogenetic analyses in angiosperms or eudicots only included one to three species for each family; thus, the circumscriptions of many families cannot be resolved due to the relatively limited taxon sampling. Recently, some heterogeneous families have re-delimited and several new families have been established, such as Arthropteridaceae (ferns) [56], Pteridryaceae (ferns) [57], Borthwickiaceae (angiosperms) [58], and Wightiaceae (angiosperms) [59]. To clarify familial circumscription, a relatively dense taxon sampling at the generic level, especially the inclusion of segregate genera, is necessary.

For delimiting a generic circumscription, species-level taxon sampling is essential. Many traditionally recognized genera have been re-delimited by employing molecular data. For example, Ranunculaceae, a basal eudicot family, contains approximately 60 genera of which 23 were re-delimited or adjusted [60]. Kadereit et al. [61] evaluated the monophyly of all genera of vascular plants in Germany, and identified that *c.* 140 genera are not monophyletic among the 840 genera examined, and the monophyly of *c.* 20 genera is ambiguous. We conducted a statistical survey of new genera of extant land plants published from 2015 to 2021 on the ISI Web of Science (www.webofscience.com/wos/alldb/basic-search, accessed on 12 June 2022) and International Plant Name Index (IPNI, www.ipni.org/, accessed on 10 June 2022) and found that more than 70 genera were described each year since 2015, most of which belongs into angiosperms, with the largest proportion of eudicots (Figure 2). At present, genus- and species-level phylogenetic studies of land plants are reported in almost every issue of molecular systematic journals, such as *Molecular Phylogenetics and Evolution*, *Taxon*, and *Systematic Botany*. Comprehensive understanding of genus- and species-level relationships across land plants is still a great challenge, especially for large and widespread genera. Phylogenetic research at the generic and species level will be the focus of reconstructing the land plant TOL for a long time in the future.
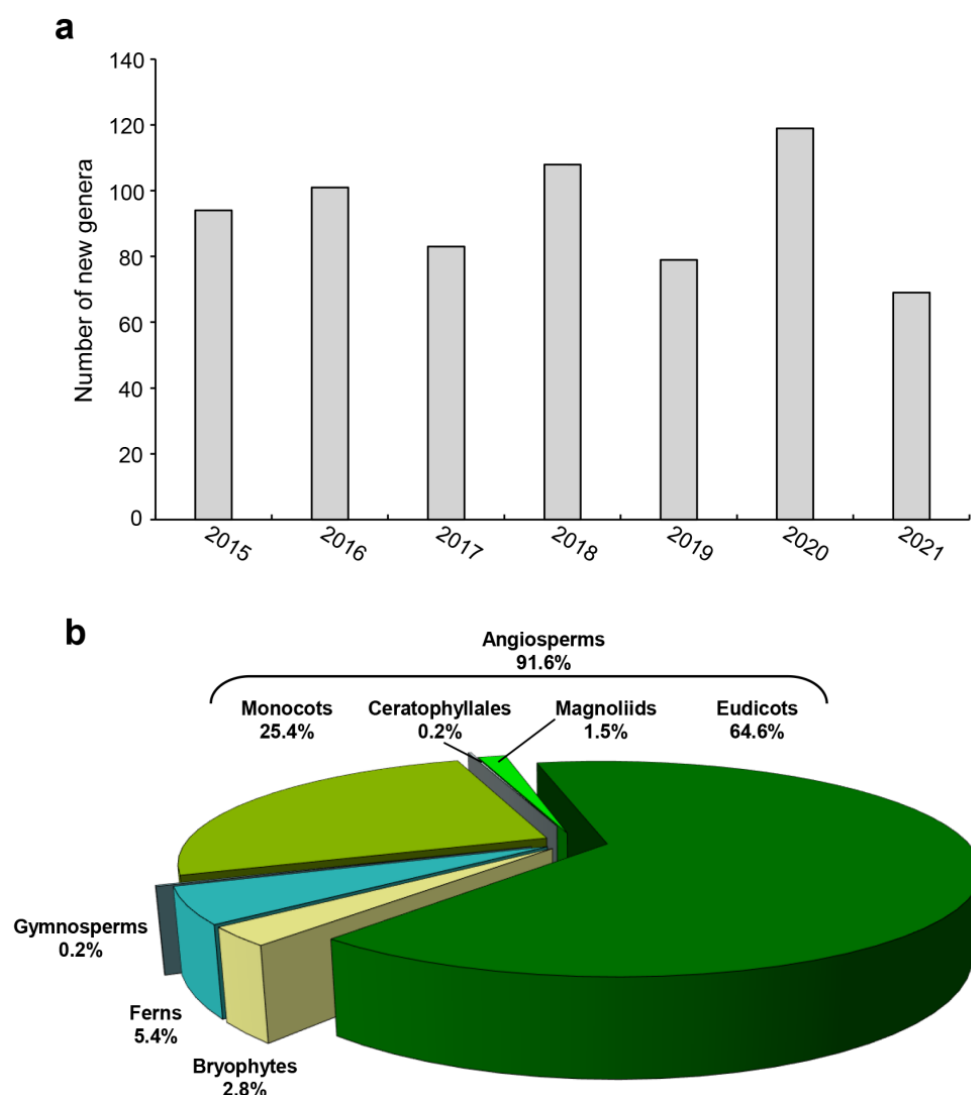

Figure 2. Cumulative number of new genera between 2015 and 2021. (**a**) The number of new genera published annually. (**b**) Proportion of new genera of major groups of land plants.

3.2. Updating the Classification Systems by Combining Morphological and Molecular Data

The APG system has been widely accepted, but it was established based only on molecular data. The orders and families in the APG system were circumscribed using only "monophyly" as criterion. From a morphological viewpoint, some of the currently recognized orders and families in the APG system are highly heterogeneous, such as Malpighiales, Cucurbitales, and Resedaceae [62]. How to determine morphological synapomorphies of those orders is a large challenge. Moreover, a key to distinguish the orders and families of angiosperms is lacking [62]. These prevent the wide implications of the APG system. In particular, the phylogenetic relationships and taxonomic status of many East Asian taxa, including Acanthochlamydaceae, Aceraceae, Bretschneideraceae, Cornaceae *sensu lato*, Hippocastanaceae, Illiciaceae, Leeaceae, Rhoipteleaceae, and Tetracentraceae, need to re-evaluated by integrating morphological and molecular data [62].

Systematics and evolutionary biology have developed for more than one century and have accumulated a wealth of morphological characters through gross morphology, anatomy, embryology, palynology, cytology, and ontogeny. Combining molecular and morphological data can greatly increase the ability to retrospect phylogenetic relationships of organisms and the process of evolution [63,64]. Yet, compared to molecular data, morphological data has been greatly ignored in reconstructing the TOL. At present, more than 95% of systematics articles use only DNA sequence data [65]. Bayesian inference, which can analyze morphological data or combined morphological and molecular data, has greatly improved the application of morphological characters [66,67]. By combing morphological and molecular data, we can determine the diagnostic characters for each lineage in the phylogenetic framework and thereby put forward a natural and reasonable classification system [68,69].

3.3. Integrating Fossil Taxa into Phylogenies Derived from Living Taxa

A complete TOL includes not only living taxa but also extinct taxa. However, more than 95% of current phylogenetic analyses include only extant taxa [65]. With current technology, it is almost impossible to obtain nucleotide sequences from extinct plant species. Therefore, morphological characters have become the only option for phylogenetic analyses of fossil plants. The widespread applications of scanning electron microscopy and synchrotron radiation X-ray tomographic microscopy in fossil taxa have enabled paleobotanists to obtain numerous accurate morphological data from fossils [70]. In contrast, these new techniques are rarely used in studies of extant plant species.

The inclusion of fossil taxa not only affects or improves topology, it is also important for a correct understanding of character evolution because fossils may have a combination of characters unlike those of any extant taxon [71]. Another role of fossil data in combination with molecular phylogenetics is to serve as calibration points based on the assumed position of fossil taxa in extant trees [72]. When multiple fossils from different periods are assigned to the same clade, the oldest fossil is usually selected as the calibration point. For a group with rich fossil records, the inclusion of fossil taxa can heavily influence its biogeographic reconstruction, particularly when the fossil taxa are outside its modern distribution [73]. If multiple fossils belong to a contemporaneous clade, directly placing them in the phylogenetic trees will result in a zero-length branch. The recently developed tip-dating (TD) Bayesian method can make use of all available fossil taxa as terminal tips to generate a timetree containing extinct and extant groups and to overcome zero-length branches to a certain extent [74]. However, this method has a wide application in zoological research. Thus, it has broad prospects in plant biogeographic studies, including living and extinct taxa.

3.4. Resolve Deep-Level and/or Rapidly Divergent Phylogenetic Relationships using Phylogenomic Data

When using a few DNA regions cannot resolve phylogenetic relationships well, more loci need to be sampled. Currently, phylogenomic analyses have been widely used in solving recalcitrant phylogenetic relationships, including relationships within rapidly radiating taxa and positions of relict taxa. The application of next-generation sequencing has greatly reduced the cost of sequencing. The plastid genomes in plants are usually a single, non-recombining locus [75] and have been widely applied in plant phylogenetic studies. Currently, most of plastid phylogenomic studies directly use protein-coding gene sequences, which may mislead phylogenetic references. Goremykin et al. [76] found that Amborellales and Nymphaeales formed a clade after removing "fast sites," which is contradicted with the prevalent view that only *Amborella* is the earliest-diverging lineage in angiosperms. Nonetheless, Drew et al. [77] revealed the "noisy" data actually supports *Amborella* as sister to all other angiosperms. To account for "noisy sites," some new nucleotide substitution models or analytic methods have also been proposed. For example, Goremykin et al. [76] proposed the CAT + GTR + Γ + covext model that considered base compositional heterogeneity. When using 82 plastid genes, Xi et al. [78] proposed posteriori data partitioning based on the Bayesian mixture mode, which resolved well the phylogenetic relationships in Malpighiales, the most recalcitrant clade in angiosperms.

Compared to plastid genome, biparental inheritance nuclear genome can not only provide more characters but can also reveal reticular evolution processes, so it has greater potential in phylogenetic studies and may be a key direction of plant phylogeny in the future. Especially, the developments of the restriction-site associated DNA sequencing, target enrichment, and genome skimming technique [79] have reduced sequencing costs and have greatly promoted nuclear phylogenomic studies of land plants, as well as other organisms. However, nuclear genomes have a more complicated evolutionary history, and may contain more evolutionary "noise," such as evolutionary saturation, base compositional heterogeneity, and synonymous codon bias. Incomplete lineage sorting is a common evolutionary phenomenon, and it may cause wrong results based on concatenated alignments. Currently, the coalescent-based method has been widely used in nuclear phylogenomic analyses, alleviating the influence of incomplete lineage sorting to a certain extent.

3.5. Building "Big Trees" Using the Supermatrix Method

With unprecedented increase of molecular data available in public databases, such as NCBI (www.ncbi.nlm.nih.gov/ accessed on 15 August 2022), TreeBASE (www.treebase.org/, accessed on 15 August 2022), and Dryad (www.datadryad.org/, accessed on 15 August 2022), massive amounts of molecular data have been generated and can be downloaded free, which provide the possibility to construct mega-phylogenies. There are two methods for constructing a mega-phylogeny: supertree and supermatrix. The supertree method compiles source trees with partially overlapping taxa into a single comprehensive tree, and the supermatrix method assembles numerous matrices with overlapping taxa into a super matrix and then reconstructs a phylogenetic tree [80]. Since the supermatrix contains genetic information of species, it has broader application for downstream analyses, such as estimating divergence times, referring ancestral ranges, and calculating diversification rates. The mega-phylogeny is widely used to explore the evolutionary dynamics of biodiversity in which the inclusion of as many taxa as possible is more important than the improvement of bootstrap values. Smith et al. [81] constructed a phylogeny of 5036 species of Caryophyllales and found that a series of diversification rate shifts occurred more recently than whole genome duplication events. Folk et al. [55] used rosids as a case to illustrate how to construct a comprehensively sampled phylogeny at the species-level on a global-scale and to discuss difficulties and opportunities of associated geographic and phenotypic resources. There were also some studies using the supermatrix approach to construct the TOL and then to explore its phylogenetic diversity, community structure, and biogeographic patterns. Thornhill et al. [82] constructed a phylogeny for California flora and discussed relative phy-

logenetic diversity, phylogenetic endemism, and neo- and paleo-endemism. Wu et al. [83] reconstructed a phylogenetic tree for 157 species of Zygophyllaceae based on four DNA markers using the supermatrix method and then integrated phylogenetic, molecular dating, biogeographic, and diversification rate methods to investigate the diversity dynamics of the family through time.

As phylogenetic trees become larger, more computer power is required, especially for generating a "big" timetree. Graphics processing unit (GPU) multi-core resources with computer clusters, cloud computing platforms, and parallel version upgrades of analysis software have started to be adopted to improve computer speed. How to visualize constructing a "big" tree is also an urgent problem [48]. In addition, providing an integrated, open, and real-time renewable TOL of land plants for science and society will be a great challenge in the future [84].

Author Contributions: Conceptualization, W.W.; writing—original draft, G.-Q.L.; writing—review and editing, G.-Q.L., L.L. and W.W.; funding acquisition, W.W. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable for studies not involving humans or animals.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Darwin, C. *The Origin of Species by Means of Natural Selection*; John Murray: London, UK, 1859.
2. Davis, C.C.; Matthews, S. *Evolution of Land Plants*; Oxford University Press: New York, NY, USA, 2019. [CrossRef]
3. 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. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E4859–E4868. [CrossRef]
4. Su, D.; Yang, L.; Shi, X.; Ma, X.; Zhou, X.; Hedges, S.B.; Zhong, B. Large-scale phylogenomic analyses reveal the monophyly of bryophytes and neoproterozoic origin of land plants. *Mol. Biol. Evol.* **2021**, *38*, 3332–3344. [CrossRef] [PubMed]
5. Donoghue, P.C.J.; Harrison, C.J.; Paps, J.; Schneider, H. The evolutionary emergence of land plants. *Curr. Biol.* **2021**, *31*, R1281–R1298. [CrossRef] [PubMed]
6. Cheng, S.; Xian, W.; Fu, Y.; Marin, B.; Keller, J.; Wu, T.; Sun, W.; Li, X.; Xu, Y.; Zhang, Y.; et al. Genomes of subaerial Zygnematophyceae provide insights into land plant evolution. *Cell* **2019**, *179*, 1057–1067.e14. [CrossRef]
7. Gitzendanner, M.A.; Soltis, P.S.; Wong, G.K.; Ruhfel, B.R.; Soltis, D.E. Plastid phylogenomic analysis of green plants: A billion years of evolutionary history. *Am. J. Bot.* **2018**, *105*, 291–301. [CrossRef]
8. One Thousand Plant Transcriptomes Initiative. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **2019**, *574*, 679–685. [CrossRef] [PubMed]
9. Ran, J.H.; Shen, T.T.; Wang, M.M.; Wang, X.Q. Phylogenomic resolves the deep phylogeny of seed plants and indicates partial convergent or homoplastic evolution between Gnetales and angiosperms. *Proc. R. Soc. B* **2018**, *285*, 20181012. [CrossRef]
10. The Pteridophyte Phylogeny Group. A community-derived classification for extant lycophytes and ferns. *J. Syst. Evol.* **2016**, *54*, 563–603. [CrossRef]
11. The Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* **2016**, *181*, 1–20. [CrossRef]
12. Yang, Y.; Ferguson, D.K.; Liu, B.; Mao, K.S.; Gao, L.M.; Zhang, S.Z.; Wan, T.; Rushforth, K.; Zhang, Z.X. Recent advances on phylogenomics of gymnosperms and a new classification. *Plant Divers.* **2022**, *44*, 340–350. [CrossRef]
13. McCourt, R.M.; Delwiche, C.F.; Karol, K.G. Charophyte algae and land plant origins. *Trends Ecol. Evol.* **2004**, *19*, 661–666. [CrossRef]
14. Karol, K.G.; McCourt, R.M.; Cimino, M.T.; Delwiche, C.F. The closest living relatives of land plants. *Science* **2001**, *294*, 2351–2353. [CrossRef] [PubMed]
15. Timme, R.E.; Bachvaroff, T.R.; Delwiche, C.F. Broad phylogenomic sampling and the sister lineage of land plants. *PLoS ONE* **2012**, *7*, e29696. [CrossRef] [PubMed]

16. Wodniok, S.; Brinkmann, H.; Glöckner, G.; Heidel, A.J.; Philippe, H.; Melkonian, M.; Becker, B. Origin of land plants: Do conjugating green algae hold the key? *BMC Evol. Biol.* **2011**, *11*, 104. [CrossRef]

17. Turmel, M.; Otis, C.; Lemieux, C. The chloroplast genome sequence of Chara vulgaris sheds new light into the closest green algal relatives of land plants. *Mol. Biol. Evol.* **2006**, *23*, 1324–1338. [CrossRef]

18. Finet, C.; Timme, R.E.; Delwiche, C.F.; Marlétaz, F. Multigene phylogeny of the green lineage reveals the origin and diversification of land plants. *Curr. Biol.* **2012**, *22*, 1456–1457. [CrossRef]

19. Laurin-Lemay, S.; Brinkmann, H.; Philippe, H. Origin of land plants revisited in the light of sequence contamination and missing data. *Curr. Biol.* **2012**, *22*, R593–R594. [CrossRef]

20. Pires, N.D.; Dolan, L. Morphological evolution in land plants: New designs with old genes. *Philos. Trans. R. Soc. B* **2012**, *367*, 508–518. [CrossRef]

21. Jill Harrison, C. Development and genetics in the evolution of land plant body plans. *Philos. Trans. R. Soc. B* **2017**, *372*, 20150490. [CrossRef]

22. Ishizaki, K. Evolution of land plants: Insights from molecular studies on basal lineages. *Biosci. Biotechnol. Biochem.* **2017**, *81*, 73–80. [CrossRef] [PubMed]

23. Szövényi, P.; Waller, M.; Kirbis, A. Evolution of the plant body plan. *Curr. Top. Dev. Biol.* **2019**, *131*, 1–34. [CrossRef]

24. Wang, Q.H.; Zhang, J.; Liu, Y.; Jia, Y.; Jiao, Y.N.; Xu, B.; Chen, Z.D. Diversity, phylogeny, and adaptation of bryophytes: Insights from genomic and transcriptomic data. *J. Exp. Bot.* **2022**, *73*, 4306–4322. [CrossRef]

25. Qiu, Y.L.; Cho, Y.; Cox, C.; Palmer, J.D. The gain of three mitochondrial introns identifies liverworts as the earliest land plants. *Nature* **1998**, *394*, 671–674. [CrossRef]

26. Qiu, Y.L.; Li, L.B.; Wang, B.; Chen, Z.D.; Knoop, V.; Groth-Malonek, M.; Dombrovska, O.; Lee, J.; Kent, L.; Rest, J.; et al. The deepest divergences in land plants inferred from phylogenomic evidence. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15511–15516. [CrossRef]

27. Chang, Y.; Graham, S.W. Bryophyte-specific primers for retrieving plastid genes suitable for phylogenetic inference. *Am. J. Bot.* **2011**, *98*, e109–e113. [CrossRef]

28. Liu, Y.; Cox, C.J.; Wang, W.; Goffinet, B. Mitochondrial phylogenomics of early land plants: Mitigating the effects of saturation, compositional heterogeneity, and codon-usage bias. *Syst. Biol.* **2014**, *63*, 862–878. [CrossRef]

29. Cox, C.J.; Li, B.; Foster, P.G.; Embley, T.M.; Civan, P. Conflicting phylogenies for early land plants are caused by composition biases among synonymous substitutions. *Syst. Biol.* **2014**, *63*, 272–279. [CrossRef]

30. Zhang, J.; Fu, X.X.; Li, R.Q.; Zhao, X.; Liu, Y.; Li, M.H.; Zwaenepoel, A.; Ma, H.; Goffinet, B.; Guan, Y.L.; et al. The hornwort genome and early land plant evolution. *Nat. Plants* **2020**, *6*, 107–118. [CrossRef]

31. De Sousa, F.; Foster, P.G.; Donoghue, P.C.J.; Schneider, H.; Cox, C.J. Nuclear protein phylogenies support the monophyly of the three bryophyte groups (Bryophyta Schimp.). *New Phytol.* **2019**, *222*, 565–575. [CrossRef]

32. Kranz, H.D.; Huss, V.A.R. Molecular evolution of pteridophytes and their relationship to seed plants: Evidence from complete 18S rRNA gene sequences. *Plant Syst. Evol.* **1996**, *202*, 1–11. [CrossRef]

33. Pryer, K.M.; Schneider, H.; Smith, A.R.; Cranfill, R.; Wolf, P.G.; Hunt, J.S.; Sipes, S.D. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* **2001**, *409*, 618–622. [CrossRef] [PubMed]

34. Rai, H.S.; Graham, S.W. Utility of a large, multigene plastid data set in inferring higher-order relationships in ferns and relatives (monilophytes). *Am. J. Bot.* **2010**, *97*, 1444–1456. [CrossRef] [PubMed]

35. Bowe, L.M.; Coat, G.; DePamphilis, C.W. Phylogeny of seed plants based on all three genomic compartments: Extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 4092–4097. [CrossRef]

36. Chaw, S.M.; Parkinson, C.L.; Cheng, Y.C.; Vincent, T.M.; Palmer, J.D. Seed plant phylogeny inferred from all three plant genomes: Monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 4086–4091. [CrossRef]

37. Liu, Y.; Wang, S.; Li, L.; Yang, T.; Dong, S.; Wei, T.; Wu, S.; Liu, Y.; Gong, Y.; Feng, X.; et al. The Cycas genome and the early evolution of seed plants. *Nat. Plants* **2022**, *8*, 389–401. [CrossRef]

38. Wan, T.; Liu, Z.M.; Li, L.F.; Leitch, A.R.; Leitch, I.J.; Lohaus, R.; Liu, Z.J.; Xin, H.P.; Gong, Y.B.; Liu, Y.; et al. A genome for gnetophytes and early evolution of seed plants. *Nat. Plants* **2018**, *4*, 82–89. [CrossRef]

39. Parkinson, C.L.; Adams, K.L.; Palmer, J.D. Multigene analyses identify the three earliest lineages of extant flowering plants. *Curr. Biol.* **1999**, *9*, 1485–1491. [CrossRef]

40. Saarela, J.M.; Rai, H.S.; Doyle, J.A.; Endress, P.K.; Mathews, S.; Marchant, A.D.; Briggs, B.G.; Graham, S.W. Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree. *Nature* **2007**, *446*, 312–315. [CrossRef]

41. Qiu, Y.L.; Dombrovska, O.; Lee, J.; Li, L.; Whitlock, B.A.; Bernasconi-Quadroni, F.; Rest, J.S.; Davis, C.C.; Borsch, T.; Hilu, K.W.; et al. Phylogenetic analyses of basal angiosperms based on nine plastid, mitochondrial, and nuclear genes. *Int. J. Plant Sci.* **2005**, *166*, 815–842. [CrossRef]

42. Amborella Genome Project; Albert, V.A.; Barbazuk, W.B.; Depamphilis, C.W.; Der, J.P.; Leebens-Mack, J.; Ma, H.; Palmer, J.D.; Rounsley, S.; Sankoff, D.; et al. The Amborella genome and the evolution of flowering plants. *Science* **2013**, *342*, 1241089. [CrossRef]

43. Zhang, L.; Chen, F.; Zhang, X.; Li, Z.; Zhao, Y.; Lohaus, R.; Chang, X.; Dong, W.; Ho, S.Y.W.; Liu, X.; et al. The water lily genome and the early evolution of flowering plants. *Nature* **2020**, *577*, 79–84. [CrossRef] [PubMed]

44. Moore, M.J.; Bell, C.D.; Soltis, P.S.; Soltis, D.E. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19363–19368. [CrossRef] [PubMed]

45. Ma, J.; Sun, P.; Wang, D.; Wang, Z.; Yang, J.; Li, Y.; Mu, W.; Xu, R.; Wu, Y.; Dong, C.; et al. The Chloranthus sessilifolius genome provides insight into early diversification of angiosperms. *Nat. Commun.* **2021**, *12*, 6929. [CrossRef] [PubMed]

46. Guo, X.; Fang, D.; Sahu, S.K.; Yang, S.; Guang, X.; Folk, R.; Smith, S.A.; Chanderbali, A.S.; Chen, S.; Liu, M.; et al. Chloranthus genome provides insights into the early diversification of angiosperms. *Nat. Commun.* **2021**, *12*, 6930. [CrossRef]

47. Chase, M.; Soltis, D.E.; Olmstead, R.G.; Morgan, D.; Les, D.H.; Mishler, B.D.; Duvall, M.R.; Price, R.A.; Hills, H.G.; Qiu, Y.L.; et al. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene rbcL. *Ann. Mo. Bot. Gard.* **1993**, *80*, 528–580. [CrossRef]

48. Soltis, D.E.; Moore, M.J.; Burleigh, G.; Soltis, P.S. Molecular markers and concepts of plant evolutionary relationships: Progress, promise, and future prospects. *Crit. Rev. Plant Sci.* **2009**, *28*, 1–15. [CrossRef]

49. The Angiosperm Phylogeny Group. An ordinal classification for the families of flowering plants. *Ann. Mo. Bot. Gard.* **1998**, *85*, 531–553. [CrossRef]

50. The Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Bot. J. Linn. Soc.* **2003**, *141*, 399–436. [CrossRef]

51. The Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* **2009**, *161*, 105–121. [CrossRef]

52. Lehtonen, S. Towards resolving the complete fern tree of life. *PLoS One* **2011**, *6*, e24851. [CrossRef]

53. Liu, Y.; Johnson, M.G.; Cox, C.J.; Medina, R.; Devos, N.; Vanderpoorten, A.; Hedenäs, L.; Bell, N.E.; Shevock, J.R.; Aguero, B.; et al. Resolution of the ordinal phylogeny of mosses using targeted exons from organellar and nuclear genomes. *Nat. Commun.* **2019**, *10*, 1485. [CrossRef] [PubMed]

54. Hinchliff, C.E.; Smith, S.A.; Allman, J.F.; Burleigh, J.G.; Chaudhary, R.; Coghill, L.M.; Crandall, K.A.; Deng, J.; Drew, B.T.; Gazis, R.; et al. Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 12764–12769. [CrossRef] [PubMed]

55. Folk, R.A.; Sun, M.; Soltis, P.S.; Smith, S.A.; Soltis, D.E.; Guralnick, R.P. Challenges of comprehensive taxon sampling in comparative biology: Wrestling with rosids. *Am. J. Bot.* **2018**, *105*, 433–445. [CrossRef]

56. Liu, H.-M.; Jiang, R.-H.; Guo, J.; Hovenkamp, P.; Perrie, L.R.; Shepherd, L.; Hennequin, S.; Schneider, H. Towards a phylogenetic classification of the climbing fern genus Arthropteris. *Taxon* **2013**, *62*, 688–700. [CrossRef]

57. Zhou, X.M.; Zhang, L.; Lu, N.T.; Gao, X.F.; Zhang, L.B. Pteridryaceae: A new fern family of Polypodiineae (Polypodiales) including taxonomic treatments. *J. Syst. Evol.* **2018**, *56*, 148–173. [CrossRef]

58. Su, J.X.; Wang, W.; Zhang, L.B.; Chen, Z.D. Phylogenetic placement of two enigmatic genera, Borthwickia and Stixis, based on molecular and pollen data, and the description of a new family of Brassicales, Borthwickiaceae. *Taxon* **2012**, *61*, 601–611. [CrossRef]

59. Liu, B.; Tan, Y.H.; Liu, S.; Olmstead, R.G.; Min, D.Z.; Chen, Z.D.; Joshee, N.; Vaidya, B.N.; Chung, R.C.K.; Li, B. Phylogenetic relationships of Cyrtandromoea and Wightia revisited: A new tribe in Phrymaceae and a new family in Lamiales. *J. Syst. Evol.* **2019**, *58*, 1–17. [CrossRef]

60. Wang, W.; Liu, Y.; Yu, S.X.; Gao, T.G.; Chen, Z.D. Gymnaconitum, a new genus of Ranunculaceae endemic to the Qinghai–Tibetan Plateau. *Taxon* **2013**, *62*, 713–722. [CrossRef]

61. Kadereit, J.W.; Albach, D.C.; Ehrendorfer, F.; Galbany-Casals, M.; Garcia-Jacas, N.; Gehrke, B.; Kadereit, G.; Kilian, N.; Klein, J.T.; Koch, M.A.; et al. Which changes are needed to render all genera of the German flora monophyletic? *Willdenowia* **2016**, *46*, 39–91. [CrossRef]

62. Wang, W.; Zhang, X.X.; Chen, Z.D.; Lu, A.M. Comments on the APG's classification of angiosperms. *Biodivers. Sci.* **2017**, *25*, 418–426. [CrossRef]

63. Nandi, O.; Chase, M.W.; Endress, P.K. A combined cladistic analysis of angiosperms using rbcL and non-molecular data sets. *Ann. Mo. Bot. Gard.* **1998**, *85*, 137–214. [CrossRef]

64. Doyle, J.A.; Endress, P.K. Morphological phylogenetic analysis of basal angiosperms: Comparison and combination with molecular data. *Int. J. Plant Sci.* **2000**, *161*, S121–S153. [CrossRef]

65. Soltis, D.E.; Moore, M.J.; Sessa, E.B.; Smith, S.A.; Soltis, P.S. Using and navigating the plant tree of life. *Am. J. Bot.* **2018**, *105*, 287–290. [CrossRef] [PubMed]

66. Glenner, H.; Hansen, A.J.; Sørensen, M.V.; Ronquist, F.; Huelsenbeck, J.P.; Willerslev, E. Bayesian inference of the metazoan phylogeny: A combined molecular and morphological approach. *Curr. Biol.* **2004**, *14*, 1644–1649. [CrossRef] [PubMed]

67. O'Reilly, J.E.; Puttick, M.N.; Pisani, D.; Donoghue, P.C.J. Probabilistic methods surpass parsimony when assessing clade support in phylogenetic analyses of discrete morphological data. *Palaeontology* **2018**, *61*, 105–118. [CrossRef] [PubMed]

68. Wang, W.; Lu, A.M.; Ren, Y.; Endress, M.E.; Chen, Z.D. Phylogeny and classification of Ranunculales: Evidence from four molecular loci and morphological data. *Perspect. Plant Ecol. Evol. Syst.* **2009**, *11*, 81–110. [CrossRef]

69. Yu, S.X.; Janssens, S.B.; Zhu, X.Y.; Lidén, M.; Gao, T.G.; Wang, W. Phylogeny of Impatiens (Balsaminaceae): Integrating molecular and morphological evidence into a new classification. *Cladistics* **2016**, *32*, 179–197. [CrossRef] [PubMed]

70. Friis, E.M.; Pedersen, K.R.; Crane, P.R. The emergence of core eudicots: New floral evidence from the earliest Late Cretaceous. *Proc. R. Soc. B* **2016**, *283*, 20161325. [CrossRef]

71. Christiansen, P. Phylogeny of the great cats (Felidae: Pantherinae), and the influence of fossil taxa and missing characters. *Cladistics* **2008**, *24*, 977–992. [CrossRef]

72. Parham, J.F.; Donoghue, P.C.; Bell, C.J.; Calway, T.D.; Head, J.J.; Holroyd, P.A.; Inoue, J.G.; Irmis, R.B.; Joyce, W.G.; Ksepka, D.T.; et al. Best practices for justifying fossil calibrations. *Syst. Biol.* **2012**, *61*, 346–359. [CrossRef]

73. Xiang, X.; Xiang, K.; Ortiz, R.D.C.; Jabbour, F.; Wang, W. Integrating palaeontological and molecular data uncovers multiple ancient and recent dispersals in the pantropical Hamamelidaceae. *J. Biogeogr.* **2019**, *46*, 2622–2631. [CrossRef]

74. Pyron, R.A. Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Syst. Biol.* **2011**, *60*, 466–481. [CrossRef] [PubMed]

75. Palmer, J.D.; Jansen, R.K.; Michaels, H.J.; Chase, M.W.; Manhart, J.R. Chloroplast DNA variation and plant phylogeny. *Ann. Mo. Bot. Gard.* **1988**, *75*, 1180–1206. [CrossRef]

76. Goremykin, V.V.; Nikiforova, S.V.; Biggs, P.J.; Zhong, B.; Delange, P.; Martin, W.; Woetzel, S.; Atherton, R.A.; McLenachan, P.A.; Lockhart, P.J. The evolutionary root of flowering plants. *Syst. Biol.* **2013**, *62*, 50–61. [CrossRef]

77. Drew, B.T.; Ruhfel, B.R.; Smith, S.A.; Moore, M.J.; Briggs, B.G.; Gitzendanner, M.A.; Soltis, P.S.; Soltis, D.E. Another look at the root of the angiosperms reveals a familiar tale. *Syst. Biol.* **2014**, *63*, 368–382. [CrossRef]

78. Xi, Z.; Ruhfel, B.R.; Schaefer, H.; Amorim, A.M.; Sugumaran, M.; Wurdack, K.J.; Endress, P.K.; Matthews, M.L.; Stevens, P.F.; Mathews, S.; et al. Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17519–17524. [CrossRef]

79. McKain, M.R.; Johnson, M.G.; Uribe-Convers, S.; Eaton, D.; Yang, Y. Practical considerations for plant phylogenomics. *Appl. Plant Sci.* **2018**, *6*, e1038. [CrossRef]

80. Smith, S.A.; Beaulieu, J.M.; Donoghue, M.J. Mega-phylogeny approach for comparative biology: An alternative to supertree and supermatrix approaches. *BMC Evol. Biol.* **2009**, *9*, 37. [CrossRef]

81. Smith, S.A.; Brown, J.W.; Yang, Y.; Bruenn, R.; Drummond, C.P.; Brockington, S.F.; Walker, J.F.; Last, N.; Douglas, N.A.; Moore, M.J. Disparity, diversity, and duplications in the Caryophyllales. *New Phytol.* **2018**, *217*, 836–854. [CrossRef]

82. Thornhill, A.H.; Baldwin, B.G.; Freyman, W.A.; Nosratinia, S.; Kling, M.M.; Morueta-Holme, N.; Madsen, T.P.; Ackerly, D.D.; Mishler, B.D. Spatial phylogenetics of the native California flora. *BMC Biol.* **2017**, *15*, 96. [CrossRef]

83. Wu, S.D.; Zhang, L.J.; Lin, L.; Yu, S.X.; Chen, Z.D.; Wang, W. Insights into the historical assembly of global dryland floras: The diversification of Zygophyllaceae. *BMC Evol. Biol.* **2018**, *18*, 166. [CrossRef]

84. Eiserhardt, W.L.; Antonelli, A.; Bennett, D.J.; Botigué, L.R.; Burleigh, J.G.; Dodsworth, S.; Enquist, B.J.; Forest, F.; Kim, J.T.; Kozlov, A.M.; et al. A roadmap for global synthesis of the plant tree of life. *Am. J. Bot.* **2018**, *105*, 614–622. [CrossRef]