New Phytologist

Review

Tansley review

The contrasting nature of woody plant species in different neotropical forest biomes reflects differences in ecological stability

Author for correspondence:
R. Toby Pennington
Tel: +44 (0)131 248 2818
Email: t.pennington@rbge.org.uk

Received: *28 May 2015*
Accepted: *7 September 2015*

R. Toby Pennington[1] **and Matt Lavin**[2]

[1]Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh, EH3 5LR, UK; [2]Department of Plant Sciences & Plant Pathology, Montana State University, PO Box 173150, Bozeman, MT 59717-3150, USA

Contents

New Phytologist (2015)
doi: 10.1111/nph.13724

Key words: Amazonia, dispersal, rain forest, seasonally dry tropical forest, speciation, species.

Summary

A fundamental premise of this review is that distinctive phylogenetic and biogeographic patterns in clades endemic to different major biomes illuminate the evolutionary process. In seasonally dry tropical forests (SDTFs), phylogenies are geographically structured and multiple individuals representing single species coalesce. This pattern of monophyletic species, coupled with their old species stem ages, is indicative of maintenance of small effective population sizes over evolutionary timescales, which suggests that SDTF is difficult to immigrate into because of persistent resident lineages adapted to a stable, seasonally dry ecology. By contrast, lack of coalescence in conspecific accessions of abundant and often widespread species is more frequent in rain forests and is likely to reflect large effective population sizes maintained over huge areas by effective seed and pollen flow. Species nonmonophyly, young species stem ages and lack of geographical structure in rain forest phylogenies may reflect more widespread disturbance by drought and landscape evolution causing resident mortality that opens up greater opportunities for immigration and speciation. We recommend full species sampling and inclusion of multiple accessions representing individual species in phylogenies to highlight nonmonophyletic species, which we predict will be frequent in rain forest and savanna, and which represent excellent case studies of incipient speciation.

I. Introduction

The goal of this paper is to review species-level phylogenetic studies of clades of woody, tropical plants that have sampled numerous individuals within species and thereby can provide insights into their history. In particular, we examine whether in a phylogenetic context species are resolved as monophyletic groups (see Table 1 glossary) and consider their inferred ages. The majority of studies we review come from genera characteristic of the seasonally dry tropical forest (SDTF) biome in Latin America, and we are able to compare these with a much more limited number of examples from the tropical rain forest biome there. Based upon these examples, we make predictions about the phylogenetic structure of species in lowland neotropical biomes. We suggest that widespread, abundant species found in the major expanses of rain forest (Amazonia and the Guianas) and savanna (the 'cerrado' of central Brazil) are more likely to be nonmonophyletic, with more range-restricted daughter species embedded within them. By contrast, species endemic to the SDTF biome, especially those from drier, succulent-rich woodlands (*sensu* Schrire *et al.*, 2005), are more likely to be monophyletic and range-restricted. If true, these predictions have considerable implications, for example in understanding the evolutionary differences between these tropical biomes and especially the nature of the suite of woody species that have been shown to dominate overwhelmingly both Amazonia (Pitman *et al.*, 2001, 2013; ter Steege *et al.*, 2013) and the Brazilian cerrado (Bridgewater *et al.*, 2004).

II. Neotropical biomes

Insights in this paper flow from a consideration of major biomes in the Neotropics as distinct evolutionary theatres. Put simply, we consider that studying tropical biogeography and diversification by considering all organisms as merely 'tropical', as is the case in many studies, especially in macroecology (e.g. Mittelbach *et al.*, 2007), can obscure patterns and processes that are distinct among tropical biomes (Fine & Ree, 2006; Pennington *et al.*, 2009; Hughes *et al.*, 2013; Oliveira-Filho *et al.*, 2013). We aim to demonstrate that it can be more fruitful to consider the influence of the ecology of different major biomes on phylogenetic and biogeographical patterns in order to gain insights into the evolutionary process. Central to our argument is a concept that we should not examine plant diversification and biogeography solely within a framework of the effect of historical climatic and geological events. It is equally important to consider the influence of intrinsic attributes of organisms that reflect the major biome that they inhabit (Baker *et al.*, 2014), such as drought tolerance and the ability to survive fire.

Against this background, an understanding of the three major lowland neotropical biomes of rain forest, SDTF and savanna is important (Pennington *et al.*, 2006, 2009; Fig. 1). From an ecological perspective, these biomes are defined by variation in climate, edaphic factors and fire regime. Rain forest is found in areas with humid climates on a variety of soil types and seasonal flooding regimes (Whitmore, 1998; Fine *et al.*, 2005), but in areas with significant dry seasons, gives way to savanna and SDTF (Malhi *et al.*, 2009). Within these drier regions, SDTF grows on soils of relatively high pH and nutrient status, while savanna is found on acid, dystrophic soils (Furley & Ratter, 1988; Pennington *et al.*, 2006). Fire has not been an important ecological factor in SDTF, whereas the abundance of C_4 grasses makes the more productive savannas flammable in the dry season (Pennington *et al.*, 2006; Lehmann *et al.*, 2011).

Table 1 Glossary

Allopatric	Occurring in separate, nonoverlapping areas. Applied to speciation, it invokes mechanisms that involve geographical isolation
F_{st}	An index that measures the degree of population genetic structure. Theoretically bounded by 0 and 1, with values of 0 indicating no population structure and higher values indicating stronger population structure
Hyperdominant	A tree or shrub species that is significantly more common than is expected for a given type of woody plant community
Monophyletic	A monophyletic group contains an ancestor and all of its descendents. The term can apply to alleles, individuals, populations, species or higher level taxa (e.g. the genus *Inga* comprises a monophyletic group of species because all species of *Inga* are more closely related to another *Inga* species than to a species of another legume genus)
Oligarchy	A limited number of tree or shrub species that dominate a particular woody plant community
Paralogous genes	Paralogous genes result from a gene duplication within a genome
Parapatric	Geographical ranges that abut but do not significantly overlap along a narrow contact zone. Applied to speciation, it invokes mechanisms that involve selection across an ecological boundary
Paraphyletic	A group of alleles, individuals, populations, species, or higher level taxa that have descended from a common evolutionary ancestor but the group does not include all descendants of that ancestor (e.g. accessions of *Andira inermis* in Fig. 3 are a paraphyletic group)
Peripatric	Describes ranges that are geographically peripheral to a larger source population. Applied to speciation, it is a form of allopatric speciation involving geographical isolation of peripheral populations
Phylogenetic diversity	The sum of the branch lengths in an ultrametric phylogenetic tree that lead from either the root node or the most recent common ancestral node to the subset of taxa sampled from a particular community. A community with high phylogenetic diversity will have high richness in species and also in taxa above the species level (e.g. genera and families)
Polyphyletic	A group of species or higher level taxa whose constituents have different evolutionary ancestral origins (e.g. tree species with winged fruits have multiple ancestral origins)
Population bottleneck	A significant reduction in the size of a population that affects genetic diversity. This may be caused by a small founding population or events that reduce population sizes
Vicariance	The geographical separation of a population via fragmentation (e.g. by continental drift, formation of a mountain range or river, or climate change causing habitat fragmentation), which often results in the evolution of new species (see 'allopatric' above)
Wright–Fisher model	A model of genetic drift that assumes that generations do not overlap and that each allele found in the new generation is drawn independently at random from all alleles at a given locus present in the previous generation

III. Coalescence

A major focus of this review is to examine how accessions representing individual species are resolved in DNA-based phylogenies. An important conceptual background in this context is coalescent theory (Kingman, 1982, 2000; Hudson, 1991). The coalescent theory was developed using the Wright–Fisher model, which occupies a central role in population genetics. For the context of this article on plant species, a very useful recent review for the nonspecialist is provided by Naciri & Linder (2015). We emphasize that we are not using the background of the coalescent theory in its more usual population genetic context, for example to estimate effective population sizes or types of selection. We are using it to understand why DNA sequences from a given locus sampled from multiple individuals across the geographical distribution of a single plant species may or may not coalesce – be resolved as monophyletic – with respect to sequences from the same locus sampled from individuals of closely related species. Diverging species inherit mixtures of alleles for different genes, and, analysed in a phylogenetic context, some genes may indicate monophyly of a species, while others may cluster populations of different species together (Naciri & Linder, 2015). In the language of the coalescent theory, the genes that indicate species monophyly have ‘coalesced’ because ancestral polymorphisms have been eliminated in daughter lineages by drift or selection. In the plant systematics and phylogenetics literature, the term ‘lineage sorting’ is often used to describe the coalescent process (Doyle, 1992; Posada & Crandall, 2001), and the genes that do not coalesce within species are in a state of ‘incomplete’ lineage sorting (Naciri & Linder, 2015). In this case, ancestor and/or daughter lineages retain ancestral polymorphism through speciation events, and this will be reflected in nonmonophyly of species. Throughout this article, we use the shorthand of ‘species monophyly’, although strictly we should be referring to the coalescence, or monophyly, of conspecific DNA sequences.

Noncoalescence can reflect several factors, including hybridization and subsequent introgression, paralogous gene copies and inadequate taxonomy, and disentangling these may be difficult (Syring *et al.*, 2007; Naciri & Linder, 2015). In the absence of these factors, the time to coalescence depends on the effective population size (N_e), which is the size of a theoretical population under the Wright–Fisher model that explains the genetic diversity observed in the population sample (Naciri & Linder, 2015). N_e will usually be smaller than the census population size (N) because of unequal reproductive success of individuals but, in simple terms, given two populations of equal genetic diversity, the time to coalescence will be much faster in small populations than in large ones.

However, N_e is complex and affected by various factors, which thereby influence time to coalescence. For example, genome organization has an effect – nuclear genome doubling by polyploidy will double the N_e of all loci, and haploid genomes such as plastids have half the N_e of the nuclear genome in hermaphrodites. In this context of genome organization, nuclear ribosomal DNA, frequently used in phylogenetics, presents additional complications, often occurring at multiple nucleolar organizing regions and subject to concerted evolution and homogenization of copy types within individuals (Claire West *et al.*, 2014). Historical demographic changes will also influence N_e. For example, a population bottleneck followed by population expansion might result in a large, geographically widespread population with low N_e. By contrast, strong spatial genetic structuring within a species can lead to higher N_e than would be expected from the species population census size N. Other factors may influence the N_e of specific loci, rather than the entire genome. For example, loci far from centromeres experience lower recombination and will have a faster coalescence time (Naciri & Linder, 2015).

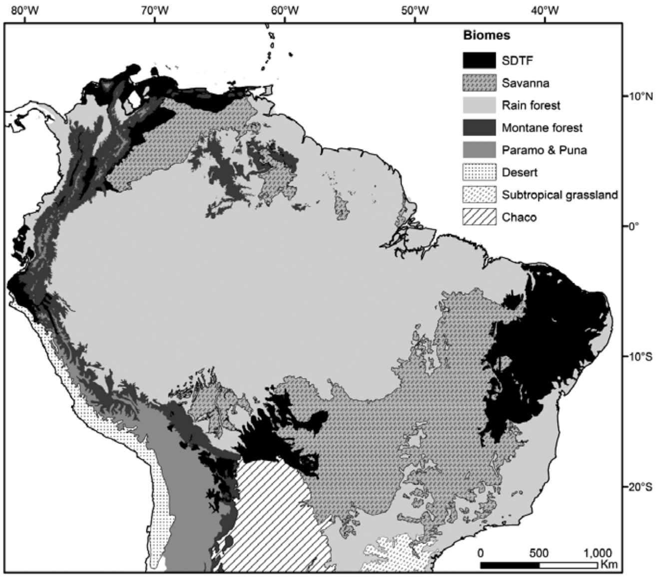

Fig. 1 Major biomes of tropical South America. Modified from Olson *et al.* (2001). SDTF, seasonally dry tropical forest.

IV. Species from seasonally dry tropical forests (SDTFs)

Numerous phylogenetic studies have examined clades containing species confined to neotropical SDTFs, many of which sample multiple individuals representing single species (Becerra, 2003, 2005; Pennington *et al.*, 2004, 2010, 2011; Duno de Stefano *et al.*, 2010; Ireland *et al.*, 2010; Govindarajulu *et al.*, 2011; Queiroz & Lavin, 2011; De Nova *et al.*, 2012; Särkinen *et al.*, 2012; Gagnon *et al.*, 2015) . Previous reviews of these studies have emphasized the geographical phylogenetic structure found in these phylogenies – that related species tend to occupy the same geographical area – and that individual species tend to be old compared with those in neighbouring biomes (Lavin, 2006; Pennington *et al.*, 2009; Särkinen *et al.*, 2012). Here, we will touch upon these topics of phylogenetic geographical structure and species ages, but will place more emphasis on how multiple accessions of individual species are resolved in phylogenies.

In these phylogenies of SDTF clades, when multiple DNA sequences representing individual species are sampled, species recognized by taxonomists on morphological grounds tend to be resolved as monophyletic groups, often on long stem lineages, and with short branches in the crown group (Fig. 2). This pattern of coalescence of multiple accessions of DNA sequences representing single species on long stem lineages is evidence for long-term evolutionary persistence of populations with small effective population sizes (cf. Barraclough, 2010).

The frequency of the pattern of conspecific coalescence of DNA sequences is exemplified by a number of studies, many of which are focused on legumes (Duno de Stefano *et al.*, 2010; Pennington *et al.*, 2010, 2011; Queiroz & Lavin, 2011; Särkinen *et al.*, 2012). Patterns of conspecific coalescence of DNA sequences and old species stem ages (often over 1 million yr (Ma); Pennington *et al.*, 2010; Särkinen *et al.*, 2012) may be frequent in the SDTF biome because of two mutually reinforcing factors: (1) the geographical distribution of the SDTF biome in small patches across the Neotropics (Fig. 1); and (2) low immigration rates into these isolated patches.

The small size of many SDTF patches, for example in interAndean valleys, means that woody species that are endemic there can never achieve the huge population sizes that are possible for rain forest species in the vastness of Amazonia. The contrast of geographical scale is massive; the SDTFs of the Marañon, Apurimac and Mantaro interAndean valleys in Peru are estimated to occupy *c.* 3100 km^2 in total (INRENA, 1995), only 0.04% of the size of the Amazon rain forest (an estimated 6 million km^2; ter Steege *et al.*, 2013). In addition, the geographical scatter of SDTFs across the Neotropics may cause dispersal rates among separate areas to be low, but we do emphasize that evidence for such long-distance dispersal events is found in phylogenies of SDTF plants (Lavin *et al.*, 2004; Pennington *et al.*, 2011; Fig. 2c (*Coursetia caribaea* var. *pacifica*)). We highlight the low rates of successful immigration into SDTF nuclei as a key factor reducing the possibility of achieving a larger effective population size – propagules may arrive by long-distance dispersal to an isolated SDTF area, but their probability of establishment is low. Low immigration is suggested by the biology of the SDTF system. In its natural state, SDTF was not frequently disturbed by fire over evolutionary timescales, as evidenced by the lack of fire adaptations in its constituent flora, and the abundance of fire-sensitive taxa such as succulent cacti (Pennington *et al.*, 2006, 2009). In addition, SDTF species are well adapted to survive severe droughts; if the rains fail, woody species simply shed their leaves and remain dormant. We suggest that SDTF has low, relatively constant mortality and recruitment rates and that rates of compositional change are slow. Low resident mortality means that the community is generally saturated (*sensu* Hubbell, 2001), which suggests that immigrant propagules are unlikely to establish as there is no space in the system. This hypothesis of low immigration rates is supported by the common finding of old age estimates for geographically confined clades of conspecific accessions of SDTF plants (Pennington *et al.*, 2009; Särkinen *et al.*, 2012). The combination of low immigration rates and scattered SDTF nuclei acts to isolate populations to the point where they have a high likelihood of becoming distinct species, which subsequently persist for long time periods.

In summary, the small patch size of SDTF at a continental scale and its ecological stability which restricts immigration act in concert to keep effective population sizes of woody SDTF species low, and this means that time to coalescence will in general be short and that the length of time of persistence of species will be long. In taxonomic terms, species recognized based upon morphology are likely to be resolved as monophyletic in DNA sequence-based phylogenies and often with an old stem age estimate (see 'VI. Ages of species' below).

In considering taxonomy, one benefit of the extensive sampling within species carried out in a broader phylogenetic context that we are advocating is that it can be an enormous aid to the taxonomic process. For example, in several clades of SDTF species, intensive sampling within species has suggested ranking at species level of populations that were previously noted as being only geographically distinct morphological forms (Duno de Stefano *et al.*, 2010; Pennington *et al.*, 2011; Queiroz & Lavin, 2011; Särkinen *et al.*, 2011). In these cases, morphologically and geographically distinct forms of a single SDTF species are each resolved as distinct monophyletic groups. The reanalysis of morphology, geography, and molecular phylogenetic data resulted in the formal description of new species or the resurrection of taxonomic synonyms to the species level, each of which were found to be endemic to a relatively small SDTF area. An example is provided in Fig. 2(a), where *Poissonia eriantha* is shown to be a well-supported monophyletic group, reciprocally monophyletic with *Poissonia hypoleuca* and *Poissonia orbicularis* and with an estimated stem age of > 5 Ma. *Poissonia eriantha* and *P. orbicularis* were tentatively considered the same species, *P. orbicularis*, by Lavin (1988), but new field observations show clear and consistent differences in flower colour and size (Fig. 2a), and *P. eriantha* is now recognized at species level (Pennington *et al.*, 2011).

V. Species from rain forest, with particular focus on Amazonia

1. Predictions

Well-sampled, species-level phylogenies that additionally sample multiple accessions of individual species of neotropical rain forest trees are infrequent. Because of this, we first make some predictions about how rain forest tree species, especially those that are abundant and widespread in Amazonia, may be resolved in phylogenies. These predictions are based upon consideration of the biology of rain forest tree species, and evidence that processes of dispersal, rather than vicariance, may underlie the biogeography of most Amazonian tree groups. We then review the admittedly fragmentary empirical evidence in order to see how it fits these predictions.

We suggest that massive population sizes, long life spans, and effective seed and pollen flow, all factors that characterize many species of rain forest trees, mean that retention of ancestral genetic polymorphism in them may be likely because time to coalescence will be very long. This suggests that many tree species with large effective population sizes will not be resolved as monophyletic in phylogenies when they are sampled using multiple accessions. Although it is not our intention to review all potential modes of speciation of Amazonian trees, which may be numerous (Bush, 1994), we suggest that noncoalescence will be common whether speciation is allopatric and driven by vicariance (e.g. by Pleistocene

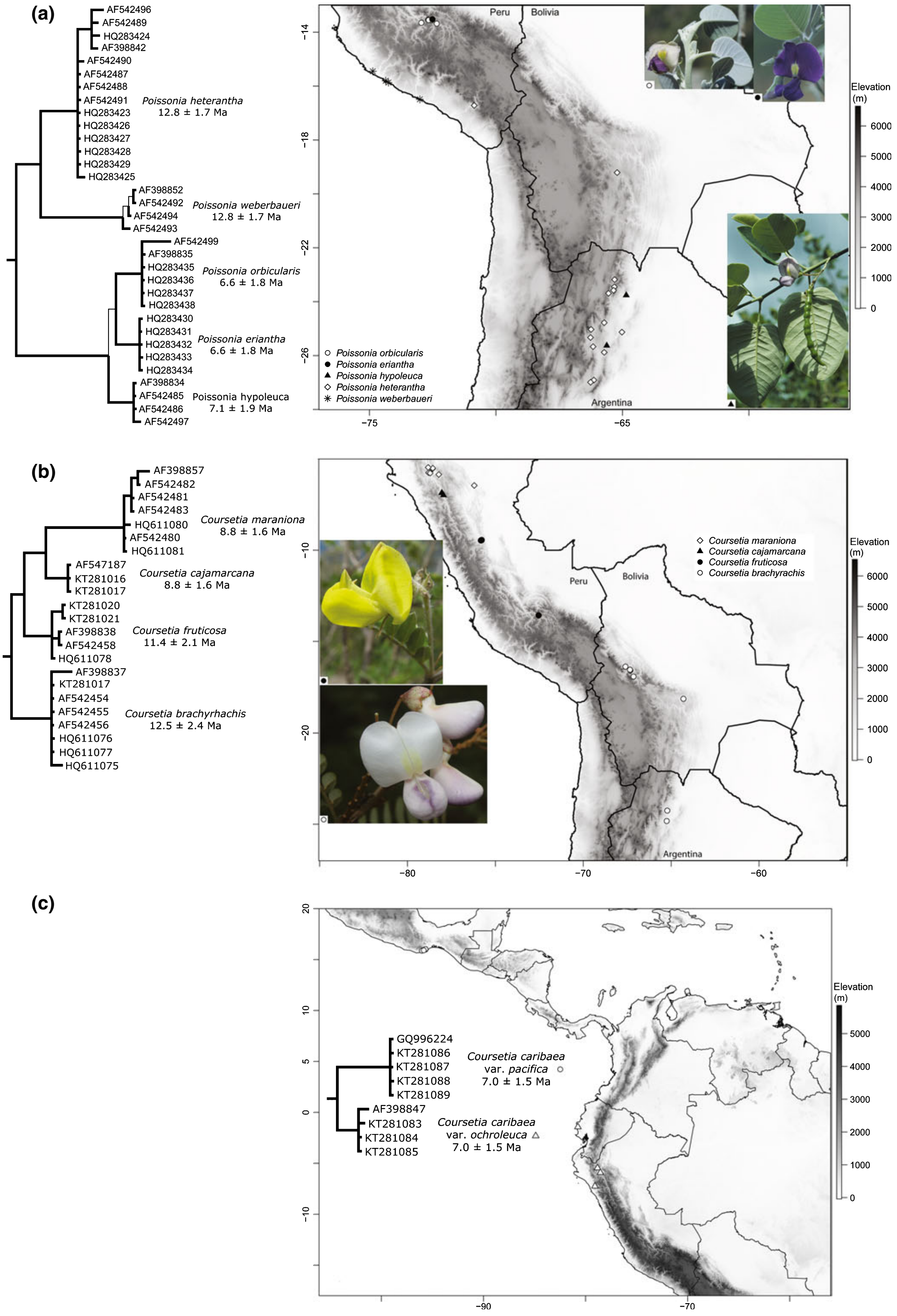

Fig. 2 Monophyly of seasonally dry tropical forest (SDTF) species. Phylogenies were inferred from a Bayesian analysis of nuclear ribosomal internal transcribed spacer (ITS) sequence data and run for over 7×10^6 generations at likelihood stationarity (thick branches indicate a posterior probability of 0.95–1.00). Note the conspecific coalescence of all accessions representing each species, and the long stem lineages and short branches within the crown group of each species. GenBank accession numbers are reported as terminal labels and these are associated with specimen voucher information (TreeBase 18275). (a) *Poissonia* (Pennington *et al.*, 2011). The map shows the distributions of five species of *Poissonia*; photographs are of *Poissonia orbicularis* (Apurimac valley, Peru; R. T. Pennington), *Poissonia eriantha* (Apurimac valley, Peru; R. T. Pennington), and *Poissonia hypoleuca* (M. Lavin) each showing the unifoliolate leaf and few flowered racemose inflorescences distinctive of this clade. (b) *Coursetia fruticosa* clade (M. Lavin *et al.*, unpublished). The map shows the distributions of four species of the *C. fruticosa* clade. Photographs: *C. fruticosa* (Apurimac valley, Peru; R. T. Pennington); *Coursetia brachyrachis* (Chiquitana, Bolivia; D. Cardoso). (c) *Coursetia caribaea* var. *pacifica* clade (M. Lavin *et al.*, unpublished). The sister relationship of *C. caribaea* var. *pacifica* (Mexico) and *C. caribaea* var. *ochroleuca* (Ecuador and Peru) is suggestive of long-distance dispersal given the wide geographical disjunction that is not consistent with any tectonic event. Ma, million yr, in reference to age estimate of stem clade.

climate changes; Haffer, 1982) or dispersal, or parapatric and driven by selection (e.g. adaptation to edaphic conditions; Fine *et al.*, 2005, 2013).

ter Steege *et al.* (2013) showed that 50% of all individual trees in Amazonia belong to just 227 'hyperdominant' species. These have a median population size of 5.79×10^8 individuals, which would suggest that time to coalescence would be long because of the influence of overall census population size (N) on N_e, which determines the rate at which gene lineages coalesce in a neutral context. As an example, a tree species with N_e of 1 million individuals and a generation time of 10 yr will require 50 million yr for monophyly to be achieved at all loci (Naciri & Linder, 2015). This may even be an underestimate given that for many tropical tree species generation times may be an order of magnitude longer (Baker *et al.*, 2014). Of course, N_e may be additionally affected by the various factors outlined in 'III. Coalescence' above, but the sheer population size of many Amazonian rain forest tree species suggests that times to coalescence across their genomes will be extremely long. If such a species is the ancestor of two daughter species because its range is split by a vicariant event, unless these daughter species are very old, they are unlikely to be resolved as monophyletic. Similarly, if such an abundant species gives rise to a daughter species via long-distance dispersal events (i.e. a peripatric mechanism), while the daughter species may be resolved as monophyletic and with young age estimates because of the population bottleneck causing an initial small effective population size, the ancestral species would not be resolved as a reciprocally monophyletic sister group to it because its time to coalescence would be long (Barraclough, 2010).

Effective gene flow in Amazonia may also be contributing to the long coalescence times by helping to ensure that effective population sizes are large. Several decades of genetic research have shown that tropical trees are outcrossing, have extensive gene flow and maintain high levels of genetic diversity (Dick *et al.*, 2008). Pollen dispersal, almost exclusively mediated biotically (wind pollination is rare in rain forest trees), has been shown to be effective, with dispersal distances of kilometres (Dick *et al.*, 2008). Although tropical trees have more genetic differentiation among populations ($F_{st} = 0.177$) than their counterparts in the temperate zone ($F_{st} = 0.116$; $n = 82$), this result was driven by a few species that display high F_{st} because of major genetic differentiation across biogeographical barriers (e.g. the Andes). No similar review is available for tropical trees in Amazonia alone, but we predict that F_{st} values will be low, which would be indicative of little population subdivision.

Population genetic estimates of seed dispersal distances are rare for tropical rain forest trees, so making inferences of contemporary gene flow via seeds is difficult. However, recent phylogenetic studies of Amazonian trees suggest that historical long-distance dispersal has been a major force shaping the composition of local tree communities in the Amazon and Guianas region. Local and regional communities of *Inga* have phylogenetic diversity that does not differ from a random expectation (Lavin, 2006; Pennington *et al.*, 2009; R. T. Pennington *et al.*, unpublished), and this is duplicated in *Swartzia*, *Protium* and *Guatteria* (R.T. Pennington *et al.*, unpublished), all of which are species-rich genera and often abundant components of neotropical rain forests. Put simply, the species found in a local or regional community are scattered across the phylogeny of each genus. Although they have not been examined quantitatively, visual inspection of the phylogenies of *Clusia* (Gustafsson & Bittrich, 2003) and *Guarea* (Pennington & Clarkson, 2013) suggests that they show the same pattern. This pattern of phylogenetic scatter is explicable only by these communities being assembled by historical dispersal (Lavin, 2006; Pennington *et al.*, 2009; Pennington & Dick, 2010). The lack of any repeated geographical patterns across phylogenies suggests that processes of historical vicariance across potential suggested barriers in Amazonia (e.g. Lake Pebas) have not been important in speciation (Pennington & Dick, 2010; Hughes *et al.*, 2013). Additionally, it is not a pattern of *in situ* regional diversification, which is what typifies SDTF phylogenies. It implies that immigration must be more feasible and more frequent into rain forest tree communities than into SDTF tree communities. Interestingly, in this context, Hubbell (2001) estimated high immigration rates using relative species abundance curves among samples from neotropical rain forests. Rain forest communities may be more invasible by immigrants than SDTF communities because of higher and more variable rates of mortality and recruitment caused by drought (da Costa *et al.*, 2010) and high levels of landscape instability over historical timescales, for example in radical changes of river courses across Amazonia (Hoorn *et al.*, 2010). These disturbance events cause tree mortality, creating space in the community, meaning that establishment is more likely. By contrast, as outlined above, the ecological stability of SDTF may be restricting successful immigration.

Because of higher rates of successful immigration and establishment in neotropical rain forests, we predict that the pattern of phylogenetic nesting of mainly recently evolved monophyletic species embedded within progenitor species will be found frequently in rain forest tree genera, reflecting a progenitor-descendent speciation process by a peripatric mechanism with a role for peripheral isolation of populations of widespread species. Distinguishing this peripatric process from one of local habitat-mediated speciation, which has been suggested to be important in Amazonia (Gentry, 1981, 1982; Fine *et al.*, 2005, 2013; Misiewicz & Fine, 2014), using phylogenetic data alone may be difficult, because the mechanism of a potentially abundant and widespread ancestor giving rise to an initially range-restricted daughter species predicts the same phylogenetic patterns.

2. Empirical patterns

It is notable that in many cases neotropical rain forest tree species are not resolved as monophyletic in phylogenies – there is not conspecific coalescence of DNA sequences. This is the pattern observed for *Andira inermis* subsp. *inermis* (Fig. 3; Simon *et al.*, 2009), *Ficus inspida* subsp. *insipida* (Coronado *et al.*, 2014), *Dussia tessmannii* (Winterton *et al.*, 2014), numerous *Inga* species (Dexter *et al.*, 2010), several *Protium* species (Fine *et al.*, 2014), and several Sapotaceae species (J. E. Richardson *et al.*, unpublished data).

Some of these examples, such as *A. inermis* (Fig. 3), *F. insipida* subsp. *insipida* (Coronado *et al.*, 2014) and some species of *Inga*

(Dexter *et al.*, 2010), show clear patterns of phylogenetic nesting, the strongest form of phylogenetic evidence for progenitor-derivative speciation (Crawford, 2010). It seems notable that Crawford's review of progenitor-derivative speciation, which focused on temperate zone examples, pinpointed just one clear-cut example of phylogenetic nesting (Baldwin, 2005), whereas, despite the currently thin empirical evidence base, examples from tropical rain forests are not difficult to uncover. The two species nested within *A. inermis*, *Andira multistipula* and *Andira jaliscensis*, were indicated as close relatives based upon morphology, with *A. inermis* suggested as a progenitor species for both of them (Pennington, 2003). So, in this case, we suggest that this may be evidence for progenitor-descendent speciation, with ancestral genetic polymorphism retained in *A. inermis*, although we note that this conclusion is based upon a single genetic locus, nuclear ribosomal internal transcribed spacers (ITS). The same pattern is also shown in *Protium subserratum* (Fine *et al.*, 2013), where a morphotype confined to nonwhite soils in Peru is a monophyletic group nested within a paraphyletic white sand morphotype. It is notable that this pattern of phylogenetic nesting of range-restricted

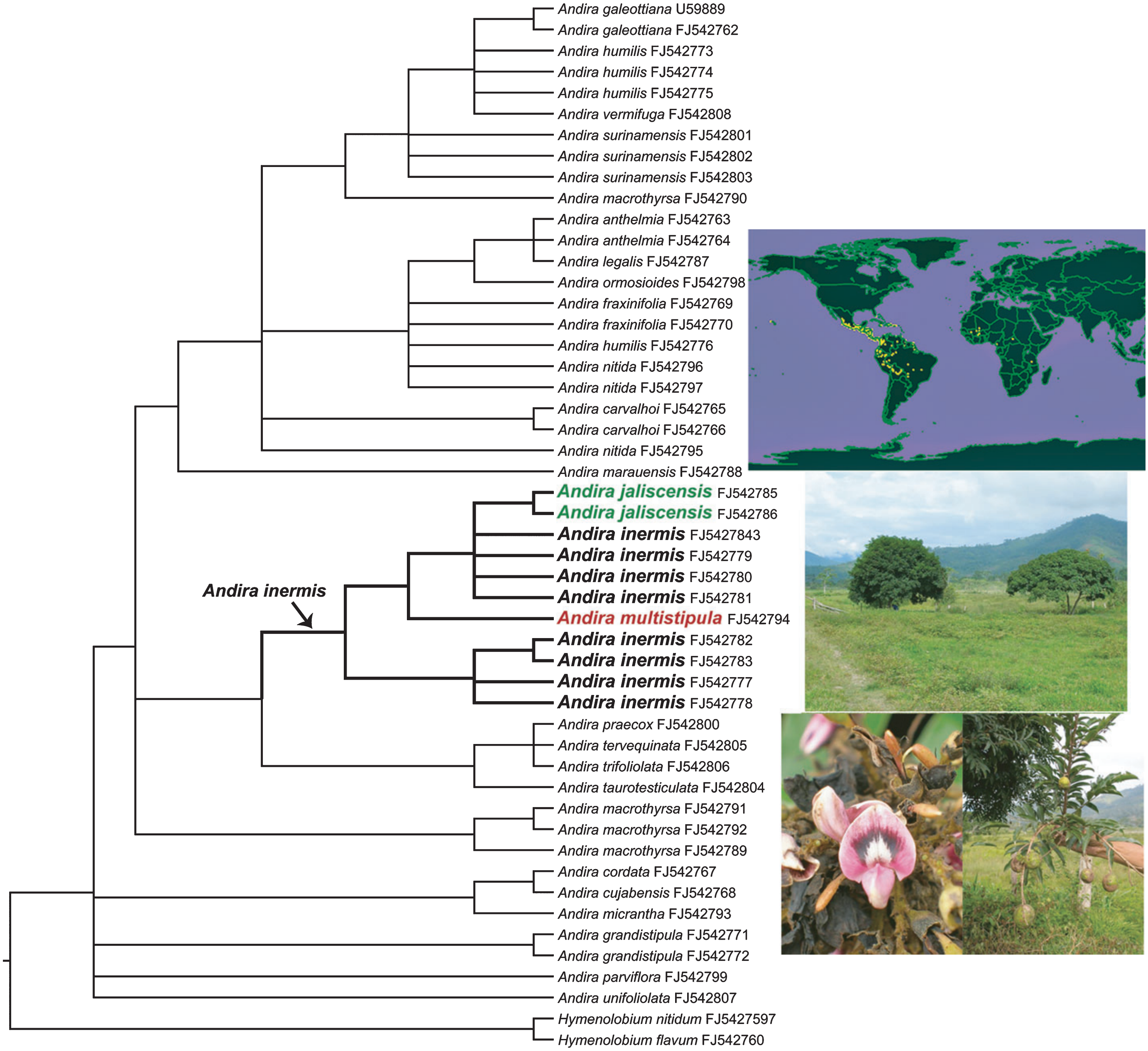

Fig. 3 Nonmonophyly of rain forest tree species, exemplified by *Andira inermis*. The phylogeny was inferred from a Bayesian analysis of nuclear ribosomal internal transcribed spacer (ITS) sequence data (TreeBase 18275; GenBank numbers and associated specimen voucher information are reported at www.pnas.org/cgi/content/full/0903410106/DCSupplemental; Simon *et al.*, 2009) and run for over 7 × 10^6 generations at likelihood stationarity (nearly all branches were resolved with a posterior probability of 0.95–1.00). Note that multiple accessions representing *A. inermis* are not resolved as a monophyletic group because accessions of *A. jaliscensis* and *A. multistipula* are nested among them. Photographs (top to bottom): distribution of *A. inermis*; trees, flowers and fruit of *A. inermis* (Satipo, Peru; R. T. Pennington).

species within widespread species, which we suggest may be characteristic of rain forest phylogenies, is not common in SDTF phylogenies, for example being seen only in one of 49 SDTF species sampled in the robinioid legume phylogeny (M. Lavin *et al.*, unpublished), in the instance where *Coursetia greenmannii* is nested within *C. caribaea* (Duno de Stefano *et al.*, 2010).

Of these rain forest phylogenies, Dexter *et al.* (2010) sampled most thoroughly across a number of species (50) using ITS and chloroplast trnDT sequences. They discovered that eight of these species were paraphyletic, six shared sequences with other species (i.e. were phylogenetically unresolved) and four were polyphyletic. This was after using a reciprocal illumination process where species identifications based upon morphology were revised in the light of the phylogeny, essentially using a species concept that emphasized monophyly as a primary criterion. Despite this bias for species monophyly, they discovered that only 32 of 50 *Inga* species were resolved as monophyletic. In addition, the *Inga* study was restricted to only 30 000 km^2 in Madre de Dios, Peru, 10 times larger in size than Peru's dry interAndean valleys, but still a tiny proportion of the entire neotropical distribution of *Inga*. We predict that sampling more widely across *Inga*'s distribution and over all *c.* 300 species of the genus would be likely to result in discovery of further nonmonophyletic species.

In other cases such as *Dussia* (Winterton *et al.*, 2014), *Inga* (Dexter *et al.*, 2010) and *Protium* (Fine *et al.*, 2014), some species are not resolved as monophyletic, but this largely reflects lack of phylogenetic resolution. As Crawford (2010) pointed out, these unresolved patterns do not contradict a progenitor-descendent speciation mechanism, but clearly more evidence is required to resolve these phylogenies. Such unresolved patterns contrast with the general prevalence of species monophyly in SDTF groups. The lack of resolution may reflect the recency of speciation in many rain forest tree clades (see 'VI. Ages of species').

Noncoalescence may be attributable to several factors, including inadequate taxonomy, hybridization and subsequent introgression, and sequencing paralogous gene copies (see 'III. Coalescence'; Syring *et al.*, 2007). Future studies using more extensive genomic data and breeding system experiments will be necessary to understand their influence, and we cannot rule out these factors as explanations in our tropical tree species examples. For example, Muellner *et al.* (2009) and Cavers *et al.* (2013), suggested that nonmonophyly of *Cedrela odorata* (Meliaceae) reflects a need for a revised taxonomy to recognize morphologically cryptic species. In other studies, hybridization has been shown to operate in *Carapa* (Meliaceae; Scotti-Saintagne *et al.*, 2013). However, while some workers consider that hybridization is an important evolutionary force in some tropical trees (Cannon & Lerdau, 2015), a breeding system study of six *Inga* species from Costa Rica demonstrated that all were inter-sterile and self-incompatible (Koptur, 1984), suggesting that frequent hybridization is unlikely. The lack of hybridization in *Inga* is also suggested by the morphological distinctiveness of sympatric species. Up to a remarkable 43 *Inga* species can be found in just 25 Ha of rain forest in Amazonian Ecuador (Valencia *et al.*, 1994). If all these local *Inga* species were hybridizing, it seems unlikely that ecologists and taxonomists would be able to distinguish these species based upon morphology, and they have little trouble in doing so (Pennington, 1997; Dexter *et al.*, 2010). It is possible that future studies using genomic data will demonstrate hybridization and introgression as a key force in tropical tree species diversification, and therefore explanations for species nonmonophyly, as recently demonstrated by the temperate herbaceous genus *Pedicularis* (Eaton & Ree, 2013). However, our purpose here is to suggest that it is important to explore other potential explanations for noncoalescence within rain forest species, or reasons why, when these species are monophyletic, most may be recently evolved. We suggest that in some cases the noncoalescence observed may be just that – a failure to achieve coalescence by lineage sorting – meaning that a species such as *A. inermis* is not necessarily a 'bad' species that needs re-delimitation.

VI. Ages of species

As alluded to above, perhaps one of the most important differences between the species from different tropical woody biomes involves age estimates of species. Phylogenies of woody SDTF plants that sample multiple individuals within species show characteristic patterns of branch lengths, with long stem lineages leading to the species crown group, in which branches are often short (Fig. 2). When the phylogenies are dated, this pattern of branch lengths is reflected in old stem ages (Figs 2, 4: the date when the species split from its nearest relative) and recent crown ages (Figs 2, 4: the age of the ancestor of all extant conspecific accessions) for individual species (Figs 2, 4; Schrire *et al.*, 2009; Duno de Stefano *et al.*, 2010; Pennington *et al.*, 2010, 2011; Queiroz & Lavin, 2011; reviewed by Särkinen *et al.*, 2012). For example, Särkinen *et al.*'s (2012) review of five legume genera showed mean stem group ages for species confined to SDTF of 6.1 Ma. Comparison of stem and crown group ages of 38 species of Robinioid legumes confined to SDTF confirms this pattern (Fig. 4). As explained above, the recent crown age estimates contrasting with much older stem age estimates for individual species is consistent with a signature of populations of small effective sizes that are persistent over evolutionary timescales and are difficult to immigrate into successfully.

Because few phylogenies of rain forest trees have sampled multiple accessions within species, making generalizations about ages of individual species is difficult. However, despite the fact that rain forest is the oldest lowland neotropical biome, the indications are that a higher proportion of species within it may be younger than species found in SDTF. Phylogenies of several species-rich genera indicate major subclades of species within them with crown group dates of 5 Ma or less (Pliocene and Quaternary; e.g. within each of *Guatteria* (Erkens *et al.*, 2007), *Protium* (Fine *et al.*, 2014), *Trichilia* and *Guarea* (Koenen *et al.*, 2015)). In all these cases except for *Protium* (Fine *et al.*, 2014), the phylogenies have the drawback of sampling single individuals per species, making estimation of species crown ages impossible. However, the young age of the nodes separating the accessions of different contemporary species suggests that these are likely to be very young.

This vision of a greater proportion of young extant species in Amazonia than is found in SDTF, which reflects the ecologically dynamic or unstable environment expected of a tropical rain forest,

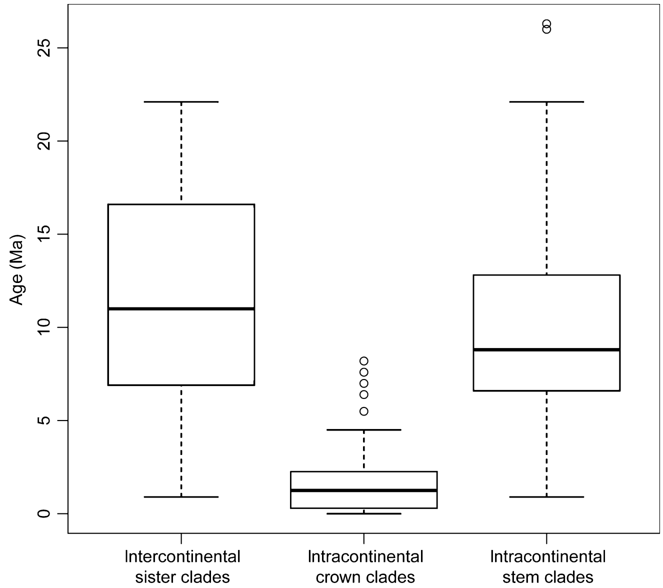

Fig. 4 Ages of clades of species of Robinieae with coalesced conspecific samples (monophyletic species) found in seasonally dry tropical forests (SDTFs) and woodlands. The crown clade ages are significantly young compared with the stem clade ages. The ages of divergence of sister species found on the same continent ('intracontinental stem clades'; e.g. Fig. 2a,b) and of sister species found on different continents ('intercontinental sister clades'; e.g. Fig. 2c) are approximately equal. These equivalent ages of sister clades that are confined to a single continent versus sister clades that have dispersed between two continents suggest that the general ability to disperse over long distances is not limiting compared with the ability to successfully colonize a patch of SDTF. The thick horizontal line represents the median age, the box encompasses 50% of the data points and the whiskers 95%, and the open circles are outlying data points (age estimates). Ma, million yr.

may appear to be in contradiction with literature that emphasizes the Palaeocene age of neotropical rain forests based upon geological evidence (Burnham & Johnson, 2004; Jaramillo *et al.*, 2006), and the dating of rate shifts in Amazonian tree diversifications in the Oligocene and Miocene (Fine *et al.*, 2014; Koenen *et al.*, 2015). However, we stress here that we are discussing the age of individual extant species, and even for clades such as Protieae, which have a late Oligocene/early Miocene (38–21 Ma) crown group age, many extant species are young (post-Pliocene), even if other species are older. Indeed, in many instances the dated phylogenies of rain forest trees suggest that many extant species date to the Pleistocene (e.g. *Inga* (Richardson *et al.*, 2001), *Trichilia* (Koenen *et al.*, 2015), *Guarea* (Pennington & Clarkson, 2013; Koenen *et al.*, 2015) and *Protium* (Fine *et al.*, 2014)), which also runs counter to the view, based largely on early vertebrate studies (reviewed by Moritz *et al.*, 2000), that tropical rain forest species in general pre-date the Pleistocene (Colinvaux *et al.*, 2001).

We suspect that these young ages may reflect greater opportunities for allopatric species formation within large expanses of rain forest such as Amazonia. A clue to the source of these opportunities comes from the lack of phylogenetic geographical structure in phylogenies of Amazonian rain forest trees outlined in the previous section. This indicates high invasibility of rain forest communities by immigrants in this ecologically dynamic environment, which is suggestive that their speciation may involve a peripatric mechanism, where founder populations are geographically isolated just sufficiently for speciation to initiate.

The idea of a high proportion of young species in Amazonia may also appear to contradict the one available study that attempts to date the ages of a series of extant neotropical rain forest species (Dick *et al.*, 2013). Dick *et al.* (2013) used ITS and chloroplast DNA sequence data to infer ages for 12 widespread Amazonian tree species, sampling multiple individuals of each (9–20 individuals, and 50 in the case of *Symphonia globulifera*) using molecular clock techniques. They demonstrated that 9 of 12 widespread Amazon tree species have Pliocene or earlier lineages (> 2.6 Ma), with seven dating from the Miocene (> 5.6 Ma) and three > 8 Ma. It should be noted that, because Dick *et al.* (2013) studied single species in isolation without sampling congeneric species, they could not infer species stem ages and that these dates represent crown group ages only. With the exception of two of these species, which come from genera where they are the single extant species (*Symphonia globulifera* and *Poulsenia armata*), we suggest that it is feasible that these widespread species are progenitors of younger daughter species that may be phylogenetically nested within them and potentially much younger. In the case of *A. inermis*, one of the Dick *et al.* (2013) study species, this is the case (Fig. 3), because *A. jaliscensis* and *A. multistipula* are nested within it. We do not deny that there will be some very old species in Amazonia. However, we suggest, based upon empirical evidence for young rain forest species in many species-rich genera (e.g. *Inga*, *Trichilia* and *Protium*) and our suggestion that long-distance dispersal may be driving frequent peripatric speciation, that many rain forest tree species may be young and potentially nested within older, widespread species.

Population genetic studies of single widespread rain forest species may be somewhat limited in their inferences if they fail to sample relevant phylogenetic diversity. This diversity is represented by other species, recognized as taxonomically distinct, which may be nested within the widespread species. Delimiting the correct scope of population genetic or phylogeographic studies will be done best in the context of a prior phylogenetic framework sampling numerous individuals of as many species as possible because it will identify these patterns of phylogenetic nestedness. The old ages that Dick *et al.* (2013) estimated for the widespread Amazonian rain forest species are accurate, but any daughter species nested within them will be younger if they originated by a progenitor-derivative speciation mechanism. Dick *et al.*'s general conclusion is that, because these widespread species arose in the Miocene when global climates were far warmer, they should be pre-adapted to future higher temperatures. We do not dispute this conclusion for the species sampled, but it may not be the case for any species derived from them, which may be much younger, therefore originating in cooler climates, and within which selection may have changed traits conferring temperature tolerance.

VII. Species from neotropical savannas

There are two main areas of South American tropical savanna (Fig. 1), the 'cerrado' which occupies 2 million km^2 of central

Brazil, and the Llanos which covers *c.* 300 km^2 of Colombia and Venezuela (Blydenstein, 1967; Huber *et al.*, 2006). The cerrado is much more species-rich, with a remarkable 10 000 vascular plant species, of which an estimated 35% are endemic (Ratter *et al.*, 2006). Fossil evidence for C_4 grasses, which dominate the ground layer of savannas, indicates that, compared with both rain forests and SDTF, savannas are geologically young, rising to dominance globally only *c.* 4 Ma (Edwards *et al.*, 2010). This picture of geological recency for the cerrado is supported by dated phylogenies that show that cerrado-confined clades are never older than 10 million yr old (Myr), with most lineages diversifying at 4 Ma or less (Simon *et al.*, 2009; Pennington & Hughes, 2014). Therefore, entire cerrado clades, which can number 50 species or more (in *Mimosa*; Simon *et al.*, 2009), are younger than many individual dry forest species.

None of these cerrado plant phylogenies sample multiple individuals of species, but we predict that patterns will mirror those described above for rain forest. Like Amazonia, the cerrado is dominated by an oligarchy of widespread, ecologically dominant species (Bridgewater *et al.*, 2004). Bridgewater *et al.* demonstrated that, from a total flora of *c.* 1000 tree species, 121 dominate the cerrado, contributing 66% of the total species composition and 75% of the Importance Value Index (a sum of relative frequency, plus relative dominance plus relative density) at any one site. The oligarchy of dominant cerrado species will have large population sizes because of the combination of their local abundance and large geographical ranges. Just as described for widespread, abundant rain forest species, time to coalescence may be slow within such species. Given that speciation events within the cerrado must be young – < 4 Myr in many cases – we predict that cerrado species will often be nonmonophyletic, especially those that are widespread and that may be progenitors for other species. Because the cerrado is disturbed by regular fire, which may cause mortality of plants, we predict that, like rain forest, it may represent a system open for successful establishment of immigrant propagules. We therefore predict that the phylogenetic geographical structure of cerrado clades may be low, like the pattern described above for rain forests – species from a given genus found growing together in local and regional communities are likely to be unrelated. Contrary to this, a phylogeny for *Calliandra*, a genus with high species endemism in the 'campos ruprestres' ('rocky fields'), a fire-prone, savanna-like vegetation found at higher elevations at the eastern margins of the cerrado, does show significant geographical structure of clades, with several to many endemics confined to specific areas (de Souza *et al.*, 2013). This may reflect the island-like distribution of campos rupestres, and it underscores the need for better sampled species phylogenies from the main area of the cerrados in central Brazil.

VIII. Conclusions and ways forward

1. A better understanding of tropical plant species, especially widespread, dominant ones

The purpose of this review has been to use DNA-based phylogenies to investigate the nature of woody, tropical species. It leads to hypotheses that can be tested in future studies, for example that species from SDTF will more frequently be resolved as monophyletic with old age estimates, whereas widespread tropical rain forest and savanna species may often be nonmonophyletic as a result of noncoalescence of conspecific DNA sequences, or monophyletic but with younger stem group ages. We explain these differences here as a function of the expected evolutionary stability of a localized patch of SDTF in contrast to the frequent drought, physical disturbance or fire experienced by a patch of rain forest or savanna.

A clear route forward for phylogenetic studies of tropical plants, whether the purpose is investigating evolutionary relationships, biogeography or taxonomy, is to include multiple accessions representing multiple individuals of all species. Ironically, to understand tropical plant speciation it may be most important to sample thoroughly across common species, which are those most frequently ignored by field collectors, whose focus is on rarities. However, some of these abundant species are playing a disproportionate role in the earth system – 50% of carbon storage and cycling in Amazonia is concentrated in only 1% of its tree species (Fauset *et al.*, 2015). Therefore, a better understanding of the nature of abundant Amazonian rain forest tree species is of great relevance, but few of them have been investigated in a population genetic, phylogeographic or phylogenetic context. We suggest a fruitful route to understand the evolutionary history and nature of these species will be a phylogenetic approach, sampling multiple individuals representing them and their congeners. Not all the 227 'hyperdominant' Amazonian tree species listed by ter Steege *et al.* (2013) are geographically widespread, although most are (cf. the 'oligarchic' species of Pitman *et al.*, 2001, 2013). It is especially these widespread, dominant species that we predict will be shown to be nonmonophyletic, with other, often more range-restricted species nested within them, as demonstrated by our examples in 'V. Species from rain forest' above.

Our main purpose has not been to discuss how DNA sequence data may be used in the process of species delimitation, but there are clear implications in this area. As explained in 'IV. Species from seasonally dry tropical forests' above, in SDTF clades, intensive sampling within species has resulted in taxonomic changes (Duno de Stefano *et al.*, 2010; Särkinen *et al.*, 2011; Gagnon *et al.*, 2015). However, in this context of taxonomy, we emphasize that just because species are resolved as nonmonophyletic in phylogenies, with other species nested within them, does not mean that they are 'bad' species that will necessarily require taxonomic redelimitation. For example, as pointed out in 'V. Species from rain forest' above, there are good reasons to suspect that they may be held together by highly effective gene flow. This genetic cohesion, allied to their morphological distinctiveness, suggests that they correspond to the lineage species concept of De Queiroz (2007) and they may well be good biological species (e.g. *sensu* Mayr, 1982 – i.e. with no genetic exchange with species they give rise to). If contemporary gene flow is not making them cohesive across their entire wide ranges, they may be on the road to disintegrating into a number of daughter species, but, because of large Ne and long generation times, the length of time required for widespread progenitor species to be resolved as monophyletic in phylogenies may be very long indeed. The concept of nonmonophyletic species in general is not new or controversial (Rieseberg & Brouillet, 1994; Knowles & Carstens, 2007; Naciri & Linder, 2015), and it calls into question the definition of species based upon some level of genetic exclusivity such as monophyly (Knowles & Carstens, 2007). Certainly, for some of the cases we have highlighted, such as *A. inermis* (Fig. 3), a species definition based on exclusivity seems inappropriate and, as a consequence, identifying this species by unique DNA sequence variation (a 'barcode') may not be possible.

With some exceptions (Govindarajulu *et al.*, 2011; Fine *et al.*, 2014), the majority of the phylogenies discussed in this paper are based on just a few loci, and require further exploration with multiple loci, preferably numerous nuclear loci. Such data sets are becoming more widely available for plants (Weitemier *et al.*, 2014), and, with such multilocus data in hand, we will be able to explore patterns of coalescence within species across a much greater number of loci. In addition to this, we believe that a novel approach to investigating species boundaries in tropical plants is offered by coalescent methods (Knowles & Carstens, 2007; Fujita *et al.*, 2012; Aydin *et al.*, 2014). Coalescent-based species delimitation methods use probabilistic models to test alternative hypotheses of species delimitation. Their background assumption is that species are independent evolutionary lineages (cf. De Queiroz, 2007), and such lineages do not need to be resolved as monophyletic in gene trees, which is why they may be especially appropriate to deploy for woody, tropical clades. Traditional taxonomists may be resistant to applying coalescent species delimitation approaches, but, as emphasized by Fujita *et al.* (2012), they would not replace morphological approaches to taxonomy. In fact, a morphology-based taxonomy is the starting point for erecting the alternative hypotheses of species delimitation for coalescent-based testing (e.g. in Fig. 3, should *A. inermis*, *A. jaliscensis* and *A. multistipula* be one species, or three?). Indeed, speaking as two traditional taxonomists, we suspect that using a coalescent framework for species delimitation may result in better agreement with traditional morphological delimitations than methods based on genetic exclusivity criteria in widespread Amazonian tree species, in which time to coalescence will be very long.

2. A better understanding of ecological processes across tropical biomes

Hypotheses about differences in the ecology of rain forests, SDTF and savanna, such as higher and more variable rates of compositional change in rain forest and savanna, are central to some of the ideas presented in this review. These hypotheses are, however, speculative and will be hard to test until there is consistent long-term ecological monitoring across biomes in the tropics. Currently, such monitoring using long-term ecological plots is well established in rain forests (e.g. the Amazon Forest Inventory Network (RAINFOR) and the Center for Tropical Forest Science and Forest Global Earth Observatory (CTFS-ForestGEO)), but it is less widespread and needs more development in SDTF and savannas.

3. A better understanding of biomes and biogeography

Densely sampled phylogenies, containing multiple accessions of all species, will certainly be a powerful tool for a better understanding of the nature of tropical plant species. An associated, perhaps more fundamental insight they can provide is into how patterns in plant phylogeny and biogeography differ among biomes. We have argued that the patterns found in densely sampled phylogenies of SDTF woody plants – a predominance of geographical structure and coalesced conspecific accessions with old stem age estimates – reflect that the SDTF biome is highly dispersal limited. By contrast, the lack of coalescence in conspecific accessions, young species stem ages and lack of geographical structure found in rain forest tree phylogenies reflect the distinct ecology of this biome. Therefore, a key outcome of this review has been to suggest that ecological processes operating over evolutionary timescales play a role in tropical biogeography, a view that is supported by recent studies of Amazonian trees (Baker *et al.*, 2014) and birds (Smith *et al.*, 2014). We suggest that such an approach emphasizing ecology might be more productive than that provided by the dispersal-vicariance debate (Nelson & Platnick, 1981), which has been the traditional background against which to discuss geographical structure in phylogenies. Patterns of reciprocally monophyletic clades of conspecific samples, generally interpreted as a signature of vicariance, may not suggest a single historical event but rather the on-going effect of ecological stability.

Acknowledgements

We thank Tim Baker, Paul Fine, Colin Hughes, Kyle Dexter, Alex Twyford and Pete Hollingsworth for critical comments and discussion, Tiina Särkinen for preparing Fig. 1, Domingos Cardoso for photographs and Julia Weintritt for editorial help.

References

Aydin Z, Marcussen T, Alaattin SE, Oxelman B. 2014. Marginal likelihood estimate comparisons to obtain optimal species delimitations in Silene sect. Cryptoneurae (Caryophyllaceae). *PLoS ONE* **9**: e116266.

Baker TR, Pennington RT, Magallon S, Laurance WF, Alexiades M, Alvarez E, Araujo A, Arets EJMM, Aymard G, de Oliveira AA *et al.* 2014. Rapid turnover promotes high diversification of Amazonian trees. *Ecology Letters* **17**: 527–536.

Baldwin BG. 2005. Origin of the serpentine-endemic herb *Layia discoidea* from the widespread *L. glandulosa* (Compositae). *Evolution* **59**: 2473–2479.

Barraclough TG. 2010. Evolving entities: towards a unified framework for understanding diversity at the species and higher levels. *Philosophical Transactions of the Royal Society, Biological Sciences* **365**: 1801–1813.

Becerra JX. 2003. Synchronous coadaptation in an ancient case of herbivory. *Proceedings of the National Academy of Sciences, USA* **100**: 12804–12807.

Becerra JX. 2005. Timing the origin and expansion of the Mexican tropical dry forest. *Proceedings of the National Academy of Sciences, USA* **102**: 10919–10923.

Blydenstein J. 1967. Tropical savanna vegetation of the llanos of Colombia. *Ecology* **48**: 1–15.

Bridgewater S, Ratter JA, Ribeiro JF. 2004. Biogeographic patterns, β-diversity and dominance in the cerrado biome of Brazil. *Biodiversity and Conservation* **13**: 2295–2318.

Burnham RJ, Johnson KR. 2004. South American palaeobotany and the origins of neotropical rain forests. *Philosophical Transactions of the Royal Society of London series B* **359**: 1595–1610.

Bush MB. 1994. Amazonian speciation: a necessarily complex model. *Journal of Biogeography* **21**: 5–17.

Cannon CH, Lerdau MT. 2015. Variable mating behaviors and the maintenance of tropical biodiversity. *Frontiers in Genetics* **6**: 183

Cavers S, Telford A, Arenal Cruz F, Castañeda AJP, Valencia R, Navarro C, Buonamici A, Lowe AJ, Vendramin GG. 2013. Cryptic species and phylogeographical structure in the tree *Cedrela odorata* L. thoughout the Neotropics. *Journal of Biogeography* **40**: 732–746.

Claire West C, James SA, Davey RP, Dicks J, Roberts IN. 2014. Ribosomal DNA sequence heterogeneity reflects intraspecies phylogenies and predicts genome structure in two contrasting yeast species. *Systematic Biology* **63**: 543–554.

Coronado E, Dexter KG, Poelchau MF, Hollingsworth PM, Phillips OL, Pennington RT. 2014. *Ficusinsipida* subsp. *insipida* (Moraceae) reveals the role of ecology in the phylogeography of widespread neotropical rain forest tree species. *Journal of Biogeography* **41**: 1697–1709.

da Costa LAC, Galbraith D, Almeida S, Tanaka Portela BT, da Costa M, de Athaydes Silva Junior J, Braga AP, de Goncalves PHL, de Oliveira AAR, Fisher R *et al.* 2010. Effect of 7 yr of experimental drought on vegetation dynamics and biomass storage of an eastern Amazonian rainforest. *New Phytologist* **187**: 579–591.

Colinvaux PA, Irion G, Räsänen ME, Bush MB. 2001. A paradigm to be discarded: Geological and paleoecological data falsify the Haffer & Prance refuge hypothesis of Amazonian speciation. *Amazoniana* **16**: 609–646.

Crawford DJ. 2010. Progenitor-derivative species pairs and plant speciation. *Taxon* **59**: 1413–1423.

De Nova JA, Medina R, Montero JC, Weeks A, Rosell JA, Olson ME, Eguiarte LE, Magallón S. 2012. Insights into the historical construction of species-rich Mesoamerican seasonally dry tropical forests: the diversification of *Bursera* (Burseraceae, Sapindales). *New Phytologist* **193**: 276–287.

De Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology* **56**: 879–886.

Dexter KG, Pennington TD, Cunningham CW. 2010. Using DNA to assess errors in tropical tree identifications: how often are ecologists wrong and when does it matter? *Ecological Monographs* **80**: 267–286.

Dick CW, Jones FA, Hardy OJ, Petit R. 2008. Spatial scales of seed and pollen-mediated gene flow in tropical forest trees. *Tropical Plant Biology* **1**: 20–33.

Dick CW, Lewis S, Maslin M, Bermingham E. 2013. Neogene origins and implied warmth tolerance of Amazon tree species. *Ecology and Evolution* **3**: 162–169.

Doyle JJ. 1992. Gene trees and species trees: molecular systematics as one character taxonomy. *Systematic Botany* **17**: 144–163.

Duno de Stefano R, Fernández-Concha GC, Can-Itza LL, Lavin M. 2010. The morphological and phylogenetic distinctions of *Coursetia greenmanii* (Leguminosae): taxonomic and ecological implications. *Systematic Botany* **35**: 289–295.

Eaton DAR, Ree RH. 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (Pedicularis: Orobanchaceae). *Systematic Biology* **62**: 689–706.

Edwards EJ, Osborne CP, Strömberg CAE, Smith SA, C_4 Grasses Consortium. 2010. The origins of C_4 grasslands: integrating evolutionary and ecosystem science. *Science* **328**: 587–591.

Erkens RHJ, Chatrou LW, Maas JW, van der Niet T, Savolainen V. 2007. A rapid diversification of rainforest trees (*Guatteria*; Annonaceae) following dispersal from Central into South America. *Molecular Phylogenetics and Evolution* **44**: 399–411.

Fauset S, Johnson MO, Gloor M, Baker TR, Monteagudo MA, Brienen RJW, Feldpausch TR, Lopez-Gonzalez G, Malhi Y, ter Steege H *et al.* 2015. Hyperdominance in Amazonian forest carbon cycling. *Nature Communications* **6**: 6857.

Fine PVA, Daly DC, Villa Muñoz G, Mesones I, Cameron KM. 2005. The contribution of edaphic heterogeneity to the evolution and diversity of Burseraceae trees in the western Amazon. *Evolution* **59**: 1464–1478.

Fine PVA, Metz MR, Lokvam J, Mesones I, Ayarza Zuniga JM, Lamarre GPA, Vasquez Pilco M, Baraloto C. 2013. Insect herbivores, chemical innovation and the evolution of habitat specialization in Amazonian trees. *Ecology* **94**: 1764–1775.

Fine PVA, Ree R. 2006. Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *American Naturalist* **168**: 796–804.

Fine PVA, Zapata F, Daly DC. 2014. Investigating processes of Neotropical rain forest tree diversification by examining the evolution and historical biogeography of Protieae (Burseraceae). *Evolution* **69**: 1988–2004.

Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C. 2012. Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology and Evolution* **27**: 480–488.

Furley PA, Ratter JA. 1988. Soil resources and plant communities of the central Brazilian cerrado and their development. *Journal of Biogeography* **15**: 97–108.

Gagnon E, Hughes CE, Lewis GP, Bruneau A. 2015. A new cryptic species in a new cryptic genus in the Caesalpinia group (Leguminosae) from the seasonally dry inter-Andean valleys of South America. *Taxon* **63**: 468–490.

Gentry AH. 1981. Distributional patterns and an additional species of the *Passiflora vitifolia* complex: Amazonian species diversity due to edaphically differentiated communities. *Plant Systematics and Evolution* **137**: 95–105.

Gentry AH. 1982. Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? *Annals of the Missouri Botanical Garden* **69**: 557–593.

Govindarajulu R, Hughes CE, Bailey CD. 2011. Phylogenetic and population genetic analyses of diploid *Leucaena* (Leguminosae) reveal cryptic species diversity and patterns of allopatric divergent speciation. *American Journal of Botany* **98**: 2049–2063.

Gustafsson M, Bittrich V. 2003. Evolution of morphological diversity and resin secretion in flowers of *Clusia* L. (Clusiaceae): insights from ITS sequence variation. *Nordic Journal of Botany* **22**: 183–203.

Haffer J. 1982. General aspects of refuge theory. In: Prance GT, ed. *Biological diversification in the tropics*. New York, NY, USA: Columbia, 6–24.

Hoorn C, Vonhof H, Wesselingh F, eds. 2010. *Amazonia, landscape and species evolution: a look into the past*. Oxford, UK: Wiley-Blackwell.

Hubbell SP. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton, NJ, USA: Princeton University Press.

Huber O, Duno de Stefano R, Aymard G, Riina R. 2006. Flora and vegetation of the Venezuelan Llanos: a review. In: Pennington RT, Lewis GP, Ratter JA, eds. *Neotropical savannas and seasonally dry forests: plant diversity, biogeography and conservation*. Boca Raton, FL, USA: CRC Press, 95–120.

Hudson RR. 1991. Gene genealogies and the coalescent process. *Oxford Surveys in Evolutionary Biology* **7**: 1–44.

Hughes CE, Pennington RT, Antonelli A. 2013. Neotropical plant evolution – assembling the big picture. *Botanical Journal of the Linnean Society* **171**: 1–18.

INRENA (Instituto Nacional de Recursos Naturales). 1995. *Mapa forestal del Peru escala 1:1000000 con guia explicativa*. Lima, Peru: INRENA.

Ireland HE, Kite GC, Veitch NC, Chase MW, Schrire B, Lavin M, Pennington RT. 2010. Biogeographic, ecological, and morphological structure of a phylogenetic analysis of *Ateleia* (Swartzieae-Leguminosae) derived from combined molecular and morphological/chemical data. *Botanical Journal of the Linnean Society* **162**: 39–53.

Jaramillo C, Rueda MJ, Mora G. 2006. Cenozoic plant diversity in the Neotropics. *Science* **311**: 1893–1896.

Kingman JFC. 1982. The coalescent. *Stochastic Processes and Their Applications* **13**: 235–248.

Kingman JFC. 2000. Origins of the coalescent 1974–1982. *Genetics* **156**: 1461–1463.

Knowles LL, Carstens BC. 2007. Delimiting species without monophyletic gene trees. *Systematic Biology* **56**: 887–895.

Koenen EJM, Clarkson JJ, Pennington TD, Chatrou LW. 2015. Recently evolved diversity and convergent radiations of rainforest mahoganies (Meliaceae) shed new light on the origins of rainforest hyperdiversity. *New Phytologist* **207**: 327–339.

Koptur S. 1984. Outcrossing and pollinator limitation of fruit set: breeding systems of neotropical Inga trees. *Evolution* **38**: 1130–1143.

Lavin M. 1988. Systematics of *Coursetia* (Leguminosae-Papilionoideae). *Systematic Botany Monographs* **21**: 1–167.

Lavin M. 2006. Floristic and geographical stability of discontinuous seasonally dry tropical forests explains patterns of plant phylogeny and endemism. In: Pennington RT, Lewis GP, Ratter JA, eds. *Neotropical savannas and seasonally dry forests: plant diversity, biogeography and conservation*. Boca Raton, FL, USA: CRC Press, 433–447.

Lavin M, Schrire BD, Lewis G, Pennington RT, Delgado-Salinas A, Thulin M, Hughes CE, Beyra-Matos A, Wojciechowski MF. 2004. Metacommunity processes rather than continental tectonic history better explain geographically structured phylogenies in legumes. *Philosophical Transactions of the Royal Society, Biological Sciences* **359**: 1509–1522.

Lehmann CER, Archibald SA, Hoffmann WA, Bond WJ. 2011. Deciphering the distribution of the savanna biome. *New Phytologist* **191**: 197–209.

Malhi Y, Aragão LEOC, Galbraith D, Huntingford C, Fisher R, Zelazowski P, Sitche S, McSweeney C, Meir P. 2009. Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proceedings of the National Academy of Sciences, USA* **106**: 20610–20615.

Mayr E. 1982. *The growth of biological thought.* Cambridge, MA, USA: Belnap Press.

Misiewicz TM, Fine PVA. 2014. Evidence for ecological divergence across a mosaic of soil types in an Amazonian tropical tree: *Protium subserratum* (Burseraceae). *Molecular Ecology* **23**: 2543–2558.

Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MB, Harrison SP, Hurlbert AH, Knowlton N, Lessios HA *et al.* 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters* **10**: 315–333.

Moritz C, Patton JL, Schneider CJ, Smith TB. 2000. Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics* **31**: 533–563.

Muellner AN, Pennington TD, Chase MW. 2009. Molecular phylogenetics of Neotropical Cedreleae (mahogany family, Meliaceae) based on nuclear and plastid DNA sequences reveal multiple origins of "*Cedrela odorata*". *Molecular Phylogenetics and Evolution* **52**: 461–469.

Naciri Y, Linder HP. 2015. Species delimitation and relationships: the dance of the seven veils. *Taxon* **64**: 3–16.

Nelson G, Platnick NI. 1981. *Systematics and biogeography: cladistics and vicariance.* New York, NY, USA: Columbia University Press.

Oliveira-Filho AT, Cardoso D, Schrire BD, Lewis GP, Pennington RT, Brummer TJ, Rotella J, Lavin M. 2013. Stability structures tropical woody plant diversity more than seasonality: Insights into the ecology of high legume-succulent-plant biodiversity. *South African Journal of Botany* **89**: 42–57.

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'amico JA, Itoua I, Strand HE, Morrison JC *et al.* 2001. Terrestrial ecoregions of the world: a new map of life on earth. *BioScience* **51**: 933–938.

Pennington RT. 2003. A monograph of *Andira* (Leguminosae-Papilionoideae). *Systematic Botany Monographs* **64**: 1–145.

Pennington RT, Daza A, Reynel C, Lavin M. 2011. *Poissonia eriantha* (Leguminosae) from Cuzco, Peru: an overlooked species underscores a pattern of narrow endemism common to seasonally dry neotropical vegetation. *Systematic Botany* **36**: 59–68.

Pennington RT, Dick CW. 2010. Diversification of the Amazonian flora and its relation to key geological and environmental events: a molecular perspective. In: Hoorn C, Vonhof H, Wesselingh F, eds. *Amazonia, landscape and species evolution: a look into the past.* Oxford, UK: Wiley-Blackwell, 373–385.

Pennington RT, Lavin M, Hughes C, Sarkinen T, Lewis G, Klitgaard B. 2010. Differing diversification histories in the Andean biodiversity hotspot. *Proceedings of the National Academy of Sciences, USA* **107**: 13783–13787.

Pennington RT, Hughes CE. 2014. The remarkable congruence of New and Old World savannah origins. *New Phytologist* **204**: 4–6.

Pennington RT, Lavin M, Oliveira-Filho A. 2009. Woody plant diversity, evolution and ecology in the tropics: perspectives from seasonally dry tropical forests. *Annual Review of Ecology, Evolution, and Systematics* **40**: 437–457.

Pennington RT, Lavin M, Prado DE, Pendry CA, Pell S, Butterworth C. 2004. Neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Philosophical Transactions of the Royal Society, Biological Sciences* **359**: 515–538.

Pennington RT, Lewis GP, Ratter JA. 2006. An overview of the plant diversity, biogeography and conservation of neotropical savannas and seasonally dry forests. In: Pennington RT, Lewis GP, Ratter JA, eds. *Neotropical savannas and seasonally dry forests: plant diversity, biogeography and conservation.* Boca Raton, FL, USA: CRC Press, 1–29.

Pennington TD. 1997. *The genus* Inga*: botany.* London, UK: Royal Botanic Gardens, Kew.

Pennington TD, Clarkson J. 2013. A revision of *Guarea* (Meliaceae). *Edinburgh Journal of Botany* **70**: 179–362.

Pitman NC, Silman MR, Terborgh JW. 2013. Oligarchies in Amazonian tree communities: a ten-year review. *Ecography* **36**: 114–123.

Pitman NCA, Terborgh JW, Silman MR, Nunez P, Neill DA, Ceron CE, Palacios WA, Aulestia M. 2001. Dominance and distribution of tree species in upper Amazonian terra firme forests. *Ecology* **82**: 2101–2117.

Posada D, Crandall KA. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution* **16**: 37–45.

Queiroz LP, Lavin M. 2011. *Coursetia* (Leguminosae) from eastern Brazil: nuclear ribosomal and chloroplast DNA sequence analysis reveal the monophyly of three caatinga-inhabiting species. *Systematic Botany* **36**: 69–79.

Ratter JA, Bridgewater S, Ribeiro JF. 2006. Biodiversity patterns of the woody vegetation of the Brazilian cerrados. In: Pennington RT, Lewis GP, Ratter JA, eds. *Neotropical savannas and seasonally dry forests: plant diversity, biogeography and conservation.* Boca Raton, FL, USA: CRC Press, 31–66.

Richardson JE, Pennington RT, Pennington TD, Hollingsworth PM. 2001. Recent and rapid diversification of a species-rich genus of neotropical trees. *Science* **293**: 2242–2245.

Rieseberg LH, Brouillet K. 1994. Are many plant species paraphyletic? *Taxon* **43**: 21–32.

Särkinen T, Marcelo Peña JL, Daza Yomona A, Simon MF, Pennington RT, Hughes CE. 2011. Underestimated endemic species diversity in the dry inter-Andean valley of the Río Marañón, northern Peru: An example from *Mimosa* (Leguminosae, Mimosoideae). *Taxon* **60**: 139–150.

Särkinen T, Pennington RT, Lavin M, Simon MF, Hughes CE. 2012. Evolutionary islands in the Andes: persistence and isolation explains high endemism in Andean dry tropical forests. *Journal of Biogeography* **39**: 884–900.

Schrire BD, Lavin M, Barker NP, Forest F. 2009. Phylogeny of the tribe Indigofereae (Leguminosae-Papilionoideae): geographically structured more in succulent-rich and temperate settings than in grass-rich environments. *American Journal of Botany* **96**: 816–852.

Schrire BD, Lavin M, Lewis GP. 2005. Global distribution patterns of the Leguminosae: insights from recent phylogenies. *Biologiske Skrifter* **55**: 375–422.

Scotti-Saintagne C, Dick CW, Caron H, Vendramin GG, Guichoux E, Buonamici A, Duret C, Sire P, Valencia R, Lemes MR *et al.* 2013. Phylogeography of a species complex of lowland Neotropical rain forest trees (*Carapa*, Meliaceae). *Journal of Biogeography* **40**: 676–692.

Simon M, Grether R, de Queiroz LP, Skema C, Pennington RT, Hughes CE. 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in-situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences, USA* **106**: 20359–20364.

Smith BT, McCormack JE, Cuervo AM, Hickerson MJ, Aleixo A, Cadena CD, Pérez-Emán J, Burney CW, Xie X, Harvey MG *et al.* 2014. The drivers of tropical speciation. *Nature* **515**: 406–409.

de Souza ER, Lewis GP, Forest F, Schnadelbach AS, van den Berg C, de Queiroz LP. 2013. Phylogeny of *Calliandra* (Leguminosae: Mimosoideae) based on nuclear and plastid molecular markers. *Taxon* **62**: 1200–1219.

ter Steege H, Pitman NCA, Sabatier D, Baraloto C, Salomão RP, Guevara JE, Phillips OL, Castilho CV, Magnusson WE, Molino JF *et al.* 2013. Hyperdominance in the Amazonian Tree Flora. *Science* **342**: 1243092.

Syring J, Farrell K, Businský R, Cronn R, Liston A. 2007. Widespread genealogical nonmonophyly in species of *Pinus* subgenus *Strobus. Systematic Biology* **5**: 163–181.

Valencia R, Blaslev H, Paz-y-Miño-C G. 1994. High tree alpha-diversity in Amazonian Ecuador. *Biodiversity and Conservation* **3**: 21–28.

Weitemier K, Straub SCK, Cronn RC, Fishbein M, Schmickl R, McDonnell A, Liston A. 2014. Hyb-seq: combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences* **2**: 1400042.

Whitmore TC. 1998. *An introduction to tropical rain forests.* Oxford, UK: Oxford University Press.

Winterton C, Richardson JE, Hollingsworth M, Clark A, Zamora N, Pennington RT. 2014. Historical biogeography of the neotropical legume genus *Dussia*: the Andes, the Panama Isthmus and the Chocó. In: Stevens WD, Montiel OM, Raven PH, eds. *Paleobotany and biogeography: a festschrift for Alan Graham in his 80th year.* St Louis, MO, USA: Missouri Botanical Garden Press, 128.