

## Global hotspots of plant phylogenetic diversity

Melanie Tietje<sup>1</sup>, Alexandre Antonelli<sup>2,3,4</sup>, Félix Forest<sup>2</sup>, Rafaël Govaerts<sup>2</sup>,  
Stephen A. Smith<sup>5</sup>, Miao Sun<sup>6</sup>, William J. Baker<sup>2,8</sup>, Wolf L. Eiserhardt<sup>1,2,7,8</sup>

<sup>1</sup> Department of Biology, Aarhus University, Denmark

<sup>2</sup> Royal Botanic Gardens, Kew, Richmond, Surrey, UK

<sup>3</sup> Department of Plant Sciences, University of Oxford, UK

<sup>4</sup> Gothenburg Global Biodiversity Centre, University of Gothenburg, Sweden

<sup>5</sup> Department of Ecology & Evolutionary Biology, University of Michigan, Ann Arbor, USA

<sup>6</sup> National Key Laboratory for Germplasm Innovation & Utilization of Horticultural Crops, Huazhong Agriculture University, China

<sup>7</sup> Aarhus Institute of Advanced Studies, Aarhus University, Denmark

<sup>8</sup> shared senior authors

## Summary

- Regions harbouring high unique phylogenetic diversity (PD) are priority targets for conservation. Here, we analyse the global distribution of plant PD, which remains poorly understood despite plants being the foundation of most terrestrial habitats and key to human livelihoods.
- Capitalising on a recently completed, comprehensive global checklist of vascular plants, we identify hotspots of unique plant PD and test three hypotheses: 1) PD is more evenly distributed than species diversity; 2) areas of highest PD (often called “hotspots”) do not maximise cumulative PD; 3) many biomes are needed to maximise cumulative PD.
- Our results support all three hypotheses: more than twice as many regions are required to cover 50% of global plant PD compared to 50% of species; regions that maximise cumulative PD substantially differ from the regions with outstanding individual PD; and while (sub-)tropical moist forest regions dominate across PD hotspots, other forest types and open biomes are also essential.
- Safeguarding PD in the Anthropocene (including the protection of some comparatively species-poor areas) is a global, yet unrecognised responsibility. Having highlighted countries with outstanding unique plant PD, further analyses are now required to fully understand the global distribution of plant PD and associated conservation imperatives across spatial scales.

**Key words:** seed plants, phylogenetic diversity, complementarity, hotspot, phylogenetic endemism, conservation, macroecology

## Introduction

Compared to species diversity patterns, the global distribution of phylogenetic diversity (PD) is poorly understood. Phylogenetic diversity is a more meaningful proxy for ‘feature diversity’ than simple species counts (Faith, 1992). By accounting for the total amount of phylogenetic history represented in a biota, PD (Faith, 1992) broadly captures diversity in form and function (Srivastava *et al.*, 2012; Tucker *et al.*, 2017; Owen *et al.*, 2019), and is connected to the resilience of key ecosystem functions (Mazzochini *et al.*, 2019) and services (Forest *et al.*, 2007; Molina-Venegas *et al.*, 2021). Phylogenetic diversity has been established explicitly as a guiding measure for conservation efforts, i.e. to identify sets of areas that jointly carry the most

PD and hence should be priorities for conservation (Faith, 1992; Winter *et al.*, 2013; Véron *et al.*, 2019). However, this requires a robust knowledge of the spatial distribution of PD.

Crucially, the regions that are richest in species or PD are not necessarily the regions of biggest conservation concern. Instead, conservation tends to focus on hotspots of *unique* (endemic) diversity, i.e. regions that harbour many species or lineages that occur nowhere else, and consequently cannot be conserved anywhere else. For example, the classic *biodiversity hotspots for conservation priorities* (Myers *et al.*, 2000) were selected to contain at least 1,500 endemic plant species each (as well as high levels of anthropogenic threat). A similar train of thought can be applied to PD, using metrics such as PD endemism (Faith, 1994) or Phylogenetic Endemism (Rosauer *et al.*, 2009). Alternatively, complementarity analyses can be used to identify the set of areas that most efficiently maximises PD (Faith *et al.*, 2003; Kukkala & Moilanen, 2013). Such analyses can show in which parts of the world extinction would cause the largest total loss of phylogenetic (or evolutionary) history.

Considering that plants are the trophic and structural foundation of most terrestrial habitats and hence also a cornerstone of human livelihoods, exploring and explaining the distribution of their diversity is clearly a priority. However, due to limitations in both geographic and phylogenetic data (Meyer *et al.*, 2016; Rudbeck *et al.*, 2022), global plant PD remains incompletely understood. This is in stark contrast to most vertebrates, which are well-served with geographic and phylogenetic data that have facilitated thorough exploration of their global PD patterns (Davies & Buckley, 2011; Safi *et al.*, 2011; Fritz & Rahbek, 2012; Voskamp *et al.*, 2017; Daru *et al.*, 2019; Gumbs *et al.*, 2020). Previous studies on global plant PD were either limited in their taxonomic sampling and resolution (Daru *et al.*, 2019), or based on incompletely documented, closed-access datasets (Qian *et al.*, 2023). For example, (Daru *et al.*, 2019) mapped PD and other diversity metrics for 6,483 (*ca.* 50%) currently accepted angiosperm genera, revealing a highly uneven distribution of angiosperm PD with hotspots in parts of South and Central America, Madagascar, Southeast Asia and Australasia. Whether species-level analyses with a more comprehensive taxon sampling will challenge or confirm these highly valuable first insights remains to be seen. Although genus- and species-level PD patterns are similar in vertebrates (Daru *et al.*, 2019) and North American angiosperm trees (Qian & Jin, 2021), this might not be the case in all plants, as also zoology and botany diverge considerably in their use of higher taxonomic ranks, with plant genera often being larger and more variable in size (Frodin, 2004; Sigwart *et al.*, 2017). Additionally, it is also important to note that neither of these studies adequately addressed complementarity from a conservation perspective.

To establish the distribution of plant PD, its relationship to species richness, and its conservation requirements, we frame three hypotheses. First, we hypothesise that PD is more evenly distributed across the surface of the Earth than species diversity (H1). While the world's plant species are concentrated in relatively few "hyperdiverse" regions (Pimm & Joppa, 2015), PD is unlikely to behave in the same way because closely related species and lineages are spatially autocorrelated (i.e. occur in close proximity at a global scale) (Kissling *et al.*, 2012b; Eiserhardt *et al.*, 2013). Typically, "hyperdiverse" regions are dominated by local diversification and therefore include many close relatives (e.g. (Schnitzler *et al.*, 2011; Hughes & Atchison, 2015). Thus, while hyperdiverse regions capture species richness and shallow phylogenetic history, many different regions are required to represent deep phylogenetic history and maximise PD. This challenges the prevailing worldview that places conservation responsibility primarily with countries that harbour "hyperdiverse" biota.

Second, we hypothesise that the areas of highest PD, which are often designated as hotspots (e.g. Daru *et al.* 2019; Qian *et al.* 2023) do not maximise cumulative PD, and thus do not effectively guide conservation attention (H2). Because of the spatial autocorrelation of diversification, areas of highest PD are often geographically adjacent and redundant in their

composition of species and lineages. We expect that hotspots based on complementarity, i.e. explicitly maximising the cumulative PD they represent, will be substantially different from the areas that individually have the highest PD.

Third, we hypothesise (H3) that absolute phylogenetic diversity is highest in tropical and subtropical moist broadleaf forest, where species richness is overall high and diversification rates low (Igea & Tanentzap, 2020; Sun *et al.*, 2020; Tietje *et al.*, 2022) and many old lineages persist (tropical conservatism hypothesis, Wien & Donoghue (2004). However, as different biomes represent different lineages (biome conservatism hypothesis, Crisp *et al.* (2009), we expect that the highest values of PD would be attained by regions with a high proportion of tropical and subtropical moist broadleaf forest, as well as elements of other evolutionarily divergent biomes. Along a similar vein, we expect a wider range of biomes to be important in complementarity-based hotspots compared to high-PD regions, in line with the idea that diversification is not only spatially, but also environmentally autocorrelated (Wiens & Donoghue, 2004; Crisp *et al.*, 2009; Eiserhardt *et al.*, 2013).

Here, we test these hypotheses in a global analysis of phylogenetic diversity focused on seed plants (Spermatophyta). With >330,000 accepted species, seed plants constitute >90% of extant land plant diversity and dominate almost all terrestrial vegetation. Our analyses capitalise on the recently completed World Checklist of Vascular Plants (Govaerts *et al.*, 2021), a publicly accessible, comprehensive taxonomic checklist of vascular plants and their geographic distributions, in conjunction with a complete phylogeny of seed plants (Smith & Brown 2018). We extend our test of these hypotheses with a high-level evaluation of global threats (deforestation, human footprint, climate change), to facilitate comparison with existing literature on conservation hotspots, which have traditionally accounted for degree of threat (Myers *et al.*, 2000). By integrating these resources within a complementarity framework for the first time, we establish a prioritised set of regions for the conservation of global plant phylogenetic diversity.

## Materials and Methods

### *Phylogeny and distribution data*

All phylogenetic measures were derived from the phylogenetic tree of Smith & Brown (2018) to which we added missing species using TACT (Chang *et al.*, 2020). Of the different trees provided by Smith & Brown (2018), we used the one that contained only species with molecular data and a backbone from Magallón *et al.* (2015; GBMB). TACT adds missing species to a time-calibrated phylogeny using a taxonomic guide tree and birth-death models to estimate branching times. Since TACT has a stochastic component, we used averages across 100 replicate phylogenetic trees produced by TACT in all our analyses.

The geographic distribution of each species was derived from the World Checklist of Vascular Plants (Govaerts *et al.*, 2021), which provides presence and absence data for each World Geographical Scheme for Recording Plant Distributions (WGSRPD) level 3 unit (hereafter “botanical country”, (Brummitt *et al.*, 2001). These botanical country names follow in some cases alternative spellings (e.g. Sumatera for Sumatra), we follow here the names as provided in Brummitt *et al.* (2001).

Species names in the Smith & Brown phylogeny follow NCBI nomenclature. These were updated to follow the WCVF nomenclature using the taxonomy matching procedure of (Sun *et al.*, 2021), which uses the WCVF taxonomy data as authority. Our data includes 330,527 described species of seed plants. Bryophytes, clubmosses and ferns are not included in our analysis as geographic and/or phylogenetic data were unavailable.

156 Diversity indices were calculated for each botanical country. Species richness was measured as  
157 the number of species recorded in a botanical country. We estimated phylogenetic diversity  
158 (PD) as well as three measures of endemism, including PD endemism, phylogenetic endemism  
159 and weighted endemism. PD was calculated as the sum of the lengths of all branches that span  
160 members of a region (Faith, 1992), PD endemism is calculated as total amount of branch length  
161 found only in a given region (Faith, 1994), phylogenetic endemism is similar to PD but inversely  
162 weighted by species ranges (Rosauer *et al.*, 2009), and weighted endemism is species richness  
163 inversely weighted by species ranges. Indices were estimated using the R package *phyloregion*  
164 (Daru *et al.*, 2020b) functions *PD*, *phylo\_endemism* and *weighted\_endemism*. We calculated  
165 spatial correlations of diversity indices using Lee's L, an integration of Pearson's *r* and Moran's *I*  
166 (Lee, 2001).

## 167 *Hotspots*

168 We identified PD hotspots following two different approaches. First, we identified botanical  
169 countries with the highest 2.5% estimates for species richness, PD and PD endemism.  
170 Choosing the top 2.5% as hotspots is an approach that has been implemented by previous  
171 authors (e.g (Orme *et al.*, 2005; Daru *et al.*, 2019) and focuses on the total estimates for a  
172 specific area, independent from other regions. For our 368 botanical countries, the top 2.5%  
173 corresponds to the top 9.2 botanical countries, which we round up to top ten for simplicity.  
174 Second, we identified each country's contribution (=complementarity) to global species  
175 richness, PD and PD endemism. To assess complementarity we used a greedy algorithm that  
176 starts with the botanical country that has the highest PD value (or other estimate of interest) and  
177 sequentially adds botanical countries, in each step choosing the country that adds most PD to  
178 the total PD in the set. Once a group of species has been covered by adding a country to the  
179 set, these species are exempt from calculations of PD for the remaining countries. This way the  
180 algorithm identified the minimum number of countries that together contained the maximum  
181 amount of PD, which is of particular importance for conservation as it allows the optimisation of  
182 diversity in as small an area as possible. This approach focuses on the contribution of each  
183 region to the global PD, avoiding redundancy. We selected the same number of countries (ten)  
184 as was included in the top 2.5% approach to PD hotspots to facilitate comparison.

185

## 186 *Hotspot characteristics*

187 For each botanical country we calculated several anthropogenic and environmental  
188 characteristics (see for sources Table S1). We estimated the percent coverage of each  
189 hotspots' area with different biomes (Olson *et al.*, 2001) and the average impact of  
190 deforestation, human footprint index and future climate change as average threat values per  
191 botanical country. We collected diversification rates (mean root distance; the average number of  
192 edges from tip to root in a phylogeny of all species occurring in a botanical country) for each  
193 botanical country from (Tietje *et al.*, 2022). Quantitative differences in threats and diversification  
194 rates between hotspot and non-hotspot botanical countries were tested using Kruskal-Wallis  
195 rank sum tests.

196

197 Analyses were done in R version 4.2.1 (R Core Team, 2022). R packages used include  
198 *data.table* (Dowle & Srinivasan, 2021), *sf* (Pebesma, 2018), *phyloregion* (Daru *et al.*, 2020b),  
199 *terra* (Hijmans, 2022a), *ggplot2* (Wickham, 2016), *cowplot* (Wilke, 2020), *raster* (Hijmans,  
200 2022b), *exactextractr* (Daniel Baston, 2022), *castor* (Louca & Doebeli, 2018), *stringr* (Wickham,  
201 2022), *spdep* (Bivand & Wong, 2018).

202

Phylogenetic diversity (PD) was strongly correlated with species richness when accounting for spatial autocorrelation (Fig. 1a,b, Lee's  $L = 0.86$ ;  $P = 0.001$ ). However, the top 2.5% countries differed clearly between species richness and PD (Fig. 2a,b). While the majority of highest species richness values were concentrated in the Neotropics, the top 2.5% PD values were more evenly distributed between South America and parts of southern Asia. The highest PD values were found in Colombia followed by China South-Central and Peru, whereas Antarctica and small islands had the lowest PD values. These low PD regions were also characterised by low species richness (Fig. 1a,b).

The ten botanical countries that were selected based on complementarity ("complementarity hotspots") were clearly different from the ten botanical countries that had the highest individual diversities, both for species richness (Fig. 2a,c) and PD (Fig. 2b,d). Importantly, the former harboured a higher total diversity than the latter (40% vs. 33.5% for species richness and 23% vs. 19% for PD). It is noteworthy that ten botanical countries can cover 40% of global species richness, but only 23% of global PD. This difference became even more apparent when comparing numbers of countries required to contain fixed proportions of global diversity. For example, while 50% of global species richness could be included in 15 botanical countries (Fig. 3a), a minimum number of 33 countries was required to cover 50% of PD (Fig. 3b). This pattern was consistent across diversity thresholds between 10 and 90 percent (Fig. 4), showing that species richness could be captured in comparatively few areas, whereas PD was more evenly distributed.

The complementarity hotspots of PD (Fig. 2d, Table 2) were almost identical to the complementarity hotspots of species richness (Fig. 2c), only differing in the inclusion of Western Australia (species richness) and Zaïre (PD). Complementarity hotspots of PD were widespread across Central- and South America, Africa, China, Madagascar, Borneo and New Guinea. They showed a significantly higher biome diversity than non-hotspot countries (Kruskal-Wallis rank sum test,  $P < 0.001$ ). These hotspots also showed higher biome coverage proportions with (sub)tropical moist and dry broadleaf forest as well as montane grasslands and shrublands than their non-hotspot counterparts (Kruskal-Wallis rank sum test,  $P < 0.005$ , Fig. S2a). Both patterns were similar for countries selected for highest 2.5% PD values (Fig. S2b). These results were consistent with the positive correlation of biome types and total PD observed (Spearman's rank correlation  $\rho = 0.57$ ,  $P < 0.001$ , Fig. S3). Diversification rates did not differ significantly between PD complementarity hotspots and other countries, but were lower in top 2.5% PD countries than in other countries (Fig. S4).

Anthropogenic and environmental characteristics varied substantially between PD complementarity hotspots (Fig. 5). Complementarity hotspots of PD had a significantly larger area affected by deforestation compared to non-hotspots (Fig. S5, Kruskal-Wallis rank sum test,  $P = 0.016$ ). Borneo was particularly strongly affected by deforestation. In general, the relative area affected by deforestation within the last 20 years varied greatly from 47% on Borneo to 2% in the Cape Provinces and Australia. We found no significant difference between hotspots and non-hotspots for other threats (Fig. S5).

Global patterns of PD endemism were correlated with species richness (Lee's  $L = 0.42$ ;  $p = 0.003$ ; Fig. 1a,c), and closely matched patterns of phylogenetic endemism and weighted endemism (Fig. S1). Among the botanical countries harbouring most PD endemism were the Cape Provinces, followed by Western Australia, Madagascar and Borneo (Fig. 1c). Relatively few countries were required to cover a given proportion of global PD endemism compared to PD and species richness (Fig. 4); e.g., only twelve botanical countries were required to cover 50% of global PD endemism.

## Discussion

Using a recently completed, comprehensive, open access dataset of the taxonomy and geographic distributions of all vascular plant species (Govaerts *et al.*, 2021), we dissected the distribution of global seed plant phylogenetic diversity (PD) using a complementarity-based approach. We found that 1) PD is more evenly distributed across the globe than species richness, 2) absolute PD (Daru *et al.*, 2019; Qian *et al.*, 2023) is no substitute for cumulative PD derived from complementarity-based analyses, and 3) tropical rain forests are important for sustaining high levels of PD, but a variety of biomes are implicated in the conservation of global seed plant PD.

### *Phylogenetic diversity is more evenly distributed across the globe than species richness*

Our results support hypothesis H1, demonstrating that, for seed plants, PD is more evenly distributed across the globe than species richness. Because PD increases more slowly with area than species richness (Connor & McCoy, 1979; Rosenzweig, 1995; Morlon *et al.*, 2011; Helmus & Ives, 2012), more than twice as many botanical countries are needed to represent 50% of global PD than to achieve the same for species richness (Fig. 3). This discrepancy is due to the spatial autocorrelation of diversification. If the area of a given sampling unit is increased, new species are added, but those species are likely close relatives to the ones that are already in the set. Biologically, this pattern results from the limited niche evolution and dispersal of diversifying clades (Wiens *et al.*, 2010; Eiserhardt *et al.*, 2013). Well-known examples of this phenomenon are local radiation events such as those driven by the uplift of the Andes (Hughes & Eastwood, 2006; Pérez-Escobar *et al.*, 2017), where species-rich regions contain disproportionately many closely related species. These results highlight the risks of focusing purely on species richness in area prioritisation for conservation (Rodrigues *et al.*, 2005).

### *Absolute phylogenetic diversity is no substitute for complementarity*

In line with our hypothesis H2, we found that hotspots selected for PD complementarity outperformed hotspots selected for highest total PD in representing global diversity. The two approaches selected substantially different sets of regions. Evidently, many of the regions of highest PD were not only spatially adjacent, but also significantly redundant in their composition of species and lineages. For example, the botanical countries with the highest PD included clusters of adjacent countries in north-western South America and continental Asia (Fig. 2b). The complementarity approach showed that several of these were redundant in maximising PD in 10 botanical countries, instead highlighting the importance of Sub-Saharan Africa and Australasia for global PD. Of note, these areas were also identified by Qian *et al.* (2023) by ranking absolute PD, but using a much lower threshold (10%), leading to many more regions being recognized as hotspots, which makes prioritisation more difficult. Focusing on the areas with the highest individual PD as “hotspots” (Daru *et al.*, 2019; Qian *et al.*, 2023) risks attention being diverted from parts of the world that are essential to safeguarding seed plant PD globally.

### *Tropical rain forests are important, but a variety of biomes are required to conserve global PD*

The distribution of high-PD regions and complementarity hotspots (Table 1) across biomes largely supports our hypothesis H3. As anticipated, regions of high individual PD had high

coverage of (sub-)tropical moist broadleaf forest (Table 1, Fig. S3). Due to biome conservatism (Crisp *et al.*, 2009), we also expected PD to be highest in regions that harbour many other biomes in addition to (sub-)tropical moist broadleaf forest. This expectation was also confirmed (Fig. S3). However, the biome composition of high-PD regions (Table 1) suggests that the type of biomes included matters too. High-PD regions typically include several other forest types, specifically (sub-)tropical dry forest, (sub-)tropical coniferous forest, or temperate broadleaf or coniferous forest. Whether this is because forest generally harbours older and/or more divergent lineages than open vegetation, or because these forest biomes just happen to be spatially adjacent to the phylogenetically highly diverse (sub-)tropical moist broadleaf forest, is unclear and worthy of further study. Our expectation that complementarity hotspots jointly cover a wider range of biomes than high-PD regions is primarily supported by the observation that complementarity hotspots encompass more open biomes than the primarily forested high-PD regions. This is particularly evident from the inclusion of the Cape Provinces of South Africa, which are exclusively covered by open biomes (at the scale of the biome maps used here). This confirms that while (sub-)tropical moist broadleaf forest is, on its own, the most phylogenetically diverse biome, other ecologically and evolutionarily divergent biomes are required to appropriately represent global seed plant PD.

### *Threats to phylogenetic diversity*

Countries identified as PD complementarity hotspots were not consistently more or less threatened by human impact than non-hotspot areas except for deforestation, which affected hotspots more strongly. The deforestation result is plausible, since hotspots also showed on average larger proportions of (sub-)tropical moist broadleaf forest, with the notable exception of Cape Provinces. This biome is known to be under intense deforestation pressure (Lindquist *et al.*, 2012). Human footprint did not show clear results, possibly due to the large spatial scale our study was conducted on, which averages footprint over large areas, not differentiating between heavily affected urban areas and remote untouched landscapes. Since the majority of high diversity areas were located at low latitudes, anticipated future climate changes were naturally rather low since the absolute extent of climate change is predicted to be larger near the poles (Rantanen *et al.*, 2022). Threat status, usually represented as some form of habitat loss of a region, has been used as one of the defining criteria of conservation hotspots (Myers *et al.*, 2000). However, the example of New Guinea with its extraordinary flora, high contribution to global PD, but insufficient threat to qualify as conservation hotspot demonstrates potential challenges with the inclusion of threat in hotspot criteria (Cámara-Leret *et al.*, 2020), especially since anthropogenic habitat loss can be rapid (Gaveau *et al.*, 2014; Gamoga *et al.*, 2021). Hence, we define hotspots solely based on their contribution to global PD, using threats as an additional layer of information to inform conservation prioritisation.

### *Conservation prioritisation*

We believe that safeguarding phylogenetic diversity in the Anthropocene is a global responsibility. Attention is often focused on exceptionally species-rich regions, such as the Neotropics (Antonelli & Sanmartín, 2011; Cazzolla Gatti *et al.*, 2022), which also are highlighted by our analyses of absolute, country-level PD. However, complementarity-based analyses clearly show that many more biogeographic regions and biomes are needed for effective conservation of global PD, and hence global feature diversity. This includes regions of the world that are known for their comparatively low species richness, such as parts of the African continent (Couvreur, 2015). Because our analyses are conducted at the scale of botanical countries, most of which correspond to political or administrative units (Brummitt *et al.*, 2001), our findings can in principle directly inform conservation policy in those units. In our view, this advantage of botanical countries outweighs their disadvantage of being variable in size, which complicates their use in ecological and evolutionary research (but see e.g. (Kissling *et al.*, 2012a; Tietje *et al.*, 2022; Guo *et al.*, 2022) which is, however, less problematic in a

conservation context. While larger (botanical) countries are more likely to be identified as priority regions for conservation, these do also in fact hold larger proportions of global PD. Of note, the countries that were most implicated in the conservation of global PD were largely unaffected by the somewhat arbitrary division of the largest countries into lower-level administrative units (Table S2, Fig. S7–9). Importantly, we stress that countries that are not selected in our complementarity analysis can still play a major role in the conservation of global plant PD, as their PD may largely overlap with adjacent, selected countries. Not being flagged as a priority country for global PD complementarity thus does not imply that a country has no role to play in conservation.

## Prospects

Our findings are a first pass at revealing not only where centres of plant PD are located, but also how broadly plant PD is distributed across the planet. By taking complementarity into account and using a taxonomically comprehensive, open access dataset, our analysis goes substantially beyond previous studies (Daru *et al.*, 2019; Qian *et al.*, 2023). However, due to data availability, our analysis is limited to a relatively large spatial scale, and further studies are warranted to fully explore how the distribution and complementarity of plant PD is affected by spatial scaling (Morlon *et al.*, 2011; Helmus & Ives, 2012; Daru *et al.*, 2020a). While analyses at finer spatial resolution are needed to more accurately pinpoint the locations of PD hotspots, these may not necessarily improve insights from PD complementarity for conservation prioritisation (Daru *et al.*, 2019). Analyses at intermediate spatial resolution using units that do not vary in area could provide invaluable insights. However, the point occurrence datasets needed for such analyses are notoriously incomplete and biased (Meyer *et al.*, 2016), making such an analysis impractical for the time being and underlining the value of the taxonomically and geographically complete dataset used here. Further investment in the assembly of global distribution datasets is clearly needed.

Phylogenetic diversity provides substantially deeper insights into total “feature diversity” than simple species counts (Faith, 1994), capturing both the evolutionary past and possible future evolutionary potential in a region. However, it is no panacea for conservation prioritisation, which requires a full grasp of trade-offs between different diversity measures such as basic species count, phylogenetic diversity metrics, functional and trait diversity, and anthropogenic factors, including rapid changes in land use and the ubiquitous effects of climate change and plant distribution and diversity. We do not intend to judge the relative importance of these factors or suggest that PD is the most relevant. Instead, integrating multiple viewpoints, as suggested by (Pavoine & Bonsall, 2011) might be the most appropriate way to explore and understand the current biodiversity crisis and its potential future implications. We also acknowledge that complementarity analysis in conservation practice requires multivariate optimisation processes guided by spatial, political, financial and social aspects and limitations (Sarkar *et al.*, 2006; Kukkala & Moilanen, 2013). Thus, while our results provide important insights into the global distribution of plant PD, future studies need to integrate these findings with other aspects of plant diversity as well as the political and socioeconomic context of real-world conservation.

## Acknowledgements

MT’s postdoc was funded by VILLUM FONDEN (grant 00025354 to WLE) and an Aarhus University Research Foundation research grant (to WLE). AA acknowledges financial support from the Swedish Research Council (2019-05191) and the Royal Botanic Gardens, Kew. SAS was supported by the NSF under grant agreement DBI 1930030. Some of the computing for this project was performed on the GenomeDK cluster.



## Author contributions

MT, WJB and WLE conceived the research ideas and designed the project. RG collected the data. MT analysed the data. MT, AA, FF, WJB, SAS and WLE interpreted the results. MT, WJB and WLE wrote the manuscript with contributions by AA, SAS, FF and in consultation with RG and MS.

## Data availability

Code and data to repeat the analysis of this study are available in Zenodo at [[link to zenodo repository](https://github.com/Eryops1/phylogenetic_hotspots) upon publication. Currently a git repo: [https://github.com/Eryops1/phylogenetic\\_hotspots](https://github.com/Eryops1/phylogenetic_hotspots)]. All data sources are referenced in Table S1.

## References

- Antonelli A, Sanmartín I. 2011.** Why are there so many plant species in the Neotropics? *Taxon* **60**: 403–414.
- Bivand RS, Wong DWS. 2018.** Comparing implementations of global and local indicators of spatial association. *Test* **27**: 716–748.
- Brummitt RK, Pando F, Hollis S, Brummitt NA. 2001.** *World geographical scheme for recording plant distributions*. International working group on taxonomic databases for plant sciences (TDWG ....
- Cámara-Leret R, Frodin DG, Adema F, Anderson C, Appelhans MS, Argent G, Arias Guerrero S, Ashton P, Baker WJ, Barfod AS, et al. 2020.** New Guinea has the world's richest island flora. *Nature* **584**: 579–583.
- Cazzolla Gatti R, Reich PB, Gamarra JGP, Crowther T, Hui C, Morera A, Bastin J-F, de-Miguel S, Nabuurs G-J, Svenning J-C, et al. 2022.** The number of tree species on Earth. *Proceedings of the National Academy of Sciences of the United States of America* **119**.
- Chang J, Rabosky DL, Alfaro ME. 2020.** Estimating Diversification Rates on Incompletely Sampled Phylogenies: Theoretical Concerns and Practical Solutions. *Systematic biology* **69**: 602–611.
- Connor EF, McCoy ED. 1979.** The statistics and biology of the species-area relationship. *The American naturalist* **113**: 791–833.
- Couvreur TLP. 2015.** Odd man out: why are there fewer plant species in African rain forests? *Plant systematics and evolution = Entwicklungsgeschichte und Systematik der Pflanzen* **301**: 1299–1313.
- Crisp MD, Arroyo MTK, Cook LG, Gandolfo MA, Jordan GJ, McGlone MS, Weston PH, Westoby M, Wilf P, Linder HP. 2009.** Phylogenetic biome conservatism on a global scale. *Nature* **458**: 754–756.
- Daniel Baston. 2022.** exactextractr: Fast Extraction from Raster Datasets using Polygons.
- Daru BH, Farooq H, Antonelli A, Faurby S. 2020a.** Endemism patterns are scale dependent. *Nature communications* **11**: 2115.
- Daru BH, Karunarathne P, Schliep K. 2020b.** phyloregion: R package for biogeographical regionalization and macroecology. *Methods in ecology and evolution / British Ecological Society* **11**: 1483–1491.
- Daru BH, le Roux PC, Gopalraj J, Park DS, Holt BG, Greve M. 2019.** Spatial overlaps between the global protected areas network and terrestrial hotspots of evolutionary diversity. *Global ecology and biogeography: a journal of macroecology* **28**: 757–766.
- Davies TJ, Buckley LB. 2011.** Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present-day richness gradients for mammals. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **366**: 2414–2425.
- Dowle M, Srinivasan A. 2021.** *data.table: Extension of 'data.frame'*.
- Eiserhardt WL, Svenning J-C, Baker WJ, Couvreur TLP, Balslev H. 2013.** Dispersal and niche

455 evolution jointly shape the geographic turnover of phylogenetic clades across continents. *Scientific*  
456 *reports* **3**: 1164.

457 **Faith DP. 1992.** Conservation evaluation and phylogenetic diversity. *Biological conservation* **61**: 1–10.

458 **Faith DP. 1994.** Phylogenetic diversity: a general framework for the prediction of feature diversity.

459 **Faith DP, Carter G, Cassis G, Ferrier S, Wilkie L. 2003.** Complementarity, biodiversity viability analysis,  
460 and policy-based algorithms for conservation. *Environmental science & policy* **6**: 311–328.

461 **Forest F, Grenyer R, Rouget M, Davies TJ, Cowling RM, Faith DP, Balmford A, Manning JC,**  
462 **Procheş S, van der Bank M, et al. 2007.** Preserving the evolutionary potential of floras in biodiversity  
463 hotspots. *Nature* **445**: 757–760.

464 **Fritz SA, Rahbek C. 2012.** Global patterns of amphibian phylogenetic diversity. *Journal of biogeography*  
465 **39**: 1373–1382.

466 **Frodin DG. 2004.** History and concepts of big plant genera. *Taxon* **53**: 753–776.

467 **Gamoga G, Turia R, Abe H, Haraguchi M, Iuda O. 2021.** The Forest extent in 2015 and the drivers of  
468 forest change between 2000 and 2015 in Papua New Guinea. *Case studies in the environment* **5**.

469 **Gaveau DLA, Sloan S, Molidena E, Yaen H, Sheil D, Abram NK, Ancrenaz M, Nasi R, Quinones M,**  
470 **Wieland N, et al. 2014.** Four decades of forest persistence, clearance and logging on Borneo. *PloS one*  
471 **9**: e101654.

472 **Govaerts R, Nic Lughadha E, Black N, Turner R, Paton A. 2021.** The World Checklist of Vascular  
473 Plants, a continuously updated resource for exploring global plant diversity. *Scientific data* **8**: 215.

474 **Gumbs R, Gray CL, Böhm M, Hoffmann M, Grenyer R, Jetz W, Meiri S, Roll U, Owen NR, Rosindell**  
475 **J. 2020.** Global priorities for conservation of reptilian phylogenetic diversity in the face of human impacts.  
476 *Nature communications* **11**: 2616.

477 **Guo Q, Qian H, Zhang J. 2022.** On the relationship between species diversity and range size. *Journal of*  
478 *biogeography* **49**: 1911–1919.

479 **Helmus MR, Ives AR. 2012.** Phylogenetic diversity–area curves. *Ecology* **93**: S31–S43.

480 **Hijmans RJ. 2022a.** terra: Spatial Data Analysis.

481 **Hijmans RJ. 2022b.** raster: Geographic Data Analysis and Modeling.

482 **Hughes CE, Atchison GW. 2015.** The ubiquity of alpine plant radiations: from the Andes to the  
483 Hengduan Mountains. *The New phytologist* **207**: 275–282.

484 **Hughes C, Eastwood R. 2006.** Island radiation on a continental scale: exceptional rates of plant  
485 diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences of the United*  
486 *States of America* **103**: 10334–10339.

487 **Igea J, Tanentzap AJ. 2020.** Angiosperm speciation cools down in the tropics. *Ecology letters* **23**: 692–  
488 700.

489 **Kissling WD, Baker WJ, Balslev H, Barfod AS, Borchsenius F, Dransfield J, Govaerts R, Svenning**  
490 **J-C. 2012a.** Quaternary and pre-Quaternary historical legacies in the global distribution of a major tropical  
491 plant lineage. *Global ecology and biogeography: a journal of macroecology* **21**: 909–921.

492 **Kissling WD, Eiserhardt WL, Baker WJ, Borchsenius F, Couvreur TLP, Balslev H, Svenning JC.**  
493 **2012b.** Cenozoic imprints on the phylogenetic structure of palm species assemblages worldwide.  
494 *Proceedings of the National Academy of Sciences of the United States of America* **109**: 7379–7384.

495 **Kukkala AS, Moilanen A. 2013.** Core concepts of spatial prioritisation in systematic conservation  
496 planning. *Biological reviews of the Cambridge Philosophical Society* **88**: 443–464.

497 **Lee S-I. 2001.** Developing a bivariate spatial association measure: An integration of Pearson's r and  
498 Moran's I. *Journal of geographical systems* **3**: 369–385.

499 **Lindquist EJ, D'Annunzio R, Gerrand A, MacDicken K, Achard F, Beuchle R, Brink A, Eva HD,**  
500 **Mayaux P, San-Miguel-Ayanz J, et al. 2012.** *Global forest land-use change 1990 - 2005*. Food and  
501 Agriculture Organization of the United Nations and European Commission Joint Research Centre. Rome,  
502 FAO.

503 **Louca S, Doebeli M. 2018.** Efficient comparative phylogenetics on large trees. *Bioinformatics* **34**: 1053–  
504 1055.

505 **Mazzochini GG, Fonseca CR, Costa GC, Santos RM, Oliveira-Filho AT, Ganade G. 2019.** Plant  
506 phylogenetic diversity stabilizes large- scale ecosystem productivity. *Global ecology and biogeography: a*  
507 *journal of macroecology* **28**: 1430–1439.

508 **Meyer C, Weigelt P, Kreft H. 2016.** Multidimensional biases, gaps and uncertainties in global plant  
509 occurrence information. *Ecology letters* **19**: 992–1006.

510 **Molina-Venegas R, Rodríguez MÁ, Pardo-de-Santayana M, Ronquillo C, Mabberley DJ. 2021.**  
511 Maximum levels of global phylogenetic diversity efficiently capture plant services for humankind. *Nature*  
512 *ecology & evolution* **5**: 583–588.

513 **Morlon H, Schwikl DW, Bryant JA, Marquet PA, Rebelo AG, Tauss C, Bohannan BJM, Green JL.**  
514 **2011.** Spatial patterns of phylogenetic diversity. *Ecology letters* **14**: 141–149.

515 **Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GA, Kent J. 2000.** Biodiversity hotspots for  
516 conservation priorities. *Nature* **403**: 853–858.

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

520 **Orme CDL, Davies RG, Burgess M, Eigenbrod F, Pickup N, Olson VA, Webster AJ, Ding T-S,**  
521 **Rasmussen PC, Ridgely RS, et al. 2005.** Global hotspots of species richness are not congruent with  
522 endemism or threat. *Nature* **436**: 1016–1019.

523 **Owen NR, Gumbs R, Gray CL, Faith DP. 2019.** Global conservation of phylogenetic diversity captures  
524 more than just functional diversity. *Nature communications* **10**: 859.

525 **Pavoine S, Bonsall MB. 2011.** Measuring biodiversity to explain community assembly: a unified  
526 approach. *Biological reviews of the Cambridge Philosophical Society* **86**: 792–812.

527 **Pebesma E. 2018.** Simple features for R: Standardized support for spatial vector data. *The R journal* **10**:  
528 439.

529 **Pérez-Escobar OA, Chomicki G, Condamine FL, Karremans AP, Bogarín D, Matzke NJ, Silvestro D,**  
530 **Antonelli A. 2017.** Recent origin and rapid speciation of Neotropical orchids in the world's richest plant  
531 biodiversity hotspot. *The New phytologist* **215**: 891–905.

532 **Pimm SL, Joppa LN. 2015.** How Many Plant Species are There, Where are They, and at What Rate are  
533 They Going Extinct? *Annals of the Missouri Botanical Garden* **100**: 170–176.

534 **Qian H, Jin Y. 2021.** Are phylogenies resolved at the genus level appropriate for studies on phylogenetic  
535 structure of species assemblages? *Plant diversity* **43**: 255–263.

536 **Qian H, Zhang J, Jiang M. 2023.** Global patterns of taxonomic and phylogenetic diversity of flowering  
537 plants: Biodiversity hotspots and coldspots. *Plant Diversity*.

538 **Rantanen M, Karpechko AY, Lipponen A, Nordling K, Hyvärinen O, Ruosteenoja K, Vihma T,**  
539 **Laaksonen A. 2022.** The Arctic has warmed nearly four times faster than the globe since 1979.  
540 *Communications Earth & Environment* **3**: 1–10.

541 **R Core Team. 2022.** R: A Language and Environment for Statistical Computing.

542 **Rodrigues ASL, Brooks TM, Gaston KJ. 2005.** Integrating phylogenetic diversity in the selection of  
543 priority areas for conservation: does it make a difference? In: *Phylogeny and Conservation*. Cambridge  
544 University Press, 101–119.

545 **Rosauer D, Laffan SW, Crisp MD, Donnellan SC, Cook LG. 2009.** Phylogenetic endemism: a new  
546 approach for identifying geographical concentrations of evolutionary history. *Molecular ecology* **18**: 4061–  
547 4072.

548 **Rosenzweig ML. 1995.** *Species Diversity in Space and Time*. Cambridge University Press.

549 **Rudbeck AV, Sun M, Tietje M, Gallagher RV, Govaerts R, Smith SA, Svenning J-C, Eiserhardt WL.**  
550 **2022.** The Darwinian shortfall in plants: phylogenetic knowledge is driven by range size. *Ecography* **2022**.

551 **Safi K, Cianciaruso MV, Loyola RD, Brito D, Armour-Marshall K, Diniz-Filho JAF. 2011.**  
552 Understanding global patterns of mammalian functional and phylogenetic diversity. *Philosophical*  
553 *transactions of the Royal Society of London. Series B, Biological sciences* **366**: 2536–2544.

554 **Sarkar S, Pressey RL, Faith DP, Margules CR, Fuller T, Stoms DM, Moffett A, Wilson KA, Williams**  
555 **KJ, Williams PH, et al. 2006.** Biodiversity Conservation Planning Tools: Present Status and Challenges  
556 for the Future. *Annual review of environment and resources* **31**: 123–159.

557 **Schnitzler J, Barraclough TG, Boatwright JS, Goldblatt P, Manning JC, Powell MP, Rebelo T,**  
558 **Savolainen V. 2011.** Causes of plant diversification in the Cape biodiversity hotspot of South Africa.  
559 *Systematic biology* **60**: 343–357.

560 **Sigwart JD, Sutton MD, Bennett KD. 2017.** How big is a genus? Towards a nomothetic systematics.  
561 *Zoological journal of the Linnean Society* **183**: 237–252.

562 **Smith SA, Brown JW. 2018.** Constructing a broadly inclusive seed plant phylogeny. *American journal of*  
563 *botany* **105**: 302–314.

564 **Srivastava DS, Cadotte MW, MacDonald AAM, Marushia RG, Mirotchnick N. 2012.** Phylogenetic  
565 diversity and the functioning of ecosystems. *Ecology letters* **15**: 637–648.

566 **Sun M, Eiserhardt W, Tietje M. 2021.** *taxonomy\_matcher: tax match 1.4*.

567 **Sun M, Folk RA, Gitzendanner MA, Soltis PS, Chen Z, Soltis DE, Guralnick RP. 2020.** Recent  
568 accelerated diversification in rosids occurred outside the tropics. *Nature communications* **11**: 3333.

569 **Tietje M, Antonelli A, Baker WJ, Govaerts R, Smith SA, Eiserhardt WL. 2022.** Global variation in  
570 diversification rate and species richness are unlinked in plants. *Proceedings of the National Academy of*  
571 *Sciences of the United States of America* **119**: e2120662119.

572 **Tucker CM, Cadotte MW, Carvalho SB, Davies TJ, Ferrier S, Fritz SA, Grenyer R, Helmus MR, Jin**  
573 **LS, Mooers AO, et al. 2017.** A guide to phylogenetic metrics for conservation, community ecology and  
574 macroecology. *Biological reviews of the Cambridge Philosophical Society* **92**: 698–715.

575 **Véron S, Saito V, Padilla-García N, Forest F, Bertheau Y. 2019.** The use of phylogenetic diversity in  
576 conservation biology and community ecology: A common base but different approaches. *The Quarterly*  
577 *review of biology* **94**: 123–148.

578 **Voskamp A, Baker DJ, Stephens PA, Valdes PJ, Willis SG. 2017.** Global patterns in the divergence  
579 between phylogenetic diversity and species richness in terrestrial birds. *Journal of biogeography* **44**: 709–  
580 721.

581 **Wickham H. 2016.** ggplot2: Elegant Graphics for Data Analysis.

582 **Wickham H. 2022.** stringr: Simple, Consistent Wrappers for Common String Operations.

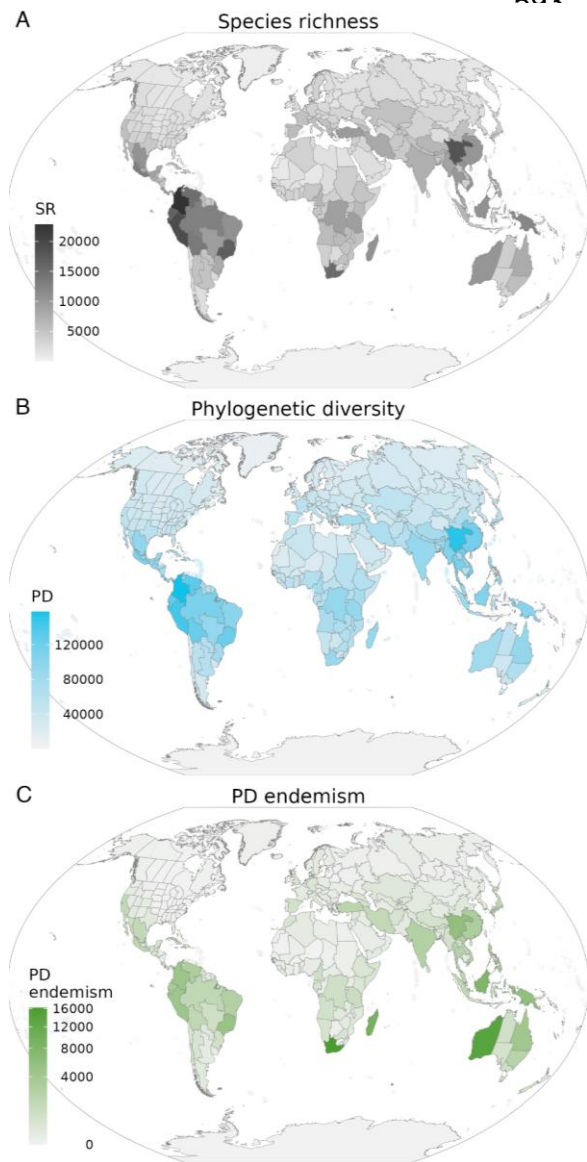
583 **Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Jonathan**  
584 **Davies T, Grytnes J-A, Harrison SP, et al. 2010.** Niche conservatism as an emerging principle in  
585 ecology and conservation biology. *Ecology letters* **13**: 1310–1324.

586 **Wiens JJ, Donoghue MJ. 2004.** Historical biogeography, ecology and species richness. *Trends in*  
587 *ecology & evolution* **19**: 639–644.

588 **Wilke CO. 2020.** cowplot: Streamlined Plot Theme and Plot Annotations for ‘ggplot2’.

589 **Winter M, Devictor V, Schweiger O. 2013.** Phylogenetic diversity and nature conservation: where are  
590 we? *Trends in ecology & evolution* **28**: 199–204.

591



624 Fig. 1. Three aspects of diversity: A) Species richness (SR); B) Phylogenetic diversity (PD) as per (Faith, 1992); C)  
625 Phylogenetic diversity endemism (PD endemism). Maps in Winkel tripel projection.  
626

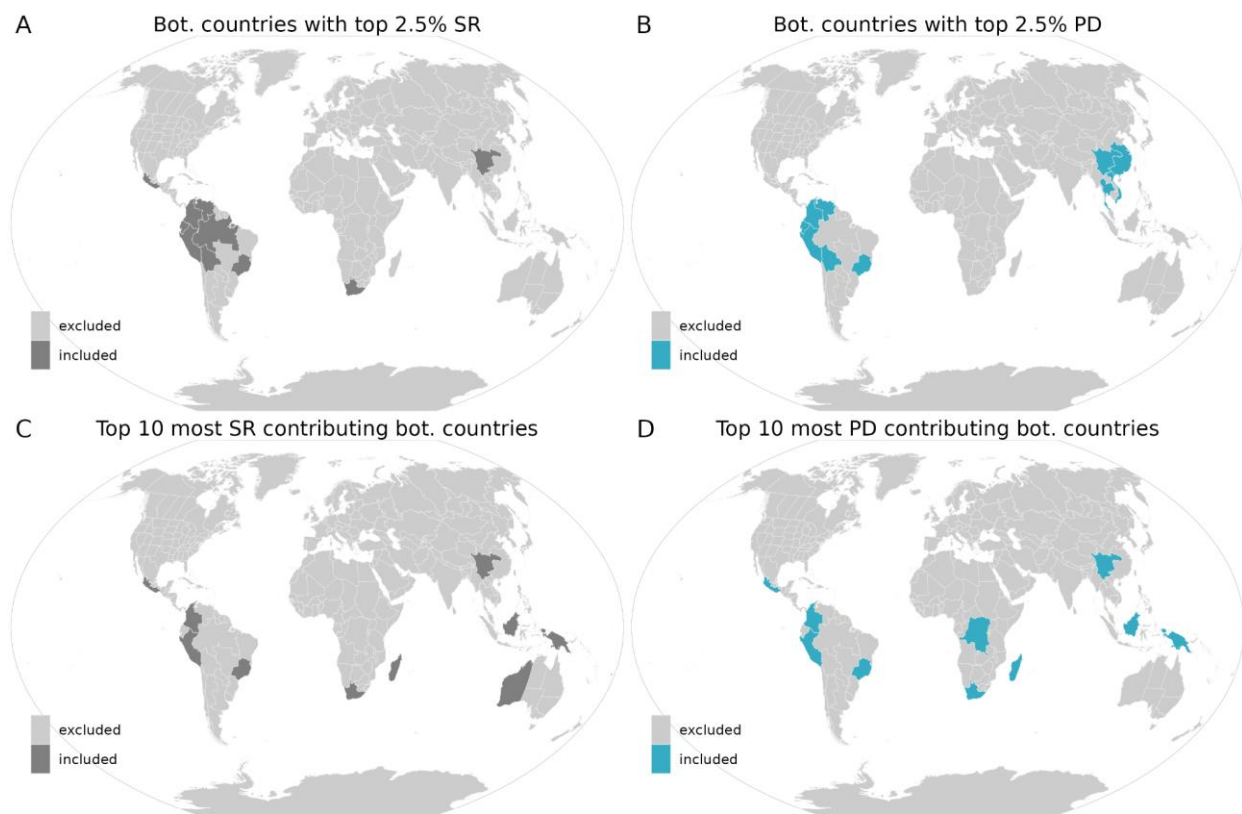


Fig. 2. Absolute values and complementarity for species richness (SR) and phylogenetic diversity (PD). Botanical countries with the top 2.5% total species richness (A) or phylogenetic diversity (B), and the top ten botanical countries with the highest contribution (complementarity) to global species richness (C) or phylogenetic diversity (D). Complementarity was assessed using a greedy algorithm that identifies the minimum number of countries containing the maximum number of species richness or phylogenetic diversity. The algorithm starts with the highest SR and PD value and subsequently adds countries with the next highest remaining contribution to SR and PD to the set.

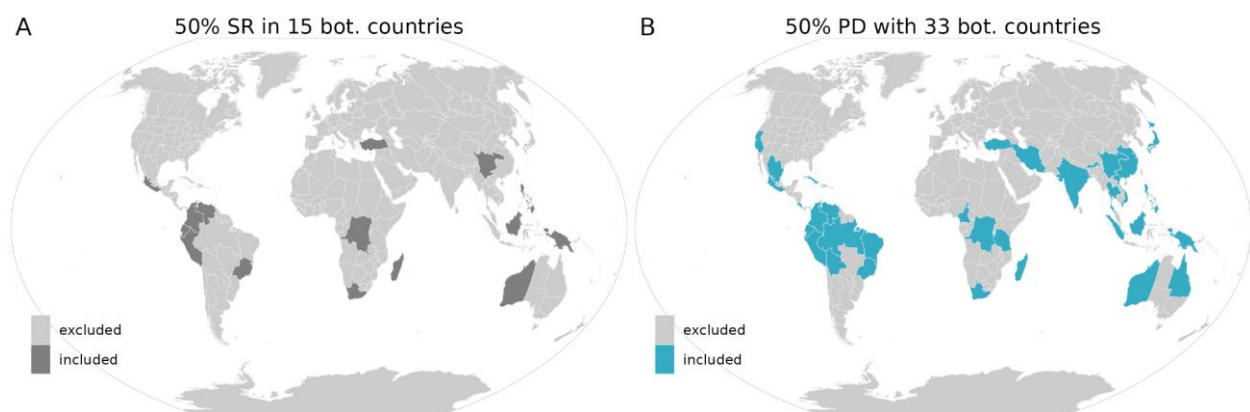


Fig. 3. The minimum number of botanical countries needed to capture 50% of global species richness (SR, A, 15 botanical countries) and phylogenetic diversity (PD, B, 33 botanical countries). Botanical countries were identified using a greedy algorithm that starts with the highest SR and PD values and subsequently adds countries with the next highest remaining contribution to SR and PD to the set.

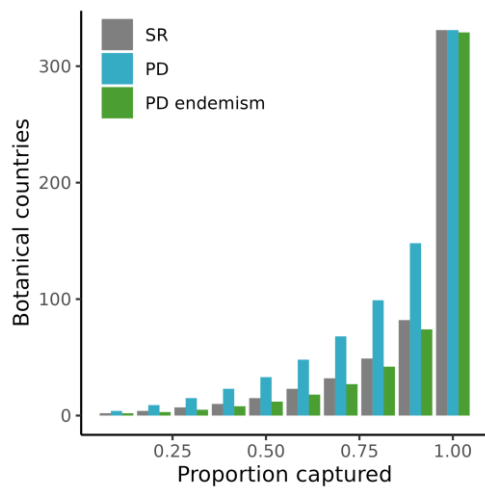


Fig 4. The number of botanical countries required to capture different percentages of species richness, phylogenetic diversity and endemic phylogenetic diversity. Half (50%) of plant species richness, phylogenetic diversity or PD endemism can be captured in either 15, 33 or 12 botanical countries, respectively.

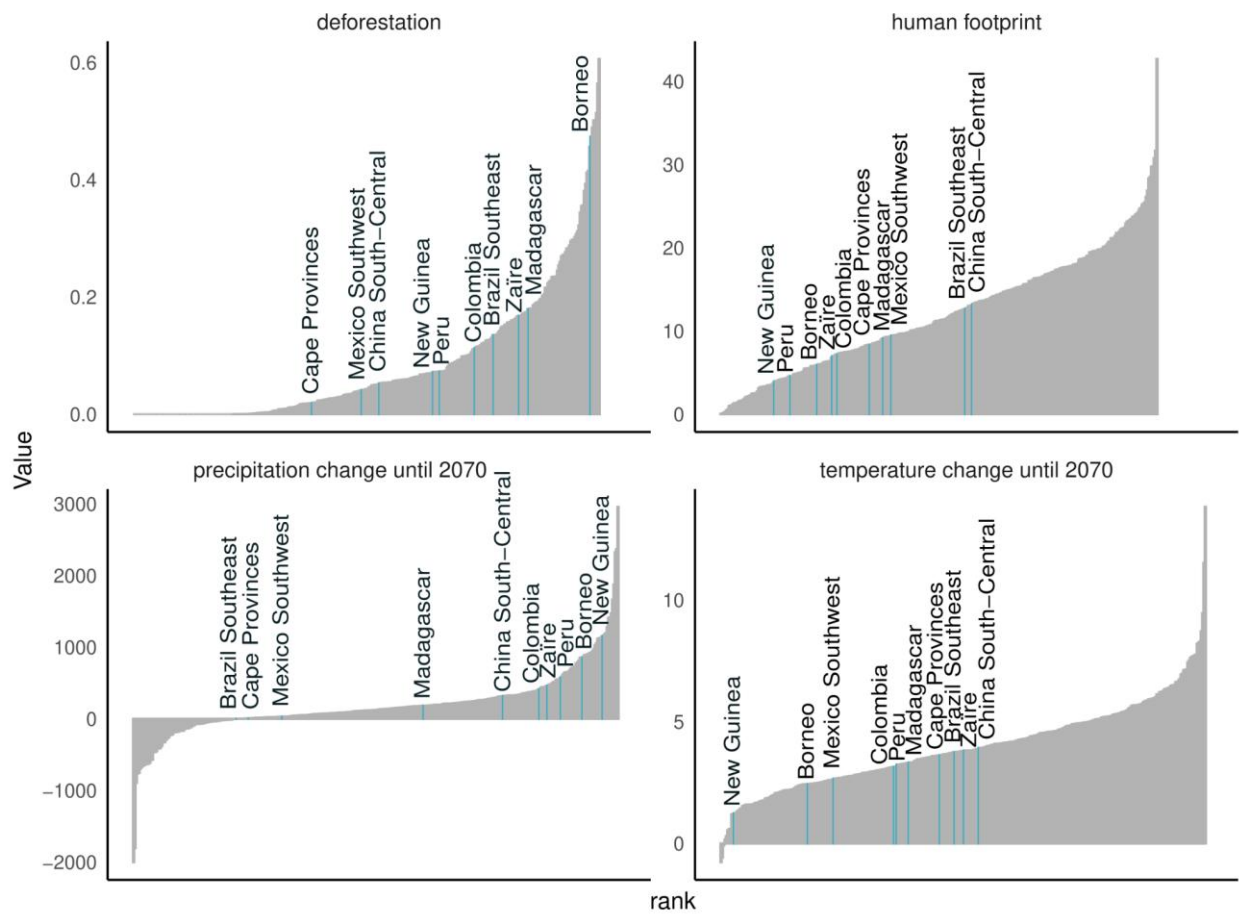


Fig. 5. Four types of threat for complementary PD hotspots. Rank of each hotspot (x-axis) with the corresponding threat value for deforestation, human footprint, predicted future change in precipitation and temperature, ordered by increasing threat values.



656  
657  
658  
659  
660  
661  
662

Table 1. Phylogenetic diversity complementarity hotspot characteristics. Hotspots were identified using a greedy algorithm that starts with the highest PD value and subsequently adds countries with the next highest remaining contribution to PD to the set. The first 10 selected countries are defined as hotspots. Top 2.5% columns indicate if the country is among the top 2.5% total PD or species richness (SR) countries (0=no, 1=yes). Complementarity top 10 columns indicate if the country is among the top 10 first selected countries using species richness (SR) or PD. Grey rows show countries with top 2.5% PD values that have not been picked as PD complementarity hotspot. The biome column depicts the biome composition of each hotspot country.

Level name	Top 2.5% PD	Top 2.5% SR	SR complementarity top 10	PD complementarity top 10	Species richness	PD	PD endemism	Biome composition
Borneo	0	0	1	1	10782	106890	8161	
Brazil Southeast	1	1	1	1	16960	121481	4913	
Cape Provinces	0	1	1	1	15362	88425	16197	
China South-Central	1	1	1	1	18238	153301	5772	
Colombia	1	1	1	1	22833	157886	4591	
Madagascar	0	0	1	1	10720	89042	9482	
Mexico Southwest	0	1	1	1	13133	111444	1754	
New Guinea	0	0	1	1	12042	103369	5874	
Peru	1	1	1	1	19235	143769	4414	
Zaire	0	0	0	1	9094	90968	763	
Bolivia	1	1	0	0	13568	123978	1642	
China Southeast	1	0	0	0	10214	117100	3061	
Ecuador	1	1	0	0	16599	128453	3295	
Thailand	1	0	0	0	10048	113039	2262	
Venezuela	1	1	0	0	14957	130885	3012	
Vietnam	1	0	0	0	10539	118236	2453	

(Sub)Tropical moist broadleaf forests

(Sub)Tropical dry broadleaf forests

(Sub)Tropical coniferous forests

Temperate broadleaf and mixed forests

Temperate coniferous forests

(Sub)Tropical grasslands, savannas, shrublands

Montane grasslands and shrublands

Mediterranean forests, woodlands, scrub

Deserts and xeric shrublands

Mangroves