

Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes

Matt Lavin^{1*}, Brian P. Schrire², Gwilym Lewis², R. Toby Pennington³, Alfonso Delgado-Salinas⁴, Mats Thulin⁵, Colin E. Hughes⁶, Angela Beyra Matos⁷ and Martin F. Wojciechowski⁸

¹Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717, USA

²Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK

³Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh EH3 5LR, UK

⁴Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Apartado Postal 70-233, 04510, Distrito Federal, Mexico

⁵Department of Systematic Botany, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden

⁶Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK

⁷Medio Ambiente de Camagüey (CIMAC), Ministerio de Ciencias, Tecnología y Medio Ambiente, Cisneros 105 entre Pobre y Angel, Camagüey, C.P. 70 100, Cuba

⁸School of Life Sciences, Arizona State University, Tempe, AZ 85287-4501, USA

Penalized likelihood estimated ages of both densely sampled intracontinental and sparsely sampled transcontinental crown clades in the legume family show a mostly Quaternary to Neogene age distribution. The mode ages of the intracontinental crown clades range from 4–6 Myr ago, whereas those of the transcontinental crown clades range from 8–16 Myr ago. Both of these young age estimates are detected despite methodological approaches that bias results toward older ages. Hypotheses that resort to vicariance or continental history to explain continental disjunct distributions are dismissed because they require mostly Palaeogene and older tectonic events. An alternative explanation centring on dispersal that may well explain the geographical as well as the ecological phylogenetic structure of legume phylogenies is Hubbell's unified neutral theory of biodiversity and biogeography. This is the only dispersalist theory that encompasses evolutionary time and makes predictions about phylogenetic structure.

Keywords: Bayesian likelihood analysis; penalized likelihood rate smoothing; molecular age determination; intracontinental crown clades; transcontinental sister clades; unified neutral theory of biodiversity and biogeography

1. INTRODUCTION

Geographical phylogenetic structure refers to geographical proximity predicting phylogenetic relatedness (figure 1*a*; Irwin 2002). Similarly, phylogenetic relatedness can be predicted by ecological setting or community membership (figure 1*b,c*, respectively; Webb 2000). Geographical phylogenetic structure of continentally distributed clades is often attributed to vicariance events, such as the submergence of land bridges or loss of filter barriers to dispersal (e.g. Donoghue *et al.* 2001). The Early Tertiary North Atlantic land bridge and the Tertiary to Quaternary Beringian land bridge are classic examples of such land bridges and filter barriers that have supposedly

caused Northern Hemisphere vicariance patterns (e.g. Tiffney 1985; Sanmartín *et al.* 2001). Other examples include Gaarlandia, a putative Early Tertiary land bridge or filter barrier that perhaps resulted in Caribbean vicariance patterns (Iturralde-Vinent & MacPhee 1999), and the Walvis Ridge, a putative Oligocene filter barrier that may have resulted in disjunct African and South American distributions (e.g. Morley & Dick 2003) after the southern Atlantic had formed during the Cretaceous (e.g. Parrish 1993).

We question the universality of vicariance explanations for several reasons. Species diversifications endemic to large continental islands, such as Madagascar, typically show the same phylogenetic structure as those endemic to smaller oceanic islands (figures 2–5). Usually, the estimated ages of island stem and crown clades are diverse, are independent of the geological age of the island and, in the case of continental islands, often do not correspond to the

* Author for correspondence (mlavin@montana.edu).

One contribution of 16 to a Discussion Meeting Issue 'Plant phylogeny and the origin of major biomes'.

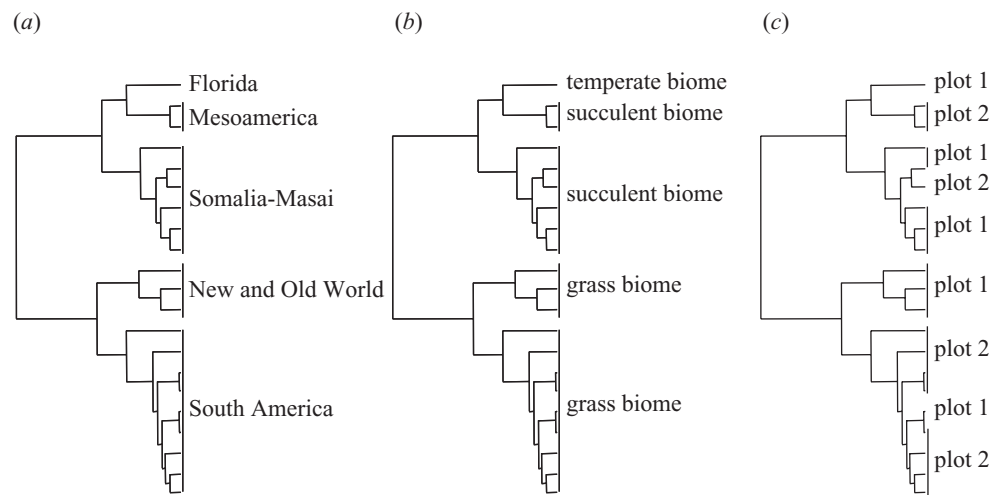


Figure 1. Different forms of phylogenetic structure. (a) Geographical phylogenetic structure arises when dispersal is limited following speciation events, such that a continental setting, for example, provides a good predictor of phylogenetic relationships. (b) Ecological phylogenetic structure arises when speciation and subsequent dispersal are confined ecologically, such that an ecological setting provides a good predictor of phylogenetic relationships. (c) Community phylogenetic structure arises when speciation and subsequent dispersal are confined ecologically, but that dispersal and ecological filtering result in closely related species residing side by side in the same ecological community (i.e. community membership becomes a good predictor of phylogenetic relationships; Webb 2000).

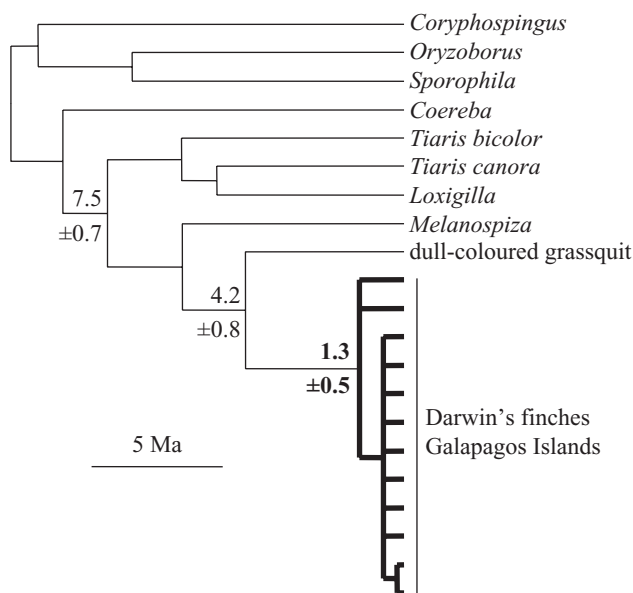


Figure 2. Phylogenetic structure associated with islands. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from data and information provided in Sato *et al.* (2001). Darwin's finches are endemic to the Galapagos Islands, oceanic islands with an estimated age of the island system at *ca.* 15 Myr (Werner *et al.* 1999). Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 20.847$, $r(\text{CG}) = 7.589$, $r(\text{AT}) = 4.146$, $r(\text{AG}) = 27.898$, $r(\text{AC}) = 3.254$, $p(\text{A}) = 0.0.270$, $p(\text{C}) = 0.287$, $p(\text{G}) = 0.178$, $p(\text{T}) = 0.266$, $\alpha = 1.749$, $iP = 0.273$. *numt3*: $3.0 - 3.1 \times 10^{-9}$ subs site⁻¹ yr⁻¹.

putative ages of alleged land bridges or filter barriers (e.g. Yoder *et al.* 2003; Renner 2004). Such geographical phylogenetic structure must be caused by sweepstake dispersal (Simpson 1940), where a low rate of immigration

co-occurs with a low rate of extinction of immigrants either because the island is relatively uninhabited or large in size (e.g. Losos & Schluter 2000).

For continental regions, similar phylogenetic structure is detected but often for widely separated geographical areas that have no historical tectonic relationship. The temperate North American genera *Robinia* and *Strophostyles*, for example, are each sister to the southern South American genera *Poissonia* and *Dolichopsis*, respectively (Lavin *et al.* 2003; Riley-Hulting *et al.* 2004). These two vicariant patterns are better explained by the mostly temperate to subtropical seasonal vegetation occupied by these genera. The genus *Arcoa* from the Dominican Republic is sister to *Tetrapterocarpon* from Madagascar (Herendeen *et al.* 2003), both inhabitants of a succulent-rich bushland characterized by Schrire *et al.* (2004) as the 'succulent biome'. The Northern Hemisphere *Gleditsia* (but with *G. amorphoides* from southern South America) and its sister *Umtiza* from South Africa (Herendeen *et al.* 2003; D. E. Prado, personal communication) commonly inhabit the succulent biome.

This study is motivated by a fruitless search for patterns of area relationships using cladistic vicariance methods among the Caribbean islands (e.g. Beyra-Matos & Lavin 1999; Lavin *et al.* 2001a), among the various South American dry forests (Pennington *et al.* 2004) and among other areas including the Somalia-Masai phytochorion in the region of the Horn of Africa (Lavin *et al.* 2000; Thulin *et al.* 2004). By contrast, the success of the cladistic vicariance analysis of a taxon–biome rather than a taxon–area cladogram of legumes (Schrire *et al.* 2004) underscored that phylogenetic structure in the legume phylogeny was caused as much by ecological setting as tectonic history. The resulting unrooted biome network produced during this analysis revealed a close relationship of Schrire *et al.*'s (2004) rainforest and grass biomes, and a similarly close relationship between the temperate regions of the

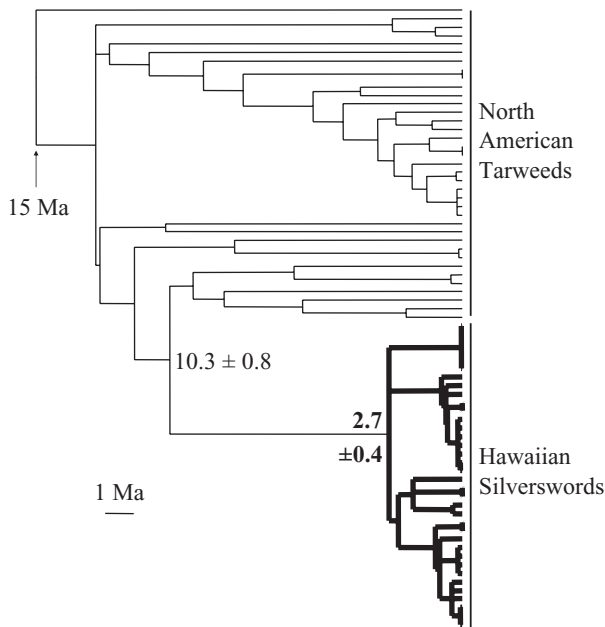


Figure 3. Phylogenetic structure associated with islands. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from data and information provided in Baldwin & Sanderson (1998). Silverswords are endemic to the Hawaiian Islands, oceanic islands with an estimated age for the entire system of *ca.* 5 Myr, but for the entire cyclical island system perhaps well over 40 Myr (Moore & Clague 1992). Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 5.178$, $r(\text{CG}) = 0.467$, $r(\text{AT}) = 1.080$, $r(\text{AG}) = 1.736$, $r(\text{AC}) = 0.786$, $p(\text{A}) = 0.234$, $p(\text{C}) = 0.229$, $p(\text{G}) = 0.247$, $p(\text{T}) = 0.290$, $\alpha = 0.619$, $iP = 0.235$. nrDNA ITS/5.8S: $5.7\text{--}6.9 \times 10^{-9}$ subs site $^{-1}$ yr $^{-1}$.

Northern and Southern Hemispheres, which comprise a general temperate biome. The succulent biome connected these two pairs, and the pattern was fractal (i.e. detected at different taxonomic levels) and thus constantly regenerated.

Molecular phylogenetic studies involving exhaustive sampling at the species level within the legume family reveal a remarkable level of geographical and ecological phylogenetic structure (e.g. Lavin *et al.* 2000, 2003; Hughes *et al.* 2003; Schrire *et al.* 2003; Pennington *et al.* 2004; Riley-Hulting *et al.* 2004; Thulin *et al.* 2004). By revealing a mostly Quaternary to Neogene age distribution of both intracontinental and transcontinental crown clades, we argue that such structure is the result of dispersal. Specifically, low but consistent historical immigration rates can add to continental biodiversity, but leave endemic species diversifications intact while ecological drift plays out, much as limited gene flow can leave an endemic gene pool intact while genetic drift plays out. Although we are not unanimous in discarding vicariance via tectonic history as a cause of geographical structure, an alternative meta-community explanation is a non-mutually exclusive interpretation of geographical structure detected in molecular phylogenies.

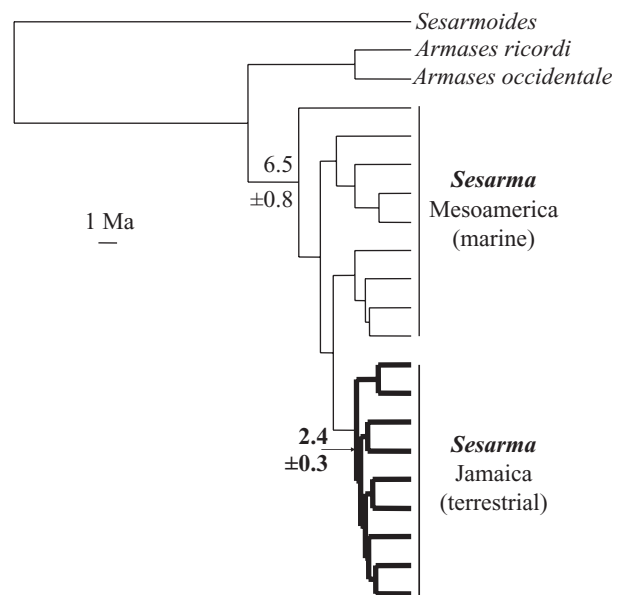


Figure 4. Phylogenetic structure associated with islands. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from data and information provided in Schubart *et al.* (1998). These crabs are endemic to Jamaica, a continental island that re-emerged during the Oligocene (Buskirk 1985). Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 34.671$, $r(\text{CG}) = 3.373$, $r(\text{AT}) = 4.159$, $r(\text{AG}) = 13.848$, $r(\text{AC}) = 1.794$, $p(\text{A}) = 0.338$, $p(\text{C}) = 0.115$, $p(\text{G}) = 0.152$, $p(\text{T}) = 0.394$, $\alpha = 1.035$, $iP = 0.560$. 16S+*cytb*: $12.1\text{--}15.6 \times 10^{-9}$ subs site $^{-1}$ yr $^{-1}$.

2. MATERIAL AND METHODS

(a) Data

The fundamental data in this study are the estimated ages in millions of years of two types of clades. One type includes exhaustively sampled crown clades, each of which is endemic to a particular continental region and has its sister group outside the continental region (e.g. Old and New World *Chapmannia*; figure 6). For each of these well-sampled intracontinental crown clades, the age of the crown is estimated to determine how long a particular diversification has been in residence within its confined area (e.g. nodes C, figure 6a). The high degree of endemism of these continental crown clades renders the geographical origin of the diversification unequivocal.

Inevitably, intracontinental crown clades over time suffer extinction and emigration such that the original geographical integrity is lost. To find the maximal age distribution of highly disintegrated continental crown clades, the age of a second type of clade is estimated: that of the transcontinental crown clade. Transcontinental crown clades comprise sister clades separated by an oceanic barrier or a large expanse of continental area, such as a given clade in North America and its sister in South America. Importantly, transcontinental crown clades are equivalent in age to intracontinental stem clades, the age of which is axiomatically the maximum for a continentally confined crown clade. Transcontinental crown clades are intentionally less well sampled and are thus prone to having their ages overestimated (figure 6b). That is, the ages of transcontinental crown clades are determined by tracing down from the terminal taxa in the legume plastid *matK* phylogeny of Wojciechowski *et al.* (2004) and Lavin *et al.* (2004) until a node is encountered which

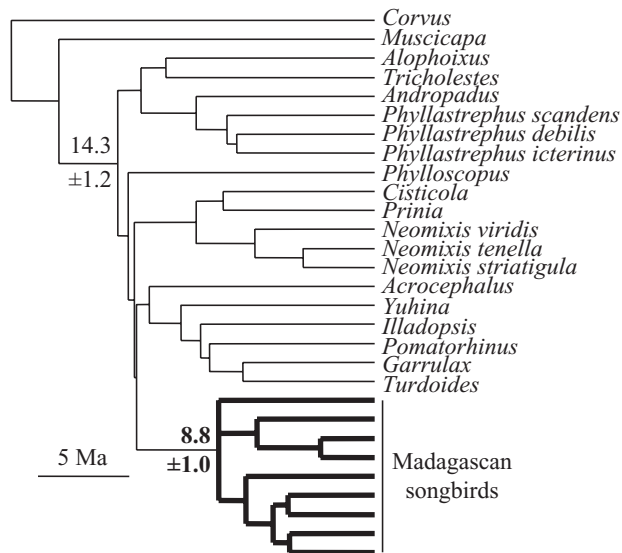


Figure 5. Phylogenetic structure associated with islands. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from data and information provided in *Cibois et al.* (2002). These songbirds are endemic to Madagascar, a continental island with an estimated age of well over 100 Myr (summarized in *Yoder et al.* 2003; *Goodman & Benstead* 2004). Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 8.776$, $r(\text{CG}) = 0.657$, $r(\text{AT}) = 1.571$, $r(\text{AG}) = 6.423$, $r(\text{AC}) = 1.643$, $p(\text{A}) = 0.316$, $p(\text{C}) = 0.403$, $p(\text{G}) = 0.097$, $p(\text{T}) = 0.185$, $1\text{st} = 0.273$, $2\text{nd} = 0.064$, $3\text{rd} = 2.663$. *cytb*: $31.9\text{--}32.7 \times 10^{-9}$ subs site⁻¹ yr⁻¹.

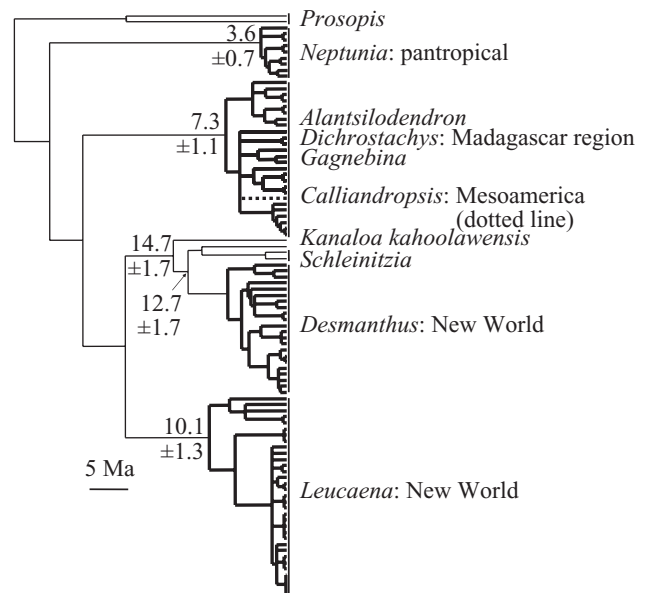


Figure 7. Phylogenetic structure on continents. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from the data and information provided in *Hughes et al.* (2003). Fixed root ages were derived from *Lavin et al.* (2005). The *Leucaena* and *Dichrostachys* groups show geographical phylogenetic structure spanning the Atlantic, but with *Kanaloa* from Hawaii and *Schleinitzia* from Southeast Asia. Pseudogene sequences were removed from the original dataset because they were not well sampled and showed great rate variation (*Bailey et al.* 2003). Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 3.481$, $r(\text{CG}) = 0.577$, $r(\text{AT}) = 1.010$, $r(\text{AG}) = 1.961$, $r(\text{AC}) = 0.829$, $p(\text{A}) = 0.198$, $p(\text{C}) = 0.337$, $p(\text{G}) = 0.333$, $p(\text{T}) = 0.132$, $\alpha = 0.832$, $iP = 0.217$. nrDNA ITS/5.8S: $3.0\text{--}4.1 \times 10^{-9}$ subs site⁻¹ yr⁻¹.

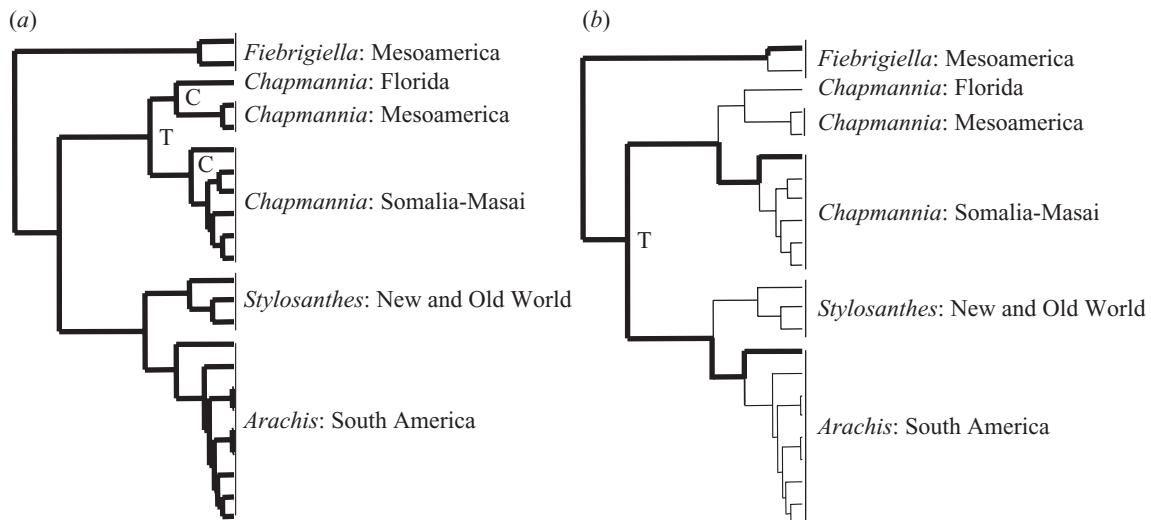


Figure 6. Transcontinental crown clades (T) and intracontinental crown clades (C). C stem clades are equivalent in age to T crown clades. The geographical origin of an exhaustively sampled C crown clade is unequivocal if all or nearly all members of the clade are confined to a single subcontinent. In this study, C crown clades were taxonomically densely sampled (a), whereas T crown clades were not (b, thick lines) and thus prone to age overestimation.

joins two lineages that are each endemic to a different continent. Because this *matK* phylogeny was sampled with respect to representation of just the major legume lineages, it is thus sparsely

sampled with respect to biogeographical questions and certainly the closest sister groups spanning different continents were not preferentially sampled.

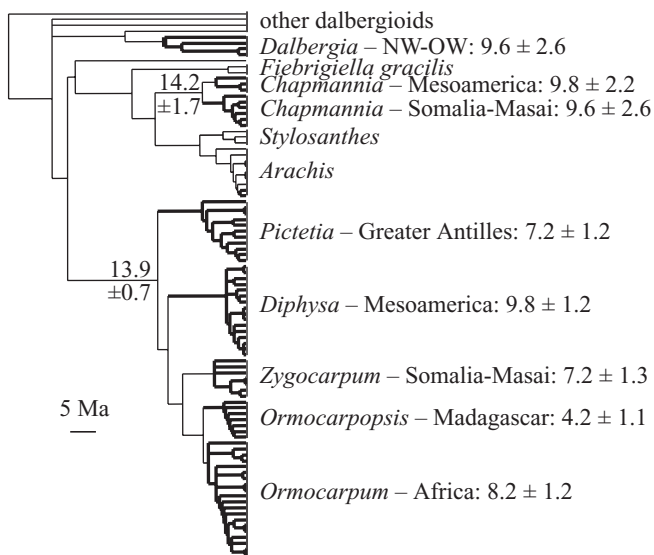


Figure 8. Phylogenetic structure on continents. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from the data and information provided in Lavin *et al.* (2000, 2004) and Thulin & Lavin (2001) (fixed root ages were derived from Lavin *et al.* (2005)). The *Chapmannia* and *Ormocarpum* groups span the Atlantic and show two independent cases of crown clades endemic to the Somalia-Masai region. Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 3.396$, $r(\text{CG}) = 0.647$, $r(\text{AT}) = 1.021$, $r(\text{AG}) = 1.938$, $r(\text{AC}) = 1.190$, $p(\text{A}) = 0.179$, $p(\text{C}) = 0.321$, $p(\text{G}) = 0.318$, $p(\text{T}) = 0.180$, $\alpha = 1.346$, $iP = 0.220$. ITS/5.8S: $3.8\text{--}8.5 \times 10^{-9}$ subs site $^{-1}$ yr $^{-1}$.

In addition to the legume-wide *matK* dataset mentioned above, many smaller but densely sampled legume datasets comprising either the nuclear ribosomal ITS/5.8S region or the plastid *matK* locus have been used in this present study. These smaller datasets can be placed in an age context because of a legume-wide estimation of absolute evolutionary rates and ages derived from the chloroplast loci *matK* and *rbcL* (Lavin *et al.* 2004). These smaller datasets subjected to rates and age estimation in this analysis include the robinoids (Lavin *et al.* 2001a, 2003), indigoferoids (Schrire *et al.* 2003), dalbergioids (Lavin *et al.* 2000, 2001b), the *Leucaena* and *Dichrostachys* groups (Hughes *et al.* 2003), the phaseoloids (Delgado-Salinas *et al.* 2004; Thulin *et al.* 2004; Riley-Hulting *et al.* 2004), Carmichaelinae (Wagstaff *et al.* 1999) and miscellaneous other groups for which dense taxon sampling can accurately estimate the position of intracontinental stem and crown clades (e.g. Wojciechowski *et al.* 2004). These studies should be consulted for the details concerning DNA isolations, primer specifications, PCR amplification conditions, DNA sequencing, data management, voucher specimen information and TREEBASE and GenBank accessions.

(b) Evolutionary rates analysis

The penalized likelihood (PL) method (Sanderson 2002) in the program r8s (Sanderson 2003) was used to estimate nucleotide substitution rates and ages of selected stem and crown clades within all of the *matK* and ITS/5.8S legume phylogenies. Via data-driven cross-validation, the penalized likelihood

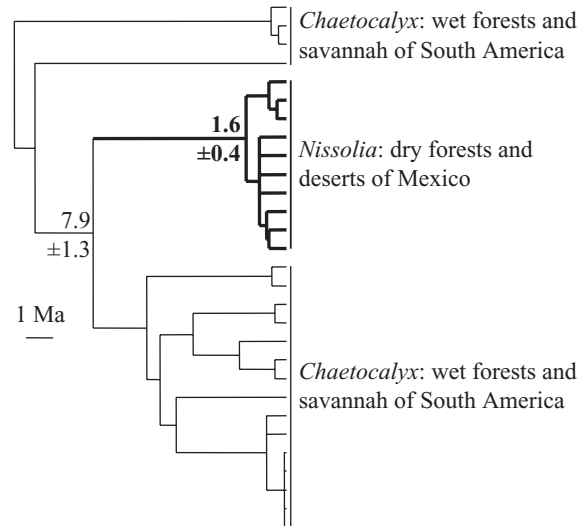


Figure 9. Phylogenetic structure on continents. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from the data and information provided in Pennington *et al.* (2004) (fixed root ages were derived from Lavin *et al.* (2005)). The mainly Mexican *Nissolia* crown is derived from within the mainly South American *Chaetocalyx* diversification. Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 15.276$, $r(\text{CG}) = 3.784$, $r(\text{AT}) = 5.522$, $r(\text{AG}) = 8.428$, $r(\text{AC}) = 3.573$, $p(\text{A}) = 0.184$, $p(\text{C}) = 0.309$, $p(\text{G}) = 0.313$, $p(\text{T}) = 0.194$, $\alpha = 1.876$, $iP = 0.236$. ITS/5.8S: $1.6\text{--}7.9 \times 10^{-9}$ subs site $^{-1}$ yr $^{-1}$.

approach finds an optimum rate-smoothing parameter for the transition of substitution rate between the ancestor and descendants. Rate smoothing by penalized likelihood generally should result in the selection of a model somewhere between the single parameter rate constant model (e.g. molecular clock; Langley & Fitch 1974) and a parameter-rich rate model (e.g. typically non-parametric rate smoothing; Sanderson 1997). For comparative purposes, these latter two approaches are also implemented in r8s.

Branch lengths were estimated using a Bayesian likelihood approach (Huelsenbeck & Ronquist 2001; Huelsenbeck *et al.* 2001). A nucleotide substitution model was selected using the Akaike information criterion, as implemented in MODELTEST (Posada & Crandall 1998) or manually for non-nested models (e.g. Johnson & Omland 2004). Multiple runs of Metropolis-coupled Monte-Carlo Markov were initiated with random starting trees and default substitution parameters. Each run comprised four chains of default temperatures and $5\text{--}10 \times 10^6$ permutations of tree parameters. The final branching-order parameter was validated with maximum parsimony using PAUP* (Swofford 2001). Tree parameters were sampled every $1\text{--}10 \times 10^5$ permutations after likelihood stationarity was attained. One hundred likelihood trees at stationarity were then systematically sampled (e.g. once every 100 000 trees) for the rates analysis to estimate means and variances for the rate and age estimates. Relative substitution rates and ages estimated with r8s were converted to absolute rates and ages by enforcing age constraints derived from the fossil record, which were derived from Lavin *et al.* (2005).

Estimating the ages of transcontinental crown clades was undertaken to obtain maximum ages of crown clades endemic to

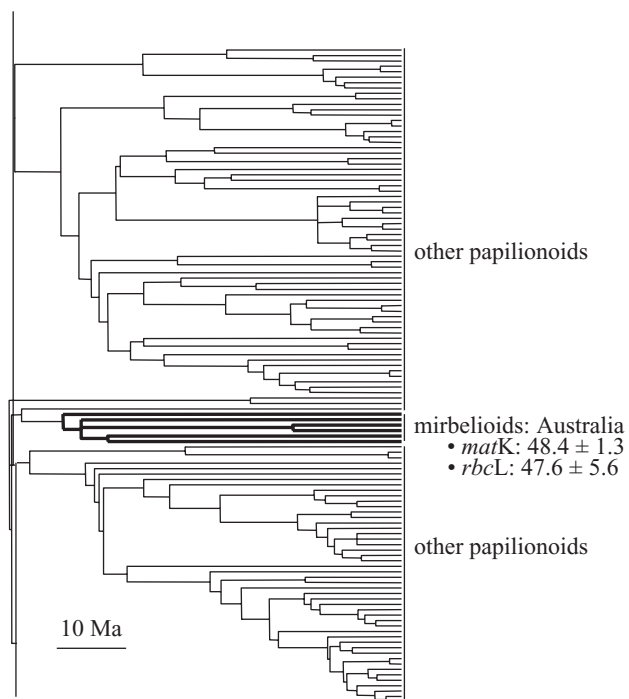


Figure 10. Phylogenetic structure on continents. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from the data and information provided in Lavin *et al.* (2005). The mirbelioid crown clade is almost entirely confined to Australia. Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 1.819$, $r(\text{CG}) = 1.356$, $r(\text{AT}) = 0.238$, $r(\text{AG}) = 1.959$, $r(\text{AC}) = 1.550$, $p(\text{A}) = 0.335$, $p(\text{C}) = 0.128$, $p(\text{G}) = 0.134$, $p(\text{T}) = 0.403$, $1\text{st} = 0.860$, $2\text{nd} = 0.695$, $3\text{rd} = 1.445$. matK : $2.1\text{--}24.6 \times 10^{-10}$ subs site $^{-1}$ yr $^{-1}$; rbcL : $1.6\text{--}8.6 \times 10^{-10}$ subs site $^{-1}$ yr $^{-1}$.

subcontinents. Estimates were biased towards older ages given our interests in revealing a young age distribution. Old-age bias was introduced by (i) inadequate sampling of transcontinental crown clades, which is prone to overestimate that age of over-water or across-continent events (figure 6*b*), and (ii) a focus by all of the co-authors on selecting the potentially oldest continental and transcontinental crown clades (i.e. we have neglected all of the young transcontinental clades, including pantropical species, in our biased past efforts to uncover the effects of Early Tertiary history on the legume family). Thus, the resulting age distribution detected in this study should be viewed as strongly biased in the direction of old-age estimates.

3. RESULTS

(a) Continental crown clades

In spite of our attempts to find intracontinental clades of early Tertiary origin (figures 7–15), only one of the 33 crown clades showed a mean age of over 25 Myr, and this was the mirbelioid crown linking Australia and South Africa (figure 10) at *ca.* 48 Myr ago (table 1). The mean age distribution of these continental crown clades has a mode at the 4–6 Myr interval (figure 16), which is true regardless of the distribution of the crown clade in tropical (figure 17) or temperate (figure 18) biomes.

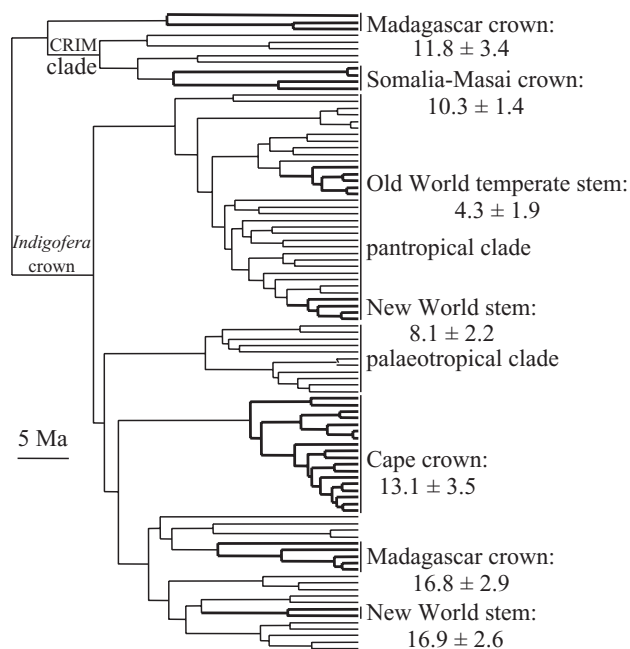


Figure 11. Phylogenetic structure on continents. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from the data and information provided in Schrire *et al.* (2003). The tribe Indigoferae is mainly centred in Africa but with distinct subcontinental and continental clades. Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 4.995$, $r(\text{CG}) = 0.779$, $r(\text{AT}) = 1.231$, $r(\text{AG}) = 3.382$, $r(\text{AC}) = 1.277$, $p(\text{A}) = 0.221$, $p(\text{C}) = 0.249$, $p(\text{G}) = 0.235$, $p(\text{T}) = 0.295$, $\alpha = 0.542$, $i\text{P} = 0.134$. $\text{ITS}/5.8\text{S}$: $2.1\text{--}4.7 \times 10^{-9}$ subs site $^{-1}$ yr $^{-1}$.

(b) Transcontinental crown clades

Transcontinental crown clades show an older mean age distribution, which is expected given that these represent in large part the corresponding stem clades of the crown clades mentioned above (table 2). The mode of these age estimates ranges from 6–16 Myr ago, and only eight of the 59 transcontinental crown clades have a mean age of over 25 Myr, or dating from the Palaeogene (figure 19). The mean age distribution is similar between tropical and temperate transcontinental crown clades (figures 20–21).

4. DISCUSSION

The modes of the estimated ages of the densely sampled intracontinental crown clades and those of the transcontinental crown clades differ on average by *ca.* 10 Myr (e.g. 6 Myr ago versus 16 Myr ago). Regardless, the predominant Neogene to Quaternary age distribution of both kinds of clades cannot be explained by Mid-Tertiary and older tectonic events, which are those implicated in the most recent major continental reconfigurations. Such events include the Cretaceous opening of the southern Atlantic (Parrish 1993), the last of the North Atlantic land bridges during the Mid-Tertiary (Tiffney 1985) and the Mid-Tertiary connection between North and South America, including the proto-Antilles (Iturralde-Vinent & MacPhee 1999). The generally young ages revealed for much of the geographical phylogenetic structure in

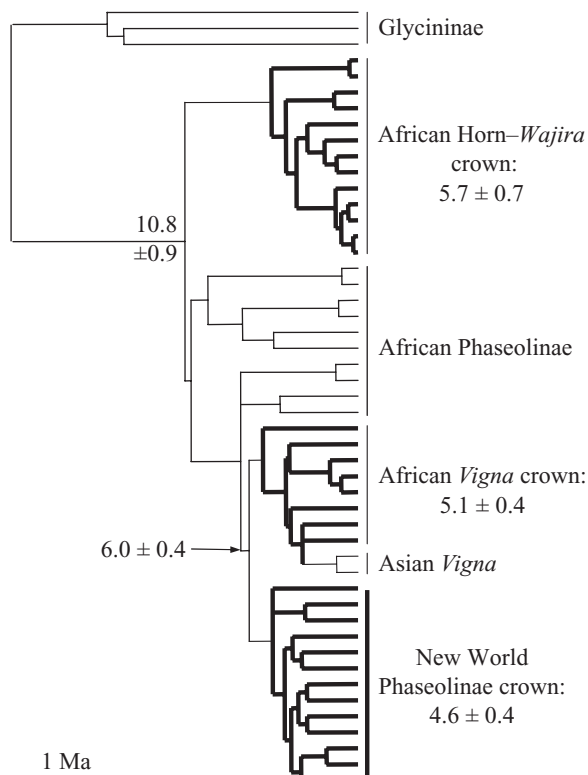


Figure 12. Phylogenetic structure on continents. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from the data and information provided in Thulin *et al.* (2004). The subtribe Phaseolinae of tribe Phaseoleae includes *Wajira* as sister to the rest. Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 0.984$, $r(\text{CG}) = 0.702$, $r(\text{AT}) = 0.274$, $r(\text{AG}) = 0.937$, $r(\text{AC}) = 1.563$, $p(\text{A}) = 0.326$, $p(\text{C}) = 0.135$, $p(\text{G}) = 0.156$, $p(\text{T}) = 0.384$, $\alpha = 0.888$, $iP = 0.192$, non-coding intron = 1.236, 1st = 0.687, 2nd = 0.689, 3rd = 1.111. cpDNA *matK*: $1.2\text{--}4.1 \times 10^{-9}$ subs site $^{-1}$ yr $^{-1}$.

legumes may be one reason why cladistic vicariance methods have not been forthcoming with robust area relationships for the Caribbean (e.g. Lavin *et al.* 2001a), South American dry forests (Pennington *et al.* 2004), and the Horn of Africa region (Lavin *et al.* 2000; Thulin *et al.* 2004).

Explanations of global legume distributions that resort to vicariance biogeography and tectonic history are considered mostly untenable in this analysis. An alternative explanation therefore must involve dispersal biology. In the realm of cladistic vicariance literature, dispersal hypotheses have been considered an anathema or prone to *ad hoc* explanations (e.g. Nelson & Platnick 1981). With the advent of the unified neutral theory of biodiversity and biogeography (Hubbell 2001), however, dispersal biology can be brought into the context of evolutionary time-scales and predictions can be made with respect to geographical phylogenetic structure.

Notably, most of the geographically structured phylogenies that we have subjected to rates analysis come from legume groups confined to tropical vegetation rich in succulent taxa (table 1), or the 'succulent biome' of Schrire *et al.* (2004). This biome is marked by erratic,

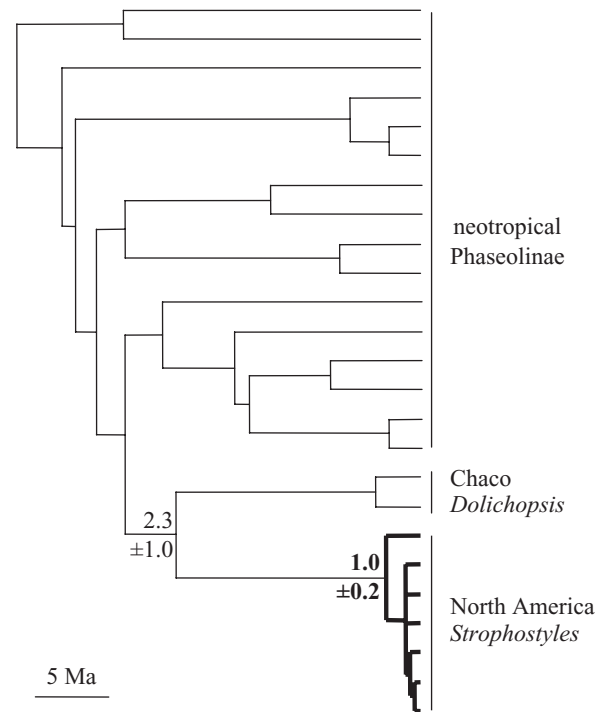


Figure 13. Phylogenetic structure on continents. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from the data and information provided in Riley-Hulting *et al.* (2004), which included a study of *Strophostyles* and related neotropical Phaseolinae. Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 5.064$, $r(\text{CG}) = 1.067$, $r(\text{AT}) = 0.947$, $r(\text{AG}) = 2.208$, $r(\text{AC}) = 0.812$, $p(\text{A}) = 0.224$, $p(\text{C}) = 0.276$, $p(\text{G}) = p(\text{T}) = 0.213$, $\alpha = 1.477$, $iP = 0.285$. cpDNA *matK*: $\pm 2.5 \times 10^{-9}$ subs site $^{-1}$ yr $^{-1}$.

unpredictable rains and does not experience natural burning (areas shaded red in figure 22 (Schrire *et al.* 2004)). Such vegetation comprises highly fragmented globally distributed patches represented, in part, by the Somalia-Masai region in and around the Horn of Africa, the Karoo-Namib of southwestern Africa (White 1983, p. 356; White & Léonard 1991), western Madagascar, the Caribbean basin, the Mexican selva baja (e.g. Rzedowski 1978) and the South American seasonally dry forests (Pennington *et al.* 2000), which includes the inter-Andean valleys, the Caatinga, Piedmont and Misiones regions (Prado & Gibbs 1993). The geographical phylogenetic structure detected in clades largely confined to these areas contrasts strongly to the absence of such structure detected in clades confined mostly to tropical savannah or wet forest vegetation (e.g. *Andira* (Pennington 1996) and *Inga* (Richardson *et al.* 2001)).

Dispersal among the patches of the succulent biome is limited, as inferred by few widespread species common among the patches of this biome. For example, few widespread legume species are shared between the Somalia-Masai and the Karoo-Namib regions of Africa, although sister species commonly occupy these two areas (Schrire *et al.* 2004). The few exceptions include *Acacia reficiens*, with subsp. *misera* in the Horn of Africa and subsp. *reficiens* from Angola and Namibia. Clades confined largely to the succulent biome in the neotropics,

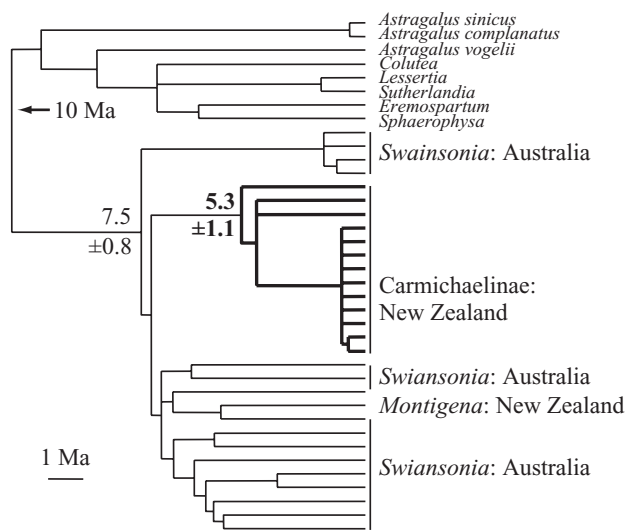


Figure 14. Phylogenetic structure on continents. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from the data and information provided in Wagstaff *et al.* (1999). Carmichaelinae are mainly of Australia but with two independent lines in New Zealand. Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 2.910$, $r(\text{CG}) = 0.737$, $r(\text{AT}) = 1.229$, $r(\text{AG}) = 2.463$, $r(\text{AC}) = 0.742$, $p(\text{A}) = 0.220$, $p(\text{C}) = 0.257$, $p(\text{G}) = 0.271$, $p(\text{T}) = 0.253$, $\alpha = 1.147$, $iP = 0.225$. nrDNA ITS/5.8S: $3.0\text{--}4.1 \times 10^{-9}$ subs site $^{-1}$ yr $^{-1}$.

(e.g. robinoid legumes) mostly contain narrow endemics and no very widespread species (e.g. Prado & Gibbs 1993) that match the natural distributions of many widespread common species found in neotropical rainforests and

savannahs, such as *Symphonia globulifera* (Clusiaceae; Dick *et al.* 2003) or *Qualea grandiflora* Mart. (Vochysiaceae; Ratter *et al.* 2003). Indeed, amphi-Atlantic and pantropical legume species are commonly found in both the rainforest and savannah (e.g. *Lonchocarpus sericeus*, *Andira inermis*, *Machaerium lunatum* and *Vigna juruana*); there are no transcontinental species confined to the succulent biome.

The relationship between a paucity of widespread species in the succulent biome and the high levels of geographical phylogenetic structure detected in the taxa confined to this vegetation can be explained in the meta-community terms of the unified neutral theory of biodiversity and biogeography (Hubbell 2001). Geographical phylogenetic structure must be ultimately caused by demographic stochasticity playing out over evolutionary time scales (i.e. ecological drift) on resident lineages, such that endemic clades of species come to occupy a distinct geographical setting. Immigration rates must be high enough to maintain or enhance resident diversity, but low enough not to replace it. Just as high levels of gene flow result in the replacement of resident alleles unique to a population and thus reduce genetic population structure, high immigration rates result in the replacement of resident species and thus reduce geographical phylogenetic structure (figure 23).

If the restricted area and highly fragmented distribution of the succulent biome has reduced evolutionary rates of immigration among the local regions of this global biome, then the following predictions can be made. The flora of the succulent biome should show little community phylogenetic structure (sensu Webb 2000). This is because reduced rates of immigration result in little

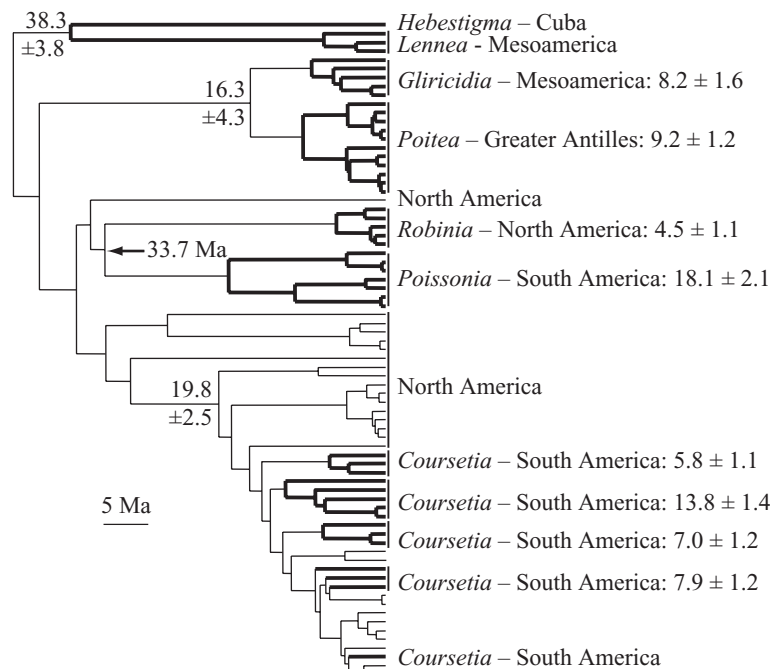


Figure 15. Phylogenetic structure on continents. This penalized likelihood rate-smoothed Bayesian consensus phylogeny and the estimated ages were derived from the data and information provided in Lavin *et al.* (2003). Robinoid legumes are centred mainly in North America but with two clades in the Greater Antilles, and several clades of South American *Coursetia*. Average nucleotide substitution parameters estimated for likelihood trees at stationarity are $r(\text{GT}) = 1.000$, $r(\text{CT}) = 6.274$, $r(\text{CG}) = 0.976$, $r(\text{AT}) = 1.792$, $r(\text{AG}) = 3.170$, $r(\text{AC}) = 1.159$, $p(\text{A}) = 0.200$, $p(\text{C}) = 0.273$, $p(\text{G}) = 0.290$, $p(\text{T}) = 0.236$, $\alpha = 1.383$, $iP = 0.217$. ITS/5.8S: $3.1\text{--}3.5 \times 10^{-9}$ subs site $^{-1}$ yr $^{-1}$; *matK*: 3.9×10^{-10} subs site $^{-1}$ yr $^{-1}$.

Table 1. Estimated ages of intracontinental crown clades having sister clades from different continents and for which taxon sampling has been exhaustive, or nearly so. Estimated mean ages and s.d., in millions of years, were derived from penalized likelihood rate smoothing of 100 Bayesian trees sampled widely at stationarity.

crown clade	area (biome; Schrire <i>et al.</i> 2004)	age (Myr)	s.d.	reference
<i>Brya</i>	Greater Antilles (succulent)	4.3	2.2	A. Beyra-Matos and M. Lavin, unpublished data
Carmichaelinae	New Zealand (temperate)	5.3	1.1	Wagstaff <i>et al.</i> (1999); this study
<i>Chapmannia</i> (Old World)	Somalia-Masai (succulent)	10.0	2.6	Lavin <i>et al.</i> (2000); this study
<i>Chapmannia</i> (New World)	Mesoamerica (succulent and temperate)	9.8	2.2	Lavin <i>et al.</i> (2000); this study
<i>Coursetia brachy-hachis</i> group	South America (succulent)	13.8	1.4	Lavin <i>et al.</i> (2003)
<i>Coursetia grandiflora</i> group	South America (succulent)	7.0	1.2	Lavin <i>et al.</i> (2003)
<i>Coursetia rostrata</i> group	South America (succulent)	5.8	1.1	Lavin <i>et al.</i> (2003)
<i>Cranocarpus</i>	South America (succulent)	5.0	2.0	A. Beyra-Matos and M. Lavin, unpublished data
<i>Desmanthus</i>	Mesoamerica (succulent, grass and temperate)	12.7	1.7	this study
<i>Diphysa</i>	Mesoamerica (succulent)	4.5	1.2	Lavin <i>et al.</i> (2000); this study
<i>Gagnebina-Alanisiiodendron</i>	Madagascar region (succulent)	7.3	1.1	this study
<i>Giricidia</i>	Mesoamerica (succulent)	8.2	1.6	Lavin <i>et al.</i> (2003)
<i>Indigofera</i>	African Cape (temperate)	13.1	3.5	Schrire <i>et al.</i> (2003)
<i>Lennea</i>	Mesoamerica (rainforest and succulent)	6.1	0.8	Lavin <i>et al.</i> (2003)
<i>Leucaena</i>	Mesoamerica (succulent)	10.1	1.3	this study
<i>Microcharis</i>	Somalia-Masai (succulent)	10.3	1.4	B. D. Schrire, M. Lavin and N. P. Barker, unpublished data
mirbelioids	Australia (temperate)	48.4	1.3	this study
<i>Neptunia</i>	panropical (rainforest and grass)	3.6	0.7	this study
New World Phaseolinae	New World (succulent and grass)	4.6	0.4	A. Delgado-Salinas, M. Lavin, M. Thulin, R. Pasquet and N. F. Weeden, unpublished data
<i>Nissolia</i>	Mesoamerica (succulent)	1.6	0.4	Pennington <i>et al.</i> (2004)
<i>Ormocarpopsis</i>	Madagascar (succulent)	4.2	1.1	Lavin <i>et al.</i> (2000); this study
<i>Ormocarpum</i>	Africa (rainforest)	8.2	1.2	Lavin <i>et al.</i> (2000); this study
<i>Phylloxylon</i>	Madagascar (rainforest and grass)	11.8	3.4	Schrire <i>et al.</i> (2003)
<i>Pictetia</i>	Greater Antilles (succulent)	7.2	1.2	Lavin <i>et al.</i> (2000); this study
<i>Poissonia</i>	South America (succulent)	18.1	2.2	Lavin <i>et al.</i> (2003)
<i>Poitea</i>	Greater Antilles (succulent)	9.2	1.2	Lavin <i>et al.</i> (2003)
<i>Robinia</i>	Appalachia (temperate)	4.5	1.1	Lavin <i>et al.</i> (2003); this study
<i>Sarphostyles</i>	Appalachia (temperate)	1.0	0.2	Riley-Hulting <i>et al.</i> (2004)
<i>Vaughania</i>	Madagascar (succulent)	16.8	2.9	Schrire <i>et al.</i> (2003)
<i>Vigna</i> s. s.	Africa (succulent and grass)	5.1	0.4	A. Delgado-Salinas, M. Lavin, M. Thulin, R. Pasquet and N. F. Weeden, unpublished data
<i>Vigna</i> subgen. <i>Ceratotropis</i>	Asia (grass)	2.0	0.2	A. Delgado-Salinas, M. Lavin, M. Thulin, R. Pasquet and N. F. Weeden, unpublished data
<i>Wajina</i>	Somalia-Masai (succulent)	5.7	0.7	Thulin <i>et al.</i> (2004)
<i>Zygocarpum</i>	Somalia-Masai (succulent)	7.2	1.3	Lavin <i>et al.</i> (2000); this study

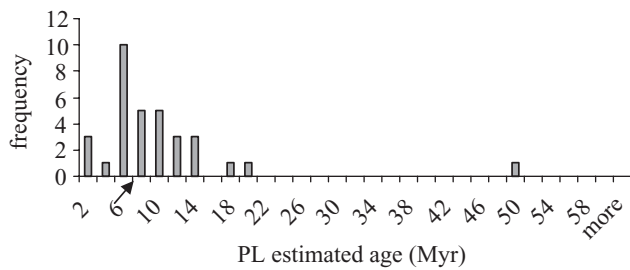


Figure 16. Age distribution (in Myr) of 33 continental crown clades listed in table 1. The arrow indicates the division between the 4–6 Myr and 6–8 Myr bins.

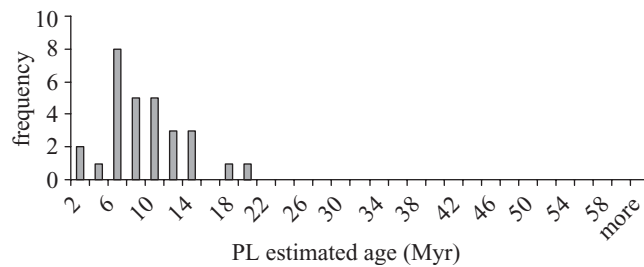


Figure 17. Age distribution (in Myr) of 29 continental crown clades listed in table 1.

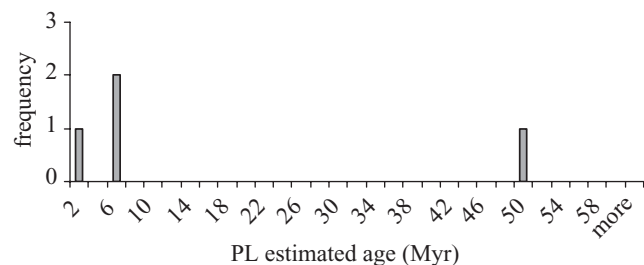


Figure 18. Age distribution (in Myr) of four continental crown clades listed in table 1. Although only four temperate continental crown clades were sampled. The distribution is suggestive of the tropical one.

sympatry and thus ecological filtering of closely related species. Furthermore, relative species abundance curves should be detected with a deficit of rare species and perhaps also a deficit of widespread common species. This would contrast to expected or perhaps elevated levels of middling abundance species. This is because endemic speciation and ecological drift would result in an abundance of locally common endemic species. A deficit in the rare tail of the relative species abundance curve would be the result of mostly resident speciation adding to standing diversity, not much augmented by immigration (e.g. compare figs 9.8 and 9.9 in ch. 9 of Hubbell (2001)). A deficit of widespread abundant species within the succulent biome would simply be a result of the relatively small area and highly fragmented distribution of this biome, which limits migration among local communities. An agreement between high levels of geographical phylogenetic structure, low levels of community phylogenetic structure, and relative species abundance curves with deficits in the widespread common or rare tails would provide strong evidence for the persistence of the succulent biome

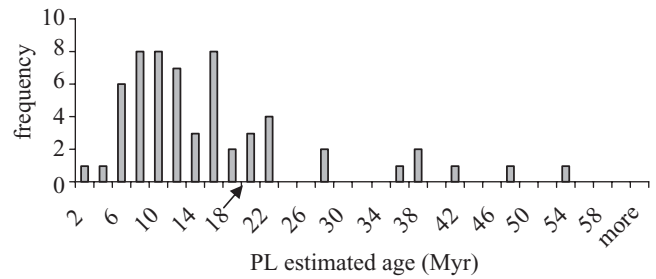


Figure 19. Age distribution (in Myr) of 59 transcontinental crown clades listed in table 2. The arrow indicates the division between the 16–18 Myr and 18–20 Myr bins.

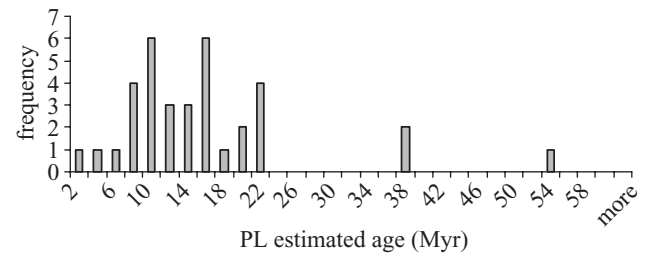


Figure 20. Age distribution (in Myr) of 35 transcontinental crown clades listed in table 2.

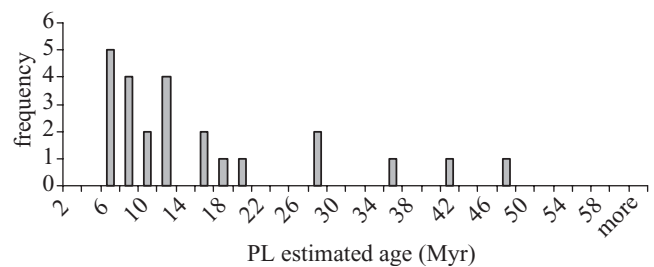


Figure 21. Age distribution (in Myr) of 24 transcontinental crown clades listed in table 2.

in the face of low levels of immigration over evolutionary time-scales.

Other predictions also follow. Rainforest and savannah biomes, where immigration rates appear to be much higher than in the succulent biome, should be occupied by resident lineages that show low levels of geographical phylogenetic structure, high levels of community phylogenetic structure and relative species abundance curves with an excess of widespread common species or a long tail at the rare end. For example, the genus *Inga* predominates almost entirely in neotropical wet forests and shows little evidence of geographical phylogenetic structure (Richardson *et al.* 2001) and potentially high levels of community phylogenetic structure (i.e. possibly up to 20 species occur in sympatry in the Peruvian Amazon; Reynel & Pennington 1997). Clades of species endemic to the succulent biome have essentially allopatric distributions that contrast with *Inga* and other wet forest taxa. For example, it is very rare to find two species growing naturally side by side that both belong to the same genus confined to the succulent biome (e.g. *Caesalpinia*, *Coursetia*, *Leucaena* or *Mimosa*). The genus *Mimosa* shows allopatric species within the succulent

Table 2. Estimated ages of transcontinental crown clades for which taxon sampling has not been exhaustive. The transcontinental crowns are defined as the most recent common ancestor (MRCA) of the two species listed, which were sampled from the *matK* data of Lavin *et al.* (2004), except where noted. Estimated mean ages and s.d., in millions of years, were derived from penalized likelihood rate smoothing of 100 Bayesian trees sampled widely at stationarity.

MRCA of:	age (Myr)	s.d.	reference
<i>Aeschynomene americana</i> and <i>Aeschynomene rudis</i>	36.7	2.1	this study
<i>Arachis pintoii</i> and <i>Stylosanthes hamata</i>	13.8	1.7	this study
<i>Arcoa gonavensis</i> and <i>Tetrapterocarpon geayi</i>	54.0	3.4	this study
<i>Astragalus canadensis</i> and <i>Astragalus nothoxys</i>	9.3	1.6	this study
<i>Baphia massaiensis</i> and <i>Baphia madagascariensis</i>	21.5	2.7	this study
<i>Bobgunnia madagascariensis</i> and <i>Swartzia simplex</i>	36.6	4.6	this study
<i>Brongniartia alamosana</i> and <i>Hovea purpurea</i>	9.9	1.8	this study
<i>Brya ebenus</i> and <i>Cranocarpus martii</i>	8.1	1.5	this study
<i>Calliandra californica</i> and <i>Calliandra surinamensis</i>	6.9	1.4	this study
<i>Calliandropsis nervosus</i> and <i>Gagnebina pervilleana</i>	7.0	1.5	this study
<i>Cercis occidentalis</i> and <i>Cercis gigantea</i>	5.1	2.2	this study
<i>Chapmannia floridana</i> and <i>Chapmannia sericea</i>	14.2	1.7	this study
<i>Cladrastis delavayi</i> and <i>Cladrastis lutea</i>	15.8	6.5	this study
<i>Cologania hintoniorum</i> and <i>Psoralea cinerea</i>	19.1	1.5	this study
<i>Dalbergia congestiflora</i> and <i>Dalbergia sissoo</i>	9.6	2.6	this study
<i>Desmodium psilocarpum</i> and <i>Campylotropis macrocarpa</i>	14.2	1.6	this study
<i>Dichrostachys cinerea</i> and <i>Calliandropsis nervosus</i>	21.0	4.1	this study
<i>Dichrostachys cinerea</i> and <i>Dichrostachys spicata</i>	0.7	0.3	this study
<i>Diphysa floribunda</i> and <i>Pictetia aculeata</i>	15.0	2.5	this study
<i>Gleditsia triacanthos</i> and <i>Gleditsia sinensis</i>	10.1	3.4	this study
<i>Glycyrrhiza lepidota</i> and <i>Callerya atropurpurea</i>	35.6	3.1	this study
<i>Gymnocladus chinensis</i> and <i>Gleditsia sinensis</i>	26.9	4.3	this study
<i>Indigofera coerulea</i> and <i>Indigofera suffruticosa</i>	16.9	2.6	Schrire <i>et al.</i> (2003)
<i>Indigofera jamaicensis</i> and <i>Indigofera spinosa</i>	13.5	2.6	B. D. Schrire, M. Lavin and N. P. Barker, unpublished data
<i>Indigofera kirilowii</i> and <i>Indigofera koreana</i>	4.3	1.9	Schrire <i>et al.</i> (2003)
<i>Indigofera lepezioides</i> and <i>Indigofera dalzellii</i>	8.1	2.2	Schrire <i>et al.</i> (2003)
<i>Inga punctata</i> and <i>Albizia julibrissin</i>	14.6	3.8	this study
<i>Inocarpus fagifer</i> and <i>Tipuana tipu</i>	20.3	3.6	this study
<i>Kanaloa kahoolawensis</i> and <i>Desmanthus balsensis</i>	14.7	1.7	this study
<i>Lathyrus sativus</i> and <i>Lathyrus latifolius</i>	6.3	1.3	this study
<i>Lonchocarpus lanceolatus</i> and <i>Paraderris elliptica</i>	10.4	1.4	this study
<i>Lotus purshianus</i> and <i>Ornithopus compressus</i>	11.6	1.2	this study
<i>Lupinus argenteus</i> and <i>Anarthrophyllum desideratum</i>	19.2	2.5	this study
<i>Lupinus consentii</i> and <i>Lupinus argenteus</i>	6.5	1.4	this study
<i>Nissolia hirsuta</i> and <i>Chaetocalyx scandens</i>	8.5	1.4	this study
<i>Onobrychis montana</i> and <i>Hedysarum boreale</i>	14.8	2.1	this study
<i>Ormocarpopsis calcicola</i> and <i>Ormocarpum kirkii</i>	7.9	1.7	this study
<i>Ormosia colombiana</i> and <i>Ormosia formosana</i>	20.7	4.3	this study
<i>Oxytropis deflexa</i> and <i>Oxytropis lambertii</i>	4.4	1.4	this study
<i>Pictetia marginata</i> and <i>Ormocarpum kirkii</i>	14.5	2.6	this study
<i>Piptanthus nepalensis</i> and <i>Thermopsis rhombifolia</i>	26.5	3.4	this study
<i>Piscidia piscipula</i> and <i>Paraderris elliptica</i>	12.5	1.6	this study
<i>Psoralea cinerea</i> and <i>Rupertia physodes</i>	6.3	0.9	this study
<i>Schleinitzia insularum</i> and <i>Desmanthus balsensis</i>	11.9	2.2	this study
<i>Sesbania emerus</i> and <i>Sesbania grandiflora</i>	4.4	0.9	this study
<i>Sesbania emerus</i> and <i>Sesbania vesicaria</i>	18.9	2.1	this study
<i>Sophora davidii</i> and <i>Sophora nuttalliana</i>	4.5	1.4	this study
<i>Styphnolobium japonicum</i> and <i>Cladrastis platycarpa</i>	47.4	2.6	this study
<i>Styphnolobium japonicum</i> and <i>Pickeringia montana</i>	40.1	0.4	this study
<i>Sutherlandia frutescens</i> and <i>Carmichaelia williamsii</i>	10.4	2.0	this study
<i>Swainsona pterostylis</i> and <i>Glianthus puniceus</i>	5.9	1.5	this study
<i>Tephrosia heckmanniana</i> and <i>Mundulea sericea</i>	7.5	1.2	this study
<i>Trifolium beckwithii</i> and <i>Trifolium repens</i>	7.3	1.5	this study
<i>Vicia ludoviciana</i> and <i>Vicia hirsuta</i>	16.5	2.2	this study
<i>Vicia sativa</i> and <i>Vicia americana</i>	11.8	1.7	this study
<i>Vigna subterranea</i> and <i>Phaseolus vulgaris</i>	8.0	0.8	this study
<i>Wajira albescens</i> and <i>Phaseolus vulgaris</i>	10.7	0.9	this study
<i>Wisteria sinensis</i> and <i>Wisteria frutescens</i>	8.2	3.0	this study
<i>Zygocarpum gillettii</i> and <i>Zygocarpum yemenense</i>	2.9	0.7	this study

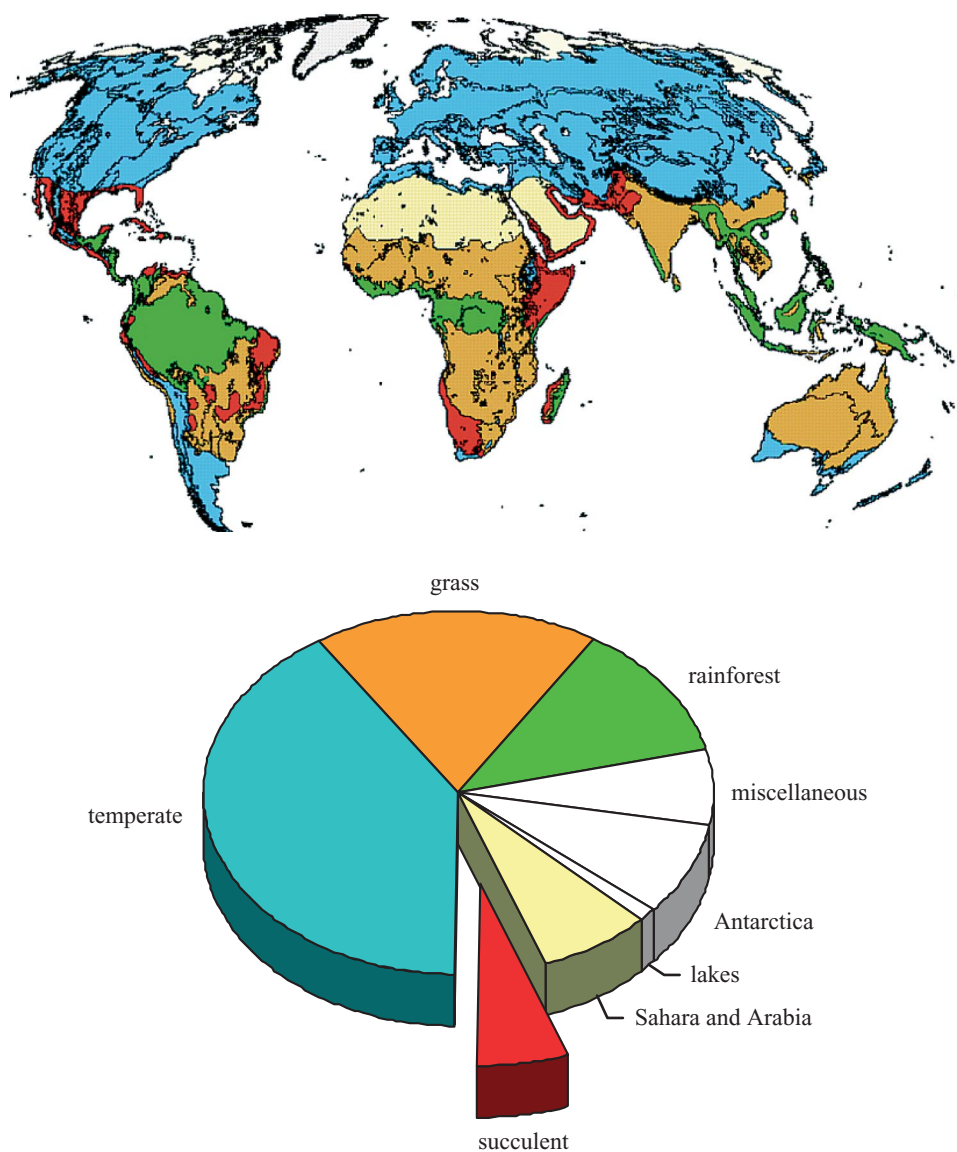


Figure 22. Distributions of the four major biomes detected in Schrire *et al.* (2004) and the relative area occupied by each. The red area indicates the small fragmented succulent biome that is rich in endemic legume diversity and harbours most of the legume groups for which continental phylogenetic structure is detected.

biome and sympatric species within the cerrado (C. E. Hughes, unpublished data). A further prediction here is that allopatric speciation in the succulent biome results in fewer barriers to gene flow other than geographical; interspecific crossability should be common among sister species endemic to the succulent biome, as in *Leucaena* (C. E. Hughes, unpublished data).

In our effort to discover the causes of geographical structure, we exposed a pattern that led us to seriously consider metacommunity process rather than vicariance. In spite of our very strong bias in sampling what we thought were old clades in legumes (i.e. by focusing on transcontinental vicariant groups), we discovered young ages of the geographical phylogenetic structure in legumes. Indeed, we propose that the geographical phylogenetic structure observed for clades endemic to oceanic islands (e.g. Baldwin & Sanderson 1998; figures 2 and 3) is not fundamentally different from such structure observed for continental islands (figures 4 and 5) and for entire continents (figures 7–15). That is, the relative isolation of islands

allows the winners of sweepstake dispersal to diversify unperturbed, protected by distance or a water barrier from high rates of immigration. For oceanic and continental islands, this process commonly results in taxonomically unbalanced biotas. For continental regions, especially with respect to the succulent biome (Schrire *et al.* 2004), low immigration rates similarly give rise to phylogenetic structure over evolutionary time-scales. Even between North and South America, transcontinental sister clades commonly pre-date the formation of the Panamanian Isthmus and postdate Gaarlandia (Engel *et al.* 1998; Pennington *et al.* 2004), so any geographical phylogenetic structure detected among New World continental clades must similarly be the result of low rates of immigration over evolutionary time-scales.

5. SUMMARY

The phylogenetic approach taken in this study allows us to identify and define global metacommunities and the extent of the constituent local communities. This is particularly the case for the succulent biome and its constituent local

- | | |
|--|--|
| (i) deficit in relative abundances of widespread common and rare species | (i) expected relative abundances of widespread common and rare species |
| (ii) geographical phylogenetic structure | (ii) little geographical phylogenetic structure |
| (iii) little community phylogenetic structure | (iii) community phylogenetic structure |

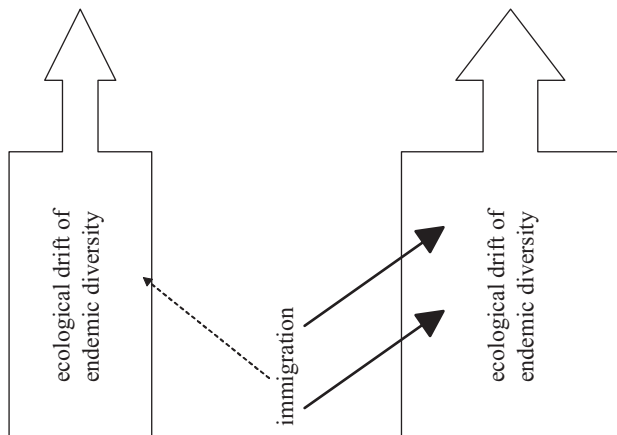


Figure 23. As ecological drift plays out on resident and immigrant diversity within the local communities of the global metacommunity, phylogenetic structure can arise if immigration rates are minimal. The left-hand box is reduced in size relative to the right-hand box, reflecting the small and highly fragmented distribution of the succulent biome. Immigration within and into this biome is thus reduced, leaving largely resident diversity in the local communities to the processes of ecological drift and speciation. Predictions pertaining to phylogenetic structure under such conditions contrast to those in global metacommunities where immigration rates are higher (e.g. the rainforest, grass and temperate biomes). Geographical phylogenetic structure results when resident species and species clades are isolated owing to low levels of immigration. Community phylogenetic structure (Webb 2000) results when immigration rates are high enough to create an abundance of sympatry of related species.

communities including the Somalia-Masai and Karoo-Namib regions of Africa, and the Mexican selva baja and the fragmented South American seasonally dry forests (Pennington *et al.* 2000, 2004). These globally distributed local regions have only recently been viewed under a single biome concept (Schrire *et al.* 2004) and the phylogenetic studies presented here support this view. When combined with neutral ecological theory (Hubbell 2001), phylogenetic analysis should be able to more accurately characterize and quantify the individual contributions of immigrant and resident lineages to standing diversity, and lead to an understanding of what balance leads to the high levels of geographical phylogenetic structure observed for island lineages or the endemic lineages inhabiting the succulent biome.

Justin Moat from the GIS unit, Royal Botanic Gardens, Kew, provided the estimates of global coverage for each of the biomes used in this study. Campbell Webb and an anonymous reviewer provided insightful comments that greatly improved the manuscript. This work was supported largely by the US National Science

Foundation (DEB-0075202) and the Leverhulme Trust (grant F/771/A to the Royal Botanic Garden Edinburgh (RBGE)).

REFERENCES

- Bailey, C. D., Carr, T. G., Harris, S. A. & Hughes, C. E. 2003 Characterisation of angiosperm nrDNA polymorphism, paralogy and pseudogenes. *Mol. Phylogenet. Evol.* **29**, 435–455.
- Baldwin, B. & Sanderson, M. J. 1998 Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proc. Natl Acad. Sci. USA* **95**, 9402–9406.
- Beyra-Matos, A. & Lavin, M. 1999 Monograph of *Pictetia* (Leguminosae: Papilionoideae) and review of the Aeschynomeneae. *Syst. Bot. Monogr.* **56**, 1–3.
- Buskirk, R. E. 1985 Zoogeographic patterns and tectonic history of Jamaica and the northern Caribbean. *J. Biogeogr.* **12**, 445–461.
- Cibois, A., Kalyakin, M. V., Lian-Xian, H. & Pasquet, E. 2002 Molecular phylogenetics of babblers (Timaliidae): reevaluation of the genera *Yuhina* and *Stachyris*. *J. Avian Biol.* **33**, 380–390.
- Delgado-Salinas, A., Bibler, R. & Lavin, M. 2004 Molecular phylogeny of *Phaseolus*. *A. Rep. Bean Improv. Coop.* **47**, 31–32.
- Dick, C. W., Abdul-Salim, K. & Bermingham, E. 2003 Molecular systematic analysis reveals cryptic Tertiary diversification of a widespread tropical rainforest tree. *Am. Nat.* **162**, 691–703.
- Donoghue, M. J., Bell, C. D. & Li, J. 2001 Phylogenetic patterns in Northern Hemisphere plant geography. *Int. J. Pl. Sci.* **162**(Suppl.), S41–S52.
- Engel, S. R., Hogan, K. M., Taylor, J. F. & Davis, S. K. 1998 Molecular systematics and paleobiogeography of the South American sigmodontine rodents. *Mol. Biol. Evol.* **15**, 35–49.
- Goodman, S. M. & Benstead, J. P. (eds) 2004 *The natural history of Madagascar*. The University of Chicago Press.
- Herendeen, P. S., Lewis, G. P. & Bruneau, A. 2003 Floral morphology in caesalpinoid legumes: testing the monophyly of the 'Umtiza clade'. *Int. J. Pl. Sci.* **164** (Suppl.), S393–S407.
- Hubbell, S. P. 2001 *The unified neutral theory of biodiversity and biogeography*. Princeton University Press.
- Huelsenbeck, J. P. & Ronquist, F. R. 2001 MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17**, 754–755.
- Huelsenbeck, J. P., Ronquist, R., Nielson, R. & Bollback, J. P. 2001 Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294**, 2310–2314.
- Hughes, C. E., Bailey, C. D., Krosnick, S. & Luckow, M. A. 2003 Relationships among genera of the informal *Dichrostachys* and *Leucaena* groups (Mimosoideae) inferred from nuclear ribosomal ITS sequences. In *Advances in legume systematics*, part 10. *Higher level systematics* (ed. B. B. Klitgaard & A. Bruneau), pp. 221–238. Kew, UK: Royal Botanic Gardens.
- Irwin, D. E. 2002 Phylogeographic breaks without geographic barriers to gene flow. *Evolution* **56**, 2383–2394.
- Iturralde-Vinent, M. A. & MacPhee, R. D. E. 1999 Paleogeography of the Caribbean region: implications for Cenozoic biogeography. *Bull. Am. Mus. Nat. Hist.* **238**, 1–95.
- Johnson, J. B. & Omland, K. S. 2004 Model selection in ecology and evolution. *Trends Ecol. Evol.* **19**, 101–108.
- 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., Thulin, M., Labat, J.-N. & Pennington, T. 2000 Africa, the odd man out: molecular biogeographic studies of dalbergioid legumes (Fabaceae) suggest otherwise. *Syst. Bot.* **25**, 449–467.
- Lavin, M., Wojciechowski, M. F., Richman, A., Rotella, J., Sanderson, M. J. & Beyra-Matos, A. 2001a Identifying Tertiary radiations of Fabaceae in the Greater Antilles: alternatives to cladistic vicariance analysis. *Int. J. Pl. Sci.* **162**(Suppl.), S53–S76.

- Lavin, M., Pennington, R. T., Klitgaard, B. B., Sprent, J. I., De Lima, H. C. & Gasson, P. 2001b The dalbergioid legumes (Fabaceae): delimitation of a pantropical monophyletic clade. *Am. J. Bot.* **88**, 503–533.
- Lavin, M., Wojciechowski, M. F., Gasson, P., Hughes, C. H. & 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.
- Lavin, M., Wojciechowski, M. F. & Herendeen, P. H. 2005 Evolutionary rates analysis of Leguminosae implicates a rapid diversification of the major family lineages immediately following an Early Tertiary emergence. *Syst. Biol.* (In the press.)
- Losos, J. B. & Schluter, D. 2000 Analysis of an evolutionary species-area relationship. *Nature* **408**, 847–850.
- Moore, J. G. & Clague, D. A. 1992 Volcano growth and evolution of the island of Hawaii. *Geol. Soc. Am. Bull.* **104**, 1471–1484.
- Morley, R. J. & Dick, C. W. 2003 Missing fossils, molecular clocks and the origin of the Melastomataceae. *Am. J. Bot.* **90**, 1638–1645.
- Nelson, G. & Platnick, N. 1981 *Systematics and biogeography: cladistics and vicariance*. New York: Columbia University Press.
- Parrish, J. T. 1993 The palaeogeography of the opening South Atlantic. In *The Africa–South America connection* (ed. W. George & R. Lavocat), pp. 8–27. Oxford: Clarendon.
- Pennington, R. T. 1996 Molecular and morphological data provide phylogenetic resolution at different hierarchical levels in *Andira*. *Syst. Biol.* **45**, 496–515.
- Pennington, R. T., Prado, D. E. & Pendry, C. A. 2000 Neotropical seasonally dry forests and Quaternary vegetation changes. *J. Biogeogr.* **27**, 261–273.
- Pennington, R. T., Lavin, M., Prado, D. E., Pendry, C. A., Pell, S. & Butterworth, C. 2004 Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Phil. Trans. R. Soc. Lond. B* **359**, 515–538. (doi:10.1098/rstb.2003.1435)
- 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.
- Ratter, J. A., Bridgewater, S. & Ribeiro, J. F. 2003 Analysis of the floristic composition of the Brazilian Cerrado vegetation. III. Comparison of the woody vegetation of 376 areas. *Edinburgh J. Bot.* **60**, 57–109.
- Renner, S. S. 2004 Multiple miocene melastomataceae dispersal between Madagascar, Africa, and India. *Phil. Trans. R. Soc. Lond. B* **359**. (In the press.) (doi: 10.1098/rstb.2004.1530)
- Reynel, C. & Pennington, T. D. 1997 *El genero Inga en el Peru. Morfologia, Distribucion y Usos*. Kew, UK: Royal Botanic Gardens.
- 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.
- Riley-Hulting, E. T., Delgado-Salinas, A. & Lavin, M. 2004 Phylogenetic systematics of *Strophostyles* (Fabaceae): a North American temperate genus within a neotropical diversification. *Syst. Bot.* **29**, 627–653.
- Rzedowski, J. 1978 *Vegetación de México*. Mexico City: Editorial Limusa.
- 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. 2002 Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* **19**, 101–109.
- Sanderson, M. J. 2003 *R8s, v. 1.6, user's manual (April 2003)*. Distributed by the author (<http://ginger.ucdavis.edu/r8s/>). Davis, CA: University of California.
- Sanmartín, I., Enghoff, H. & Ronquist, F. 2001 Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biol. J. Linn. Soc.* **73**, 345–390.
- Sato, A., Tichy, H., O'hUigin, C., Grant, P. R., Grant, B. R. & Klein, J. 2001 On the origin of Darwin's finches. *Mol. Biol. Evol.* **18**, 299–311.
- Schubert, C. D., Diesel, R. & Hedges, S. B. 1998 Rapid evolution to terrestrial life in Jamaican crabs. *Nature* **393**, 363–365.
- Schrire, B. D., Lavin, M., Barker, N. P., Cortes-Burns, H., von Senger, I. & Kim, J.-H. 2003 Towards a phylogeny of *Indigofera* (Leguminosae: Papilionoideae): identification of major clades and relative ages. In *Advances in legume systematics*, part 10, *Higher level systematics* (ed. B. B. Klitgaard & A. Bruneau), pp. 269–302. Kew, UK: Royal Botanic Gardens.
- Schrire, B. D., Lavin, M. & Lewis, G. P. 2004 Global distribution patterns of the Leguminosae: insights from recent phylogenies. *Biologiske Skrifter*. (In the press.)
- Simpson, G. G. 1940 Mammals and land bridges: address delivered before the Washington Academy of Sciences, February 15, 1940. *J. Washington Acad. Sci.* **30**, 137–163.
- Swofford, D. 2001. *PAUP*, phylogenetic analysis using parsimony (* and other methods)*, v. 4.0β10. Sunderland, MA: Sinauer.
- Thulin, M. & Lavin, M. 2001 Phylogeny and biogeography of the *Ormocarpum* group (Fabaceae): a new genus *Zygocarpum* from the Horn of Africa region. *Syst. Bot.* **26**, 299–317.
- Thulin, M., Lavin, M., Pasquet, R. & Delgado-Salinas, A. 2004 Phylogeny and biogeography of *Wajira* (Leguminosae): a monophyletic segregate of *Vigna* centered in the Horn of Africa region. *Syst. Bot.* **29**. (In the press.)
- Tiffney, B. 1985 The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. *J. Arnold Arboretum, Harvard Univ.* **66**, 243–273.
- Wagstaff, S. J., Heenan, P. B. & Sanderson, M. J. 1999 Classification, origins, and patterns of diversification in New Zealand Carmichaelinae (Fabaceae). *Am. J. Bot.* **86**, 1346–1356.
- Webb, C. O. 2000 Exploring the phylogenetic structure of ecological communities: an example for rainforest trees. *Am. Nat.* **156**, 145–155.
- Werner, R., Hoernle, K., Van den Bogaard, P., Ranero, C. & von Huene, R. 1999 Drowned 14-m.y.-old Galapagos Archipelago off the coast of Costa Rica: implications for tectonic and evolutionary models. *Geology* **27**, 499–502.
- White, F. 1983 *The vegetation of Africa, a descriptive memoir to accompany the UNESCO/AEFTAT/UNSO vegetation map of Africa*. Paris: UNESCO.
- White, F. & Léonard, J. 1991 Phytogeographical links between Africa and Southwest Asia. *Flora et Vegetatio Mundi* **9**, 229–246.
- Wojciechowski, M. F., Lavin, M. & Sanderson, M. J. 2004 Plastid *matK* phylogeny resolves many well-supported subclades within the legume family (Leguminosae). *Am. J. Bot.* **91**. (In the press.)
- Yoder, A. D., Burns, M. M., Zehr, S., Delefosse, T., Veron, G., Goodman, S. M. & Flynn, J. J. 2003 Single origin of Malagasy Carnivora from an African ancestor. *Nature* **421**, 734–737.