Dissecting the difference in tree species richness between Africa and South America

Pedro Luiz Silva de Miranda[a,1] 📷, Kyle G. Dexter[b] 📷, Michael D. Swaine[c] 📷, Ary Teixeira de Oliveira-Filho[d] 📷, Olivier J. Hardy[e], and Adeline Fayolle[a] 📷

Differences in species diversity over continental scales represent imprints of evolutionary, ecological, and biogeographic events. Here, we investigate whether the higher tree species richness in South America relative to Africa is due to higher richness in certain taxonomic clades, irrespective of vegetation type, or instead due to higher richness in specific biomes across all taxonomic clades. We used tree species inventory data to address this topic and began by clustering inventories from each continent based on species composition to derive comparable vegetation units. We found that moist forests in South America hold approximately four times more tree species than do moist forests in Africa, supporting previous studies. We also show that dry vegetation types in South America, such as tropical dry forests and savannas, hold twice as many tree species as do those in Africa, even though they cover a much larger area in Africa, at present and over geological time. Overall, we show that the marked species richness difference between South America and Africa is due primarily to a key group of families in the South American Amazon and Atlantic moist forests, which while present and speciose in Africa, are markedly less diverse there. Moreover, we demonstrate that both South American and African tree floras are organized similarly and that speciose families on one continent are likely speciose on the other. Future phylogenetic and functional trait work focusing on these key families should provide further insight into the processes leading to South America's exceptional plant species diversity.

taxonomic diversity | tropical trees | tropical moist forest | tropical dry forest | savanna

Plant diversity is not evenly distributed across the biosphere—the tropics are more species rich than other regions, and moist tropical forests have more plant species than the dry tropics (1, 2). Historically, the distribution of biodiversity has been investigated from a broadscale historical perspective (pattern description over large geographic scales) or from a local ecological perspective (hypothesis testing at community scales) (3), both leading to key findings. Among the numerous hypotheses that have been proposed, high plant species richness in tropical moist regions has been associated with high water availability and reduced climatic seasonality (4, 5), while the decrease in plant species richness toward the poles has been linked (among several other factors) to current plant clades having a tropical origin and lacking the adaptations required to inhabit temperate or boreal zones (6–8). From a different perspective, according to the species-area relationship (9), species richness and area are positively correlated (an increase in area will likely lead to an increase in richness), an effect which is modulated by how environmentally stable an area has been over geological time (10). Importantly, in the tropics, this effect may be partly enhanced by higher speciation and/or lower extinction rates [(11) and references therein]. Moreover, it has been hypothesized that favorable environmental conditions may enhance biotic interactions, which in turn favor higher diversification rates (12, 13). Though differences in plant richness between temperate and tropical zones are well understood, differences in plant species richness among tropical regions remain largely unexplored [though see (4, 14–17)]. Tropical Africa's depauperate forest tree flora, in comparison to tropical South America and Southeast Asia—the odd-man-out pattern—has been investigated relatively rarely despite awareness of the pattern since at least 1973 (18, 19).

Among regions, differences in timing of the origin and diversification of lineages, along with differences in dispersal and extinction, may lead to substantial discrepancies in species richness known as diversity anomalies (6). These anomalies are imprints of past evolutionary and ecological events and are, therefore, key evidence to unravel how communities were assembled over time. By comparing regional floras, Richards (18) first showed that African tropical moist forests held fewer species than similar forests elsewhere and suggested that its depauperate flora is linked to harsher past and present-day climatic conditions, as well as to differences in human occupation. Latin America as a whole has 3.8 times as many plant species as tropical Africa (20), and much of this

Significance

Our full-scale comparison of Africa and South America's lowland tropical tree floras shows that both Africa and South America's moist and dry tree floras are organized similarly: plant families that are rich in tree species on one continent are also rich in tree species on the other continent, and these patterns hold across moist and dry environments. Moreover, we confirm that there is an important difference in tree species richness between the two continents, which is linked to a few families that are exceptionally diverse in South American moist forests, although dry formations also contribute to this difference. Plant families only present on one of the two continents do not contribute substantially to differences in tree species richness.

Author affiliations: [a]Gembloux Agro-Bio Tech, University of Liege, 5030 Gembloux, Belgium; [b]School of GeoSciences, University of Edinburgh, Edinburgh EH8 9YL, United Kingdom; [c]Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen AB24 3FX, United Kingdom; [d]Departamento de Botânica, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-901, Brazil; and [e]Evolutionary Biology and Ecology, Faculté des Sciences, Université Libre de Bruxelles, 1050 Brussels, Belgium

[1]To whom correspondence may be addressed. Email: plsilvademiranda@gmail.com.

diversity difference is due to the exceptional diversity of South America, which is notable for being greater than that of Africa even though South America is 59% of the size. Further research has also linked differences in tree species richness at the plot level (i.e., alpha diversity) between moist forests in Western and Central Africa and the Amazon to differences in current climatic and ecological conditions [water availability (4)], turnover rates (15), stem density (21), and the presence of elephants and other megaherbivores in Africa (22). These studies focused only on moist tropical forests, particularly parts of the Amazon forest and Western and Central African (Guineo-Congolian) forests. Drier biomes and vegetation types have been scarcely considered. This is likely due to lack of comparable assimilated data from the dry tropics, but sufficient data are now available [e.g., (23) for Africa and (24) for South America]. Therefore, the full extent of anomalies in plant species richness between the two continents, considering all tropical biomes, is ripe for dissection and explanation [e.g., (20)].

The striking floristic similarity between Africa and South America has already been highlighted by Gentry (25) and recently confirmed using phylogenetic approaches (16, 26). In most accounts, the high number of shared families and genera (27) between the two continents is attributed to their common geological past—Western Gondwana—whose split presumably led to vicariance-driven divergence events, though this view has been firmly contested [e.g., evidence for vicariance is easily distorted or lost by sampling errors (28)]. The rise of angiosperms roughly coincides with the Western Gondwanan split (start: ~130 Mya, end: ~90 Mya), when the Gondwanan flora was dominated by gymnosperms (e.g., *Araucauria* and *Podocarpus*) and seed ferns (e.g., *Komlopteris* and *Pachypteris*) (29). Angiosperms only dominated African and South American floras after the mass extinction event marking the Cretaceous–Paleogene (K-Pg) boundary [~65 Mya (30–32)], when both continents were already isolated from one another, as well as from other land masses. The K-Pg boundary extinction event was followed by an increase in the diversity of plant genera (33, 34) and by the origin and diversification of important pantropical and speciose plant families [e.g., Fabaceae (35)]. Therefore, the tree species richness difference between Africa and South America most likely results from biogeographic events that took place after the end of the Cretaceous. Consequently, the observed taxonomic and phylogenetic similarities are as likely or more likely to be related to (long-distance) dispersal events via various routes than to vicariance (18, 36–39).

Observed differences in tree species richness between Africa and South America have often been attributed to mechanisms that would impact net diversification rates, primarily in moist tropical forests. In addition to differences in current climate (4), South American tropical moist forests cover a larger area and have been subjected to weaker expansion/contraction cycles than their African counterparts (40). In addition, the Amazon forest, which potentially harbors ~16,000 species of trees (41, 42), may also act as a biodiversity pump by being the source of lineages of plants and animals found in other South American biomes (43). Much less research attention has been given to drier biomes, which in Africa cover most of the continent, extend over large environmental gradients (44), and have done so over geological time. Furthermore, since the beginning of the Pliocene, Africa has become increasingly arid due to changes in ocean currents (34), while in South America, the Andes limited continental desiccation during glacial periods (40, 45). Given the larger area occupied by dry biomes in Africa, we may expect that they are more diverse than the dry

biomes of South America. Therefore, the greater overall tree species richness of South America in relation to Africa may be linked entirely to its moist forests, which must also hold sufficient tree species in order to surpass any potential richness difference in favor of Africa in the dry tropics. Conversely, as plant clades and families are fundamentally different in their net diversification rates and their biogeographic histories (46, 47), differences in tree species richness between Africa and South America might be linked simply to plant families that are entirely (or almost entirely) restricted to South America [e.g., Malpighiaceae (48), Vochysiaceae (18, 49), and to a lesser degree Arecaceae (50, 51)]. However, how much individual families contribute to the overall tree species richness difference between the two continents remains unknown, especially in drier environments.

Here, we sought to test whether South America's high tree species richness compared to Africa is driven solely by higher species richness in moist forest vegetation or whether there are meaningful differences in species richness between the two continents' dry formations as well. We compare floristic profiles, the distribution of species richness per family, for Africa and South America as the means to understand how their tree floras are taxonomically organized. We hypothesize that differences in tree species richness between the two continents are mainly linked to their moist forests, though dry formations may meaningfully contribute to total tree species counts. Furthermore, we hypothesize that much of the diversity difference will be due to the families that are restricted, or nearly so, to South America.

Results

African and South American Vegetation Clusters. We assembled tree species checklists for both Africa and South America from various sources (see *Materials and Methods*). In order to develop standardized units for comparison, we first delimited 11 vegetation clusters in Africa and 12 in South America (Fig. 1 and *SI Appendix*, Fig. S1) via hierarchical clustering based on floristic turnover among tree species inventories (computed for each pair of sites using the Simpson index of beta diversity). We conducted the clustering analyses for both continents separately, given that they share few tree species. This allowed us to achieve equivalent and comparable floristic clusters via a standardized methodology. Overall, Africa is dominated by drier vegetation clusters (7 out of 11, Fig. 2A and *SI Appendix*, Fig. S2A), whereas moist forest clusters are more prevalent in South America (7 out of 12, Fig. 2B and *SI Appendix*, Fig. S2 A and C). Comparatively, African clusters are drier than the ones in South America, with the Sahel being the driest cluster across the two continents (Fig. 2C and *SI Appendix*, Fig. S2 B and D). South America also holds moist forests with colder temperature regimes (Amazonian-Andean foothill forests, Amazonian Guiana shield forests, and subtropical Atlantic forest; Fig. 2C) that have no analog in the African dataset.

To compare the climatic space occupied by vegetation clusters in both Africa and South America, we used a principal component analysis (PCA) on gridded climatic variables for each of our sites. The first axis of the PCA collated precipitation-related information and explained ~36% of the variance in the climatic data (Fig. 2E). The second axis encompassed temperature-related information and explained ~33% of the variance. By following the mean score of each vegetation type along the precipitation gradient, we divided the vegetation clusters into two categories, moist and dry vegetation clusters
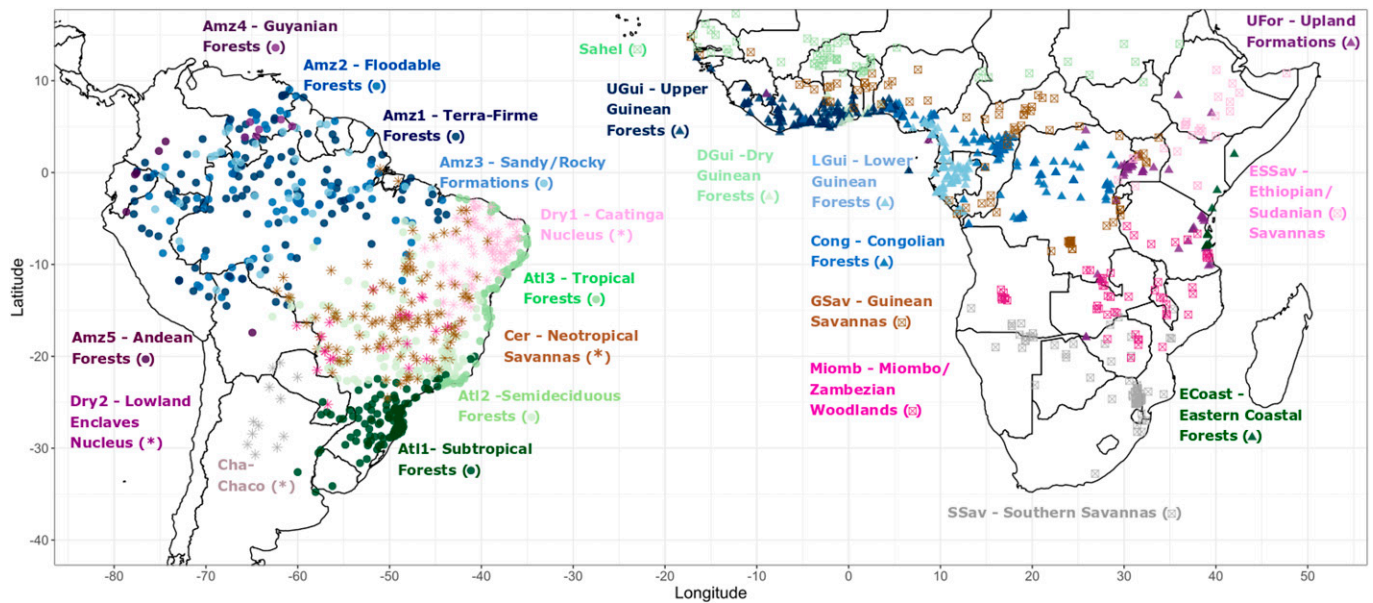
Fig. 1. Map of Africa and South America indicating the main vegetation clusters present in each continent identified via a hierarchical clustering analysis (UPGMA) based on tree species turnover (Simpson beta-diversity index). Each point corresponds to a georeferenced tree species checklist (n = 722 per continent).

(*SI Appendix,* Fig. S2 *A and B*), which would place the mean annual precipitation threshold dividing these two categories between 1,150 mm y^{-1} (5% quantile of the moist group) and 1,786 mm y^{-1} (95% quantile of the dry group).

African and South American Tree Species Diversity and Floristic Relatedness. Both continents are strikingly similar in how their floras are organized (Table 1, Fig. 3, and *SI Appendix,* Fig. S3). If a given family is species rich on one continent, it will most likely be species rich on the other continent as well (Fig. 3), a pattern that does not change when focusing only on moist or dry formations. However, vegetation clusters within the same continent are more similar among themselves with regard to tree species richness per genus and family than they are with clusters present on the opposite continent (*SI Appendix,* Fig. S4). Nevertheless, Africa and South America share a total of 99 tree families in our dataset, while their moist and dry vegetation clusters share 93 and 81 families, respectively. On average, the families present on both continents hold around 95% of the total observed tree species richness (Table 1). Therefore, families found exclusively on one continent only account for ~5% of the total tree species richness. Importantly, ~50% of the existing tree species richness in the whole of Africa and South America across both moist and dry vegetation clusters belongs to a group ranging from seven to nine families (Fig. 4 and *SI Appendix,* Tables S1–S3). Fabaceae is by far the most species rich and ecologically diversified family across the two continents, with numerous species in each vegetation cluster, moist and dry (Fig. 4 and *SI Appendix,* Fig. S5). Apart from a few families, such as Combretaceae, Phyllanthaceae, and Sapindaceae being relatively more important in Africa and Lauraceae, Melastomataceae, and Myrtaceae being more prevalent in South America, the most speciose families on each continent are largely the same (*SI Appendix,* Fig. S5). Importantly, most of the families shared by the two continents are younger than the Gondwanan split and the K-Pg boundary extinction event (crown node age younger than 90 and 65 Mya, respectively, *SI Appendix,* Table S7). Only a few families, such as Annonaceae,

Arecaceae, and Lauraceae, are older than the Gondwanan split (*SI Appendix,* Table S7). Moreover, tree species richness per family is not correlated with family age (*SI Appendix,* Fig. S6).

The difference in tree species richness between the two continents is substantial (Table 1). While our dataset for Africa contains 3,048 species distributed across 816 genera and 131 families over 722 sites, the South American dataset holds 8,842 tree species across 1,083 genera and 152 families for the same number of sites (we subset the larger South American dataset by using spatially stratified random sampling to enable a fair comparison with Africa; see *Materials and Methods*). Intercontinental ranked correlations of tree species richness per family between the two continents' entire tree floras and moist and dry floras separately yielded correlation coefficients around 0.62 (all highly significant, $P < 0.0001$, Fig. 3). However, the moist and dry floras of each continent are still more correlated to one another than they are to their intercontinental counterparts. African moist and dry vegetation formations at the family level (93 families in common) are highly correlated in terms of number of species per family ($r_{Spearman} = 0.79$, $P < 0.0001$, Fig. 3C), and this correlation is even more striking between South America moist and dry vegetation formations (123 families shared, $r_{Spearman} = 0.92$, $P < 0.0001$, Fig. 3D). On both continents, families present in both moist and dry clusters tend to be more species rich in moist vegetation than in dry vegetation. At the genus level, intercontinental comparisons are limited due to the relatively low number of genera shared between the two continents (*SI Appendix,* Fig. S4). Among the shared genera, Africa and South America's dry vegetation are more correlated in species richness (94 shared genera, $r_{Spearman} = 0.43$, $P < 0.0001$, *SI Appendix,* Fig. S4C) than the moist vegetation clusters (111 shared genera, $r_{Spearman} = 0.27$, $P < 0.005$, *SI Appendix,* Fig. S4B). Once again, moist and dry vegetation formations of the same continent are more correlated to one another than to their intercontinental counterparts (*SI Appendix,* Fig. S4D, Africa: 358 shared genera, $r_{Spearman} = 0.57$, $P < 0.0001$; *SI Appendix,* Fig. S4E, South America: 664 shared genera, $r_{Spearman} = 0.74$, $P < 0.0001$).

Fig. 2. Climatic space of all clusters identified in Africa (*A* and *B*, 722 checklists in total) and South America (*C* and *D*, 722 checklists in total) represented by the first two axes of a PCA. (*A*) Africa, moist clusters; (*B*) Africa, dry clusters; (*C*) South America, moist clusters; and (*D*) South America, dry clusters, generated via the same PCA, which was then subdivided into four panels as the means to show the four main climatic clusters encountered in Africa and South America. (*E*) Variable correlation circle generated via the same PCA. Climatic variables included in the PCA are the 19 variables provided by CHELSA plus climatic water deficit (89). The larger points represent the mean of each group. For vegetation cluster names, see Fig. 1. Climatic variables' names are as follows: MAT, mean annual temperature; MDR, mean diurnal range; Tiso, isothermality; Tsea, temperature seasonality; Tmwm, maximum temperature of warmest month; Tmcm, minimum temperature of coldest month; Tar, temperature annual range; Tmweq, mean temperature of wettest quarter; Tmdq, mean temperature of driest quarter; Tmwaq, mean temperature of warmest quarter; Tmcq, mean temperature of coldest quarter; MAP, annual precipitation; Pwm, precipitation of wettest month; Pdm, precipitation of driest month; Psea, precipitation seasonality; Pweq, precipitation of wettest quarter; Pdq, precipitation of driest quarter; Pwaq, precipitation of warmest quarter; Pcq, precipitation of coldest quarter; CWD, climatic water deficit. PC, Principal Component.

Differences in tree species richness between Africa and South America. Contrary to our expectation that differences in tree species richness between Africa and South America would be driven by families restricted or almost restricted to South America, we found that tree species richness differences are driven by a few families that are speciose on both continents but exceptionally more speciose in South America. While other families have detectable species richness differences (42 families are significantly more species rich in South America than Africa; only 3 families are significantly more species rich in Africa), relatively few families drive the large difference in overall species richness totals (Fig. 4 *A–C* and *SI Appendix,* Tables S1–S3). The families that are restricted (or nearly so) to South America account for a small proportion of the total species richness in South America (5%) and of the diversity difference (~0.08%). Null model simulations show that if South America and Africa had the same taxonomic diversity profiles, the much

higher species richness in South America should be linked to having more families than observed. South America has 11 fewer families than expected given its very high species richness ($P < 0.003$), indicating that the diversity difference is driven not by having more families per se (*SI Appendix, SI Materials and Methods*) but by having more species in a key group of families. When considering the observed proportion of overall tree species richness in South America (SA) relative to Africa (AF) as the baseline expectation in binomial tests (AF 0.26/SA 0.74), it is possible to see that both Africa and South America have a similar number of families that are more speciose than expected (*SI Appendix,* Table S1, AF 18/SA 19). However, 3 of the top 5 and 6 of the top 20 most speciose families are significantly more speciose in South America, while only 1 of the top 5 and 3 of the top 20 are significantly more speciose in Africa. It is overall less speciose families (with fewer than 100 species in both continents combined) that tend to be more speciose in

Table 1. Summary of the total number of botanical families and tree species present in lowland tropical South America and Africa

	Total family richness	Total species richness	No. of families present in both continents	No. of species belonging to the shared families	No. of families only found in one continent	No. of species belonging to the unique families
South America, all vegetation clusters (722 sites)	152	8,842	99	8,393 (95%)	53	449 (5%)
Africa, all vegetation clusters (722 sites)	131	3,048	99	2,954 (97%)	32	94 (3%)
South America, moist vegetation clusters (407 sites)	151	7,979	93	7,501 (94%)	58	478 (6%)
Africa, moist vegetation clusters (410 sites)	116	2,148	93	2,092 (97%)	23	56 (3%)
South America, dry vegetation clusters (315 sites)	125	3,498	81	3,032 (86%)	44	466 (14%)
Africa, dry vegetation clusters (312 sites)	109	1,570	81	1,505 (96%)	38	65 (4%)

In parentheses, we report the relative proportion of species present in each fraction of families according to the total tree species pool. Numbers given for South America refer to the subsampled dataset. Percentages given in columns 4 and 6 refer to the total species richness values reported in column 2.

Africa relative to a 0.26/0.74 baseline expectation. Interestingly, the three families that are among the top 20 most speciose families overall and are more speciose in Africa than in South America, given a baseline 0.26/0.74 expectation, are Fabaceae, Malvaceae, and Sapindaceae, families that have successfully radiated in moist and dry environments (52–54).

Our intercontinental comparison among diversity profiles of moist and dry vegetation clusters, along with binomial tests, shows that the high number of tree species in South America is mainly due to high richness in moist vegetation clusters (Amazon forest + moist Atlantic forest clusters, Fig. 4B and *SI Appendix*, Table S2). Out of the 93 families shared by the moist formations of the two continents, 41 hold significantly more species in South America than in Africa, while only Putranjivaceae holds significantly more species in Africa than in South America (Fig. 4 and *SI Appendix*, Table S2). When taking the observed differences in tree species richness between moist forests on the two continents as the baseline (*SI Appendix*, Table S2, AF 0.21/SA 0.79), 19 families hold more tree species than expected in Africa, while 13 families hold more tree species than expected in South America. Meanwhile, concerning the dry vegetation, out of the 81 shared families, 22 hold significantly more species in South America than in Africa, whereas three families present the opposite pattern (Fig. 4C and *SI Appendix*, Table S3). When considering the difference in tree species richness between dry clusters (*SI Appendix*, Table S3, baseline expectation = AF 0.33/SA 0.67), both continents have nine families that are more speciose than expected. Similar results were obtained via post hoc χ^2 tests (*SI Appendix*, Tables S4–S6). In total, the families which are significantly more species rich in South America than in Africa account for a total of 5,422 more tree species (61% of South America's tree species pool in 42 families), 4,739 (59% in 41 families) when only moist vegetation is taken into account, and 1,161 (33% in 22 families) when comparing the two continents' dry floras. Importantly, Fabaceae, Lauraceae, Melastomataceae, and Myrtaceae are the main families driving the tree species difference, as they alone account for 2,837 tree species in South America while only having 657 tree species in Africa, in our dataset subsampled for South America.

Discussion

Our findings confirm the meaningful difference in tree species richness between Africa and South America, which helped confer Africa the title of odd man out (18) and has been documented to a limited degree in other research efforts (4, 15, 20). Here, we were able to demonstrate that such intercontinental differences in tree species richness are clearly driven by variation in the species richness of moist vegetation clusters. We show that South America's moist vegetation holds most of the tree species accounting for intercontinental differences (four times more species in South America than in Africa), even though South America's dry formations also hold more tree species overall than their African counterparts (two times more species in South America than in Africa). Concerning whether South America's high tree species richness could be linked to specific families, we were able to identify a restricted group of families that have more species in South America's moist forests than anywhere else. Importantly, these families are species rich in Africa as well, just less so there. Moreover, 9 of these families account for around 50% of the tree floras within both continents, totaling 16 families across all diversity profiles (*SI Appendix*, Fig. S5), with Fabaceae, Rubiaceae, and Malvaceae being among the most species rich overall. Importantly, when overall differences in tree species richness between the two continents are accounted for, it is possible to see that Africa, which has proportionally more dry geographic area, also holds highly diversified plant families, such as Anacardiaceae, Combretaceae, Fabaceae, and Malvaceae, families which are notable for radiating in both moist and dry environments.
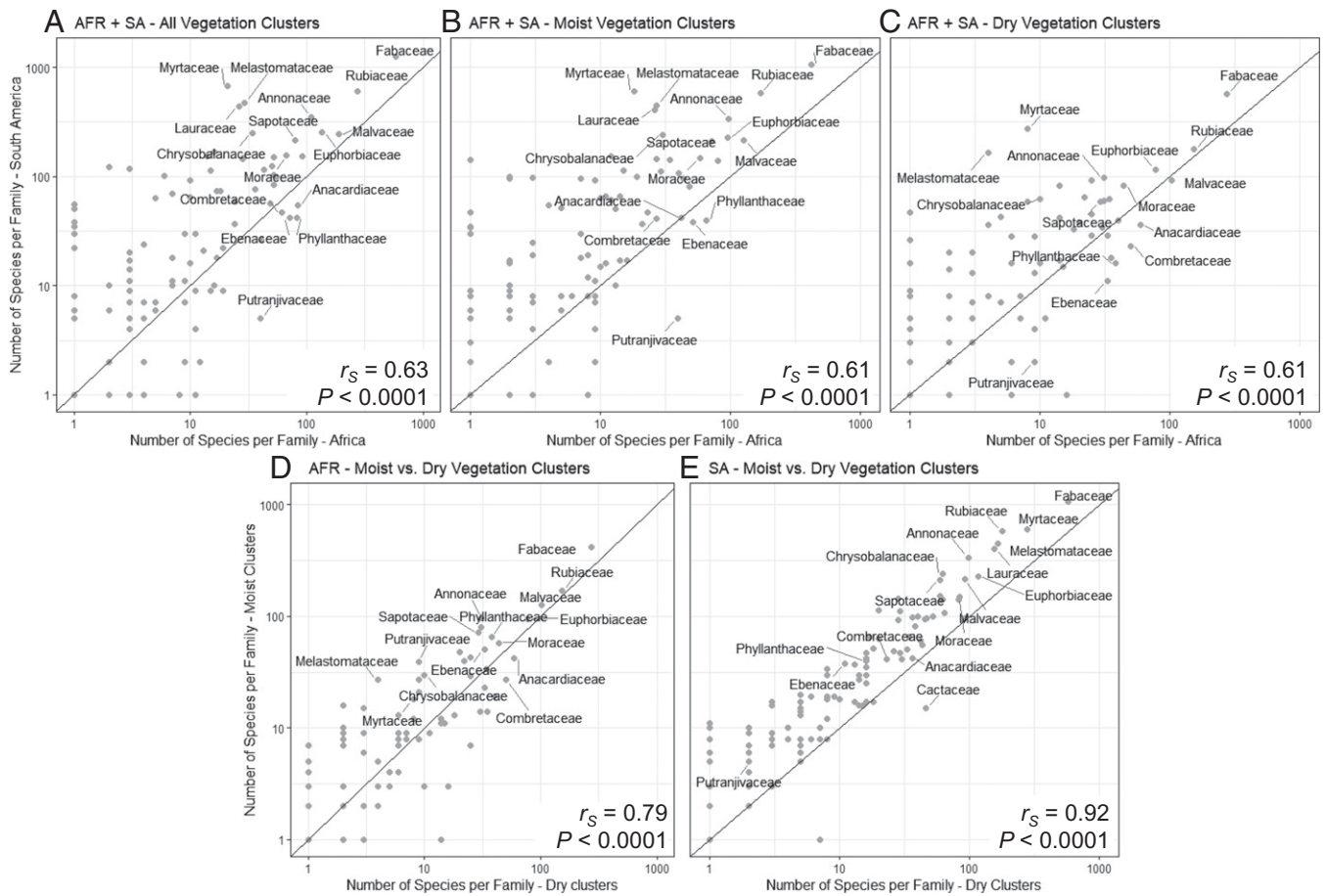
Fig. 3. Correlation plots among African and South American tree floras concerning number of species per family. (*A*) All vegetation clusters are included. (*B*) Only moist vegetation clusters. (*C*) Only dry vegetation clusters. (*D*) African moist and dry vegetation clusters are compared. (*E*) South American moist and dry vegetation clusters are compared. Spearman's rank correlation coefficients (r_S) and significance levels (*P*) are given within each panel. For information on which vegetation clusters were classified as moist or dry, see Fig. 2. AFR - Africa, SA - South America.

South America's Moist Vegetation Clusters Account For Most of the Difference in Tree Species Richness between the Two Continents. Numerous hypotheses have been proposed to explain why the Amazon and the Atlantic forests possess such high tree species richness. The diversity of environments in the Amazon, spanning such a broad area and ranging from terra firme, to seasonally flooded forests, to forests growing on white sands, has been put forward as one of the reasons why this forest is so diverse (55–58). Climatic stability and high water availability have also been considered as possible drivers of this high biodiversity (4, 10, 59, 60). With regard to the Atlantic forest's tree species diversity, it has been hypothesized that its high species richness is linked, at least in part, to its broad elevational (0 to ~2,000 m) and latitudinal ranges (61). Moreover, a recent study has shown the high tree species diversity of the Atlantic forest (at the regional level) is linked to the plethora of environments, giving a diversity of habitats (62).

Even though the two continents share a common geological past, the African flora has been subjected to different environmental pressures (e.g., increased aridity, forest cover reduction, and fragmentation due to glaciation cycles), which may hold the answer as to why the African continent holds fewer species than South America. At present, South America has a greater forest cover than Africa (63) and about 1.5 times more individual trees (64). Over evolutionary time, the African moist forest flora has been subjected to stronger contraction and expansion events due to climatic variation during the Pleistocene and

Miocene than the moist vegetation in South America, which has been sheltered, to some degree, from drier climatic conditions by the Andes (19, 40, 65). Also, during forest contraction events, Africa may have provided fewer and less extensive refugia for its moist forest flora due to fewer mountain ranges, although the exact location of forest refugia in Africa (66) is still in debate (67). In our spatially and environmentally comprehensive dataset, South America has nearly triple the number of tree species as Africa (SA/AF = 2.9). We suggest that the current differences in forest cover and number of trees and the historical differences in contraction/expansion dynamics may be sufficient to explain current differences in tree species richness (15) without the need to invoke differences in lineage speciation rates between the two continents. However, plant families such as Combretaceae, Ebenaceae, Fabaceae, Phyllanthaceae, and Rubiaceae in Africa and Arecaceae, Chrysobalanaceae, Lauraceae, Malpighiaceae, Melastomataceae, and Myrtaceae in South America challenge this perspective, given that they are surprisingly speciose in one of the two continents, even when accounting for the overall difference in species richness between them.

Ranked Family Richness Is Conserved between Africa and South America Despite Differences in Tree Species Richness. Africa and South America, during most of Earth's geological history, were joined together and formed the bulk of a continent known as Gondwana, a fact that led past botanical
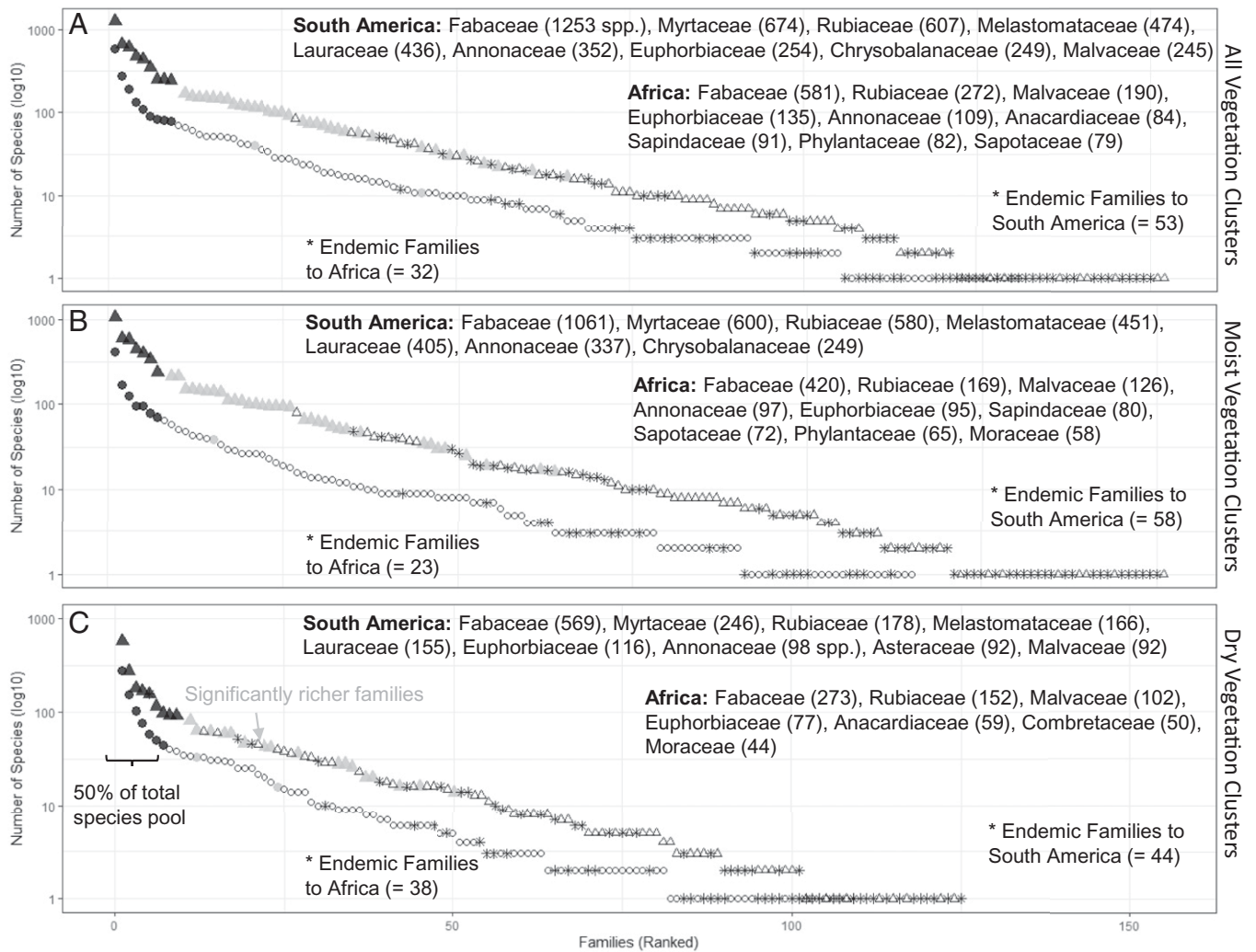
Fig. 4. Families ranked in decreasing order of species richness (log10) in both Africa (circles) and South America (triangles). (*A*) All vegetation clusters combined. (*B*) Only the moist clusters. (*C*) Only the dry clusters. Families in black hold 50% of the total species pool in each continent. Families in gray hold a significantly higher number of species in relation to that same family in the opposing continent. Asterisks represent families that can only be found on the continent they are in. Text boxes in each panel lists the families represented by the black symbols and shows the number of tree species they hold in each configuration.

researchers to associate the two continent's floristic similarities (at family and genus levels) to their shared geological past (18, 68). However, during the past ∼90 My, these two continents split and drifted away from one another, so observed similarities are unlikely to be linked to a shared geological past. Though some of the families shared by the two continents, such as Annonaceae (69), have had their origins dated to times prior to the Gondwanan split, the majority of both African and South American tree floras is composed of families that originated after the Gondwanan split, with most of them only appearing and diversifying after the K-Pg extinction event (*SI Appendix*, Table S7). Therefore, the tree community assembly of Africa and South America is the outcome of both long-distance dispersal, via transoceanic dispersal, and ancient land bridges and speciation events intrinsic to each continent and their vegetation clusters (36). The correlated species richness within families and genera across the two continents is most likely related to climatic niches and diversification rates being relatively well conserved at the family level (70), although further studies are needed in order to understand and test this pattern.

It is surprising to observe that tree species richness across families on the two continents is remarkably conserved, regardless of the overall richness difference between them. Our findings indicate ∼50% of each continent's tree species richness is formed by a restricted group of families that are mostly present on both continents, a pattern also found at the global level (49). Given that how plant families have been circumscribed has been well established over the years [e.g., Angiosperm Phylogeny Group (APG) III (71) and APG IV (72)], it is unlikely that this finding will change in the future or that it is a direct bias of how the classification system is structured. When investigating the taxonomic structure of several sites in moist forests in Africa, Asia, and South America as the means to test the role of neutral processes in community assembly on different continents, the same pattern that we highlight here—families that are species rich on one continent are most likely species rich on the other continent—was observed as well (14). The striking result that a set of 16 families accumulates 50% of each continent's total tree species richness across both moist and dry vegetation clusters results from each family's biogeographic history, along with features that would ensure adaptive advantages over

continental scales (12, 13). For example, most clades from the Fabaceae family can fix nitrogen (73), enabling this group to adapt to a variety of harsh environmental conditions, particularly high seasonality (52). Meanwhile, most Rubiaceae clades have a complex biogeographic history, and their high diversity in South America seems to be linked to the rise of the Andes (74). Families like Myrtaceae and Lauraceae are more commonly found in moist environments and, in South America, have acquired adaptations to colder temperatures, enabling them to diversify in higher elevations (75), which could explain their high diversity in South America and low diversity in Africa. In contrast, Ebenaceae, particularly the genus *Diospyros,* is known for its morphological and species diversity in Africa, coupled with a geographically wide distribution (76, 77). Future studies unveiling the biogeographic and evolutionary history of key clades encompassing multiple growth habits (e.g., herbs and lianas) will provide more information on the comparative evolution of African and South American floras.

Conclusion

Here, we show that the previously observed difference in tree species richness between Africa and South America is the result of species richness anomalies in a restricted group of families that are exceptionally diverse in South American moist forests. Surprisingly, these same families are also species rich in African moist vegetation and in African and South American dry vegetation; they are just much more speciose in South American moist forests. We also show that both African and South American tree floras have similar taxonomic organizations regardless of differences in tree species richness: families that are speciose on one continent are speciose on the opposite continent as well. However, our findings also point to each continent having its own floristic identity and intrinsic patterns of tree species richness and distribution, evidenced by strong within-continent correlations in richness between moist and dry vegetation formations. Therefore, intracontinental dynamics seem to have a more prominent role in biome assembly than intercontinental lineage dispersal or migration.

Materials and Methods

Tree Species Inventories. We analyzed two datasets: one for Africa and one for South America. The African dataset (AfroTropTree) is the union of the datasets employed for the biogeography of forest (78) and savanna (23) trees and was first jointly analyzed by Aleman et al. (44), while the South American dataset (NeoTropTree, http://www.neotroptree.info) was finished in 2018 and has been fully available online ever since. Both are collections of georeferenced tree species checklists compiled from published (e.g., scientific articles) and unpublished (e.g., master's and PhD theses) sources that have been carefully compiled, checked, and incorporated over the years. In both AfroTropTree and NeoTropTree, we define trees as woody plants capable of growing 3 m in height and that are freestanding. Importantly, the compilation of these resources was made by constantly verifying species identifications via contacting taxonomists and specialists and by performing yearly taxonomic updates. Only valid and accepted species names are included in both datasets; we checked name validity by consulting Tropicos (https://tropicos.org/), the African Plant Database (curated by the Conservatoire et Jardin Botanique de la Ville de Genève), and the Lista de Espécies da Flora do Brasil (Brazil only). Moreover, when possible, species inclusion in the dataset was verified by evaluating herbaria vouchers. Both African (23, 44, 78) and South American (79–81) datasets have been explored and validated in previous research aiming to investigate macroecological, biogeographic, and evolutionary research questions within continents. Further details on how both datasets were assembled can be found in the references (23, 44, 78–81) and in the *SI Appendix, SI Materials and Methods*. Here, we only included checklists of

frost-free areas [fourth criterion of (82)] and below 1,750 m of elevation. In the case of South America, we also excluded inventories from the Andes and from the Pacific Coast.

Delimiting Vegetation Clusters. As the means to create a framework enabling tree species diversity comparisons between the two continents for analogous vegetation clusters, we employed a hierarchical clustering approach based on species turnover in order to delimit the main vegetation clusters in lowland tropical Africa and South America. By working with species assemblages of vegetation clusters instead of collections of individual checklists, we reduce the pseudoreplication effect generated by species cooccurrence over broad geographic spaces (83). We conducted two clustering analyses–one for Africa and one for South America–since the two continents share only a small fraction of species (only 31 shared species in the combined dataset). The final African matrix included 3,048 tree species (816 genera and 131 families) distributed along 722 sites. The final South American matrix included 10,268 tree species (1,197 genera and 158 families) distributed along 4,980 sites. These occurrence tables were then used to build pairwise floristic-distance matrices showing how similar or dissimilar each site is to the other sites in relation to their tree species composition. To this end, we employed the Simpson index of dissimilarity [Betasim (84), available on the R package recluster (85)]. This index has been shown to produce unbiased results even when the data holds 1) differences in sampling effort (uneven sampling) and 2) meaningful differences in species richness.

We then grouped the sites according to their pairwise floristic distances by employing the unweighted pair group method with arithmetic mean (UPGMA) clustering algorithm, as recommended by Kreft and Jetz (86), as it was proven to consistently have the best performance among other algorithm options for biogeographic delimitation purposes when analyzing occurrence data (presence/absence). We repeated this procedure 100 times by randomizing the order of rows in the community matrix and assimilated the resulting dendrograms into a final dendrogram by following the majority-rule consensus approach. Therefore, the two final dendrograms only portray groups/branches present in a majority of dendrograms. In order to obtain fully resolved dendrograms (all nodes are bifurcations), we employed the RogueNaRok algorithm, a tool commonly used to build fully resolved phylogenies (87). Here, we used this algorithm to detect sites across the dendrograms with high instability in placement, which prevent the determination of a fully resolved final solution. This led to the removal of 40 sites in Africa and 80 sites in South America. We built the pairwise distance matrices and the final dendrograms using functions in the recluster package (88) in R software. We compiled and ran the RogueNaRok algorithm in C language on a Linux (Ubuntu) machine.

We inspected the two final dendrograms for Africa and South America and manually delimited their vegetation clusters by observing the following criteria: 1) overall branching pattern of each dendrogram and 2) main vegetation types present on each branch. We then investigated how these vegetation clusters were distributed in multivariate, compositional, and geographic spaces. For the former, we performed a nonmetric multidimensional scaling (NMDS) ordination and plotted the sites according to their scores on the first and second axes. For the latter, we plotted the sites onto maps and investigated the limits of their geographic distribution. We performed the NMDS analyses by applying the metaMDS function from the vegan package in the R software. Importantly, prior to performing further analysis and due to differences in dataset size, we selected 722 sites from South America in a geographically stratified, but otherwise random, fashion to maintain the proportion of sites per vegetation type present in the South American dataset and to guarantee full geographic coverage of sampling. The combined dataset for all analyses described in the following sections contained 1,444 sites (722 for each continent).

Environmental Affinities. To investigate the climatic space occupied by the delimited vegetation clusters and compare overlaps/partitions within and between continents, we performed a PCA for all sites in Africa and the subsampled sites for South America, based on their values for climatic variables. We retrieved all 19 climatic variables available from Climatologies at High Resolution for the Earth's Land Surface (CHELSA), which portrays yearly variation patterns in precipitation and temperature. We also included the climatic water deficit from Chave et al. (89). We then compared the average PCA scores per vegetation cluster on the first two axes of the combined climatic PCA. Based on

these results, we categorized all vegetation clusters into two broad categories: moist climate versus dry climate vegetation clusters. We conducted the PCA with the ade4 package (90). PCA biplots were built with the factorextra package (91), also in R.

Comparing Diversity Profiles. Here, we compared the taxonomic organization of the tree floras of Africa and South America. Firstly, we counted the number of families and genera present on the two continents and in their moist and dry clusters. We then proceeded to calculate the proportion of the total species pool the shared families and genera hold (between the whole continents and between moist and dry vegetation clusters). Secondly, to assess the extent of congruence in the taxonomic composition of the two continents, we performed a series of Spearman rank correlations, where high correlation values are obtained when families are ranked in the same position according to their tree species richness. We did the following correlations: 1) African and South American whole floras, 2) African and South American moist floras, and 3) African and South American dry floras. We also investigated intracontinental correlations between moist and dry floras at the family level 4) for African moist and dry floras and 5) for South American moist and dry floras. Following the same protocol, we also did Spearman rank correlations at the genus level (number of species per genus).

To compare African and South American diversity profiles, we built ranked-richness curves, ordering families according to their species richness in decreasing order and therefore following the same logic applied to the construction of ranked-abundance plots in community ecology [e.g., (41)]. To this end, we used the same subsampled dataset we employed to build the PCA and built diversity profiles for Africa and South America in three different ways: 1) including all vegetation clusters, 2) including only moist vegetation clusters, and 3) including only dry vegetation clusters. In each profile, we identified 1) the most species-rich families that make up ~50% of the total tree species richness of each curve, following the logic used to identify species that are hyperdominant in terms of abundance (41); 2) the families only present in one continent; and 3) the

families with statistically higher species richness in one continent in comparison to the other. As the means to detect the latter families, we applied two complementary rounds of binomial tests to the species richness per family in each diversity profile (whole continent, moist vegetation clusters, and dry vegetation clusters). Only families present in the two continents were considered. In the first round, we set an equal null expectation of a family having the same number of tree species on both continents (AF 0.50/SA 0.50), allowing for detection of meaningful differences in species richness between the two continents. In the second round, expected species richness per family was conditioned by the existing overall richness proportions of the two continents (whole continent, AF 0.26/SA 0.74; moist vegetation clusters, AF 0.21/SA 0.79; and dry vegetation clusters, AF 0.33/SA 0.67), therefore allowing the detection of families that are speciose even when accounting for the chief trend of South America being more species rich than Africa. We confirmed the results obtained via binomial tests with χ^2 tests (followed by post hoc χ^2 tests) to identify families that were more or less speciose than expected. We obtained significance measures in the χ^2 tests by applying a Monte Carlo simulation procedure, as not all assumptions of the test were met (no expected values are lower than 1 and at least 80% of expected values are higher than 5). We applied Bonferroni correction to determine thresholds for significance due to multiple testing on both binomial and χ^2 post hoc tests.

1. A. Esquivel-Muelbert *et al.*, Seasonal drought limits tree species across the Neotropics. *Ecography* **40**, 618–629 (2017).
2. V. Droissart *et al.*, Beyond trees: Biogeographical regionalization of tropical Africa. *J. Biogeogr.* **45**, 1153–1167 (2018).
3. R. E. Ricklefs, A comprehensive framework for global patterns in biodiversity. *Ecol. Lett.* **7**, 1–15 (2004).
4. I. Parmentier *et al.*, The odd man out? Might climate explain the lower tree α-diversity of African rain forests relative to Amazonian rain forests? *J. Ecol.* **95**, 1058–1071 (2007).
5. O. L. Phillips, P. Hall, A. H. Gentry, S. A. Sawyer, R. Vásquez, Dynamics and species richness of tropical rain forests. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2805–2809 (1994).
6. R. E. Latham, R. E. Ricklefs, Global patterns of tree species richness in moist forests: Energy-diversity theory does not account for variation in species richness. *Oikos* **67**, 325–333 (1993).
7. J. J. Wiens, M. J. Donoghue, Historical biogeography, ecology and species richness. *Trends Ecol. Evol.* **19**, 639–644 (2004).
8. D. M. Neves *et al.*, The adaptive challenge of extreme conditions shapes evolutionary diversity of plant assemblages at continental scales. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021132118 (2021).
9. O. Arrhenius, Species and area. *J. Ecol.* **9**, 95–99 (1921).
10. P. V. A. Fine, R. H. Ree, Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *Am. Nat.* **168**, 796–804 (2006).
11. G. G. Mittelbach *et al.*, Evolution and the latitudinal diversity gradient: Speciation, extinction and biogeography. *Ecol. Lett.* **10**, 315–331 (2007).
12. N. M. Nürk *et al.*, Diversification in evolutionary arenas-assessment and synthesis. *Ecol. Evol.* **10**, 6163–6182 (2020).
13. M. J. Donoghue, M. J. Sanderson, Confluence, synnovation, and depauperons in plant diversification. *New Phytol.* **207**, 260–274 (2015).
14. R. E. Ricklefs, S. S. Renner, Global correlations in tropical tree species richness and abundance reject neutrality. *Science* **335**, 464–467 (2012).
15. J. W. F. Slik *et al.*, An estimate of the number of tropical tree species. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7472–7477 (2015).
16. J. W. F. Slik *et al.*, Phylogenetic classification of the world's tropical forests. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 1837–1842 (2018).
17. Y. Kubota, B. Kusumoto, T. Shiono, W. Ulrich, Environmental filters shaping angiosperm tree assembly along climatic and geographic gradients. *J. Veg. Sci.* **29**, 607–618 (2018).
18. P. Richards, "Are the 'odd man out' in *Tropical Forest Ecosystems of Africa and South America: A Comparative Review*, B. J. Meggers, E. S. Ayensu, W. D. Duckworth, Eds. (Smithsonian Institution Press, 1973), pp. 21–26.
19. T. L. P. Couvreur, Odd man out: Why are there fewer plant species in African rain forests? *Plant Syst. Evol.* **301**, 1299–1313 (2015).
20. P. H. Raven *et al.*, The distribution of biodiversity richness in the tropics. *Sci. Adv.* **6**, eabc6228 (2020).
21. S. L. Lewis *et al.*, Above-ground biomass and structure of 260 African tropical forests. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120295 (2013).
22. J. Terborgh *et al.*, The African rainforest: Odd man out or megafaunal landscape? African and Amazonian forests compared. *Ecography* **39**, 187–193 (2016).

23. A. Fayolle *et al.*, A sharp floristic discontinuity revealed by the biogeographic regionalization of African savannas. *J. Biogeogr.* **46**, 454–465 (2019).
24. K. Banda-R *et al.*; DRYFLOR, Plant diversity patterns in neotropical dry forests and their conservation implications. *Science* **353**, 1383–1387 (2016).
25. A. H. Gentry, Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Ann. Mo. Bot. Gard.* **75**, 1–34 (1988).
26. Ş. Proches, S. Ramdhani, A global regionalisation based on the present-day distribution of broad plant lineages. *Phytotaxa* **442**, 20–26 (2020).
27. K. G. Dexter *et al.*, Floristics and biogeography of vegetation in seasonally dry tropical regions. *Int. For. Rev.* **17**, 10–32 (2015).
28. P. Upchurch, Gondwanan break-up: Legacies of a lost world? *Trends Ecol. Evol.* **23**, 229–236 (2008).
29. S. McLoughlin, The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Aust. J. Bot.* **49**, 271–300 (2001).
30. M. D. Crisp, L. G. Cook, Cenozoic extinctions account for the low diversity of extant gymnosperms compared with angiosperms. *New Phytol.* **192**, 997–1009 (2011).
31. C. J. Cleal, B. Cascales-Miñana, Composition and dynamics of the great Phanerozoic Evolutionary Floras. *Lethaia* **47**, 469–484 (2014).
32. M. R. Carvalho *et al.*, Extinction at the end-Cretaceous and the origin of modern Neotropical rainforests. *Science* **372**, 63–68 (2021).
33. C. Jaramillo, M. J. Rueda, G. Mora, Cenozoic plant diversity in the neotropics. *Science* **311**, 1893–1896 (2006).
34. T. L. P. Couvreur *et al.*, Tectonics, climate and the diversification of the tropical African terrestrial flora and fauna. *Biol. Rev. Camb. Philos. Soc.* **96**, 16–51 (2021).
35. E. J. M. Koenen *et al.*, The origin of the legumes is a complex paleopolyploid phylogenomic tangle closely associated with the cretaceous-paleogene (K-Pg) mass extinction event. *Syst. Biol.* **70**, 508–526 (2021).
36. C. W. Dick, R. T. Pennington, History and geography of neotropical tree diversity. *Annu. Rev. Ecol. Evol. Syst.* **50**, 279–301 (2019).
37. S. Renner, Plant dispersal across the tropical Atlantic by wind and sea currents. *Int. J. Plant Sci.* **165**, S23–S33 (2004).
38. R. T. Pennington, C. W. Dick, The role of immigrants in the assembly of the South American rainforest tree flora. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **359**, 1611–1622 (2004).
39. M. J. M. Christenhusz, M. W. Chase, Biogeographical patterns of plants in the Neotropics – Dispersal rather than plate tectonics is most explanatory. *Bot. J. Linn. Soc.* **171**, 277–286 (2013).
40. D. Anhuf *et al.*, Paleo-environmental change in Amazonian and African rainforest during the LGM. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **239**, 510–527 (2006).
41. H. ter Steege *et al.*, Hyperdominance in the Amazonian tree flora. *Science* **342**, 1243092 (2013).
42. H. Ter Steege *et al.*, Biased-corrected richness estimates for the Amazonian tree flora. *Sci. Rep.* **10**, 10130 (2020).
43. A. Antonelli *et al.*, Amazonia is the primary source of Neotropical biodiversity. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6034–6039 (2018).
44. J. C. Aleman *et al.*, Floristic evidence for alternative biome states in tropical Africa. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 28183–28190 (2020).

45. T. F. Rangel *et al.*, Modeling the ecology and evolution of biodiversity: Biogeographical cradles, museums, and graves. *Science* **361**, eaar5452 (2018).
46. J. P. Scholl, J. J. Wiens, Diversification rates and species richness across the Tree of Life. *Proc. Biol. Sci.* **283**, 20161334 (2016).
47. J. J. Wiens, What explains patterns of biodiversity across the Tree of Life?: New research is revealing the causes of the dramatic variation in species numbers across branches of the Tree of Life. *BioEssays* **39**, 1600128 (2017).
48. C. C. Davis, C. D. Bell, S. Mathews, M. J. Donoghue, Laurasian migration explains Gondwanan disjunctions: Evidence from Malpighiaceae. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6833–6837 (2002).
49. E. Beech, M. Rivers, S. Oldfield, P. P. Smith, GlobalTreeSearch: The first complete global database of tree species and country distributions. *J. Sustain. For.* **36**, 454–489 (2017).
50. R. Muscarella *et al.*, The global abundance of tree palms. *Glob. Ecol. Biogeogr.* **29**, 1495–1514 (2020).
51. J. Dransfield *et al.*, *Genera Palmarum. The Evolution and Classification of Palms* (University of Chicago Press, Chicago, IL, 2008).
52. M. Lavin, M. Thulin, J.-N. Labat, R. T. Pennington, Africa, the odd man out: Molecular biogeography of dalbergioid legumes (Fabaceae) suggests otherwise. *Syst. Bot.* **25**, 449–467 (2000).
53. A. Zizka *et al.*, Transitions between biomes are common and directional in Bombacoideae (Malvaceae). *J. Biogeogr.* **47**, 1310–1321 (2020).
54. S. Buerki *et al.*, An evaluation of new parsimony-based versus parametric inference methods in biogeography: A case study using the globally distributed plant family Sapindaceae. *J. Biogeogr.* **38**, 531–550 (2011).
55. C. M. Hulshof, M. J. Spasojevic, The edaphic control of plant diversity. *Glob. Ecol. Biogeogr.* **29**, 1634–1650 (2020).
56. C. A. Quesada *et al.*, Basin-wide variations in Amazon forest structure and function are mediated by both soils and climate. *Biogeosciences* **9**, 2203–2246 (2012).
57. H. ter Steege *et al.*, Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* **443**, 444–447 (2006).
58. A. T. Oliveira-Filho *et al.*, On the floristic identity of Amazonian vegetation types. *Biotropica* **53**, 767–777 (2021).
59. A. Antonelli, I. Sanmartín, Why are there so many plant species in the Neotropics? *Taxon* **60**, 403–414 (2011).
60. J. Terborgh, On the notion of favorableness in plant ecology. *Am. Nat.* **107**, 481–501 (1973).
61. A. T. Oliveira-Filho, M. A. L. Fontes, Patterns of floristic differentiation among Atlantic Forests in Southeastern Brazil and the influence of climate. *Biotropica* **32**, 793–810 (2000).
62. D. M. Neves *et al.*, Dissecting a biodiversity hotspot: The importance of environmentally marginal habitats in the Atlantic Forest Domain of South America. *Divers. Distrib.* **23**, 898–909 (2017).
63. M. C. Hansen *et al.*, High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).
64. T. W. Crowther *et al.*, Mapping tree density at a global scale. *Nature* **525**, 201–205 (2015).
65. W. D. Kissling *et al.*, Cenozoic imprints on the phylogenetic structure of palm species assemblages worldwide. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7379–7384 (2012).
66. J. Maley, "Late Quaternary climatic changes in the African rain forest: Forest refugia and the major role of sea surface temperature variations" in *Paleoclimatology and Paleometeorology: Modern and Past Patterns of Global Atmospheric Transport*, M. Leinen, M. Sarnthein, Eds., NATO ASI Series (Springer Netherlands, 1989), pp. 585–616.
67. J. Duminil *et al.*, Late Pleistocene molecular dating of past population fragmentation and demographic changes in African rain forest tree species supports the forest refuge hypothesis. *J. Biogeogr.* **42**, 1443–1454 (2015).
68. P. H. Raven, D. I. Axelrod, Angiosperm biogeography and past continental movements. *Ann. Mo. Bot. Gard.* **61**, 539 (1974).
69. T. L. P. Couvreur *et al.*, Early evolutionary history of the flowering plant family Annonaceae: Steady diversification and boreotropical geodispersal: Evolutionary history of the Annonaceae. *J. Biogeogr.* **38**, 664–680 (2011).
70. O. J. Hardy, P. Couteron, F. Munoz, B. R. Ramesh, R. Pélissier, Phylogenetic turnover in tropical tree communities: Impact of environmental filtering, biogeography and mesoclimatic niche conservatism. *Glob. Ecol. Biogeogr.* **21**, 1007–1016 (2012).
71. The Angiosperm Phylogeny Group, An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* **161**, 105–121 (2009).
72. The Angiosperm Phylogeny Group *et al.*, An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* **181**, 1–20 (2016).
73. J. I. Sprent, J. Ardley, E. K. James, Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytol.* **215**, 40–56 (2017).
74. A. Antonelli, J. A. A. Nylander, C. Persson, I. Sanmartín, Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9749–9754 (2009).
75. R. S. Bergamin *et al.*, Elevational shifts in phylogenetic diversity of angiosperm trees across the subtropical Brazilian Atlantic Forest. *Austral Ecol.* **46**, 486–495 (2021).
76. F. White, The taxonomy, ecology and chorology of African Ebenaceae I. The Guineo-Congolian species. *Bull. Jard. Bot. Natl. Belg. Bull. Van Natl. Plantentuin Van Belg.* **48**, 245–358 (1978).
77. F. White, The taxonomy, ecology and chorology of African Ebenaceae II. The non-Guineo-Congolian species of Diospyros (Excluding sect. Royena). *Bull. Jard. Bot. Natl. Belg. Bull. Van Natl. Plantentuin Van Belg.* **58**, 325–448 (1988).
78. A. Fayolle *et al.*, Patterns of tree species composition across tropical African forests. *J. Biogeogr.* **41**, 2320–2331 (2014).
79. R. A. Segovia *et al.*, Freezing and water availability structure the evolutionary diversity of trees across the Americas. *Sci. Adv.* **6**, eaaz5373 (2020).
80. D. M. Neves *et al.*, Evolutionary diversity in tropical tree communities peaks at intermediate precipitation. *Sci. Rep.* **10**, 1188 (2020).
81. P. L. Silva de Miranda *et al.*, Using tree species inventories to map biomes and assess their climatic overlaps in lowland tropical South America. *Glob. Ecol. Biogeogr.* **27**, 899–912 (2018).
82. K. J. Feeley, J. T. Stroud, Where on Earth are the "tropics"? *Front. Biogeogr.* **10**, e38649 (2018).
83. B. A. Hawkins *et al.*, Structural bias in aggregated species-level variables driven by repeated species co-occurrences: A pervasive problem in community and assemblage data. *J. Biogeogr.* **44**, 1199–1211 (2017).
84. P. Koleff, K. J. Gaston, J. J. Lennon, Measuring beta diversity for presence–absence data. *J. Anim. Ecol.* **72**, 367–382 (2003).
85. A. Baselga, Partitioning the turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.* **19**, 134–143 (2010).
86. H. Kreft, W. Jetz, A framework for delineating biogeographical regions based on species distributions: Global quantitative biogeographical regionalizations. *J. Biogeogr.* **37**, 2029–2053 (2010).
87. A. J. Aberer, D. Krompass, A. Stamatakis, Pruning rogue taxa improves phylogenetic accuracy: An efficient algorithm and webservice. *Syst. Biol.* **62**, 162–166 (2013).
88. L. Dapporto *et al.*, recluster: An unbiased clustering procedure for beta-diversity turnover. *Ecography* **36**, 1070–1075 (2013).
89. J. Chave *et al.*, Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Change Biol.* **20**, 3177–3190 (2014).
90. J. Thioulouse *et al.*, *Multivariate Analysis of Ecological Data with ade4* (Springer New York, 2018).
91. A. Kassambara, F. Mundt, *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*, R package version 1.0.7. https://CRAN.R-project.org/package=factoextra (2020).
92. A. T. Oliveira-Filho, Data from "NeoTropTree, Flora arbórea da Região Neotropical: Um banco de dados envolvendo biogeografia, diversidade e conservação." Universidade Federal de Minas Gerais. Available at http://www.neotroptree.info. Deposited 22 November 2017.