Model species to investigate the origin of flowers

Charles P Scutt

Chapter 5: Model species to investigate the origin of flowers.

Charles P. Scutt

Submitted for inclusion in: Flower Development: Methods and Protocols
(edited by Cristina Ferrandiz and Jose Luis Riechmann)
Third version, submitted 27/04/2022

Laboratoire Reproduction et Développement des Plantes, Univ. Lyon, ENS de Lyon, UCB Lyon-1, CNRS, INRA, F-69342 Lyon, France

E-mail charlie.scutt@ens-lyon.fr
Tel +33 4 72 72 86 03
Fax +33 4 72 72 86 00

Key words: ANA grade, magnoliids, gymnosperms, origin of angiosperms, origin of flowers, evo-devo, *Amborella trichopoda, Nymphaea thermarum, Trithuria,* Austrobaileyales

Abstract

The angiosperms, or flowering plants, arose at least 135 million years ago (MYA) and rapidly diversified to form over 300 000 species alive today. This group appears, however, to have separated from its closest living relatives, the extant gymnosperms, much earlier: over 300 MYA. Representatives of basally-diverging angiosperm lineages are of key importance to studies aimed at reconstructing the most recent common ancestor of living angiosperms, including its morphological, anatomical, eco-physiological and molecular aspects. Furthermore, evo-devo comparisons of angiosperms with living gymnosperms may help to determine how the many novel aspects of angiosperms, including those of the flower, first came about. This chapter reviews literature on the origin of angiosperms and focusses on basally-diverging angiosperms and gymnosperms that show advantages as potential experimental models, reviewing information and protocols for the use of these species in an evo-devo context. The final section suggests a means by which data from living and fossil groups could be integrated to better elucidate evolutionary events that took place on the long stem-lineage that apparently preceded the radiation of living angiosperms.

Introduction

What is a flower?

For the purposes of this chapter, flowers are defined as the sexual reproductive axes of the flowering plants, or angiosperms. In most angiosperms, flowers take the form of compact, bisexual axes that bear ovule-containing carpels centrally, surrounded in turn by pollen-producing stamens and sterile perianth organs: typically petals and sepals. This *bauplan* is sometimes reduced or simplified, such as in the many dioecious or monoecious species that contain unisexual flowers, or in taxa that lack a perianth, such as *Salix* (willows) and *Ulmus* (elms). In extreme cases, the flower may be reduced to a single floral organ, as in some Chloranthaceae.

The rise of flowering plants: when, where, and why?

The flowering plants are the latest major clade of land plants to emerge. A recent molecular clock-based study by Barba-Montoya *et al*. [1] gives five possible ranges of dates for the most recent common ancestor (MRCA) of living flowering plants, depending on the fossil calibration strategy used, which correspond to a combined range of 256-149 million years ago (Ma). The more recent end of this range, near the start of the Cretaceous (~145 Ma), is perhaps not too incongruent with the oldest known fossil pollen grains of clear angiosperm affinity, which date from around 135 Ma [2]. However, recent molecular-clock studies that give estimates for the radiation of living angiosperms long before the start of the Cretaceous (e.g. 214 Ma [3], 275 Ma [4], and the earlier part of the range given by Barba-Montoya *et al*. [1]) are clearly highly incongruous with the pollen fossil evidence. The reasons for this discrepancy are not yet entirely clear, but molecular studies that give very early dates for the MRCA of living angiosperms may have been biased by factors related to the origin of very large clades [5]. Meanwhile, the possible angiosperm affinities of numerous Jurassic flower-like macrofossils, including the recently discovered *Nanjinganthus* from 174 Ma [6], have been widely called into question [7][8][9]. Indeed, the oldest known convincing fossilized angiosperm flower dates from only ~125 Ma, and has been interpreted as a relative of the living genus *Ceratophyllum* [10], which is likely sister to eudicots (Fig. 1).

The geographic radiation and evolutionary diversification of the flowering plants can be traced efficiently through the appearance of novel pollen types, as discussed by Coiro *et al*. [8]. The earliest known angiosperm pollen is monosulcate: having one pore for the emergence of the pollen tube. This type of pollen first appears in N. Africa, the Middle East and Western Europe, territories which, in the early Cretaceous, were centred around northern Gondwana and situated at tropical latitudes of the northern hemisphere. However, within approximately 10 Ma, novel pollen types, including the tricolpate pollen characteristic of eudicots (the group comprising the majority of living angiosperms) begin to appear in sediments expanding away from the northern paleotropics. These detailed paleogeographic data lend strong support to a late Jurassic/early Cretaceous origin of angiosperms in Northern Gondwana, followed by a rapid geographic radiation and biological diversification.

According to most recent molecular phylogenies (e.g. [1][11]), the closest living relatives of the angiosperms comprise a clade containing all extant gymnosperms (Fig. 1). Whatever absolute date-estimates such studies propose, they consistently indicate that the angiosperm and gymnosperm lineages separated considerably before the most recent common ancestor of living angiosperms. Accordingly, the angiosperms appear to have radiated from a stem lineage of perhaps some 150 Ma in length, which means there are no living close relatives of the flowering plants that could help elucidate the origin of the flower.

Evidence exists that a whole genome duplication, termed epsilon, preceded the diversification of living angiosperms [12], though a study by Zwaenepoel and Van der Peer [13] has questioned this finding. The extra sequences generated in the epsilon duplication (if it occurred) may have formed the raw material for the large-scale neofunctionalization of developmental regulators and thereby enabled the rapid evolution of angiosperm-specific features, including those of the flower.

Several biotic and abiotic factors have been proposed to have played a role in the rapid initial expansion and diversification of the angiosperms, as reviewed by Willis and McElwain [2]. These include coevolution with insect pollinators, more efficient photosynthesis as a response to falling atmospheric CO_2 concentration, a reduction in genome size, faster growth rates, and a shorter juvenile phase. This proposed acceleration of the angiosperm life cycle may have enabled these plants to adapt to new ecological niches formed in shaded, damp and disturbed habitats [14], or even to more efficiently escape from newly-evolved groups of low-browsing, herbivorous dinosaurs, a hypothesis critically discussed by Barrett and Willis [15].

A family tree of the flowering plants

The phylogenetic tree of living flowering plants is remarkably asymmetrical [16]. It contains a grade of three lineages, termed the ANA grade after the initials of its component orders Amborellales, Nymphaeales and Austrobaileyales, which diverge basally from a remaining angiosperm lineage (Fig. 1). The ANA grade contains a total of only around 200 living species, while the remaining angiosperm lineage has

diversified to form a clade of over 300 000 living species, termed the euangiosperms or mesangiosperms. The euangiosperms include an early-diverging magnoliid clade of five orders, which contain a total of around 11 000 species. The remaining euangiosperms include the two major clades of eudicots and monocots, which together contain the vast majority of living angiosperms. The small aquatic order Ceratophyllales, containing only *Ceratophyllum,* is probably sister to eudicots.

A portrait of the ancestral flower

Numerical reconstruction methods, based on present-day character states, have enabled the reconstruction of many features of the MRCA of living angiosperms. According to these analyses, this ancestral species was probably a woody plant, perhaps a scrambling shrub with possible liana-like tendencies and a relatively short life cycle, that grew in shaded and disturbed environments such as the banks of fast-flowing streams in dense forest [14]. Its flowers were probably small [17], bisexual [18], protogynous [19], actinomorphic, and contained an undifferentiated perianth of tepals [19][20][18]. Sauquet *et al*. [18] suggest that the perianth and androecium of the MRCA of living angiosperms were whorled, but that its gynoecium was of spiral phyllotaxy. However, other authors have argued that such transitions in symmetry between the androecium and gynoecium are not found in present-day angiosperms and may be prevented by developmental constraints [21][22].

The gynoecium of the MCRA of living angiosperms was almost certainly superior and contained several free, ascidiate (bottle-shaped) carpels. These carpels probably contained an aperture or canal at the apex through which pollen tubes could grow, which was filled by substances secreted from the adjacent cells [23]. Each of these carpels probably contained a single pendant ovule, or a small number of such ovules [19], and these were likely to have been of either anatropous or orthotropous symmetry [24]. The ovules of the MRCA of living angiosperms almost certainly possessed two integuments and an extensive nucellus, or maternal nutritive tissue. These ovules also contained an embryo sac which probably consisted of four cells, and double fertilization would have led to the production of a zygote and a biparental endosperm, both of which were diploid [25].

Models and molecular approaches to study the origin of flowers

Most of what is currently known of the molecular mechanisms of flower development derives from the study of well-adapted plant models from the eudicots and monocots, examples of which are given in Fig. 1. To better understand the origin and early evolution of flowers, a wider spectrum of model species is required, chosen for their key phylogenetic positions. Of particular importance are the living lineages that diverged before and shortly after the initial radiation of the living angiosperms, which correspond to the gymnosperms and basally-diverging angiosperms, respectively.

Models chosen for molecular studies should be amenable to standard molecular biology procedures such as nucleic acid extraction and RNA *in situ* hybridization. It is helpful if study species can be cultivated easily, providing simple access to material at all developmental stages for laboratory work. It is also a huge advantage if the species chosen are amenable to functional-genetic approaches such as the production and long-term storage of mutant collections, and/or stable genetic transformation and the use of RNAi and gene-editing technologies [26]. Recent advances using developmental transcription factors to facilitate transformation and gene-editing in a wide range of species may prove particularly useful [27]. Most functional-genetic approaches are much simplified in species that are diploid, rather than polyploid. Large-scale approaches typically further require plants to be self-fertile, have short generation times and small physical size at maturity. The species chosen should ideally produce copious amounts of seed with orthodox storage characteristics and simple germination requirements. An alternative functional-genetic approach is available in species amenable to Virus Induced Gene Silencing, in which gene knock-downs can be performed using transgenic viruses modified to contain a fragment of a plant gene of interest [28].

It was previously a significant advantage for models to have small genomes, which greatly facilitated whole-genome sequencing and assembly, though recent technical advances have made possible the sequencing and assembly of very large genomes such as those of numerous gymnosperms. Nonetheless, compact genomes with relatively short intergenic regions and introns remain advantageous in model species.

Long intergenic regions can complicate both *in-vitro* and *in-vivo* analyses, partly because it is difficult to estimate, in such circumstances, the size of functionally significant upstream and/or downstream cis-regulatory regions. Similarly, very long introns (e.g. an average of 30.8 kb in the recently sequenced *Cycas* gymnosperm [29]), which may or may not contain important cis-regulatory sequences, inevitably complicate the use of entire gene sequences in experiments involving plant transformation. Fixed publications such as book chapters are not very efficient means for cataloguing genomic or transcriptomic resources. Therefore, besides the resources and information specifically mentioned in this chapter, the reader is directed to regularly updated web-sites dedicated to genomics, transcriptomics and the collation of other useful information, such as those listed in Tab. 1.

Species that are not readily amenable to functional genetics can be studied using a wide range of *in-vitro* and heterologous *in-planta* methods. These include the complementation-testing of mutants in well-established model systems using orthologous coding sequences from species of interest [30], the study of protein-protein interactions by a wide range of methods, and the study of protein-DNA interactions using SELEX, Protein Binding Microarrays [31], or DAP-seq [32]. Conservation of epigenetic marks can also be compared between established models and other species chosen for their phylogenetic positions, conditionally on the availability of sequenced genomes [31]. The following sections detail some of the most promising plant models for study among basally diverging angiosperm and gymnosperm lineages. By comparing the molecular networks that control reproductive development in these two groups, it may be possible to identify key molecular changes that were responsible for the origin and early evolution of the flower and its component organs.

Amborellales

Amborella trichopoda (referred to below as *Amborella*) is the only known member of Amborellales and hence probable sister to all other living flowering plants ([16] , Fig. 1). This dioecious, scrambling shrub (Fig. 2A), endemic to the understorey of sub-tropical cloud forests in New Caledonia, has been widely used in studies aimed at

establishing the pleisiomorphic features of living flowering plants and their underlying molecular mechanisms [33]. Likely pleisiomorphic features of the *Amborella* flower include its undifferentiated perianth of free tepals (Fig. 2B and C) and its gynoecium of free, ascidiate carpels (Fig. 2C), each of which contains an apical secretion-filled canal for pollen-tube growth (Fig. 2D) and a pendent, bitegmic ovule. Probably derived features of *Amborella* include its dioecy [18][32], which recent studies [34] indicate to be due to the presence of a pair of ZW sex chromosomes (ZW = female, ZZ = male) that likely originated less than 16.5 Ma, i.e. much more recently that the origin of flowering plants. The non-recombining, sex-determining region is located on Chromosome 9 of *Amborella* and extends over ~4 Mb [34]. Another derived feature of *Amborella* is its unique eight-celled embryo sac that gives rise, after double fertilization, to a triploid endosperm [35], this type of endosperm ploidy having apparently evolved twice independently in Amborellales and euangiosperms. Features of *Amborella* that correspond to ambiguous character states in the MRCA of living angiosperms include its near-orthotropous ovule symmetry [24] and entirely spiral floral phyllotaxy.

Amborella can be propagated by seed or cuttings and thrives in a temperate, moist atmosphere at relatively low light-intensities [36]. Its seed, which is enclosed within a drupe, typically matures over several months. Germination of *Amborella* seed is furthermore subject to a period of morphophysiological dormancy of over 3 months, though this can be shortened by mechanical or chemical weakening of the pericarp [37]. Such morphophysiological seed dormancy is an inferred ancestral characteristic in angiosperms. Though partially resistant to dessication, *Amborella* seed loses viability through long-term storage, as reviewed by Poncet *et al*. [33].

Amborella is readily amenable to standard molecular-biology procedures such as *in situ* hybridization, and this technique has been used in several studies to assess conservation of functions of flower development genes since the MRCA of living angiosperms (e.g. [38][39][40][41]). However, *Amborella*'s seed characteristics, dioecious breeding system and woody habit make it poorly adapted to functional-genetic studies, and it seems relatively unlikely that a large-scale mutagenesis programme could be undertaken in this species. Few data are available on tissue-

culture in *Amborella*, and no protocols for its stable genetic transformation have yet been reported.

The ~870 Mb-long genome of *Amborella* has been entirely sequenced [42], and more recently a chromosome-level assembly has been made available (Tab. 2). Interestingly, the individual used for genomic sequencing was phenotypically male, though had been reported occasionally to change sex during vegetative propagation [43] and is now known to be genetically female [34], carrying both the Z and W sex-chromosome haplotypes. Analyses of this genome sequence indicate that no whole genome duplications have occurred in the *Amborella* lineage since the hypothesized epsilon event. The phylogenetic position of *Amborella*, and the absence of lineage-specific duplications, make this genome a very important reference for the study of genome evolution in the angiosperms as a whole, and it has already made a major contribution to studies aimed at reconstructing gene-content and gene-order in the MRCA of living angiosperms [3].

Nymphaeales

Nymphaeales, the probable second-earliest diverging ANA-grade order, form a widely distributed group consisting of three families of aquatic and semi-aquatic herbs: Nymphaeaceae (water-lilies), Cabombaceae and Hydatellaceae, of which the latter is basally diverging ([16], Fig. 1). Nymphaeaceae have a cosmopolitan distribution and include the five genera *Nymphaea, Nuphar, Barclaya, Victoria* and *Euryale*. However, a phylogeny by Löhne *et al*. [44] indicates that *Victoria* and *Euryale* may form a clade nested within *Nymphaea,* suggesting a possible need for taxonomic revision. The genome sequence of the unique species of *Euryale, E. ferox,* has recently been published [45]. Cabombaceae include five species of *Cabomba,* which are native to the Americas, and the unique species of *Brasenia, B. schreberi,* which is found in N. America, Africa and Asia, among other territories. Hydatellaceae contain the single genus *Trithuria* (>12 species [46][47]), and are native to Australasia and the southern tip of India.

Within Nymphaeaceae, several species or cultivars of *Nuphar* and *Nymphaea* have been used as evolutionary-developmental models to assess the conservation of expression patterns of developmental regulators with established model species (e.g. [39][48]). In addition, genome sequences have been published from two *Nymphaea* species: *N. colorata (*see PlaBiDP, Tab. 1) and *N. thermarum* [49]. Of these, *N. thermarum,* perhaps uniquely in the ANA grade, regroups a wide range of characteristics that make it highly suitable as a molecular-genetic model.

N. thermarum is the smallest known water lily, with adult plants measuring around 10-20 cm in diameter. The species was known from only one location in Rwanda, but is now considered extinct in the wild [50] and is conserved mainly in botanic gardens and research collections. Flower structure in *N. thermarum* has been described by Fischer and Magdalena Rodriguez [50], and by Povilus *et al*. [51]. Mature, open flowers of *N. thermarum* measure up to 2 cm in diameter (Fig. 2E). These are of whorled phyllotaxy and contain 4 greenish outer tepals and 6-8 whitish inner tepals. The *N. thermarum* gynoecium consists of 7-9 carpels which fuse together basally during flower development. Angiospermy (enclosure of the ovules) in all *Nymphaea* spp. is completed by post-genital fusion at the carpel margins, rather than by secretion of substances into an apical aperture or canal [23]. Numerous bitegmic, anatropous ovules form in each carpel of *N. thermarum, e*ach of which contains a four-celled embryo sac that contributes, after double fertilization, to produce a diploid endosperm [51]. The bisexual axis, bitegmic ovule, four-celled embryo sac and diploid endosperm of *N. thermarum* are likely pleisiomorphic features of living angiosperms, whereas the complete closure of its carpels by cellular structures, its partially syncarpic gynoecium, and its high number of ovules-per-carpel all appear to be derived features. Anatropous ovule symmetry in *N. thermarum* contrasts with the near-orthotropous symmetry in *Amborella*, and one of these two symmetry types is likely to be ancestral in living angiosperms. *N. thermarum* therefore possesses a list of likely pleisiomorphic features that is highly complementary to those of *Amborella.*

N. thermarum is a perennial herb that can be grown from seed to flowering in around 2-3 months [50]. It thrives in a warm, humid environment under high light intensities and can be easily cultivated in pots of wet compost (Fig. 2F) as the plants do not need to grow completely immersed in water. The small size of *N. thermarum* flowers and

their component organs make these highly practical for *in-situ* hybridization and other microscopic procedures. A recent study included genetic crosses in *N. thermarum,* demonstrating the practicality of this species for such standard genetic procedures [52]. *N. thermarum* is self-fertile, and approximately 150 seeds are produced in each fruit. Optimal seed storage conditions have not yet been published, though seeds of some other water lilies are known to be of the orthodox type (Kew Millennium Seed Bank database, Tab. 1). *N. thermarum* seeds germinate rapidly under water, but the young seedlings, which are filiform in structure, should then be placed horizontally on the surface of a substrate, covered by a thin film of water, until the first true leaves appear [50]. Unlike some water lilies, *N. thermarum* cannot be propagated by rhizomes and no methods of vegetative propagation have yet been reported.

The scientific literature on *N. thermarum* is currently very limited, but this is an obvious candidate for full-scale development as an ANA-grade angiosperm model. Its diploid genome is among the smallest in Nymphaeaceae and among the ANA-grade angiosperms as a whole (Tab. 2), and some tissue/stage-specific transcriptomic data are also available (Tab. 2). Growth of *N. thermarum* in tissue culture should be facilitated by the ready availability of seed that can easily be surface-sterilized and germinated *in vitro,* and further studies in this species may pave the way to the development of protocols for its stable genetic transformation.

Trithuria, the only genus of Hydatellaceae, includes very small aquatic/semi-aquatic annuals and perennials [46]. These plants produce tufts of linear leaves, each containing a single vascular bundle, and apical reproductive units (RUs), also termed “non-flowers”, that consist of compact axes on which reproductive organs are surrounded by a small number of bract-like phyllomes [53]. The genus includes dioecious species in which male and female RUs occur on separate individuals, monoecious species in which both male and female RUs form on the same individuals, and bisexual species in which both carpels and stamens form on the same RUs. In bisexual *Trithuria* spp., the carpels generally form externally to the stamens (Fig. 2G), and it is possible that the *Trithuria* RU represents an inflorescence, each flower of which is reduced to a single reproductive organ, as discussed by Rudall *et al*. [53]. *Trithuria* appears therefore to have either a highly derived or radically simplified floral structure, compared to the reconstructed MRCA of living angiosperms. *Trithuria*’s

carpels, however, appear relatively underived. These occur separately, are acidiate in shape, contain a short apical canal for pollen growth [54], and enclose a single pendant, bitegmic, anatropous ovule. The *Trithura* ovule includes a large perisperm, derived from the nucellus, which has been suggested to be a relictual character that has persisted from gymnosperm-like ancestors [55]. The embryo sac structure and endosperm ploidy in *Trithuria* [55] are similar to those of *N. thermarum,* described above. As in *Amborella*, the likely pleisiomorphic condition of morphophysiological seed dormancy is present in *Trithuria* [56], in contrast to the more recently-evolved physiological dormancy present in other Nymphaeales [37]. The relatively underived structure of the individual floral organs of *Trithuria* make this a potentially useful model for the origin and early evolution of angiosperms, whereas its highly derived organization into RUs make it an essential model for studies of the evolution of apparently ”inside-out” flower-like structures, which are very rare among the angiosperms.

Few molecular data have yet been reported in *Trithuria*, and the tractability of these species as potential molecular-genetic models remains uncertain. Plants of the bisexual *T. submersa* can easily be grown to maturity under *in-vitro* conditions ([57], Fig. 2H), though seed production under these conditions has not been reported. *T. konkanensis* has also been grown *in vitro* [58]. Genome sizes in *Trithuria* are modest (Kew Millennium Seed Bank database, Tab. 1), though *T. submersa*, which might otherwise make a potential genetic model, appears to be octaploid [47], a feature which would likely complicate any attempt at functional genetics. A trancriptomic analysis of entire *T. submersa* plants has been performed (Tab. 2), though no full genomic sequence has yet been published.

Cabomba is probably better adapted to represent Cabombaceae in molecular-based evo-devo studies than its sister taxon *Brasenia schreberi,* as *Cabomba* flowers are the simpler in structure, and because the leaves, stems and floral organs of *B. schreberi* are coated in a layer of highly viscous mucilage, which might complicate some molecular-biology procedures. *Cabomba* plants form horizontal rhizomes and sparsely branched submerged stems that give rise to submerged leaves that are highly dissected to reduce water-drag (Fig. 2I). Adventitious roots can emerge at the submerged leaf nodes [59]. After induction to flowering, however, a series of floating

leaves are produced at the *Cabomba* stem apex (Fig. 2J), each of which subtends a floral bud. These floating leaves are sagittate to orbicular in shape, with entire margins, and function to support the subtended flowers at the water's surface. *Cabomba* flowers contain a largely undifferentiated perianth of two trimerous whorls of tepals, a doubled trimerous androecium of six stamens, and a gynoecium, typically of two or three free carpels (Fig. 2K). The *Cabomba* carpel includes the likely pleisiomorphic angiosperm features of an ascidiate shape, an apical canal for pollen tube growth, and a low number of pendant, bitegmic ovules [23]. These ovules are anatropous in symmetry. Embryo sac structure and endosperm ploidy have not been investigated in Cabombaceae.

Cabomba plants can be grown without difficulty in a medium-sized aquarium and induced to flower by adding far-red light (~740 nm, from light-emitting diodes, for example) to standard lighting for aquarium plants [59]. Vegetative propagation of *Cabomba* can be easily performed by planting bunches of cut stems in the aquarium substrate to encourage the production of adventitious roots. *C. caroliniana,* the most temperate species of the genus, is a significant invasive weed in many regions and is now subject to strict controls in several jurisdictions [60]. However, less hardy species such as *C. aquatica* may substitute for *C. caroliniana* in evo-devo studies. *Cabomba* flowers and other tissues are readily amenable to standard molecular biology procedures such as *in situ* hybridization (e.g. [38][59][61]). A modest-sized *Cabomba* flower EST database is available [59], though no large-scale RNA-seq data have been produced. Genetic transformation has not been reported in *Cabomba*, while its aquatic habit, fragile stems and uncharacterized seed-storage characteristics make the genus unlikely to be suitable for large-scale molecular-genetic procedures such as the generation of mutant collections.

Austrobaileyales

Austrobaileyales, the probable third-earliest diverging ANA-grade order, includes three families of woody plants containing a total of around 100 species in five genera [16]. The most basally-diverging of these families, Austrobaileyaceae, contains only *Austrobaileya scandens,* a liana endemic to Northern Queensland. Trimeniaceae

contains only the small genus *Trimenia*, whose members are small trees, shrubs, or lianas, and are native to Eastern Australia, SE Asia and islands in the Pacific. Schisandraceae, the third family of Austrobaileyales, contains the three small-to-medium sized genera *Illicium* (Fig. 2L), *Schisandra* and *Kadsura*, whose members are shrubs or lianas, native mainly to SE Asia, though with some species present in N. America and the Caribbean.

Austrobaileyales include plants with bisexual and/or unisexual flowers and a range of breeding systems (discussed by Endress [17]), including protogyny, e.g. in *Austrobaileya scandens and Illicium spp.*; andromonoecy, e.g. in *Trimenia* spp.; dioecy, e.g. in *Schisandra chinensis*; mixed monoecy and dioecy, e.g. in *Kadsura japonica;* and self-incompatibility (SI) in *Illicium floridanum* [62]*, Austrobaileya scandens* [63] and *Trimenia moorii* [64]. Of these various breeding systems, bisexuality with protogyny is thought to be the pleisiomorphic condition of living angiosperms, and SI may also have been present very early in angiosperm evolution [65]. Indeed, almost all of the likely pleisiomorphic features of flowers can be found in Austrobaileyales, including bisexuality (e.g. *Austrobaileya scandens* and *Illicium*), an undifferentiated perianth of tepals (all taxa), an unfused gynoecium of carpels, each containing a canal or aperture for pollen tube growth (all taxa except *Illicium*), and a single or low number of bitegmic ovules (all taxa). The presumed pleisiomorphic arrangement of a 4-celled embryo sac and diploid endosperm is known to occur in *Illicium* [66] and may be present throughout Austrobaileyales. An interesting and potentially ancient phenomenon revealed in *Trimenia* shows that multiple embryo sacs may persist in each ovule, and compete for fertilization by growing towards the oncoming pollen tubes [67]. Austrobaileyales have spiral floral phyllotaxy (though this is pseudo-whorled in *Illicium*), which may or may not be a pleisiomorphic character in angiosperms (as discussed in the Introduction, above).

Austrobaileyales have been extensively included in analyses to reconstruct morphological character states in the MRCA of living angiosperms. However, members of this order have been little used to investigate the molecular bases of flower evolution. Their woody habit makes these plants, like *Amborella*, relatively poorly adapted as potential molecular-genetic models, and there is no published literature on methods for their stable genetic transformation. Some Austrobaileyales, such as *Austrobaileya*

scandens, are known to flower unpredictably in cultivation, though others, such as species of *Schisandra*, *Kadsura* and *Illicium,* are ornamental or commercially important plants that flower reliably and can be readily obtained from botanic gardens or commercial sources.

Genome sizes of Austrobaileyales are larger than that of *Amborella* and much larger than those of most Nymphaeales [59], and no genome of this order has yet been sequenced. Transcriptomic resources are, however, currently available for *Austrobaileya scandens, Illicium* spp., *Kadsura heteroclita* (OneKP database, Tab. 1) and *Schisandra* spp. [68, 69]. It is likely, as studies of ANA-grade angiosperms progress, that more attention will be focused on Austrobaileyales. Despite their large genomes and lack of ideal features as molecular-genetic models, many of these species retain a high proportion of likely pleisiomorphic angiosperm characters and, as woody, terrestrial plants, may have undergone a more conservative evolutionary history than did the aquatic and herbaceous Nymphaeales [17]. In addition, *Amborella,* the only other woody, non-aquatic member of the ANA-grade, is likely to be derived in numerous respects (described above), underlining the need for multiple models and a comparative approach in evolutionary analyses.

Magnoliids

The magnoliids form a diverse clade of around 11 000 species, classified in five orders (Fig. 1), and include both woody and herbaceous taxa[16]. The magnoliid stem lineage emerges within the living angiosperm tree, shortly after those of the ANA grade (Fig 1). Herbaceous magnoliids might make more convenient, if slightly less basally-diverging models than the woody plants or fully aquatic herbs that make up most of the ANA grade. In addition to potentially providing further insights into the origin of angiosperms, magnoliid models can be expected to constitute an external reference to better reconstruct the origin of the later-emerging major clades of the eudicots and monocots (Fig. 1).

One species suggested as a promising herbaceous magnoliid model is *Aristolochia fimbriata* [70]. *Aristolochia* (Piperales, Aristolochiaceae; also known as pipevines) is a

large genus containing both woody and herbaceous species. *A. fimbriata* is a vine that attracts pollinating flies into a chamber formed by its bilaterally symmetrical perianth of fused sepals. Bilateral perianth symmetry is a novel feature that has arisen several times independently in the angiosperms and may have occurred first within magnoliids [70]. Other floral features of *Aristolochia* that are distinct from those of earlier-diverging lineages include its inferior ovary, in which the stamens are developmentally fused to the style to form a gynostemium.

A. fimbriata is highly suitable as a model magnoliid, in part for its small size at maturity, rapid life cycle, self-compatibility and small genome ([70],Tab. 2), which is now entirely sequenced [71]. Interestingly, *A. fimbriata* is thought, like *Amborella,* to have undergone no lineage-specific whole genome duplication event since the epsilon duplication at the base of the angiosperm cade. Crucially for its use as a model basal angiosperm, tissue culture procedures have been established in *A. fimbriata* [72], leading to protocols for its stable genetic transformation using *Agrobacterium tumefaciens* [70]. *A. fimbriata* is readily amenable to standard molecular procedures such as *in-situ* hybridization, and this technique has already been used to study the expression of numerous classes of MADS-box floral homeotic genes [73]. Interactions between the corresponding proteins have also been studied in yeast-two-hybrid experiments [74].

In addition to that of *A. fimbriata,* 13 further magnoliid genome sequences are now available, details of which are given on the PlaBiDB database (Tab. 1). Among the first magnoliid species with published genomes were *Persea americana* (Laurales, Lauraceae; avocado; [75]), a woody crop species of high economic importance, and *Liriodendron chinense* (Magnoliales, Magnoliaceae; Chinese tulip tree; [76]), an ornamental tree from E. Asia. In contrast to *A. fimbriata,* analyses of the *P. americana* genome suggest the presence of two lineage-specific whole genome duplications that occurred subsequently to the hypothesized epsilon duplication [75], one of which appears to be shared with the *L. chinense* lineage [29, 76]. An analysis of the expression patterns of *P. americana* MADS-box floral regulators has been performed [39], while numerous floral transcriptomic resources are available for *L. chinense* and its unique sister species *L. tulipifera* (Tab. 2).

Gymnosperms

The gymnosperms form the closest living out-group to the flowering plants and therefore constitute a vital external reference to understand the evolutionary events that contributed to the flower and other angiosperm-specific traits. The reproductive structures of gymnosperms differ from most angiosperm flowers by, among other features, being unisexual and lacking a perianth. Gymnosperm ovules have a single integument, rather than the two integuments present in most angiosperm groups, and are not enclosed within carpels. However, further tissue layers surround the ovules in certain gymnosperms (e.g. *Taxus, Juniperus* and Gnetales), forming fleshy structures that persist around the seed. Pollination in gymnosperms occurs mostly by wind, though some Cycadales and Gnetales are pollinated by both wind and insects [77]. Double fertilization to produce a biparental endosperm is absent in gymnosperms, though a supernumary zygote is formed through a second fertilization event in Gnetales [78].

Most recent molecular phylogenies of seed plants indicate the extant gymnosperms to be sister to angiosperms, though two studies, dating from some years ago, suggested an alternative topology in which a clade containing Cycadales and angiosperms was sister to remaining gymnosperms [79][80]. By contrast, the internal topology of the gymnosperm clade remains unresolved, and several alternative topologies have been proposed (reviewed by Doyle [81]). Most of these topologies include Cycadales and Ginkgoales in basally-diverging positions relative to a remaining clade containing all other living gymnosperms, as shown in Fig 1. Within this remaining gymnosperm clade, however, the position of the small order Gnetales appears highly labile, as various studied have placed this as sister to Pinaceae, sister to all conifers, or sister to all conifers excluding Pinaceae. Some phylogenies have even shown Gnetales as basally-diverging within living gymnosperms (e.g. [82]), though such conclusions may reflect bias due to the inclusion of particularly fast-evolving Gnetales sequences in phylogenomic analyses (see Zhong *et al*. [83]). Topologies showing Gnetales in a basal position in gymnosperms have also been criticized as inconsistent with the fossil record, which indicates a relatively recent origin of this group [81]. The lack of complete phylogenetic resolution in gymnosperms leads to problems for their classification into higher taxa. A classification has, however, been proposed which circumvents this

problem by placing several gymnosperm families, including Pinaceae and the three families within Gnetales (Fig. 1), within separate monotypic orders [84].

Gymnosperms form a vital external reference to enable the reconstruction of the molecular-evolutionary processes that generated the angiosperm flower. One potentially useful aim might be to reconstruct gene-order in the genome of the MRCA of living seed plants (angiosperms+gymnosperms), from which the angiosperm stem lineage originally emerged. This goal would no doubt benefit from a representative taxonomic sampling of genomes, and indeed complete genome sequences are now available from all major gymnosperm clades. Some of these genomes are listed in Tab. 2, along with transcriptomic resources in reproductive tissues, while publications describing the genomes of *Abies alba, Sequoia sempervirens, Sequoiadendron giganteum*, *Taxus chinensis and T. walichiana* (conifers), *Gnetum montanum and Welwitschia mirabilis* (Gnetales), *and Gingko biloba* (Ginkgoales), are given in the PlaBiDB database (Tab 1). The recent publication of a chromosome-level assembly of the *Cycas panzhihuaensis* genome [29] has revealed numerous interesting features, including a likely whole-genome duplication at the base of the living gymnosperm clade and a potentially common molecular basis of chromosomal sex-determination shared between *Cycas* and *Gingko.* Transcriptomic resources that include reproductive tissues are currently available from a number of gymnosperms (Tab. 2). Of particular interest, analyses of a complete developmental series of male and female cones have been performed in *Pinus tabuliformis* [85], while a recently developed spatially-resolved transcriptome profiling technique has been successfully applied to female cones of *Picea abies* [86].

Gymnosperm orthologs of numerous flower development genes have been studied using a wide range of methods. Major elements of conservation between mechanisms of development in gymnosperm cones and angiosperm flowers have been concluded from *in situ* hybridization data (e.g. [87][88]) protein-DNA interaction data (e.g. [88][89] and from the complementation of angiosperm floral homeotic mutants using orthologous sequences from gymnosperms (e.g. [87][90]).

One practical problem for transcriptomics relates to the slow growth of many gymnosperm reproductive tissues, which can complicate sampling procedures. However, certain taxa such as *Welwitschia mirabilis* (Gnetales, Welwitschiaceae)

possess a developmental series in each cone, with younger stages occurring nearer the apex [88], while others such as *Ephedra* spp. (Gnetales, Ephedraceae) complete their reproductive development relatively quickly, facilitating sampling from all developmental stages in each reproductive season. Sampling from *Picea abies* is complicated by the form of the adult trees, on which cones typically emerge at a height of several meters on slender and fragile branches. However, the *acrocona* mutant of *P. abies* [91] produces cones at low level and can be grown from seed to reproductive maturity in under three years. The molecular basis of the *acrocona* mutation is not yet known, though the wild-type locus may act to downregulate a specific MADS-box gene [92].

Numerous gymnosperms are amenable to genetic transformation using *Agrobacterium tumefaciens* through the generation of somatic embryos (reviewed by Uddenberg *et al*. [91], Fig. 1). However, transformation methods have not yet been used to analyse the roles of genes controlling reproductive development in gymnosperms, probably due to the typically long juvenile phase of these plants (e.g. over 25 years in wild-type *P. abies*). It is possible, however, that the *acrocona* mutant of *P. abies*, and perhaps in the future equivalent mutants in other gymnosperms [91] may provide ways to speed-up this process. The use of gymnosperms with shorter juvenile phases, such as *Welwitschia mirabilis* and *Ephedra* spp., could also help to accelerate functional-genetic analyses of gymnosperm reproductive development.

Missing links and new approaches

This chapter mainly concerns living species of basally diverging angiosperms and gymnosperms that can be used to better understand the origin of the flower. However, the MRCA of all living seed plants is separated from that of living angiosperms by a stem lineage of perhaps some 150 Ma, from which no species have survived other than the angiosperms themselves. In the absence, therefore, of close living relatives of the flowering plants, the epsilon whole genome duplication may provide a point of reference to help reconstruct the lost ancestors of the angiosperms (notwithstanding the recent conclusions of Zwaenepoel and Van der Peer [13] on the absence of this duplication!).

The combined analysis of molecular phylogeny and synteny in multiple angiosperm genomes (e.g. [3]), may soon enable the systematic identification of paralogs whose gene lineages separated at the (hypothesized) epsilon duplication. Certain gene lineages are known, from model angiosperms, to be necessary for the development of angiosperm-specific characters such as the perianth, bisexual axis, carpel and outer ovule integument (reviewed by Scutt [93]). Protein sequences from nodes in any molecular phylogeny can be reconstructed using ancestral sequence reconstruction or "protein resurrection", as discussed by Vialette-Guiraud *et al*. [31], and the corresponding genes and proteins can then be generated in the laboratory for both *in-vitro* and *in-vivo* studies [31]. Reconstructed regulators of reproductive development from the time of the epsilon duplication may, therefore, help to shed light on the morphological features of the plants in which these molecules functioned. Notably, information on the biochemical properties of ancestralized regulatory molecules, and a comparison of their rates and modes of sequence evolution from before and after the hypothesized epsilon event, may provide clues on the order-of-acquisition of novel angiosperm features, including the perianth, bisexual axis, carpel and outer integument.

Though no living, close relative of the angiosperms is available for study, there is an extensive fossil record of extinct gymnosperm groups, some of which show features that suggest affinities with angiosperms. Several groups of gymnosperms have been suggested as potential stem-lineage relatives of the angiosperms and tentatively placed, using phylogenetic analyses based on morphological features, on the seed-plant phylogenetic tree ([81, 94][95], Fig 1). Interestingly, several of these potential angiosperm relatives show contrasting sets of angiosperm-like features. For example, *Caytonia* (Caytoniales) are fossilized female reproductive structures in which multiple unitegmic ovules are enclosed within cupules that are pinnately arranged on a radially symmetrical rachis (Fig. 2M). The *Caytonia* cupule and rachis have been suggested to be potentially homologous to the angiosperm outer integument and carpel, respectively (reviewed by Doyle [95], and see also Shi *et al.* [94]). Like Caytoniales, Glossopteridales have unisexual female reproductive structures composed of ovule-containing cupules, though these typically form in the axil of a fertile leaf or sporophyll (Fig. 2N). In the case of Glossopteridales, the cupule and sporophyll may be homologous to the angiosperm outer integument and carpel, respectively.

Bennettitales, by contrast, do not possess cupules, but some species of this order have a bisexual axis that bears male organs externally to female organs, as in the angiosperm flower (Fig. 2O). The reproductive axes of Bennettitiales also possess a perianth-like structure, giving them a distinctly flower-like appearance.

The predicted order-of-acquisition of angiosperm-specific features will vary depending on which fossil gymnosperm is regarded as the closest relative to angiosperms. If, for example, Bennettitales were considered as the closest angiosperm relatives, it would be logical to conclude that the origin of the angiosperm perianth and bisexual axis preceded that of the outer integument and carpel. If, by contrast, Caytoniales or Glossopteridales were regarded as closest to angiosperms, the opposite order-of-acquisition of the above-listed features would be concluded. Such contrasting predictions could be compared with the timeline (notably, before and after epsilon) of molecular signatures of neofunctionalization in gene lineages known to be involved in the development of the angiosperm characters under consideration. In this way, it may be possible to combine molecular data from living groups to choose between alternative potential ancestors of the angiosperms, and thus shine some light into the dark ages surrounding the origin of the flower.

Acknowlegements

Thanks are due to Amélie Andres-Robin for help producing plant photos and to the editors for inviting this contribution.

Literature cited

1. Barba-Montoya J, dos Reis M, Schneider H, et al (2018) Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous Terrestrial Revolution. New Phytol 218:819–834. https://doi.org/10.1111/nph.15011

2. Willis K, McElwain J (2013) The Evolution of Plants, 2nd ed. OUP, Oxford

3. Murat F, Armero A, Pont C, et al (2017) Reconstructing the genome of the most recent common ancestor of flowering plants. Nat Genet 49:490-+. https://doi.org/10.1038/ng.3813

4. Salomo K, Smith JF, Feild TS, et al (2017) The Emergence of Earliest Angiosperms May be Earlier than Fossil Evidence Indicates. Syst Bot 42:. https://doi.org/10.1600/036364417X696438

5. Budd GE, Mann RP (2018) History is written by the victors: The effect of the push of the past on the fossil record. Evolution 72:2276–2291. https://doi.org/10.1111/evo.13593

6. Fu Q, Bienvenido Diez J, Pole M, et al (2018) An unexpected noncarpellate epigynous flower from the Jurassic of China. Elife 7:e38827. https://doi.org/10.7554/eLife.38827

7. Bateman RM (2009) Hunting the Snark: the flawed search for mythical Jurassic angiosperms. J Exp Bot in press

8. Coiro M, Doyle JA, Hilton J (2019) How deep is the conflict between molecular and fossil evidence on the age of angiosperms? New Phytol 223:83–99. https://doi.org/10.1111/nph.15708

9. Sokoloff DD, Remizowa MV, El ES, et al Supposed Jurassic angiosperms lack pentamery, an important angiosperm-specific feature. New Phytol 2019. https://doi.org/doi: 10.1111/nph.15974

10. Gomez B, Daviero-Gomez V, Coiffard C, et al (2015) Montsechia, an ancient aquatic angiosperm. Proc Natl Acad Sci U S A 112:10985–10988. https://doi.org/10.1073/pnas.1509241112

11. Magallon S, Hilu KW, Quandt D (2013) LAND PLANT EVOLUTIONARY TIMELINE: GENE EFFECTS ARE SECONDARY TO FOSSIL CONSTRAINTS IN RELAXED CLOCK ESTIMATION OF AGE AND SUBSTITUTION RATES. Am J Bot 100:556–573. https://doi.org/10.3732/ajb.1200416

12. Jiao Y, Wickett NJ, Ayyampalayam S, et al (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473:97-U113. https://doi.org/10.1038/nature09916

13. Zwaenepoel A, Van de Peer Y (2019) Inference of Ancient Whole-Genome Duplications and the Evolution of Gene Duplication and Loss Rates. Mol Biol Evol 36:1384–1404. https://doi.org/10.1093/molbev/msz088

14. Feild TS, Arens NC, Doyle JA, et al (2004) Dark and disturbed: a new image of early angiosperm ecology. Paleobiology 30:82–107. https://doi.org/10.1666/0094-8373(2004)030<0082:DADANI>2.0.CO;2

15. Barrett PM, Willis KJ (2001) Did dinosaurs invent flowers? Dinosaur-angiosperm coevolution revisited. Biol Rev 76:411–447. https://doi.org/10.1017/S1464793101005735

16. Byng JW, Chase MW, Christenhusz MJM, et al (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot J Linn Soc 181:1–20. https://doi.org/10.1111/boj.12385

17. Endress PK (2001) The flowers in extant basal angiosperms and inferences on ancestral flowers. Int J Plant Sci 162:1111–1140. https://doi.org/10.1086/321919

18. Sauquet H, von Balthazar M, Magallon S, et al (2017) The ancestral flower of angiosperms and its early diversification. Nat Commun 8:16047. https://doi.org/10.1038/ncomms16047

19. Endress PK, Doyle JA (2015) Ancestral traits and specializations in the flowers of the basal grade of living angiosperms. Taxon 64:1093–1116. https://doi.org/10.12705/646.1

20. Endress PK, Doyle JA (2009) RECONSTRUCTING THE ANCESTRAL ANGIOSPERM FLOWER AND ITS INITIAL SPECIALIZATIONS. Am J Bot 96:22–66. https://doi.org/10.3732/ajb.0800047

21. De-Paula OC, Assis LCS, de Craene LPR (2018) Unbuttoning the Ancestral Flower of Angiosperms. Trends Plant Sci 23:551–554. https://doi.org/10.1016/j.tplants.2018.05.006

22. Sokoloff DD, Remizowa MV, Bateman RM, Rudall PJ (2018) Was the ancestral angiosperm flower whorled throughout? Am J Bot 105:5–15. https://doi.org/10.1002/ajb2.1003

23. Endress PK, Igersheim A (2000) Gynoecium structure and evolution in basal angiosperms. Int J Plant Sci 161:S211–S223. https://doi.org/10.1086/317572

24. Endress PK (2011) Angiosperm ovules: diversity, development, evolution. Ann Bot 107:1465–1489. https://doi.org/10.1093/aob/mcr120

25. Friedman WE, Ryerson KC (2009) Reconstructing the ancestral female gametophyte of angiosperms: insights from Amborella and other ancient lineages of flowering plants. Am J Bot 96:129–143. https://doi.org/10.3732/ajb.0800311

26. Mao Y, Botella JR, Liu Y, Zhu J-K (2019) Gene editing in plants: progress and challenges. Natl Sci Rev 6:421–437. https://doi.org/10.1093/nsr/nwz005

27. Maher MF, Nasti RA, Vollbrecht M, et al (2020) Plant gene editing through de novo induction of meristems. Nat Biotechnol 38:84–89. https://doi.org/10.1038/s41587-019-0337-2

28. Di Stilio VS (2011) Empowering plant evo-devo: Virus induced gene silencing validates new and emerging model systems. Bioessays 33:711–718. https://doi.org/10.1002/bies.201100040

29. Liu Y, Wang S, Li L, et al (2022) The Cycas genome and the early evolution of seed plants. Nat Plants 8:389–401. https://doi.org/10.1038/s41477-022-01129-7

30. Kramer EM (2015) A stranger in a strange land: the utility and interpretation of heterologous expression. Front Plant Sci 6:734. https://doi.org/10.3389/fpls.2015.00734

31. Vialette-Guiraud ACM, Andres-Robin A, Chambrier P, et al (2016) The analysis of Gene Regulatory Networks in plant evo-devo. J Exp Bot 67:2549–2563. https://doi.org/10.1093/jxb/erw119

32. Bartlett A, O'Malley RC, Huang SC, et al (2017) Mapping genome-wide transcription-factor binding sites using DAP-seq. Nat Protoc 12:1659–1672. https://doi.org/10.1038/nprot.2017.055

33. Poncet V, Birnbaum P, Burtet-Sarramegna V, et al (2019) Amborella– Bearing Witness to the Past? Annu Plant Rev Online 1–41

34. Kafer J, Bewick A, Andres-Robin A, et al (2022) A derived ZW chromosome system in Amborella trichopoda, representing the sister lineage to all other extant flowering plants. New Phytol 233:1636–1642. https://doi.org/10.1111/nph.17662

35. Friedman WE (2006) Embryological evidence for developmental lability during early angiosperm evolution. Nature 441:337–340. https://doi.org/10.1038/nature04690

36. Grosse-Veldmann B, Korotkova N, Reinken B, et al (2011) Amborella trichopoda- Cultivation of the most ancestral angiosperm in botanic gardens. SIBBALDI 9:143–155

37. Fogliani B, Gateble G, Villegente M, et al (2017) The morphophysiological dormancy in Amborella trichopoda seeds is a pleisiomorphic trait in angiosperms. Ann Bot 119:581–590. https://doi.org/10.1093/aob/mcw244

38. Fourquin C, Vinauger-Douard M, Fogliani B, et al (2005) Evidence that CRABS CLAW and TOUSLED have conserved their roles in carpel development since the ancestor of the extant angiosperms. Proc Natl Acad Sci U S A 102:4649–4654. https://doi.org/10.1073/pnas.0409577102

39. Kim S, Koh J, Yoo MJ, et al (2005) Expression of floral MADS-box genes in basal angiosperms: implications for the evolution of floral regulators. Plant J 43:724–744. https://doi.org/10.1111/j.1365-313X.2005.02487.x

40. Vialette-Guiraud ACM, Adam H, Finet C, et al (2011) Insights from ANA-grade angiosperms into the early evolution of CUP-SHAPED COTYLEDON genes. Ann Bot 107:1511–1519. https://doi.org/10.1093/aob/mcr024

41. Arnault G, Vialette ACM, Andres-Robin A, et al (2018) Evidence for the Extensive Conservation of Mechanisms of Ovule Integument Development Since the Most Recent Common Ancestor of Living Angiosperms. Front Plant Sci 9:1352. https://doi.org/10.3389/fpls.2018.01352

42. Albert VA, Barbazuk WB, dePamphilis CW, et al (2013) The Amborella Genome and the Evolution of Flowering Plants. Science 342:1467-+. https://doi.org/10.1126/science.1241089

43. Buzgo M, Soltis PS, Soltis DE (2004) Floral developmental morphology of Amborella trichopoda (Amborellaceae). Int J Plant Sci 165:925–947. https://doi.org/10.1086/424024

44. Loehne C, Borsch T, Wiersema JH (2007) Phylogenetic analysis of nymphaeales using fast-evolving and noncoding chloroplast markers. Bot J Linn Soc 154:141–163. https://doi.org/10.1111/j.1095-8339.2007.00659.x

45. Yang Y, Sun P, Lv L, et al (2020) Prickly waterlily and rigid hornwort genomes shed light on early angiosperm evolution. Nat Plants 6:215–222. https://doi.org/10.1038/s41477-020-0594-6

46. Iles WJD, Rudall PJ, Sokoloff DD, et al (2012) MOLECULAR PHYLOGENETICS OF HYDATELLACEAE (NYMPHAEALES): SEXUAL-SYSTEM HOMOPLASY AND A NEW SECTIONAL CLASSIFICATION. Am J Bot 99:663–676. https://doi.org/10.3732/ajb.1100524

47. Sokoloff DD, Marques I, Macfarlane TD, et al (2019) Cryptic species in an ancient flowering-plant lineage (Hydatellaceae, Nymphaeales) revealed by molecular and micromorphological data. Taxon 68:1–19. https://doi.org/10.1002/tax.12026

48. Yamada T, Ito M, Kato M (2003) Expression pattern of INNER NOOUTER homologue in Nymphaea (water lily family, Nymphaeaceae). Dev Genes Evol 213:510–513. https://doi.org/10.1007/s00427-003-0350-8

49. Povilus RA, DaCosta JM, Grassa C, et al (2020) Water lily (Nymphaea thermarum) genome reveals variable genomic signatures of ancient vascular cambium losses. Proc Natl Acad Sci U S A 117:8649–8656. https://doi.org/10.1073/pnas.1922873117

50. Fischer E, Rodriguez C (2010) 690. NYMPHAEA THERMARUM. Curtiss Bot Mag 27:318–327. https://doi.org/10.1111/j.1467-8748.2010.01711.x

51. Povilus RA, Losada JM, Friedman WE (2015) Floral biology and ovule and seed ontogeny of Nymphaea thermarum, a water lily at the brink of extinction with potential as a model system for basal angiosperms. Ann Bot 115:211–226. https://doi.org/10.1093/aob/mcu235

52. Povilus RA, Diggle PK, Friedman WE (2018) Evidence for parent-of-origin effects and interparental conflict in seeds of an ancient flowering plant lineage. Proc R Soc B-Biol Sci 285:20172491. https://doi.org/10.1098/rspb.2017.2491

53. Rudall PJ, Remizowa MV, Prenner G, et al (2009) NONFLOWERS NEAR THE BASE OF EXTANT ANGIOSPERMS? SPATIOTEMPORAL ARRANGEMENT OF ORGANS IN REPRODUCTIVE UNITS OF HYDATELLACEAE AND ITS BEARING ON THE ORIGIN OF THE FLOWER. Am J Bot 96:67–82. https://doi.org/10.3732/ajb.0800027

54. Rudall PJ, Sokoloff DD, Remizowa MV, et al (2007) Morphology of Hydatellaceae, an anomalous aquatic family recently recognized as an early-divergent angiosperm lineage. Am J Bot 94:1073–1092. https://doi.org/10.3732/ajb.94.7.1073

55. Friedman WE, Bachelier JB, Hormaza JI (2012) EMBRYOLOGY IN TRITHURIA SUBMERSA (HYDATELLACEAE) AND RELATIONSHIPS BETWEEN EMBRYO, ENDOSPERM, AND PERISPERM IN EARLY-DIVERGING FLOWERING PLANTS. Am J Bot 99:1083–1095. https://doi.org/10.3732/ajb.1200066

56. Tuckett RE, Merritt DJ, Rudall PJ, et al (2010) A new type of specialized morphophysiological dormancy and seed storage behaviour in Hydatellaceae, an early-divergent angiosperm family. Ann Bot 105:1053–1061. https://doi.org/10.1093/aob/mcq062

57. Kynast RG, Joseph JA, Pellicer J, et al (2014) CHROMOSOME BEHAVIOR AT THE BASE OF THE ANGIOSPERM RADIATION: KARYOLOGY. Am J Bot 101:1447–1455. https://doi.org/10.3732/ajb.1400050

58. Adsul AA, Patil SM, Yadav SR, Bapat VA (2012) In vitro culture of Trithuria konkanensis, one of the smallest angiosperms. Curr Sci 103:979–980

59. Vialette-Guiraud ACM, Alaux M, Legeai F, et al (2011) Cabomba as a model for studies of early angiosperm evolution. Ann Bot 108:589–598. https://doi.org/10.1093/aob/mcr088

60. Scheers K, Denys L, Jacobs I, et al (2019) Cabomba caroliniana Gray (Cabombaceae) invades major waterways in Belgium. Knowl Manag Aquat Ecosyst 22. https://doi.org/10.1051/kmae/2019014

61. Yamada T, Yokota S, Hirayama Y, et al (2011) Ancestral expression patterns and evolutionary diversification of YABBY genes in angiosperms. Plant J 67:26–36. https://doi.org/10.1111/j.1365-313X.2011.04570.x

62. THIEN L, WHITE D, YATSU L (1983) THE REPRODUCTIVE-BIOLOGY OF A RELICT ILLICIUM-FLORIDANUM ELLIS. Am J Bot 70:719–727. https://doi.org/10.2307/2443126

63. Prakash N, Alexander JH (1984) Self-incompatibility in Austrobaileya scandens. In: Williams EG, Knox RB (eds) 'Pollination '84'. University of Melbourne, Melbourne, pp 214–216

64. Bernhardt P, Sage T, Weston P, et al (2003) The pollination of Trimenia moorei (Trimeniaceae): Floral volatiles, insect/wind pollen vectors and stigmatic self-incompatibility in a basal angiosperm. Ann Bot 92:445–458. https://doi.org/10.1093/aob/mcg157

65. Allen AM, Hiscock SJ (2008) Evolution and phylogeny of self-incompatibility systems in Angiosperms. In: V. E.Franklin-Tong (ed) Self-Incompatibility in Flowering Plants. Springer Verlag, Berlin-Heidelberg, pp 73–101

66. Williams JH, Friedman WE (2004) The four-celled female gametophyte of Illicium (Illiciaceae; Austrobaileyales): Implications for understanding the origin and early evolution of monocots, eumagnoliids, and eudicots. Am J Bot 91:332–351. https://doi.org/10.3732/ajb.91.3.332

67. Bachelier JB, Friedman WE (2011) Female gamete competition in an ancient angiosperm lineage. Proc Natl Acad Sci U S A 108:12360–12365. https://doi.org/10.1073/pnas.1104697108

68. Qing-tian Z, Shu-tian F, Yi-ming Y, et al (2016) De novo transcriptome assembly of Schisandra chinensis Turcz. (Baill.). Genomics Data 10:153–154. https://doi.org/10.1016/j.gdata.2016.11.008

69. Wang S, Zhang Z, Song W-L, et al (2018) TRANSCRIPTOME ANALYSIS OF SCHISANDRA SPHENANTHERA DISCOVERS PUTATIVE LIGNAN BIOSYNTHESIS GENES AND GENETIC MARKERS. Pak J Bot 50:1047–1059

70. Bliss BJ, Wanke S, Barakat A, et al (2013) Characterization of the basal angiosperm Aristolochia fimbriata: a potential experimental system for genetic studies. Bmc Plant Biol 13:13. https://doi.org/10.1186/1471-2229-13-13

71. Qin L, Hu Y, Wang J, et al (2021) Insights into angiosperm evolution, floral development and chemical biosynthesis from the Aristolochia fimbriata genome. Nat Plants 7:1239-+. https://doi.org/10.1038/s41477-021-00990-2

72. Bliss BJ, Landherr L, dePamphilis CW, et al (2009) Regeneration and plantlet development from somatic tissues of Aristolochia fimbriata. Plant Cell Tissue Organ Cult 98:105–114. https://doi.org/10.1007/s11240-009-9543-9

73. Pabon-Mora N, Suarez-Baron H, Ambrose BA, Gonzalez F (2015) Flower Development and Perianth Identity Candidate Genes in the Basal Angiosperm Aristolochia fimbriata (Piperales: Aristolochiaceae). Front Plant Sci 6:. https://doi.org/10.3389/fpls.2015.01095

74. Perez-Mesa P, Suarez-Baron H, Ambrose BA, et al (2019) Floral MADS-box protein interactions in the early diverging angiosperm Aristolochia fimbriata Cham. (Aristolochiaceae: Piperales). Evol Dev 21:96–110. https://doi.org/10.1111/ede.12282

75. Rendon-Anaya M, Ibarra-Laclette E, Mendez-Bravo A, et al (2019) The avocado genome informs deep angiosperm phylogeny, highlights introgressive hybridization, and reveals pathogen-influenced gene space adaptation. Proc Natl Acad Sci U S A 116:17081–17089. https://doi.org/10.1073/pnas.1822129116

76. Chen J, Hao Z, Guang X, et al (2019) Liriodendron genome sheds light on angiosperm phylogeny and species-pair differentiation. Nat Plants 5:18–25. https://doi.org/10.1038/s41477-018-0323-6

77. Nepi M, Little S, Guarnieri M, et al (2017) Phylogenetic and functional signals in gymnosperm ovular secretions. Ann Bot 120:923–936. https://doi.org/10.1093/aob/mcx103

78. Friedman WE, Carmichael JS (1996) Double fertilization in gnetales: Implications for understanding reproductive diversification among seed plants. Int J Plant Sci 157:S77–S94. https://doi.org/10.1086/297405

79. Magallon S, Sanderson MJ (2002) Relationships among seed plants inferred from highly conserved genes: Sorting conflicting phylogenetic signals among ancient lineages. Am J Bot 89:1991–2006. https://doi.org/10.3732/ajb.89.12.1991

80. Mathews S, Clements MD, Beilstein MA (2010) A duplicate gene rooting of seed plants and the phylogenetic position of flowering plants. Philos Trans R Soc B-Biol Sci 365:383–395. https://doi.org/10.1098/rstb.2009.0233

81. Doyle JA (2012) Molecular and Fossil Evidence on the Origin of Angiosperms. In: Jeanloz R (ed) Annual Review of Earth and Planetary Sciences, Vol 40. pp 301–326

82. Lee EK, Cibrian-Jaramillo A, Kolokotronis S-O, et al (2011) A Functional Phylogenomic View of the Seed Plants. Plos Genet 7:e1002411. https://doi.org/10.1371/journal.pgen.1002411

83. Zhong B, Deusch O, Goremykin VV, et al (2011) Systematic Error in Seed Plant Phylogenomics. Genome Biol Evol 3:1340–1348. https://doi.org/10.1093/gbe/evr105

84. Christenhusz MJM, Reveal JL, Farjon A, et al (2011) A new classification and linear sequence of extant gymnosperms. Phytotaxa 19:55–70. https://doi.org/10.11646/phytotaxa.19.1.3

85. Niu S, Yuan H, Sun X, et al (2016) A transcriptomics investigation into pine reproductive organ development. New Phytol 209:1278–1289. https://doi.org/10.1111/nph.13680

86. Giacomello S, Salmen F, Terebieniec BK, et al (2017) Spatially resolved transcriptome profiling in model plant species. Nat Plants 3:17061. https://doi.org/10.1038/nplants.2017.61

87. Zhang PY, Tan HTW, Pwee KH, Kumar PP (2004) Conservation of class C function of floral organ development during 300 million years of evolution from gymnosperms to angiosperms. Plant J 37:566–577. https://doi.org/10.1046/j.1365-313X.2003.01983.x

88. Moyroud E, Monniaux M, Thevenon E, et al (2017) A link between LEAFY and B-gene homologues in Welwitschia mirabilis sheds light on ancestral mechanisms prefiguring floral development. New Phytol 216:469–481. https://doi.org/10.1111/nph.14483

89. Wang Y-Q, Melzer R, Theissen G (2010) Molecular interactions of orthologues of floral homeotic proteins from the gymnosperm Gnetum gnemon provide a clue to the evolutionary origin of “floral quartets.” Plant J 64:177–190. https://doi.org/10.1111/j.1365-313X.2010.04325.x

90. Winter KU, Saedler H, Theissen G (2002) On the origin of class B floral homeotic genes: functional substitution and dominant inhibition in Arabidopsis by expression of an orthologue from the gymnosperm Gnetum. Plant J 31:457–475. https://doi.org/10.1046/j.1365-313X.2002.01375.x

91. Uddenberg D, Akhter S, Ramachandran P, et al (2015) Sequenced genomes and rapidly emerging technologies pave the way for conifer evolutionary developmental biology. Front Plant Sci 6:970. https://doi.org/10.3389/fpls.2015.00970

92. Uddenberg D, Reimegard J, Clapham D, et al (2013) Early Cone Setting in Picea abies acrocona Is Associated with Increased Transcriptional Activity of a MADS Box Transcription Factor. Plant Physiol 161:813–823. https://doi.org/10.1104/pp.112.207746

93. Scutt CP (2018) The origin of angiosperms. Evol Dev Biol Ref Guide 1–20

94. Shi G, Herrera F, Herendeen PS, et al (2021) Mesozoic cupules and the origin of the angiosperm second integument. Nature 594:223-+. https://doi.org/10.1038/s41586-021-03598-w

95. Doyle JA (2008) Integrating molecular phylogenetic and paleobotanical evidence on origin of the flower. Int J Plant Sci 169:816–843. https://doi.org/10.1086/589887

96. Leebens-Mack JH, Barker MS, Carpenter EJ, et al (2019) One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574:679-+. https://doi.org/10.1038/s41586-019-1693-2

97. Flores-Tornero M, Proost S, Mutwil M, et al (2019) Transcriptomics of manually isolated Amborella trichopoda egg apparatus cells. Plant Reprod 32:15–27. https://doi.org/10.1007/s00497-019-00361-0

98. Rivarola Sena AC, Andres-Robin A, Vialette AC, et al Custom methods to identify conserved genetic modules applied to novel transcriptomic data from Amborella trichopoda. J Exp Bot. https://doi.org/10.1093/jxb/erac044

99. Leitch IJ, Hanson L (2002) DNA C-values in seven families fill phylogenetic gaps in the basal angiosperms. Bot J Linn Soc 140:175–179. https://doi.org/10.1046/j.1095-8339.2002.00096.x

100. Povilus RA, Friedman WE Transcriptomes across fertilization and seed development in the water lily Nymphaea thermarum (Nymphaeales): evidence for epigenetic patterning during reproduction. Plant Reprod. https://doi.org/10.1007/s00497-022-00438-3

101. Pellicer J, Kelly LJ, Magdalena C, Leitch IJ (2013) Insights into the dynamics of genome size and chromosome evolution in the early diverging angiosperm lineage Nymphaeales (water lilies). Genome 56:437–449. https://doi.org/10.1139/gen-2013-0039

102. Wu Q, Wu J, Li S-S, et al (2016) Transcriptome sequencing and metabolite analysis for revealing the blue flower formation in waterlily. Bmc Genomics 17:897. https://doi.org/10.1186/s12864-016-3226-9

103. Yoo M-J, Chanderbali AS, Altman NS, et al (2010) Evolutionary trends in the floral transcriptome: insights from one of the basalmost angiosperms, the water lily Nuphar advena (Nymphaeaceae). Plant J 64:687–698. https://doi.org/10.1111/j.1365-313X.2010.04357.x

104. Marques I, Montgomery SA, Barker MS, et al (2016) Transcriptome-derived evidence supports recent polyploidization and a major phylogeographic division in Trithuria submersa (Hydatellaceae, Nymphaeales). New Phytol 210:310–323. https://doi.org/10.1111/nph.13755

105. BHARATHAN G, LAMBERT G, GALBRAITH D (1994) NUCLEAR-DNA CONTENT OF MONOCOTYLEDONS AND RELATED TAXA. Am J Bot 81:381–386. https://doi.org/10.2307/2445466

106. Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Report 9:208–218

107. Yang Y, Xu M, Luo Q, et al (2014) De novo transcriptome analysis of Liriodendron chinense petals and leaves by Illumina sequencing. Gene 534:155–162. https://doi.org/10.1016/j.gene.2013.10.073

108. Wang K, Li M, Gao F, et al (2012) Identification of Conserved and Novel microRNAs from Liriodendron chinense Floral Tissues. Plos One 7:e44696. https://doi.org/10.1371/journal.pone.0044696

109. Lorenz WW, Ayyampalayam S, Bordeaux JM, et al (2012) Conifer DBMagic: a database housing multiple de novo transcriptome assemblies for 12 diverse conifer species. Tree Genet Genomes 8:1477–1485. https://doi.org/10.1007/s11295-012-0547-y

110. 0'Brien IEW, Smith DR, Gardner RC, Murray BG (1996) Flow cytometric determination of genome size in Pinus. Plant Sci 115:91–99

111. Joyner KL, Wang XR, Johnston JS, et al (2001) DNA content for Asian pines parallels New World relatives. Can J Bot-Rev Can Bot 79:192–196. https://doi.org/10.1139/b00-151

112. Niu S, Li J, Bo W, et al (2022) The Chinese pine genome and methylome unveil key features of conifer evolution. Cell 185:204-+. https://doi.org/10.1016/j.cell.2021.12.006

113. Du J, Zhang Z, Zhang H, Junhong T (2017) EST-SSR marker development and transcriptome sequencing analysis of different tissues of Korean pine (Pinus koraiensis Sieb. et Zucc.). Biotechnol Biotechnol Equip 31:679–689. https://doi.org/10.1080/13102818.2017.1331755

114. Fuchs J, Jovtchev G, Schubert I (2008) The chromosomal distribution of histone methylation marks in gymnosperms differs from that of angiosperms. Chromosome Res 16:891–898. https://doi.org/10.1007/s10577-008-1252-4

115. Nystedt B, Street NR, Wetterbom A, et al (2013) The Norway spruce genome sequence and conifer genome evolution. Nature 497:579–584. https://doi.org/10.1038/nature12211

116. Little SA, Boyes IG, Donaleshen K, et al (2016) A transcriptomic resource for Douglas-fir seed development and analysis of transcription during late megagametophyte development. Plant Reprod 29:273–286. https://doi.org/10.1007/s00497-016-0291-9

117. Neale DB, McGuire PE, Wheeler NC, et al (2017) The Douglas-Fir Genome Sequence Reveals Specialization of the Photosynthetic Apparatus in Pinaceae. G3-Genes Genomes Genet 7:3157–3167. https://doi.org/10.1534/g3.117.300078

118. Hizume M, Kondo T, Shibata F, Ishizuka R (2001) Flow cytometric determination of genome size in the Taxodiaceae, Cuprssaceae sensu stricto and Sciadopityaceae. Cytologia (Tokyo) 66:307–311

119. Zhou W, Chen Q, Wang X-B, et al (2019) De novo assembly of the Platycladus orientalis (L.) Franco transcriptome provides insight into the development and pollination mechanism of

female cone based on RNA-Seq data. Sci Rep 9:10191. https://doi.org/10.1038/s41598-019-46696-6

120. Pirone-Davies C, Prior N, von Aderkas P, et al (2016) Insights from the pollination drop proteome and the ovule transcriptome of Cephalotaxus at the time of pollination drop production. Ann Bot 117:973–984. https://doi.org/10.1093/aob/mcw026

121. Wang L, Lu Z, Li W, et al (2016) Global comparative analysis of expressed genes in ovules and leaves of Ginkgo biloba L. Tree Genet Genomes 12:29. https://doi.org/10.1007/s11295-016-0989-8

122. Wang L, Zhao J, Luo K, et al (2016) Deep sequencing discovery and profiling of conserved and novel miRNAs in the ovule of Ginkgo biloba. Trees-Struct Funct 30:1557–1567. https://doi.org/10.1007/s00468-016-1389-2

123. Smarda P, Vesely P, Smerda J, et al (2016) Polyploidy in a "living fossil' Ginkgo biloba. New Phytol 212:11–14. https://doi.org/10.1111/nph.14062

124. Guan R, Zhao Y, Zhang H, et al (2016) Draft genome of the living fossil Ginkgo biloba. Gigascience 5:49. https://doi.org/10.1186/s13742-016-0154-1

125. Scutt CP, Vandenbussche M (2014) Current trends and future directions in flower development research. Ann Bot 114:1399–1406. https://doi.org/10.1093/aob/mcu224

Table 1. Online databases of genomic and transcriptomic resources and other information useful for studies of the origin of flowers.

Name	Web address	Description
PlaBiPD	https://www.plabipd.de/	Continually updated listings of published plant genomes, presented as cladograms with links to the relevant publications, among other resources.
OneKP Database [96]	https://db.cngb.org/onekp/	Transcriptomes of >1000 plant species available for downloading or BLAST searching.
CoGe	https://genomevolution.org/coge/	Comparative genomics platform covering numerous plant genomes and containing many useful features.
PHYTOZOME	https://phytozome-next.jgi.doe.gov/	Comparative plant genomics platform with many useful features for data-searching and bulk-downloading etc.
PLAZA	https://bioinformatics.psb.ugent.be/plaza/	Comparative plant genomics platform with many useful features including micro-synteny analyses and phylogenetic pipeline construction. Separate instances for dicots, monocots, gymnosperms etc.
PlantGenie	https://plantgenie.org/	Plant genomics/transcriptomics platform including several gymnosperms.

World Flora Online	http://www.worldfloraonline.org/	Authoritative and comprehensive resource for plant taxonomy.
Angiosperm Phylogeny Group [16]	https://www.mobot.org/MOBOT/research/APweb/	Authoritative, comprehensive, periodically updated consensus view of angiosperm phylogeny.
Kew C-value database	https://cvalues.science.kew.org/	World-leading database of plant genome size estimations.
Kew Millennium Seed Bank database	https://data.kew.org/sid/sidsearch.html	World-leading database of seed dormancy types and optimum storage conditions.
TimeTree	http://www.timetree.org/	Convenient resource to generate dated species phylogenies.

Table 2. Transcriptomic resources including reproductive tissues from basally-diverging angiosperms and gymnosperms, with corresponding genomic data where available.

Species	Taxonomy	Transcriptomic resources	Nuclear genome size	Nuclear genome sequence
Amborella trichopda	ANA-grade angiosperms, Amborellales, Amborellaceae	RNA-seq data of tepal, leaf, root and egg-apparatus [97], male and female flower buds [34], and laser-microdissected female flower tissues [98].	01c =.89 pg [99]; 870 Mb [42]	Genome V1 [42] https://phytozome-next.jgi.doe.gov/ Genome V6. (chromosome-level assembly) https://genomevolution.org/coge/
Nymphaea thermarum	ANA-grade angiosperms, Nymphaeales, Nymphaeaceae	RNA-seq of mature ovules and developing seeds [100]	1c = 0.51 pg [101]; 497 Mb [49]	Povilus *et al*. [49]
Nymphaea 'King of Siam'	ANA-grade angiosperms, Nymphaeales, Nymphaeaceae	RNA-seq of petals [102]		
Nuphar advena	ANA-grade angiosperms, Nymphaeales, Nymphaeaceae	Flower EST microarray analyses [103]	1c = 2.78 pg [101]	
Trithuria submersa	ANA-grade angiosperms, Nymphaeales, Hydatellaceae	RNA-seq of whole plants, including reproductive tissues [104]	1c = 1.37 pg [57], but probably octaploid [47]	
Aristolochia fimbriata	Magnollids, Piperales, Aristolochiaceae	RNA-seq of dissected floral organs and other tissues [71]	1c = 0.50 [105] ; 258 Mb [71]	Qin *et al*. [71]
Persea americana	Magnollids, Laurales, Lauraceae	Microarrays of flower ESTs [103]	1c = 0.92 [106]; 840 Mb [75]	Rendon-Anaya *et al*. [75]
Liriodendron chinense	Magnollids, Laurales, Magnoliaceae	Petal and leaf transcriptome [107]; Flower-expressed microRNAs [108]	1.75 Gb [76]	Chen *et al*. [76]
Pinus taeda	Gymnosperms, Pinaceae	RNA-seq of mixed tissues including female cones [109];	1c = 22.10 [110] ; 11.6 Gb	https://plantgenie.org/

		https://plantgenie.org/		
Pinus tabuliformis	Gymnosperms, Pinaceae	RNA-seq and microarrays of male and female cone stages [85]	1c = 25.70 [111] ; 25.4 Gb [112]	Niu *et al*. [112]
Pinus koraiensis	Gymnosperms, Pinaceae	RNA-seq of mixed tissues including cones [113]	1c = 28.20 [111] ;	
Picea abies	Gymnosperms, Pinaceae	Transcriptome profiling of female cones [86]	1c = 20.01 [114] ; 19.8 Gb [115]	https://plantgenie.org/; Nystedt *et al*. [115]
Pseudotsuga menziesii	Gymnosperms, Pinaceae	RNA-seq of megagametophyte stages [116]	1c = 19.05 [110] ; 15.7 Gb [117]	Neale *et al*. [117]
Cryptomeria japonica	Gymnosperms, Cupressaceae	RNA-seq including cone tissues https://db.cngb.org/onekp/	1c = 11.05 [118]	
Platycladus orientalis	Gymnosperms, Cupressaceae	RNA-seq of female cones [119]	1c = 10.46 [118]	
Cephalotaxus sinensis	Gymnosperms, Cephalotaxaceae	Ovule RNA-seq (Pirone-[120]		
Ginkgo biloba	Gymnosperms, Ginkgoales, Ginkgoaceae	Ovule RNA-seq [121]; Ovule small RNA-seq [122]	1c = 9.39 [123] ; 11.75 Gb [124]	Guan *et al*. [124]
Cycas panzhihuaensis	Gymnosperms, Cycadales, Cycadaceae	RNA-seq of both reproductive and vegetative tissues [29]	10.5 Gb [29]	Lui *et al*. [29]

Figures

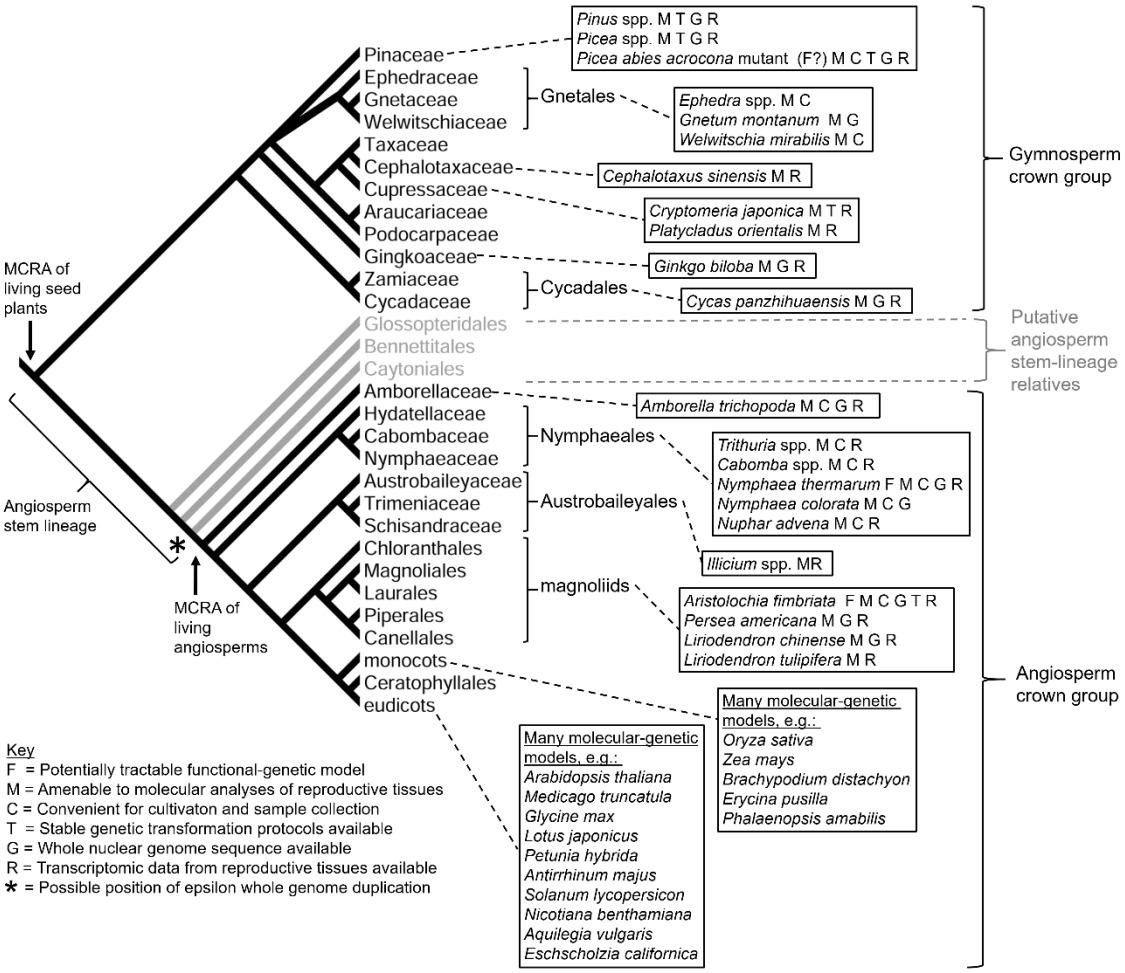

Figure 1. Experimental features of basally-diverging angiosperms and gymnosperms. Useful features (see key) for experimental purposes are listed after boxed species names. The schematic phylogeny is based on Doyle [95] and Byng *et al*. [16], but collapses Gnetales, Pinaceae and other conifers to a polytomy. The placement of putative angiosperm stem-lineage relatives (shown in grey) from the fossil record is based on Doyle [95].

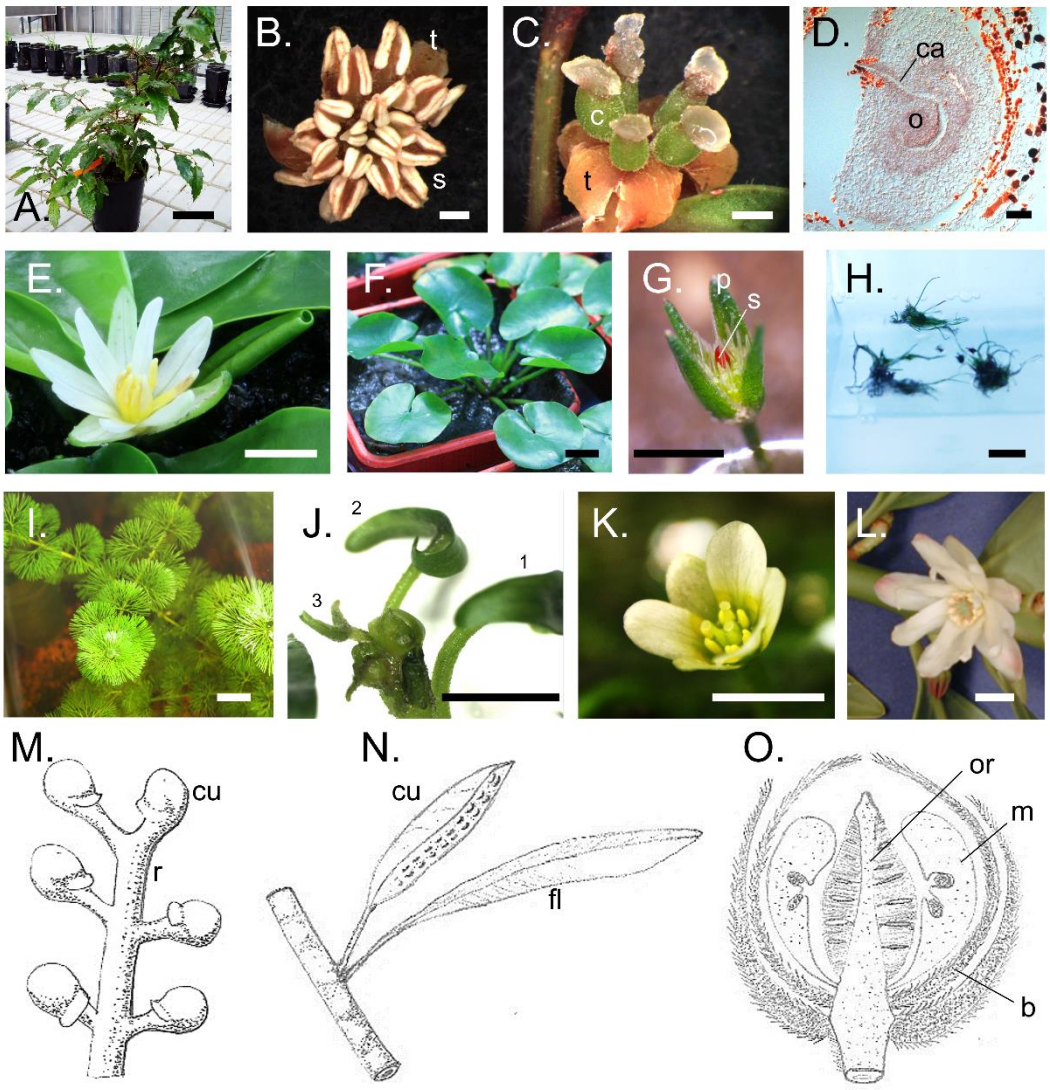

Figure 2. ANA-grade angiosperms and putative angiosperm stem-lineage relatives.

A-D. *Amborella trichopoda:* 3-year old plant in flower (A.), male flower (B.), female flower (C.) and median longitudinal section of developing carpel (D.). E-F. *Nymphaea thermarum:* newly opened flower (E.) and young plant (F.). G-H. *Trithuria submersa:* reproductive unit (G.) and adult plants growing on agar (F.). I-K. *Cabomba caroliniana*: submerged leaves and stems (I.), numbered succession of young floating leaves subtending flower buds present in the mass of tissues at the centre (J.), newly opened flower (K.). L. *Illicium anisatum*, mature flower. M. *Caytonia* (Caytoniales) female cupule-bearing rachis. N. *Dictyopteridium* (Glossopteridales), female phytomer, O. *Williamsoniella* (Bennettitales), flower-like bisexual reproductive axis in longitudinal section. Panels C., D., G. [125], I-K. [59] and M-0. [93] are reproduced with permission from published sources. b = bract, c = carpel, ca = canal (for pollen tube growth), cu = cupule, fl = fertile leaf, m = microsporophyll, o = ovule, or = ovuliferous receptacle, p =

bract-like phyllome, r = rachis, t = tepal. Scale bars: 10 cm in A.; 1 mm in B., C. and G.; 50 µm in D.; 1cm in E., F., H. and I-L.