Phylogenetic patterns of rarity in a regional species pool of tropical woody plants

M. Isabel Loza, Iván Jiménez, Peter M. Jørgensen, Gabriel Arellano, Manuel J. Macía, Vania W. Torrez, and Robert E. Ricklefs

M. Isabel Loza: Department of Biology, University of Missouri, St. Louis, MO, 63121, USA and Herbario Nacional de Bolivia, Campus Universitario Cota-Cota, calle 27, Correo Central Cajón Postal 10077, La Paz, Bolivia. belarivera23@gmail.com
Iván Jiménez: Center for Conservation and Sustainable Development, Missouri Botanical Garden, P.O. Box 299, St. Louis, Missouri 63166, USA; and Department of Biology, University of Missouri, St. Louis, MO, 63121, USA.
Ivan.Jimenez@mobot.org
Peter M. Jørgensen: Missouri Botanical Garden, P.O. Box 299, St. Louis, Missouri 63166, USA. peter.jorgensen@mobot.org
Gabriel Arellano: Center for Tropical Forest Science – ForestGEO, Smithsonian Tropical Research Institute.NMNH-MRC 166, West Loading Dock, 10th and Constitution Ave. NW, Washington, DC 20560, USA.
gabriel.arellano.torres@gmail.com
Manuel J. Macía: Departamento de Biología, Área de Botánica, Universidad Autónoma de Madrid, Calle Darwin 2, ES–28049 Madrid, Spain.
Manuel.macia@uam.es
Vania W.Torrez: Division of Plant Conservation and Population Biology, Department of Biology, University of Leuven, B-3001 Leuven, Belgium.torflorvania@gmail.com
Robert E. Ricklefs: Department of Biology, University of Missouri, St. Louis, MO, 63121, USA. ricklefs@umsl.edu

Running title: Phylogenetic patterns of rarity

Keywords: Andean floras, Bolivia, habitat breadth, geographical range size, local abundance, Madidi, rarity, phylogenetic conservatism, phylogenetic signal.

Send all correspondence to: M. Isabel Loza. belarivera23@gmail.com

Type of article: Research paper
Words in abstract: 294
Words in main body (from the Introduction through Biosketch): 5409
Words in references: 1354 (52 references)
Number of figures: 7
Number of tables: 0
Number of appendices: 3

ABSTRACT

Aim Rarity, which is believed to influence extinction risk, can be defined in terms of local abundance, geographic range size, and habitat breadth. Phylogenetic patterns in these attributes provide insight into the extent to which rarity and extinction risk are conserved during evolution, and the potential for species-level heritability. We evaluated phylogenetic signal (i.e., related species resembling each other more than species drawn at random) and evolutionary conservatism (similarity among related species exceeding that expected from a Brownian model of evolution) in three axes of rarity (local abundance, geographic range size, and habitat breadth) among species in a regional pool of tropical woody plants.

Location The Madidi region in Bolivia.

Time period 2001 to 2010.

Major taxa studied Lignophyta clade.

Methods We used a network of 48 1-ha forest plots and 442 0.1-ha forest plots to measure local abundance and habitat breadth of 1,700+ woody plant species (from 100+ plant families). We estimated geographic range size from occurrence records of individual species across the Neotropics. We characterized overall phylogenetic patterns of rarity using Blomberg's *K* and applied variance partitioning among taxonomic levels, as well as disparity analysis, to describe patterns of trait distribution at different depths in the phylogeny.

Results We found phylogenetic signal, but not evolutionary conservatism, in the three axes of rarity. The variance in rarity among supra-specific taxa, particularly families and genera, exceeded that calculated from random draws of species from the Madidi region. Phylogenetic signal, estimated by the proportion of variance among supra-specific taxonomic levels, varied between 23 and 36% for local abundance and geographic range size, and between 9 and 10% for habitat breadth.

Main conclusions The regional pool of woody plant species in Madidi exhibits phylogenetic signal in rarity that is consistent with biologically significant species-level heritability.

INTRODUCTION

Many authors have defined rarity in terms of the local abundance of individuals and geographic range size (see Gaston 1994 for a review), with habitat breadth often included as a third axis (Rabinowitz *et al.*, 1986; McGill 2011). Each of these three axes of rarity is a species-level trait that describes a key aspect of abundance and distribution thought to influence extinction risk, though at different scales: small local populations are vulnerable to demographic stochasticity and environmental variation (Lande 1993; Aitken *et al.*, 2008); species with small geographic ranges are particularly susceptible to adverse conditions occurring simultaneously across their entire extent (McKinney 1997; Gaston 2003; Harnik *et al.*, 2012); species that occupy fewer habitats are more vulnerable to environmental change (McKinney 1997; Biesmeijer *et al.*, 2006; Colles *et al.*, 2009).

Studies concerning the determinants of species abundance and distribution often focus on ecological factors, leaving aside the evolution of traits that may determine rarity. Phylogenetic patterns in local abundance, geographic range size, and habitat breadth provide insight into the evolution of traits that determine rarity and the factors that influence extinction risk (Jones *et al.*, 2005). Furthermore, the extent to which the axes of rarity are phylogenetically conserved or labile determines, in part, how extinction risk is distributed across a given phylogeny, and assesses the amount of uniquely shared evolutionary history under threat (Purvis 2008). Likewise, because phylogenetic signal includes species-level heritability of traits (*sensu* Housworth *et al.*, 2004), phylogenetic patterns of rarity indicate the extent to which rarity is heritable at the species-level (Jablonski 1987; 2008; Ricklefs & Latham 1992; Waldron 2007; Borregaard *et al.*, 2012).

Some previous analyses have shown that rarity is more similar among closely related species than expected by chance (Jablonski 1987; Ricklefs & Latham 1992; Jones *et al.*, 2005; Waldron 2007; Menken *et al.*, 2009; Leao *et al.*, 2014; Mouillot & Gaston 2009; Machac *et al.*, 2011), while others have supported the opposite conclusion (Gaston 2003; Ricklefs, 2010; 2011). These different outcomes may be explained in at least two ways, each deserving of further inquiry. First, phylogenetic patterns of rarity might be largely idiosyncratic, reflecting peculiarities of different study systems (Jones *et al.*, 2005). To examine this possibility, studies of phylogenetic patterns of rarity in major groups of organisms other than well-studied vertebrates as birds and carnivores, would be informative. We know of only three analyses of the phylogenetic structure of distribution and abundance in tropical plants (Ricklefs 2010; Leao *et al.*, 2014; Dexter & Chave 2016). Yet, tropical plants represent a sizable portion of diversity on Earth, and include many species with poor dispersal abilities and small geographic ranges (Gaston 2003; Ghalambor *et al.*, 2006). These characteristics might promote similarity in rarity among closely related species (Jones *et al.*, 2005), in which case phylogenetic patterns of rarity in tropical plants might differ from those in other organisms, such as birds, carnivores, or extra-tropical plants.

Second, most previous analyses have not been designed to examine general phylogenetic patterns of rarity, particularly when these patterns might vary with the phylogenetic or taxonomic level at which they are measured (Jones *et al.*, 2005; Machac *et al.*, 2011). Most have used metrics that quantify patterns of rarity across entire phylogenies, such as Blomberg's K (Blomberg *et al.*, 2003) or Pagel's λ (Pagel 1999). In addition, few studies have addressed the hierarchical distribution of rarity

within a phylogeny, as may be done using taxonomically nested analyses or node-by-node analyses (but see Jones *et al.*, 2005; Machac *et al.*, 2011). Combining these different metrics might reveal whether particular evolutionary patterns of rarity are common only at certain taxonomic levels or phylogenetic depths. Likewise, phylogenetic patterns of rarity might depend on the particular axis of rarity being examined. This possibility could be directly evaluated by estimating phylogenetic patterns in the three axes of rarity in the same study system(s). Yet, few empirical studies have done so (Ricklefs 2010; 2011).

Here, we tested the hypothesis that the three axes of rarity (local abundance, geographic range size, and habitat breadth) in tropical plants exhibit phylogenetic signal and evolutionary conservatism. We assessed phylogenetic signal at different levels of phylogenetic relationship using a whole-phylogeny metric as well as two approaches that describe patterns at different phylogenetic depths. We used an extensive network of forest plots (48 large plots and 442 small plots) located in the Madidi region of the Bolivian Andes, which encompasses extensive topographic and environmental variability and hosts rare and widespread species. We measured local abundance and habitat breadth of 1,700+ woody species; we also use occurrence records across the Neotropics to quantify geographic range size. We found phylogenetic signal in the three axes of rarity: closely related species had more similar population characteristics than expected from a null model that randomly assigned rarity values to species. These results suggest that extinction risk may be elevated in some clades of the Madidi species phylogeny.

METHODS

Field sampling

We focused on woody plants (including trees, shrubs, lianas, palms, and tree ferns) of the Madidi region of Bolivia, located in the northeastern slope of the Bolivian Andes. The Madidi region encloses a wide range of vegetation types and environmental conditions (see Fuentes 2005 for a review), which goes from 200−6000 m in elevation.

Between 2001 and 2010, 490 vegetation plots were established at 200−4500 m of elevation distributed in many tropical forest habitats (Fig. 1). Forty-eight large plots were 1-ha (100×100 m) in area and 442 small plots were 0.1-ha (100×10 m or 50×20 m) in area. In each large plot, all plant stems with a diameter at breast height (DBH, at 1.3 m above ground level) ≥10 cm were recorded and tagged. For the small plots, all stems with DBH≥2.5 cm were recorded, but not tagged. All plots were within closed-canopy mature forest with no sign of recent disturbance, and each plot was at least 250 m from the nearest other plot. Each stem was assigned to a morphospecies in the field. Vouchers of every morphospecies were deposited at multiple herbaria (including the Herbario Nacional de Bolivia in La Paz [LPB] and the Missouri Botanical Garden in St. Louis [MO]), and voucher information was digitized in Tropicos® (http://tropicos.org/).

Taxonomy and dated phylogeny

In 2011, we identified vouchers and undertook a comprehensive appraisal with the help of specialists to ensure that species names were applied consistently across all plots. The taxonomy followed the Angiosperm Phylogeny Website (APW, Stevens 2001 onwards) and Tropicos®. Our analyses included only individuals that were

determined to currently accepted species, and excluded morphospecies as well as stems determined only to supra-specific taxonomic ranks. Altogether the large plots contained 28,409 individuals representing 806 accepted species in 100 families, and the small plots contained 112,492 individuals representing 1,729 accepted species in 134 families. We excluded 692 morphospecies (29% of all species) from the small plots and 224 (22%) from the large plots, because they could not be assigned to named species. Additionally, 246 individuals from the large plots (1%) and 5,801 from the small plots (5%) were excluded because they remained unidentified.

A phylogeny resolved to the species level was not available for the species in the study region. We therefore constructed a phylogeny that included all extant Lignophyta families (angiosperms, gymnosperms, and ferns) as terminal taxa (Fig. S1 in appendix S1), based on tree R20120829 in Phylomatic version 3 (Webb & Donoghue 2005), revised following APW. We compiled published estimates of node ages from APW for 74% of the 426 nodes in the Lignophyta family-level phylogeny. Because single nodes were often assigned different ages by different studies (Table S2 in appendix S1), we explored the effect of uncertainty in node age estimates by conducting analyses using three ultrametric trees based on the minimum (youngest), median, and maximum (oldest) ages for each node. For each of these three alternatives, we obtained pseudo-chronograms using the BLADJ function of Phylocom version 4.2 (Webb *et al.*, 2011).

Measuring the three axes of rarity

For each species, we estimated local abundance, geographic range size, and habitat breadth (Tables S3−4 in appendix S2), using data from large and small plots separately because the species pools included in each type of plot did not fully overlap. Local abundance was calculated as the sum of all individuals divided by the number of all same-sized plots in which a species occurred. Geographic range size was calculated as the extent of occurrence (EOO), a measure of the spatial spread of the occurrences of a species. This is not a measure of the area over which a species actually occurs, but it is rather the geographic extent of the species (Gaston & Fuller 2009). We estimated EOO as the area of the minimum convex polygon enclosing all known occurrence points (following IUCN 2014, Joppa *et al.*, 2015). For each species, all georeferenced records of occurrence in the Neotropics were downloaded from Tropicos®. Because Tropicos houses the Madidi data set, the taxonomic match between vegetation plot data and specimen data is well curated. We excluded records for which the geographic coordinates of the specimen and the description of the collecting locality did not match the country of origin. Habitat breadth was quantified as the number of habitats in which a species was known to occur within the Madidi region, following standardized vegetation types delineated to represent major biotic responses to bioclimatic, geomorphological, and edaphic features in the study region (Navarro & Maldonado 2002, Fig. 1).

Because the distributions of local abundance, geographic range size, and habitat breadth showed marked positive skew (Table S5 in appendix S3), we log-transformed these variables to approximate normal distributions for further analysis. The three axes of rarity were weakly correlated, with Spearman's correlation coefficients less than 0.21 (Fig. S6 in Appendix S3).

Measuring phylogenetic patterns of rarity

We defined phylogenetic signal in each axis of rarity as the tendency of related species to resemble each other more than species drawn at random from the same phylogeny (Blomberg *et al.*,2003; Münkemüller *et al.*, 2012). The null model used to test this hypothesis, known as "tip randomization", randomly assigns species trait values across a given phylogeny, so that trait values have no memory of their ancestry and any similarity among species due to shared ancestry is eliminated. Following Losos (2008), we defined phylogenetic conservatism (and, conversely, lability) as more (less) phylogenetic trait clustering than expected from a Brownian (random evolution) null model of evolution (Felsenstein 1985). Using these definitions, phylogenetic signal is necessary but insufficient to demonstrate phylogenetic conservatism.

We described phylogenetic patterns of rarity by combining metrics from three approaches: (i) a whole-phylogeny metric (Blomberg's *K*, Blomberg *et al.*, 2003), (ii) an analysis designed to describe patterns at different phylogenetic depths (variance partitioning among taxonomic levels, Hadfield & Nakagawa 2010), and (iii) an analysis designed to describe patterns on a node-by-node basis (disparity analysis, Harmon *et al.*, 2003). The first approach tested for both phylogenetic signal and phylogenetic conservatism, while the others tested only for phylogenetic signal.

The first approach was based on Blomberg's *K*, a metric that is scaled by the expected *K* value under a Brownian model of evolution, so that $K<1$ ($K>1$) implies that relatives resemble each other less (more) than expected under a Brownian model of evolution (Blomberg *et al.*, 2003). We calculated *K* based on the family-level phylogeny described above, and random samples of one species per family. We draw 1,000 random samples to obtain a distribution of 1,000 observed *K* values. To test for phylogenetic signal and conservatism, we compared each of the 1,000 observed *K* values against 10,000 iterations of the tip randomization and Brownian null models. We emphasize that this first approach, based on Blomberg's *K*, quantifies phylogenetic patterns in rarity among species that belong to different plant families, but not among confamilial species.

Our second approach quantified patterns of rarity among species within genera and did not require a dated phylogeny. Rather it was based on partitioning variance among nested random effects representing hierarchical taxonomic levels (Hadfield & Nakagawa 2010). The proportion of variance explained by supra-specific taxonomic levels provides an estimate of "phylogenetic heritability" (*sensu* Housworth *et al.*, 2004) that is analogous to Pagel's λ (Hadfield & Nakagawa 2010). Following Prinzing *et al.* (2001), we tested for phylogenetic signal by comparing observed variation within hierarchical taxonomic levels to expected values according to a tip randomization null model. We calculated 95% confidence intervals for expected variation within each hierarchical taxonomic level as the interval between the 2.5 and 97.5 percentiles of 10,000 iterations of the null model. Then, we examined whether observed values fell within these confidence intervals. Additionally, to know which genera and families were more common or rarer than expected by the tip randomization null model (in terms of the three axes of rarity), for each genus and family we compared observed fitted values to the respective distribution of fitted values from the 10,000 null model iterations.

The third approach involved disparity analysis, which tested for phylogenetic signal at each node of the phylogeny, using all the species in the datasets. For this

analysis, we added polytomies for genera and species at the tips of the family-level phylogeny described above. Disparity is the average Euclidean distance in trait space between pairs of species within a clade (Harmon *et al.*, 2003). Relative disparity for a clade descending from a given node within a broader phylogeny is calculated by dividing its disparity by the average disparity across the whole phylogeny. Relative disparity values <1 imply that clades contain relatively little of the variation present across the phylogeny as a whole and, consequently most variation is found between clades (rather than within clades). Conversely, values >1 imply that most of the variation across the phylogeny is contained within clades. To test for phylogenetic signal, we calculated observed relative disparity at each node of the phylogeny, and compared it to expected values derived from 50,000 iterations of the tip randomization null model.

All the comparisons between distributions of observed and expected values were based on Holm-Bonferroni *p*-adjusted values to account for multiple comparisons. All analyses were performed in R version 3.0.2 (R Development Core Team, 2013).

RESULTS

The first approach we used to describe phylogenetic patterns of rarity, based on Blomberg's *K*, quantified phylogenetic patterns in rarity only among species that belong to different plant families. It did not reveal statistically significant phylogenetic signal or conservatism (Fig. 2 and Figs. S7−8 in Appendix S3). Observed Blomberg's *K* values were always indistinguishable from those generated by the tip randomization null model and, except in the case of local abundance in the large plots, always lower than the values generated by the Brownian model of evolution (Fig. 2 and Figs. S7−8).

Our second approach to describing phylogenetic patterns of rarity was variance partitioning among taxonomic levels, including confamilial and congeneric species. It revealed statistically significant phylogenetic signal in all three axes of rarity. Even though most of the variance in each of the axes of rarity was concentrated among species within genera, variance among supra-specific taxonomic levels always exceeded that produced by the tip randomization null model (Fig. 3). Phylogenetic heritability, estimated by the proportion of variance explained by supra-specific taxonomic levels, was 23−36% for local abundance and geographic range size, and 9−10% for habitat breadth (Fig. 3). For the three axes of rarity, the proportion of variance associated with a taxonomic family was always greater than expected by the tip randomization null model. The same was true for genera with the exception of habitat breadth in the large plots.

The comparisons between the observed fitted values and the expected fitted values revealed that both large and small plots included families and genera that were more common or rarer than expected by chance. For the large plots, only 4 % of the families and 9.6 % of the genera were rarer than expected by chance in at least in one of the three axes of rarity examined here (Table S9, Figs S10−11 in appendix S3). Conversely, for the small plots 32.8 % of the families and 22 % of the genera were found to be rarer than expected by chance on at least one of the rarity axes, and 2.2% of the families and 2.3 % of the genera were rarer than expected by chance on all three axes (Table S12, Figs S13−14 in appendix S3). Interestingly, most families in the asterid clade were rarer than expected by chance on at least one rarity axis (Figs.

4a−5a and Figs. S10−11 and S13−14), and some (e.g., Gesneriaceae, Campanulaceae, and Macgraviaceae) were rare on all three axes (Fig. 5a). Rarity within the asterid clade exhibits a strong pattern of phylogenetic signal in our study region.

Disparity analysis indicated that clades originating at various depths in the phylogeny exhibited phylogenetic signal in habitat breadth and geographic range size (disparity < 0.8; $p \leq 0.05$). For the large plots, habitat breadth exhibited phylogenetic signal in the clade formed by the families Boraginaceae, Lamiaceae, Bignoniaceae, and Verbenaceae (Fig. 6). For the small plots, geographic range size exhibited phylogenetic signal in several nodes of the Rosales clade (involving Rhamnaceae, Cannabaceae, Urticaceae, Moraceae, and Ulmaceae), in the nodes subtending Laurales and Magnoliales (including Monimiaceae, Hernandiaceae, Lauraceae, Siparunaceae, Myristicaceae, Magnoliaceae, and Annonaceae), and in the node subtending two monocot families Arecaceae and Poaceae (Fig. 7). These patterns were partly consistent with those revealed by the variance partitioning analysis: the clades that revealed phylogenetic signal with the disparity analysis (described above) also contained families that were more common than expected by chance with respect to geographic range size (Fig. 5) and rarer than expected by chance with respect to habitat breadth (Fig. 6), reaffirming that these clades exhibit phylogenetic signal.

Disparity analysis also revealed some clades that exhibit high phylogenetic lability in local abundance (Fig. 7). In particular, analyses based on the small plots showed that relative disparity was higher than expected by the tip randomization null model in the nodes subtending the clade formed by two monocot families (Arecaceae and Poaceae; disparity = 8.5; $p \leq 0.05$; Fig 7).

DISCUSSION

We examined phylogenetic patterns of rarity within a regional pool of tropical woody plants and found phylogenetic signal in local abundance, habitat breadth, and geographic range size, in the sense that variation among supra-specific taxa exceeded that expected by chance. Embedded in this overall pattern, a few clades exhibited high lability in local abundance. Overall, our results support the hypothesis that local abundance, geographic range, and habitat breadth are more similar among closely related species than expected by chance. Below, we explore implications of these findings and discuss caveats with respect to our analyses.

Detection of phylogenetic patterns of rarity

Null models are central to our ability to infer phylogenetic patterns of rarity. Metrics that compare rarity (local abundance, habitat breadth and geographic range size in this case) across all species in a phylogeny (e.g., Blomberg's *K*), or between sister species, are compared to values expected under various null models (Krasnov *et al.*, 2011; Machac *et al.*, 2011; Waldron 2007). However, to our knowledge this is the first use of an explicit null model (tip-randomization) to assess phylogenetic patterns of rarity based on variance partitioning across taxonomic levels. Previous studies have suggested that rarity is phylogenetically labile because it is most variable among species within genera (Gaston 2003; Ricklefs 2010, 2011), a result similar to this study. However, previous studies did not explicitly use null models to test the significance of phylogenetic signal or conservatism. Our variance partitioning analysis showed that the three axes of rarity were most variable among species within genera, but supra-specific taxonomic levels always explained more variance than expected by

the tip randomization null model (Fig. 3), indicating phylogenetic signal. Thus, our study highlights the usefulness of interpreting variance partitioning across taxonomic levels with explicit null models when one is testing for phylogenetic patterns of rarity.

While the analysis of variance across taxonomic levels detected phylogenetic signal (Fig. 3), the analysis based on Blomberg's *K* failed to detect phylogenetic signal or conservatism (Fig. 2). Two considerations might contribute to this difference. One is that the analysis based on Blomberg's *K* did not include all species. In particular, we generated 1,000 samples by randomly choosing one species per family and, based on each of those samples, measured Blomberg's *K* (see Methods). This, in turn, might have decreased the precision of estimates of Blomberg's *K*, which is thought to decrease with the degree of incomplete sampling in phylogenies (Münkemüller *et al.*, 2012). Moreover, incomplete sampling in our analysis was not random. Instead, we designed our analysis to include only one species per family. Thus, the analysis based in Blomberg's *K* could not detect phylogenetic signal in rarity among confamilial and congeneric species. Yet, the phylogenetic signal detected by the analysis of variance is largely due to similarity among confamilial species (Fig. 3).

A second potential issue is that, with whole-phylogeny metrics such as Blomberg's *K*, opposite patterns at different phylogenetic depths might obscure each other (Machac *et al.*, 2011). For example, analyses of local abundance in the small plots detected phylogenetic signal among confamilial and congeneric species (Fig. 3), but also high lability at deeper levels including a few clades corresponding to families or groups of families (Fig 7). This shows that phylogenetic patterns of rarity may depend on the phylogenetic scope over which they are measured (Jones *et al.*, 2005; Machac *et al.*, 2011).

Biological significance of phylogenetic patterns of rarity

We detected significant phylogenetic signal in rarity, but not phylogenetic conservatism as defined here relative to a Brownian model of evolution (Losos 2008). More detailed phylogenetic data might uncover conservatism in rarity among con-familial species, and this should be a goal for future studies. However, the level of phylogenetic signal detected here might be biologically significant in the context of species-level heritability (Jablonski 2008), even in the absence of phylogenetic conservatism. Absence of phylogenetic conservatism does not imply absence of species-level heritability. Indeed, species-level heritability can be defined as the tendency of closely related species to be more similar than expected by chance (Jablonski 2008), which matches the definition of "phylogenetic heritability" (Housworth *et al.*, 2004; Hadfield & Nakagawa 2010) as well as the definition of phylogenetic signal adopted here.

The processes underlying this pattern of phylogenetic signal are poorly understood. At least three hypotheses could explain the pattern. (i) Phylogenetic signal in rarity could follow phylogenetic signal in other biological variables such as body size, dispersal ability, habitat requirements, and other life history traits (Gaston 2003). It might also be partly determined by phylogenetic signal in vulnerability to negative-density dependence imposed by pathogens (Gilbert & Webb 2007), although these interactions could potentially evolve rapidly (Ricklefs 2010; 2011). (ii) Closely related species may share broad-scale geographic domains and, thus, temporal and spatial environmental templates that may largely determine species rarity (Mouillot &Gaston 2009; Machac *et al.*, 2011). (iii) Given directional trends between range size and

405 species age (Pigot *et al.*, 2012), clades characterized by high levels of recent
406 diversification would include a higher proportion of species with small ranges than
407 clades characterized by lower levels of recent diversification. Thus phylogenetic signal
408 in geographic range would emerge in regional species pools that include clades that
409 diversified extensively and recently as well as clades that did not. This third hypothesis
410 is consistent with a negative relationship between species diversity, on one hand, and
411 range size and abundance, on the other (Dexter & Chave 2016).
412 **Phylogenetic scope and patterns of rarity in tropical plants**
413 The strength of phylogenetic signal or conservatism may depend on the
414 phylogenetic scope used to describe patterns of rarity (see discussion above, Jones *et*
415 *al.*, 2005; Machac *et al.*, 2011). Thus, it is important to describe the phylogenetic
416 scope of studies focused on regional species pools so that the results are interpreted
417 accordingly, and proper comparisons established. The most basal nodes in the dated,
418 family-level phylogenies that we used reach into the Paleozoic, separating ferns from
419 seed plants and gymnosperms from angiosperms (Figs. 4a and 5a). However, the
420 regional pool of woody species in Madidi includes few non-angiosperm species, and it
421 is mostly composed of clades that diverged from each other no more than 110 Ma
422 (Figs. 4b and 5b), i.e., postdating the crown age of angiosperms according to most
423 estimates (Magallón *et al.*, 2015). Point estimates for the age of the most derived nodes
424 in the phylogenies that we used in this analysis are largely >20 Ma (Figs. 4b and 5b),
425 reflecting the fact that divergence of angiosperm families is generally thought to
426 predate the Neogene (Magallón *et al.*, 2015). Nonetheless, we detected phylogenetic
427 signal due to similarity of confamilial species of unknown age. In lieu of information
428 on the ages of many angiosperm families, current appraisals of the origin of
429 Neotropical plant diversity suggest that a large fraction of the extant species diversity,
430 and perhaps genus diversity, arose during the Mid-to-Late Miocene or more recently
431 (Hughes *et al.*, 2013), including the Pliocene and Pleistocene in the case of Andean
432 species (Hughes & Atchison 2015). Hence, our findings likely describe phylogenetic
433 patterns of rarity that range mostly from the Late Cretaceous through the Neogene.
434 We are aware of only three previous studies of phylogenetic patterns of rarity
435 in tropical plants that span a phylogenetic scope similar to that of our study. These
436 studies focused on one or two axes of rarity (as opposed to three in our study). Leao et
437 al., (2014) found evidence of phylogenetic signal in geographic range size measured as
438 extent of occurrence and area of occupancy among 6,929 species from the Atlantic
439 forest of Brazil, representing 112 angiosperm families (Pagel's λ 0.54, 0.57
440 respectively). Dexter & Chave (2016) demonstrated phylogenetic signal in geographic
441 range size and also local abundance among 631 Amazonian angiosperm tree genera
442 (Pagel's λ 0.37, 0.32 respectively). The third study employed variance partitioning
443 across taxonomic levels to examine phylogenetic signal in local abundance among 250
444 tree species representing 53 angiosperm families in a 50-ha plot in Panama (Ricklefs
445 2010). Although this latter study did not use an explicit null model to test the
446 significance of phylogenetic signal (see above discussion), it found that 32% of the
447 variance occurred among supra-specific taxonomic levels. Further studies are needed
448 before generalizations about phylogenetic patterns of rarity can be drawn for regional
449 pools of tropical plant species. Nonetheless, the studies currently available suggest that
450 regional pools of tropical plant species might often be characterized by phylogenetic
451 signal in rarity.

Implications for conservation

Some evidence (e.g., for mammals, crinoids, and foraminifera) suggests that species distributed across more types of habitats and that tolerate wider ranges of environmental conditions have higher persistence (Colles et al 2009). Also, geographic range size, followed by habitat breadth, appears to play an important role in the survival of marine animal species (Harnik *et al.*, 2012; Saupe *et al.*, 2015); extant species have significantly larger geographic ranges than extinct species (Saupe *et al.*, 2015). In this sense, rarity has been associated with differential persistence of species during times of background levels of extinction as well as during mass extinction events (Jablonski 1986; 2008; Colles *et al.*, 2009; Harnik *et al.*, 2014; Saupe *et al.*, 2015). Thus, the pattern of phylogenetic signal described here (Fig. 3 and Fig. 5−7) may partly determine the outcomes of biotic impacts of ongoing climate change on Andean floras. Temperature in the study region has increased by 0.2−0.3 °C per decade during the last 30 years, and these rates increase at higher elevations (Vuille & Bradley 2000). Species with small population sizes, narrow habitat breadths, and restricted geographic distributions are likely to be more threatened by climate change than species with larger populations and broader habitat breadths and geographic ranges (Aitken et al., 2008). If that were the case, then our results showing phylogenetic signal in rarity would suggest that extinction risk from climate change would be phylogenetically clustered across Madidi woody plants. Accordingly, some clades, such as Asterids (Figs. 4a−5a and Figs. S10−11 and S13−14), would have a particularly high proportion of threatened species, so the amount of evolutionary history under threat could be higher than expected by chance (Purvis 2008).

Caveats

Our analyses might have been affected by the exclusion of morphospecies (species without a specific epithet) from the analysis. Morphospecies accounted for 29% and 22% of all named plus unnamed species in the 0.1 ha and 1 ha plots, respectively. The potential impact of this exclusion depends on the distribution of morphospecies across the phylogeny. If the morphospecies were randomly distributed across the phylogeny, then we would not expect the analysis to be strongly biased. Alternatively, if morphospecies were not randomly distributed across the phylogeny, and if they tended to have low (or high) values of rarity, then the potential for bias in our analysis would be higher. Additionally, we excluded from the analysis individuals that were not determined to the species level. Some of these individuals may have belonged to species included in the analysis, in which case, habitat breadth and local abundance might have been underestimated for some species. Geographic range size was estimated across the Neotropics and therefore would be less likely to be affected by these exclusions. Regardless, the proportion of unidentified individuals was relatively small (1% and 5% for 1-ha and small plots, respectively) and likely had only a minor effect on our analyses.

We measured geographic range size as the extent of occurrence (EOO), calculated as the area encompassed by the minimum convex polygon enclosing all known occurrences of a species. This measure is affected by incomplete botanical sampling, even though it may be relatively robust to this kind of sampling error (Joppa *et al.*, 2015). The effect of this error on estimates of phylogenetic signal depends on how it is distributed across the phylogeny. Under the reasonable assumption that error in measures of EOO is random relative to phylogenetic relationships, error in estimates

of EOO would introduce negative bias in estimates of phylogenetic signal (Housworth *et al.*, 2004). Thus, phylogenetic signal in geographic range size would be stronger than we documented here. Deviations from the assumption above could introduce negative or positive bias in estimates of phylogenetic signal, depending on the particular relationship between measurement error and the phylogeny.

Phylogenetic patterns of rarity might depend on the spatial scale at which they are examined (Mouillot & Gaston 2009; Machac *et al.*, 2011; Krasnov *et al.*, 2011). It is therefore important to highlight that we measured habitat breadth and local abundance at a regional scale (across the Madidi region), and geographic range size at a continental scale (across the Neotropics). It is difficult to predict how our results would change if habitat breadth and local abundance were measured at a continental scale. Krasnov *et al.*, (2011) suggested that phylogenetic patterns of habitat breadth and abundance are stronger when measured at larger spatial extents. If, so, we would expect that phylogenetic signal in habitat breadth and abundance measured at a continental scale would be even higher than we documented here.

ACKNOWLEDGEMENTS

We are very grateful to Alfredo Fuentes and Leslie Cayola, who coordinate most of the work in Bolivia, and to Alejandro Araujo-Murakami, Javier Quisbert, Maritza Cornejo, Tatiana Miranda, Renate Seidel, Narel Paniagua, and Carla Maldonado, who provided valuable plot data for this study. We also thank all the taxonomic experts, students, and local guides involved in the collection of the field data and in the identification of plant specimens. Peter Stevens provided valuable advice for constructing phylogenies and conducting the analysis, as well as comments on the manuscript. We also thank H. ter Steege, B. Enquist, B. McGill, M. Dornelas, R. Field and an anonymous reviewer for helpful comments on the manuscript. We thank Dirección General de Biodiversidad, the Bolivian Park Service (SERNAP), the Madidi National Park and local communities for permits, access and collaboration in Bolivia. This study was funded by the National Science Foundation (DEB 0101775, DEB 0743457), the Missouri Botanical Garden, the National Geographic Society, the Comunidad de Madrid (Spain), Universidad Autónoma de Madrid, and the Taylor and Davidson families.

Appendix S1: Dated global family phylogeny source information
Appendix S2: Dataset source information
Appendix S3: Supplementary results

BIOSKETCH

M. Isabel Loza is a graduate student at the University of Missouri-Saint Louis. She is part of the Madidi Project, a long-term collaborative project to document patterns in plant diversity and distribution in the Bolivian Andes (www.mobot.org/madidi/). This work was part of her MSc. dissertation. She is interested in the study of community assembly, and of rarity patterns and their determinants in the Neotropics.

REFERENCES

1. Aitken, S.N., Yeaman, S., Holliday, J.A., Wang, T. & Curtis-McLane, S. (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, 1(1), 95–111.
2. Biesmeijer, J., Roberts, C., Reemer, S. P. M., Ohlemuller, M. R., Edwards, M. & Peeters, T. (2006) Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science*, 313, 351–354.
3. Blomberg, S. P., Garland, T. & Ives, A. T. (2003) Testing for phylogenetic signal in comparative data, behavioral traits are more labile. *Evolution*, 57, 717–745.
4. Borregaard, M. K., Gotelli, J. G. & Rahbek, C. (2012) Are range size distributions consistent with species–level heritability? *Evolution*, 66, 2216–2226.
5. Colles, A., Liouw, L. H. & Prinzing, A. (2009) Are specialists at risk under environmental change? Neoecological, paleoecological and phylogenetic approaches. *Ecology Letters*, 12, 849–863.
6. Dexter, K. & Chave, J. (2016) Evolutionary patterns of range size, abundance and species richness in Amazonian angiosperm trees. *PeerJ*, 4:e2402;DOI 10.7717/peerj.2402
7. Fuentes, A. (2005) Una introducción a la vegetación de la región de Madidi. *Ecología en Bolivia*, 40,1–31
8. Gaston, K. J. (1994) Rarity. Chapman and Hall, London. 207 pp.
9. Gaston, K. J. (2003) The structure and dynamics of geographic ranges. Oxford University Press, Oxford, UK. 266 pp.
10. Gaston, K. J. & Fuller, R. A. (2009) The sizes of species' geographic ranges. *Journal of Applied Ecology*, 46, 1–9.
11. Ghalambor, C. K., Huey, R. B., Martin, P. R., Tewksbury, J. J., & Wang, G. (2006). Are mountain passes higher in the tropics? Janzen's hypothesis revisited. *Integrative and Comparative Biology*, *46*, 5–17.
12. Gilbert, G. S. & Webb, C. O. (2007) Phylogenetic signal in plant pathogen–host range. *Proceedings of the National Academy of Sciences*, 104, 4979–4983
13. Hadfield, J. D. & Nakagawa S. (2010) General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *Journal of Evolutionary Biology*, 23, 494–508.
14. Harmon, L. J., Schulte, J. A., Larson, A. & Losos, J. B. (2003) Tempo and mode of evolutionary radiation in iguanian lizards. *Science*, 301, 961–964.
15. Harnik, P. G., Simpson, C. & Payne, J. L. (2012) Long-term differences in extinction risk among the seven forms of rarity. *Proceedings of The Royal Society*, 279, 4969–4976.
16. Harnik, P. G., Fitzgerald, P. C., Payne, J. L. & Carlson S. J. (2014) Phylogenetic signal in extinction selectivity in Devonian terebratulide brachiopods. *Paleobiology*, 40, 675–692.
17. Housworth, E.A., Martins, E.P. & Lynch, M. 2004. The phylogenetic mixed model. *American Naturalist*. 163: 84–96.

18. Hughes, C. E., Pennington, R. T., & Antonelli, A. (2013). Neotropical plant evolution: assembling the big picture. *Botanical Journal of the Linnean Society*, *171*(1), 1-18
19. Hughes, C. E., & Atchison, G. W. (2015). The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytologist*, 207, 275-282
20. Felsenstein, J. (1985) Phylogenies and the comparative method. *The American Naturalist* 125, 1–15.
21. IUCN Standards and Petitions Working Group (2014) Guidelines for using the IUCN Red List categories and criteria. Version 11.IUCN SSC Biodiversity Assessments Subcommittee, Gland, Switzerland.
22. Jablonski, D. (1986) Background and mass extinctions, the alternation of macroevolutionary regimes. *Science,* 231, 129–133.
23. Jablonski, D. (1987) Heritability at the species level, analysis of geographic ranges of Cretaceous mollusks. *Science,* 238, 360–363.
24. Jablonski, D. (2008) Species selection, theory and data. *Annual Review of Ecology and Systematics*, 39, 501–524.
25. Jones, K. E., Sechrest, W., & Gittleman, J. L. (2005) Age and area revisited, identifying global patterns and implications for conservation. *Phylogeny and Conservation* (eds A. Purvis, J. L. Gittleman & T. Brooks), pp. 141–165, Cambridge University Press, Cambridge.
26. Joppa, L. N., Butchart, S. H., Hoffmann, M. M., Bachman, S. P., Akçakaya, H. R., Moat, J. F., Böhm, M., Holland, R.A., Newton, A., Polidoro, B., & Hughes A. (2015) Impact of alternative metrics on estimates of extent of occurrence for extinction risk assessment. *Conservation Biology* 00, 1–9
27. Krasnov, R., Poulin, R. & Mouillot, D. (2011) Scale-dependence of phylogenetic signal in ecological traits of ectoparasites. *Ecography,* 34, 114–122.
28. Lande, R. (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *The American Naturalist*, 142, 911–927
29. Leao, T. C., Fonseca, C. R., Peres, C. A. & Tabarelli, M. (2014) Predicting extinction risk of Brazilian Atlantic Forest Angiosperms. Conservation Biology, 28:1349–1359
30. Losos, J. B. (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters,* 11, 995–1003.
31. Machac, A., Zrzavy, J. & Storch, D. (2011) Range size heritability in carnivore is driven by geographic constraints. *The American Naturalist*, 177, 767–779.
32. Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, LL. & Hernández-Hernández, T. (2015). A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist* 207: 437 − 453.
33. McGill, B. J. (2011). Linking biodiversity patterns by autocorrelated random sampling. *American journal of botany*, 98(3), 451-502.
34. McKinney, M. L. (1997) Extinction vulnerability and selectivity, combining ecological and paleontological views. *Annual Review of Ecology and Systematics,* 28, 495–516.

35. Menken, S. B., Boomsma, J. J. & van Nieukerken, E. J. (2009) Large-scale evolutionary patterns of host plant associations in the Lepidoptera. *Evolution*, 64, 1098–1119.
36. Mouillot, D. & Gaston, K. (2009) Spatial overlap enhances geographic range size conservatism. *Ecography*, 32, 672–675
37. Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombar, T., Schiffers, K. & Thuiller, W. (2012) How to measure and test phylogenetic signal. *Methods in Ecology and Evolution*, 3, 743–756.
38. Navarro, G. & Maldonado, M. (2002) Vegetación y unidades biogeográficas. *Geografía Ecológica de Bolivia, Vegetación y Ambientes Acuáticos* (eds. G. Navarro & H. Maldonado), pp. 200–220, Centro de Ecología Difusión Simón I. Patiño. Santa Cruz Bolivia.
39. Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature* 401:877–884
40. Pigot, A. L., Owens, I. P. & Orme, C.D.L. (2012) Speciation and extinction drive the appearance of directional range size evolution in phylogenies and the fossil record. *PLoS Biol*, *10*(2), p.e1001260.
41. Prinzing, A., Durka, W., Klotz S. & Brandl, R. (2001) The niche of higher plants: evidence for phylogenetic conservatism. *Proccedings of the Royal Society of London B*, 268: 2383–2389.
42. Purvis, A. (2008) Phylogenetic approaches to the study of extinction. *Annual Review of Ecology and Systematics*, 39:301–319.
43. Rabinowitz, D., Cairns, S. & Dillon, T. (1986) Seven forms of rarity and their frequency in the flora of the British Isles. *Conservation biology, the science of scarcity and diversity* (ed. M.E. Soulé), pp 182–204, Sunderland Massachusetts.
44. Ricklefs, R. E., and R. E. Latham. (1992) Intercontinental correlation of geographical ranges suggests stasis in ecological traits of relict genera of temperate perennial herbs. *American Naturalist* 139: 1305-1321.
45. Ricklefs, R. E. (2010) Evolutionary diversification, coevolution between populations and their antagonists, and the filling of niche space. *Proceedings of the National Academy of Sciences USA*, 107, 1265–1272.
46. Ricklefs, R. E. (2011) Applying a regional community concept to forest birds of eastern North America. *Proceedings of the National Academy of Sciences USA*, 108, 2300–2305.
47. Saupe, E. E., Qiao, H., Hendricks, J. R., Portell, R. W., Sthepen, J. H., Soberón, J. & Liberman, B. S. (2015) Niche breadth and geographic range size as determinants of species survival on geological time scales. *Global Ecology and Biogeography*, 24,1159-1169
48. Stevens, P. F. (2001 onwards). Angiosperm Phylogeny Website. Version 12, July 2012 [and more or less continuously updated since]." http://www.mobot.org/MOBOT/research/APweb/.
49. Vuille, M. & Bradley, R. (2000) Mean annual temperature trends and their vertical structure in the tropical Andes. *Geophysical Research Letters*, 27, 3885–3888.
50. Waldron A. (2007) Null models of geographic range size evolution reaffirm its heritability. *The American Naturalist*, 170, 221–231.

51. Webb, C. & Donoghue, M. J. (2005) Phylomatic, tree assembly for applied phylogenetics. *Molecular Ecology Notes,* 5, 181–183.
52. Webb, C., Ackerly, D. & Kembel, S. (2011) Phylocom, Software for the analysis of phylogenetic community structure and character evolution (with phylomatic and ecovolve). pp. 38.

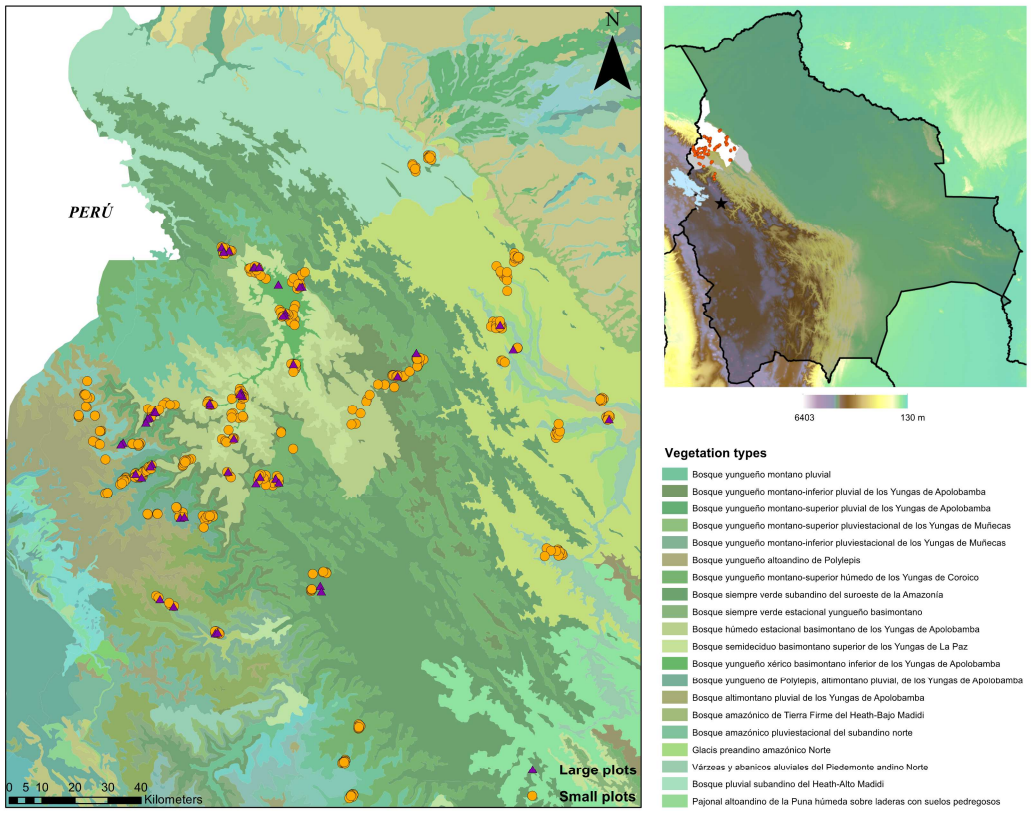

Figure 1. Geographic distribution of plots in the Madidi Region of Bolivia. The map on the rigth shows the vegetation types according to Navarro & Maldonado (2002), large plots (1.0 ha, triangles) and small plots (0.1 ha, circles). The map on the top right shows the elevation gradient for Bolivia, the Madidi region in white and the plots in red.

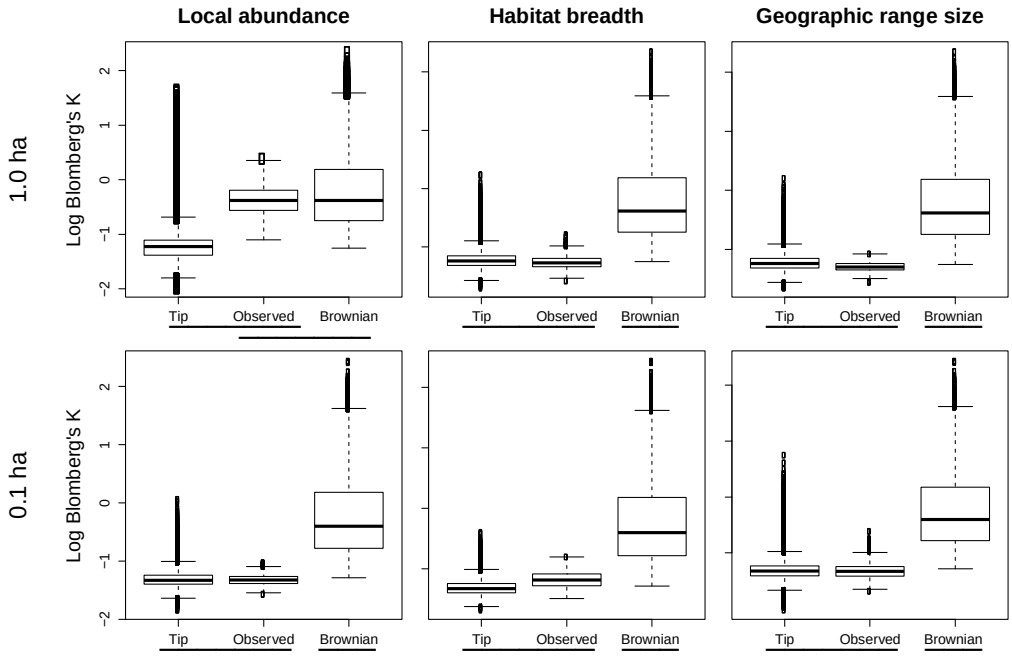

Figure 2. Distribution of Blomberg's *K* values for observed data and the two null models (Tip randomization and Brownian motion) using the ultrametric tree based on the median ages for the large (1.0 ha) and small plots (0.1 ha). Boxplots show the median (solid line), the interquartile range (box) and whiskers extending to the most extreme values within 1.5 × interquartile ranges from the box. The distribution of observed values was generated by repeatedly sub-sampling the species pool to obtain one species per family. The horizontal lines at the bottom join observed and null model distributions when they do not differ statistically (using adjusted $p \leq 0.05$).

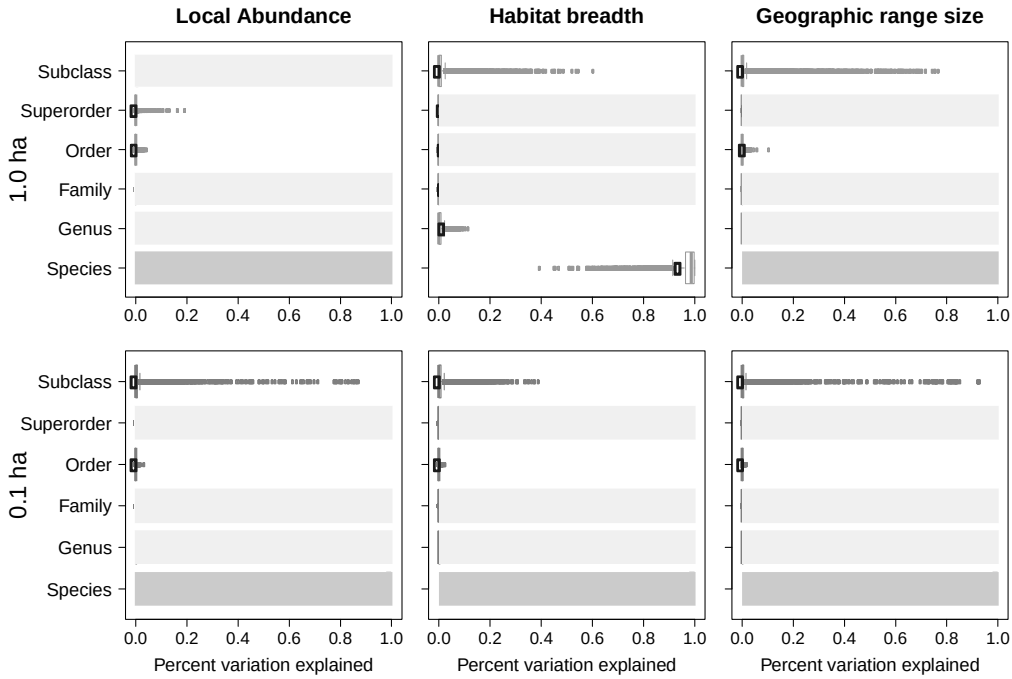

Figure 3. Percent of variation in rarity axes explained by different taxonomic levels according to a variance partitioning analysis for the large (1.0 ha) and small plots (0.1 ha). Large circles represent observed values and boxplots the distribution of values generated by the tip randomization null model. Boxplots show the median (solid line), the interquartile range (box) and whiskers extending to the most extreme values within 1.5 × interquartile ranges from the box. Dark and light gray backgrounds indicate that observed values fall below or above the 95% confidence interval for the null model, respectively.

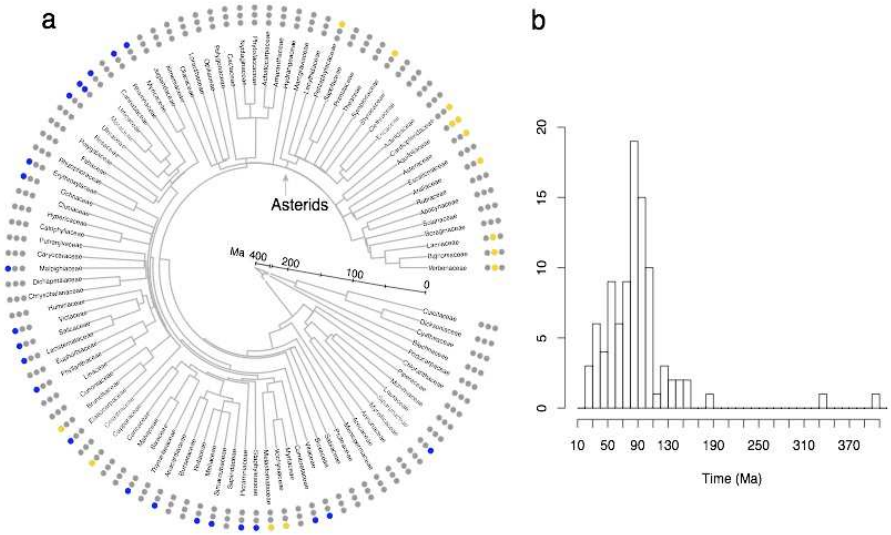

Figure 4. **a.** Phylogeny calibrated at the family level for large plots (1.0 ha), showing which families are more common (blue dots) or rare (yellow dots) than expected by chance. The closest ring to the phylogeny represents local abundance, the next habitat breadth and the third geographic range size. **b.** Histogram showing the frequency of node ages in the phylogeny in relationship with the evolutionary time.

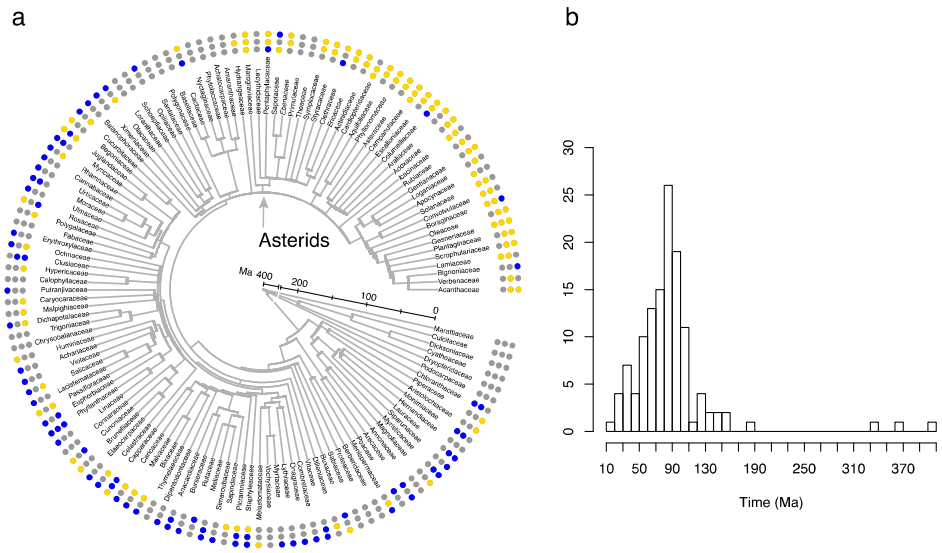

Figure 5. **a.** Phylogeny calibrated at the family level for small plots (0.1 ha), showing which families are more common (blue dots) or rare (yellow dots) than expected by chance. The closest ring to the phylogeny represents local abundance, the next habitat breadth and the third geographic range size. **b.** Histogram showing the frequency of node ages in the phylogeny in relationship with the evolutionary time.

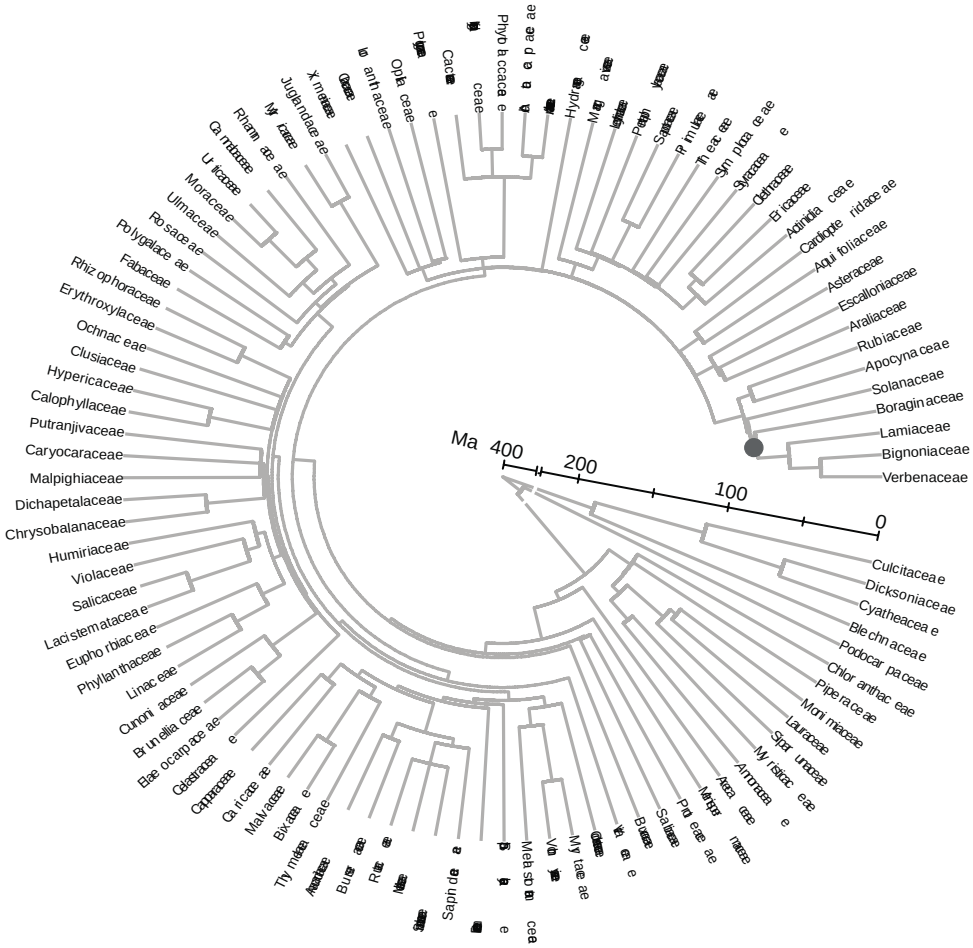

Figure 6. Phylogeny calibrated at the family level for the large plots (1.0 ha), showing a node with phylogenetic signal for habitat breadth (grey dot) according to disparity analysis.

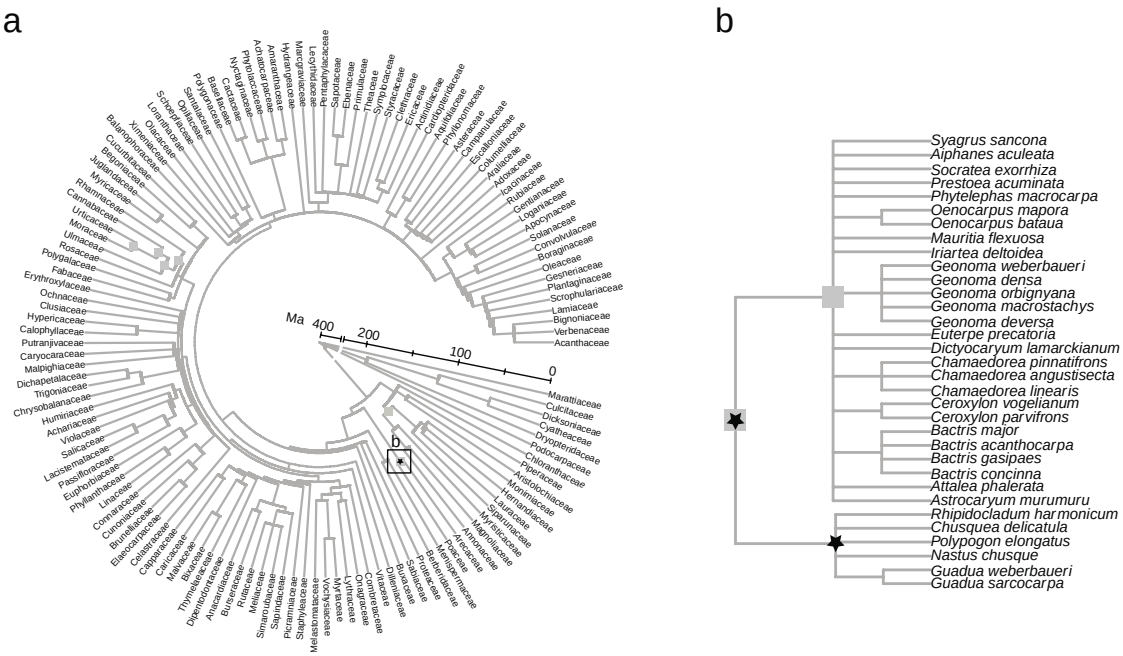

Figure 7. a. Phylogeny calibrated at the family level for the small plots (0.1 ha), showing nodes with phylogenetic signal for geographic range size (grey squares) and phylogenetic lability for abundance (black stars) according to disparity analysis. **b.** Detail of the phylogeny focused on clades characterized by phylogenetic signal in geographic range size and phylogenetic lability in abundance.