

# Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification

R. Toby Pennington<sup>1\*</sup>, Matt Lavin<sup>2</sup>, Darién E. Prado<sup>3</sup>, Colin A. Pendry<sup>1</sup>, Susan K. Pell<sup>4</sup> and Charles A. Butterworth<sup>5</sup>

<sup>1</sup>Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh EH3 5LR, UK

<sup>2</sup>Department of Plant Sciences, Montana State University, Bozeman, MT 59717, USA

<sup>3</sup>Cátedra de Botánica, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, S2125ZAA, Zavalla, Argentina

<sup>4</sup>Lewis B. and Dorothy Cullman Program for Molecular Systematics Studies, New York Botanical Garden, Bronx, NY 10458, USA

<sup>5</sup>Department of Botany, Iowa State University, Ames, IA 50011, USA

Historical climate changes have had a major effect on the distribution and evolution of plant species in the neotropics. What is more controversial is whether relatively recent Pleistocene climatic changes have driven speciation, or whether neotropical species diversity is more ancient. This question is addressed using evolutionary rate analysis of sequence data of nuclear ribosomal internal transcribed spacers in diverse taxa occupying neotropical seasonally dry forests, including *Ruprechtia* (Polygonaceae), robinoid legumes (Fabaceae), *Chaetocalyx* and *Nissolia* (Fabaceae), and *Loxopterygium* (Anacardiaceae). Species diversifications in these taxa occurred both during and before the Pleistocene in Central America, but were primarily pre-Pleistocene in South America. This indicates plausibility both for models that predict tropical species diversity to be recent and that invoke a role for Pleistocene climatic change, and those that consider it ancient and implicate geological factors such as the Andean orogeny and the closure of the Panama Isthmus. Cladistic vicariance analysis was attempted to identify common factors underlying evolution in these groups. In spite of the similar Mid-Miocene to Pliocene ages of the study taxa, and their high degree of endemism in the different fragments of South American dry forests, the analysis yielded equivocal, non-robust patterns of area relationships.

**Keywords:** seasonally dry tropical forests; Pleistocene refugia; Amazon; speciation; cladistic biogeography

## 1. INTRODUCTION

The neotropics have an estimated 90 000 plant species, more than any other continental area (Thomas 1999). The fossil record suggests that most of these species must have originated during the past 65 Myr of the Cenozoic era. Although eudicot taxa, which account for 75% of extant angiosperm species, have a fossil record stretching back 125 Myr to the Lower Cretaceous, there is little record of many of the species-rich clades until the Upper Cretaceous. For example, the earliest putative fossils of Leguminosae, the family now dominating many neotropical forest ecosystems, appear in Maastrichtian strata of the Late Cretaceous (Magallón *et al.* 1999). These are, however, pollen fossils assignable at best to the relatively species-poor subfamily Caesalpinioideae. Pollen and macrofossils of the species-rich subfamilies Mimosoideae and Papilionoideae, as well as macrofossils of Caesalpinioideae,

are not well documented until the Eocene, *ca.* 50 Myr ago (Herendeen *et al.* 1992).

The evolutionary basis for the Cenozoic angiosperm diversification that led to the high species numbers in the neotropics and other tropical areas remains uncertain. Two hypotheses have been proposed. The first held sway for much of the twentieth century, and sees the tropics as a 'museum', where relatively stable climates through the Cenozoic resulted in low extinction rates, allowing species to accumulate over time (e.g. Fischer 1960; Stebbins 1974). This hypothesis was later challenged by evidence that tropical climates had not been stable, especially over the past 2 million years of the Pleistocene. This led to hypotheses of more recent speciation, the most popular of which was a 'refuge' model invoking allopatric differentiation in populations of rainforest species that became isolated from one another by vegetation adapted to more xeric conditions during times of cool-dry climate (e.g. Haffer 1969; Prance 1973). The refuge hypothesis was popular in the 1970s and 1980s, but has more recently attracted strong criticism, especially for neotropical biota, from three sources. First, from palaeoecologists who see no evidence from pollen cores for reduced cover of

\* Author for correspondence (t.pennington@rbge.org.uk).

One contribution of 17 to a Theme Issue 'Tropical forests and global atmospheric change'.

rainforest in the Amazon Basin (e.g. Colinvaux *et al.* 2001; Mayle *et al.* 2004). Second, dynamic vegetation model simulations also reject the hypothesis of widespread savannah in Amazonia at the LGM (Cowling *et al.* 2001; Mayle *et al.* 2004). Third, genetic divergence data show speciation in tropical rainforest animals to generally pre-date the Pleistocene (summarized by Moritz *et al.* 2000; Glor *et al.* 2001). These animal genetic data alone are sufficient for some workers to dismiss the possibility of Pleistocene speciation by any model (Colinvaux *et al.* 2001, p. 611). More recent molecular divergence literature, however, shows some evidence for Pleistocene speciation in neotropical rainforest trees (Richardson *et al.* 2001). Complete dismissal of the specific refuge model of recent speciation might be premature because pollen data are consistent with the presence of SDTF in the Amazon basin during the Pleistocene, and this vegetation might have isolated refugial islands of rainforest (Pennington *et al.* 2000; but see Mayle *et al.* 2004).

In the debate of the museum and recent speciation hypotheses, it is important to consider that they are not mutually exclusive. An explanation for high tropical diversity may lie in some combination of ancient and recent speciation, as predicted, for example, by models of random birth and death of species in a biota (e.g. Nee *et al.* 1994). For this issue, understanding whether recent Pleistocene climatic changes did provide a significant evolutionary engine in the neotropics may help us to understand how species might react to future anthropogenic climate changes.

#### (a) *Seasonally dry tropical forests and climatic change*

In this paper, we address the age of species in the neotropics, but by using the novel approach of focusing upon SDTF rather than rainforests. Indirectly, this allows an assessment of how species occupying this vegetation reacted to climatic change in the Pleistocene. SDTF has been relatively neglected, but may offer a better system than rainforests for testing whether Pleistocene speciation occurred. In contrast to wet forests, the disjunct areas of present-day distribution of neotropical SDTF (figure 1) may represent actual present-day refugia for SDTF species. This is because these forests grow in strongly seasonal areas where rainfall is less than 1600 mm yr<sup>-1</sup>, with at least five to six months receiving less than 100 mm (Gentry 1995; Graham & Dilcher 1995). We are in a wet interglacial period, but glacial climates in the neotropics were drier (Clapperton 1993; Van der Hammen & Absy 1994; Hooghiemstra 1997; Mayle *et al.* 2000), suggesting that areas suitable for growth of SDTF are likely to be at their minimum, but may have expanded in arid glacial times. Recent fossil pollen data (Mayle *et al.* 2004) show, however, that in Bolivia, the Chiquitano SDTF did not increase in area during the LGM, but moved northwards. Even if regarding all current areas of SDTF as remnants of a more widespread seasonal woodland does not prove entirely accurate, identifying areas of endemism for SDTF for biogeographic analysis is straightforward because of their geographical separation. By contrast, much debate of Amazonian biogeography has not proceeded beyond postulating the locations of ever-wet refugia when past climates were drier (e.g. Nelson *et al.* 1993).

A full discussion of SDTFs, their location, ecology and floristic composition is given by Pennington *et al.* (2000), and this will only be summarized here. These forests have a smaller stature and lower basal area than tropical rainforests (Murphy & Lugo 1986), and thorny species are often prominent. Net primary productivity is lower than in rainforests because growth only takes place in the wet season. Leaf litter builds up in the dry season because the vegetation is mostly deciduous (figure 2) and sunlight penetrates to the forest floor and reduces decomposition by lowering the relative humidity. Wetter forests are semi-deciduous, and deciduousness tends to increase as rainfall declines. However, in the driest forests there is an increase in evergreen and succulent species (Mooney *et al.* 1995). Our definition of SDTF is distinctly general, and includes formations as diverse as tall forest on moister sites to cactus scrub on the driest. Many different names are used for SDTF over its range (e.g. tropical and subtropical dry forests, caatinga, mesotrophic, mesophilous or mesophytic forest, semideciduous or deciduous forest, bosque caducifolio, bosque espinoso). The Leguminosae and Bignoniaceae dominate the woody flora with the Anacardiaceae, Burseraceae, Myrtaceae, Rubiaceae, Sapindaceae, Euphorbiaceae, Flacourtiaceae and Capparidaceae also more or less strongly represented (Gentry 1995). The Cactaceae are prominent, particularly at the formation's latitudinal extremes, and are an important element in the diversity of these forests (figure 3; Gentry 1995). SDTFs usually have a closed canopy, with a sparse ground flora consisting of rather few grasses, with Bromeliaceae, Compositae, Malvaceae and Marantaceae also represented.

The largest areas of SDTF in South America are found in northeastern Brazil (the 'caatingas', extending south to eastern Minas Gerais), in two areas defined by Prado & Gibbs (1993) as the 'Misiones' and 'Piedmont' nuclei (figure 1) and on the Caribbean coasts of Colombia and Venezuela. Other smaller and more isolated areas of SDTF occur in dry valleys in the Andes in Bolivia, Peru, Ecuador and Colombia, coastal Ecuador and northern Peru, the 'Mato Grosso de Goiás' in central Brazil and scattered throughout the Brazilian cerrado biome on areas of fertile soils (Ratter *et al.* 1978). In Central America, SDTFs are concentrated along the Pacific coast from Guanacaste in northern Costa Rica, to just north of the Tropic of Cancer in the Mexican state of Sonora. Over much of their neotropical range, SDTFs have been virtually entirely destroyed, with the largest intact block remaining in the Chiquitano region of Bolivia (Mayle *et al.* 2004).

Tropical savannahs such as the cerrado of central Brazil (figure 1) and the chaco woodlands of the plains of northern Argentina, western Paraguay and southeastern Bolivia (figure 1) both grow in areas of low precipitation and high seasonality. They are also likely to cover less area currently than in past glacial times. Cerrado and SDTF areas have similar climates, but cerrado is fire-prone and its soils are acid, dystrophic, with low calcium and magnesium availability, and often with high levels of aluminium (Furley & Ratter 1988; Ratter *et al.* 1997). By contrast, SDTF grows on fertile soils with moderate to high pH and nutrient levels, and is less subject to burning, as evidenced by the abundance of cacti, which are generally not adapted to survive fire. Chaco is subjected to low soil moisture and

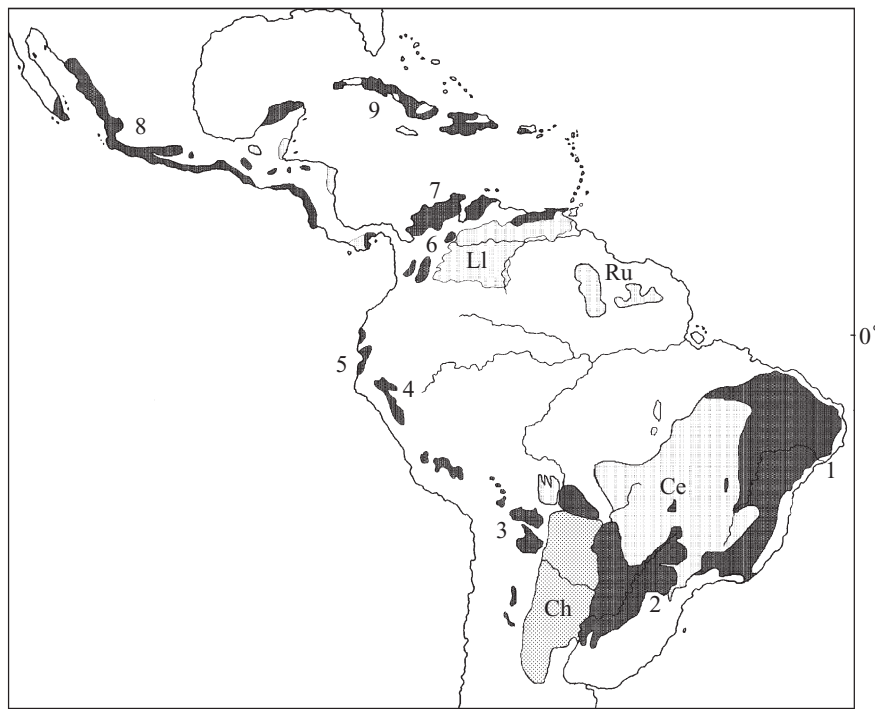


Figure 1. The distribution of seasonally dry vegetation in the neotropics highlighting areas of endemism. Dark grey, SDTF; white, savannah; light grey, Chaco. Seasonally dry forest: 1, Caatingas; 2, Misiones Nucleus; 3, Piedmont Nucleus (including northern Bolivian SDTF); 4, northern Peruvian inter-Andean valleys; 5, Pacific coastal Ecuador and Peru; 6, Colombian and Venezuelan inter-Andean valleys; 7, Caribbean coast of Colombia and Venezuela; 8, Mexico and Central America; 9, Caribbean Islands. Savannahs: Ce, Cerrado; Ll, Llanos; Ru, Rupununi; Ch, Chaco. Modified after Pennington *et al.* (2000).

freezing in the dry season and waterlogging and air temperatures up to 49 °C in the wet season. Because of the regular frosts that it receives, its floristic links are to temperate Monte and Andean Prepuna formations (Cabrera 1976) and it can be considered a subtropical extension of a temperate formation.

The debate of what may have replaced moist forest in the neotropics during glacial periods has been strongly influenced by African studies that showed savannahs increased in extent in dry ice age climates. For many, there has been an implicit acceptance that the dominant vegetation of ice age Amazonia was a form of cerrado. Tropical savannahs such as the cerrado differ from SDTF in a xeromorphic, fire-tolerant grass layer. For some palaeoecologists, the failure to find grass pollen in Amazon lake cores (Colinvaux *et al.* 1996) or the offshore sediments at the mouth of the Amazon River (Haberle 1997; Hoorn 1997) means that refuge theory can be entirely rejected because cerrado must never have been present in the Amazon Basin. This, and the appearance of cool-adapted taxa such as *Alnus*, *Podocarpus* and Ericaceae in pollen cores (Liu & Colinvaux 1985; Bush *et al.* 1990; Colinvaux *et al.* 1996), has led to the idea that climatic cooling was the dominant force in the ice age Amazon, and that aridification of the climate was less significant (Colinvaux *et al.* 1996, 2001). More recent work (Cowling *et al.* 2001; Huang *et al.* 2001; Mayle *et al.* 2004) has also implicated changes in CO<sub>2</sub> concentrations in driving vegetation changes. Lowered CO<sub>2</sub> concentrations coupled with less precipitation in glacial periods are thought to favour grasses such as those that dominate neotropical savannahs, and which photosynthesize using the C<sub>4</sub> pathway.

Much work, however, neglects the possibility that SDTF, a drought-adapted, tree-dominated ecosystem with a more or less continuous canopy and in which grasses are a minor element (Mooney *et al.* 1995), may have spread into Amazonia in times of drier climate. Study of the contemporary distributions of species inhabiting SDTF lends support to the notion that these formations were more widespread in cooler and drier periods of the Pleistocene (Prado & Gibbs 1993; Pennington *et al.* 2000). At least 104 plant species from a wide range of families and often with limited dispersal capabilities are each found in two or more of the isolated areas of SDTF scattered across the neotropics (figures 4 and 5). It is much more parsimonious to assume that these repeated patterns of distribution imply a former, more continuous extent of these forests rather than a series of long-distance dispersal events (Prado & Gibbs 1993; Pennington *et al.* 2000).

In this paper, we focus upon a series of taxonomically unrelated groups (*Ruprechtia* (Polygonaceae), *Coursetia* (Leguminosae–Papilionoideae–Robinieae), *Nissolia* plus *Chaetocalyx* (Leguminosae–Papilionoideae–Aeschynomeneae), *Loxopterygium* (Anacardiaceae) and *Pereskia* (Cactaceae)) whose individual species show a different distribution pattern: high levels of endemism in the separate areas of neotropical SDTF (e.g. figure 6). Pennington *et al.* (2000) suggested that the endemic species in these groups might have been produced by an allopatric process driven by wet interglacial climates fragmenting South American SDTF. It might be expected that any effect of Pleistocene vicariance of these forests on speciation would be most pronounced in South America, where they are currently distributed in disjunct areas, as opposed to



Figure 2. SDTF near Loja, Ecuador, showing almost complete deciduousness in the dry season.



Figure 3. SDTF in Mexico, with abundant cacti.

Central America where they are more continuous. We test this hypothesis of Pleistocene speciation by using molecular biogeographic approaches and thereby test indirectly whether SDTF did expand and contract in response to Quaternary climatic changes. Further, we also use cladistic vicariance analysis to test this hypothesis of speciation caused by climate-driven vicariance of SDTF.

## 2. MATERIAL AND METHODS

### (a) *Evolutionary rate analysis*

Evolutionary rate analysis places a time-dimension on a molecular phylogeny. A date gained from the fossil record or a geological event is placed on a particular node of a phylogenetic tree, and the distribution of DNA base substitutions along the branches of the rest of the tree allows the ages of all other nodes to be inferred from this single point of calibration. New methods of analysis such as those used in this paper can allow for variation in the rate of substitution, and therefore do not rely upon the assumption of a constant 'molecular clock'.

Evolutionary rates were analysed for robinoid legumes, *Chaetocalyx/Nissolia* and *Ruprechtia*. In all cases, phylogenies

were based upon nucleotide sequence data, largely from the nuclear ribosomal ITS and intervening 5.8S sequences. This region is commonly informative for species-level comparisons (see, for example, Delgado-Salinas *et al.* 1999; Lavin *et al.* 2001*b*). Ideally, phylogenetic estimates from the nuclear genome should be verified by independent data from the chloroplast genome. However, readily amplifiable chloroplast loci such as the *trnL-F* or *matK/trnK* regions often do not have sufficient sequence divergence for species-level studies. This proved true for most of the taxa studied here, and in these cases, only phylogenies based upon sequences from the ITS region are presented. An effort was made to detect potentially confounding paralogous or introgressed loci (e.g. Buckler *et al.* 1997; Li *et al.* 2002). The strategy here included direct sequencing of PCR products in both forward and reverse orientations, and using only those sequences that were resolved unambiguously. In addition, sequences from multiple accessions of individual species were obtained when possible. Regardless, there was no precedence from previous published and unpublished studies that the targeted taxa in this study should be suspected of having confounding paralogous or introgressed sequences from the ITS region.

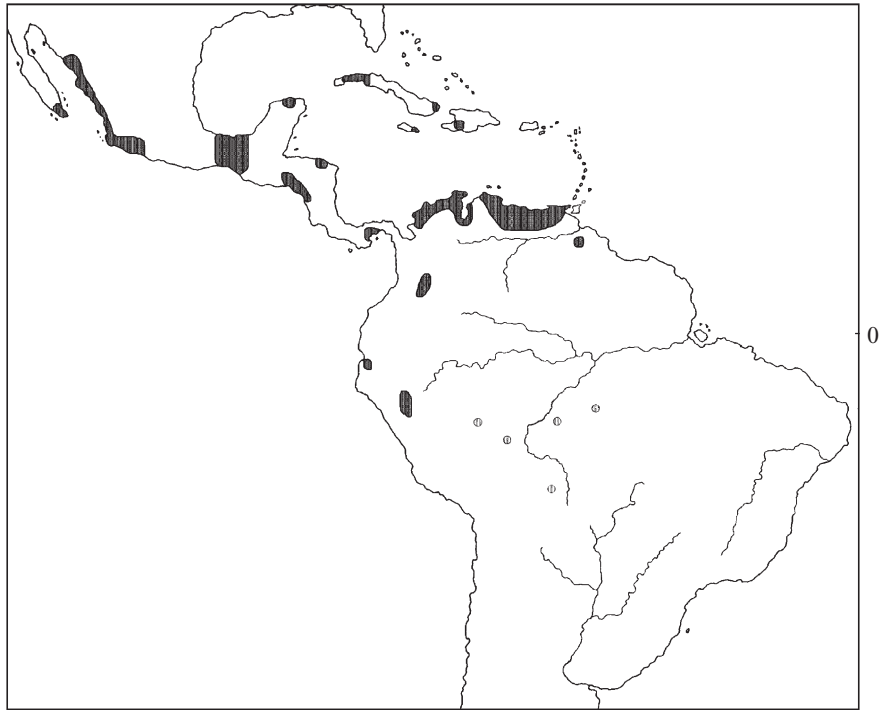


Figure 4. Distribution of *Chloroleucon mangense* (Jacquin) Britton & Rose *sensu lato* (dark-grey patches). Redrawn from Barneby & Grimes (1996) and reproduced with permission from Pennington *et al.* (2000, fig. 2; Blackwell Publishing).

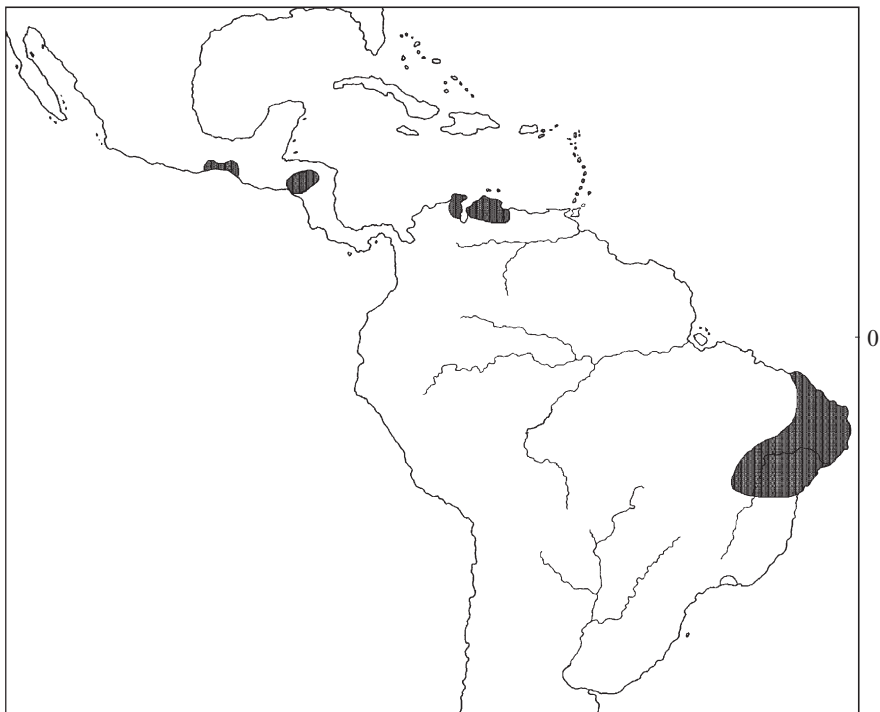


Figure 5. Distribution of *Mimosa tenuiflora* (Willdenow) Poirlet (dark-grey patches). Redrawn from Barneby & Grimes (1996) and reproduced with permission from Pennington *et al.* (2000, fig. 3; Blackwell Publishing).

For the phylogenetic analysis, branch lengths were estimated during a maximum-likelihood analysis of the 5.8S/ITS datasets. This likelihood analysis involved a search of tree parameter space by using a Bayesian approach. This involves a Metropolis-coupled Markov Chain Monte Carlo permutation of tree parameters, an initial random tree, over one million permutations of tree parameters, and four chains (Huelsenbeck & Ronquist 2001; Huelsenbeck *et al.* 2001). The Aikake information

criterion model selection approach in MODELTEST (Posada & Crandall 1998) was used to validate the estimated nucleotide substitution parameters, whereas maximum parsimony implemented in PAUP\* (Swofford 2000) was used to validate branching order. All estimated parameters (i.e. branching order, branch lengths and nucleotide substitution) were then used to generate parametric bootstrap replicates using SEQ-GEN (Rambaut 1996). Each replicate was subjected to likelihood and

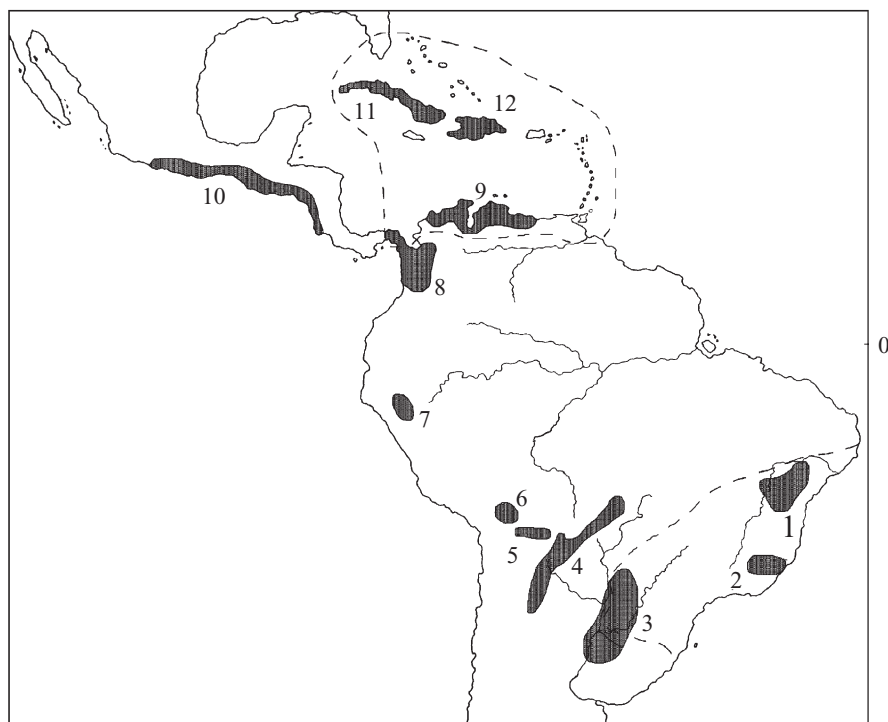


Figure 6. Distribution of *Pereskia*. The dashed line indicates the limits of *P. aculeata* Miller, found in the Caribbean region and eastern South America. The other species are distributed as follows: 1, *P. bahiensis* Gürke, *P. aureifolia* Ritter, *P. stenantha* Ritter; 2, *P. grandifolia* Haworth; 3, *P. nemorosa* Rojas Acosta; 4, *P. sacharosa* Griseb.; 5, *P. diaz-romeroana* Cárdenas; 6, *P. weberiana* K. Schumann; 7, *P. humboldtii* Britton & Rose; 8, *P. bleo* (Kunth) DC.; 9, *P. guamacho* F. A. C. Weber; 10, *P. lychnidiflora* DC.; 11, *P. zinniiiflora* DC.; 12, *P. portulacifolia* (L.) DC., *P. quisqueyana* Liogier. Redrawn from Leuenberger (1986) and reproduced with permission from Pennington *et al.* (2000, fig. 8; Blackwell Publishing).

rate smoothing analyses (see below) to obtain mean and standard deviations of evolutionary substitution rates and ages of clades. Datasets of aligned nucleotide sequences and trees are available at TREEBASE (study accessions S928) and the Web site <http://gemini.oscs.montana.edu/~mlavin/data/dryforest.htm>.

The rate smoothing program r8s (Sanderson 2001) was used to assess variance in evolutionary substitution rates in the datasets involving the ITS region, and incorporate such variance into the estimation of ages of lineages (Sanderson 1997, 1998, 2001, 2002). This program uses a PL approach to identify an optimal rate smoothing parameter that constrains the amount of change from ancestors to descendants. The optimal smoothing parameter is determined by a cross validation approach whereby the rate smoothing parameter chosen best predicts the overall terminal branch lengths in a saturated rate model. The results of PL analysis are compared with those of an autocorrelated rates approach, NPRS (Sanderson 1997), and a rate constant model (Langley & Fitch (1974), referred to as LF). NPRS and LF theoretically define the extremes of the continuum from the saturated to the clock-like rates model. We consider PL to provide the best estimates for ages in the face of variation in substitution rates, so PL results are reported in the text. Values from NPRS and LF are reported in tables 1–3.

Relative substitution rates and ages estimated with r8s were converted to absolute rates and ages by enforcing time constraints at certain nodes, where available. Such time constraints derived from plate tectonic or fossil data are discussed during the enumeration of the taxa analysed in this study.

(i) *Ages of diversifications (crown groups) and ages of lineages (stem groups)*

Hennig (1966) recognized the distinction between the age of origin of a lineage and the age of diversification of the modern

members of a group. The age of the lineage is the date that the modern group split from its sister clade. The age of the modern group is the age of its diversification into modern descendants, or the most recent common ancestor of all of the extant members of the group. The age of the diversification of a modern group may be considerably younger than the age of the lineage. Palaeontologists later termed the diversification of modern groups ‘crown groups’ and the lineages leading to them (comprising extinct taxa) ‘stem groups’ or ‘stem lineages’ (Jefferies 1979; Ax 1987).

When assigning ages to clades by using fossils, one needs to consider the distinction of stem and crown group, for example, in cases where fossils possess apomorphies that diagnose only the crown group. The age of the crown group and the average age of the constituent species within a crown group are most relevant to the question of the age of origin of extant species that is being answered in this paper. The age of the stem group is, however, important when a geological event such as the closure of the Panama Isthmus may have allowed a lineage to enter and radiate in a new area. The age of a geological event such as the rise of an isthmus that enabled dispersal, or the rise of a barrier that impeded dispersal, are assigned to the stem group because such events are likely to have pre-dated the diversification of the crown group.

(b) *Cladistic vicariance methods*

Cladistic vicariance analyses seek common patterns of relationships of areas occupied by unrelated organisms. The discovery of common area patterns implies that diversification in unrelated groups may have been driven by the same underlying process. In this case, we used phylogenies of *Coursetia* and *Poissonia* (robinoid legumes), *Chaetocalyx/Nissolia*, *Ruprechtia*,

Table 1. Age (million years ago) and rate estimates for selected clades of *Coursetia* based upon PL, LF and NPRS analyses. Age constraints for calibration are marked in figure 7. Crown clades are those labelled in figure 7.

crown clade	PL (age and rate)	LF (age and rate)	NPRS (age and rate)
A	38.3 ± 3.8 0.0031 ± 0.0012	38.4 ± 3.8 0.0035 ± 0.0002	37.9 ± 4.1 0.0055 ± 0.0022
B	16.3 ± 4.3 0.0034 ± 0.0004	16.3 ± 4.3 0.0035 ± 0.0002	26.4 ± 3.9 0.0060 ± 0.0011
C	18.1 ± 2.1 0.0034 ± 0.0002	17.7 ± 2.1 0.0035 ± 0.0002	25.8 ± 2.2 0.0036 ± 0.0007
D	19.8 ± 2.5 0.0034 ± 0.0002	19.6 ± 2.4 0.0035 ± 0.0002	28.2 ± 3.0 0.0046 ± 0.0008
D1	15.7 ± 1.8 0.0034 ± 0.0002	15.6 ± 1.8 0.0035 ± 0.0002	23.8 ± 2.9 0.0037 ± 0.0006
D2	4.1 ± 1.0 0.0034 ± 0.0002	4.1 ± 1.0 0.0035 ± 0.0002	15.2 ± 3.8 0.0020 ± 0.0003
D3	5.8 ± 1.1 0.0034 ± 0.0002	5.8 ± 1.1 0.0035 ± 0.0002	11.6 ± 3.0 0.0022 ± 0.0004
D4	7.0 ± 1.2 0.0035 ± 0.0002	6.9 ± 1.1 0.0035 ± 0.0002	10.0 ± 2.0 0.0034 ± 0.0007
D5	7.9 ± 1.2 0.0035 ± 0.0002	7.8 ± 1.2 0.0035 ± 0.0002	11.8 ± 2.1 0.0034 ± 0.0006
D6	4.8 ± 0.9 0.0035 ± 0.0002	4.8 ± 0.9 0.0035 ± 0.0002	8.1 ± 1.9 0.0026 ± 0.0006

*Pereskia* and *Loxopterygium*, all of which have similar distribution patterns of endemic species in the disjunct areas of SDTF. For each group, determining the relationships of component species enables inferences to be made of the historical relationships between the areas of SDTF in which they occur. If the patterns suggested by these different, remotely related groups are similar, the implication is that the same process must have driven speciation in each genus. This common process could be climate-driven vicariance of a wider expanse of SDTF.

#### (i) Areas of endemism

Most areas of endemism for SDTF are easy to define because of their clear geographical separation (figure 1). However, the inter-Andean valleys of Bolivia, Peru, Ecuador, Colombia and Venezuela collectively contain a series of small SDTF areas that are variably spaced and difficult to group objectively because of the lack of floristic data for areas such as northwestern Bolivia and Peru. We consider the inter-Andean SDTF of northwestern Bolivia, for example those in the Yungas and Apolo valley, as an extension of the Piedmont area. This is justified by the study of Kessler & Helme (1999), who showed the flora of the central Tuichi valley in northwestern Bolivia to have affinities with SDTF areas in southeastern Bolivia and other areas to the south rather than with Peruvian forest to the north. The inter-Andean SDTF of northern Peru (principally in the Marañón valley) and southern Ecuador are treated as an area of endemism separate from the SDTF on the Pacific coasts of these countries. This is justified by the strong floristic similarities shown by these inter-Andean areas by Bridgewater *et al.* (2003). The northern inter-Andean valleys of Colombia and Venezuela are treated as a single area, separate from the SDTF of the adjacent Caribbean coast. This is reasonable given the very high degree of phenotypic similarity (and thus sister relationship) of *Coursetia andina* from inter-Andean Venezuela and *C. intermontana* from inter-Andean Colombia (Lavin 1988). These two species are distantly related to *C. ferruginea* from coastal Venezuela, which is a phylogenetically isolated species within the *Coursetia* radiation (Lavin *et al.* 2003).

Species occupying other habitats are also represented in our study groups. These habitats are Monte, Peruvian desert, Mexican desert, restinga (coastal vegetation in Brazil), Brazilian Atlantic coastal rainforest, rainforest of the Amazon and Guianas, and rainforest in Central America. The three rainforest areas were treated separately because they represent distinct areas of endemism for rainforest plants, and cannot be considered a single historical entity (e.g. Mori *et al.* 1981; Wendt 1993).

#### (ii) Construction of area cladograms for individual taxa

Taxon cladograms were generated using maximum-parsimony analyses of ITS (*Coursetia*, *Poissonia*, *Ruprechtia*, *Chaetocalyx/Nissolia*, *Loxopterygium*) and chloroplast sequences (*Pereskia*, *Loxopterygium*), which in all cases except *Chaetocalyx/Nissolia* were already published (see § 2c). Strict consensus cladograms were converted into area cladograms.

The terminal taxa in molecular systematic studies are individual organisms rather than species. This creates difficulties in converting molecular cladograms to area cladograms because we wish to make inferences about the relationships of distribution areas occupied by entire species. Where a species is represented by a single accession, or where multiple accessions from a single species form a monophyletic group, it is straightforward to assume that the accession(s) represent the entire species, and can be treated as a single terminal coded for all areas where the species occurs. Where multiple accessions of a single species do not form a monophyletic group, the following conventions were adopted. If accessions from separate areas of endemism were collectively non-monophyletic, but implying an informative relationship between these areas of endemism, they were treated separately and coded for the area occupied by each accession (e.g. different accessions of *C. hasslerii* in the Misiones and Piedmont areas; figure 7). When multiple accessions of a single species from different areas of endemism were unresolved, nothing could be implied about the relationships of the different areas that they occupy. Thus, such accessions were collapsed to a

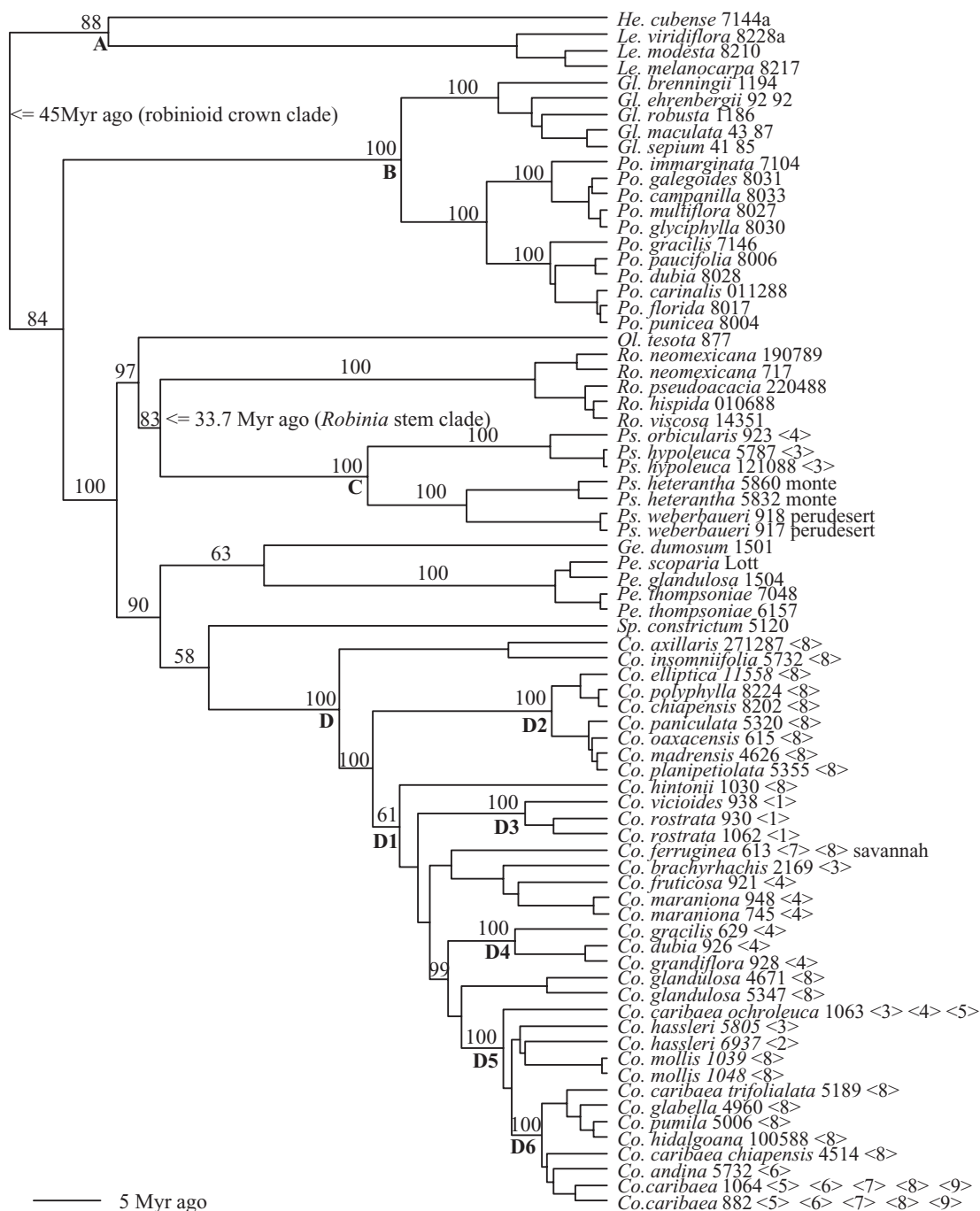


Figure 7. Chronogram for robinoid legumes derived from PL rate smoothing of a Bayesian likelihood consensus tree, which was estimated with sequences from the ITS region. Codes for crown clades (e.g. A, B, C and D) are those used in table 1 and the text. Numerical codes after terminal taxa labels are accession numbers (see Lavin *et al.* 2003). Numerical codes in angled brackets after terminal taxa labels are the SDTF areas of endemism numbered in figure 1. Non-SDTF areas are: savannah, savannah areas in Central America and northern South America; perudesert, Peruvian coastal desert; and monte, arid Andean vegetation. The 45 Myr ago maximum age constraint at the basal node is derived from an analysis of large-scale rates of all legumes (Wojciechowski 2003; M. Lavin, M. F. Wojciechowski, P. Herendeen and M. J. Sanderson, unpublished data). The 33.7 Myr ago minimum age constraint at the second lowest node is derived from the fossil wood record (see § 2). Average nucleotide substitution parameters estimated for 10 000 likelihood trees at stationarity are  $r(\text{GT}) = 1.00$ ,  $r(\text{CT}) = 6.379$ ,  $r(\text{CG}) = 0.980$ ,  $r(\text{AT}) = 1.758$ ,  $r(\text{AG}) = 3.018$ ,  $r(\text{AC}) = 1.155$ ,  $p(\text{A}) = 0.204$ ,  $p(\text{C}) = 0.271$ ,  $p(\text{G}) = 0.293$ ,  $p(\text{T}) = 0.233$ ,  $\alpha = 1.418$ ,  $iP = 0.216$ . See table 1 for the estimated ages and rates of substitution. Numbers above branches are Bayesian posterior probabilities. *He.*, *Hebestigma*; *Le.*, *Lennea*; *Gl.*, *Gliricidia*; *Po.*, *Poitea*; *Ol.*, *Olneya*; *Ro.*, *Robinia*; *Ps.*, *Poissonia*; *Ge.*, *Genistidium*; *Pe.*, *Peteria*; *Sp.*, *Sphinctospermum*; *Co.*, *Coursetia*.

single terminal and coded for all areas occupied by the species (e.g. the accessions of *C. caribaea* var. *caribaea*; figure 7).

Two of the clades used in the vicariance analysis, *Coursetia* and *Poissonia*, come from single dataset of aligned sequences

encompassing all robinoid legumes (figure 7; Lavin *et al.* 2003). These two clades were converted into separate area cladograms, rather than constructing a single larger area cladogram from the entire robinoid phylogeny. This avoids coding taxa that occur



in areas not otherwise represented in the analysis, such as the temperate USA occupied by *Robinia* species. Addition of areas that are not represented in all datasets reduces resolution in the vicariance analyses.

### (iii) Vicariance analyses

Component analysis (Page 1993) and three-area statement analysis (Nelson & Ladiges 1994), both considered standard techniques (Morrone & Carpenter 1995), were employed. All three assumptions for treating widespread taxa (0, 1, 2; Nelson & Ladiges 1991) were investigated.

COMPONENT 2.0 (Page 1990, 1993) performs branch-swapping on the taxon area cladogram to produce a fundamental area cladogram that minimizes hypotheses of dispersal, sympatric speciation and extinction. The dataset analysed using COMPONENT 2.0 is in Appendix A. Search options used subtree-pruning-regrafting, treating missing areas as uninformative and mapping (assumption 0) or not mapping (assumption 1) widespread associates. The criteria minimized were leaves added and losses.

The program TASS (three-area-statements analysis (Nelson & Ladiges 1994)) automatically implements assumption 2. This is generally regarded as more reasonable than assumptions 0 and 1 as a general empirical procedure that does not assume dispersal, vicariance or extinction events, but, at the same time, never denies that they occur (Humphries 1992; Humphries & Parenti 1999). Importantly, TASS also aims to minimize the effect of paralogous area relationships in biogeographic analysis. It achieves this by reducing a complex area cladogram to one or more subtrees that are paralogy free (i.e. in which each area occurs only once). This is analogous to identifying paralogous clades in molecular systematics, and using the evidence of each as a separate, independent estimate of phylogeny. The output file from TASS was analysed using HENNIG86 (Farris 1988). Because the representation of geographical data associated with nodes of subtrees entails numerous missing data entries, conventional parsimony programs such as HENNIG86 produce over-resolved trees (i.e. with resolution not supported by data (Nelson & Ladiges 1996)). The least resolved minimal trees were identified with TAXUTIL (Nelson & Ladiges 1995), where the trees with the minimum number of independent three item statements were selected.

## (c) Taxa

### (i) *Coursetia* and *Poissonia*

*Coursetia* and *Poissonia* (Leguminosae, Papilionoideae, tribe Robinieae) contain 35 and four species, respectively, of mostly shrubs that primarily inhabit SDTFs, and other dry or desert environments (Lavin 1988). The two genera were considered collectively monophyletic (e.g. Lavin 1998). However, a recent multiple dataset phylogeny of all robinoid genera (Lavin *et al.* 2003) sampled 33 species from these two genera and demonstrated they do not form a monophyletic clade. Four *Coursetia* species from Peru, Argentina and Bolivia are placed in a separate monophyletic group, *Poissonia*, which is sister to *Robinia* rather than most closely related to the remaining species of *Coursetia*. We thus have to consider the entire phylogeny of the robinoid legumes for molecular biogeographic analysis, and the two monophyletic groups, *Coursetia* and *Poissonia*, in the cladistic vicariance analyses.

The likelihood and rate smoothing analysis used to estimate ages for the robinoid phylogeny is already reported by Lavin *et al.* (2003). The 45 Myr ago maximum age constraint at the basal

node is derived from an analysis of large-scale rates of all legumes (Wojciechowski 2003; M. Lavin, M. F. Wojciechowski, P. Herendeen and M. J. Sanderson, unpublished data). The 33.7 Myr ago minimum age constraint for the stem lineage of *Robinia* is derived from fossils of *Robinia* wood from the Late Eocene (Wheeler & Landon 1992; Wheeler 2001).

### (ii) *Ruprechtia*

*Ruprechtia* (Polygonaceae, tribe Triplariaceae) comprises 38 species of trees, shrubs and lianas (Pendry 2004). It is primarily a genus of SDTF, but nine species are confined to more moist conditions in rainforests or gallery forests in seasonally dry areas. The SDTF species show high levels of endemism in single, disjunct SDTF areas, whereas the moist forest species are more widespread. An ITS phylogeny for 18 species of *Ruprechtia* was presented by Pennington *et al.* (2004). The species sampled included all of the most common SDTF species, plus five moist forest species, which collectively cover the full geographical range of the genus.

Ages were estimated by using a fossil *Ruprechtia* fruit from Miocene sediments of the Cuenca Basin of southern Ecuador. These sediments are inexactly dated at 8–12 Myr ago (Burnham & Graham 1999), but are most likely at the younger end of this range (R. J. Burnham, personal communication). The fossil is clearly from a *Ruprechtia* species because it possesses a calyx tube that is shorter than the achene, an apomorphy defining the genus. It appears to be most similar to *R. aperta*, but does not correspond exactly with the fruits of any existing species (Pendry 2004). We assume that this fossil provides a minimum age of 8 Myr ago for the entire *Ruprechtia* crown group because: (i) *R. aperta* belongs to a clade that is relatively basal in *Ruprechtia*; and (ii) the fossil may be the sister group of the clade containing *R. aperta* and *R. albida* because it does not possess unequivocal synapomorphies that place it within a subclade of this *R. aperta/R. albida* clade (see Sanderson 1998).

### (iii) *Chaetocalyx* and *Nissolia*

*Chaetocalyx* and *Nissolia* (Leguminosae, Papilionoideae, tribe Aeschynomeneae) are twining herbaceous to woody vines characteristic of neotropical SDTF. *Chaetocalyx* includes ca. 13 species centred in South America (Rudd 1958, 1972, 1996), and *Nissolia* ca. 13 species centred in Mesoamerica (Rudd 1956, 1970, 1975). Chloroplast *matK/trnK* and nuclear ribosomal ITS/5.8S sequences both indicated *Nissolia* to be derived from within *Chaetocalyx* (Lavin *et al.* 2001a). However, *Nissolia* plus *Chaetocalyx* were shown to be monophyletic, and are therefore suitable as the ingroup (Nixon & Carpenter 1993) for the cladistic analysis presented here. *Chaetocalyx* and *Nissolia* are placed within the *Adesmia* subclade of the dalbergioid clade of papilionoid legumes (Lavin *et al.* 2001a). We used other members of this clade (species of *Adesmia*, *Poiretia*, *Amicia* and *Zornia*) as multiple outgroups as an additional test of the monophyly of the ingroup (Nixon & Carpenter 1993).

Nuclear ribosomal ITS/5.8S sequences for 21 accessions representing eight species of *Chaetocalyx* and 10 accessions representing seven species of *Nissolia* (Appendix B) are reported for the first time in this study. Sequences were generated by using protocols described by Beyra-M. & Lavin (1999), Delgado-Salinas *et al.* (1999) and Lavin *et al.* (2001a). The ITS1 region proved difficult to sequence for *Nissolia* species, so the analyses here involve 5.8S/ITS2 sequences only. The age of the *Chaetocalyx/Nissolia* stem lineage is set at 10 Myr ago, an age estimated from a global rate smoothing analysis of the entire

Table 2. Age (million years ago) and rate estimates for selected clades of *Ruprechtia* based upon PL, LF and NPRS analyses. The age constraint for calibration is marked in figure 8. Crown clades are those labelled in figure 8.

crown clade	PL (age and rate)	LF (age and rate)	NPRS (age and rate)
A	6.6 ± 0.8 0.00127 ± 0.00059	5.1 ± 1.4 0.00231 ± 0.00028	7.4 ± 0.3 0.00265 ± 0.00121
B	4.1 ± 0.5 0.00142 ± 0.00016	2.8 ± 0.5 0.00231 ± 0.00028	6.0 ± 0.3 0.00291 ± 0.00021
C	3.7 ± 0.5 0.00135 ± 0.00016	2.5 ± 0.5 0.00231 ± 0.00028	5.6 ± 0.2 0.00176 ± 0.00015
D	1.2 ± 0.2 0.00145 ± 0.00018	0.8 ± 0.1 0.00231 ± 0.00028	3.1 ± 0.3 0.00212 ± 0.00014

Table 3. Age (million years ago) and rate estimates for selected clades of *Chaetocalyx* and *Nissolia* based upon PL, LF and NPRS analyses (based upon 10 Myr ago root). Crown clades are those labelled in figure 9.

crown clade	PL (age and rate)	LF (age and rate)	NPRS (age and rate)
A	7.9 ± 1.3 0.00479 ± 0.00271	7.9 ± 1.3 0.00606 ± 0.00054	8.7 ± 0.9 0.00736 ± 0.00422
B	5.9 ± 1.7 0.00547 ± 0.00205	5.9 ± 1.7 0.00606 ± 0.00054	7.0 ± 1.4 0.00785 ± 0.00311
C	1.6 ± 0.4 0.00605 ± 0.00053	1.6 ± 0.4 0.00606 ± 0.00054	5.0 ± 0.8 0.00525 ± 0.00130

Fabaceae using cpDNA *matK* sequences (M. Lavin, M. F. Wojciechowski, P. Herendeen and M. J. Sanderson, unpublished data).

#### (iv) *Pereskia*

*Pereskia* (Cactaceae) comprises 16 species of trees, herbs and scramblers that are unusual in Cactaceae for their regular, slightly succulent leaves (Leuenberger 1986). *Pereskia* species are found in the SDTF of the Caribbean, Central America and South America. A phylogeny generated using chloroplast DNA restriction site mutations (Wallace 1995), sampling all species of *Pereskia*, showed it to be paraphyletic because *Leptocereus* (11 species, Caribbean) is nested within it. This result is corroborated by chloroplast *rpl16* intron sequences (C. A. Butterworth, unpublished data), which also show *Calymmanthium* (one species, northern Peru) to be sister to *Leptocereus*. It is this clade of *Pereskia*, *Calymmanthium* and *Leptocereus* that is considered in the cladistic vicariance analyses presented here. The *rpl16* phylogeny better resolves relationships between subclades of *Pereskia* species than the restriction site data, but shows exactly the same relationships among *Pereskia* species within subclades.

*Pereskia* has no fossil record, and the Panama Isthmus closure was not appropriate for a calibration because there is no clear imprint of this event present in the phylogeny. Only *P. lychnidiflora* is endemic to Central America, and its relationships with South American taxa are not well resolved.

#### (v) *Loxopterygium*

The most recent revision of *Loxopterygium* (Anacardiaceae) recognized five species (Barkley 1962). However, it now comprises only three species because *L. gutierrezii* has been transferred to *Cyrtocarpa* (Mitchell & Daly 1991), and *L. gardneri* to *Apterokarpos* (Rizzini 1975). *Loxopterygium huasango* grows in the SDTF of northern Peru and Ecuador, and *L. grisebachii* in

the Piedmont nucleus of SDTF in northern Argentina and southern Bolivia. *Loxopterygium sagotii* is found in the moist forests of the Guianas and Venezuelan Guayana. A phylogeny of genera of Anacardiaceae based upon chloroplast *trnL-F* intron-spacer sequences, which sampled all three species of *Loxopterygium*, shows it to be monophyletic with strong bootstrap support, with a monophyletic group of *Apterokarpos* (one species, Brazilian caatingas) and *Cardenasiodendron* (one species, Piedmont nucleus), as sister group (S. K. Pell, unpublished data. Accession information is given in Appendix C). A phylogeny based upon chloroplast *rps16* intron sequences (Mosse 2000) did not sample *Apterokarpos* and *Cardenasiodendron*, but revealed the same relationships among the three *Loxopterygium* species, with *L. huasango* and *L. grisebachii* a monophyletic group, and *L. sagotii* as their sister group.

Because of low levels of sequence divergence in chloroplast loci, for evolutionary rates analysis, nuclear ribosomal ITS/5.8S sequences were generated for eight accessions representing all three *Loxopterygium* species, *Cardenasiodendron*, plus three other Anacardiaceae genera (Appendix B). DNA extractions and PCR purifications were done by using Qiagen Kits (DNeasy Plant Mini Kit and QIAquick PCR Purification Kit; Qiagen Inc., 28159 Avenue Standord, Valencia, CA 91355, USA) and PCR amplification was accomplished with the *TaKaRa Ex Taq* Hot Start Version kit (Takara Shuzo Co., Biomedical Group, Seta 3-4-1, Otsu, Shiga, 520-2193, Japan). Cycle sequencing of PCR products was performed on an ABI 377 automated sequencer in the Molecular Systematics Laboratory at the New York Botanical Garden. Ages were estimated using a fossil *Loxopterygium* fruit from the same deposits of the Cuenca Basin as the *Ruprechtia* fossil (Burnham & Graham 1999). Photographs of the fossil fruit were compared with illustrations and descriptions of extant species in Barkley's (1962) account. The fossil fruit matches those of *L. huasango* and *L. grisebachii*, which are similar

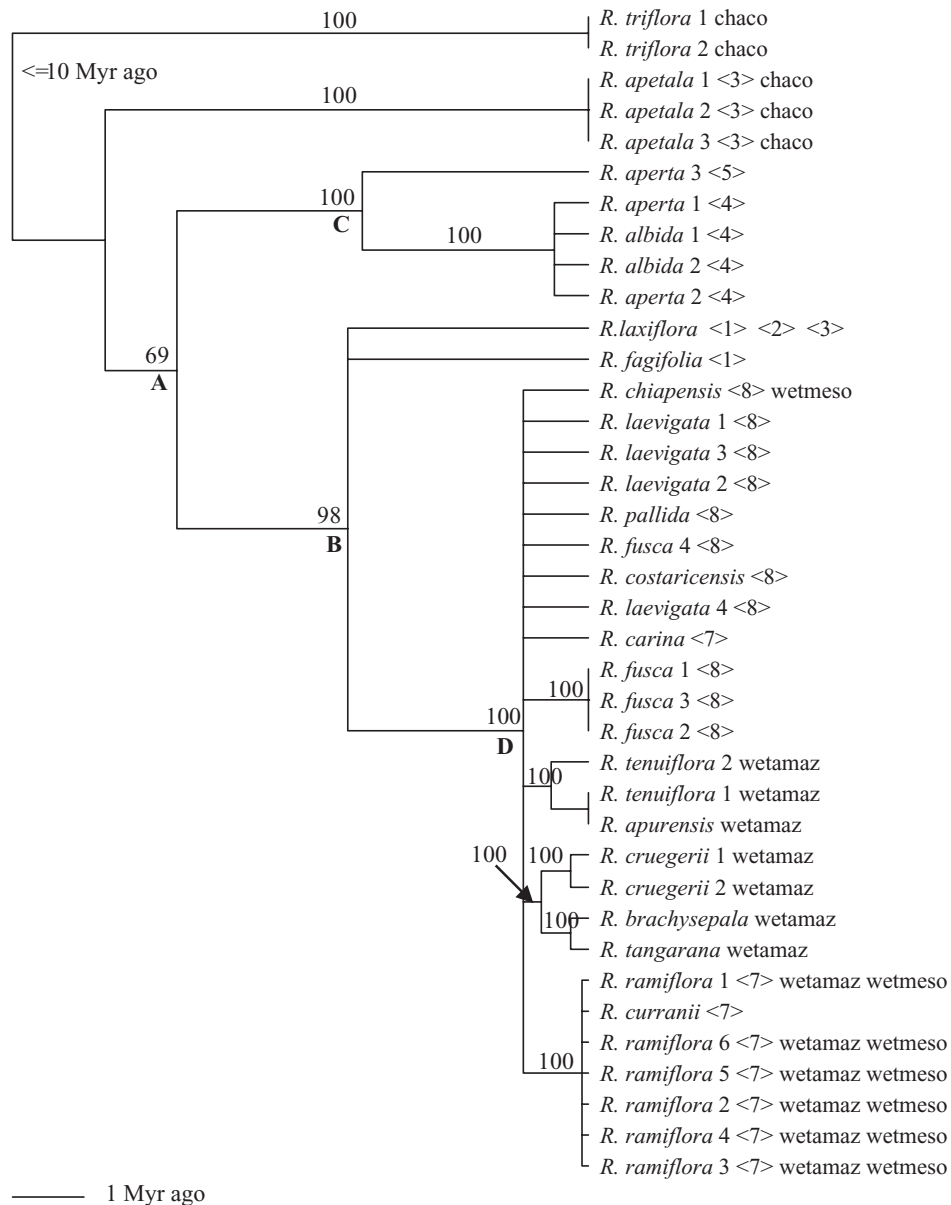


Figure 8. Chronogram for *Ruprechtia* derived from PL rate smoothing of a Bayesian likelihood tree, which was estimated with sequences from the ITS region. Codes for crown clades (e.g. A, B, C and D) are those used in table 2 and the text. Numerical codes after terminal taxa labels are accession numbers (see Pennington *et al.* 2004). Numerical codes in brackets after terminal taxa labels are the SDTF areas of endemism numbered in figure 1. Non-SDTF areas are: *wetmeso*, Central American rainforest; *wetamaz*, Amazonian rainforest; and *chaco*, chaco woodlands of Argentina, Paraguay and Bolivia. The 8 Myr ago maximum age constraint at the basal node is derived from fossil data (see § 2). Average nucleotide substitution parameters estimated for 10 000 likelihood trees that were sampled from over one million at stationarity are  $r(\text{GT}) = 1.000$ ,  $r(\text{CT}) = 6.199$ ,  $r(\text{CG}) = 0.892$ ,  $r(\text{AT}) = 1.812$ ,  $r(\text{AG}) = 5.440$ ,  $r(\text{AC}) = 1.547$ ,  $p(\text{A}) = 0.193$ ,  $p(\text{C}) = 0.302$ ,  $p(\text{G}) = 0.299$ ,  $p(\text{T}) = 0.205$ ,  $\alpha = 0.132$ ,  $iP = 0.708$ . See table 2 for the estimated ages and rates of substitution. Numbers above branches are Bayesian posterior probabilities. *R.*, *Ruprechtia*.

morphologically, but is much smaller than the fruit of *L. sagotii*. It is, therefore, used to assign a date of 8 Myr ago to the stem group of *L. huasango* and *L. grisebachii*.

### 3. RESULTS

#### (a) Evolutionary rate analysis

##### (i) Coursetia and Poissonia

The branching order produced during the Bayesian analysis is essentially the same as that produced by maximum parsimony (Lavin *et al.* 2003). By setting the root of the robinoid tree at 45 Myr ago and the *Robinia*

stem lineage at 33.7 Myr ago, rate smoothing through PL and imposing a globally constant rate (LF) yielded similar results, generally much younger than ages estimated by NPRS (table 1). Two clades confined primarily to Caribbean and Mesoamerican dry forests (clades A and B figure 7; table 1) involve stem and crown clades that are all Tertiary in age. More relevant to this study is the estimated age of the entirely South American *Poissonia* crown clade (C) as at least 18 Myr ago, and the *Coursetia* crown clade (D) as at least 19 Myr ago (figure 7; table 1). The estimated age of the diversification of lineages in the crown clade *Poissonia*, starting at 18 Myr ago and continuing until

Table 4. Age (million years ago) and rate estimates for selected clades of *Loxopterygium* and outgroups based upon PL, LF and NPRS analyses. The age constraint for calibration is marked in figure 10. Crown clades are those labelled in figure 10.

crown clade	PL (age and rate)	LF (age and rate)	NPRS (age and rate)
A	11.4 ± 1.6 0.0085 ± 0.0015	10.0 ± 0.9 0.0092 ± 0.0014	11.2 ± 1.5 0.0091 ± 0.0016
B	2.6 ± 0.5 0.0122 ± 0.0019	2.8 ± 0.5 0.0092 ± 0.0014	2.8 ± 0.5 0.0152 ± 0.0031

5 Myr ago, coincides approximately with Alpers & Brimhall's (1988) estimate of 9–15 Myr ago for the onset of aridification of the Atacama Desert and adjacent Andes where the species of this clade occur. Within the *Coursetia* clade, all the subclades of South American species occur within the crown group labelled D1 (figure 7; table 1), which started to diversify by at least 15 Myr ago. The crown groups D3 and D4 contain only South American species, whereas the basal lineages of the crown clade D5 include South American taxa (e.g. *C. caribaea* var. *ochroleuca* and *C. hassleri*). Collectively these three clades began diversifying ca. 6 Myr ago, although a few constituent species are Pleistocene in age (e.g. *C. dubia*, *C. grandiflora*; figure 7). Notably, many species in the two primarily Central American diversifications, the crown clades D2 and D6, originated during the Pleistocene (table 1; figure 7). These age estimates are probably accurate because the estimated substitution rates for the ITS region of ca.  $3 \times 10^{-9}$  to  $4 \times 10^{-9}$  substitutions site<sup>-1</sup> yr<sup>-1</sup> (see table 1, especially the PL and LF estimates) for the robinoid legumes are close the range for other legume groups (Richardson *et al.* 2001).

#### (ii) Ruprechtia

The topology produced by the Bayesian analysis (figure 8) is identical to that produced by the maximum-parsimony analysis of Pennington *et al.* (2004). All South American SDTF lineages are basal and paraphyletic with the exception of species from the Caribbean coast of Colombia and Venezuela. These are placed in a more apical monophyletic group with all Central American species and two subclades of South American rainforest species.

The accuracy of the fossil calibration can be assessed by the estimated age of the stem lineage subtending the monophyletic group that contains all Central American species (clade D in figure 8; table 2). The diversification of this clade could reflect a scenario of radiation in Central America, which *Ruprechtia* reached after the closure of the Panama Isthmus, because it contains all the Central American species sampled. If this is correct, it predicts an age of ca. 3 Myr ago for the stem lineage of the monophyletic group containing all Central American species, because this is when the Isthmus of Panama closed (Coates & Obando 1996). With an 8 Myr ago calibration at the base of the *Ruprechtia* phylogeny, the estimated date of this stem group (labelled B in figure 8) using PL is  $4.1 \pm 0.5$  Myr ago, suggesting that this stem clade is probably slightly older than the age of closure of the Panamanian Isthmus, and perhaps implying dispersal when the over-water gap was narrow. Certainly, the estimated age of the diversification of this crown group (labelled D in

figure 8) is less than 3 Myr ago regardless of the rate smoothing procedure. This gives some confidence in the fossil calibration. What is certainly clear is that the basally divergent South American SDTF lineages of *Ruprechtia* must pre-date the Pleistocene, and the only evidence for Pleistocene speciation is in northern Peru within the *R. aperta/R. albida* clade. It should, however, be noted that several South American dry forest species were not sampled in this analysis, and Pleistocene origin for these cannot be discounted. The diversification of the Central American species of *Ruprechtia* is clearly post-Pleistocene. Intriguingly, the few species of *Ruprechtia* sampled that are characteristic of moist forests are some of the youngest in the phylogeny.

#### (iii) Chaetocalyx and Nissolia

The topology produced by the Bayesian analysis (figure 9) shows *Chaetocalyx* to be paraphyletic with respect to *Nissolia*, confirming the maximum-parsimony result of Lavin *et al.* (2001a) except that here there is a more thorough sampling of species within each genus. Calibrating the root of the *Chaetocalyx*–*Nissolia* crown group as 10 Myr ago gives a date of diversification for the *Nissolia* crown group as ca. 1.6 Myr ago (using either PL or LF; table 3), the *Chaetocalyx* subclade that is its sister as ca. 5.9 Myr ago. The age of the *Nissolia* stem lineage is estimated at ca. 7.9 Myr ago. This suggests that much of the South American *Chaetocalyx* diversification took place well before the Pleistocene, although some individual species (e.g. *C. longiflora*, *C. glaziovii*) are Pleistocene in age. By contrast, *Nissolia* diversified entirely during Pleistocene, even though it became an isolated lineage well before.

The geographical structure of the tree is consistent with the diversification of the *Nissolia* lineage in Central America after the closure of the Panama Isthmus. All *Nissolia* species belong to a well-supported monophyletic group confined to Central America (except *N. fruticosa* with a range that also includes South America), with the sister clade a group of *Chaetocalyx* species, all of which are endemic to South America (except for *C. scandens* that is widespread in South America, the Caribbean and Central America). However, the date of 7.9 Myr ago for the *Nissolia* stem group is well before the closure of the Panama Isthmus ca. 3 Myr ago, implying that some other event caused the divergence.

The 10 Myr ago calibration of the *Chaetocalyx*/*Nissolia* crown clade derived from a global rates analysis of all legumes (M. Lavin, M. F. Wojciechowski, P. Herendeen and M. J. Sanderson, unpublished data) is considered conservative if only because this calibration yields a range of substitution rates at ca.  $5 \times 10^{-9}$  to  $8 \times 10^{-9}$

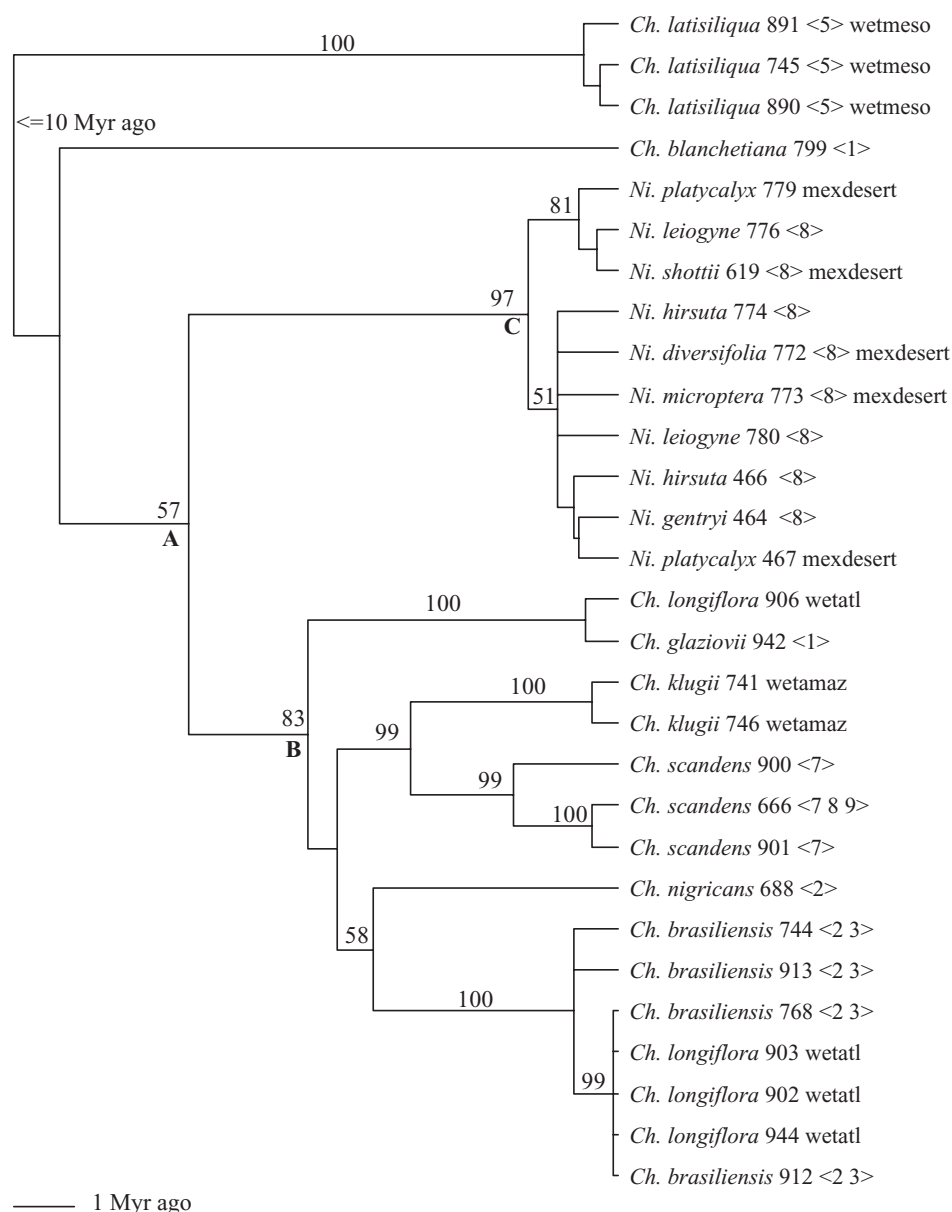


Figure 9. Chronogram for *Chaetocalyx/Nissolia* derived from PL rate smoothing of a Bayesian likelihood tree, which was estimated with sequences from the ITS region. Codes for crown clades (e.g. A, B and C) are those used in table 3 and the text. Numerical codes after terminal taxa labels are accession numbers (see Appendix B). Numerical codes in brackets after terminal taxa labels are the SDTF areas of endemism numbered in figure 1. Non-SDTF areas are: *wetmeso*, Central American rainforest; *mexdesert*, Mexican deserts; *wetamaz*, Amazonian rainforest; and *wetatl*, eastern Brazilian Atlantic coastal rainforest. The 10 Myr ago maximum age constraint at the basal node is derived from an analysis of large-scale rates of all legumes (Wojciechowski 2003; M. Lavin, M. F. Wojciechowski, P. Herendeen and M. J. Sanderson, unpublished data). Average nucleotide substitution parameters estimated for 10 000 likelihood trees that were sampled from over one million at stationarity are  $r(\text{GT}) = 1.000$ ,  $r(\text{CT}) = 15.276$ ,  $r(\text{CG}) = 3.784$ ,  $r(\text{AT}) = 5.522$ ,  $r(\text{AG}) = 8.428$ ,  $r(\text{AC}) = 3.573$ ,  $p(\text{A}) = 0.184$ ,  $p(\text{C}) = 0.309$ ,  $p(\text{G}) = 0.313$ ,  $p(\text{T}) = 0.194$ ,  $\alpha = 1.876$ ,  $iP = 0.236$ . See table 3 for the estimated ages and rates of substitution. Numbers above branches are Bayesian posterior probabilities. *Ch.*, *Chaetocalyx*; *Ni.*, *Nissolia*.

substitutions  $\text{site}^{-1} \text{yr}^{-1}$  (table 3). This is on the high end of rates reported for the ITS region for leguminous shrubs and herbs by Richardson *et al.* (2001). If we are truly estimating too fast a rate, then this can only mean that our estimated ages are too young and the true ages of the clades reported above are even older. If we consider only relative ages because of difficulties in calibration of the *Chaetocalyx/Nissolia* phylogeny, it is entirely clear that the Central American radiation of *Nissolia* is more recent than the divergence of most *Chaetocalyx* lineages.

#### (iv) *Loxopterygium*

The topology produced by the Bayesian analysis of ITS data (figure 10), with *Loxopterygium* resolved as monophyletic, is identical to that produced by parsimony analysis of chloroplast DNA sequences. The divergence of the two SDTF species of *Loxopterygium*, *L. huasango* (Peru/Ecuador) and *L. grisebachii* (Bolivia/Argentina) is dated as 2.6 Myr ago (clade B in figure 10; table 4), just before the Pleistocene. The nucleotide substitution rates of  $ca. 1 \times 10^{-9}$  are somewhat lower than those reported for the

ITS region in the woody, neotropical genus *Inga* (Richardson *et al.* 2001), but *Loxopterygium* species are woody and probably slow growing with a long generation time, which would suggest that substitution rates should be less. It therefore seems likely that these substitution rates are a reasonable estimate.

#### (b) Cladistic vicariance analysis

Fundamental area cladograms produced by COMPONENT 2.0 applying assumption 0 (i.e. mapping widespread associates) and minimizing leaves added and losses are presented in figures 11 and 12. Area cladograms applying assumption 1 and minimizing leaves added and losses are presented in figures 12 and 13. The analyses are clearly sensitive to the criterion minimized: there are no area relationships of SDTF areas that are common to all the area cladograms. Moreover, the analyses were also sensitive to minor re-coding of areas such as the addition or removal of a single area of endemism from a widespread species. The fundamental area cladograms produced when losses are minimized appear more plausible under both assumptions 0 and 1 in that the Caribbean island SDTFs are grouped with geographically adjacent areas (Mesoamerica and the Caribbean coastal area of Colombia and Venezuela) rather than with SDTF areas in Ecuador, Peru, Bolivia and Argentina.

The strict consensus of 20 least resolved area cladograms (CI = 0.63, RI = 0.71) found by TASS and TAXUTIL is presented in figure 14. This shows the northern SDTF areas (Caribbean islands, Mesoamerica and the Caribbean coastal area of Colombia and Venezuela) to group together with all the rainforest areas and the savannah area. The remaining South American SDTF areas are paraphyletic, and the caatingas occupy a basal, isolated position as sister to the restinga area.

## 4. DISCUSSION

### (a) Miocene/Pliocene patterns in South American seasonally dry tropical forest

The distribution patterns of species of *Pereskia*, *Ruprechtia*, *Coursetia*, *Loxopterygium* and *Chaetocalyx/Nissolia* in South American SDTF, where different species are endemic to separate areas, suggested that they might have been produced allopatrically by Pleistocene vicariance of these forests (Pennington *et al.* 2000). This hypothesis is largely rejected by the analysis of evolutionary rates analysis presented here. In no case, irrespective of the reconstruction method (LF, NPRS, PL) or the means of calibration (fossil or geological event), are crown groups of South American species indicated to have started diversifying during the Pleistocene (1.8 Myr ago or less). Furthermore, very few individual species are shown to be Pleistocene in age. This finding is bolstered by estimated substitution rates that approximate those published for the ITS region from other taxa (e.g. Richardson *et al.* 2001). In no case did we estimate substitution rates on the slow end of the reported range. Thus, our findings are unlikely to be biased towards older age estimates.

For the *Coursetia* and *Ruprechtia* phylogenies, we can be reasonably confident of the estimated dates because fossil evidence was verified independently using geological events. These South American diversifications started *ca.*

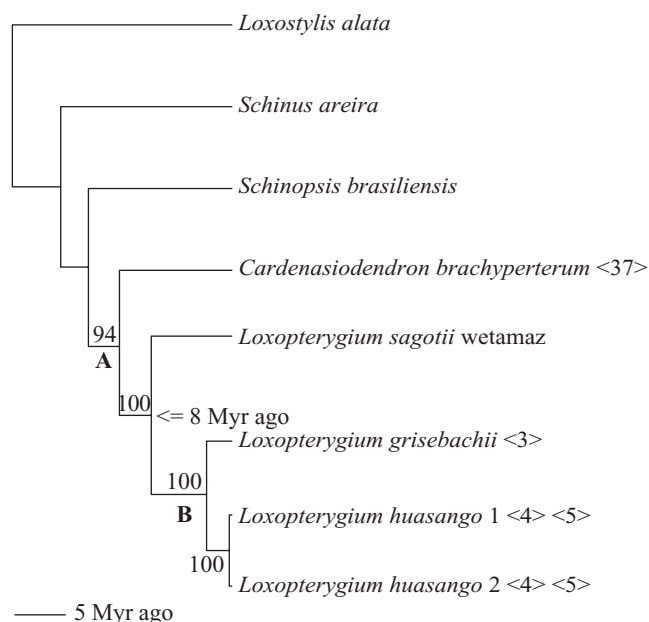


Figure 10. Chronogram for *Loxopterygium* derived from PL rate smoothing of a Bayesian likelihood tree, which was estimated with sequences from the ITS region. Codes for crown clades (e.g. A, B) are those used in table 3 and the text. Numerical codes after terminal taxa labels are accession numbers (see Appendix C). Numerical codes in brackets after the *Loxopterygium* labels are the SDTF areas of endemism numbered in figure 1. The non-SDTF area is wetamaz, Amazonian rainforest. The 8 Myr ago maximum age constraint at the node marking the *Loxopterygium* crown clade is derived from fossil data (see § 2). Average nucleotide substitution parameters estimated for 10 000 likelihood trees that were sampled from over one million at stationarity are  $r(\text{GT}) = 1.000$ ,  $r(\text{CT}) = 6.908$ ,  $r(\text{CG}) = 1.590$ ,  $r(\text{AT}) = 1.973$ ,  $r(\text{AG}) = 2.504$ ,  $r(\text{AC}) = 0.708$ ,  $p(\text{A}) = 0.181$ ,  $p(\text{C}) = 0.335$ ,  $p(\text{G}) = 0.291$ ,  $p(\text{T}) = 0.193$ ,  $\alpha = 0.389$ ,  $iP = 0.000$ . See table 4 for the estimated ages and rates of substitution. Numbers above branches are Bayesian posterior probabilities.

20 Myr ago (*Coursetia*), 18 Myr ago (*Poissonia*), 8–12 Myr ago (*Ruprechtia*) and *ca.* 11 Myr ago (*Loxopterygium*). For *Chaetocalyx/Nissolia*, assigning an age to the base of the clade was more problematic. Using a 10 Myr ago age constraint derived from an analysis of rates of all legumes (M. Lavin, M. F. Wojciechowski, P. Herendeen and M. J. Sanderson, unpublished data) yields a fairly fast estimate of substitution. Regardless, a pre-Pleistocene age of most of the South American *Chaetocalyx* diversification is estimated.

In summary, these genera diversified during roughly the same time-frame in South America, from the Mid-Miocene to the Pliocene. This gives confidence in a critical assumption of the cladistic vicariance analysis for these areas because the branching patterns of the independent phylogenies were caused by events of approximately the same age. This strongly suggests that such taxa should be amenable to cladistic vicariance analysis.

### (b) Recent patterns in Central American seasonally dry tropical forest and South American rainforests

In contrast to the more ancient origins of species in South American SDTF, Central American dry forest

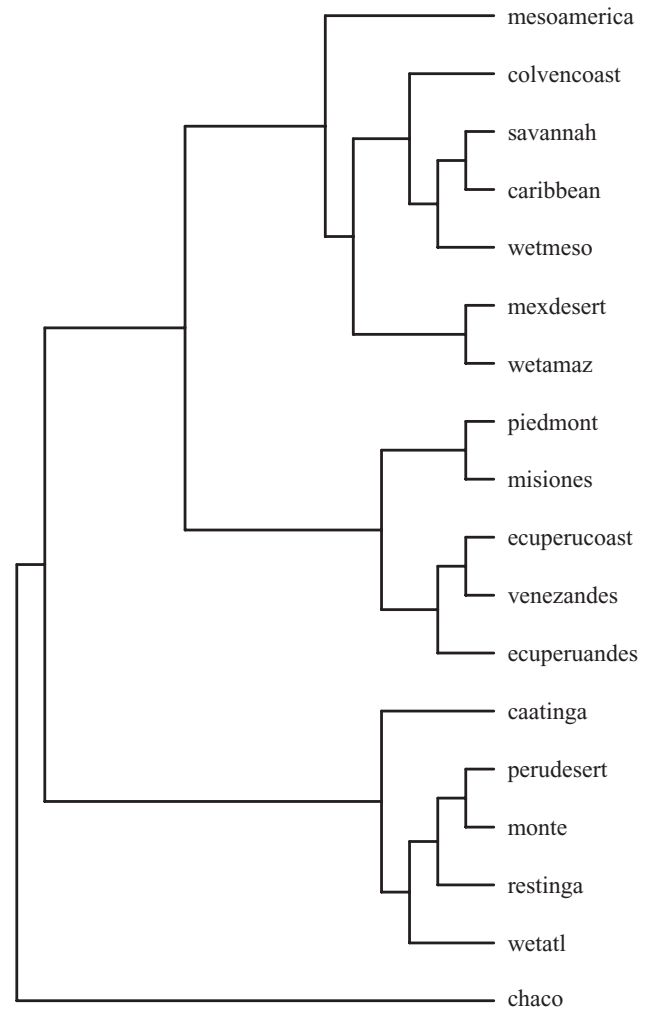
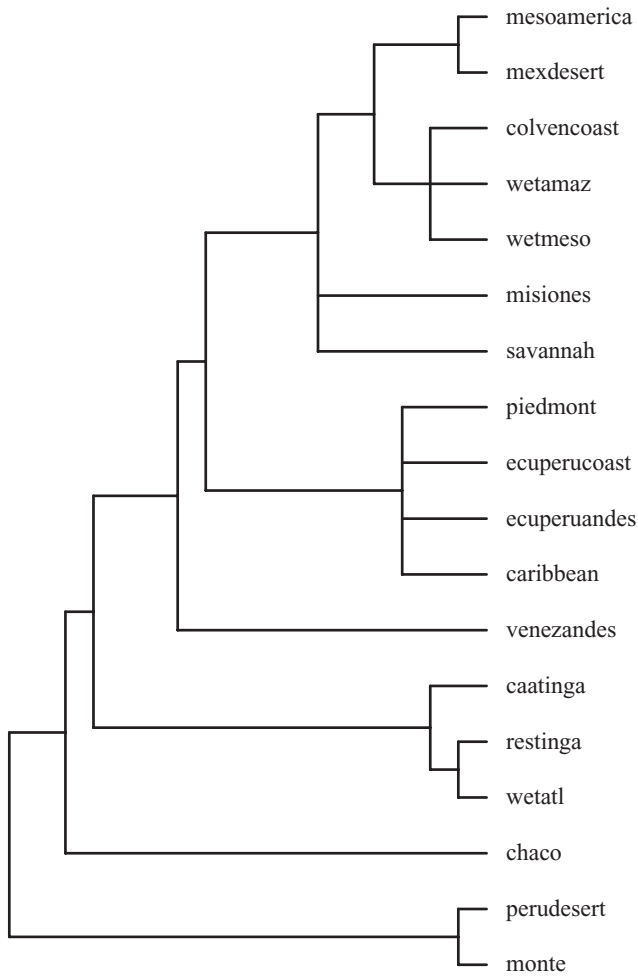


Figure 11. Strict consensus area cladogram derived from 20 trees generated by a component analysis minimizing leaves added (minimal value 497) and implementing assumption 0. For map of areas of endemism see figure 1. SDTF areas: mesoamerica, Central American SDTF; colvencoast, SDTF of Caribbean coast of Colombia and Venezuela; caribbean, SDTF of Caribbean islands; piedmont, SDTF of Piedmont Nucleus; misiones, SDTF of Misiones Nucleus (including northern Bolivian dry forests); ecuperucoast, SDTF of Pacific coastal Ecuador and Peru; venezandes, SDTF of Venezuelan inter-Andean valleys; ecuperuandes, SDTF of northern Peruvian and southern Ecuadorian inter-Andean valleys; caatinga, SDTF of Caatingas. Non-SDTF areas: savannah, savannah areas in Central America and northern South America; wetmeso, Central American rainforest; mexdesert, Mexican deserts; wetamaz, Amazonian rainforest; perudesert, Peruvian coastal desert; monte, arid Andean vegetation; restinga, eastern Brazilian coastal restinga vegetation; wetatl, eastern Brazilian Atlantic coastal rainforest; chaco, chaco woodlands of Argentina, Paraguay and Bolivia.

Figure 12. Area cladogram generated by a component analyses minimizing losses and implementing assumption 0 (minimal value 175) and assumption 1 (minimal value 167). For map of areas of endemism see figure 1. SDTF areas: mesoamerica, Central American SDTF; colvencoast, SDTF of Caribbean coast of Colombia and Venezuela; caribbean, SDTF of Caribbean islands; piedmont, SDTF of Piedmont Nucleus; misiones, SDTF of Misiones Nucleus (including northern Bolivian dry forests); ecuperucoast, SDTF of Pacific coastal Ecuador and Peru; venezandes, SDTF of Venezuelan inter-Andean valleys; ecuperuandes, SDTF of northern Peruvian and southern Ecuadorian inter-Andean valleys; caatinga, SDTF of Caatingas. Non-SDTF areas: savannah, savannah areas in Central America and northern South America; wetmeso, Central American rainforest; mexdesert, Mexican deserts; wetamaz, Amazonian rainforest; perudesert, Peruvian coastal desert; monte, arid Andean vegetation; restinga, eastern Brazilian coastal restinga vegetation; wetatl, eastern Brazilian Atlantic coastal rainforest; chaco, chaco woodlands of Argentina, Paraguay and Bolivia.

species have in many cases originated more recently. In *Ruprechtia*, all Central American species are more recent than 1.2 Myr ago, and most of the 12 Central American taxa in two *Coursetia* crown clades (D2 and D6; figure 7) are Pleistocene in age. Similarly, the primarily Central American *Nissolia* crown clade began to diversify well after *Chaetocalyx* did in South America, and its species probably have a Pleistocene origin.

One intriguing aspect of the results is the recent origin of rainforest species of *Ruprechtia*. The two clades of rainforest

species (figure 8) are dated as *ca.* 1 Myr ago or less. This recent origin is corroborated by species such as *R. tenuiflora* where individual accessions are resolved as paraphyletic with respect to *R. apurensis*, which is consistent with the recent origin of *R. apurensis* from populations of *R. tenuiflora*. Similar recent speciation, evidenced by low sequence divergence for ITS between species, was found in the neotropical rainforest genus *Inga* (Richardson *et al.* 2001).

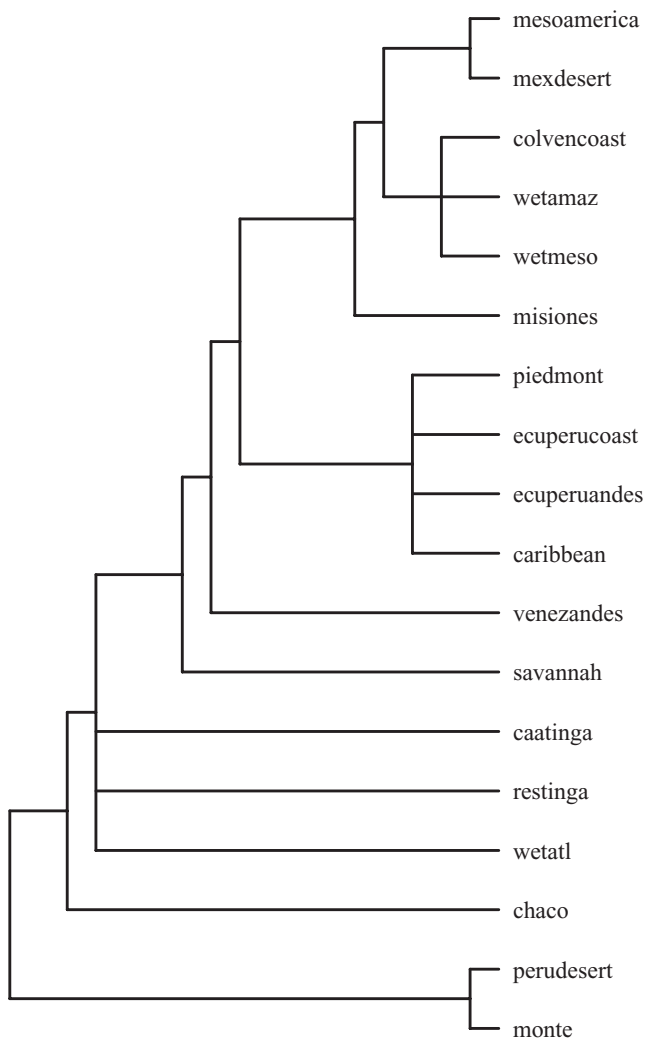


Figure 13. Strict consensus area cladogram derived from 180 trees generated by a component analysis minimizing leaves added (minimal value 461) and implementing assumption 1. For map of areas of endemism see figure 1. SDTF areas: mesoamerica, Central American SDTF; colvencoast, SDTF of Caribbean coast of Colombia and Venezuela; caribbean, SDTF of Caribbean islands; piedmont, SDTF of Piedmont Nucleus; misiones, SDTF of Misiones Nucleus (including northern Bolivian dry forests); ecuperucoast, SDTF of Pacific coastal Ecuador and Peru; venezandes, SDTF of Venezuelan inter-Andean valleys; ecuperuandes, SDTF of northern Peruvian and southern Ecuadorian inter-Andean valleys; caatinga, SDTF of Caatingas. Non-SDTF areas: savannah, savannah areas in Central America and northern South America; wetmeso, Central American rainforest; mexdesert, Mexican deserts; wetamaz, Amazonian rainforest; perudesert, Peruvian coastal desert; monte, arid Andean vegetation; restinga, eastern Brazilian coastal restinga vegetation; wetatl, eastern Brazilian Atlantic coastal rainforest; chaco, chaco woodlands of Argentina, Paraguay and Bolivia.

The clear evidence for Pleistocene speciation in Central American SDTF and rainforests contradicts most studies of genetic divergence in rainforest animals, which demonstrate species to have originated before the Pleistocene in virtually all cases (reviewed by Moritz *et al.* 2000). The conclusion of Colinvaux *et al.* (2001, p. 611), that 'data from molecular genetics are probably sufficient by themselves to discount the [Pleistocene refuge] hypothesis', is clearly overstated.

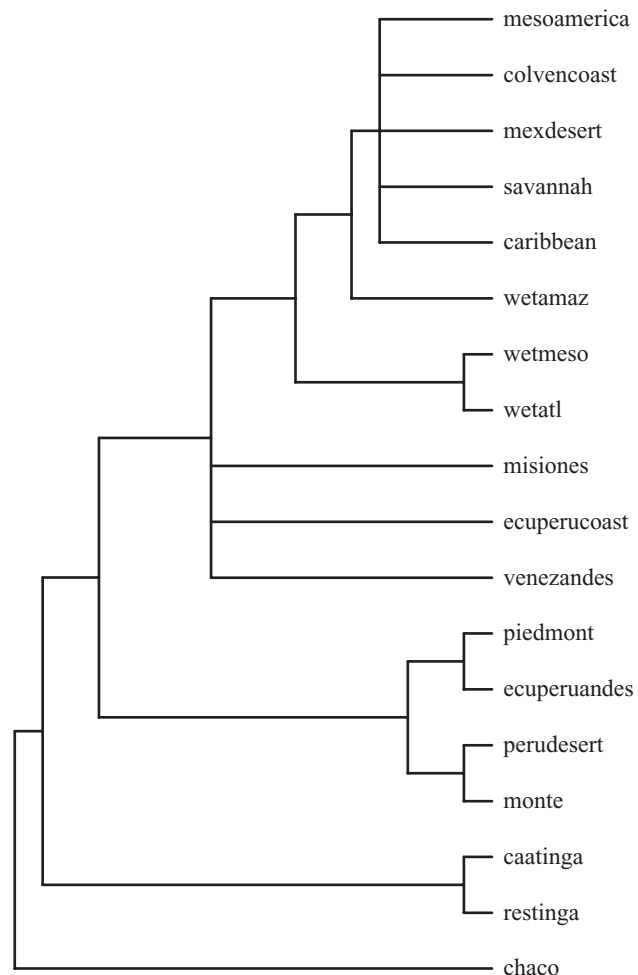


Figure 14. Strict consensus of 20 minimally resolved trees generated by a three-area statements analysis implementing Assumption 2. For map of areas of endemism see figure 1. SDTF areas: mesoamerica, Central American SDTF; colvencoast, SDTF of Caribbean coast of Colombia and Venezuela; caribbean, SDTF of Caribbean islands; piedmont, SDTF of Piedmont Nucleus; misiones, SDTF of Misiones Nucleus (including northern Bolivian dry forests); ecuperucoast, SDTF of Pacific coastal Ecuador and Peru; venezandes, SDTF of Venezuelan inter-Andean valleys; ecuperuandes, SDTF of northern Peruvian and southern Ecuadorian inter-Andean valleys; caatinga, SDTF of Caatingas. Non-SDTF areas: savannah, savannah areas in Central America and northern South America; wetmeso, Central American rainforest; mexdesert, Mexican deserts; wetamaz, Amazonian rainforest; perudesert, Peruvian coastal desert; monte, arid Andean vegetation; restinga, eastern Brazilian coastal restinga vegetation; wetatl, eastern Brazilian Atlantic coastal rainforest; chaco, chaco woodlands of Argentina, Paraguay and Bolivia.

### (c) Vicariance analyses

The vicariance analyses do poorly in uncovering common patterns in area relationships. At least, the results of different patterns analyses implementing assumptions 0, 1 and 2 have little in common. The results of the individual analyses are also weakly supported, as evidenced by the low CI and RI of the assumption 2 analysis using TASS. Perhaps more critically, the individual analyses were highly sensitive to the most minor recoding of areas for a single taxon. Furthermore, the analyses are complicated by the need to include non-dry forest areas.



Proponents of vicariance techniques might argue that these unresolved results reflect the different histories of the different groups analysed. It might also be argued that better vicariance results would be gained by analysing more monophyletic radiations whose species show a greater degree of endemism in South American SDTF. However, the angiosperm groups chosen for analysis here were the result of a wide literature survey (Pennington *et al.* 2000), and we doubt that there are better candidates. Other SDTF species are commonly widespread (Prado & Gibbs 1993; Pennington *et al.* 2000) and come from genera that contain species characteristic of many different vegetation types. Such groups would be more likely to confound rather than clarify the outcomes of vicariance analyses. We chose taxonomic groups that have the best potential of having had similar histories, and the evolutionary rate analysis shows them to have diversified over approximately the same time period, which implies that they should have been influenced by similar environmental variables.

An alternative explanation of the apparent weakness of vicariance approaches is that we have used too few taxon phylogenies, and that results may improve with the addition of more phylogenies, in analogy with the increase in phylogenetic accuracy with increased number of characters in phylogenetic analysis (Hillis 1996). However, there is no precedence in the literature for having faith that addition of taxon cladograms will aid resolution, and cladistic vicariance analysis may well suffer from being an inconsistent method like parsimony in general (e.g. Felsenstein 1982).

In summary, we agree with Lavin *et al.* (2001*b*) that the failure of vicariance methods may lie in the rarity of cases where modern distributions have been minimally influenced by extinction, dispersal and sympatric speciation.

(i) *Seasonally dry tropical forest and Pleistocene climate change*

If all the South American SDTF species in our study group had been shown to have Pleistocene origins, this would have represented substantial evidence that ice age climates had strongly influenced this vegetation. However, the finding that these species are primarily more ancient cannot refute the hypothesis that SDTF was more widespread in glacial times. Perhaps the time of isolation in SDTF areas of endemism (figure 1) has not been long enough to promote allopatric speciation in most cases. Evidence from the distribution of species that are widespread in neotropical SDTFs (Prado & Gibbs 1993; Pennington *et al.* 2000) is still suggestive that these forests were historically more widespread. This is partly supported by the vegetation modelling of Mayle *et al.* (2004), which shows deciduous broad-leaf forests covered the southern half of Amazonia at the LGM. Elsewhere, however, their model simulates northwestern Amazonia to have retained evergreen broadleaved forest, and the SDTF Chiquitano nucleus in Bolivia to have moved north rather than expanded. If this model is correct, then some of the inferences made from plant distribution patterns by Prado & Gibbs (1993) and Pennington *et al.* (2000) may be erroneous. We aim to test whether these widespread distribution patterns are the result of long-distance dispersal or caused by the vicariance of a more widespread SDTF, by examining the population genetic variation within individual widespread species across all SDTF nuclei.

We also encourage future workers to abandon the simple rainforest/savannah dichotomy when considering the effects of past and future climatic change on neotropical vegetation, and to also consider SDTF.

(d) *The future of evolution of seasonally dry tropical forest plants in the face of anthropogenic climate change*

Some authors (e.g. Myers & Knoll 2001; Woodruff 2001) have pointed out that conservation policies fail to consider the 'future of evolution': the long-term evolutionary aspects of biodiversity loss, habitat alteration and climatic change. Although predicting the products of evolution is impossible given that it is a stochastic process, these authors argue that we should attempt to allow the processes of evolution to proceed as they would without human intervention. A difficulty of this approach is the disparity between evolutionary and human time-scales, which is emphasized by the findings of this paper. Only a few South American SDTF species are 'recent', but even this is 1–2 Myr ago, 5000 times greater than a time-scale considered 'long' by humans (200 years). Persuading governments to plan for such 'short' evolutionary time-scales would be doomed to failure. Arguments about extinction, which will certainly be a more rapid process, will perhaps be more persuasive, especially if they are framed in terms of the extinction of biomes as well as of species. SDTFs are already so heavily impacted across the neotropics, and in many cases surrounded by agricultural landscapes, that it is hard to envisage SDTF species expanding their ranges in the face of hotter and drier climates. Either adequate species populations in SDTF nuclei do not exist, or migration would be impossible because of lack of habitat in an anthropogenic landscape. If the case demonstrated for the Bolivian Chiquitano region in the face of drier climates (Mayle *et al.* 2004)—that SDTF species merely shift their ranges rather than expand them—is general, then the future is bleaker still because they will die where they stand. Clearly, conservation planning must allow for migration across ecotones, such as that advocated for the forests of the Andean flanks by Bush (2002). In the case of SDTF, this would require the protection of relatively intact, large SDTF areas, plus other intact biomes surrounding them. This ecotonal approach might indirectly address the future evolutionary concerns of authors such as Myers & Knoll (2001) because speciation associated with ecotones is indicated to be widespread in our study. It is clear that most SDTF species come from genera containing species characteristic of several other vegetation types, which were recognized as non-SDTF areas of endemism in the vicariance analyses. This indicates that ecotonal speciation may be generally important for the neotropical flora. Furthermore the individual phylogenies often indicate multiple, rather than single, diversification events across ecotonal boundaries. For example, in the *Ruprechtia* phylogeny, there are two recent rainforest radiations in this primarily SDTF genus.

## 5. CONCLUSION

The debate about when neotropical species arose and how climatic changes have affected their evolution has focused principally upon rainforests, and especially those

of the Amazon Basin. Our study differs in its focus on SDTF, especially those in South America. Evolutionary rate analysis demonstrates that in monophyletic radiations of species with high levels of endemism in these forests, diversification took place beginning in the Late Miocene and Pliocene. This contradicts our own earlier assertions (Lavin 1998; Pennington *et al.* 2000) that the geographical patterns in these genera were likely to reflect Pleistocene vicariance of a once wider area of SDTF. It indicates that Pleistocene climatic changes were not a major force driving the speciation of woody, South American SDTF plants.

Many readers may interpret the principal finding of this paper—Miocene and Pliocene speciation in South American SDTFs—as confirming the now popular rejection of neotropical Pleistocene speciation, especially that of the specific model of allopatric speciation in rainforest refuges (e.g. Colinvaux *et al.* 2001). We, however, emphasize that our data do show Pleistocene diversification in Central American SDTF in cases where taxa have an ancestral area of South America, and may have reached Central America after the Panama Isthmus closed. Pleistocene climatic change as a factor in neotropical speciation cannot be entirely discounted. Furthermore, there is clear

evidence for Pleistocene speciation in South American rainforest species of *Ruprechtia*, just as there is for the species-rich genus *Inga* (Richardson *et al.* 2001). Our data, perhaps unsurprisingly, show that a mixture of ancient and recent diversification explains the extant diversity in five genera of plants centred in neotropical SDTF ecosystems.

This work was supported largely by the Leverhulme Trust (grant F/771/A to the Royal Botanic Garden Edinburgh (RBGE)) and the US National Science Foundation (DEB-0075202). D.E.P. is a researcher of the Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET); he also received travel grants from the Universidad Nacional de Rosario (Programa Viajes al Exterior) to partly support visits to Montana State University and RBGE. Sequencing of Anacardiaceae was supported by the Lewis B. and Dorothy Cullman Program for Molecular Systematics Studies at the New York Botanical Garden. We thank Carlos Reynel, Aniceto Daza, Monica Moraes, Stephan Beck, Mario Saldias, Felipe Ribeiro, Jim Ratter, Sam Bridgewater, Mario Sousa and Alberto Reyes for help with fieldwork. We thank Michelle Hollingsworth, Alex Ponge, Stephen O'Sullivan, Will Goodall-Copestake and Ben Mosse for assistance in the laboratory at RBGE. We thank Chris Dick, Francis Mayle and Oliver Phillips for valuable comments on the manuscript.

#### APPENDIX A: DATASET FOR ANALYSIS WITH COMPONENT 2.0

#NEXUS

[Final COMPONENT file for SDTF analysis]

BEGIN TAXA;

DIMENSIONS N TAX=18;

TAXLABELS

caatinga [1]  
 misiones [2]  
 piedmont [3]  
 ecuperuandes [4]  
 ecuperucoast [5]  
 venezandes [6]  
 colvencoast [7]  
 mesoamerica [8]  
 caribbean [9]  
 perudesert  
 monte  
 mexdesert  
 savanna  
 wetamaz  
 restinga  
 chaco  
 wetmeso  
 wetatl

;

ENDBLOCK;

BEGIN DISTRIBUTION;

TITLE='Coursetia';

N TAX=97;

RANGE

Ccaribaeatri : mesoamerica [8],  
 Cpumila : mesoamerica [8],  
 Chidalgoana : mesoamerica [8],  
 Cglabella : mesoamerica [8],  
 Ccaribaea : mesoamerica [8] colvencoast [7] ecuperucoast [5] venezandes [6] caribbean [9] wetmeso,  
 Ccaribaeachiap: mesoamerica [8],  
 Candina : venezandes [6],

Ccaribaeaoch : ecuperucoast [5] ecuperuandes [4] piedmont [3],  
 Cmollis1 : mesoamerica [8],  
 ChassleriMi : misiones [2],  
 ChassleriPi : piedmont [3],  
 Cglandulosa : mesoamerica [8],  
 Cgrandiflora1 : ecuperuandes [4],  
 Cgracilis : ecuperuandes [4],  
 Cgrandiflora2 : ecuperuandes [4],  
 Cdubia : ecuperuandes [4],  
 Cmaranoniana : ecuperuandes [4],  
 Cfruticosa : ecuperuandes [4],  
 Cbrachyrachis : piedmont [3],  
 Cferruginea : colvencoast [7] mesoamerica [8] savanna,  
 Chintonii : mesoamerica [8],  
 Cmollis2 : mesoamerica [8],  
 Celliptica : mesoamerica [8],  
 Cchiapensis : mesoamerica [8],  
 Cpolyphylla : mesoamerica [8],  
 Cmadrensis : mesoamerica [8],  
 Coaxacensis : mesoamerica [8],  
 Cpaniculata : mesoamerica [8],  
 Cplanipetiolata : mesoamerica [8],  
 Crostrata : caatinga [1],  
 Cvicioides : caatinga [1],  
 Cweberbauri : perudesert,  
 Cheterantha : monte,  
 Corbicularis : ecuperuandes [4],  
 Chypoleuca : piedmont [3],  
 Chblanchetiana : caatinga [1],  
 Chlatisiliqua : ecuperucoast [5] wetmeso,  
 Chscandens1 : colvencoast [7],  
 Chscandens2 : mesoamerica [8] colvencoast [7] savanna caribbean [9],  
 Chklugii : wetamaz,  
 Chnigricans : misiones [2],  
 Chbrasiliensis : piedmont [3] misiones [2] wetamaz wetmeso wetatl,  
 Chlongiflora : wetatl,  
 Cglaziovii : caatinga [1],  
 Ngentryi : mesoamerica [8],  
 Nplatycalyx : mexdesert,  
 Nhirsuta : mesoamerica [8],  
 Ndiversifolia : mesoamerica [8] mexdesert,  
 Nmicroptera : mesoamerica [8] mexdesert,  
 Nleiogyne1 : mesoamerica [8],  
 Nleiogyne2 : mesoamerica [8],  
 Nshotii : mesoamerica [8] mexdesert,  
 Pportulacifolia : caribbean [9],  
 Pziniiflora : caribbean [9],  
 Pquisqueyana : caribbean [9],  
 Pbleo : wetmeso,  
 Paureiflora : caatinga [1],  
 Pguamacho : colvencoast [7],  
 Pgrandgrand : wetatl restinga,  
 Pbahiensis : caatinga [1],  
 Pstenatha : caatinga [1],  
 Pgrandviol : wetatl,  
 Psacharosa : piedmont [3] misiones [2],  
 Pnemorosa : misiones [2],  
 Pweberiana : piedmont [3],  
 Pdiazromeriana : piedmont [3],  
 Phumbhumb : ecuperuandes [4],  
 Phumbrauh : ecuperuandes [4],  
 Paculeata : colvencoast misiones [2] caribbean [9] wetatl restinga,  
 Plychnidiflora : mesoamerica [8],

Leptoqua : caribbean [9],  
 Calymmalthium: ecuperuandes [4],  
 Rcrugeri : wetamaz,  
 Rbrachysepala : wetamaz,  
 Rtangarana : wetamaz,  
 Rramiflora : colvencoast [7] wetamaz wetmeso,  
 Rcurranii : colvencoast [7],  
 Rtenuiflora : wetamaz,  
 Rapurensis : wetamaz,  
 Rcarina : colvencoast [7],  
 Rchiapensis : mesoamerica [8] wetmeso,  
 Rcostaricensis : mesoamerica [8],  
 Rpallida : mesoamerica [8],  
 Rfusca : mesoamerica [8],  
 Rlaevigata : mesoamerica [8],  
 Rlaxiflora : misiones [2] caatinga [1] piedmont [3],  
 Rfagifolia : caatinga [1],  
 RapertaPe : ecuperuandes [4],  
 Ralbida : ecuperuandes [4],  
 RapertaEc : ecuperucoast [5],  
 Rapetala : piedmont chaco,  
 Rtriflora : chaco,  
 Agardneri : caatinga [1],  
 Cgrisebachii : piedmont [3],  
 Lhuasango : ecuperucoast [5] ecuperuandes [4],  
 Lgrisebachii : piedmont [3],  
 Lsagotii : wetamaz;

## TREE

T1=(((12,(8,(5,6,7,(1,(4,(2,3))))),(11,(9,10))),((13,14),(15,16))),((20,(19,(17,18))),((21,22))),((23,(24,25)),(28,29,  
 (26,27))),((30,31)),((32,33),(34,35)),(36,37,((44,(40,(38,39))),((41,(42,43))),((45,46,47,48,49,50,(51,52))))),(((56,(55,  
 (53,54))),((57,58))),((62,(59,(60,61))),((63,64))),((70,(71,72))),((69,((65,66),(67,68))))),((((((73,(74,75)),(76,77)),  
 (78,79),80),81,82,83,84,85),86,87),((88,89),90)),91,92),((93,94),((95,96),97)));  
 ENDBLOCK;

**APPENDIX B: VOUCHER SPECIMEN INFORMATION FOR THE SAMPLING OF MOLECULAR DATA FOR  
 PHYLOGENETIC ANALYSIS OF *CHAETOCALYX* AND *NISSOLIA***

species	locality	voucher	GenBank accession
<i>Adesmia corymbosa</i> Clos	Argentina: Chubut	Lavin 8283 (MONT)	AF183494
<i>Adesmia volckmannii</i> 1	Argentina: Neuquén	Lavin 8258 (MONT)	U59893
<i>Amicia medicaginea</i> 737	Argentina: Salta	Lavin 5773 (MONT)	AF183501
<i>Amicia glandulosa</i> 476	Ecuador	Pennington 654 (E)	AF183502
<i>Chaetocalyx blanchetiana</i> 799	Brazil: Bahia	Coradin 6585 (NY)	AY253887
<i>Chaetocalyx blanchetiana</i> 893	Brazil: Bahia	Hatschbach 56922 (F)	AY253896
<i>Chaetocalyx brasiliensis</i> 584	Mexico: Veracruz	Ventura 14977 (MO)	AY253877
<i>Chaetocalyx brasiliensis</i> 744	Argentina: Corrientes	Vanni 76 (MO)	AY253888
<i>Chaetocalyx brasiliensis</i> 768	Argentina: Misiones	Prado s.n. (MONT)	AF183505
<i>Chaetocalyx brasiliensis</i> 912	Brazil: Goiás	Anderson 9595 (MO)	AY253874
<i>Chaetocalyx brasiliensis</i> 913	Peru: Madre de Dios	Funk 8327 (NY)	AY253871
<i>Chaetocalyx glaziovii</i> 942	Brazil: Bahia	Coradin 5741 (NY)	AF183507
<i>Chaetocalyx klugii</i> 741	Brazil: Acre	Daly 6778 (NY)	AF183506
<i>Chaetocalyx klugii</i> 746	Ecuador: Napo	Neill 9462 (MO)	AY253875
<i>Chaetocalyx latisiliqua</i> 745	Ecuador: Guayas	Dodson 12993 (MO)	AY253889
<i>Chaetocalyx latisiliqua</i> 890	Panama: Cerro Campana	D'Arcy 9554 (MO)	AY253902 and AY253879
<i>Chaetocalyx latisiliqua</i> 891	Costa Rica: Puntarenas	Kernan 122 (MO)	AF203566 and AY253903
<i>Chaetocalyx longiflora</i> 902	Brazil: Rondônia	Silva 5823 (NY)	AY253890
<i>Chaetocalyx longiflora</i> 903	Brazil: Distrito Federal	Silva 177 (MO)	AY253876
<i>Chaetocalyx longiflora</i> 904	Brazil: São Paulo	Norris 224 (NY)	AY253901
<i>Chaetocalyx longiflora</i> 906 (probably <i>C. glaziovii</i> )	Brazil: Bahia	Harley 21608 (NY)	AY253872

<i>Chaetocalyx longiflora</i> 944	Brazil: Sabara	Mello Bareto 5704 (F)	AY253873
<i>Chaetocalyx nigricans</i> 688	Argentina: Misiones	Vanni 2955 (F)	AF183508
<i>Chaetocalyx scandens</i> 666	Brazil: Roraima	CIAT20536 (MONT)	AF183509
<i>Chaetocalyx scandens</i> 897	Caribbean: Dominican Republic	Liogier 13528 (NY)	AY253898 and AY253892
<i>Chaetocalyx scandens</i> 898	Caribbean: Grenadines, Carriacou	Howard 10845 (NY)	AY253899 and AY253893
<i>Chaetocalyx scandens</i> 900	Venezuela: Distrito Federal	Guerere 36 (MO)	AY253900 and AY253894
<i>Chaetocalyx scandens</i> 901	Venezuela: Distrito Federal	Ramírez 2641 (NY)	AY253897 and AY253895
<i>Nissolia diversifolia</i> 772	Mexico: Nuevo Leon	Montery F-1505682 (F)	AY253880
<i>Nissolia gentryi</i> 464	Mexico: Sonora	Van Devender 93-189 (NY)	AY253881
<i>Nissolia hirsuta</i> 466	Mexico: Morelos	Flores 228 (NY)	AY253882
<i>Nissolia hirsuta</i> 774	Mexico: Ixtapan	Roe 1904 (F)	AY253883
<i>Nissolia leiogyne</i> 776	Mexico: Tapalapa	Magallanes 2902 (F)	AF203563
<i>Nissolia leiogyne</i> 780	Mexico: Jalisco	Wilbur 2383 (NY)	AY253878
<i>Nissolia microptera</i> 773	Mexico: Nayarit	Flores 1302 (NY)	AY253884
<i>Nissolia platycalyx</i> 467	Mexico: Coahuila	Johnston 11028 (NY)	AY253885
<i>Nissolia platycalyx</i> 779	Mexico: Coahuila	Johnston 12065 (NY)	AY253886
<i>Nissolia schottii</i> 619	Mexico: Sonora	Joyal 2094 (NY)	AF183510
<i>Poiretia angustifolia</i> 820	Brazil	Fonseca <i>et al.</i> 1419 (MO)	AF183503
<i>Poiretia punctata</i> 823	Ecuador	Madsen 63491 (MO)	AF183504
<i>Zornia</i> sp. 858	Mexico	Lavin 5039 (MONT)	AF183500

#### APPENDIX C: VOUCHER SPECIMEN INFORMATION FOR THE SAMPLING OF MOLECULAR DATA FOR PHYLOGENETIC ANALYSIS OF *LOXOPTERYGIUM*

species	locality	voucher	GenBank accession
<i>Cardenasiodendron brachypterum</i>	Bolivia: Santa Cruz	Pendry 691, RBGE	AY531204
<i>Loxopterygium grisebachii</i>	Bolivia: Santa Cruz	Pendry 678, RBGE	AY531207
<i>Loxopterygium huasango</i> 1	Peru: Cajamarca	Pennington 813, RBGE	AY531205
<i>Loxopterygium huasango</i> 2	Peru: Piura	Pennington 816, RBGE	AY531206
<i>Loxopterygium sagotii</i>	Guyana	Polak 309, RBGE	AY531208
<i>Loxostylis alata</i>	grown from seed collected in South Africa	NYBG living collection 488/84	AY531201
<i>Schinopsis brasiliensis</i>	Brazil	RBGE living collection 1990880A	AY531203
<i>Schinus areira</i>	Bolivia: Santa Cruz	Pendry 737, RBGE	AY531202

#### REFERENCES

- Alpers, C. N. & Brimhall, G. H. 1988 Middle Miocene climatic change in the Atacama Desert, northern Chile: evidence from supergene mineralization at La Escondida. *Bull. Geol. Soc. Am.* **100**, 1640–1656.
- Ax, P. 1987 *The phylogenetic system*. Chichester, UK: John Wiley.
- Barkley, F. A. 1962 Anacardiaceae: Rhoideae: *Loxopterygium*. *Lloydia* **25**, 109–122.
- Barneby, R. C. & Grimes, J. W. 1996 Silk tree, Guanacaste, monkey's earring. A generic system for the synandrous Mimosaaceae of the Americas. Part I. *Abarema*, *Albizia* and allies. *Mem. New York Bot. Gard.* **74**, 1–292.
- Beyra-M, A. & Lavin, M. 1999 Monograph of *Pictetia* (Leguminosae–Papilionoideae) and review of the Aeschynomeneae. *Syst. Bot. Monogr.* **56**, 1–93.
- Bridgewater, S., Pennington, R. T., Reynel, C. A., Daza, A. & Pennington, T. D. 2003 A preliminary floristic and phytogeographic analysis of the woody flora of seasonally dry forests in northern Peru. *Candollea* **58**, 129–148.
- Buckler IV, E. S., Anthony, A. & Holtsford, T. P. 1997 The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics* **145**, 821–832.
- Burnham, R. J. & Graham, A. 1999 The history of neotropical vegetation: new developments and status. *Ann. Missouri Bot. Gard.* **86**, 546–589.
- Bush, M. B. 2002 Distributional change and conservation on the Andean flank: a palaeoecological perspective. *Global Ecol. Biogeogr.* **11**, 463–473.
- Bush, M. B., Colinvaux, P. A., Weinmann, M. C., Piperno, D. R. & Liu, K.-B. 1990 Late Pleistocene temperature depression and vegetation change in Ecuadorian Amazon. *Quatern. Res.* **34**, 330–345.
- Cabrera, A. L. 1976 *Regiones Fitogeográficas Argentinas*, 2nd edn. Buenos Aires: ACME S.A.C.I., Enciclopedia Argentina de Agricultura y Jardinería.
- Clapperton, C. M. 1993 Nature of environmental changes in South America at the last glacial maximum. *Palaeogeogr. Palaeoclimat. Palaeoecol.* **101**, 189–208.
- Coates, A. G. & Obando, J. A. 1996 The geologic evolution of the Central American Isthmus. In *Evolution and environment in tropical America* (ed. J. B. C. Jackson, A. F. Budd & A. G. Coates), pp. 21–56. University of Chicago Press.
- Colinvaux, P. A., de Oliveira, P. E., Moreno, J. E., Miller, M. C. & Bush, M. B. 1996 A long pollen record from lowland Amazonia: forest and cooling in glacial times. *Science* **274**, 85–87.

- Colinvaux, P. A., Irion, G., Räsänen, M. E. & Bush, M. B. 2001 A paradigm to be discarded: geological and paleoecological data falsify the Haffer & Prance refuge hypothesis of Amazonian speciation. *Amazoniana* **16**, 609–646.
- Cowling, S. A., Maslin, M. A. & Sykes, M. T. 2001 Paleovegetation simulations of lowland Amazonia and implications for neotropical allopatry and speciation. *Quatern. Res.* **55**, 140–149.
- Delgado-Salinas, A., Turley, T., Richman, A. & Lavin, M. 1999 Phylogenetic analysis of the cultivated and wild species of Phaseolus (Fabaceae). *Syst. Bot.* **24**, 438–460.
- Farris, J. S. 1988 HENNIG86, v. 1.5 (computer software and manual). Available from <http://www.cladistics.org/education/hennig86.html>.
- Felsenstein, J. 1982 Numerical methods for inferring evolutionary trees. *Q. Rev. Biol.* **57**, 379–404.
- Fischer, A. G. 1960 Latitudinal variations in organic diversity. *Evolution* **14**, 64–81.
- Furley, P. A. & Ratter, J. A. 1988 Soil resources and plant communities of the central Brazilian cerrado and their development. *J. Biogeogr.* **15**, 97–108.
- Gentry, A. H. 1995 Diversity and floristic composition of neotropical dry forests. In *Seasonally dry tropical forests* (ed. S. H. Bullock, H. A. Mooney & E. Medina), pp. 146–194. Cambridge University Press.
- Glor, R. E., Vitt, L. J. & Larson, A. 2001 A molecular phylogenetic analysis of diversification in Amazonian *Anolis* lizards. *Mol. Ecol.* **10**, 2661–2668.
- Graham, A. & Dilcher, D. 1995 The Cenozoic record of tropical dry forest in northern Latin America and the southern United States. In *Seasonally dry tropical forests* (ed. S. H. Bullock, H. A. Mooney & E. Medina), pp. 124–145. Cambridge University Press.
- Haberle, S. 1997 Upper Quaternary vegetation and climate history of the Amazon Basin: correlating marine and terrestrial pollen records. In *Proceedings of the Ocean Drilling Program, Scientific Results*, vol. 155 (ed. R. D. Flood, D. J. W. Piper, A. Klaus & L. C. Peterson), pp. 381–396. Ocean Drilling Program, Texas A&M University.
- Haffer, J. 1969 Speciation in Amazonian forest birds. *Science* **165**, 131–137.
- Hennig, W. 1966 *Phylogenetic systematics*. Chicago, IL: University of Illinois Press.
- Herendeen, P. S., Crepet, W. L. & Dilcher, D. L. 1992 The fossil history of the Leguminosae: phylogenetic and biogeographic implications. In *Advances in legume systematics, part 4: the fossil record* (ed. P. S. Herendeen & D. L. Dilcher), pp. 303–316. Kew, London: Royal Botanic Gardens.
- Hillis, D. M. 1996 Inferring complex phylogenies. *Nature* **383**, 140–141.
- Hooghiemstra, H. 1997 *Tropical rain forest versus savanna; two sides of a precious medal*, pp. 32–43. NWO/Huygenslezing (1997). The Hague, The Netherlands: The Netherlands Foundation for Scientific Research NWO.
- Hoorn, C. 1997 Palynology of the Pleistocene glacial/interglacial cycles of the Amazon fan (holes 940A, 944A, and 946A). In *Proceedings of the Ocean Drilling Program, Scientific Results*, vol. 155 (ed. R. D. Flood, D. J. W. Piper, A. Klaus & L. C. Peterson), pp. 397–409. Ocean Drilling Program, Texas A&M University.
- Huang, Y., Street-Perrott, F. A., Metcalfe, S. E., Brenner, M., Moreland, M. & Freeman, K. H. 2001 Climate change as the dominant control on glacial–interglacial variations in C<sub>3</sub> and C<sub>4</sub> plant abundance. *Science* **293**, 1647–1651.
- Huelsenbeck, J. P. & Ronquist, F. R. 2001 MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17**, 754–755.
- Huelsenbeck, J., Ronquist, F. R., Nielson, R. & Bollback, J. P. 2001 Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294**, 2310–2314.
- Humphries, C. J. 1992 Cladistic biogeography. In *Cladistics: a practical course in systematics*. Systematics Association Publication 10 (ed. P. L. Forey, C. J. Humphries, I. J. Kitching, R. W. Scotland, D. J. Siebert & D. M. Williams), pp. 137–159. Oxford University Press.
- Humphries, C. J. & Parenti, L. R. 1999 *Cladistic biogeography*, 2nd edn. Oxford University Press.
- Jefferies, R. P. S. 1979 The origin of the chordates: a methodological essay. In *The origin of the major vertebrate groups* (ed. M. R. House), pp. 443–477. London: Academic.
- Kessler, M. & Helme, N. 1999 Floristic diversity and phylogeography of the central Tuichi Valley, an isolated dry forest locality in the Bolivian Andes. *Candollea* **54**, 341–366.
- Langley, C. H. & Fitch, W. 1974 An estimation of the constancy of the rate of molecular evolution. *J. Mol. Evol.* **3**, 161–177.
- Lavin, M. 1988 Systematics of *Coursetia* (Leguminosae-Papilionoideae). *Syst. Bot. Monogr.* **21**, 1–167.
- Lavin, M. 1998 Origin, diversity and biogeography of neotropical Fabaceae. *Monogr. Syst. Bot. Missouri Bot. Gard.* **68**, 131–145.
- Lavin, M., Pennington, R. T., Klitgaard, B. B., Sprent, J. I., de Lima, H. C. & Gasson, P. E. 2001a The dalbergioid legumes (Fabaceae): delimitation of a pantropical monophyletic clade. *Am. J. Bot.* **88**, 503–533.
- Lavin, M., Wojciechowski, M. F., Richman, A., Rotella, J., Sanderson, M. J. & Beyra Matos, A. 2001b Identifying Tertiary radiations of Fabaceae in the Greater Antilles: alternatives to cladistic vicariance analysis. *Int. J. Pl. Sci.* **162**(Suppl. 6), S53–S76.
- Lavin, M., Wojciechowski, M. F., Gasson, P., Hughes, C. E. & Wheeler, E. 2003 Phylogeny of robinoid legumes (Fabaceae) revisited: *Coursetia* and *Gliricidia* recircumscribed, and a biogeographical appraisal of the Caribbean endemics. *Syst. Bot.* **28**, 387–409.
- Leuenberger, B. E. 1986 *Pereskia* (Cactaceae). *Mem. New York Bot. Gard.* **41**, 1–141.
- Li, J., Huang, H. & Sang, T. 2002 Molecular phylogeny and infrageneric classification of *Actinidia* (Actinidiaceae). *Syst. Bot.* **27**, 408–415.
- Liu, K.-B. & Colinvaux, P. A. 1985 Forest changes in the Amazon during the last glacial maximum. *Nature* **318**, 556–557.
- Magallón, S., Crane, P. R. & Herendeen, P. S. 1999 Phylogenetic pattern, diversity and diversification of eudicots. *Ann. Missouri Bot. Gard.* **86**, 297–372.
- Mayle, F. E., Burbidge, R. & Killeen, T. J. 2000 Millennial-scale dynamics of southern Amazonian rain forests. *Science* **290**, 2291–2294.
- Mayle, F. E., Beerling, D. J., Gosling, W. D. & Bush, M. B. 2004 Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum. *Phil. Trans. R. Soc. Lond. B* **359**, 499–514. (DOI 10.1098/rstb.2003.1434.)
- Mitchell, J. D. & Daly, D. C. 1991 *Cyrtocarpa* Kunth (Anacardiaceae) in South America. *Ann. Missouri Bot. Gard.* **78**, 184–189.
- Mooney, H. A., Bullock, S. H. & Medina, E. 1995 Introduction. In *Seasonally dry tropical forests* (ed. S. H. Bullock, H. A. Mooney & E. Medina), pp. 1–8. Cambridge University Press.
- Mori, S. A., Boom, B. M. & Prance, G. T. 1981 Distribution patterns and conservation of eastern Brazilian coastal forest tree species. *Brittonia* **33**, 233–245.
- Moritz, C., Patton, J. L., Schneider, C. J. & Smith, T. B. 2000 Diversification of rainforest faunas: an integrated molecular approach. *A. Rev. Ecol. Syst.* **31**, 553–563.

- Morrone, J. J. & Carpenter, J. M. 1995 In search of a method for cladistic biogeography: an empirical comparison of component analysis, Brooks parsimony analysis, and three area statements. *Cladistics* **10**, 99–153.
- Mosse, B. 2000 The molecular phylogeny of *Loxopterygium*. Honours thesis, University of Edinburgh, UK.
- Murphy, P. & Lugo, A. E. 1986 Ecology of tropical dry forest. *Ann. Rev. Ecol. Syst.* **17**, 67–88.
- Myers, N. & Knoll, A. H. 2001 The biotic crisis and the future of evolution. *Proc. Natl Acad. Sci. USA* **98**, 5389–5392.
- Nee, S., May, R. M. & Harvey, P. H. 1994 The reconstructed evolutionary process. *Phil. Trans. R. Soc. Lond. B* **344**, 305–311.
- Nelson, B. W., Ferreira, C. A. C., da Silva, M. F. & Kawasaki, M. L. 1993 Endemism centres, refugia and botanical collection density in the Brazilian Amazonia. *Nature* **345**, 714–716.
- Nelson, G. & Ladiges, P. 1991 Standard assumptions for biogeographic analysis. *Aust. Syst. Bot.* **4**, 41–58.
- Nelson, G. & Ladiges, P. 1994 *TASS. Three area subtrees*. New York and Melbourne: published by the authors.
- Nelson, G. & Ladiges, P. 1995 *TAX: MSDos computer programs for systematics*. New York and Melbourne: published by the authors.
- Nelson, G. & Ladiges, P. 1996 Paralogy in cladistic biogeography and analysis of paralogy-free subtrees. *Am. Mus. Novit.* **3167**, 1–58.
- Nixon, K. C. & Carpenter, J. M. 1993 On outgroups. *Cladistics* **9**, 413–426.
- Page, R. D. M. 1990 Component analysis: a valiant failure? *Cladistics* **6**, 119–136.
- Page, R. D. M. 1993 *COMPONENT*, v. 2. London: The Natural History Museum.
- Pendry, C. A. 2004 A monograph of *Ruprechtia*. *Syst. Bot. Monogr.* **67**, 1–113.
- Pennington, R. T., Prado, D. A. & Pendry, C. 2000 Neotropical seasonally dry forests and Pleistocene vegetation changes. *J. Biogeogr.* **27**, 261–273.
- Pennington, R. T., Pendry, C. A., Goodall-Copstake, W. & O'Sullivan, S. 2004 In A monograph of *Ruprechtia* (ed. C. A. Pendry). *Syst. Bot. Monogr.* **67**, 12–17.
- Posada, D. & Crandall, K. A. 1998 MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.
- Prado, D. E. & Gibbs, P. E. 1993 Patterns of species distributions in the dry seasonal forests of South America. *Ann. Missouri Bot. Gard.* **80**, 902–927.
- Prance, G. T. 1973 Phytogeographic support for the theory of Pleistocene forest refuges in the Amazon Basin, based upon evidence from distribution pattern in Caryocaraceae, Chrysobalanaceae, Dichapetalaceae and Lecythidaceae. *Acta Amazon.* **3**, 5–28.
- Rambaut, A. 1996 SEQ-GEN v. 1.0 alfa 1: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Computer Appl. Biosci.* **13**, 235–238.
- Ratter, J. A., Askew, G. P., Montgomery, R. & Gifford, D. R. 1978 Observations on forests of some mesotrophic soils in central Brazil. *Revista Bras. Bot.* **1**, 47–58.
- Ratter, J. A., Ribeiro, J. F. & Bridgewater, S. 1997 The Brazilian cerrado vegetation and threats to its biodiversity. *Ann. Bot.* **80**, 223–230.
- Richardson, J. A., Pennington, R. T., Pennington, T. D. & Hollingsworth, P. 2001 Recent and rapid diversification of a species-rich genus of neotropical trees. *Science* **293**, 2242–2245.
- Rizzini, C. T. 1975 Novitates florae Brasiliensis. *Leandra* **6**, 40–43.
- Rudd, V. E. 1956 A revision of the genus *Nissolia*. *Contrib. US Natl Herb.* **32**, 173–206.
- Rudd, V. E. 1958 A revision of the genus *Chaetocalyx*. *Contrib. US Natl Herb.* **32**, 207–243.
- Rudd, V. E. 1970 Revival of *Nissolia microptera* (Leguminosae). *Phytologia* **20**, 324.
- Rudd, V. E. 1972 Supplementary studies in *Chaetocalyx*. I. (Leguminosae). Including a new species from Brazil. *Phytologia* **24**, 295–297.
- Rudd, V. E. 1975 *Nissolia chiapensis*, a new species of Leguminosae from Mexico. *Phytologia* **31**, 427–430.
- Rudd, V. E. 1996 *Chaetocalyx longiloba* (Fabaceae, Papilionoideae), a new species from Peru. *Novon* **6**, 119.
- Sanderson, M. J. 1997 A non-parametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* **14**, 1218–1231.
- Sanderson, M. J. 1998 Estimating rate and time in molecular phylogenies: beyond the molecular clock. In *Molecular systematics of plants* (ed. D. Soltis, P. Soltis & J. J. Doyle), pp. 108–236. New York: Chapman & Hall.
- Sanderson, M. J. 2001 *R8s, v. 1(beta), user's manual (June 2001)*. Davis, CA: University of California: distributed by M. J. Sanderson. See [http://ginger.ucdavis.edu/r8s/](http://ginger.ucdavis.edu/r8s/http://ginger.ucdavis.edu/r8s/).
- Sanderson, M. J. 2002 Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* **19**, 101–109.
- Stebbins, G. L. 1974 *Evolution above the species level*. Cambridge, MA: Harvard University Press.
- Swofford, D. L. 2000 *PAUP\*. Phylogenetic analysis using parsimony (\* and other methods)*, v. 4.0b4a. Sunderland, MA: Sinauer.
- Thomas, W. W. 1999 Conservation and monographic research on the flora of Tropical America. *Biodiv. Conserv.* **8**, 1007–1015.
- Van der Hammen, T. & Absy, M. L. 1994 Amazonia during the last glacial. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **109**, 247–261.
- Wallace, R. S. 1995 Molecular systematic study of Cactaceae: using chloroplast DNA variation to elucidate Cactus phylogeny. *Bradleya* **13**, 1–12.
- Wendt, T. 1993 Composition, floristic affinities and origins of the Mexican Atlantic slope rain forests. In *Biological diversity of Mexico, origins and distribution* (ed. T. P. Ramamoorthy, R. Bye, A. Lot & J. Fa), pp. 595–680. Oxford University Press.
- Wheeler, E. A. 2001 Fossil dicotyledonous woods from the Florissant Fossil Beds National Monument, Colorado. In *Proc. Denver Mus. Sci. Ind., ser. 4* (ed. E. Evanoff, K. M. Gregory-Wodziki & K. R. Johnson), pp. 197–213. Denver, CO: Denver Museum of Science and Industry.
- Wheeler, E. A. & Landon, J. 1992 Late Eocene (Chadronian) dicotyledonous woods from Nebraska: evolutionary and ecological significance. *Rev. Palaeobot. Palynol.* **74**, 267–282.
- Wojciechowski, M. F. 2003 Reconstructing the phylogeny of legumes (Fabaceae): an early 21st century perspective. In *Advances in legume systematics, part 10* (ed. A. Bruneau & B. Klitgaard), pp. 5–35. Kew, UK: Royal Botanic Gardens.
- Woodruff, D. S. 2001 Declines of biomes and biotas and the future of evolution. *Proc. Natl Acad. Sci. USA* **98**, 5471–5476.

## GLOSSARY

- CI: consistency index  
 ITS: internal transcribed spacer  
 LGM: last glacial maximum  
 LF: Langley–Fitch  
 NPRS: non-parametric rate smoothing  
 PL: penalized likelihood  
 RI: retention index  
 SDTF: seasonally dry tropical forest