Review

Pervasive genome duplications across the plant tree of life and their links to major evolutionary innovations and transitions

Xin Qiao [a], Shaoling Zhang [a], Andrew H. Paterson [b],*

[a] Centre of Pear Engineering Technology Research, State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China
[b] Plant Genome Mapping Laboratory, University of Georgia, Athens, GA 30605, USA

A B S T R A C T

Whole-genome duplication (WGD) has occurred repeatedly during plant evolution and diversification, providing genetic layers for evolving new functions and phenotypes. Advances in long-read sequencing technologies have enabled sequencing and assembly of over 1000 plant genomes spanning nearly 800 species, in which a large set of ancient WGDs has been uncovered. Here, we review the recently reported WGDs that occurred in major plant lineages and key evolutionary positions, and highlight their contributions to morphological innovation and adaptive evolution. Current gaps and challenges in integrating enormous volumes of sequenced plant genomes, accurately inferring WGDs, and developing web-based analysis tools are emphasized. Looking to the future, ambitious genome sequencing projects and global efforts may substantially recapitulate the plant tree of life based on broader sampling of phylogenetic diversity, reveal much of the timetable of ancient WGDs, and address the biological significance of WGDs in plant adaptation and radiation.

© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

* Corresponding author.
  E-mail addresses: qiaoxin@njau.edu.cn (X. Qiao), nnzsl@njau.edu.cn (S. Zhang), paterson@uga.edu (A.H. Paterson).

1. Introduction

The evolution of green plants to become the world's dominant flora has been marked by prevalent whole-genome duplication (WGD) events [1–8]. WGD can abruptly double the entire set of chromosomes in the nucleus, conferring polyploidy [9]. Occurring by fusion of unreduced gametes or doubling of chromosomes in zygote [10,11], WGD has long been considered to provide 'layers' of redundant genetic information that can be co-opted for the evolution of genome complexity, phenotypic innovation and adaptability [12–20].

Genome sequencing propelled the identification of ancient WGD (paleopolyploidy) events in a range of taxa. Analysis of the first plant genome–*Arabidopsis thaliana*–revealed the remnants of three ancient WGDs referred to as the alpha (α)–beta (β)–gamma (γ) series shared by Brassicaceae [21]. Interestingly, rice, sorghum and other monocot genome analyses also unraveled three rounds of ancient WGDs predating the divergence of Poaceae species, termed the rho (ρ)–sigma (σ)–tau (τ) series [22,23]. Paleopolyploidy has not been limited to duplication, with cases of whole-genome triplication (WGT) and even penta-plication contributing to the astonishing genome complexity of plants [24–26]. About 24 %-35 % of extant vascular plants were inferred to be of recent polyploid origin, with 15 % of speciation events in angiosperms related to ploidy increase as well as 31 % in ferns [27–29]. Polyploidy has also been suggested to be associated with crop domestication, with domesticated crops (e.g. wheat, canola, cotton, peanut, coffee, oat) having experienced more polyploidy events than their wild relatives [30]. Understanding polyploid speciation and archival of its products may be of growing importance for maintaining plant diversity and global food security, given the expectation that plant extinction will exceed plant speciation in the near future [31].

Considerable progress in long-read sequencing technologies has propelled the global wave of genome sequencing with over 1000 plant genomes representing ∼ 800 species available now [32,33]. Remarkably, the quality of genome assembly has been elevated from 'draft' level to 'chromosome-level', and even to 'telomere-to-telomere' or 'gap-free' or 'complete' for many plants [34–40]. The large set of sequenced plant genomes provides valuable resources for deeply resolving ancient WGDs and tracing genome changes linked to major evolutionary transitions and variable adaptations [7].

Here, we detail WGDs identified in early-diverging seed-free plants and other major plant lineages. Ancient WGDs inferred at important phylogenetic positions were also highlighted, and their roles in phenotypic innovations, adaptive transitions, and species richness were addressed. The increasing challenges in processing and integrating plant genome data, and accurately identifying ancient WGDs are discussed. Collaborative research projects and global efforts remain required to fill current gaps in phylogenetic sampling, in inferring and positioning ancient WGDs across plant phylogeny and in deeply resolving the consequences of WGDs in plant adaptation and rapid expansion.

2. Paleopolyploidy events discovered across major plant lineages and their links to evolutionary innovations

2.1. Streptophyte algae–the closest sister lineage to land plants

Terrestrialization is a landmark of plant evolution and radiation, which had a profound influence on Earth's ecosystems [41–47]. The genome sequencing of streptophyte algae–the sister lineage to land plants (Embryophyta)–has reinforced our understanding of genome and gene changes of early plants during the evolutionary transition from water to land [48–53]. The WGD or polyploidy event is prevalent in land plant evolution, but is rare in green (chlorophyte and streptophyte) algae lineages (Fig. 1). Genome analysis of the earliest-diverging unicellular streptophyte algae, *Mesostigma viride*, showed evidence of WGD which contributed to relatively large genome size (442 Mb) and expansion of specific gene families involved in multicellularity and terrestrial pre-adaptation [53]. However, in a parallel study, evidence of WGD was not found in *M. viride*, perhaps due to fragmented and incomplete genome assembly with N50 = 113 Kb [52]. The occurrence of WGD or large-scale segmental duplications in basal streptophyte alga may be linked to the evolutionary transition from salt water to freshwater.

Intriguingly, a whole-genome triplication (WGT) event was deciphered in the unicellular algae *Spirogloea muscicola* with sub-aerial habitat, belonging to another lineage of streptophyte algae–Zygnematophyceae, and representing a closer living relative to land plants [49]. This WGT event increased the gene sets in *S. muscicola*, conferring genetic redundancy useful to adapt to sub-aerial stress and symbiotic interaction. Intragenomic synteny analysis also suggested that *S. muscicola* may have undergone a more ancient WGD (4dTv peak ∼ 0.85). Although WGD was not detected in *Penium margaritaceum* (1C = 4.7 Gb), the closest relative to land plants sequenced so far, a large number of segmental duplications was discovered within its genome [50]. The roles of WGD or WGT events in the molecular changes and evolution from freshwater to terrestrial environment remain to be deeply resolved.

2.2. Bryophytes–the pioneer of land plants

As the representative of early land plants, bryophytes including hornworts, liverworts and mosses, emerged about 460–506 Mya [54,55], and have aroused great interest to reveal key genome features involved in adaptation to terrestrial habitat [56–64]. Multiple ancient WGD events has been uncovered in the model moss *Physcomitrium patens* [60,65]. The remnants of ancient WGDs were also unraveled in other moss species including those from Sphagnales order, and three subclasses of Bryopsida (Timmiidae, Dicranidae, and Bryidae) [7,66]. The timing of WGD identified in moss *Calohypnum plumiforme* has been dated to 8 Mya (Ks ∼ 0.13–0.18) regardless of evolutionary rate correction [61]. The preferred retention of regulation- and stress-related genes after WGD in some moss genomes may have conferred the capability to survive adverse conditions such as cold, UV-B radiation, desiccation and heat [64,65,67]. In contrast, the currently sequenced species belonging to hornworts and liverworts lack evidence of WGD. The consequences of higher frequencies of WGDs in relation to the adaptation and diversification of mosses still need to be elucidated.

2.3. Lycophytes and Ferns–the early-diverging vascular plants

Lycophytes are thought to be representative of primitive vascular plants, and their origination has been dated to 392–451 Mya [68]. Strikingly, the evolution of a model lycophyte, *Selaginella moellendorffii*, has been influenced by at least two WGD events [69]. However, the contributions of repeated WGDs to genome changes and origin of the vascular system remain largely unclear. In contrast to desiccation sensitive *S. moellendorffii*, some *Selaginella* species showed extreme desiccation tolerance [70,71]. The roles of WGDs and subsequent species-specific diploidization in the divergent evolution of desiccation tolerance need to be investigated. In addition, an ancient WGD (Ks = 1.8) was revealed in the aquatic lycophyte *Isoetes taiwanensis*, implying its potential roles in the evolution of aquatic adaptation and CAM [72].

Ferns present great diversity with over 10,000 species, and generally possess large genome size and high chromosome number
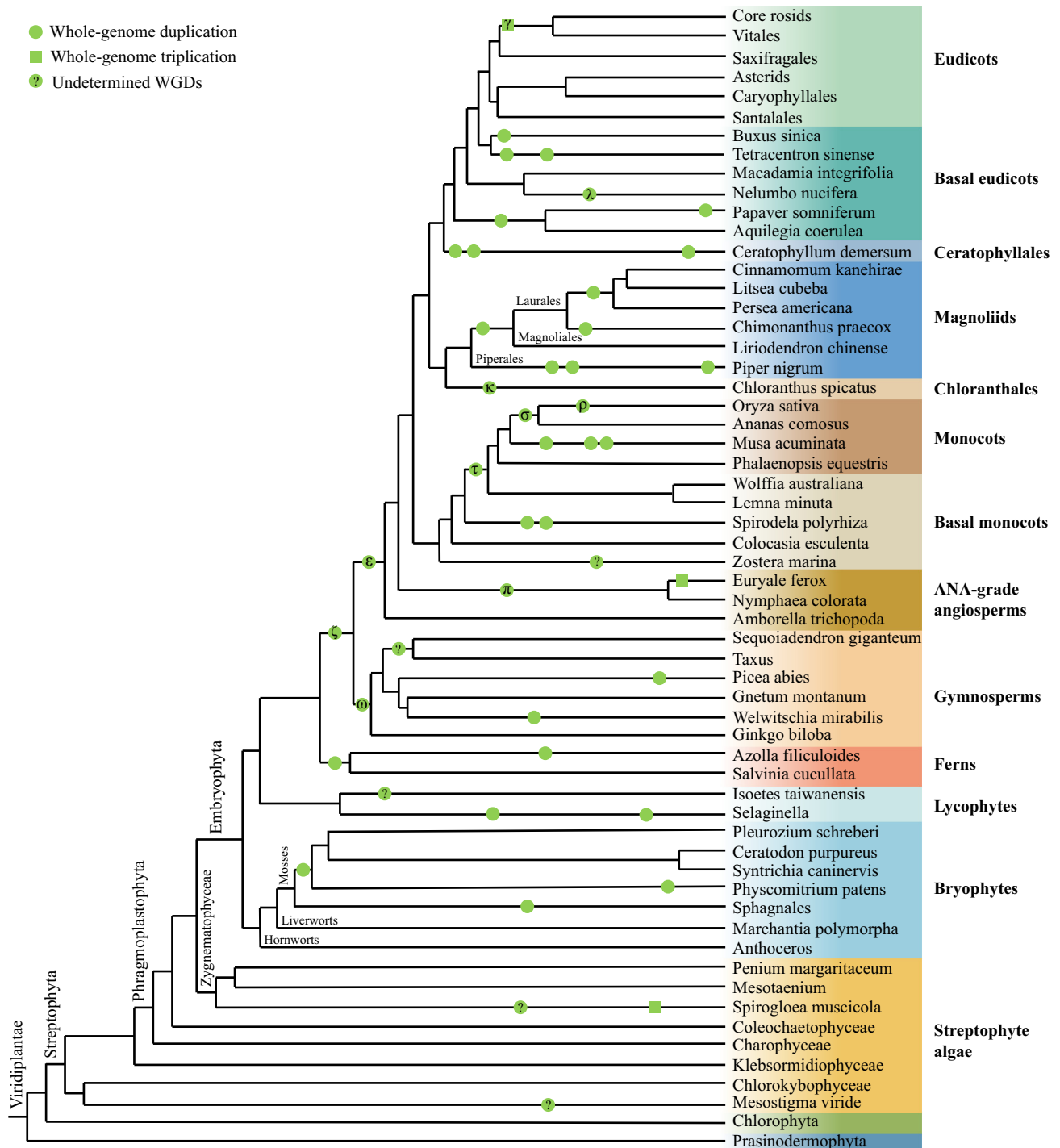
Fig. 1. The distribution of reported ancient WGDs across major lineages of green plants. WGDs (circle) and WGTs (square) were placed on the different branches of phylogenetic tree based on previous studies. The branch length is not proportional to the timescale. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

[73–75]. The mean genome size of ferns is estimated to be 14 Gb [76], and the most common chromosome number (n) is 41, compared to 9 in eudicots and 7 in monocots [74]. It has been hypothesized that recurrent WGD events have led to an unusual increase of chromosome number, genome size, and fern species richness [5,27,74,77,78]. Indeed, a high-frequency WGD events were unraveled across the fern phylogeny, with 66 %–97 % of extant ferns inferred to be derived from ancient polyploids [5,66]. Chromosome number analysis also suggested that over 30 % of speciation events

occurring in ferns are related to polyploidy events [27]. In contrast to their high species richness, only four fern genomes are available now, of which one is chromosomal-scale (*Alsophila spinulosa*) [79], two (*Azolla filiculoides* and *Salvinia cucullata*) are draft [80] and one is fragmented and incomplete (*Ceratopteris richardii*) [81]. The tree fern *A. spinulosa* was found to have experienced two ancient WGDs, the more recent one shared by Cyatheaceae species (termed 'Cyatheaceae WGD', Ks = 0.3), and the ancient one in the common ancestor of Cyatheales (termed 'Cyatheales WGD', Ks = 1.5) [79].

The large proportion of syntenic blocks are well preserved in genome of *A. spinulosa* following 'Cyatheaceae WGD' (~100 Mya), which was considered to be due to slow genome reshuffling and diploidization but responsible for the large genome size (6.2 Gb) and high chromosome number (n = 69). However, the associations between successive WGD events and the evolution of xylem and vascular tissues remain elusive. Genome analysis of *A. filiculoides* and *S. cucullata* revealed an ancient WGD occurring in the common ancestor of core leptosporangiates, and a more recent Azolla-specific WGD [80]. Analysis of the model fern *C. richardii* also revealed a lineage-specific WGD [81]. Future efforts are highly desired to improve the current fern genomes and sequence more fern species in poorly represented lineages, accelerating the investigation of genome changes related to the evolutionary switch between homospory and heterospory which is regarded to be important for the evolution of seed [82].

2.4. Gymnosperms–the appearance of seed

Gymnosperms are the pioneers of seed plants, and appeared around 300 Mya [68]. The genome size of gymnosperms is commonly large with average 18 Gb [76]. Gymnosperms also show longevity, and are prone to occur in harsh environments such as cold, mountainous areas at high altitudes [83,84]. Contradicting prior hypothesis that WGD is largely absent in gymnosperm lineages, important roles of WGDs in the evolution and radiation of gymnosperms have recently been suggested. Based on phylogenomic analysis of 24 transcriptomes, three ancient WGD events were revealed during the radiation of gymnosperms, with one in the common ancestor of Pinaceae, one in the common ancestor of Cupressaceae and Taxaceae (cupressophytes), and one in *Welwitschia mirabilis* [85]. The ancient WGD event (denoted as zeta, ζ) [1] inferred to have occurred in the common ancestor of seed plants was not shared with ferns [85]. However, analysis of chromosome-level genomes of three species from Taxaceae showed no evidence of recent WGD [86–88], countering the inference from [85]. The deeper investigation of WGDs in the chromosome-scale genome of *Sequoiadendron giganteum* (Cupressaceae) is required to verify the cupressophytes-specific WGD [89]. A broader sampling of gymnosperm transcriptomes uncovered a more ancient WGD occurring in the common ancestor of all extant gymnosperms and corroborated the Pinaceae-specific and *Welwitschia*-specific WGD event, but the contribution of WGDs to the diversification of gymnosperms was found to be limited [90]. *Welwitschia*-specific WGD was further confirmed and dated to ~86 Mya (Ks = ~1) based on genome analysis of *W. mirabilis* [91], but not discovered in assembling and analyzing a genome of the same species in another parallel study [92]. The chromosome-level genome assembly for *Cycas panzhihuaensis* (cycads) enables deep investigation of occurrence of ancient WGDs during gymnosperm evolution [93].

An ancient 'gymnosperm-wide WGD' (denoted as ω, Ks = ~0.85) was revealed based on a combination of phylogenomic and syntenic analysis of *C. panzhihuaensis* and other plant genomes, which contributed to the evolution of lineage-specific genes in gymnosperms. In addition, a Pinaceae-specific WGD (Ks = ~0.6) [90] was reconfirmed based on genome analysis of high-quality assembly of *Pinus tabuliformis* [94]. It has been suggested that duplicated genes provide important layers of redundancy for convergent evolution of conifer adaptation to cold tolerance [95] and morphological innovation [90]. Some questions remain to be resolved regarding potential links between polyploidy or small-scale gene duplication and stress tolerance adaptation of gymnosperms [84], the origin of protective layers surrounding seeds [96], and extreme longevity of some gymnosperm species [91,97].

2.5. Basal angiosperms–the starting point towards the largest plant group

Flowering plants (termed angiosperms), the crown-group of plants, have experienced an extraordinarily high-frequency of WGD events which shaped their genomes and morphological complexity [1,4,12,16,98,99]. WGDs has long been regarded as an important driving force to render the first appearance of flowering plants and subsequent rapid expansion during an incredibly short period of time, a phenomenon referred to as the "abominable mystery" by Darwin [100,101]. Over recent years, understanding of the early evolution of angiosperms has been enhanced owing to the genome sequencing of the earliest-diverging angiosperm lineages including Amborellales, Nymphaeales and Austrobaileyales, collectively referred to as ANA-grade angiosperms. The decoding of the Amborella genome showed no evidence of recent lineage-specific WGD but revealed an ancient polyploidy event (denoted as epsilon, ε) preceding the diversification of angiosperms [102]. Genome analysis of water lily (*Nymphaea colorata*) unveiled a lineage-specific WGD (denoted π, Ks = 0.9) shared by Nymphaeaceae, which has been linked to the evolution of key genes underpinning flower development and attractive features [103]. The roles of genome and gene duplication in the evolutionary retrogressions of water lily from terrestrial to aquatic environment and from woody to herbaceous habit are worthy of being explored [104,105]. Interestingly, prickly waterlily (*Euryale ferox*, Nymphaeaceae) experienced an additional WGT event after divergence with water lily [106], which was correlated to evolution of specific leaf development and aquatic adaptation [107].

2.6. Monocots

Monocots (termed monocotyledons) were estimated to have over 70,000 species [108,109], including economically important food crops and fruits such as wheat, rice, maize, barley, banana, pineapple, sugarcane. The major source of human nutrients and calories come from monocots. The genome sequencing and assembly of basal monocots open new door for studying the early evolution of monocots. The Lemnaceae family (known as duckweed) belonging to the Alismatales has a basal phylogenetic position relative to other monocot families, and evolved to readapt to aquatic habitat [110]. The high-quality genomes of several Lemnaceae species have been published including *Spirodela polyrhiza* [111,112], *Spirodela intermedia* [113], *Lemna minuta* [114], *Wolffia australiana* [115,116]. The *S. polyrhiza* has undergone two ancient WGDs (~95 Mya) but showed considerable reduction in genome size (~158 Mb) and gene number (~19,000) which may be associated with the simplification of its plant body (lacking stem) [117,118]. *Wolffia australiana* has a smaller and simpler plant body lacking roots, and has further reduced gene number (~15,000) but a larger genome (~400 Mb) with signature of an ancient WGD (Ks peak = ~1.6). An ancient WGD was also found in *L. minuta*, presenting a close Ks peak (~1.7) to that found in the sister species-*W. australiana*, implying that the tau (τ) WGD may have occurred before the divergence of Lemnaceae and other major monocot lineages [114]. The ancient WGDs and subsequent extensive gene loss and genome readjustment may have contributed to the evolution of specialized gene sets underpinning fast growth rate, specific body plan, and aquatic lifestyle of these tiny and basal monocots. Seagrass (Alismatales) is sister to Lemnaceae with basal position in monocot lineage, and is the only marine flowering plants. The genome sequencing of *Zostera marina* [119,120] and *Zostera muelleri* [121] has laid a solid foundation for better understanding of angiosperm adaption to the sea. The remnants of a lineage-specific WGD (64–72 Mya) were detected in *Zostera* genomes. The gene gain and loss have been linked to the degeneration and

acquisition of specific traits of seagrasses to readapt from freshwater to marine environment [122].

The evolution and diversification of major monocot groups has been influenced by recurrence of lineage-specific and species-specific WGDs [123–127]. Intragenomic and intergenomic synteny analysis of rice (*Oryza sativa*) and sorghum (*Sorghum bicolor*) genome revealed two ancient WGD events in the common ancestor of grass lineage (Poaceae), denoted as ρ and σ respectively [22,128]. A combination of phylogenomic and syntenic analysis further uncovered a more ancient WGD (denoted τ) common to commelinid monocots including Arecales, Poales, Zingiberales and Commelinales, terming as a WGD series of rho (ρ)-sigma (σ)-tau (τ) [23]. The ρ WGD was positioned before the divergence of the Pharoideae and the core grasses based on evolutionary analysis of *Pharus latifolius* genome [129]. The assembly of pineapple (*Ananas comosus*), and detailed synteny and phylogenetic analysis of pineapple and other monocot plants placed the σ WGD (100–120 Mya) before the divergence of Poales, and placed τ WGD (110–135 Mya) after the split of Alismatales and commelinids but before the split of commelinids and Asparagales [124,130]. The Orchidaceae (commonly called orchids) is one of the largest families of flowering plants and the largest family of monocots, with more than 20,000 species. The genome analysis of *Apostasia shenzhenica* [130] and *Vanilla planifolia* [127] jointly provided evidence for a 'Orchidaceae-wide' WGD (or pan-orchid α^O WGD), which was inferred to occur shortly before orchid radiation and to be associated with the great diversity of extant orchids.

2.7. Chloranthales and magnoliids

Genome and phylogenomic analysis of two Chloranthales species (*Chloranthus spicatus* and *C. sessilifolius*) provides evidence for an ancient WGD in the common ancestor of Chloranthales (denoted as kappa, κ), and supported Chloranthales as the sister lineage to magnoliids [131,132]. This Chloranthales-specific WGD was also resolved to have a large role in the evolution of high volatile content. Magnoliids, comprised of four orders (Laurales, Magnoliales, Piperales, and Canellales) and containing over 9,000 species, has attracted interest for genome sequencing due to its uncertain phylogenetic relationships to eudicots and monocots [133–135]. Interestingly, magnoliid evolution has been influenced by frequent WGD events. An ancient WGD event preceding the divergence of Laurales and Magnoliales [3,136] was termed lambda (λ) [2] although this symbol had been used to indicate the *Nelumbo*-specific WGD in a prior study [137]. Two Laurales families, Lauraceae and Calycanthaceae, each experienced an additional lineage-specific WGD event [136,138]. In addition, *Cassytha filiformis* (Lauranceae) underwent a more recent species-specific WGD [136]. A WGD was also suggested to take place in the common ancestor of Gomortegaceae and Atherospermataceae (Laurales) [136]. Remarkably, genome comparison with *Aristolochia fimbriata* indicated that *Piper nigrum* (black pepper, Piperales) has experienced three successive WGDs [3], contradicting a previous report of only one WGD [139]. Their high-frequency of WGDs has been recognized as an important driving force for the evolution of unique features of magnoliids species such as rich volatile compounds and specific floral morphology.

2.8. Ceratophyllales and eudicots

Ceratophyllales has been inferred as the sister lineage to all extant eudicots [106,131]. The relics of three WGD events were discovered in an aquatic Ceratophyllales species–rigid hornwort (*Ceratophyllum demersum*) [106]. Recently, genome sequences of a number of species from early-branching eudicot lineages or basal eudicots (Ranunculales, Proteales, Trochodendrales and Buxales)

shed light on the early evolution of eudicots [140]. *Aquilegia* (Ranunculaceae, Ranunculales) possesses a basal phylogenetic position relative to other eudicots, and has been corroborated to experience a paleo-tetraploidy event shared with *Coptis chinensis* (Ranunculaceae) [141], although incongruent views about the placement of this event have been proposed [142–145]. Papaveraceae (Ranunculales) is the closest sister lineage to Ranunculaceae. Genome analysis of three *Papaver* (Papaveraceae) species showed that *P. setigerum* shared a WGD with *P. somniferum* (opium poppy) (∼7.2 Mya), and experienced an additional more recent WGD (∼4 Mya) after its split with opium poppy [146]. These WGD events play important roles in evolution of the biosynthetic pathway of benzylisoquinoline alkaloids (BIAs) which are the medically important constituents of opium poppy. *Akebia trifoliata* (Lardizabalaceae, Ranunculales), a successor lineage to *Aquilegia*, showed evidence of a lineage-specific WGD, but whether this event is shared with *Aquilegia* remains unclear [147]. Another Ranunculales species *Kingdonia uniflora* (Circaeasteraceae) likewise presented evidence of WGD but with phylogenetic placement remaining to be determined [148]. Genome sequencing of sacred lotus (*Nelumbo nucifera*) [137] and *Macadamia integrifolia* [149], belonging to Proteales, indicated that they have experienced independent lineage-specific WGDs. As the closest and successive lineages to core eudicots, Trochodendrales (*Tetracentron sinense*) and Buxales (*Buxus sinica*) genomes analysis revealed that they each underwent one or more lineage-specific WGDs [140,150,151]. More WGDs discovered in core eudicots have been well documented [2–4,8,18,19].

3. Links between paleopolyploidy and plant survival in extreme environments

Ancient WGDs have been found in some extremophytes such as halophytes [152], resurrection plants [60,153], alpine plants [154–156], desert plants [64,157,158], and invasive plants [159–161], suggesting important roles of WGD in the evolution of plant adaptation to extreme environments [64,162]. Interestingly, the occurrence of paleopolyploidy events has been suggested to coincide with periods of environmental upheaval, such as mass extinction events or extreme climate change [4,17]. Several 'waves' of WGD have been suggested to be associated with different epochs or geologic time periods, including the Cretaceous–Paleogene (K–Pg) extinction event (∼66 Mya) [4,163], early Aptian epoch within the Cretaceous period (∼120 Mya), and Neogene period (23.0–2.6 Mya) [2,18]. Moreover, some genetic evidence supports significant contributions of ancient WGDs to plant adaptation and survival under abrupt climate changes [18]. Duplicated genes involved in resistance to cold, heat, darkness, and other adverse conditions were preferentially retained after independent WGDs near period of dramatic climate change [18,164,165]. In addition, polyploidy has been found with higher frequency in colder and higher latitude regions, consistent with the notion that polyploids have a propensity to survive in and inhabit extreme and fluctuating environments [166].

4. Summary and outlook

Over 1000 genomes representing 798 plant species have been sequenced and assembled in the past twenty years [32,33]. A large number of ancient WGD events have been deciphered among diverse lineages [4,18,19], with 244 inferred and placed across plant phylogeny based on analysis of 1000 plant transcriptomes [7,8]. Nonetheless, resolving dates and phylogenetic placements of ancient WGDs remains challenging due to plant genome complexity. For instance, reanalysis of available genomes of Cucurbitaceae species revealed a paleo-tetraploidy predating the

divergence of Cucurbitaceae that was overlooked in previous studies [167]. Reinvestigation confirmed that the WGD found in durian (*Durio zibethinus*) is lineage-specific [168] rather than shared with cotton [169]. Moreover, new WGD events were uncovered based on reanalysis of *Selaginella moellendorffii* [69] and *Aquilaria sinensis* [170] genomes respectively. More effort remains necessary to reconfirm reported WGDs and rediscover overlooked WGDs based on high-quality and chromosome-level genomes.

Driven by the rapid development of long-read sequencing technologies (e.g. PacBio HiFi, Nanopore), telomere-to-telomere and gap-free genome assemblies have been produced for some model plants and important crops such as Arabidopsis [37,39], tomato [40], maize [34], rice [36,38] and banana [35]. These highly contiguous genome assemblies provide a valuable opportunity to perform more accurate genome comparisons and synteny analysis, refine the resolution of evolutionary analysis, and recheck ancient WGDs reported in previous studies. Moreover, growing software resources have been developed to infer date and place ancient WGDs across plant phylogeny such as MAPS [85], Whale [171], Beluga [172], wgd [173], WGDI [145], in which synteny-based or phylogenomic-based strategies are implemented. A combination of synteny- and phylogenomic-based methods to correctly identify and verify ancient polyploidy events has an important role in ongoing and upcoming genome sequencing projects.

Several ambitious projects including the Earth BioGenome Project [174], Darwin Tree of Life Project [175], and 10KP [176] have been launched to sequence the genomes of all eukaryotic species on earth, and eventually to document the biodiversity on our planet. Such efforts pose considerable challenges for integrating enormous plant genome data and developing interactive web tools for performing genome comparison and synteny analysis, identifying genome and gene duplications, and visualizing results. Some pioneering databases including Phytozome [177], Ensembl Plants [178], PLAZA [179], CoGe [180], PGDD [181] incorporating genome integration, synteny analysis, WGD inference and comparative genomic analysis tools, have facilitated research in plant genome evolution. Nonetheless, the capacity of current computational tools and storage systems remain limited toward dealing with the huge volumes of 400,000 plant genomes [109]. Apart from well-known databases such as NCBI, EMBL-EBI and DDBJ, some new platforms have been built to provide prompt and convenient access to sequenced plant genomes such as Genome Warehouse [182], CNGBdb (https://db.cngb.org) and Plant GARDEN (https://plantgarden.jp/en/index).

Moving forward, global efforts by the plant science community may substantially reconstruct the genome-level green plant phylogeny based on broader sampling and sequencing of plant diversity, and anchor/position WGDs on different branches along the evolutionary timeline.

CRediT authorship contribution statement

Xin Qiao: Conceptualization, Writing – original draft, Writing – review & editing, Funding acquisition. **Shaoling Zhang:** Writing – review & editing, Supervision, Funding acquisition. **Andrew H. Paterson:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, et al. Ancestral polyploidy in seed plants and angiosperms. Nature 2011;473:97–100.

[2] Zhang L, Wu S, Chang X, Wang X, Zhao Y, Xia Y, et al. The ancient wave of polyploidization events in flowering plants and their facilitated adaptation to environmental stress. Plant Cell Environ 2020;43:2847–56.

[3] Qin L, Hu Y, Wang J, Wang X, Zhao R, Shan H, et al. Insights into angiosperm evolution, floral development and chemical biosynthesis from the Aristolochia fimbriata genome. Nat Plants 2021;7:1239–53.

[4] Van de Peer Y, Mizrachi E, Marchal K. The evolutionary significance of polyploidy. Nat Rev Genet 2017;18:411–24.

[5] Huang CH, Qi X, Chen D, Qi J, Ma H. Recurrent genome duplication events likely contributed to both the ancient and recent rise of ferns. J Integr Plant Biol 2020;62:433–55.

[6] Qiao X, Li Q, Yin H, Qi K, Li L, Wang R, et al. Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. Genome Biol 2019;20:38.

[7] Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA, Graham SW, et al. One thousand plant transcriptomes and the phylogenomics of green plants. Nature 2019;574:679–85.

[8] Li Z, Barker MS. Inferring putative ancient whole-genome duplications in the 1000 Plants (1KP) initiative: access to gene family phylogenies and age distributions. GigaScience 2020;9. https://doi.org/10.1093/gigascience/giaa004.

[9] Yant L, Bomblies K. Genome management and mismanagement—cell-level opportunities and challenges of whole-genome duplication. Genes Dev 2015;29:2405–19.

[10] Moghe GD, Shiu SH. The causes and molecular consequences of polyploidy in flowering plants. Ann N Y Acad Sci 2014;1320:16–34.

[11] Kreiner JM, Kron P, Husband BC. Evolutionary Dynamics of Unreduced Gametes. Trends Genet 2017;33:583–93.

[12] Landis JB, Soltis DE, Li Z, Marx HE, Barker MS, Tank DC, et al. Impact of whole-genome duplication events on diversification rates in angiosperms. Am J Bot 2018;105:348–63.

[13] Roman-Palacios C, Molina-Henao YF, Barker MS. Polyploids increase overall diversity despite higher turnover than diploids in the Brassicaceae. Proc Biol Sci 2020;287:20200962.

[14] Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng CF, et al. Polyploidy and angiosperm diversification. Am J Bot 2009;96:336–48.

[15] Soltis DE, Visger CJ, Soltis PS. The polyploidy revolution then...and now: Stebbins revisited. Am J Bot 2014;101:1057–78.

[16] Soltis PS, Soltis DE. Ancient WGD events as drivers of key innovations in angiosperms. Curr Opin Plant Biol 2016;30:159–65.

[17] Van de Peer Y, Ashman TL, Soltis PS, Soltis DE. Polyploidy: an evolutionary and ecological force in stressful times. Plant Cell 2021;33:11–26.

[18] Wu S, Han B, Jiao Y. Genetic Contribution of Paleopolyploidy to Adaptive Evolution in Angiosperms. Mol Plant 2020;13:59–71.

[19] Ren R, Wang H, Guo C, Zhang N, Zeng L, Chen Y, et al. Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. Mol Plant 2018;11:414–28.

[20] Jiao Y, Paterson AH. Polyploidy-associated genome modifications during land plant evolution. Philos Trans R Soc Lond B Biol Sci 2014;369:20130355.

[21] Bowers JE, Chapman BA, Rong J, Paterson AH. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 2003;422:433–8.

[22] Tang HB, Bowers JE, Wang XY, Paterson AH. Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. PNAS 2010;107:472–7.

[23] Jiao Y, Li J, Tang H, Paterson AH. Integrated syntenic and phylogenomic analyses reveal an ancient genome duplication in monocots. Plant Cell 2014;26:2792–802.

[24] Wang X, Guo H, Wang J, Lei T, Liu T, Wang Z, et al. Comparative genomic deconvolution of the cotton genome revealed a decaploid ancestor and widespread chromosomal fractionation. New Phytol 2016;209:1252–63.

[25] Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, et al. The genome of the mesopolyploid crop species Brassica rapa. Nat Genet 2011;43:1035–9.

[26] Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IA, et al. The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes. Nat Commun 2014;5:3930.

[27] Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH. The frequency of polyploid speciation in vascular plants. Proc Natl Acad Sci U S A 2009;106:13875–9.

[28] Mayrose I, Zhan SH, Rothfels CJ, Magnuson-Ford K, Barker MS, Rieseberg LH, et al. Recently formed polyploid plants diversify at lower rates. Science 2011;333:1257.

[29] Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA. On the relative abundance of autopolyploids and allopolyploids. New Phytol 2016;210:391–8.

[30] Salman-Minkov A, Sabath N, Mayrose I. Whole-genome duplication as a key factor in crop domestication. Nat Plants 2016;2:16115.

[31] Gao JG, Liu H, Wang N, Yang J, Zhang XL. Plant extinction excels plant speciation in the Anthropocene. BMC Plant Biol 2020;20:430.

[32] Marks RA, Hotaling S, Frandsen PB, VanBuren R. Representation and participation across 20 years of plant genome sequencing. Nat Plants 2021;7:1571–8.

[33] Sun Y, Shang L, Zhu QH, Fan L, Guo L. Twenty years of plant genome sequencing: achievements and challenges. Trends Plant Sci 2022;27:391–401.

[34] Liu J, Seetharam AS, Chougule K, Ou S, Swentowsky KW, Gent JI, et al. Gapless assembly of maize chromosomes using long-read technologies. Genome Biol 2020;21:121.

[35] Belser C, Baurens FC, Noel B, Martin G, Cruaud C, Istace B, et al. Telomere-to-telomere gapless chromosomes of banana using nanopore sequencing. Commun Biol 2021;4:1047.

[36] Li K, Jiang W, Hui Y, Kong M, Feng LY, Gao LZ, et al. Gapless indica rice genome reveals synergistic contributions of active transposable elements and segmental duplications to rice genome evolution. Mol Plant 2021. https://doi.org/10.1016/j.molp.2021.06.017.

[37] Naish M, Alonge M, Wlodzimierz P, Tock AJ, Abramson BW, Schmücker A, et al. The genetic and epigenetic landscape of the Arabidopsis centromeres. Science 2021;374:eabi7489.

[38] Song JM, Xie WZ, Wang S, Guo YX, Koo DH, Kudrna D, et al. Two gap-free reference genomes and a global view of the centromere architecture in rice. Mol Plant 2021. https://doi.org/10.1016/j.molp.2021.06.018.

[39] Wang B, Yang X, Jia Y, Xu Y, Jia P, Dang N, et al. High-quality Arabidopsis thaliana genome assembly with nanopore and HiFi long reads. Genom. Proteom. Bioinform. 2021. https://doi.org/10.1016/j.gpb.2021.08.003.

[40] van Rengs WMJ, Schmidt MH, Effgen S, Le DB, Wang Y, Zaidan M, et al. A chromosome scale tomato genome built from complementary PacBio and Nanopore sequences alone reveals extensive linkage drag during breeding. Plant J 2022. https://doi.org/10.1111/tpj.15690.

[41] Gao JG. Tracking the evolutionary innovations of plant terrestrialization. Gene 2021;769:145203.

[42] Rensing SA. How plants conquered land. Cell 2020;181:964–6.

[43] Donoghue P, Paps J. Plant evolution: assembling land plants. Curr Biol 2020;30:R81–3.

[44] Martin WF, Allen JF. An algal greening of land. Cell 2018;174:256–8.

[45] Furst-Jansen JMR, de Vries S, de Vries J. Evo-physio: on stress responses and the earliest land plants. J Exp Bot 2020;71:3254–69.

[46] de Vries J, Archibald JM. Plant evolution: landmarks on the path to terrestrial life. New Phytol 2018;217:1428–34.

[47] Rensing SA. Great moments in evolution: the conquest of land by plants. Curr Opin Plant Biol 2018;42:49–54.

[48] Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, Ullrich KK, et al. The Chara Genome: Secondary Complexity and Implications for Plant Terrestrialization. Cell 2018;174(448–464):e24.

[49] Cheng S, Xian W, Fu Y, Marin B, Keller J, Wu T, et al. Genomes of subaerial zygnematophyceae provide insights into land plant evolution. Cell 2019;179 (1057–1067):e14.

[50] Jiao C, Sorensen I, Sun X, Sun H, Behar H, Alseekh S, et al. The penium margaritaceum genome: hallmarks of the origins of land plants. Cell 2020;181(1097–1111):e12.

[51] Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, et al. Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. Nat Commun 2014;5:3978.

[52] Wang S, Li L, Li H, Sahu SK, Wang H, Xu Y, et al. Genomes of early-diverging streptophyte algae shed light on plant terrestrialization. Nat Plants 2020;6:95–106.

[53] Liang Z, Geng Y, Ji C, Du H, Wong CE, Zhang Q, et al. Mesostigma viride genome and transcriptome provide insights into the origin and evolution of streptophyta. Adv Sci (Weinh) 2020;7:1901850.

[54] Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, et al. The timescale of early land plant evolution. Proc Natl Acad Sci U S A 2018;115:E2274–83.

[55] Su D, Yang L, Shi X, Ma X, Zhou X, Hedges SB, et al. Large-scale phylogenomic analyses reveal the monophyly of bryophytes and neoproterozoic origin of land plants. Mol Biol Evol 2021;38:3332–44.

[56] Li FW, Nishiyama T, Waller M, Frangedakis E, Keller J, Li Z, et al. Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. Nat Plants 2020;6:259–72.

[57] Zhang J, Fu X-X, Li R-Q, Zhao X, Liu Y, Li M-H, et al. The hornwort genome and early land plant evolution. Nat Plants 2020. https://doi.org/10.1038/s41477-019-0588-4.

[58] Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, Ishizaki K, et al. Insights into land plant evolution garnered from the marchantia polymorpha genome. Cell 2017;171:287–304.e15.

[59] Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, et al. The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science 2008;319:64–9.

[60] Lang D, Ullrich KK, Murat F, Fuchs J, Jenkins J, Haas FB, et al. The Physcomitrella patens chromosome-scale assembly reveals moss genome structure and evolution. Plant J 2018;93:515–33.

[61] Mao L, Kawaide H, Higuchi T, Chen M, Miyamoto K, Hirata Y, et al. Genomic evidence for convergent evolution of gene clusters for momilactone biosynthesis in land plants. Proc Natl Acad Sci U S A 2020;117:12472–80.

[62] Pederson ERA, Warshan D, Rasmussen U. Genome sequencing of Pleurozium schreberi: The assembled and annotated draft genome of a pleurocarpous feather moss. G3 (Bethesda) 2019;9:2791–7.

[63] Yu J, Li L, Wang S, Dong S, Chen Z, Patel N, et al. Draft genome of the aquatic moss Fontinalis antipyretica (Fontinalaceae, Bryophyta). Gigabyte 2020;2020:1–9.

[64] Silva AT, Gao B, Fisher KM, Mishler BD, Ekwealor JTB, Stark LR, et al. To dry perchance to live: Insights from the genome of the desiccation-tolerant biocrust moss Syntrichia caninervis. Plant J 2021;105:1339–56.

[65] Gao B, Chen M, Li X, Liang Y, Zhang D, Wood AJ, et al. Ancestral gene duplications in mosses characterized by integrated phylogenomic analyses. J Syst Evol 2020. https://doi.org/10.1111/jse.12683.

[66] Szovenyi P, Gunadi A, Li FW. Charting the genomic landscape of seed-free plants. Nat Plants 2021. https://doi.org/10.1038/s41477-021-00888-z.

[67] Liu S, Fang S, Liu C, Zhao L, Cong B, Zhang Z. Transcriptomics integrated with metabolomics reveal the effects of ultraviolet-B radiation on flavonoid biosynthesis in antarctic moss. Front Plant Sci 2021;12:788377.

[68] Barba-Montoya J, Dos Reis M, Schneider H, Donoghue PCJ, Yang Z. Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous Terrestrial Revolution. New Phytol 2018;218:819–34.

[69] Wang J, Yu J, Sun P, Li C, Song X, Lei T, et al. Paleo-polyploidization in Lycophytes. Genom Proteom Bioinform 2020;18:333–40.

[70] VanBuren R, Wai CM, Ou S, Pardo J, Bryant D, Jiang N, et al. Extreme haplotype variation in the desiccation-tolerant clubmoss Selaginella lepidophylla. Nat Commun 2018;9:13.

[71] Xu Z, Xin T, Bartels D, Li Y, Gu W, Yao H, et al. Genome analysis of the ancient tracheophyte selaginella tamariscina reveals evolutionary features relevant to the acquisition of desiccation tolerance. Mol Plant 2018;11:983–94.

[72] Wickell D, Kuo LY, Yang HP, Dhabalia Ashok A, Irisarri I, Dadras A, et al. Underwater CAM photosynthesis elucidated by Isoetes genome. Nat Commun 2021;12:6348.

[73] I, PPG. A community-derived classification for extant lycophytes and ferns. J System Evol 2016;54:563–603.

[74] Rice A, Glick L, Abadi S, Einhorn M, Kopelman NM, Salman-Minkov A, et al. The Chromosome Counts Database (CCDB) - a community resource of plant chromosome numbers. New Phytol 2015;206:19–26.

[75] Pellicer J, Leitch IJ. The plant DNA C-values database (release 7.1): an updated online repository of plant genome size data for comparative studies. New Phytol 2020;226:301–5.

[76] Pellicer J, Hidalgo O, Dodsworth S, Leitch IJ. Genome size diversity and its impact on the evolution of land plants. Genes (Basel) 2018;9. https://doi.org/10.3390/genes9020088.

[77] Fujiwara T, Liu H, Meza-Torres EI, Morero RE, Vega AJ, Liang Z, et al. Evolution of genome space occupation in ferns: linking genome diversity and species richness. Ann Bot 2021. https://doi.org/10.1093/aob/mcab094.

[78] Hidalgo O, Pellicer J, Christenhusz M, Schneider H, Leitch AR, Leitch IJ. Is there an upper limit to genome size? Trends Plant Sci 2017;22:567–73.

[79] Huang X, Wang W, Gong T, Wickell D, Kuo LY, Zhang X, et al. The flying spider-monkey tree fern genome provides insights into fern evolution and arborescence. Nat Plants 2022. https://doi.org/10.1038/s41477-022-01146-6.

[80] Li FW, Brouwer P, Carretero-Paulet L, Cheng S, de Vries J, Delaux PM, et al. Fern genomes elucidate land plant evolution and cyanobacterial symbioses. Nat Plants 2018;4:460–72.

[81] Marchant DB, Sessa EB, Wolf PG, Heo K, Barbazuk WB, Soltis PS, et al. The C-Fern (Ceratopteris richardii) genome: insights into plant genome evolution with the first partial homosporous fern genome assembly. Sci Rep 2019;9:18181.

[82] Banks JA. Fern genomes finally here. Nat Plants 2018;4:404–5.

[83] Jin WT, Gernandt DS, Wehenkel C, Xia XM, Wei XX, Wang XQ. Phylogenomic and ecological analyses reveal the spatiotemporal evolution of global pines. Proc Natl Acad Sci U S A 2021;118. https://doi.org/10.1073/pnas.2022302118.

[84] Hancock AM. How conifers adapt to the cold. Science 2016;353:1362–3.

[85] Li Z, Baniaga AE, Sessa EB, Scascitelli M, Graham SW, Rieseberg LH, et al. Early genome duplications in conifers and other seed plants. Sci Adv 2015;1:e1501084.

[86] Cheng J, Wang X, Liu X, Zhu X, Li Z, Chu H, et al. Chromosome-level genome of Himalayan yew provides insights into the origin and evolution of the paclitaxel biosynthetic pathway. Mol Plant 2021;14:1199–209.

[87] Xiong X, Gou J, Liao Q, Li Y, Zhou Q, Bi G, et al. The Taxus genome provides insights into paclitaxel biosynthesis. Nat Plants 2021;7:1026–36.

[88] Song C, Fu F, Yang L, Niu Y, Tian Z, He X, et al. Taxus yunnanensis genome offers insights into gymnosperm phylogeny and taxol production. Commun Biol 2021;4. https://doi.org/10.1038/s42003-021-02697-8.

[89] Scott AD, Zimin AV, Puiu D, Workman R, Britton M, Zaman S, et al. A reference genome sequence for Giant Sequoia. G3 (Bethesda) 2020;10:3907–19.

[90] Stull GW, Qu XJ, Parins-Fukuchi C, Yang YY, Yang JB, Yang ZY, et al. Gene duplications and phylogenomic conflict underlie major pulses of phenotypic evolution in gymnosperms. Nat Plants 2021;7:1015–25.

[91] Wan T, Liu Z, Leitch IJ, Xin H, Maggs-Kölling G, Gong Y, et al. The Welwitschia genome reveals a unique biology underpinning extreme longevity in deserts. Nat. Commun 2021;12. https://doi.org/10.1038/s41467-021-24528-4.

[92] Han Y, Zhang W, Zhou B, Zeng P, Tian Z, Cai J. Chromosome-level genome assembly of Welwitschia mirabilis, a unique Namib Desert species. Mol Ecol Resour 2022;22:391–403.

[93] Liu Y, Wang S, Li L, Yang T, Dong S, Wei T, et al. The Cycas genome and the early evolution of seed plants. Nat Plants 2022;8:389–401.

[94] Niu S, Li J, Bo W, Yang W, Zuccolo A, Giacomello S, et al. The Chinese pine genome and methylome unveil key features of conifer evolution. Cell 2022;185(204–217):e14.

[95] Yeaman S, Hodgins KA, Lotterhos KE, Suren H, Nadeau S, Degner JC, et al. Convergent local adaptation to climate in distantly related conifers. Science 2016;353:1431–3.

[96] Soltis DE. Ancient seeds spill secrets about the evolution of flowering plants. Nature 2021;594:185–6.

[97] Bowers JE, Paterson AH. Chromosome number is key to longevity of polyploid lineages. New Phytol 2021. https://doi.org/10.1111/nph.17361.

[98] Jiao Y, Leebens-Mack J, Ayyampalayam S, Bowers JE, McKain MR, McNeal J, et al. A genome triplication associated with early diversification of the core eudicots. Genome Biol 2012;13:R3.

[99] Murat F, Armero A, Pont C, Klopp C, Salse J. Reconstructing the genome of the most recent common ancestor of flowering plants. Nat Genet 2017;49:490–6.

[100] Friedman WE. The meaning of Darwin's "abominable mystery". Am J Bot 2009;96:5–21.

[101] Buggs RJA. The origin of Darwin's "abominable mystery". Am J Bot 2021. https://doi.org/10.1002/ajb2.1592.

[102] Albert VA, Barbazuk WB, Der JP, Leebens-Mack J, Ma H, Palmer JD, et al. The Amborella genome and the evolution of flowering plants. Science 2013;342:1241089.

[103] Zhang L, Chen F, Zhang X, Li Z, Zhao Y, Lohaus R, et al. The water lily genome and the early evolution of flowering plants. Nature 2020;577:79–84.

[104] Tang H, Zhang L, Chen F, Zhang X, Chen F, Ma H, et al. Nymphaea colorata (Blue-Petal Water Lily). Trends Genet 2020;36:718–9.

[105] Povilus RA, DaCosta JM, Grassa C, Satyaki PRV, Moeglein M, Jaenisch J, et al. Water lily (Nymphaea thermarum) genome reveals variable genomic signatures of ancient vascular cambium losses. Proc Natl Acad Sci U S A 2020;117:8649–56.

[106] Yang Y, Sun P, Lv L, Wang D, Ru D, Li Y, et al. Prickly waterlily and rigid hornwort genomes shed light on early angiosperm evolution. Nat Plants 2020. https://doi.org/10.1038/s41477-020-0594-6.

[107] Wu P, Zhang L, Zhang K, Yin Y, Liu A, Zhu Y, et al. The adaptive evolution of Euryale ferox to the aquatic environment through paleo-hexaploidization. Plant J 2022. https://doi.org/10.1111/tpj.15717.

[108] Christenhusz MJM, Byng JW. The number of known plants species in the world and its annual increase. Phytotaxa 2016;261. https://doi.org/10.11646/phytotaxa.261.3.1.

[109] Lughadha EN, Govaerts R, Belyaeva I, Black N, Lindon H, Allkin R, et al. Counting counts: revised estimates of numbers of accepted species of flowering plants, seed plants, vascular plants and land plants with a review of other recent estimates. Phytotaxa 2016;272. https://doi.org/10.11646/phytotaxa.272.1.5.

[110] Acosta K, Appenroth KJ, Borisjuk L, Edelman M, Heinig U, Jansen MAK, et al. Return of the Lemnaceae: duckweed as a model plant system in the genomics and postgenomics era. Plant Cell 2021;33:3207–34.

[111] Harkess A, McLoughlin F, Bilkey N, Elliott K, Emenecker R, Mattoon E, et al. Improved Spirodela polyrhiza genome and proteomic analyses reveal a conserved chromosomal structure with high abundance of chloroplastic proteins favoring energy production. J Exp Bot 2021;72:2491–500.

[112] An D, Zhou Y, Li C, Xiao Q, Wang T, Zhang Y, et al. Plant evolution and environmental adaptation unveiled by long-read whole-genome sequencing of Spirodela. Proc Natl Acad Sci U S A 2019;116:18893–9.

[113] Hoang PTN, Fiebig A, Novak P, Macas J, Cao HX, Stepanenko A, et al. Chromosome-scale genome assembly for the duckweed Spirodela intermedia, integrating cytogenetic maps, PacBio and Oxford Nanopore libraries. Sci Rep 2020;10:19230.

[114] Abramson BW, Novotny M, Hartwick NT, Colt K, Aevermann BD, Scheuermann RH, et al. The genome and preliminary single-nuclei transcriptome of Lemna minuta reveals mechanisms of invasiveness. Plant Physiol 2022;188:879–97.

[115] Michael TP, Ernst E, Hartwick N, Chu P, Bryant D, Gilbert S, et al. Genome and time-of-day transcriptome of Wolffia australiana link morphological minimization with gene loss and less growth control. Genome Res 2021;31:225–38.

[116] Park H, Park JH, Lee Y, Woo DU, Jeon HH, Sung YW, et al. Genome of the world's smallest flowering plant, Wolffia australiana, helps explain its specialized physiology and unique morphology. Commun Biol 2021;4:900.

[117] Wang W, Haberer G, Gundlach H, Glasser C, Nussbaumer T, Luo MC, et al. The Spirodela polyrhiza genome reveals insights into its neotenous reduction fast growth and aquatic lifestyle. Nat Commun 2014;5:3311.

[118] Hoang PNT, Michael TP, Gilbert S, Chu P, Motley ST, Appenroth KJ, et al. Generating a high-confidence reference genome map of the Greater Duckweed by integration of cytogenomic, optical mapping, and Oxford Nanopore technologies. Plant J 2018;96:670–84.

[119] Olsen JL, Rouze P, Verhelst B, Lin YC, Bayer T, Collen J, et al. The genome of the seagrass Zostera marina reveals angiosperm adaptation to the sea. Nature 2016;530:331–5.

[120] Ma X, Olsen JL, Reusch TBH, Procaccini G, Kudrna D, Williams M, et al. Improved chromosome-level genome assembly and annotation of the seagrass, Zostera marina (eelgrass). F1000Research 2021;10. https://doi.org/10.12688/f1000research.38156.1.

[121] Lee H, Golicz AA, Bayer PE, Jiao Y, Tang H, Paterson AH, et al. The Genome of a Southern Hemisphere Seagrass Species (Zostera muelleri). Plant Physiol 2016;172:272–83.

[122] Williams SL. From sea to sea. Nature 2016;530:290–1.

[123] Wang S, Xiao Y, Zhou ZW, Yuan J, Guo H, Yang Z, et al. High-quality reference genome sequences of two coconut cultivars provide insights into evolution of monocot chromosomes and differentiation of fiber content and plant height. Genome Biol 2021;22:304.

[124] Ming R, VanBuren R, Wai CM, Tang H, Schatz MC, Bowers JE, et al. The pineapple genome and the evolution of CAM photosynthesis. Nat Genet 2015;47:1435–42.

[125] Peng Z, Lu Y, Li L, Zhao Q, Feng Q, Gao Z, et al. The draft genome of the fast-growing non-timber forest species moso bamboo (Phyllostachys heterocycla). Nat Genet 2013;45:456–61,:461e1-2.

[126] Harkess A, Zhou J, Xu C, Bowers JE, Van der Hulst R, Ayyampalayam S, et al. The asparagus genome sheds light on the origin and evolution of a young Y chromosome. Nat Commun 2017;8:1279.

[127] Hasing T, Tang H, Brym M, Khazi F, Huang T, Chambers AH. A phased Vanilla planifolia genome enables genetic improvement of flavour and production. Nature Food 2020;1:811–9.

[128] Paterson AH, Bowers JE, Chapman BA. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci U S A 2004;101:9903–8.

[129] Ma PF, Liu YL, Jin GH, Liu JX, Wu H, He J, et al. The Pharus latifolius genome bridges the gap of early grass evolution. Plant Cell 2021;33:846–64.

[130] Zhang GQ, Liu KW, Li Z, Lohaus R, Hsiao YY, Niu SC, et al. The Apostasia genome and the evolution of orchids. Nature 2017;549:379–83.

[131] Guo X, Fang D, Sahu SK, Yang S, Guang X, Folk R, et al. Chloranthus genome provides insights into the early diversification of angiosperms. Nature. Communications 2021;12. https://doi.org/10.1038/s41467-021-26922-4.

[132] Ma J, Sun P, Wang D, Wang Z, Yang J, Li Y, et al. The Chloranthus sessilifolius genome provides insight into early diversification of angiosperms. Nature. Communications 2021;12. https://doi.org/10.1038/s41467-021-26931-3.

[133] Soltis DE, Soltis PS. Nuclear genomes of two magnoliids. Nat Plants 2019;5:6–7.

[134] Li Z, Van de Peer Y. A non-duplicated magnoliid genome. Nat Plants 2021;7:1162–3.

[135] Palmer JD, Soltis DE, Chase MW. The plant tree of life: an overview and some points of view. Am J Bot 2004;91:1437–45.

[136] Chen YC, Li Z, Zhao YX, Gao M, Wang JY, Liu KW, et al. The Litsea genome and the evolution of the laurel family. Nat Commun 2020;11:1675.

[137] Ming R, VanBuren R, Liu Y, Yang M, Han Y, Li L-T, et al. Genome of the long-living sacred lotus (Nelumbo nucifera Gaertn.). Genome Biol 2013;14:1–11.

[138] Shen Z, Li W, Li Y, Liu M, Cao H, Provart N, et al. The red flower wintersweet genome provides insights into the evolution of magnoliids and the molecular mechanism for tepal color development. Plant J 2021;108:1662–78.

[139] Hu L, Xu Z, Wang M, Fan R, Yuan D, Wu B, et al. The chromosome-scale reference genome of black pepper provides insight into piperine biosynthesis. Nat Commun 2019;10:4702.

[140] Chanderbali AS, Jin L, Xu Q, Zhang Y, Zhang J, Jian S, et al. Buxus and Tetracentron genomes help resolve eudicot genome history. Nat Commun 2022;13:643.

[141] Liu Y, Wang B, Shu S, Li Z, Song C, Liu D, et al. Analysis of the Coptis chinensis genome reveals the diversification of protoberberine-type alkaloids. Nat Commun 2021;12. https://doi.org/10.1038/s41467-021-23611-0.

[142] Akoz G, Nordborg M. The Aquilegia genome reveals a hybrid origin of core eudicots. Genome Biol 2019;20:256.

[143] Shi T, Chen J. A reappraisal of the phylogenetic placement of the Aquilegia whole-genome duplication. Genome Biol 2020;21:295.

[144] Xie J, Zhao H, Li K, Zhang R, Jiang Y, Wang M, et al. A chromosome-scale reference genome of Aquilegia oxysepala var. kansuensis. Hortic Res 2020;7:113.

[145] Sun P, Jiao B, Yang Y, Shan L, Li T, Li X, et al. WGDI: A user-friendly toolkit for evolutionary analyses of whole-genome duplications and ancestral karyotypes. BioRxiv 2021:2021.04.29.441969.

[146] Yang X, Gao S, Guo L, Wang B, Jia Y, Zhou J, et al. Three chromosome-scale Papaver genomes reveal punctuated patchwork evolution of the morphinan and noscapine biosynthesis pathway. Nature Commun 2021;12. https://doi.org/10.1038/s41467-021-26330-8.

[147] Huang H, Liang J, Tan Q, Ou L, Li X, Zhong C, et al. Insights into triterpene synthesis and unsaturated fatty-acid accumulation provided by chromosomal-level genome analysis of Akebia trifoliata subsp. australis. Hortic Res 2021;8:33.

[148] Sun Y, Deng T, Zhang A, Moore MJ, Landis JB, Lin N, et al. Genome sequencing of the endangered Kingdonia uniflora (Circaeasteraceae, Ranunculales)

reveals potential mechanisms of evolutionary specialization. IScience 2020;23:101124.

[149] Lin J, Zhang W, Zhang X, Ma X, Zhang S, Chen S, et al. Signatures of selection in recently domesticated macadamia. Nat Commun 2022;13:242.

[150] Liu PL, Zhang X, Mao JF, Hong YM, Zhang RG, E, Y., et al. The Tetracentron genome provides insight into the early evolution of eudicots and the formation of vessel elements. Genome Biol 2020;21:291.

[151] Li M, Yang Y, Xu R, Mu W, Li Y, Mao X, et al. A chromosome-level genome assembly for the tertiary relict plant Tetracentron sinense oliv (trochodendraceae). Mol Ecol Resour 2021;21:1186–99.

[152] Yuan F, Wang X, Zhao B, Xu X, Shi M, Leng B, et al. The genome of the recretohalophyte Limonium bicolor provides insights into salt gland development and salinity adaptation during terrestrial evolution. Mol Plant 2022. https://doi.org/10.1016/j.molp.2022.04.011.

[153] VanBuren R, Man Wai C, Pardo J, Giarola V, Ambrosini S, Song X, et al. Desiccation tolerance evolved through gene duplication and network rewiring in lindernia. Plant Cell 2018;30:2943–58.

[154] Soza VL, Lindsley D, Waalkes A, Ramage E, Patwardhan RP, Burton JN, et al. The Rhododendron genome and chromosomal organization provide insight into shared whole-genome duplications across the heath family (Ericaceae). Genome Biol Evol 2019;11:3353–71.

[155] Wang X, Gao Y, Wu X, Wen X, Li D, Zhou H, et al. High-quality evergreen azalea genome reveals tandem duplication-facilitated low-altitude adaptability and floral scent evolution. Plant Biotechnol J 2021;19:2544–60.

[156] Zhang T, Qiao Q, Novikova PY, Wang Q, Yue J, Guan Y, et al. Genome of Crucihimalaya himalaica, a close relative of Arabidopsis, shows ecological adaptation to high altitude. Proc Natl Acad Sci U S A 2019;116:7137–46.

[157] Hu Q, Ma Y, Mandakova T, Shi S, Chen C, Sun P, et al. Genome evolution of the psammophyte Pugionium for desert adaptation and further speciation. Proc Natl Acad Sci U S A 2021;118. https://doi.org/10.1073/pnas.2025711118.

[158] Zhang J, Wu F, Yan Q, John UP, Cao M, Xu P, et al. The genome of Cleistogenes songorica provides a blueprint for functional dissection of dimorphic flower differentiation and drought adaptability. Plant Biotechnol J 2021;19:532–47.

[159] Ramachandran D, Huebner CD, Daly M, Haimovitz J, Swale T, Barrett CF. Chromosome level genome assembly and annotation of highly invasive Japanese stiltgrass (Microstegium vimineum). Genome Biol Evol 2021;13. https://doi.org/10.1093/gbe/evab238.

[160] Oh DH, Kowalski KP, Quach QN, Wijesinghege C, Tanford P, Dassanayake M, et al. Novel genome characteristics contribute to the invasiveness of Phragmites australis (common reed). Mol Ecol 2022;31:1142–59.

[161] Qian Z, Li Y, Yang J, Shi T, Li Z, Chen J. The chromosome-level genome of a free-floating aquatic weed Pistia stratiotes provides insights into its rapid invasion. Mol Ecol Resour 2022. https://doi.org/10.1111/1755-0998.13653.

[162] Huang L, Ma Y, Jiang J, Li T, Yang W, Zhang L, et al. A chromosome-scale reference genome of Lobularia maritima, an ornamental plant with high stress tolerance. Hortic Res 2020;7:197.

[163] Vanneste K, Baele G, Maere S, Van de Peer Y. Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous-Paleogene boundary. Genome Res 2014;24:1334–47.

[164] Song X-M, Wang J-P, Sun P-C, Ma X, Yang Q-H, Hu J-J, et al. Preferential gene retention increases the robustness of cold regulation in Brassicaceae and other plants after polyploidization. Hortic Res 2020;7. https://doi.org/10.1038/s41438-020-0253-0.

[165] Li Z, Van de Peer Y. "Winter Is Coming": How did polyploid plants survive? Mol Plant 2020;13:4–5.

[166] Leitch AR, Leitch IJ. Genomic plasticity and the diversity of polyploid plants. Science 2008;320:481–3.

[167] Wang J, Sun P, Li Y, Liu Y, Yang N, Yu J, et al. An overlooked paleotetraploidization in cucurbitaceae. Mol Biol Evol 2018;35:16–26.

[168] Wang J, Yuan J, Yu J, Meng F, Sun P, Li Y, et al. Recursive paleohexaploidization shaped the durian genome. Plant Physiol 2019;179:209–19.

[169] Teh BT, Lim K, Yong CH, Ng CCY, Rao SR, Rajasegaran V, et al. The draft genome of tropical fruit durian (Durio zibethinus). Nat Genet 2017;49:1633–41.

[170] Meng F, Chu T, Tang Q, Chen W. A tetraploidization event shaped the Aquilaria sinensis genome and contributed to the ability of sesquiterpenes synthesis. BMC Genom 2021;22:647.

[171] Zwaenepoel A, Van de Peer Y. Inference of ancient whole-genome duplications and the evolution of gene duplication and loss rates. Mol Biol Evol 2019;36:1384–404.

[172] Zwaenepoel A, Van de Peer Y. Model-based detection of whole-genome duplications in a phylogeny. Mol Biol Evol 2020;37:2734–46.

[173] Zwaenepoel A, Van de Peer Y. wgd-simple command line tools for the analysis of ancient whole-genome duplications. Bioinformatics 2019;35:2153–5.

[174] Lewin HA, Richards S, Lieberman Aiden E, Allende ML, Archibald JM, Balint M, et al. The Earth BioGenome Project 2020: Starting the clock. Proc Natl Acad Sci U S A 2022;119. https://doi.org/10.1073/pnas.2115635118.

[175] Darwin Tree of Life Project, Consortium. Sequence locally, think globally: The Darwin Tree of Life Project. Proc Natl Acad Sci U S A 2022;119. https://doi.org/10.1073/pnas.2115642118.

[176] Cheng S, Melkonian M, Smith SA, Brockington S, Archibald JM, Delaux PM, et al. 10KP: A phylodiverse genome sequencing plan. GigaScience 2018;7:1–9.

[177] Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 2012;40:D1178–86.

[178] Yates AD, Allen J, Amode RM, Azov AG, Barba M, Becerra A, et al. Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. Nucleic Acids Res 2022;50:D996–D.

[179] Van Bel M, Silvestri F, Weitz EM, Kreft L, Botzki A, Coppens F, et al. PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. Nucleic Acids Res 2022;50:D1468–74.

[180] Lyons E, Pedersen B, Kane J, Alam M, Ming R, Tang HB, et al. Finding and comparing syntenic regions among arabidopsis and the outgroups papaya, poplar, and grape: CoGe with rosids. Plant Physiol 2008;148:1772–81.

[181] Lee TH, Tang HB, Wang XY, Paterson AH. PGDD: a database of gene and genome duplication in plants. Nucleic Acids Res 2013;41:D1152–8.

[182] Chen M, Ma Y, Wu S, Zheng X, Kang H, Sang J, et al. Genome warehouse: A public repository housing genome-scale data. Genom Proteom Bioinform 2021. https://doi.org/10.1016/j.gpb.2021.04.001.