



Review

Evolution and biogeography of gymnosperms



Xiao-Quan Wang*, Jin-Hua Ran

State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

ARTICLE INFO

Article history:

Received 30 June 2013

Revised 6 February 2014

Accepted 10 February 2014

Available online 22 February 2014

Keywords:

Phylogeny
Biogeography
Evolution
Gnetales
Conifers
Gymnosperms

ABSTRACT

Living gymnosperms comprise only a little more than 1000 species, but represent four of the five main lineages of seed plants, including cycads, ginkgos, gnetophytes and conifers. This group has huge ecological and economic value, and has drawn great interest from the scientific community. Here we review recent advances in our understanding of gymnosperm evolution and biogeography, including phylogenetic relationships at different taxonomic levels, patterns of species diversification, roles of vicariance and dispersal in development of intercontinental disjunctions, modes of molecular evolution in different genomes and lineages, and mechanisms underlying the formation of large nuclear genomes. It is particularly interesting that increasing evidence supports a sister relationship between Gnetales and Pinaceae (the Gnepine hypothesis) and the contribution of recent radiations to present species diversity, and that expansion of retrotransposons is responsible for the large and complex nuclear genome of gymnosperms. In addition, multiple coniferous genera such as *Picea* very likely originated in North America and migrated into the Old World, further indicating that the center of diversity is not necessarily the place of origin. The Bering Land Bridge acted as an important pathway for dispersal of gymnosperms in the Northern Hemisphere. Moreover, the genome sequences of conifers provide an unprecedented opportunity and an important platform for the evolutionary studies of gymnosperms, and will also shed new light on evolution of many important gene families and biological pathways in seed plants.

© 2014 Elsevier Inc. All rights reserved.

Contents

1. Introduction	25
2. Diversity and classification of gymnosperms	25
3. Phylogeny and evolution of gymnosperms	25
3.1. Origin and diversification	25
3.2. Phylogenetic reconstruction	28
3.2.1. The phylogenetic position of Gnetales	28
3.2.2. Phylogenetic studies of other gymnosperms	30
3.3. Molecular and genome evolution	31
3.3.1. Chromosomal and genome size variation	31
3.3.2. Composition and evolutionary patterns of the three genomes of gymnosperms	32
4. Biogeography of gymnosperms	33
4.1. Disjunctive distribution in the two hemispheres	33
4.2. Disjunctive distribution in the Southern Hemisphere	34
4.3. Disjunctive distribution in the Northern Hemisphere	34
4.4. Biogeographic difference between the two hemispheres	35
5. Concluding remarks	35
Acknowledgments	36

* Corresponding author. Address: State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, 20 Nanxincun, Xiangshan, Beijing 100093, China. Fax: +86 10 62590843.

E-mail address: xiaoq_wang@ibcas.ac.cn (X.-Q. Wang).

Appendix A. Supplementary material	36
References	36

1. Introduction

Gymnosperms are of great ecological and economic importance, although this ancient and widespread plant lineage currently comprises only a little more than 1000 species that are two to three orders of magnitude lower than the approximately 300,000 species of extant angiosperms. Also, from the evolutionary perspective, studies of angiosperms depend a lot on our knowledge of gymnosperms given the sister relationship between the two groups. However, the evolutionary study of gymnosperms is still facing great difficulties in the genomic era due to the large genome size, high heterozygosity and long generation time of this group, although a draft assembly of the Norway spruce (*Picea abies*) genome was generated by Nystedt et al. (2013) and several comparative genomics projects involving gymnosperms are being carried out, such as the 1000 Plants (oneKP or 1KP) Initiative (<http://www.onekp.com/>), the Plant Genomics Consortium (<http://sciweb.nybg.org/science2/GenomicsLab.asp>), and the Conifer Genome Network (<http://www.pinegenome.org/index.php>). Nevertheless, in recent years, fascinating progress has been made in our understanding of evolution and biogeography of gymnosperms, which inspires us to write this review. For a better understanding of the content, we first give a brief introduction of the diversity and classification of gymnosperms. Then, we focus on: (1) Phylogeny and evolution of gymnosperms, including evolutionary history, phylogenetic relationships, and molecular and genome evolution; (2) Historical biogeography of gymnosperms.

2. Diversity and classification of gymnosperms

Living gymnosperms are distributed in all continents except Antarctica, of which two-thirds are conifers, a group that constitutes over 39% of the world's forests (Armenise et al., 2012). The gymnosperms play major roles in global carbon cycles, provide important sources of timber, resins and even drugs and foods (Zonneveld, 2012c; Murray, 2013), and are crucial to preventing soil erosion. Additionally, they are a mainstay of gardening.

Gymnosperms represent four of the five main lineages of seed plants, i.e., cycads, ginkgos, gnetophytes and conifers (including cupressophytes and Pinaceae), and were recently classified into four subclasses (Ginkgoideae, Cycadidae, Pinidae and Gnetidae) under the class Equisetopsida (Chase and Reveal, 2009). They comprise 12 families, 83 genera (Christenhusz et al., 2011), and about 1000 species (Table 1), including ca 297–331 species of cycads in 10 genera, one extant ginkgophyte, 80–100 gnetophytes in three genera, and ca 614 species of conifers in 69 genera (Farjón, 2010; Christenhusz et al., 2011). Among these genera, 34 (40.96%) are monotypic, 22 (26.5%) have only two to five species, and only three (*Cycas*, *Pinus* and *Podocarpus*) harbor near or more than 100 species (Table 1, and Fig. 1). It is interesting that half (45) of the genera occur in Asia and 31 in Australia (continent), and the vast majority of the monotypic genera are found in these two continents (Fig. 2).

As the largest lineage of gymnosperms, conifers were divided into seven families by Pilger (1926), including Taxaceae, Podocarpaceae, Araucariaceae, Cephalotaxaceae, Pinaceae, Taxodiaceae and Cupressaceae. However, Eckenwalder (1976) proposed a merger of Taxodiaceae and Cupressaceae based on the phenetic analysis and Hayata (1931) proposed to place *Sciadopitys* in a separate family (Sciadopityaceae), and these views have been adopted in most of the following classification schemes of gymnosperms (e.g., Farjón, 2001, 2005; Christenhusz et al., 2011)

and supported by most non-molecular and molecular phylogenetic studies (e.g., Hart, 1987; Price and Lowenstein, 1989; Brunsfeld et al., 1994; Gadek et al., 2000; Yang et al., 2012). Currently, it is still controversial whether Taxaceae and Cephalotaxaceae should be merged into a single family (e.g., Quinn et al., 2002; Hao et al., 2008; Christenhusz et al., 2011; Ghimire and Heo, 2014). Recently, Eckenwalder (2009) and Farjón (2010) published two very valuable books on all conifers, recognizing 546 and 615 species, respectively. Although both books provided an identification guide to each species, Farjón incorporated more recent advances in the systematics of conifers and recognized more species, genera, and even families than Eckenwalder. For example, Cephalotaxaceae and Phyllocladaceae were recognized by Farjón (2010), but were put into Taxaceae and Podocarpaceae, respectively, by Eckenwalder (2009). In addition, Farjón (2010) recognized three extra genera, *Pilgerodendron* and *Xanthocyparis* in Cupressaceae and *Sundacarpus* in Podocarpaceae. At the family level, Farjón (1990, 2005) published two excellent monographs on Pinaceae and Cupressaceae s.l., respectively. An interesting thing is that recent phenotypic and molecular phylogenetic studies do not support the monophyly of *Cupressus* (Cupressaceae). Adams et al. (2009) divided this genus into two lineages, including *Cupressus* s.s. comprising the Old World species and a new genus (*Hesperocyparis*) comprising the New World species that are closely related to two small controversial genera, i.e., *Callitropsis* from northwestern North America and *Xanthocyparis* from northern Vietnam (Little, 2006; Yang et al., 2012). However, except that *Xanthocyparis* was accepted by Farjón (2010), the other three genera (*Callitropsis*, *Hesperocyparis* and *Xanthocyparis*) were not accepted by Eckenwalder (2009), Farjón (2010) and Christenhusz et al. (2011). At present, it is widely accepted that conifers comprise two major clades, Pinaceae and the remaining non-Pinaceae conifers (Conifer II or Cupressophytes) (see Section 3.2, phylogenetic reconstruction), with Pinaceae and Podocarpaceae representing the first and second largest families (Farjón, 2001; Knopf et al., 2012).

The gnetophytes comprise three families (Ephedraceae, Gnetaceae and Welwitschiaceae), each containing a single genus (Table 1). Compared to gnetophytes and conifers, there were more debates on the classification of cycads. Initially, all living species of cycads were placed in a single family, the Cycadaceae (see reviews by Stevenson, 1990 and Jones, 2002). However, afterwards, three to four families, including Cycadaceae, Stangeriaceae, Zamiaceae and Boweniaceae, were recognized by different authors (Johnson, 1959; Stevenson, 1981, 1990, 1992). The Boweniaceae was erected by Stevenson (1981), but was merged into Stangeriaceae by Stevenson (1992). Recently, molecular phylogenetic studies support a division of the 10 cycad genera into two families (Cycadaceae and Zamiaceae) (Treutlein and Wink, 2002; Hill et al., 2003; Chaw et al., 2005; Zgurski et al., 2008; Nagalingum et al., 2011; Salas-Leiva et al., 2013), although the genus status of *Chigua* is still accepted by some researches (<http://plantnet.rbg-syd.nsw.gov.au/PlantNet/cycad/>) (see review by Osborne et al., 2012).

3. Phylogeny and evolution of gymnosperms

3.1. Origin and diversification

Based on fossil evidence and molecular clock calibration, the divergence between gymnosperms and angiosperms could be dated to about 300–350 million years ago (Mya) in the

Table 1
Extant genera of gymnosperms and their biological and genome information.

Families	Genera	Number of species		Genome size (pg/2C)	No. of chromosomes (1n)	Distribution
		Farjón (2010) ^e	Eckenwalder (2009) ^e			
Pinaceae						
	<i>Abies</i>	47	40	35.28 ± 2.38	12	AF (N), AS, EU, NA
	<i>Cathaya</i>	1	1	49.5	12	AS
	<i>Cedrus</i>	3	2	32.16 ± 0.65	12	AF (N), AS, EU
	<i>Keteleeria</i>	3	2	48.4	12	AS
	<i>Larix</i>	11	10	25.17 ± 3.64	12	AS, EU, NA
	<i>Nothotsuga</i>	1	1		12	AS
	<i>Picea</i>	38	29	36.20 ± 2.76	12	AS, EU, NA
	<i>Pinus</i>	113	97	52.86 ± 8.21	12	AF (N), AS, EU, NA
	<i>Pseudolarix</i>	1	1	52.2	22	AS
	<i>Pseudotsuga</i>	4	4	38.1	12, 13	AS, NA
	<i>Tsuga</i>	9	8	38.14 ± 3.46	12	AS, NA
	Total	231	195			
Cupressaceae						
	<i>Actinostrobus</i>	3	3	21.30	11	AU
	<i>Athrotaxis</i>	3	2	20.17 ± 0.06	11	AU
	<i>Austrocedrus</i>	1	1	21.80	11	SA
	<i>Callitris</i>	15	17	18.43 ± 2.18	11	AU
	<i>Calocedrus</i>	4	3	33.80 ± 3.21	11	AS, NA
	<i>Chamaecyparis</i>	5	5	21.16 ± 2.88	11	AS, NA
	<i>Cryptomeria</i>	1	1	22.48 ± 0.55	11	AS
	<i>Cunninghamia</i>	2	2	26.92 ± 2.01	11	AS
	<i>Cupressus</i>	15	17	23.64 ± 2.06	11	AF (N), AS, EU, NA
	<i>Diselma</i>	1	1	18.1	11	AU
	<i>Fitzroya</i>	1	1	35	22	SA
	<i>Fokienia</i>	1	1	25.05	11	AS
	<i>Glyptostrobus</i>	1	1	19.97	11	AS
	<i>Juniperus</i>	53	54	27.33 ± 8.51	11, 22	AF, AS, EU, NA
	<i>Libocedrus</i>	5	6	35.70 ± 6.60	11	AU
	<i>Metasequoia</i>	1	1	22.08	11	AS
	<i>Microbiota</i>	1	1	20.34	11	AS
	<i>Neocallitropsis</i>	1	1	25.5	11	AU
	<i>Papuacedrus</i>	1	1	24	11	AU
	<i>Pilgerodendron</i> ^a	1	0		11	SA
	<i>Platycladus</i>	1	1	21.86 ± 1.34	11	AS
	<i>Sequoia</i>	1	1	64.27	33	NA
	<i>Sequoiadendron</i>	1	1	19.85	11	NA
	<i>Taiwania</i>	1	1	30.22 ± 6.83	11	AS
	<i>Taxodium</i>	2	2	19.00 ± 1.27	11	NA
	<i>Tetraclinis</i>	1	1	25.7	11	AF (N), EU
	<i>Thuja</i>	5	5	25.07 ± 1.20	11	AS, NA
	<i>Thujopsis</i>	1	1	27.93	11	AS
	<i>Widdringtonia</i>	4	4	21.35 ± 0.50	11	AF
	<i>Xanthocyparis</i> ^{a,b}	2	0	22.80	11	AS, NA
	Total	135	136			
Taxaceae						
	<i>Amentotaxus</i>	6	2	60.4	18	AS
	<i>Austrotaxus</i>	1	1		–	AU
	<i>Pseudotaxus</i>	1	1	34.6	12	AS
	<i>Taxus</i>	10	8	23.06 ± 0.77	12	AS, EU, NA
	<i>Torreya</i>	6	6	43.97 ± 0.71	11, 12	AS, NA
	Total	24	18			
Cephalotaxaceae ^a						
	<i>Cephalotaxus</i>	8	5	50.7	12	AS
Sciadopityaceae						
	<i>Sciadopitys</i>	1	1	41.6	10	AS
Araucariaceae						
	<i>Agathis</i>	17	15	29.25 ± 3.32	13	AS, AU
	<i>Araucaria</i>	19	19	33.33 ± 6.27	13	AU, SA
	<i>Wollemia</i>	1	1	27.87	13	AU
	Total	37	35			
Podocarpaceae						
	<i>Acmopyle</i>	2	2	15.65 ± 2.62	10	AU
	<i>Afrocarpus</i>	5	2	10.90 ± 0.99	12	AF
	<i>Dacrycarpus</i>	9	9	23.45 ± 18.03	10	AS, AU
	<i>Dacrydium</i>	22	21	17.40 ± 9.12	10	AS, AU

Table 1 (continued)

Families	Genera	Number of species		Genome size (pg/2C)	No. of chromosomes (1n)	Distribution
		Farjón (2010) ^e	Eckenwalder (2009) ^e			
	<i>Falcatifolium</i>	6	5	22.40	10	AS, AU
	<i>Halocarpus</i>	3	3	18.27 ± 4.45	9, 11, 12	AU
	<i>Lagarostrobos</i>	1	1	30.40	15	AU
	<i>Lepidothamnus</i>	3	3	12.00 ± 2.17	14, 15	AU, SA
	<i>Manoao</i>	1	1	27.60	10	AU
	<i>Microcachrys</i>	1	1	8.30	15	AU
	<i>Nageia</i>	5	5	11.20	10, 13	AS, AU
	<i>Parasitaxus</i>	1	1	18	18	AU
	<i>Pherosphaera/Microstrobos</i>	2	2	8.50 ± 0.14	13	AU
	<i>Podocarpus</i>	97	82	18.72 ± 2.58	10, 11, 17–19	AF, AS, AU, SA
	<i>Prumnopitys</i>	9	8	14.13 ± 2.22	18, 19	AU, SA
	<i>Retrophyllum</i>	5	4	11.80	10	AS, AU, SA
	<i>Saxegothaea</i>	1	1	10.20	12	SA
	<i>Sundacarpus</i> ^a	1	0	–	–	AS, AU
Phyllocladaceae ^{a,b}	<i>Phyllocladus</i>	4	5	20.18 ± 3.11	9	AS, AU
	Total	178	156			
Gnetaceae	<i>Gnetum</i>	39 ^f		6.73 ± 1.94	11, 22	AF, AS, SA
Welwitschiaceae	<i>Welwitschia</i>	1 ^f		14.40	21	AF
Ephedraceae	<i>Ephedra</i>	50 ^f		31.37 ± 5.91	7, 14, 28	AF, AS, EU, NA, SA
Ginkgoaceae	<i>Ginkgo</i>	1	1	23.50	12	AS
Cycadaceae	<i>Cycas</i>	107 ^c	103 ^d	26.83 ± 1.37	11	AF, AS, AU
Zamiaceae	<i>Bowenia</i>	2 ^c	2 ^d	33.48 ± 11.20	9	AU
	<i>Ceratozamia</i>	27 ^c	18 ^d	63.25 ± 0.07	8	NA
	<i>Chigua</i> ^g	0	2 ^d			SA
	<i>Dioon</i>	14 ^c	11 ^d	49.50 ± 1.13	9	NA
	<i>Encephalartos</i>	65 ^c	64 ^d	58.36 ± 5.00	9	AF
	<i>Lepidozamia</i>	2 ^c	2 ^d	57.80 ± 3.54	9	AU
	<i>Macrozamia</i>	41 ^c	40 ^d	53.40 ± 0.57	9	AU
	<i>Microcycas</i>	1 ^c	1 ^d	41.20	13	NA
	<i>Stangeria</i>	1 ^c	1 ^d	29.64	8	AF
	<i>Zamia</i>	71 ^c	53 ^d	35.76 ± 6.92	8–14	NA, SA
	Total	224	194			

AF, Africa; AF (N), North Africa; AS, Asia; AU, Australia (continent); NA, North America; SA, South America; EU, Europe.

Information of chromosome numbers mainly comes from Plant DNA C-values Database (<http://data.kew.org/cvalues>) and Murray (2013). The average genome size of each genus was also calculated based on information from the Plant DNA C-values Database (before May 4, 2013).

^a Genus or family accepted by Farjón (2010) but not by Eckenwalder (2009).

^b Genus or family accepted by Farjón (2010) but not by Christenhusz et al. (2011).

^c Species number from Osborne et al. (2012).

^d Species number from the Cycad Page (CP) (<http://plantnet.rbg.gov.au/PlantNet/cycad/wlist.html>).

^e Only numbers of conifer species were listed.

^f Species number from Price (1996).

^g Genus not accepted by Eckenwalder (2009) and Osborne et al. (2012) but by CP.

Carboniferous (Hedges et al., 2006; Won and Renner, 2006; Clarke et al., 2011; Crisp and Cook, 2011; Magallón et al., 2013). The five main lineages of gymnosperms (cycads, ginkgos, cupressophytes, Pinaceae and gnetophytes) separated from each other during the Late Carboniferous to the Late Triassic (311–212 Mya), which were much earlier than the occurrence of the earliest extant angiosperms (Magallón et al., 2013). Although only about 1000 extant species of gymnosperms inhabit our planet, gymnosperms were dominant through much of the Mesozoic.

Considering the earliest fossil record of gymnosperms from the Palaeozoic (reviewed in Won and Renner, 2006) and the highly conserved morphology of this group, with a number of species being referred to as “living fossils”, such as *Metasequoia glyptostroboides*, *Ginkgo biloba* (Niklas, 1997) and cycads (Norstog and

Nicholls, 1997), scientists have long assumed that much of the species diversity of living gymnosperms is relictual, representing the last remnants of their prosperous past (Florin, 1963; Hill and Brodribb, 1999; Miller, 1999; McLoughlin and Vajda, 2005; Keppel et al., 2008; Xiao et al., 2010; Alvarez-Yepiz et al., 2011). However, the gymnosperm diversity has experienced interesting pulses of extinction and speciation (Davis and Schaefer, 2011). Most of extant gymnosperm species, even genera, are much younger than we thought before (Crisp and Cook, 2011; Nagalingum et al., 2011). For example, cycads were often cited as an example of living fossils, because it was believed that this lineage of gymnosperms reached their greatest diversity during the Jurassic–Cretaceous and then declined to the present diversity of around 300 species. Unexpectedly, however, Nagalingum et al. (2011) found that living

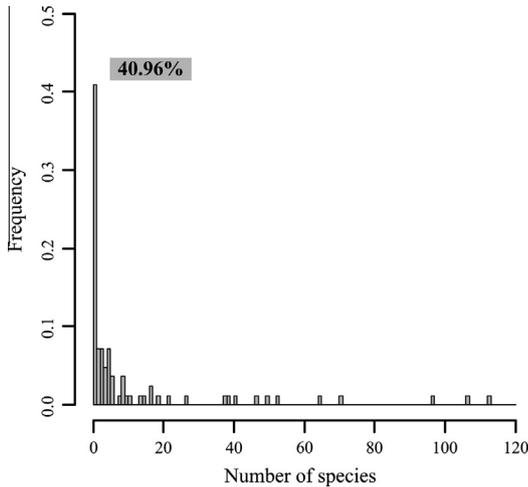


Fig. 1. Frequency distribution of species numbers of all gymnospermous genera. The species numbers of coniferous genera follow Farjón (2010), while those of cycad genera follow Osborne et al. (2012). The detailed information of each genus is shown in Table 1.

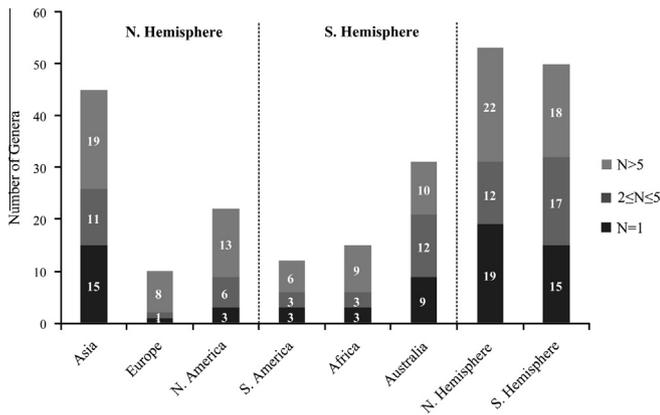


Fig. 2. Number of gymnospermous genera in different continents and their species richness. The distribution and species number of each genus correspond to Table 1. N, number of species in a genus.

cycad species are not much older than about 12 million years, suggesting recent synchronous radiation of a living fossil and coevolution between living cycads and their insect pollinators. In addition, Crisp and Cook (2011) estimated the ages and diversification rates of gymnosperm lineages, and found that living gymnosperm groups are not ancient, occupy diverse habitats and some probably survived after making adaptive shifts. Actually, the radiative speciation in the middle to late Cenozoic with a very low interspecific genetic differentiation has been reported in most of the studied gymnospermous genera with multiple species, such as *Agathis*, *Araucaria*, *Cycas*, *Ephedra*, *Gnetum*, *Juniperus*, *Picea*, *Pinus*, and *Podocarpus* (Hill, 1995; Wagstaff, 2004; Ran et al., 2006; Won and Renner, 2006; Willyard et al., 2007; Ickert-Bond et al., 2009; Biffin et al., 2010; Mao et al., 2010; Crisp and Cook, 2011; Nagalingum et al., 2011; Gaudeul et al., 2012; Leslie et al., 2012).

The low species diversity of extant gymnosperms could be attributed to the Cenozoic extinctions, although recent radiations have occurred in some lineages such as cycads (Crisp and Cook, 2011; Nagalingum et al., 2011). The extinctions at the Cretaceous–Palaeogene boundary affected major clades of gymnosperms and angiosperms indifferently (Macphail et al., 1994; Niklas, 1997; Wing, 2004; Crepet and Niklas, 2009). It was suggested that the

sharp cooling and drying of the global climate at the end of the Eocene caused the extinction of several conifer and cycad lineages because gymnosperms probably occupied warmer and wetter aseasonal environments during much of their early history (Hill, 1998, 2004; Hill and Brodribb, 1999). The climatic changes also stimulated some lineages that survived to undergo successful adaptive shifts and rediversify in new environments (Hill and Brodribb, 1999; Pittermann et al., 2012), such as *Callitris* (Paull and Hill, 2010) and *Macrozamia* (Carpenter, 1991). Leslie et al. (2012) found that the lineages of gymnosperms that diversified mainly in the Southern Hemisphere show a significantly older distribution of divergence ages than their counterparts in the Northern Hemisphere, which could be attributed to the fundamental differences between the two hemispheres in the distribution of oceans and landmasses. They also inferred that the complex patterns of migration and range shifts during climatic cycles in the Neogene could have led to elevated rates of speciation and extinction and resulted in the abundance of recent divergence in northern clades, whereas the scattered persistence of mild, wetter habitats in the Southern Hemisphere might have favored the survival of older lineages (Leslie et al., 2012).

In the long history of diversification, the morphology of gymnosperms has also undergone frequent parallel or convergent evolution, such as erect cones, deciduous needles and seed scale abscission in Pinaceae (Wang et al., 2000), quadrangular leaves linked to drought-tolerance in *Picea* (Ran et al., 2006), winged seeds adapted to wind-dispersal in *Pinus*, small and imbricate leaves that can minimize water loss in Cupressaceae *s.l.* (Pittermann et al., 2012), colorful and fleshy cone bracts adapted to animal dispersal in *Ephedra* (Hollander and Vander Wall, 2009). In particular, Lovisetto et al. (2012) found that a series of different MADS-box genes are involved in the formation of gymnosperm fruit-like structures, and the same gene types have been recruited in the two phylogenetically distant species *Ginkgo biloba* and *Taxus baccata* to make fleshy structures with different anatomical origins. A further study showed that the gymnosperm B-sister genes may have a main function in ovule/seed development and a subsidiary role in the formation of fleshy fruit-like structures that have an ovular origin, as in *Ginkgo* (Lovisetto et al., 2013). Therefore, these non-synapomorphic characters should be used very carefully in phylogenetic reconstruction and infrageneric classification.

3.2. Phylogenetic reconstruction

A number of molecular phylogenies of gymnosperms have been published since the first molecular study that supports the sister relationship between extant gymnosperms and angiosperms was conducted by Hori et al. (1985) using 5S rRNA sequences. While some interfamilial and intergeneric relationships have been resolved, more interesting phylogenetic hypotheses, especially on the evolutionary relationships of deep branches of gymnosperms, have been proposed and hotly debated, such as the phylogenetic position of gnetophytes (Chaw et al., 1997, 2000; Bowe et al., 2000; Mathews, 2009; Ran et al., 2010; Yang et al., 2012). In recent years, low-copy nuclear genes and EST sequences have been used in phylogenetic reconstruction of gymnosperms (Lee et al., 2011; Yang et al., 2012; Xi et al., 2013), but most of previous studies still used cytoplasmic DNA markers and/or nuclear ribosomal DNA (nrDNA). The followings are the main progresses on phylogenetic reconstruction.

3.2.1. The phylogenetic position of Gnetales

In the plant tree of life, Palmer et al. (2004) presented six major unsolved problems, of which the most radical, shocking, and controversial one was the placement of Gnetales, a small and morphologically unique group of gymnosperms. After nine years, the

phylogenetic position of Gnetales still remains enigmatic, although many efforts have been made to resolve this problem.

The Gnetales includes the three isolated genera *Ephedra*, *Gnetum*, and *Welwitschia*. Based on the early cladistic analyses using morphological characters, Gnetales was considered to be the sister of angiosperms (the anthophyte hypothesis, Fig. 3A) (Doyle and Donoghue, 1986; Rothwell and Serber, 1994; reviewed by Doyle, 1998). The anthophyte hypothesis seemed to be supported by some morphological, anatomical and reproductive characters, such as net-veined leaves, vessels in the wood, double fertilization and the simple, unisexual, flower-like structures, but later studies revealed that these character similarities between Gnetales and angiosperms are not really homologous or due to parallel evolution (Winter et al., 1999; Donoghue and Doyle, 2000; Mundry and Stuetzel, 2004). Moreover, this hypothesis has been questioned by most molecular studies (Supplementary material S1) (e.g., Magallón and Sanderson, 2002; Soltis et al., 2002; Burleigh and Mathews, 2004, 2007a; Ran et al., 2010; Zhong et al., 2010; Lee et al., 2011; Wu et al., 2011; Xi et al., 2013). Although angiosperms and Gnetales were grouped together in several studies using ribosomal DNA or the third codon of the *PHY* gene, the results were very weakly supported statistically (Stefanović et al., 1998; Rydin et al., 2002; Schmidt and Schneider-Poetsch, 2002), and could be caused by some limitations of ribosomal DNA such as many paralogous copies (Zhang et al., 2012) or the substitutional saturation in *PHY* (Xia et al., 2003). Nevertheless, it remains interesting that the anthophyte hypothesis is supported by the slowly-evolving ribosomal DNA. In particular, Burleigh and Mathews (2007b) found bias against recovering the anthophyte hypothesis in the molecular data.

Some molecular studies supported Gnetales as sister to the other seed plants (Supplementary material S1), especially when some chloroplast DNA (cpDNA) sequences (all or the third codon positions) were used and the maximum parsimony (MP) method was used in phylogenetic reconstruction (e.g., Sanderson et al., 2000; Rydin et al., 2002; Soltis et al., 2002; Rai et al., 2003, 2008). This topology was known as the Gnetales-sister hypothesis by Burleigh and Mathews (2004) or the gnetales-sister II hypothesis by Braukmann et al. (2009). For clarity, we suggest that it be called “the gnetales-other seed plants hypothesis” (Fig. 3B). Moreover, some studies indicated Gnetales as sister to the rest of gymnosperms by analyzing cytoplasmic or nuclear genes (Hasebe et al., 1992; Fröhlich and Estabrook, 2000; Schmidt and Schneider-Poetsch, 2002; Burleigh and Mathews, 2004; also see Supplementary material S1). It was referred as the Gnetales-sister I hypothesis (Braukmann et al., 2009), and is hereafter called “the gnetales-other gymnosperms hypothesis” (Fig. 3C). This hypothesis was also strongly supported by a couple of recent phylogenomic studies that used combined sequences mostly from expressed sequenced tags (ESTs) (Cibrián-Jaramillo et al., 2010; Lee et al., 2011). Currently, neither the gnetales-other seed plants hypothesis nor the gnetales-other gymnosperms hypothesis is widely accepted. One of the most important reasons is that the MP method is more easily affected by long branch attraction (LBA) than maximum likelihood (ML) and Bayesian inference (BI) (Felsenstein, 1978; Hendy and Penny, 1989; Huelsenbeck, 1995). Another reason is that the datasets used in the phylogenomic analyses contained too many missing data, which might perturb phylogenetic inference (Roure et al., 2013). However, while recognizing the shortcomings in data analysis, we should keep in mind that these factors are also problems

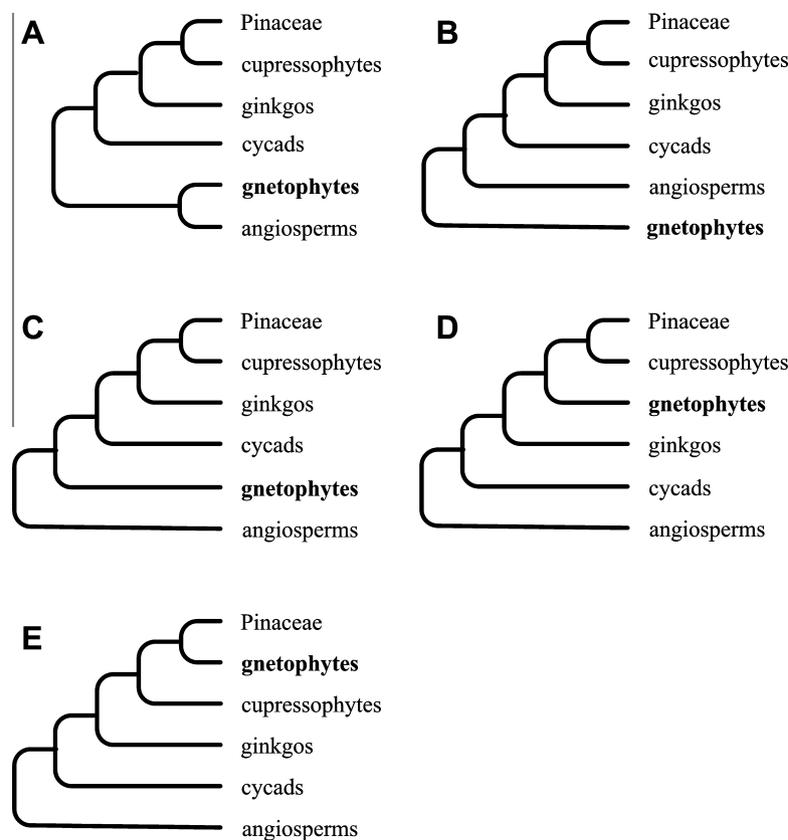


Fig. 3. Five main hypotheses of the phylogenetic position of Gnetales. (A) the anthophyte hypothesis; (B) the gnetales-other seed plants hypothesis; (C) the gnetales-other gymnosperms hypothesis; (D) the Gnetifer hypothesis; and (E) the Gnepine hypothesis.

underlying the phylogenetic inference of seed plants, a group experiencing a long evolutionary history and many extinction events.

Most molecular phylogenetic studies indicate that Gnetales has a close relationship with conifers. Some of them support the Gnepine hypothesis, i.e., Gnetales sister to conifers (Fig. 3D), based mainly on analyses of nrDNA (Chaw et al., 1997; Soltis et al., 1999) and the mitochondrial *rps3* gene (Ran et al., 2010). However, the Gnepine hypothesis, i.e., Gnetales sister to Pinaceae (Fig. 3E, Supplementary material S1), is supported by many more studies after eliminating bias in data analyses (Bowe et al., 2000; Chaw et al., 2000; Gugerli et al., 2001; Magallón and Sanderson, 2002; Burleigh and Mathews, 2004; Hajibabaei et al., 2006; Zhong et al., 2010; Wu et al., 2011; Burleigh et al., 2012; Xi et al., 2013), although some results are locus-dependent. Moreover, the Gnepine hypothesis is supported by the loss of all *ndh* genes, *rpl16* and two introns of *clpP* in the chloroplast genomes shared between Gnetales and Pinaceae (Tatsuya et al., 1994; Braukmann et al., 2009; Wu et al., 2009). Interestingly, Zhong et al. (2010) obtained a tree topology supporting a sister relationship between Gnetales and cupressophytes (the Gnecup hypothesis) when all chloroplast genes were used (Supplementary material S1). However, when they discarded some fast-evolving genes and three genes with many parallel amino acid substitutions between Gnetales and cupressophytes, the topology changed to support the Gnepine hypothesis. Although the nearly complete loss of one of the inverted repeats (IRs) in all conifer chloroplast genomes seemed to support the monophyly of conifers (Raubeson and Jansen, 1992; Tsudzuki et al., 1992; Hirao et al., 2008; McCoy et al., 2008), this structural mutation occurred in different IRs of Pinaceae and Cupressophytes, and thus could not be used as a homologous character or a synapomorphy (Wu et al., 2011; Wu and Chaw, 2013).

Recent advances in genome sequencing technologies, especially next generation sequencing, have facilitated rapid sequencing of a complete genome, transcriptome or large number of cDNA sequences from non-model organisms, providing a massive amount of information for phylogenetic analyses. Phylogenomic analyses have been widely used to reconstruct the tree of life, including to resolve the phylogenetic position of Gnetales (de la Torre-Barcelona et al., 2009; Cibrián-Jaramillo et al., 2010; Zhong et al., 2010; Lee et al., 2011; Wu et al., 2011, 2013; Burleigh et al., 2012). It is interesting that most studies based on concatenated protein-coding nuclear genes (ESTs) support the gnetales-other gymnosperms hypothesis (de la Torre-Barcelona et al., 2009; Cibrián-Jaramillo et al., 2010; Lee et al., 2011) whereas the studies based on chloroplast genes support the Gnepine hypothesis after data treatment (Zhong et al., 2010; Wu et al., 2011). This incongruence could have two explanations. First, for gymnosperms, the nuclear genome sequence is currently unavailable, and the available EST sequences cannot well represent the nuclear genome except in some species of Pinaceae (Mackay et al., 2012). Therefore, it is difficult to obtain enough real orthologous nuclear genes from all main lineages of gymnosperms for phylogenetic reconstruction. For example, Lee et al. (2011) analyzed 12,469,970 amino acid sites from 150 species across land plants. However, according to our reanalysis of the data, none of the sites was left after removing the poorly aligned positions in the alignment using Gblocks (Talavera and Castresana, 2007). Moreover, only 3688 and 1864 sites can be well aligned between Gnetophytes and angiosperms, and between Gnetophytes and Pinaceae, respectively, and in particular the two alignments have only about 130 overlapped sites (our unpublished data). Second, the incongruence might have resulted from systematic errors (Jeffroy et al., 2006), although phylogenomic analyses could sometimes improve the resolution of trees and solve difficult phylogenetic questions.

Systematic errors cannot be removed by increasing data, and, on the contrary, may grow with increased size of datasets (Rodríguez-Ezpeleta et al., 2005), owing to sequence composition biases among lineages and sequence heterotachy (Wu et al., 2011). Burleigh and Mathews (2004) found that both MP and ML trees supported the Gnepine hypothesis when fast-evolving positions were removed from a 13-locus concatenated seed plant dataset, and Zhong et al. (2010) obtained a sister relationship between Gnetales and Pinaceae after excluding fast-evolving genes and parallel sites. In addition, Wu et al. (2011) got congruent and robust tree topologies supporting the gnetophytes-Pinaceae clade as sister of cupressophytes when the removal of highly heterotachous genes alleviated the artifact of LBA. Moreover, different rootings will directly influence the topology of phylogenetic trees, especially for the inference of relationships among major lineages of seed plants. For instance, Donoghue and Doyle (2000) investigated the effect of alternative rootings on inferred relationships of Gnetales, angiosperms and conifers, and discussed the anthophyte hypothesis. Mathews et al. (2010) explored this issue using a duplicate gene rooting in analyses of phytochrome amino acids of seed plants, and yielded trees that unite cycads and angiosperms in a clade.

3.2.2. Phylogenetic studies of other gymnosperms

Besides the great effort in resolving the phylogenetic position of Gnetales, many phylogenetic studies have been conducted on other gymnosperms. For instance, the monotypic genus *Ginkgo* is the sole survivor of ginkgos that originated at least 270 Mya, and its systematic position has been controversial for a long time (see review by Wu et al., 2013). It was placed in the coniferophyte clade *sensu* Chamberlain (1935), comprising conifers, cordaites, ginkgophytes, and gnetophytes. Some studies suggested that *Ginkgo* is closer to conifers than cycads based on comparative development of the spermatozoids (Norstog et al., 2004) or intermediate between these two lineages based on embryogenesis (Wang et al., 2011). However, most molecular phylogenetic studies based on single or a few genes support *Ginkgo* as sister to a clade comprising conifers and gnetophytes (e.g., Chaw et al., 2000; Hajibabaei et al., 2006; Mathews, 2009; Ran et al., 2010). It is particularly interesting that most recent phylogenomic analyses support a sister relationship between *Ginkgo* and cycads (e.g., Cibrián-Jaramillo et al., 2010; Finet et al., 2010; Wu et al., 2013; Xi et al., 2013), which is consistent with the morphological characters shared between the two groups, such as an haustorial pollen tube (Friedman, 1993) and multiflagellated sperms (Ikeno and Hirase, 1897; Brenner et al., 2003).

For cycads, molecular phylogenies were previously constructed using two chloroplast genes and nrDNA (Hill et al., 2003), 17 chloroplast genes and associated noncoding regions (Rai et al., 2003; Zgurski et al., 2008), *matK*+ITS/5.8S rDNA (Chaw et al., 2005), *matK*+26S rDNA (Crisp and Cook, 2011), and *PHYP* (Nagalingum et al., 2011), respectively. All of the phylogenies support the basal position of *Cycas* and the division of cycads into two families, which have been recognized in the new book of Osborne et al. (2012). That is, Cycadaceae comprises the single genus *Cycas*, while Zamiaceae includes the rest of nine genera. However, some intergeneric relationships within the family Zamiaceae, especially the phylogenetic positions of *Bowenia*, *Dioon* and *Stangeria*, still remain unresolved, although the genus *Dioon* was placed in a basal position in most phylogenies (Hill et al., 2003; Rai et al., 2003; Chaw et al., 2005; Zgurski et al., 2008; Crisp and Cook, 2011). Very recently, a relatively solid phylogeny of cycads was reconstructed using five single-copy nuclear genes, in which, for Zamiaceae, *Dioon* diverged first, followed by *Bowenia*, and then an encephalartoid clade (*Macrozamia*–*Lepidozamia*–*Encephalartos*) sister to a zamoid clade (Salas-Leiva et al., 2013).

Pinaceae is the largest and economically most important family in conifers (Farjón, 1990). The first molecular phylogeny of all its 11 genera was constructed by Wang et al. (2000), using paternal chloroplast, maternal mitochondrial and biparental low-copy nuclear genes. They found that different gene trees are largely identical in topology except the phylogenetic position of *Cedrus* and an unresolved trichotomy formed by *Cathaya*, *Picea* and *Pinus*, and the three-genome phylogeny supports the division of the pine family into two major groups *Abies–Keteleeria–Nothotsuga–Tsuga–Pseudolarix–(Cedrus)* and *Cathaya–Picea–Pinus–Pseudotsuga–Larix* that correspond well with the distribution (number and position) of resin canals in the central vascular cylinder of the young taproot. Gernandt et al. (2008) investigated the phylogenetic history of Pinaceae using two chloroplast genes (*matK* and *rbcl*) and non-molecular characters, and also found that the position of *Cedrus* is inconsistent among different analyses. Recently, based on comparative chloroplast genomics, Lin et al. (2010) concluded that *Cedrus* was sister to the *Abies–Keteleeria* clade, and *Cathaya* was closer to *Pinus* than to *Picea*. However, it is interesting that *Cathaya* looks like a hybrid between *Picea* and *Pinus* in sequences of many genes (our unpublished data). Actually, *Cathaya* is also morphologically intermediate between the two genera. Like *Picea*, the needles of *Cathaya* do not form bundles and are spirally arranged. In particular, some species of *Picea* also have *Cathaya*-like flattened leaves that may represent a symplesiomorphic character. On the other hand, *Cathaya* is similar to *Pinus* in having axillary seed cones and in reproductive characteristics (reviewed in Wang et al., 1998). Therefore, the relationships among the three genera still need more investigations.

For Conifer II, the interfamilial relationships were consistently revealed by previous molecular phylogenetic studies. That is, the two families Araucariaceae and Podocarpaceae diverged first, followed by Sciadopityaceae, and then Taxaceae–Cephalotaxaceae that is sister to Cupressaceae *s.l.* comprising Taxodiaceae and Cupressaceae *s.s.* (Chaw et al., 1997, 2000; Bowe et al., 2000; Rai et al., 2008; Ran et al., 2010; Crisp and Cook, 2011; Burleigh et al., 2012; Yang et al., 2012). Yang et al. (2012) reconstructed the phylogeny of gymnosperms using two sister single-copy nuclear genes *LFY* and *NLY* that originated from a gene duplication in the common ancestor of seed plants, and further used the two genes, together with chloroplast *matK* and mitochondrial *rps3* genes to reconstruct the phylogeny of Cupressaceae *s.l.* represented by all its 32 genera. The different gene trees generated are topologically highly congruent, supporting the basal position of *Cunninghamia* and the division of Cupressaceae *s.l.* into six or seven subfamilies, as recognized in Gadek et al. (2000) and Farjón (2005), respectively. However, there are still some topological conflicts among different trees that need to be resolved, such as the positions of *Papuacedrus* and *Tetraclinis*, and the relationship between *Fokienia–Chamaecyparis* and *Thuja–Thujopsis*. The conflicts could be attributed to insufficient resolution of the molecular markers or historical hybridization. For example, the inconsistent relationships among the three genera *Sequoia*, *Sequoiadendron* and *Metasequoia* revealed in different gene trees suggest an allopolyploid origin for the hexaploid *Sequoia* by hybridization between *Metasequoia* and *Sequoiadendron* or an extinct taxodiaceae plant. This inference is also supported by the reticulation among the three genera shown in the network analysis of the *LFY* and *NLY* genes (Yang et al., 2012). The published molecular phylogenies of other families of Conifer II include Podocarpaceae (Kelch, 1998; Conran et al., 2000; Knopf et al., 2012), Araucariaceae (Setoguchi et al., 1998; Liu et al., 2009), and Taxaceae (Cheng et al., 2000; Wang and Shu, 2000; Hao et al., 2008). All molecular studies support *Phyllocladus* as a genus of Podocarpaceae rather than as an independent family (Phyllocladaceae), and some intergeneric relationships within Podocarpaceae, the second largest family of conifers

comprising 19 genera, have not been well resolved (Kelch, 1998; Conran et al., 2000; Knopf et al., 2012).

It is also exciting that phylogenies of some genera of gymnosperms have been reconstructed based on extensive species sampling and multiple gene markers, such as *Picea* (Ran et al., 2006), *Pinus* (Gernandt et al., 2005; Parks et al., 2009, 2012), *Cedrus* (Qiao et al., 2007), *Larix* (Wei and Wang, 2004), *Tsuga* (Havill et al., 2008), *Pseudotsuga* (Wei et al., 2010), *Juniperus* (Mao et al., 2010), *Thuja* (Peng and Wang, 2008), *Ephedra* (Rydin and Korall, 2009) and *Podocarpus* (Knopf et al., 2012). For instance, Wei et al. (2010) unveiled the interspecific relationships within the genus *Pseudotsuga* and found that the Taiwanese species *P. wilsoniana* might have originated by hybridization between two lineages from mainland China, using five cpDNA and two mtDNA fragments as well as the nuclear gene *LFY*. However, surprisingly, most species of the large genera such as *Picea*, *Pinus*, *Podocarpus* and *Cycas* originated from recent radiation. For example, Ran et al. (2006) investigated the evolutionary history of 33 species of *Picea* using two chloroplast and one mitochondrial genes, and fossil evidence. They found that this genus experienced at least two radiative speciation events that occurred in northeastern Asia and southwestern China, respectively. Therefore, it is still very tough to resolve evolutionary relationships of closely related species of gymnosperms due to radiative speciation and frequent interspecific hybridization.

3.3. Molecular and genome evolution

3.3.1. Chromosomal and genome size variation

Gymnosperms are probably the best studied group of land plants by far with regard to chromosome number and genome size (Fig. 4). Among the 83 genera of gymnosperms, the genome size has been estimated for 344 species representing all genera except *Nothotsuga*, *Parasitaxus*, *Pilgerodendron* and *Sundacarpus*, while the chromosome number has been reported from a number of species covering all genera except *Austrotaxus* and *Sundacarpus* (Zonneveld, 2012a,b; Leitch and Leitch, 2013; Murray, 2013). Considering that several reviews have discussed the genome size variation (e.g., Morgante and De Pauli, 2011; Leitch and Leitch, 2013), or the karyotype variation and evolution in gymnosperms (e.g., Nkongolo

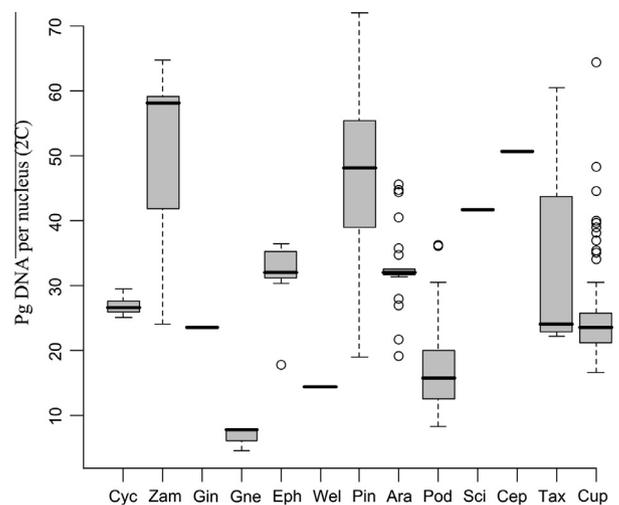


Fig. 4. Variation of genome size in 13 gymnosperm families. Cyc, Cycadaceae; Zam, Zamiaceae; Gin, Ginkgoaceae; Gne, Gnetaceae; Eph, Ephedraceae; Wel, Welwitschiaceae; Pin, Pinaceae; Ara, Araucariaceae; Pod, Podocarpaceae; Sci, Sciadopityaceae; Cep, Cephalotaxaceae; Tax, Taxaceae; Cup, Cupressaceae. The data were calculated based on information from the Plant DNA C-values Database (<http://data.kew.org/cvalues>) before May 4, 2013.

and Mehes-Smith, 2012; Murray, 2013), here we only summarize the main characteristics of gymnosperm genomes as follows.

One remarkable feature is the narrow variation range of chromosome numbers, from $2n = 14$ to 66 in wild plants. In Pinaceae, all studied species except *Pseudolarix amabilis* ($2n = 44$) and *Pseudotsuga menziesii* ($2n = 26$) have the same chromosome number of $2n = 2x = 24$. In Cupressaceae s.l., all 30 genera have the same basic chromosome number of $x = 11$, and all studied species show $2n = 2x = 22$ except $2n = 6x = 66$ in *Sequoia semperviens* and $2n = 4x = 44$ in *Juniperus chinensis* 'Pfitzeriana' and *Fitzroya cupressoides*. The basic chromosome number shows the widest variation in Podocarpaceae, but only ranges from $x = 9$ to $x = 19$ (Table 1). The second remarkable feature is that all congeneric species have the same basic chromosome number except *Halocarpus*, *Lepidothamnus*, *Nageia*, *Podocarpus* and *Prumnopitys* in Podocarpaceae, *Pseudotsuga* in Pinaceae, *Torreya* in Taxaceae, and *Zamia* in Zamiaceae. Moreover, intraspecific variation of basic chromosome number was very rarely reported, such as in *Zamia loddigesii* (Vovides and Olivares, 1996). According to the study of Zhou et al. (2000), $x = 7$ and $x = 20$ reported in *Amentotaxus argotaenia* (Chuang and Hu, 1963; Guan et al., 1993) could be wrong counts. The third feature is the remarkable karyotype conservation across species and genera (e.g., Hizume et al., 2002; Shibata and Hizume 2008; Nkongolo and Mehes-Smith, 2012). For the fertile interspecific hybrids in *Larix*, *Pinus*, *Dacrydium* and *Podocarpus*, no chromosomal translocations were observed (Sax, 1932, 1960; Saylor and Smith, 1966; Quinn and Rattenbury, 1972; Wardle, 1972; Williams et al., 2002; Eckenwalder, 2009). Comparative genetic mapping also revealed a remarkable interspecific and intergeneric conservation of gene distribution and order (Krutovsky et al., 2004; Pelgas et al., 2006). The fourth feature is that chromosomes are extremely large. For example, the length of metaphase chromosomes ranges from 6.4 to 16.2 μm in *Pinus*, 5.4 to 14.5 μm in *Picea*, and 4.4 to 11.6 μm in *Larix* (Hizume, 1988). The fifth feature is that genomes of gymnosperms are, on average, larger than those of other land plant groups (Murray, 1998; Leitch et al., 2001; Ahuja, 2005; Ahuja and Neale, 2005). The mean genome size of gymnosperms is $1C = 18.08$ pg, a value much larger than that of angiosperms ($1C = 5.9$ pg) (Leitch and Leitch, 2013). The smallest and largest genomes were found in *Gnetum* ($1C = 2.25$ pg) and *Pinus* ($1C = 36$ pg), respectively (Table 1). The sixth feature is that polyploidy is exceedingly rare in gymnosperms (Ahuja, 2005; Williams, 2009; Fawcett et al., 2013; Murray, 2013), with a frequency of about 4–5% (Khoshoo, 1959; Delevoryas, 1980; Wood et al., 2009). There are only three natural polyploids in conifers, including the hexaploid *Sequoia semperviens* and the two tetraploids *Juniperus chinensis* 'Pfitzeriana' and *Fitzroya cupressoides* (Ahuja, 2005), although recent studies indicate that polyploidy seems to have played a more important role than earlier envisioned in the speciation of conifers (Zonneveld, 2012a). The rest of natural polyploids are all from *Ephedra*, accounting for about 50–65% of the studied species of the genus (Khoshoo, 1959; Huang et al., 2005).

3.3.2. Composition and evolutionary patterns of the three genomes of gymnosperms

3.3.2.1. Nuclear genome. Aided by the development of sequencing technologies, fully sequenced plant genomes are rapidly growing in number. However, none of them is from gymnosperms owing to the extremely large genome sizes in this group. For instance, conifers have a genome size of 8–70 gigabases (Gb) ($2C$), which is among the largest of any nonpolyploid plant species (<http://data.kew.org/cvalues/>). This, together with large amounts of repetitive DNA, has limited efforts to produce a conifer reference genome (Lorenz et al., 2012; Ritland, 2012), and early studies strongly suggest that the genome of conifers is very different from that of angiosperms in composition and structure (see review by

Mackay et al., 2012). Fortunately, a draft assembly of the Norway spruce (*Picea abies*) genome, approximately 19.6 Gb, has been generated very recently based on a hierarchical strategy combining fosmid pools with both haploid and diploid whole genome shotgun (WGS) data, and RNA sequencing (RNA-Seq) data, and it is surprising that this large genome harbors only about 28,354 genes (Nystedt et al., 2013). At the same time, Birol et al. (2013) reported an assembling of the 20.8 Gb white spruce (*Picea glauca*) genome based on the WGS data. Moreover, the genome sequence of the loblolly pine (*Pinus taeda*, around 22 Gb ($1C$)), is expected to be reported soon (Sederoff, 2013). The available data indicate that different conifers (Norway spruce, white spruce, Sitka spruce, and loblolly pine) have very similar numbers of unigenes, less than 30,000 as in many angiosperms, including *Arabidopsis thaliana* (Ritland, 2012). This finding clearly does not support the previous hypotheses that there are more than 225,000 genes in the genome of Pinaceae (Kinlaw and Neale, 1997; Rabinowicz et al., 2005). The overestimation by early researchers could be caused by including degenerate retroelements (Morgante and De Pauli, 2011), and large numbers of pseudogenes (Rabinowicz et al., 2005; Garcia-Gil, 2008; Kovach et al., 2010; Rigault et al., 2011; Nystedt et al., 2013).

Why do gymnosperms have a so large genome? It is well known that polyploidy or chromosome duplication is an important mechanism for generating large genomes in angiosperms. However, the gymnosperm genomes show no evidence of recent whole-genome or chromosome duplication (Ohri and Khoshoo, 1986; Kovach et al., 2010; Nystedt et al., 2013), and few polyploid species are found in the gymnosperm lineages except *Ephedra* (Khoshoo, 1959, 1961; Ahuja, 2005). Moreover, the basic chromosome number variation in diploid gymnosperms is very likely resulted from chromosome fission or fusion, such as in Douglas fir (Khoshoo, 1961; Fuchs et al., 1995; Pelgas et al., 2006; Pavy et al., 2012). Although a whole-genome duplication was inferred to have occurred in the common ancestor of seed plants around 350 Mya (Jiao et al., 2011; Nystedt et al., 2013), it should not have only contributed to the increase of genome size in gymnosperms.

The large and complex genome of gymnosperms could have developed by expansion of retrotransposons (Morse et al., 2009; Nystedt et al., 2013), which can increase the genome size in very short time by dispersing to new locations in a copy-and-paste fashion through an RNA intermediate. Based on reassociation kinetics, Rake et al. (1980) found that 75% of a conifer genome could be repetitive elements, and suggested that dramatic amplification of noncoding DNA might have contributed to the origin of large genome sizes in gymnosperms. Recently, the large amount of repetitive elements, especially the LTR (long terminal repeat) retrotransposons such as *Ty1-copia*, *Ty3-gypsy* and *Gymny*, have been recognized as the primary factor causing inflation of gymnosperm genomes (Friesen et al., 2001; Morse et al., 2009; Plomion et al., 2011; Nystedt et al., 2013). The *gypsy* and *copia* elements in Pinaceae are distributed at the chromosome ends or are associated with 18S rDNA and centromeric regions (Friesen et al., 2001; Morse et al., 2009; Grover and Wendel, 2010; Plomion et al., 2011), which is fundamentally different from angiosperms (Leitch and Leitch, 2013). It is particularly interesting that the transposable element diversity is shared among extant conifers. In addition, it has been confirmed that 24-nucleotide short RNAs (sRNAs), a class of sRNA that can silence transposable elements by the establishment of DNA methylation, are present in gymnosperms, but they are highly specific to reproductive tissues and at substantially lower levels than in angiosperms (Nystedt et al., 2013). Furthermore, it has been found that pseudogenes occur much more frequently than functional protein-coding genes in the genome of gymnosperms (Kovach et al., 2010; Nystedt et al., 2013).

Many species of gymnosperms, especially conifers, have wide distributions. Long generation time and large effective population

size are often hypothesized to be the causes for low synonymous polymorphism in conifers (Savolainen and Pyhäjärvi, 2007). Buschiazzi et al. (2012) reported a slower evolutionary rate and a higher d_N/d_S value in conifers than in angiosperms, and found a d_N/d_S ratio >1 in many pair-wise comparisons of orthologs. However, Chen et al. (2012) found that the mean d_N/d_S value of gymnosperms is similar to that of angiosperms. Although they got a lower synonymous substitution rate per year in gymnosperms, most of this difference disappeared when generation time was taken into account. It means that evolutionary constraints could be similar between gymnosperms and angiosperms.

3.3.2.2. Chloroplast and mitochondrial genomes. Chloroplast and mitochondrial genomes are predominantly maternally inherited in angiosperms (Birky, 2008), except for a few species such as *Silene vulgaris* (McCauley et al., 2005; reviewed by Crosby and Smith, 2012). However, the inheritance patterns of organelle genomes are quite diverse in gymnosperms. The chloroplast genomes are predominantly paternally inherited in conifers but maternally inherited in other gymnosperms, while the mitochondrial genomes are maternally inherited in Pinaceae, Taxaceae, Cycadales and Gnetales but paternally inherited in Araucariaceae, Cupressaceae *s.l.* and Podocarpaceae (reviewed by Mogensen, 1996).

The chloroplast genome of most land plants consists of four parts, including a large single copy (LSC) region, a small single copy (SSC) region, and two copies of large inverted repeats (IRs) that may be important for maintaining conserved gene orders (Palmer and Thompson, 1982). Interestingly, different contraction of IRs has been documented in gymnosperms (Raubeson and Jansen, 1992; Wu et al., 2011; Zhou et al., 2012). *Ginkgo biloba* has slightly reduced IRs (Zhou et al., 2012). The extremely reduced IRs are found in conifers (Lin et al., 2010; Wu et al., 2011), and the whole chloroplast genome data suggest that Pinaceae and Conifer II have lost different IR copies (Wu et al., 2011; Wu and Chaw, 2013). However, gnetophytes have uncontracted IRs, although they have the smallest compact chloroplast genomes in gymnosperms (Wu et al., 2009). The occurrence of numerous structural rearrangements in the chloroplast genomes of conifers strongly supports the hypothesis that the reduction or deletion of IR could have made the genome less stable (Hirao et al., 2008). It is particularly interesting that the paternally inherited chloroplast genomes are significantly smaller than the maternally inherited ones except in gnetophytes (Crosby and Smith, 2012). Therefore, it is likely that the contraction of IRs is correlated with the shift to paternal inheritance of chloroplast in conifers. Furthermore, the *ndh* genes have been lost from the chloroplast genomes of Pinaceae and gnetophytes, but still remain in conifer II (Braukmann et al., 2009; Wu et al., 2009, 2011).

Compared to dozens of chloroplast genome sequences that are available, the mitochondrial genome has been completely sequenced only for one gymnosperm, i.e., *Cycas taitungensis*, with a size of 414.9 kb that is similar to the situation in angiosperms (Chaw et al., 2008). The mitochondrial genome of cycad shows many features different from that of angiosperms, including a much higher proportion of tandem repeat ($\approx 5\%$), a particularly lower A + T content (53.1%), more RNA editing sites (1084) and cis-spliced introns, and fewer gene losses (Chaw et al., 2008). However, it is still unclear whether these features are shared among gymnosperms. For example, Regina et al. (2005) reported a novel additional group II intron from the mitochondrial *rps3* gene of *Cycas revoluta*, and considered this intron as a signature of gymnosperms. However, Ran et al. (2010) and Regina and Quagliariello (2010) found that this intron was lost in different lineages of gymnosperms. In particular, Ran et al. (2010) found a dramatically high variation in both length and sequence of a *rps3* exon in Conifer II that could be attributed to the intron loss. In addition, Jaramillo-

Correa et al. (2013) reported an ancient microsatellite hotspot in the conifer mitochondrial genome. They found that sequence tandem repeats (STRs) can accumulate and be retained in short regions of the mtDNA genome over long periods of time and between distantly related lineages. It is unexpected that the draft mitochondrial genome of Norway spruce is larger than 4 Mb, being among the largest reported for plants and rich in short open-reading frames (Nystedt et al., 2013). Therefore, more mitochondrial genomes should be sequenced for understanding the evolutionary pattern of mtDNA of gymnosperms.

4. Biogeography of gymnosperms

Historical biogeographic reconstruction relies largely on phylogenetic studies (Crisp et al., 2011; Yang et al., 2012). Extinction is a key determinant of observed biogeographic patterns, but was often considered intractable and ignored (Lamm and Redelings, 2009). In recent years, developing synergies between phylogenetics, biogeography, ecology, molecular dating and palaeontology have provided novel data and opportunities for testing biogeographic hypotheses (Crisp et al., 2011; Ronquist and Sanmartín, 2011; Gillespie et al., 2012; Mao et al., 2012; Yang et al., 2012; Wen et al., 2013). The biogeography of gymnosperms has further been studied with the help of new techniques and methods as well as abundant microfossils and megafossils, shedding new light on the origin and development of intercontinental disjunctions.

4.1. Disjunctive distribution in the two hemispheres

Some families and genera of gymnosperms are distributed in both hemispheres (Table 1), which may provide an opportunity to unveil the break-up history of Pangea. However, based on molecular clock analysis, the crown ages of these genera, such as *Cycas*, *Ephedra*, *Gnetum* and *Podocarpus*, could only be dated to the Tertiary (mostly in or after the Oligocene), a time much later than the separation between Gondwana and Laurasia (Won and Renner, 2006; Ickert-Bond et al., 2009; Nagalingum et al., 2011; Leslie et al., 2012). Therefore, their current distributions can only be explained by long-distance dispersal among continents. It is very likely that the genus *Ephedra* originated in Eurasia, then dispersed into North America in the Oligocene by the Bering Land Bridge, and further into South America approximately 25 Mya, well before the closure of the Panamanian Isthmus (Ickert-Bond et al., 2009). In contrast, the genus *Gnetum* possibly originated in South America, then dispersed by seawater to West Africa in the Oligocene/Miocene, and finally to tropical and subtropical Asia, because its seeds have special structures (Won and Renner, 2006). Although a member of the family Taxaceae, *Austrotaxus*, is endemic to New Caledonia, this monotypic genus separated from its sisters in the Northern Hemisphere after the middle Cretaceous (Leslie et al., 2012).

The coniferous family Cupressaceae *s.l.* does provide an excellent example for studying the break-up history of Pangea. This family occurs in all continents except Antarctica (Farjón, 2005), and diversified into seven subfamilies during the Late Triassic and Jurassic (Gadek et al., 2000; Mao et al., 2012; Yang et al., 2012), predating or coinciding with the separation of Gondwana and Laurasia. Two of the seven subfamilies, Athrotaxidoideae and Callitroideae, are confined to the Southern Hemisphere. Consistent with fossil evidence, the divergence between the Gondwanan Callitroideae and its sister subfamily in the Northern Hemisphere, Cupressoideae, could be dated back to the Jurassic (around 150–180 Mya) by molecular clock analysis, providing strong evidence for the vicariance between the two subfamilies by the spilt of Laurasia and Gondwana (Li and Yang, 2002; Mao et al., 2012; Yang

et al., 2012). However, a relatively younger divergence between Callitroideae and Cupressoideae (in the Cretaceous) was reported in Crisp and Cook (2011) and Leslie et al. (2012). The subfamily Athrotaxidoideae includes only one genus *Athrotaxis*, which comprises three extant species in Tasmania, with the most recent common ancestor dated to the Tertiary (Leslie et al., 2012; Mao et al., 2012). However, fossils of this group have been reported from North and South America in the lower Cretaceous sediments (Miller and LaPasha, 1983; Del Fueyo et al., 2008), with an age younger than the break-up of Pangea. Hence, the present distribution of *Athrotaxis* could be the result of a southward expansion from North America (Mao et al., 2012).

4.2. Disjunctive distribution in the Southern Hemisphere

Biogeographical patterns in the Southern Hemisphere have drawn great interest from biologists and geologists (Sanmartin and Ronquist, 2004; Knapp et al., 2005; Barker et al., 2007; Upchurch, 2008; Yang et al., 2012), and were usually explained by vicariance scenarios. However, recent studies suggest that these patterns have also been shaped by dispersal events (McLoughlin, 2001; Givnish and Renner, 2004; Sanmartin and Ronquist, 2004; Crisp et al., 2011). There are three lineages of conifers mainly distributed in the Southern Hemisphere, the subfamily Callitroideae of Cupressaceae, Araucariaceae, and Podocarpaceae (Leslie et al., 2012).

The Callitroideae is an ideal taxon for studying Southern Hemisphere biogeography. It comprises 10 genera, each of which, except *Callitris* and *Libocedrus*, is endemic to a single continent or even a single island. These genera have relatively ancient origins (mostly in the Cretaceous) and some reliable fossils (Mao et al., 2012; Yang et al., 2012), although the divergence time estimates by Leslie et al. (2012) are generally younger. Using multiple calibrations, Yang et al. (2012) performed a relaxed molecular clock analysis for Cupressaceae *s.l.* based on nuclear, chloroplast and combined gene datasets, respectively, and reconstructed ancestral distributions for the Callitroideae genera. They found that the separation of East and West Gondwana at 165–130 Mya led to the divergence between the two clades *Callitris*–*Actinostrobus*–*Neocallitropsis* and *Widdringtonia*–*Fitzroya*, and the split between *Widdringtonia* and *Fitzroya*–*Diselma* that occurred at least 95 Mya as suggested by fossil evidence is generally consistent with the final separation of Africa from South America around 105 Mya (McLoughlin, 2001). Therefore, vicariance could be mainly responsible for the current distribution pattern of the Gondwanan Callitroideae. However, the other two lineages, Araucariaceae and Podocarpaceae, have different biogeographic histories.

The Araucariaceae comprises three genera, *Agathis*, *Araucaria* and *Wollemia*. Except two species of *Araucaria* (*A. araucana* and *A. angustifolia*) in South America, all the other species of the family are distributed in Australia and its adjacent regions (Farjón, 2010). Phylogenetic analysis and molecular dating indicate that the two South American species diverged from their sisters *Araucaria bidwillii* in Australia and *A. hunsteinii* in New Guinea at Oligocene or Miocene (Setoguchi et al., 1998; Leslie et al., 2012), implying a dispersal event from Australia to South America because South America was connected with Australia through Antarctica during 52–35 Mya (McLoughlin, 2001; Sanmartin and Ronquist, 2004). Based on the phylogeny and divergence times of Podocarpaceae (Knopf et al., 2012; Leslie et al., 2012), the dispersal from Australia to South America through Antarctica could also have occurred in three genera of the family, i.e., *Lepidothamnus*, *Podocarpus*, and *Prumnopitys*. However, *Saxegothaea*, a monotypic genus of Podocarpaceae in South America, originated around late Jurassic to early Cretaceous (Leslie et al., 2012), and thus might represent a remnant of ancient lineages. Although *Podocarpus* is

widely distributed in Asia, South America, Africa, and Australia and its neighboring islands, this genus and its main lineages diverged in the Paleogene and Neogene (Biffin et al., 2011; Leslie et al., 2012). Therefore, its present wide distribution could also be attributed to long-distance dispersal.

4.3. Disjunctive distribution in the Northern Hemisphere

The biogeographic patterns in the Northern Hemisphere are complex due to effects of both vicariance and frequent dispersal by the Bering Land Bridge (BLB) and the North Atlantic Land Bridge (NALB), and the Eastern Asian–North American disjunction is particularly interesting (Ian Milne, 2006; Wen, 1999). Previous studies suggest that the Asia-to-New World migration was common for angiosperm lineages (Wen, 1999; Wen et al., 2010). However, current data imply that more gymnosperm lineages could have experienced a history of the New World-to-Old World migration (Ran et al., 2006; Wei et al., 2010; Wen et al., 2010).

All the thirteen gymnospermous genera that are disjunctly distributed between North America and Eurasia or Asia (Table 1, *Cupressus* is not included) dispersed by the BLB one or more times during their evolutionary history (Fig. 5). Seven of them likely originated in North America and migrated into the Old World, including *Abies* (Xiang et al., 2009), *Chamaecyparis* (Wang et al., 2003), *Larix* (Wei and Wang, 2003), *Picea* (Ran et al., 2006; Klymiuk and Stockey, 2012), *Pseudotsuga* (Wei et al., 2010), *Taxus* (Li et al., 2001; Hao et al., 2008), and *Thuja* (Peng and Wang, 2008). In contrast, an Eurasian origin and dispersal to North America were inferred for only four genera, including *Ephedra* (Ickert-Bond et al., 2009), *Juniperus* (Mao et al., 2010), *Pinus* (Eckert and Hall, 2006; Ryberg et al., 2012), and *Torreya* (Hao et al., 2008). The biogeographic history of the rest two genera, *Calocedrus* and *Tsuga*, has not been well resolved, although vicariance between East Asia and North America was suggested for them (Chen et al., 2009; Havill et al., 2008).

According to the molecular phylogeny of extant species, *Tsuga* could have a North American origin. However, likelihood-based biogeographic inference using Lagrange, which incorporated phylogeny, divergence times, fossil data, and rates of lineage dispersal and extinction, inferred an Eocene basal crown group diversification and an initial widespread circumpolar distribution with subsequent vicariance and extinction events for the genus (Havill et al., 2008). Lockwood et al. (2013) reported a molecular phylogeny of *Picea*, and suggested an Asian origin for the genus based mainly on signal from large motifs in mitochondrial gene introns. It should be reminded that this kind of introns often shows intra-specific variation in both sequence and structure, and thus is not suitable for the reconstruction of interspecific relationships if only a couple of individuals are sampled for each species. Based on our recent studies (unpublished data), all cytoplasmic and nuclear gene trees still support the basal position of the western North American *Picea breweriana* and a North American origin of the genus. Our results are also consistent with the fossil record (LePage, 2001), especially the earliest fossil of *Picea* from northern Vancouver Island dated to 136 Ma (Klymiuk and Stockey, 2012).

Some lineages of *Juniperus* and *Pinus* possibly migrated from Eurasia to North America by the NALB (Eckert and Hall, 2006; Mao et al., 2010). Moreover, multiple dispersal events between Eurasia and North America could have occurred in *Pinus* (Eckert and Hall, 2006), *Picea* (Ran et al., 2006), and *Juniperus* (Mao et al., 2010). Although some genera that originated in Eurasia, such as *Juniperus* and *Ephedra*, have a distribution in high latitudes and alpine regions, their ancestors possibly occurred in warmer habitats and might represent remnants of the Madrean–Tethyan vegetation belts (Eckert and Hall, 2006; Ickert-Bond et al., 2009; Mao et al., 2010).

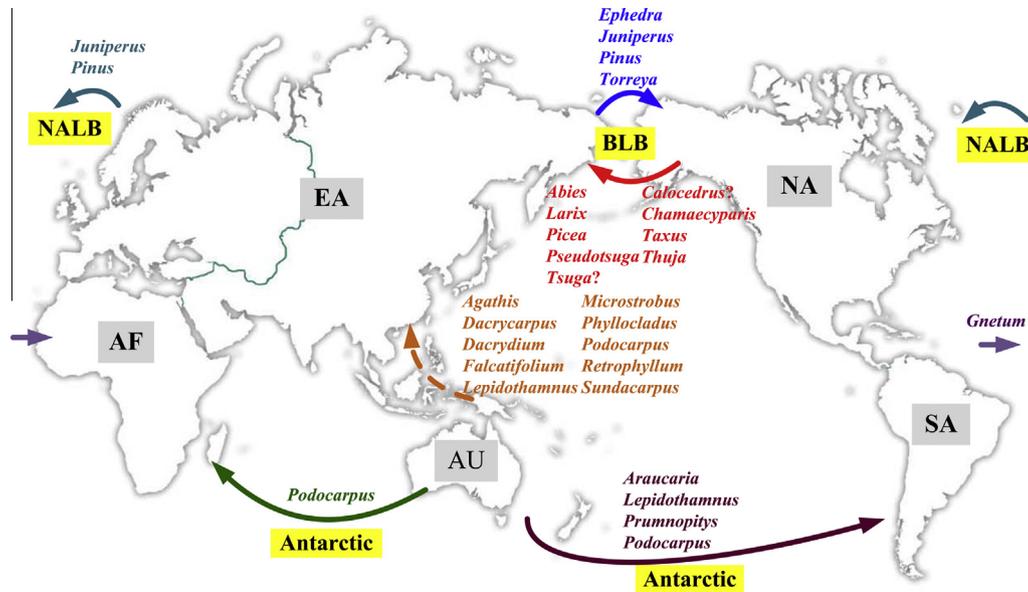


Fig. 5. Inferred dispersal routes of some gymnospermous genera with intercontinental distributions. BLB, Bering Land Bridge; NALB, North Atlantic Land Bridge; AF, Africa; AU, Australia (continent); EA, Eurasia; NA, North America; SA, South America. The question marks indicate migration routes of the taxa have not been resolved, while the dash arrow indicates a possible route of migration.

4.4. Biogeographic difference between the two hemispheres

As summarized above, both vicariance and dispersal events played important roles in shaping the present distribution patterns of the Southern Hemisphere conifers while dispersal contributed much more than vicariance in developing intercontinental disjunctions of gymnosperms in the Northern Hemisphere. The direction of historical intercontinental dispersal is also different between the Southern and Northern Hemisphere gymnosperms. Simply, dispersal in the Southern Hemisphere shows an “out-of-Australia” mode, but that in the Northern Hemisphere can be generalized as “bidirectional”, i.e., “out-of-America” and “out-of-Eurasia”.

For the distribution pattern of gymnosperms, most genera mainly distributed in the Southern Hemisphere differ greatly from the genera in the Northern Hemisphere. A total of 41 and 43 genera are distributed (or mainly distributed) in the Southern Hemisphere and the Northern Hemisphere, respectively, of which 15 occur widely in both Eurasia and North America while only eight live in more than one continents in the Southern Hemisphere (Table 1). That is, the Southern Hemisphere genera are more isolated than the Northern Hemisphere ones. This distribution difference could be caused by the fundamental difference between the two hemispheres in geography. The three main continents, Africa, South America and Australia, separated at least 83 Mya, preventing plant dispersal among the continents in the Southern Hemisphere. Although South America was connected to Australia through Antarctica during 52–35 Mya (McLoughlin, 2001; Sanmartin and Ronquist, 2004), only some cold-tolerant plants could spread by this land bridge (Hill and Brodribb, 1999). Unfortunately, most of the gymnosperm lineages in the Southern Hemisphere prefer relatively wet and warm environments (Hill and Brodribb, 1999; Leslie et al., 2012; Pittermann et al., 2012), and could not disperse by Antarctic. They might tend to expand into some wet and warm places, like the ten genera that have a Southern Hemisphere origin but now occur in the tropic areas of both Australia and Asia (Table 1; and Fig. 5). On the contrary, due to the climatic cooling in the Cenozoic, especially the extensive glaciations, the middle and high-latitude land masses in the Northern Hemisphere shifted to colder, drier, and more seasonal climates from warm subtropical and temperate climates since the Oligocene (Wolfe, 1994; Ivany

et al., 2000; Zachos et al., 2001; Moran et al., 2006; Dupont-Nivet et al., 2007; Zanazzi et al., 2007; Eldrett et al., 2009). Some lineages that better adapted to cooler and/or drier conditions could have replaced older lineages. Furthermore, Beringia was not covered by ice during Quaternary glaciations, and could have played dual roles as both a glacial refugium and a route of colonization of plants across the continents of Eurasia and North America (Hopkins, 1967). Therefore, repeated instances of species migration, and range contraction and expansion responding to glacial cycles could also have contributed to the complex biogeographic patterns of the Northern Hemisphere gymnosperms (Leslie et al., 2012).

Cycads are distributed in most subtropical and tropical regions of the world. However, except that *Cycas*, the biggest genus of cycads, has a relatively wide distribution in Asia, Australia and Africa, all the other genera are confined to a single continent or two continents once connected, such as *Zamia* in the tropical areas of South America and North America. Molecular dating indicates that Cycadaceae and Zamiaceae diverged from each other in the Jurassic and all genera of the two families originated at least before Oligocene (Nagalingum et al., 2011). According to fossil evidence, cycads were diverse in the Mesozoic, but experienced extinctions toward the end of the Mesozoic (Norstog and Nicholls, 1997; Hermsen et al., 2006; Taylor et al., 2009). This group underwent a nearly synchronous global species rediversification that began in the late Miocene, followed by a slowdown toward the recent (Nagalingum et al., 2011). During the secondary species diversification, why did the distribution of the genera not expand largely? It would be interest to investigate the speciation of cycads, which may provide answers for why all species of this group are endemic and endangered.

5. Concluding remarks

In the past two decades, there have been exciting advances in our understanding of gymnosperm evolution and biogeography, including reconstructed phylogenetic relationships at different taxonomic levels based on comprehensive evidence, patterns of species diversification such as recent radiation in most lineages, relative roles of vicariance and dispersal in development of

intercontinental disjunctions, modes of molecular evolution in different genomes and groups, and mechanisms underlying the formation of large nuclear genome. In particular, most evidence supports the Gnepine hypothesis; Multiple coniferous genera such as *Picea* very likely originated in North America and migrated into the Old World, further indicating that the center of diversity is not necessarily the place of origin; Conifers have very similar numbers of unigenes as angiosperms in the nuclear genome, and expansion of retrotransposons is responsible for their large and complex genomes. However, the Gnepine hypothesis has not been completely accepted by botanists, and it is still difficult to understand the sister relationship between Gnetales and Pinaceae from morphological characters. In addition, the evolutionary relationships of congeneric species that originated from recent radiation are difficult to be resolved due to frequent interspecific gene flow and incomplete lineage sorting. The molecular phylogenies with low resolution should be cautiously used in biogeographic inference. It also should be mentioned that most fossil evidence used in previous biogeographic studies of the Northern Hemisphere gymnosperms is from Europe and North America. More fossil evidence from Asia is particularly helpful to test the biogeographic hypotheses.

The draft genome sequences of conifers provide an unprecedented opportunity and an important platform for the evolutionary studies of gymnosperms, especially for unraveling the mechanisms of genome evolution and the genetic basis of morphological characters. It will also shed new light on evolution of many important gene families and biological pathways in seed plants, even land plants. However, these genome data should be carefully used in phylogenetic reconstruction of gymnosperms before the gene orthology/paralogy is clarified. According to our experience, using some genes with a clear evolutionary history is much better than using all genes in the genome that have not been well studied. Phylogenomics does not always work very well. Moreover, it would be very valuable to investigate the speciation mechanisms, the adaptive value of large nuclear genomes, and the response of genome to environmental changes in gymnosperms using next-generation sequencing techniques. Also, more phylogeographical studies are encouraged to reveal the response of gymnosperms to climatic oscillations and the locations of glacial refugia, which need a wide population sampling of closely related species. This kind of knowledge is important for the conservation of gymnosperms.

Acknowledgments

The authors thank the anonymous reviewer's insightful comments and suggestions on the manuscript. We also thank Dr. Dong-Mei Guo for providing some information of Gnetales. This work was supported by the National Natural Science Foundation of China (Grant Numbers: 31170197, 31330008).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ympvev.2014.02.005>.

References

Adams, P.R., Bartel, J.A., Price, R.A., 2009. A new genus, *Hesperocyparis*, for the cypresses of the Western Hemisphere (Cupressaceae). *Phytologia* 91, 160–185.

Ahuja, M.R., 2005. Polyploidy in gymnosperms: revisited. *Silvae Genet.* 54, 59–69.

Ahuja, M.R., Neale, D., 2005. Evolution of genome size in conifers. *Silvae Genet.* 54, 126–137.

Alvarez-Yepiz, J.C., Dovicak, M., Burquez, A., 2011. Persistence of a rare ancient cycad: effects of environment and demography. *Biol. Conserv.* 144, 122–130.

Armenise, L., Simeone, M., Piredda, R., Schirone, B., 2012. Validation of DNA barcoding as an efficient tool for taxon identification and detection of species diversity in Italian conifers. *Eur. J. Forest Res.* 131, 1337–1353.

Barker, N.P., Weston, P.H., Rutschmann, F., Sauquet, H., 2007. Molecular dating of the 'Gondwanan' plant family Proteaceae is only partially congruent with the timing of the break-up of Gondwana. *J. Biogeogr.* 34, 2012–2027.

Biffin, E., Hill, R.S., Lowe, A.J., 2010. Did kauri (*Agathis*: Araucariaceae) really survive the Oligocene drowning of New Zealand? *Syst. Biol.* 59, 594–601.

Biffin, E., Conran, J.G., Lowe, A.J., 2011. Podocarp evolution: a molecular phylogenetic perspective. In: Turner, B., Cernusak, L. (Eds.), *Ecology of Podocarpaceae in Tropical Forests*. Smithsonian Contributions to Botany No. 95. Smithsonian Scholarly Press, Washington, DC.

Birky Jr., C.W., 2008. Uniparental inheritance of organelle genes. *Curr. Biol.* 18, R692–R695.

Biról, I., Raymond, A., Jackman, S.D., Pleasance, S., Coope, R., Taylor, G.A., Yuen, M.M.S., Keeling, C.I., Brand, D., Vandervalk, B.P., Kirk, H., Pandoh, P., Moore, R.A., Zhao, Y., Mungall, A.J., Jaquish, B., Yanchuk, A., Ritland, C., Boyle, B., Bousquet, J., Ritland, K., MacKay, J., Bohlmann, J., Jones, S.J.M., 2013. Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinformatics* 29, 1492–1497.

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

Braukmann, T.W., Kuzmina, M., Stefanovic, S., 2009. Loss of all plastid *ndh* genes in Gnetales and conifers: extent and evolutionary significance for the seed plant phylogeny. *Curr. Genet.* 55, 323–337.

Brenner, E.D., Stevenson, D.W., Twigg, R.W., 2003. Cycads: evolutionary innovations and the role of plant-derived neurotoxins. *Trends Plant Sci.* 8, 446–452.

Brunsfeld, S.J., Soltis, P.S., Soltis, D.E., Gadek, P.A., Quinn, C.J., Strenge, D.D., Ranker, T.A., 1994. Phylogenetic relationships among the genera of Taxodiaceae and Cupressaceae: evidence from *rbcL* sequences. *Syst. Bot.* 19, 253–262.

Burleigh, J.G., Mathews, S., 2004. Phylogenetic signal in nucleotide data from seed plants: implications for resolving the seed plant tree of life. *Am. J. Bot.* 91, 1599–1613.

Burleigh, J.G., Mathews, S., 2007a. Assessing among-locus variation in the inference of seed plant phylogeny. *Int. J. Plant Sci.* 168, 111–124.

Burleigh, J.G., Mathews, S., 2007b. Assessing systematic error in the inference of seed plant phylogeny. *Int. J. Plant Sci.* 168, 125–135.

Burleigh, J.G., Barbazuk, W.B., Davis, J.M., Morse, A.M., Soltis, P.S., 2012. Exploring diversification and genome size evolution in extant gymnosperms through phylogenetic synthesis. *J. Bot.* 2012, 1–6.

Buschiazio, E., Ritland, C., Bohlmann, J., Ritland, K., 2012. Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. *BMC Evol. Biol.* 12, 8.

Carpenter, R., 1991. *Macrozamia* from the early Tertiary of Tasmania and a study of the cuticles of extant species. *Aust. J. Bot.* 4, 433–444.

Chamberlain, C.J., 1935. *Gymnosperms: Structure and Evolution*. University of Chicago Press, Chicago, Illinois.

Chase, M.W., Reveal, J.L., 2009. A phylogenetic classification of the land plants to accompany APG III. *Bot. J. Linn. Soc.* 161, 122–127.

Chaw, S.M., Zharkikh, A., Sung, H.M., Lau, T.C., Li, W.H., 1997. Molecular phylogeny of extant gymnosperms and seed plant evolution: analysis of nuclear 18S rRNA sequences. *Mol. Biol. Evol.* 14, 56–68.

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

Chaw, S.M., Walters, T.W., Chang, C.C., Hu, S.H., Chen, S.H., 2005. A phylogeny of cycads (Cycadales) inferred from chloroplast *matK* gene, *trnK* intron, and nuclear rDNA ITS region. *Mol. Phylogenet. Evol.* 37, 214–234.

Chaw, S.M., Shih, A.C., Wang, D., Wu, Y.W., Liu, S.M., Chou, T.Y., 2008. The mitochondrial genome of the gymnosperm *Cycas taitungensis* contains a novel family of short interspersed elements, *Bpu* sequences, and abundant RNA editing sites. *Mol. Biol. Evol.* 25, 603–615.

Chen, C.H., Huang, J.P., Tsai, C.C., Chaw, S.M., 2009. Phylogeny of *Calocedrus* (Cupressaceae), an eastern Asian and western North American disjunct gymnosperm genus, inferred from nuclear ribosomal nrITS sequences. *Bot. Stud.* 50, 425–433.

Chen, J., Uebbing, S., Gyllenstrand, N., Lagercrantz, U., Lascoux, M., Kallman, T., 2012. Sequencing of the needle transcriptome from Norway spruce (*Picea abies* Karst L.) reveals lower substitution rates, but similar selective constraints in gymnosperms and angiosperms. *BMC Genom.* 13, 589.

Cheng, Y., Nicolson, R.G., Tripp, K., Chaw, S.-M., 2000. Phylogeny of Taxaceae and Cephalotaxaceae genera inferred from chloroplast *matK* gene and nuclear rDNA ITS region. *Mol. Phylogenet. Evol.* 14, 353–365.

Christenhusz, M.J.M., Reveal, J.L., Farjón, A., Gardner, M.F., Mill, R.R., Chase, M.W., 2011. A new classification and linear sequence of extant gymnosperms. *Phytotaxa* 19, 55–70.

Chuang, T.-I., Hu, W., 1963. Study of *Amentotaxus argotaenia* (Hance) Pilger. *Bot. Bull. Acad. Sin.* 4, 10–14.

Cibrián-Jaramillo, A., De la Torre-Bárcena, J.E., Lee, E.K., Katari, M.S., Little, D.P., Stevenson, D.W., Martienssen, R., Coruzzi, G.M., DeSalle, R., 2010. Using phylogenomic patterns and gene ontology to identify proteins of importance in plant evolution. *Genome Biol. Evol.* 2, 225–239.

Clarke, J.T., Warnock, R.C., Donoghue, P.C., 2011. Establishing a time-scale for plant evolution. *New Phytol.* 192, 266–301.

- Conran, J.G., Wood, G.M., Martin, P.G., Dowd, J.M., Quinn, C.J., Gadek, P.A., Price, R.A., 2000. Generic relationships within and between the gymnosperm families Podocarpaceae and Phyllocladaceae based on an analysis of the chloroplast gene *rbcl*. *Aust. J. Bot.* 48, 715–724.
- Crepet, W.L., Niklas, K.J., 2009. Darwin's second "abominable mystery": why are there so many angiosperm species? *Am. J. Bot.* 96, 366–381.
- Crisp, M.D., Cook, L.G., 2011. Cenozoic extinctions account for the low diversity of extant gymnosperms compared with angiosperms. *New Phytol.* 192, 997–1009.
- Crisp, M.D., Trewick, S.A., Cook, L.G., 2011. Hypothesis testing in biogeography. *Trends Ecol. Evol.* 26, 66–72.
- Crosby, K., Smith, D.R., 2012. Does the mode of plastid inheritance influence plastid genome architecture. *PLoS ONE* 7, e46260.
- Davis, C.C., Schaefer, H., 2011. Plant evolution: pulses of extinction and speciation in gymnosperm diversity. *Curr. Biol.* 21, R995–R998.
- de la Torre-Barcelona, J.E., Kolokotronis, S.O., Lee, E.K., Stevenson, D.W., Brenner, E.D., Katari, M.S., Coruzzi, G.M., DeSalle, R., 2009. The impact of outgroup choice and missing data on major seed plant phylogenetics using genome-wide EST data. *PLoS ONE* 4, e5764.
- Del Fueyo, G.M., Archangelisky, S., Llorens, M., Cuneo, R., 2008. Coniferous ovulate cones from the Lower Cretaceous of Santa Cruz Province, Argentina. *Int. J. Plant Sci.* 169, 799–813.
- Delevoyas, T., 1980. Polyploidy in gymnosperms. In: Lewis, W.H. (Ed.), *Polyploidy*. Plenum Press, New York, pp. 215–218.
- Donoghue, M.J., Doyle, J.A., 2000. Seed plant phylogeny: demise of the anthophyte hypothesis? *Curr. Biol.* 10, R106–R109.
- Doyle, J.A., 1998. Phylogeny of vascular plants. *Annu. Rev. Ecol. Syst.* 29, 567–599.
- Doyle, J.A., Donoghue, M.J., 1986. Seed plant phylogeny and the origin of angiosperms – an experimental cladistic approach. *Bot. Rev.* 52, 321–431.
- Dupont-Nivet, G., Krijgsman, W., Langereis, C.G., Abels, H.A., Dai, S., Fang, X., 2007. Tibetan plateau aridification linked to global cooling at the Eocene–Oligocene transition. *Nature* 445, 635–638.
- Eckenwalder, J.E., 1976. Re-evaluation of Cupressaceae and Taxodiaceae: a proposed merger. *Madrono* 23, 237–256.
- Eckenwalder, J.E., 2009. *Conifers of the World: The Complete Reference*. Timber Press, London.
- Eckert, A.J., Hall, B.D., 2006. Phylogeny, historical biogeography, and patterns of diversification for *Pinus* (Pinaceae): phylogenetic tests of fossil-based hypotheses. *Mol. Phylogenet. Evol.* 40, 166–182.
- Eldrett, J.S., Greenwood, D.R., Harding, I.C., Huber, M., 2009. Increased seasonality through the Eocene to Oligocene transition in northern high latitudes. *Nature* 459, 969–991.
- Farjón, A., 1990. Pinaceae: drawings and descriptions of the genera *Abies*, *Cedrus*, *Pseudolarix*, *Keteleeria*, *Nothotsuga*, *Tsuga*, *Cathaya*, *Pseudotsuga*, *Larix* and *Picea*. In: *Regnum Vegetabile*. Koeltz Scientific Books, Königstein, Germany.
- Farjón, A., 2001. *World Checklist and Bibliography of Conifers*, second ed. Royal Botanic Gardens, Kew, England.
- Farjón, A., 2005. *A Bibliography of Cupressaceae and Sciadopityes*. Royal Botanic Gardens, Kew, England.
- Farjón, A., 2010. *A Handbook of the World Conifers*, vol. 1,2. Brill Press, Leiden.
- Fawcett, J.A., Van de Peer, Y., Maere, S., 2013. Significance and biological consequences of polyploidization in land plants. In: Leitch, I.J., Greilhuber, J., Dolezel, J., Wendel, J.F. (Eds.), *Plant Genome Diversity, Physical Structure, Behaviour and Evolution of Plant Genomes*. Springer-Verlag, Wien, pp. 277–293.
- Felsenstein, J., 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27, 401–410.
- Finet, C., Timme, R.E., Delwiche, C.F., Marlétaz, F., 2010. Multigene phylogeny of the green lineage reveals the origin and diversification of land plants. *Curr. Biol.* 20, 2217–2222.
- Florin, R., 1963. The distribution of conifer and taxid genera in time and space. *Acta Hort. Berg.* 20, 121–312.
- Friedman, W.E., 1993. The evolutionary history of the seed plant male gametophyte. *Trends Ecol. Evol.* 8, 15–21.
- Friesen, N., Brandes, A., Heslop-Harrison, J.S., 2001. Diversity, origin, and distribution of retrotransposons (*gypsy* and *copla*) in conifers. *Mol. Biol. Evol.* 18, 1176–1188.
- Frohlich, M.W., Estabrook, G.F., 2000. Wilkinson support calculated with exact probabilities: an example using *Floricaula/LEAFY* amino acid sequences that compares three hypotheses involving gene gain/loss in seed plants. *Mol. Biol. Evol.* 17, 1914–1925.
- Fuchs, J., Brandes, A., Schubert, I., 1995. Systematics and evolution telomere sequence localization and karyotype evolution in higher plants. *Plant Syst. Evol.* 196, 227–241.
- Gadek, P.A., Alpers, D.L., Heslewood, M.M., Quinn, C.J., 2000. Relationships within Cupressaceae *sensu lato*: a combined morphological and molecular approach. *Am. J. Bot.* 87, 1044–1057.
- García-Gil, M.R., 2008. Evolutionary aspects of functional and pseudogene members of the phytochrome gene family in Scots pine. *J. Mol. Evol.* 67, 222–232.
- Gaudeul, M., Rouhan, G., Gardner, M.F., Hollingsworth, P.M., 2012. AFLP markers provide insights into the evolutionary relationships and diversification of New Caledonian *Araucaria* species (Araucariaceae). *Am. J. Bot.* 99, 68–81.
- Gernandt, D.S., López, G.G., García, S.O., Liston, A., 2005. Phylogeny and classification of *Pinus*. *Taxon* 54, 29–42.
- Gernandt, D.S., Magallón, S., López, G.G., Flores, O.Z., Willyard, A., Liston, A., 2008. Use of simultaneous analyses to guide fossil-based calibrations of Pinaceae phylogeny. *Int. J. Plant Sci.* 169, 1086–1099.
- Ghimire, B., Heo, K., 2014. Cladistic analysis of Taxaceae s. l. *Plant Syst. Evol.* 300, 217–223.
- Gillespie, R.G., Baldwin, B.G., Waters, J.M., Fraser, C.I., Nikula, R., Roderick, G.K., 2012. Long-distance dispersal: a framework for hypothesis testing. *Trends Ecol. Evol.* 27, 52–61.
- Givnish, T.J., Renner, S.S., 2004. Tropical intercontinental disjunctions: Gondwana breakup, immigration from the boreotropics, and transoceanic dispersal. *Int. J. Plant Sci.* 165, S1–S6.
- Grover, C., Wendel, J.F., 2010. Recent insights into mechanisms of genome size change in plants. *J. Bot.* 2010, 1–8.
- Guan, Q.-L., Yu, Z.-L., Feng, Y.-L., 1993. Studies of chromosomes in *Amentotaxus argotaenai*. *Acta Bot. Yunn.* 15, 385–391.
- Gugerli, F., Sperisen, C., Buchler, U., Brunner, I., Brodbeck, S., Palmer, J.D., Qiu, Y.L., 2001. The evolutionary split of Pinaceae from other conifers: evidence from an intron loss and a multigene phylogeny. *Mol. Phylogenet. Evol.* 21, 167–175.
- Hajibabaei, M., Xia, J., Drouin, G., 2006. Seed plant phylogeny: gnetophytes are derived conifers and a sister group to Pinaceae. *Mol. Phylogenet. Evol.* 40, 208–217.
- Hao, D., Xiao, P., Huang, B., Ge, G., Yang, L., 2008. Interspecific relationships and origins of Taxaceae and Cephalotaxaceae revealed by partitioned Bayesian analyses of chloroplast and nuclear DNA sequences. *Plant Syst. Evol.* 276, 89–104.
- Hart, J.A., 1987. A cladistic analysis of conifers: preliminary results. *J. Arn. Arb.* 68, 269–307.
- Hasebe, M., Kofuji, R., Ito, M., Kato, M., Iwatsuki, K., Ueda, K., 1992. Phylogeny of gymnosperms inferred from *rbcl* gene sequences. *Bot. Mag. Tokyo* 105, 673–679.
- Havill, N.P., Campbell, C.S., Vining, T.F., LePage, B., Bayer, R.J., Donoghue, M.J., 2008. Phylogeny and biogeography of *Tsuga* (Pinaceae) inferred from nuclear ribosomal ITS and chloroplast DNA sequence data. *Syst. Bot.* 33, 478–489.
- Hayata, B., 1931. The Sciadopityaceae represented by *Sciadopitys verticillata* Sieb. et Zucc., an endemic species of Japan. *Bot. Mag. Tokyo* 45, 567–569.
- Hedges, S.B., Dudley, J., Kumar, S., 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* 22, 2971–2972.
- Hendy, M.D., Penny, D., 1989. A framework for the quantitative study of evolutionary trees. *Syst. Zool.* 38, 297–309.
- Hermesen, E.J., Taylor, T.N., Taylor, E.L., Stevenson, D.W., 2006. Cataphylls of the Middle Triassic cycad *Antarcticycas schopfi* and new insights into cycad evolution. *Am. J. Bot.* 93, 724–738.
- Hill, R.S., 1995. Conifer origin, evolution, diversification in the Southern Hemisphere. In: Enright, N.J., Hill, R.S. (Eds.), *Ecology of the Southern Conifers*. Melbourne Univ. Press, Melbourne, pp. 10–29.
- Hill, R.S., 1998. The fossil record of cycads in Australia. In: McCarthy, P.M. (Ed.), *Flora of Australia, Ferns, Gymnosperms, and Allied Groups*, vol. 48. CSIRO Publishing, Melbourne, Australia, pp. 539–544.
- Hill, R.S., 2004. Origins of the southeastern Australian vegetation. *Phil. Trans. Roy. Soc. B* 359, 1537–1549.
- Hill, R.S., Brodribb, T.J., 1999. Southern conifers in time and space. *Aust. J. Bot.* 47, 639–696.
- Hill, K.D., Chase, M.W., Stevenson, D.W., Hills, H.G., Schutzman, B., 2003. The families and genera of cycads: a molecular phylogenetic analysis of cycadophyta based on nuclear and plastid DNA sequences. *Int. J. Plant Sci.* 164, 933–948.
- Hirao, T., Watanabe, A., Kurita, M., Kondo, T., Takata, K., 2008. Complete nucleotide sequence of the *Cryptomeria japonica* D. Don. chloroplast genome and comparative chloroplast genomics: diversified genomic structure of coniferous species. *BMC Plant Biol.* 8, 70.
- Hizume, M., 1988. Karyomorphological studies in the family Pinaceae. *Mem. Fac. Educ. Ehime Univ. Nat. Sci.* 8, 1–108.
- Hizume, M., Shibata, F., Matsusaki, Y., Garajova, Z., 2002. Chromosome identification and comparative karyotypic analyses of four *Pinus* species. *Theor. Appl. Genet.* 105, 491–497.
- Hollander, J.L., Vander Wall, S.B., 2009. Dispersal syndromes in North American *Ephedra*. *Int. J. Plant Sci.* 170, 323–330.
- Hopkins, D.M., 1967. *The Bering Land Bridge*. Stanford University Press, Stanford.
- Hori, H., Lim, B.-L., Osawa, S., 1985. *Evolution of green plants as deduced from 5S rRNA sequences*. *Proc. Natl. Acad. Sci. USA* 82, 820–823.
- Huang, J., Giannasi, D.E., Price, R.A., 2005. Phylogenetic relationships in *Ephedra* (Ephedraceae) inferred from chloroplast and nuclear DNA sequences. *Mol. Phylogenet. Evol.* 35, 48–59.
- Huelsenbeck, J.P., 1995. Performance of phylogenetic methods in simulation. *Syst. Biol.* 44, 17–48.
- Ian Milne, R., 2006. Northern Hemisphere plant disjunctions: a window on Tertiary land bridges and climate change? *Ann. Bot.* 98, 465–472.
- Ickert-Bond, S.M., Rydin, C., Renner, S.S., 2009. A fossil-calibrated relaxed clock for *Ephedra* indicates an Oligocene age for the divergence of Asian and New World clades and Miocene dispersal into South Africa. *J. Syst. Evol.* 47, 444–456.
- Ikeno, S., Hirase, S., 1897. Spermatozoids in gymnosperms. *Ann. Bot.* 11, 344–345.
- Ivany, L.C., Patterson, W.P., Lohmann, K.C., 2000. Cooler winters as a possible cause of mass extinctions at the Eocene/Oligocene boundary. *Nature* 407, 887–890.
- Jaramillo-Correa, J., Aguirre-Planter, E., Eguiarte, L., Khasa, D., Bousquet, J., 2013. Evolution of an ancient microsatellite hotspot in the conifer mitochondrial genome and comparison with other plants. *J. Mol. Evol.* 76, 146–157.
- Jeffroy, O., Brinkmann, H., Delsuc, F., Philippe, H., 2006. Phylogenomics: the beginning of incongruence? *Trends Genet.* 22, 225–231.

- Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E., Tomsho, L.P., Hu, Y., Liang, H., Soltis, P.S., Soltis, D.E., Clifton, S.W., Schlarbaum, S.E., Schuster, S.C., Ma, H., Leebens-Mack, J., dePamphilis, C.W., 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473, 97–100.
- Johnson, L.A.S., 1959. The families of cycads and the Zamiaceae of Australia. *Proc. Linn. Soc. NSW* 84, 64–117.
- Jones, D.L., 2002. Cycads of the World—Ancient Plant in Today's Landscape, second ed. Smithsonian Institution Press, Washington, DC.
- Kelch, D.G., 1998. Phylogeny of Podocarpaceae: comparison of evidence from morphology and 18S rDNA. *Am. J. Bot.* 85, 986–996.
- Keppel, G., Hodgskiss, P.D., Plunkett, G.M., 2008. Cycads in the insular South-west Pacific: dispersal or vicariance? *J. Biogeogr.* 35, 1004–1015.
- Khosho, T.N., 1959. Polyploidy in gymnosperms. *Evolution* 13, 24–39.
- Khosho, T.N., 1961. Chromosome numbers in gymnosperms. *Silvae Genet.* 10, 1–7.
- Kinlaw, C.S., Neale, D.B., 1997. Complex gene families in pine genomes. *Trends Plant Sci.* 2, 356–359.
- Klymiuk, A.A., Stockey, R.A., 2012. A lower Cretaceous (Valanginian) seed cone provides the earliest fossil record for *Picea* (Pinaceae). *Am. J. Bot.* 99, 1069–1082.
- Knapp, M., Stockler, K., Havell, D., Delsuc, F., Sebastiani, F., Lockhart, P.J., 2005. Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (southern beech). *PLoS Biol.* 3, e14.
- Knopf, P., Schulz, C., Little, D.P., Stützel, T., Stevenson, D.W., 2012. Relationships within Podocarpaceae based on DNA sequence, anatomical, morphological, and biogeographical data. *Cladistics* 28, 271–299.
- Kovach, A., Wegrzyn, J.L., Parra, G., Holt, C., Bruening, G.E., Loopstra, C.A., Hartigan, J., Yandell, M., Langley, C.H., Korf, I., Neale, D.B., 2010. The *Pinus taeda* genome is characterized by diverse and highly diverged repetitive sequences. *BMC Genom.* 11, 420.
- Krutovsky, K.V., Troggio, M., Brown, G.R., Jermstad, K.D., Neale, D.B., 2004. Comparative mapping in the Pinaceae. *Genetics* 168, 447–461.
- Lamm, K.S., Redelings, B.D., 2009. Reconstructing ancestral ranges in historical biogeography: properties and prospects. *J. Syst. Evol.* 47, 369–382.
- Lee, E.K., Cibrian-Jaramillo, A., Kolokotronis, S.-O., Katari, M.S., Stamatakis, A., Ott, M., Chiu, J.C., Little, D.P., Stevenson, D.W., McCombie, W.R., Martienssen, R.A., Coruzzi, G., DeSalle, R., 2011. A functional phylogenomic view of the seed plants. *PLoS Genet.* 7, e1002411.
- Leitch, I.J., Leitch, A.R., 2013. Genome size diversity and evolution in land plants. In: Leitch, I.J., Greilhuber, J., Doležel, J., Wendel, J.F. (Eds.), *Plant Genome Diversity Volume 2, Physical Structure, Behaviour and Evolution of Plant Genomes*. Springer-Verlag, Wien, pp. 307–322.
- Leitch, I.J., Hanson, L., Winfield, M., Parker, J., Bennett, M.D., 2001. Nuclear DNA C-values complete familial representation in gymnosperms. *Ann. Bot.* 88, 843–849.
- LePage, B.A., 2001. New species of *Picea* A. Dietrich (Pinaceae) from the middle Eocene of Axel Heiberg Island, Arctic Canada. *Bot. J. Linn. Soc.* 135, 137–167.
- Leslie, A.B., Beaulieu, J.M., Rai, H.S., Crane, P.R., Donoghue, M.J., Mathews, S., 2012. Hemisphere-scale differences in conifer evolutionary dynamics. *Proc. Natl. Acad. Sci. USA* 109, 16217–16221.
- Li, C., Yang, Q., 2002. Divergence time estimates for major lineages of Cupressaceae(s). *Acta Phytotaxon. Sin.* 40, 323–333.
- Li, J.H., Davis, C.C., Del Tredici, P., Donoghue, M.J., 2001. Phylogeny and biogeography of *Taxus* (Taxaceae) inferred from sequences of nuclear ribosomal ITS region. *Harvard Pap. Bot.* 6, 267–274.
- Lin, C.P., Huang, J.P., Wu, C.S., Hsu, C.Y., Chaw, S.M., 2010. Comparative chloroplast genomics reveals the evolution of Pinaceae genera and subfamilies. *Genome Biol. Evol.* 2, 504–517.
- Little, D.P., 2006. Evolution and circumscription of the true cypresses (Cupressaceae: *Cupressus*). *Syst. Bot.* 31, 461–480.
- Liu, N., Zhu, Y., Wei, Z., Chen, J., Wang, Q., Jian, S., Zhou, D., Shi, J., Yang, Y., Zhong, Y., 2009. Phylogenetic relationships and divergence times of the family Araucariaceae based on the DNA sequences of eight genes. *Chin. Sci. Bull.* 54, 2648–2655.
- Lockwood, J.D., Aleksić, J.M., Zou, J., Wang, J., Liu, J., Renner, S.S., 2013. A new phylogeny for the genus *Picea* from plastid, mitochondrial, and nuclear sequences. *Mol. Phylogenet. Evol.* 69, 717–727.
- Lorenz, W.W., Ayyampalayam, S., Bordeaux, J., Howe, G., Jermstad, K., Neale, D., Rogers, D., Dean, J.D., 2012. Conifer DBMagic: a database housing multiple de novo transcriptome assemblies for 12 diverse conifer species. *Tree Genet. Gen.* 8, 1477–1485.
- Lovisetto, A., Guzzo, F., Tadiello, A., Toffali, K., Favretto, A., Casadoro, G., 2012. Molecular analyses of MADS-box genes trace back to gymnosperms the invention of fleshy fruits. *Mol. Biol. Evol.* 29, 409–419.
- Lovisetto, A., Guzzo, F., Busatto, N., Casadoro, G., 2013. Gymnosperm B-sister genes may be involved in ovule/seed development and, in some species, in the growth of fleshy fruit-like structures. *Ann. Bot.* 112, 535–544.
- Mackay, J., Dean, J.F., Plomion, C., Peterson, D.G., Canovas, F.M., Pavy, N., Ingvarsson, P.K., Savolainen, O., Guevara, M.A., Fluch, S., Vinceti, B., Abarca, D., Diaz-Sala, C., Cervera, M.T., 2012. Towards decoding the conifer giga-genome. *Plant Mol. Biol.* 80, 555–569.
- Macphail, M.K., Alley, N.F., Truswell, E.M., Sluiter, I.R.K., 1994. Early tertiary vegetation: evidence from spores and pollen. In: Hill, R.S. (Ed.), *History of the Australian Vegetation: Cretaceous to Recent*. Cambridge University Press, Cambridge, UK, pp. 189–261.
- Magallón, S., Sanderson, M.J., 2002. Relationships among seed plants inferred from highly conserved genes: sorting conflicting phylogenetic signals among ancient lineages. *Am. J. Bot.* 89, 1991–2006.
- Magallón, S., Hilu, K.W., 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.
- Mao, K., Hao, G., Liu, J., Adams, R.P., Milne, R.I., 2010. Diversification and biogeography of *Juniperus* (Cupressaceae): variable diversification rates and multiple intercontinental dispersals. *New Phytol.* 188, 254–272.
- Mao, K., Milne, R.I., Zhang, L., Peng, Y., Liu, J., Thomas, P., Mill, R.R., Renner, S.S., 2012. Distribution of living Cupressaceae reflects the breakup of Pangea. *Proc. Natl. Acad. Sci. USA* 109, 7793–7798.
- Mathews, S., 2009. Phylogenetic relationships among seed plants: persistent questions and the limits of molecular data. *Am. J. Bot.* 96, 228–236.
- Mathews, S., Clements, M.D., Beilstein, M.A., 2010. A duplicate gene rooting of seed plants and the phylogenetic position of flowering plants. *Phil. Trans. Roy. Soc. B* 365, 383–395.
- McCauley, D.E., Bailey, M.F., Sherman, N.A., Darnell, M.Z., 2005. Evidence for paternal transmission and heteroplasmy in the mitochondrial genome of *Silene vulgaris*, a gynodioecious plant. *Heredity* 95, 50–58.
- McCoy, S.R., Kuehl, J.V., Boore, J.L., Raubeson, L.A., 2008. The complete plastid genome sequence of *Welwitschia mirabilis*: an unusually compact plastome with accelerated divergence rates. *BMC Evol. Biol.* 8, 130.
- McLoughlin, S., 2001. The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Aust. J. Bot.* 49, 271–300.
- McLoughlin, S., Vajda, V., 2005. Ancient wollemi pines resurgent. *Am. Sci.* 93, 540–547.
- Miller Jr, C.N., 1999. Implications of fossil conifers for the phylogenetic relationships of living families. *Bot. Rev.* 65, 239–277.
- Miller Jr, C.N., LaPasha, C.A., 1983. Structure and affinities of *Athrotaxites berryi* Bell, an Early Cretaceous conifer. *Am. J. Bot.* 70, 772–779.
- Mogensen, H.L., 1996. The hows and whys of cytoplasmic inheritance in seed plants. *Am. J. Bot.* 83, 383–404.
- Moran, K., Backman, J., Brinkhuis, H., Clemens, S.C., Cronin, T., Dickens, G.R., Eynaud, F., Gattacceca, J., Jakobsson, M., Jordan, R.W., Kaminski, M., King, J., Koc, N., Krylov, A., Martinez, N., Matthiessen, J., McInroy, D., Moore, T.C., Onodera, J., O'Regan, M., Palike, H., Rea, B., Rio, D., Sakamoto, T., Smith, D.C., Stein, R., St John, K., Suto, I., Suzuki, N., Takahashi, K., Watanabe, M., Yamamoto, M., Farrell, J., Frank, M., Kubik, P., Jokat, W., Kristoffersen, Y., 2006. The Cenozoic palaeoenvironment of the Arctic Ocean. *Nature* 441, 601–605.
- Morgante, M., De Pauli, E., 2011. Toward the conifer genome sequence. In: Plomion, C., Bousquet, J., Kole, C. (Eds.), *Genetics, Genomics and Breeding of Conifer Trees*. CABI International Publishers and CRC Press, New York.
- Morse, A.M., Peterson, D.G., Islam-Faridi, M.N., Smith, K.E., Magbanua, J., Garcia, S.A., Kubisiak, T.L., Amerson, H.V., Carlson, J.E., Nelson, C.D., Davis, J.M., 2009. Evolution of genome size and complexity in *Pinus*. *PLoS ONE* 4, e4332.
- Mundry, M., Stutzel, T., 2004. Morphogenesis of the reproductive shoots of *Welwitschia mirabilis* and *Ephedra distachya* (Gnetales), and its evolutionary implications. *Org. Div. Evol.* 4, 91–108.
- Murray, B.G., 1998. Nuclear DNA amounts in gymnosperms. *Ann. Bot.* 82 (Supplement A), 3–15.
- Murray, B.G., 2013. Karyotype variation and evolution in gymnosperms. In: Greilhuber, J., Dolezel, J., Wendel, J.F. (Eds.), *Plant Genome Diversity*. Springer Vienna, pp. 231–243.
- Nagalingum, N.S., Marshall, C.R., Quental, T.B., Rai, H.S., Little, D.P., Mathews, S., 2011. Recent synchronous radiation of a living fossil. *Science* 334, 796–799.
- Niklas, K.J., 1997. *The Evolutionary Biology of Plants*. University of Chicago press, Chicago.
- Nkongolo, K.K., Mehesh-Smith, M., 2012. Karyotype evolution in the Pinaceae: implication with molecular phylogeny. *Genome* 55, 735–753.
- Norstog, K.J., Nicholls, T.J., 1997. The fossil Cycadophytes. In: Norstog, K.J., Nicholls, T.J. (Eds.), *The Biology of Cycads*. Cornell University Press, New York, pp. 169–201.
- Norstog, K.J., Gifford, E.M., Stevenson, D.W., 2004. Comparative development of the spermatozooids of cycads and *Ginkgo biloba*. *Bot. Rev.* 70, 5–15.
- Nystedt, B., Streit, N.R., Wetterbom, A., Zuccolo, A., Lin, Y.-C., Scofield, D.G., Vezzi, F., Delhomme, N., Giacomello, S., Alexeyenko, A., Vicedomini, R., Sahlin, K., Sherwood, E., Elfstrand, M., Gramzow, L., Holmberg, K., Hallman, J., Keech, O., Klasson, L., Koriabine, M., Kucukoglu, M., Kaller, M., Luthman, J., Lysholm, F., Niittyta, T., Olson, A., Rilakovic, N., Ritland, C., Rossello, J.A., Sena, J., Svensson, T., Talavera-Lopez, C., Theissen, G., Tuominen, H., Vanneste, K., Wu, Z.-Q., Zhang, B., Zerbe, P., Arvestad, L., Bhalerao, R., Bohlmann, J., Bousquet, J., Garcia Gil, R., Hvidsten, T.R., de Jong, P., MacKay, J., Morgante, M., Ritland, K., Sundberg, B., Lee Thompson, S., Van de Peer, Y., Andersson, B., Nilsson, O., Ingvarsson, P.K., Lundeberg, J., Jansson, S., 2013. The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584.
- Ohri, D., Khosho, T.N., 1986. Genome size in gymnosperms. *Plant Syst. Evol.* 153, 119–132.
- Osborne, R., Calonje, M.A., Hill, K.D., Stanberg, L., Stevenson, D.W., 2012. The world list of cycads. *Mem. New York Bot. Gard.* 106, 480–510.
- Palmer, J.D., Thompson, W.F., 1982. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequences is lost. *Cell* 29, 537–550.
- Palmer, J.D., Soltis, D.E., Chase, M.W., 2004. The plant tree of life: an overview and some points of view. *Am. J. Bot.* 91, 1437–1445.

- Parks, M., Cronn, R., Liston, A., 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biol.* 7, 84.
- Parks, M., Cronn, R., Liston, A., 2012. Separating the wheat from the chaff: mitigating the effects of noise in a plastome phylogenomic data set from *Pinus* L. (Pinaceae). *BMC Evol. Biol.* 12, 100.
- Paull, R., Hill, R.S., 2010. Early oligocene *Callitris* and *Fitzroya* (Cupressaceae) from Tasmania. *Am. J. Bot.* 97, 809–820.
- Pavy, N., Pelgas, B., Laroche, J., Rigault, P., Isabel, N., Bousquet, J., 2012. A spruce gene map infers ancient plant genome reshuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers. *BMC Biol.* 10, 84.
- Pelgas, B., Beauseigle, S., Achere, V., Jeandroz, S., Bousquet, J., Isabel, N., 2006. Comparative genome mapping among *Picea glauca*, *P. mariana* x *P. rubens* and *P. abies*, and correspondence with other Pinaceae. *Theor. Appl. Genet.* 113, 1371–1393.
- Peng, D., Wang, X.-Q., 2008. Reticulate evolution in *Thuja* inferred from multiple gene sequences: implications for the study of biogeographical disjunction between eastern Asia and North America. *Mol. Phylogenet. Evol.* 47, 1190–1202.
- Pilger, R., 1926. Gymnospermae 6: Klasse Coniferae. In: Engler, A., Prantler, E. (Eds.), *Die Naturliche Pflanzfamilien*, second ed. Englemann, Leipzig, pp. 121–403.
- Pittermann, J., Stuart, S.A., Dawson, T.E., Moreau, A., 2012. Cenozoic climate change shaped the evolutionary ecophysiology of the Cupressaceae conifers. *Proc. Natl. Acad. Sci. USA* 109, 9647–9652.
- Plomion, C., Bousquet, J., Kole, C., 2011. Genetics, Genomics, and Breeding of Conifers. CRC Press, Boca Raton.
- Price, R.A., 1996. Systematics of the Gnetales: a review of morphological and molecular Evidence. *Int. J. Plant Sci.* 157 (6 Suppl.), S40–S49.
- Price, R.A., Lowenstein, J.M., 1989. An immunological comparison of the Sciadopiaceae, Taxodiaceae, and Cupressaceae. *Syst. Bot.* 14, 141–149.
- Qiao, C.-Y., Ran, J.-H., Li, Y., Wang, X.-Q., 2007. Phylogeny and biogeography of *Cedrus* (Pinaceae) inferred from sequences of seven paternal chloroplast and maternal mitochondrial DNA regions. *Ann. Bot.* 100, 573–580.
- Quinn, C.J., Rattenbury, J.A., 1972. Structural hybridity in New Zealand *Dacrydium*. *New Zeal. J. Bot.* 10, 427–436.
- Quinn, C.J., Price, R.A., Gadek, P.A., 2002. Familial concepts and relationships in the conifers based on *rbcl* and *matK* sequence comparisons. *Kew. Bull.* 57, 513–531.
- Rabinowicz, P.D., Citek, R., Budiman, M.A., Nunberg, A., Bedell, J.A., Lakey, N., O'Shaughnessy, A.L., Nascimento, L.U., McCombie, W.R., Martienssen, R.A., 2005. Differential methylation of genes and repeats in land plants. *Genome Res.* 15, 1431–1440.
- Rai, H.S., O'Brien, H.E., Reeves, P.A., Olmstead, R.G., Graham, S.W., 2003. Inference of higher-order relationships in the cycads from a large chloroplast data set. *Mol. Phylogenet. Evol.* 29, 350–359.
- Rai, H.S., Reeves, P.A., Peakall, R., Olmstead, R.G., Graham, S.W., 2008. Inference of higher-order conifer relationships from a multi-locus plastid data set. *Botany* 86, 658–669.
- Rake, A.V., Mijsche, J.P., Hall, R.B., Hansen, K.M., 1980. DNA reassociation kinetics for four conifers. *Can. J. Genet. Cytol.* 22, 69–79.
- Ran, J.-H., Wei, X.-X., Wang, X.-Q., 2006. Molecular phylogeny and biogeography of *Picea* (Pinaceae): implications for phylogeographical studies using cytoplasmic haplotypes. *Mol. Phylogenet. Evol.* 41, 405–419.
- Ran, J.-H., Gao, H., Wang, X.-Q., 2010. Fast evolution of the retroprocessed mitochondrial *rps3* gene in Conifer II and further evidence for the phylogeny of gymnosperms. *Mol. Phylogenet. Evol.* 54, 136–149.
- Raubeson, L.A., Jansen, R.K., 1992. A rare chloroplast-DNA structural mutation is shared by all conifers. *Biochem. Syst. Ecol.* 20, 17–24.
- Regina, T.M.R., Quagliariello, C., 2010. Lineage-specific group II intron gains and losses of the mitochondrial *rps3* gene in gymnosperms. *Plant Physiol. Biochem.* 48, 646–654.
- Regina, T.M.R., Picardi, E., Lopez, L., Pesole, G., Quagliariello, C., 2005. A novel additional group II intron distinguishes the mitochondrial *rps3* gene in gymnosperms. *J. Mol. Evol.* 60, 196–206.
- Rigault, P., Boyle, B., Lepage, P., Cooke, J.E., Bousquet, J., MacKay, J.J., 2011. A white spruce gene catalog for conifer genome analyses. *Plant Physiol.* 157, 14–28.
- Ritland, K., 2012. Genomics of a phylum distant from flowering plants: conifers. *Tree Genet. Gen.* 8, 573–582.
- Rodríguez-Ezpeleta, N., Brinkmann, H., Burey, S.C., Roure, B., Burger, G., Löffelhardt, W., Bohnert, H.J., Philippe, H., Lang, B.F., 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr. Biol.* 15, 1325–1330.
- Ronquist, F., Sanmartin, I., 2011. Phylogenetic methods in biogeography. *Annu. Rev. Ecol. Syst.* 42, 441–464.
- Rothwell, G.W., Serber, R., 1994. Lignophyte phylogeny and the evolution of spermatophytes: a numerical cladistic analysis. *Syst. Bot.* 19, 443–482.
- Roure, B., Baurain, D., Philippe, H., 2013. Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. *Mol. Biol. Evol.* 30, 197–214.
- Ryberg, P.E., Rothwell, G.W., Stockey, R.A., Jason, H., Mapes, G., Riding, J.B., 2012. Reconsidering relationships among stem and crown group Pinaceae: oldest record of the genus *Pinus* from the Early Cretaceous of Yorkshire, United Kingdom. *Int. J. Plant Sci.* 173, 917–932.
- Rydin, C., Korall, P., 2009. Evolutionary relationships in *Ephedra* (Gnetales), with implications for seed plant phylogeny. *Int. J. Plant Sci.* 170, 1031–1043.
- Rydin, C., Kallersjö, M., Friest, E.M., 2002. Seed plant relationships and the systematic position of Gnetales based on nuclear and chloroplast DNA: conflicting data, rooting problems, and the monophyly of conifers. *Int. J. Plant Sci.* 163, 197–214.
- Salas-Leiva, D.E., Meerow, A.W., Calonje, M., Griffith, M.P., Francisco-Ortega, J., Nakamura, K., Stevenson, D.W., Lewis, C.E., Namoff, S., 2013. Phylogeny of the cycads based on multiple single-copy nuclear genes: congruence of concatenated parsimony, likelihood and species tree inference methods. *Ann. Bot.* 112, 1263–1278.
- Sanderson, M.J., Wojciechowski, M.F., Hu, J.M., Khan, T.S., Brady, S.G., 2000. Error, bias, and long-branch attraction in data for two chloroplast photosystem genes in seed plants. *Mol. Biol. Evol.* 17, 782–797.
- Sanmartin, I., Ronquist, F., 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Syst. Biol.* 53, 216–243.
- Savolainen, O., Pyhajarvi, T., 2007. Genomic diversity in forest trees. *Curr. Opin. Plant Biol.* 10, 162–167.
- Sax, H.J., 1932. Chromosome pairing in *Larix* species. *J. Arn. Arbor.* 13, 368–373.
- Sax, K., 1960. Meiosis in interspecific pine hybrids. *Forest Sci.* 6, 135–138.
- Saylor, L.C., Smith, B.W., 1966. Meiotic irregularity in species and interspecific hybrids of *Pinus*. *Am. J. Bot.* 53, 453–468.
- Schmidt, M., Schneider-Poetsch, H.A., 2002. The evolution of gymnosperms redrawn by phytochrome genes: the Gnetales appear at the base of the gymnosperms. *J. Mol. Evol.* 54, 715–724.
- Sederoff, R., 2013. Genomics: a spruce sequence. *Nature* 497, 569–570.
- Setoguchi, H., Osawa, T.A., Pintaud, J.-C., Jaffré, T., Veillon, J.-M., 1998. Phylogenetic relationships within Araucariaceae based on *rbcl* gene sequences. *Am. J. Bot.* 85, 1507–1516.
- Shibata, F., Hizume, M., 2005. Comparative FISH karyotype analysis of 11 *Picea* species. *Cytologia* 73, 203–211.
- Soltis, P.S., Soltis, D.E., Wolf, P.G., Nickrent, D.L., Chaw, S.-M., Chapman, R.L., 1999. The phylogeny of land plants inferred from 18S rDNA sequences: pushing the limits of rDNA signal? *Mol. Biol. Evol.* 16, 1774–1784.
- Soltis, D.E., Soltis, P.S., Zanis, M.J., 2002. Phylogeny of seed plants based on evidence from eight genes. *Am. J. Bot.* 89, 1670–1681.
- Stefanovic, S., Jager, M., Deusch, J., Broutin, J., Masselot, M., 1998. Phylogenetic relationships of conifers inferred from partial 28S rRNA gene sequences. *Am. J. Bot.* 85, 688.
- Stevenson, D.W., 1981. Observations on ptyxis, phenology, and trichomes in the Cycadales and their systematic implications. *Am. J. Bot.* 68, 1104–1114.
- Stevenson, D.W., 1990. Morphology and systematics of the Cycadales. *Mem. NY Bot. Gard.* 57, 8–55.
- Stevenson, D.W., 1992. A formal classification of the extant cycads. *Brittonia* 44, 220–223.
- Talavera, G., Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577.
- Tatsuya, W., Junko, T., Shiho, I., Keiko, N., Takahiko, T., Masahiro, S., 1994. Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. *Proc. Natl. Acad. Sci. USA* 91, 9794–9798.
- Taylor, T.N., Taylor, E.L., Krings, M., 2009. *Paleobotany: The Biology and Evolution of Fossil Plants*. Academic Press, Burlington, MA.
- Treutlein, J., Wink, M., 2002. Molecular phylogeny of cycads inferred from *rbcl* sequences. *Naturwissenschaften* 89, 221–225.
- Tsudzuki, J., Nakashima, K., Tsudzuki, T., Hiratsuka, J., Shibata, M., Wakasugi, T., Sugiura, M., 1992. Chloroplast DNA of black pine retains a residual inverted repeat lacking rRNA genes: nucleotide sequences of *trnQ*, *trnK*, *psbA*, *trnI* and *trnH* and the absence of *rps16*. *Mol. Gen. Genet.* 232, 206–214.
- Upchurch, P., 2008. Gondwanan break-up: legacies of a lost world? *Trends Ecol. Evol.* 23, 229–236.
- Vovides, A.P., Olivares, M., 1996. Karyotype polymorphism in the cycad *Zamia loddigesii* (Zamiaceae) of the Yucatan peninsula, Mexico. *Bot. J. Linn. Soc.* 120, 77–83.
- Wagstaff, S.J., 2004. Evolution and biogeography of the austral genus *Phyllocladus* (Podocarpaceae). *J. Biogeogr.* 31, 1569–1577.
- Wang, X.-Q., Shu, Y.-Q., 2000. Chloroplast *matK* gene phylogeny of Taxaceae and Cephalotaxaceae, with additional reference to the systematic position of *Nageia*. *Acta Phytotax. Sin.* 38, 201–210.
- Wang, X.-Q., Han, Y., Hong, D.-Y., 1998. A molecular systematic study of *Cathaya*, a relic genus of the Pinaceae in China. *Plant Syst. Evol.* 213, 165–172.
- Wang, X.-Q., Tank, D.C., Sang, T., 2000. Phylogeny and divergence times in Pinaceae: evidence from three genomes. *Mol. Biol. Evol.* 17, 773–781.
- Wang, W.P., Hwang, C.Y., Lin, T.P., Hwang, S.Y., 2003. Historical biogeography and phylogenetic relationships of the genus *Chamaecyparis* (Cupressaceae) inferred from chloroplast DNA polymorphism. *Plant Syst. Evol.* 241, 13–28.
- Wang, L., Wang, D., Lin, M.M., Lu, Y., Jiang, X.X., Jin, B., 2011. An embryological study and systematic significance of the primitive gymnosperm *Ginkgo biloba*. *J. Syst. Evol.* 49, 353–361.
- Wardle, P., 1972. *Podocarpus totara* var. *waihoensis* var. *nov.*: the result of introgressive hybridization between *P. totara* and *P. acutifolius*. *New Zeal. J. Bot.* 10, 195–201.
- Wei, X.-X., Wang, X.-Q., 2003. Phylogenetic split of *Larix*: evidence from paternally inherited cpDNA *trnT-trnF* region. *Plant Syst. Evol.* 239, 67–77.
- Wei, X.-X., Wang, X.-Q., 2004. Recolonization and radiation in *Larix* (Pinaceae): evidence from nuclear ribosomal DNA paralogues. *Mol. Ecol.* 13, 3115–3123.
- Wei, X.-X., Yang, Z.-Y., Li, Y., Wang, X.-Q., 2010. Molecular phylogeny and biogeography of *Pseudotsuga* (Pinaceae): insights into the floristic relationship between Taiwan and its adjacent areas. *Mol. Phylogenet. Evol.* 55, 776–785.
- Wen, J., 1999. Evolution of eastern Asian and eastern North American disjunct pattern in flowering plants. *Annu. Rev. Ecol. Syst.* 30, 421–455.

- Wen, J., Ickert-Bond, S., Nie, Z.-L., Li, R., 2010. Timing and modes of evolution of eastern Asian – North American biogeographic disjunctions in seed plants. In: Long, M., Gu, H., Zhou, Z. (Eds.), Darwin's Heritage Today: Proceedings of the Darwin 2010 Beijing International Conference. Higher Education Press, Beijing, pp. 252–269.
- Wen, J., Ree, R.H., Ickert-Bond, S.M., Nie, Z., Funk, V., 2013. Biogeography: where do we go from here? *Taxon* 62, 912–927.
- Williams, C.G., 2009. *Conifer Reproductive Biology*. Springer, Dordrecht.
- Williams, C.G., Joyner, K.L., Auckland, L.D., Johnston, S., Price, H.J., 2002. Genomic consequences of *Pinus* spp. hybridization. *Biol. J. Linn. Soc.* 75, 503–508.
- Willyard, A., Syring, J., Gernandt, D.S., Liston, A., Cronn, R., 2007. Fossil calibration of molecular divergence infers a moderate mutation rate and recent radiations for *Pinus*. *Mol. Biol. Evol.* 24, 90–101.
- Wing, S.L., 2004. Mass extinctions in plant evolution. In: Taylor, T.N. (Ed.), *Extinctions in the History of Life*. Cambridge University Press, Cambridge, UK.
- Winter, K.U., Becker, A., Munster, T., Kim, J.T., Saedler, H., Theissen, G., 1999. MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants. *Proc. Natl. Acad. Sci. USA* 96, 7342–7347.
- Wolfe, J.A., 1994. Tertiary climatic changes at middle latitudes of western North America. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 108, 195–205.
- Won, H., Renner, S., 2006. Dating dispersal and radiation in the gymnosperm *Gnetum* (Gnetales)-Clock calibration when outgroup relationships are uncertain. *Syst. Biol.* 55, 610–622.
- Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoon, P.B., Rieseberg, L.H., 2009. The frequency of polyploid speciation in vascular plants. *Proc. Natl. Acad. Sci. USA* 106, 13875–13879.
- Wu, C.-S., Chaw, S.-M., 2013. Highly rearranged and size-variable chloroplast genomes in conifers II clade (cupressophytes): evolution towards shorter intergenic spacers. *Plant Biotech. J.* <http://dx.doi.org/10.1111/pbi.12141>.
- Wu, C.S., Lai, Y.T., Lin, C.P., Wang, Y.N., Chaw, S.M., 2009. Evolution of reduced and compact chloroplast genomes (cpDNAs) in gnetophytes: selection toward a lower-cost strategy. *Mol. Phylogenet. Evol.* 52, 115–124.
- Wu, C.S., Wang, Y.N., Hsu, C.Y., Lin, C.P., Chaw, S.M., 2011. Loss of different inverted repeat copies from the chloroplast genomes of Pinaceae and cupressophytes and influence of heterotachy on the evaluation of gymnosperm phylogeny. *Genome Biol. Evol.* 3, 1284–1295.
- Wu, C.S., Chaw, S.M., Huang, Y.Y., 2013. Chloroplast phylogenomics indicates that *Ginkgo biloba* is sister to cycads. *Genome Biol. Evol.* 5, 243–254.
- Xi, Z., Rest, J.S., Davis, C.C., 2013. Phylogenomics and coalescent analyses resolve extant seed plant relationships. *PLoS ONE* 8, e80870.
- Xia, X., Xie, Z., Salemi, M., Chen, L., Wang, Y., 2003. An index of substitution saturation and its application. *Mol. Phylogenet. Evol.* 26, 1–7.
- Xiang, Q.-P., Xiang, Q.-Y., Guo, Y.-Y., Zhang, X.-C., 2009. Phylogeny of *Abies* (Pinaceae) inferred from nrITS sequence data. *Taxon* 58, 141–152.
- Xiao, L.Q., Moller, M., Zhu, H., 2010. High nrDNA ITS polymorphism in the ancient extant seed plant *Cycas*: incomplete concerted evolution and the origin of pseudogenes. *Mol. Phylogenet. Evol.* 55, 168–177.
- Yang, Z.-Y., Ran, J.-H., Wang, X.-Q., 2012. Three genome-based phylogeny of Cupressaceae s.l.: further evidence for the evolution of gymnosperms and Southern Hemisphere biogeography. *Mol. Phylogenet. Evol.* 64, 452–470.
- Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292, 686–693.
- Zanazzi, A., Kohn, M.J., MacFadden, B.J., Terry, D.O., 2007. Large temperature drop across the Eocene–Oligocene transition in central North America. *Nature* 445, 639–642.
- Zgurski, J.M., Rai, H.S., Fai, Q.M., Bogler, D.J., Francisco-Ortega, J., Graham, S.W., 2008. How well do we understand the overall backbone of cycad phylogeny? New insights from a large, multigene plastid data set. *Mol. Phylogenet. Evol.* 47, 1232–1237.
- Zhang, N., Zeng, L., Shan, H., Ma, H., 2012. Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in angiosperms. *New Phytol.* 195, 923–937.
- Zhong, B., Yonezawa, T., Zhong, Y., Hasegawa, M., 2010. The position of gnetales among seed plants: overcoming pitfalls of chloroplast phylogenomics. *Mol. Biol. Evol.* 27, 2855–2863.
- Zhou, Q.-X., Gu, Z.-J., Yue, Z.-S., 2000. Karyomorphology and relationships of *Amentotaxus* Pilg. *Acta Phytotaxon. Sin.* 36, 522–527.
- Zhou, L., Su, Y.C.F., Thomas, D.C., Saunders, R.M.K., 2012. 'Out-of-Africa' dispersal of tropical floras during the Miocene climatic optimum: evidence from *Uvaria* (Annonaceae). *J. Biogeogr.* 39, 322–335.
- Zonneveld, B.J.M., 2012a. Conifer genome sizes of 172 species, covering 64 of 67 genera, range from 8 to 72 picogram. *Nor. J. Bot.* 30, 490–502.
- Zonneveld, B.J.M., 2012b. Genome sizes for all genera of Cycadales. *Plant Biol.* 14, 253–256.
- Zonneveld, B.J.M., 2012c. The pine nut syndrome: a simple test for genome size of twelve pine nut producing trees links the bitter aftertaste to nuts of *Pinus armandii*. *Plant Syst. Evol.* 297, 201–206.