SYSTEMATICS AND PHYLOGENY

# Node ages, relationships, and phylogenomic incongruence in an ancient gymnosperm lineage – Phylogeny of *Ephedra* revisited

**Catarina Rydin,[1,2] Ruben Blokzijl,[1] Olle Thureborn[1] & Niklas Wikström[1,2]**

1 *Department of Ecology, Environment and Plant Sciences, Stockholm University, 106 91 Stockholm, Sweden*
2 *Bergius Foundation, The Royal Swedish Academy of Sciences, Box 50005, 104 05 Stockholm, Sweden*
Address for correspondence: *Catarina Rydin, Catarina.Rydin@su.se*

**DOI** https://doi.org/10.1002/tax.12493

**Abstract** Knowledge of relationships among plants has improved dramatically but for many groups, it still rests solely on information from the plastid genome. Consequently, only parts of the organisms' evolutionary history are revealed. For the ancient gymnosperms of *Ephedra* (Gnetales), previous conclusions were based on both plastid and nuclear ribosomal DNA data, but results were typically poorly resolved and supported, presumably because of information poverty in the utilized gene regions. With the aim of resolving phylogenetic questions using more data, we sequenced the plastid genome and the nuclear ribosomal cistron for 50 specimens of *Ephedra*, largely covering the diversity of the genus. Phylogeny and node ages were estimated using maximum likelihood and Bayesian methods. However, instead of clarifying a few remaining uncertainties, we were left with new questions and incongruent results. Age estimates of the crown group of *Ephedra* vary considerably depending on utilized software and specified tree prior. Furthermore, previous estimates of the phylogeny of *Ephedra* have largely reflected information from nrITS, despite utilization of plastid gene regions as well. With analyses based on the entire plastid genome, completely new results emerge. Earlier conclusions of deep divergences in the genus are not supported, and there are ample examples of phylogenetic incongruence. Our study overturns conclusions in previous work and highlights that we still know fairly little about evolution in the ancient *Ephedra* lineage. How many species of *Ephedra* are there, and how are they related? How old is the crown group? Many species appear affected by a history of hybridization/introgression and/or polyploidy, but other processes may result in similar patterns and reasons for the detected incongruences must be further analyzed, preferably using population-level sampling and low-copy nuclear data.

**Keywords** Cretaceous; cytonuclear discordance; fossils; Gnetales; Pliocene; phylogenomics

**Supporting Information** may be found online in the Supporting Information section at the end of the article.

## ■ INTRODUCTION

Phylogenetic relationships of Earth's organisms are quickly becoming better and better understood. Not so long ago this would have seemed impossible to accomplish, but in 2004, Donoghue & Cracraft (2004) wrote "we believe that it is now realistic to conceive of reconstructing the entire Tree of Life—eventually to include all of the living and extinct species". Now, well over a decade later, with accumulating achievements and new sequencing techniques, this goal is approaching faster than ever before. However, it is equally true that many phylogenetic questions have remained unresolved, or at least ambiguously answered, despite considerable effort. Well-known examples in the green clade of the Tree of Life are the sister group of land plants, relationships among liverworts, hornworts, and mosses, the position of Equisetophyta, and seed plant relationships (see, e.g., Burleigh & Mathews, 2007; Mathews, 2009; Ruhfel & al., 2014; Wickett & al., 2014; Shen & al., 2017; Cox, 2018; Puttick & al., 2018; Ran & al., 2018; Smith & Brown, 2018; de Sousa & al., 2019; Liu & al., 2019). In recent years, studies using genomic data have revealed that difficult phylogenetic questions often are associated with discordance among gene trees founded in biological processes such as incomplete lineage sorting, introgression and hybridization, and recurrent cycles of polyploidy (e.g., Jarvis & al., 2014; Wendel, 2015; Barker & al., 2016; Folk & al., 2017; Vargas & al., 2017; Cox, 2018; Morales-Briones & al., 2018; Pease & al., 2018; Puttick & al., 2018; Lee-Yaw & al., 2019; Liu & al., 2019; Qiao & al., 2019), and phylogenomic studies have substantially contributed to resolving some longstanding questions on land plant evolution.

However, for an average land plant genus, contemporary knowledge on interspecific relationships may still be based primarily on information from the plastid genome (Zimmer & Wen, 2012; Davis & al., 2014; Ruhfel & al., 2014; Rothfels & al., 2015; Rydin & al., 2017; Vargas & al., 2017; Johnson & al., 2019), and perhaps only from a handful of genes. While information from the plastome clearly has revolutionized

**Article history:** Received: 21 Aug 2020 | returned for (first) revision: 9 Oct 2020 | (last) revision received: 22 Jan 2021 | accepted: 5 Feb 2021 | published online: 4 May 2021 | **Associate Editor:** Mary E. Endress | 

plant systematics, a strong dependence on plastid data means a risk that complex evolutionary histories involving, for example, reticulations and/or polyploidy are undetected (e.g., Rieseberg & Soltis, 1991; Rothfels & al., 2013; Folk & al., 2017; Morales-Briones & al., 2018; Johnson & al., 2019; Lee-Yaw & al., 2019). Other factors such as ancestral gene polymorphism at different loci (caused by gene duplication) or at a single locus (allelic variation) may also cause problems in phylogenetic reconstruction (Fitch, 1970; Takahata & Nei, 1985). Deep coalescence (temporal mismatch between alleles uniting at a common ancestral gene and species uniting at a common ancestor) may result in a gene tree that differs from the species tree. The divergence time of a pair of allelic genes sampled from different species may be substantially greater than that of the speciation (Takahata & Nei, 1985), at least when effective population sizes are large and the time between speciation and the preceding unification of the genes is short (Maddison, 1997).

A solution to such problems is to utilize data from many loci (Rieseberg & Soltis, 1991), as is done in multigene studies and phylogenomic studies, particularly important in studies of closely related species/specimens (Takahata & Nei, 1985; Maddison, 1997). However, the value of multigene studies as a buffer against problems with conflicting gene trees and species trees is low if the utilized gene regions are thought to be linked, inherited as a unit (Rieseberg & Soltis, 1991). To test hypotheses of phylogeny based on plastid data, using information from the other two genomic compartments, is thus highly desirable. It is, however, equally important to base assessments of genomic discordance on sufficient amounts of data. Given the new time- and cost-efficient methods for DNA sequencing, it is now both possible and relevant to base phylogenetic studies on whole genome data, at least for the organellar genomes (e.g., Rydin & al., 2017; Vargas & al., 2017; Lee-Yaw & al., 2019).

Here we study the gymnosperm genus *Ephedra* L. of the Gnetales. The Gnetales (comprising the three extant genera *Ephedra*, *Gnetum* L. and *Welwitschia* Hook.f.) have a long, diverse and intriguing evolutionary history that extends at least to the Early Cretaceous (among many, Krassilov, 1982, 1986; Krassilov & Bugdaeva, 1982; Crane & Upchurch, 1987; Osborn & al., 1993; Crane, 1996; Mohr & Friis, 2000; Rydin & al., 2003, 2004, 2006b, 2010; Y. Yang & al., 2005, 2013, 2018, 2020; Friis & al., 2007, 2013, 2014, 2019; Rydin & Friis, 2010; X. Wang & Zheng, 2010; Ickert-Bond & Rydin, 2011; Norbäck Ivarsson, 2013; Rothwell & Stockey, 2013; Bolinder & al., 2015, 2016a,b; Hou & al., 2015; Rydin & Bolinder, 2015; Han & al., 2016; Rydin & Hoorn, 2016). Extant diversity is limited to some 100 species, most of them in *Ephedra* (Ickert-Bond & Renner, 2016). The order is nevertheless well-known to systematists because of its central role in the longstanding question of seed plant phylogeny (e.g., Arber & Parkin, 1908; Thompson, 1918; Crane, 1985; J.A. Doyle & Donoghue, 1986; Chase & al., 1993; Goremykin & al., 1996; Bowe & al., 2000; Chaw & al., 2000; Rydin & al., 2002; Burleigh & Mathews, 2004, 2007; Hilton & Bateman, 2006; Mathews, 2009; Zhong & al., 2010, 2011; Lee & al., 2011; Burleigh & al., 2012; Mathews & Kramer, 2012; Xi & al., 2013; Ruhfel & al., 2014; Coiro & al., 2018; Ran & al., 2018; Smith & Brown, 2018). The *Ephedra* lineage can be traced back to the Early Cretaceous, where fossils sharing unique features with the extant clade have been discovered from low paleolatitudes (Rydin & al., 2004, 2006a). These fossil seeds with *in situ* pollen (Rydin & al., 2004, 2006a) are exquisitely preserved and display anatomical, cellular and functional (pollen germination) details otherwise only found in living species of *Ephedra*. They could thus potentially have been produced by members of the extant clade (i.e., the *Ephedra* crown). However, analyses of their reproductive morphology indicated that they are stem relatives to the extant clade (Rydin & al., 2010), meaning that the *Ephedra* crown could be much younger. Analyses of divergence times based on molecular data have resulted in relatively young estimated mean ages of the *Ephedra* crown of about 30 Myr (Ickert-Bond & al., 2009), corresponding to the earliest Oligocene or slightly younger (i.e., 32–8 Ma; Huang & Price, 2003), apparently matching the indications (Rydin & al., 2010) based on morphological data.

A few years later, new discoveries once again challenged the age estimate of the *Ephedra* crown. Information from fossil and extant pollen (Norbäck Ivarsson, 2013; Bolinder & al., 2016b) showed two distinct pollen types in *Ephedra*, an ancestral and a derived type, of which the derived type is not present among the early-diverging lineages of the *Ephedra* crown group (Bolinder & al., 2016b). In the fossil record, the ancestral pollen type is documented since the late Paleozoic (Wilson, 1959, 1962; Z. Wang, 2004). The derived pollen type is much younger but still ancient; the oldest documentation we have encountered is from the late Mesozoic (further explained in the Discussion). This information would for the first time permit a dating analysis of *Ephedra* with a calibration point present within the clade, and using this new knowledge on *Ephedra* pollen would push back the resulting age of the *Ephedra* crown to about 100 Myr.

Phylogenetic relationships in the crown group of *Ephedra* have been repeatedly addressed using molecular data and modern analytical tools (Ickert-Bond & Wojciechowski, 2004; Rydin & al., 2004; Huang & al., 2005; Rydin & Korall, 2009; Kakiuchi & al., 2011; Loera & al., 2012, 2015; Thureborn, 2014). Results have, however, typically displayed relationships with little or no statistic support, and finding the root of the phylogeny has proven particularly challenging. A striking dearth of information in most utilized gene regions and a long evolutionary distance to *Ephedra*'s closest living relatives, hamper phylogenetic inference. The most comprehensive studies to date, including representative species from all major subgroups of the genus, indicate geographical grouping of species (Ickert-Bond & Wojciechowski, 2004; Rydin & al., 2004; Huang & al., 2005; Rydin & Korall, 2009; Thureborn, 2014; Ickert-Bond & al., 2020) and post-Pangean long-distance dispersal as responsible for the observed phylogenetic pattern, although the lack of support for some of the deepest splits in

the genus has left evolutionary interpretations uncertain. Species-level relationships are partly resolved and supported in these studies, but there is poor resolution among at least some species in all the major clades.

The original aim of the present study was to (1) resolve phylogenetic questions about *Ephedra* that remain ambiguously answered using large amounts of data produced by next-generation sequencing techniques, and (2) conduct a dating analysis of the clade using the new knowledge on fossil *Ephedra* pollen for calibration of the results to absolute times. Among important questions to address were the age of the *Ephedra* crown group, monophyly/polyphyly of the Asian clade (sensu Rydin & Korall, 2009), and the sister relationship of *E. foeminea* to the remaining species of the genus, a result indicated in most recent studies although with poor statistic support. It is vital to achieve stable results on these matters because the phylogeny constitutes an important framework for almost any further work on the evolutionary history, biogeography, ecology, and reproductive biology of the Gnetales. As the work with the present study progressed, it became increasingly clear that we would not reach our initial goal. Instead, the present study reveals some surprising findings that highlight challenges to phylogeny reconstruction and node age estimation, findings we believe are of relevance not only to students of the Gnetales, but to scientists generally interested in the evolution of plants.

## ■ MATERIALS AND METHODS

**Taxon sampling and data production. —** Fifty *Ephedra* specimens, representing the vast majority of the species diversity in *Ephedra* and all major clades as identified in previous work (Ickert-Bond & Wojciechowski, 2004; Rydin & al., 2004; Rydin & Korall, 2009) were selected for the present study (Appendix 1 and suppl. Table S1). The plastid genome has already been published for specimen KB543 of *E. foeminea* (NC_029347; Hou & al., 2017), but its nuclear ribosomal DNA was newly produced for the present study. In addition, the published plastid genome of *E. equisetina* (NC_011954; C.S. Wu & al., 2009) was included. Analyses of plastid data are thus based on 51 ingroup samples, whereas analyses of nuclear data are based on 50 ingroup samples. Outgroup information from up to 10 representatives of the remaining seed plants were utilized in some analyses for rooting purposes (for details, see below). DNA was extracted from herbarium, live, or silica-dried specimens using a cetyltrimethylamonium bromide CTAB protocol (J.J. Doyle & Doyle, 1987; J.J. Doyle, 1991). Extracted DNA was cleaned using the QIAquick PCR cleaning kit from Qiagen (Qiagen, Hilden, Germany) following the protocol specified by the manufacturer. High-throughput sequencing was carried out at Science for Life Laboratory (SciLifeLab, Stockholm, Sweden) following the manufacturer's instructions for the Illumina HiSeq2500 platform (Illumina, San Diego, California, U.S.A.). Pair-end runs with 350-bp insert size fragments and 2 × 125 bp read lengths were performed. Each sample was multiplexed with 93 other samples and run in three different lanes. Library preparation at the SciLifeLab was done using the ThruPLEX DNA-seq library preparation kit from Rubicon (Rubicon Genomics, Ann Arbor, Michigan, U.S.A.). Demultiplexing and conversion was conducted using bcl2fastq v.2.17 from the CASAVA software suite (Illumina).

**Plastid sequence assembly. —** All sequences were assembled using a reference-based approach. Plastid sequences were isolated from the original reads through a BLAT (BLAST-like alignment tool v.36; Kent, 2002) search of forward and reverse reads against an initial database of three previously published plastid genomes of taxa from the Gnetales. The initial database included the plastid genomes of *Ephedra equisetina* (NC_011954; C.S. Wu & al., 2009), *Gnetum gnemon* (NC_026301; Zhu & al., 2016), and *Welwitschia mirabilis* (NC_010654; McCoy & al., 2008). Forward and reverse reads were both saved if either showed at least 70% similarity to any of the reference genomes. Following the BLAT search, reads were extracted from the original fastq data files using pullseq v.1.0.1 (github.com/bcthomas/pullseq) into new forward and reverse fastq data files representing a "plastid" subset. *De novo* assembly of the plastid subset was performed for each taxon using ABySS v.1.5.2 (Simpson & al., 2009) and eight different k-mer lengths (49, 55, 61, 67, 73, 85, 91, 97). Generated contigs were pooled and mapped onto the *E. equisetina* (NC_011954; C.S. Wu & al., 2009) reference genome using bwa v.0.7.5a-r405 (Li & Durbin, 2009). This resulted in complete or near-complete draft genomes. All original reads were subsequently mapped onto the draft genomes using bwa v.0.7.5a-r405 allowing sequencing depths to be evaluated and unfinished gaps to be filled. Average sequencing depth was estimated for each genome in this final mapping step using the depth option in SAMtools v.1.7 (Li & al., 2009). Assemblies were reviewed and edited using gap5 from the Staden Package v.2.0.0b10 (Staden, 1996; Staden & al., 2000). Protein coding (CDS), transfer RNA (tRNA), and ribosomal DNA (rDNA) genes were annotated using Sequin v.15.50 (available at www.ncbi.nlm.nih.gov/Sequin/) by transferring the annotation of *E. foeminea* (NC_029347; Hou & al., 2017).

**Assembly of nuclear rDNA data. —** Nuclear rDNA data were assembled in much the same way. Unlike the plastid genome, no previously published reference sequence from the Gnetales was available including a complete or near-complete rDNA cistron. An initial reference sequence of *Ephedra americana* was therefore generated by using the near-complete rDNA cistron from *Asclepias syriaca* (JF312046; Straub & al., 2011) as the initial reference. The *Asclepias syriaca* reference sequence comprised a total of 7261 bp including ETS (bp 1–764), 18S rDNA (bp 765–2572), ITS-1 (bp 2573–2807), 5.8S rDNA (bp 2808–2971), ITS-2 (bp 2972–3210), 28S rDNA (bp 3211–6603), and an NTS region (bp 6604–7261). Contigs generated by the *de novo* assembly were mapped onto the *Asclepias* reference yielding an initial draft sequence of the *E. americana* rDNA cistron. Original reads were subsequently mapped in an iterative process onto the

generated draft sequence using bwa v.0.7.5a-r405 (Li & Durbin, 2009) allowing for unfinished gaps to be filled and for the 5′ ETS and 3′ NTS regions of the cistron to be extended. The final near-complete rDNA cistron sequence of *E. americana* comprised a total of 7950 bp including ETS (bp 1–792), 18S rDNA (bp 793–2605), ITS-1 (bp 2606–3726), 5.8S rDNA (bp 3727–3888), ITS-2 (bp 3889–4135), 28S rDNA (bp 4136–7552), and an NTS region (bp 7553–7950). The generated *E. americana* rDNA cistron sequence was subsequently used as the reference sequence in the BLAT searches of remaining specimens. Average sequencing depth for each assembly was estimated during the final mapping step in the same way as done for the plastid genomes. All assemblies were reviewed and edited using gap5 from the Staden Package v.2.0.0b10 (Staden, 1996; Staden & al., 2000), and sequences were annotated using Sequin v.15.50.

**Alignment. —** Protein coding (CDS), tRNA, plastid rDNA, intron, and intergenic spacer (IGS) gene regions were individually extracted from the annotated GenBank files of each assembled plastid genome using an in-house Python script built on Biopython v.1.63 (Cock & al., 2009). Individual gene regions were aligned using MUSCLE v.3.8.31 (Edgar, 2004) and concatenated into a CDS set comprising 67 gene regions (62,418 characters), a tRNA + rDNA set (tRNA 2226 characters, rDNA 4808 characters), an intron set (6578 characters), and an IGS set (17,734 characters). The final plastid dataset included 51 ingroup (*Ephedra*) samples and 10 outgroup (*Welwitschia*, four *Gnetum*, and five non-angiosperm seed plants) samples (Appendix 1, suppl. Table S1), and 93,764 aligned characters (suppl. Appendix S1).

Nuclear rDNA regions were extracted from the annotated GenBank files of the rDNA cistrons in a corresponding way. Individual regions were aligned using MUSCLE v.3.8.31 (Edgar, 2004) and concatenated into an 18S + 5.8S rDNA set (18S: 1818 characters, 5.8S: 162 characters), a 26S rDNA set (3461 characters), an ITS-1 + ITS-2 set (ITS-1: 1146 characters, ITS-2: 248 characters), and an ETS set (797 characters). The final nuclear rDNA dataset included 50 ingroup (*Ephedra*) samples and 5 outgroup (*Welwitschia*, four *Gnetum*) samples (Appendix 1, suppl. Table S1), and 7632 aligned characters (suppl. Appendix S2). The variable regions of ITS-1, ITS-2 and ETS were not possible to align between ingroup and outgroup and were all scored as missing in the outgroup.

**Phylogenetic analyses. —** Phylogenetic analyses were conducted with Markov chain Monte Carlo methods (Z. Yang & Rannala, 1997; Larget & Simon, 1999; Mau & al., 1999) in MrBayes v.3.2.6 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003), and BEAST v.1.8.4 (Drummond & al., 2012). The complete partitioned and annotated plastid and nuclear rDNA datasets are available in Nexus format (suppl. Appendices S1, S2). In addition to this analytical work (specified below), we further explored the information in the datasets by performing maximum likelihood analyses, as well as Bayesian analyses where the respective datasets were partitioned according to evolutionary rate (Cummins & McInerney, 2011; Rota & al., 2017). For details on maximum likelihood analyses and Bayesian analyses employing an alternative partitioning scheme, see supplementary Appendix S3.

*Analyses of plastid data.* – The CDS was split into three separate partitions based on codon position in all nucleotide-based analyses. The GTR+I+Γ substitution model was used for the three protein-coding partitions CDS_1pos, CDS_2pos, CDS_3pos, and for the IGS partition. The GTR+Γ substitution model was used for the tRNA + rDNA partition, and the HKY+I+Γ substitution model for the intron partition. Substitution models were selected based on the corrected Akaike information criterion (AICc) as calculated using the programs MrAIC v.1.4.6 (Nylander, 2004) and PHYML v.3.0 (Guindon & al., 2010), and were unlinked across all partitions. Two sets of analyses were run in MrBayes: (1) non-clock analyses including 61 terminals (51 ingroup and 10 outgroup samples) with results rooted on the outgroup; (2) relaxed-clock analyses using the independent gamma rate (IGR; LePage & al., 2007) relaxed-clock model as implemented in MrBayes, including the ingroup only and rooted following results of the analysis. The purpose of this analysis was to infer topology and rooting position, not absolute node ages, and no calibration was used.

MrBayes was run for 20 million generations in the non-clock analyses and for 50 million generations in the relaxed-clock analyses. Two independent runs, each with four chains and heating parameters set to default values, were conducted for all analyses. Trees and parameter estimates were sampled to yield a total of 10,000 trees, and parameter estimates in each of the two runs and posterior probability values were calculated after discarding the first 50% of the trees and parameters as burn-in. This was well beyond the burn-in phase of the chains based on the potential scale reduction factor (PSRF) convergence diagnostic (Gelman & Rubin, 1992).

*Analyses of nuclear rDNA data.* – Based on the AICc criterion, the GTR+I substitution model was selected for the 18S + 5.8S partition, and GTR+I+Γ for the other three. Two sets of analyses were conducted: (1) non-clock analyses including 55 terminals (50 ingroup and 5 outgroup samples) with results rooted on the outgroup; (2) relaxed-clock analyses including the ingroup only and rooted following results of the analyses. Non-clock and relaxed-clock runs were conducted and summarized as described for analyses of plastid data.

**Node ages. —** To infer the age of the crown group of *Ephedra*, a set of calibrated analyses were run in BEAST and MrBayes. The analyses were calibrated to absolute time using the fossil taxon and sister to *Welwitschia*, *Cratonia cotyledon* (Rydin & al., 2003), discovered from the lower part of the Crato Formation of the Araripe basin of northern Gondwana. The formation is assigned to late Aptian age (e.g., Heimhofer & Hochuli, 2010), and we used a normally distributed prior probability that the split between *Gnetum* and *Welwitschia* was at least at 113 Ma, corresponding to the upper boundary of the Aptian (Gradstein & al., 2012). Following Ickert-Bond & al. (2009), standard deviation of the distribution was set to 2.5 Ma yielding a 95% confidence interval of 10 Ma (123–113 Ma). Two sets of analyses were run: one using a Yule (Yule, 1925) and a second using a birth-death (Kendall, 1948) tree prior.

A Yule tree prior is not explicitly included as an option in MrBayes and was emulated under the birth-death prior by setting the extinction rate to zero.

**Silent substitution rates. —** Using the CDS dataset and a constrained topology consistent with the results from the relaxed-clock analyses (see above), maximum likelihood estimates of branch-specific silent substitution ($d_s$) branch lengths were estimated with HyPhy v.2.2.6 (Pond & al., 2005) under the MG94W9 codon model (Muse & Gaut, 1994).

## ■ RESULTS

**Assembled plastids. —** Plastid sequences were successfully assembled from 49 accessions of *Ephedra*, which were analyzed together with existing plastid genomes of the ingroup: *E. equisetina* (NC_011954; C.S. Wu & al., 2009) and *E. foeminea* (NC_029347; Hou & al., 2017). Assembled sequences comprise the large-single copy region (LSC), the inverted repeat region (IR) and the small single-copy region (SSC), and range in length from 88,898 bp in *E. alata* to 91,056 in *E. californica*. The total number of sequenced fragments varied across samples from $4.8 \times 10^6$ fragments in *E. viridis* (CR262) to $11.7 \times 10^6$ fragments in *E. trifurca* with an average number of $8.8 \times 10^6$. Average sequencing depth varied from 46 in *E. aphylla* (AK154) to 1647 in *E. tweediana* with an average depth of 392. Assembled sequences are deposited in GenBank. For voucher information and accessions, see Appendix 1. Extended information, also including taxon distribution and area and year of collection, is available in supplementary Table S1.

**Assembled nuclear rDNA cistrons. —** Partial ribosomal DNA (rDNA) cistrons were successfully assembled from 50 accessions of *Ephedra*, 4 accessions of *Gnetum* and 1 accession of *Welwitschia*. Assembled sequences comprise external transcribed spacer region (ETS), small subunit ribosomal DNA (18S rDNA), internal transcribed spacer 1 (ITS-1), 5.8S subunit ribosomal DNA (5.8S rDNA), internal transcribed spacer 2 (ITS-2), large subunit ribosomal DNA (26S rDNA), and non-transcribed spacer (NTS). Total lengths of the assembled sequences range from 6326 bp in *Welwitschia mirabilis* to 8091 bp in *E. viridis* (CR262). Average sequencing depth varied from 76 in *E. aphylla* (AK154) to 1118 in *E. likiangensis* with an average depth of 438. Assembled sequences are deposited in GenBank. For voucher information and accessions, see Appendix 1. Extended information, also including taxon distribution and area and year of collection, is available in supplementary Table S1.

**Non-clock analysis of plastid data. —** In the non-clock analysis of plastid data (Fig. 1), *Ephedra* is divided into three major clades supported by a Bayesian posterior probability (BPP) of 1.00: a Mediterranean clade, a New World clade, and an Asian clade. It should be noted though, that some species have distributions that extend beyond these geographical areas. For example, *E. distachya* is clearly most closely related to Asian species, but has a distribution range that extends from eastern central temperate Asia to southern Europe. The Mediterranean/western temperate Asian species *E. major* falls outside of these clades and is resolved as sister to all other species of *Ephedra* (BPP = 0.94). The Mediterranean clade and the New World clade are sisters (BPP = 0.96), but the position of the root of the phylogeny is not well supported. *Ephedra foeminea* is sister to the remaining species of the Mediterranean clade (BPP = 1). The New World clade comprises two clades of North American taxa (both BPP = 1), which are successive sisters to a clade of South American taxa (BPP = 1). The North American species *E. pedunculata*, *E. antisyphilitica*, and *E. coryi* are sister to the South American clade. *Ephedra milleri* is nested in the Asian clade, sister to a clade comprising *E. distachya* + *E. transitoria* (BPP = 1) in all our analyses of plastid data. *Ephedra minuta* is sister to *E. gerardiana* (BPP = 1). *Ephedra likiangensis* is included in a clade also comprising *E. monosperma*, *E. fedtschenkoi* and the specimen of *E. equisetina* downloaded from GenBank (BPP = 1).

**Relaxed-clock analysis of plastid data. —** The relaxed-clock analysis of plastid data (Fig. 2) results in a conflicting position of the *Ephedra* root with the Mediterranean clade (BPP = 0.97) now resolved as sister to the remaining species of *Ephedra* (BPP = 1). Results are otherwise entirely congruent with those from the non-clock analysis. *Ephedra foeminea* is sister to the remaining species of the Mediterranean clade (BPP = 1). Species not included in the Mediterranean clade comprise two sister groups: the New World clade (BPP = 1) and a clade where *E. major* is resolved as sister to the Asian clade (BPP = 1). These last two relationships result as a direct consequence of the altered position of the root compared to the non-clock analysis. Within the New World clade, North American species form two groups (both BPP = 1). One of them, comprising *E. pedunculata*, *E. antisyphilitica*, and *E. coryi*, is sister to a South American clade (BPP = 1). *Ephedra milleri* is nested in the Asian clade, sister to *E. distachya* and *E. transitoria* (BPP = 1). *Ephedra minuta* is sister to *E. gerardiana* (BPP = 1). *Ephedra likiangensis* is included in a clade also comprising *E. monosperma*, *E. fedtschenkoi* and the specimen of *E. equisetina* downloaded from GenBank (BPP = 1).

**Non-clock analysis of nuclear rDNA data. —** Results of the non-clock analysis of nuclear rDNA data (suppl. Fig. S1) are mostly the same as those retrieved from the relaxed-clock analysis of nuclear rDNA data (see below), but typically slightly less well supported.

**Relaxed-clock analysis of nuclear rDNA data. —** In the relaxed-clock analysis of nuclear rDNA (Fig. 3), results among major clades are generally poorly supported and the deepest splits in *Ephedra* are not resolved. The Mediterranean species except *E. foliata* and *E. ciliata* form a group (BPP = 1), but the Arabian Peninsula species *E. milleri* is nested in this clade, sister to *E. alata* (BPP = 1). The New World clade is well supported (BPP = 1), with *E. pedunculata* sister to remaining species (BPP = 1). South American species are monophyletic (BPP = 1), as are North American species excluding *E. pedunculata* (BPP = 1). Asian species are resolved in three clades, one comprising *E. minuta* and *E. likiangensis* (BPP = 1), a second


**Fig. 1.** Phylogram resulting from the non-clock analysis of plastid data (including outgroup taxa). The tree was rooted on *Ginkgo*. **A**, Overview, showing branch lengths and relationships among genera; **B**, Same tree but branch lengths upscaled more than 100 times to show detailed results within *Ephedra*. Nodes indicated by a black dot are well supported and have a BPP equal to or greater than 0.95 (Alfaro & al., 2003; Erixon & al., 2003). Support values are 1.00 unless otherwise stated in the figure. Taxon distribution (see suppl. Table S1 for details) and the country where the accession was collected (in parenthesis) are indicated to the right of the taxon name. Terminals in red are discussed in the text.

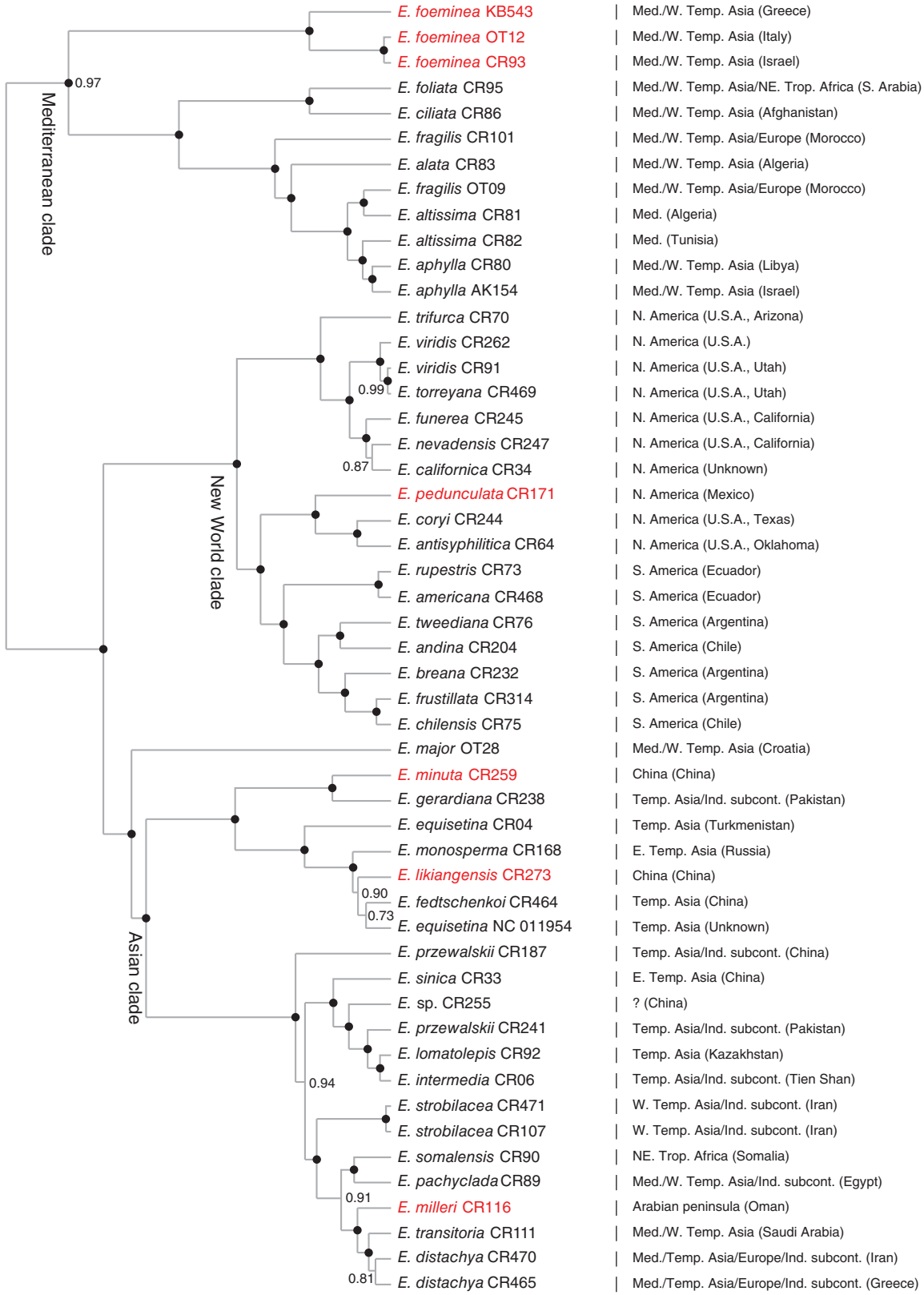


**Fig. 2.** Chronogram (relative times only) resulting from the relaxed-clock analysis of plastid data of the ingroup (*Ephedra*). The analysis was conducted in MrBayes using the independent gamma rate (IGR; LePage & al., 2007) relaxed-clock model, and the tree was rooted as a result of the analysis. Nodes indicated by a black dot are well supported and have a BPP equal to or greater than 0.95 (Alfaro & al., 2003; Erixon & al., 2003). Support values are 1.00 unless otherwise stated in the figure. Taxon distribution (see suppl. Table S1 for details) and the country where the accession was collected (in parenthesis) are indicated to the right of the taxon name. Terminals in red are discussed in the text.

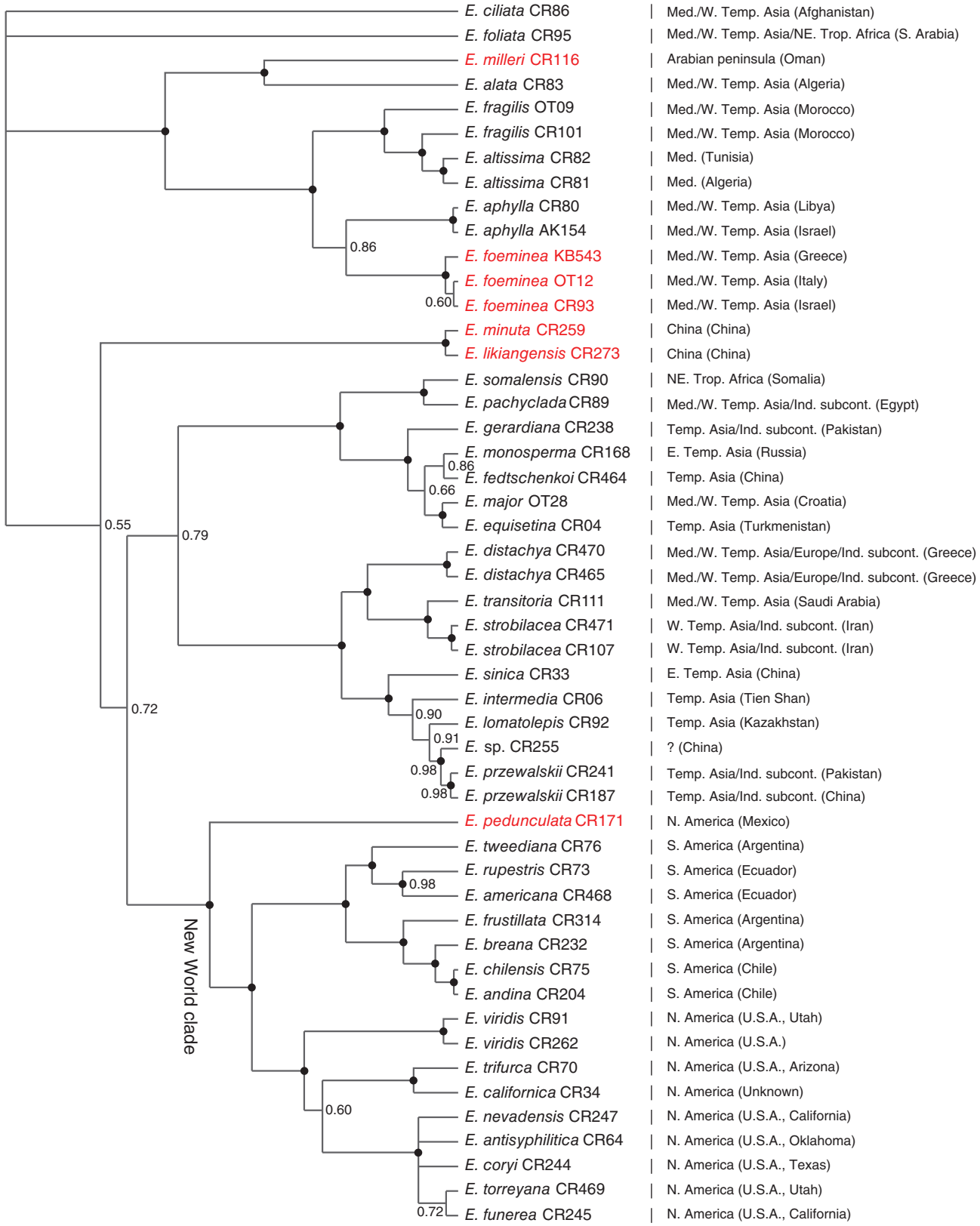


**Fig. 3.** Chronogram (relative times only) resulting from the relaxed-clock analysis of nuclear rDNA data of the ingroup (*Ephedra*). The analysis was conducted in MrBayes using the independent gamma rate (IGR; LePage & al., 2007) relaxed-clock model, and the tree was rooted as a result of the analysis. Nodes indicated by a black dot are well supported and have a BPP equal to or greater than 0.95 (Alfaro & al., 2003; Erixon & al., 2003). Support values are 1.00 unless otherwise stated in the figure. Taxon distribution (see suppl. Table S1 for details) and the country where the accession was collected (in parenthesis) are indicated to the right of the taxon name. Terminals in red are discussed in the text.

comprising *E. somalensis*, *E. pachyclada*, *E. gerardiana*, *E. monosperma*, *E. fedtschenkoi*, *E. major* and *E. equisetina* (BPP = 1), and a third comprising *E. distachya*, *E. transitoria*, *E. strobilacea*, *E. sinica*, *E. intermedia*, *E. lomatolepis* and *E. przewalskii* (BPP = 1). Relationships among these clades are not supported.

**Maximum likelihood analyses and Bayesian analyses with data partitioned according to evolutionary rate. —** None of these exercises had any impact on the ingroup topology as reported above, and will not be further discussed. For plastid data, the position of the root of the phylogeny may differ among results from analyses using different partitioning schemes, but differences are unsupported. Details and results of the ML analyses and the Bayesian analyses with data partitioned according to evolutionary rate are available in supplementary Appendix S3.

**Node ages and silent substitution rates. —** The analysis using an uncorrelated lognormal clock and a birth-death tree prior, run in BEAST (Fig. 4A), returned a 95% highest posterior density of the *Ephedra* crown of 5–2 Ma. The analysis using the IGR relaxed-clock model and a birth-death tree prior, run in MrBayes (Fig. 4B), returned a 95% highest posterior density of the *Ephedra* crown of 3–2 Ma. The analysis using an uncorrelated lognormal clock and a pure birth (Yule) tree prior, run in BEAST (Fig. 4C), returned a 95% highest posterior density of the *Ephedra* crown of 74–13 Ma. The analysis using the IGR relaxed-clock model and a pure birth (Yule) tree prior, run in MrBayes (Fig. 4D), returned a 95% highest posterior density of the *Ephedra* crown of 4–3 Ma.

Assuming an age of the Gnetales crown of 150 Myr (125 Myr is the minimum age based on fossils; Rydin & al., 2004, 2006a) and an age of the *Ephedra* crown between 10 and 100 Myr will yield mean absolute silent substitution rates between 0.8 and 0.08 silent substitutions per site and billion years (SSB) for the *Ephedra* crown, and between 3.5 and 9.9 SSB for the *Ephedra* stem (Table 1). More details are available in supplementary Appendix S4.

## ■ DISCUSSION

**Cytonuclear discordance. —** Previous studies aiming to resolve phylogenetic relationships in *Ephedra* have typically used information from nuclear ribosomal DNA as well as from the plastid genome. Rydin & al. (2004) used the plastid gene regions *rbcL* and *rps4* and the nuclear ribosomal regions 18S, 26S, and nrITS, Ickert-Bond & Wojciechowski (2004) used plastid *rps4* and nrITS1, and Rydin & Korall (2009) used plastid *rbcL*, *rps4*, *rpl16* intron, *trnS*$^{UGA}$-*trnfM*$^{CAU}$ intergenic spacer, and the nuclear 18S, 26S, and nrITS. In those studies, no supported conflicts between results from the two genomic compartments were detected. However, in the light of results of the present study, it becomes clear that the previously used plastid gene regions did not provide a significant amount of information on the phylogeny of *Ephedra*. Results based on the few plastid regions included were poorly supported and partly unresolved, and did not reveal the consensus information of the plastid genome. Instead, results of combined analyses conducted in those studies largely reflected the phylogenetic signal in nrITS. Thus, in contrast with most plant groups, in which phylogenetic knowledge often is based (mostly) on plastid data, previous knowledge on the phylogeny of *Ephedra* is based almost exclusively on information from the nuclear ribosomal DNA (i.e., nrITS). Our results of the nuclear rDNA cistron are in general consistent with conclusions in previous studies (e.g., Rydin & Korall, 2009), whereas many of our results based on plastid data differ from what has previously been assumed on the phylogeny of *Ephedra*.

Here, based on substantially larger amounts of data, considerable incongruence between results retrieved from the plastid genome (Figs. 1, 2) and the nuclear rDNA cistron (Fig. 3, suppl. Fig. S1) is revealed. There is also a rooting conflict

**Table 1.** Maximum likelihood estimates of absolute silent substitution rates per site and billion years (SSB).

| *Ephedra* crown group age (Ma) | SSB of *Ephedra* crown group: min–max (average) | SSB of *Ephedra* stem lineage |
|---|---|---|
| 10 | 0.38–1.05 (0.80) | 3.52 |
| 20 | 0.19–1.52 (0.40) | 3.80 |
| 30 | 0.13–0.35 (0.27) | 4.11 |
| 40 | 0.10–0.26 (0.20) | 4.49 |
| 50 | 0.08–0.21 (0.16) | 4.93 |
| 60 | 0.06–0.17 (0.13) | 5.48 |
| 70 | 0.05–0.15 (0.11) | 6.17 |
| 80 | 0.05–0.13 (0.10) | 7.05 |
| 90 | 0.04–0.12 (0.09) | 8.22 |
| 100 | 0.04–0.10 (0.08) | 9.87 |

Estimates are given for the *Ephedra* crown group and the *Ephedra* stem lineage at alternative ages for the *Ephedra* crown group and an age for the crown group of the Gnetales of 150 Ma. Branch lengths were estimated in HyPhy v.2.2.6 (Pond & al., 2005) using the plastid CDS dataset, a constrained topology where the topology and rooting of *Ephedra* were consistent with results from the relaxed-clock analyses, and the MG94W9 codon model (Muse & Gaut, 1994). See Supporting Information found online (suppl. Appendix S4) for all maximum likelihood estimates of silent substitution branch lengths in Gnetales.

**Fig. 4.** Chronograms showing estimated absolute ages of the crown group of *Ephedra* based on plastid data and a fossil-based constraint (see text for details). Node heights are median ages. Confidence intervals (95% highest posterior density) of node ages are indicated as bars and in absolute numbers for major clades. Nodes indicated by a dot have a BPP of 1.00. **A**, Analysis conducted in BEAST (uncorrelated lognormal relaxed-clock model) and a birth-death tree prior; **B**, Analysis conducted in MrBayes using the independent gamma rate relaxed-clock model (IGR) and a birth-death tree prior; **C**, Analysis conducted in BEAST (uncorrelated lognormal relaxed-clock model) and a pure birth (Yule) tree prior. **D**, Analysis conducted in MrBayes using the independent gamma rate relaxed-clock model (IGR) and a pure birth (Yule) tree prior.

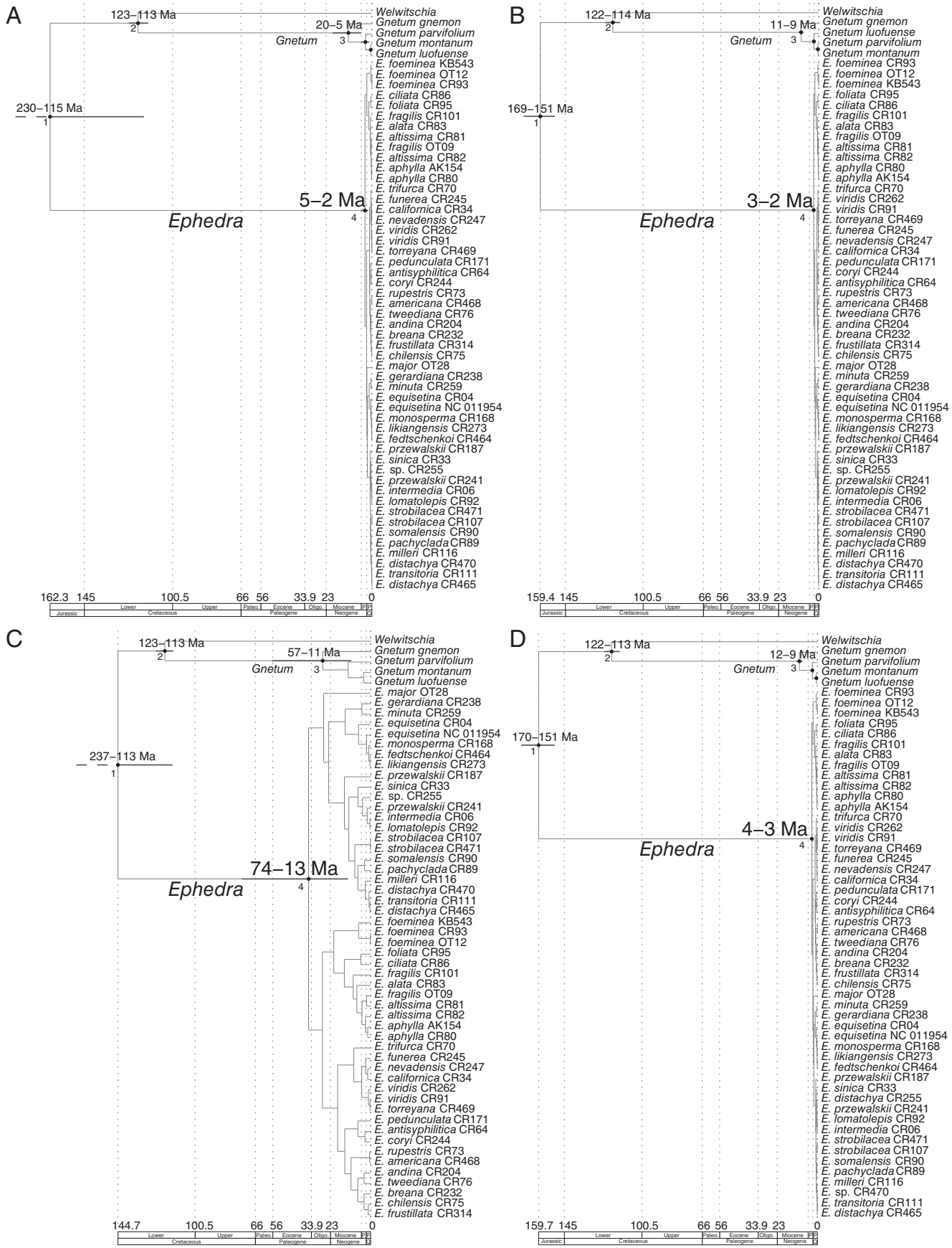


**Fig. 4.** Caption on previous page.

between results retrieved from non-clock and relaxed-clock analyses, where the non-clock analysis of plastid data places the root of *Ephedra* on *E. major* (Fig. 1), whereas relaxed-clock analyses of the same data support a root with Mediterranean taxa sister to the rest of the genus (Fig. 2). Concerning relationships among major subgroups of *Ephedra*, a single result is consistent in all our analyses: New World taxa are always monophyletic. Otherwise, relationships among the major geographical groups are somewhat uncertain and/or conflicting, partly resulting from the rooting conflict between the results from non-clock and relaxed-clock analyses of plastid data.

Reasons for the cytonuclear discordance detected here needs additional research. Data from a large set of presumably independently evolving low-copy nuclear genes could be used as a basis for a species tree estimate, to be contrasted against the two "gene trees" (comprising presumably clonally evolving gene regions) we present here. Results of the present study indicate, however, that historical speciation in *Ephedra* has been affected by processes of reticulation. Many species show supported but highly incongruent phylogenetic positions in results of the analyses of plastid data compared to those based on nuclear data. Even some deep divergences may have resulted from ancient reticulation. Hybridization, often in association to polyploidy, has repeatedly been suggested to be an important driving force in evolution (e.g., Rieseberg & Soltis, 1991; Mallet, 2007; Soltis & Soltis, 2009). Polyploidy has recently been shown to be exceptionally common in *Ephedra* (Ickert-Bond & al., 2020), and it is possible that the incongruent phylogenetic patterns documented here indicate that many of these polyploids are allopolyploids. However, other biological processes can result in similar phylogenetic patterns (Folk & al., 2017, 2018; Vargas & al., 2017; Lee-Yaw & al., 2019). Future studies of *Ephedra*, aiming to distinguish polyploidy, hybridization and introgression from processes such as incomplete lineage sorting, would benefit from population-level sampling and low-copy genomic data, and a combination of concatenation-based and coalescent-based analytical approaches.

**The age of the crown group of *Ephedra*. —** The age of the *Ephedra* crown is another difficult evolutionary question. Our results range from a crown group age of a few million years to more than 30 million years (with confidence limits extending as far back in time as 74 million years), depending on software utilized and selected tree prior (Fig. 4; further discussed below). The question is interesting because these gymnosperms have the potential to be of truly ancient origin compared to most angiosperm genera, and the answer has decisive impact on interpretations of evolutionary events in the group. The lineage (stem group) is clearly documented in fossils from the Early Cretaceous (i.e., Rydin & al., 2004, 2006a) that share unique features with the living species. The extant clade (the crown group) has been estimated to be much younger, based on morphology (Rydin & al., 2010) as well as dating analyses using molecular data (Huang & Price, 2003; Ickert-Bond & al., 2009; Loera & al., 2015).

However, our results show that previous estimates of the age of the *Ephedra* crown can only be reproduced in analyses using a Yule tree prior as implemented in BEAST (see Fig. 4). Analyses using a birth-death tree prior return a very young (Pliocene-Pleistocene) age of the *Ephedra* crown in all our analyses. Furthermore, using a pure birth (Yule) tree prior yields incompatible age estimates depending on analytical approach. When run in MrBayes (using the IGR relaxed-clock model), the Yule tree prior provides largely the same Pliocene age of the *Ephedra* crown as does the analyses using a birth-death prior; however, when the exact same dataset is run with the Yule tree prior as implemented in BEAST (using the uncorrelated lognormal relaxed-clock model), the *Ephedra* crown is estimated to be much older, with a confidence interval ranging from the Late Cretaceous to the mid Miocene (74–13 Ma). The latter result (based on analysis in BEAST using a pure birth tree prior) largely corresponds with that reported by Ickert-Bond & al. (2009), who also used a Yule tree prior in their analyses as this was the only option implemented in BEAST at the time.

We can offer no explanation for the inconsistent results found using a pure birth (Yule) prior as implemented in BEAST versus MrBayes. The Yule prior would, in general, be expected to yield an older age for the *Ephedra* crown compared to the birth-death prior because the Yule prior assumes a constant birth rate and zero extinction resulting in nodes being more evenly spread over the tree. Why this theoretical expectation seems to be met in our analyses using BEAST and the uncorrelated lognormal relaxed-clock model (also shown for cycads by Condamine & al., 2015), but not using MrBayes and the independent gamma rate relaxed-clock model, we do not know. Condamine & al. (2015) suggest the use of Bayes factors for determination of the relative fit to data of different evolutionary models, an approach we generally agree on. For *Ephedra*, however, stepping stone analyses are difficult to run to completion, probably because of the unusual information composition in the data. There is on average very little genetic variation among species of *Ephedra*, both in the plastome and the nuclear ribosomal DNA. Consequently, estimated silent substitution rates are extremely low, particularly if the *Ephedra* crown is estimated to be of Cretaceous origin (Table 1, suppl. Appendix S4). Assuming an age of the *Ephedra* crown of 100 Myr yields a mean absolute silent substitution rate of 0.08 (0.04–0.10) SSB in the crown (Table 1), which is much lower than rates estimated for the plastome of angiosperms (Richardson & al., 2013). Even assuming a much younger crown group of *Ephedra*, of 10 Myr, will yield a low substitution rate of 0.8 (0.38–1.05) SSB (Table 1). Furthermore, these very low substitution rates estimated for the *Ephedra* crown group stand in sharp contrast to those estimated for the stem, which are much greater, ranging between approximately 3.5 and 10 SSB (depending on the assumed respective ages of the Gnetales crown and the *Ephedra* crown; Table 1). It seems questionable whether analytical tools used to estimate node ages can handle the extreme rate shifts that apparently has occurred during the evolutionary history of *Ephedra* (see also discussion in Rydin & al., 2006a, and references therein). The Bayesian random local clocks model (Drummond & Suchard,

2010) was designed to approach problems of rate variation among lineages, but difficulties in reaching analytical convergence prevented us from yielding results based on this model.

Considering the deviating and uncertain analytical results concerning a clade like *Ephedra*, with a very long evolutionary history, other information should be taken into consideration as well, for example indications provided by the tree topology, biogeography, morphology and the fossil record. Fossils from the Early Cretaceous document that many vegetative and reproductive morphological features of modern-day *Ephedra* were already present around 125 million years ago (e.g., Rydin & al., 2004, 2006a; Bolinder & al., 2016b; Han & al., 2016). The Early Cretaceous held a substantial global diversity of plants with an indeterminate but clear affinity to ephedroids/gnetaleans, at least at low paleolatitudes (e.g., Krassilov, 1982, 1986; Krassilov & al., 1998; Y. Yang & al., 2005, 2013, 2018, 2020; Rydin & al., 2006b, 2010; Friis & al., 2007, 2019; X. Wang & Zheng, 2010; Ricardi-Branco & al., 2013). Most of the *Ephedra*-like fossils are difficult to place phylogenetically, partly because of their relatively poor preservation state but also because there are few diagnostic characters that clearly distinguish *Ephedra* from the remaining Gnetales. Many features of *Ephedra* are, in fact, ancestral in the Gnetales. Only a few well-preserved fossil seeds share uniquely derived features with the living clade, and the similarity between these fossil seeds (Rydin & al., 2004, 2006a) and those of living species, to the smallest anatomical and cellular detail, is baffling. However, detailed morphological investigations of living plants excluded these fossil seeds from the extant clade of *Ephedra*, instead placing them along the stem to *Ephedra* based on character optimization (Rydin & al., 2010). Similarly, micromorphological seed surface structures, suggested by some to be a direct link (i.e., homologous features) between certain modern species and a Cretaceous fossil, were shown to differ substantially upon closer examination (Ickert-Bond & Rydin, 2011). Comparative morphological studies and dating analyses using molecular data and fossil calibration outside of *Ephedra* have thus unanimously refuted an ancient (Early Cretaceous) origin of the *Ephedra* crown. A fossil that can be used as a calibration point within the *Ephedra* crown, i.e., a fossil that can be unambiguously phylogenetically placed within the crown group of *Ephedra*, has not been available.

Recently, new indications of a Cretaceous origin of the *Ephedra* crown have emerged based on information from fossil pollen. A derived ephedroid pollen type shown to have features unique to some but not all living species of *Ephedra* (Norbäck Ivarsson, 2013; Bolinder & al., 2015, 2016a,b) is present in sediments up to about 90 million years of age. The earliest report is from the North American Raritan Formation (Steeves & Barghoorn, 1959). Based on examination of slides from this section (James A. Doyle, pers. comm.), this pollen is from the upper part of the Raritan Formation (Zone V of J.A. Doyle & Robbins, 1977), which is of either latest Turonian or Coniacian age (Massoni & al., 2015), i.e., approximately 90–86 Myr (following Gradstein & al., 2012). Using this information for calibration of dating analyses based on molecular data would push back the age of the *Ephedra* crown to about 100 Ma. We find it, however, premature to use this new knowledge for dating purposes. The conclusions on morphology and evolution of *Ephedra* pollen made in previous work (e.g., Bolinder & al., 2016a,b) appear clear, but the topological uncertainties revealed in the present study show that much more work is needed before pollen information can be used as calibration point within *Ephedra* for the purpose of estimating divergence times in the genus.

Here, we instead chose to further explore the information provided by molecular data when used with an indisputable calibration point, which is available for the minimum age of the *Gnetum-Welwitschia* clade based on the welwitschioid fossil *Cratonia cotyledon* from the late-Aptian Crato Formation (Rydin & al., 2003). Results of dating analyses of *Ephedra* are, however, here shown to be highly uncertain. Not only are they dependent on the selection of the tree prior, but they are also dependent on other aspects of the evolutionary model and/or how the model is implemented in the software used for conducting the analyses. For example, the best age estimate to date for the *Ephedra* crown (Ickert-Bond & al., 2009), yielding a median age of about 30 Ma (95% confidence interval of about 74–20 Ma), is mimicked only in one of our analyses (Fig. 4C), and the result is clearly only valid for analyses using a pure birth (Yule) tree prior as implemented in BEAST. Other choices of tree prior and/or software return confidence intervals, spanning only of a few million years thus indicating an extremely young crown group of *Ephedra* (Fig. 4A,B,D).

Turning instead to other data provides a similar uncertainty. Topological results and biogeographic indications do not seem to support a mid-late Cretaceous age of the *Ephedra* crown as is indicated by the earliest occurrence of fossil pollen of the derived type and included in the confidence interval of the dating analyses in BEAST using the Yule tree prior (Fig. 4C). If the *Ephedra* crown is of Cretaceous origin, divergences in the group should reflect the break-up of the Gondwana and/or Laurasia continents. But instead, we see monophyly of New World taxa (North and South American species forming a clade), a consistent topological result of the present study. However, topology and biogeography do not seem to support an extremely young crown group of *Ephedra* either. A nearly worldwide distribution, accomplished only in a few million years would indicate very strong dispersal capabilities and should conceivably result in inferred repeated dispersal events, perhaps an almost random geographical pattern as optimized on the phylogeny. Instead, we see geographically distinct groups of *Ephedra*, apparently originating from a single or few dispersal events. Similar questions arise regarding the silent substitution rates estimated for *Ephedra* (Table 1, suppl. Appendix S4). Variation among species and plant groups is substantial (Richardson & al., 2013) making it difficult to compare *Ephedra* to angiosperms. It should be noted, however, that extremely low estimated silent substitution rates, like those estimated for the *Ephedra* crown if the group is

approaching 100 Ma (Table 1), are unusual and would require that considerable correction of the DNA exists since mutations are random events that do occur at some average level of speed. On the other hand, extremely low silent substitution rates do occur, not least in the plant mitochondrion (Richardson & al., 2013), and a migration of *Ephedra* to higher paleolatitudes during the Paleogene, from a previous equatorial–low-latitude distribution, is conceivable, for example as a consequence of climate change and/or increasing competition from angiosperms at low paleolatitudes.

Thus, although Bayes factors could determine the relative fit of different evolutionary models to data and aid, for example, in the choice between a birth-death or pure birth tree prior, the deviating results and indications from different sources of data make it difficult to speculate about the age of the *Ephedra* crown. Accepting the results of our analyses and taking all current knowledge into consideration, the age of the *Ephedra* crown could be anything between 2 Myr and 100 Myr. To progress in the field, the most important task at present is to verify the uniqueness of *Ephedra* pollen of the derived type (Norbäck Ivarsson, 2013; Bolinder & al., 2015, 2016a,b). Confidently placing the Cretaceous fossils of this type among extant *Ephedra* would settle the minimum age of the *Ephedra* crown to the mid-Cretaceous, whereas a rejection of homology of all pollen of the derived type would lend some support in favor of a "young *Ephedra* crown group" hypothesis.

**Species of possible hybrid/polyploid origin. —** With information from the entire plastome and from a large proportion of the nuclear ribosomal DNA cistron, as used here, cytonuclear discordance is apparent at all levels of the phylogeny. We count to at least 18 of our 50 samples displaying different phylogenetic positions as retrieved from plastid data vs. nuclear rDNA data. A striking example concerns *Ephedra milleri*. Little material has ever been collected of this species, which was described relatively recently (Freitag & Maier-Stolte, 1992). We have studied a paratype, a male specimen (*Miller 7667A* [E00216912]), in several studies (Rydin & Korall, 2009; Thureborn, 2014; Bolinder & al., 2016b). In concatenated analyses of plastid and nuclear rDNA data conducted in previous studies, it has typically been resolved as sister to "core *Ephedra*" (sensu Rydin & Korall, 2009) (but see Appendix Fig. A1 in Loera & al., 2015). Here, it is clearly shown that whereas nuclear ribosomal DNA places *E. milleri* among Mediterranean species and sister to *E. alata*, plastid data instead places it well nested in the Asian clade. Considering the suggested strictly maternal inheritance of the chloroplast in *Ephedra* (Mogensen, 1996), the most straightforward explanation is that the result from plastid and nuclear data reflects the maternal and paternal ancestry, respectively, either of a hybrid specimen, or of a species affected by reticulation and/or polyploidy. Further support for the assumption comes from the fact that *E. milleri* was collected in Oman on the Arabian Peninsula. It is not necessary to infer long-distance dispersal of pollen to explain hybridization between a Medi-terranean and an Asian species because some *Ephedra* species have a broad distribution that may overlap with species of evolutionary origin from other continents. An example among living species is *E. distachya* of the Asian clade and *E. foeminea* of the Mediterranean clade/species assemblage, which grow in sympatry in eastern Europe and western-most Asia. According to the results of the present study, *E. milleri* shares a maternal ancestor with a clade including the broadly distributed (Kakiuchi & al., 2011) *E. distachya* (Figs. 1, 2). Interpreting the results based on nuclear rDNA (Fig. 3, suppl. Fig. S1), the paternal ancestor of *E. milleri* may be *E. alata*, which has a distribution extending from the Mediterranean (northern Africa) into the Arabian deserts (Freitag & Maier-Stolte, 1994), i.e., apparently including the type locality of *E. milleri*. The "intermediate" placement of *E. milleri*, as sister to the "core clade of *Ephedra*" in previous work (Rydin & Korall, 2009), is most likely a consequence of conflicting information between the relatively few (seven) nuclear ribosomal and chloroplast molecular regions used in that study.

Another equally striking example concerns *Ephedra pedunculata*, native to Texas and northern Mexico. Its phylogenetic position has varied in previous studies and has often been poorly supported, statistically. In our results based on nuclear rDNA, it is strongly supported as sister to all other New World species (Fig. 3, suppl. Fig. S1), which is congruent with results of previous studies employing a dense sampling of the entire genus (e.g., Ickert-Bond & Wojciechowski, 2004; Rydin & Korall, 2009). However, in our analyses of plastid data, the same specimen of *E. pedunculata* is strongly supported as sister to the North American taxa *E. coryi* and *E. antisyphilitica*. It is thus possible that the previous difficulties in resolving the phylogenetic position of *E. pedunculata* with statistical support can be explained by a history of hybridization/introgression or an allopolyploid origin of the species.

Yet other examples concern the phylogenetic positions of the Chinese species *Ephedra likiangensis* and *E. minuta*, which are inconsistently positioned in results based on plastid data and nuclear rDNA data. In Rydin & Korall (2009), they formed a strongly supported clade, sister to the remaining Asian species. Here they are sisters in the tree based on nuclear rDNA data (Fig. 3, suppl. Fig. S1), but the phylogenetic position of this *E. likiangensis–E. minuta* clade is unsupported. In contrast, plastid data resolve and support both species as nested well inside the Asian clade and not as each other's closest relative (Figs. 1, 2). Additional incongruent results are found for the Asian *E. pachyclada*, *E. somalensis* and *E. strobilacea* as well as the New World *E. californica*, *E. trifurca*, *E. viridis*. These examples of possible reticulate evolution are perhaps not so surprising (although the apparent extent of it may be). Hybridization and polyploidy have been suggested to frequently occur in *Ephedra*, demonstrated, e.g., based on morphology (Wendt, 1993), pollen data (Ickert-Bond & al., 2003), seed micromorphology (Ickert-Bond & Rydin, 2011), as well as chromosome data (Mehra, 1946; Choudry, 1984; Ickert-Bond & al., 2020) in combination with molecular data (H. Wu & al., 2016).

Finally, the position of the Mediterranean species *Ephedra foeminea* is strongly challenged by the results of the

present study. In earlier work, it was resolved as sister to the remaining species of *Ephedra* based on molecular data (Rydin & Korall, 2009), a result supported by reproductive biology/morphology and pollination mechanisms (Bolinder & al., 2015, 2016a; Rydin & Bolinder, 2015). In the present study, there is no support for a sister relationship between *E. foeminea* and the remaining species of *Ephedra*. Our non-clock analysis of plastid data (Fig. 1), as well as our analyses of nuclear ribosomal DNA data (Fig. 3, suppl. Fig. S1), fail to find a well-supported position for the root in *Ephedra*, but none of the alternative placements of the root imply a sister relationship between *E. foeminea* and the remaining species of the genus.

## ■ CONCLUSIONS

Instead of a comprehensive understanding of the phylogeny of *Ephedra*, as we aimed for, we are left with inconsistencies that call for further investigations. We find no support for the previously proposed sister relationship between *E. foeminea* and the remaining species of the genus, although the morphological and functional differences between this species and other species of *Ephedra* (Bolinder & al., 2015, 2016a; Rydin & Bolinder, 2015) remain clear, with the features of *E. foeminea* seemingly being ancestral. Furthermore, a large proportion of the specimens used in the present study display incongruent phylogenetic positions as assessed by analyses of the plastid genome vs. the nuclear ribosomal DNA. To fully explore the reasons behind the cytonuclear discordance requires additional studies. The most straightforward hypothesis seems to us to involve a history of hybridization/introgression and/or allopolyploidy, but other processes such as incomplete lineage sorting cannot be ruled out at this point. Another striking example of uncertain and inconsistent results concerns the age of the *Ephedra* crown. Assessing its age based on information from fossil pollen that seems to share unique features with pollen of some, but not all, species of the crown group, yields an estimated age of the *Ephedra* crown of about 100 Myr. Dismissing this fossil information as uncertain (perhaps representing parallel evolution of similar features), instead relying on calibration outside of *Ephedra*, results in node age estimates of the *Ephedra* crown between 2 Myr and 74 Myr depending on analytical approach. A Cretaceous origin of the crown group of *Ephedra* would indicate extremely low silent substitution rates in *Ephedra* and the tree topology does not seem to reflect, for example, a Gondwanan or Laurasian distribution. On the other hand, accepting a recent origin of the *Ephedra* crown, as indicated by all the molecular estimates using a birth-death tree prior, would leave us with a crown group that must have dispersed very quickly over large parts of the world. Yet, *Ephedra* shows no signs of repeated dispersal events. The tree topology displays distinct geographical groups, all seemingly the result of a single or few dispersal events. Would such patterns be expected if the crown group of *Ephedra* originated as late as during the Pliocene-Pleistocene?

## ■ AUTHOR CONTRIBUTIONS

CR and NW designed the research. RB and OT conducted laboratory work. NW, RB and OT analyzed the data. All authors interpreted the results. CR and NW wrote the manuscript with comments from RB and OT. — CR, https://orcid.org/0000-0002-3347-7820; RB, https://orcid.org/0000-0002-5963-3537; OT, https://orcid.org/0000-0002-9609-4245; NW, https://orcid.org/0000-0002-4276-9366

## ■ ACKNOWLEDGEMENTS

We thank Anbar Khodabandeh for technical assistance, and Professor Bertil Ståhl (Uppsala University) and the herbaria E, HA, L, MO, O, S, UC, UPS, WU and Z for access to plant material. We acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure, NGI, and Uppmax for providing assistance in massive parallel sequencing and computational infrastructure. The project was funded by grants from the Royal Academy of Sciences, Stockholm University, and the Swedish Research Council to CR.

## ■ LITERATURE CITED

**Alfaro, M.E., Zoller, S. & Lutzoni, F.** 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molec. Biol. Evol.* 20: 255–266. https://doi.org/10.1093/molbev/msg028

**Arber, E.A.N. & Parkin, J.** 1908. Studies on the evolution of angiosperms: The relationship of the angiosperms to the Gnetales. *Ann. Bot. (Oxford)* 22: 489–515. https://doi.org/10.1093/oxfordjournals.aob.a089185

**Barker, M.S., Husband, B.C. & Pires, J.C.** 2016. Spreading Winge and flying high: The evolutionary importance of polyploidy after a century of study. *Amer. J. Bot.* 103: 1139–1145. https://doi.org/10.3732/ajb.1600272

**Bolinder, K., Niklas, K.J. & Rydin, C.** 2015. Aerodynamics and pollen ultrastructure in *Ephedra* (Gnetales). *Amer. J. Bot.* 102: 457–470. https://doi.org/10.3732/ajb.1400517

**Bolinder, K., Humphreys, A.M., Ehrlén, J., Alexandersson, R., Ickert-Bond, S.M. & Rydin, C.** 2016a. From near extinction to diversification by means of a shift in pollination mechanism in the gymnosperm relict *Ephedra* (Ephedraceae, Gnetales). *Bot. J. Linn. Soc.* 180: 461–477. https://doi.org/10.1111/boj.12380

**Bolinder, K., Norbäck Ivarsson, L., Humphreys, A.M., Ickert-Bond, S.M., Han, F., Hoorn, C. & Rydin, C.** 2016b. Pollen morphology of *Ephedra* (Gnetales) and its evolutionary implications. *Grana* 55: 24–51. https://doi.org/10.1080/00173134.2015.1066424

**Bowe, L.M., Coat, G. & de Pamphilis, 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. U.S.A.* 97: 4092–4097. https://doi.org/10.1073/pnas.97.8.4092

**Burleigh, J.G. & Mathews, S.** 2004. Phylogenetic signal in nucleotide data from seed plants: Implications for resolving the seed plant tree of life. *Amer. J. Bot.* 91: 1599–1613. https://doi.org/10.3732/ajb.91.10.1599

**Burleigh, J.G. & Mathews, S.** 2007. Assessing among-locus variation in the inference of seed plant phylogeny. *Int. J. Pl. Sci.* 168: 111–124. https://doi.org/10.1086/509586

**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: 292857. https://doi.org/10.1155/2012/292857

**Chase, M.W., Soltis, D.E., Olmstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., Duvall, M.R., Price, R.A., Hills, H.G., Qui, Y.-L., Kron, K.A., Rettig, J.H., Conti, E., Palmer, J.D., Manhart, J.R., Sytsma, K.J., Michaels, H.J., Kress, W.J., Karol, K.G., Clark, W.D., Hedrén, M., Gaut, B.S., Jansen, R.K., Kim, K.-J., Wimpee, C.F., Smith, J.F., Furnier, G.R., Strauss, S.H., Xiang, Q.-Y., Plunkett, G.M., Soltis, P.S., Swensen, S.M., Williams, S.E., Gadek, P.A., Quinn, C.J., Eguiarte, L.E., Golenberg, E., Learns, G.H., Graham, S.W., Barrett, S.C.H., Dayanandan, S. & Albert, V.** 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbcL*. *Ann. Missouri Bot. Gard.* 80: 528–580. https://doi.org/10.2307/2399846

**Chaw, S.-M., Parkinson, C.L., Cheng, Y., 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. U.S.A.* 97: 4086–4091. https://doi.org/10.1073/pnas.97.8.4086

**Choudry, A.S.** 1984. Karyomorphological and cytogenetical studies in *Ephedra*. *J. Sci. Hiroshima Univ. Ser. B, Div. 2, Bot.* 19(1): 57–109.

**Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B. & de Hoon, M.J.L.** 2009. Biopython: Freely available python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25: 1422–1423. https://doi.org/10.1093/bioinformatics/btp163

**Coiro, M., Chomicki, G. & Doyle, J.A.** 2018. Experimental signal dissection and method sensitivity analyses reaffirm the potential of fossils and morphology in the resolution of the relationship of angiosperms and Gnetales. *Paleobiology* 44: 490–510. https://doi.org/10.1017/pab.2018.23

**Condamine, F.L., Nagalingum, N.S., Marshall, C.R. & Morlon, H.** 2015. Origin and diversification of living cycads: A cautionary tale on the impact of the branching process prior in Bayesian molecular dating. *B. M. C. Evol. Biol.* 15: 65. https://doi.org/10.1186/s12862-015-0347-8

**Cox, C.J.** 2018. Land plant molecular phylogenetics: A review with comments on evaluating incongruence among phylogenies. *Crit. Rev. Pl. Sci.* 37: 113–127. https://doi.org/10.1080/07352689.2018.1482443

**Crane, P.R.** 1985. Phylogenetic analysis of seed plants and the origin of angiosperms. *Ann. Missouri Bot. Gard.* 72: 716–793. https://doi.org/10.2307/2399221

**Crane, P.R.** 1996. The fossil history of the Gnetales. *Int. J. Pl. Sci.* 157: 50–57. https://doi.org/10.1086/297403

**Crane, P.R. & Upchurch, G.R.** 1987. *Drewria potomacensis* gen. et sp. nov., an Early Cretaceous member of Gnetales from the Potomac Group of Virginia. *Amer. J. Bot.* 74: 1722–1736. https://doi.org/10.2307/2444143

**Cummins, C.A. & McInerney, J.O.** 2011. A method for inferring the rate of evolution of homologous characters that can potentially improve phylogenetic inference, resolve deep divergence and correct systematic biases. *Syst. Biol.* 60: 833–844. https://doi.org/10.1093/sysbio/syr064

**Davis, C.C., Xi, Z. & Mathews, S.** 2014. Plastid phylogenomics and green plant phylogeny: Almost full circle but not quite there. *B. M. C. Biol.* 12: 11. https://doi.org/10.1186/1741-7007-12-11

**De Sousa, F., Foster, P.G., Donoghue, P.C.J., Schneider, H. & Cox, C.J.** 2019. Nuclear protein phylogenies support the monophyly of the three bryophyte groups (Bryophyta Schimp.). *New Phytol.* 222: 565–575. https://doi.org/10.1111/nph.15587

**Donoghue, M.J. & Cracraft, J.** 2004. Charting the tree of life. P. 3 in: Cracraft, J. & Donoghue, M.J. (eds.), *Assembling the tree of life*. New York: Oxford University Press.

**Doyle, J.A. & Donoghue, M.J.** 1986. Seed plant phylogeny and the origin of angiosperms: An experimental cladistic approach. *Bot. Rev. (Lancaster)* 52: 321–431. https://doi.org/10.1007/bf02861082

**Doyle, J.A. & Robbins, E.I.** 1977. Angiosperm pollen zonation of the continental Cretaceous of the Atlantic Coastal Plain and its application to deep wells in the Salisbury Embayment. *Palynology* 1: 43–78. https://doi.org/10.1080/01916122.1977.9989150

**Doyle, J.J.** 1991. DNA protocols for plants. Pp. 283–293 in: Hewitt, G., Johnson, A.W.B. & Young, J.P.W. (eds.), *Molecular techniques in taxonomy*. Berlin: Springer.

**Doyle, J.J. & Doyle, J.L.** 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull. Bot. Soc. Amer.* 19: 11–15.

**Drummond, A.J. & Suchard, M.A.** 2010. Bayesian random local clocks, or one rate to rule them all. *B. M. C. Biol.* 8: 114. https://doi.org/10.1186/1741-7007-8-114

**Drummond, A.J., Suchard, M., Xie, D. & Rambaut, A.** 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molec. Biol. Evol.* 29: 1969–1973. https://doi.org/10.1093/molbev/mss075

**Edgar, R.C.** 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucl. Acids Res.* 32: 1792–1797. https://doi.org/10.1093/nar/gkh340

**Erixon, P., Svennblad, B., Britton, T. & Oxelman, B.** 2003. Reliability of Bayesian posterior probabilities and bootstrap frequences in phylogenetics. *Syst. Biol.* 52: 665–673. https://doi.org/10.1080/10635150390235485

**Fitch, W.M.** 1970. Distinguishing homologous from analogous proteins. *Syst. Zool.* 19: 99–113. https://doi.org/10.2307/2412448

**Folk, R.A., Mandel, J.R. & Freudenstein, J.V.** 2017. Ancestral gene flow and parallel organellar genome capture result in extreme phylogenomic discord in a lineage of angiosperms. *Syst. Biol.* 66: 320–337. https://doi.org/10.1093/sysbio/syw083

**Folk, R.A., Soltis, P.S., Soltis, D.E. & Guralnick, R.** 2018. New prospects in the detection and comparative analysis of hybridization in the tree of life. *Amer. J. Bot.* 105: 364–375. https://doi.org/10.1002/ajb2.2018

**Freitag, H. & Maier-Stolte, M.** 1992. A new species and a new combination in the genus *Ephedra* from Arabia. *Edinburgh J. Bot.* 49: 89–93. https://doi.org/10.1017/S0960428600001517

**Freitag, H. & Maier-Stolte, M.** 1994. Characterization of areas, Ephedraceae. Pp. 5–16 in: Browicz, K. (ed.), *Chorology of trees and shrubs in South-West Asia and adjacent regions*. Kornik: Polish Academy of Sciences, Institute of Dendrology.

**Friis, E.M., Crane, P.R., Pedersen, K.R., Bengtson, S., Donoghue, P.C.J., Grimm, G.W. & Stampanoni, M.** 2007. Phase-contrast X-ray microtomography links Cretaceous seeds with Gnetales and Bennettitales. *Nature* 450: 549–552. https://doi.org/10.1038/nature06278

**Friis, E.M., Pedersen, K.R. & Crane, P.R.** 2013. New diversity among chlamydospermous seeds from the Early Cretaceous of Portugal and North America. *Int. J. Pl. Sci.* 174: 530–558. https://doi.org/10.1086/668250

**Friis, E.M., Pedersen, K.R. & Crane, P.R.** 2014. Welwitschioid diversity in the Early Cretaceous: Evidence from fossil seeds with pollen from Portugal and eastern North America. *Grana* 53: 175–196. https://doi.org/10.1080/00173134.2014.915980

**Friis, E.M., Crane, P.R. & Pedersen, K.R.** 2019. Chlamydospermous seeds document the diversity and abundance of extinct gnetalean relatives in Early Cretaceous vegetation. *Int. J. Pl. Sci.* 180: 643–666. https://doi.org/10.1086/704356

**Gelman, A. & Rubin, D.B.** 1992. Inference from iterative simulation using multiple sequences. *Statist. Sci.* 7: 457–472. https://doi.org/10.1214/ss/1177011136

**Goremykin, V., Bobrova, V.K., Pahnke, J., Troitsky, A.V., Antonov, A. & Martin, W.** 1996. Noncoding sequences from the slowly evolving chloroplast inverted repeat in addition to *rbcL* data do not support gnetalean affinities with angiosperms. *Molec. Biol. Evol.* 13: 383–396.

**Gradstein, F.M., Ogg, J.G., Schmitz, M. & Ogg, G.** 2012. *The geologic time scale 2012*. Oxford, U.K.: Elsevier.

**Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O.** 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59: 307–321. https://doi.org/10.1093/sysbio/syq010

**Han, F., Rydin, C., Bolinder, K., Dupont-Nivet, G., Abels, H.A., Koutsodendris, A., Zhang, K. & Hoorn, C.** 2016. Steppe development on the Northern Tibetan Plateau inferred from Paleogene ephedroid pollen. *Grana* 55: 71–100. https://doi.org/10.1080/00173134.2015.1120343

**Heimhofer, U. & Hochuli, P.A.** 2010. Early Cretaceous angiosperm pollen from a low-latitude succession (Araripe Basin, NE Brazil). *Rev. Palaeobot. Palynol.* 161: 105–126. https://doi.org/10.1016/j.revpalbo.2010.03.010

**Hilton, J. & Bateman, R.M.** 2006. Pteridosperms are the backbone of seed-plant phylogeny. *J. Torrey Bot. Soc.* 133: 119–168. https://doi.org/10.3159/1095-5674(2006)133[119:patbos]2.0.co;2

**Hou, C., Humphreys, A.M., Thureborn, O. & Rydin, C.** 2015. New insights into the evolutionary history of *Gnetum* (Gnetales). *Taxon* 64: 239–253. https://doi.org/10.12705/642.12

**Hou, C., Wikström, N. & Rydin, C.** 2017. The chloroplast genome of *Ephedra foeminea* (Ephedraceae, Gnetales), an entomophilous gymnosperm endemic to the Mediterranean area. *Mitochondrial DNA, A*, 28: 330–331. https://doi.org/10.3109/19401736.2015.1122768

**Huang, J. & Price, R.A.** 2003. Estimation of the age of extant *Ephedra* using chloroplast *rbcL* sequence data. *Molec. Biol. Evol.* 20: 435–440. https://doi.org/10.1093/molbev/msg049

**Huang, J., Giannasi, D.E. & Price, R.A.** 2005. Phylogenetic relationships in *Ephedra* (Ephedraceae) inferred from chloroplast and nuclear DNA sequences. *Molec. Phylogen. Evol.* 35: 48–59. https://doi.org/10.1016/j.ympev.2004.12.020

**Huelsenbeck, J.P. & Ronquist, F.R.** 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755. https://doi.org/10.1093/bioinformatics/17.8

**Ickert-Bond, S.M. & Renner, S.S.** 2016. The Gnetales: Recent insights on their morphology, reproductive biology, chromosome numbers, biogeography, and divergence times. *J. Syst. Evol.* 54: 1–16. https://doi.org/10.1111/jse.12190

**Ickert-Bond, S.M. & Rydin, C.** 2011. Micromorphology of the seed envelope of *Ephedra* L. (Gnetales) and its relevance for the timing of evolutionary events. *Int. J. Pl. Sci.* 172: 36–48. https://doi.org/10.1086/657299

**Ickert-Bond, S.M. & Wojciechowski, M.F.** 2004. Phylogenetic relationships in *Ephedra* (Gnetales): Evidence from nuclear and chloroplast DNA sequence data. *Syst. Bot.* 29: 834–849. https://doi.org/10.1600/0363644042451143

**Ickert-Bond, S.M., Skvarla, J.J. & Chissoe, W.F.** 2003. Pollen dimorphism in *Ephedra* L. (Ephedraceae). *Rev. Palaeobot. Palynol.* 124: 325–334. https://doi.org/10.1016/S0034-6667(03)00002-2

**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 America. *J. Syst. Evol.* 47: 444–456. https://doi.org/10.1111/j.1759-6831.2009.00053.x

**Ickert-Bond, S.M., Sousa, A., Min, Y., Loera, I., Metzgar, J., Pellicer, J., Hidalgo, O. & Leitch, I.J.** 2020. Polyploidy in gymnosperms - Insights into the genomic and evolutionary consequences of polyploidy in *Ephedra*. *Molec. Phylogen. Evol.* 147: 106786. https://doi.org/10.1016/j.ympev.2020.106786

**Jarvis, E.D., Mirarab, S., Aberer, A.J., Li, B., Houde, P., Li, C., Ho, S.Y.W., Faircloth, B.C., Nabholz, B., Howard, J.T., Suh, A., Weber, C.C., da Fonseca, R.R., Li, J.W., Zhang, F., Li, H., Zhou, L., Narula, N., Liu, L., Ganapathy, G., Boussau, B., Bayzid, M.S., Zavidovych, V., Subramanian, S., Gabaldon, T., Capella-Gutierrez, S., Huerta-Cepas, J., Rekepalli, B., Munch, K., Schierup, M., Lindow, B., Warren, W.C., Ray, D., Green, R.E., Bruford, M.W., Zhan, X.J., Dixon, A., Li, S.B., Li, N., Huang, Y.H., Derryberry, E.P., Bertelsen, M.F., Sheldon, F.H., Brumfield, R.T., Mello, C.V., Lovell, P.V., Wirthlin, M., Schneider, M.P.C., Prosdocimi, F., Samaniego, J.A., Velazquez, A.M.V., Alfaro-Nunez, A., Campos, P.F., Petersen, B., Sicheritz-Ponten, T., Pas, A., Bailey, T., Scofield, P., Bunce, M., Lambert, D.M., Zhou, Q., Perelman, P., Driskell, A.C., Shapiro, B., Xiong, Z.J., Zeng, Y.L., Liu, S.P., Li, Z.Y., Liu, B.H., Wu, K., Xiao, J., Yinqi, X., Zheng, Q.M., Zhang, Y., Yang, H.M., Wang, J., Smeds, L., Rheindt, F.E., Braun, M., Fjeldsa, J., Orlando, L., Barker, F.K., Jonsson, K.A., Johnson, W., Koepfli, K.P., O'Brien, S., Haussler, D., Ryder, O.A., Rahbek, C., Willerslev, E., Graves, G.R., Glenn, T.C., McCormack, J., Burt, D., Ellegren, H., Alström, P., Edwards, S.V., Stamatakis, A., Mindell, D.P., Cracraft, J., Braun, E.L., Warnow, T., Jun, W., Gilbert, M.T.P. & Zhang, G.J.** 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346: 1320–1331. https://doi.org/10.1126/science.1253451

**Johnson, M.G., Pokorny, L., Dodsworth, S., Botigué, L.R., Cowan, R.S., Devault, A., Eiserhardt, W.L., Epitawalage, N., Forest, F., Kim, J.T., Leebens-Mack, J.H., Leitch, I.J., Maurin, O., Soltis, D.E., Soltis, P.S., Wong, G.K.-S., Baker, W.J. & Wickett, N.J.** 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Syst. Biol.* 68: 594–606. https://doi.org/10.1093/sysbio/syy086

**Kakiuchi, N., Mikage, M., Ickert-Bond, S.M., Maier-Stolte, M. & Freitag, H.** 2011. A molecular phylogenetic study of the *Ephedra distachya/E. sinica* complex in Eurasia. *Willdenowia* 41: 203–215. https://doi.org/10.3372/Wi.41.41201

**Kendall, D.G.** 1948. On the generalized "birth-and-death" process. *Ann. Math. Statist.* 19: 1–15. https://doi.org/10.1214/aoms/1177730285

**Kent, J.W.** 2002. BLAT—The BLAST-like alignment tool. *Genome Res*. 12: 656–664. https://doi.org/10.1101/gr.229202

**Krassilov, V.A.** 1982. Early Cretaceous flora of Mongolia. *Palaeontographica, Abt. B, Paläophytol.* 181: 1–43.

**Krassilov, V.A.** 1986. New floral structures from the Lower Cretaceous of Lake Baikal area. *Rev. Palaeobot. Palynol.* 47: 9–16. https://doi.org/10.1016/0034-6667(86)90003-5

**Krassilov, V.A. & Bugdaeva, E.V.** 1982. Achene-like fossils from the Lower Cretaceous of the Lake Baikal area. *Rev. Palaeobot. Palynol.* 36: 279–295. https://doi.org/10.1016/0034-6667(82)90023-9

**Krassilov, V.A., Dilcher, D.L. & Douglas, J.G.** 1998. New ephedroid plant from the Lower Cretaceous Koonwarra Fossil Bed, Victoria, Australia. *Alcheringa* 22: 123–133. https://doi.org/10.1080/03115519808619195

**Larget, B. & Simon, D.L.** 1999. Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Molec. Biol. Evol.* 16: 750–759. https://doi.org/10.1093/oxfordjournals.molbev.a026160

**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. https://doi.org/10.1371/journal.pgen.1002411

**Lee-Yaw, J.A., Grassa, C.J., Joly, S., Andrew, R.L. & Rieseberg, L.H.** 2019. An evaluation of alternative explanations for widespread cytonuclear discordance in annual sunflowers (*Helianthus*). *New Phytol.* 221: 515–526. https://doi.org/10.1111/nph.15386

**LePage, T., Bryant, D.H.P. & Lartillot, N.** 2007. A general comparison of relaxed molecular clock models. *Molec. Biol. Evol.* 24: 2669–2680. https://doi.org/10.1093/molbev/msm193

**Li, H. & Durbin, R.** 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

**Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & 1000 Genome Project**

**Data Processing Subgroup** 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

**Liu, Y., Johnson, M.G., Cox, C.J., Medina, R., Devos, N., Vanderpoorten, A., Hedenäs, L., Bell, N.E., Shevock, J.R., Aguero, B., Quandt, D., Wickett, N.J., Shaw, A.J. & Goffinet, B.** 2019. Resolution of the ordinal phylogeny of mosses using targeted exons from organellar and nuclear genomes. *Nature, Commun.* 10: 1485. https://doi.org/10.1038/s41467-019-09454-w

**Loera, I., Sosa, V. & Ickert-Bond, S.M.** 2012. Diversification in North American arid lands: Niche conservatism, divergence and expansion of habitat explain speciation in the genus *Ephedra*. *Molec. Phylogen. Evol.* 65: 437–450. https://doi.org/10.1016/j.ympev.2012.06.025

**Loera, I., Ickert-Bond, S.M. & Sosa, V.** 2015. Ecological consequences of contrasting dispersal syndromes in New World *Ephedra*: Higher rates of niche evolution related to dispersal ability. *Ecography* 38: 1187–1199. https://doi.org/10.1111/ecog.01264

**Maddison, W.P.** 1997. Gene trees in species trees. *Syst. Biol.* 46: 523–536. https://doi.org/10.1093/sysbio/46.3.523

**Mallet, J.** 2007. Hybrid speciation. *Nature* 446: 279–283. https://doi.org/10.1038/nature05706

**Massoni, J., Doyle, J. & Sauquet, H.** 2015. Fossil calibration of Magnoliidae, an ancient lineage of angiosperms. *Palaeontol. Electronica* 18.1.2FC. https://doi.org/10.26879/435

**Mathews, S.** 2009. Phylogenetic relationships among seed plants: Persistent questions and the limits of molecular data. *Amer. J. Bot.* 96: 228–236. https://doi.org/10.3732/ajb.0800178

**Mathews, S. & Kramer, E.M.** 2012. The evolution of reproductive structures in seed plants: A re-examination based on insights from developmental genetics. *New Phytol.* 194: 910–923. https://doi.org/10.1111/j.1469-8137.2012.04091.x

**Mau, B., Newton, M.A. & Larget, B.** 1999. Bayesian phylogenetic inference via Markov chain Monte Carlo methods. *Biometrics* 55: 1–12. https://doi.org/10.1111/j.0006-341X.1999.00001.x

**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. *B. M. C. Evol. Biol.* 8: 130. https://doi.org/10.1186/1471-2148-8-130

**Mehra, B.** 1946. A study of the karyotypes and the occurrence of diploid male gametophytes in some species of the genus *Ephedra*. *Proc. Natl. Acad. Sci. India, B*, 16: 259–286.

**Mogensen, H.L.** 1996. The hows and whys of cytoplasmic inheritance in seed plants. *Amer. J. Bot.* 83: 383–404. https://doi.org/10.1002/j.1537-2197.1996.tb12718.x

**Mohr, B. & Friis, E.M.** 2000. Early angiosperms from the Lower Cretaceous Crato Formation (Brazil), a preliminary report. *Int. J. Pl. Sci.* 161(S6): S155–S167. https://doi.org/10.1086/317580

**Morales-Briones, D.F., Liston, A. & Tank, D.C.** 2018. Phylogenomic analyses reveal a deep history of hybridization and polyploidy in the Neotropical genus *Lachemilla* (Rosaceae). *New Phytol.* 218: 1668–1684. https://doi.org/10.1111/nph.15099

**Muse, S.V. & Gaut, B.S.** 1994. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. *Molec. Biol. Evol.* 11: 715–724. https://doi.org/10.1093/oxfordjournals.molbev.a040152

**Norbäck Ivarsson, L.** 2013. *Pollen morphology in Ephedra (Gnetales) and implications for understanding fossil ephedroid pollen from the Tibetan Plateau, using a phylogenetic approach*. M.Sc. Thesis. Stockholm University, Stockholm, Sweden.

**Nylander, J.A.A.** 2004. MrAIC.pl. Program distributed by the author. Evolutionary Biology Centre, Uppsala University. https://github.com/nylander

**Osborn, J.M., Taylor, T.N. & de Lima, M.R.** 1993. The ultrastructure of fossil ephedroid pollen with gnetalean affinities from the Lower Cretaceous of Brazil. *Rev. Palaeobot. Palynol.* 77: 171–184. https://doi.org/10.1016/0034-6667(93)90003-D

**Pease, J.B., Brown, J.W., Walker, J.F., Hinchliff, C.E. & Smith, S.A.** 2018. Quartet sampling distinguishes lack of support from conflicting support in the green plant tree of life. *Amer. J. Bot.* 105: 385–403. https://doi.org/10.1002/ajb2.1016

**Pond, S.L., Frost, S.D. & Muse, S.V.** 2005. HyPhy: Hypothesis testing using phylogenies. *Bioinformatics* 21: 676–679. https://doi.org/10.1093/bioinformatics/bti079

**Puttick, M.N., Morris, J.L., Williams, T.A., Cox, C.J., Edwards, D., Kenrick, P., Pressel, S., Wellman, C.H., Schneider, H., Pisani, D. & Donoghue, P.C.J.** 2018. The interrelationships of land plants and the nature of the ancestral embryophyte. *Curr. Biol.* 28: 733–745. https://doi.org/10.1016/j.cub.2018.01.063

**Qiao, X., Li, Q., Yin, H., Qi, K., Li, K., Wang, R., Zhang, S. & Paterson, A.H.** 2019. Gene duplication and evolution in recurring polyploidization-diploidization cycles in plants. *Genome Biol.* 20: 38. https://doi.org/10.1186/s13059-019-1650-2

**Ran, J.H., Shen, T.T., Wang, M.M. & Wang, X.Q.** 2018. Phylogenomics resolves the deep phylogeny of seed plants and indicates partial convergent or homoplastic evolution between Gnetales and angiosperms. *Proc. Roy. Soc. Lond., Ser. B, Biol. Sci.* 285: 20181012. https://doi.org/10.1098/rspb.2018.1012

**Ricardi-Branco, F., Torres, M., Tavares, S.S., De Souza Carvalho, I., Tavares, P.G.E. & Campos, A.C.A.** 2013. *Itajuba yansanae* gen. and sp. nov. of Gnetales, Araripe Basin (Albian–Aptian) in Northeast Brazil. Chapter 7 in: Ray, P. & Zhang, Y. (eds.), *Climate change and regional/local responses*. Rijeka: InTechOpen. https://doi.org/10.5772/55704

**Richardson, A.O., Rice, D.W., Young, G.J., Alverson, A.J. & Palmer, J.D.** 2013. The "fossilized" mitochondrial genome of *Liriodendron tulipifera*: Ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. *B. M. C. Biol.* 11: 29. https://doi.org/10.1186/1741-7007-11-29

**Rieseberg, L.H. & Soltis, D.E.** 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evol. Trends Pl.* 5: 65–84.

**Ronquist, F. & Huelsenbeck, J.P.** 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

**Rota, J., Malm, T., Chazot, N., Peña, C. & Wahlberg, N.** 2017. A simple method for data partitioning based on relative evolutionary rates. *PeerJ* 6: e5498. https://doi.org/10.7717/peerj.5498

**Rothfels, C.J., Larsson, A., Li, F.-W., Sigel, E.M., Huiet, L., Burge, D.O., Ruhsam, M., Graham, S.W., Stevenson, D.W., Wong, G.K.-S., Korall, P. & Pryer, K.M.** 2013. Transcriptome-mining for single-copy nuclear markers in ferns. *PLoS ONE* 8: e76957. https://doi.org/10.1371/journal.pone.0076957

**Rothfels, C.J., Li, F.-W., Sigel, E.M., Huiet, L., Larsson, A., Burge, D.O., Ruhsam, M., Deyholos, M., Soltis, D.E., Stewart, C.N., Shaw, S.W., Pokorny, L., Chen, T., dePamphilis, C., DeGironimo, L., Chen, L., Wei, X., Sun, X., Korall, P., Stevenson, D.W., Graham, S.W., Wong, G.K.-S. & Pryer, K.M.** 2015. The evolutionary history of ferns inferred from 25 low-copy nuclear genes. *Amer. J. Bot.* 102: 1089–1107. https://doi.org/10.3732/ajb.1500089

**Rothwell, G.W. & Stockey, R.A.** 2013. Evolution and phylogeny of gnetophytes: Evidence from the anatomically preserved seed cone *Protoephedrites eamesii* gen. et sp. nov. and the seeds of several bennettitalean species. *Int. J. Pl. Sci.* 174: 511–529. https://doi.org/10.1086/668688

**Ruhfel, B.R., Gitzendanner, M.A., Soltis, P.S., Soltis, D.E. & Burleigh, J.G.** 2014. From algae to angiosperms –Inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *B. M. C. Evol. Biol.* 14: 23. https://doi.org/10.1186/1471-2148-14-23

**Rydin, C. & Bolinder, K.** 2015. Moonlight pollination in the gymnosperm *Ephedra* (Gnetales). *Biol. Lett.* 11: 20140993. https://doi.org/10.1098/rsbl.2014.0993

**Rydin, C. & Friis, E.M.** 2010. A new Early Cretaceous relative of Gnetales: *Siphonospermum simplex* gen. et sp. nov. from the Yixian Formation of northeast China. *B. M. C. Evol. Biol.* 10: 183. https://doi.org/10.1186/1471-2148-10-183

**Rydin, C. & Hoorn, C.** 2016. The Gnetales: Past and present. *Grana* 55: 1–4. https://doi.org/10.1080/00173134.2015.1118530

**Rydin, C. & Korall, P.** 2009. Evolutionary relationships in *Ephedra* (Gnetales), with implications for seed plant phylogeny. *Int. J. Pl. Sci.* 170: 1031–1043. https://doi.org/10.1086/605116

**Rydin, C., Källersjö, M. & Friis, 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. Pl. Sci.* 163: 197–214. https://doi.org/10.1086/338321

**Rydin, C., Mohr, B.A.R. & Friis, E.M.** 2003. *Cratonia cotyledon* gen. et sp. nov.: A unique Cretaceous seedling related to *Welwitschia*. *Proc. Roy. Soc. Lond., Ser. B, Biol. Sci.* 270(Suppl. 1): S29–S32. https://doi.org/10.1098/rsbl.2003.0044

**Rydin, C., Pedersen, K.R. & Friis, E.M.** 2004. On the evolutionary history of *Ephedra*: Cretaceous fossils and extant molecules. *Proc. Natl. Acad. Sci. U.S.A.* 101: 16571–16576. https://doi.org/10.1073/pnas.0407588101

**Rydin, C., Pedersen, K.R., Crane, P.R. & Friis, E.M.** 2006a. Former diversity of *Ephedra* (Gnetales): Evidence from Early Cretaceous seeds from Portugal and North America. *Ann. Bot. (Oxford)* 98: 123–140. https://doi.org/10.1093/aob/mcl078

**Rydin, C., Wu, S.Q. & Friis, E.M.** 2006b. *Liaoxia* (Gnetales): Ephedroids from the Early Cretaceous Yixian Formation in Liaoning, northeastern China. *Pl. Syst. Evol.* 262: 239–265. https://doi.org/10.1007/s00606-006-0481-2

**Rydin, C., Khodabandeh, A. & Endress, P.K.** 2010. The female reproductive unit of *Ephedra* (Gnetales): Comparative morphology and evolutionary perspectives. *Bot. J. Linn. Soc.* 163: 387–430. https://doi.org/10.1111/j.1095-8339.2010.01066.x

**Rydin, C., Wikström, N. & Bremer, B.** 2017. Conflicting results from mitochondrial genomic data challenge current views of Rubiaceae phylogeny. *Amer. J. Bot.* 104: 1522–1532. https://doi.org/10.3732/ajb.1700255

**Shen, H., Jin, D.M., Shu, J.P., Zhou, X.L., Lei, M., Wei, R., Shang, H., Wei, H.J., Zhang, R., Liu, L., Gu, Y.F., Zhang, X.C. & Yan, Y.H.** 2017. Large-scale phylogenomic analysis resolves a backbone phylogeny in ferns. *GigaScience* 7(2): gix116. https://doi.org/10.1093/gigascience/gix116

**Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J.M. & Birol, I.** 2009. ABySS: A parallel assembler for short read sequence data. *Genome Res.* 19: 1117–1123. https://doi.org/10.1101/gr.089532.108

**Smith, S.A. & Brown, J.W.** 2018. Constructing a broadly inclusive seed plant phylogeny. *Amer. J. Bot.* 105: 302–314. https://doi.org/10.1002/ajb2.1019

**Soltis, P.S. & Soltis, D.E.** 2009. The role of hybridization in plant speciation. *Annual Rev. Pl. Biol.* 60: 561–588. https://doi.org/10.1146/annurev.arplant.043008.092039

**Staden, R.** 1996. The Staden sequence analysis package. *Molec. Biotechnol.* 5: 233–241. https://doi.org/10.1007/BF02900361

**Staden, R., Beal, K.F. & Bonfield, J.K.** 2000. The Staden Package, 1998. Pp. 115–130 in: Misener, S. & Krawets, S.A. (eds.), *Methods in molecular biology*, vol. 132, *Bioinformatics methods and protocols*. New Jersey: Humana Press Inc.

**Steeves, M.W. & Barghoorn, E.S.** 1959. The pollen of *Ephedra*. *J. Arnold Arbor.* 40: 221–255.

**Straub, S.C.K., Fishbein, M., Livshultz, T., Foster, Z., Parks, M., Weitemier, K., Cronn, R.C. & Liston, A.** 2011. Building a model: developing genomic resources for common milkweed (*Asclepias syriaca*) with low coverage genome sequencing. *B. M. C. Genomics* 12: 211. https://doi.org/10.1186/1471-2164-12-211

**Takahata, N. & Nei, M.** 1985. Gene genealogy and variance of interpopulational nucleotide differences. *Genetics* 110: 325–344.

**Thompson, W.P.** 1918. Independent evolution of vessels in Gnetales and angiosperms. *Bot. Gaz.* 65: 83–90.

**Thureborn, O.** 2014. *New insights into the deep divergences of Ephedra (Gnetales) using molecular data*. M.Sc. Thesis. Stockholm University, Stockholm, Sweden.

**Vargas, O.M., Ortiz, E.M. & Simpson, B.B.** 2017. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). *New Phytol.* 214: 1736–1750. https://doi.org/10.1111/nph.14530

**Wang, X. & Zheng, S.L.** 2010. Whole fossil plants of *Ephedra* and their implications on the morphology, ecology and evolution of Ephedraceae (Gnetales). *Chin. Sci. Bull.* 55: 1511–1519. https://doi.org/10.1007/s11434-010-3069-8

**Wang, Z.** 2004. A new Permian gnetalean cone as fossil evidence for supporting current molecular phylogeny. *Ann. Bot. (Oxford)* 94: 281–288. https://doi.org/10.1093/aob/mch138

**Wendel, J.F.** 2015. The wondrous cycles of polyploidy in plants. *Amer. J. Bot.* 102: 1753–1756. https://doi.org/10.3732/ajb.1500320

**Wendt, T.** 1993. A new variety of *Ephedra torreyana* (Ephedraceae) from West Texas and Chihuahua, with notes on hybridization in the *E. torreyana* complex. *Phytologia* 74: 141–150.

**Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., Ayyampalayam, S., Barker, M.S., Burleigh, J.G., Gitzendanner, M.A., Ruhfel, B.R., Wafula, E., Der, J.P., Graham, S.W., Mathews, S., Melkonian, M., Soltis, D.E., Soltis, P.S., Miles, N.W., Rothfels, C.J., Pokorny, L., Shaw, A.J., DeGironimo, L., Stevenson, D.W., Surek, B., Villarreal, J.C., Roure, B., Philippe, H., dePamphilis, C.W., Chen, T., Deyholos, M.K., Baucom, R.S., Kutchan, T.M., Augustin, M.M., Wang, J., Zhang, Y., Tian, Z.J., Yan, Z.X., Wu, X.L., Sun, X., Wong, G.K.S. & Leebens-Mack, J.** 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc. Natl. Acad. Sci. U.S.A.* 111: E4859–E4868. https://doi.org/10.1073/pnas.1323926111

**Wilson, L.R.** 1959. Geological history of the Gnetales. *Oklahoma Geol. Notes* 19: 35–40.

**Wilson, L.R.** 1962. Permian plant microfossils from the Flowerpot Formation Oklahoma, *Oklahoma Geol. Surv. Circ.* 49: 5–50.

**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. *Molec. Phylogen. Evol.* 52: 115–124. https://doi.org/10.1016/j.ympev.2008.12.026

**Wu, H., Ma, Z., Wang, M.-M., Qin, A.-L., Ran, J.-H. & Wang, X.-Q.** 2016. A high frequency of allopolyploid speciation in the gymnospermous genus *Ephedra* and its possible association with some biological and ecological features. *Molec. Ecol.* 25: 1192–1210. https://doi.org/10.1111/mec.13538

**Xi, Z.X., Rest, J.S. & Davis, C.C.** 2013. Phylogenomics and coalescent analyses resolve extant seed plant relationships. *PLoS ONE* 8: e80870. https://doi.org/10.1371/journal.pone.0080870

**Yang, Y., Geng, B.Y., Dilcher, D.L., Chen, Z.D. & Lott, T.A.** 2005. Morphology and affinities of an Early Cretaceous *Ephedra* (Ephedraceae) from China. *Amer. J. Bot.* 92: 231–241. https://doi.org/10.3732/ajb.92.2.231

**Yang, Y., Lin, L.B. & Wang, Q.** 2013. *Chengia laxispicata* gen. et sp nov., a new ephedroid plant from the Early Cretaceous Yixian Formation of western Liaoning, Northeast China: Evolutionary, taxonomic, and biogeographic implications. *B. M. C. Evol. Biol.* 13: 72. https://doi.org/10.1186/1471-2148-13-72

**Yang, Y., Lin, L.B., Ferguson, D.K. & Wang, Y.W.** 2018. Macrofossil evidence unveiling evolution of male cones in Ephedraceae (Gnetidae). *B. M. C. Evol. Biol.* 18: 125. https://doi.org/10.1186/s12862-018-1243-9

**Yang, Y., Wang, Y. & Ferguson, D.K.** 2020. A new macrofossil ephedroid plant with unusual bract morphology from the Lower Cretaceous Jiufotang Formation of northeastern China. *B. M. C. Evol. Biol.* 20: 19. https://doi.org/10.1186/s12862-019-1569-y

**Yang, Z. & Rannala, B.** 1997. Bayesian phylogenetic inference using DNA sequences: A Markov Chain Monte Carlo method. *Molec. Biol. Evol.* 14: 717–724. https://doi.org/10.1093/oxfordjournals.molbev.a025811

**Yule, G.U.** 1925. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F. R. S. *Philos. Trans., Ser. B.* 213: 21–87. https://doi.org/10.1098/rstb.1925.0002

**Zhong, B.J., Yonezawa, T., Zhong, Y. & Hasegawa, M.** 2010. The position of Gnetales among seed plants: Overcoming pitfalls of chloroplast phylogenomics. *Molec. Biol. Evol.* 27: 2855–2863. https://doi.org/10.1093/molbev/msq170

**Zhong, B.J., Deusch, O., Goremykin, V.V., Penny, D., Biggs, P.J., Atherton, R.A., Nikiforova, S.V. & Lockhart, P.J.** 2011. Systematic error in seed plant phylogenomics. *Genome Biol. Evol.* 3: 1340–1348. https://doi.org/10.1093/gbe/evr105

**Zhu, A.D., Guo, W.H., Gupta, S., Fan, W.S. & Mower, J.P.** 2016. Evolutionary dynamics of the plastid inverted repeat: The effects of expansion, contraction, and loss on substitution rates. *New Phytol.* 209: 1747–1756. https://doi.org/10.1111/nph.13743

**Zimmer, E.A. & Wen, J.** 2012. Using nuclear gene data for plant phylogenetics: Progress and prospects. *Molec. Phylogen. Evol.* 65: 774–785. https://doi.org/10.1016/j.ympev.2012.07.015

**Appendix 1.** Taxon sampling and GenBank accessions.

The following information is given (separated by commas): ***Taxon***, *DNA voucher*, location data (country), laboratory identity, plastid accession number, rDNA cistron accession number. Missing information is indicated by –; sequences downloaded from GenBank are indicated by *. Additional information, including taxon distribution and collection area and year, is available in supplementary Table S1.

***Ephedra alata*** Decne., *Anderberg 480* (S), Algeria, CR83, MG594447, MG594395, ***E. altissima*** Desf., *Samuelsson 6227* (S), Algeria, CR81, MG594448, MG594396, ***E. altissima*** Desf., *Bot. Dept. SU S-C-7688* (S), Tunisia, CR82, MG594449, MG594397, ***E. americana*** Humb. & Bonpl. ex Willd., *Ståhl 8049* (S), Ecuador, CR468, MG594450, MG594398, ***E. andina*** Poepp. ex C.A.Mey., *Eggli & Leuenberger 1711* (Z), Chile, CR204, MG594451, MG594399, ***E. antisyphilitica*** Berland. ex C.A.Mey., *Hoggard 451* (S), U.S.A. (Oklahoma), CR64, MG594452, MG594400, ***E. aphylla*** Forssk., *Amdurski & Segal 402* (S), Israel, AK154, MG594453, MG594401, ***E. aphylla*** Forssk., *Anderberg 853* (S), Libya, CR80, MG594454, MG594402, ***E. breana*** Phil., *Hunziker 1852* (S), Argentina, CR232, MG594455, MG594403, ***E. californica*** S.Watson, *Stedje s.n.* (O 68-154; from Cult.), CR34, MG594495, MG594404, ***E. chilensis*** C.Presl., *Forbes s.n.* (UC 49.0542; from Cult.), Chile, CR75, MG594456, MG594405, ***E. ciliata*** Fisch. & C.A.Mey., *Rechinger 16183* (S), Afghanistan, CR86, MG594457, MG594406, ***E. coryi*** Reed, *Correll 32785* (S), U.S.A. (Texas), CR244, MG594458, MG594407, ***E. distachya*** L., *Bolinder 720* (S), Greece, CR465, MG594460, MG594409, ***E. distachya*** L., *Rechinger 53066* (WU), Iran, CR470, MG594486, MG594436, ***E. equisetina*** Bunge, *Rydin 16* (S; from Cult.), Turkmenistan, CR04, MG594461, MG594410, ***E. equisetina*** Bunge, –, –, –, NC_011954*, –, ***E. fedtschenkoi*** Paulsen, *Smith 11694* (UPS), China, CR464, MG594462, MG594411, ***E. foeminea*** Forssk., *Bolinder 543* (S), Greece, KB543, NC_029347*, MG594412, ***E. foeminea*** Forssk., *Osvald 26068* (UPS), Israel, CR93, MG594463, MG594413, ***E. foeminea*** Forssk., *Rydberg 6* (S; from Cult.), Italy, OT12, MG594464, MG594414, ***E. foliata*** Boiss. ex C.A.Mey. *Hedberg & Hedberg 92019A* (UPS), Saudi Arabia, CR95, MG594465, MG594415, ***E. fragilis*** Desf., *Jonsell 5412* (UPS V-54673), Morocco, CR101, MG594466, MG594416, ***E. fragilis*** Desf., *T. Denk & G. Gruber 178-02* (S), Morocco, OT09, MG594467, MG594417, ***E. frustillata*** Miers, *Rydin s.n.* (S S04-482; from Cult.), Argentina, CR314, MG594468, MG594418, ***E. funerea*** Coville & C.V.Morton, *Rose 67021* (S), U.S.A. (California), CR245, MG594469, MG594419, ***E. gerardiana*** Wall. ex Klotzsch & Garcke, *Bosshard & al. 802.67* (Z), Pakistan, CR238, MG594470, MG594420, ***E. intermedia*** Schrenk & C.A.Mey., *Rydin 64a* (S 03-925; from Cult.), –, CR06, MG594471, MG594421, ***E. likiangensis*** Florin, *Rydin 273* (S; from Cult.), China, CR273, MG594472, MG594422, ***E. lomatolepis*** Schrenk, *Baitulin & al. 276414* (UPS), Kazakhstan, CR92, MG594473, MG594423, ***E. major*** Host, *Thureborn & Norbäck Ivarsson 9* (S), Croatia, OT28, MG594474, MG594424, ***E. milleri*** Freitag & Maier-St., *Miller 7667A* (E), Oman, CR116, MG594475, MG594425, ***E. minuta*** Florin, *Rydin 259* (S; from Cult.), China, CR259, MG594476, MG594426, ***E. monosperma*** J.G.Gmel. ex C.A.Mey., *Honegger 92/111* (Z), Russia, CR168, MG594477, MG594427, ***E. nevadensis*** S.Watson, *Balls 10689a* (S), U.S.A. (California), CR247, MG594478, MG594428, ***E. pachyclada*** Boiss., *Danin s.n.* (S S-2455), Egypt, CR89, MG594479, MG594429, ***E. pedunculata*** Engelm. ex S.Watson, *M.T. Edwards 385* (S), Mexico, CR171, MG594480, MG594430, ***E. przewalskii*** Stapf, *Bartholomew 8292* (Z), China, CR187, MG594481, MG594431, ***E. przewalskii*** Stapf, *Bosshard & al. 802.71* (Z), Pakistan, CR241, MG594482, MG594432, ***E. rupestris*** Benth., *Ornduff 9675* (UC 87.1368; from Cult.), Ecuador, CR73, MG594483, MG594433, ***E. sinica*** Stapf, *Schönenberger s.n.* (S), China, CR33, MG594484, MG594434, ***E. somalensis*** Freitag & Maier-St., *Thulin 10925A* (UPS), Somalia, CR90, MG594485, MG594435, ***E.* sp.**, *Söderbom 7439* (S), China, CR255, MG594459, MG594408, ***E. strobilacea*** Bunge, *Rechinger & Rechinger 2703* (S), Iran, CR107, MG594487, MG594437, ***E. strobilacea*** Bunge, *Andersen & Jensen 7298* (L 0790043), Iran, CR471, MG594488, MG594438, ***E. torreyana*** S.Watson, *Rydin 469* (S), U.S.A. (Utah), CR469, MG594489, MG594439, ***E. transitoria*** Riedl, *Collonette 9095B* (E), Saudi Arabia, CR111, MG594490, MG594440, ***E. trifurca*** Torr., *Miller & Harder 8187* (MO 04630447), U.S.A. (Arizona), CR70, MG594491, MG594441, ***E. tweediana*** Fisch. & C.A.Mey., *Forbes s.n.* (UC 66.0742; from Cult.), Argentina, CR76, MG594492, MG594442, ***E. viridis*** Coville, *Rydin 262* (S; from Cult.), U.S.A., CR262, MG594493, MG594443, ***E. viridis*** Coville, *Holmgren 1826* (UPS), U.S.A. (Utah), CR91, MG594494, MG594444. **OUTGROUP:** ***Gnetum gnemon*** L., *Hou & Pan CXS019* (HITBC), China, CH102, NC_026301*, MG594391, ***G. luofuense*** C.Y.Cheng, *Hou CHK006* (S), China, CH107, KX385193*, MG594392, ***G. montanum*** Markgr., *Hou & Pan CXS036* (HITBC), China, CH105, KX385195*, MG594393, ***G. parvifolium*** (Warb.) W.C.Cheng, *Hou & Lau CHK030* (S), China, CH109, NC_011942*, MG594394, ***Welwitschia mirabilis*** Hook.f., Stockholm University (in Cult.), Namibia, CR36, NC_010654*, MG594390, ***Ginkgo biloba*** L., –, –, AB684440*, –, ***Cycas taitungensis*** C.F.Shen, K.D.Hill, C.H.Tsou & C.J.Chen, –, –, NC_009618*, –, ***Pinus koraiensis*** Siebold & Zucc., –, –, NC_004677*, –, ***Larix decidua*** Mill., –, –, NC_016058*, –, ***Cephalotaxus wilsoniana*** Hayata, –, –, NC_016063*, –.