Phylogenetic relatedness within neotropical fern communities increases with soil fertility

Samuli Lehtonen[1]*, Mirkka M. Jones[1, 2], Gabriela Zuquim[1], Jefferson Prado[3] and Hanna Tuomisto[1]

[1] Department of Biology, University of Turku, FI-20014 Turku, Finland
[2] Ecoinformatics and Biodiversity, Department of Bioscience, Aarhus University, Ny Munkegade 114, DK-8000 Aarhus C., Denmark
[3] Herbarium SP, Instituto de Botânica, C.P. 68041, CEP 04045-972, São Paulo, SP, Brazil

*Correspondence: Samuli Lehtonen, Department of Biology, University of Turku, FI-20014 Turku, Finland. E-mail: samile@utu.fi

Running title: Phylogenetic structure of neotropical fern communities

ABSTRACT

Aim To examine the relative importance of gradients in soil fertility and rainfall for the phylogenetic structure of neotropical forest fern communities, and to quantify how much the results are affected by phylogenetic resolution.

Location Tropical lowland forests in Brazil (central Amazonia) and Panama (the Panama canal watershed).

Methods We inventoried local fern communities at a total of 87 sites to model how their species richness and phylogenetic relatedness varied along gradients in soil fertility and rainfall. We produced a time-calibrated species-level molecular phylogeny of ferns, and quantified the phylogenetic relatedness of species within local communities using Faith's phylogenetic diversity (PD) and mean phylogenetic diversity (MPD). Calculations were compared for phylogenies resolved to the species and genus levels.

Results (1) We found significant and consistent effects of soil nutrient status on local phylogenetic community structure in both regions. In contrast, phylogenetic structure showed only weak or no relationships with rainfall. (2) In both regions, MPD declined with increasing soil fertility, which means that fern communities on poor soils consisted, on average, of less closely related lineages than fern communities on rich soils. (3) Different fern genera were over-represented in different sections of the soil nutrient gradient. (4) Qualitatively similar results were obtained whether phylogenies were resolved to the species or the genus level.

Main conclusions Our results highlight the importance of edaphic variation in structuring plant communities over evolutionary time-scales. Within the tropical forests studied, the effects of soil variation on local phylogenetic community structure seem to outweigh those of climate. Several fern genera show strong edaphic niche conservatism to either poor or rich soils, whereas many other genera have radiated to span a rather broad edaphic range. PD-based measures are so dominated by deep relationships that genus-level phylogenies are sufficient to investigate these patterns in tropical fern communities.

Keywords
Community phylogenetics, edaphic variation, niche conservatism, phylogenetic overdispersion, phylogenetic underdispersion, precipitation, Pteridophyta

INTRODUCTION

Within neotropical forests, plant species richness and species composition show strong relationships with rainfall regimes (e.g. ter Steege *et al*., 2003; Punyasena *et al.*, 2008). Turnover in plant community composition with rainfall also occurs at higher taxonomic levels (Punyasena *et al*., 2008), which suggests that climatic niche conservatism plays a role in structuring tropical forest plant communities. Adaptive radiation along ecological gradients other than climate is also likely to have contributed to the diversification of rainforest floras (e.g. Fine *et al*., 2005, 2013; Tuomisto, 2006). There is, for example, ample evidence for edaphic determinism of community composition (e.g., Tuomisto & Poulsen, 1996; Tuomisto *et al*., 2003a, 2003b; Phillips *et al*., 2003; Jones *et al*., 2013; ter Steege *et al.*, 2006), but the degree of conservatism in the edaphic niches of tropical plants remains little explored (but cf. Schreeg *et al*., 2010; Fine & Kembel, 2011).

It has become increasingly common to investigate the phylogenetic relatedness of co-occurring species in order to understand the links between evolution, historical biogeography, and ecological processes in shaping local community structure (e.g. Webb *et al*., 2002; Emerson & Gillespie, 2008; Vamosi *et al.*, 2009). Phylogenetic clustering is observed when species in a community are more closely related to each other than expected by chance. This has generally been interpreted as evidence of niche conservatism: close relatives are constrained to occur in similar environments (Webb *et al.*, 2002). Phylogenetic overdispersion is observed when species in a local community are less closely related to each other than expected by chance. This has been suggested to reflect competitive interactions between close relatives (Webb *et al*., 2002). Alternative explanations for phylogenetic overdispersion include the idea that older environments contain proportionately more representatives of old lineages than more recent environments do (Kooyman *et al*., 2011). In general, stable environmental conditions may promote regional species accumulation through reduced extinction probabilities, enabling the persistence of old lineages, and perhaps also by promoting speciation. However, speciation may also be augmented by environmental variability (Doebeli & Dieckmann, 2003).

Relatively few studies have examined trends in the phylogenetic structure of tropical plant communities along environmental gradients other than climate. These have mostly focused on tropical trees or palms at relatively broad spatial scales and across coarse habitat categories (Fine & Kembel, 2011; Eiserhardt *et al*., 2012; Umaña *et al*., 2012). Studies including quantitative soil data have mostly been restricted to local spatial scales (e.g. Schreeg *et al*., 2010), although Eiserhardt *et al*. (2012) used soil sand content as a surrogate for fertility at broader scales. Most importantly, these studies illustrate that habitat specialization, in association with habitat evolution over long time periods, can leave a signal in phylogenetic community structure. For example, there is some evidence that past geophysical processes are reflected in the phylogenetic structure of Western Amazonian tree communities (Fine & Kembel, 2011).

Rainforests with a modern plant family composition were apparently established in the Neotropics in the late Paleocene (Wing *et al*., 2009) or early Eocene (Wilf *et al*., 2003). Geological history and molecular systematic data suggest that in Amazonia these early forests may have been adapted to poor soils derived from the Guiana shield (Kubitzki, 1989; Frasier *et al*., 2008). More fertile soils with a higher clay content began to develop in Amazonia after the uplift of the Andes and consequent fluvial inputs of nutrient-rich sediments from the mountains into the lowlands (Kubitzki, 1989). These became widely distributed after the drainage of the west Amazonian Pebas wetland system in the Late Miocene (Räsänen et al. 1995). Within Amazonia, poor soils are both older and more widespread than rich soils, poor soils might be expected to have also accumulated more species there. However, a recent study on pteridophytes (ferns and lycophytes) found the

opposite trend (Tuomisto *et al*., 2014). In contrast to the Amazon region, Central America comprises a mosaic of volcanic and sedimentary soils derived from geological formations of generally more recent origin, produced by its gradual evolution from a chain of volcanic islands to a continuous land mass upon the closure of the isthmus of Panama c. 3 Ma ago (Kirby *et al.*, 2008).

Here we use a phylogenetic hypothesis of unusually high quality to compare the phylogenetic community structure of ferns across soil and rainfall gradients in two Neotropical forest regions with different geological histories: central Amazonia and central Panama. The younger age and more dynamic recent history of the isthmus of Panama can be predicted to have resulted in more phylogenetically clustered community assemblages at a regional level than in central Amazonia, where long-term geological stability is likely to have resulted in a proportianately greater representation of older lineages, producing less clustered communities. We also test if the diversity patterns observed in Amazonia (more species in the younger and less widespread substrates) are repeated at a broader scale, and whether phylogenetic information can shed light on the factors behind these patterns. Within each region, we also assess if (and if so, how) phylogenetic community structure varies with local environmental conditions. We furthermore compare the results obtained with phylogenies resolved to the species and genus levels to assess to what degree phylogenetic inferences are affected by the accuracy of information on interspecific relationships.

METHODS

Field sampling and environmental variables

We inventoried understory ferns in lowland tropical forests in two geographical areas, central Amazonia (the states of Amazonas and Roraima in Brazil) and Panama (the Panama Canal watershed; Fig. 1). All 32 Amazonian sites were at less than 200 m above sea level in non-inundated old-growth forests. Most sites had tall forests on clayey to loamy soils (*terra firme*), but two had white sand soils and forests with a reduced stature and relatively simple canopy structure (*campinarana*). The 55 Panamanian sites were predominantly at less than 300 m a.s.l. (with one site at 530 m and one at 660 m) and the forests were of variable ages, from secondary to old growth. Each Panamanian site was represented by a sampling unit of 0.1 ha that consisted of two parallel transects (100 m x 5 m) separated by about 35 m. Each Amazonian site was originally sampled using a transect of 500 m x 5 m, but to match the sampling unit size used in Panama, only the central 200 m were used in the present analyses.

Understory ferns were defined as ferns growing either on the ground or as epiphytes or climbers on the lower parts of tree trunks. To be included in the inventory, a fern individual had to have at least one green leaf longer than 10 cm, and this had to be less than 2 m above ground if the individual was epiphytic or climbing. In the data analyses, no separation was made between individuals on different substrates. All individuals were identified to species or given a field name if the species name was not known. To verify the identifications, representative voucher specimens were collected of all entities with a different field name. Duplicates of the Amazonian material are deposited in two herbaria in Brazil (SP: full set; INPA: fertile specimens) and one in Finland (TUR). Duplicates of the Panamanian material are deposited in two herbaria in Panama (PMA: full set; STRI) and one in Finland (TUR). The specimens were first sorted into morphospecies that were thought to correspond to biological species, and then assigned species names with the help of taxonomic literature and by comparisons with existing neotropical fern material at TUR. Some closely related species were lumped before analysis, because they were not consistently distinguished in the field. Unnamed morphospecies were treated as good species if they were found to be both monophyletic and genetically distinct from their nearest relative (details explained

below).

Geographical coordinates for each sampling unit were obtained using a hand-held GPS. Estimates of annual mean precipitation for all sites were extracted from the WorldClim database (Hijmans *et al*., 2005; http://www.worldclim.org/bioclim) at a resolution of 2.5 arc minutes using the DIVA-GIS software (Hijmans *et al*., 2001).

Composite surface soil samples were collected in each sampling unit to determine the concentration of extractable bases in the soil (Ca, Mg and K). The samples were air-dried and sieved through a 2-mm mesh before analysis. In Amazonia, three soil samples were taken along each transect, one close to each end and one in the middle (see Tuomisto *et al*., 2003a for further details on the methodology). The cations were extracted with 1M ammonium acetate at pH 7 at MTT Agrifood Research (Jokioinen, Finland). In Panama, two soil samples were taken at each site within a square 1-ha plot that contained both fern inventory transects (see Jones *et al*., 2013 for further details). The cations were extracted with the Mehlich-III method at the Smithsonian Tropical Research Institute soil laboratory (STRI, Panama City, Panama). Before numerical analyses, the cation concentrations of samples from the same site were averaged to obtain a single cation value for each sampling unit.

Because the MTT and STRI soil laboratories used different protocols for extracting the cations, 23 of the 110 Panamanian soil samples were used as calibration samples and analyzed in both laboratories. These samples were selected to represent the entire range of soil cation concentrations in the Panama data. A linear regression model was parameterized with the calibration samples to predict MTT-equivalent values of log-transformed soil cation concentration on the basis of the STRI-measured values (R^2=0.996). The regression equation was applied to those soil samples that were only analysed at STRI; for the 23 samples analysed at both laboratories, the measured values from MTT were used.

In most analyses, we used annual rainfall and log-transformed soil cation concentration as continuous variables. However, for those analyses requiring an environmental classification the soil cation concentration gradient was divided into four classes and the precipitation gradient into three classes (Table 2). Precipitation class limits were chosen to coincide with breaks in the distribution of the observed values in the data, whereas the log-transformed soil cation gradient was simply divided evenly into four classes. The poorest-soil class consisted of Amazonian sites only and the richest-soil class of Panamanian sites only, but the two intermediate classes contained sites from both areas. The rainfall classes were geographically less biased, although the wettest class was dominated by Panamanian sites.

DNA extraction and sequencing

For each species we extracted DNA from one to three specimens. We amplified and sequenced the plastid *rbcL* gene from every extraction for phylogenetic reconstruction and supplemented these data with some plastid *atpB*, *atpA*, and *rps4* sequences produced from these extractions for other purposes. Laboratory work followed established protocols (Lehtonen, 2011). Altogether, this resulted in 413 *rbcL*, 17 *atpB*, 5 *atpA*, and 30 *rps4* sequences (GenBank accession numbers KJ628500–KJ628963).

Phylogenetic analyses

To produce a phylogeny of as high a quality as possible, we used the large dataset analysed in Lehtonen (2011) and added to it both the new DNA sequences described above and sequences made

available in GenBank release 184 (June 15, 2011). Details on dataset construction are given in Lehtonen (2011).

We tested the monophyly of our species by performing an equally weighted parsimony analysis based on 500 ’new technology’ search replications in TNT 1.1 (Goloboff *et al*., 2008). Whenever a species was resolved as monophyletic, we discarded any duplicate samples and retained only one specimen, except when the same species was present in both Panama and Brazil. In this case, a specimen from each region was retained. If unnamed morphospecies were not diagnosable from their DNA sequences, they were lumped until they formed a genetically distinguishable monophyletic unit. We lacked DNA samples of six observed species. For five of these, we were able to use a sequence from Genbank, but one species that occurred in a single transect in central Amazonia had to be excluded from the analyses.

After removing duplicate samples, the final data set consisted of 3,137 operational terminal units (OTU), which will here be referred to as 'species' for simplicity. These represent about 25% of the global fern flora (Moran 2008) and 4,392 aligned base pairs of molecular data (*rbcL* 1,316 bp; *rps4* 379 bp; *atpB* 1,188 bp; *atpA* 1,509 bp). The final data set was analysed under the maximum likelihood (ML) criterion using parallel Pthreads-version of the computer program RAxML 7.3.0 (Stamakis, 2006; Ott *et al*., 2007). Search was initiated with 500 rapid bootstrap replications followed by a thorough ML search on the full data matrix (-T 16 -f a -x 12345 -p 12345 -# 500 -m GTRGAMMA). Free model parameters were estimated by RAxML under the GTR+Γ model.

The obtained ML phylogeny was pruned to retain only those species that occurred in our Panamanian and Amazonian field data using the command *sampleprune* in Phylocom (Webb *et al*., 2008). Pruning preserved the topological relationships and branch lengths of the original ML tree, but greatly reduced the number of terminal taxa and hence allowed time calibration. The pruned phylogeny was time calibrated with a penalized likelihood method in r8s 1.70 (Sanderson, 2003). The root node age was fixed to 354 Ma following Pryer *et al*. (2004), and the ages of a further 64 internal nodes were fixed according to their age estimates as given by Schuettpelz and Pryer (2009). The final ultrametric tree was used in all subsequent analyses.

Community structural analyses

Phylogenetic structure of the fern community in each sampling unit was quantified using three different metrics. Species richness is simply a count of the terminal taxa in the final ultrametric tree that were encountered in the sampling unit. Phylogenetic diversity (PD) is the total length of the branches connecting these terminal taxa (Faith, 1992), and mean phylogenetic diversity (MPD) is PD divided by species richness (Webb *et al*., 2002). Because we used a dated phylogeny, both PD and MPD are in practice measured in units of time (millions of years). In fact, MPD/2 can be interpreted as the average time that has passed since a randomly chosen species pair from the phylogeny diverged from a common ancestor. Each metric was calculated separately for each sampling unit using Phylocom version 4.2 (Webb *et al*., 2008).

We tested for two kinds of phylogenetic bias in the fern communities. Firstly, we tested whether particular clades were significantly over-represented or under-represented in different parts of the environmental gradients. For each sampling unit we generated a pool of 999 reference communities randomly drawn from the final ultrametric tree, each consisting of the number of species originally observed, using the ”nodesig” command in Phylocom (Webb *et al*., 2008). Then we tested for clade over- or under-representation in sampling units belonging to each of the four soil cation concentration classes and three rainfall classes. Because small clades have a high probability of

being absent from random communities, under-representation is in practice only detectable at deeper phylogenetic levels, even if it occurs throughout the phylogeny (Parra *et al*., 2010). For this reason, we focus mostly on over-representation.

The second phylogenetic bias we tested for was whether MPD in the sampling units differed from random expectation within each study region. This was done by randomizing species identities within each of the two regional species lists separately. Local communities were considered significantly clustered if their MPD values were within the lowest 2.5% of the values obtained for randomized communities, and significantly dispersed if the values were within the highest 2.5%.

Linear and second order polynomial regression analyses were run to test whether species richness, PD and MPD were significantly related to gradients in soil cation concentration or annual rainfall.

All analyses were done using species presence-absence data. The sensitivity of the analyses to phylogenetic resolution was tested by collapsing all intra-generic nodes in the phylogeny and repeating the analyses using genus-level resolution.

RESULTS

The total number of species encountered (gamma richness) was 106 in our Panamanian data and 66 in our central Amazonian data. The average number of genera per sampling unit was 7 for Panama and 6 for central Amazonia; on average, 1.5 species per genus were observed per sampling unit in both regions. The average number of species per sampling unit (alpha richness) was 10.5 for Panama and 8.9 for central Amazonia (Table 1). Consequently, beta richness (= gamma/alpha) was 10.1 for the Panamanian data and 7.4 for the central Amazonian data. The higher total number of species encountered in Panama therefore reflected both higher species density at the local scale and higher heterogeneity among the sampling units. The higher heterogeneity, in turn, was consistent with the fact that the number of sampling units was larger in Panama, and that both the rainfall gradient and especially the soil cation concentration gradient were longer (Table 1). The geographical extent of sampling, however, was far smaller in Panama than in central Amazonia (Fig. 1).

Analyses of both PD and MPD gave very similar results whether the phylogeny was resolved to the species level or to the genus level (Fig. 2). In the following, we focus on the species resolution results.

We found that PD was strongly correlated with species richness. Both species richness and PD had a hump-shaped response to soil cation concentration in Panama and when data from both regions were combined. In central Amazonia, the trends were not statistically significant (Fig. 3, Table 2). MPD decreased with increasing soil cation concentration in both Panama and central Amazonia, as well as in the combined data. Soil and rainfall gradients were not correlated with each other (correlation coefficient -0.12 in Panama and -0.06 in central Amazonia), hence their effects were modelled separately. All diversity metrics were better explained (higher R^2 value) by the regression model with soil cation concentration than by the one with annual rainfall (Table 2). In Panama, PD and MPD increased with annual rainfall, and when both areas were considered together, species richness and PD also increased with rainfall. Within central Amazonia, no significant trends with rainfall were observed.

Two central Amazonian sampling units had phylogenetically clustered fern communities, and these were the ones with the highest soil cation concentrations (Fig. 3). Similarly, phylogenetically

clustered communities in Panama were mostly on high-cation soils, wheres dispersed communities were on intermediate soils. Patterns in relation to rainfall were less clear, but phylogenetically clustered communities in Panama were absent from the wettest sites and phylogenetically dispersed communities from the driest ones. No phylogenetically dispersed communities were observed in central Amazonia.

Mean MPD of the fern communities in our sampling units averaged 311 Ma in Panama and 380 Ma in central Amazonia (Table 3). The difference (69 Ma) is significantly larger than expected by chance (49 Ma). An even larger difference was observed between the sampling units belonging to the cation-richest of the four edaphic classes (average MPD 284 Ma) and those belonging to the cation-poorest class (389 Ma). This difference (105 Ma) was more than twice as large as the difference expected by chance (49 Ma). In other words, species in Amazonian communities and on poor soils were found to be more distantly related than species in Panamanian communities and on cation-rich soils.

The nodal significance test revealed that some fern genera were concentrated to specific parts of the soil cation gradient (Fig. 4). *Lindsaea*, *Trichomanes* and especially *Triplophyllum* were over-represented in communities growing on the cation-poorest soils. *Adiantum* was common in all soil categories but had its highest representation on cation-rich soils, whereas *Thelypteris* and *Tectaria* were clearly over-represented in the cation-richest soil class. Species in the other lineages were more broadly distributed across the soil cation concentration gradient.

The relative species richness of poor-soil specialist genera (*Lindsaea*, *Trichomanes*, *Triplophyllum*) steadily decreased towards cation-rich soils whereas the relative species richness of rich-soil specialist genera (*Tectaria*, *Thelypteris*) increased (Fig. 5). The relative species richness of rich-soil specialist genera varied considerably even in the richest soils, whereas the poorest soils were largely dominated by poor-soil specialist genera. Relatively few biases were identified in relation to annual rainfall, with the clearest pattern at the genus level being the over-representation of *Thelypteris* in the intermediate rainfall class.

DISCUSSION

Impact of soil and rainfall gradients on phylogenetic community structure

We found significant and consistent effects of soil nutrient status on fern phylogenetic community structure. These results are in striking contrast with the view that soil niches are too labile to leave a detectable phylogenetic signal (Swenson, 2013). Within both regions, fern communities in richer soil habitats exhibited a higher degree of phylogenetic relatedness than those in poorer soil habitats. In Panama, phylogenetic relatedness tended to decline with increasing rainfall, although the effect was weaker than that of soil cation concentration. Significant phylogenetic turnover with rainfall has also been detected in Panamanian tree communities (Hardy *et al.*, 2012), but trends in the phylogenetic relatedness of trees have not been examined along either rainfall or edaphic gradients. We found no detectable effects of rainfall on fern phylogenetic community structure in central Amazonia. This difference between the two regions is not surprising, given that all of the central Amazonian sites were in evergreen forests, whereas some of the Panamanian sites were in semideciduous forests. It may also reflect the poorer quality of digital climate data in Amazonia, where meteorological stations are extremely sparse (Hijmans et al. 2005).

Eiserhardt *et al.* (2012) found that palm communities on extremely nutrient-poor white sand soils in Western Amazonia tended to be phylogenetically overdispersed when compared to the regional

species pool, whereas communities on richer upland and floodplain soils tended to be clustered. This result parallels the trend we observed in mean MPD, even though the MPD values of our individual poor-soil communities did not deviate from random expectation. Fine and Kembel (2011) found that the choice of the phylogenetic diversity index that was used determined whether phylogenetic overdispersion or random structure was found in the tree flora of white sand soils in Peruvian Amazonia. Kluge & Kessler (2011) investigated the phylogenetic structure of tropical fern communities along an elevational gradient in Costa Rica. They did not find any clear pattern, and hypothesized that communities may be structured more with respect to local habitat conditions than the elevational gradient itself. Our results provide support for the importance of local conditions.

Poor soils are inhabited by more distantly related species than rich soils

Our results from both regions are consistent with the idea that older lineages tend to be disproportionately represented in fern communities on poor soils relative to those in richer soils, as has been found for tropical trees in Peruvian Amazonia (Fine & Kembel, 2011). Furthermore, we observed that the lineages making up local communities are older, on average, in central Amazonia than in Panama. Because the poorest soils were found in central Amazonia and the richest soils in Panama, it is difficult to cleanly separate between the effects of soils vs. geographical location. Nevertheless, the fact that the difference in average lineage relatedness was clearly greater between the poorest and richest soils (105 Ma) than between the two study regions (69 Ma) suggests that soils may be more important than geographic region for the observed pattern.

The differences in MPD are largely driven by the basal phylogenetic position of the poor-soil specialist genera *Lindsaea* and *Trichomanes*, which were more consistently present in poor-soil sites than the derived poor-soil specialist genus *Triplophyllum* was. All three genera were relatively poorly represented in Panama, as has previously been noted for *Triplophyllum* (Prado & Moran, 2008). Whether this is caused by the lack of suitable soil conditions or other factors, such as historical dispersal limitation, remains to be clarified. Similarly, rich-soil specialist genera (*Tectaria*, *Thelypteris*) were very species poor in our central Amazonian sites, as has been observed in earlier inventories in the region (Tuomisto & Poulsen, 1996; Zuquim et al., 2009). This is in agreement with the generally low cation concentrations of central Amazonian soils. In conclusion, the difference between Panama and central Amazonia in the commonness of these genera seems partly attributable to different soil properties, which in turn are determined by the distinct geological histories of the two regions.

White sands have been suggested to represent ancestral habitat for Amazonian species (e.g. Kubitzki, 1989; Frasier *et al*., 2008), with much recent speciation occurring in younger and more fertile soils. However, in our phylogeny there is no evidence that crown-group diversification differs temporally between poor-soil and rich-soil specialist groups. Poor-soil specialists do tend to belong to older lineages, but the main diversification within these lineages appears no more ancient than that of the rich-soil specialists. Further studies that specifically address the timing of diversification, and also take into account taxa from other geographical areas, are needed to clarify this point.

Poor soils in both regions had relatively low fern species richness, as has been found in earlier studies (Tuomisto et al. 2014). Overall, the richness-soil fertility relationship was hump-shaped across all sites, and also within Panama. Low species richness in the richest soils in Panama may, however, be due to limited water availability, since these include shallow soils of low water-holding capacity on limestone, sandstone and other sedimentary rocks. Even the tree layer on these shallow soils is deciduous and consists of drought-adapted species in spite of relatively high annual rainfall

at some sites (Bohlman, 2010). It would therefore be interesting to repeat these analyses using actual measurements of plant-available moisture, rather than using annual rainfall as a surrogate.

Edaphic niche evolution and conservatism

Different fern genera were over-represented in different sections of the soil nutrient gradient in agreement with patterns reported previously (Tuomisto & Poulsen, 1996). Hence, for several genera our results are consistent with phylogenetic conservatism of edaphic niches. However, for at least as many genera the observed pattern was more consistent with niche evolution or adaptive radiation along the soil fertility gradient: a given lineage was represented by different species on different kinds of soils.

Although species turnover also occurs along a rainfall gradient in both regions (Jones *et al*., 2013, Zuquim *et al*., 2014), our results in relation to the rainfall gradient were less conclusive. Little evidence was found for either climatic niche conservatism or niche evolution within the rainfall range covered by our sampling. This perhaps reflects the fact that the rainfall gradient was relatively shorter than the soil fertility gradient when compared to the physiological tolerance of fern species.

In this respect it is noteworthy that the proportion of the local flora that belongs to poor-soil specialist genera closely followed soil cation concentration: sites in the cation-poorest quarter of the soil gradient were largely dominated by such species. Similarly, the proportion of species belonging to rich-soil specialist genera tended to increase with increasing soil cation concentration, but in this case the variation was much larger. In addition, rich-soil specialist genera never attained as high a dominance in the rich-soil sites as the poor-soil specialist genera did in the opposite end of the gradient. In other words, most fern species on very poor soils belong to specialist genera, whereas rich soils support communities with a mixture of rich-soil specialist genera and generalist genera. This indicates that poor-soil environments may be harder for ferns to cope with and adapt to. As a result, there may be a bias in the prevalence of niche shifts within lineages, with more lineages shifting from poor to rich soils than vice versa. This could explain the higher observed species diversity in rich soils despite the older age of poor soils in Amazonia (Tuomisto *et al*., 2014), and perhaps also the decrease in MPD towards rich soils observed in our data. Since the specialist genera recognised here are pantropical in their distribution, a more rigorous analysis of their niche evolution is needed at the global extent.

Phylogenetic resolution and community phylogenetic analyses

It has been suggested that highly resolved and accurate phylogenies are critical in studying the interface between community ecology and evolution (Kress *et al*., 2009). However, our analyses produced similar results with genus-resolution and full-resolution phylogenies. This is partly explained by the fact that we observed an average of only 1.5 congeneric species per community, and partly by the PD measure being dominated by deeper relationships, such that the short terminal branches make little difference to its value at the community level. Similarly, Fine & Kembel (2011) reported that phylogenies resolved to the family level gave very similar results to genus-level phylogenies in the case of tropical trees. Simulations have also shown that loss of terminal resolution in a phylogeny has little effect on measures of community phylogenetic diversity (Swenson, 2009). The implication is that future studies using PD-based measures can be conducted with well-resolved backbone phylogenies even when species-level information is not available.

CONCLUSION

Our results suggest that soil nutrient gradients are more important than rainfall gradients in defining the phylogenetic structure of fern communities within neotropical lowland forests. Several genera were concentrated to specific parts of the soil cation concentration gradient, which is consistent with the idea of phylogenetic niche conservatism. Two of the three poor-soil specialist genera were phylogenetically basal, and communities on poor soils were composed, on average, of more distantly related species than communities on rich soils were. The proportion of species belonging to genera that were soil specialists was higher in poor-soil communities than in rich-soil communities. Sites on the poorest quarter of the soil gradient were largely dominated by poor-soil specialist genera, in contrast to the richest quarter, where the majority of species belonged to generalist genera.

Acknowledgements

This study was financially supported by grants from the Finnish Cultural Foundation and Academy of Finland to S.L., H.T. and M.M.J., and from Oskar Öflunds Stiftelse to M.M.J. For scientific and logistic support we thank Kalle Ruokolainen, Flávia R.C. Costa, Richard Condit, Ben Turner, Salomon Aguilar, Rolando Pérez and many others at the Smithsonian Tropical Research Institute and Instituto Nacional de Pesquisas da Amazônia. We are grateful to our field assistants and thank the villages and private landowners for permission to work in their territory. Gabriel Moulatlet help to draw the map. The research and collection permits in Brazil were granted by the Brazilian National Council for Scientific and Technological Development and the Instituto Chico Mendes de Conservação da Biodiversidade and in Panama by the Autoridad Nacional del Ambiente. Willi Hennig Society is acknowledged for making TNT freely available.

References

Bohlman, S.A. (2010) Landscape patterns and environmental controls of deciduousness in forests of central Panama. *Global Ecology and Biogeography*, **19**, 376–385.

Doebeli, M. & Dieckmann, U. (2003) Speciation along environmental gradients. *Nature*, **421**, 259–264.

Eiserhardt, W.L., Svenning, J.-C., Borchsenius, F., Kristiansen, T. & Balslev, H. (2012) Separating environmental and geographical determinants of phylogenetic community structure in Amazonian palms (Arecaceae). *Botanical Journal of the Linnean Society*, **171**, 244–259.

Emerson, B.C. & Gillespie, R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology and Evolution*, **23**, 619–630.

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**, 1–10.

Fine, P.V.A., Daly, D.C., Muñoz, G.V., Mesones, I. & Cameron, K.M. (2005) The contribution of edaphic heterogeneity to the evolution and diversity of Burseraceae trees in the Western Amazon. *Evolution*, **59**, 1464–1478.

Fine, P.V.A. & Kembel, S.W. (2011) Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography*, **34**, 552–565.

Fine, P.V.A., Zapata, F., Daly, D.C., Mesones, I., Misiewicz, T.M., Cooper, H.F. & Barbosa, C.E.A. (2013) The importance of environmental heterogeneity and spatial distance in generating phylogeographic structure in edaphic specialist and generalist tree species of *Protium* (Burseraceae) across the Amazon Basin. *Journal of Biogeography*, **40**, 646-661.

Frasier, C.L., Albert, V.A. & Struwe, L. (2008) Amazonian lowland, white sand areas as ancestral regions for South American biodiversity: biogeographic and phylogenetic patterns in *Potalia*

(Angiospermae: Gentianaceae). *Organisms, Diversity & Evolution*, **8**, 44–57.
Goloboff, P.A., Farris, J.S. & Nixon, K.C. (2008) TNT, a free program for phylogenetic analysis. *Cladistics*, **24**, 774–786.
Hardy, O.J., Couteron, P., Munoz, F., Ramesh, B.R. & Pélissier, R. (2012) Phylogenetic turnover in tropical tree communities: impact of environmental filtering, biogeography and mesoclimatic niche conservatism. *Global Ecology and Biogeography*, **21**, 1007–1016.
Hijmans, R.J., Guarino, L., Cruz, M. & Rojas, E. (2001). Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter*, **127**, 15–19.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
Jones, M.M., Ferrier, S., Condit, R., Manion, G., Aguilar, S. & Pérez, R. (2013) Strong congruence in tree and fern community turnover in response to soil chemistry and climate in central Panama. *Journal of Ecology*, **101**, 506–516.
Kirby, M.X., Jones, D.S. & MacFadden, B.J. (2008) Lower Miocene stratigraphy along the Panama Canal and its bearing on the Central American peninsula. *PLoS ONE*, **3**, e2791.
Kluge, J. & Kessler, M. (2011) Phylogenetic diversity, trait diversity and niches: species assembly of ferns along a tropical elevational gradient. *Journal of Biogeography*, **38**, 394–405.
Kooyman, R., Rossetto, M., Cornwell, W. & Westoby, M. (2011) Phylogenetic tests of community assembly across regional to continental scales in tropical and subtropical rain forests. *Global Ecology and Biogeography*, **20**, 707–716.
Kress, W.J., Erickson, D.L., Jones, F.A., Swenson, N.G., Perez, R., Sanjur, O. & Bermingham, E. (2009) Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proceedings of the National Academy of Sciences, USA*, **106**, 18621–18626.
Kubitzki, K. (1989) Amazon lowland and Guayna highland – historical and ecological aspects of the development of their floras. *Amazoniana*, **11**, 1–12.
Lehtonen, S. (2011) Towards resolving the complete fern tree of life. *PLoS ONE*, **6**, e24851.
Moran, R.C. (2008). Diversity, biogeography, and floristics. *Biology and Evolution of Ferns and Lycophytes* (ed. by T.A. Ranker and C.H. Haufler), pp. 367–394. Cambridge University Press, Cambridge, UK.
Ott, M., Zola, J., Aluru, S. & Stamatakis, A. (2007) Large-scale maximum likelihood-based phylogenetic analysis on the IBM BlueGene/L. Proceedings of ACM/IEEE Supercomputing conference.
Parra, J.L, McGuire, J.A. & Graham, C.H. (2010) Incorporating clade identity in analysis of phylogenetic community structure: an example with hummingbirds. *The American Naturalist*, **176**, 573–587.
Phillips, O. L., Vargas, P. N., Monteagudo, A. L., Cruz, A. P., Zans, M. –E. C., Sánchez, W. G., Yli-Halla, M. & Rose, S. (2003) Habitat association among Amazonian tree species: a landscape-scale approach. *Journal of Ecology*, **91**, 757–775.
Prado, J. & Moran, R.C. (2008) Revision of the neotropical species of *Triplophyllum* (Tectariaceae) *Brittonia*, **60**, 103–130.
Pryer, K.M., E. Schuettpelz, P.G. Wolf, H. Schneider, A.R. Smith, & Cranfill, R. (2004) Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. *American Journal of Botany*, **91**, 1582–1598.
Punyasena, S.W., Eshel, G. & McElwain, J.C. (2008) The influence of climate on the spatial patterning of Neotropical plant families. *Journal of Biogeography*, **35**, 117–130.
Räsänen, M.E., Linna, A.M., Santos, J.C.R. & Negri, F.R. (1995) Late Miocene tidal deposits in the Amazonian Foreland Basin. *Science*, **269**, 386–390.
Sanderson, M. (2003) R8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*, **19**, 301–302.

Schreeg, L.A., Kress, W.J., Erickson, D.L. & Swenson, N.G. (2010) Phylogenetic analysis of local-scale tree soil associations in a lowland moist tropical forest. *PLoS ONE*, **5**, e13685.
Schuettpelz, E. & Pryer, K.M. (2009) Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. *Proceedings of the National Academy of Sciences, USA*, **106**, 11200–11205.
Stamatakis A. (2006) RAxML-V1-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.
Swenson, N.G. (2009) Phylogenetic resolution and quantifying the phylogenetic diversity and dispersion of communities. *PLoS ONE*, **4**, e4390.
Swenson, N.G. (2013) The assembly of tropical tree communities - the advances and shortcomings of phylogenetic and functional trait analyses. *Ecography*, **36**, 264–276.
ter Steege, H. (and 22 others). (2003) A spatial model of tree α-diversity and tree density for the Amazon. *Biodiversity and Conservation*, **12**, 2255–2277
ter Steege, H. (and 11 others). (2006) Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, **443**, 444–447.
Tuomisto, H. (2006) Edaphic niche differentiation among *Polybotrya* ferns in Western Amazonia: implications for coexistence and speciation. *Ecography*, **29**, 273–284.
Tuomisto, H. & Poulsen, A. D. (1996) Influence of edaphic specialization on pteridophyte distribution in neotropical rain forests. *Journal of Biogeography*, **23**, 283–293.
Tuomisto, H., Poulsen, A. D., Ruokolainen, K., Moran, R. C., Quintana, C., Celi, J. & Cañas, G. (2003a) Linking floristic patterns with soil heterogeneity and satellite imagery in Ecuadorian Amazonia. *Ecological Applications*, **13**, 352–371.
Tuomisto, H., Ruokolainen, K. & Yli-Halla, M. (2003b) Dispersal, environment, and floristic variation of Western Amazonian forests. *Science*, **299**, 241–244.
Tuomisto, H., Zuquim, G. & Cárdenas, G. (2014) Species richness and diversity along edaphic and climatic gradients in Amazonia. *Ecography*, DOI: 10.1111/ecog.00770.
Umaña, M.N., Norden, N., Cano, Á. & Stevenson, P.R. (2012) Determinants of Plant Community Assembly in a Mosaic of Landscape Units in Central Amazonia: Ecological and Phylogenetic Perspectives. *PLoS ONE*, **7**, e45199.
Vamosi, S.M., Heard, S.B., Vamosi, C. & Webb, C.O. (2009) Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular Ecology*, **18**, 572–592.
Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.
Webb, C.O., Ackerly, D.D. & Kembel, S.W. (2008) Phylocom: software for the analysis of phylogenetic community structure and character evolution. *Bioinformatics*, **24**, 2098–2100.
Wilf, P., Cúneo, N.R. Johnson, K.R., Hicks, J.F., Wing, S.L. & Obradovich, J.D. (2003) High plant diversity in Eocene South America: evidence from Patagonia. *Science*, **300**, 122–125.
Wing, S.L., Herrera, F., Jaramillo, C.A., Gómez-Navarro, C., Wilf, P. & Labandeira, C.C. (2009) Late Paleocene fossils from the Cerrejón Formation, Colombia, are the earliest record of Neotropical rainforest. *Proceeding of the National Academy of Sciences, USA*, **106**, 18627–18632.
Zuquim, G., Prado, J. & Costa, F.R.C. (2009). An annotated checklist of ferns and lycophytes from the Biological Reserve of Uatumã, an area with patches of rich-soils in Central Amazonia, Brazil. *Fern Gazette*, **18**, 286–306.
Zuquim, G., Tuomisto, H., Jones, M.M., Prado, J., Figueiredo, F.O.G., Moulatlet, G.M., Costa, F.R.C., Quesada, C.A. & Emilio, T. (2014) Predicting environmental gradients with fern species composition in Brazilian Amazonia. *Journal of Vegetation Science*, **25**, 1195–1207.

Biosketch

S.L., M.M.J., G.Z. and H.T. are members of the University of Turku Amazon Research Team

(http://www.utu.fi/amazon). All authors share an interest in tropical biodiversity. Author contributions: M.M.J. conceived the original idea, M.M.J., H.T., G.Z. and J.P. collected the data, S.L., H.T. and M.M.J. performed the data analyses, S.L. and M.M.J. led the writing with contribution from all others.

Table 1. General description of study sites. Soil cation concentrations for Panama have been adjusted to ensure comparability with values for central Amazonia (see Field sampling and environmental variables for details). Means±SD are given, with the total range in parentheses.

	Panama	Amazonia
Number of 0.1-ha plots	55	32
Total species richness	106	66
Species richness per plot	11±5 (3–22)	9±4 (2–18)
Genus richness per plot	7±3 (2–15)	6±2 (2–12)
Soil cation concentration (cmol[+]/kg)	23.03±18.50 (0.55–65.62)	0.30±0.43 (0.07–2.58)
Annual precipitation (mm)	2582±403 (1866–3421)	2338±261 (1827–2974)

Table 2. Explanatory power (adjusted R^2) of log-transformed soil cation concentration and total annual rainfall in linear and second-order polynomial regression models with fern species richness, phylogenetic diversity (PD) or mean phylogenetic diversity (MPD) as the dependent variable. When two values are shown, the first value refers to analyses at the species-level resolution and the second value to analyses at the genus-level resolution. Regressions were run both for the two geographical regions separately and for all data combined. Statistical significance: *** p<0.001; ** p<0.01; * p<0.05.

Dependent variable	Region	log(Soil cations) linear	log(Soil cations) quadratic	Annual rainfall linear	Annual rainfall quadratic
Richness	Amazonia	0.07 / 0.03	0.03 / 0.04	-0.03	-0.01 / -0.02
Richness	Panama	0.02 / 0.04	0.30***	0.07* / 0.16**	0.07 / 0.15
Richness	All	0.01	0.16*** / 0.20***	0.07** / 0.13**	0.06 / 0.12
PD	Amazonia	-0.01	0.03 / 0.05	-0.03	-0.04
PD	Panama	0.11**	0.35** / 0.36***	0.16** / 0.14**	0.14 / 0.12
PD	All	0.00	0.20**	0.09** / 0.07**	0.08 / 0.06
MPD	Amazonia	0.29*** / 0.28**	0.30**	-0.03	-0.04
MPD	Panama	0.35*** / 0.36***	0.38 / 0.39	0.06* / 0.05	0.07 /0.05
MPD	All	0.42*** / 0.43***	0.47** / 0.48**	-0.01	0.00

Table 3. Breakdown of the 87 sample plots by geographical region, soil cation concentration and annual rainfall, with the average values of fern species richness, phylogenetic diversity (PD), and mean phylogenetic diversity (MPD) shown for the plots belonging to each category. MPDrnd gives the random expectation for MPD generated by randomly shuffling species labels across the phylogeny 999 times. Statistical significance of departures of MPD from MPDrnd: *** $p<0.001$; ** $p<0.01$; * $p<0.05$. For geographical locations of the plots, see Fig. 1.

	Plots	Species	PD	MPD	MPDrnd
Central Amazonia	32	8.9	1281	380	384
Panama	55	10.5	1313	311**	335
Soil cation concentration					
0.07–0.28 cmol(+)/kg	27	8.6	1281	388	384
0.42–1.87 cmol(+)/kg	9	8.9	1281	365	356
2.58–11.47 cmol(+)/kg	15	13.5	1686	350	338
11.97–65.62 cmol(+)/kg	36	9.7	1160	284***	335
Annual rainfall					
1827–2024 mm	9	6.9	1098	355	357
2152–2668 mm	62	9.8	1254	331**	356
2840–3421 mm	16	12.3	1596	344	341

Figure 1. Locations of 55 fern inventory plots in Panama and 32 plots in central Amazonia (black triangles). Gray scale represents altitude according to Shuttle Radar Topographic Mission (SRTM) data.

Figure 2. Relationships between measures quantifying different aspects of phylogenetic community structure in fern communities in central Amazonia (32 sites) and Panama (55 sites). **A, C**. The phylogenetic measures PD (phylogenetic distance, panel A) and MPD (Mean Phylogenetic Diversity, panel C) plotted against species richness. **B, D**. The phylogenetic measures calculated from a genus-resolution phylogeny plotted against the same measure calculated from a full resolution phylogeny. Symbols are sized in proportion to log-transformed soil cation concentration. Pearson correlation coefficients (r) are shown for each comparison.

Figure 3. Species richness (**A-B**) and mean phylogenetic distance (**C-D**) of fern communities in central Amazonia (32 sites) and Panama (55 sites) plotted against soil cation concentration (**A, C**) and annual rainfall (**B, D**). Linear and second-order polynomial regression lines are shown if statistically significant at the $P<0.05$ level (dashed line shows the regression obtained when both regions are combined). R^2 values and P levels are in Table 2. Note the logarithmic scale of the soil cation axis. Triangles: phylogenetically dispersed communities (MPD larger than expected), circles: phylogenetically clustered communities (MPD smaller than expected), diamonds: MPD not different from random expectation.

Figure 4. A dated phylogeny of the taxa observed in 55 sites in Panama and 32 sites in central Amazonia. The topology and branch lengths are identical to those in a larger phylogeny containing 3,137 terminals. The coloured bars to the right of the phylogeny represent the four classes of soil cation concentration (from brown = most cation-poor to green = most cation-rich), and the presence of a given colour indicates that the species in question was observed at a site belonging to that class. Branches that were significantly over-represented in sites belonging to the lowest or highest soil cation concentration classes are labeled and highlighted in colour. Brown: lineages over-represented in the most cation-poor sites. Green: lineages over-represented in the most cation rich sites. Branch thickness indicates how often the lineage was significantly over-represented in the set of communities belonging to the soil class in question, as determined by randomly shuffling species labels across the phylogeny 999 times. For the cation concentration limits of the soil classes, see Table 3. Numbered lineages are: 1) *Danaea*, 2) *Schizaea* + *Lygodium*, 3) tree ferns, 4) *Saccoloma*, 5) *Pityrogramma* + *Pteris*, 6) *Adiantopsis*, 7) *Vittaria* + *Polytaenium* + *Anetium*, 8) Aspleniaceae, 9) *Diplazium*, 10) Blechnaceae, 11) Dryopteridaceae, 12) Lomariopsidaceae, 13) Nephrolepidaceae, 14) Polypodiaceae.

Figure 5. The percentage of species belonging to (A) genera over-represented in the cation-poorest quarter of the soil gradient (*Lindsaea*, *Trichomanes*, *Triplophyllum*) and (B) genera over-represented in the cation-richest quarter of the soil gradient (*Tectaria*, *Thelypteris*) out of total observed species richness in sites across the soil cation gradient.

Fig. 1.

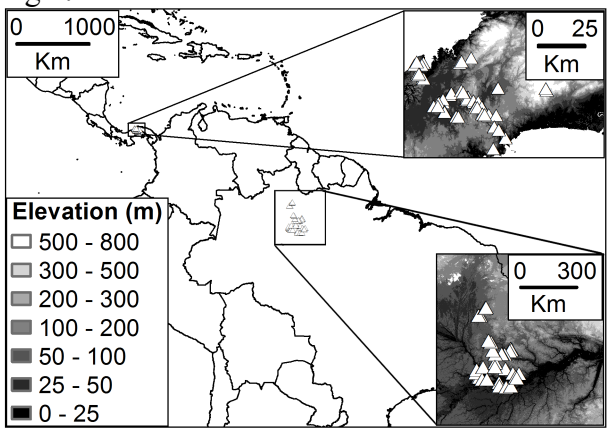


Fig. 2

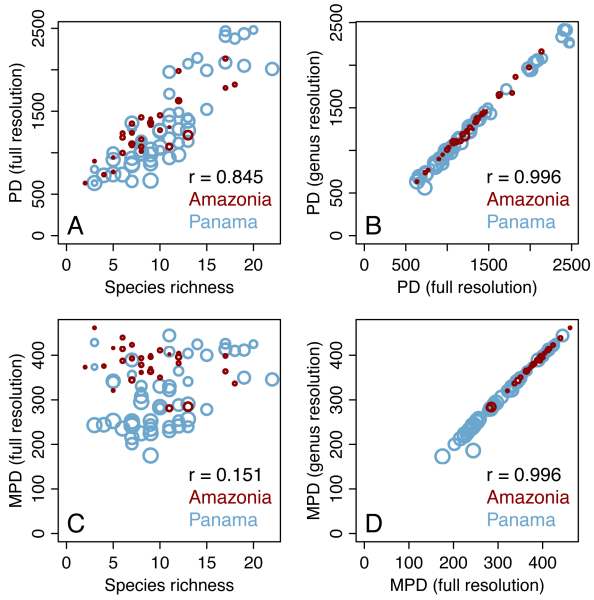


Fig. 3

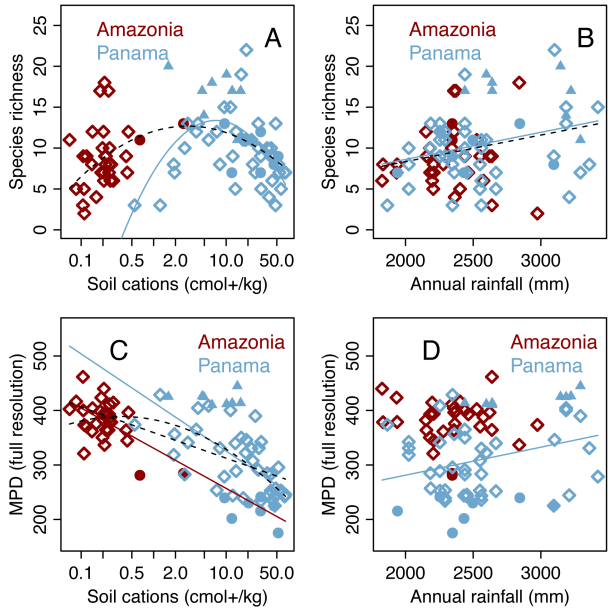


Fig. 4

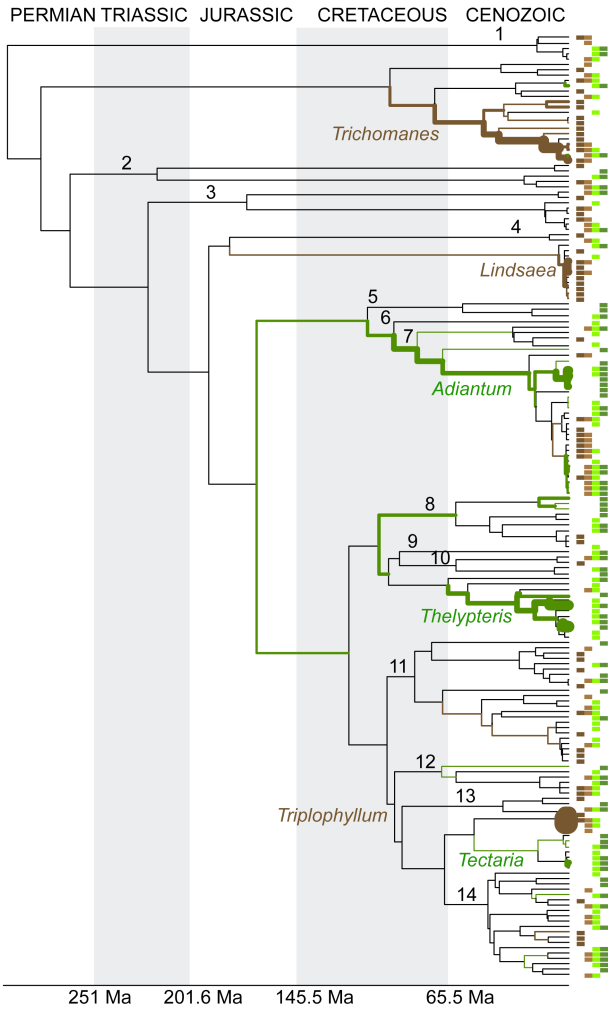
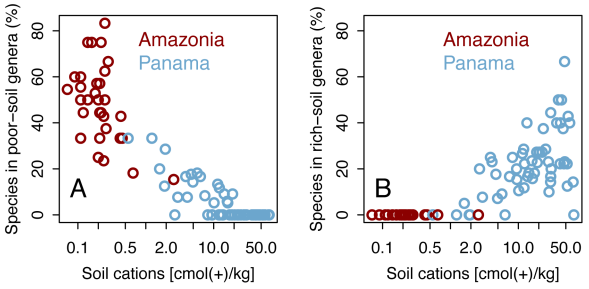

PERMIAN
TRIASSIC
JURASSIC
CRETACEOUS
CENOZOIC
1
Trichomanes
2
3
4
Lindsaea
5
6
7
Adiantum
8
9
10
Thelypteris
11
12
13
Triplophyllum
Tectaria
14
251 Ma
201.6 Ma
145.5 Ma
65.5 Ma

Fig. 5


Species in poor−soil genera (%)
0
20
40
60
80
Amazonia
Panama
A
0.1
0.5
2.0
10.0
50.0
Soil cations [cmol(+)/kg]
Species in rich−soil genera (%)
0
20
40
60
80
Amazonia
Panama
B
0.1
0.5
2.0
10.0
50.0
Soil cations [cmol(+)/kg]