Distribution of living Cupressaceae reflects the breakup of Pangea

Citation for published version:

Digital Object Identifier (DOI):
10.1073/pnas.1114319109

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Proceedings of the National Academy of Sciences

Publisher Rights Statement:
Freely available online through the PNAS open access option.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Distribution of living Cupressaceae reflects the breakup of Pangea

Kangshao Mao,a,b,c 1 Richard I. Milne,a,b,c 1 Libing Zhangd,e, Yanling Peng,1 Jianquan Liua,2 Philip Thomas,c Robert R. Millia and Susanne S. Renner1

a State Key Laboratory of Grassland Agro-Ecosystem, School of Life Sciences, Lanzhou University, Lanzhou, Gansu 730000, People’s Republic of China; b Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh EH3 3LR, United Kingdom; c Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, Sichuan 610041, People’s Republic of China; d Missouri Botanical Garden, St. Louis, MO 63166; e Systematic Botany and Mycology, Department of Biology, University of Munich, 80638 Munich, Germany

Most extant genus-level radiations in gymnosperms are of Oligocene age or younger, reflecting widespread extinction during climate cooling at the Oligocene/Miocene boundary [~23 million years ago (Ma)]. Recent biogeographic studies have revealed many instances of long-distance dispersal in gymnosperms as well as in angiosperms. Acting together, extinction and long-distance dispersal are likely to erase historical biogeographic signals. Notwithstanding this problem, we show that phylogenetic relationships in the gymnosperm family Cupressaceae (162 species, 32 genera) exhibit patterns expected from the Jurassic/Cretaceous breakup of Pangea. A phylogeny was generated for 122 representatives covering all genera, using up to 10,000 nucleotides of plastid, mitochondrial, and nuclear sequence per species. Relying on 16 fossil calibration points and three molecular dating methods, we show that Cupressaceae originated during the Triassic, when Pangea was intact. Vicariance between the two subfamilies, the Laurasian Cupressoideae and the Gondwanan Callitroideae, occurred around 153 Ma (124–183 Ma), when Gondwana and Laurasia were separating. Three further intercontinental disjunctions involving the Northern and Southern Hemisphere are coincident with or immediately followed the breakup of Pangea.

ancstral areas reconstruction | molecular clock

Between the Early Triassic and the Middle Jurassic, virtually all continents were joined to form the supercontinent Pangaea (1–3). Around 160–138 million years ago (Ma) (1, 3), Pangaea broke up into two supercontinents: Laurasia, comprising land that eventually gave rise to North America, Europe, and much of Asia, and Gondwana, made up of land that subsequently gave rise to South America, Africa, India, Antarctica, and Australia. Bi stratigraphic data suggest that Late Triassic and Early Jurassic Pangaea had a warm and equable climate without glaciation or sea ice and that it lacked significant geographic barriers from pole to pole (4). However, because of Pangaea’s great latitudinal expanse, faunal provinces already had developed before its break-up, and dated molecular phylogenies of reptiles, amphibians, and mammals have made clear that subsequent lineage divergence within these groups matches the separation and fragmentation of Laurasia and Gondwana (5–10). Until now, there has been no equivalent evidence for any plant family.

The fossil record shows that gymnosperms dominated the vegetation of Pangaea but declined in dominance and abundance from the Mid-Cretaceous onwards (11, 12). Perhaps because of the extinction of entire clades, molecular-clock studies of gymnosperms consistently have inferred young, usually Oligocene, ages for the crown groups of living genera, e.g., Phyllocladus (13), Cetina (14), Cedrus (15), Agathis (16, 17), Ephedra (18), Juniperus (19), Pseudotsuga (20), Podocarpus, Nageia, Dacrydium, Dacrycarpus (21), and Pinus subgenus Pinus and Strobos (22, 23). Radiations are especially young in the cycads (24–26).

Among the few spermatophyte clades that still may reflect events related to the break-up of Pangaea is the conifer family Cupressaceae (including the former Taxodiaceae) (27–31). Cupressaceae occur on all continents except Antarctica and comprise 162 species in 32 genera (see Table S2 for subfamilies, genera, and species numbers). The family has a well-studied fossil record going back to the Jurassic (32–36). Using ancient fossils to calibrate genetic distances in molecular phylogenies can be problematic, because the older a fossil is, the more likely it is to represent an extinct lineage that diverged somewhere along the line leading to the extant taxon with which it is being compared (37). However, probability distributions on fossil calibration ages allow some manipulation of this uncertainty (38), and judicious use of multiple fossils also may help circumvent calibration pitfalls (39).

Here we present a phylogeny for 122 species from the 32 genera of Cupressaceae (plus 22 species representing relevant outgroups) and use 16 fossil calibration points and three dating approaches to estimate divergence times in the Cupressaceae. We then perform ancestral area reconstructions (AARs) using maximum likelihood based on datasets with or without incorporated fossils. Possible changes in diversification rates were inferred with an approach that accounts for nonrandom taxon sampling in molecular phylogenies (40). Nonrandom sampling arises when phylogenies include at least one species per genus but not all congeneric, thereby over representing deep nodes (diversification events) in the tree. Experiments have confirmed the theoretical expectation that such sampling leads to the erroneous inference of diversification rate downturns (41, 42). We aimed to test the hypothesis that, given their fossil record, the deepest Cupressaceae divergences should reflect the break-up of Pangaea and that evolution of the family then continued on the separating continental landmasses.

Results

Cupressaceae Phylogenetics. After sequence alignment and removal of ambiguous regions, we obtained two datasets, one of 56 taxa and 10,472 aligned nucleotides from plastid, mitochondrial, and nuclear DNA, the other of 144 taxa and 7,171 nucleotides from plastid DNA only. Maximum likelihood, parsimony, and Bayesian optimization inferred similar topologies from both datasets. Support

www.pnas.org/cgi/doi/10.1073/pnas.1114319109

PNAS | May 15, 2012 | vol. 109 | no. 20 | 7793–7798


This article is a PNAS Direct Submission. C.C.D. is a guest editor invited by the Editorial Board.

Free download online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (JF725972–JF725993). GenBank accession numbers and provenance of sequenced samples are provided in Table S1.

1K.M. and R.I.M. contributed equally to this work.

2To whom correspondence should be addressed. E-mail: liujq@nwipb.ac.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1114319109/-/DCSupplemental.
values for major groups are high, with three-quarters of the genus-level nodes having >95% posterior probability (Fig. S1).

**Divergence Times.** Bayesian coestimation of topology and divergence time (BEAST, using uniform prior distributions on calibration ages) (43) generally gave the oldest ages, and Penalized likelihood (44, 45) gave the youngest (Table S3). An alternative Bayesian approach, which used a topological constraint (MULTI-DIVTIME) (46), yielded ages for short-branched nodes (most nodes within Cupressoideae; Fig. S1) that were similar to or younger than those obtained with BEAST; ages for long-branched nodes (most nodes within Callitroideae; Fig. S1) were similar to or older than those obtained with BEAST. Confidence intervals around estimates from the two Bayesian approaches overlapped (Table S3). With all three dating approaches, the more densely sampled 144-taxon dataset produced slightly older age estimates (compare Fig. S2 A and B), a result that is consistent with the effects of undersampling observed elsewhere (47). Because BEAST allows more complex nucleotide-substitution models than do the other two dating approaches, and because dates from the 56-taxon matrix might be less accurate because of undersampling, the following discussion focuses on the results obtained with the 144-taxon matrix analyzed using BEAST (Fig. 1 and Fig. S2B).

Cupressaceae split from their sister lineage during the late Permian and early Triassic (209–282 Ma; node 1 in Fig. 1, Table 1, and Table S3) and began to diversify into seven major lineages (commonly ranked as subfamilies) during the Triassic (184–254 Ma; node 2 in Fig. 1, Table 1, and Table S3) and the Cretaceous/ Tertiary boundary (65.5 ± 0.3 Ma) (48) (Fig. S2B); the only exception is *Chamaecyparis*, which is dated to 61–108 Ma (node 23 in Fig. S2B; note that this is the crown age for the *Chamaecyparis-Fokienia* clade).

BEAST analyses with different uniform distribution priors on the calibration closest to the root (calibration point P; Figs. S2 and S3) yielded largely overlapping 95% highest posterior density (HPD) age ranges for all nodes of interest, indicating that the chosen maximum constraint (the only such constraint used in the analysis) had no overall strong effect on the remaining dates (Fig. S3; compare run 1 with runs 2–4). BEAST analyses with different subsets of calibration points, all with uniform priors, showed that calibration P plus calibrations A, B, E, F, G, J, K, and L (subset VND; *SI Text* and Table S4) (Fig. S3, run 5) yielded node ages similar to those obtained with calibrations A through P (Table S4) (Fig. S3, run 1), whereas calibration P plus calibrations C, D, H, I, M, N, O, and P (subset NVND; *SI Text* and Table S4) gave much younger ages (Fig. S3, run 6).

A BEAST analysis that used lognormal prior distributions on the ages of calibration P and subset VND (and uniform priors for subset NVND) (run 7) generated age estimates younger than but largely overlapping those obtained with uniform priors for calibration A through P (run 1) (Fig. S3, Table 1, and Tables S3 and S4; see *SI Text* for a detailed comparison). In all nine BEAST analyses, effective sample sizes for each parameter were well above 200.

**Ancestral Areas and Diversification Rate Changes.** Likelihood AARs were implemented under the dispersal-extinction-cladogenesis model in LAGRANGE (49). We defined eight continent pairs (NS, NE, SF, NA, AE, FE, SU, NF; area codes are explained in *Materials and Methods* and are illustrated in Fig. 2A) and one continent group (NAE), which reflect continental connections known from plate tectonics (1, 3). The most likely scenarios (Fig. 2B) required 31 dispersal events, 21 vicariance events, and five local extinctions. Likelihood AAR for living Cupressaceae (Fig. S2B) suggested that the family originated in Asia and its early members expanded to North America from where Callitroideae and *Athrotaxis* entered Gondwanan South America. The integration of fossil Cupressaceae (Fig. 2C–E, Fig. S4, and Table...
Austrohamia minuta
Austrosequoia
138 Ma)
B
– Athrotaxis
C and its extant sister lineage, comprising
183 Ma (node 7 in Fig. 1 and Table 1; mean: 153
7 in Fig. 1, – 282 Ma), while Pangea was still
E
and its putative sister species
– Sewardiodendron laxum
currently is endemic
6i n PNAS
100 Ma, as judged from the mid-Cretaceous
– dispersed from North America (via South America) to
(Fig. 2
fi
| 126 Ma) (48) of North America (53). The Gondwanan
– 26), and a recent meta-analysis
†
S5 in likelihood AAR resulted in a similar scenario but with two
more incursions from Laurasian to Gondwanan continents
(Austrostroesequia and Austrohamia minuta). Reconstructions using
alternative placements of certain fossil taxa (two Austrohamia
species and Sewardiodendron laxum) affected only the ancestral
areas inferred for the nodes closest to them (compare Fig. 2 B
and C–E).
Because the 122 species that we sequenced are not a random
sample of the family’s 162 species but instead overrepresent deep
nodes, we fitted birth/death diversification models to the maxi-
mum likelihood topology after statistical completion, assuming
nonrandom species sampling (40). When the nonsequenced 40
species were added to the tree under a constant-rate birth-death
model, assuming they originated during the past 10 million y, the
best fit to the 1,000 simulated completed trees was a two-rate
model with a decrease in diversification rate at 1.37 Ma. Of the
alternative models [constant-rate pure-birth (CR-PB), constant-
rate birth-death, logistic density dependence, and exponential
density dependence], the CR-PB model provided the second-
best fit to the completed trees.

**Discussion**
The dense taxon sampling and large amount of sequence data used
here yielded a solidly supported phylogeny for the Cupressaceae
(Fig. S1), which are monophyletic and sister to the Taxaceae sensu
lato (i.e., Taxaceae sensu stricto plus Cephalotaxaceae and Amen-
totaxaceae) (31), as found previously (50, 51). The divergence of
Cupressaceae from their sister lineage occurred >200 Ma (node
1 in Fig. 1 and Table 1; 209–282 Ma), while Pangea was still
intact, matching fossil evidence of Cupressaceae in the Jurassic of
Europe (Hughmilleries juddii) (35), Asia (Sewardiodendron laxum
and Austrohamia acanthobractea) (32, 36), and South America
(Austrohamia minuta) (34). Cupressaceae diversified into seven
major lineages (subfamilies) during the Triassic and Jurassic (nodes
2–7 in Fig. 1, Fig. S2, Table 1, and Table S3), predating or coinciding
with the separation of Gondwana and Laurasia (orange column in
Fig. 1). Furthermore, AAR (with or without fossil taxa) yielded Asia
as the ancestral area for the family (Fig. 2). Cunninghamhamioidae may
have originated in Asia (Fig. 2), and the divergence of Taiwanoii-
deae from their sister lineage (Fig. 2B) may match the separation of
Asia from North America at ∼200 Ma (3); the three subfamilies,
Athrortoxioideae, Sequoioidae, and Taxidoideae (nodes 4–6 in
Fig. 1) probably diverged from their sister lineage in North America
(Fig. 2 C–E); the divergence of Callitriodeae from Cupressaceae was
dated to 124–183 Ma (node 7 in Fig. 1 and Table 1; mean: 153
Ma), an age range almost coinciding with the separation of Gond-
wana from Laurasia (Fig. 1) during the Late Jurassic (160–138 Ma)
(1, 3). Living members of Cupressaceae occur mainly in former
Laurasian continents (Fig. 1 Lower, B), whereas Callitriodeae are endemic to fragments of Gondwana (30) (Fig. 1 Lower, E). African
Cupressaceae apparently derived from a series of southward expansions during the middle and late Tertiary (Fig. 1). Un-
ambiguous fossils (with reproductive organs) of Cupressaceae are
known only from former Laurasia and those of Callitriodeae from
Gondwana (33). Overland connections between Laurasian and
Gondwanan continents were severed from the Late Jurassic until
the middle Tertiary, when India connected with Eurasia, followed by
the subsequent reconnection of Africa to Eurasia and South
America with North America (1, 3). It is clear from our results that
the divergence between Cupressoideae and Callitriodeae correlates
with the break-up of Pangea (Fig. 1) and most likely was caused by
it, as shown in the likelihood AARs (Fig. 2B).
In the remaining five subfamilies, we further inferred three
intercontinental disjunctions between the Northern and Southern
Hemispheres (Fig. 2 and Fig. S4). The most recent involves the
extinct Austrostroesequia and its extant sister lineage, comprising
Sequoia and Sequoiadendron (Fig. 2 C–E and Fig. S4). Austro-
sequoia dispersed from North America (via South America) to
Australia around 94–100 Ma, as judged from the mid-Cretaceous
(Cenomanian) fossil remains in Australia (33, 52). The second
disjunction involves Athrotaxis and its putative sister species
Athrortoxies berryi (53) (Fig. S5F). Athrotaxis currently is endemic
to Australia (30) but is known from fossils in Argentina (Athro-
toxies ungerii) (54), and Athrotaxis berryi is known from the Aptian
(ca. 111–126 Ma) (48) of North America (53). The Gondwanan
taxon Athrotaxis probably originated from a southward expansion
from North America, as suggested by likelihood AARs (Fig. 2 C–
E). The third inferred intercontinental disjunction involves the extinct Austrohamia, with one species (Austrohamia minuta)
from the Jurassic of southern Argentina (34) and the other (Austro-
hamia acanthobractea) from the late Jurassic of northern China
(30). The South American A. minuta might have arrived there
following dispersal from the Laurasian North America, as sug-
gested by likelihood AAR (Fig. 2 C–E). These three instances of
intercontinental disjunctions all involve north-to-south expan-
sion. We found no instance of range expansion from the Southern
to the Northern Hemisphere but detected a clear signal of dis-
persal or overland expansion among the Southern Gondwanan
continents themselves (Fig. S4).
Previous studies of gymnosperm radiations mostly have inferred
Oligocene-age crown groups (14–26), and a recent meta-analysis
found a median crown age for gymnosperm genera of 32 Ma,
younger than that found for angiosperm genera (25). Our dating of
Oligocene and younger than that found for angiosperm genera (25). Our dating of

**Table 1. Divergence times for the Cupressaceae obtained under a Bayesian relaxed clock as implemented in the program BEAST**

<table>
<thead>
<tr>
<th>Node</th>
<th>Node description</th>
<th>Uniform priors (P, subset VND)†</th>
<th>Lognormal priors (P, subset VND)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Stem lineage of Cupressaceae</td>
<td>245 (209–282)</td>
<td>242 (194–293)</td>
</tr>
<tr>
<td>2</td>
<td>Crown lineage of Cupressaceae (Stem of Cunninghamhamioidae)</td>
<td>219 (184–254)</td>
<td>211 (168–259)</td>
</tr>
<tr>
<td>3</td>
<td>Stem of Taiwanioideae</td>
<td>204 (170–238)</td>
<td>195 (157–240)</td>
</tr>
<tr>
<td>4</td>
<td>Stem of Athrotaxioideae</td>
<td>190 (157–224)</td>
<td>182 (145–222)</td>
</tr>
<tr>
<td>5</td>
<td>Stem of Sequoioidae</td>
<td>183 (150–215)</td>
<td>174 (139–213)</td>
</tr>
<tr>
<td>6</td>
<td>Stem of Taxidoideae</td>
<td>170 (140–201)</td>
<td>159 (128–194)</td>
</tr>
<tr>
<td>7</td>
<td>Divergence between Cupressoideae and Callitriodeae</td>
<td>153 (124–183)</td>
<td>143 (114–175)</td>
</tr>
<tr>
<td>8</td>
<td>Crown lineage of Callitriodeae</td>
<td>128 (98–159)</td>
<td>121 (92–152)</td>
</tr>
<tr>
<td>9</td>
<td>Stem lineage of Widdringtonia (endemic in southern Africa)</td>
<td>65 (42–92)</td>
<td>62 (40–86)</td>
</tr>
<tr>
<td>10</td>
<td>Crown lineage of Cupressoideae</td>
<td>134 (104–164)</td>
<td>123 (93–154)</td>
</tr>
</tbody>
</table>

*Million year ranges in parentheses denote the 95% HPD.
*BEAST run 1 as described in **SI Text** and Fig. S3 in which uniform priors were applied to calibration P, subset VND (C, D, H, I, M, N, and O) and subset NVND (A, B, E, F, G, J, K, and L).
*BEAST run 7 as described in **SI Text** and Fig. S3 in which lognormal priors were assigned to calibration P and subset VND, whereas uniform priors were retained for subset NVND; see **SI Text** for a full explanation of all runs.
with Sequoioideae also found in Australia, but now are reduced to two species each in southern North America and one and two species, respectively, in East Asia (Fig. 2 C–E and Fig. S6 B and C). Athrotaxidoideae were present in both North and South America during the Cretaceous (53, 54) but today consist of three species in Australia (Fig. 2 C–E). Finally, the genera
Austrocedrus (56), Calocedrus (57), Chamaecyparis (58), Fitzroya (59), Pseudotsuga (60), and Tetraclinus (61) all had wider distributions in the past. The ice ages of the past 2 million years further contributed to population extinction and reductions in species range, as inferred from our diversification modeling, which indicated a downturn in Cupressaceae diversification rates at 1.37 Ma. Relatively few Cupressaceae lineages have adapted to the strongly seasonal and semiarid climatic conditions that became more widespread with the global cooling during the Oligocene/Miocene (30). Those that did, such as Junipers, experienced a diversification burst during the Miocene (19).

Besides throwing light onto the diversification of a Triassic/Jurassic Pangean spermatophyte lineage, our findings confirm and illustrate the power of incorporating fossils directly into AARs, rather than using them only for molecular-clock calibration (62, 63). Specifically, it was the incorporation in the AARs of up to 29 fossil taxa (groups) (Fig. S4 and Table S5), most from areas where the respective lineage no longer occurs, that provided insights about range changes, but with the direction of range expansion (predominantly north-to-south) being inferred less from the fossil record than the molecular topology.

Materials and Methods

Plant Material, DNA Isolation and Sequencing, and Sequence Alignment. Table S1 lists all plant materials used in this study, with species name and author, geographic provenance, herbarium voucher and deposition, and GenBank accession numbers. A total of 290 sequences were newly produced. Phylogenetic and dating analyses were conducted on two datasets. The 56-taxon dataset comprised 35 ingroup species, 21 outgroups, and 10,472 aligned nucleotides from 10 DNA regions (see below). The 144-taxon dataset comprised 122 ingroup species, 22 outgroups, and 7,171 aligned nucleotides from six plastid DNA regions. The ingroup species represented all 32 Cupressaceae genera; outgroups represented the other conifer families, Cycas, Ginkgo, and a basal angiosperm for rooting purposes. We sequenced the mitochondrial regions atpA and cox1, the nuclear regions 18S and 26S, and the plastid regions matK, rbcL, psbA, psbB, psbD, rps4, and trn-L-F (for primer sequences, see refs. 19 and 64). For DNA extraction, PCR, and sequencing procedures, see Mao et al. (19). The sequences produced were aligned using ClustalX version 1.83 (65), followed by manual adjustments in Mega4 (66).

Phylogenetic Analyses. Phylogenetic relationships were reconstructed using parsimony, Bayesian, and maximum likelihood inference. Parsimony analyses relied on PAUP version 4.10b (67) and the University of Oslo Biportal (http://www.bioportal.uio.no) using heuristic searching, starting trees obtained via stepwise addition, tree-bisection-reconnection branch swapping, steepest descent, and the MulTrees and Collapse options in effect, with no upper limit for the number of trees held in memory; support values for all nodes (on a 50% majority rule bootstrap tree) were calculated with the same settings as above for 1,000 replicates; 10 searches with random taxon additions were conducted for each replicate, and the strict consensus tree of all shortest trees were saved. Bayesian analysis relied on MrBayes version 3.1.2 (68) and the GTR+G model as suggested by MrModeltest version 2.3 (70). We used the default of one cold and three heated Markov chain Monte Carlo chains, starting from random initial trees, and chains were run for 6,000,000 generations, sampling every 200th. The default options in MrBayes were used for chain heating and mixing. We discarded a burn-in of the first 2,000,000 generations and used 20,000 trees from the posterior distribution to obtain a majority rule consensus tree. Maximum likelihood analyses relied on Garli version 1.0 (71) with the GTR+K substitution model, starting from random trees and using 5,000,000 generations per search; 30 independent searches were performed, and the best tree was saved. Separate phylogenetic analyses of the nuclear, plastid, and mitochondrial datasets did not yield statistically supported (>75% likelihood bootstrap support) topological contradictions (data available upon request). Therefore we combined the three data partitions in the 56-taxon dataset.

Molecular-Clock Models and Calibration. A likelihood ratio test in PAUP 4.10b (67, 72) suggested that the 56-taxon and 144-taxon datasets reject a strict molecular clock (P < 0.01), and we therefore used relaxed molecular-clock approaches: Bayesian coalescent of branch lengths and topology with uncorrelated lognormally distributed rates in BEAST 1.5.3 (48), Bayesian estimation with an input phylogeny in MULTIDIVTIME (46), and penalized likelihood rate smoothing in R8S (44, 45). In each case, genetic distances were transformed into absolute time (in million years) by using 16 fossil calibration points, of which 12 were within Cupressaceae (Fig. S2 and Table S4). Fossils were assigned to the stem of their most closely related lineages; Table S4 lists the morphological features used for each fossil taxonomic assignment. For BEAST analyses, we used uniform prior distributions for minimum constraint (calibration points A–O), with the younger bound set by the youngest date of the respective fossil and the older bound set to 366.8 Ma (maximum constraint for calibration point P place near the root) (Table S4). Calibration point P was restricted to fall between 306.2 and 366.8 Ma (Table S4). We tested the effects of the outer intervals at calibration point P, calibrations with different subsets of fossils, and different distribution prior for calibrations by carrying out eight additional BEAST runs (SI Text and Fig. S3). BEAST analyses were run on the Cyberinfrastructure for Phylogenetic Research (CIPRES) Science Gateway (http://www.phylo.org/portal2) (73). For MULTIDIVTIME and penalized likelihood, constraints were as in BEAST, except that these programs do not allow specific prior distributions on fossil constraints.

Ancestral Area Reconstructions (AAR). AAR relied on the likelihood dispersal-extinction-cladogenesis approach implemented in LAGRANGE (49). The matrix of migration probabilities among continents in LAGRANGE (SI Text and Table S6) allowed dispersal between six operational geographic areas: E, Europe, north Africa, and northern Arabia; A, Asia; N, North America, Caribbean, and Central America; S, South America; F, south to middle Africa and southern Arabia; and U, Australia, New Guinea, New Caledonia, and New Zealand (see Fig. 2A). Boundaries between A, E, and F were defined to minimize the number of species that fell in two areas. The northern boundary between A and E was defined by the Ural Mountains, which is the conventional boundary between Europe Russian and Asian Russia. The boundary between E and F is the Tropic of Cancer, which runs along the middle of a broad belt of very low precipitation (<100 mm y−1) stretching across all of North Africa and most of Arabia (74); this belt of low precipitation is a significant biogeographic barrier for Cupressaceae (30).

For details on the selection of fossils, the inference of calibration fossils’ phylogenetic position, cross validation of calibration fossils (Fig. S7), Ancestral Area Reconstructions, and diversification modeling, see SI Text.

ACKNOWLEDGMENTS. We thank Damon Little, Elena Conti, Richard Abbott, the editor, and three anonymous reviewers for their constructive suggestions. The University of Oslo Biportal, CIPRES Science Gateway, and the William Hennig Society provided computation and software resources. This research was supported by the China Scholarship Council for 1 y study at the University of Edinburgh. The Royal Botanic Garden Edinburgh is supported by the Scottish Government’s Rural and Environment Science and Analytical Services Division.


Mao et al.