Syst. Biol. 61(2):289–313, 2012

DOI:10.1093/sysbio/syr116
Advance Access publication on December 26, 2011

Testing the Impact of Calibration on Molecular Divergence Times Using a Fossil-Rich Group: The Case of *Nothofagus* (Fagales)

HERVÉ SAUQUET[1,*], SIMON Y. W. HO[2,3], MARIA A. GANDOLFO[4], GREGORY J. JORDAN[5], PETER WILF[6], DAVID J. CANTRILL[7], MICHAEL J. BAYLY[8], LINDELL BROMHAM[2], GILLIAN K. BROWN[7,8], RAYMOND J. CARPENTER[9], DAPHNE M. LEE[10], DANIEL J. MURPHY[7], J. M. KALE SNIDERMAN[11], AND FRANK UDOVICIC[7]

[1]*Laboratoire Écologie, Systématique, Évolution, Université Paris-Sud, CNRS UMR 8079, 91405 Orsay, France;* [2]*Centre for Macroevolution and Macroecology, Research School of Biology, Australian National University, Canberra, ACT 0200, Australia;* [3]*School of Biological Sciences, University of Sydney, Sydney, NSW 2006, Australia;* [4]*L.H. Bailey Hortorium, Department of Plant Biology, Cornell University, Ithaca, NY 14853, USA;* [5]*School of Plant Science, University of Tasmania, Private bag 55, Hobart, TAS 7001, Australia;* [6]*Department of Geosciences, Pennsylvania State University, University Park, PA 16802, USA;* [7]*National Herbarium of Victoria, Royal Botanic Gardens Melbourne, Private Bag 2000, South Yarra, VIC 3141, Australia;* [8]*School of Botany, The University of Melbourne, Melbourne, VIC 3010, Australia;* [9]*Department of Ecology and Environmental Biology, School of Earth and Environmental Sciences, University of Adelaide, Adelaide, SA 5005, Australia;* [10]*Department of Geology, University of Otago, PO Box 56, Dunedin 9054, New Zealand; and* [11]*School of Geography and Environmental Science, Monash University, Melbourne, VIC 3800, Australia;*
**Correspondence to be sent to: Laboratoire Écologie, Systématique, Évolution, Université Paris-Sud, CNRS UMR 8079, 91405 Orsay, France; E-mail: herve.sauquet@u-psud.fr.*

Hervé Sauquet and Simon Y. W. Ho have contributed equally to this work.

Received 11 February 2011; reviews returned 25 April 2011; accepted 17 November 2011
Associate Editor: Frank E. Anderson

Abstract.—Although temporal calibration is widely recognized as critical for obtaining accurate divergence-time estimates using molecular dating methods, few studies have evaluated the variation resulting from different calibration strategies. Depending on the information available, researchers have often used primary calibrations from the fossil record or secondary calibrations from previous molecular dating studies. In analyses of flowering plants, primary calibration data can be obtained from macro- and mesofossils (e.g., leaves, flowers, and fruits) or microfossils (e.g., pollen). Fossil data can vary substantially in accuracy and precision, presenting a difficult choice when selecting appropriate calibrations. Here, we test the impact of eight plausible calibration scenarios for *Nothofagus* (Nothofagaceae, Fagales), a plant genus with a particularly rich and well-studied fossil record. To do so, we reviewed the phylogenetic placement and geochronology of 38 fossil taxa of *Nothofagus* and other Fagales, and we identified minimum age constraints for up to 18 nodes of the phylogeny of Fagales. Molecular dating analyses were conducted for each scenario using maximum likelihood (RAxML + r8s) and Bayesian (BEAST) approaches on sequence data from six regions of the chloroplast and nuclear genomes. Using either ingroup or outgroup constraints, or both, led to similar age estimates, except near strongly influential calibration nodes. Using "early but risky" fossil constraints in addition to "safe but late" constraints, or using assumptions of vicariance instead of fossil constraints, led to older age estimates. In contrast, using secondary calibration points yielded drastically younger age estimates. This empirical study highlights the critical influence of calibration on molecular dating analyses. Even in a best-case situation, with many thoroughly vetted fossils available, substantial uncertainties can remain in the estimates of divergence times. For example, our estimates for the crown group age of *Nothofagus* varied from 13 to 113 Ma across our full range of calibration scenarios. We suggest that increased background research should be made at all stages of the calibration process to reduce errors wherever possible, from verifying the geochronological data on the fossils to critical reassessment of their phylogenetic position. [Relaxed molecular clock; calibration; age constraints; confidence intervals; fossil record; geochronology; molecular dating; *Nothofagus*; Fagales.]

Methods for estimating evolutionary timescales from DNA sequence data have become increasingly important over the past few decades (Kumar 2005; Hedges and Kumar 2009). Although a range of molecular dating methods is now available, they all share a vital dependence on age calibrations. DNA sequences only record the number of substitutions that have taken place, inferred from the degree of genetic divergence among taxa. These sequences do not provide separate information about the amount of time over which genetic variation accumulated or about the rate of substitution. Therefore, an absolute timescale is required in order to calibrate molecular date estimates, and geology usually represents the ultimate source of all absolute age data. In most analyses above the species level, the fossil record or particular geological events are used for this calibration (for a recent review, see Ho and Phillips 2009). For example, fossils considered to be phylogenetically close to the bird–mammal divergence have often been used to calibrate estimates of vertebrate evolutionary timescales (e.g., Hedges and Kumar 2003). Although most available molecular dating methods allow the use of multiple calibrations in a single analysis, the practice of using a single calibration point is still widespread. This is often due to the difficulty of finding reliable calibrations for the taxa of interest.

Some general observations about the influence of calibrations have been made in recent studies. First, calibrations at the root of the tree appear to be the most effective for obtaining precise (but not necessarily accurate) date estimates (Hug and Roger 2007; Sanders and Lee 2007; Ho and Phillips 2009). Second, calibrations at

nodes within the ingroup are generally preferred over calibrations within the outgroup (van Tuinen and Hedges 2004), especially in population-level studies, for which deep fossil calibrations are generally inappropriate (Ho and Larson 2006; Ho et al. 2008). Finally, calibrations based on molecular estimates (i.e., "secondary" calibrations) should only be employed cautiously, with particular attention given to the uncertainty associated with the original estimate (Graur and Martin 2004).

The use of fossils as calibration points involves both taxonomic and geological uncertainties. These fundamental issues have not always been acknowledged by biologists using molecular dating methods (for reviews of these issues from the biological perspective, see Magallón 2004; Gandolfo et al. 2008; Forest 2009).

The choice of fossils for calibration often includes a trade-off between how well the fossils are represented and sampled in the record and how precisely fossils can be identified (taxonomic resolution) (Burnham 2008). Taxonomic resolution is dependent on the number and distinctiveness of the characters that can be observed in the fossils as well as whether the character states are phylogenetically informative (synapomorphic states). The use of plant fossils clearly illustrates the compromise between representation and taxonomic resolution. Fossil pollen often has very high representation, especially for wind-pollinated species, and therefore, pollen produced by members of a given clade can appear in the fossil record relatively soon after the origin of that clade. However, fossil pollen is typically identified with relatively low taxonomic resolution. This is exacerbated by the fact that the identification of fossil pollen types is rarely supported by synapomorphies and is often based solely on gross similarity (see Sauquet, Weston, Anderson, et al. 2009 for an exception). As a result, extinct plesiomorphic types can be confused with extant groups. In contrast, macro- and mesofossils (e.g., leaves, flowers, and fruits) can have high taxonomic resolution that is sometimes based on analytical phylogenetic placement and therefore can be attributed to clades with high confidence. However, this is dependent on adequate preservation, and well-preserved macro- and mesofossils typically have much lower representation in the fossil record than pollen (and consequently are less likely to appear close to the origin of the clade). These extremes can be referred to as "safe but late" and "early but risky" options for calibration. Although not all potential calibration points fall into these categories, the dichotomy appears to be widespread. Furthermore, fossils that do not fall into these categories do not pose problems: It is reasonable to expect that old "safe" fossils will be chosen and young "risky" fossils rejected.

In addition to taxonomic uncertainty, there can be large variation in the age control for the fossils themselves, depending on the quality of their associated stratigraphic information in terms of accurate correlation to the geologic timescale (GTS; Gradstein et al. 2004). The absolute timescale (geochronology) is usually based on the parent–daughter ratios of radioactive isotope pairs observed in minerals of volcanic origin, especially from the ^{40}Ar–^{39}Ar (most reliable from sanidine but also from biotite and other minerals) and the ^{238}U–^{206}Pb plus ^{235}U–^{207}Pb series (usually from zircon). Fossils are typically dated by correlating the rocks in which they occur with radiometrically dated rocks and thus to the GTS using a variety of tools, including lithostratigraphic, biostratigraphic, paleomagnetic, and isotopic correlation (Gradstein et al. 2004). The accuracy of geochronologic science and the GTS is improving extremely rapidly and now achieves age precision of <0.1% in many cases. However, especially in older literature from areas not actively investigated for many years, stratigraphic and especially geochronologic information published for fossils is often out of date or inadequately reported. Poor age control can lead to very large uncertainties, and for this reason, it is widely understood by geologists and paleontologists that published ages, especially from older work, require careful geological review, and sometimes a primary reinvestigation, of the bases and criteria for age assignments. Well-studied areas are usually subject to ongoing highly detailed stratigraphic revisions (example given below).

Recent Advances in Molecular Dating Methods and the Application of Age Constraints

In recent years, there have been significant advances in methods for estimating molecular divergence dates. The most important has been the development of relaxed-clock methods, which allow molecular dates to be estimated in the presence of rate heterogeneity among lineages (for reviews, see Sanderson et al. 2004; Welch and Bromham 2005; Rutschmann 2006). A number of novel calibration techniques were introduced with these dating methods, improving upon the earlier practice of fixing nodal ages to errorless point values. In the nonparametric rate smoothing (NPRS) algorithm developed by Sanderson (1997), calibrations could be incorporated in the form of minimum (or maximum) age constraints on nodes. This was a significant advance because it explicitly took into account the fact that fossil evidence is only able to indicate the first confirmed appearance of a lineage, postdating the true, but unknown, divergence age. Subsequent relaxed-clock methods, including penalized likelihood (PL) (Sanderson 2002), the generalized mean path length algorithm (Britton et al. 2007), and Bayesian relaxed clocks (Thorne et al. 1998) also permit the incorporation of multiple age constraints in the tree.

Further developments in calibration techniques came with new Bayesian relaxed-clock methods, which allowed age calibrations to be modeled using parametric distributions (Drummond et al. 2006; Yang and Rannala 2006). These probabilistic calibrations enable the uncertainty in fossil calibrations to be taken into account. For example, lognormal priors allow a higher probability to be given to ages slightly older than that of the fossil (thereby modeling the sampling gap), while at the same time implementing a soft maximum bound with a diminishing tail of probability towards much older ages (thereby restricting the sampling gap). However, the

fossil record seldom offers a reliable quantification of calibration uncertainty (Ho 2007; Gandolfo et al. 2008; Ho and Phillips 2009), unless particular assumptions are made, such as random fossilization (e.g., Marshall 2008). Therefore, although probabilistic calibrations are an intuitively attractive way to incorporate fossil evidence into a molecular dating analysis, they can be difficult to implement in practice.

Although the importance of using suitable calibration techniques is now widely acknowledged, many outstanding questions remain. For instance, in studies calibrated with multiple fossils, how do age constraints interact with one another and what is the influence of individual constraints? Are more age constraints always better? Is outgroup calibration an effective substitute when ingroup calibration is unavailable? And when both outgroup and ingroup calibrations are available, is it better to use both simultaneously? Finally, how reliable and precise is secondary calibration when there is no alternative?

An Example: Nothofagus

To answer some of these questions, we designed an empirical study using the flowering plant genus *Nothofagus* (southern beeches), which comprises 35 species of trees confined to the southern hemisphere and is now assigned to the monogeneric family Nothofagaceae in order Fagales (APG 2009). *Nothofagus* provides an ideal case study for assessing the impact of calibration on molecular dating methods for several reasons. First, large amounts of suitable DNA data are available for phylogenetic analysis (Martin and Dowd 1993; Manos 1997; Setoguchi et al. 1997; Knapp et al. 2005). Second, *Nothofagus* has a rich fossil record that provides numerous potential calibration points representing the full range from safe but late macrofossils to early but risky pollen records. Each of the four subgenera has a distinctive pollen type that is abundantly represented in the fossil record of the southern hemisphere (Dettmann et al. 1990), and many macrofossils of leaves and fruits have been described and assigned to clades within the genus (e.g., Hill 1991). In particular, phylogenetic analyses have been used to place some very well-preserved leaf fossils within the phylogeny of living species (Jordan and Hill 1999). Finally, phylogenetic analyses have unambiguously supported *Nothofagus* as the sister group of all remaining Fagales (Nixon 1989; Manos and Steele 1997; Qiu et al. 1998; Hilu et al. 2003; Li et al. 2004; Herbert et al. 2006). The latter is a large and widespread clade, hereafter referred to as core Fagales, and includes six families (sensu APG 2009), three of which (Betulaceae, Fagaceae, and Juglandaceae) have a notably rich and well-dated fossil record. Furthermore, the relationships of some of these fossils have been explicitly analyzed using phylogenetic methods (Crepet and Nixon 1989; Manos et al. 2007). Core Fagales thus provide a vast choice of potential outgroup calibration points for estimating divergence times in *Nothofagus*.

Nothofagus is also ecologically important, being dominant in many temperate rainforests of Chile, Argentina, Australia, and New Zealand, and some tropical rainforests in New Guinea and New Caledonia. *Nothofagus* has been used as a test case for the classic hypothesis that Gondwanan vicariance can explain major biogeographic patterns and, as a result, its evolutionary history has received intense scrutiny from studies based on its fossil record (e.g., Dettmann et al. 1990; Hill 1991, 2001) and cladistic biogeography (Humphries 1981; Linder and Crisp 1995; Swenson et al. 2001).

Phylogenetic relationships in the genus now appear to be well supported and stable and are consistent with the current classification of *Nothofagus* in four monophyletic subgenera. In particular, all recently inferred phylogenies (Hill and Jordan 1993; Manos 1997; Jordan and Hill 1999; Knapp et al. 2005) support the monophyly of the four subgenera identified by Hill and Read (1991) using leaf and reproductive morphology and by Dettmann et al. (1990) using pollen. There have also been several attempts to use molecular dating methods to test the Gondwanan vicariance hypothesis for *Nothofagus* (Martin and Dowd 1993; Cook and Crisp 2005; Knapp et al. 2005). In each case, estimated ages were considerably too young for Gondwanan vicariance to explain either of the two disjunctions between Australia and New Zealand or the disjunction between New Guinea and New Caledonia. Nevertheless, these studies were all subject to the potential limitations discussed above, in particular with regard to age calibration.

Aims of this Study

The present study has two main objectives: (i) to draw attention to all the background research needed to set up reliable age constraints for molecular dating analyses, and the many questions arising from this research, and (ii) to test empirically the effect of various calibration scenarios on age estimates, reflecting a diversity of situations encountered when applying molecular dating methods. Specifically, with the latter, we aim to test: (i) "ingroup" versus "outgroup" versus "ingroup + outgroup" calibration; (ii) "safe but late" versus "safe but late + early but risky" fossil age constraints; (iii) vicariance versus fossil calibration; and (iv) secondary versus fossil calibration. We also compare the behavior of the two most widely used relaxed-clock dating methods (i.e., PL implemented in r8s and the uncorrelated lognormal [UCLN] clock implemented in BEAST) across our range of calibration scenarios.

Materials and Methods

Molecular Sequence Data

DNA sequences for all taxa were obtained from GenBank. Accession numbers are listed in Appendix S1 (Dryad doi 10.5061/dryad.qq106tm4). Two protein-coding genes (*atpB* and *rbcL*) and three noncoding regions (*atpB–rbcL* intergenic spacer, *trnL* intron, and *trnL–trnF* intergenic spacer) of chloroplast DNA and one noncoding region, the internal transcribed spacers (ITS) of nuclear DNA were used for phylogenetic analysis.

TABLE 1. Summary of the molecular partitions used in this paper

Locus	*atpB*	*rbcL*	*atpB-rbcL*	*trnL* intron	*trnL–trnF*	ITS
Number of sequences						
Nothofagus	12	27	24	12	12	22
Outgroups	19	21	14	19	17	23
Total	31	48	38	31	29	45
Model (BEAST)						
All	GTR + G	TrN + I + G	TVM + G	K81uf + I	TVM	GTR + G
Codon1	TrN + I	F81 + I + G				
Codon2	HKY + I	JC + I + G				
Codon3	TVM + I	K81uf + G				

Twenty-seven species of *Nothofagus* were sampled as the ingroup, with an additional 21 outgroup species representing core Fagales (Table 1; Appendix S1). Two families, Fagaceae and Juglandaceae, were sampled more densely because they have numerous fossils available for use as calibration points (see below). *Cucumis* (Cucurbitales: Cucurbitaceae) and *Lotus* and *Phaseolus* (Fabales: Fabaceae) were used as additional outgroup taxa to root our phylogeny of Fagales. Current evidence tends to support Cucurbitales as the sister group of Fagales (Soltis et al. 2007; Wang et al. 2009; Moore et al. 2010), although some conflict remains regarding the relative positions of Rosales, Cucurbitales, and Fagales (e.g., Burleigh et al. 2009). Therefore, all trees in this study are presented with Fabales as the most distant outgroup.

Sequence Alignment

Sequences were manually aligned in BioEdit v.7.0.1 (Hall 1999) and Se-Al v. 2.0a11 (Rambaut 2002). All DNA regions were aligned separately and concatenated before analyses. In ITS and *trnL* intron, regions of uncertain alignment (thus doubtful homology) were excluded from analysis. Four regions were excluded from the ITS (between the final alignment positions 123 + 124, 224 + 225, 399 + 400, and 502 + 503) and three from the *trnL* intron (between the final alignment positions 242 + 243, 274 + 275, and 347 + 348). As a result, the alignments of ITS and *trnL* intron were reduced in length by 26.5% and 17.5%, respectively. The concatenated alignment was partitioned into 10 subsets: three codon positions of *atpB*; three codon positions of *rbcL*; *atpB–rbcL* spacer; *trnL–trnF* spacer; *trnL* intron; and ITS. This partitioning scheme was used for all phylogenetic analyses conducted in this study. We found overwhelming support for this scheme compared with a simpler scheme in which there were only four subsets (first and second codon positions of chloroplast protein-coding genes, third codon positions of chloroplast protein-coding genes, noncoding chloroplast DNA and ITS), with $\log_{10}$ Bayes factors of approximately 22–23 across our various analyses.

Phylogenetic Analysis

The phylogeny was inferred using maximum likelihood (ML) in RAxML version 7.0.4 (Stamatakis 2006). Each subset of the concatenated sequence alignment was assigned a separate GTR + I + G substitution model. The inclusion of a parameter representing the proportion of invariant sites led to an increase in the log likelihood of the data (GTR + I + G: −21,382; GTR + G: −21,533). Six categories were used for the gamma model of among-site rate heterogeneity. The best-known likelihood tree was found using default search parameters and with 10 search replicates. Support for nodes in the inferred phylogeny was assessed using bootstrap analysis with 1000 pseudoreplicates. This ML tree (Fig. 1; Fig. S0; TreeBASE accession number TB2:S12119) was used to discuss fossil relationships and set up the calibration scenarios (see below).

Calibration

The complete list of fossils considered in this study, with details justifying their identifications and geologic ages, is provided in Appendix S2 (see Table 2 for a summary). It includes 4 stem Fagales, 24 core Fagales, and 10 *Nothofagus* fossils. This list is not intended to be exhaustive, and many fossils assigned to these groups are not included because of too much uncertainty about their phylogenetic placement or geologic age.

In order to obtain age constraints, we employed the following protocol: (i) consider all well-identified fossils available both within and immediately outside the group of interest, emphasizing those representing early clade occurrences; (ii) review the proposed relationships of each fossil in light of the most recent phylogeny of extant taxa available and assign each one to a specific branch of the phylogeny; (iii) investigate the geological age given in the paper describing the fossil and revise the age if necessary; (iv) use the upper (younger) absolute bound of the age range of each fossil as a minimum age for the stem node of the extant clade to which it is assumed to be related; (v) use the oldest of all minimum age constraints available for each given node (i.e., eliminate younger and redundant age constraints for setting up the analysis—but keep these on file as additional corroborative evidence for the age of this node).

Fossil relationships.—Three methods are commonly used to assess the relationships of fossils to extant taxa. Here, we refer to these three methods as the "intuitive," "apomorphy-based," and "phylogenetic" methods. These are described in detail below.

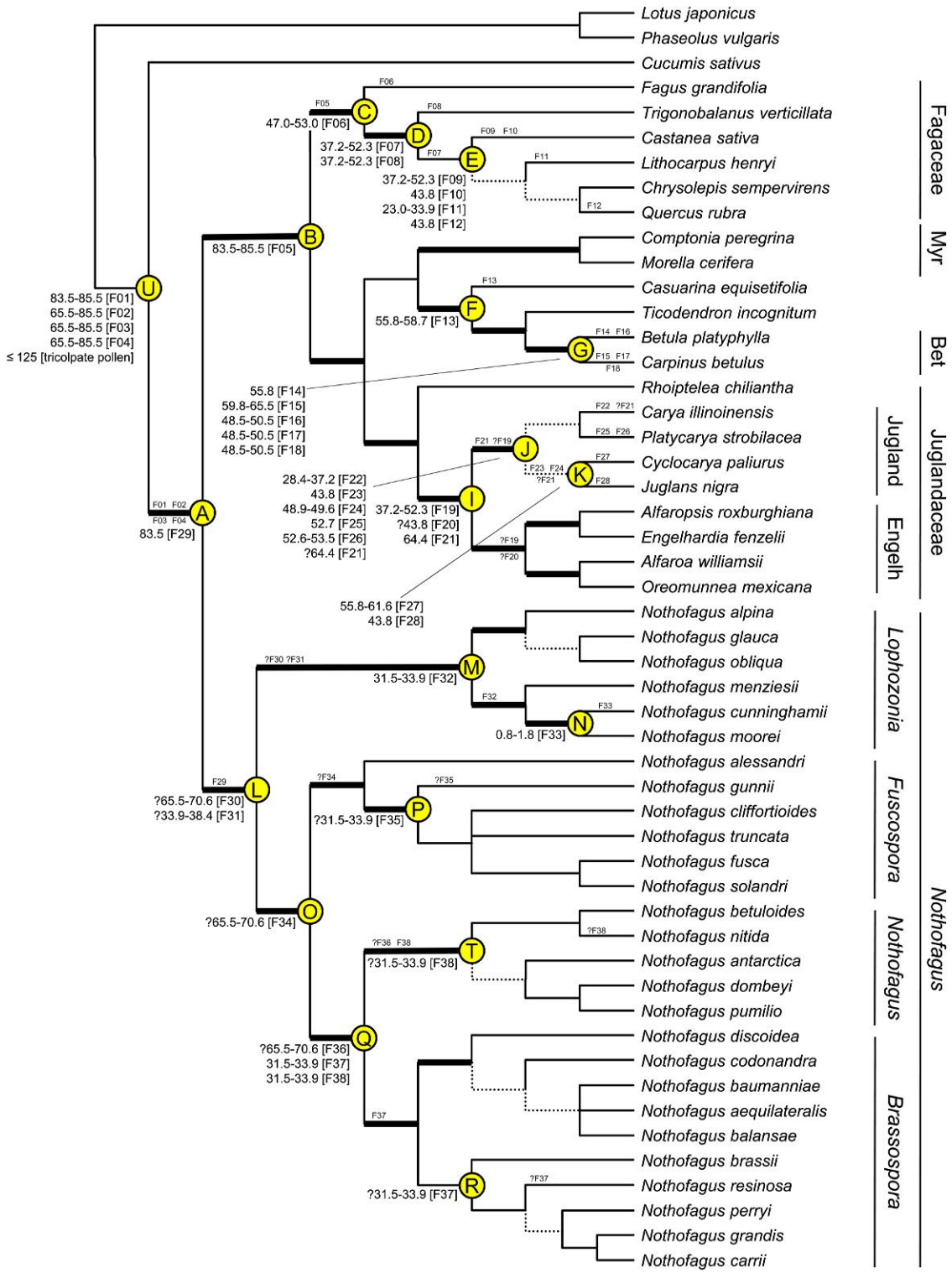

FIGURE 1. Reference tree used to set up the calibration scenarios, obtained from an ML analysis (RAxML) of combined sequence data from six regions of the chloroplast and nuclear genomes. Putative fossil (F) relationships are indicated (see Table 2 and Appendix S2 for details). Question marks denote early but risky fossil assignments, whereas the rest are considered safe but late (see text). Nodes on which age constraints were applied in at least one of the scenarios are marked with a circled letter. Below each node is the absolute age range of the oldest stratigraphic record of each fossil, following the GTS of Gradstein et al. (2004). Bootstrap support (bs) is indicated with branch thickness as follows: thick (bs $\geqslant$ 95%), plain (95% > bs $\geqslant$ 60%), dotted (bs < 60%). See Appendix S3 for details of branch support in this and other (BEAST) analyses. Family names follow APG (2009). Abbreviations: Myr = Myricaceae; Bet = Betulaceae; Jugland = Juglandoideae; Engelh = Engelhardioideae (sensu Manos et al. 2007). This figure is available in black and white in print and in color at *Systematic Biology* online.

TABLE 2. Fossils used in the calibration scenarios (for complete details, see Appendix S2)

	Fossil identity				Node assignment		Geologic age	
F#	Fossil taxon	Source	Organ category	Locality	Node[a]	Method	Age[b] (Ma)	Basis of age
FAGALES (stem relatives)								
F01	*Protofagacea allonensis* Herend., P.R. Crane & Drinnan	Herendeen et al. (1995)	Infl., flowers, pollen, fruits, cupules	Allon Quarry, GA (Gaillard Fm.)	U	Apomorphy-based	83.5–85.5	Palynology (Herendeen et al. 1995 but also see Burnett 1996)
CORE FAGALES: Fagaceae								
F05	*Antiquacupula sulcata* H. Sims, Herend. & P.R. Crane	Sims et al. (1998)	Flowers, pollen, fruits, cupules	Allon Quarry, GA (Gaillard Fm.)	B	Apomorphy-based	83.5–85.5	See F01
F06	*Fagus langevinii* Manchester & Dilhoff	Manchester and Dillhoff (2004)	Cupules/nuts, leaves, pollen	McAbee, British Columbia, Canada	C	Apomorphy-based	47.0–53.0	K-Ar dates (Ewing 1981)
F07	*Castanopsoidea columbiana* Crepet & Nixon	Crepet and Nixon (1989)	Infl., fruits, pollen	Buchanan, TN (Claiborne Fm.)	D	Phylogenetic (Crepet and Nixon 1989)	37.2–52.3	Claiborne Fm. ranges from late early to late middle Eocene; floras isolated and not placed stratigraphically
F10	*Castanopsis crepetii* Manchester	Manchester (1994)	Fruits	Nut Beds, OR (Clarno Fm.)	E	Apomorphy-based	43.8	^{40}Ar–^{39}Ar weighted mean age of 43.8 ± 0.3 Ma, and Bridgerian mammal fossils (Manchester 1994)
CORE FAGALES: Casuarinaceae								
F13	*Gymnostoma antiquum* Scriven & R.S. Hill	Scriven and Hill (1995)	Fruits, articles, cuticles	Lake Bungarby, New South Wales, Australia	F	Apomorphy-based	55.8–58.7	Late Paleocene Upper *Lygistepollenites balmei* Zone of Stover and Partridge (1973)
CORE FAGALES: Betulaceae								
F15	*Palaeocarpinus* P.R. Crane (several species)	Pigg et al. (2003), Manchester et al. (2004)	Fruits	Joffre Bridge, Alberta, Canada (Paskapoo Fm.)	G	Apomorphy-based	59.8–65.5	Associated Ti3 mammals (Fox 1990); the locality cited refers to the oldest well-dated occurrence
CORE FAGALES: Juglandaceae								
F21	*Polyptera manningii* Manchester & Dilcher	Manchester and Dilcher (1982, 1997)	Fruits, infructescences, catkins, pollen, leaves	United States (Fort Union Fm.)	I/?J	Phylogenetic (Manos et al. 2007)	64.4	^{40}Ar–^{39}Ar dates (Belt et al. 2004)
F25	*Paleoplatycarya wingii* Manchester	Wing and Hickey (1984), Manchester (1987)	Infructescences, fruits, leaves	Wyoming, US (Willwood and Wasatch Fms.)	J	Phylogenetic (Manos et al. 2007)	52.7	^{40}Ar–^{39}Ar date of 52.59 ± 0.12 Ma for the Willwood ash, 13 m above the plant bed (Wing et al. 1991, 2000; Smith et al. 2004), and stratigraphic interpolation to the nearest magnetic reversal (Currano et al. 2010)
F27	*Cyclocarya brownii* Manchester & Dilcher	Manchester and Dilcher (1982)	Fruits	United States (Fort Union Fm.)	K	Apomorphy-based	55.8–61.6	Stratigraphic correlation
NOTHOFAGUS								
F29	*Nothofagidites senectus* M.E. Dettmann & G. Playford	Dettmann and Playford (1968), Stover and Partridge (1973), Macphail (2007)	Pollen	Bass and Gippsland Basins, Australia	A	Apomorphy-based	83.5	Palynostratigraphy, cross-correlated to the marine record
F30	*Nothofagidites asperus* (Cookson) E.J. Romero	Romero (1973), Dettmann and Thomson (1987)	Pollen	Antarctica	?L	Intuitive	65.5–70.6	Dinocysts and the revised Helby et al. scheme dates (Crame et al. 2004)
F32	*Nothofagus tasmanica* R.S. Hill	Hill (1983)	Leaf with cuticle	Little Rapid River, Tasmania, Australia	M	Phylogenetic (Jordan and Hill 1999)	31.5–33.9	Palynostratigraphy

(*Continued*)

TABLE 2. (*Continued*)

F#	Fossil identity				Node assignment		Geologic age	
	Fossil taxon	Source	Organ category	Locality	Node[a]	Method	Age[b] (Ma)	Basis of age
F33	*Nothofagus cunninghamii* (Hook. f.) Oerst.	Ørsted (1871), Jordan (1999)	Leaf with cuticle	Regatta Point, Tasmania, Australia	N	Phylogenetic (Jordan and Hill 1999)	0.8–1.8	Palynostratigraphy
F34	*Nothofagidites brachyspinulosus* (Cookson) W.K. Harris	Harris (1965), Macphail (2007)	Pollen	Australia	?O	Intuitive	65.5–70.6	Palynostratigraphy
F35	*Nothofagus gunnii* (Hook. f.) Oerst.	Hill (1991)	Leaf with cuticle	Cethana, Tasmania, Australia	?P	Intuitive	31.5–33.9	Palynostratigraphy
F36	*Nothofagidites lachlaniae* (Couper) Milden. & Pockn.	Mildenhall and Pocknall (1984), Dettmann et al. (1990)	Pollen	Antarctica	?Q	Intuitive	65.5–70.6	See F30
F37	*Nothofagus macronata* R.S. Hill and *N. serrata* R.S. Hill	Hill (1991)	Leaf with cuticle	Little Rapid River, Tasmania, Australia	Q/?R	Phylogenetic (Jordan and Hill 1999)	31.5–33.9	Palynostratigraphy
F38	*Nothofagus lobata* R.S. Hill	Hill (1991)	Leaf with cuticle	Little Rapid River, Tasmania, Australia	Q/?T	Phylogenetic (Jordan and Hill 1999)	31.5–33.9	Palynostratigraphy

Notes: Only the oldest fossils of each clade are listed here. For a complete list of fossils reviewed in this study (including those providing younger or redundant age constraints), see Appendix S2.

[a]A question mark denotes a risky assignment, used only in scenarios 4–6 (for position of nodes, see Fig. 1).

[b]For calibration, the younger bound of any range provided was used as a minimum age constraint.

The "intuitive" method considers all the characters observed on a fossil and looks for the extant taxon that best matches the fossil. Although characters might be weighted implicitly in the process (based on knowledge of their variation), this approach is best described as a "global similarity" method. It is used extensively in fossil identification, especially because in many cases relevant morphological phylogenetic data are not available. This method has two drawbacks for molecular dating studies. First, it can lead to incorrect phylogenetic relationships if the similarities are shared ancestral characters. Second, it might not distinguish between stem lineage and crown group, an issue that is critical for calibrating molecular dating analyses (Doyle and Donoghue 1993).

The "apomorphy-based" method, on the other hand, is explicitly phylogenetic. This approach, which relies only on derived characters to assess fossil relationships, requires the phylogeny of extant taxa to be at least partly known. The apomorphy-based method has been widely used to justify fossil calibrations in molecular dating analyses, even though the putative apomorphies in question are rarely spelled out and discussed in either the molecular dating paper or in the literature cited for each fossil. A potential shortcoming of this method is the risk of giving an arbitrarily high weight to some characters while overlooking others.

The phylogenetic method specifically aims to assess fossil relationships using phylogenetic analysis of a data set that includes both extant and fossil taxa. Originally, such approaches used only morphological data (e.g., Crane 1985; Donoghue et al. 1989; Doyle and Donoghue 1992; Nixon et al. 1994), but in some recent studies, molecular data have also been incorporated using a "total evidence" approach (e.g., Eernisse and Kluge 1993; Jordan and Hill 1999; Clarke et al. 2007; Magallón 2007, 2010; Manos et al. 2007; Hermsen and Hendricks 2008; Wiens et al. 2010) or a "molecular backbone" approach (e.g., Springer et al. 2001; Magallón 2007; Manos et al. 2007; Hermsen and Hendricks 2008; Sauquet, Weston, Anderson, et al. 2009; Doyle and Endress 2010). Because phylogenetic methods explore all of the most parsimonious positions for a fossil, including some that might not be readily identified using a handful of selected apomorphies, they are potentially more objective and can help identify alternative positions missed by other methods. In spite of this advantage, phylogenetic methods are rarely used to assess fossil relationships (Crane et al. 2004), and few studies have relied on them to calibrate molecular dating analyses (Gernandt et al. 2008; Lee et al. 2009; Sauquet, Weston, Anderson, et al. 2009; Sauquet, Weston, Barker, et al. 2009).

The safe but late fossils used throughout the present study are ones whose position has been addressed with phylogenetic methods (Crepet and Nixon 1989; Jordan and Hill 1999; Manos et al. 2007) or could be justified using an apomorphy-based approach (Table 2). However, in order to assess the effect of including early but risky fossil age constraints in dating analyses, we also performed some specific analyses using fossils whose relationships had only been discussed

with an intuitive approach. We note that the relationship we make here between the method for assessing the position of a fossil and the confidence in the assignment can seem simplistic. In particular, an important factor weighing on the confidence in the assignment of a fossil is the number of discriminatory features (characters) known for the fossil, whichever method is used. When this number is high, the three approaches may converge. However, when this number is low, as is often the case with fossil pollen, an intuitive approach may be misleading, whereas apomorphy-based or phylogenetic approaches will tend to suggest very ambiguous assignments (i.e., many possible branching positions on the tree) and can therefore be considered less risky. It is also important to distinguish safe from precise assignments (i.e., with few possible branching positions on the tree). Evidently, an ideal fossil for calibration should have both a safe and precise phylogenetic placement. Finally, safe and risky are only employed with respect to phylogenetic assignment throughout this study. Because we carefully revised the geologic age of each fossil (see below) and used a conservative approach to calibration (using only confirmed minimum ages), the fossil age constraints included in this study can be considered safe in their geologic age.

Calibration: two examples.—Herein, we illustrate the protocol that we followed for identifying fossil age constraints for this study using an example from the outgroup (Juglandaceae) and one from the ingroup (*Nothofagus).* First, the fossil species *Paleoplatycarya wingii* Manchester (fossil F25 in Table 2) was previously described as the fruit of *Platycarya castaneopsis* (Lesquereux) Wing and Hickey, which was based on an assemblage of co-occurring leaves (which carried the basionym), pistillate inflorescences, fruits, and pollen from various late early Eocene deposits of several basins in western Wyoming, United States (Wing and Hickey 1984). Using a phylogenetic analysis of morphological characters of extant and fossil taxa, Wing and Hickey (1984) proposed that this fossil was more closely related to extant *Platycarya strobilacea* than to any other living species of Juglandaceae. Later, rejecting Wing and Hickey's expanded definition of the extant *Platycarya* genus to accommodate characters only seen in this extinct species, Manchester (1987) reassigned the fruits of *Platycarya castaneopsis* to the extinct genus *Paleoplatycarya* and the extinct species *Paleoplatycarya wingii*, while not revising the nomenclature for the other associated organs.

Using various analyses that integrated morphological and molecular data (including the total evidence and molecular backbone approaches), Manos et al. (2007) confirmed support for the relationship of *Paleoplatycarya wingii* with extant *Platycarya* (bootstrap = 77%). However, the phylogenetic position of extant *Platycarya* itself is uncertain. Using molecular data alone, *Platycarya* appears to be nested within Juglandoideae (Manos et al. 2007). Our analyses, based on limited taxon sampling, also suggest this result, but with a different relationship (*Platycarya* sister to *Carya* rather than sister to *Juglans* + *Cyclocarya*) and low support (Fig. 1; Appendix S3). On the other hand, the integration of molecular and morphological data suggests that *Platycarya* is sister to the rest of Juglandoideae (Manos et al. 2007). Because of this uncertainty, we decided that *Paleoplatycarya wingii* provides a safe minimum age only for the last common ancestor of *Platycarya* and the rest of Juglandoideae (Fig. 1: node J).

Using the 2004 GTS (Gradstein et al. 2004), an early Eocene age translates into an absolute age of 48.6–55.8 Ma. The fossil *Paleoplatycarya wingii* therefore provides a safe minimum age of 48.6 Ma for node J. However, in this particular case, more precise geologic dating of the rocks in which the fossil has been found justifies an older safe absolute age, while also illustrating some of the challenges in high-precision dating of fossils even in well-studied areas. For the Western Interior US basins, land mammal zonation has great biostratigraphic importance because of intensive fossil sampling and geologic correlation of the zones (Woodburne 2004). The land mammal zonation for the Paleocene and Eocene has been, and continues to be, extensively studied and revised via correlation of land mammal stratigraphic ranges to reliable datums including volcanic tuffs and paleomagnetic reversals (Woodburne 2004; Secord et al. 2006; Clyde et al. 2007). As reviewed by Wing and Hickey (1984) and Wilf (2000), all occurrences of *Paleoplatycarya wingii/Platycarya castaneopsis* can be correlated with strata containing mammal fossils of the Wasatchian land mammal zone, and in particular to the short Lostcabinian subzone of the Wasatchian. The best age resolution of *Palaeoplatycarya* occurrences comes from the Fifteen Mile Creek flora, upper Willwood Formation of the Bighorn Basin, where the taxon occurs 13 m below the Willwood Ash (Wing et al. 1991, 2000), which was recently redated to 52.59 ± 0.12 Ma (2 sigma) from ^{40} Ar–^{39}Ar analyses of sanidine (Smith et al. 2004). The stratigraphically closest datum in the Bighorn Basin that underlies the fossils is the base of magnetic polarity chron C24n (Clyde et al. 2007), which was dated to 53.8 Ma from calibration of seafloor magnetic anomalies using astronomically tuned cyclostratigraphy, available radiometric dates, and other data (Luterbacher et al. 2004; Ogg and Smith 2004). However, one of the radiometric constraints used was in fact the older date for the Willwood Ash of 52.8 ± 0.3 Ma (Wing et al. 1991). The seafloor magnetic anomaly record is not yet recalibrated using the revised age for the Willwood Ash of Smith et al. (2004), and more critically, the duration of C24 is undergoing some revision from new marine cyclostratigraphic data (Westerhold and Röhl 2009). These issues, conservatively, produce no more than approximately 0.2 Ma uncertainty for the fossils of interest, especially given their close stratigraphic proximity to the Willwood Ash. Using the published 53.8 Ma datum for the base of C24 allows the age of the fossiliferous horizon to be estimated by linear interpolation as 52.65 Ma (Currano et al. 2010), producing our minimum age assignment for node J of 52.7 ± 0.2 Ma.

Four other fossils considered in our study also provide safe minimum ages for node J, ranging from 28.4 to 52.6 Ma (fossils F22–24 and F26 in Appendix S2; Fig. 1). Because they are younger than or have ages not distinguishable from *Paleoplatycarya wingii* (F25), these fossils provide redundant age constraints and were not included in the analyses presented in this paper. Nevertheless, we suggest that redundant age constraints are worth reporting in molecular dating studies (as we do here in Fig. 1 and Appendix S2) because they provide additional evidence for the age of a given node. Multiple occurrences of lineages can also be used to bracket divergence times statistically in more exhaustive studies of a dense fossil record (Marshall 2008).

Dettmann et al. (1990) reviewed the entire fossil pollen record of *Nothofagus* and placed several fossil species into the *fusca* (b) group. This group of pollen is represented today only by the extant species of subgenus *Nothofagus*. Using the intuitive approach, the plants producing *fusca* (b) pollen would be more closely related to the extant species of subgenus *Nothofagus* than to any other extant species of *Nothofagus*. Thus, *fusca* (b) pollen would be placed on the stem lineage or in the crown group of subgenus *Nothofagus* and therefore provide a minimum age for the stem node of this subgenus (Fig. 1: node Q). However, because this relationship has not been investigated using apomorphy-based or phylogenetic methods, we treated this constraint as early but risky. Although Dettmann et al. (1990) and Macphail et al. (1994), claimed Campanian occurrences of *fusca* (b) pollen, a lack of detail on the source of these fossils means that these records cannot be confirmed at present. Therefore, the oldest documented fossil occurrence of this group known to us is *Nothofagidites lachlaniae* (Couper) Pocknall & Mildenhall (fossil F36 in Table 2) (Dettmann and Thomson 1987) from James Ross Island, Antarctica. The age of these sediments was recently revised from Late Campanian to Maastrichtian, based on strontium isotope dating (Crame et al. 2004). Using the 2004 GTS (Gradstein et al. 2004), this translates into a conservative minimum age of 65.5 Ma (rather than 70.6 Ma if we had used the Campanian age originally quoted). In our calibration scenarios restricted to safe but late fossil age constraints, this early but risky age constraint was removed. Instead, two separate lines of evidence, both based on fossil leaves with cuticles that had been analyzed phylogenetically by Jordan and Hill (1999), provide safe minimum age constraints of only 31.5 Ma for node Q (fossils F37 and F38 in Table 2; Fig. 1).

Calibration scenarios.—To provide an empirical test for the impact of various approaches to calibration of *Nothofagus*, we devised eight distinct calibration scenarios (Table 3), reflecting the diversity of situations faced by biologists when trying to obtain divergence times for a group using molecular dating methods. In the first series (scenarios 1–3), we considered only safe but late fossil age constraints. All such available constraints were applied in scenario 1 (Fig. S3), whereas only ingroup constraints or outgroup constraints were used in scenarios 2 (Fig. S4) and 3 (Fig. S5), respectively. In the second series (scenarios 4–6), we considered both safe but late and early but risky fossil age constraints. These were again applied to the entire tree (scenario 4; Fig. S6), only the ingroup (scenario 5; Fig. S7) or only the outgroup (scenario 6; Fig. S8). In doing so, some fossils placed conservatively in the safe but late scenarios were shifted to more risky nested positions (fossils F21, F37, and F38 in Table 2). In one case, this also meant that a safe but late constraint (52.7 Ma on node J from fossil F25) became uninformative because an older risky constraint applied to the same node (64.4 Ma from fossil F21).

In scenario 7, we replaced fossil age constraints with assumptions of vicariance for the main biogeographic disjunctions in *Nothofagus*. The area cladogram implicit in our phylogeny shows a simple pattern that is consistent with all molecular phylogenies of the genus (e.g., Manos 1997). Six nodes are calibrated by this pattern, which involves New Zealand being closer to Australia/Tasmania than South America in both subgenera *Fuscospora* and *Lophozonia* and New Caledonia being closer to New Guinea than to South America in the clade comprising subgenera *Brassospora* and *Nothofagus* (Fig. S9). New Guinea is part of the Australian plate, and New Caledonia sits on the sunken landmass that includes New Zealand (Zealandia; Mortimer 2004). The split between Australia and New Zealand is often taken as 84 Ma, the timing of the oldest seafloor magnetic anomaly in the Tasman Sea. However, the Tasman Sea opened in the south first and progressively unzipped northwards, with the continent of Zealandia remaining pinned to the Australian margin with seafloor spreading ceasing in the earliest Eocene (Gaina et al. 1998). The timing of physical separation is poorly constrained as it is linked to the subsidence history of the crustal blocks that make up the continent of Zealandia (Ladiges and Cantrill 2007). Limited seismic studies, and core material, give some insights with terrestrial sediments occurring over several crustal blocks in the earliest Eocene (55.8 Ma) that are replaced with marine deposits shortly after (Exon et al. 2006). This provides a younger bound of vicariance between Australia and New Zealand of 55.8 Ma. The relationship between Australia/New Zealand and South America is complicated by the role that Antarctica played, the lack of Antarctic biota that has been integrated into phylogenies, and the high latitude that might have played an important role in filtering species exchange between landmasses. The geological separation of South America from Antarctica/Australia/New Zealand has a broad range of estimates due to uncertainties in the geological reconstructions and formation of seaways (Barker et al. 2007). Recent studies suggest shallow seaways by middle Eocene times (40 Ma; Livermore et al. 2005; but see Barker et al. 2007 for a summary) but older ages are possible. Because genetic divergence and isolation might have predated physical separation of the areas considered, and this time lapse is hardly quantifiable, we treated these constraints as hard minimum bounds just as we did for fossils (but see below).

TABLE 3. Summary of the calibration scenarios

Scenario	Number of age constraints			Number of informative age constraints			Smoothing (PL)	Coefficient of variation of rates (BEAST)
	Total	Ingroup	Outgroup	Total	Ingroup	Outgroup		Mean (95% credibility interval)
0: No calibration (root = 125 Ma)	0	0	0	0	0	0	1000	0.6 (0.4–0.8)
Fossil calibration (safe only)								
1: Default	14	4	10	9	3	6	20	1.1 (0.9–1.4)
2: Ingroup	4	4	0	4	4	0	1000	0.8 (0.6–1.0)
3: Outgroup	10	0	10	6	0	6	13	1.0 (0.8–1.3)
Fossil calibration (safe and risky)								
4: Risky	18	9	9	12	6	6	160	1.4 (1.1–1.7)
5: Ingroup risky	9	9	0	7	7	0	13	1.2 (1.0–1.5)
6: Outgroup risky	9	0	9	6	0	6	13	1.1 (0.8–1.3)
Geologic calibration								
7: Vicariance	6	6	0	3	3	0	400	1.5 (1.1–1.8)
Secondary calibration								
8: Wikström et al. (2001) (B = 60–61 Ma)	1	0	1	1	0	1	1000	0.6 (0.4–0.8)
8a: Wikström et al. (2001) (B = 57–65 Ma)	1	0	1	1	0	1	1000	0.6 (0.4–0.8)
8b: Magallón and Castillo (2009) (B = 93.5 Ma)	1	0	1	1	0	1	1000	0.6 (0.4–0.8)
8c: Wang et al. (2009) (B = 91–100 Ma)	1	0	1	1	0	1	790	0.6 (0.4–0.8)
8d: Bell et al. (2010) (B = 43–68 Ma)	1	0	1	1	0	1	250	0.6 (0.4–0.8)
8e: This study, scenario 1 (B = 100.8–118.3 Ma)	1	0	1	1	0	1	16	0.6 (0.4–0.8)
8f: This study, scenario 1 (L = 53.4–93.2 Ma)	1	1	0	1	1	0	1000	0.7 (0.5–0.9)

Notes: The maximum age constraint on the root (node U $\leqslant$ 125.0; fixed in scenario 0) applied in all scenarios is not counted here. The stem node of *Nothofagus* (A) is treated as an ingroup node.

In the final scenario 8, we used a secondary calibration point derived from a previous study, a common practice when no fossils are known for the group of interest. For flowering plants, the main source of such secondary calibration has been the 567-taxon three-gene NPRS analysis of Wikström et al. (2001). Unfortunately, *Nothofagus* was not sampled in this analysis, so we used the most recent common ancestor of the rest of Fagales (node 89 of Wikström et al. 2001) as a proxy for the age of core Fagales in our study (Fig. 1: node B). Wikström et al. (2001) used three methods of branch length optimization before analysis with NPRS: parsimony with ACCTRAN, parsimony with DELTRAN, and ML. The ages estimated for core Fagales were 61, 60, and 61 Ma, respectively. Therefore, we used a range of 60–61 Ma to calibrate node B in scenario 8 (Fig. S10).

However, we also experimented with a number of alternative secondary calibration settings (scenarios 8a–f; Table 3). First, we used the error around ACCTRAN age estimates calculated by Wikström et al. (2001) using 100 bootstrap replicates. For core Fagales, this standard error was estimated to 4 Ma, translating to a more cautious range of 57–65 Ma for the calibration of this node (scenario 8a; Fig. S11). Second, we used three recent molecular dating studies, based on more advanced relaxed-clock methods and multiple calibrations applied as minimum ages (contrary to the single point calibration of Wikström et al. 2001). Magallón and Castillo (2009) obtained an estimated age of 93.5 Ma for core Fagales in their "relaxed" PL analysis (scenario 8b; Fig. S12). This is actually the 95% confidence interval (based on bootstrapping) reported for this node, and thus, no alternative to a fixed secondary calibration is possible here. Wang et al. (2009) obtained a 95% credibility interval of 91–100 Ma for the same node in their "BRC-1" BEAST analysis (scenario 8c; Fig. S13). Bell et al. (2010) obtained a 95% credibility interval of 43–68 Ma for this node in their second BEAST analysis with fossil constraints implemented as lognormal priors (scenario 8d; Fig. S14). Last, we experimented with two secondary calibrations derived from within our study. In doing so, we wanted to answer a simple question: In the best-case scenario, assuming we have the correct age for at least one node, what happens to the rest of the tree when there are no fossils? This was also intended as a more direct test of the accuracy and reliability of secondary calibration in general. We used the 95% credibility intervals from the BEAST analysis of scenario 1 as minimum and maximum age constraints to calibrate node B (100.8–118.3 Ma; scenario 8e; Fig. S15) or node L (53.4–93.2 Ma; scenario 8f; Fig. S16), which is the crown node of *Nothofagus* (Fig. 1).

In addition, a 125 Ma maximum age constraint was applied to the root (the divergence of Cucurbitales and

Fagales; node U) in all our dating analyses. This date is justified by the well-documented first appearance of tricolpate pollen in the fossil record in the earliest Aptian (Hughes and McDougall 1990; Doyle and Hotton 1991; Doyle 1992) and its reported absence in older sediments worldwide. Since tricolpate pollen is a synapomorphy of eudicots as a whole and has been preserved in most of its early diverging lineages (Doyle and Endress 2000; Furness and Rudall 2004; Doyle 2005), it is unlikely that crown eudicots are older than the origin of this character in the fossil record. It seems therefore reasonable to apply a maximum age constraint of 125 Ma (the current estimate for the age of the Barremian/Aptian boundary being 125 ± 1.0 Ma) to crown eudicots, or any clade nested in them. This assumption has been widely applied in other studies (Anderson et al. 2005; Bell and Donoghue 2005; Davis et al. 2005; Magallón and Castillo 2009; Sauquet, Weston, Anderson, et al. 2009, Sauquet, Weston, Barker, et al. 2009; Wang et al. 2009; but see Results and Discussion below).

Finally, we ran a "calibration-free" analysis by fixing the age of the root to 125 Ma and using no other constraints (scenario 0; Fig. S17). A fixed calibration point was necessary because using a maximum age constraint alone does not represent sufficient calibrating information for estimating divergence times. This scenario provides a benchmark for comparing the effects of using multiple age constraints.

Divergence Time Estimation

ML analysis.—Using the ML tree with branch lengths (phylogram) obtained with RAxML (Fig. S0), divergence times were estimated with r8s (Sanderson 2003). The most distant outgroup clade (*Phaseolus* + *Lotus*) was pruned from the tree prior to all analyses. Dates were estimated using PL with an additive penalty. The truncated-Newton algorithm was employed, with 10 independent starts. The optimal smoothing parameter was determined using cross-validation analysis, testing values in increments of $10^{0.1}$ across a range spanning $10^{0.1}$–10^{3} (Table 3). Cross-validation was performed by pruning terminal branches and calculating prediction error. A separate analysis was performed for each calibration scenario. For scenarios in which the optimal smoothing parameter was found to be at the upper end of the tested range (10^3), using higher values did not have measurable effects on estimates of divergence times. All fossil and vicariance age constraints (scenarios 1–7) were applied as minimum ages, except for the 125 Ma maximum age constraint applied to the root in all analyses. Age ranges applied to specific nodes in secondary calibrations (scenarios 8 and 8a–f) were enforced as a combination of a minimum and a maximum age constraint on the same node (e.g., for scenario 8a, node B was constrained with "minage = 57" and "maxage = 65"), except for scenario 8b where node B was fixed to 93.5 Ma. To obtain confidence intervals for the divergence date estimates, a bootstrap-based approach was taken. Using the ML tree as a fixed topology, sets of branch lengths were estimated from 1000 bootstrap replicates using the software RAxML. Divergence dates were estimated from the resulting 1000 trees using the software r8s, with settings as described above. The smoothing parameter was not optimized for each replicate but was instead fixed to the value obtained for the original data set.

Bayesian analysis.—The phylogeny and divergence times were estimated simultaneously using the Bayesian software BEAST version 1.4.8 (Drummond and Rambaut 2007). The optimal substitution model was selected for each data subset by comparison of Bayesian information criterion scores using ModelGenerator (Keane et al. 2006; Table 1). A Yule prior (which includes a parameter for describing the net rate of speciation) was specified for the tree, whereas rate heterogeneity among lineages was modeled using a UCLN relaxed clock (Drummond et al. 2006). Posterior distributions of parameters, including the tree, were estimated by Markov chain Monte Carlo (MCMC) sampling. Samples from the posterior were drawn every 2000 steps over 30,000,000 steps of a single chain, with the first 10% of samples discarded by default as burn-in. To reduce the length of the burn-in period, an ultrametric tree satisfying the calibration constraints was used as the starting tree for the MCMC. Acceptable sampling from the posterior was checked for each analysis using the software Tracer version 1.5 (Rambaut and Drummond 2007). All chronograms from BEAST presented in this paper (e.g., Figs. 2 and 3) are maximum clade credibility trees with node heights rescaled to match posterior mean estimates, compiled from the posterior using TreeAnnotator.

A separate analysis was performed for each of the calibration scenarios described above. To allow comparison with the estimates made using ML, calibrations were specified in the form of uniform priors rather than exponential, lognormal, or gamma priors. Another problem arising with the use of nonuniform priors is that their meaningful parameterization should take into account the multiple sources of uncertainty outlined above, including uncertainty in phylogenetic position, uncer-tainty in the time lapse between divergence from the extant lineage and fossilization, and uncertainty in geological dating. Quantitative assessments of the fossil record can help the measurement and modeling of sampling uncertainty, but phylogenetic uncertainty is more difficult to quantify, especially given the disparate source of the studies used to assign fossil taxa to specific clades here. Therefore, fossil and vicariance age constraints (scenarios 1–7) were all implemented as hard minimum bounds using uniform priors with arbitrarily large maximum bounds (1000 Ma), except for the maximum age constraint on the root in all analyses, implemented using a uniform prior with a minimum bound of 0 Ma and a maximum bound of 125 Ma. Age ranges for secondary calibrations (scenarios 8 and 8a–f) were all implemented as uniform priors (93.49–93.51 Ma for scenario 8b because ages cannot be fixed in BEAST).

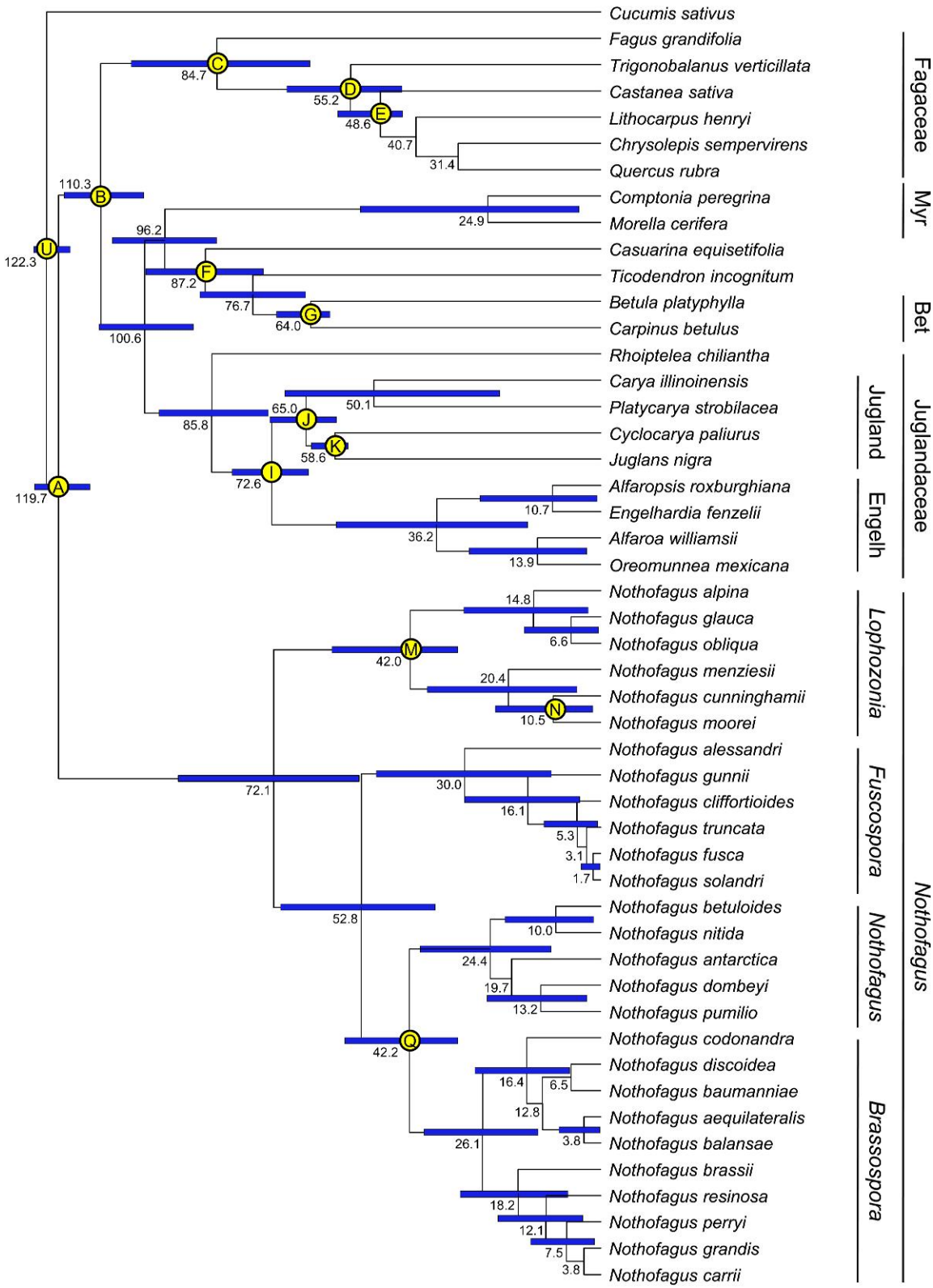

FIGURE 2. Chronogram obtained with the UCLN relaxed-clock method (implemented in BEAST) with age constraints from scenario 1 (default scenario: safe fossil minimum age constraints). Circled letters denote nodes with age constraints (for details, see Table 2 and Fig. S3). Numbers correspond to mean age estimates (in million years) and blue bars represent 95% credibility intervals (for details, see Appendix S4). Abbreviations as in Fig. 1. This figure is available in black and white in print and in color at *Systematic Biology* online.

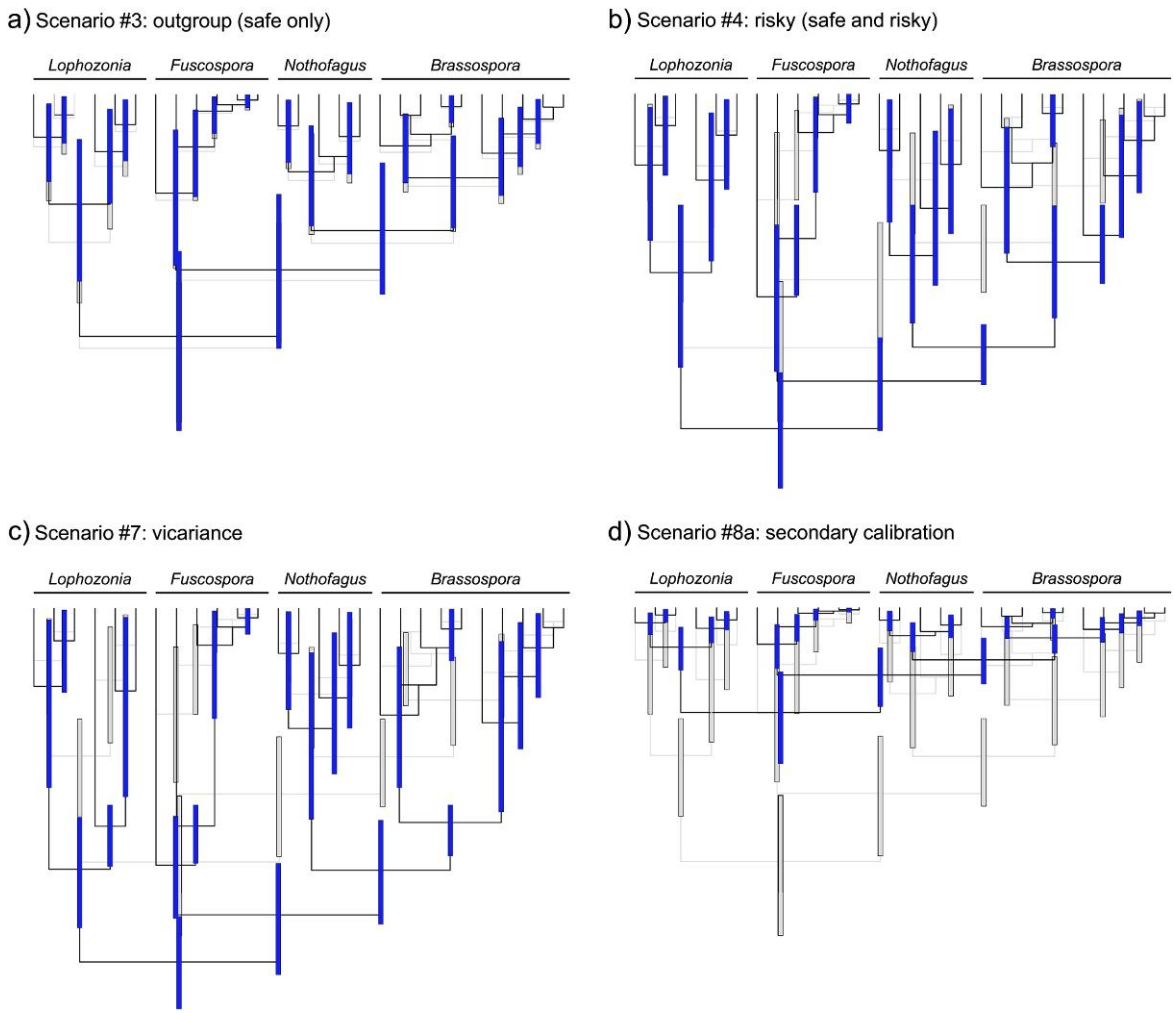

FIGURE 3. Four representative chronograms illustrating the variation in age estimates obtained with BEAST across calibration scenarios. For comparison, the chronogram and 95% credibility intervals from scenario 1 (Fig. 2) are redrawn in light gray in the background of each chronogram. These chronograms have been edited to only show the *Nothofagus* subtree (for complete chronograms of these and the remaining scenarios, see Figs. S19–S32; for details of calibration, see Table 2 and Figs. S4–S17; for all age estimates, see Appendix S4). This figure is available in black and white in print and in color at *Systematic Biology* online.

Because nonuniform priors, especially the exponential and lognormal priors, are widely used to model age constraints in BEAST analyses, we also ran additional sets of analyses for the fossil and vicariance scenarios using exponential priors (scenarios 1a–7a) and lognormal priors (scenarios 1b–7b) for all age constraints except the root (where the uniform prior above was maintained). Since we could not quantify all the uncertainties listed above, we parameterized these priors as follows. For exponential priors, the offset (minimum age constraint) matched the hard minimum bound as in the uniform priors, and the mean was set to be 10% older than this value. For lognormal priors, the offset (minimum age constraint) was set to be 10% younger than the hard minimum bound implemented in the uniform priors, the median value matched the hard minimum bound of the uniform priors, and the standard deviation was adjusted so that the 95% upper limit of the distribution (soft maximum) would be 10% older than the hard minimum bound of the uniform priors. This means that our results from analyses with lognormal priors are not directly comparable to those from our analyses with uniform priors, but it allows us to take advantage of the lognormal distribution by allowing a small probability that a fossil might actually be older than the clade to which it was assigned (and thus misplaced).

RESULTS AND DISCUSSION

Topology

The ML analysis with RAxML yielded a well-resolved phylogenetic tree (Fig. 1; Fig. S0). Bootstrap support (bs) was generally very strong, except for some relationships within Fagaceae and within each subgenus of *Nothofagus* (Appendix S3). All families of Fagales (sensu APG 2009) were monophyletic (bs = 100%), and the relationships among these families were overall strongly supported and consistent with previous studies (Manos and Steele 1997; Soltis et al. 2000, 2007; Hilu et al. 2003;

Li et al. 2004; Herbert et al. 2006). The only exception was the position of Myricaceae, here weakly supported as sister of [Casuarinaceae + [Ticodendraceae + Betulaceae]] (bs = 67%), and for which previous studies have given inconsistent results. Within *Nothofagus*, the four subgenera were monophyletic (bs = 100%), and the relationships among them were also strongly supported (bs = 100%) and consistent with previous studies (Martin and Dowd 1993; Manos 1997; Setoguchi et al. 1997; Jordan and Hill 1999; Cook and Crisp 2005; Knapp et al. 2005).

Our various Bayesian analyses with BEAST yielded essentially the same topology (Fig. 2; Figs. S19–S32) and consistent posterior probability support (Appendix S3), which was generally higher than ML bootstrap support as expected (Zander 2004). This was true not only of our calibration-free scenario (scenario 0) but also of all remaining scenarios (scenarios 1–8) for which the specification of age constraints could potentially have had an impact on phylogenetic inference in BEAST, since calibrations can carry intrinsic assumptions about the ordering of divergence events in the tree. Likewise, the use of different priors for age constraints in BEAST analyses of scenarios 1–7 did not affect the topology, with posterior probability values being very similar in each scenario. Therefore, in this study, calibrations did not seem to have a noticeable influence on topology nor branch support. It is possible that the topological impact of calibrations, as described by Ho and Phillips (2009), is more problematic for very uninformative data sets.

Finally, although branch support remained weak overall within each subgenus of *Nothofagus*, the topology itself, summarized by the ML RAxML tree (Fig. 1) and the maximum clade credibility BEAST trees (Fig. 2; Figs. S19–S32), was remarkably stable across all analyses. In fact, the only variation concerned the relative positions of *Nothofagus codonandra* and *N. discoidea* within subgenus *Brassospora*.

Age and Rate Variation among Scenarios

In general, we found that age estimates and confidence intervals varied substantially from one scenario to another, regardless of the dating method used (Table 4; Figs. 2, 3, 5, and 6; Appendix S4; Figs. S19–S32). For instance, the age of crown group *Nothofagus* (Fig. 1: node L) was estimated to 53.4–93.2 Ma (mean = 72.1 Ma) with BEAST for our default (safe only) scenario (scenario 1, Fig. 2). Using a secondary calibration derived from the study by Wikström et al. (2001) led to a drastically younger estimate of 16.7–39.5 Ma (mean = 27.4 Ma) for the same node (scenario 8, Fig. S25). This variation was observed at all levels of the tree. However, age estimates for nodes close to the root (e.g., node A) or the tips (e.g., node N) of the tree tended to be less variable, but this was expected because the ages of the root and tips were bounded by the same hard limits in all the analyses.

The maximum age constraint on the root (125 Ma, node U) was strongly influential in almost all the calibration scenarios, with the estimated age of the root being at or close to the constraint in almost all scenarios (Appendix S4). The only exceptions are the BEAST (but not PL) analyses of scenarios 8, 8a, and 8d (upper bound of 95% credibility interval = 114.8, 116.1, and 109.6 Ma, respectively). When this maximum age constraint was removed, the analyses failed to converge on a unique solution. This emphasizes the need for at least some maximal age information among the calibrations, whether in the form of a maximum bound or a point calibration (Hug and Roger 2007). Given the influence of the 125 Ma maximum age constraint on our analyses, two recent angiosperm dating studies are especially relevant here. Neither Bell et al. (2010) nor Smith et al. (2010) assumed that crown eudicots had a maximum age of 125 Ma. Using 36 minimum age constraints (all nested in angiosperms), Bell et al. (2010) estimated the age of crown eudicots to 123–139 Ma and the split of Cucurbitales and Fagales to 96–110 Ma. Using 32 minimum age constraints (including 5 outside of angiosperms), Smith et al. (2010) obtained 138–172 Ma for the crown group age of eudicots and no more than 110 Ma for the split of Cucurbitales and Fagales. Therefore, although both studies suggest that crown eudicots might in fact be older than has been widely assumed, both also agree that clades nested further up in the eudicot tree are indeed younger than 125 Ma as we assumed in our analyses. Furthermore, previous work with sensitivity analyses of the same age constraint has shown that the exact value given to the maximum age specified for the root might have little impact on age estimates of more shallow nodes in the tree (Sauquet, Weston, Barker, et al. 2009).

Some minimum age constraints were always influential (most notably node K, but also nodes E, G, I, M, and Q), whereas others never were (nodes B, C, and N). The calibration for node K (minimum age constraint of 55.8 Ma) was very significant. When the calibration was removed, the age of the node was estimated at around 10–20 Ma (Table 4). These results are not surprising. When using multiple age constraints, it is usually expected that at least some will influence the analysis because molecular data alone cannot always disentangle among-lineage rate heterogeneity. In these cases, increasing the number of calibrations might lead to better estimates of rate variation (Ho and Phillips 2009). On the other hand, because the fossil record is incomplete and fossil taxa can only be attributed to an internal lineage of a clade after distinctive apomorphic characters of this lineage have evolved, it is also possible that some nodes would actually be much older than the minimum age constraints applied to them in light of available fossil evidence. For instance, crown group Fagaceae (node C) were estimated to be 64.2–103.6 Ma old in scenario 1, whereas the oldest fossil attributable to an internal lineage of Fagaceae (F06 = *Fagus langevinii*) could only justify a minimum age constraint of 47 Ma for this node (Table 2).

The PL estimates were nearly always younger than mean BEAST estimates, except for the secondary calibration scenarios (scenarios 8 and 8a–f) where ages

TABLE 4. Summary of age estimates (in Ma) for selected nodes according to each calibration scenario (for full details, see Appendix S4)

	A (Fagales)		K (Juglandinae)		L (*Nothofagus*)		N		T (subgenus *Nothofagus*)	
Scenario	ML–PL	BEAST	ML–PL	BEAST	ML–PL	BEAST	ML–PL	BEAST	ML–PL	BEAST
0	79.7 (70.5–92.5)	101.6 (73.5–125.0)	7.1 (5.1–9.5)	9.0 (3.4–15.6)	20.7 (17.2–25.4)	34.8 (18.5–53.7)	2.6 (1.5–3.8)	4.0 (1.1–7.5)	5.4 (4.0–7.2)	9.3 (4.2–14.7)
1	108.3 (102.7–115.1)	119.7 (112.7–124.8)	55.8 (55.8–55.8)	58.6 (55.8–63.8)	51.1 (46.2–56.8)	72.1 (53.4–93.2)	8.1 (4.6–13.1)	10.5 (1.8–23.2)	16.3 (11.1–22.3)	24.4 (11.1–39.8)
2	92.4 (84.5–103.6)	109.2 (90.8–123.7)	9.1 (6.6–12.0)	12.9 (4.3–22.9)	44.1 (41.1–47.7)	61.8 (43.9–81.2)	4.3 (2.5–6.3)	7.4 (1.2–15.2)	8.7 (6.1–11.6)	16.7 (7.6–27.3)
3	106.7 (101.0–114.0)	119.5 (112.5–124.8)	55.8 (55.8–55.8)	58.5 (55.8–63.5)	30.0 (24.7–39.3)	68.7 (44.5–95.5)	3.8 (2.2–6.0)	8.5 (1.4–18.8)	8.0 (5.8–11.9)	21.9 (8.8–37.3)
4	111.3 (106.5–116.8)	120.2 (114.1–124.9)	55.8 (55.8–55.8)	60.4 (55.8–67.3)	73.8 (70.6–77.1)	95.1 (79.1–112.0)	7.9 (4.7–11.6)	11.6 (1.5–27.1)	31.5 (31.5–31.5)	45.9 (31.5–65.1)
5	101.7 (95.4–110.6)	115.3 (104.0–124.6)	9.8 (6.9–14.4)	16.3 (3.9–30.4)	72.9 (70.0–76.3)	93.1 (77.8–108.0)	10.0 (5.0–16.9)	11.2 (1.7–23.3)	44.8 (31.5–54.5)	41.5 (31.5–57.2)
6	107.2 (101.5–114.4)	119.7 (112.9–124.8)	55.8 (55.8–56.2)	59.8 (55.8–65.7)	30.1 (24.7–39.8)	68.6 (45.4–93.0)	3.8 (2.2–6.2)	8.7 (1.3–19.2)	8.0 (5.9–12.8)	22.4 (8.9–38.0)
7	105.8 (99.9–113.6)	116.4 (105.5–124.7)	12.7 (9.3–17.0)	17.8 (3.3–33.6)	73.6 (70.0–78.2)	100.3 (87.4–113.6)	11.4 (6.4–17.3)	23.4 (2.6–53.4)	17.6 (11.5–24.7)	34.1 (12.6–59.8)
8	79.0 (71.9–85.1)	77.1 (64.8–91.5)	7.1 (5.0–8.9)	7.0 (2.7–12.0)	20.5 (16.9–24.2)	27.4 (16.7–39.5)	2.6 (1.5–3.8)	3.1 (0.8–5.7)	5.4 (3.8–6.9)	7.3 (3.8–11.2)
8a	79.7 (71.1–89.1)	77.6 (63.7–93.5)	7.1 (5.1–9.3)	7.0 (2.7–11.7)	20.7 (17.0–24.5)	27.5 (16.8–41.1)	2.6 (1.5–3.8)	3.1 (1.0–5.8)	5.4 (3.9–7.1)	7.3 (3.9–11.6)
8b	103.4 (99.7–108.4)	110.9 (100.3–121.3)	8.9 (6.4–11.6)	10.4 (4.0–17.8)	26.0 (21.7–30.9)	40.7 (25.4–59.4)	3.2 (1.9–4.8)	4.7 (1.3–8.6)	6.7 (4.9–8.8)	10.9 (5.8–16.6)
8c	101.6 (97.6–107.2)	111.3 (100.5–122.3)	8.8 (6.3–11.3)	10.6 (3.8–18.0)	25.6 (21.4–30.5)	42.5 (25.5–61.4)	3.2 (1.9–4.7)	4.8 (1.4–8.7)	6.6 (4.8–8.7)	11.1 (5.7–17.2)
8d	79.9 (55.6–96.8)	67.3 (48.9–88.6)	7.2 (3.8–11.3)	6.0 (2.1–10.3)	20.9 (15.0–62.7)	23.5 (13.3–36.3)	2.6 (1.1–17.8)	2.7 (0.7–5.2)	5.5 (2.3–23.0)	6.1 (5.9–3.0)
8e	108.8 (105.8–113.0)	117.3 (110.1–124.0)	12.4 (8.2–17.3)	11.2 (4.2–19.5)	29.7 (24.1–37.2)	44.8 (26.6–64.4)	3.7 (2.1–5.6)	5.1 (1.2–9.6)	7.8 (5.6–10.9)	12.2 (6.2–18.8)
8f	92.4 (85.3–102.1)	102.7 (79.3–123.5)	8.3 (6.0–11.0)	10.3 (3.1–17.9)	53.4 (53.4–53.4)	59.5 (53.4–71.7)	3.3 (2.0–4.9)	5.1 (1.3–9.7)	6.7 (4.8–8.8)	11.9 (5.2–20.0)

obtained from the two methods appeared more similar (Table 4; Appendices S4 and S5; Figs. S18 and S33–S35). However, the confidence intervals of the PL estimates often overlapped with the credibility intervals of the BEAST estimates, except in a few cases (Appendix S5). For instance, PL age estimates for crown group *Nothofagus* (node L) were significantly younger than BEAST ages in five scenarios (scenarios 3–7). Because the UCLN relaxed-clock model implemented in BEAST does not assume direct autocorrelation of molecular rates between parent and descendants, and because BEAST age estimates take into account phylogenetic uncertainty, we will focus on BEAST ages rather than PL ages for the rest of our discussion. However, all our conclusions below also apply to the comparison of PL ages across the various scenarios (Table 4; Appendix S4).

The use of alternative priors on age constraints in BEAST analyses of scenarios 1–7 yielded different age estimates in each scenario, although credibility intervals widely overlapped in most cases (Appendix S6; Figs. S36–S42). Ages obtained with exponential priors were nearly always younger than ages obtained with uniform priors, and ages obtained with lognormal priors were nearly always younger than ages obtained with either uniform or exponential priors. For instance, the age of crown group *Nothofagus* (node L) was estimated to 53.4–93.2 Ma using uniform priors in scenario 1, 42.2–81.5 Ma using exponential priors in the same scenario, and 39.3–69.4 Ma using lognormal priors (Fig. S36). This was expected because of the way we parameterized these priors. Using the same hard minimum bound as in the uniform priors, exponential priors put more probability on younger ages. Using 10% younger hard minimum bounds than in the uniform or exponential priors, lognormal priors also put more probability on younger ages. A related effect was that 95% credibility intervals were often narrower in analyses using exponential priors and even more narrow using lognormal priors, in comparison with analyses using uniform priors (Appendix S6). Although this result was also expected, it has implications on the choice of priors applied to age constraints in molecular dating analyses using BEAST. Using exponential or lognormal priors can decrease the uncertainty in age estimates, and therefore, particular care must be taken in choosing and justifying the parameters of these priors.

The BEAST estimate of the mean substitution rate was fairly similar across different calibration scenarios, with the noticeable exception of scenarios 2, 8, and 8a–f where the estimated rate was much higher (Fig. 4). In the UCLN relaxed-clock model implemented in BEAST, the coefficient of rate variation was also estimated. This statistic measures the degree of rate heterogeneity among lineages, with a value of zero indicating the presence of a strict molecular clock. In each of the scenarios, there was departure from a strict clock, but to varying degrees (Table 3). The rates among branches were more clocklike under scenarios 2, 8, and 8a–f. These scenarios only involved four (scenario 2) and one (scenarios 8 and 8a–f) calibrations. These results suggest that rate

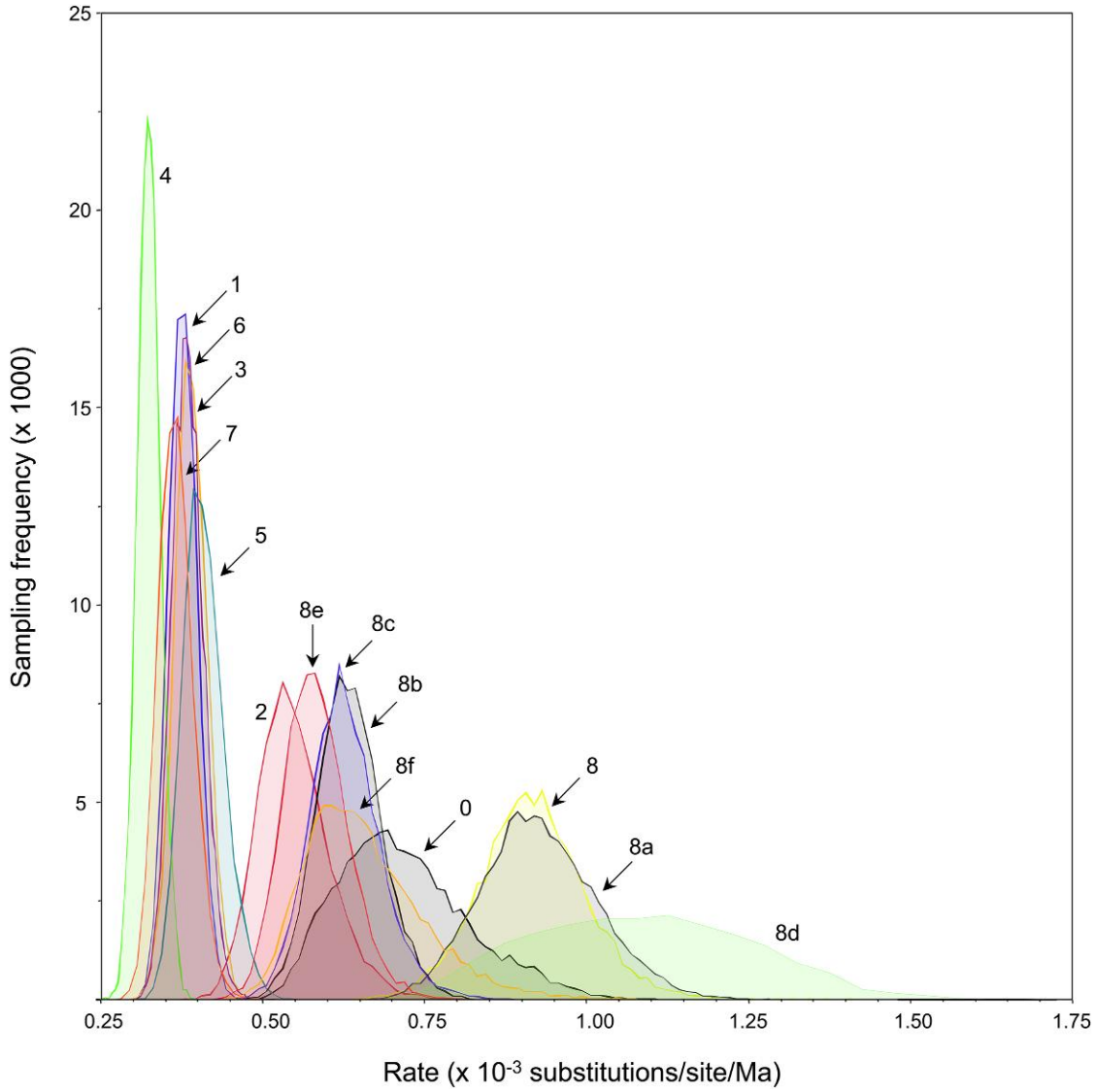

FIGURE 4. Posterior densities of mean substitution rates estimated in BEAST across calibration scenarios. This figure is available in black and white in print and in color at *Systematic Biology* online.

variation throughout the tree can be underestimated when the number of calibrations is small. One of the interesting consequences of this observation is that the use of calibrations individually might lead to quite disparate estimates of divergence times.

In PL analyses implemented by r8s, the optimized smoothing parameter can also provide an indication of how clocklike the sequences are (Sanderson 2003). Specifically, arbitrarily high smoothing parameters reflect clock-like behavior, whereas very low values are suggestive of rate heterogeneity. In our results, there did not appear to be a strong correspondence between the Bayesian coefficient of rate variation and the optimized smoothing parameter values. However, the calibration scenarios producing the highest PL smoothing parameters also produced some of the lowest BEAST estimates of the coefficient of rate variation (Table 3).

Ingroup versus Outgroup Calibration

Removing age constraints in either the outgroup (scenario 2) or the ingroup (scenario 3) did not have a drastic influence on age estimates across the tree (Table 4; Fig. 3a; Figs. S19 and S20). For instance, the estimated age for crown group *Nothofagus* (node L) was 53.4–93.2 Ma with scenario 1, 43.9–81.2 Ma with scenario 2, and 44.5–95.5 Ma with scenario 3. This observation also applies to analyses conducted using exponential or lognormal priors instead of uniform priors (Appendix S6). These results suggest that, in the absence of suitable internal age constraints for the ingroup, increasing the sampling of outgroups to include outgroup age constraints might be an appropriate solution. A recent example of this calibration strategy is the study by Michalak et al. (2010), who used fossils of Lauraceae and Monimiaceae to calibrate divergence

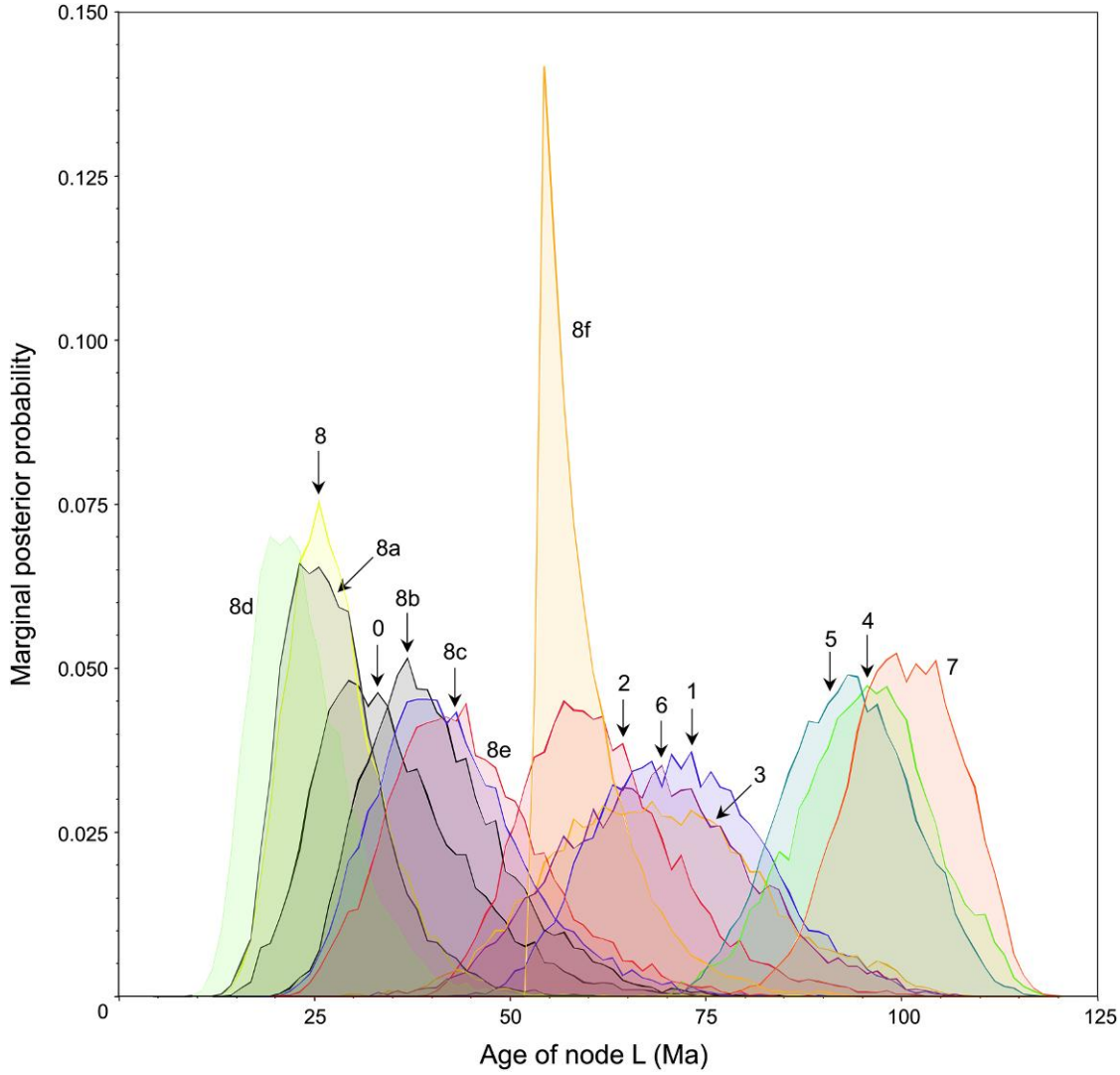

FIGURE 5. Posterior densities of estimated crown group ages for *Nothofagus* (node L) obtained with BEAST across calibration scenarios (see also Table 4). This figure is available in black and white in print and in color at *Systematic Biology* online.

times within the closely related Hernandiaceae. Removing age constraints in general (within or outside the ingroup) resulted in younger ages overall (this is consistent with the above observations on mean substitution rate and coefficient of variation of rates). This suggests that, even when suitable internal age constraints for the ingroup are available, more accurate age estimates can be obtained by combining both ingroup and outgroup calibrations.

Safe versus Risky Calibration

Adding early but risky age constraints generally led to older age estimates (Table 4; Fig. 3b; Appendix S4; Fig. S21). For instance, the crown group age of subgenus *Nothofagus* (node T) was estimated to 11.1–39.8 Ma in scenario 1 (safe constraints only) and 31.5–65.1 Ma in scenario 4 (safe and risky constraints). This observation also applies to analyses conducted using exponential or lognormal priors instead of uniform priors (Appendix S6). This result is not surprising: The risky constraints in this case study were derived from older fossils than those considered to be safe for the same nodes. This does not necessarily mean that risky age constraints were incorrectly placed on the phylogeny of extant taxa, but it draws further attention to the influence of such constraints on the whole dating analysis.

The difference between the risky and safe scenarios is greater within *Nothofagus*, due to the greater number of risky constraints in this clade (Fig. S2). Interestingly, when removing all outgroup calibrations (scenarios 2 and 5), the difference becomes even greater (Appendix S4). Conversely, when removing all ingroup calibrations (scenarios 3 and 6), the effect of risky age constraints became negligible. However, the latter is probably due to the fact that there was only one risky age constraint in the outgroup.

Our results show that the choices made in the common dilemma of choosing between early but risky and

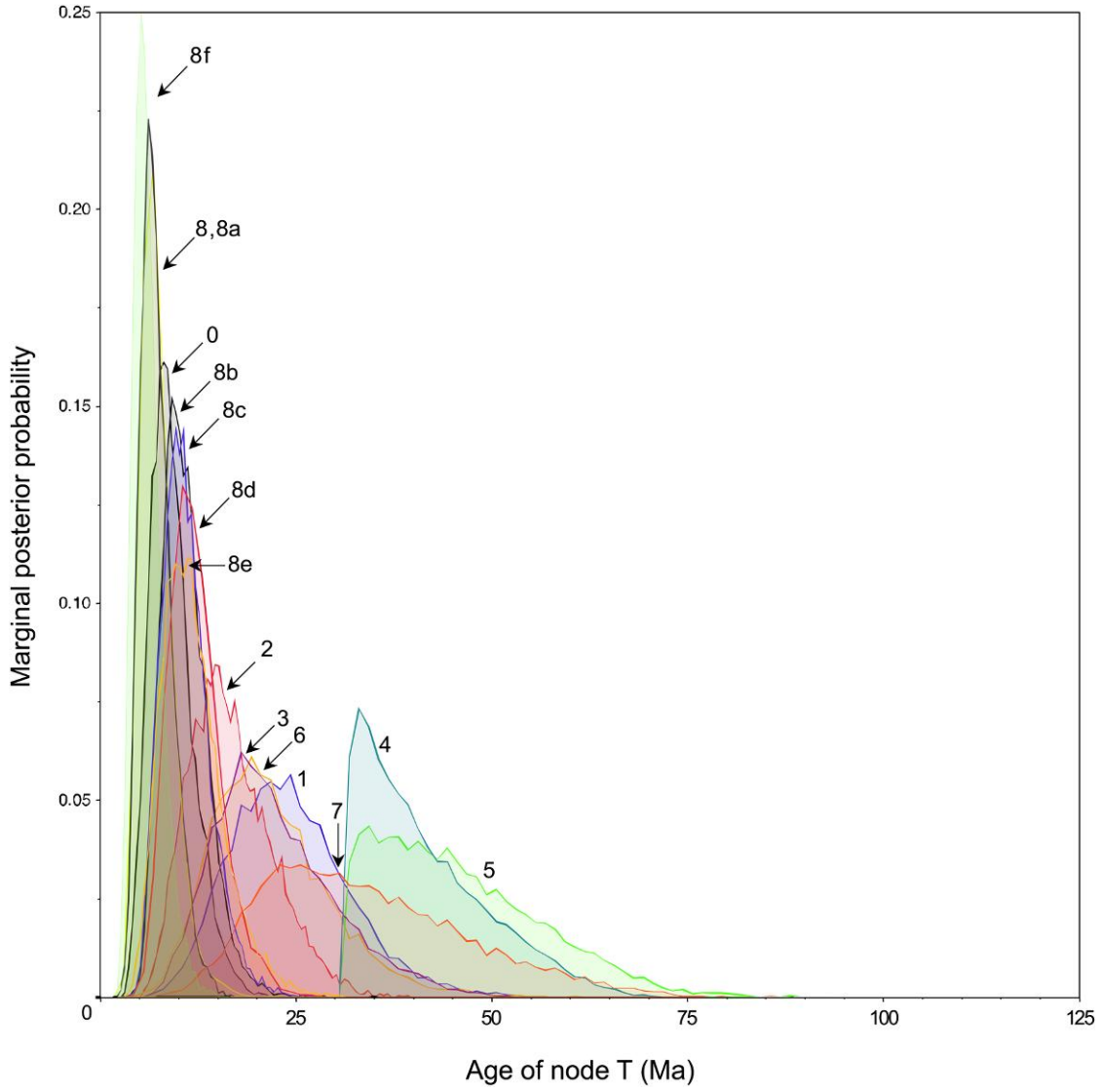

FIGURE 6. Posterior densities of estimated crown group ages for *Nothofagus* subgenus *Nothofagus* (node T) obtained with BEAST across calibration scenarios (see also Table 4). This figure is available in black and white in print and in color at *Systematic Biology* online.

safe but late fossils can strongly influence estimated ages. Therefore, risky fossils should clearly be identified when used for calibration in molecular dating studies. Similarly, when studies only use safe but late fossils, the presence of older but risky alternatives should be made clear. Ideally, both alternatives should be experimented with and their results compared when such dilemmas arise.

Fossil versus Vicariance Calibration

Using assumptions of vicariance to calibrate the dating analyses (scenario 7; Fig. S9) led to considerably older ages in the ingroup than those based on safe fossil calibration points (Table 4; Fig. 3c; Appendix S4; Fig. S24). For instance, the crown group age of subgenus *Brassospora* (node X) was estimated to 55.8–70.2 Ma in scenario 7 (calibrated with vicariance, including a minimum age of 55.8 Ma on this node) and 13.9–39.8 Ma in scenario 1 (calibrated with safe fossils but with no constraint on this particular node). Furthermore, when using safe calibration points, the 95% credibility intervals for nodes V (Australia vs. New Zealand in subgenus *Lophozonia*), P (Australia vs. New Zealand in subgenus *Fuscospora*), and X (New Guinea vs. New Caledonia in subgenus *Brassospora*) all resulted in ages less than 39 Ma, which is considerably younger than the youngest plausible ages for land connectivity between Australia/Antarctica and New Zealand (~55.8 Ma). These results are consistent with previous molecular dating studies (Martin and Dowd 1993; Cook and Crisp 2005; Knapp et al. 2005), which inferred that the geographic disjunctions within *Nothofagus* between Australia and New Zealand were better explained by long-distance

dispersal than vicariance. In contrast, when risky fossils were used in calibration (scenario 4), the estimated ages for nodes P (31.5–57.3 Ma) and X (31.8–63.6 Ma) were compatible with a vicariance age of ~55.8 Ma, although the disjunction within subgenus *Lophozonia* (node V, 5.3–47.5 Ma) remained too young for vicariance.

However, several alternative explanations could weaken inferences of long-distance dispersal in *Nothofagus* made in previous studies (Martin and Dowd 1993; Cook and Crisp 2005; Knapp et al. 2005) and the safe but late scenarios of the current study. First, the maximum age constraint of 125 Ma on the root (node U) might be an incorrect assumption, possibly resulting in the ages of some nodes being underestimated (but see above). Second, there might have been systematic changes in the rates of evolution, with generally higher rates of evolution early in the diversification of the group than at later stages. Third, the risky fossils might have provided a more accurate calibration of the phylogeny. In this sense, using early but risky calibrations for dating might be a way to provide more secure tests of hypotheses (such as vicariance) where one is aiming to find maximum plausible ages for nodes.

Finally, using assumptions of vicariance to provide calibrations in *Nothofagus* (scenario 7) surprisingly led to younger ages in the outgroup (core Fagales) than using fossil age constraints in both the ingroup and outgroup (scenario 1). For instance, the crown group age of Juglandaceae (node I) was estimated to 31.7–73.1 Ma in scenario 7 and 64.4–81.3 Ma in scenario 1 (Appendix S4; Fig. S24). This is surprising because one would expect that enforcing older ages in the ingroup would have a similar effect on the outgroup. Although this is probably an effect of the maximum age constraint on the root, it highlights the interdependency of molecular rates across the entire tree, even when using a method (uncorrelated relaxed clock) that does not assume direct autocorrelation. Using a method assuming autocorrelation (PL) resulted in even greater discrepancy (e.g., 25.8–35.4 Ma vs. 64.4–64.9 Ma for node I).

All the observations above also apply to analyses conducted using exponential or lognormal priors (Appendix S6). We note that when assuming vicariance to explain a particular divergence, the timings of vicariance and divergence should coincide, rather than the divergence being at least as old as the vicariance event (Crisp et al. 2011). Thus, in a scenario using assumptions of vicariance as calibrations, a normal or lognormal prior centered around the presumed timing of vicariance seems more appropriate than a uniform prior. Here, we opted for the lognormal prior (scenario 7b) because the median we used (e.g., 55.8 Ma for the vicariance between Australia and New Zealand) was already the youngest proposed date for this event (therefore, it seemed more likely that the divergence would be slightly older rather than younger than this date, as modeled with the lognormal prior). However, all the differences outlined above still hold when comparing scenario 7b (lognormal prior) with fossil-calibrated scenario 1, whichever way fossil age constraints were modeled (uniform, exponential, or lognormal).

Fossil versus Secondary Calibration

Age estimates obtained in molecular dating analyses relying on a secondary calibration point (scenarios 8 and 8a–f) were always younger than those obtained from our default scenario calibrated with safe fossils (scenario 1). This difference was particularly pronounced in the two scenarios using a secondary calibration point derived from the Wikström et al. (2001) study (scenarios 8 and 8a) and the one using a secondary calibration from the Bell et al. (2010) study (scenario 8d). In this case, 95% credibility intervals rarely overlapped and mean age estimates for nodes nested in *Nothofagus* were all at least three times younger than in the fossil-calibrated scenario (Table 4; Fig. 3d; Appendix S4; Figs. S25, S26, and S29). For instance, the crown group age of subgenus *Nothofagus* (node T) was estimated to 3.8–11.2 Ma in scenario 8 and 11.1–39.8 Ma in scenario 1. When using secondary calibration from two other recent angiosperm dating studies (scenario 8b: Magallón and Castillo 2009; scenario 8c: Wang et al. 2009), this difference was reduced to an average ratio of 2:1 in *Nothofagus*. All these observations were also true of the comparison of the secondary calibration scenarios with analyses of scenario 1 where age constraints were modeled with exponential (scenario 1a) or lognormal (scenario 1b) priors, despite the small age differences noticed between these analyses and the default one obtained with uniform priors (Appendix S6).

It is possible that the inconsistency between secondary-calibrated ages and fossil-calibrated ages is a result of the calibration schemes used in the higher-level analyses that provided secondary calibrations. Importantly, the Wikström et al. (2001) analysis used a single calibration point, based on the fossil taxa *Protofagacea* and *Antiquacupula*, both appearing in the Late Santonian of Georgia, US, and showing typical cupules, which are assumed to be a synapomorphy of Fagales. The authors consequently used these fossils to argue that the stem node of Fagales (i.e., the split between Cucurbitales and Fagales) is at least 84 Ma old. This is consistent with our analysis, where we used *Protofagacea* (F01) to put a similar minimum age constraint of 83.5 Ma on the same node (node U; Table 2; Fig. 1). However, given that Wikström et al. (2001, p. 2212) used a point calibration, and not minimum or maximum age constraints, all their age estimates were to be interpreted as minimum ages (as specified by the authors). In spite of this, secondary calibration points have been derived extensively from their study to fix the age of a particular node in molecular dating studies of angiosperm clades lacking suitable fossils (e.g., Goldblatt et al. 2002).

Magallón and Castillo (2009) used 49 minimum age constraints distributed across the angiosperm tree, including 1 on crown group Fagales, which were forced to be at least 93.5 Ma based on Late Cenomanian *Normapolles* pollen. *Nothofagus* was not included in their data

set. The minimum age constraint of 93.5 Ma was thus in effect applied to crown group core Fagales (node B of our tree, Fig. 1) in their analysis and the age estimated for this node was the same (93.5 Ma). Yet our analysis, calibrated with multiple (younger) minimum age constraints from the fossil record (scenario 1), suggested that this node (crown group core Fagales) is actually older (100.8–118.3 Ma with BEAST, 93.6–104.5 Ma with PL). Bell et al. (2010) used 36 minimum age constraints, including the same 84 Ma calibration as Wikström et al. (2001) applied to the stem node of Fagales, and obtained even younger ages in this clade than any of the other previous studies considered here. Thus, calibrating node B to a younger age than we found in scenario 1, whether using the estimates of Wikström et al. (2001: B = 60–61 Ma, scenario 8, or B = 57–65 Ma, scenario 8a), Magallón and Castillo (2009; B = 93.5 Ma, scenario 8b), or Bell et al. (2010; B = 43–68 Ma, scenario 8d), might explain why the ages obtained in these secondary calibration scenarios are much younger than in scenario 1.

On the other hand, Wang et al. (2009) used three different fossil taxa, assumed to be nested in the core Fagales, to argue for a minimum age constraint of 85 Ma on the crown group node of this clade (one of seven age constraints used in their dating analysis of the rosid clade). This is consistent with our own use of an 83.5 Ma minimum age constraint on the same node B, justified by one of their three fossils, *Antiquacupula* (F05 in our study), which we assumed to be a stem relative of Fagaceae (Table 2; Fig. 1). In spite of applying a younger age constraint on crown group core Fagales than Magallón and Castillo (2009), Wang et al. (2009) obtained a range of older estimates for the age of this node (e.g., 91–100 Ma for their BRC-1 analysis), which are more consistent with those we obtained in scenario 1. Therefore, the age of the secondary calibration itself used in the scenario based on Wang et al. (2009; B = 91–100 Ma, scenario 8c) is not sufficient to explain why ages obtained in this scenario are still much younger than in scenario 1.

In order to further test whether this discrepancy is due to the age (or age range) set for the secondary calibration, we conducted two additional experiments. In the first one (scenario 8e), we constrained the age of node B to be in the 95% credible range of 100.8–118.3 Ma obtained for the age of this node in scenario 1. Age estimates obtained across the tree remained much younger than in scenario 1, sometimes drastically so, and by an average ratio of 2:1 in *Nothofagus* (Table 4; Appendix S4; Fig. S30). When using a different node for this secondary calibration based on scenario 1 (crown group node of *Nothofagus*, node L = 53.4–93.2 Ma, scenario 8f), the result was the same (Fig. S31). Therefore, these experiments suggest that it is not only the age set for the secondary calibration that is responsible for the large differences observed but also the use of a single calibration point.

A potential explanation is that there have been significant departures from the molecular clock in this group and that, with only one internal calibration, the relaxed-clock dating methods used are unable to reconstruct the pattern of molecular rate change accurately. This is consistent with our observation of lower coefficients of rate variation for all the secondary calibration scenarios (Table 3).

Conclusion and Perspectives

Previous work showed that molecular date estimates can depend on the choice of genes (Bell and Donoghue 2005; Bell et al. 2005; Magallón and Sanderson 2005; Endicott and Ho 2008; Goodall-Copestake et al. 2009), models of nucleotide substitution (Schenk and Hufford 2010), partitioning strategies (Brandley et al. 2011), or relaxed-clock methods (Bell et al. 2005; Goodall-Copestake et al. 2009). However, an even greater source of variance in molecular dates appears to lie in the calibration scheme (Forest et al. 2005; Ho et al. 2008; Goodall-Copestake et al. 2009; Inoue et al. 2010). Our study was specifically designed to address this point, by estimating and comparing molecular dates from eight different calibration scenarios drawn from a range of plausible situations encountered by researchers using molecular dating methods. As illustrated in Figure 5, depicting the full range of estimates obtained for the crown group age of *Nothofagus*, we found that the choice of a specific calibration scheme can have a drastic influence on the outcome of the molecular dating analysis. In this case, mean age estimates varied from 23.5 (scenario 8d) to 100.3 Ma (scenario 7). When taking into account the full extent of 95% credibility intervals, possible ages for this node span an even greater range, from 13.3 to 113.6 Ma (Table 4).

Our results do not imply that molecular dating methods are unreliable and will systematically fail to estimate divergence times accurately. When properly taking into account the uncertainty around molecular dates (i.e., when estimating and considering 95% confidence or credibility intervals rather than point or mean estimates), many of our scenarios provided overlapping estimates. Therefore, the estimates did not completely contradict each other, despite being derived from very different calibration situations. In addition, the variance highlighted in Figure 5 does not necessarily apply to all nodes in the tree. For instance, estimated ages for nodes nested higher up in the tree such as the crown group node of subgenus *Nothofagus* (Fig. 6: node T) were comparatively more consistent across calibration scenarios than the previous example.

The main message we would like to convey with this study is instead the following: Calibration is critical and highly influential on molecular dating inference and consequently should represent an essential and integral aspect of any research using molecular dating methods. Errors due to calibration are probably much larger than typically acknowledged, and keeping this in mind, every effort should be made to reduce all sources of error that can be controlled. Setting up a well-supported and accurate calibration scheme should involve research and

critical scrutiny of literature at all stages, including the vetting of all geochronological data on the fossiliferous rocks and using the latest GTS, reassessing relationships of fossil taxa to extant taxa in light of current phylogenetic hypotheses, and providing a clear rationale for the conversion of stratigraphic into absolute ages and the use of minimum versus maximum age constraints (or more complex priors). Each of these steps should be documented, supported by the appropriate references, and made available in publications to allow replication (see also Clarke et al. 2011; Parham et al., 2011). In-depth preparation and documentation of the calibration procedure will not only improve the quality of estimated divergence times but will also allow a critical assessment by anyone wishing to take into account results from a molecular dating contribution.

Compared with previous assessments of the impact of calibration on molecular dating inference, our study was conducted on a plant group with a particularly rich fossil record, thus providing an empirical qualitative test of different calibration schemes. The extent to which the conclusions from this comparative case study are applicable to molecular dating analyses in general remains to be tested in other taxonomic groups. In particular, the directionality of age differences between scenarios may depend on the specific characteristics of each group. However, the following conclusions may be more general and help with designing and assessing the results from future molecular dating studies:

1. A greater number of age constraints leads to improved date estimates. Removing age constraints in general (within or outside the ingroup) results in younger ages overall.
2. Even when many fossil age constraints are available and incorporated, large uncertainties remain, and therefore, age estimates should be taken with all the necessary caution (i.e., avoid the illusory precision of mean ages and report 95% confidence intervals instead).
3. Using a single calibration point, for example, a secondary calibration derived from a previous higher-level study, can lead to biased estimates. Sampling more outgroup taxa to include external fossil age constraints is a better option than relying on secondary calibration when no suitable fossils are available to calibrate the group of interest.

We therefore make some general recommendations (see also Gandolfo et al. 2008):

1. When citing a fossil to justify an age constraint, provide separate references for: (i) the original description of the fossil; (ii) the assignment of the fossil on the phylogeny of extant taxa; and (iii) the revised stratigraphy and geochronology of the sediments where the fossil has been found.
2. If the phylogenetic relationship of the fossil to extant taxa has not been tested explicitly using phylogenetic methods, mention whether the fossil assignment considered in the calibration scheme is intuitive or based on shared apomorphies. If based on shared apomorphies, mention explicitly the character(s) involved as well as the phylogenetic framework where these characters are inferred to be apomorphic.
3. Always check the latest stratigraphic and geochronological revisions of the rocks involved and mention explicitly the GTS of reference used in the calibration scheme. This procedure will allow more precise and accurate calibration than available from a literal reading of most original fossil descriptions.

Supplementary Material

Supplementary material, including data files and online-only appendices, can be found at the Dryad data repository (Dryad doi 10.5061/dryad.qq106tm4).

Funding

This work was initiated through a workshop funded by the Australian Research Council–New Zealand Research Network for Vegetation Function (Working Group 32, "Calibrating the Evolutionary History of Southern Hemisphere Plant Clades"), held in 2008 at the Royal Botanic Gardens, Melbourne, and organized by M.A.G., D.J.C., and P.W. We thank Mark Westoby for his dynamic leadership of the Network and encouragement of this Working Group. Additional support came from NSF [DEB-0919071 to P.W. and M.A.G., DEB-0345750 to P.W. and M.A.G., DEB-0918932 to M.A.G.].

Acknowledgments

We thank Andy Anderson and two anonymous reviewers for very helpful comments on an earlier draft of this paper.

References

Anderson C.L., Bremer K., Friis E.M. 2005. Dating phylogenetically basal eudicots using *rbcL* sequences and multiple fossil reference points. Am. J. Bot. 92:1737–1748.

APG. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot. J. Linn. Soc. 161:105–121.

Barker P.F., Filippelli G.M., Florindo F., Martin E.E., Scher H.D. 2007. Onset and role of the Antarctic Circumpolar Current. Deep Sea Res. 54:2388–2398.

Bell C.D., Donoghue M.J. 2005. Dating the Dipsacales: comparing models, genes, and evolutionary implications. Am. J. Bot. 92: 284–296.

Bell C.D., Soltis D.E., Soltis P.S. 2005. The age of the angiosperms: a molecular timescale without a clock. Evolution. 59:1245–1258.

Bell C.D., Soltis D.E., Soltis P.S. 2010. The age and diversification of the angiosperms re-revisited. Am. J. Bot. 97:1296–1303.

Belt E.S., Hartman J.H., Diemer J.A., Kroeger T.J., Tibert N.E., Curran H.A. 2004. Unconformities and age relationships, Tongue River and older members of the Fort Union Formation (Paleocene), western Williston Basin, U.S.A. Rocky Mt. Geol. 39:113–140.

Brandley M.C., Wang Y., Guo X., Montes de Oca A.N., Feria-Ortiz M., Hikida T., Ota H. 2011. Accommodating heterogenous rates of evolution in molecular divergence dating methods: an example using intercontinental dispersal of *Plestiodon (Eumeces)* lizards. Syst. Biol. 60:3–15.

Britton T., Anderson C.L., Jacquet D., Lundqvist S., Bremer K. 2007. Estimating divergence times in large phylogenetic trees. Syst. Biol. 56:741–752.

Burleigh J.G., Hilu K., Soltis D. 2009. Inferring phylogenies with incomplete data sets: a 5-gene, 567-taxon analysis of angiosperms. BMC Evol. Biol. 9:61.

Burnett J.A. 1996. Nannofossils and Upper Cretaceous (sub-)stage boundaries-state of the art. J. Nannoplankton Res. 18:23–32.

Burnham R.J. 2008. Hide and go seek: what does presence mean in the fossil record? Ann. Mo. Bot. Gard. 95:51–71.

Clarke J.A., Ksepka D.T., Stucchi M., Urbina M., Giannini N., Bertelli S., Narvaez Y., Boyd C.A. 2007. Paleogene equatorial penguins challenge the proposed relationship between biogeography, diversity, and Cenozoic climate change. Proc. Natl. Acad. Sci. U.S.A. 104:11545–11550.

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

Clyde W.C., Hamzi W., Finarelli J.A., Wing S.L., Schankler D., Chew A. 2007. Basin-wide magnetostratigraphic framework for the Bighorn Basin, Wyoming. Geol. Soc. Am. Bull. 119:848–859.

Cook L., Crisp M. 2005. Not so ancient: the extant crown group of *Nothofagus* represents a post-Gondwanan radiation. Proc. R. Soc. Lond. B. Biol. Sci. 272:2535–2544.

Crame J.A., Francis J.E., Cantrill D.J., Pirrie D. 2004. Maastrichtian stratigraphy of Antarctica. Cretaceous Res. 25: 411–423.

Crane P.R. 1985. Phylogenetic analysis of seed plants and the origin of angiosperms. Ann. Mo. Bot. Gard. 72:716–793.

Crane P.R., Herendeen P., Friis E.M. 2004. Fossils and plant phylogeny. Am. J. Bot. 91:1683–1699.

Crepet W.L., Nixon K.C. 1989. Earliest megafossil evidence of Fagaceae: phylogenetic and biogeographic implications. Am. J. Bot. 76:842–855.

Crisp M.D., Trewick S.A., Cook L.G. 2011. Hypothesis testing in biogeography. Trends Ecol. Evol. 26:66–72.

Currano E.D., Labandeira C.C., Wilf P. 2010. Fossil insect folivory tracks paleotemperature for six million years. Ecol. Monogr. 80:547–567.

Davis C.C., Webb C.O., Wurdack K.J., Jaramillo C.A., Donoghue M.J. 2005. Explosive radiation of Malpighiales supports a mid-Cretaceous origin of modern tropical rain forests. Am. Nat. 165:E36–E65.

Dettmann M.E., Playford G. 1968. Taxonomy of some Cretaceous spores and pollen grains from Eastern Australia. Proc. R. Soc. Victoria. 81:69–93.

Dettmann M.E., Pocknall D.T., Romero E.J., Zamaloa M.C. 1990. *Nothofagidites* Erdtman ex Potonie, 1960: a catalogue of species with notes on the palaeogeographic distribution of *Nothofagus* Bl. (Southern Beech). New Zeal. Geol. Surv. Paleontol. Bull. 60:1–79.

Dettmann M.E., Thomson M.R.A. 1987. Cretaceous palynomorphs from the James Ross Island area, Antarctica - a pilot study. Brit. Antarct. Surv. Bull. 77:13–59.

Donoghue M.J., Doyle J.A., Gauthier J., Kluge A.G., Rowe T. 1989. The importance of fossils in phylogeny reconstruction. Annu. Rev. Ecol. Syst. 20:431–460.

Doyle J.A. 1992. Revised palynological correlations of the lower Potomac Group (USA) and the Cocobeach sequence of Gabon (Barremian-Aptian). Cretaceous Res. 13:337–349.

Doyle J.A. 2005. Early evolution of angiosperm pollen as inferred from molecular and morphological phylogenetic analyses. Grana. 44:227–251.

Doyle J.A., Donoghue M.J. 1992. Fossils and seed plant phylogeny reanalyzed. Brittonia. 44:89–106.

Doyle J.A., Donoghue M.J. 1993. Phylogenies and angiosperm diversification. Paleobiology. 19:141–167.

Doyle J.A., Endress P.K. 2000. Morphological phylogenetic analysis of basal angiosperms: comparison and combination with molecular data. Int. J. Plant Sci. 161:S121–S153.

Doyle J.A., Endress P.K. 2010. Integrating Early Cretaceous fossils into the phylogeny of living angiosperms: Magnoliidae and eudicots. J. Syst. Evol. 48:1–35.

Doyle J.A., Hotton C.L. 1991. Diversification of early angiosperm pollen in a cladistic context. Pollen et Spores. 44: 169–195.

Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4:e88.

Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.

Eernisse D.J., Kluge A.G. 1993. Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. Mol. Biol. Evol. 10:1170–1195.

Endicott P., Ho S. 2008. A Bayesian evaluation of human mitochondrial substitution rates. Am. J. Hum. Genet. 82:895–902.

Ewing T.E. 1981. Regional stratigraphy and structural setting of the Kamloops Group, south-central British Columbia. Can. J. Earth Sci. 18:1464–1477.

Exon N.F., Hill P.J., Lafoy Y., Hiene C., Bernadel G. 2006. Kenn Plateau off northeastern Australia: a continental fragment in the southwest Pacific jigsaw. Aust. J. Earth Sci. 53:541–564.

Forest F. 2009. Calibrating the Tree of Life: fossils, molecules and evolutionary timescales. Ann. Bot. 104:789–794.

Forest F., Savolainen V., Chase M.W., Lupia R., Bruneau A., Crane P.R. 2005. Teasing apart molecular- versus fossil-based error estimates when dating phylogenetic trees: a case study in the birch family (Betulaceae). Syst. Bot. 30:118–133.

Fox R.C. 1990. The succession of Paleocene mammals in western Canada. Geol. Soc. Am. Spec. Pap. 243:51–70.

Furness C.A., Rudall P.J. 2004. Pollen aperture evolution - a crucial factor for eudicot success? Trends Plant Sci. 9: 154–158.

Gaina C., Müller D.R., Royer J.-Y., Stock J., Hardebeck J., Symonds P. 1998. The tectonic history of the Tasman Sea: a puzzle with 13 pieces. J. Geophys. Res. 103:12413–12433.

Gandolfo M.A., Nixon K.C., Crepet W.L. 2008. Selection of fossils for calibration of molecular dating models. Ann. Mo. Bot. Gard. 95:34–42.

Gernandt D.S., Magallon S., Geada Lopez G., Zeron Flores O., Willyard A., Liston A. 2008. Use of simultaneous analyses to guide fossil-based calibrations of Pinaceae phylogeny. Int. J. Plant Sci. 169:1086–1099.

Goldblatt P., Savolainen V., Porteous O., Sostaric I., Powell M., Reeves G., Manning J.C., Barraclough T.G., Chase M.W. 2002. Radiation in the Cape flora and the phylogeny of peacock irises *Moraea* (Iridaceae) based on four plastid DNA regions. Mol. Phylogenet. Evol. 25:341–360.

Goodall-Copestake W.P., Harris D.J., Hollingsworth P.M. 2009. The origin of a mega-diverse genus: dating *Begonia* (Begoniaceae) using alternative datasets, calibrations and relaxed clock methods. Bot. J. Linn. Soc. 159:363–380.

Gradstein F.M., Ogg J.G., Smith A.G., Agterberg F.P., Bleeker W., Cooper R.A., Davydov V., Gibbard P., Hinnov L., House M.R., Lourens L., Luterbacher H.-P., McArthur J., Melchin M.J., Robb L.J., Shergold J., Villeneuve M., Wardlaw B.R., Ali J., Brinkhuis H., Hilgen F.J., Hooker J., Howarth R.J., Knoll A.H., Laskar J., Monechi S., Powell J., Plumb K.A., Raffi I., Röhl U., Sanfilippo A., Schmitz B., Shackleton N.J., Shields G.A., Strauss H., Dam J.V., Veizer J., Kolfschoten T.v., Wilson D. 2004. A geologic time scale 2004. Cambridge, UK: Cambridge University Press.

Graur D., Martin W. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. Trends Genet. 20:80–86.

Hall T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41:95–98.

Harris W.K. 1965. Basal Tertiary microfloras from the Princetown area, Victoria, Australia. Palaeontogr. Abt. B. Palaophytol. 115:75–106.

Hedges S.B., Kumar S. 2003. Genomic clocks and evolutionary timescales. Trends Genet. 19:200–206.
Hedges S.B., Kumar S. 2009. Discovering the timetree of life. In: Hedges S.B., Kumar S., editors. The timetree of life. New York: Oxford University Press. p. 3–18.
Herbert J., Chase M.W., Möller M., Abbott R.J. 2006. Nuclear and plastid DNA sequences confirm the placement of the enigmatic *Canacomyrica monticola* in Myricaceae. Taxon. 55:349–357.
Herendeen P.S., Crane P.R., Drinnan A.N. 1995. Fagaceous flowers, fruits and cupules from the Campanian (Late Cretaceous) of Central Georgia, USA. Int. J. Plant Sci. 156:93–116.
Hermsen E.J., Hendricks J.R. 2008. W(h)ither fossils? Studying morphological character evolution in the age of molecular sequences. Ann. Mo. Bot. Gard. 95:72–100.
Hill R.S. 1983. *Nothofagus macrofossils* from the Tertiary of Tasmania. Alcheringa. 7:169–183.
Hill R.S. 1991. Tertiary *Nothofagus* (Fagaceae) macrofossils from Tasmania and Antarctica and their bearing on the evolution of the genus. Bot. J. Linn. Soc. 105:73–112.
Hill R.S. 2001. Biogeography, evolution and palaeoecology of *Nothofagus* (Nothofagaceae): the contribution of the fossil record. Aust. J. Bot. 49:321–332.
Hill R.S., Jordan G.J. 1993. The evolutionary history of *Nothofagus* (Nothofagaceae). Aust. Syst. Bot. 6:111–126.
Hill R.S., Read J. 1991. A revised infrageneric classification of *Nothofagus* (Fagaceae). Bot. J. Linn. Soc. 105:37–72.
Hilu K.W., Borsch T., Muller K., Soltis D.E., Soltis P.S., Savolainen V., Chase M.W., Powell M.P., Alice L.A., Evans R., Sauquet H., Neinhuis C., Slotta T.A.B., Rohwer J.G., Campbell C.S., Chatrou L.W. 2003. Angiosperm phylogeny based on *matK* sequence information. Am. J. Bot. 90:1758–1776.
Ho S.Y.W. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. J. Avian Biol. 38:409–414.
Ho S.Y.W., Larson G. 2006. Molecular clocks: when times are a-changin'. Trends Genet. 22:79–83.
Ho S.Y.W., Phillips M.J. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst. Biol. 58:367–380.
Ho S.Y.W., Saarma U., Barnett R., Haile J., Shapiro B. 2008. The effect of inappropriate calibration: three case studies in molecular ecology. PLoS One. 3:e1615.
Hug L.A., Roger A.J. 2007. The impact of fossils and taxon sampling on ancient molecular dating analyses. Mol. Biol. Evol. 24:1889–1897.
Hughes N.F., McDougall A.B. 1990. Barremian-Aptian angiospermid pollen records from southern England. Rev. Palaeobot. Palynol. 65:145–151.
Humphries C.J. 1981. Biogeographical methods and the southern beeches (Fagaceae: *Nothofagus*). In: Forey P.L., editor. The evolving biosphere: chance, change and challenge. London: Cambridge University Press. p. 283–297.
Inoue J., Donoghue P.C.J., Yang Z. 2010. The impact of the representation of fossil calibrations on Bayesian estimation of species divergence times. Syst. Biol. 59:74–89.
Jordan G. 1999. A new Early Pleistocene species of *Nothofagus* and the climatic implications of co-ocurring *Nothofagus* fossils. Aust. Syst. Bot. 12:757–765.
Jordan G.J., Hill R.S. 1999. The phylogenetic affinities of *Nothofagus* (Nothofagaceae) leaf fossils based on combined molecular and morphological data. Int. J. Plant Sci. 160:1177–1188.
Keane T., Creevey C., Pentony M., Naughton T., McLnerney J. 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. BMC Evol. Biol. 6:29.
Knapp M., Stockler K., Havell D., Delsuc F., Sebastiani F., Lockhart P.J. 2005. Relaxed molecular clock provides evidence for long-distance dispersal of Nothofagus (southern beech). PLoS Biol. 3:38–43.
Kumar S. 2005. Molecular clocks: four decades of evolution. Nat. Rev. Genet. 6:654–662.
Ladiges P.Y., Cantrill D. 2007. New Caledonia-Australian connections: biogeographic patterns and geology. Aust. Syst. Bot. 20:383–389.
Lee M.S.Y., Oliver P.M., Hutchinson M.N. 2009. Phylogenetic uncertainty and molecular clock calibrations: a case study of legless lizards (Pygopodidae, Gekkota). Mol. Phylogenet. Evol. 50:661–666.
Li R.Q., Chen Z.D., Lu A.M., Soltis D.E., Soltis P.S., Manos P.S. 2004. Phylogenetic relationships in Fagales based on DNA sequences from three genomes. Int. J. Plant Sci. 165:311–324.
Linder H.P., Crisp M.D. 1995. *Nothofagus* and Pacific biogeography. Cladistics. 11:5–32.
Livermore R.A., Nankivell A., Eagles G., Morris P. 2005. Paleogene opening of Drakes Passage. Earth Planet Sci. Lett. 236:459–470.
Luterbacher H., Ali J., Brinkhuis H., Gradstein F., Hooker J., Monechi S., Ogg J., Powell J., Röhl U., SanFilippo A., Schmitz B. 2004. The Paleogene period. In: Gradstein et al., editors. A geologic time scale 2004. Cambridge, UK: Cambridge University Press. p. 384–408.
Macphail M.K. 2007. Australian palaeoclimates: Cretaceous to Tertiary. A review of palaeobotanical and related evidence to the year 2000. Bentley (WA): CRC LEME.
Macphail M.K., Alley N.F., Truswell E.M., Sluiter I.R.K. 1994. Early Tertiary vegetation: evidence from spores and pollen. In: Hill R.S., editor. History of the Australian vegetation: Cretaceous to Recent. Cambridge, UK: Cambridge University Press. p. 189–261.
Magallón S. 2007. From fossils to molecules: phylogeny and the core eudicot floral groundplan in Hamamelidoideae (Hamamelidaceae, Saxifragales). Syst. Bot. 32:317–347.
Magallón S. 2010. Using fossils to break long branches in molecular dating: a comparison of relaxed clocks applied to the origin of angiosperms. Syst. Biol. 59:384–399.
Magallón S., Castillo A. 2009. Angiosperm diversification through time. Am. J. Bot. 96:349–365.
Magallón S.A. 2004. Dating lineages: molecular and paleontological approaches to the temporal framework of clades. Int. J. Plant Sci. 165:S7–S21.
Magallón S.A., Sanderson M.J. 2005. Angiosperm divergence times: the effect of genes, codon positions, and time constraints. Evolution. 59:1653–1670.
Manchester S.R. 1987. The fossil history of the Juglandaceae. Mo. Bot. Gard. Monogr. 21:1–137.
Manchester S.R. 1994. Fruits and seeds of the Middle Eocene Nut Beds Flora, Clarno Formation, Oregon. Palaeontogr. Am. 58:1–205.
Manchester S.R., Dilcher D.L. 1982. Pterocaryoid fruits (Juglandaceae) in the Paleogene of North America and their evolutionary and biogeographic significance. Am. J. Bot. 69:275–286.
Manchester S.R., Dilcher D.L. 1997. Reproductive and vegetative morphology of *Polyptera* (Juglandaceae) from the Paleocene of Wyoming and Montana. Am. J. Bot. 84:649–663.
Manchester S.R., Dillhoff R.M. 2004. *Fagus* (Fagaceae) fruits, foliage, and pollen form the Middle Eocene of Pacific Northwestern North America. Can. J. Bot. 82:1509–1517.
Manchester S.R., Pigg K., Crane P.R. 2004. *Palaeocarpinus dakotaensis* sp. nov. (Betulaceae: Coryloideae) and associated staminate catkins, pollen and leaves from the Paleocene of North Dakota. Int. J. Plant Sci. 165:1135–1148.
Manos P.S. 1997. Systematics of *Nothofagus* (Nothofagaceae) based on rDNA spacer sequences (ITS): taxonomic congruence with morphology and plastid sequences. Am. J. Bot. 84:1137–1155.
Manos P.S., Soltis P.S., Soltis D.E., Manchester S.R., Oh S.-H., Bell C.D., Dilcher D.L., Stone D.E. 2007. Phylogeny of extant and fossil Juglandaceae inferred from the integration of molecular and morphological data sets. Syst. Biol. 56:412–430.
Manos P.S., Steele K.P. 1997. Phylogenetic analyses of "higher" Hamamelididae based on plastid sequence data. Am. J. Bot. 84: 1407–1419.
Marshall C.R. 2008. A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. Am. Nat. 171:726–742.
Martin P.G., Dowd J.M. 1993. Using sequences of rbcL to study phylogeny and biogeography of *Nothofagus* species. Aust. Syst. Bot. 6:441–447.
Michalak I., Zhang L.-B., Renner S.S. 2010. Trans-Atlantic, trans-Pacific and trans-Indian Ocean dispersal in the small Gondwanan Laurales family Hernandiaceae. J. Biogeogr. 37:1214–1226.
Mildenhall D.C., Pocknall D.T. 1984. Paleobotanical evidence for changes in Miocene and Pliocene climates in New Zealand. In: Vogel J.C., editor. Late Cenozoic paleoclimates of the Southern Hemisphere. Rotterdam (The Netherlands): A.A. Balkema. p. 159–171.

Moore M.J., Soltis P.S., Bell C.D., Burleigh J.G., Soltis D.E. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proc. Natl. Acad. Sci. U.S.A. 107: 4623–4628.
Mortimer N. 2004. New Zealand's geological foundations. Gondwana Res. 7:261–272.
Nixon K.C. 1989. Origins of Fagaceae. In: Crane P.R., Blackmore S., editors. Evolution, systematics, and fossil history of the Hamamelidae. Oxford: Clarendon Press. p. 23–43.
Nixon K.C., Crepet W.L., Stevenson D., Friis E.M. 1994. A reevaluation of seed plant phylogeny. Ann. Mo. Bot. Gard. 81: 484–533.
Ogg J., Smith A. 2004. The geomagnetic polarity time scale. In: Gradstein et al., editors. A geologic time scale 2004. Cambridge, UK: Cambridge University Press. p. 63–86.
Ørsted A.S. 1871. Bidrag til Kundshab om Egefamilien i Fortid og Nutid. Geogr. Tidsskr. 9:331–538.
Parham J.F., Donoghue P.C.J., Bell C.J., Calway T.D., Head J.J., Holroyd P.A., Inoue J.G., Irmis R.B., Joyce W.G., Ksepka D.T., Patané J.S.L., Smith N.D., Tarver J.E., van Tuinen M., Yang Z., Angielczyk K.D., Greenwood J.M., Hipsley C.A., Jacobs L., Makovicky P.J., Müller J., Smith K.T., Theodor J.M., Warnock R.C.M., Benton M.J. 2011. Best practices for justifying fossil calibrations. Syst. Biol. 61:346–359.
Pigg K., Manchester S.R., Wehr W. 2003. *Corylus, Carpinus* and *Palaeocarpinus* (Betulaceae) from the middle Eocene Klondike Mountain and Allenby Formations of northwestern North America. Int. J. Plant Sci. 164:807–822.
Qiu Y.-L., Chase M.W., Hoot S.B., Conti E., Crane P.R., Sytsma K.J., Parks C.R. 1998. Phylogenetics of the Hamamelidae and their allies: parsimony analyses of nucleotide sequences of the plastid gene Oxford, UK: University of Oxford. *rbcL*. Int. J. Plant Sci. 159:891–905.
Rambaut A. 2002. Se-Al: sequence alignment editor. Version 2.0a11. Oxford, UK: University of Oxford. Available from: http://tree.bio.ed.ac.uk/software/seal/.
Rambaut A., Drummond A.J. 2007. Tracer, version 1.5. Oxford, UK: University of Oxford.
Romero E.J. 1973. Polen fósil de *"Nothofagus" ("Nothofagidites")* del Cretácico y Paleoceno de Patagonia. Rev. Mus. La Plata. 7:291–303.
Rutschmann F. 2006. Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times. Divers. Distrib. 12:35–48.
Sanders K.L., Lee M.S.Y. 2007. Evaluating molecular clock calibrations using Bayesian analyses with soft and hard bounds. Biol. Lett. 3:275–279.
Sanderson M.J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14:1218–1231.
Sanderson M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19:101–109.
Sanderson M.J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics. 19:301–302.
Sanderson M.J., Thorne J.L., Wikstrom N., Bremer K. 2004. Molecular evidence on plant divergence times. Am. J. Bot. 91: 1656–1665.
Sauquet H., Weston P.H., Anderson C.L., Barker N.P., Cantrill D.J., Mast A.R., Savolainen V. 2009. Contrasted patterns of hyperdiversification in Mediterranean hotspots. Proc. Natl. Acad. Sci. U.S.A. 106:221–225.
Sauquet H., Weston P.H., Barker N.P., Anderson C.L., Cantrill D.J., Savolainen V. 2009. Using fossils and molecular data to reveal the origins of the Cape proteas (subfamily Proteoideae). Mol. Phylogenet. Evol. 51:31–43.
Schenk J.J., Hufford L. 2010. Effects of substitution models on divergence time estimates: simulations and an empirical study of model uncertainty using Cornales. Syst. Bot. 35: 578–592.
Scriven L.J., Hill R.S. 1995. Macrofossil Casuarinaceae: their identification and the oldest macrofossil record, *Gymnostoma antiquum* sp. nov., from the Late Paleocene of New South Wales, Australia. Aust. Syst. Bot. 8:1035–1053.
Secord R., Gingerich P.D., Smith M.E., Clyde W.C., Wilf P., Singer B.S. 2006. Geochronology and mammalian biostratigraphy of middle and upper Paleocene continental strata, Bighorn Basin, Wyoming. Am. J. Sci. 306:211–245.
Setoguchi H., Ono M., Doi Y., Koyama H., Tsuda M. 1997. Molecular phylogeny of *Nothofagus* (Nothofagaceae) based on the atpB-rbcL intergenic spacer of the chloroplast DNA. J. Plant Res. 110: 469–484.
Sims H.J., Herendeen P.S., Crane P.R. 1998. New genus of fossil Fagaceae from the Santonian (Late Cretaceous) of Central Georgia. Int. J. Plant Sci. 159:391–404.
Smith M.E., Singer B.S., Carroll A.R. 2004. Ar-40/Ar-19 geochronology of the Eocene Green River Formation, Wyoming. Geol. Soc. Am. Bull. 116:253–256.
Smith S.A., Beaulieu J.M., Donoghue M.J. 2010. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. Proc. Natl. Acad. Sci. U.S.A. 107:5897–5902.
Soltis D.E., Gitzendanner M.A., Soltis P.S. 2007. A 567-taxon data set for angiosperms: the challenges posed by Bayesian analyses of large data sets. Int. J. Plant Sci. 168:137–157.
Soltis D.E., Soltis P.S., Chase M.W., Mort M.E., Albach D.C., Zanis M., Savolainen V., Hahn W.H., Hoot S.B., Fay M.F., Axtell M., Swensen S.M., Prince L.M., Kress W.J., Nixon K.C., Farris J.S. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. Bot. J. Linn. Soc. 133:381–461.
Springer M.S., Teeling E.C., Madsen O., Stanhope M.J., de Jong W.W. 2001. Integrated fossil and molecular data reconstruct bat echolocation. Proc. Natl. Acad. Sci. U.S.A. 98:6241–6246.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.
Stover L.E., Partridge A.D. 1973. Tertiary and Late Cretaceous spores and pollen from the Gippsland Basin, southeastern Australia. Proc. R. Soc. Victoria. 85:237–286.
Swenson U., Backlund A., McLoughlin S., Hill R.S. 2001. *Nothofagus* biogeography revisited with special emphasis on the enigmatic distribution of subgenus *Brassospora* in New Caledonia. Cladistics. 17:28–47.
Thorne J.L., Kishino H., Painter I.S. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15: 1647–1657.
van Tuinen M., Hedges S.B. 2004. The effect of external and internal fossil calibrations on the avian evolutionary timescale. J. Paleontol. 78:45–50.
Wang H., Moore M.J., Soltis P.S., Bell C.D., Brockington S.F., Alexandre R., Davis C.C., Latvis M., Manchester S.R., Soltis D.E. 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. Proc. Natl. Acad. Sci. U.S.A. 106:3853–3858.
Welch J.J., Bromham L. 2005. Molecular dating when rates vary. Trends Ecol. Evol. 20:320–327.
Westerhold T., Röhl U. 2009. High resolution cyclostratigraphy of the early Eocene – new insights into the origin of the Cenozoic cooling trend. Clim. Past. 5:309–327.
Wiens J.J., Kuczynski C.A., Townsend T., Reeder T.W., Mulcahy D.G., Sites J.W. 2010. Combining phylogenomics and fossils in higher-level squamate reptile phylogeny: molecular data change the placement of fossil taxa. Syst. Biol. 59: 674–688.
Wikström N., Savolainen V., Chase M.W. 2001. Evolution of the angiosperms: calibrating the family tree. Proc. R. Soc. Lond. B. Biol. Sci. 268:2211–2220.
Wilf P. 2000. Late Paleocene-early Eocene climate changes in southwestern Wyoming: paleobotanical analysis. Geol. Soc. Am. Bull. 112:292–307.
Wing S., Bao H., Koch P. 2000. An early Eocene cool period? Evidence from continental cooling during the warmest part of the Cenozoic. In: Huber B., MacLeod K., Wing S., editors. Warm climates in earth history. Cambridge (UK): Oxford Univ Press. p. 197–237.
Wing S.L., Bown T.M., Obradovich J.D. 1991. Early Eocene biotic and climatic change in interior western North America. Geology. 19:1189–1192.

Wing S.L., Hickey L.J. 1984. The *Platycarya* perplex and the evolution of the Juglandaceae. Am. J. Bot. 71:388–411.

Woodburne M.O. 2004. Late Cretaceous and Cenozoic mammals of North America: biostratigraphy and geochronology. New York: Columbia University Press.

Yang Z., Rannala B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Mol. Biol. Evol. 23:212–226.

Zander R.H. 2004. Minimal values for reliability of bootstrap and jackknife proportions, decay index, and Bayesian posterior probability. Phyloinformatics. 2:1–13.