EVOLUTIONARY HISTORY OF SIMAROUBACEAE (SAPINDALES): SYSTEMATICS, BIOGEOGRAPHY AND DIVERSIFICATION

By

JOSHUA WILLIAM CLAYTON

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2008

To WilsonPhillips-X, for all the good times

ACKNOWLEDGMENTS

I would like to thank my committee, Doug Soltis, Pam Soltis, Walter Judd, Steve Manchester, and Gustav Paulay. I am very grateful to everyone who provided material for this study: Christopher Quinn at the Sydney Botanic Gardens, Frieda Billiet at the National Botanic Garden of Belgium, Libby MacMillan at the Royal Botanic Gardens, Edinburgh, Andrew Ford at CSIRO, Bruce Wannan of EPA, Queensland, Richard Abbott at FLMNH, Wang Hong at the Shishuangbanna Tropical Botanical Garden, China, Serena Lee at Singapore Botanic Gardens, Kew DNA Bank, MOBOT DNA Bank, and curators of MO, E, AAU, WAG, NY and CAY. Thanks also to Rick Ree (Field Museum, Chicago, USA) and Stephen Smith (Yale University, USA) for help with Lagrange, Alexei Drummond for comments regarding BEAST, Aleksej Hvalj (Komorov Institute, St. Petersburg, Russia) for providing information on *Leitneria* fossils, and Wayt Thomas, Michael Moore, Scot Kelchner, Susanne Renner and three anonymous reviewers for their helpful comments. Funding was provided by the National Science Foundation (angiosperm AtoL EF-0431266 to DES, PSS et al.; DDIG DEB-0710202 to JWC, DES), Botanical Society of America, and American Society of Plant Taxonomists. Finally, I would like to thank all my friends in the Soltis Lab, Crysta, Andrew, and my family, for their support.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A	Mainland Asia and southeast Asia (in AAR)
AAR	Ancestral area reconstruction
AIC	Akaike Information Criterion
BS	Bootstrap support
D_1 or D_2	Dispersal scenario 1 or 2
DIVA	Dispersal-vicariance Analysis
E	Europe (in AAR)
F	Tropical Africa (in AAR)
HPD	Highest posterior density
LRT	Likelihood ratio test
N	North and Central America and Caribbean Islands (in AAR)
M	Madagascar (in AAR)
M1, M2 etc…	Biogeographic likelihood model 1, 2 etc…
Ma	Million years ago
MCMC	Markov chain Monte Carlo
MPR	Most parsimonious reconstruction
Myr	Million years
NALB	North Atlantic land bridge
NW	New World
OW	Old World
PP	Posterior probability
S	South America (in AAR)
SE Asia	Southeast Asia
sp.	Species (singular)

spp.	Species (plural)
U	Australia, New Guinea, Papua New Guinea, New Caledonia and the Pacific Islands (in AAR)
UCLN	Uncorrelated lognormal

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

EVOLUTIONARY HISTORY OF SIMAROUBACEAE (SAPINDALES): SYSTEMATICS, BIOGEOGRAPHY AND DIVERSIFICATION

By

Joshua William Clayton

August, 2008

Chair: Douglas E. Soltis
Major: Botany

Simaroubaceae (Sapindales) are a geographically widespread and ecologically diverse clade of pantropical and temperate trees and shrubs, consisting of 22 genera and ca. 100. Phylogenetic relationships within the family have been poorly understood, and detailed studies of the evolutionary history of pantropical clades are still relatively few. Some of the underlying causes of shifts in diversification rates are investigated in Simaroubaceae, both historical movements and morphological innovation. Molecular and morphological data are used to reconstruct the phylogeny of Simaroubaceae and clarify generic limits, employing Maximum Parsimony (MP) and Bayesian approaches. Both MP and Bayesian analyses of combined data from four gene regions produce well-supported phylogenies, and these data support the monophyly of most traditionally circumscribed genera. Morphological data also support most genera, but provide little backbone resolution compared to molecular data. Bayesian uncorrelated rates analyses and robust fossil calibrations are used in divergence date estimation, and a recently developed likelihood method for reconstructing ancestral areas (the Dispersal-Extinction-Cladogenesis (DEC) model) is employed. Simaroubaceae exhibit an early history of range expansion between major continental areas in the Northern Hemisphere, but reconstruction of ancestral areas for lineages diverging in the early Tertiary are sensitive to the parameters of the

model used. A North American origin is suggested for the family, with migration via Beringia by ancestral taxa. In contrast to traditional views, long-distance dispersal events are inferred to be common, particularly in the late Oligocene and later. Morphological data reconstruct clades of genera, but show a high degree of homoplasy. Character-state reconstructions reveal a number of putative morphological synapomorphies for major clades in Simaroubaceae, and a complex history of character-state transitions. Shifts in diversification rate were examined using a number of statistical techniques, both temporal and topological. Four nodes in the phylogeny represent a point of diversification rate increase, and all can be attributed to a biogeographical shift, namely the arrival of the ancestor of *Simaba*, *Simarouba* and *Pierreodendron* in the New World, the movement of *Castela* from North to South America, the expansion of the ancestor of *Brucea*, *Soulamea* and *Amaroria* into Africa, and the arrival of *Soulamea* on New Caledonia. The topological imbalance demonstrated at the root node of *Leitneria* may be the result of extinction, as the fossil history of *Leitneria* indicates this clade was more diverse in the past. No morphological innovations of putative evolutionary significance were associated with increases in diversification rate, but analytical techniques in this area need further development.

CHAPTER 1
INTRODUCTION

Opportunities to undertake in-depth studies of the evolutionary history of groups of organisms are ever-increasing, with advances in technology for collecting molecular data, and constant improvements and refinements to analytical techniques. The search for general patterns in biome evolution and biodiversity gradients is reliant on studies of individual clades. One such clade with potential for a detailed treatment is the pantropical angiosperm family Simaroubaceae. Simaroubaceae (in the Sapindales) are a geographically widespread and ecologically diverse clade of trees and shrubs, consisting of 22 genera and ca. 100 species. Previous phylogenetic analyses of the family were limited to a single gene and seven genera (Fernando et al., 1995), and relationships among the genera, until this study, remain poorly understood. The family has a history of taxonomic uncertainty, and has only recently been recircumscribed in a phylogenetic framework (Fernando and Quinn, 1995). Molecular data are explored using a variety of methods for phylogenetic reconstruction, and special attention will be given to model selection in partitioning the data. A molecular phylogeny has also provided the opportunity to revisit and consolidate data on the natural history of the group, and produce the first general synopsis of the family since its recircumscription. Several pantropical clades have been examined in terms of their extant and ancestral geographic distributions (e.g. Lavin et al., 2000, 2004; Renner et al., 2001; Davis et al., 2002; Richardson et al., 2004; Zerega et al., 2005), including other members of the Sapindales (Weeks et al., 2005; Muellner et al., 2006), but historical biogeography in the Simaroubaceae had not been addressed, until the investigation presented herein. The family's current and fossil geographic distribution may be the culmination of a number of historical dispersal events, both long-distance and overland migrations via putative land bridges, as in studies of other pantropical clades. With a robust phylogeny and divergence date estimates,

hypotheses of past movements can be looked at through ancestral area reconstruction. A number of angiosperm clades have also been investigated in terms of variation in species diversity, and possible correlates to this diversity (e.g. Eriksson and Bremer, 1991; Hodges and Arnold, 1995; Richardson et al., 2001a, 2001b; von Hagen and Kadereit, 2003; Becerra, 2005; Kay et al., 2005; Hughes and Eastwood, 2006; Good-Avila et al., 2006; Moore and Donoghue, 2007). Sufficient variation in diversification rates has been found in Simaroubaceae, and so a similar approach to some of these studies was undertaken. This study allowed the examination of factors both intrinsic to the organisms, that is, key morphological innovations, and the external environment, ecology and historical dispersal events, that might influence speciation rates, and thus covary with diversification rate shifts in the family.

In this study I apply the latest techniques in data analysis to tackle the following questions:

- What are the phylogenetic relationships within Simaroubaceae?
- Over what timescale has the family and its constituent subclades evolved?
- How have historical migration events influenced the extant and fossil geographic distributions of the family?
- What accounts for the variation in species diversity of clades within the family?

CHAPTER 2
OVERVIEW OF SIMAROUBACEAE (SAPINDALES)

Simaroubaceae A.P. de Candolle, Prodromus I. 733 (1824), nom. cons.

Ailanthaceae (Arnott) J. Agardh (1858). Castelaceae J. Agardh (1858). Holacanthaceae Jadin (1901). Leitneriaceae Benth. & Hook.f. (1880). Quassiaceae Bertolini (1827). Simabaceae Horaninow (1847). Soulameaceae Endlicher (1874).

Family Description

Trees and shrubs, occasionally with thorns. Pith conspicuous, with triterpenoid compounds of the quassinoid type throughout vegetative tissues. Leaves alternate, spirally arranged, exstipulate (stipules found in *Picrasma*), pinnately compound or unifoliolate (rarely trifoliolate); leaflets entire, coarsely toothed, serrate or basally lobed, sometimes with conspicuous pitted or flattened glands beneath or above, venation pinnate, brochidodromus or occasionally reticulate. Hairs mostly simple, unicellular or multicellular, sometimes glandular-capitate. Inflorescences terminal or axillary, determinate thyrses, sometimes appearing raceme-like, pseudo-umbellate, catkin-like or flowers clustered in leaf axils. Flowers bisexual, polygamous or unisexual, actinomorphic, bracteate (bracts large and surrounding flowers in *Leitneria*), pedicels bracteolate, occasionally jointed; sepals 4–5 (0 in *Leitneria*), fused below, calyx sometimes splitting unevenly, occasionally bearing glands; petals 4–5(–8) (0 in *Leitneria*), free; stamens 4–10(–18), free, filaments often with hairy appendage, anthers dorsifixed, basifixed or versatile, opening by 2 longitudinal slits, introrse (occasionally extrorse to latrorse); ovary superior, (1–)2–5 carpels, free or fused basally, occasionally fused axially and deeply lobed, axile placentation, one ovule per locule, anatropous; stylodia free to connate, occasionally absent, stigmas divergent, stellately spreading, or a single slightly lobed or capitate stigma; fruit 1–5 samaroid or drupaceous

mericarps, exocarp thin, fleshy, occasionally dry, nut-like, often carinate, endocarp reticulate; testa membranaceous, cotyledons planoconvex, endosperm mostly lacking.

A family of 22 genera and about 109 species, mainly tropical and subtropical but some temperate species.

Vegetative Morphology

The family is woody, composed of large trees up to 50 m, shrubs, subshrubs, and occasionally suffructescent plants with all the leaves basal in *Simaba*. The wood is pithy or fistulous (Cronquist, 1944*d*), making it lightweight, and the bark and twigs are often striated. The family is typified by a bitter taste to the bark and twigs, on account of quassinoid compounds in scattered secretory cells throughout the vegetative structures (Cronquist, 1981). Thorns are present in *Castela* and in *Holacantha*, where they occur at the tips of all branches (Cronquist, 1944*d*).

Leaves are predominantly once-pinnately compound, arranged spirally around cylindrical stems. Unifoliolate leaves have evolved multiple times (based on studies of character evolution, Clayton, unpublished data), and are characteristic of *Castela*, *Leitneria, Amaroria* and *Samadera*, and are found in six species of *Soulamea* (Jaffre and Fambart, 2002) and two species of *Simaba* (*S. monophylla* and *S. obovata*). The leaves of *Holacantha* are reduced to scales or absent entirely, except in the seedlings (Cronquist, 1944*d*). Leaflets are alternate, subopposite or opposite, but always opposite in *Quassia*, which has a distinctive winged and jointed rachis. Leaflet shape is diverse, but strongly asymmetrical leaf bases are common in compound leaves. Leaf margins are predominantly entire, but are serrate or coarsely toothed in temperate species of *Ailanthus*, *Picrasma* and *Brucea*. Stipules are found in *Picrasma* (Nooteboom, 1962*b*), in which they are triangular, ovate or orbicular, and early caducous.

Vegetative Anatomy

Wood anatomy is described in detail by Webber (1936) and Record and Hess (1943; New World genera). Growth rings are present but indistinct, and diffuse porous or ring-porous (*Ailanthus*, *Leitneria*). Wood is dominated by fiber-tracheids, except in *Holacantha* and *Castela*, in which wood fibers are libriform (Webber, 1936). Vessels have spiral thickenings in *Castela, Holacantha* and *Leitneria*, but these are rare or absent in the rest of the family. Normal wood parenchyma cells are sparse to moderately abundant, the cells often septate and crystalliferous (Webber, 1936; Record and Hess, 1943). Vertical secretory canals are common in stems (Spiekerkoetter, 1924), and in *Leitneria* resin ducts are described as present in the margin of the pith (Record and Hess, 1943). Nodes are tri- or multilacunar, and calcium oxalate crystals are present in parenchymatous tissues (Cronquist, 1981). Flattish or concave glands are common on leaf surfaces, typically towards the margin, and often associated with teeth if present. Multicellular secretory glands are found on the abaxial surface of the sepals of *Samadera* (Nair and Joseph, 1957). Primarily unicellular, but also multicellular and glandular hairs are common on the inflorescence axes and floral organs (Nair and Joseph, 1957; Nair and Joshi, 1958; Nair and Sukumaran, 1960; Nooteboom, 1962*b*).

Inflorescence Structure

Inflorescences can be axillary or terminal, and are determinate thyrses, with the dichasia often appearing fasciculate or reduced to a single flower, giving the appearance of a panicle (*sensu* Weberling, 1989). Thyrses vary between open and spreading (e.g. *Ailanthus, Eurycoma, Picrolemma*), and narrow, elongate and sparsely branched (e.g. *Brucea, Soulamea, Amaroria*). In *Simarouba* and *Picrolemma* the staminate thyrses are larger and have more flowers than the carpellate thyrses (Cronquist, 1944*b*). *Picrasma* has a short, broad, rounded thyrse with a long peduncle (often described as a cyme), and in *Samadera* the inflorescence axis is condensed to

form a pseudo-umbel (Nair and Joseph, 1957). *Quassia amara* has a distinctive long raceme-like thyrse, occasionally branched at the base, and in *Castela* the flowers are occasionally solitary or clustered in the leaf axils (Cronquist, 1944*d*), as in some *Samadera*. In *Leitneria* the inflorescence is an erect or occasionally pendulous catkin-like thyrse: in the staminate inflorescence the flowers cluster in cymules of three in the axils of large, spirally-arranged bracts; in the carpellate inflorescence the flowers are solitary in the bract axils (Abbe and Earle, 1940).

Flower Structure

Flowers in Simaroubaceae are small, actinomorphic, open and 4- or 5-merous (trimerous in *Soulamea*), with an intrastaminal nectary disc. Petals are usually red, pink, yellow, pale green or white. Unlike the majority of the family, *Quassia* has flowers with elongate, glabrous petals (sometimes with hairs at the base) that are coherent, forming a tube, and the stamens and style are exserted. *Leitneria* is unusual in having asepalous and apetalous flowers, although Abbe and Earle (1940) observed vestigial perianth structures. *Leitneria* also has a unicarpellate gynoecium, with vascular bundles suggesting reduction from a bicarpellate gynoecium (Abbe and Earle, 1940). The androecium in the family is most commonly obdiplostemonous, although it is reduced to haplostemony in *Picrasma, Brucea, Picrolemma* and *Eurycoma*. In the latter two genera the stamens alternate with staminodes in the staminate flowers. In *Pierreodendron* the outer whorl of stamens is doubled. Adaxial scale-like appendages on the filaments occur in eleven genera, and are taxonomically significant at the species level, varying in shape, length, pubescence and bifurcation. In unisexual flowers, vestigial staminodes and pistillodes are common. Filaments are inserted at the base of the nectary disc, which can vary between strongly lobed, cushion-like, tall and cylindrical, conical or inconspicuous. The disc usually enlarges in

fruit. The gynoecium of *Soulamea* is reduced to two or three fused carpels, and is a single carpel in *Amaroria*.

Embryology

Embryology for the family was reviewed by Mauritzon (1935). Detailed studies of embryo anatomy are available for *Ailanthus* (Narayana, 1957), *Samadera* (Nair and Joseph, 1957), *Brucea* (Nair and Sukumaran, 1960) and *Leitneria* (Pfeiffer, 1912), and the following characteristics should be considered typical for the family: the anther wall consists of an epidermis, a fibrous endothecium, two to three middle layers and a multinucleate secretory tapetum (binucleate in *Ailanthus excelsa*); microsporogenesis is simultaneous; pollen tetrads are tetrahedral and decussate, shed at the two-celled stage; ovules are anatrapous or hemi-anatropous, crassinucellate and bitegmic, the inner integument forming the zig-zag micropyle; the nucellus is multinucleate, and the nucellar epidermis divides to form a cap; the archesporium can be multicellular or unicellular (*Ailanthus*), only one archesporial cell developing further; megaspores are arranged linearly (a solitary T-shaped tetrad is reported for *Ailanthus integrifolia*); the chalazal megaspore develops into a *Polygonum*-type embryo sac; fertilization may be chalazogamous, mesogamous or porogamous (Wiger, 1935), but only porogamy is confirmed in *Samadera* and *Brucea*; endosperm development precedes embryo development, and is of the Nuclear type.

Pollen Morphology

Basak (1963, 1967) and Moncada and Machado (1987) used light microscopy to survey pollen morphology in *Quassia*, *Samadera*, *Simarouba*, *Simaba*, *Eurycoma, Soulamea*, *Ailanthus, Brucea, Castela* and *Picrasma*, and Zavada and Dilcher (1986) examined *Leitneria* with SEM and TEM. Pollen grains are 3-zonocolporate, typically 20–35 μm long by 13–30 μm wide, prolate in equatorial view (sometimes subspheroidal in *Castela* and spheroidal in *Samadera* and

Leitneria) and planaperturate, with distinctly lalongate endoapertures; however, *Quassia* pollen grains are suboblate and angulaperturate, with a square type of endoaperture. Exine is 2–3μm thick, and the surface pattern finely to coarsely reticulate, sometimes verrucate, in most genera. The exine is striate in *Soulamea* and *Brucea*, and striato-reticulate in *Quassia*.

Karyology

Simaroubaceae have a base chromosome number of 8–13 (Stevens, 2006). Bennett and Leitch (2005) record $2n = 64$ in *Ailanthus integrifolia*, which would suggest the plant is octoploid. Raven (1975) reports $x = 16$ for *Leitneria*, and *Castela coccinea* has $2n = 26$ (Bernardello et al., 1990).

Reproductive Biology

Simaroubaceae can be hermaphroditic, monoecious or dioecious. The extent of self-compatability is unknown; however, flowers of *Quassia amara* have been shown to self-fertilize (Roubik et al., 1985). Insect-pollination predominates in the family, the flowers typically being small, actinomorphic, open, fragrant and borne in thyrses, attracting generalist small insects such as small bees and moths (e.g. Aubréville, 1962; Hardesty, 2005). *Quassia amara* is hummingbird-pollinated, as suggested by the raceme-like inflorescences bearing deep pink or red tubular flowers. Roubik et al. (1985) observed the role of nectar robbers in reproductive fitness of *Q. amara*, revealing that flowers were visited by nectar robbing bees (*Trigona*) and hummingbirds, as well as the primary hummingbird pollinator. *Leitneria* shows strong morphological divergence towards wind-pollination in that the flowers lack a perianth and nectary disk, and are borne in catkin-like inflorescences that develop before the leaves emerge.

Fruits and Seeds

Fruits in the family are predominantly schizocarpous with drupaceous mericarps, and typically only 1–3 carpels reach maturity. The drupes have a thin pericarp, in which the exocarp

can be fleshy (e.g. *Hannoa*, *Quassia*, *Simaba*), woody and fibrous (*Samadera*) or thin and dry (*Eurycoma*, *Leitneria*, some *Brucea*). The fleshy fruits can be pale yellow to red to deep purple-black, with a bitter taste, globose, obovoid, ovoid or ellipsoid, and between 0.3 cm and 10 cm long. The drupes are often carinate or bicarinate and flattened, and in *Samadera indica* are strongly laterally dorso-ventrally compressed with a narrow unilateral thinner edge in the apical half. In *Ailanthus* each carpel develops into a samaroid mericarp, elliptic in shape and tapering at each end. Variation in samara morphology is discussed in some detail by Nooteboom (1962b) and Corbett and Manchester (2004). In *Soulamea* the carpels remain fused in fruit, forming a dry, narrowly- to broadly-winged, obcordate fruit. Fernando and Quinn (1992) discuss variation in pericarp anatomy in the family in detail. The exocarp varies in thickness and lignification, and in *Ailanthus* is lacking except for the epidermal layer. Fernando and Quinn (1992) describe the endocarp as consisting of "a broad homogeneous zone of irregularly arranged isodiametric sclereids" with a strongly lignified inner epidermis. *Castela* and *Picrasma* lack the typical lightly lignified mesocarp and parenchymatous outer mesocarp. *Nothospondias* has an unusual *Spondias*-type endocarp, similar to that found in the Anacardiaceae (Fernando and Quinn, 1992).

The embryo is straight or curved, and consists of two large planoconvex cotyledons and a short plumule. Most Simaroubaceae have little or no endosperm, except for *Brucea* (Nair and Sukumaran, 1960) and some *Soulamea* (Nooteboom, 1962*b*). Fatty oil and aleuron bodies are the most common seed storage products in the family, but starch is also reported from seeds of *Simaba* and *Perriera* (Linsbauer, 1926) and *Leitneria* (Pfeiffer, 1912), and reserve celluloses also occur (Czaja, 1978; Stevens, 2006). The seed coat is thin and hard, undistinguished or with scattered lignified cells (Stevens, 2006), and is described as membranaceous in some genera.

Dispersal

Fleshy drupaceous fruits of Simaroubaceae are dispersed by fruit-eating birds and mammals, often primates (e.g. Hardesty et al., 2005). The samaroid mericarps of *Ailanthus* disperse over small distances by wind. Fruits of *Samadera indica*, a species that frequents alluvial and swamp forest, and *Soulamea amara*, a littoral species, are dispersed by water (Nooteboom, 1962*b*), which may account for their broad geographical distributions. *Leitneria* is also suspected to be water dispersed, typically growing in freshwater and brackish swamps. In all cases buoyancy is provided by an air cavity between seed and endocarp.

Phytochemistry

Simaroubaceae are characterized by their quassinoid chemistry. Quassinioids are triterpenoid derivatives, biosynthetically related to the limonoids of Rutaceae and Meliaceae (Da Silva and Gottlieb, 1987), and are considered by some (Dreyer, 1983; Waterman, 1983) to be further steps down the oxidative pathway of limonoids. Quassinoid structural and chemical characteristics are summarized in Waterman and Grundon (1983) and Da Silva and Gottlieb (1987), who report 35 different structural types in *Picrasma* alone. Pentacyclic triterpenes are also common (Hegnauer, 1983). Alkaloids have been reported in nine Simaroubaceae genera (Mester, 1983), most commonly tryptophan derived, but also a quinolone alkaloid is reported in *Ailanthus*. Only a single simple coumarin has been detected in the family, in *Picrasma* and *Ailanthus* (Gray, 1983). Of the flavonoid groups, flavonol glycosides and glycoflavone are reported in *Ailanthus* (Harborne, 1983) and flavonols and flavones in *Leitneria* (Giannasi, 1986). Essential oils contained within secretory cells and resin canals contain a low proportion of volatile compounds compared to Rutaceae and Meliaceae, and are in smaller amounts (Hegnauer, 1983). Tannin content is low to considerable, and with relatively high levels of gallic and ellagic acid (Hegnauer, 1983), although *Leitneria* lacks ellagic acid (Giannisi, 1986).

Distribution and Habitats

Simaroubaceae have a primarily pantropical distribution; however, some species of *Brucea, Castela*, *Holacantha*, *Ailanthus* and *Picrasma* are subtropical, and *Ailanthus altissima*, *Picrasma quassioides* and *Leitneria floridana* grow in temperate climates. Generic diversity is split evenly among the New World, Africa, and Asia and Australasia; however, half of the species in the family occur in the New World. *Picrasma* is disjunct among Asia, SE Asia and Central and South America, *Brucea* is disjunct between Africa and SE Asia, and *Soulamea* has one species in the Seychelles, one widespread in Malesia and Polynesia, and the remainder endemic to New Caledonia. *Samadera* is primarily Australian and SE Asian, but *S. indica* occurs as far west as India and Madagascar. Several genera in the Simaroubaceae consist of one or two species with restricted geographic ranges, the majority of these genera being in Africa. *Simaba* (25 spp.) is the most species-rich genus and is restricted to Central and South America.

Simaroubaceae are found in moist lowland tropical forest (although *Brucea mollis* is recorded as a high as 1800m in the Philippines, and *Odyendea gabonensis* at 2500m in Gabon), dry deciduous forest, and open sandy or savannah-type vegetation. *Soulamea amara* is a littoral species, *Castela* and *Holacantha* are found in desert and dry scrub environments, and *Leitneria*, *Samadera indica* and occasionally *Pierreodendron* inhabit swamp forest. *Eurycoma* is classified as silicicolous, showing a preference for acidic, leached sandy soils (Nooteboom, 1962*b*).

Dating and biogeographic analyses (see Chapter 4) suggest the family originated in North America in the early Tertiary. However, ancient vicariant and dispersal patterns in the family are obscured by a multitude of more recent migration events, within and between the continents, post-Oligocene.

Fossil History

Fossils of the distinctive samaroid fruits of *Ailanthus* are found across the entire Northern Hemisphere, dating from the early Eocene up to the Pleistocene (Corbett and Manchester, 2004). Three extinct species have been recognised, with the earliest occurrence a samara of *A. confucii* from the Green River Formation, Wyoming. Leaf fossils are also known with reasonable certainty from the Oligocene of Germany and Miocene of China, exhibiting distinctive basal teeth with enlarged glands on the leaflets, characteristic of extant *A. altissima* (Corbett and Manchester, 2004). *Leitneria* has fossil fruits in western Siberia from the Oligocene and in Europe from the Miocene to the Pliocene (Dorofeev, 1994; Nikitin, 2006a), but no fossil record from North America. Fossils fruits in transverse section have identical endocarp anatomy to extant *Leitneria floridana* (Dorofeev, 1994). Less well understood are fossil fruits of *Chaneya*, an extinct genus from the Tertiary of North America, Europe and Eastern Asia (Wang and Manchester, 2000; Teodoridis and Kvacek, 2005). Teodoridis and Kvacek (2005) suggest an affiliation with the extant genus *Picrasma*, based on gynoecial morphology and persistent wing-like petals; however, the fossil has distinctive oil cells typical of Rutaceae. Fossil leaves formerly reported as *Leitneria* from the Eocene of Tennessee (Berry, 1916) were subsequently reassigned to Rubiaceae, based on stipule configuration, epidermal anatomy and leaf architecture (Roth and Dilcher, 1979). Fossil pollen of Simaroubaceae has been reported for *Ailanthus* (Song et al., 2004) and *Leitneria* (Machen, 1971), but given the lack of distinctive morphological characteristics in extant Simaroubaceae pollen (Basak, 1963, 1967; Moncada and Machado, 1987), these are considered unreliable.

Affinities

In the traditional circumscription, Simaroubaceae *s.l.* comprised six subfamilies (Engler, 1931). However, molecular work by Fernando et al. (1995; also see Gadek et al., 1996) showed

the family to be polyphyletic, with subfamilies originating in several places within eurosids I and II (*sensu* APG II, 2003). Members of subfamily Simarouboideae, however, form a well-supported monophyletic group (excluding *Harrisonia*) within Sapindales. *Leitneria*, a genus traditionally segregated into the monotypic family Leitneriaceae on account of its wind-pollinated flowers (Cronquist, 1981; Takhtajan, 1996), was also found to be part of the Simarouboideae clade. Hence, the subfamily was recircumscribed as Simaroubaceae *s.s.*, a clade of 20 genera and approximately 95 species (Fernando and Quinn 1995). *Nothospondias*, a monotypic genus sometimes placed in Anacardiaceae (Engler, 1905), is a member of the family (Van der Veken, 1960; Chapter 3). Also included is *Laumoniera*, a monotypic genus from Sumatra (Nooteboom, 1987) that was omitted from the family recircumscription of Fernando and Quinn (1995).

Simaroubaceae are well supported as a member of a Simaroubaceae+Rutaceae+Meliaceae clade in Sapindales (Gadek et al., 1996; Källersjö et al., 1998; Savolainen et al., 2000*b*; Soltis et al., 2000), but the family's sister group is still undetermined, with data supporting three alternative topologies – Rutaceae sister to Simaroubaceae (Gadek et al., 1996); Meliaceae sister to Simaroubaceae (Chase et al., 1999; Muellner et al., 2006); Rutaceae sister to Meliaceae (Fernando et al., 1995; Stevens, 2006). Traditional morphological and phytochemical classifications typically suggest an affiliation with Rutaceae and Meliaceae (e.g. Cronquist, 1981; Takhtajan, 1996).

Relationships within the Family

Engler's (1931) classification of Simaroubaceae *s.l.* divided subfamily Simarouboideae (Simaroubaceae *s.s.*) into three tribes: Simaroubeae, Picrasmeae and Soulameae. Tribes were delimited by presence or absence of filament appendages and the degree of fusion of carpels. Of Engler's tribes, molecular data (Chapter 3) show only Soulameae, composed of *Soulamea* and

Amaroria, to be monophyletic. Relationships based on DNA sequence data from the chloroplast genome (*rbcL*, *atpB*, *matK*) and nuclear genome (1kb of *phyC*) produced a well-resolved and well-supported phylogeny (excluding *Laumoniera* and *Iridosma*; Chapter 3), with all genera except *Soulamea* being monophyletic. For details of phylogenetic relationships see Chapter 3. Putative synapomorphies for genera are italicized in genus descriptions.

Uses and Economic Importance

A range of biological properties has been demonstrated by the quassinoids of Simaroubaceae, including antimalarial, antileukemic, antiviral, insecticidal and amoebicidal properties (Polonsky, 1983; Klocke et al., 1985), and correspondingly, many genera are used locally as medicinal plants. *Quassia amara* and *Picrasma quassioides* have been used to aid digestion, and treat chronic dyspepsia. Fruits of *Brucea javanica* were imported into Europe as a drug (Nooteboom, 1962*b*), and the plant is used locally in Malaysia to treat malaria and dysentry. *Eurycoma* is used to treat malaria, diabetes, hypertension and stomach ache, typically by boiling the roots for drinking. *Ailanthus* is known in traditional Chinese and Korean medicine as a treatment for digestive complaints, haemorrhoids and mastitis. Simaroubaceae are not commercially harvested for timber but are used locally in building in some areas of both the Old and New World. *Leitneria* (corkwood) is one of the lightest known woods, and has been used traditionally by fisherman for net floats. *Ailanthus* (Tree of Heaven), *Simarouba* (paradise tree) and *Quassia amara* are cultivated and planted as ornamentals.

Key to the New World Genera

1. Petals and sepals 0 or vestigial; flowers surrounded by large, hirsute bracts **5. *Leitneria***
1. Petals +, sepals +; bracts not large, not surrounding flowers 2
2. Stamens with appendaged filaments 3

2.	Filaments lacking appendage	5
3.	Leaf rachis distinctly winged and jointed	**12. *Quassia***
3.	Leaf rachis not winged, not jointed	4
4.	Flowers unisexual; stigmas as long as style or longer, stellately spreading; leaflets alternate	**21. *Simarouba***
4.	Flowers bisexual; stigmas capitate or lobed; leaflets typically opposite or subopposite	**22. *Simaba***
5.	Leaves unifoliolate or absent; plant often armed with thorns; stamens twice as many as petals	6
5.	Leaves pinnately compound; plant without thorns; stamens equal in number to petals	7
6.	Plant with leaves; petals 4–5; stamens 8–10	**2. *Castela***
6.	Plant leafless, or leaves reduced to scales; petals 6–8; stamens 12–16	**3. *Holacantha***
7.	Staminodes present in staminate flowers; inflorescence elongate, narrowing above; fruit ellipsoid, elongate, 20–30 mm long	**11. *Picrolemma***
7.	Staminodes absent or in carpellate flowers only; inflorescence short, broad and rounded; fruit globose, less than 15mm long	**1. *Picrasma***

Key to the Old World Genera

1.	Stamens with appendaged filaments	2
1.	Filaments lacking appendage	8
2.	Leaves unifoliolate; inflorescence a pseudo-umbel	**13. *Samadera***
2.	Leaves pinnately compound; inflorescence not umbellate	3
3.	Leaf rachis jointed and often narrowly winged	**12. *Quassia***

3. Leaf rachis not jointed, not winged 4

4. Stamens alternating with outer whorl of staminodes or staminodal scales in staminate flowers; induplicate–valvate aestivation; Indomalesia **14. *Eurycoma***

4. Staminodes absent in staminate flowers; contorted, imbricate, occasionally valvate aestivation; tropical Africa 5

5. Petals 7–8, valvate in bud; stamens 12–13 **19. *Iridosma***

5. Petals 4–5, imbricate or contorted in bud; stamens 8–10(–15) 6

6. Stamens (10–)15(–18); leaves with 11–31 leaflets, up to 1 m long; leaflets apex with hard, pointed gland **20. *Pierreodendron***

6. Stamens 8–10; leaves with 3–15 leaflets, less than 60 cm long; leaflets without hard pointed gland at apex 7

7. Calyx in bud irregularly rupturing into 2–3 lobes; 5 petals; 10 stamens; 5 carpels; fruits 15–35 mm long **17. *Hannoa***

7. Calyx fused with 4(–5) very short obtuse lobes; 4(–5) petals; 8(–10) stamens; 4 carpels; fruits 50–70 mm long **18. *Odyendea***

8. Gynoecium 1 or 2(–3) fused carpels 9

8. Gynoecium (2)3–5 carpels (if 2 then carpels free) 10

9. Gynoecium a single carpel; fruit ovoid, not winged; flowers 4- or 5-merous **7. *Amaroria***

9. Gynoecium two or three carpels; fruit obcordate, winged; flowers predominantly 3-merous **6. *Soulamea***

10. Fruit samaroid **4. *Ailanthus***

10. Fruit drupaceous, fleshy or dry and nut-like 11

11. Stamens equal in number to petals 12

11. Stamens twice as many as petals 14

12. Inflorescence short, broad and rounded; sepals and petals persistent in fruit, accrescent; fruit globose **1. *Picrasma***

12. Inflorescence mostly unbranched, elongate; petals caducous in fruit; fruit ovoid or ellipsoid, or nut-like with 2 ribs when mature 13

13. Leaves imparipinnate; stigmas free, recurving; fruit 7–18 mm long **8. *Brucea***

13. Leaves paripinnate; stigmas connate, discoid; fruit 45–60 mm long **9. *Laumoniera***

14. Carpels 2; inflorescence axillary; Madagascar **16. *Perriera***

14. Carpels 4–5; inflorescence typically terminal; African tropics 15

15. Leaves with 19–43 leaflets; flowers 4-merous; fruits up to 45 mm in length; tropical west Africa **10. *Nothospondias***

15. Leaves with 13–25 leaflets; flowers 5-merous; fruits about 100 mm in length; Cote D'Ivoire endemic **15. *Gymnostemon***

Descriptions of Genera

1. *Picrasma* Blume

Picrasma Blume, Bijdr. Fl. Ned. Ind. 247 (1825); Cronquist, Brittonia 5:128–147 (1944).

Small trees, sometimes to 20 m, or shrubs; monoecious or dioecious. Leaves imparipinnate, *stipules present*, early caducous; leaflets opposite to subopposite, petiolulate, entire or serrate-crenate, glabrous or nearly so, without glands. Flowers in axillary, *short and broad, rounded* determinate thyrses with puberulent axes; sepals 4(–5), free or basally fused; petals 4(–5), *valvate*, mostly glabrous; stamens 4(–5), filaments lacking appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, sometimes conical, glabrous or hairy; carpels (2–)4(–5), free, stylodia fused above, sometimes free, stigmatic branches filiform,

recurved. Fruit 1–3(–5) drupaceous mericarps, *globose*, not carinate, 5–12 mm long, exocarp red to blue-black at maturity, pericarp fleshy. Eight species, two in Asia and SE Asia, six in Mexico to Argentina, and Caribbean islands.

2. *Castela* Turpin

Castela Turpin, Ann. Mus. Natl. Hist. Nat. 7:78 (1806); Cronquist, J. Arnold Arbor. 25:122–128 (1944).

Shrubs, erect or trailing, or small trees to 5 m; dioecious; armed with (occasionally branching) thorns. Leaves unifoliolate, petiolate, entire, glabrous to tomentose-pubescent, without glands. Flowers solitary, clustered in leaf axils, or in axillary, sparsely-flowered determinate thyrses, with typically a single, sparsely to densely hairy axis; sepals 4(–5), basally fused; petals 4(–5), imbricate, glabrous to occasionally pubescent; stamens 8 or 10, filaments lacking appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, ring-like, glabrous; carpels 4(–5), weakly united or free, stylodia fused at base, stigmatic branches linear, divergent or recurved, occasionally circinately rolled. Fruit 1–2(–4) drupaceous mericarps, lenticular, bicarinate, 6–12 mm long, exocarp red at maturity, pericarp fleshy. Twelve species from southern United States to Argentina, the Caribbean islands and the Galápagos.

3. *Holacantha* A.Gray

Holacantha A.Gray, Pl. Nov. Thurb. 310 (1854); Cronquist, Brittonia 5:128–147 (1944).

Depressed, ascending or erect shrubs or small trees to 5 m; dioecious; armed with thorns at branch tips. Essentially leafless or leaves scale-like. Flowers in axillary, short, densely-flowered determinate thyrses with one or two strongly hirsute axes, or appearing fasciculate in leaf axils; sepals 5–8, basally fused; petals 6–8, imbricate, strigose on abaxial surface; *stamens 12–16*, filaments lacking appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc narrow and ring-like, densely hairy to glabrous; *carpels 6–8*, weakly united, style short and

broad, stigmatic branches stellately spreading. Fruit 1–4 drupaceous mericarps, ovoid and slightly compressed, sometimes carinate on abaxial side, 5–9 mm long, exocarp red or greenish at maturity, pericarp fleshy. Two species from southern California, southern and western Arizona to Mexico.

4. *Ailanthus* Desf.

Ailanthus Desf., Mém. Acad. Roy. Sci. (Paris) I, 8:265 (1786); Nooteboom, Fl. Males., Ser. 1, Spermat. 6:193–226 (1962*b*).

Large trees to 60 m; dioecious or monoecious. Leaves imparipinnate or paripinnate; leaflets opposite, subopposite or alternate below, petiolulate, entire to coarsely toothed, glabrous to densely pubescent, with sometimes large abaxial glands, occasionally domatia present as hair tufts at leaf base. Flowers in axillary or terminal determinate thyrses with glabrous to sparsely hairy axes; sepals 5(–6), fused basally or calyx cupular with very short lobes; petals 5(–6), *induplicate-valvate*, glabrous to pubescent; stamens 10, filaments lacking appendage, anthers ± ventrifixed, staminodes absent in staminate flowers; disc fleshy, glabrous; carpels 2–5, stylodia free to fused, stigmatic branches peltate, stellately spreading, sometimes recurved. Fruit 1–5 *samaroid mericarps with elongate, membranous wings tapering towards the ends*, 25–220 mm long, exocarp brown at maturity, pericarp dry. Five species from Turkestan, India, China, SE Asia, and northern Australia.

5. *Leitneria* Chapm.

Leitneria Chapm., Fl. South. U.S., 428 (1860).

Small tree to 6 m; typically dioecious. Leaves *unifoliolate*, petiolate, entire, villous, without glands. Flowers in axillary, elongate, *highly reduced thyrses appearing catkin-like*, with cymules of 1–3 flowers in staminate inflorescence, *flowers solitary in carpellate inflorescence, surrounded by densely hirsute bracts* and arranged on an single glabrous axis; *perianth 0 in*

staminate flowers, vestigial in carpellate flowers; stamens (1–)4 per flower in bract axil, filaments lacking appendage, anthers basifixed to dorsifixed, staminodes absent in staminate flowers; *disc absent or rudimentary*; *carpel 1*, stigmatic branch distally expanded, recurved. Fruit a drupe, narrowly ellipsoid, conspicuously flattened, bicarinate, 12–30 mm long, exocarp brown at maturity, pericarp dry to occasionally fleshy. One species, *L. floridana*, in SE United States.

6. *Soulamea* Lam.

Soulamea Lam., Encycl. 1:449 (1783); Jaffré & Fambart, Adansonia 24:159–168 (2002).

Shrubs or small trees to 5(–15) m; dioecious or bisexual (S. amara). Leaves unifoliolate or imparipinnate; leaflets opposite, petiolulate, leaves petiolate, entire and often revolute, densely pubescent or glabrous on adaxial surface, sometimes with glands. Flowers in axillary, elongate determinate thyrses, typically with a single, often ferruginous-tomentose major axis; sepals 3(–5), basally fused; petals 3(–5), glabrous to pubescent towards the base; stamens 6(–10), filaments lacking appendage, anthers basifixed to dorsifixed, staminodes absent in staminate flowers; disc fleshy, glabrous; carpels 2(–3), connate, stylodia free, flattened, horizontally appressed to carpel, stigma fleshy, rarely reniform. Fruit samaroid, 2-celled, obcordate, flattened, with a distinct wing, 10–20 mm long, exocarp brown at maturity, pericarp dry. Thirteen species, one widespread in SE Asia and Polynesia (*S. amara*), one endemic to the Seychelles (*S. terminalioides*), and eleven species endemic to New Caledonia.

7. *Amaroria* A.Gray

Amaroria A.Gray, Bot. U. St. Expl. Exped. 1:356. I. 40 (1854); Smith, Fl. Vit. Nova Vol. 3:479–487 (1985).

Small tree to 15(–20) m; dioecious. Leaves unifoliolate, petiolate, entire, glands unknown. Flowers in axillary. elongate determinate thyrses with a single major axis; sepals 4–5, basally

fused; petals 4–5, glabrous or sometimes short strigillose along adaxial midline; stamens 8 or 10, filaments lacking appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, globose; *carpel 1*, stigma sessile. Fruit a drupe, ovoid to subglobose, slightly flattened, sometimes inconspicuously carinate, 17–30 mm long, exocarp greenish yellow, becoming white at maturity, pericarp fleshy. One species, *A. soulameoides*, endemic to Fiji.

8. *Brucea* J.F.Mill

Brucea J.F.Mill, Icon. Anim. Pl. t. 25 (1780).

Shrubs or small trees to 12 m; dioecious or polygamous. Leaves imparipinnate; leaflets opposite, petiolulate to subsessile, entire or crenate-serrate, ferruginous-pubescent to glabrous, *with dotted glands associated with peripheral secondary venation underneath*. Flowers in axillary, elongate determinate thyrses, typically with a single glabrous to densely pubescent major axis; sepals (3)4(–5), basally fused; petals (3–)4(–5), imbricate, glabrous to densely pubescent; stamens (3)4(–5), protruding between disc lobes, filaments lacking appendage, anthers basifixed to dorsifixed, staminodes absent in staminate flowers; disc fleshy, glabrous; carpels (3–)4(–5), free or united at the base, stylodia fused at base, stigmatic branches linear, recurved or bending inwards. Fruit 1–2(–4) drupaceous mericarps, *ovoid*, bicarinate, 4–18 mm long, exocarp red to black at maturity, pericarp dry to thinly fleshy. Six to seven species, tropical Africa to tropical and subtropical Asia and northern Australia.

9. *Laumoniera* Noot.

Laumoniera Noot., Blumea 32:383–384 (1987).

Small tree to 16 m; dioecious. Leaves *paripinnate*; leaflets petiolulate, entire, glands unknown. Flowers in axillary determinate thyrses, typically with a single pubescent axis; sepals 4, basally fused; petals 4, sparsely pubescent; stamens 4, filaments lacking appendage, staminodes absent in staminate flowers; disc fleshy, slightly hairy; carpels 4, free, stigmas sessile, connate, *discoid,*

covering top of ovaries. Fruit 1–4 drupaceous mericarps, ellipsoid, 45–60 mm long, exocarp yellow at maturity, pericarp fleshy. One species, *L. bruceadelpha*, Indonesia.

10. *Nothospondias* Engl.

Nothospondias Engl., Bot. Jahrb. Syst. 36:216 (1905); Van der Veken, Bull. Jardin Bot. État Bruxelles 30:105–109 (1960).

Tree to 25 m; dioecious. Leaves imparipinnate; leaflets opposite to alternate, petiolulate, entire, glabrous, without glands. Flowers in axillary or terminal determinate thyrses, with multiple densely pubescent axes; sepals 4, basally fused; petals 4, slightly imbricate, glabrous to puberulent; stamens 8, filaments lacking appendage, anthers basifixed, staminodes absent in staminate flowers; disc fleshy, glabrous; carpels 4, free, style fused below. Fruit 1–4 drupaceous mericarps, ovoid-ellipsoid, 20–45 mm long, exocarp yellow to orange at maturity, pericarp fleshy. One species, *N. staudtii*, in tropical west Africa. No apparent morphological synapomorphies, but a *Spondias*-type endocarp is unique to *Nothospondias*.

11. *Picrolemma* Hook.f.

Picrolemma Hook.f., Gen. Pl. (Bentham & Hooker f.) i. 312 (1862); Cronquist, Brittonia 5:128–147 (1944).

Small shrubs, up to 6 m; dioecious. Leaves imparipinnate; leaflets opposite to sometimes alternate below, petiolulate, entire, glabrous, punctate glands *associated with secondary venation underneath*. Flowers in terminal determinate thyrses, with multiple glabrous axes; sepals (4–)5, basally fused; petals (4–)5, imbricate, glabrous; stamens 5, filaments lacking appendage, anthers dorsifixed, *staminodes alternating with petals in staminate flowers*; disc fleshy, glabrous; carpels (4–)5, free, stylodia free but cohering, *stigmatic branches fleshy, club-like*. Fruit 1–2 drupaceous mericarps, *ellipsoid and slightly elongate, not carinate*, 20–30 mm long, exocarp brown to red at maturity, pericarp fleshy. Two species from Peru and Brazil.

12. *Quassia* L.

Quassia L., Sp. Pl. ed. 2, 1:553 (1762); Engler, Nat. Pflanzenfam. (Engler & Prantl) 19a:377–379 (1931).

Shrub or small tree to 8 m; bisexual. Leaves imparipinnate, *rachis and petiole conspicuously winged* in *Q. amara*, *narrowly winged* or wingless in *Q. africana*, *articulated*; leaflets *opposite, sessile*, entire, glabrous, with punctate glands towards leaf apex adaxially. Flowers in axillary or terminal determinate thyrses, appearing raceme-like in *Q. amara*, with puberulent axes; sepals 5, free, overlapping at base; petals 5, *contorted*, glabrous or basally pubescent inside, cohering into a tube in *Q. amara*; stamens 10, filaments with basal appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, narrowing towards base, glabrous; carpels 5, free, stylodia fused, stigma capitate or slightly lobed. Fruit 1–2 drupaceous mericarps, obovoid to ellipsoid, bicarinate, 10–25mm long, exocarp dark red at maturity, pericarp fleshy. Two species, 1 neotropical, 1 in tropical west Africa.

13. *Samadera* Gaertn.

Samadera Gaertn., Fruct. Sem. Pl. 2:352 (1791).

Small tree, occasionally up to 20 m; bisexual. Leaves *unifoliolate*, entire, glabrous, with scattered punctate glands. Flowers in axillary or terminal *pseudo-umbels*, axes glabrous to puberulent, or clustered in leaf axils; sepals (3–)4(–5), free or mostly fused with short lobes, occasionally with a concave gland; petals (3–)4(–5), imbricate or contorted, glabrous to pubescent abaxially; stamens 8 or 10, filaments with basal appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, conical or cylindrical, glabrous; carpels 4–5, stylodia fused, stigma capitate or slightly lobed. Fruit 1(–5) drupaceous mericarps, ovoid, ellipsoid or semicircular and flattened, slightly to strongly carinate, 5–50 mm long, exocarp orange to red or brown at

maturity, pericarp fleshy or dry and woody. Five to six species from Madagascar, Indo-China, SE Asia and Australia.

14. *Eurycoma* Jack

Eurycoma Jack, Malay. Misc. ii. 7:44 (1822); Nooteboom, Fl. Males., Ser. 1, Spermat. 6:193–226 (1962*b*).

Small trees to 10 m, or rarely shrubs; monoecious or dioecious. Leaves imparipinnate; leaflets opposite to subopposite, sessile or nearly so, sometimes appearing articulated, entire, glabrous, without glands. Flowers in axillary determinate thyrses, multiple axes with thick, *capitate-glandular hairs*; sepals 5(–6), basally fused, with *capitate-glandular hairs*; petals 5(–6), *induplicate-valvate*, pubescent, with *capitate-glandular hairs*; stamens 5(–6), filaments with very small appendage near base, anthers dorsifixed, *5(–6) staminodes alternating with stamens in staminate flowers*; disc inconspicuous; carpels 5(–6), free, stylodia connate or cohering, stigma lobed, peltate. Fruits 1–5 nut-like mericarps, ovoid, bicarinate, 10–20 mm long, exocarp brown at maturity, pericarp dry. Three species, tropical SE Asia, Sumatra, Malay peninsula, Borneo, S. Philippines.

15. *Gymnostemon* Aubrév. & Pellegr.

Gymnostemon Aubrév. & Pellegr., Bull. Soc. Bot. France, 84:181–184 (1937).

Large tree; bisexual or polygamous. Leaves imparipinnate; leaflets opposite to subopposite, subsessile, entire, glabrous, with punctate glands regularly spaced towards apex adaxially. Flowers in axillary or terminal determinate thyrses, with multiple densely short-hairy axes; sepals 5, fused with short lobes; petals 5, slightly imbricate, villous; stamens 10, filaments lacking appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, pubescent; carpels 5, free, stylodia fused, stigma simple or slightly lobed.. Fruit typically a single

drupaceous mericarp, ovoid, *up to 100 mm long*, pericarp fleshy, fibrous. One species, *G. zaizou*, endemic to Cote D'Ivoire.

16. *Perriera* Courchet

Perriera Courchet, Bull. Soc. Bot. France 52:284 (1905); Perrier de la Bathie, Fl. Madagasc. 105:1–7 (1950).

Tree to 30 m; typically bisexual. Leaves imparipinnate; leaflets opposite to subopposite, subsessile, entire, pubescent when young, becoming glabrous, with punctate glands regularly spaced towards apex adaxially. Flowers in axillary determinate thyrses, with multiple pubescent axes; sepals 5, basally fused; petals 5, *induplicate-valvate*, slightly villous; stamens 10, filaments lacking appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy; *carpels 2*, slightly united at base, stylodia fused, stigmatic branches divergent. Fruit typically a single drupaceous mericarp, ovoid, up to 50 mm long, exocarp pale yellow at maturity, pericarp fleshy. One or two species endemic to Madagascar.

17. *Hannoa* Planch.

Hannoa Planch., London J. Bot. 5:566 (1846).

Trees to 50 m or shrubs, sometimes suffrutescent; typically bisexual. Leaves imparipinnate; leaflets opposite to alternate, subsessile to petiolulate, entire, glabrous, with punctate glands on upper surface, more so towards margins. Flowers in terminal or occasionally axillary determinate thyrses, with multiple glabrous to sparsely pubescent axes; sepals 5, or *often calyx rupturing into 2–3 irregular lobes*; petals 5, imbricate, puberulent to densely tomentose; stamens 10, filaments with appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, sometimes with gynoecium sunken within, glabrous; carpels 5, free, stylodia fused, stigmatic branches short, spindly lobes. Fruit 1–3 drupaceous mericarps, ellipsoid or ovoid, slightly

bicarinate, 15–35mm long, exocarp red to purplish brown at maturity, pericarp fleshy. Five to seven species in tropical Africa.

18. *Odyendea* (Pierre) Engl.

Odyendea (Pierre) Engler, Nat. Pflanzenfam. (Engler & Prantl) III 4:215 (1896); Aubrév. & Pellegr. Fl. Gabon 3:33–52 (1962).

Tree to 30m; bisexual. Leaves imparipinnate; leaflets opposite to subopposite, petiolulate, entire, glabrous, with punctate glands on upper surface, more so towards margins. Flowers in terminal or axillary determinate thyrses, with multiple glabrous axes; *sepals 4*(–5), calyx cupular with short or absent lobes; *petals 4*(–5), imbricate, puberulent adaxially; stamens 8(–10), filaments with densely hairy appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, subcylindrical, with gynoecium slightly immersed within, glabrous; *carpels 4*, free or united at base, stylodia fused, stigmatic branches very short, divergent. Fruit a single drupaceous mericarp, obovoid to ellipsoid, *strongly carinate, up to 70 mm long*, exocarp red at maturity, pericarp fleshy. One species, *O. gabonensis*, endemic to Gabon and Cameroon.

19. *Iridosma* Aubrév. & Pellegr.

Iridosma Aubrév. & Pellegr., Fl. Gabon 3:47–49 (1962).

Tree; bisexual. Leaves imparipinnate; leaflets opposite to subopposite, subsessile, entire, glabrous, glands unknown. Flowers in determinate thyrses, with multiple pubescent axes; calyx cupular, irregularly undulating; *petals (7–)8, valvate*, villous; *stamens 12–13*, filaments with appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, pubescent; *carpels 4*, free, stylodia spirally twisted to form single column, stigma peltate, stellate. Fruit unknown. One species, *I. letestui*, endemic to Gabon and Cameroon.

20. *Pierreodendron* A.Chev.

Pierreodendron A.Chev., Vég. Utiles Afrique Trop. Franç. 9:257 (1917).

Tree to 15 m; bisexual. Leaves imparipinnate; leaflets subopposite to alternate, petiolulate, entire, glabrous to sparsely pubescent below, without glands. Flowers in axillary or terminal determinate thyrses, with one or two major axes; sepals 5, calyx cupular with short lobes; petals 5, imbricate or contorted; *stamens (10–)15(–18)*, filament appendage short with small free tip, anthers basifixed, staminodes absent in staminate flowers; disc fleshy, sometimes with gynoecium sunken within, glabrous; carpels 5, free, stylodia fused, stigmatic branches short, divergent, or stigma discoid. Fruits 1–5 drupaceous mericarps, oblong-ellipsoid, laterally compressed, 70–80 mm long, exocarp yellow at maturity, pericarp fleshy, fibrous. Two species from tropical Africa.

21. *Simarouba* Aubl.

Simarouba Aubl., Hist. Pl. Guiane 2:859 (1775); Cronquist, Bull. Torrey Bot. Club 71:226–234 (1944).

Shrubs and trees to 35 m; *dioecious*. Leaves paripinnate or imparipinnate; *leaflets alternate* to occasionally subopposite, petiolulate, entire, glabrous or densely tomentose below, with punctate glands scattered on upper surface, more so towards apex. Flowers in terminal determinate thyrses, with multiple glabrous axes; sepals 5, basally fused; petals 5, imbricate or contorted, glabrous; stamens 10, filaments with glabrous to pubescent appendage, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, short, glabrous to pubescent; carpels 5, free or weakly united, stylodia fused below, stigmatic branches stellately spreading, recurved. Fruit 1–3 drupaceous mericarps, ovoid or ellipsoid, slightly flattened, bicarinate, 10–25 mm long, exocarp orange-red to black at maturity, pericarp fleshy. Six species in Central and South America, the Caribbean islands and south Florida.

22. *Simaba* Aubl.

Simaba Aubl., Hist. Pl. Guiane. 1:409 (1775); Cronquist, Lloydia 7:81–92 (1944); Thomas, Brittonia 36:244–247 (1984).

Trees to 30m, shrubs, rarely suffructescent with all leaves basal; bisexual. Leaves paripinnate or imparipinnate, trifoliolate or rarely unifoliolate; leaflets usually opposite, petiolulate to sessile, entire, glabrous to occasionally pubescent, with punctate glands on upper surface, and occasionally with conspicuous apical gland. Flowers in terminal or axillary determinate thyrses, with multiple glabrous to densely pubescent axes, or occasionally reduced to axillary clusters; sepals (4–)5, basally fused; petals (4–)5, imbricate, puberulent to densely pubescent; stamens (8–)10, filaments with appendage, degree of fusion between filament and appendage variable, anthers dorsifixed, staminodes absent in staminate flowers; disc fleshy, cylindrical, glabrous to densely pubescent; carpels (4–)5, free or weakly united, stylodia fused, stigma capitate or slightly lobed. Fruit 1(–5) drupaceous mericarps, ellipsoid to obovoid, lenticular, slighty carinate or occasionally strongly winged, 10–80(–100) mm long, exocarp orange, red, brown, black or yellow at maturity, pericarp fleshy. 25 species in tropical South America with one species *S. cedron*, extending into Central America.

CHAPTER 3
MOLECULAR PHYLOGENY OF SIMAROUBACEAE BASED ON CHLOROPLAST AND NUCLEAR MARKERS[1]

Introduction

Simaroubaceae *s.s.* are a small yet morphologically diverse angiosperm family of tropical and temperate trees and shrubs in the Sapindales. In the traditional circumscription, Simaroubaceae *s.l.* comprised six subfamilies (Engler, 1931) and encompassed considerable diversity in secondary chemistry and macromorphology (Cronquist, 1981). However, molecular work by Fernando et al. (1995; also see Gadek et al., 1996) showed the family to be polyphyletic, with subfamilies originating in several places within eurosids I and II (*sensu* APG II, 2003). Members of subfamily Simarouboideae, represented in Fernando et al. (1995) by eight taxa, form a well-supported monophyletic group (excluding *Harrisonia*) within the Sapindales. *Leitneria*, a genus traditionally segregated into the monotypic family Leitneriaceae on account of its wind-pollinated flowers (Cronquist, 1981; Takhtajan, 1996), was also found to be part of the Simarouboideae clade. Hence, the subfamily was recircumscribed as Simaroubaceae *s.s.*, a clade of 20 genera and approximately 95 species (Fernando and Quinn, 1995). Two additional genera, *Laumoniera* and *Nothospondias*, may also need to be included in the family (Fernando and Quinn, 1992; Nooteboom, 1987). Tribal, generic and species relationships within Simaroubaceae *s.s.* (henceforth referred to as Simaroubaceae), however, remain unresolved.

Synapomorphies for the Simaroubaceae are triterpenoid compounds of the quassinoid type, five carpels united only by their styles and separating in fruit and one ovule per locule (Judd et al., 2002). Floral diversity ranges from wind-pollinated catkin-like inflorescences in *Leitneria* (Cronquist, 1981) to hummingbird-pollinated flowers in *Quassia amara* (Roubik et al., 1985) to

[1] Reproduced with permission from Clayton et al., 2007. International Journal of Plant Sciences 168:1325-1339.

insect pollination in the remaining genera. Habit ranges from large rainforest and temperate forest trees, to forest understory, coastal and desert shrubs.

Simaroubaceae comprise three tribes (in Simarouboideae *sensu* Engler, 1931): Simaroubeae, Picrasmeae and Soulameae (see Table 3-1), delimited by the presence or absence of filament appendages and the degree of fusion of the carpels. The three tribes were divided into seven subtribes based primarily on the nature of the androecium, particularly variation in the number of stamens and presence of staminodes (Engler, 1931). The only recognized subgeneric classification in the family is the division of the largest genus *Simaba* into three sections, *Tenuiflorae*, *Floribundae* and *Grandiflorae*, based primarily on flower size (Engler, 1874; Cronquist, 1944*c*). A recircumscription of *Quassia* (Nooteboom, 1962*a*) included all members of Simaroubeae except *Eurycoma*, but many authors have not accepted this expanded view of *Quassia* (e.g. Feuillett, 1983; Fernando and Quinn, 1992). *Laumoniera*, a monotypic genus from Sumatra, was described as a close relative of *Brucea* (Nooteboom, 1987); however, the status of *Laumoniera* as a new genus in the family has not been investigated further, and thus was omitted from the more recent family circumscription (Fernando and Quinn, 1995). Based on the morphological description (Nooteboom, 1987), we consider *Laumoniera* a distinct member of the family. *Nothospondias* is a monotypic genus originally described in the Anacardiaceae (Engler, 1905) and has a *Spondias*-type endocarp (Fernando and Quinn, 1992). However, it was referred to Simaroubaceae by Van der Veken (1960) based on morphological characters, and its phylogenetic placement is yet to be confirmed.

Simaroubaceae are well supported as a member of a Simaroubaceae+Rutaceae+Meliaceae clade in Sapindales (Gadek et al., 1996; Källersjö et al., 1998; Savolainen et al., 2000*b*; Soltis et al., 2000), but the family's sister group is still undetermined, with data supporting all three

alternative topologies – Rutaceae sister to Simaroubaceae (Gadek et al., 1996; Cronquist, 1981); Meliaceae sister to Simaroubaceae (Chase et al., 1999; Muellner et al., 2006); Rutaceae sister to Meliaceae (Stevens, 2001).

Simaroubaceae have a primarily pantropical distribution (Table 3-1); however, *Leitneria*, *Castela*, *Holacantha*, *Ailanthus*, *Picrasma* and *Brucea* have subtropical and temperate members. *Ailanthus* has an extensive fossil record dating from the early Eocene, across the entire Northern Hemisphere (Corbett and Manchester, 2004), *Leitneria* has fossil fruits dating from the Oligocene of western Siberia (Dorofeev, 1994; Nikitin, 2006a) and *Chaneya*, an extinct genus from the tertiary of North America, Europe and Eastern Asia (Wang and Manchester, 2000; Teodoridis and Kvacek, 2005), is suggested to have an affiliation with fruits of the extant genus *Picrasma*.

The goal of this study is to reconstruct phylogenetic relationships within the Simaroubaceae, using broad taxon sampling and multiple genetic markers. Sequences of the chloroplast genes *rbcL, atpB* and *matK* have repeatedly been of utility in family- and genus-level phylogeny reconstruction (e.g. Hoot et al., 1995; Chase et al., 1999; Muellner et al., 2003; Kathriarachchi et al., 2005; Wilson, 2005). More recently, the low-copy nuclear gene *phyC* has also proven useful in phylogenetic reconstruction at this level (Mathews et al., 1995; Davis et al. 2002; Kathriarachchi et al., 2005; Saarela et al., 2007), provided homologous copies are analyzed. Previous molecular phylogenetic studies of the family (Fernando et al., 1995; Gadek et al., 1996) were based only on *rbcL* sequence data, which failed to resolve relationships among genera, except to show that *Ailanthus* was sister to the rest of the family. Furthermore, sampling was restricted to just seven of the 22 genera, each represented by a single species. We therefore conducted a phylogenetic study employing nearly complete generic sampling and sequence data

from four genes, to elucidate relationships in the family. The resulting phylogeny is used to examine generic limits in the family, with particular reference to Nooteboom's (1962*a*) controversial broad recircumscription of *Quassia*. The position of *Nothospondias,* a genus that has been affiliated with both Simaroubaceae and Anacardiaceae, is also examined using a broader sample of Sapindales. In addition, the effects of five different data partitioning strategies in Bayesian analyses focused on Simaroubaceae are explored using Bayes factors, likelihood scores and resulting topologies and clade support.

Methods

Taxon Sampling

We conducted a broad analysis of Sapindales to ascertain the familial placement of the monotypic genus *Nothospondias.* Sequences of *rbcL* and *atpB* for 64 Sapindales and three Malvales outgroups were assembled from Genbank as well as from this study, including 36 Simaroubaceae and three Anacardiaceae taxa.

For the focused study of Simaroubaceae, 19 of the 20 genera *sensu* Fernando and Quinn (1995) were sampled, plus *Nothospondias*, for a total of 67 ingroup accessions encompassing 58 species (Appendix 1). Material was taken from herbarium specimens obtained from E, MO, NY, AAU and CAY, from silica-dried leaf material from the Royal Botanic Gardens, Edinburgh, Singapore Botanic Gardens, National Botanic Garden of Belgium, MOBOT DNA Bank and from wild plants in Florida, northern Australia and China. Extracted DNA was obtained from the Kew DNA Bank and previous studies of Simaroubaceae (Fernando et al., 1995). We were unable to obtain material of *Laumoniera* (omitted from Fernando and Quinn's (1995) recircumscription, but considered a member of the Simaroubaceae by Nooteboom (1987)) and *Iridosma* (a monotypic genus from Gabon, included in Fernando and Quinn (1995)). Three species each of Meliaceae and Rutaceae (Appendix 1) were used as outgroups based on Fernando et al. (1995)

and Gadek et al. (1996). In addition, a more distant outgroup from Sapindaceae (*Acer*) was included. Outgroup sequences were determined for this study or obtained from GenBank.

DNA Extraction, Amplification and Sequencing

Total genomic DNA was extracted from silica-dried and herbarium leaf material using CTAB extraction (Doyle and Doyle, 1990) or DNeasy Plant Mini Kits (Qiagen Inc., Valencia, CA), with the addition of Proteinase K for problematic accessions. Extractions that were difficult to amplify were cleaned using Promega Wizard Cleanup kits. Three plastid loci, *rbcL, atpB* and *matK* with ~500 additional bp of *trnK* intron on either side of the *matK* exon, and one nuclear locus, ~1000 bp of *phyC* starting ~550 bp downstream in exon 1, were amplified and sequenced. Plastid loci were amplified using the primers shown in Table 3-2. Amplification of DNA was conducted with a Biometra T3 Thermocycler (Biometra, Göttingen, Germany) or an Eppendorf Mastercycler (Brinkmann Inc., Westbury, NY) using PCR reactions in 25µl volumes, containing 2.5 units of *Taq* polymerase, 0.5 µM of each primer, 0.1 mM of each dNTP in an equimolar ratio, 10X buffer containing 1.5 mM $MgCl_2$, 10–30 ng of genomic DNA, 1M Betaine and dH_2O. PCR cycling conditions for chloroplast loci were as follows: 1) an initial heating step at 95°C for 5 min, 2) 94°C for 1 min, 3) the initial annealing temperature was 58°C for one minute and 4) elongation at 72°C for 2.5 min. Steps 2–4 were repeated for 6 cycles, except for dropping the annealing temperature by 1°C in each of the six cycles until 52°C, and then the annealing temperature was maintained at 52°C and steps 2–4 were repeated for a total of 34 cycles. PCR products were cleaned using ExoSap or Promega Wizard Cleanup kits. Weak PCR products were combined and concentrated using the Promega Wizard Cleanup kit.

DNA sequencing was performed on a Beckman-Coulter CEQ 8000 Automated Sequencer with DTCS chemistry (Beckman Coulter, Fullerton, CA) or an ABI 3730xl DNA Analyzer

(Applied Biosystems, Foster City, CA) with the T7 primer and BigDye Terminator Cycle Sequencing chemistry, using amplification primers plus additional sequencing primers (Table 3-2). Sequencher™v4.2 (Gene Codes Corp., Ann Arbor, MI) was used to assemble complementary sequences, and sequences were deposited in GenBank (see Appendix 1).

Partial *phyC* sequence was amplified using primers designed from GenBank sequences (see Table 3-2). The reaction mixture was as for chloroplast loci, using the following PCR conditions: 1) an initial heating step at 95°C for 5 min, 2) 94°C for 1 min, 3) annealing at 53°C for 1 min and 4) elongation at 72°C for 2.5 min. Steps 2–4 were repeated for 5 cycles, except for dropping the annealing temperature by 1°C in each of the five cycles until 48°C, and then the annealing temperature was maintained at 48°C and steps 2–4 were repeated for a total of 35 cycles. PCR products were sequenced as above. Sequences obtained for three *Ailanthus* accessions and one *Soulamea* accession showed overlapping peaks (we were unable to sequence *phyC* for two *Ailanthus* species). These PCR products were therefore cloned using TOPO TA Cloning Kits (Invitrogen Corp., Carlsbad, CA, USA). PCR products of species from seven other genera were also cloned to confirm their status as single copies. Sixteen clones per sample were chosen (to ensure multiple *phyC* copies were sequenced if present) from colonies grown on agar plates containing kanamycin, and amplified and sequenced using M13F and M13R bacterial primers on an ABI 3730xl DNA Analyzer.

Alignment and Indel Coding

Nucleotide alignments for each locus were assembled using Clustal X v1.83 (Thompson et al., 1997). Alignment of coding regions in the analysis focused on Simaroubaceae was straightforward, with few manual adjustments; however, the *trnK* intron had areas of ambiguity caused by large insertions in individual species. A total of 332 ambiguous positions was

excluded from phylogenetic analyses, 273 bp of which were large insertions in up to 12 species. Difficulties were encountered in assessing homology of the majority of indels within the *trnK* intron and thus coding them, due to ambiguity in the alignment. Indel coding was therefore not used in further analyses; however, easily coded indel synapomorphies are noted in the Results. Alignment of *rbcL* and *atpB* in the broader Sapindales analysis was also straightforward, and no indels were present.

Maximum Parsimony Analyses and Data Congruence

Gene partitions for the focused study of Simaroubaceae were analyzed separately to identify areas of incongruence between genes and between plastid and nuclear data (Huelsenbeck et al., 1996; Johnson and Soltis, 1998; Barker and Lutzoni, 2002; Hipp et al., 2004). Maximum parsimony (MP) analyses were conducted using PAUP* 4.0b10 (Swofford, 2002). MP analyses for each partition used heuristic searches of 1,000 random addition replicates, with TBR branch swapping, MulTrees in effect, and saving all trees. Bootstrap support (Felsenstein, 1985) was estimated from 1000 bootstrap replicates using 10 random additions per replicate, with TBR branch swapping, MulTrees in effect, and saving all trees.

A parsimony analysis of *phyC* clones revealed two copy types in *Ailanthus*; all other species from eight genera (including *Soulamea*) that were cloned had a single copy. The placement of one *Ailanthus* copy was congruent with the plastid phylogenies, and therefore a single clone from this copy was selected at random for each *Ailanthus* species, to be included in further phylogenetic analyses. The second *Ailanthus* copy was found to be sister to a clade of *Leitneria+Soulamea+Amaroria+Brucea* (i.e., the equivalent of clade III in figures 1 and 2) in the phylogeny of clones. Its incongruent position suggests it is a paralog resulting from a duplication event in *phyC*, and is under different evolutionary constraints than the first copy. Thus it was removed from further analyses.

Congruence between all data partitions, and between combined plastid and nuclear partitions, was assessed with the ILD test (Farris et al., 1994), implemented in PAUP as the partition homogeneity test, using 100 replicates and search strategies as for MP analyses. However, ILD results must be interpreted with caution as the test is sensitive, particularly in data sets with rate heterogeneity between partitions (Darlu and Lecointre, 2002). Therefore, Wilcoxon sum of rank tests, which take into account relative levels of clade support when assessing congruence, were also implemented following Zerega et al. (2005). Comparisons were made between all pairs of individual gene partitions and between combined plastid and nuclear (*phyC*) data, for bootstrap constraint trees at 70%, 80% and 90% thresholds.

For the focused study of Simaroubaceae, combined analysis of all data was conducted as described above for the separate analyses, except with bootstrap support assessed using 10,000 bootstrap replicates. For the broader analysis of Sapindales, the combined *rbcL* and *atpB* data were analyzed as above, using 1000 bootstrap replicates to assess support.

Bayesian Analyses and Partitioning Strategies

A Bayesian analysis of the combined data set focused on Simaroubaceae was performed, with the data partitioned by gene (with the *trnK* intron, excluding the *matK* exon, a single separate partition) and by codon position for the coding regions, totalling 13 separate partitions. Models of nucleotide substitution for each partition were determined using Modeltest v3.6 (Posada and Crandall, 1998). The Akaike information criterion (AIC) was used to selected an appropriate model, based on the relative informational distance between the ranked models in the Modeltest output. In all cases there were no topological differences among the five highest ranked models, therefore the highest ranked model was chosen for each partition, as follows: *rbcL* codon position 1(p1) - GTR+I+G, p2 - F81+I, p3 - GTR+G; *atpB* p1 - TrN+G, p2 - TrN+I+G, p3 - K81uf+G; *matK* exon p1 - TVM+G, p2 - GTR+G, p3 - TVM+G; *phyC* p1 -

HKY+G, p2 - TVM+G, p3 - TVM+G; *trnK* intron – TIM+I+G. Analyses were implemented in MrBayes v3.1.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Two independent analyses each ran for 5 million generations using four Markov chains and with all other parameters at default values; trees were sampled every 1000th generation with a burn-in of 100,000 generations; model parameters were unlinked across partitions. Stationarity of the MCMC was determined by the average standard deviation of split frequencies between runs and by examination of the distribution of the posterior in Tracer v1.3 (Rambaut and Drummond, 2003). Majority rule consensus trees were created in PAUP using the resulting posterior distribution of trees.

For the broader analysis of the Sapindales, data were partitioned by gene and codon position, with the following substitution models determined using Modeltest: *rbcL* codon position 1(p1) - GTR+I+G, p2 - F81+I+G, p3 - GTR+G; *atpB* p1 - GTR+G, p2 - K81uf+I+G, p3 - TVM+G. Bayesian analyses were conducted following the methods described above, with a burn-in of 100,000 generations.

For the focused study of Simaroubaceae, concerns about model specification and over-parameterization of the data (Rannala, 2002; Lemmon and Moriarty, 2004; Nylander et al., 2004; Sullivan and Joyce, 2005; Kelchner and Thomas, 2007) were addressed by running five separate partitioning strategies: 1) no partitioning of the data (P_1); 2) one chloroplast partition and one nuclear partition (P_2); 3) partitioning by codon position only plus a separate intron partition (P_4); 4) partitioning by gene and intron only (P_5); 5) full partitioning by gene and codon position plus intron (P_{13}). Number of data partitions for each partitioning strategy (P) is indicated by a numerical subscript. Partitions were decided upon *a priori*, under the assumption that different genes and the three codon positions may evolve at different rates, both within and between

plastid and nuclear genomes. Competing partitioning strategies were assessed with Bayes factors, comparison of the distribution of likelihood scores and resulting posterior distributions of trees, following the methods described in Brandley et al. (2005). Harmonic means of likelihoods were estimated using the *sump* command in MrBayes. Brandley et al. (2005) concluded that this method was adequate when compared to manual calculation of harmonic means from the posterior distribution of likelihoods. Likelihood scores for their partitioning strategies, and the differences between them, are comparable in magnitude to likelihood scores in this study. Bayes factors were calculated as the ratio of the harmonic means of the two competing hypotheses (partitioning strategies). The traditional cutoff for the support of a competing hypothesis over the null hypothesis is 2ln Bayes factor $\geq$ 10 (Kass and Raftery, 1995), indicating strong support for the alternative hypothesis, so this cutoff was used here. Arithmetic means and 95% credibility intervals of likelihood scores were calculated and compared among partitioning strategies, and posterior probabilities of clades were assessed across the five topologies generated by the different partitioning strategies.

Results

Data Congruence

Initial observation of separate and combined MP tree topologies and Bayesian topologies for the focused study of Simaroubaceae indicated all incongruence was between poorly supported clades, due to short branches or conflicting signal, and no hard incongruence existed among partitions. ILD tests (Farris et al., 1994) showed significant incongruence among all four loci ($p = 0.01$), and between plastid and nuclear data ($p = 0.01$), but not among the three plastid loci ($p=0.12$). Results of the Wilcoxon sum of rank tests (Table 3-3) showed no significant incongruence between any pairs of data partitions, except when the *phyC* tree search was constrained by the 70% *atpB* bootstrap consensus. However, significant incongruence was

observed between combined plastid and *phyC* data with all constraints, except when the *phyC* tree search was constrained by the 90% plastid bootstrap consensus. Inspection of the plastid and *phyC* strict consensus trees suggests the major sources of conflict are the positions of *Nothospondias* and *Quassia* in the phylogeny. In the *phyC* tree, *Nothospondias* appears in a polytomy with *Picrolemma* and a clade of *Samadera*+*Quassia*+clade V, whereas in the plastid tree *Nothospondias* is found in the same position as in the combined analysis (see figure 1). *Quassia* is sister to clade V in the *phyC* tree, whereas in the plastid tree *Quassia* is sister to a clade of *Samadera* and clade V, as in the Bayesian topology.

Phylogenetic Analyses and Tree Topology

Results of MP analyses of individual partitions and combined data for the family are provided (Table 3-4). MP analysis of the combined data set yielded eight most-parsimonious trees (figure 1). Results of the fully partitioned Bayesian analysis yielded a tree topology that was nearly identical to the MP strict consensus (figure 2), with the exception of the position of the *Quassia* clade, which diverges before *Samadera* in the Bayesian analysis and after *Samadera* in the MP strict consensus. Neither position received high posterior probability (PP = 0.91) or bootstrap support (BS = 62%).

Most relationships within Simaroubaceae have high PP and BS values in the combined MP analysis and fully partitioned Bayesian analysis. An early-diverging clade of *Picrasma*+*Holacantha*+*Castela* (clade I, figure 1 and 2) has low PP (0.74) and BS support (78%). A clade of *Leitneria*+*Soulamea*+*Amaroria*+*Brucea* (clade III) has high PP (1.0) and BS support (100%). *Soulamea* is rendered paraphyletic by the inclusion of *Amaroria*, which is sister to *S. amara*+*S. terminalioides* (PP = 0.99; BS = 84%). The position of *Nothospondias staudtii* as diverging after clade III has low BS support (<50%) in the MP analysis, but high PP (0.98) in the fully partitioned Bayesian analysis. Relationships among the four major lineages of clade V

(figures 1 and 2) have low PP (< 0.50) and BS (< 50%) values. Within clade V, there is high PP (1.0) for a clade of *Hannoa*, *Gymnostemon, Perriera, Odyendea zimmermannii* and *Brucea tenuifolia*, but low BS support (58%). A clade of *Simaba*, *Simarouba* and *Pierreodendron* has high PP (1.0) but BS support of only 71%. The monophyly of *Simarouba* has high PP (1.0) and BS (100%) values, but low support values for *Pierreodendron* as sister (PP = 0.69; BS = 69%). The monophyly of *Simaba* has low a PP (0.75) and BS (53%) value, but consists of two subclades with high support values (PP = 1.0, BS = 99% and PP = 1.0, BS = 100%, respectively).

Indels within the *matK* coding region were all specific to closely related species: indels support the monophyly of *Picrasma*, *Samadera* (excluding *S. indica*), *Samadera bidwillii+S.*spB, and a clade of *Simaba cedron+S. ferruginia+S. insignis*, respectively. Unambiguous *trnK* intron indels support a clade of *Picrasma+Holacantha+Castela,* a clade of *Simaba cedron+S. ferruginia+S. insignis* and the monophyly of *Eurycoma* and *Picrasma*, respectively. For the other easily coded *trnK* indels, there was an obvious tendency towards homoplasy.

In both MP and Bayesian analyses of the broader Sapindales data set, a clade of Simaroubaceae has high BS (100%) and PP (1.0) values; there is high PP (0.98), but low BS support (65%) for Meliaceae as sister to Simaroubaceae. *Nothospondias* was found to be nested within the Simaroubaceae clade in the same position as in the focused Bayesian analysis (trees not shown).

Partitioning Strategies

Absolute differences in 2ln Bayes factors between partitioning strategies (Table 3-5) showed that the most complex model (P_{13}) was a much better explanation of the data than other partitioning strategies. This was also evident in the mean $-\ln L$ of P_{13}, which was over 400 likelihood units better than the next best partitioning strategy, P_5. Partitioning by gene (P_5) and

by codon position (P_4) showed very similar mean likelihood scores, but partitioning by gene was determined to be a significantly better model than codon partitioning (2ln Bayes factor = 14.60). No partitioning of the data resulted in a lower mean –ln*L* than the next most complex partitioning strategy (P_2) and was decisively rejected as the best method of data partitioning according to all Bayes factors.

Topologies generated from the different partitioning strategies were identical. Differences in posterior probabilities between competing topologies were only apparent in weakly supported nodes (see Table 3-5 for examples). A general trend of increasing clade support with more complex partitioning was only observed in the positions of *Nothospondias* (node B) and *Quassia* (node C). Support for the sister relationship of *Pierreodendron* and *Simarouba* (node D) showed an inverse relationship of partitioning complexity to clade support, as did support for *Simaba* (node E).

Discussion

Data Congruence and Partitioning Strategies

In the focused study of Simaroubaceae, soft incongruence between the four loci was relatively rare and was typically due to a paucity of informative characters in *rbcL* and *atpB*, both of which made it difficult to resolve finer-scale relationships at the species level and therefore yielded high numbers of shortest trees for these two genes (Table 3-4). This problem was particularly evident in clade V, with several branches having a single substitution or a length of zero. Sequences of *matK*, including a portion of the *trnK* intron, provided much better resolution than either *rbcL* or *atpB* due to the high number of parsimony-informative characters (Table 3-4). *PhyC* also proved highly informative, particularly in resolving relationships within clade V.

There were no instances of hard incongruence observed between the trees of each partition, so despite the results of the ILD test and Wilcoxon sum of rank tests for nuclear versus plastid data, a combined approach was favored, due to the synergistic effects of combining data (Olmstead and Sweere, 1994; Huelsenbeck et al., 1996; Sullivan, 1996; Soltis et al., 1998). This approach is particularly effective with the broad level of taxon sampling employed in this study (Lecointre et al., 1993; Graybeal, 1998; Hedtke et al., 2006) and the range of evolutionary rates across the data sets being combined (Olmstead and Sweere, 1994; Sullivan, 1996). In particular, combining data resolved and greatly improved support for several finer-scale groupings that were not observed or poorly supported in the trees from separate data partitions (e.g. the *Simaba+Simarouba+Pierreodendron* clade and the position of *Perriera+Gymnostemon* as sister to *Hannoa*). The incongruent position of *Quassia* between parsimony and Bayesian analyses is likely due to the long-branch to *Quassia* in the parsimony analysis (see Bergsten, 2005).

Analyses of different partitioning strategies tested the relative importance of model specification and concerns about correct model selection in a complex multigene data set. Over-parameterization of the data, by using a fully partitioned, complex substitution model, was not evident, as the most complex partitioning strategy proved to be the best choice for this data set, as in other studies (Nylander et al., 2004; Brandley et al., 2005). This result, however, should be viewed with caution, given that little is known about how tree topology and clade support are affected by the random error associated with increasing number and decreasing size of partitions (Rannala, 2002; Cummings et al., 2003; Lemmon and Moriarty, 2004; Nylander et al., 2004; Brandley et al., 2005; Lewis et al., 2005; Sullivan and Joyce, 2005; Kelchner and Thomas, 2007). The increases in clade support with decreasing model complexity seen in the position of *Pierreodendron* (node D) and support for the monophyly of *Simaba* (node E; Table 3-5)

exemplify the difficulties in assessing model performance, and the problems of nonidentifiability and random error in measures of clade support for weakly supported nodes (Rannala, 2002; Sullivan and Joyce, 2005). This is especially true for short internodes (Alfaro et al., 2003; Erixon et al., 2003; Brandley et al., 2005; Lewis et al., 2005) such as those observed at nodes D and E (figure 2), because simpler models can produce inflated estimates of clade support (Suzuki et al., 2002; Lemmon and Moriarty, 2004; Nylander et al., 2004; Brandley et al., 2005).

Nylander et al. (2004) noted the strong influence of correctly modelling within- as well as between-partition rate variation, a factor that was not explored in this study, other than assuming that the Akaike information criterion (Akaike, 1974) determined the substitution model best suited to each data partition (Posada and Buckley, 2004). Nylander et al. (2004) and Lemmon and Moriarty (2004) found that failure to model among-site rate heterogeneity correctly had the greatest influence on likelihood scores, and caused the largest decrease in likelihood values when rate heterogeneity was not accounted for. This problem is unlikely to have affected our results, as all partition models in this study (except for *rbcL* codon position 2) estimated a gamma shape parameter and therefore incorporated among-site rate heterogeneity. However, base frequencies and transition/transversion rates may have been more complex than necessary, and thus the within-partition models could have used excessive parameters (Nylander et al., 2004).

Despite the concerns reviewed above, topologies generated from the different partitioning strategies were identical, with just a few differences in support values for weakly supported clades, suggesting that the underlying phylogenetic signal is robust and resistant to model misspecification (Kelchner and Thomas, 2007). The superiority of parameterizing by gene over codon position (Table 3-5) may be due to the range of rates of evolution across the four genes used; for example, the *PHY* gene family is reported to evolve about 10 times faster than *rbcL*

(Mathews et al., 1995). This variation in evolutionary rates is therefore more accurately modelled by gene partitioning, rather than by combining first, second and third codon positions across genes that evolve differently.

The differences in posterior probabilities of weakly supported nodes among all five topologies are most likely due to a lack of informative characters, borne out by the short branches of these nodes (figure 2). Resolution of and support for these nodes will likely improve, not with a better-fitting model, but with the addition of more sequence data, and in some instances more taxa.

Systematics of the Simaroubaceae

Molecular data have clarified evolutionary relationships among the genera and many species of the Simaroubaceae. Our discussion of relationships will focus on results of the combined, fully partitioned Bayesian analysis of the focused study (figure 2). The inclusion of *Picrasma*, *Castela* and *Holacantha* has revealed that *Ailanthus* is not sister to the rest of the family, as previously suggested in earlier studies (Fernando et al., 1995; Gadek et al., 1996). Instead, *Ailanthus* diverges after the *Picrasma+Castela+Holacantha* clade. Engler's (1931) tribes Picrasmeae and Simaroubeae are not monophyletic; however, Soulameae (consisting of *Soulamea* and *Amaroria*) has high PP (1.0) and BS (100%) support.

Although the clade of *Picrasma+Holacantha+Castela* has low PP and BS values in the combined analysis, it does appear in the MP strict consensus of most partitions and the combined analysis. A single 6-bp gap in the 5' *trnK* intron also supports this grouping. Cronquist (1944*a*) considered leafless *Holacantha* a "specialized offshoot" of *Castela*, an observation supported here by the early divergence of this ditypic genus from *Castela*. Cronquist (1944*a*,*d*) also considered *Picrasma* and *Castela* to be closest relatives in the Simaroubaceae, based on leaflet morphology and having similar drupaceous fruits. Biogeographically, *Picrasma*, *Castela* and

Holacantha are primarily New World, with just *P. quassioides* and *P. javanica* found in mainland and southeast Asia.

Fernando et al. (1995) confirmed *Leitneria* as a member of the Simaroubaceae, but its relationship to other genera was unresolved. The present study clearly places *Leitneria* within clade III (figures 1 and 2) as sister to a clade of *Brucea*, *Soulamea* and the monotypic *Amaroria*. Morphological synapomorphies for clade III may be difficult to detect due to the dramatic floral divergence of *Leitneria* compared to other clade members: *Leitneria* is anemophilous whereas other clade members are insect-pollinated. *Leitneria* shares simple leaves with a long petiole with several species of *Soulamea,* and *Leitneria* has a single carpel, suspected to be reduced from a bicarpellate gynoecium (Trelease, 1895), the condition found in *Soulamea*. In contrast, *Brucea* has a four-carpellate gynoecium and pinnately compound leaves. However, the androecium of *Leitneria* is reduced to four stamens (Judd et al., 2002) as in *Brucea*. Hence, a reduction in numbers of reproductive parts characterizes clade III. Further investigation is needed to identify morphological synapomorphies for the clade, which may be in the form of cryptic characters. The placement of the monotypic *Amaroria* within a paraphyletic *Soulamea* is in agreement with Nooteboom's (1962*b*) suggestion that *Amaroria* should not be considered separate from *Soulamea*. The genera differ chiefly in carpel number (one in *Amaroria*; two in *Soulamea*). However, only *rbcL* and *atpB* sequence data were available for *Amaroria*, and only three of the 14 species of *Soulamea* were sampled here, so additional data are needed to confirm this placement of *Amaroria*.

Nothospondias is found to belong within the Simaroubaceae based on a broad analysis of Sapindales. In the analyses focused on the Simaroubaceae, *Nothospondias* diverges after clade III in the Bayesian analysis, but is in an unresolved position in the parsimony analysis. The

similarity of the endocarp anatomy in fruits of *Nothospondias staudtii* to the distinctive *Spondias*-type endocarp found in the Anacardiaceae (Fernando and Quinn, 1992) may therefore be considered similarity due to homoplasy.

Quassia, the first genus name published for the Simaroubaceae, by Linnaeus in 1762 (Cronquist, 1944*d*), has been ascribed to various members of the tribe Simaroubeae, notably as a synonym of *Samadera*, *Simaba*, *Hannoa* and *Odyendea* (Pierre, 1896; Nooteboom, 1962*a*, *b*; Hewson, 1985; Mabberley, 1997). Nooteboom (1962*a*) considered *Quassia*, *Samadera*, *Simarouba*, *Simaba*, *Hannoa*, *Odyendea* and *Pierreodendron* to constitute a broadly defined *Quassia s.l.*, prompted by examination of a new species, *Q. borneensis*, which had traits in common with both New World *Simaba* and African genera *Hannoa* and *Odyendea*. In his recircumscription of *Quassia s.l.*, Nooteboom (1962*a*) recognized four sections: sect. *Quassia,* sect. *Samadera,* sect. *Simarouba* and sect. *Simaba*.

Within this sectional classification of *Quassia s.l.* (Nooteboom, 1962*a*), sect. *Quassia* consisted only of *Q. amara*, a New World species with a winged, jointed leaf rachis, a raceme-like inflorescence, pedicels articulated in the middle and erect, glabrous petals (with hairs at the base). However, *Q. africana* was also observed to have very narrow wings on a jointed leaf rachis (Nooteboom, 1962*a*) and glabrous petals (with hairs at the base), although it typically has a mulit-branched inflorescence and spreading petals. Furthermore, the short pedicels of *Q. africana* can make it difficult to distinguish the position of articulation (Cronquist, 1944*d*). For these reasons Nooteboom (1962*a*) and Cronquist (1944*c*) referred *Q. africana* to sect. *Simaba* in *Quassia s.l.* and the genus *Simaba*, respectively. However, Engler (1931) considered both *Q. amara* and *Q. africana* to constitute the genus *Quassia s.s*. Our data show a well-supported clade of *Q. amara* and *Q. africana* (PP = 1.0; BS = 100%). Pubescence of the petals and wings on the

rachis, considered unimportant in generic delimitation by Cronquist (1944*c*) and Nooteboom (1962*a*), may be synapomorphies for a narrowly defined *Quassia*. *Quassia amara* has unique features including a raceme-like inflorescence, pedicel articulation and erect corolla forming a tube-like flower, which are likely newly acquired traits associated with hummingbird pollination (Roubik et al., 1985). Furthermore, Basak (1967) noted pollen grains of *Q. amara* to be an entirely different type to that of *Samadera*, *Simaba* and *Simarouba*.

Following Nooteboom (1962*a*), sect. *Samadera* consisted of just two species, *Q. indica* and *Q. harmandiana*, ranging from Madagascar to southeast Asia. The two species are diagnosed as having large simple leaves and pseudo-umbellate or racemose inflorescences. However, Nooteboom placed two other species, *Q. baileyana* and *Q. bidwillii*, in sect. *Simaba*, despite both having simple leaves and *Q. baileyana* having a "stalked pseudo-umbel" (*Q. bidwillii* is described as having "flowers in clusters in the axils of the leaves"). Nooteboom (1962*a*) also acknowledged that *Q. baileyana*, although placed in sect. *Simaba*, "forms the connection with sect. *Samadera* as regards the structure of inflorescence and flowers". *Quassia baileyana* and *Q. bidwillii* occur in Australia, suggesting a closer association to sect. *Samadera* (southeast Asia and Madagascar) than to the other New World members of sect. *Simaba*. Molecular data revealed that *Q. indica, Q. baileyana* and *Q. bidwillii* (*sensu* Nooteboom, 1962*a*) form a clade with high PP (1.0) and BS (100%) values. This supports their recognition as the separate genus, *Samadera* (Bennett, 1872; Engler, 1931; Backer and Van den Brink, 1965), along with two other Australian species, *Samadera* sp.B and *Samadera* sp.C, that are yet to be formally described. In the circumscription of the genus *Samadera sensu* Bennett (1872), Engler (1931) and Backer and Van den Brink (1965), large simple leaves and umbellate/pseudo-umbellate inflorescences may be synapomorphies. Although *Q. harmandiana* was not sampled, based on its morphological

similarity to *Q. indica* (Nooteboom, 1962*a*), it is expected to be part of the *Samadera* clade. *Quassia borneensis* (for which molecular data were unobtainable), although found in Malaysia, is described as having similarities to the African members of sect. *Simaba*, notably *Q. gabonensis* (*sensu* Nooteboom 1962*a*). However, a single specimen of *Q. borneensis* observed for this study has simple leaves, so the position of this species remains unclear.

In clade V, no resolution was found among the four subclades. This lack of resolution can be attributed to very short branches, consisting of just one or two nucleotide substitutions supporting the basal branching pattern (figure 2), despite the use of 6 kb of sequence data. One possible explanation for these short branches is that clade V underwent a rapid radiation. More sequence data or more rapidly evolving genes are clearly needed to resolve the branching pattern among the subclades of clade V.

The resolution of four subclades in clade V refutes Nooteboom's (1962*a*) broad circumscription of *Quassia*, with three well-supported and geographically distinct groups, along with a single segregate species, *Odyendea gabonensis*. *Eurycoma*, identified by Nooteboom (1962*a*) as the only genus in this clade deserving of recognition outside *Quassia s.l.*, has high PP (1.0) and BS (100%) values and is diagnosed by its induplicate-valvate aestivation and outer whorl of staminodes (Nooteboom 1962*a*). *Hannoa*, *Gymnostemon* and *Perriera* also form a clade with a high PP (1.0) in Bayesian analyses, with the inclusion of *Odyendea zimmermannii* and *Brucea tenuifolia*. The clade does not receive a high BS value (58%) in the MP analysis and is subtended by a short internode (figure 2); therefore, Bayesian analyses may be over-estimating support for this clade (Suzuki et al., 2002; Alfaro et al., 2003; Erixon et al., 2003; Brandley et al., 2005; Lewis et al., 2005). However, the clade is composed only of species found in Africa and Madagascar. *Brucea tenuifolia* appears as sister to *O. zimmermannii*, and from the limited

herbarium material available for these two species, leaf morphology between the two is very similar, so a misidentification of *B. tenuifolia* is suspected. Regardless of this possible misidentification, *O. zimmermannii* may need to be reassigned to *Hannoa*. In Nooteboom's (1962*a*) recircumscription of *Quassia*, *H. undulata, H. chlorantha*, *H. klaineana* and *O. zimmermannii* were synonymous under *Q. undulata* in sect. *Simaba*, as the characters distinguishing the different species, such as number and shape of leaflets, length of lateral petiolules, and flower size, were regarded as highly variable and thus not useful. Based on results of our phylogenetic analyses of Simaroubaceae, Nooteboom's observations may be corroborated by the very short branches within *Hannoa* and low PP and BS values within the clade (Figs. 1 and 2). A re-examination of morphological characters in the *Hannoa* clade will be beneficial to clarify species limits.

The well-supported clade of *Perriera* (a genus of two species endemic to Madagascar) and *Gymnostemon* (a monotypic endemic of Côte d'Ivoire) was not surprising, as Aubréville and Pellegrin (1937) suggested the two genera were closely related. *Perriera* was traditionally placed in tribe Picrasmeae due to the lack of a filament appendage (Engler, 1931; Perrier de la Bathie, 1950). *Gymnostemon*, which was not included in Engler's tribal classification, also lacks a filament appendage (Aubréville and Pellegrin, 1937; Hutchinson and Dalziel, 1954). Given that these are the only two genera in clade V without a filament appendage, loss of the appendage is likely a synapomorphy for *Perriera* and *Gymnostemon*. *Perriera* and *Gymnostemon* were not considered by Nooteboom in his broad circumscription of *Quassia*. The placement of *Odyendea gabonensis* outside the African clade in the basal polytomy of clade V (figure 2) is surprising, given its morphological similarity to members of *Hannoa*. More sequence data and a re-examination of morphology may clarify the position of this species in clade V.

The largest subclade in clade V, which consists of *Simaba, Simarouba* and *Pierreodendron*, has a high PP (1.0) but relatively low BS support (71%). A close relationship of *Pierreodendron* to *Simarouba* and *Simaba* in particular has never previously been suggested based on morphology. The monophyly of *Simarouba* is supported by high PP (1.0) and BS (100%) values, and its status as a genus distinct from *Simaba* is also suggested. However, the monophyly of *Simaba* is not well supported (PP = 0.75; BS = <50%). *Simaba* is distinct from *Simarouba* based on its perfect flowers with capitate or lobed stigmas, and typically opposite leaflets, as compared to the unisexual flowers, long divergent stigmas and offset leaflets of *Simarouba* (Cronquist, 1944*b,c,d*). *Simaba* consists of two subclades, one of which corresponds to Engler's (1874) section *Tenuiflorae* (containing the *S. guianensis* complex (Thomas, 1985; Franceschinelli and Thomas, 2000)), which is characterized by small flowers with puberulent petals and is primarily found in the moist Amazon basin (Cronquist, 1944*c*). The second subclade consists of sections *Floribundae* and *Grandiflorae*; these are larger-flowered species with villous-tomentose petals, primarily found in the drier regions of southern and eastern Brazil and Paraguay (Cronquist, 1944*c*). Section *Floribundae* is paraphyletic with respect to *Grandiflorae*. However, only *S. cedron* was sampled from section *Grandiflorae*, so more taxa are needed to assess the monophyly of the two sections. The variable positions of accessions of *S. guianensis* and low PP and BS values within section *Tenuiflorae* reflect the difficulties of delimiting species and subspecies in this clade (Thomas, 1985, pers. comm.). More taxa and a more rapidly evolving gene region will help to clarify relationships in this group.

Overall the generic limits in the Simaroubaceae are supported by molecular data. *Amaroria* is nested within *Soulamea*, as suggested by Nooteboom (1962*b*), and *Nothospondias*, sometimes placed in Anacardiaceae, is supported as included in Simaroubaceae. Molecular data reveal

several clades, with high PP and BS values, that correspond to traditional generic limits in tribe Simaroubeae, and thus we consider a broad circumscription of *Quassia* (Nooteboom, 1962*a*) unnecessarily conservative, in agreement with previous authors (Porter, 1973; Cavalcante, 1983; Feuillett, 1983; Fernando and Quinn, 1992). *Quassia* should be limited to just two species, *Q. amara* and *Q. africana*, and simple-leaved Old World *Quassia* species should be recognized as *Samadera*. *Simaba, Simarouba, Pierreodendron, Hannoa, Gymnostemon, Perriera, Eurycoma* and *Odyendea gabonensis* are recognized as distinct entities based on our molecular data, in agreement with quassinoid chemistry (Da Silva and Gottlieb, 1987), pericarp anatomy (Fernando and Quinn, 1992) and diagnostic morphological characters (Engler, 1931; Cronquist, 1944*b,c,d*).

The relative importance that Engler (1931) placed on characters such as appendaged filaments and the nature of the androecium in delimiting tribes and subtribes is not borne out by the molecular phylogeny. Engler placed *Perriera* with genera of tribe Picrasmeae (e.g. *Brucea, Ailanthus, Picrasma*), because they all lack appendaged filaments. However, molecular data place *Perriera* within tribe Simaroubeae (e.g. *Simaba, Quassia, Hannoa, Eurycoma*) in the phylogeny. Although variation in the androecium characters was commonly used by Engler (1931) to delimit subtribes, the phylogeny reveals lability in the nature of the androecium, with typically diplostemonous flowers in the family, but haplostemonous flowers in *Picrasma*, *Brucea, Picrolemma* and *Eurycoma*, pleiostemonous flowers in *Pierreodendron*, and staminodes occurring in the staminate flowers of *Picrolemma* and *Eurycoma*. Thus, haplostemony and staminodes evolved multiple times in the family.

This study demonstrates the effectiveness of combining data from two different genomes and genes with a variety of rates of evolution (Olmstead and Sweere, 1994; Sullivan, 1996; Huelsenbeck et al., 1996). A total evidence approach, coupled with broad taxonomic sampling,

has produced a well-resolved and well-supported phylogeny at all levels from the species to the family. Comparisons among different strategies for partitioning the data in combined Bayesian analyses have revealed that the most complex model provides the best fit to the data using the Bayes factor criterion and examination of likelihood scores. Partitioning analysis has also highlighted the importance of exploring the effects of modelling on resulting posterior probability distributions and tree topologies (Cummings et al., 2003). Results should be interpreted cautiously, due to the lack of data concerning the effects of random error on tree topology and clade support (Lewis et al., 2005), and further work is needed to explore within-partition modelling (Lemmon and Moriarty, 2004; Nylander et al., 2004) and signs of nonidentifiability among model parameters (Rannala, 2002; Sullivan and Joyce, 2005). Despite this caution, the identical topologies across partitioning strategies suggest an underlying phylogenetic signal that is strong enough to allay some of the concerns about incorrect model choice and over-parameterization, particularly when the goal is to produce a robust topology, and place less importance on accurately estimating branch lengths (Lemmon and Moriarty, 2004; Kelchner and Thomas, 2007).

Table 3-1. List of the 22 genera of the Simaroubaceae grouped by Engler's (1931) tribal classification, with number of species per genus, number of species sampled in this study and geographic distribution. Taxa labeled unknown are those not included in Engler's circumscription of the subfamily Simarouboideae.

TRIBE Genus	Clade number (this study)	Number of species	Number of species sampled	Geographic Distribution
SIMAROUBEAE				
Eurycoma	V	3	2	Southeast Asia
Hannoa	V	4	3	Tropical Africa
Odyendea	V	2	2	Tropical west Africa
Pierreodendron	V	2	1	Tropical Africa
Quassia	-	2	2	New World, tropical Africa
Samadera	IV	7	5	Old World tropics
Simaba	V	25	8	Central and South America
Simarouba	V	6	4	Central and South America
PICRASMEAE				
Ailanthus	II	5	4	Asia to Australia
Brucea	III	7 – 8	6	Old World tropics
Castela	I	12	5	SW U.S., Mexico, Central and South America
Holacantha	I	2	1	SW U.S., Mexico
Perriera	V	1 – 2	1	Madagascar
Picrasma	I	8	5	New World, Asia, southeast Asia
Picrolemma	-	2	1	South America
SOULAMEAE				
Amaroria	III	1	1	Fiji
Soulamea	III	13	3	Seychelles, southeast Asia, New Caledonia and Pacific Islands
UNKNOWN TRIBE				
Gymnostemon	V	1	1	Côte d'Ivoire
Iridosma	V	1	0	Tropical west Africa
Laumoniera	-	1	0	Sumatra
Leitneria	III	1	1	SE U.S.
Nothospondias	-	1	1	Tropical west Africa
Total		~109[a]	57	

[a] Note total number of species is estimated to be higher than Fernando and Quinn (1995).

Table 3-2. Primers used for PCR amplification and sequencing.

Locus	Primer name	Amp/Seq	Sequence	Reference
rbcL	Z1	A+S	ATGTCACCACAAACAGAAACTAAAGCAAGT	Zurawski et al. 1984
	3’	A+S	CTCGGAGCTCCTTTTAGTAAAAGATTGGGCCGA	Zurawski et al. 1984
	346F	S	ATGTTTACTTCCATTGTGGGTAATGTATTT	Zurawski et al. 1984
	895R	S	ACCATGATTCTTCTGTCTATCAATAACTGC	Zurawski et al. 1984
atpB	S2	A+S	TATGAGAATCAATCCTACTACTTCT	Hoot et al. 1995
	S1494R	A+S	TCAGTACACAAAGATTTAAGGTCAT	Hoot et al. 1995
	S335	S	ACGTGCTTGGGGAGCCTGTTGATAA	Hoot et al. 1995
	S1186R	S	TGTCCTGAAGTTCTTTGTAACGTTG	Hoot et al. 1995
matK	trnK-3932F	A+S	CCACGACTGATCCTGAAAGG	This study
	trnK-2R*	A+S	ATCCCCGTGTCAACCAATAG	This study
	501F	A+S	GTTCAACCCCTTCGCTACTG	This study
	935R	A+S	TTTCCATTTAGTCATCAGAAGAGG	This study
	147F	S	TCATTGGAAAATGGGGGTTA	This study
	1388R	S	TTTACGAACCAAACTTTTAACACA	This study
phyC	PHYC-smbF1	A+S	GGCAYTGAARTCATAYAARCTTGC	This study
	PHYC-smbR1	A+S	CCRCCCCACTTGATCTCYTT	This study
	PHYC-304F	S	CGCAAGCTTCCAGATTTCTT	This study
	PHYC-699R	S	AGCATGTCACAAAGCACAGTTT	This study

Table 3-3. Wilcoxon sum of rank test results showing pairwise comparisons between data partitions for three bootstrap constraint topologies, ranging from most stringent (70%) to least stringent (90%). Comparisons with significant incongruence ($p < 0.05$) are indicated with an asterisk.

Partition	70% Bootstrap constraint topology				
	rbcL	*atpB*	*matK*	*phyC*	plastid
rbcL	-	*0.279*	*0.070*	*0.056*	-
atpB	*0.488*	-	*0.315*	*0.185*	-
matK	*0.352*	*0.383*	-	*0.072*	-
phyC	*0.415*	**0.013*	*0.204*	-	**0.005*
plastid	-	-	-	**0.014*	-

	80% Bootstrap constraint topology				
	rbcL	*atpB*	*matK*	*phyC*	plastid
rbcL	-	*0.446*	*0.094*	*0.056*	-
atpB	*0.421*	-	*0.342*	*0.258*	-
matK	*0.347*	*0.394*	-	*0.074*	-
phyC	*0.608*	*0.553*	*0.204*	-	**0.007*
plastid	-	-	-	**0.014*	-

	90% Bootstrap constraint topology				
	rbcL	*atpB*	*matK*	*phyC*	plastid
rbcL	-	*0.458*	*0.111*	*0.075*	-
atpB	*0.438*	-	*0.341*	*0.223*	-
matK	*0.347*	*0.329*	-	*0.128*	-
phyC	*0.546*	*0.686*	*0.396*	-	*0.284*
plastid	-	-	-	**0.031*	-

Table 3-4. Results of maximum parsimony analyses for individual data partitions and combined analyses of Simaroubaceae. [a] indicates MaxTrees limit was reached.

	Number of ingroup taxa	Aligned length	% missing data	Variable characters	Informative characters	Number of shortest trees	Length of best tree	Consistency index	Retention Index
rbcL	67	1394	1.12	260	153	100,000[a]	524	0.608	0.831
atpB	65	1438	4.92	276	160	524	446	0.722	0.854
matK + partial *trnK* intron	66	2184	3.54	874	526	8	1569	0.705	0.854
phyC	62	954	5.63	452	285	360	969	0.635	0.813
Plastid combined	67	5016	4.76	1410	839	8	2557	0.684	0.846
All Combined	67	5970	5.93	1859	1132	8	3544	0.667	0.835

Table 3-5. Summary of results from data partitioning analyses. From left: 2ln Bayes factors comparing the five different partitioning strategies; arithmetic mean –ln*L* and credibility intervals for each partition strategy; comparisons between posterior probability clade support for five exemplar nodes that are weakly supported in the phylogeny (Figure 2).

Partitioning Strategy	2ln Bayes Factors for alternative hypotheses				Likelihood scores			Posterior probabilities of nodes A – E (Figure 2)				
	P_{13}	P_5	P_4	P_2	Mean –ln*L*	Upper 95% CI	Lower 95% CI	*A*	*B*	*C*	*D*	*E*
P_1	1391.34	550.00	535.40	58.88	31031.49	31013.72	31051.20	0.74	0.91	0.69	0.94	0.87
P_2	1332.46	491.12	476.52	-	30999.67	30981.38	31020.84	0.83	0.92	0.78	0.94	0.88
P_4	855.94	14.60	-		30756.37	30736.61	30777.39	0.56	0.90	0.71	0.88	0.83
P_5	841.34	-			30754.10	30733.42	30775.83	0.80	0.93	0.93	0.76	0.81
P_{13}	-				30324.82	30301.60	30348.62	0.74	0.98	0.91	0.69	0.75

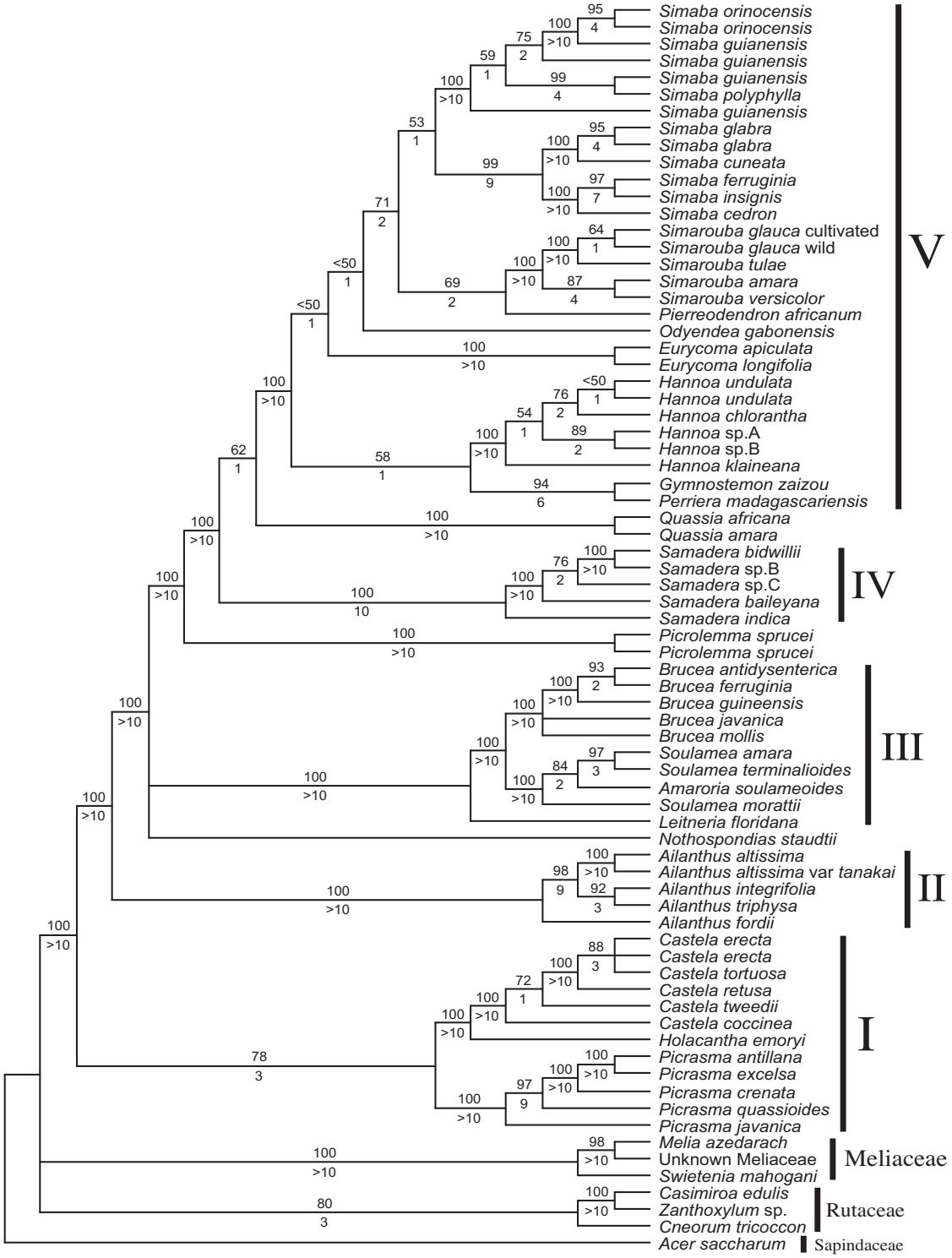


Figure 3-1. Strict consensus of eight most parsimonious trees recovered from a combined analysis of plastid genes *rbcL*, *atpB* and *matK* (including partial *trnK* intron) and 1kb of nuclear gene *phyC* for Simaroubaceae. Bootstrap support (BS) greater than 50% is shown above branches, decay indices are shown below. Major clades (I – V) of Simaroubaceae are indicated on the right hand side.

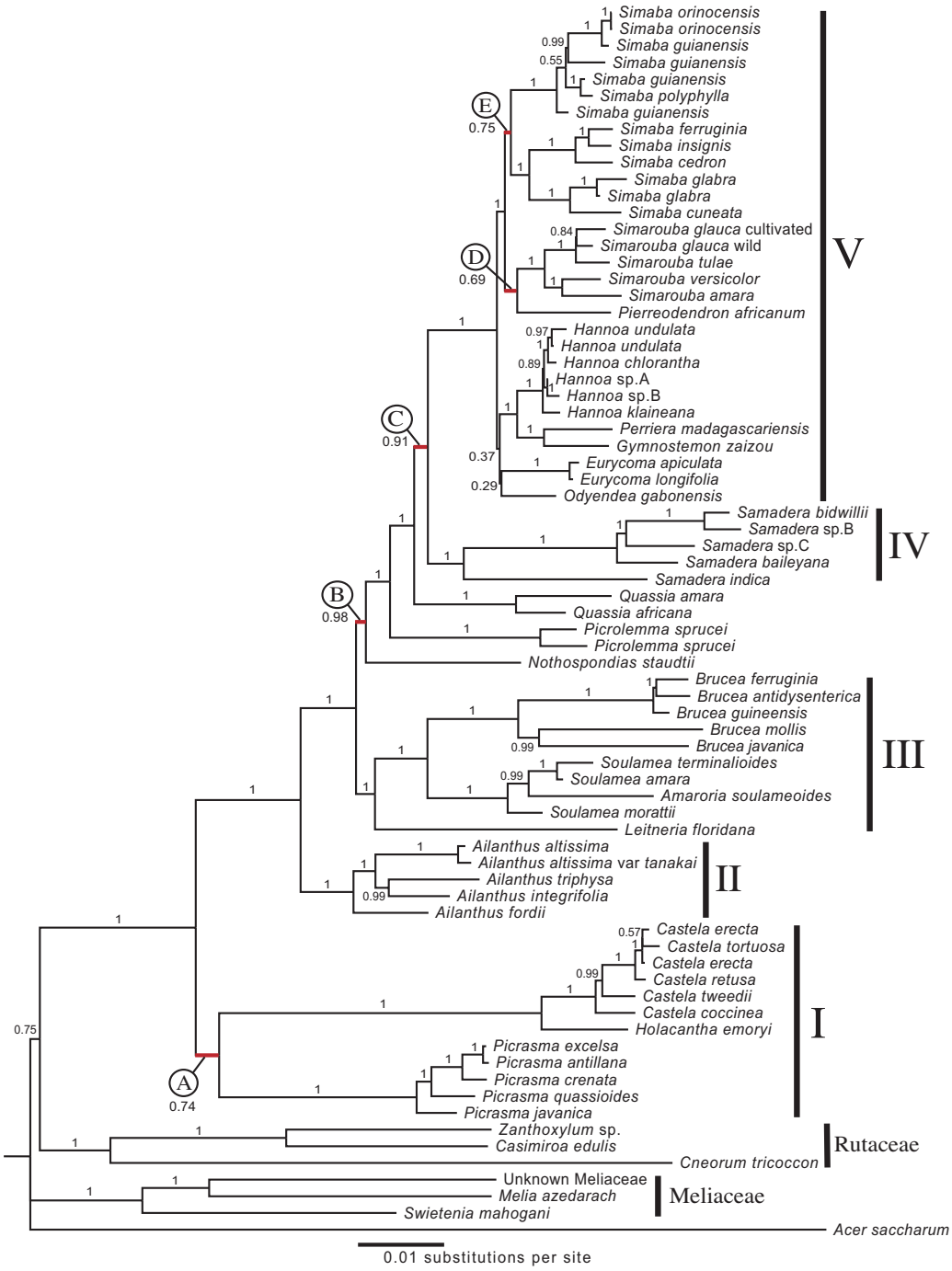


Figure 3-2. Phylogram of the majority rule consensus of trees for Simaroubaceae, sampled from the posterior distribution in the Bayesian analysis partitioned by gene and codon position (5,000,000 generations, sampled every 1000 generations; burnin = 100,000 generations). Posterior probabilities (PP) are shown on branches. Branch lengths are proportional to the mean number of nucleotide substitutions per site, calculated from the posterior distribution density. Five nodes compared between different partitioning strategies are highlighted (A – E). Major clades (I – V) of Simaroubaceae are indicated on the right hand side.

CHAPTER 4
LIKELIHOOD ANALYSIS OF GEOGRAPHIC RANGE EVOLUTION IN SIMAROUBACEAE

Introduction

The number of phylogeny-based biogeographical treatments for tropical groups has grown considerably in recent years (Renner et al., 2001; Richardson et al., 2004; Davis et al., 2002; Lavin et al., 2000, 2004; Weeks et al., 2005; Zerega et al., 2005; Muellner et al., 2006), with improved phylogeny reconstruction and dating techniques, and a diversity of methods for ancestral area reconstruction (AAR). However, resolved and well-supported phylogenies of pantropical clades, that is, clades with taxa distributed in each of the neotropics, tropical Africa and tropical Asia and Australia, are still relatively few (e.g. Renner et al., 2001; Davis et al., 2002; Lavin et al., 2004; Weeks et al., 2005; Zerega et al., 2005; Muellner et al., 2006). Simaroubaceae are an ideal pantropical candidate for exploring hypotheses of intercontinental migration, given their well-supported relationships, taxon sampling and broad geographical distribution (Chapter 3). Although the family is much smaller in species diversity (~109 species) than many well-studied pantropical clades, a number of intriguing disjunctions warrant a detailed investigation. All three major tropical disjunctions, between Asia and the New World, between Africa and the New World, and between Africa and Southeast Asia, occur in the family, and occur multiple times. Furthermore, Simaroubaceae have tropical, subtropical and temperate elements, including both mesophytic and dry-adapted taxa. These factors add an ecological dimension to the interpretation of biogeographical patterns (Riddle, 1996; Wiens and Donoghue, 2004) and may provide evidence for migration and speciation as a result of climatic change. Recent studies have shown the biogeographical history of a number of tropical angiosperm clades (e.g. Renner et al., 2001; Davis et al., 2002; Weeks et al., 2005; Zerega et al., 2005; Muellner et al., 2006) to be the result of interactions between climate change, migration via land-

bridges and long-distance dispersal, and the synthesis of data such as these has provided an analytical framework for the study of the origins and maintenance of tropical biomes as a whole (Pennington et al., 2004a, 2006; Davis et al., 2005). Additionally, the timing of radiations of strictly tropical taxa from subtropical and temperate ancestors, may reveal how adaptive shifts contribute to the modern tropical community (Lavin et al., 2004), and Simaroubaceae provide an opportunity for such a study. Although the family does not constitute a significant component of biome diversity in terms of species numbers as in, for example, Fabaceae (Lavin et al., 2004) or Melastomataceae (Renner et al., 2001), it may provide data on how smaller families enhance the phylogenetic diversity (PD; Faith, 1992a; Barker, 2002) of species-rich environments.

Simaroubaceae have a variety of fruit types and morphology; fruits can be small and nut-like (e.g. *Brucea*, *Eurycoma*), small and fleshy (e.g. *Picrasma*, *Castela*), large and fleshy (e.g. *Gymnostemon*, *Pierreodendron*), dry and winged (samaras; e.g. *Ailanthus* and *Soulamea*) and woody and floating (*Samadera indica*). Reconstructing dispersal events in the history of the family may provide some insight into how dispersal capabilities are manifest in fruit type.

The goals of this study are to: 1) reconstruct divergence dates within Simaroubaceae using recently developed molecular rates analyses and fossil calibration; 2) propose a center of origin for the Simaroubaceae based on molecular rates analyses and AARs; 3) use suitable inter- and intracontinental dispersal routes, their maintenance and disappearance, to test hypotheses of vicariant cladogenesis in the history of the family; and 4) identify potential long-distance dispersal events when vicariant hypotheses of speciation seem doubtful, and relate these to fruit morphology. Consideration will be given to the effects of past episodes of climate change, the ecological requirements of extant species, and how these might be reflected in ancestral distributions. Maximum-likelihood analysis of geographic range evolution (Ree et al., 2005; Ree

and Smith, 2008) has yet to be applied to a pantropical group such as Simaroubaceae. The complex history of inter- and intracontinental migration that is suggested by phylogenetic relationships between the extant species, and the areas they inhabit, should be well suited to a model-based method, rather than parsimony or event-based methods (dispersal-vicariance analysis (DIVA); Ronquist, 1997) that are currently used. A DIVA analysis is included in the study for comparative purposes.

Methods

Taxon Sampling, DNA Sequencing and Phylogenetic Analyses

Taxa used in this study were the same as those of Clayton et al. (2007), with the addition of four species of *Simaba* (plus a second accession of *S. cuneata*) and one species of *Simarouba*. Sequences for these new taxa were determined using the same methods as Clayton et al. (2007) for four gene regions, *rbcL*, *atpB*, *matK* plus partial *trnK* intron, and *phyC* (Genbank accession numbers EU546227 – EU546249). These taxa were then added to the 67 accessions used in the previous study, totalling 73 ingroup accessions overall, and 62 out of ca. 109 species. Phylogenetic analyses consisted of Bayesian inference. Sequence data were partitioned by gene region and codon position for coding regions. This partitioning strategy was determined to be the best-fitting model using Bayes Factors (Chapter 3). Substitution models for each partition were determined with the Akaike information criterion (AIC) in Modeltest, version 3.6 (Posada and Crandall, 1998). Two independent analyses were run for 10,000,000 generations each, using four Markov chains and default parameters; trees were sampled every thousandth generation, with a burn-in of 150,000 generations, and stationarity of the Markov chain Monte Carlo (MCMC) was determined using the average standard deviation of split frequencies between runs. The data matrix and resulting tree were submitted to TreeBASE (matrix accession number M3859; study accession number S2060). Compared to analyses in Clayton et al. (2007), no changes to the

topology, its resolution, or clade support were found with the additional species. A likelihood-ratio test (LRT) was performed on the combined data in PAUP* 4.0b10 (Swofford, 2002), to determine if the regions sequenced are evolving in a clock-like fashion using maximum likelihood. The substitution model for the combined data set (GTR+I+G) was determined using AIC in Modeltest. Results of this test ($\chi^2 = 245$ and d.f. = 79, $p < 0.0001$) rejected the molecular clock hypothesis, and a relaxed clock method was used.

Divergence Date Estimation

Divergence date estimation was performed using the uncorrelated lognormal (UCLN) model of molecular rates estimation (Drummond et al., 2006), implemented in BEAST (v1.4.7, Drummond and Rambaut, 2007), a Bayesian approach to molecular rates analysis. This method uses a global sampling strategy (Markov Chain Monte Carlo) to maximize both phylogenetic inference and molecular rate estimation on branches simultaneously (Drummond et al., 2006), thus eliminating the problems associated with selecting a tree topology and branch lengths a priori. The method also allows for calibration date priors to have a probability distribution, to reflect uncertainties in fossil dates.

Fossil calibration

The phylogeny was calibrated with four fossils: samaras of two different fossil species of *Ailanthus* (Corbett and Manchester, 2004), *Leitneria* fruit endocarps (Dorofeev, 1994; Nikitin, 2006a) and *Chaneya* fruits (Wang and Manchester, 2000; Teodoridis and Kvacek, 2005). *Ailanthus* samaras are common in deposits throughout the Tertiary across the Northern Hemisphere (Fig. 1), with earliest occurrences in the Middle to Early Eocene. The earliest example is from the Fossil Butte fish quarries of the Green River formation, Wyoming, dated at 52.2 – 52.7 Ma. Fossil *Ailanthus* samaras share all the fruit synapomorphies of the extant genus that are discernible in the preserved remains. However, attributing the fossil fruits specifically to

one or another extant species is difficult, due to the limited number of characters available (Corbett and Manchester, 2004). Three fossil species are recognized: *A. confucii*, the most commonly occurring species in the genus' fossil history, and closest morphologically to extant *A. altissima*; *A. tardensis*, known from a single locality in Hungary in the Early Oligocene, and closest in form to extant *A. triphysa*; *A. gigas*, known from a single locality in Croatia in the Early Oligocene, and described as being "intriguingly similar" to extant *A. integrifolia* (Corbett and Manchester, 2004). Although the three fossil species have affinities to extant species, there is variation in characteristics such as fruit size and position of the ventral vein. *Ailanthus confucii* specimens are the earliest example of the genus; however, *A. fordii* is the sister to all other extant species (although *A. excelsa* has not been sampled; Chapter 3).

For these reasons three different *Ailanthus* calibration points were tested in the molecular rates analyses. Firstly, the earliest fossil occurrence (*A. confucii*, 52.7 Ma) was used to date the node at which *Ailanthus* diverges from the rest of the Simaroubaceae (the *Ailanthus* stem lineage). Secondly, the earliest fossil occurrence (*A. confucii*, 52.7 Ma) was used to date the node at which *A. altissima* diverges from its sister group (*A. integrifolia*+*A. triphysa*). Thirdly, the *Ailanthus gigas* fossil (Early Oligocene) was used to date the node at which *A. integrifolia* diverges from its sister species (*A. triphysa*). Given the numerous occurrences of *Ailanthus* fossils in deposits across the Northern Hemisphere, and their consistent appearance throughout the Tertiary, up until the Pleistocene of Japan (Corbett and Manchester, 2004), the date for *A. confucii* was considered a robust calibration, as in previous dating analyses (Muellner et al., 2006, 2007). This robustness is considered in terms of how closely the earliest known appearance of the fossil approximates to the first appearance of the genus: *Ailanthus* samaras are readily preserved, straightforward to identify and relatively common, and the sudden appearance

of samaras in the fossil record in the Early Eocene may indicate that the genus did not exist before that time. Proximity of the assigned fossil date to the initial cladogenesis event in geological time is critical to accurate divergence date estimation.

Leitneria has fossilized endocarps from the Early Oligocene to the Pliocene of Western Siberia and the Miocene and Pliocene of Europe (Fig. 1). Several species are described (Dorofeev, 1994; Nikitin, 2006a) and show a striking similarity to the endocarp anatomy of extant *L. floridana*. Nikitin (2006a) recorded a putative *Leitneria* sp. from the Late Eocene (Priabonian), but its precise affinities are questioned by the author, therefore Early Oligocene was used in calibrations. As extant *Leitneria* is a monotypic genus, the only logical position for the calibration point is the node at which *L. floridana* diverges from its sister clade.

A third fossil is the extinct genus *Chaneya* (Wang and Manchester, 2000; Teodoridis and Kvacek, 2005), which has certain affinities with extant *Picrasma*. Although the fossil appears to have oil cells in the perianth, which are lacking in Simaroubaceae, an accrescent pentamerous calyx surrounding one, two or three globose mericarps when mature, is consistent with *Picrasma* (Wang and Manchester, 2000). A *Chaneya* calibration point at the *Picrasma* stem group was therefore tested for compatibility with the other fossils, using its earliest occurrences in the Middle Eocene of western North America and Asia. However, the potential utility of *Chaneya* is limited by its uncertain relationship.

Fossil pollen of Simaroubaceae was not considered for this study, given the lack of distinctive morphological characteristics in extant Simaroubaceae pollen (Basak, 1963, 1967; Moncada and Machado, 1987). Reports of *Ailanthus* fossil pollen (Song et al., 2004) and *Leitneria* fossil pollen (Machen, 1971) were considered unreliable.

In BEAST v1.4.7 (Drummond and Rambaut, 2007), calibration points take the form of priors with a probability distribution. Each prior tested (Table 4-1) took a lognormal distribution, with the means and standard deviations designated to reflect our confidence in the fossils used. The *Ailanthus confucii* fossil had a narrow lognormal distribution, given its well-characterized fossils and more precise stratigraphic age. *Leitneria* was given a broader probability distribution than *A. confucii* to account for the questionable Priabonian species. Both *A. gigas* and *Chaneya* calibrations had larger standard deviations because of the uncertain affiliations of these fossils. The hard lower bounds represent the youngest dates (Donoghue and Benton, 2007) for stratigraphic ages of epoch boundaries (Gradstein and Ogg, 2004), and in the case of *Ailanthus confucii*, the more precise dating of the Fossil Butte fish quarries at Green River, Wyoming. In total, eight different combinations of fossil placements (here referred to as calibration schemes; Table 4-2) were tested, to determine how well fossils and divergence dates correlated with each other. Calibration at the family level (e.g. Wikström et al., 2001, 2004; Muellner et al., 2007) was not used, to avoid compounding the error inherent in secondary calibration dates from other molecular rates analyses (Graur and Martin, 2004). However, results for the age of the Simaroubaceae are discussed with reference to divergence dates from Muellner et al. (2007).

Molecular rates analyses

Analyses for each calibration scheme (Table 2) were performed in BEAST v1.4.7 (Drummond and Rambaut, 2007) using the various calibration points described above. Each analysis was run using the four-gene, 62-species data set plus outgroups. A combined data set of the four coding markers, with the exclusion of the non-coding *trnK* intron, used GTR+I+G as the model of nucleotide substitution, determined with the AIC in Modeltest v3.6 (Posada and Crandall, 1998). The non-coding intron was excluded because it was more difficult to align than the coding region (Chapter 3) and contained large indels in some species; misalignment could

cause discrepancies in branch lengths that in turn would affect estimates of molecular rates along those branches. The coding regions were partitioned into three codon positions, based on results of partitioning strategies examined in Clayton et al. (2007; partitioning by gene region is not currently implemented in BEAUti v1.4.7 (Drummond and Rambaut, 2007)). Both the substitution model and rate heterogeneity were unlinked across codon positions. Priors other than the calibration points were uniform and set at default values, and the tree root prior was set as a uniform distribution, 0 – 132 Myr (the estimated age of the angiosperms (Hughes, 1994)). In calibration scheme 2 this prior produced a tree with zero likelihood and so was removed. Each BEAST analysis used a starting tree with initial branch lengths satisfying the upper and lower bounds for the calibration priors being used, constructed using pathd8 (Britton et al., 2007). An initial analysis was run using the fossil placement in calibration scheme 1 and no topological constraints, producing a tree with an identical topology to the original phylogenetic analysis (Fig. 2), except that *Quassia* was sister to Clade V (PP = 0.86; also seen in parsimony analyses of Clayton et al. (2007)), and all outgroups formed a clade (PP = 0.41) sister to Simaroubaceae. We therefore constrained the topology of all further BEAST analyses to have *Samadera* as sister to Clade V, and Rutaceae sister to Simaroubaceae, as per previous Bayesian analyses (Fig. 1; Clayton at al., 2007). Although there is some evidence to suggest application of a relaxed molecular clock leads to improved phylogeny estimation (Drummond et al., 2006), we preferred to follow the more thorough phylogenetic analyses from the previous chapter, particularly with the single outgroup clade seen in the unconstrained analysis. For each calibration scheme, two independent analyses ran for 10,000,000 generations each, sampling every 1000 generations. Stationarity of the MCMC was determined by comparing standard deviations of node times between the two independent runs, and observing burn-in plots for the posterior and all model

parameter values in Tracer v1.4 (Rambaut and Drummond, 2003). A burn-in of 1,000,000 generations was removed from each run. For all estimated parameters, an effective sample size (ESS) of > 150 was obtained to ensure the parameter space had been sufficiently sampled (except in the case of calibration scheme 2, where two further runs were needed for an ESS > 150 for all parameters). Once stationarity had been determined, the two independent runs for each scheme were combined using LogCombiner v1.4.7 (Drummond and Rambaut, 2007). The calibration scheme ultimately selected for biogeographic analyses, out of the eight combinations of fossils tested (see Results), was generated a second time with four separate 10-million-generation runs and with scale operators adjusted to recommendations in the BEAST output from the original run for that calibration scheme (BEAST input .xml in Supplementary Information). The four runs were combined using LogCombiner v1.4.7 (Drummond and Rambaut, 2007), ensuring ESS > 600 for all parameters.

Biogeographic Analyses Using a Likelihood Approach

Biogeographic data were compiled from species distributions in the literature (e.g. Cronquist, 1944a,b,c,d; Nooteboom, 1962a,b), and GIS mapping of specimen localities in ArcGIS v9.0 (http://www.esri.com/software/arcgis/). The major areas in which Simaroubaceae are distributed were categorized broadly into seven areas for AAR (Fig. 1). Areas were delimited by continental divisions based on present and past separation of major landmasses, with the exception of Asia and Europe. Interpretation of the results, however, incorporates more specific geographic information, especially for island taxa. Figure 4 shows geographical distributions assigned to terminals for AAR, based on the following categories: 1) **N** – North and Central America (north of the Panamanian Isthmus), and Caribbean Islands; 2) **S** – South America; 3) **F** – Tropical Africa; 4) **M** – Madagascar; 5) **A** – Mainland Asia, including but not restricted to India, China, Bhutan, Myanmar, Japan and Russia (fossil localities), plus insular SE Asia,

including but not restricted to Malaysia, Indonesia and the Philippines; 6) **U** – Australia, New Guinea, Papua New Guinea, New Caledonia and the Pacific Islands; 8) **E** – Europe (fossil localities). Widespread species were assigned to more than one area.

The program Lagrange (Ree and Smith, 2008) was used for AAR. Unlike DIVA (Ronquist, 1997), it incorporates an explicit model of dispersal routes available at historical intervals, and estimates vagility and extinction parameters, correlating stochastic events with lineage persistence, as part of the DEC model (Dispersal-Extinction-Cladogenesis; Ree and Smith, 2008). Five different combinations of parameters (here referred to as models M1-M5) were tested to assess how model complexity and alteration of certain parameters would affect the resulting AARs.

The model of dispersal route availability was developed based on geologic history and the presence and dissolution of land bridges and island chains, and climatic data. In particular, geological events such as the presence of the North Atlantic Land Bridge, the Southern track between Australia and South America, island chains between North and South America, the collision of Australian and Asian plates, and the closure of the Panamanian Isthmus (Tiffney, 1985a; Morley, 2003; Scotese, 2001) are accounted for in dispersal probabilities. Two different dispersal scenarios were tested. The first dispersal scenario (D_1) uses probabilities ranging from 0.1 for well-separated areas to 1.0 for contiguous landmasses (Table 4-3). The second scenario (D_2) uses the same connections but imposes much smaller dispersal probabilities (0.01) for areas not directly connected (Table 4-3).

The five models are as follows:

1. M1 – Taxa are free to disperse between any area at any time, but are restricted to the following ancestral ranges (based on potentially plausible and extant ranges): NS, NF, NE, NA, SU, FM, FA, FE, AU, AE, MAU;

2. M2 – Ancestral ranges are unrestricted and taxa are subject to dispersal probabilities of scenario D_1;
3. M3 – Ancestral ranges are restricted as in M1 and subject to dispersal probabilities of scenario D_1;
4. M4 – Ancestral ranges are restricted as in M1 and subject to dispersal probabilities of scenario D_2;
5. M5 – Ancestral ranges are restricted as in M1 (excluded E, NE, FE, AE) and subject to dispersal probabilities of scenario D_1; fossil taxa are removed from the analysis.

An additional analysis was performed with no constraints on the model, in which taxa are free to disperse at any time and ancestral ranges are unrestricted. However, this analysis proved too computationally expensive, and no results were obtained. Fossil areas were integrated by manually adding a terminal for *Ailanthus confucii* (coded as present in E, N and A) as the sister to extant *Ailanthus*, and a *Leitneria* terminal (coded as present in E and A) as sister to extant *Leitneria*. *Chaneya* was not included in the analysis. The non-overlapping ranges of the fossil and extant lineages of *Leitneria* led us to place the fossil lineage as diverging from the extant lineage at 28.4 Ma, the latest stratigraphic date for the Early Oligocene, when the fossils first appeared. *Ailanthus* fossils were more difficult to integrate, given the presence of fossils in Asia, where extant species also live, and the possibility of these fossils representing part of the *Ailanthus* stem lineage. We tested three different points of divergence for the fossil lineage: diverging at 52.2 Myr, i.e. 2.7 Myr from the base of the stem branch, halfway along the stem branch, and 2 Myr from the tip of the stem branch. All three produced near identical likelihood scores and AARs (data not shown). Therefore, the divergence at 52.2 Ma was chosen, as this had the highest likelihood score of the three positions. Input files for Lagrange are published in Supplementary Information.

Model 3 (M3) was chosen for further analysis and discussion of the biogeographical implications of AAR in Simaroubaceae. This choice was not made on statistical grounds, as M3

has a much lower likelihood than M1, M2 and M5 (Table 4-4), with only a slight increase in the number of parameters. A less constrained model may fit the data well, but at the expense of realism. This is the case here, where freedom to disperse anywhere at any time, ancestral ranges that encompass non-contiguous and sometimes well-separated areas, and exclusion of fossils, fail to adequately represent the biology of the organisms. We optimized the chosen model (M3) by examining the effects of different ancestral ranges between the Old World and New World (namely NA, NE, NF and SU) on the likelihood score of the model. These potential migration routes are important to pantropical clades, and in Simaroubaceae, disjunctions between the Old World and New World are common throughout the family. Each ancestral range was tested alone, along with all possible combinations (Table 4-5). It is not clear how to compare likelihood scores between models with changes in the specified ancestral ranges, because the models are not nested, and the additional parameters involved in having fewer ancestral ranges are fixed at a boundary value of zero (R. Ree, pers. comm.). The resulting likelihood scores were therefore compared directly.

A DIVA analysis (Ronquist, 1997) was also performed on the data, to compare AARs between this event-based approach and the explicit model-based approach of Lagrange (Ree and Smith, 2008). Tree topology and area assignments from Lagrange were also used in DIVA, which ran with default program parameters. Two analyses were conducted, with maxareas unlimited and constrained to two.

Results

Divergence Date Estimation

Tracer output statistics for BEAST analyses show the data are not evolving in a clock-like manner (coefficient of variation for calibration scheme 7 = 0.31 [0.22, 0.40]), and substitution rates between ancestor-descendant lineages are uncorrelated (covariance for

calibration scheme 7 = 0.0071 [-0.16, 0.15]). Therefore, the UCLN model of rate variation was the most appropriate method for estimating divergence dates, as opposed to methods such as r8s (Sanderson, 1997, 2002) and multidivtime (Thorne and Kishino, 2002), which use an a priori assumption of autocorrelation between ancestor-descendent lineages. However, the covariance statistic appears to bias results against finding autocorrelation (A. Drummond, pers. comm.). Results of the different calibration schemes are shown with dates for fossil nodes as well as other example nodes from the tree (Table 2), and lower and upper 95% highest posterior densities are given in square brackets. Fossils of *Leitneria* and *Ailanthus confucii* (calibrating the *Ailanthus* stem group) corresponded reasonably well with each other (calibration schemes 1 and 5); therefore, these two fossils were used together (calibration scheme 7) to produce the chronogram for biogeographic analyses (Fig. 3). In calibration scheme 7, as well as in scheme 1, the divergence date of *Leitneria* was estimated to be older than the fossil dates. Using *A. confucii* at the divergence of *A. altissima* (scheme 2) produced very old dates throughout the tree, with the genus diverging from its sister clade at 139.3 [95.0,190.0] Ma. *Ailanthus gigas* gave higher age estimates on its own (family crown group 87.3 [67.9,105.3] Ma), and in the final chronogram (calibration scheme 7) the divergence of *A. integrifolia* was considerably younger than the *A. gigas* fossil, hence its exclusion from the final analysis. *Chaneya* alone produced younger dates than the other calibrations, estimating the divergence of *Ailanthus* at 35.2 [27.8,42.9] Ma.

Biogeographic Analyses

Likelihood scores and estimates of dispersal and extinction rates for the five models tested are shown in Table 4-4, along with AARs for selected nodes in the phylogeny, and results of DIVA analyses. Models M1 and M2 showed numerous alternative reconstructions for nodes throughout the tree (within the confidence window of a 2 log likelihood unit difference (Edwards, 1992; Ree and Smith, 2008)). For example, the root node in M2 had 33 alternative

reconstructions. Dispersal and extinction rate estimates were lower for M1 and M2 than the more constrained models. In M3, M4 and M5, alternative reconstructions were common, but typically involved the same areas, and only differed in how the ranges evolved at the node. DIVA results showed unconstrained analyses had numerous equally optimal reconstructions at deeper nodes in the phylogeny (Table 4-4). With ancestral ranges restricted to two areas, results were similar to Lagrange, particularly the less constrained models. Optimal reconstruction required 30 dispersal events for the unconstrained analysis, and 32 dispersals when maxareas was limited to two.

The effects of different ancestral range combinations in M3 are shown in Table 4-5, ordered by decreasing likelihood. Analyses including SU and NE typically had lower likelihoods and higher dispersal and extinction rates than other combinations. Although the likelihood scores could not be compared using an objective criterion such as the AIC (Akaike, 1974), the top four ancestral range permutations had scores over 8 log likelihood units better than the next best model, and all produced near-identical AARs (Table 4-5), except for the root node. The combination with the highest likelihood was NA + NF, and this is shown in Figure 4-4. For clarity, the chronogram has been pruned of selected terminals that have a sister taxon in the same area. For nodes where likelihood scores were not significantly different between multiple reconstructions, the relative probability of the global likelihood is shown. Figure 4 shows three circumstances of range inheritance. Firstly, there are instances of dispersal resulting in range expansion, which are common throughout the tree, especially in the Miocene. Secondly, there are local extinction events (three are hypothesized), which are inferred when a daughter lineage inherits a different range from its parent (a range expansion prior to extinction is inferred). Thirdly, there is vicariance by cladogenesis, where the ancestral range encompassing two or more areas subdivides between daughter lineages. Also seen in Figure 4-4 are numerous

occurrences of incipient speciation, whereby an ancestral range of two areas subdivides, with one daughter inheriting the widespread ancestral range and one daughter inheriting just one of the two ancestral areas. An example is the root node of clade III (Fig. 4-4).

Discussion

Divergence Date Estimation

The different placements of the three fossil constraints strongly affected the divergence dates estimated, especially dates prior to the divergence of *Ailanthus*, that is, the divergence of *Picrasma*, *Castela* and *Holacantha*, and the age of the family. Prior to the divergence of *Ailanthus*, however, only minor differences were seen between different fossil placements in Clades III, IV and V. There are many sources of error in divergence time estimation (Sanderson, 2002; Graur and Martin, 2004; Near and Sanderson, 2004). These include, but are not limited to: accuracy of dating fossil strata; correctly identifying fossils (e.g. the ambiguity around the *Chaneya* assignment); assigning fossil dates to the most appropriate nodes in the phylogeny; poorly supported nodes in the phylogeny; justifying the nature of the prior probability distribution of dates around the fossils; errors in estimating rates of molecular evolution, such as the problems inherent in decoupling rates and time, correctly estimating branch lengths on the topology, and assuming certain characteristics of molecular evolution (e.g. the uncorrelated lognormal distribution prior for substitution rates). However, these potential sources of error do not necessarily inhibit our ability to make general statements about migration patterns and dispersal events, in a system which encompasses large-scale and relatively slow processes, such as continental drift and land bridge formation and dissolution, over long periods of time.

Cross-calibration of fossils showed some compatibility between *Ailanthus confucii* and *Leitneria* fossils, but only when *A. confucii* was associated with the *Ailanthus* stem group. Calibration scheme 7 was chosen as the best chronogram for biogeographic analyses, as it

included the two fossil placements that had the best fossil records and showed some overlap in the probability distributions of node ages. In calibration schemes with multiple calibrations, *A. confucii* at the *Ailanthus* stem lineage was the most influential calibration point. The older ages for the *Leitneria* divergence produced in calibration schemes 1, 4, 7 and 8 could be accounted for by the *Leitneria* sp. described by Nikitin (2006a) as questionably Priabonian, that is, older than the conservative lower bound at the end of the Early Oligocene. Although the *Chaneya* calibration (scheme 6) produced dates younger than those derived from the more extensive fossil records of *Ailanthus* and *Leitneria*, the mean dates produced for the *Picrasma* stem and crown in calibration scheme 7 were 67.4 [57.8,77.8] Ma and 14.2 [9.2,20.0] Ma, respectively, and the *Chaneya* fossils fall between these dates. Therefore, an association between *Chaneya* and *Picrasma* remains a possibility.

The final chronogram (Fig. 3) illustrates when the major clades of Simaroubaceae arose. Muellner et al. (2007) estimated the age of the Simaroubaceae crown group at 52 Myr, as this was the most conservative placement for the *Ailanthus* fossil, given their limited taxon sampling. With the addition of *Picrasma*, *Castela* and *Holacantha* in this study, the age of the crown group is estimated to be Maastrichtian. The mean age of the stem group (88.4 [75.9,101.4] Ma) is difficult to compare to dates provided by Muellner et al. (2007), given that the sister relationship of the family remains poorly resolved.

Biogeographic Analyses

Likelihood models

Reconstructions have a tendency to become more ambiguous deeper in the phylogeny (the so-called widespread ancestor problem; Ree et al., 2005), which is seen in the number of ambiguous nodes in the AARs. This is an inevitable result given the amount of time that has passed since the Simaroubaceae first arose, and the extrapolations made from geographic

distributions of the extant tips. The highly ambiguous reconstructions for M2 throughout the tree, especially at the deeper nodes, suggested the need for a biogeographic model that better explains the data. The impact of the fossils was primarily the inclusion of Europe as the seventh area in analyses, but they only influenced AARs in analyses which did not include NA (Table 4-5). In reality the connection between Europe and North America was an integral part of the NALB, but with only two terminals assigned to Europe in the tree, the NE ancestral range did not feature in the chosen analysis.

The restriction of ancestral ranges to one or two (and in one case three) areas helps to counterbalance the widespread ancestor problem (Ree et al., 2005), and is an assumption based on the low likelihood of an ancestral species having a range across three or more of the geographic areas described above. Given the extant species distributions, only *Samadera indica* has a presence in three areas. Specification of ancestral ranges was an influential factor in how ancestral areas were reconstructed, particularly the inclusion or exclusion of NA, NF or NE. These differences indicate the critical nature of the path between the Old World and the New World, and which is more plausible for Simaroubaceae. Comparison of likelihood scores between NA-, NE-, NF- or SU-constrained models (Table 4-5) shows that NA has the highest likelihood for a single ancestral range, and in combination NA and NF have the highest likelihood (Fig. 4). This suggests trans-Beringial migration to be an important determining factor in Simaroubaceae biogeography, more so than the NALB. Exclusively trans-Atlantic disjunctions (excluding *Ailanthus* fossil species which could have migrated via Beringia or the NALB) were only hypothesized for three nodes, one of which arose after the NALB was available (the root node of *Simaba, Simarouba* and *Pierreodendron*), and one of which the timing is unknown (the dispersal of *Nothospondias* to Africa). Models that included NF and NE,

but excluded NA, increased probability of migration via the NALB as opposed to Beringia a priori, and may be more biologically realistic for tropical groups. However, the early-diverging clades of Simaroubaceae, including genera such as *Picrasma*, *Ailanthus* and *Leitneria*, continue to maintain a temperate element, supporting the Beringial track. The Southern track (SU) produced significantly lower likelihoods than the other three ranges, due to how disconnected South America and Australia were through much of the Tertiary. North America was typically the optimal reconstruction for the root node in the ancestral range permutations with the highest likelihoods. Ree and Smith (2008) note that in their analysis of *Psychotria*, widespread ranges are rapidly reduced to single areas by cladogenesis or local extinction. However, in Simaroubaceae, many nodes retain widespread distributions, especially deeper in the phylogeny. Most of these widespread distributions break down through vicariance, with few lineages going extinct locally (*Picrolemma* and *Nothospondias*). Whether these ancestral ranges are plausible over a significant period of time is debatable, but the assumption of trans-Atlantic and trans-Beringial disjunctions in the early history of the family, has been suggested for other tropical angiosperm families. The likelihood model could be streamlined further by introducing stratified ancestral ranges, such that ancestral distributions can vary with different periods in the history of the group.

The differences in AARs between dispersal scenarios D_1 and D_2 were small, for example, a possible European-North American ancestor for *Ailanthus* and clade III in D_2 (Table 4-4). However, AAR of almost all nodes was the same in the optimal reconstruction between D_1 and D_2. As expected, the inclusion of more stringent dispersal probabilities in D_2 produced a lower overall likelihood score and much higher estimated rates of dispersal and extinction. The impact of more stringent dispersal was seen in very few clades, but notably for dispersals between North

and South America in *Holacantha* and *Castela*, due to the reduced probability of movement between these two areas in D_2 (data not shown). The fact that both scenarios produced near-identical AARs leads us to believe that only extreme changes in dispersal probabilities, such as a reduction to zero between areas, will result in AARs that affect our interpretation of the data. Because of the similarities between D_1 and D_2, results for the more relaxed dispersal probabilities in D_1 (which had a much higher likelihood score) are illustrated in Figure 4-4 and discussed in the context of historical biogeography.

Ultimately we must assess whether the model-based approach is a valid one, given the ambiguity in the results and the number of model parameters that have been introduced to fit the data. In any likelihood analysis there is a trade-off between a model too simplistic to describe the data accurately, and one that is overly complex, leading to overconfidence in one particular result. Ambiguity is seen in AARs for all models, and there are differences among the five models (Table 4-4) and ancestral range permutations (Table 4-5), especially for deeper nodes. This ambiguity is likely a reflection of a simplistic model attempting to describe the multitude of historical events that have shaped the family's history. Despite this limitation, there is an overall biogeographical signal concurrent among all of the different models tested. Optimal AARs are congruent among models for lineages diverging in the Late Oligocene and later, and the signal is one of multiple recent range shifts. These recent shifts appear to overshadow reconstruction of events deeper in the family's history, and as a result, Paleocene, Eocene and Early Oligocene reconstructions are much more sensitive to the specification of the model. However, when alternative scenarios within each model have very similar likelihoods, the areas involved are typically the same (Table 4-4). It is how these areas are inherited, whether through vicariance, incipient speciation or dispersal followed by extinction, that distinguishes the alternative AARs.

The likelihood method encourages a more fluid approach to geographic range evolution and historical biogeography, a discipline that has traditionally been dominated by the vicariance-dispersal dichotomy.

DIVA produced results similar to Lagrange when ancestral ranges were confined to two areas, but South America was more commonly involved in DIVA AARs than in likelihood analyses (Table 4-4). Without the maxareas constraint, most deeper nodes in the phylogeny had implausibly widespread distributions. Some improbable ancestral ranges were observed, such as SA and FU, demonstrating the limitation of this method in a complex system. Also there is no objective method to favor one reconstruction over another at any particular node. Freedom from temporal constraints (Donoghue and Moore, 2003) and limited ancestral ranges removes extinction as a factor in shaping distributions, as alternative vicariant events will always be "cheaper" in the DIVA cost matrix.

Origin and early history of Simaroubaceae

Given the small size of Simaroubaceae (ca. 109 spp.), which is more comparable to a large genus than previously studied tropical families (Melastomataceae: 3000 spp.; Malpighiaceae:1200 spp., Meliaceae: 550 spp.; Burseraceae: 500 spp.; Moraceae: 1500 spp. including temperate elements; Annonaceae: 2300 spp.; Dalbergioid legumes: 1100 spp. (species numbers from Judd et al., 2007; Lavin et al., 2000)), the complex reconstruction of biogeographical events proposed, and the resulting extant geographical heterogeneity, is perhaps more surprising. This is especially true given the origin of the family in the Late Cretaceous, a similar timeframe to many of the families mentioned above. Additionally, although attempts were made to include extinct species in AAR, further extinctions in the family are hypothesized, especially for depauperate and isolated lineages, such as *Nothospondias* (arising 46.7 [41.8,51.5]

Ma, a single species in Africa) and *Picrolemma* (diverging 39.1 [33.7,44.4] Ma, two species in South America).

Based on the results of models including fossil data, the family originated in North America, with expansion to Asia via Beringia early on in the Tertiary. This is hypothesized if we assume ancestral species were adapted to temperate climates, and several extant species from early-diverging clades retain this characteristic. However, if we assume ancestral species were tropical, we could rule out NA and SU as ancestral ranges, and the results for NF + NE (the highest likelihood ancestral range permutation without NA or SU) support movement via the NALB, which has been proposed for other tropical groups (Lavin et al., 2000; Renner et al., 2001; Davis et al., 2002; Richardson et al., 2004; Weeks et al., 2005; Zerega et al., 2005; Muellner et al., 2006). The origin of clade I is North American only; however, the relationship of *Picrasma* to *Castela* and *Holacantha* is not strongly supported in phylogenetic analyses (Chapter 3), which could impact AAR for the earliest-diverging nodes in the phylogeny. Clade III is hypothesized to have originated in Asia and North America, with subsequent dispersals to Africa and SE Asia. The presence of *Nothospondias* in Africa is the result of a migration via the NALB followed by local extinction in North America, although a later episode of long-distance dispersal cannot be ruled out. Similarly, *Picrolemma* dispersed to South America with subsequent local extinction, possibly via island chains, or potentially much later after the closure of the Panamanian Isthmus.

Long-distance dispersal

The break-up of Gondwanaland, traditionally the most parsimonious explanation for tropical disjunctions (Raven and Axelrod, 1974), has now been replaced by hypotheses incorporating knowledge of phylogenetic relationships, more realistic divergence dates, availability of dispersal routes (Morley, 2003), and a general acceptance of long-distance

dispersal as a major driving force in extant plant distributions (Givnish and Renner, 2004; Lavin et al., 2004; Renner, 2004a; De Queiroz, 2005). The African-Mesoamerican/South American disjunctions in Simaroubaceae mostly occurred after the NALB was viable for tropical groups (Tiffney, 1985a). The ancestor of *Quassia amara* and *Q. africana* supposedly maintained a trans-Atlantic disjunction for 22 Myr, which underwent vicariance, an unlikely scenario without the NALB for much of that period. NF and NE ancestors are unlikely in the Miocene when tropical vegetation was restricted to lower latitudes; therefore, long-distance dispersal is the best explanation for these disjunctions, and dispersals between Africa and South America are hypothesized for a number of well-studied clades (Lavin et al., 2004; Renner, 2004a; De Queiroz, 2005). *Leitneria* may be another candidate for trans-Atlantic or trans-Pacific long-distance dispersal, but this is dependent on how the fossils are integrated, and if they are considered part of the extant lineage. In our interpretation, migration via Beringia is most likely, but the presence of fossils in Europe and Russia, with no definitive fossils in North America, might suggest a late arrival to North America via long-distance dispersal.

Disjunctions between mainland Asia and North America occur in a number of temperate genera (Gray, 1859; Li, 1952; Donoghue et al., 2001; Xiang and Soltis, 2001; Donoghue and Smith, 2004), with phylogenetic relationships between the two areas showing a variety of patterns. The disjunction seen in *Picrasma*, between two Asian/SE Asian species and six Central and South American and Caribbean species, is less common (e.g. *Magnolia*, Azuma et al., 2001; *Ehretia*, Gottschling et al., 2004; *Hedyosmum*, Zhang and Renner, 2003). Reconstructions suggest *Picrasma* arrived in Asia from North America, but the timing of this migration is unknown, as ancestral *Picrasma* could have maintained a North American-Asian ancestral range beginning at any time along the stem branch of the genus. This ancestral distribution may be

reflected in fossils of *Chaneya*, which are widespread across the Northern Hemisphere (Fig. 1). Trans-Beringial migration is plausible for cool-adapted species (Sanmartin et al., 2001; Donoghue et al., 2001), and *Picrasma quassioides* is a temperate species that occurs as far north as northern Japan. Range expansion into the Asian tropics may have prompted the incipient speciation of *P. javanica*, with the ancestral NA range eventually undergoing a vicariant split ~9.9 [6.1,14.0] Ma, and the New World species moving south. Alternatively, a long-distance dispersal event by the ancestor of *P. quassioides* could be inferred, but this would be more plausible if it were adapted to a climate similar to that inhabited by the New World species.

Dispersal between Africa and Asia has been proposed for clades such as Crypteroniaceae (Conti et al., 2002) under the hypothesis of rafting on the Indian sub-continent, and *Exacum* (Yuan et al., 2005) and Melastomataceae (Renner, 2004b) through long-distance dispersal. *Eurycoma*, nested within the African species, and sister clades *Brucea* and *Soulamea*, fit with the idea of long-distance dispersal, given the significant ocean barrier between Africa and mainland and SE Asia during the Oligocene (Scotese, 2001; Morley, 2003), coupled with climatic cooling. Long-distance dispersal to Madagascar has been suggested for many taxa, both animals (Poux et al., 2005) and plants (e.g. *Adansonia* (Baum et al., 1998), *Exacum* (Yuan et al., 2005), *Nepenthes* (Meimberg et al., 2001), Melastomataceae (Renner, 2004b)). Simaroubaceae are represented on Madagascar by two species, both of which arrived there by long-distance dispersal in the Miocene. With the presence of numerous island chains among mainland Asia, the SE Asian islands and Australia during later epochs (Meimberg et al., 2001; Cannon and Manos, 2003; Morley, 2003), island-hopping dispersal explains the widespread distributions of species of *Ailanthus*, *Picrasma*, *Brucea*, and *Samadera*. *Soulamea*, like a number of other genera e.g. New Caledonian Sapotaceae (Bartish et al., 2005), *Cyrtandra* (Cronk et al., 2005), and *Weinmannia*

(Bradford, 2002), has readily dispersed between New Caledonia and surrounding islands, notably Fiji and Tuvalu, reaching as far as the Seychelles (*S. terminalioides*).

In the New World, *Castela* and *Holacantha* show evidence of multiple dispersals between North and Central and South America, with speciation on several Caribbean islands and the Galápagos. This pattern reflects the amphitropical disjunctions of other dry-adapted clades such as *Tiquilia* subg. *Tiquilia* (Moore et al., 2006) and *Hoffmannseggia* (Simpson et al., 2005). New World disjunctions are also seen in *Simarouba* and *Picrasma*, and both genera have at least three species distributed on Caribbean islands, plus *Simarouba glauca* in south Florida, implying one or more overwater dispersal events within the last ten million years (Lavin et al., 2003; Santiago-Valentin and Olmstead, 2004; Morris et al., 2007).

There is no obvious link between dispersability and diaspore size in Simaroubaceae, especially in the tropical clades, which tend to have drupaceous, bird- or mammal-dispersed fruits (e.g. Hardesty et al., 2005). The fruits of *Castela*, *Picrasma* and *Simarouba* are small, bird-dispersed drupes, and so north-south dispersal may be facilitated by the migratory patterns of fruit-eating birds in the New World. Similarly, dispersal of the large fleshy drupes of *Gymnostemon* and *Perriera* between mainland Africa and Madagascar, may involve migrating birds (Renner, 2004b). Wind-dispersed *Ailanthus* was widespread across the Northern Hemisphere based on fossil evidence (Corbett and Manchester, 2004), and extant *A. altissima* is a weedy species well known for its dispersability (Benvenuti, 2007). The fruits of *Soulamea* and *Samadera indica* are dry, with an air cavity allowing them to float (Nooteboom, 1962b), which potentially explains the widespread distribution of these taxa around the Indian Ocean basin, in conjunction with ocean currents, such as monsoon circulation seen today (Schott and McCreary,

2001). *Leitneria*, which inhabits swamp and coastal forests, also has floating fruits, which could have facilitated an oceanic migration to the New World from Europe or Asia.

Table 4-1. Fossil calibrations used in BEAST analyses, for testing compatibility between fossils, and constructing the final chronogram for biogeographic analyses.

Calibration point	Fossil	Prior distribution	Mean / SD of lognormal distribution	Hard lower bound / Mean / Soft upper bound (95%) (Ma)
A	*Ailanthus confucii*	lognormal	1.0 / 0.5	52.2 / 55.3 / 58.4
B	*Ailanthus gigas*	lognormal	1.0 / 1.0	28.4 / 32.9 / 42.5
C	*Leitneria* spp.	lognormal	1.5 / 0.75	28.4 / 34.3 / 43.8
D	*Chaneya* spp.	lognormal	1.0 / 1.0	40.4 / 44.9 / 54.5

Table 4-2. Divergence dates (in Ma) resulting from molecular rates analyses for eight different calibration schemes. Numbers in bold are dates for calibrated nodes, using priors described in Table 4-1. Lower and upper 95% highest posterior densities are shown in square brackets. The latter four nodes are examples taken from the phylogeny for comparative purposes.

	Ailanthus stem group (*A. confucii* fossil)	*A. altissima* divergence (*A. confucii* fossil)	*A. integrifolia* divergence (*A. gigas*)	Leitneria divergence	*Picrasma* stem group (*Chaneya* fossil)	*Picrasma* crown group	Clade V crown	Family crown	*Quassia* crown
1	**55.1[52.9,58.1]**	22.3[14.5,30.1]	17.5[10.2,25.1]	42.8[36.8,48.0]	68.6[59.3,79.3]	14.3[9.4,19.9]	20.9[17.0,25.4]	71.0[62.7,79.8]	13.6[7.9,19.9]
2	139.3[95.0,190.0]	**55.2[52.9,58.3]**	43.3[30.2,54.0]	107.6[71.1,148.4]	173.6[112.8,236.8]	36.2[20.2,54.6]	53.0[35.6,73.7]	179.5[118.1,243.1]	34.1[16.3,52.2]
3	68.9[53.7,84.3]	35.0[29.6,41.9]	**30.4[28.5,33.5]**	53.2[40.5,66.3]	84.3[65.0,104.0]	18.5[11.1,26.8]	26.3[19.4,33.4]	87.3[67.9,105.3]	17.0[8.4,25.6]
4	**55.8[53.0,59.7]**	33.2[29.2,37.7]	**30.0[28.5,32.3]**	43.3[37.3,49.2]	70.2[58.4,81.6]	15.0[9.9,21.0]	21.8[17.9,26.2]	72.7[63.5,83.4]	14.1[7.9,20.4]
5	43.1[35.0,53.9]	17.3[10.6,25.0]	13.5[7.1,20.4]	**33.3[28.8,40.4]**	53.7[41.6,69.0]	11.1[6.9,16.0]	16.4[12.4,21.5]	55.5[43.5,70.3]	10.6[5.8,16.2]
6	35.2[27.8,42.9]	14.1[8.8,19.9]	11.0[6.2,16.3]	27.2[20.9,33.8]	**44.0[40.5,51.0]**	9.2[5.6,13.1]	13.5[10.2,17.3]	45.4[40.5,53.6]	8.6[4.8,12.9]
7	**54.9[52.8,57.4]**	21.9[14.3,29.8]	17.1[10.3,25.1]	**40.7[34.3,46.4]**	67.4[57.8,77.8]	14.2[9.2,20.0]	20.6[16.8,24.6]	69.8[62.1,78.6]	13.4[7.7,19.4]
8	**55.3[53.0,58.4]**	33.1[29.1,37.6]	**30.0[28.5,32.3]**	**41.2[34.5,47.4]**	69.1[58.7,81.5]	15.0[9.6,20.6]	21.6[17.7,26.2]	71.5[63.0,82.2]	13.9[7.8,20.4]

Table 4-3. Dispersal network model implemented in Lagrange, showing probabilities of dispersal between areas for two different scenarios (see text), based on data from Tiffney (1985a), Morley (2003) and Scotese (2001). Probabilities are for movement in both directions.

Area connections	Time period (Ma)	D_1	D_2
All overwater dispersals without island chains	70 – 0	0.1	0.01
Areas not adjacent between 70Ma and present	70 – 0	0.1	0.01
North America – South America	70 – 45	0.25	0.01
	45 – 5	0.75	0.01
	5 – 0	1.0	1.0
North America – Europe	70 – 30	0.75	0.1
	30 – 0	0.1	0.01
North America – Mainland and SE Asia	70 – 30	0.5	0.1
	30 – 0	0.1	0.01
North America – Africa	70 – 30	0.5	0.1
	30 – 0	0.1	0.01
South America – Australia	70 – 45	0.5	0.1
	45 – 0	0.1	0.01
Africa – Madagascar	70 – 0	0.5	0.1
Africa – Mainland and SE Asia	70 – 45	0.75	0.1
	45 – 30	0.75	0.01
	30 – 0	0.1	0.01
Africa – Europe	70 – 30	0.75	0.1
	30 – 5	0.5	0.01
	5 – 0	0.25	0.01
Mainland and SE Asia – Australia	70 – 45	0.1	0.01
	45 – 30	0.5	0.01
	30 – 0	0.75	1.0
Mainland Asia – Europe	70 – 0	1.0	1.0

Table 4-4. Results for biogeographic models tested, including likelihood scores (-lnL) and estimates of dispersal (D) and extinction (E) rates (events per Myr). In AARs for example nodes, areas shown indicate the range inherited by the daughter lineages. In cases where two ranges are separated by a bar, the first area is inherited by the upper branch on Figure 4, the second area is inherited by the lower branch. Relative probability of each AAR is shown in parentheses for multiple reconstructions. For DIVA results, all equally optimal reconstructions are shown, separated by commas.

Model	Constraints	-lnL	D	E	Examples of reconstructions for ancestral taxon of the following clades:						
					Whole family	Clade I	Sister clade to clade I	*Ailanthus*	Clade III	*Nothospondias*	Clade V
M1	Ancestral areas restricted (see text)	137.3	0.011	0.005	N (0.24) NA | N (0.23) A | NA (0.15) A | N (0.13)[a]	N (0.59) N | NA (0.17)	NA | A (0.29) A (0.29) A | NA (0.09) N (0.08)[a]	A (0.64) A | NA (0.13) N (0.09)	A (0.49) A | NA (0.31) A | N (0.09)	N (0.46) A (0.24)	N | F (0.47) N | A (0.35) NF | F (0.07)
M2	Ancestral areas unrestricted; dispersal probabilities stratified by age using D_1	133.2	0.023	0.000	NSFAU | N (0.06) SFAU | N (0.06) NSAU | N (0.03) SFAU | S (0.03)[a]	N (0.58) S (0.24)	F | NAE(0.10) F | NA (0.05) N | NA (0.03) NSFU | A (0.03)[a]	A | NA (0.45) A (0.18) A | NE (0.17) A | NAE (0.15)	A | NA (0.45) A | NE (0.17) A | NAE (0.15) F | NE (0.10)[a]	F (0.45) SFU | F(0.09) NSFU | F (0.07) NF | F (0.07)	NSF | F(0.35) F | FA (0.32) SF | F (0.16) S | FA (0.06)
M3	Full constraint (M1+M2); dispersal probabilities set to D_1	143.9	0.045	0.005	NA | N (0.31) N (0.28) A | NA (0.11) N | NA (0.07)[a]	N (0.65) N | NA (0.23)	NA | A (0.34) N | NA (0.21) N | A (0.10) A (0.09)[a]	A (0.51) A | NA (0.28) A | N (0.13)	A | NA	N (0.53) NF | F (0.18) F (0.10)	F | FA (0.63) N | A (0.14) NF | F (0.12)
M4	Full constraint (M1+M2); dispersal probabilities set to D_2	203.6	0.087	0.011	NA | N (0.34) N | NS (0.25) A | NA (0.13) N | NA (0.07)[a]	N (0.46) N | NA (0.24) NS | S (0.19)	NA | A (0.28) N | NA (0.23) N (0.13) A | NA (0.09)[a]	A | NA (0.44) A (0.36) E | NE (0.08)	A | NA (0.70) FA | A (0.18) E | NE (0.12)	NF | F (0.53) N (0.25) F (0.18)	N | A (0.44) F | FA (0.29) NF | F (0.22)
M5	Full constraint (M1+M2) with fossils excluded; dispersal probabilities set to D_1	120.5	0.065	0.006	A | NA (0.20) A | N (0.18) N (0.17) NA | N (0.15)[a]	N (0.56) N | NA (0.26)	A (0.22) NA | A (0.22) F | A (0.22) N | A (0.18)[a]	A	A | N (0.21) F (0.15) A (0.15) F | N (0.11)[a]	F (0.40) N (0.27) NF | F (0.11)	F | FA (0.64) NF | F (0.13) N | A (0.12)
DIVA 1	Unconstrained	-	-	-	NSFAUE	N, S, NA, SA, NSA	FAUE, NFAUE, SFAUE, NSFAUE	A	UE, NUE, FUE, NFUE, AUE, NAUE, FAUE, NFAUE	F, SF	F, SF
DIVA 2	Maxareas = 2	-	-	-	A, NA, SA	N, S, NA, SA	A	A	A	F, SF	F, SF

[a] Further alternative reconstructions (not shown) are within 2 log likelihood units of the optimal reconstruction.

Table 4-5. Ancestral range permutations for the chosen model (M3), ordered by decreasing likelihood. Likelihood scores (-lnL), and estimates of dispersal (D) and extinction (E) rates (events per Myr) are shown. For example nodes, AARs follow the format of Table 4-4; only the AAR with the highest relative probability is shown.

Ancestral ranges	-lnL	D	E	Examples of reconstructions for ancestral taxon of the following clades:						
				Whole family	Clade I	Sister clade to clade I	*Ailanthus*	Clade III	*Nothospondias*	Clade V
NA + NF	142.8	0.047	0.005	N (0.37)	N (0.74)	NA | A (0.36)	A (0.55)	A | NA	N (0.63)	F | FA (0.61)
NA + NF + SU	142.9	0.047	0.005	N (0.37)	N (0.74)	NA | A (0.36)	A (0.55)	A | NA	N (0.63)	F | FA (0.61)
NA + NE + NF	143.8	0.045	0.005	NA | N (0.31)	N (0.66)	NA | A (0.34)	A (0.51)	A | NA	N (0.53)	F | FA (0.63)
All possible [a]	143.9	0.045	0.005	NA | N (0.31)	N (0.65)	NA | A (0.34)	A (0.51)	A | NA	N (0.53)	F | FA (0.63)
NE + NF + SU	152.0	0.072	0.011	N | S (0.44)	S (0.48)	NE | E (0.25)	E (0.31)	E | NE (0.66)	N (0.53)	F | FA (0.48)
NE + NF	154.4	0.074	0.013	N (0.37)	N (0.58)	NE | E (0.22)	E (0.34)	E | NE (0.62)	N (0.47)	F | FA (0.44)
NA + NE	156.1	0.063	0.010	N (0.31)	N (0.63)	NA | A (0.23)	A (0.40)	A | NA	N (0.20)	N | A (0.25)
NA + NE + SU	156.2	0.063	0.010	N (0.31)	N (0.62)	NA | A (0.23)	A (0.40)	A | NA (0.70)	N (0.20)	N | A (0.25)
NA	158.8	0.056	0.008	N (0.37)	N (0.74)	NA | A (0.41)	A (0.59)	A | NA (0.83)	A (0.53)	NA | A (0.46)
NA + SU	158.9	0.055	0.008	N (0.37)	N (0.73)	NA | A (0.41)	A (0.59)	A | NA (0.83)	A (0.53)	NA | A (0.45)
NE + SU	161.4	0.092	0.016	N | S (0.39)	S (0.44)	NE | E (0.29)	E (0.33)	E | NE	E (0.34)	NE | E (0.49)
NE	163.3	0.095	0.018	N (0.44)	N (0.60)	NE | E (0.27)	E (0.35)	E | NE	E (0.34)	NE | E (0.50)
NF + SU	170.6	0.072	0.011	A | U (0.31)	S (0.28)	NF | F (0.38)	A | FA (0.29)	F | NF (0.53)	F (0.47)	F | FA (0.48)
NF	177.5	0.069	0.012	F (0.29)	F (0.35)	F | FA (0.26)	A (0.34)	F | NF (0.38)	F (0.63)	F | FA (0.49)
SU	206.0	0.121	0.022	S (0.24)	S (0.44)	U (0.48)	U (0.76)	U | SU (0.50)	U (0.51)	U | A

[a] Equivalent to M3 in Table 4-4.

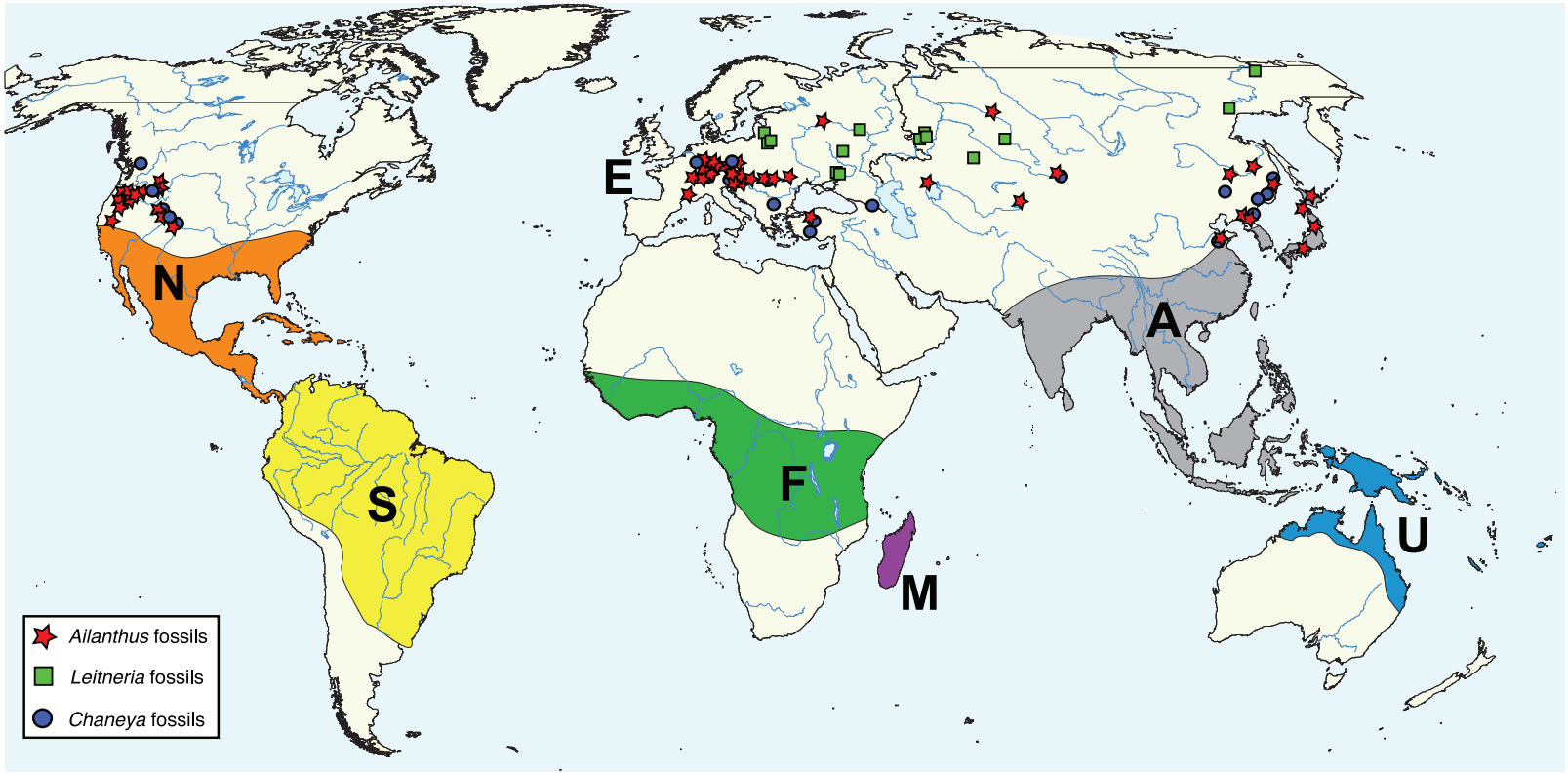


Figure 4-1. Map showing extant geographic distribution and approximate fossil localities for Simaroubaceae. Colors and letters refer to the seven areas assigned to terminals in biogeographic analyses and shown in Figure 4-4.

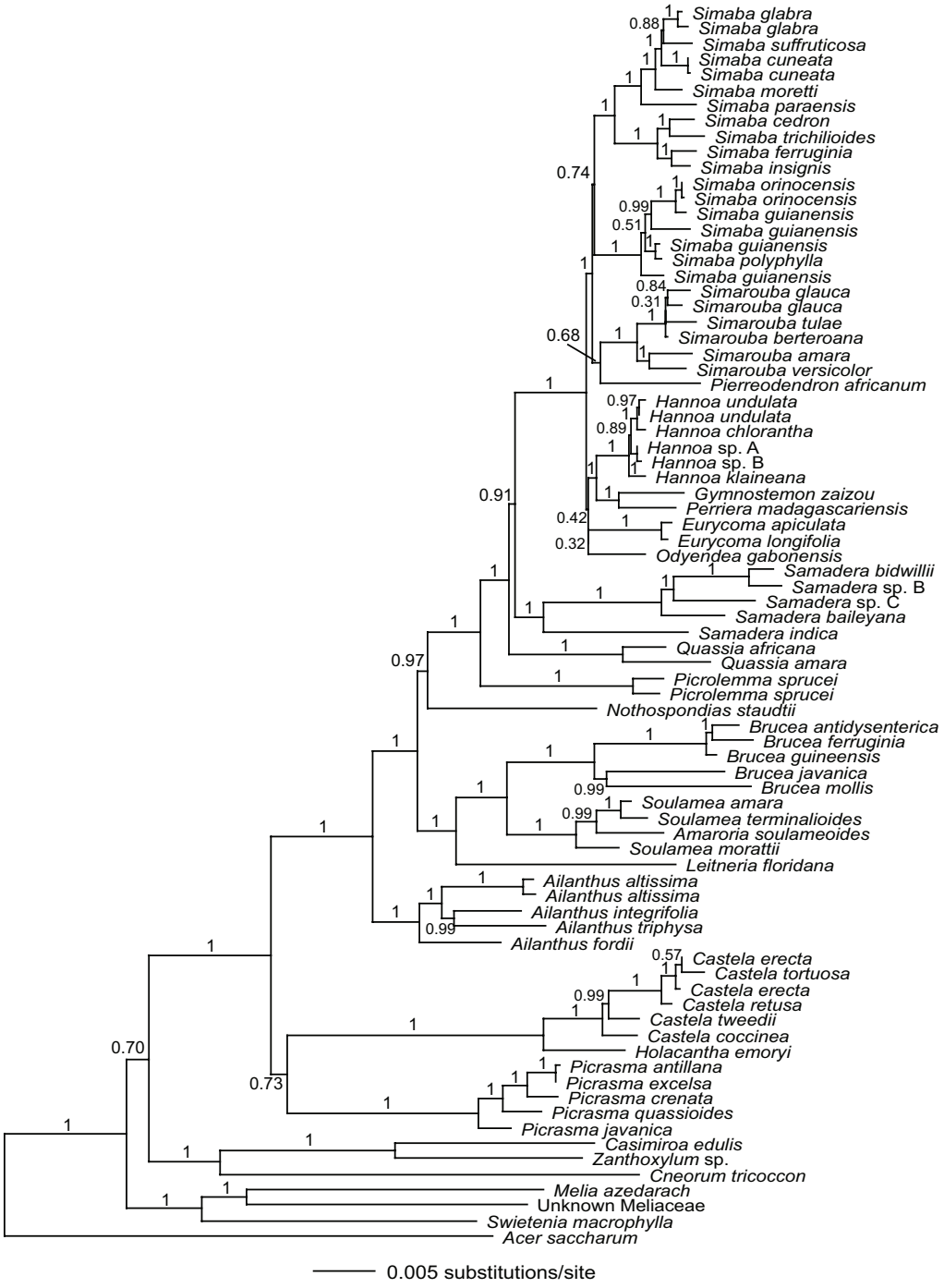


Figure 4-2. Phylogram of the majority rule consensus of trees for Simaroubaceae, sampled from the posterior distribution in the Bayesian analysis partitioned by gene and codon position (combination of two runs of 10,000,000 generations, sampled every 1000 generations; burn-in = 150,000 generations). Posterior probabilities (PP) are shown on branches. Branch lengths are proportional to the mean number of nucleotide substitutions per site, calculated from the posterior distribution density.

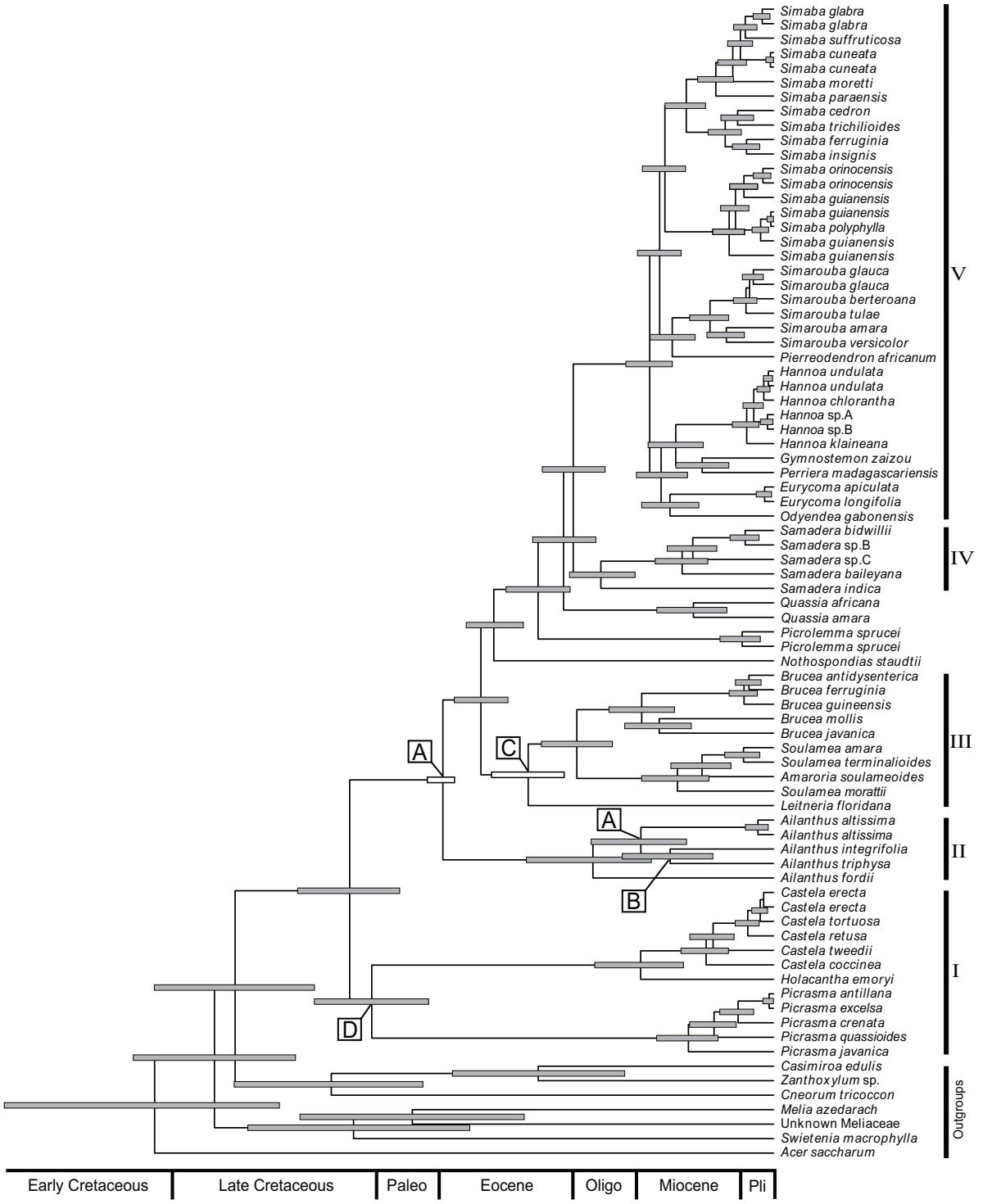


Figure 4-3. Chronogram with 95% HPD bars, based on BEAST analyses using two fossil calibrations, *Ailanthus confucii* and *Leitneria* sp., at nodes indicated with an open bar. Nodes labelled A, B, C and D refer to the positions of the four fossils used to assess different calibration schemes (Table 4-1). Geological time scale (Gradstein and Ogg, 2004) is shown at the bottom, and major clades of Simaroubaceae are indicated on the right hand side.

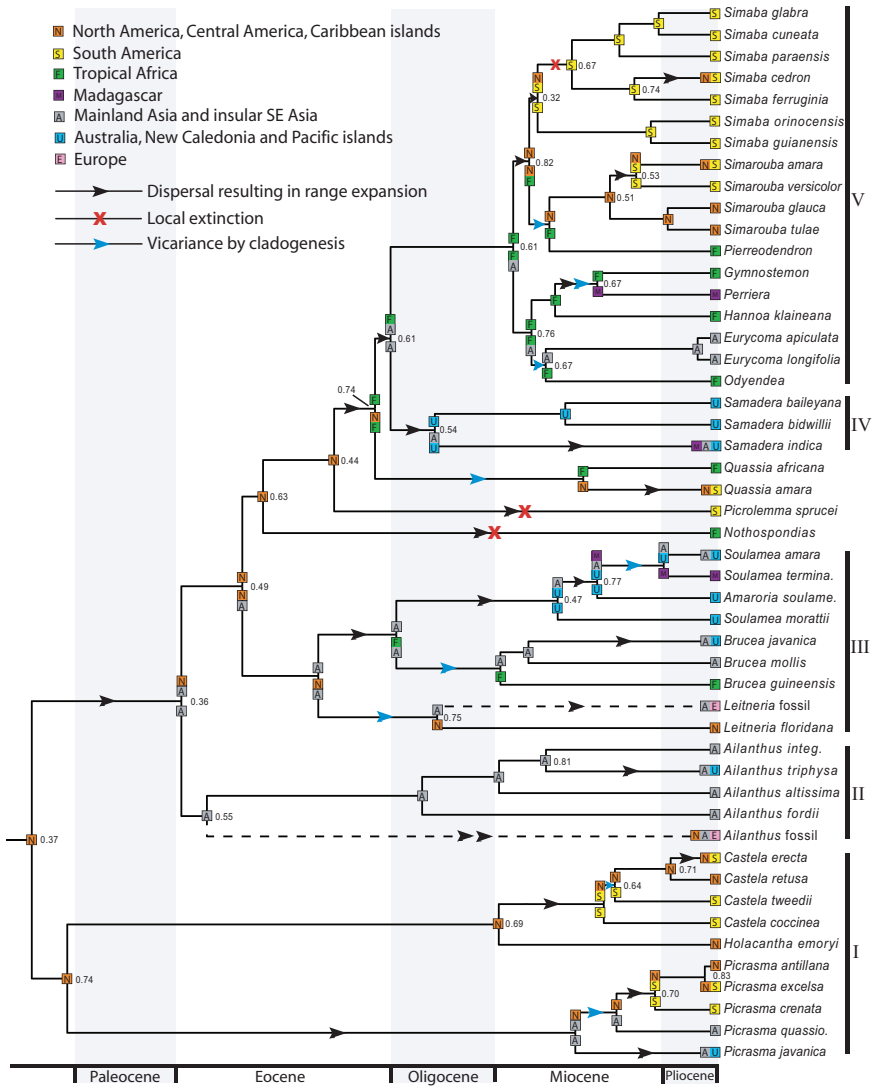


Figure 4-4. Ancestral area reconstruction for Simaroubaceae using model M3, with ancestral ranges NS, NF, NA, FM, FA, FE, AU, AE and MAU, and stratified dispersal probabilities between areas (D_1). AARs with highest likelihood are shown as colored boxes at each node. Single area boxes indicate an ancestor confined to a single geographic area; combined boxes indicate an ancestor with a distribution encompassing two or more areas; two boxes separated by a space indicate the ancestral ranges inherited by each of the daughter lineages arising from the node. For nodes with alternative reconstructions (within 2 log likelihood units of the maximum), the relative probability of the global likelihood for the optimal reconstruction is given. Three modes of range inheritance (range expansion, local extinction and vicariance by cladogenesis) are indicated as symbols on branches in the phylogeny (see Results). For clarity, selected sister terminals from the same single area have been pruned from the chronogram. Fossil lineages are shown with a dotted line. Geological time scale (Gradstein and Ogg, 2004) is shown at the bottom, and major clades of Simaroubaceae are indicated on the right hand side.

CHAPTER 5
DIVERSIFICATION AND MORPHOLOGICAL EVOLUTION IN SIMAROUBACEAE

Introduction

Wiens and Donoghue (2004) suggest there is a large void between ecology and historical biogeography in the interpretation of biodiversity gradients. Although studies encompassing major biomes (Cracraft, 1985; Dick and Wright, 2005; Pennington and Dick, 2004; Pennington et al. 2004a; 2004b; 2006; Mittlebach et al., 2007) provide an overview, the origins of biodiversity cannot be explored without detailed studies of clades with well-established phylogenetic relationships. The underlying causes of biodiversity are often a complex interaction of factors (de Queiroz, 2002; Lavin et al., 2004; Wiens and Donoghue, 2004), and so pinpointing specific determinants may be best suited to a data exploration approach. In the context of a clade such as Simaroubaceae, this can be done by detecting shifts in species diversification rates within a phylogenetic framework, and examining biological correlates to these shifts. Diversification rate shifts, that is, net changes in rates of speciation or extinction between clades in a phylogenetic tree, can be considered both in a temporal (Nee et al., 1992; Pybus and Harvey, 2000; Magallón and Sanderson, 2001) and a topological (Chan and Moore, 2002) context. The correlates of species diversity can be historical, ecological and intrinsic to the organism (Cracraft, 1985; Herrera, 1989; Eriksson and Bremer, 1992; Dodd et al., 1999; Lavin et al., 2004; Pennington and Dick, 2004; Pennington et al. 2004a; 2004b; 2006; Dick and Wright, 2005; Wiens and Donoghue, 2004), and for the present study this translates to two identifiable variables: dispersal into new environments (and subsequent adaptive radiation across environmental gradients), and key morphological innovation.

Reconstruction of the biogeographic history of Simaroubaceae (Chapter 4) has allowed the identification of long-distance, island-hopping and overland dispersal events ancestrally within

specific clades. Dispersal events are significant because they drive cladogenesis, and the isolation of one of the daughter lineages in the newly colonized environment may or may not promote diversification. However, ecological adaptation is manifest in morphological innovation, and even subtle changes in local environmental factors, rather than a major geographical isolation event, can cause niche specialization and promote speciation (Funk et al., 2006), for example, by diversifying floral form to take advantage of pollinator diversity, resulting in reproductive isolation (Hodges and Arnold, 1995; Waser, 1998; Kay et al., 2005; von Hagen and Kadereit, 2003). Thus we might expect to find links to diversification rate shifts between either morphological characters considered important to niche specialization, such as reproductive morphology, or a historical dispersal event, or both, for a given lineage in which a diversification rate shift is detected. Historical dispersal events and environmental variation have been linked to patterns of diversification in studies of specific clades (Richardson et al., 2001a, 2001b; von Hagen and Kadereit, 2003; Becerra, 2005; Moore and Donoghue, 2007; Alfaro et al., 2007; Wiens et al., 2007), and even experimentally in a controlled environment (Fukami et al., 2007). Similarly, morphological correlates to species diversity have been examined for a number of angiosperm groups (Herrera, 1989; Eriksson and Bremer, 1991; Eriksson and Bremer, 1992; Dodd et al., 1999; Good-Avila et al., 2006), a commonly cited example being the evolution of nectar spurs (Hodges and Arnold, 1995; Ree, 2005). Testing hypotheses of key innovations has only recently been undertaken with a more critical approach in a phylogenetic framework (de Queiroz, 2002; Ree, 2005; Leschen and Buckley, 2007; Moore and Donoghue, 2007), to attempt to overcome the problems associated with simple measures of clade diversity through species numbers. When parsing out the relative effects of the two major factors hypothesized to be involved in diversification rate increases, there is a subsequent issue of contingency (de Queiroz,

2002; Davies et al., 2004), for example, the contribution of innovative dispersal mechanisms to future chances of dispersal (Moore and Donoghue, 2007), or rapid morphological diversification of the daughter population from the parent population in order to take advantage of new ecological opportunities (von Hagen and Kadereit, 2003; Moore and Donoghue, 2007).

This study will take a data exploration approach to address the following questions:

- Does the phylogenetic history of Simaroubaceae show significant shifts in diversification rates between lineages?
- Do any significant diversification rate shifts correlate to morphological character-state changes, historical biogeographic movements, both of these factors, or neither?

As a necessary precursor to answering these questions, an analysis of morphological variation and evolution in Simaroubaceae was undertaken. Morphological data extracted from the literature and herbarium material was firstly used to determine if any phylogenetic signal is present for the selected characters. Particular consideration was given to coding of the characters. These were compared to the molecular phylogeny for the family (Chapter 3) to see which clades are supported by morphological data. A second analysis was performed on a combined matrix of molecular and morphological data. The purpose of this was to see if the morphological data have any impact on the molecular phylogeny, in changing topology and support for clades (Pennington, 1996; Nylander et al., 2004). It also serves to determine the phylogenetic placement of numerous taxa for which no molecular data have been obtained, particularly the unsampled monotypic genera *Laumoniera* and *Iridosma*. Potential morphological synapomorphies for major clades of Simaroubaceae were also identified. Selected morphological traits are used to reconstruct character evolution in Simaroubaceae, to examine the lability of various traits deemed of putative evolutionary significance, and their pattern of inheritance. A number of different techniques for diversification rate analysis have been applied to the previous molecular phylogeny, and to phylogenetic data obtained in this study. Ultimately, any diversification rate

shifts are examined in the context of character-state transitions and historical biogeographical events (Chapter 4).

Methods

Phylogeny Estimation and Character Evolution

Morphological and anatomical characters were recorded from species descriptions in the literature (Trelease, 1895; Engler, 1931; Aubréville and Pellegrin, 1937; Abbe and Earle, 1940; Cronquist, 1944a,b,c,d; Perrier de la Bathie, 1950; Hutchinson and Dalziel, 1954; Gilbert, 1958; Aubréville, 1962; Nooteboom, 1962a, 1962b, 1987; Basak, 1963, 1967; Wild and Phipps, 1963; Backer and Van den Brink, 1965; Porter, 1971, 1973; Li, 1977; Cavalcante, 1983; Feuillett, 1983; Thomas, 1984; Hewson, 1985; Smith, 1985; Thomas, 1985; Moncada and Machado, 1987; Fernando and Quinn, 1992; Pirani, 1987, 1997; Franceschinelli and Yamamoto, 1999; Franceschinelli et al., 1999; Franceschinelli and Thomas, 2000; Stannard, 2000; Hahn and Thomas, 2001; Jaffre and Fambart, 2002; Thomas, 2002) and 120 herbarium sheets were examined (see Appendix B and C) macroscopically and with a binocular microscope. The tabulated characters were then converted into a data matrix of coded characters in MacClade v.4.08 (Maddison and Maddison, 2004). The following criteria were used for coding characters in discrete states for maximum parsimony and Bayesian analysis:

- Species were considered the operational taxonomic units (OTUs) in the matrix. Character data were sythesized from multiple species in the literature and herbarium specimens. For continuous characters the total variation was used (maximum and minimum values), and variation in meristic and discrete characters was coded as ambiguous. See Appendix D for details.
- If more than 11 of the 22 genera had missing data for a character, it was not used.
- Continuous characters were only included if two or more non-overlapping character states could be assigned. Non-overlapping was defined as any gap between the highest and lowest measurements of the character. See Appendix D for details.
- All characters were coded as unordered.

- All characters had equal weight.
- Character states were coded as missing, rather than absent, if they are dependent on a second character in the matrix which is also absent, to avoid non-independence e.g. petal characters in *Leitneria*, which has no petals.

Morphological characters were not coded for outgroups, because of the difficulty involved in selecting suitable outgroup taxa. Both Rutaceae and Meliaceae are families with a large amount of morphological variation, even more so than Simaroubaceae, and the position of the root of Simaroubaceae would be biased towards whichever taxa were selected as outgroups. Reconstruction of an ancestral Rutaceae or Meliaceae taxon, that encompassed the ancestral traits of these families, would be the most suitable outgroup in this case, but these data are unavailable. Therefore, for morphological analyses, *Picrasma* and *Castela+Holacantha* were together designated as a functional outgroup, based on their position as sister to the rest of the family in previous molecular analyses (Chapter 3), and each genus was also tested individually as an outgroup.

A maximum parsimony (MP) analysis was conducted on the morphological data alone for 105 taxa. The analysis was performed in PAUP* 4.0b10 (Swofford, 2002), using a heuristic search of 100 random addition replicates with TBR branch swapping, MulTrees in effect, and saving all trees, with a maxtrees limit set to 1000 trees for each replicate. A second heuristic search was made using the shortest trees resulting from the first search, with a maxtrees of 100,000. ACCTRAN character optimization was used. A strict consensus of most-parsimonious trees was made, and support for clades was assessed by bootstrapping (Felsenstein, 1985), using 1000 bootstrap replicates with one random addition per replicate, maxtrees set to 1000 for each replicate, TBR branch swapping, MulTrees in effect. A Bayesian analysis of morphological data was also performed under the Mk1 model of character evolution (Lewis, 2001), that is, one symmetrical transition rate between character states, with a gamma rate distribution. The analysis

was implemented in MrBayes v3.1.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Two independent analyses each ran for 10 million generations using four Markov chains and with all other parameters at default values; trees were sampled every 1000th generation with a burn-in of 100,000 generations. Coding bias was set to variable for the morphological data, to account for the fact that no constant characters were included in the morphological matrix. Stationarity of the MCMC was determined by the average standard deviation of split frequencies between runs and by examination of the distribution of the posterior in Tracer v1.4 (Rambaut and Drummond, 2003). A majority-rule consensus tree was created using the posterior distribution of trees. The resulting estimations of phylogeny for morphology alone were compared to the molecular phylogeny for the family (Chapter 3) to see which clades, if any, were recovered by the morphological data.

A second MP analysis was performed on a combined matrix of molecular and morphological data, using the same heuristic search criteria as the morphology alone, except with DELTRAN character optimization. Bootstrap support was assessed as above. A combined Bayesian analysis was also performed with the molecular data partitioned by codon position and gene region using models of nucleotide evolution determined using the AIC in Modeltest (Posada and Crandall, 1998; see Chapter 3). Morphological data were included as a separate partition, modeled under the Mk1 model of character evolution (Lewis, 2001). Analyses were performed in MrBayes v3.1.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Two independent analyses each ran for 40 million generations using four Markov chains and with all other parameters at default values; trees were sampled every 1000th generation with a burn-in of 10 million generations; model parameters were unlinked across partitions, and the prior for the rate parameter was allowed to vary between partitions. Coding bias was fixed for

molecular data and set to variable for the morphological data. Stationarity of the MCMC was determined as above.

Character evolution was examined using parsimony (Swofford and Maddison, 1987). Morphological characters were not investigated in these analyses if they were unique to fewer than four taxa, or were potentially subject to phenotypic variation. Also excluded were characters with significant amounts of missing data, and those in which coding was deemed too subjective, especially continuous quantitative traits. Characters of particular interest were those that might be considered evolutionarily significant (e.g. Herrera, 1989; Eriksson and Bremer, 1991, 1992; Waser, 1998; Dodd et al., 1999; Moore and Donoghue, 2007), and characters that have evolved multiple times across the phylogeny. We therefore focused on reproductive/floral traits, dispersal biology and habit. A total of 22 characters were examined, plus four additional characters not used in phylogeny reconstruction: pollination syndrome (insect vs. bird vs. wind), mode of dispersal (biotic vs. abiotic) and general ecology, which was broken up into two characters – temperate vs. subtropical/tropical and arid vs. mesophytic vs. swamp/riverine/coastal. Parsimony mapping of characters was performed in MacClade v.4.08 (Maddison and Maddison, 2004). Characters were mapped on a phylogeny that was randomly chosen from most-parsimonious tree topologies of combined molecular and morphological data that were compatible with a 70% majority-rule consensus and clades having PP > 0.95 and BS > 70%. The outgroups were removed from this tree, along with four taxa that appeared sister to the rest of their respective genera due to a lack of morphological data: *Picrasma selleana*, *Castela jacquinifolia*, *Simarouba laevis* and *Soulamea fraxinifolia*. *Simaba cavalcantei* was also removed as it appeared sister to *Samadera*, likely due to a lack of informative characters. Although some have suggested circularity in mapping characters on a phylogeny which has been constructed using these

characters (Coddington, 1988; Armbruster, 1992), the large amount of molecular data should limit the effects of non-independence (Luckow and Bruneau, 1997). Furthermore, the full data set provides the opportunity to use over 100 species. Characters were mapped using the MPR function in MacClade, and minimum number of unambiguous state changes, and number of most-parsimonious reconstructions were calculated.

Shifts in Diversification Rates

Temporal methods for detecting shifts in diversification rate (*r*) are numerous, and two of the most commonly used statistics were chosen. The first is the relative cladogenesis statistic (Nee et al., 1992), which is calculated as the likelihood that at any given time slice across the phylogeny, each lineage at that time slice will give rise to a number of species, summing to the total number of species, such that each lineage has diversified at the same stochastic rate. Lineages deviating from this null expectation are therefore considered significantly more diverse than expected. The second statistic follows methods employed by Magallón and Sanderson (2001) to calculate *r*, which are less sensitive to taxon sampling, because calculations of diversification rate use only clade diversity and time since origination of the stem or crown lineage. We calculated *r* for major clades of Simaroubaceae using both crown and stem lineages. Confidence limits are set as a function of the background diversification rate of the family as a whole, and clades that fall outside the 95% confidence interval are considered hyperdiverse or significantly depauperate in species. The method also requires the user to specify an extinction rate, and therefore a range of values from 0 to 0.9 was tested (Alfaro et al., 2007). Taxon sampling is important for this method with respect to measuring rates for stem and crown lineages, because using the crown group assumes the sampling is sufficient to encompass the most inclusive node for the clade. The crown group method was used for most clades in Simaroubaceae because most unsampled taxa are limited to the three large genera, *Soulamea*,

Castela and *Simaba*. In *Simaba* the deepest node has probably been sampled because the genus consists of two sister clades, which are morphologically distinct, and members of both clades have been sampled in the phylogeny. In the case of *Castela* and *Soulamea*, it may be unreasonable to assume the deepest node has been sampled, and in this case we also calculated diversification rates with the stem group method, along with a rate for the New Caledonian *Soulamea* (for which only a single species was sampled). The stem group method also had to be applied to monotypic genera *Leitneria*, *Nothospondias* and *Odyendea*, which have no crown group, and *Picrolemma* and *Holacantha*, which have two species but only one was sampled. All temporal methods were performed in R using the laser package (Rabosky, 2006) and the geiger package (Harmon et al., 2008). For both the relative cladogenesis statistic and the method of Magallón and Sanderson, the most parsimonious interpretation is a shift in diversification rate implied at the node ancestral to the root of the significantly more diverse clade (Purvis et al., 1995).

A topological method was also used for detecting shifts in diversification rate, as this would allow the use of the combined molecular and morphological data set, and does not suffer the trickle down effect seen in the temporal methods (Moore et al., 2004). Trees resulting from the combined data had much better taxon sampling than molecular data only, but could not be included in temporal methods described above because divergence dates are estimated using nucleotide substitutions only. The topological method relies on tree shape, and detects nodes that are significantly imbalanced using a shift statistic (Δ, Chan and Moore, 2002; Moore et al., 2004; Moore and Donoghue, 2007). The shift statistic (Δ) is calculated for all nodes, using the difference between likelihood ratios at each node. Nodes are treated as three-taxon trees, and a likelihood ratio is calculated for the ingroup-outgroup diversity partition, and the diversity

partition between the two ingroup clades. The likelihood ratio is a ratio between a one-rate model, where both clades are expected to have the same branching rate, and a two-rate model, where both clades are expected to have different branching rates (see Chan and Moore, 2002; Moore et al., 2004; Moore and Donoghue, 2007). The program SymmeTREE (Chan and Moore, 2005) implements the topological method (Chan and Moore, 2002), and returns shift statistics and their signifcance level for whole tree imbalance and for individual nodes. The software uses a tree as the only input, and so the tree used in parsimony character mapping was used.

Correlates of Shifts in Diversification Rate

Both biogeography and morphological evolution were considered for correlation to diversification rate shifts. Biogeographical shifts are taken directly from analyses performed in Chapter 4. All nodes after which a range expansion or vicariant event occurred were used as points of biogeographic shift (see Chapter 4). Most of these range shifts were categorized as long-distance dispersal events, with few considered overland migrations. These nodes are identified in Figure 5-7. The biogeographic shifts were compared visually to diversification rate shifts determined in analyses described earlier. Shifts along terminal lineages were excluded (e.g. range expansion in *Samadera indica* or *Brucea javanica*), as no shift in diversification rate can be inferred. Morphological traits used to examine character evolution were also tested as potential key innovations. The method of Ree (2005) was used to test whether the morphological traits used for character mapping covary with diversification rate. This method involves calculating a test statistic by mapping characters on the Simaroubaceae chronogram using stochastic mapping under a continuous-time Markov model (Huelsenbeck et al., 2003). Bayesian stochastic mapping allows for multiple state changes along branches, thus better dealing with long branches than parsimony methods (Cunningham et al., 1998; Huelsenbeck et al., 2003). A diversification rate is calculated simply as the number of historical branching events with one

state of the character divided by the time spent in that state (based on the stochastic mapping of the character). The difference between the rates for one state (the putative key innovation) versus the second state (the ancestral state) is then calculated. This process is repeated for a number of mappings and the mean difference in diversification rates between the two states is used as the observed test statistic. For the null distribution, Monte Carlo simulation is used to generate any number of trees with the same topology as the Simaroubaceae chronogram but with branch lengths generated under a pure birth process. For each of these trees, the test statistic is generated as it was for the Simaroubaceae chronogram, and the resulting values constitute the null distribution, against which the observed test statistic can be compared to determine its posterior predictive p value (Ree et al., 2005; Moore and Donoghue, 2007). The process is implemented in keyinnotest (http://www.phylodiversity.net/rree/software.html). The characters chosen were recoded as binary traits if they were not already, with multistate characters being broken down into presence/absence of a particular state. Currently keyinnotest does not accommodate missing data, so missing states were recoded based on most-parsimonious reconstructions determined in MacClade. Given the temporal component of the analysis, only taxa present in the molecular rates data set could be included (Chapter 4). To generate the observed test statistic for each of the 36 binary character traits, 100 stochastic mappings were performed on the Simaroubaceae chronogram. For each character the tree length was scaled to the minimum number of character-state changes needed to observe the distribution of states at the terminals (Ree, 2005). The null distribution was generated with 1000 pure birth trees, and 10 stochastic mappings on each simulated tree.

Results

Phylogeny Estimation and Character Evolution

The 71 morphological characters used for phylogeny estimation are provided in Appendix D. Summary statistics for both morphology only and combined MP analyses are shown in Table 5-1. A strict consensus for morphological data only is shown in Figure 5-1, with clade support shown on branches at cutoffs of 70% BS and 0.95 PP. Several genera and some major clades of Simaroubaceae were resolved by morphological data, but backbone structure was poorly supported. For combined morphological and molecular data, a strict consensus is shown (Fig. 5-2), with bootstrap support and Bayesian posterior probabilities. The topology was similar to the previous molecular phylogeny (Chapter 3); however, support within genera and for some backbone nodes was greatly reduced. The previously unplaced genus *Laumoniera* was sister to *Nothospondias* in the analysis of morphology only, but its position was unresolved in the combined analysis. *Iridosma* was found in a clade with *Gymnostemon* and *Perriera* in the combined analyses, but was unresolved within Clade V for morphology only. Table 5-2 shows morphological synapomorphies for the major clades in Simaroubaceae, including taxa in which the characters have reversed or evolved in parallel elsewhere. Table 5-3 shows details for MPR for the 26 characters used for character mapping, including the number of unambiguous character-state changes required to explain the data, the number of most-parsimonious reconstructions for each character, and ancestral states for major clades of Simaroubaceae.

Diversification Analyses

The relative cladogenesis statistic (Nee et al., 1992) identified six nodes in the phylogeny with significant shifts in diversification rate. All but one of these are attributed to the 'trickle-down' effect, where a particularly diverse clade causes the successively deeper clades in which it is nested to appear more diverse than expected (Moore et al., 2004). In this case, the shallowest

clade is assumed to be the clade with significantly elevated levels of diversity, which in Simaroubaceae is the New World subclade of Clade V (plus *Pierreodendron*).

The diversification rates calculated using methods of Magallón and Sanderson (2001) are illustrated in Figures 5-5 and 5-6, with 95% confidence interval boundary lines for three different extinction rates (E). For crown groups (Fig. 5-5), no clades fall outside the 95% bounds when E = 0.9. Eight crown clades are significantly diverse for E = 0 and three clades for E = 0.5. Of the eight clades, six are equivalent to the shifts detected by the relative cladogenesis statistic, and thus five of these are eliminated under the assumption of the 'trickle-down' effect (Moore et al., 2004), leaving *Simaba* as the least inclusive clade. The other two crown clades above the 95% bound for E = 0 were *Soulamea* and *Castela*, although these assume the true crown node was sampled for each, which may not be the case. When these two clades were sampled as stem groups, only *Castela* fell outside the 95% CI for E = 0. At the lower 95% bound, *Quassia, Perriera*+*Gymnostemon*, and Asian *Brucea* had significantly fewer species than expected for E = 0. For the stem groups (Figure 5-6) only one clade, New Caledonian *Soulamea*, was significantly more diverse than expected, for E = 0 and E = 0.5. Monotypic genera *Leitneria*, *Nothospondias* and *Odyendea* fall below the lower 95% bound for all three extinction rates, but it is difficult to evaluate diversification in clades of a single species (Wiens et al., 2007).

Results from SymmeTREE analyses (Chan and Moore, 2005) showed only one node to have a significant shift in diversification rates, the root node of Clade III, that is, the divergence of *Leitneria* from *Brucea* and *Soulamea*. One other node was considered near-significant (p=0.11), the root node of the family, that is, an imbalance between Clade I and the rest of the Simaroubaceae. SymmeTREE also detected significant variation in diversification rates across the whole tree ($p < 0.01$). Nodes at which a diversification rate shift was detected, using either

temporal or topological methods described above, are indicated in Figure 5-7. Also shown in Fig. 5-7 are nodes at which biogeographical shifts are proposed to have occurred. Results for analysis of key innovations are shown in Table 5-4. Of the 36 binary characters tested, four showed significant correlation, with $p < 0.05$, but none at $p < 0.01$.

Discussion

Morphological data exhibited high levels of homoplasy, but also showed a number of characters were phylogenetically informative, more so at the generic level than deeper in the tree. In both MP and Bayesian analyses, clade support was generally low; however, the MP strict consensus was resolved for several genera, and bootstrap support (> 70%) and posterior probabilities (> 0.95) supported most of this generic resolution (Fig. 5-1). Genera of Simaroubaceae are well-supported because each has unique characteristics, for example, the winged and jointed rachis of *Quassia*, the trimerous flowers and dry, obcordate fruits of *Soulamea*, or the cymose inflorescence and globose fruits of *Picrasma*. Morphology also provided some resolution for larger clades found in molecular analyses (Chapter 3), such as a clade of *Picrolemma* plus its sister group. The sister relationship of *Picrasma* and *Castela+Holacantha* to the rest of the family could be an artifact of the lack of outgroup taxa (see Methods). However, regardless of which of the two genera are selected to be the outgroup (based on results of molecular data), the other is always the next-diverging clade in the family, supporting their early divergence in the family. Without outgroup taxa it is impossible to say if their forming a clade is supported by morphology, or if morphology supports the *Picrasma* lineage diverging before the *Castela* lineage or vice versa. There are no synapomorphies for a clade of *Picrasma*, *Castela* and *Holacantha* that could not alternatively be interpreted as pleisomorphies in the absence of outgroup taxa, e.g. the absence of leaf glands; flowers with four carpels. Furthermore, both *Picrasma* and *Castela+Holacantha* have synapomorphies unique to

the family, that could be shared with outgroup taxa, e.g. cymes in *Picrasma*; spines in *Castela* and *Holacantha*.

The combined analyses of molecular and morphological data produced a phylogeny (Fig. 5-2) similar to the previous molecular phylogeny (Chapter 3). The introduction of taxa without molecular data reduced support within genera substantially, but only significantly impacted the topology of Clade V. The backbone remains poorly supported, and based on analyses of morphology alone, there is a considerable amount of homoplasy in the data (CI = 0.29). Of particular note in both combined and morphology-only analyses is the genera *Iridosma* (Aubréville, 1962) and *Laumoniera* (Nooteboom, 1987). *Iridosma* appears closely related to its African counterparts in Clade V; however, herbarium material was unavailable for examination, so a phylogenetic position with improved support could be found with more data. *Laumoniera* was initially described as closely related to *Brucea* (Nooteboom, 1987), as they shared tetramerous flowers, a narrow thyrsoid inflorescence and lack of a filament appendage, and both are found in SE Asia. However, *Laumoniera* being sister to *Nothospondias* is hypothesized under critical analysis of morphological data. *Nothospondias*, although found only in Africa, also shares tetramerous flowers, a narrow thyrse and an absence of a filament appendage with *Laumoniera*, and additionally has large, fleshy, yellow, ellipsoid fruits, similar to those of *Laumoniera*, and quite unlike the small, ovoid, scarcely fleshy drupes of *Brucea*. It is also curious that in his description of *Laumoniera*, Nooteboom did not mention any prominent submarginal glands on the undersurface of the leaflets, glands that are characteristic of *Brucea*, and which *Nothospondias* lacks. Again, no herbarium material of *Laumoniera* was available for examination, but based on these data, *Laumoniera* is more closely related to *Nothospondias* than *Brucea*. More data will determine whether differences between the two are sufficient to maintain

generic status, or whether *Laumoniera* should be placed in *Nothospondias* under the new combination *Nothospondias bruceadelpha* (Noot.) J.W.Clayton.

Analyses of character evolution reveal a complex history of trait evolution, particularly for a family with only ca. 100 species, although character evolution data from related families, for comparison, are scarce. A high rate of transition, and high numbers of most-parsimonious reconstructions, are seen between states for quantitative characters such as size and color of flowers and fruits, and degree of corolla pubescence. This is partly due to the number of states assigned to these characters, and the subjectivity of coding such characters, with some taxa showing conditions close to the boundaries of the different states. It may also be due to the influence of the environment, and local adaptation to shifts in pollinators, dispersers and the immediate surroundings that support a particular growth form. Certain reproductive characters also show lability, such as the types of inflorescence, the number of stamens and the number of carpels. For example, the ancestor of the family is reconstructed to have had four carpels, but in Clade III there are examples of reduction to one, two and three carpels. Characters which show a low amount of change and few most-parsimonious reconstructions are strong synapomorphies for defining some of the major clades in the phylogeny. The presence of a filament appendage was traditionally used to delimit tribes (Engler, 1931), and it is still upheld as a good synapomorphy for all taxa diverging after *Picrolemma*, with a single reversion in *Perriera* and *Gymnostemon*. Punctate glands have received little attention in the literature, except in individual species descriptions, and when Aubréville and Pellegrin (1937) suggested a relationship between *Perriera* and *Gymnostemon* on the basis of punctate glands. They are, however, only present in all taxa diverging after *Picrolemma*, except for two apparent losses. Clade III can be defined as having reduced flowers (if we also code *Leitneria*'s absent perianth as "reduced"), and this clade

also shows a tendency towards having spiciform inflorescences. Ancestral states for the family correspond to ancestral states for the Sapindales (e.g. pinnately compound leaves), but interestingly, Sapindales are hypothesized to have pentamerous flowers ancestrally (Ronse de Craene, 2008), depending on the rosid topology, whereas Simaroubaceae are ancestrally tetramerous.

Diversification rate shifts in the family are few and dependent on the methods used, but the family as a whole is suggested to have significant nodal imbalance based on tree shape alone. The diversification rate has increased in *Castela*, a clade of *Brucea*, *Soulamea* and *Amaroria*, New Caledonian *Soulamea*, and either the New World members of Clade V or *Simaba* (depending on the method). Although the strength of the Magallón and Sanderson (2001) method is its applicability to clades with unsampled taxa, the reliance on species diversity may fail to detect rate shifts in clades that show a recent upturn in diversification rate, which has yet to manifest itself in increased clade diversity (Ree, 2005), or in clades where a rapid increase in diversification rate is followed by a period of stasis. Methods based on clade diversity, measured in species numbers, are phylogenetically non-independent, unlike methods based on tree shape (Chan and Moore, 2002; Moore et al., 2004), which look for rate shifts relative to parent and daughter lineages. This may not be a problem when species-rich clades appear recently, such as those detected here, but it is the cause of the 'trickle-down' effect (Moore et al., 2004) seen in the elevated diversity of more inclusive clades containing *Simaba*.

The current study found four significant correlations between diversification rate shifts and putative key innovations, that is, presence of thorns, evolution of haplostemony from diplostemony, having a short style, and movement into a temperate habitat. There is always potential for other morphological characters to impact diversification, particularly cryptic

characters for which the ecological impacts are poorly understood. The significance and near-significance of some of the characters is, however, questionable based on visual inspection of the tree, and may be an artifact of insufficient stochastic mappings, insufficient sampling to generate the null distribution, or imbalance in character-state frequencies. In all significant cases, simulated trees create a very narrow null distribution, and the significance of thorns may be an artifact of the long branch leading to *Castela*, which would result in a wide variance around the mean time spent having thorns, leading to an anomalous result. Regardless, none of the characters were significant at $p < 0.01$, whereas in previous examples where this method has shown significance (nectar spurs in *Aquilegia*: Ree, 2005; stamen number in Valerianaceae: Moore and Donoghue, 2007), the predictive posterior p value was < 0.001. A conservative approach is taken in this study, given uncertainty around the sampling, and no significant key innovations are proposed. As stated by Moore and Donoghue (2007), the statistical properties of this test need further exploration. Normalizing the null distribution for different states depending on their frequency of occurrence may be necessary (Ree, 2005). Furthermore, these types of studies have an important caveat. In order for character-state transitions to covary with rate shifts with true statistical significance, a larger sample size would be needed, with repeated examples of covariance across a number of more distantly related clades (Ree, 2005). In most cases, a single character is shown to correlate to a single rate shift (Ree, 2005; Moore et al., 2006; Leschen and Buckley, 2007; Moore and Donoghue, 2007; Wiens et al., 2007), resulting in a single data point, although it could be argued that highly homoplasious characters may represent more data points (Ree, 2005). Without a certain level of statistical confidence, the conclusions drawn from such a study must account for the possibility of random chance producing the results seen, and the myriad of interactions of key innovations with other organismal traits and the

environment (de Queiroz, 2002). Very few studies such as these have been undertaken (Moore et al., 2006; Leschen and Buckley, 2007; Moore and Donoghue, 2007), and Moore and Donoghue found biogeographical movements to be most commonly correlated with diversification rate increases rather than putative key innovations. Looking at morphological characters across a broader phylogenetic spectrum, for example, within the Sapindales, would reveal more about diversification in the Simaroubaceae, and stochastic effects.

Biogeographical movements appear connected to increased species diversity in Simaroubaceae, but with the number of dispersal events occurring in the recent history of the family, it is difficult to parse out significant events if they are surrounded by events showing no apparent connection to diversification. In Simaroubaceae diversification is linked in two instances to dispersal to and movement within the New World, one instance to arrival on New Caledonia, and one instance to movement into Africa from Asia. There are, however, clades which also show movement between North and South America and no accompanying significant increase in diversification (*Picrasma, Picrolemma* and *Simarouba*). Because exploring historical biogeography on a global scale limits the "state transitions" that are examined, these patterns can be compared to other studies in which increased diversification was also associated, especially with movement within the New World (Moore and Donoghue, 2007; Hughes and Eastwood, 2006; Wiens et al., 2007), but this is typically coupled with movement across an environmental gradient. This is contrary to statistical tests for key innovations, of which few have been undertaken, and which have potentially hundreds of morphological state transitions to examine. In the case of *Castela*, the increased diversity of the clade correlates with a movement of the ancestor into South America. However, repeated dispersal between arid environments of North and South America (and to the Galápagos Islands) is more likely to be the driver of this

increased speciation rate (Lavin et al., 2004). If the initial adaptation to arid environments of North America in the Oligocene and Miocene were the cause of an adaptive radiation, we would expect to see significantly increased speciation rates including *Holacantha*, as well as a significant test statistic for the aridity character (Table 5-4, character 36), which is not the case. Moore and Donoghue (2007) describe dispersal into Latin America as a promoter of diversification rate increase, and it is altitudinal environmental fluctation within the Andes that has caused allopatric speciation, leading to high diversification rates (also see Hughes and Eastwood, 2006; Wiens et al., 2007). Long-distance dispersal events are the initial promoter insofar as they expose the founder to a variety of new niches in which smaller-scale processes can operate towards speciation (von Hagen and Kadereit, 2003; Kay et al., 2005; Funk et al., 2006). Lavin et al. (2004) argue that metacommunity processes are the driving force behind current legume distributions, with historical dispersal events frequent enough to suggest that ecological preadaptation determines extant geographical patterns, which in turn promotes speciation due to isolation by distance. Although Simaroubaceae are far less diverse in total species numbers compared to legumes, their deep history and multiple post-Oligocene dispersal events might be expected to generate a similar pattern of speciation through dispersal. Recent speciation has occurred in clades with species adapted to seasonally dry or arid environments, such as previously mentioned *Holacantha* and *Castela*, African *Brucea* species, and a number of species in Clade V, e.g. *Hannoa* in African savannah habitat, and *Simaba* sect. *Floribundae* in the drier areas of Brazil and Paraguay (Cronquist, 1944c). In *Simaba*, the dry-adapted *Floribundae* clade, the much more species-rich clade of *Simaba*, is sister to the Amazonian species, and character mapping suggests primary rainforest is the ancestral habitat in Clade V. A

more fine-scale approach to categorizing ecological traits, such as seasonal rainfall averages, might provide some insight into the dynamics of dry forest speciation (Pennington et al., 2004b).

New Caledonian *Soulamea* were not resolved as a clade outside the other species of *Soulamea* and *Amaroria* in morphological analyses, and none of the characters examined in ancestral-state reconstruction mapped strictly to the New Caledonian species. It is more likely that *Soulamea* experienced an adaptive radiation associated with new arrivals to an island environment (Whittaker, 1998). Conversely, *Soulamea amara* has a very widespread distribution, and is the only bisexual species in *Soulamea*. All other species are confined to a single island or group of islands, and are dioecious. The ability to self is critical to establishment in a new environment distant from the parent population (Whittaker, 1998), whereas in the other species dioecy may confer an advantage to maintaining heterozygosity in limited space and resources.

Shifts in diversification rate are always considered in a relative context, compared either to background levels of diversification in the family or to more and less inclusive nodes. Temporal methods are usually interpreted from the point of increased, rather than decreased, rates of diversification, as these are of interest when looking at biodiversity gradients, adaptive radiations, and typical cause and effect relationships (Alfaro et al., 2007; Moore and Donoghue, 2007; Wiens et al., 2007). Clades with reduced levels of diversity, such as *Leitneria*, *Nothospondias*, *Picrolemma* and *Quassia*, are not as easily interpreted, because low diversity can be attributed just as easily to extinction as to a slow-down in diversification rate. The correlating factors of biogeography and morphological change could be implicated in an increased extinction rate, but more typically this is attributed to factors extrinsic to the organism (Cracraft, 1985), such as climate change. Although *Leitneria* registers as a significantly depauperate

lineage, based on the methods of Magallón and Sanderson (2001) and implied by node imbalance (Chan and Moore, 2006), its fossil history shows a greater diversity of species in the past (Dorofeev, 1994; Nikitin, 2006a). It is interesting to note that this lineage has such a striking shift in floral and pollen morphology from insect- to wind-pollination, yet a single extant species remains. Wind pollination has evolved many times independently in the angiosperms (Soltis et al., 2005), and a number of studies (e.g. Eriksson and Bremer, 1992; Dodd et al., 1999) suggest wind-pollinated groups have lower diversification rates. Shifts to wind pollination are typically coupled with temperate environments (Culley et al., 2002), which rarely support the significant levels of species diversity seen in the tropics (Wiens and Donoghue, 2004), so a slow-down in diversification might be expected. Pollinator shifts have been hypothesized as drivers of tropical forest diversity (Gentry, 1982) due to reproductive isolation. *Quassia amara* demonstrates a shift from insect to hummingbird pollination, yet this lineage consists of a single New World species, diverging from its African sister species approximately 13 Ma. *Costus* subg. *Costus* was demonstrated to show rapid speciation under the influence of shifts between hummingbird and bee pollination (Kay et al., 2005), which was coupled with local environmental fluctuation. However, empirical examples in a phylogenetic context such as *Costus* are few, and no conclusions can be drawn regarding *Quassia*. It should also be noted that extinct species are more likely to be recovered as fossils in temperate lineages, given preservational and collecting biases, and tropical lineages such as *Nothospondias* and *Picrolemma* could be the remnants of more diverse clades. The aridification of Africa has been hypothesized as a major cause of extinction in the Miocene (Davis et al., 2002b; Morley et al., 2003; Plana, 2004) so we might expect reduced species diversity in ancient African lineages (of which *Nothospondias* is the

oldest in Simaroubaceae). Conversely, *Picrolemma* is much less diverse than expected compared to the other incursions of Simaroubaceae into South America (*Simarouba*+*Simaba* and *Castela*).

Whether we are examining diversification as a function of biogeography, morphological evolution or stochastic processes, we are faced with the problem of phylogeny as a whole. For every clade studied, there is a unique evolutionary history that cannot be replicated in an experimental design, and inferences made must take this into account. Only when results can be synthesized across a number of plant clades and patterns are repeated, can we begin to draw firm conclusions about the origins of extant biodiversity in the Simaroubaceae. Despite this, our study has suggested recent dispersal events, rather than morphological innovation, may have contributed significantly to increased diversification rates in the family.

Table 5-1. Results of maximum parsimony analyses for morphological data and combined morphological and molecular data of Simaroubaceae. [a] indicates MaxTrees limit was reached.

	Number of ingroup taxa	Number of characters	Variable characters	Informative characters	Number of shortest trees	Length of best tree	Consistency index	Retention Index
Morphological	105	71	71	67	100,000[a]	404	0.290	0.776
Combined	115	6035	1924	1208	100,000[a]	4036	0.614	0.822

Table 5-2. Synapomorphies for major clades of Simaroubaceae, including taxa inside the clade showing a reversal in the character, and parallel acquisitions or losses outside the clade. Only characters have a CI greater than 0.1 and unambiguous reconstruction on the tree are included.

Clade	Synapomorphies
Clade I (*Picrasma*, *Castela* and *Holacantha*)	• No unambiguous synapomorphies
Castela and *Holacantha*	• Thorns present
Clade III (*Leitneria*, *Brucea*, *Soulamea* and *Amaroria*)	• Narrow, elongate thyrse with a single major axis • Corolla length less than 3 mm (parallel evolution in *Ailanthus fordii, Castela erecta, C. retusa, C. macrophylla, Picrasma crenata, Samadera bidwillii*) • Reticulately wrinkled endocarp • Striate pollen exine (except for *Leitneria* in which it is minutely verrucate)
Clade V (*Simaba* to *Odyendea* inclusively)	• No unambiguous synapomorphies
Clade V to *Quassia* inclusively	• Punctate glands on lamina • Bisexual plants (reversal in *Eurycoma* and *Simarouba*; parallel evolution in *Soulamea amara*) • Filament appendage present • Stigma inconspicuous or capitate (reversal in *Simarouba*)
Clade V to *Picrolemma* inclusively	• Five carpels (reversals in *Iridosma*, *Perriera*, *Odyendea* and *Samadera indica*; parallel in *Ailanthus*) • Crystalliferous cells present; parallel evolution in *Castela coccinea*; data is limited for this character) • Fused portion of style much longer than free stylodia branches
Ailanthus to Clade V inclusively	• No unambiguous synapomorphies

Table 5-3. Character evolution on the Simaroubaceae phylogeny for 26 characters under maximum parsimony reconstructions (MPR). Ancestral states for selected nodes are shown for ACCTRAN optimization. Asterisk indicates an alternative reconstruction under DELTRAN optimization.

No.	Character	No. of states	No. of most parsimonious recons.	No. of changes under MPR	Ancestral state for family	Ancestral state for Clade I	Ancestral state for Clade III	Ancestral state for Clade V
1	Height (life form)	2	2	9	Shrub or small tree < 12 m	Shrub or small tree < 12 m	Shrub or small tree < 12 m	Shrub or small tree < 12 m
2	Thorns	2	1	1	Absent	Absent	Absent	Absent
3	Leaf type	3	6	9	Pinnate	Pinnate	Unifoliolate*	Pinnate
4	Marginal teeth	2	4	6	Entire	Entire	Entire	Entire
5	Leaf glands	2	12	5	Equivocal	Absent	Present	Present
6	Leaf gland type	2	1	6	Absent	-	Prominent, dark	Punctate, light
7	Inflorescence type	4	15	5	Equivocal	Equivocal	Typical thyrse	Typical thyrse
8	Thyrsoid structure	2	15	6	Equivocal	-	Narrow, elongate	Open, multiple axes
9	Sex	2	1	4	Unisexual	Unisexual	Unisexual	Bisexual
10	Merosity	4	5	8	Tetramerous	Tetramerous	Tetramerous	Pentamerous
11	Corolla hairs	4	340	18	Glabrous	Glabrous	Glabrous	Pubescent on both sides
12	Corolla length	4	90	15	2.1 – 8 mm	2.1 – 8 mm	0 – 2 mm	2.1 – 8 mm
13	Corolla color	2	6	9	White/yellow/green	White/yellow/green	White/yellow/green	White/yellow/green
14	Shape of nectar disc	3	37	8	Fleshy, broad	Fleshy, broad	Fleshy, broad	Columnar*
15	Nature of androecium	3	4	7	Equivocal	Equivocal	Haplostemonous*	Diplostemonous
16	Filament appendage	2	1	2	Absent	Absent	Absent	Present
17	Number of carpels	6	6	10	Equivocal	Four	One*	Five
18	Style length	2	1	3	Free portion longer	Free portion longer	Free portion longer	Free portion shorter
19	Stigma shape	4	2	7	Linear, recurved	Linear, recurved	Linear, recurved	Inconspicuous
20	Fruit type	6	1	5	Ellipsoid drupe	Ellipsoid drupe	Ellipsoid drupe	Ellipsoid drupe
21	Drupe size	3	36	12	Equivocal	≤ 13 mm	13.1 – 25 mm	13.1 – 25 mm
22	Drupe color when ripe	2	12	4	Orange/red/black	Orange/red/black	Orange/red/black	Orange/red/black
23	Pollinator	3	1	2	Insect	Insect	Insect	Insect
24	Mode of dispersal	2	1	4	Biotic	Biotic	Biotic	Biotic
25	Temperate vs. tropical habitat	2	2	4	Tropical	Tropical	Tropical	Tropical
26	Moisture in habitat	3	4	8	Mesophytic	Mesophytic	Mesophytic	Mesophytic

Table 5-4. Characters tested for correlations with diversification rate shift. Binary character states are shown, along with posterior predictive p value (values in bold are significant ($p < 0.05$)).

	Character	State 0	State 1	P
1	Maximum height (life form)	Shrub < 12 m	Tree > 15 m	0.42
2	Thorns	Absent	Present	**0.03**
3	Leaf type	Pinnately compound	Unifoliolate or leafless	0.89
4	Margin teeth	Entire	Crenulate, serrate or shallowly undulating	0.06
5	Punctate glands	Absent	Present	1.00
6	Prominent, dark glands	Absent	Present	0.12
7	Typical thyrse inflorescence	Not typically thyrsoid	Typical thyrse	0.88
8	Fascicle-like inflorescence	Not fasciculate	Fasciculate	0.07
9	Umbellate inflorescence	Not umbellate	Umbellate	0.88
10	Cymose inflorescence	Not cymose	Cymose	0.06
11	Thyrse structure (narrow)	Not spiciform	Spiciform	0.67
12	Thyrse structure (open)	Absent or narrow	Broad, compound, open,	0.85
13	Sex	Unisexual (or androdioecious)	Bisexual	1.00
14	Trimerous	No	Yes	0.12
15	Tetramerous	No	Yes	0.06
16	Pentamerous	No	Yes	0.92
17	Petal hair	Glabrous, or occasional hairs	Pubescent	0.99
18	Small flowers	Flowers > 2 mm	Flowers ≤ 2 mm	0.09
19	Medium flowers	≤ 2 mm or > 8 mm	2.1 – 8 mm	0.55
20	Large flowers	< 9 mm	≥ 9 mm	1.00
21	Petal color (abaxial)	White/yellow/green	Pink/red/violet	0.18
22	Disc shape	Ring-like, fleshy, lobed or absent	Columnar	0.98
23	Androecium	Diplostemonous (or pleiostemonous)	Haplostemonous	**0.02**
24	Filament appendage	Absent	Present	0.97
25	Number of carpels	More than one	One	0.45
26	Short style	Style long, stigma inconspicuous	Style short, stigma long, or discoid	**0.04**
27	Long style	Style short, stigma long or discoid	Long style, inconspicuous stigma	0.99
28	Fruit type	Fleshy, drupaceous	Dry, samaroid	0.14
29	Drupe size (excl. wing, keel)	≤ 25 mm	29 mm or more	1.00
30	Red drupes	Green, yellow or n/a	Orange to red to blue-black	0.42
31	Yellow or green drupes	Orange to red to blue-black or n/a	Yellow or green	0.95
32	Pollinator	Insect	Bird or wind	0.14
33	Dispersal	Biotic	Abiotic	0.21
34	Habitat1	Tropical	Temperate	**0.03**
35	Habitat2	Not swamp	Swamp, riverine, coastal	0.77
36	Habitat3	Not arid	Arid	0.11

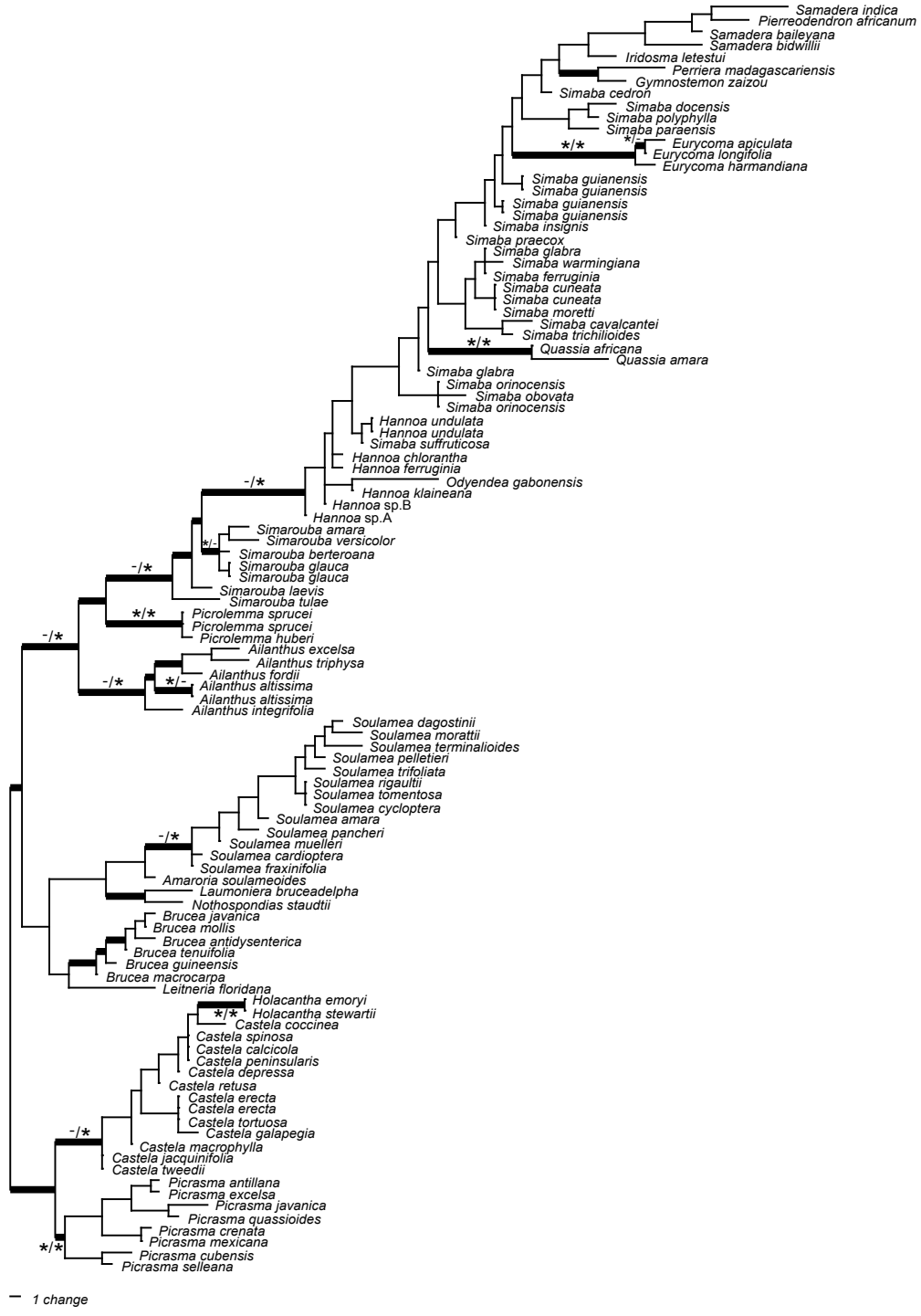


Figure 5-1. Phylogram randomly selected from 100,000 most-parsimonious trees recovered from 71 morphological characters for Simaroubaceae. Thick branches appear in the strict consensus. Asterisks indicate >70% bootstrap support and > 0.95 posterior probability, respectively.

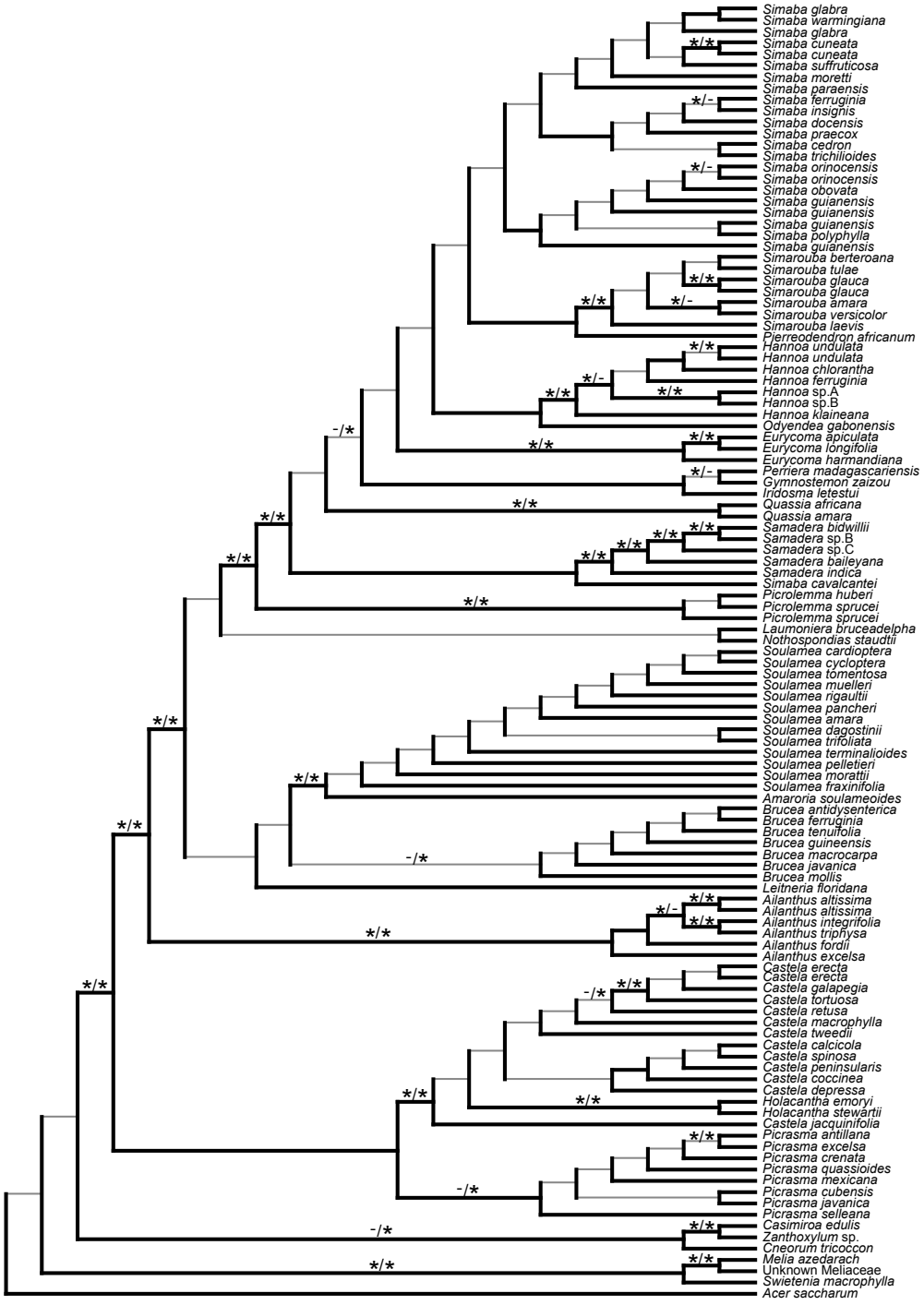


Figure 5-2. Phylogeny randomly selected of 100,000 most parsimonious trees recovered from a combined analysis of *rbcL*, *atpB*, *matK* (including partial *trnK* intron) and *phyC*, and 71 morphological characters for Simaroubaceae. Grey branches collapse in the strict consensus. Asterisks indicate >70% bootstrap support and > 0.95 posterior probability, respectively.

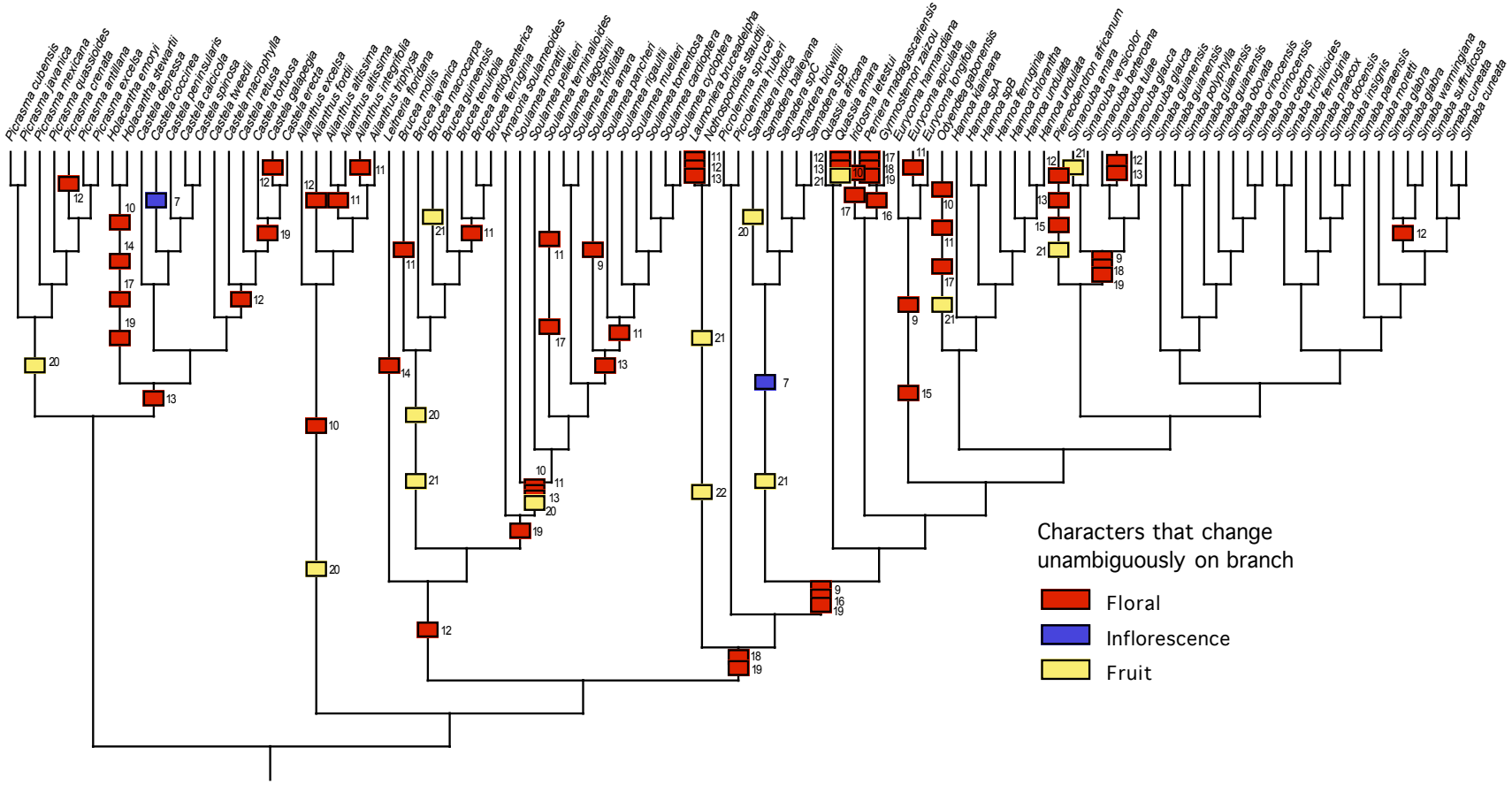


Figure 5-3. Unambiguous character-state changes based on maximum parsimony reconstruction for reproductive traits. Numbers by character-state transitions refer to character numbers in Table 5-3.

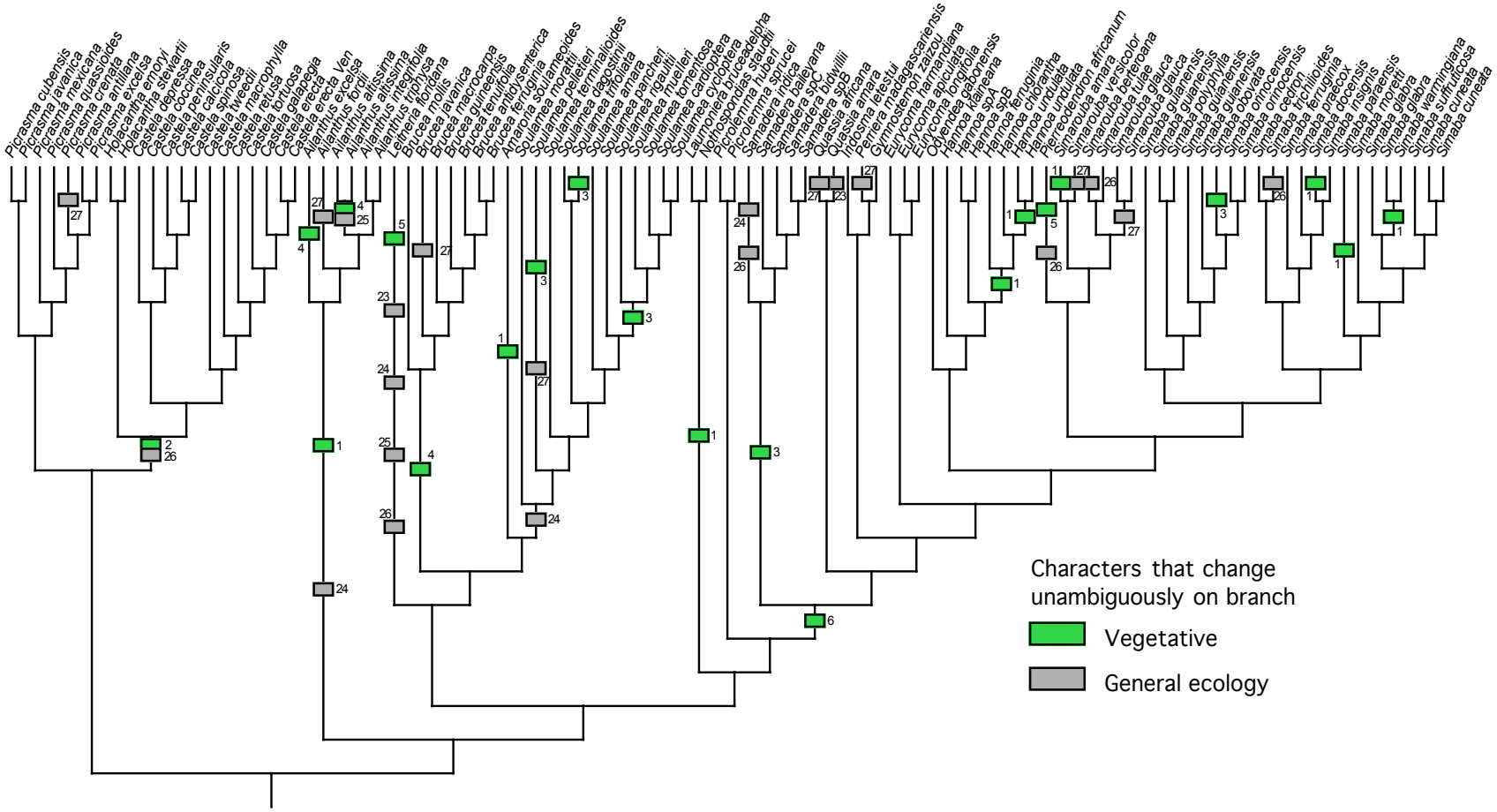


Figure 5-4. Unambiguous character-state changes based on maximum parsimony reconstruction for vegetative traits and general ecology. Numbers by character-state transitions refer to character numbers in Table 5-3.

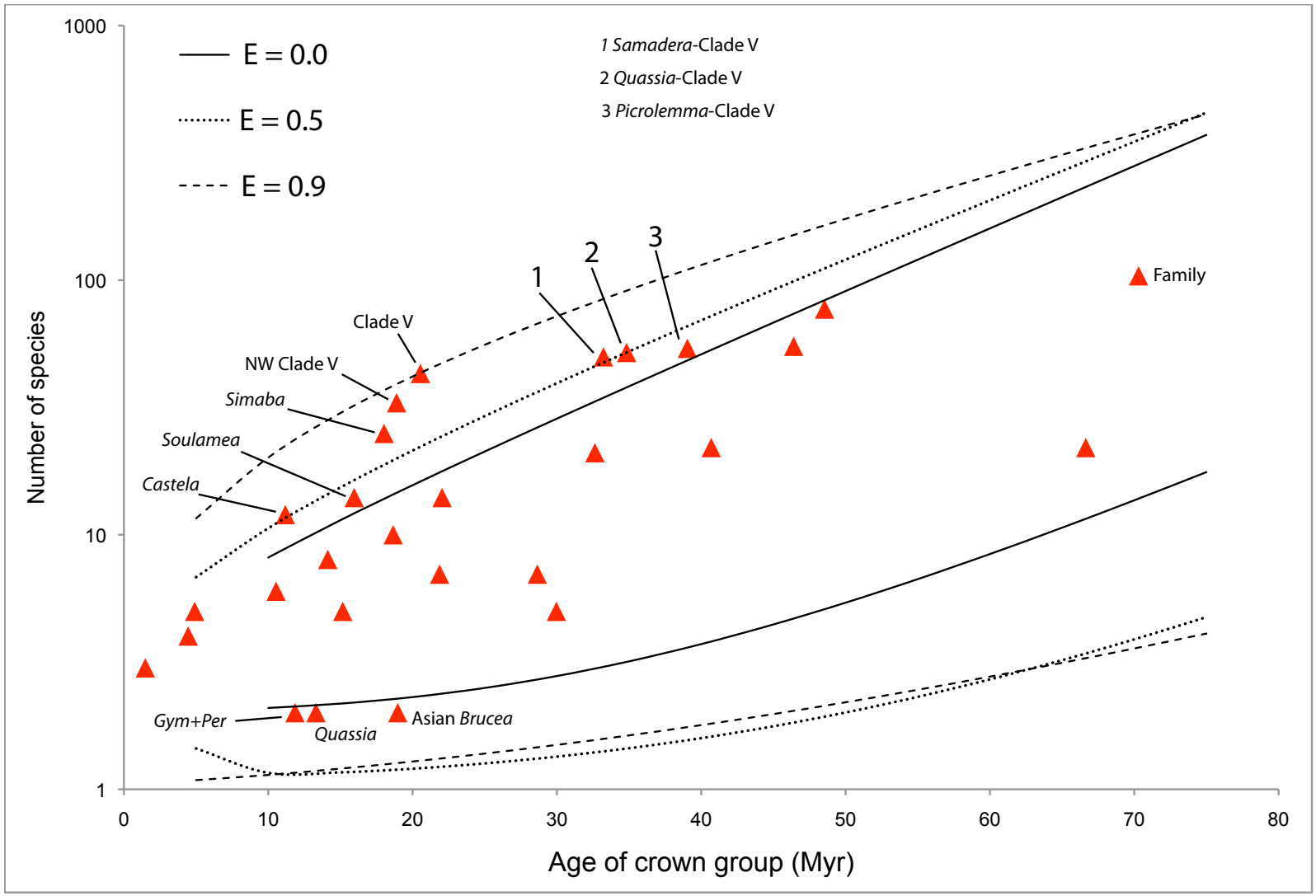


Figure 5-5. Crown group analysis of clades of Simaroubaceae using methods of Magallón and Sanderson (2001). Confidence intervals for three extinction rates are shown.

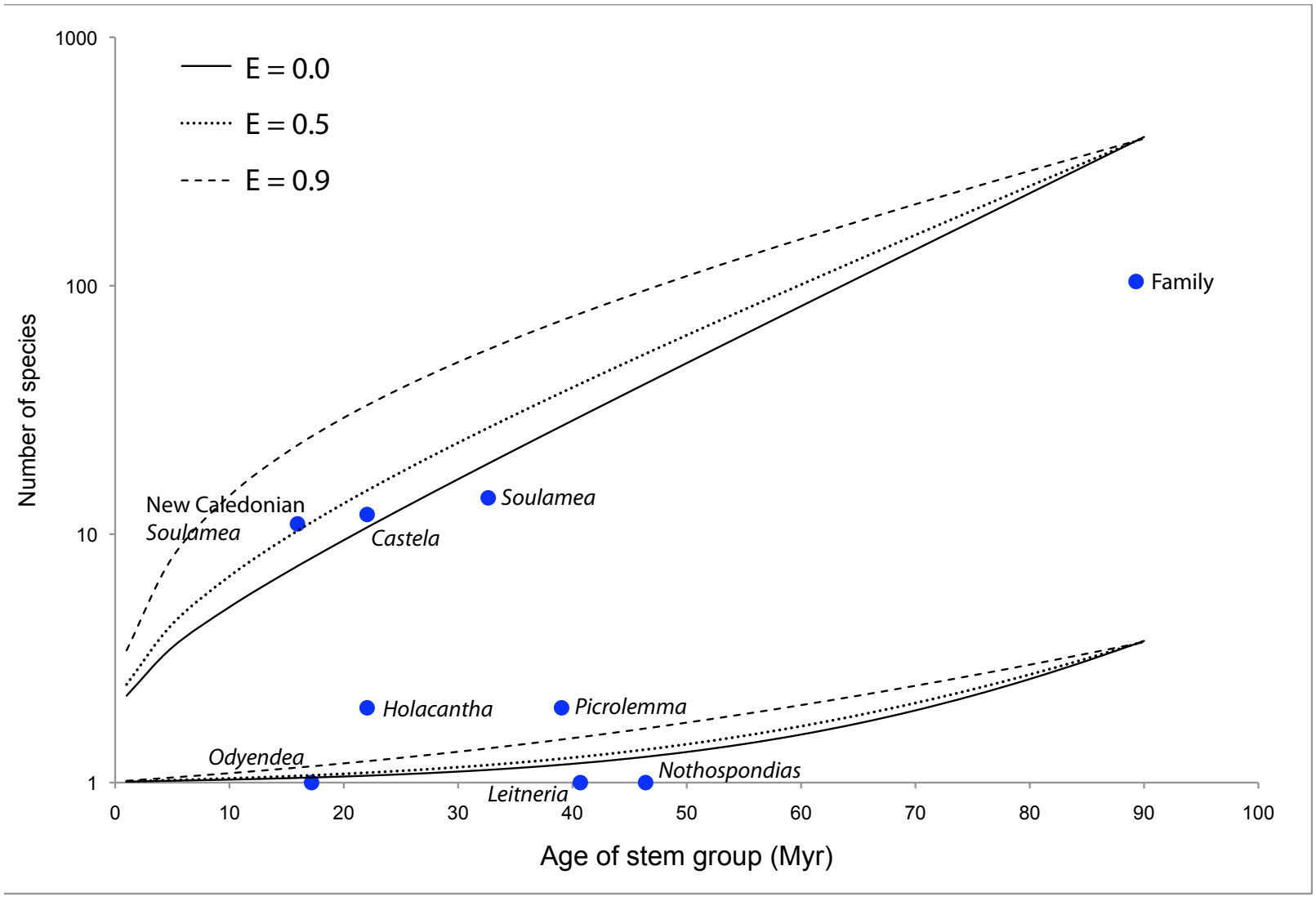


Figure 5-6. Stem group analysis of clades of Simaroubaceae using methods of Magallón and Sanderson (2001). Confidence intervals for three extinction rates are shown.

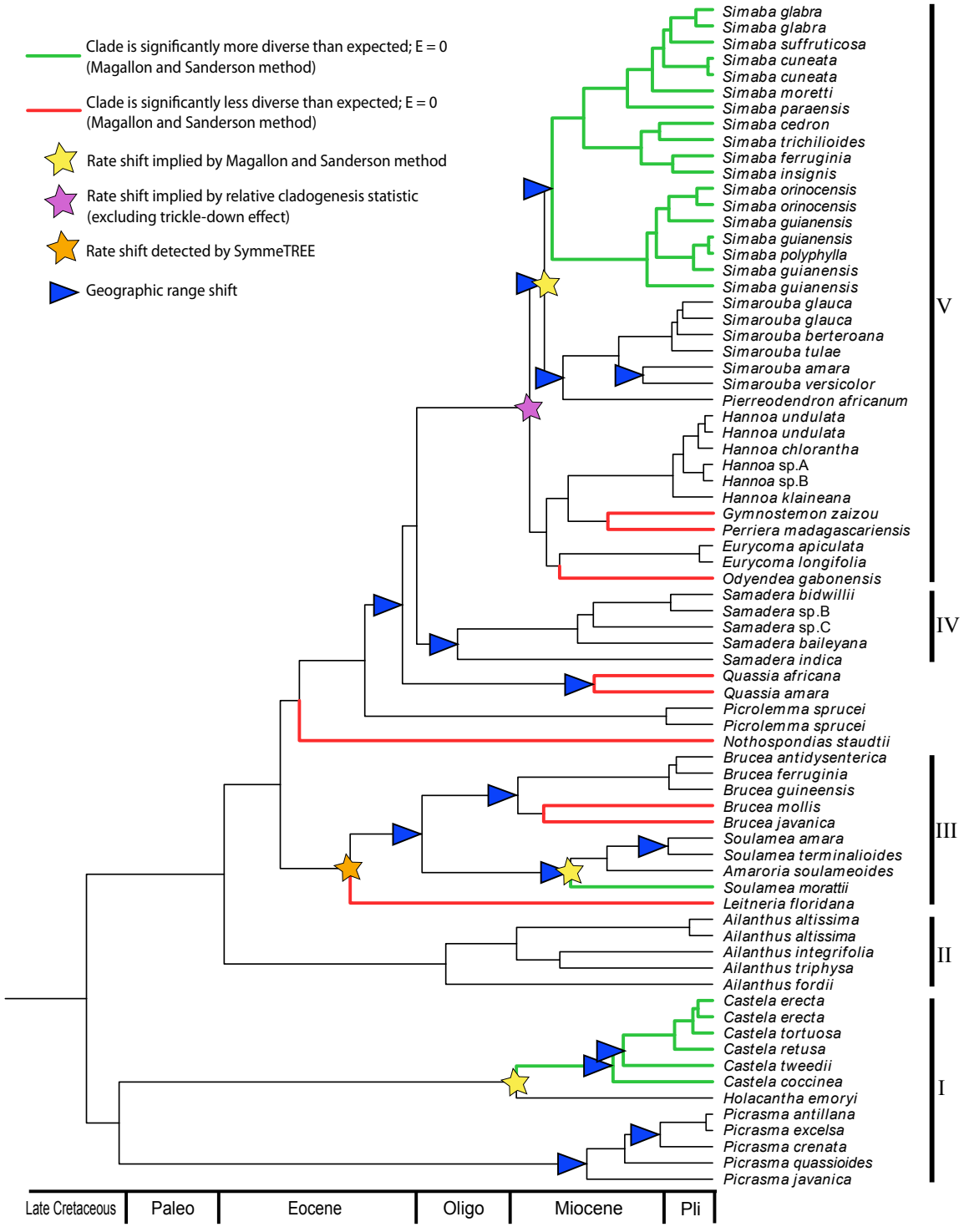


Figure 5-7. Chronogram for Simaroubaceae (reproduced from Figure 4-3) showing shifts in diversification rates based on the methods of Magallón and Sanderson (2001), the relative cladogenesis statistic (Nee et al., 1992), and SymmeTREE (Chan and Moore, 2005). Also shown are nodes at which biogeographic shifts between two areas are hypothesized, based on biogeographic analyses in Chapter 4.

CHAPTER 6
CONCLUSIONS

Application of the latest techniques in systematics has been critical to improving our understanding of the evolutionary history of Simaroubaceae, particularly as the family has received little attention in recent years. Phylogenetic reconstruction has been aided by the family's small size, and successful procurement of plant material, which has been obtained for almost all genera and the entire geographic range of the family. Sequencing a combination of both well-established and less often used gene regions, and applying the latest approaches to modeling sequence evolution in a Bayesian framework, has produced a robust phylogeny. Combining this phylogeny with fossils of *Ailanthus* and *Leitneria*, and a recently developed model of uncorrelated Bayesian molecular rates analysis (Drummond et al., 2006), has generated divergence date estimations for the family. In turn these dates, along with geographical data for extant and extinct species, have been used to test hypotheses of historical migration and long-distance dispersal events, under a very recently developed likelihood model of geographic range evolution (Ree and Smith, 2008). Finally, the phylogeny has been used to examine character evolution in the family, and search for correlates of diversification rate shifts.

Not since Engler's (1931) classification has the entire family been dealt with in a comprehensive overview. Since then, five of Engler's six subfamilies, and *Harrisonia* from Simarouboideae, have been found to be more closely related to various taxa throughout the rosids (Fernando et al., 1995) than to subfamily Simarouboideae. Cronquist produced synopses of *Castela* (1944a), *Simarouba* (1944b) and *Simaba* (1944c), and an overview of the American genera (1944d), and Nooteboom (1962a) gave an excellent description of the genera and species found in Malaysia, with reference to the surrounding areas. He also recircumscribed several genera as the single genus *Quassia*, a classification not favored by other systematists of

Simaroubaceae. The African genera and species in Simaroubaceae have received little attention except in Floras of individual regions (Aubréville, 1962; Stannard, 2000). Chapter 2 is the first time the family and its constituent genera have been described as a whole, since their phylogenetic recircumscription (Fernando and Quinn, 1995). The chapter provides a key to the genera, plus morphological descriptions, approximate species numbers and geographical distribution. The natural history of the group, including anatomy, phytochemistry, reproductive biology, dispersal mechanisms, and important economic uses, is also summarized.

Chapter 3 focuses on reconstructing the molecular phylogeny of Simaroubaceae. Approximately six thousand base pairs of sequence data have proved sufficient in resolving most nodes in the phylogeny, and these data were tested thoroughly for the effects of how the genetic data are partitioned. Model selection is an increasingly important aspect of phylogeny reconstruction, as molecular data from multiple gene regions across all genomes become more accessible, and analytical techniques have to adapt to larger data sets. This study showed that the partitioning of the four gene regions did little to influence the underlying phylogenetic signal. By selecting commonly used genes (*rbcL, atpB, matK*), the study has allowed for potential integration of Simaroubaceae into larger scale Tree of Life and bar-coding projects. Additionally, with the use of *phyC*, a less well-studied gene, the phylogeny has improved our understanding of the utility of this nuclear marker.

Historical biogeography is a field undergoing considerable change, so the application of the latest techniques used in Chapter 4 is paramount to maximizing the potential of geographic and phylogenetic data. Pantropical clades are often logistically more difficult to sample, given their widespread nature, and for Simaroubaceae this is particularly true with the relatively low number of species, leading to isolated lineages and little sympatry. However, this was not a

limiting factor, with successful DNA extraction from herbarium material and contributions of silica-dried leaves from botanists worldwide. Robust fossil calibration with a combination of *Ailanthus* and *Leitneria* fossil fruits has produced a chronogram that proposes a Late Cretaceous origin for the family, a time scale similar to other pantropical families. Despite the many uncertainties associated with biogeographical analyses for geographically widespread and ancient groups, such as errors in molecular rate estimation and the relative simplicity of ancestral area reconstruction, many clear patterns in the biogeographic history of Simaroubaceae can be inferred from phylogenetic and fossil data. The family shows patterns of overland migration via putative land bridges, coupled with a number of inferred long-distance dispersal events. The break-up of Gondwanaland, traditionally the most parsimonious explanation for tropical disjunctions (Raven and Axelrod, 1974), has now been replaced by hypotheses incorporating knowledge of phylogenetic relationships, more realistic divergence dates, availability of dispersal routes (Morley, 2003), and a general acceptance of long-distance dispersal as a major driving force in extant plant distributions (Givnish and Renner, 2004; Lavin et al., 2004; Renner, 2004a; De Queiroz, 2005). The North Atlantic Land Bridge is often implicated in pantropical disjunctions, but testing a variety of potential ancestral ranges between the Old World and the New World has illustrated the importance of a trans-Beringial connection for Simaroubaceae. The biogeographic model requires further refinement to accommodate stratified ancestral ranges and more realistic fossil integration, but is an excellent first step in moving analyses towards a model-based approach, which can include testing hypotheses of a variety of biogeographic scenarios.

Chapter 5 takes a data exploration approach towards patterns of diversification in the family, similar to the recent study by Moore and Donoghue (2007). Morphological data do

contain a phylogenetic signal which recovers some of the deep-level relationships found with molecular data, and also shows many of the genera to be well-characterized morphologically. Up to four significant diversification rate increases are found, depending on the methods used. These shifts are typically associated with biogeographic dispersal events, but a number of dispersals have not resulted in diversification rate increases; therefore, a variety of factors, both environmental and intrinsic to the organism, are probably involved (de Queiroz, 2002). The technique used to examine putative key innovations is only recently developed, and requires further simulation-based studies to determine its statistical properties and behavior in a variety of phylogenetic systems. However, it has allowed the exploration of a number of morphological characters and how state changes affect the frequency of cladogenesis. By a conservative estimation, none of the 36 characters examined had a significant association, but this part of the study would benefit from improved sampling for stochastic mapping. Extending this approach to the Sapindales would hopefully yield a wealth of interesting data. The order would provide a broad phylogenetic spectrum across which more recent character-state changes could be considered independent, whilst the types of morphological characters examined are comparable, for example, floral structure, fruit morphology, etc.

In conclusion, only with these types of detailed studies of individual clades can we begin to create an overall picture of the evolutionary history of tropical and temperate biodiversity. For such a small and relatively understudied family, Simaroubaceae has produced a wealth of insights into the modeling of molecular data, pantropical historical biogeography, morphological evolution and diversification.

APPENDIX A
SPECIMEN DATA FOR MOLECULAR ANALYSES

Sources of DNA and GenBank accession numbers for sequence data. Taxon names in parentheses are original determinations.

Taxon	Voucher	*rbcL*	*atpB*	*matK*	*phyC*
Ailanthus altissima	J.W.Clayton 14 (FLAS)	EU042978	EU042770	EU042840	EU042911
A. altissima var. *tanakai*	Chase 16982 (K)	EU042979	EU042771	EU042841	EU042912
A. fordii	Lau 038 (MO)	EU042980	EU042772	EU042842	-
A. integrifolia	B.Hyland 15229 (QRS)	EU042981	EU042773	EU042843	-
A. triphysa	Fernando 1540 (UNSW)	EU042982	EU042774	EU042844	EU042913
Amaroria soulameoides	No voucher, ex SUVA, Fiji	U38923	AF066856	-	-
Brucea antidysenterica	Simon et al 1140 (MO)	EU042983	EU042775	EU042845	EU042914
B. ferruginia	Acc. 19073747 NBG Belgium	EU042984	EU042776	EU042846	EU042915
B. guineensis	McPherson 18015 (MO)	EU042985	EU042777	EU042847	EU042916
B. javanica	JTH 134	EU042986	EU042778	EU042848	EU042917
B. mollis	J.R.I.Wood 6887 (E)	EU042987	-	EU042849	EU042918
B. tenuifolia	Mwangoka and Kalage 2687 (MO)	EU042988	EU042779	EU042850	EU042919
Castela coccinea	A.Krapovickas and C.L.Cristobal 44166 (MO)	EU042989	EU042780	EU042851	EU042920
C. erecta	Webster and Armbruster 23577 (MO)	EU042991	EU042782	EU042853	EU042922
C. erecta subsp. *texana*	P.Tenorio 20405 (MO)	EU042990	EU042781	EU042852	EU042921
C. retusa	Abisai Garcia M. 2830 (MO)	EU042992	EU042783	EU042854	EU042923
C. tortuosa	Juan Torres 00172 (MO)	EU042993	EU042784	EU042855	EU042924
C. tweedii	G.Hatschbach et al. 72435 (MO)	EU042994	EU042785	EU042856	EU042925
Eurycoma apiculata	No voucher ex FRIM, Malaysia	EU042995	EU042786	EU042857	EU042926
E. longifolia	Gwee and Samsuri GW19 (SING)	EU042996	EU042787	EU042858	EU042927
Gymnostemonn zaizou	L. Ake 19117 (MO)	EU042997	EU042788	EU042859	EU042928
Hannoa chlorantha	D.K.Harder et al. 3732 (MO)	EU042998	EU042789	EU042860	EU042929
H. klaineana	M.Merello et al. 1584 (MO)	EU042999	EU042790	EU042861	EU042930
H. undulata	Madsen 4075 (MO)	EU043001	EU042792	EU042863	EU042932
H. undulata	H.H.Schmidt et al. 3336 (MO)	EU043000	EU042791	EU042862	EU042931
Holacantha emoryi	T.K. Lowrey and C.J. Quinn,1876 (UNM)	EU043002	EU042793	EU042864	EU042933
Leitneria floridana	J.R.Abbott 14212 (FLAS)	EU043003	EU042794	EU042865	EU042934
Nothospondias staudtii	H.P.Bourobou Bourobou 474 (MO)	EU043004	EU042795	EU042866	EU042935
Odyendea gabonensis	G.Walters et al. 577 (MO)	EU043005	EU042796	EU042867	EU042936
O. zimmermannii	L.B. Mwasumbi 14191 (MO)	EU043006	EU042797	EU042868	EU042937

Perriera madagascariensis	Noyes et al 1079 (MO)	EU043007	EU042798	EU042869	EU042938
Picrasma antillana	P.Aceredo Rdgz, A.Siaca 4262 (MO)	EU043009	EU042800	EU042871	-
P. crenata	A.C.Cervi and R. Spichiger 6861 (NY)	EU043010	EU042801	EU042872	EU042940
P. javanica	No voucher, wild collected, China.	EU043011	EU042802	EU042873	EU042941
P. quassioides	Acc. 19510406 RBGE cultivated	EU043008	EU042799	EU042870	EU042939
P. excelsa	Hill 25443 (MO)	EU043012	EU042803	EU042874	EU042942
Picrolemma sprucei (*P. pseudocoffea*)	G. Bourdy GB2988 (CAY)	EU043013	EU042804	EU042875	EU042943
P. sprucei	C.Grandez and N.Jaramillo 2871 (MO)	EU043014	-	EU042876	EU042944
Pierreodendron africanum	Terese Butler Hart 1386 (MO)	EU043015	EU042805	EU042877	EU042945
Quassia africana	McPherson 16672 (MO)	EU043016	EU042806	EU042878	EU042946
Q. amara	Chase 18959 (K)	EU043017	EU042807	EU042879	EU042947
Samadera baileyana	UNSW 22894 (UNSW)	EU043018	EU042808	EU042880	EU042948
S. bidwillii	Craven and Walker 9339	EU043019	EU042809	EU042881	EU042949
S. indica	JTH 138	EU043020	EU042810	EU042882	EU042950
S. sp. B	Fernando 1538 (UNSW)	EU043021	EU042811	EU042883	EU042951
S. sp. C *(Quassia* Barong*)*	A. Ford 4679	EU043022	EU042812	EU042884	EU042952
Simaba cedron	N.P.Taylor 690 (MO)	EU043024	EU042814	EU042886	EU042953
S. cuneata	R.P. Lyra-Lemos 4082 (NY)	EU043025	EU042815	EU042887	EU042954
S. cuneata	M.R. Barbosa s.n.	EU546244	EU546227	EU546232	EU546238
S. cf *ferruginia*	Pott and Franco 6177 (E)	EU043027	EU042817	EU042889	EU042956
S. glabra (*S.* cf *blanchetii*)	Ratter et al R8004 (E)	EU043023	EU042813	EU042885	-
S. glabra	J.A.Ratter and Valdir P. de Lima 6719 (MO)	EU043028	EU042818	EU042890	EU042957
S. guianensis (*S. cuspidata*)	Korning and Thomsen 47632 (AAU)	EU043026	EU042816	EU042888	EU042955
S. guianensis	Martin Timana 3683 (MO)	EU043030	EU042820	EU042892	EU042959
S. guianensis	L.Barrabe and M.Pechberty 185 (CAY)	EU043029	EU042819	EU042891	EU042958
S.guianensis (*S. orinocensis*)	P. Grenand 3267 (CAY)	EU043034	EU042824	EU042896	EU042963
S. insignis	Pirani et al. 4517 (NY)	EU043031	EU042821	EU042893	EU042960
S. morettii	M. Prevost and D. Sabathier 2987	EU546245	EU546228	EU546233	EU546239
S. orinocensis (*S. multiflora*)	Rimanchi 10347 (MO)	EU043032	EU042822	EU042894	EU042961
S. orinocensis (*S. multiflora*)	Gillespie 2576 (MO)	EU043033	EU042823	EU042895	EU042962
S. aff *paraensis*	G.L. Farias 301	EU546246	-	EU546234	EU546240
S. polyphylla	Molino J.-F. et al 1998 (CAY)	EU043035	EU042825	EU042897	EU042964
S. suffruticosa	J. Ratter et al. 3532	EU546247	EU546229	EU546235	EU546241
S. trichilioides	R.C. Forzza et al. 455	EU546248	EU546230	EU546236	EU546242
Simarouba amara	A.Araujo et al. 353 (MO)	EU043036	EU042826	EU042898	EU042965

S. berteroana	WT 14671	EU546249	EU546231	EU546237	EU546243
S. glauca (cultivated)	J.R.Abbott 19605 (FLAS)	EU043037	EU042827	EU042899	EU042966
S. glauca (wild collected)	Acc. 19843216 NBG Belgium	EU043038	EU042828	EU042900	EU042967
S. tulae	Taylor 10589 (MO)	EU043039	EU042829	EU042901	EU042968
S. versicolor	Nee 39012 (MO)	EU043040	EU042830	EU042902	EU042969
Soulamea amara	Chambers 78 (MO)	EU043041	EU042831	EU042903	EU042970
S. terminalioides	Robertson 2529 (MO)	EU043043	EU042833	EU042905	EU042972
S. sp.	J.Munziger and McPherson 589 (MO)	EU043042	EU042832	EU042904	EU042971
OUTGROUPS					
Acer saccharum	Chase 106 (NCU)	L13181	AF035893	AY724265	-
Melia azaderach	J.R.Abbott 8456 (FLAS)	EU042973	EU042764	EU042834	EU042906
Swietenia mahogani/ macrophylla	J.W.Clayton 12 (FLAS)	AY128241	EU042765	EU042835	EU042907
Casimiroa edulis	J.R.Abbott 8156 (FLAS)	EU042975	EU042767	EU042837	EU042909
Zanthoxylum sp.	J.W.Clayton 15 (FLAS)	EU042976	EU042768	EU042838	EU042910
Cneorum tricoccon	J.W.Clayton 13 (FLAS)	EU042977	EU042769	EU042839	-
Unknown Meliaceae	Acc. 19514782 NBG Belgium	EU042974	EU042766	EU042836	EU042908

APPENDIX B
SOURCES OF MORPHOLOGICAL DATA

Morphological data obtained from herbarium specimens (see Appendix C), botanical literature, or both.

Species	Herbarium specimen	Literature
Ailanthus altissima	X	X
A. excelsa	X	X
A. fordii	X	X
A. integrifolia		X
A. triphysa	X	X
Amaroria soulameoides		X
Brucea antidysenterica	X	X
B. guineensis	X	X
B. javanica	X	X
B. macrocarpa		X
B. mollis	X	X
B. tenuifolia		X
Castela calcicola		X
C. coccinea	X	X
C. depressa		X
C. erecta	X	X
C. galapegia	X	X
C. jacquinifolia		X
C. macrophylla		X
C. peninsularis		X
C. retusa	X	X
C. spinosa		X
C. tortuosa	X	X
C. tweedii	X	X
Eurycoma apiculata		X
E. harmandiana		X
E. longifolia	X	X
Gymnostemon zaizou	X	X
Hannoa chlorantha	X	X
H. ferruginia	X	X
H. klaineana	X	X
H. sp. A	X	
H. sp. B	X	
H. undulata	X	X
Holacantha emoryi	X	X
H. stewartii	X	X
Iridosma letestui		X
Laumoniera bruceadelpha		X
Leitneria floridana	X	X
Nothospondias staudtii	X	X
Odyendea gabonensis	X	X
Picrasma antillana	X	X
P. crenata	X	X
P. cubensis		X
P. excelsa	X	X
P. mexicana		X
P. quassioides	X	X
P. selleana	X	X

Perriera madagascariensis	X	X
Picrolemma huberi		X
P. sprucei	X	X
Pierreodendron africanum	X	X
Quassia africana	X	X
Q. amara	X	X
Samadera baileyana	X	X
S. bidwillii	X	X
S. indica	X	X
Simaba cavalcantei		X
S. cedron	X	X
S. cuneata	X	X
S. docensis		X
S. ferruginia	X	X
S. glabra	X	X
S. guianensis	X	X
S. insignis	X	
S. obovata	X	X
S. orinocensis	X	X
S. paraensis	X	X
S. polyphylla	X	X
S. praecox		X
S. suffruticosa		X
S. trichilioides		X
S. warmingiana	X	
Simarouba amara	X	X
S. berteroana	X	X
S. glauca		X
S. laevis		X
S. tulae	X	X
S. versicolor	X	X
Soulamea amara	X	X
S. cardioptera	X	
S. cycloptera		X
S. dagostinii		X
S. fraxinifolia		X
S. morattii		X
S. muelleri	X	X
S. pancheri	X	X
S. pelletieri		X
S. rigaultii		X
S. terminalioides	X	
S. tomentosa		X
S. trifoliata		X

APPENDIX C
SPECIMEN DATA FOR MORPHOLOGICAL ANALYSES

Voucher specimens used in examination of morphological characters

Species	Voucher
Ailanthus altissima	M.S.Franc 75 (FLAS); D.R.Windler, M.Burch & P.H.Kennan 3698 (FLAS); A.E. Radford 33944 (FLAS); E. Prendes (FLAS);A Cuthbert (FLAS)
A. excelsa	E.West (FLAS)
A. fordii	C.P.Lau 038 (MO)
A. triphysa	J.F.Maxwell 89-244 (MO)
Brucea antidysenterica	W. Kindeketa 309 (MO); G.Simon, N.Senti & Y. Raphael 1140 (MO); G.J.H.Amshoff 1970 (MO)
B. guineensis	H.H.Schmidt et al. 3594 (MO); G.McPherson 18015(MO)
B. javanica	P.J.Martin 36966 (MO); Paul Chai et al. 33226 (MO); Hu & But 22470 (MO)
B. mollis	J.R.I.Wood 6887 (MO)
Castela coccinea	A. Krapovickas & C.L.Cristóbal 44166 (MO); T. L. Gragson 159 (MO); M.Nee 48131 (NY)
C. erecta	T.B. Croat & D.P. Hannon 66020 (MO); P. Tenorio L. 20405 (MO); G.L. Webster & W. S. Armbruster 23577 (MO)
C. galapegia	R.D. Suttkus 66-42-2 (FLAS)
C. retusa	Abisai Garcia M. et al. 2830 (MO)
C. tortuosa	Juan Torres 00172 (MO)
C. tweedii	G. Hatschbach, R. Goldenberg & J.M. Silva 72435 (MO)
Eurycoma longifolia	K. Larsen et al. 43047 (MO)
Gymnostemon zaizou	L. Aké 19117 (MO)
Hannoa chlorantha	D.K.Harder, M.Bingham, B.Luwiika & N. Zimba 3732 (MO)
H. ferruginia	D.W.Thomas & H.L.Mcleod 5301 (MO)
H. klaineana	M.Merello, H.H.Schmidt, J.Amponsah, M.Chintoh & K.Baah 1584 (MO)
H. sp. A *(Brucea tenuifolia)*	M.A.Mwangoka & A.Kalage 2687 (MO)
H. sp. B *(Odyendea zimmermannii)*	L.B.Mwasumbi 14191 (MO)
H. undulata	J.E.Madsen 2909 (MO); J.E.Madsen 2909 (AAU); J.E.Madsen 4075 (AAU); C.C.Jongkind & C.M.J.Nieuwenhuis 1887 (MO)
Holacantha emoryi	R.F. Thorne 45120 (FLAS)
H. stewartii	M.C.Johnston, T.L.Wendt & F. Chiang C. 11391 (FLAS)
Leitneria floridana	R.K.Godfrey 62862 (FLAS); D.Demaree 71628 (FLAS); R.K.Godfrey & A.F.Clewell 63248 (FLAS); A. Gholson (FLAS); W.Judd 3323 (FLAS); J.R.Abbott 9047 (FLAS)
Nothospondias staudtii	H.P.Bourobou Bourobou 474 (MO)

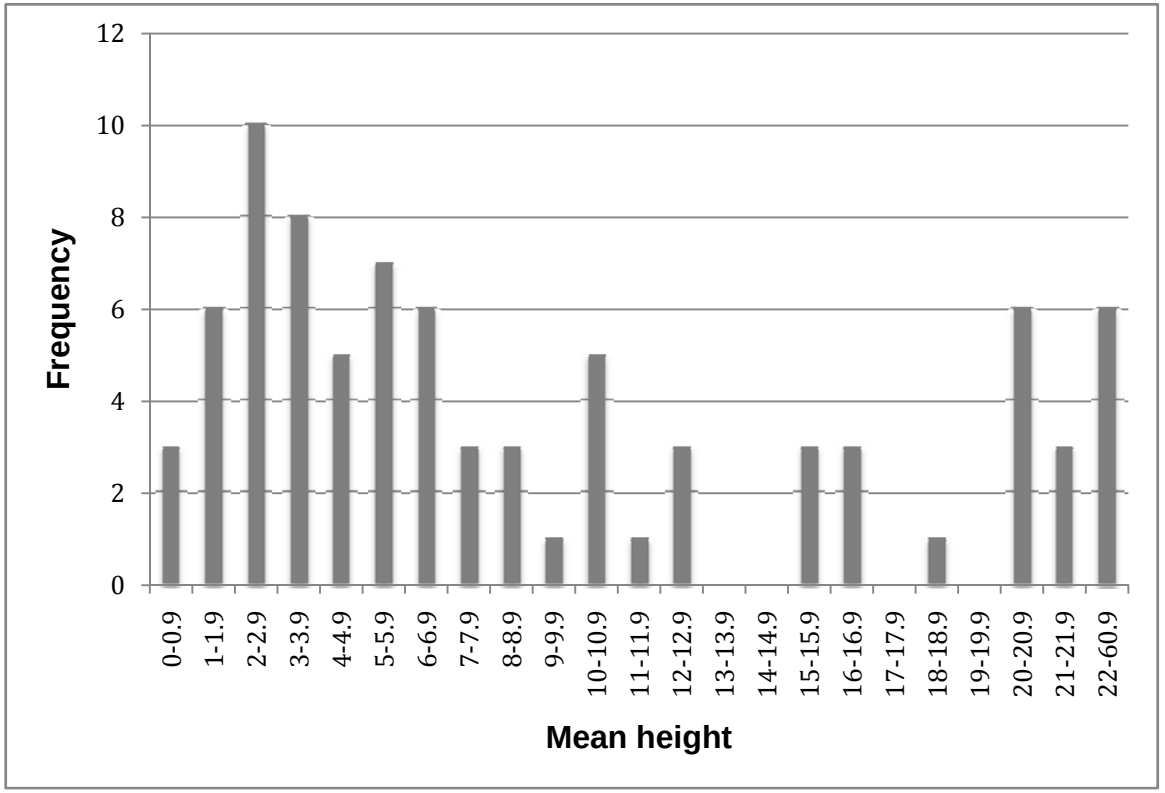
Odyendea gabonensis	J.J. de Wilde 8470 (MO); G. McPherson 16947 (MO); G. Walter, J.Stone, G.N.Essouma, A.Mintsa & L.Ndong 577 (MO)
Picrasma antillana	P. Acevedo Rdgz. A. Siaca 4262 (MO)
P. crenata	J.C.Lindeman & J.H.de Haas 992 (NY); A.C.Cervi & R. Spichiger 6861 (NY); W.W.Thomas, J.Kallunki & J.Jardim 11919 (NY)
P. excelsa	M.Nee, R. Vásquez, G.Coimbra & A. Becerra 49143 (NY), G.R.Proctor 45915 (MO), J. Betancur 1338 (MO), S. R. Hill 25443 (MO), A.Grijalva & Mario Sousa 30453 (MO), G.A.Goodfriend (FLAS)
P. quassioides	K.Sohma 3001(MO); S.Tsugaru & T.Takahashi 18049 (MO); S.Tsugaru & M.Sawada 18551 (MO)
P. selleana	W.S.Judd 4400 (FLAS)
Perriera madagascariensis	R.D.Noyes et al. 1079 (MO)
Picrolemma sprucei	C.Grández & N. Jaramillo 2871 (MO); R. Vásquez and N. Jaramillo 5674(MO);G.Bourdy 2988 (CAY); M.-F.Prévost, D.Sabatier & J.-F.Molino 4440 (CAY)
Pierreodendron africanum	T. Butler Hart 1386 (MO)
Quassia africana	F.J.Breteler 6597 (MO);G.McPherson 16672 (MO); X.M. van der Burgt, D.Ndoum, B.S.van Gemerden & S. Gideon 529 (WAG); J.J.F.E. de Wilde 8296 (MO)
Q. amara	A. Welsing, M, Merello & H. Schmidt 10 (MO); M.Nee 8638 (MO); R.Rueda & H. Cuadros 357 (MO); J.González 511 (MO)
Samadera baileyana	B.Gray 1732 (MO)
S. bidwillii	D.Halford Z155 (MO)
S. indica	R.Rabevohitra 2170 (MO); W.Meijer 10134 (MO); S.Waas 378 (MO)
Simaba blanchetii	J.A.Ratter, S. Bridgewater & J.Batista 8004 (E)
S. cedron	A.Gentry & A.Perry 77869 (MO); N.P.Taylor 690 (MO)
S. cuneata	R.P.Lyra-Lemos 4082 (NY)
S. glabra	J.A.Ratter & V.P. de Lima 6719 (MO)
S. guianensis	M.Timana 3683 (MO); J.J.Pipoly et al. 12746 (MO); D.C.Daly et al. 8303 (NY); R.Vásquez et al. 14174 (MO); L. Barrabé & M.Pechberty 185 (CAY); J.Korning & K.Thomsen 47632 (AAU); J.J. de Granville 16641 (CAY)
S. insignis	J.R.Pirani et al. 4517 (NY)
S. obovata	J.J.Wurdack & L.S.Adderley 42989 (MO)
S. orinocensis	L.J.Gillespie 2576 (MO); M. Rimanchi 10347 (MO);R. Liesner & A.C. González 5613 (MO); C.C.Berg & A.J.Henderson BG672 (MO)
S. paraensis	P.Vinha 844 (NY); S.Espinoza & C.Gualinga 827 (MO)
S. polyphylla	M. Alexiades 189 (NY); M.Aulestia 2227 (MO); J.-F. Molino, M.-F.Prévost & D. Sabatier 1998 (CAY)
S. warmingiana	J.G.Jardim et al. 912 (NY)
Simarouba amara	P.Moreno & J.C.Sandino 6481 (MO); P.Moreno & J.C.Sandino 6167 (MO); R.Liesner & A. González 11470 (MO); Quevedo 2399 (NY)
S. berteroana	T. Zanoni & F. Jiménez 45005 (MO); T.Zanoni, J. Pimentel & R. García 38070 (FLAS); T.Zanoni, J. Pimentel &

	R. García 38071 (FLAS)
S. tulae	C.M.Taylor 10589 (MO)
S. versicolor	L.P. de Queiroz & N.S.Nascimento 4133 (NY); M.Nee 39012 (MO);J.G.Jardim et al. 888 (NY); A.M. de Carvalho 4062 (MO); S.Tsugaru & Y. Sano B192 (NY)
Soulamea amara	A.F. Chambers 78 (MO)
S. cardioptera	G. McPherson 3292 (MO)
S. muelleri	J. Munzinger 345 (MO)
S. pancheri	G.McPherson 2080 (MO); J. Munzinger 955 (MO)
S. terminalioides	S.A. Robertson 2529(MO)

APPENDIX D
MORPHOLOGICAL CHARACTERS USED IN PHYLOGENETIC ANALYSES

Morphological characters and character states used in phylogenetic analyses. Justification is provided where characters and character state delimitations might be considered subjective.

	Character	States
1	Mean height (life form)	0 Shrub or small tree < 12 m; 1 Large tree > 15m Mean height showed a gap between 12 m and 15 m, at which the boundary between character states was drawn.



2	Thorns	0 Absent; 1 Present
3	Branchlet hair	0 Glabrous; 1 Pubescent towards branch ends, especially young shoots Glabrous shoots may have shown very occasional hairs; pubescence when present was typically dense

4 Pinnation

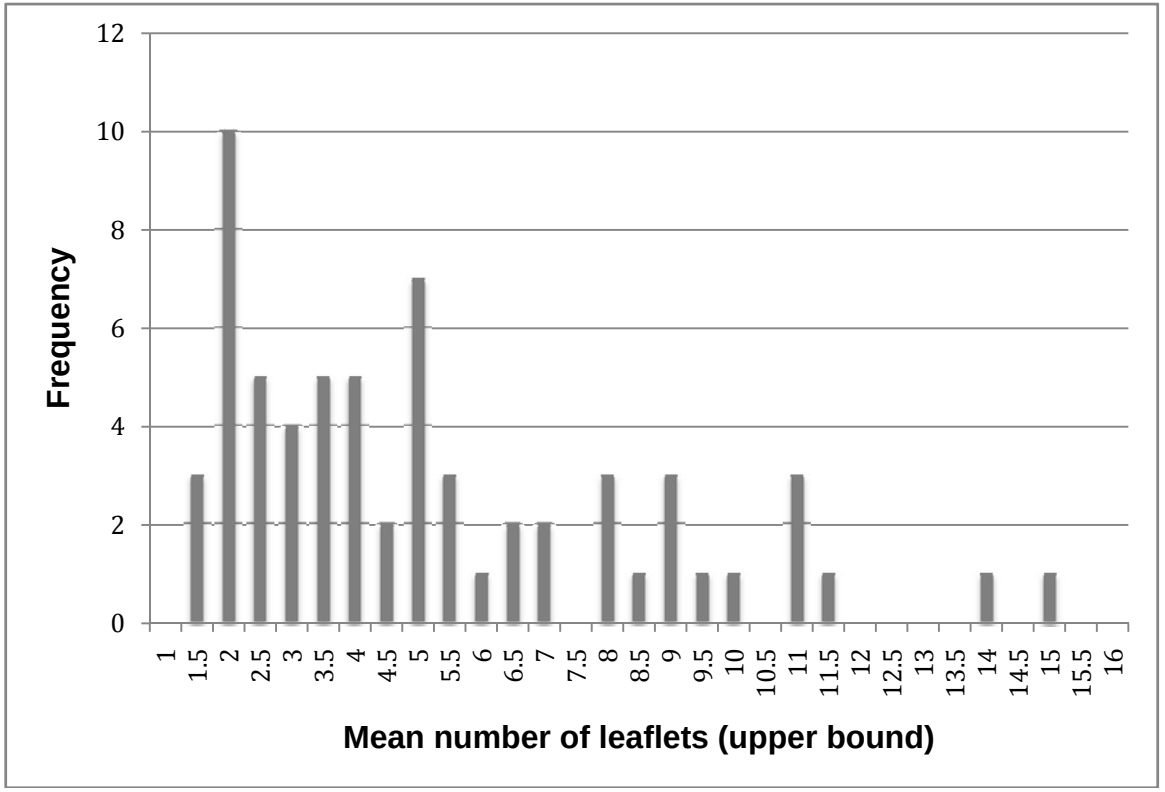
0 Imparipinnate; Paripinnate
Certain specimens of *Ailanthus* and *Simarouba* showed both conditions, and were coded as ambiguous

5 Leaflet phyllotaxy

0 Opposite to subopposite above; 2 Alternate above
Most genera had some variation in degree of alternation on lower leaflets, but this character refers specifically to leaflets in *Simarouba*, which are distinctly alternating along the entire length of the rachis

6 Mean no. lflt prs

0 One to seven; 1 Eight or more
The break between the character states was drawn based on a gap in the range of mean number of leaflet pairs



7 Leaflet shape

0 Primarily ovate to elliptic; 1 Primarily oblong; 2 Primarily obovate
Oblong and obovate leaflets were a distinct departure from the other leaflet shapes, which were more generally categorised as ovate to elliptic (see Judd et al. (2001) for character definitions). "Primarily" refers to the dominant shape seen in the lower leaflets.

8 Leaf shape

0 Primarily ovate to elliptic; 1 Primarily oblong; 2 Primarily obovate
See character 7.

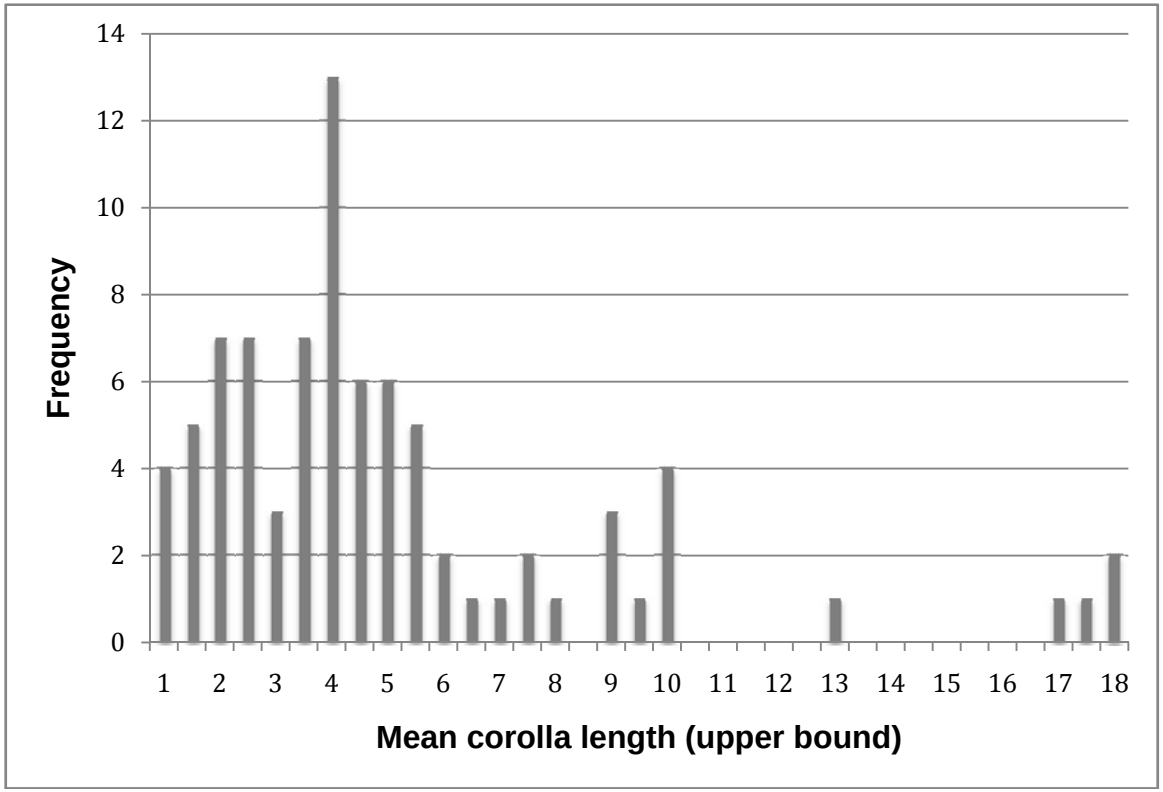
9 Leaf type

0 Leafless; 1 Simple or unifoliolate; 2 Pinnately compound

		Leafless refers to *Holacantha*, which is essentially leafless but may have scale-like structures homologous to leaves.
10	Leaflet apex	0 Rounded, truncate or emarginate; 1 Acute to shortly acuminate; 2 Long acuminate Long acuminate refers to the apex seen in *Picrolemma*, *Quassia*, and one species of *Picrasma*, in which the significantly narrowed portion of the tip has a visible length along which the sides are parallel.
11	Leaf apex	0 Rounded, truncate or emarginate; 1 Acute to shortly acuminate; 2 Long acuminate See character 10.
12	Margin teeth	0 Entire; 1 Crenulate or serrate or shallowly undulating Shallowly undulating was seen in species of *Brucea*, in which the margin is wavy on the same plane as the leaf blade (as oppose to undulations perpendicular to the leaf blade in character 14).
13	Margin revolute	0 Flat or slightly revolute; 1 Strongly revolute Strongly revolute is when the margin curls at least 180° i.e. completely back on itself or more.
14	Margin undulation	0 Flat; 1 Finely undulating Only *Picrolemma* showed the finely undulating character state, where the frequency of vertical undulation was much higher than any other taxon.
15	Mature leaf hairs abaxially	0 Glabrous to sparsely hairy; 1 Densely hairy on lamina For the very few taxa in which there was not a clear distinction between the states due to moderate and varying degrees of pubescence, the character was coded as ambiguous.
16	Mature leaf hairs adaxially	0 Glabrous to sparsely pubescent; 1 Densely pubescent See mature leaf hairs abaxially.
17	Leaf axes hairs	0 Glabrous or occasional hairs; 1 Hairy Leaf axes refers to petioles, petiolules and the rachis. See characters 15 and 16.
18	Glands	0 Absent; 1 Present Refers to all types of glands seen.
19	Gland type	0 Punctate, light colored; 1 Prominent, usually dark colored Prominent refers to dot-like or raised glands, rather than the circular, pitted glands which usually had a lighter coloration.
20	Punctate leaf gland location	0 Scattered below; 1 In vein forks below; 2 Above, marginal towards apex; 3 Above, marginal towards base; 4 Above, around entire margin; 5 Underneath, at base Vein forks are where secondary venation running out from the mid-rib divides in two, typically as the vein loops near the margin. This is opposed to marginal glands which were not obviously associated with the venation. At the base underneath refers to two or three cases where three to four glands were situated in the attenuate portion of the lamina base.
21	Petiolule	0 Sessile or subsessile; 1 Petiolulate
22	Rachis jointed	0 Not jointed; 1 Jointed
23	Rachis wing	0 Absent; 1 Present
24	Secondaries prominence above	0 Indistinct; 1 Sulcate or flush but still visible; 2 Prominent or clearly visible Flush refers to faint lines which appear neither sunken in the lamina or raised, but are darker or lighter than the lamina.
25	Secondaries prominence	0 Indistinct; 1 Sulcate; 2 Slightly sulcate, flush or slightly raised; 3 Prominent

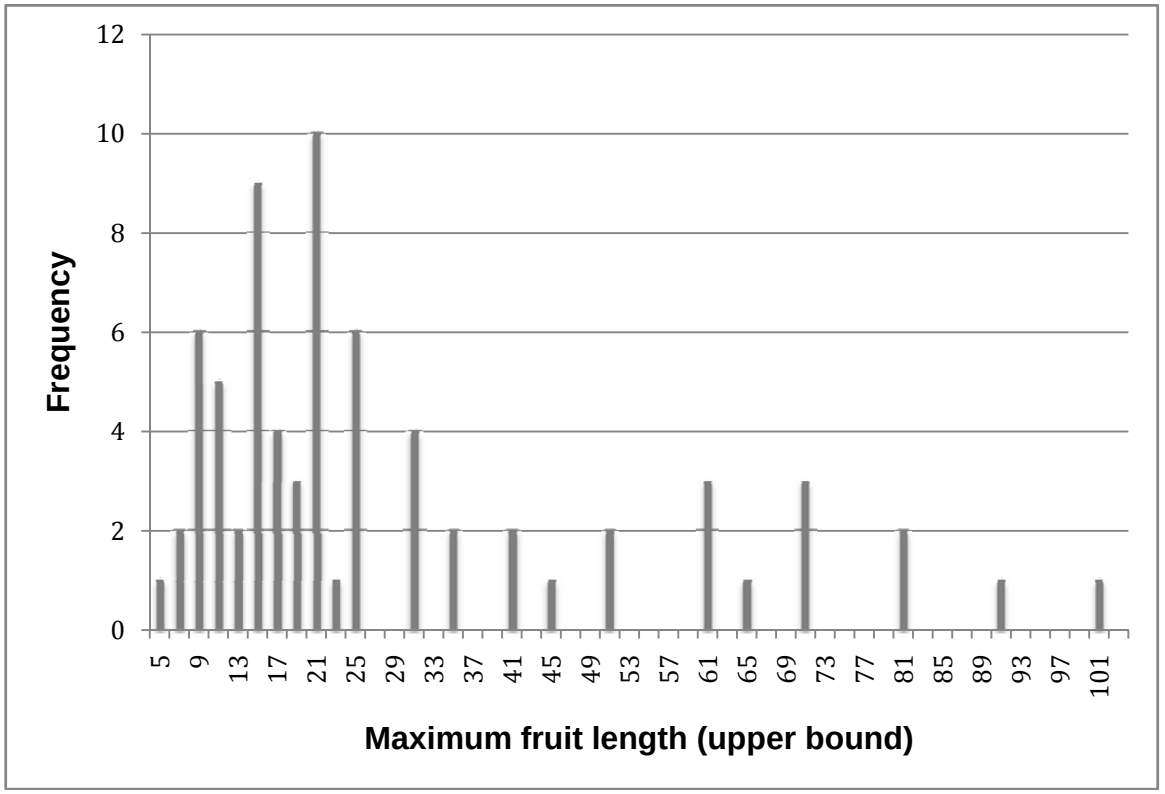
	beneath	
		See character 24.
26	Stipules	0 Absent; 1 Present
27	Inflorescence type	0 Typical thyrse; 1 Fascicle-like in leaf axils; 2 Umbellate; 3 Short, broad, rounded thyrse
28	Typical thyrse structure	0 Second axes below absent (flowers appearing fasciculate) or short and pedicel-like; 1 Second axes similar to primary axis
29	Position of inflorescence	0 Axillary; 1 Terminal
30	Inflorescence axis hairs (incl. pedicel)	0 Glabrous or very sparsely hairy; 1 Moderately to densely hairy
		See character 15.
31	Hair type	0 Simple; 1 Glandular Glandular refers solely to the glandular-capitate hairs of *Eurycoma*.
32	Pedicel articulation	0 None or at base; 1 Halfway along pedicel
33	Sex	0 Bisexual; 1 Unisexual (or androdioecious) Although some unisexual plants were reported as having perfect flowers or being androdioecious in the literature, this character was coded on herbarium specimens where possible, and unisexual was used if perfect flowers were rare.
34	Merosity	0 Trimerous; 1 Tetramerous; 2 Pentamerous; 3 Other Other refers to *Castela* and *Iridosma*.
35	Calyx shape	0 Slightly fused at base to halfway; 1 Almost entirely fused, triangular very short sepals; 2 Ripping into uneven lobes
36	Calyx glands	0 Absent; 1 Present
37	Calyx hairs	0 Glabrous; 1 Glabrous except for marginal hairs, ciliolate; 2 Pubescent abaxially only; 3 Pubescent adaxially only; 3 Pubescent on both sides and margins Pubescence here refers to any type of hair, as taxa are reported to by villous, strigose, puberulent and tomentose as well as pubescent.
38	Petal hair	0 Glabrous, or occasional hair; 1 Pubescent adaxially; 2 Pubescent abaxially; 3 Pubescent on both sides See character 37.
39	Petal shape	0 Oblong, linear; 1 Ovate, obovate Oblong or linear is considered a special case of petal shape, where petals are strap-like. Ovate, obovate refers to all other shapes observed.
40	Corolla length	0 ‘0 – 2 mm'; 1 '2.1 – 8 mm'; 2 '9 – 14 mm'; 3 '17+ mm' States 2 and 3 were delimited as gaps in the distribution of corolla length, with *Quassia amara* flowers reaching a maximum of 50 mm in length. States 0 and 1 had no obvious break, however, on observation of the taxa making up the lower end of the distribution, 15 of the 16 taxa with mean corolla lengths equal to or less than 2 mm were members of

Clade III. Only two members of Clade III had corollas longer than 2.5 mm, and never did they exceed 3 mm. Four members of Clade III had no data on flower length.



41	Corolla aestivation	0 Valvate; 1 Induplicate-valvate; 2 Imbricate; 3 Contorted
42	Petal color (abaxial)	0 White-yellow-green; 1 Pink-red-violet Petal color could typically vary between white, creamy white, pale yellow or green between specimens of the same species and between descriptions of the same species, similarly for pink, red and violet. However, only in one or two instances did color vary between states 1 and 2, and these were coded as ambiguous.
43	Disc hair	0 Glabrous; 1 Pubescent
44	Disc shape	0 Very short, ring-like, or absent; 1 Fleshy, lobed, broader than tall; 2 Columnar, or occ. gynoecium embedded inside
45	Androecium	0 Haplostemonous; 1 Diplostemonous; 2 Pleiostemonous,

46	Filament hairiness	0 Glabrous; 1 Sparsely to densely hairy below, sparse or glabrous above
47	Filament appendage	0 Absent; 1 Present
48	Appendage fusion	0 Mostly adnate, just the tip free; 1 Substantial free portion Free portion free over half of the total length of the filament or more.
49	Dehiscence	0 Latrorse to extrorse; 1 Introrse
50	Staminodes in staminate flower	0 Absent; 1 Present
51	Number of carpels	0 Six or more; 1 One; 2 Two; 3 Three; 4 Four; 5 Five
52	Fusion of ovary	0 Free or fused at base; 1 Fused except at apex Fusion up to the apex refers to the gynoecia of *Soulamea* only.
53	Style length	0 Free portion shorter, fused along length; 1 Free portion equal or longer, fused at top of ovary In taxa with free portion shorter, styles were typically much longer than in state 1, and were associated with capitate or very shortly lobed stigmatic branches.
54	Stigmatic branch shape	0 Reniform, fleshy, thick; 1 Linear, elongate, upright; 2 Linear, elongate, recurved; 3 Inconspicuous to slightly lobed
55	Fruit shape	0 Ellipsoid, obovoid or slightly ovoid; 1 Strongly ovoid; 2 Globose; 3 *Samadera* shape; 4 Elongate (*Ailanthus*); 5 Obcordate (*Soulamea*) Strongly ovoid was restricted to the fruits of *Brucea*.
56	Pericarp	0 Dry; 1 Fleshy
57	Fruiting gynophore	0 Not or slightly enlarging; 1 Enlarging considerably, fleshy
58	Maximum drupe size	0 '≤ 13 mm'; 1 '13.1 – 25 mm'; 2 '> 29 mm', Maximum drupe size was used rather than a mean because fruits in a number of specimens may not have reached maturity, and the measurement was preferred to be as close to mature size as possible. The break between states 1 and 2 was according to a gap in the distribution, and segregated the very large fruits such as *Gymnostemon* and *Pierreodendron*. State delimitation between 0 and 1 was more ambiguous. However, it was useful to break up the large range (0 – 25 mm) into smaller ranges for observation of fruit evolution in character mapping. These states should not be regarded as robust in terms of phylogenetic analysis.



59	Bicarination	0 None; 1 Present but fruit not flattened; 2 Present, fruit flattened
60	Drupe color when ripe	0 Orange to red to blue-black; 1 Green to yellow
61	Hairs on fruit	0 Glabrous or with occasional hairs; 1 Pubescent
62	Fruit vascular bundle	0 Towards seed; 1 Intramarginal *Ailanthus* only.
63	Stylar scar position	0 Level with seed centre; 1 Above seed; 2 Below seed *Ailanthus* only.
64	Endocarp surface	0 Crustaceous, smooth; 1 Reticulately wrinkled
65	Endocarp type	0 One layer of sclereids; 1 Two layer of sclereids

		Data from Fernando and Quinn (1992).
66	Crystalliferous cells	0 Absent; 1 Present
		Data from Fernando and Quinn (1992).
67	Mesocarp secretory canals	0 Absent; 1 Present
		Data from Fernando and Quinn (1992).
68	Pollen amb shape	0 Circular to subcircular; 1 Trangular to subtriangular
		Data from Basak (1963, 1967), Moncada and Machado (1987) and Zavada and Dilcher (1986)
69	Pollen exine	0 Reticulate; 1 Striato-reticulate; 2 Striate; 3 Smooth
		Data from Basak (1963, 1967), Moncada and Machado (1987) and Zavada and Dilcher (1986)
70	Pollen ora	0 Lalongate; 1 Square
		Data from Basak (1963, 1967), Moncada and Machado (1987) and Zavada and Dilcher (1986)
71	Pollen aperturate	0 Planaperturate; 1 Colpi in middle of flat angles; 2 Angulaperturate
		Data from Basak (1963, 1967), Moncada and Machado (1987) and Zavada and Dilcher (1986)

LIST OF REFERENCES

Abbe, E. C., and T. T. Earle. 1940. Inflorescence, floral anatomy and morphology of *Leitneria floridana*. Bull. Torrey Bot. Club 67:173–193.

Akaike, H. 1974. A new look at statistical model identification. IEEE Transactions on Automatic Control 19:716–723.

Alfaro, M. E., S. Zoller, and F. Lutzoni. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol. Biol. Evol. 20:255–266.

Alfaro, M., F. Santini, and C. D. Brock. 2007. Do reefs drive diversification in marine teleosts? Evidence from the pufferfish and their allies. Evolution 61:2104–2126.

APG II. 2003. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. 141:399–436.

Armbruster, W. S. 1992. Phylogeny and the evolution of plant–animal interactions. Bioscience 42:12–20.

Aubréville, A., and F. Pellegrin. 1937. *Gymnostemon* A. et P., genre nouveau de la Côte d'Ivoire, voisin d'un endémique de Madagascar. Bul. Soc. Bot. Fr. 84:181–184.

Aubréville, A. 1962. Simaroubacées. Pages 33–52 *in* Flore du Gabon (A. Aubréville, N. Hallé, B. Descoings, J. Koechlin, and A. Cavaco, eds.), volume 3. Muséum National d'Histoire Naturelle, Paris, France.

Azuma, H., J. G. Garcia-Franco, V. Rico-Gray, and L. B. Thien. 2001. Molecular phylogeny of the Magnoliaceae: the biogeography of tropical and temperate disjunctions. Am. J. Bot. 88:2275–2285.

Backer, C. A., and R. C. B. Van den Brink. 1965. Flora of Java Vol II Angiospermae. Families 111–160. N.V.P. Noordhoff, Groningen, Netherlands.

Barker, G. M. 2002. Phylogenetic diversity: a quantitative framework for measurement of priority and achievement in biodiversity conservation. Bot. J. Linn. Soc. 76:165–194.

Barker, F. K., and F. M. Lutzoni. 2002. The utility of the incongruence length difference test. Syst. Biol. 51:625–637.

Bartish, I. V., U. Swenson, J. Munzinger, and A. A. Anderberg. 2005. Phylogenetic relationships among New Caledonian Sapotaceae (Ericales): Molecular evidence for generic polyphyly and repeated dispersal. Am. J. Bot. 92:667–673.

Basak, R. K. 1963. Pollen morphology of Indian Simaroubaceae. Bull. Bot. Surv. India 5:381–397.

Basak, R. K. 1967. Studies on the pollen morphology of the Simaroubaceae. Bull. Bot. Surv. India 9:63–67.

Baum, D. A., R. L. Small, and J. F. Wendel. 1998. Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. Syst. Biol. 47:181–207.

Becerra, J. 2005. From The Cover: Timing the origin and expansion of the Mexican tropical dry forest. Proc. Natl. Acad. Sci. USA 102:10919–10923

Bennett, A. W. 1872. Simarubeae. Pages 517–523 *in* Flora of British India (J.D. Hooker, ed.) volume 1, Ranunculaceae to Sapindaceae. L. Reeve and Co, London, UK.

Bennett, M. D., and I. J. Leitch. 2005. Nuclear DNA amounts in angiosperms: progress, problems and prospects. Annals Bot. 95:45–90.

Benvenuti, S. 2007. Weed seed movement and dispersal strategies in the agricultural environment. Weed Biol. Manag. 7:141–157.

Bergsten, J. 2005. A review of long-branch attraction. Cladistics 21:163–193.

Bernadello, L. M., L. B. Stiefkens, and M. A. Piovano. 1990. Números cromosómicos en dicotiledóneas Argentinas. Boletín de la Sociedad Argentina de Botánica. 26:149–157

Berry, E. W. 1916. The physical conditions and age indicated by the flora of the Alum Bluff Formation. Profess. Pap. U.S. Geol. Surv. 98E:41–59.

Bradford, J. C. 2002. Molecular phylogenetics and morphological evolution in Cunonieae (Cunoniaceae). Ann. Mo. Bot. Gard. 89:491–503.

Brandley, M. C., A. Schmitz, and T. W. Reeder. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of Scincid Lizards. Syst. Biol. 54:373–390.

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

Cannon, C. H., and P. S. Manos. 2003. Phylogeography of the Southeast Asian stone oaks (*Lithocarpus*). J. Biogeogr. 30:211–226.

Cavalcante, P. 1983. Revisão taxonômica do gênero *Simaba* Aublet (Simaroubaceae) na América do Sul. Publicações Avulsas do Museu Goeldi 37. Belém, Pará, Brasil.

Chan, K., and B. R. Moore. 2002. Whole-tree methods for detecting differential diversification rates. Syst. Biol. 51:855–865.

Chan, K., and B. R. Moore. 2005. SymmeTREE: whole-tree analysis of differential diversification rates. Bioinformatics 21:1709–1710.

Chase, M. W., Morton, C. M., and J. A. Kallunki. 1999. Phylogenetic relationships of Rutaceae: A cladistic analysis of the subfamilies using evidence from *rbcL* and *atpB* sequence variation. Am. J. Bot. 86:1191–1199.

Clayton, J. W., E. S. Fernando, P. S. Soltis, and D. E. Soltis. 2007. Molecular phylogeny of the Tree-of-Heaven family (Simaroubaceae) based on chloroplast and nuclear markers. Int. J. Plant Sci. 168:1325–1339.

Coddington, J. 1988. Cladistic tests of adaptational hypotheses. Cladistics 10:229–258.

Conti, E., T. Eriksson, J. Schönenberger, K. J. Sytsma, and D. A. Baum. 2002. Early Tertiary out-of-India dispersal of Crypteroniaceae: evidence from phylogeny and molecular dating. Evolution 56:1931–1942.

Corbett, S. L., and S. R. Manchester. 2004. Phytogeography and fossil history of *Ailanthus* (Simaroubaceae). Int. J. Plant Sci. 165:671–690.

Cracraft, J. 1985. Biological diversification and its causes. Ann. Mo. Bot. Gard. 72:794–822.

Cronk, Q. C. B., M. Kiehn, W. L. Wagner, and J. F. Smith. 2005. Evolution of *Cyrtandra* (Gesneriaceae) in the Pacific Ocean: the origin of a supertramp clade. Am. J. Bot. 92:1017–1024.

Cronquist, A. 1944a. Studies in the Simaroubaceae I: The genus *Castela*. J. Arnold Arboretum 25:122–128.

Cronquist, A. 1944b. Studies in the Simaroubaceae II: The genus *Simarouba*. Bull. Torrey Bot. Club 71:226–234.

Cronquist, A. 1944c. Studies in the Simaroubaceae III: The genus *Simaba*. Lloydia 7:81–92.

Cronquist, A. 1944d. Studies in the Simaroubaceae IV: Resume of American genera. Brittonia 5:128–147.

Cronquist, A. 1981. An integrated system of classification of flowering plants. Columbia University Press, New York, USA.

Cummings, M. P., S. A. Handley, D. S. Myers, D. L. Reed, A. Rokas, and K. Winka. 2003. Comparing bootstrap and posterior probability values in the four-taxon case. Syst. Biol. 52:477–487.

Cunningham, C. W., K. E. Omland, and T. H. Oakley. 1998. Reconstructing ancestral character states: a critical reappraisal. Trends Ecol. Evol. 13:361–366.

Czaja, A. T. 1978. Stärke und stärkespeicherung bei Gefässpflanzen. Gustav Fischer Verlag, Stuttgart, Germany.

Da Silva, M. F. D. G. F., and O. R. Gottlieb. 1987. Evolution of quassinoids and limonoids in the Rutales. Biochem. Syst. Ecol. 15:85–103.

Darlu, P., and G. Lecointre. 2002. When does the incongruence length difference test fail? Mol. Biol. Evol. 19:432–437.

Davis, C. C., C. D. Bell, S. Mathews, and M. J. Donoghue. 2002a. Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. Proc. Natl. Acad. Sci. USA 99:6833–6837.

Davis, C. C., C. D. Bell, P. W. Fritsch, and S. Mathews. 2002b. Phylogeny of *Acridocarpus-Brachylophon* (Malpighiaceae): Implications for Tertiary tropical floras and Afroasian biogeography. Evolution 56:2395–2404.

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

De Queiroz, A. 2002. Contingent predictability in evolution: key traits and diversification. Syst. Biol. 51:917–929.

De Queiroz, A. 2005. The resurrection of oceanic dispersal in historical biogeography. Trends Ecol. Evol. 20:68–73.

Dick, C., and S. Wright. 2005. Tropical mountain cradles of dry forest diversity. Proc. Natl. Acad. Sci. USA 102:10757–10758.

Dodd, M. E., J. Silvertown, and M. W. Chase. 1999. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. Evolution 53:732–744.

Donoghue, M. J., C. D. Bell, and J. Li. 2001. Phylogenetic patterns in Northern Hemisphere plant geography. Int. J. Pl. Sci. 162:S41–S52.

Donoghue, M. J., and B. R. Moore. 2003. Toward an integrative historical biogeography. Integr. Comp. Biol. 43:261–270.

Donoghue, M. J., and S. A. Smith. 2004. Patterns in the assembly of temperate forests around the Northern Hemisphere. Philos T. Roy. Soc. B 359:1633–1644.

Donoghue, P. C. J., and M. J. Benton. 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. Trends Ecol. Evol. 22:424–431.

Dorofeev, P. I. 1994. *Leitneria*. Pages 8–12 *in* Fossil flowering plants of Russia and adjacent states (L. Budantsev, ed.), volume 3, Leitneriaceae-Juglandaceae. Kamarov Botanical Institute, Russian Academy of Sciences, St. Petersburg, Russia.

Doyle, J. J., and J. L. Doyle. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13–15.

Dreyer, D. L. 1983. Limonoids of the Rutaceae. Pages 215–246 *in* Chemistry and chemical taxonomy of the Rutales (P. G. Waterman, and M. F. Grundon, eds.). Academic Press Inc., London, UK.

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

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

Edwards, A. W. F. 1992. Likelihood. John Hopkins University Press, Baltimore, USA.

Engler, A. 1874. Simaroubaceae. Pages 197–246, tabs. 40–46 *in* Flora Brasiliensis (C. F. P. de Martius, and I. Urban, eds.), volume 12(2). Frid. Fleischer in comm., Lipsiae, Germany.

Engler, A. 1905. Anacardiaceae africanae III. Botanische Jahrbücher Syst. 36:216–217.

Engler, A. 1931. Simarubaceae. Pages 359–405 *in* Die natürlichen Planzenfamilien 2nd ed. (A. Engler, and K. Prantl, eds.), volume 19a. Engelmann, Leipzig, Germany.

Eriksson, O., and B. Bremer, B. 1991. Fruit characteristics, life forms, and species richness in the plant family Rubiaceae. Am. Nat. 138:751–761.

Eriksson, O., and B. Bremer, B. 1992. Pollination systems, dispersal modes, life forms, and diversification rates in angiosperm families. Evolution 46:258–266.

Erixon, P., B. Svennblad, T. Britton, and B. Oxelman. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst. Biol. 52:665–673.

Faith, D. P. 1992a. Conservation evaluation and phylogenetic diversity. Biol. Conserv. 61:1–10.

Farris, J. S., M. Kallersjo, A. G. Kluge, and C. Bult. 1994. Testing significance of incongruence. Cladistics 10:315–319.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791.

Fernando, E. S., and Quinn, C. J. 1992. Pericarp anatomy and systematics of the Simaroubaceae s.l. Aust. J. Bot. 40:263–289.

Fernando, E. S., and Quinn, C. J. 1995. Picramniaceae, a new family, and recircumscription of Simaroubaceae. Taxon 44:177–181.

Fernando, E. S., Gadek, P. A., and Quinn, C. J. 1995. Simaroubaceae, an artificial construct: Evidence from *rbcL* sequence variation. Am. J. Bot. 82:92–103.

Feuillet, C. 1983. Le statut des genres *Quassia* L., *Samadera* Gaertn., *Simaba* Aubl. et *Simarouba* Aubl. (Simaroubaceae). Bull. Jardin Bot. Nat. Belg. 53:510–511.

Franceschinelli, E. V. and K. Yamamoto. 1999. *Simaba docensis*, a new Brazilian species of Simaroubaceae. Novon 9:345–348.

Franceschinelli, E. V., K. Yamamoto and G. J. Shepherd. 1999. Distinctions among three *Simarouba* species. Syst. Bot. 23:479–488.

Franceschinelli, E. V., and W. W. Thomas. 2000. *Simaba guianensis* subsp. *huberi*, a new Venezuelan taxon of Simaroubaceae. Brittonia 52:311–314.

Fukami, T., H. Beaumont, X. Zhang, and P. Rainey. 2007. Immigration history controls diversification in experimental adaptive radiation. Nature 446:436–439.

Funk, D., P. Nosil, and W. Etges. 2006. Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. Proc. Natl. Acad. Sci. USA 103:3209–3213.

Gadek P. A., E. S. Fernando, C. J. Quinn, S. B. Hoot, T. Terrazas, M. C. Sheahan, and M. W. Chase. 1996. Sapindales: Molecular delimitation and infraordinal groups. Am. J. Bot. 83:802–811.

Gentry, A. H. 1982. Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? Ann. Mo. Bot. Gard. 69: 557–593.

Giannasi, D. E. 1986. Phytochemical aspects of phylogeny in Hamamelidae. Ann. Mo. Bot. Gard. 73:417–437.

Gilbert, G. 1958. Simaroubaceae. Pages 119–131 *in* Flore du Congo Belge et du Ruanda-Urundi (Comité exécutif de la Flore du Congo belge et le Jardin botanique de l'État), volume VII. I.N.É.A.C., Brussels, Belgium.

Givnish, T., and S. Renner. 2004. Tropical intercontinental disjunctions: Gondwana breakup, immigration from the boreotropics, and transoceanic dispersal. Int. J. Plant Sci. 165:S1–S6.

Good-Avila, S., V. Souza, B. Gaut, and L. Eguiarte. 2006. Timing and rate of speciation in *Agave* (Agavaceae). Proc. Natl. Acad. Sci. USA 103:9124–9129.

Gottschling, M., N. Diane, H. H. Hilger, and M. Weigend. 2004. Testing hypotheses on disjunctions present in the primarily woody Boraginales: Ehretiaceae, Cordiaceae, and Heliotropiaceae, inferred from ITS1 sequence data. Int. J. Plant Sci. 165:S123–S135.

Gradstein, F. M., and J. G. Ogg. 2004. Geologic Time Scale 2004 – why, how, and where next! Lethaia 37:175–181.

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

Gray A. 1859. Diagnostic characters of phanerogamous plants collected by Charles Wright, botanist of the U.S. North Pacific Exploring Expedition, with observations upon the relations of the Japanese flora to that of North America, and other parts of the Northern Temperate Zone. Mem. Am. Acad. Arts Sci. NS 6:377–453.

Gray, A. I. 1983. Structural diversity and distribution of coumarins and chromones in the Rutales. Pages 97–146 *in* Chemistry and chemical taxonomy of the Rutales (P. G. Waterman, and M. F. Grundon, eds.). Academic Press Inc., London, UK.

Graybeal, A. 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? Syst. Biol. 47:9–17.

Hahn, W. J., and W. W. Thomas. 2001. Simaroubaceae DC. Pages 2368–2372 *in* Flora de Nicaragua (W. D. Stevens, C. U. Ulloa, A. Pool, and O. M. Montiel, eds.), volume 85, part III. Missouri Botanical Garden Press, Missouri, USA.

Harborne, J. B. 1983. The flavonoids of the Rutales. Pages 147–174 *in* Chemistry and chemical taxonomy of the Rutales (P. G. Waterman, and M. F. Grundon, eds.). Academic Press Inc., London, UK.

Hardesty, B. D., C.W. Dick, A. Kremer, S. Hubbell, and E. Bermingham. 2005. Spatial genetic structure of *Simarouba amara* Aubl. (Simaroubaceae), a dioecious, animal-dispersed Neotropical tree, on Barro Colorado Island, Panama. Heredity 95:290–297.

Harmon, L., J. Weir, C. Brock and R. Glor. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics 24:129–131.

Hedtke, S. M., T. M. Townsend, and D. M. Hillis. 2006. Resolution of phylogenetic conflict in large data sets by increased taxon sampling. Syst. Biol. 55:522–529.

Hegnauer, R. 1983. Chemical characters and the classification of the Rutales. Pages 401–440 *in* Chemistry and chemical taxonomy of the Rutales (P. G. Waterman, and M. F. Grundon, eds.). Academic Press Inc., London, UK.

Herrera, C. 1989. Seed dispersal by animals: a role in angiosperm diversification? Am. Nat. 133:309–322.

Hewson, H. J. 1985. Simaroubaceae. Pages 188–197 *in* Flora of Australia (A. S. George, ed.), volume 25. Australian Government Publishing Service Canberra, Australia.

Hipp, A. L., J. C. Hall, and K. J. Sytsma. 2004. Congruence versus phylogenetic accuracy: Revisiting the incongruence length difference test. Syst. Biol. 53:81–89.

Hodges, S., and M. Arnold. 1995. Spurring Plant Diversification: Are Floral Nectar Spurs a Key Innovation? Proc. Royal Soc. London B 262:343–348.

Hoot, S.B, A. Culham, and P.R. Crane. 1995. The utility of *atpB* gene sequences in resolving phylogenetic relationships: Comparison with *rbcL* and 18S ribosomal DNA sequences in the Lardizabalaceae. Ann. Mo. Bot. Gard. 82:194–207.

Huelsenbeck, J.P., J. J. Bull, and C.W. Cunningham. 1996. Combining data in phylogenetic analysis. Trends Ecol. Evol. 11:152–158.

Huelsenbeck, J.P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754–755.

Huelsenbeck, J. P., R. Nielsen, and J. P. Bollback. 2003. Stochastic Mapping of Morphological Characters. Syst. Biol. 52:131–158.

Hughes, N. F. 1994. The enigma of angiosperm origins. Cambridge University Press, Cambridge, UK.

Hughes, C., and R. Eastwood. 2006. Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. Proc. Natl. Acad. Sci. USA 103:10334–10339.

Hutchinson, J., and J. M. Dalziel. 1954. Flora of west tropical Africa, 2nd edition, volume 1. Crown Agents, London, UK.

Jaffre, T., and J. Fambart. 2002. Quatre nouvelles espèces de *Soulamea* (Simaroubaceae) de Nouvelle-Calédonie. Adansonia 24:159–168.

Johnson, L.A., and D.E. Soltis. 1998. Assessing congruence: Empirical examples from molecular data. Pages 297–348 *in* Molecular systematics of plants II: DNA sequencing (D.E. Soltis, P.S. Soltis, and J.J. Doyle, eds.). Kluwer, Norwell, Massachusetts, USA.

Judd, W. S., C. S. Campbell, E. A. Kellogg, P. F. Stevens, and M. J. Donoghue. 2008. Plant Systematics: A Phylogenetic Approach, 3rd edition. Sinauer Associates Inc., Sunderland, Massachusetts, USA.

Källersjö, M., J. S. Farris, M. W. Chase, B. Bremer, M. F. Fay, C. J. Humphries, G. Petersen, O. Seberg, and K. Bremer. 1998. Simultaneous parsimony jackknife analysis of 2538 *rbcL* DNA sequences reveals support for major clades of green plants, land plants, and flowering plants. Plant. Syst. Evol. 213:259–287.

Kass, R.E., and A.E. Raftery. 1995. Bayes factors. J. Am. Stat. Assoc. 90:773–795.

Kathriarachchi, H., P. Hoffmann, R. Samuel, K. J. Wurdack, and M. W. Chase. 2005. Molecular phylogenetics of Phyllanthaceae inferred from five genes (plastid *atpB, matK*, 3'*ndhF*, *rbcL*, and nuclear *PHYC*). Mol. Phylogenet. Evol. 36:112–134.

Kay, K. M., P. A. Reeves, R. G. Olmstead, and D. W. Schemske. 2005. Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. Am. J. Bot. 92: 1899–1910.

Kelchner, S.A., and M. A. Thomas. 2007. Model use in phylogenetics: nine key questions. Trends Ecol. Evol. 22:87–94.

Klocke, J.A., M. Arisawa, S. S. Handa, A. D. Kinghorn, G. A. Cordell, and N. R. Farnsworth. 1985. Growth inhibitory, insecticidal and antifeedant effects of some antileukemic and cytotoxic quassinoids on two species of agricultural pests. Cell. Mol. Life. Sci. 4:379–382.

Lavin, M., M. Thulin, J. -N. Labat, and R.T. Pennington. 2000. Africa, the odd man out: Molecular biogeography of dalbergioid legumes (Fabaceae) suggests otherwise. Syst. Bot. 25:449–467.

Lavin, M., M. F. Wojciechowski, P. Gasson, C. H. Hughes, and E. Wheeler. 2003. Phylogeny of robinioid legumes (Fabaceae) revisited: *Coursetia* and *Gliricidia* recircumscribed, and a biogeographical appraisal of the Caribbean endemics. Syst. Bot. 28:387–409.

Lavin, M., B. P. Schrire, G. Lewis, R. T. Pennington, A. Delgado-Salinas, M. Thulin, C. E. Hughes, A. B. Matos, and M. F. Wojciechowski. 2004. Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. Philos T. Roy. Soc. B 359:1509–1522.

Lecointre, G., H. Philippe, H. L. Van Le, and H. Guyader. 1993. Species sampling has a major impact on phylogenetic inference. Mol. Phylogenet. Evol. 2:205–224.

Lemmon, A. R., and E. C. Moriarty. 2004. The importance of proper model assumption in Bayesian phylogenetics. Syst. Biol. 53:265–277.

Leschen, R. A. B., and T. R. Buckley. 2007. Multistate Characters and Diet Shifts: Evolution of Erotylidae (Coleoptera). Syst. Biol. 56:97–112.

Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50:913–925.

Lewis, P. O., M. T. Holder, and K. E. Holsinger. 2005. Polytomies and Bayesian phylogenetic inference. Syst. Biol. 54:241–253.

Li, H-.L. 1952. Floristic relationships between eastern Asia and eastern North America. Trans. Am. Phil. Soc. 42:371–429.

Li, H-.L. 1977. Simarubaceae. Pages 538–543 *in* Flora of Taiwan (H-.L. Li, T-. S. Liu, T-. C. Huang, T. Koyama, and C. E. DeVol, eds.), volume 3. Epoch Publishing Co., Ltd., Taipei, Taiwan, Republic of China.

Luckow, M., and A. Bruneau. 1997. Circularity and Independence in Phylogenetic Tests of Ecological Hypotheses. Cladistics 13:145–151.

Mabberley, D. 1997. The Plant Book. Cambridge University Press, Cambridge, UK.

Machen, J. 1971. Plant microfossils from Tertiary deposits of the Isle of Wight. New Phytol. 70:851–872.

Maddison, D. R., and W. P. Maddison. 2004. MacClade 4: Analysis of phylogeny and character evolution. Version 4.08. Sinauer Associates, Sunderland, MA.

Magallón, S., and M. J. Sanderson. 2001. Absolute diversification rates in angiosperm clades. Evolution 55:1762–1780.

Mathews, S., M. Lavin, and R. A. Sharrock. 1995. Evolution of the phytochrome gene family and its utility for phylogenetic analyses of angiosperms. Ann. Mo. Bot. Gard. 82:296–321.

Mauritzon, J. 1935. Kritic von J. Wiger's Abhandlung "Embryological studies on the families Buxaceae, Meliaceae, Simarubaceae and Burseraceae. Bot. Not. 1935:490–502.

Meimberg, H., A. Wistuba, P. Dittrich, and G. Heubl. 2001. Molecular phylogeny of Nepenthaceae based on cladistic analysis of plastid trnK intron sequence data. Plant Biology 3:164–175.

Mester, I. 1983. Structural diversity and distribution of alkaloids in the Rutales. Pages 31–96 *in* Chemistry and chemical taxonomy of the Rutales (P. G. Waterman, and M. F. Grundon, eds.). Academic Press Inc., London, UK.

Mittelbach, G., D. Schemske, H. Cornell, and A. Allen. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecol. Lett. 10:315–331.

Moncada, M., and S. Machado. 1987. Los granos de polen de Simarubaceae. Acta. Bot. Cub. 45:1–7.

Moore, B. R., K. Chan, and M. J. Donoghue. 2004. Detecting diversification rate variation in supertrees. Pages 487–533 *in* Phylogenetic supertrees: combining information to reveal the Tree of Life (O. R. P. Bininda-Edmonds, ed.). Kluwer Academic Publishers, Dordrecht, The Netherlands.

Moore, B. R., Q. -Y. Xiang, J. Thorne, D. Thomas, and M. J. Donoghue. 2006. Biogeographic and biological correlates of diversification in dogwoods: Dispersal and evolution of fruits and inflorescences in space and time. Evolution Meeting (Stonybrook, NY).

Moore, B. R., and M. J. Donoghue. 2007. Correlates of diversification in the plant clade Dipsacales: geographic movement and evolutionary innovations. Am. Nat. 170:S28–S55.

Moore, M. J., A. Tye, and R. K. Jansen. 2006. Patterns of long-distance dispersal in *Tiquilia* subg. *Tiquilia* (Boraginaceae): implications for the origins of amphitropical disjuncts and Galapagos Islands endemics. Am. J. Bot. 93:1163–1177.

Morley, R. J. 2003. Interplate dispersal paths for megathermal angiosperms. Perspect. Plant Ecol. 6:5–20.

Morris, A. B., C. D. Bell, J. W. Clayton, W. S. Judd, D. E. Soltis, and P. S. Soltis. 2007. Phylogeny and divergence time estimation in *Illicium* with implications for New World biogeography. Syst. Bot. 32:236–249.

Muellner, A. N., R. Samuel, S. A. Johnson, M. Cheek, T. D. Pennington, and M.W. Chase. 2003. Molecular phylogenetics of Meliaceae (Sapindales) based on nuclear and plastid DNA sequences. Am. J. Bot. 90:471–480.

Muellner, A. N., V. Savolainen, R. Samuel, and M. W. Chase. 2006. The mahogany family "out-of-Africa": Divergence time estimation, global biogeographic patterns inferred from plastid *rbcL* DNA sequences, extant, and fossil distribution of diversity. Mol. Phylogenet. Evol. 40:236–250.

Muellner, A. N., D. D. Vassiliades, and S. Renner. 2007. Placing Biebersteiniaceae, a herbaceous clade of Sapindales, in a temporal and geographic context. Plant Syst. Evol. 266:233–252.

Nair, N. C., and T. C. Joseph. 1957. Floral morphology and embryology of *Samadera indica*. Bot. Gaz. 119:104–115.

Nair, N. C., and R. K. Joshi. 1958. Floral morphology of some members of the Simaroubaceae. Bot. Gaz. 120:88–99.

Nair, N. C., and N. P. Sukumaran. 1960. Floral morphology and embryology of *Brucea amarissima*. Bot. Gaz. 121:175–185.

Narayana, L. L. 1957. Embryology of Simaroubaceae. Curr. Sci. 26:323–324.

Near, T. J., and M. J. Sanderson. 2004. Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. Philos. T. Roy. Soc. B 359:1477–1483.

Nee, S., A. Ø. Mooers, and P. H. Harvey. 1992. Tempo and mode of evolution revealed from molecular phylogenies. Proc. Natl. Acad. Sci. USA 89:8322–8326.

Linsbauer, K. 1926. Handbuch der Pflanzenanatomie, volume 2(X). Borntraeger, Berlin, Germany.

Nikitin, V. P. 2006a. Paleocarpology and stratigraphy of the Paleogene and Neogene strata in Asian Russia. Academic Publishing House "Geo", Novosibirsk, Russia.

Nooteboom, H. P. 1962a. Generic delimitation in Simaroubaceae Tribus Simaroubeae and a conspectus of the genus *Quassia* L. Blumea 11:509–528.

Nooteboom, H. P. 1962b. Simaroubaceae. Flora Malesiana Series I 6:193–226.

Nooteboom, H. P. 1987. *Laumoniera*, a new genus of Simaroubaceae from Sumatra. Blumea 32:383–384.

Nylander, J. A. A., F. Ronquist, J. P. Huelsenbeck, and J. L. Nieves-Aldrey. 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53:47–67.

Ollerton, J., S. D. Johnson, and A. B. Hingston. 2006. Geographical variation in diversity and specificity of pollination systems. Pages 283 – 308 *in* Plant-pollinator interactions: from specialization to generalization (N. M. Waser, and J. Ollerton, eds.). University of Chicago Press, Chicago, Illinois, USA.

Olmstead, R. G., and J. A. Sweere. 1994. Combining data in phylogenetic systematics: An empirical approach using three molecular data sets in the Solanaceae. Syst. Biol. 43:467–481.

Pennington, R.T. 1996. Molecular and morphological data provide phylogenetic resolution at different hierarchical levels in *Andira*. Syst. Biol. 45:496–515.

Pennington, R. T., and C. W. Dick. 2004. The role of immigrants in the assembly of the South American rainforest tree flora. Philos. T. Roy. Soc. B 359:1611–1622

Pennington, R. T., Q. C. B. Cronk, and J. E. Richardson. 2004a. Introduction and synthesis: plant phylogeny and the origin of major biomes. Philos. T. Roy. Soc. B 359:1455–1464.

Pennington, R. T., M. Lavin, D. E. Prado, C. A. Pendry, S. K. Pell, and C. A. Butterworth. 2004b. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. Philos. T. Roy. Soc. B 359:515–538.

Pennington, R. T., J. E. Richardson, and M. Lavin. 2006. Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. New. Phytol. 172:605–616.

Perrier de la Bathie, H. 1950. Simarubacées. Pages 1–7 *in* Flore de Madagascar (H. Humbert ed.), volume 105. Firmin-Didot and Co., Paris, France.

Pfeiffer, W. M. 1912. The morphology of *Leitneria floridana*. Bot. Gaz. 53:189–203.

Pierre, L. 1896 Plantes du Gabon. Bull Soc Linn Paris 2:1236.

Pirani, J. R. 1987. Simaroubaceae. Pages 7–27 *in* Flora del Paraguay (R. Spichiger, ed.). Missouri Botanical Garden, Missouri, USA.

Pirani, J. R. 1997. Simaroubáceas. Pages 1–45 *in* Flora Ilustrada Catarinense (A. Reis, ed.). Itajaí, Santa Catarina, Brasil.

Plana, V. 2004. Mechanisms and tempo of evolution in the African Guineo-Congolian rainforest. Philos. T. Roy. Soc. B 359:1585–1594.

Polonsky, J. 1983. Chemistry and biological activity of the quassinoids. Pages 247–266 *in* Chemistry and chemical taxonomy of the Rutales (P. G. Waterman, and M. F. Grundon, eds.). Academic Press Inc., London, UK.

Porter, D. M. 1971. Simaroubaceae. Pages 754–756 *in* Flora of the Galápagos Islands (I. L. Wiggins, and D. M. Porter, eds.). Stanford University Press, Stanford, California, USA.

Porter, D. M. 1973. Flora of Panama, Family 90, Simaroubaceae. Ann. Mo. Bot. Gard. 60:23–39.

Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818.

Posada D., and T. R. Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53:793–808.

Poux, C., O. Madsen, E. Marquard, D. R. Vieites, W. W. de Jong, and M. Vences. 2005. Asynchronous colonization of Madagascar by the four endemic clades of primates, tenrecs, carnivores, and rodents as inferred from nuclear genes. Syst. Biol. 54:719–730.

Purvis, A., S. Nee and P. H. Harvey. 1995. Macroevolutionary inferences from primate phylogeny. Proc. Royal Soc. London B 260:329–333.

Pybus, O., and P. H. Harvey. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. Proc. Royal Soc. London B 267:2267–2272.

Rabosky, D. 2006. Likelihood methods for detecting temporal shifts in diversification rates. Evolution 60:1152–1164.

Rambaut, A., and A. J. Drummond. 2003. Tracer v1.3. Available from http://evolve.zoo.ox.ac.uk/.

Rannala, B. 2002. Identifiability of parameters in MCMC Bayesian inference of phylogeny. Syst. Biol. 51:754–760.

Raven, P. H., and D. I. Axelrod. 1974. Angiosperm biogeography and past continental movements. Ann. Mo. Bot. Gard. 61:539–673.

Raven, P. H. 1975. The bases of angiosperm phylogeny: cytology. Ann. Mo. Bot. Gard. 68:724–764.

Record, S. J., and R. W. Hess. 1943. Timbers of the New World. Yale University Press, New Haven, Conneticut, USA.

Ree, R. 2005. Detecting the historical signature of key innovations using stochastic models of character evolution and cladogenesis. Evolution 59:257–265.

Ree, R. H., B. R. Moore, C. O. Webb, and M. J. Donoghue. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution 59:2299–2311.

Ree, R. H., and S. A. Smith. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57:4–14.

Renner, S. S., G. Clausing, and K. Meyer. 2001. Historical biogeography of Melastomataceae: the roles of Tertiary migration and long-distance dispersal. Am. J. Bot. 88:1290–1300.

Renner, S. S. 2004a. Plant dispersal across the tropical Atlantic by wind and sea currents. Int. J. Plant Sci. 165:S23–S33.

Renner, S. S. 2004b. Multiple Miocene Melastomataceae dispersal between Madagascar, Africa and India. Philos T. Roy. Soc. B 359:1485–1494.

Richardson, J. E., R. T. Pennington, T. Pennington, and P. Hollingsworth. 2001a. Rapid diversification of a species-rich genus of neotropical rain forest trees. Science 293: 2242–2245.

Richardson, J. E., F. M. Weitz, M. F. Fay, Q. C. B. Cronk, H. P. Linder, G. Reeves, and M. W. Chase. 2001b. Rapid and recent origin of species richness in the Cape flora of South Africa. Nature 412:181–183.

Richardson, J. E., L. W. Chatrou, J. B. Mols, R. H. J. Erkens, and M. D. Pirie. 2004. Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. Philos T. Roy. Soc. B 359:1495–1508.

Riddle, B. R. 1996. The molecular phylogeographic bridge between deep and shallow history in continental biotas. Trends. Ecol. Evol. 11:207–211.

Ronquist, F. 1997. Dispersal-Vicariance Analysis: A new approach to the quantification of historical biogeography. Syst. Biol. 46:195–203.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Ronse de Craene, L. P. 2008. Homology and evolution of petals in the core eudicots. Syst. Bot. 33:301–325.

Roth, J. L. and D. L. Dilcher. 1979. Investigations of angiosperms from the Eocene of North America: stipulate leaves of the Rubiaceae including a probable polyploid population. Am. J. Bot. 66:1194–1207.

Roubik, D. W., N. M. Holbrook, and G. V. Parra. 1985. Roles of nectar robbers in reproduction of the tropical treelet *Quassia amara* (Simaroubaceae). Oecologia 66:161–167.

Saarela, J. M., H. S. Rai, J. A. Doyle, P. K. Endress, S. Mathews, A. D. Marchant, B. G. Briggs, and S. W. Graham. 2007. Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree. Nature 446:312–315.

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.

Sanmartin, I., H. Enghoff, and F. Ronquist. 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biol. J. Linn. Soc. 73:345–390.

Santiago-Valentin, E., and R. G. Olmstead. 2004. Historical biogeography of Caribbean plants: introduction to current knowledge and possibilities from a phylogenetic perspective. Taxon 53:299–319.

Savolainen, V., M. W. Chase, S. B. Hoot, C. M. Morton, D. E., Soltis, C. Bayer, M. F Fay, A. Y. de Bruijn, S. Sullivan, and Y. -L. Qiu. 2000b. Phylogenetics of flowering plants based on a combined analysis of plastid *atpB* and *rbcL* gene sequences. Syst. Biol. 49:306–362.

Schott, F. A., and J. P. McCreary. 2001. The monsoon circulation of the Indian Ocean. Prog. Oceanogr. 51:1–123.

Scotese, C. R. 2001. Atlas of Earth History. PALEOMAP Project, Arlington, TX. www.scotese.com.

Simpson, B. B., J. A. Tate, and A. Weeks. 2005. The biogeography of *Hoffmannseggia* (Leguminosae, Caesalpinoideae, Caesalpinieae): a tale of many travels. J. Biogeogr. 32:15–27.

Smith, A. C. 1985. Simaroubaceae. Pages 479–489 *in* Flora Vitiensis Nova (A. C. Smith, ed.), volume 3. Pacific Tropical Botanical Garden, Hawaii, USA.

Soltis, D. E., P. S. Soltis, M. E. Mort, M. W. Chase, V. Savolainen, S. B. Hoot, and C. M. Morton. 1998. Inferring complex phylogenies using parsimony: An empirical approach using three large DNA data sets for angiosperms. Syst. Biol. 47:32–43.

Soltis, D. E., P. S. Soltis, M. W. Chase, M. E. Mort, D. C. Albach, M. Zanis, V. Savolainen, W. H. Hahn, S. B. Hoot, M. F. Fay, M. Axtell, S. M. Swensen, L. M. Prince, W. J. Kress, K. C. Nixon, and J. S. Farris. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. Bot. J. Linn. Soc. 133:381–461.

Soltis, D. E., P. S. Soltis, P. K. Endress, and M. W. Chase. 2005. Phylogeny and evolution of angiosperms. Sinauer Associates Inc., Sunderland, Massachusetts, USA.

Song, Z. -C., W. -M. Wang, and F. Huang. 2004. Fossil pollen records of extant angiosperms in China. Bot. Rev. 70:425–458.

Spiekerkoetter, H. 1924. Untersuchungen zur Anatomie und Systematik östafrikanischer Meliaceen, Burseraceen und Simarubaceen. Bot. Arch. 7:274–320.

Stannard, B. 2000. Simaroubaceae. Pages 1–15 *in* Flora of Tropical East Africa (H. J. Beentje, ed.). Royal Botanic Gardens, Kew, London, UK.

Stevens, P. F. 2006. Angiosperm Phylogeny Website v 7. http://www.mobot.org/MOBOT/research/APweb/.

Sullivan, J. 1996 Combining data with different distributions of among-site rate variation. Syst. Biol. 45:375–380.

Sullivan, J., and P. Joyce. 2005. Model selection in phylogenetics. Annu. Rev. Ecol. Evol. Syst. 36:445–466.

Suzuki, Y., G. V. Glazko, and M. Nei. 2002. Overcredibility of molecular phylogenetics obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. USA 99:16138–16143.

Swofford, D. L. and W.P. Maddison. 1987. Reconstructing ancestral character states under Wagner parsimony. Math. Biosci. 87:199–229.

Swofford, D. S. 2002. Phylogenetic analysis using parsimony, version 4.0b10. Sinauer Associates Inc., Sunderland, MA.

Tahktajan, A. 1996. Diversity and classification of flowering plants. Columbia University Press, New York, USA.

Teodoris, V., and Z. Kvacek. 2005. The extinct genus *Chaneya* Wang et Manchester in the Tertiary of Europe – a revision of *Porana*-like fruit remains from Öhningen and Bohemia. Rev. Palaeobot. Palyno. 134:85–103.

Thomas, W. W. 1984. A new species of *Simaba* (Simaroubaceae) from Pará, Brazil, wih a key to the species north of the Amazon river. Brittonia 36:244–247.

Thomas, W. W. 1985. The *Simaba guianensis* complex in northern South America. Acta Amazonica, supl. 15:71–79.

Thomas, W. W. 2002. Simaroubaceae. Pages 685–689 *in* Guide to the vascular plants of Central French Guiana (S. A. Mori, G. Cremers, C. A. Gracie, J-. J. de Granville, S. V. Head, M. Hoff, and J. D. Mitchell), part 2, Dicotyledones. The New York Botanical Garden Press, New York, USA.

Thomson, J., T. Gibson, F. Plewniak, F. Jeanmougin, and D. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignments aided by quality analysis tools. Nucleic Acids Res. 25:4876–4882.

Thorne, J. L., and H. Kishino. 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51:689–702.

Tiffney, B. H. 1985a. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the northern hemisphere. J. Arnold Arboretum 66:243–273.

Trelease, W. 1895. *Leitneria floridana*. Rept. Mo. Bot. Garden 6:1–26 pls. 30–44.

Van der Veken, P. 1960. *Nothospondias* Engl. Simaroubaceae africaine meconnue. Bull. Jardin Botanique Etat Bruxelles 30:105–109.

Von Hagen, K., and J. Kadereit. 2003. The diversification of *Halenia* (Gentianaceae): ecological opportunity versus key innovation. Evolution 57:2507–2518.

Wang, Y., and S. R. Manchester. 2000. *Chaneya*, a new genus of winged fruit from the Tertiary of North America and Eastern Asia. Int. J. Plant Sci. 161:167–178.

Wanntorp, L., and H. E. Wanntorp. 2003. The biogeography of Gunnera L.: vicariance and dispersal. J. Biogeogr. 30:979–987.

Waser, N. 1998. Pollination, angiosperm speciation, and the nature of species boundaries. Oikos 81:198–201.

Waterman, P. G., and M. F. Grundon. 1983. Chemistry and chemical taxonomy of the Rutales. Academic Press Inc., London, UK.

Waterman, P. G. 1983. Phylogenetic implications of the distribution of secondary metabolites within the Rutales. Pages 377–400 *in* Chemistry and chemical taxonomy of the Rutales (P. G. Waterman, and M. F. Grundon, eds.). Academic Press Inc., London, UK.

Webber, I. E. 1936. Systematic anatomy of the woods of the Simarubaceae. Am. J. Bot. 23:577–587.

Weberling, F. 1989. Morphology of flowers and inflorescences. Cambridge University Press, UK.

Weeks, A., D. C. Daly, and B. B. Simpson. 2005. The phylogenetic history and biogeography of the frankincense and myrrh family (Burseraceae) based on nuclear and chloroplast sequence data. Mol. Phylogenet. Evol. 35:85–101.

Whittaker, R. J. 1998. Island biogeography. Oxford University Press, Oxford, UK.

Wiens, J. J., and M. J. Donoghue. 2004. Historical biogeography, ecology and species richness. Trends Ecol. Evol. 19:639–644.

Wiens, J. J., G. Parra-Olea, M. García-París, and D. B. Wake. 2007. Phylogenetic history underlies elevational biodiversity patterns in tropical salamanders. Proc. Royal Soc. London B 274:919–928.

Wiger, J. 1935. Embryological studies in Buxaceae, Meliaceae, Simarubaceae and Burseraceae. Dissertation, University of Lund, Sweden.

Wikström, N., V. Savolainen, and M.W. Chase. 2001. Evolution of the angiosperms: calibrating the family tree. Proc. R. Soc. B 268:2211–2220.

Wikström, N., V. Savolainen, and M. W. Chase. 2004. Angiosperm divergence times: congruence and incongruence between fossils and sequence divergence estimates. Pages 142–165 *in* Telling the evolutionary time: molecular clocks and the fossil record (P.C.J. Donoghue, and M.P. Smith, eds.). Taylor and Francis, London, UK.

Wild, H., and J. B. Phipps. 1963. Simaroubaceae. Pages 210–220 *in* Flora Zambesiaca (A. W. Exell, A. Fernandes, and H. Wild, eds.), volume 2, part 1. Crown Agents for Oversea Governments and Administrations, London, UK.

Wilson, P. G., M. M. O'Brien, M. M. Heslewood, and C. J. Quinn. 2005. Relationships within Myrtaceae sensu lato based on a *mat*K phylogeny. Plant Syst. Evol. 251:3–19.

Xiang, Q. Y., and D. E. Soltis. 2001. Dispersal-vicariance analysis of intercontinental disjuncts: historical biogeographical implications for angiosperms in the Northern Hemisphere. Int. J. Plant Sci. 162:S29–S39.

Yuan, Y. M., S. Wohlhauser, M. Möller, J. Klackenberg, M. Callmander, and P. Küpfer. 2005. Phylogeny and biogeography of *Exacum* (Gentianaceae): a disjunctive distribution in the Indian Ocean basin resulted from long-distance dispersal and extensive radiation. Syst. Biol. 54:21–34.

Zavada, M. S., and D. L. Dilcher. 1986. Comparative pollen morphology and its relationship to phylogeny of pollen in the Hamamelidae. Ann. Mo. Bot. Gard. 73:348–381.

Zerega, N. J. C., W. L. Clement, S. L. Datwyler, and G. D. Weiblen. 2005. Biogeography and divergence times in the mulberry family (Moraceae). Mol. Phylogenet. Evol. 37:402–416.

Zhang, L. -B. and S. S. Renner. 2003. The deepest splits in Chloranthaceae as resolved by chloroplast sequences. Int. J. Plant Sci. 164:S383–S392.

Zurawski G., M. T. Clegg, and A. H. D. Brown. 1984. The nature of nucleotide sequence divergence between barley and maize chloroplast DNA. Genetics 106:735–749.

BIOGRAPHICAL SKETCH

Joshua W. Clayton received his Bachelor of Science degree in plant science from the University of Edinburgh, in June, 2001. During the following year he worked as a field botanist in the upland forests of Mauritius. He began his graduate career in the fall of 2002 at the University of Edinburgh and Royal Botanic Gardens, Edinburgh, receiving his Master of Science in the Biodiversity and Taxonomy of Plants in August, 2003. His Masters thesis focused on the historical biogeography of *Manilkara*, a genus of tropical trees. Josh started his doctoral work with Doug and Pam Soltis at the University of Florida in Gainesville, where he studied biogeographical patterns and evolutionary history of Simaroubaceae. Upon completion of his dissertation, Josh plans to continue research through a postdoctoral position in the UK.