Global patterns and determinants of multiple facets of plant diversity

Enrico Tordoni1 | Carlos Pérez Carmona1 | Aurèle Toussaint1,2 | Riin Tamme1 | Meelis Pärtel1

Abstract

Aim: Combining different biodiversity dimensions can reveal new diversity patterns disclosing the relative roles of historical, environmental and anthropogenic factors in shaping global seed plant diversity.

Location: Global.

Time period: Present.

Major taxa studied: Vascular plants.

Methods: We collated a database encompassing taxonomic (249,000 species), functional and phylogenetic information (34,694 species) of seed plants across different regions of the world. Species richness in each region was weighted accounting for their phylogenetic and functional distinctiveness, obtaining a new metric—μ-diversity—which was modelled to disentangle the relative roles of historical factors such as climate variability since the Last Glacial Maximum (LGM), environmental features (e.g. actual evapotranspiration—AET) and anthropogenic factors (past and current).

Results: Higher μ-diversity was observed in Papuasia, South East Asia, Australia and Central America, whereas the lowest values were primarily located in the Northern Hemisphere. Climate variability and AET were the most important determinants of μ-diversity and individual diversity facets, whereas the importance of past human impacts (i.e. the onset of pastoralism) equated or exceeded those of the present ones.

Main conclusions: Our integrative approach proved more sensitive in describing species diversity patterns. Few areas on Earth host high and unique proportions of multiple diversity facets and individual diversity facets contribute differently to μ-diversity across continents. Historical climate stability and water-energy dynamics strongly affect species diversity, but we also observed that past land-use legacy may have influenced current plant diversity, which is under intense anthropogenic pressure, especially in Asia as well as in Central and South America.

KEYWORDS

Anthropocene, biodiversity conservation, climate change, climate stability, functional diversity, land use change, latitudinal gradient, phylogenetic diversity, taxonomic diversity
Anthropogenic global change dramatically alters and reduces biodiversity across the globe (IPBES, 2019; Newbold et al., 2015; Trisos et al., 2020). To predict the future changes of biodiversity, it is essential to understand the factors that have shaped diversity patterns worldwide. Most of the knowledge produced so far about global plant diversity is based on taxonomic diversity (Cazzolla Gatti et al., 2022; Kier et al., 2005; Sabatini et al., 2022), while other diversity facets such as species evolutionary history (Phylogenetic Diversity—PD; Faith, 1992) and their functional structure (Functional Diversity—FD; i.e. trait variability among organisms, Petchey & Gaston, 2006) have been less explored at the global scale (but see Guo et al., 2022). While PD provides information on how past speciation and dispersal events may have shaped current species assemblages (Fritz & Rahbek, 2012) along with their “evolutionary potential” in response to environmental changes (sensu Winter et al., 2013), FD depicts ecosystem functions and services that variation in species richness and taxonomic composition might not completely disclose (Flinn et al., 2011).

Alternative methods were proposed for integrating different biodiversity dimensions, especially for biodiversity conservation and for spatial prioritization (Cadotte & Tucker, 2018). For instance, by weighting species diversity based on either phylogenetic or functional information (e.g. species distinctiveness; Violle et al., 2017) or by combining them in a composite metric (Cadotte et al., 2013; De Bello et al., 2017). The integration of different diversity facets and the incorporation of their spatial complementarity is needed for biodiversity conservation and can provide new insights into the processes shaping life on Earth such as evolutionary processes, ecosystem functioning, and community dynamics (Brum et al., 2017; Pollock et al., 2017; Tucker et al., 2019). Additionally, understanding the main determinants of different diversity facets is of utmost importance to anticipate the trajectories of global change.

Given the pivotal role of different diversity aspects on ecosystem functioning and stability (Cadotte et al., 2012; Flynn et al., 2011; van der Pias, 2019), and considering that each facet provides complementary information (Cadotte et al., 2019; Mazel et al., 2018; Owen et al., 2019; Tucker et al., 2018), integrating multiple diversity facets is an overarching issue for conservation (Pollock et al., 2020; Tordoni et al., 2021; Veron et al., 2017). Indeed, synthetic indices accounting for different aspects of diversity can even be combined across multiple taxa to provide an overall estimation of the total biodiversity in a given area, allowing conservationists not to be focused on specific aspects of diversity when developing conservation programs (Brum et al., 2017; Isaac et al., 2007). Here, we designed a new metric—μ-diversity—which weights species richness by its functional and phylogenetic counterparts. The latter two together form the ecological distinctiveness—ED—of an assemblage in a given area, whereby higher ED reflects a higher irreplaceability in terms of evolutionary history and functional strategies. Thus, a synergistic evaluation of biodiversity can highlight hidden patterns as well as disclose the role of some potential determinants, which could remain underestimated when focusing on a single biodiversity dimension.

Several hypotheses have been proposed to explain broad-scale patterns of species diversity, usually relying on species richness, but a general consensus remains elusive (see Fine, 2015 and references therein). These hypotheses (Table 1) consider different environmental determinants linked to productivity, environmental heterogeneity, as well as historical large-scale processes such as climate and geological stability (Qian & Ricklefs, 2000; Schemske & Mittelbach, 2017).

### 1.1 Environmental determinants

According to the water-energy dynamics hypothesis, species richness increases when intermediate levels of energy are coupled with high availability of water (Evans et al., 2005; Francis & Currie, 2003) which lessens competition allowing the co-existence of more closely related species (Thullier et al., 2020). Similar to species richness, higher values of functional diversity are expected in wet and warm regions endowed with high productivity (Lamanna et al., 2014; Swenson et al., 2012), even though these regions may also display a high degree of functional clustering (Carmona, Bueno, et al., 2021). These areas are also characterized by higher evolutionary potential (i.e. presence of older lineages) in agreement with ‘phylogenetic niche conservatism’ (Wiens & Donoghue, 2004; Kerkhoff et al., 2014; but see Brown, 2014), despite an overall lower speciation rate (Igea & Tanentzap, 2020). Environmental heterogeneity has also been shown to positively influence species richness (Stein et al., 2014), but the mechanisms acting behind PD and FD remain more elusive due to the contrasting effect that heterogeneity may have on these diversity facets (Meynard et al., 2011).

### 1.2 Historical determinants

Another compelling hypothesis is based on the stability of climate over evolutionary time-scales (climate stability hypothesis), which suggests a long-term persistence of higher diversity in regions characterized by stable climate (i.e. closer to the equator), thanks to reduced extinction rates providing also enough time for speciation events. Evidence for the climate stability hypothesis is based mostly on animals, even though recent research on angiosperms supports it (Feng et al., 2019), also suggesting the influence of climate stability on the ecological specialization of plants (i.e. more species tend to be ecological and physiological generalists at higher latitudes) (Lancaster & Humphreys, 2020). Finally, quaternary geological activities (e.g. volcanism, sea level changes triggered by glacial–interglacial climatic oscillations, tectonic movements) have been shown to positively influence species diversification (Couvreur et al., 2021; Kubota et al., 2015; Weigelt et al., 2016).
<table>
<thead>
<tr>
<th>Type</th>
<th>Hypothesis</th>
<th>Predictor</th>
<th>Mechanism</th>
<th>Expected relationships</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental</td>
<td>Water-energy dynamics</td>
<td>AET</td>
<td>Increased energy in the system coupled with high water availability should create more benign environments (i.e. warm and wet with high productivity, Evans et al., 2005; Lamanna et al., 2014). The latter allows the presence of different lineages but also the coexistence of closely related species promoting phylogenetic clustering. Regarding FD we expect that productivity increase the variety of functions.</td>
<td>++   +/-  +/-    +</td>
</tr>
<tr>
<td>Environmental</td>
<td>Environmental heterogeneity</td>
<td>Roughness</td>
<td>Roughness can be considered a proxy of environmental heterogeneity (Amatulli et al., 2018, Stein et al., 2014). More heterogeneous areas can sustain different habitat types and increase available niche space where more species and functions can coexist.</td>
<td>++   +/-  +/-    +</td>
</tr>
<tr>
<td>Historical</td>
<td>Climate instability</td>
<td>Median rate of change of surface temperature</td>
<td>Areas characterized by high historical climatic instability have higher extinction rates and lower diversity (Fine, 2015).</td>
<td>--  --  --  --  --</td>
</tr>
<tr>
<td>Historical</td>
<td>Geological dynamics</td>
<td>PC1 as a gradient of geological activity</td>
<td>Geological activities (e.g. volcanism, tectonic movements) positively influence species diversification rates and species radiation causing differences in species diversity and adaptations (Couvreur et al., 2021; Weigelt et al., 2016)</td>
<td>++   +    +    +</td>
</tr>
<tr>
<td>Anthropogenic</td>
<td>Current human impacts</td>
<td>Human footprint index</td>
<td>Areas under strong human disturbances are expected to exhibit lower species richness (Newbóld et al., 2015). The relationship with phylogenetic and functional diversity should be less steep because of redundancy and the arrival of new species due to biological invasions.</td>
<td>--  --  --  --  --</td>
</tr>
<tr>
<td>Anthropogenic</td>
<td>Past human impacts</td>
<td>Year of onset of pastoralism</td>
<td>Legacies of past human impacts should be present in current plant assemblages (Ellis et al., 2021; Levis et al., 2017); whereby lower diversity is expected where land-use change started earlier in time. The effect may be stronger for PD since humans may have selectively exploited specific clades.</td>
<td>--  --  --  --  --</td>
</tr>
</tbody>
</table>

Note: The plus and minus signs indicate positive and negative relationships, respectively, while the number of signs indicates the expected strength of the relationship. +/- refers to uncertain relationships.
1.3 | Anthropogenic determinants

Besides natural processes, human activities have also played a pivotal role in the last decades, with up to 60% of the terrestrial land being under moderate to intense human pressure (Williams et al., 2020; Winkler et al., 2021). Although the intensity of human activities has accelerated since the Industrial Revolution, humans have caused transformative changes on Earth since the early Holocene (Ellis et al., 2017; McMichael et al., 2023; Odone et al., 2019).

To date, few studies have investigated broad-scale patterns of vascular plant diversity considering different diversity aspects (species richness, PD, FD) together or individually in a comparative way, mostly due to the lack of evenly distributed data across the globe. As a consequence, the relative roles of environmental, historical (i.e., biogeography) and anthropogenic factors (past and present) in shaping broad-scale patterns of plant diversity (Cai et al., 2023; Feng et al., 2019; Gao & Liu, 2018; Kreft & Jetz, 2007) are still unclear across diversity facets, especially on the functional component of diversity. In addition, by investigating the global latitudinal patterns of multifaceted plant diversity, we can provide a more comprehensive picture of one of the most studied patterns in macroecology. Here, we exploit data with unprecedented coverage from the World Checklist of Vascular Plants (Govaerts et al., 2021), which includes more than 300,000 accepted vascular plant species, and the related geographic sources obtained from Plants of the World Online (https://powo.science.kew.org/). We developed a single metric (μ-diversity) that combines information on species richness with the functional and phylogenetic distinctiveness of species, and analyze it with the following specific aims: (i) to identify the regions of the globe hosting the highest plant μ-diversity and (ii) to disentangle the relative role of environmental, historical and anthropogenic drivers in explaining broad-scale patterns of μ-diversity as well as species richness, phylogenetic and functional diversity. We hypothesize a primary role of water-energy dynamics and historical factors (e.g., climatic stability) in driving global patterns of μ-diversity. Specifically, we considered six traits being the most widespread in public database and already used in several studies (e.g., Küster et al., 2016; Zhou et al., 2022): specific leaf area (sla, mm² mg⁻¹), leaf area (la, mm²), plant height (ph, m), seed mass (sm, mg), nitrogen content per unit mass (ln, mg g⁻¹) and specific stem density (ssd, g cm⁻³) (Table S1). These traits were used to build a multivariate space (i.e., functional space, see below for further details) which synthesizes the intrinsic dimensionality of plant trait variation. The first dimension of this space represents the size of individual species, whereas the second dimension reflects the use of resources (Leaf Economics Spectrum, Wright et al., 2004). This functional space has been shown to be robust to the addition of further traits (Joswig et al., 2021) as well as to the inclusion of root traits, which tend to be independent from aboveground traits (Carmona, Bueno et al., 2021).

2 | MATERIALS AND METHODS

2.1 | Plant data

Plant data were retrieved from the Plants of the World Online (POWO—Plants of the World Online, 2022), which uses The World Checklist of Vascular Plants (WCWP, Govaerts et al., 2021) as a backbone. The original file consisted of approximately 1,400,000 records; we then extracted information about seed plants filtering out ferns, hybrids and records not identified at the species level, thus leaving 956,753 records belonging to 249,175 taxa for subsequent analyses. Spatial information on WCVP was retrieved from Level 3 of the Taxonomic Databases Working Group (TDWG, Brummitt, 2001), which is based on "Botanical Countries" but does not always reflect the country's administrative units.

2.2 | Functional traits

Functional trait data for vascular plants were collected from TRY database (Kattge et al., 2020), using the data available in Carmona, Tamme, et al. (2021). The selected aboveground traits were based on the global spectrum of plant form and function (Díaz et al., 2016), and they reflect trade-offs in plant size, resource allocation and leaf construction costs. Specifically, we considered six traits being the most widespread in public database and already used in several studies (e.g., Küster et al., 2016; Zhou et al., 2022): specific leaf area (sla, mm² mg⁻¹), leaf area (la, mm²), plant height (ph, m), seed mass (sm, mg), nitrogen content per unit mass (ln, mg g⁻¹) and specific stem density (ssd, g cm⁻³) (Table S1). These traits were used to build a multivariate space (i.e., functional space, see below for further details) which synthesizes the intrinsic dimensionality of plant trait variation. The first dimension of this space represents the size of individual species, whereas the second dimension reflects the use of resources (Leaf Economics Spectrum, Wright et al., 2004). This functional space has been shown to be robust to the addition of further traits (Joswig et al., 2021) as well as to the inclusion of root traits, which tend to be independent from aboveground traits (Carmona, Bueno et al., 2021).
2.3 | Phylogeny and trait imputation

Plant phylogeny was obtained using ‘V.PhyloMaker’ R package (Jin & Qian, 2019), which exploits the mega-phylogeny for seed plants of Smith and Brown (2018) coupled with clades from Zanne et al. (2014). Species missing from the phylogenetic tree were added according to the “Scenario 3” following Qian and Jin (2016) and similarly to Cai et al. (2023). This scenario entails to add a new tip to the 1/2 point of the family branch with some exceptions (Jin & Qian, 2019; Qian & Jin, 2016). To obtain a complete trait matrix, we imputed missing traits only for species having at least one trait and phylogenetic information as described in recently published papers (average trait completeness=34.05%, range=23%-59%; see Table S1 and Carmona, Tamme, et al., 2021 for more details), which shows also the accuracy and robustness of trait imputation procedure (Carmona, Bueno, et al., 2021; Tordoni et al., 2021; Toussaint et al., 2021). In order to make a comparable effort between phylogenetic and functional dimensions, we subset from the phylogenetic tree only species with functional trait data, leaving 34,694 species for further analysis on functional and phylogenetic facets.

2.4 | Calculation of diversity metrics

The geography of vascular plant diversity was obtained by calculating diversity metrics in each botanical country. Species richness (SR) was calculated as the total number of species within each botanical country available from the full list of occurrences to maximize data coverage. Phylogenetic diversity (PD) was expressed using Faith’s PD (Faith, 1992), which reflects the evolutionary history within a given assemblage (botanical country in our case), and it was computed using the ‘picante’ R package (Kembel et al., 2010). To compute functional diversity (FD), we first log-transformed and scaled the functional trait matrix to unit variance, and then we computed a principal component analysis (PCA) in order to build the functional space. We applied Horn’s parallel analysis to assess the minimum number of axes to retain (first two principal components), using the R package ‘paran’ (Dinno, 2018). The resulting functional space has the fundamental characteristics of the global spectrum of plant form and function, as described in previous studies (Carmona et al., 2021b; Díaz et al., 2016). We then applied the TPD framework (Carmona et al., 2016), implemented in the ‘TPD’ R package (Carmona, 2019) to estimate FD within each botanical country. In particular, FD was computed as functional richness, which describes the amount of functional space occupied by a given assemblage (Carmona et al., 2016).

Given the known dependence between PD/FD and SR, we used null model simulations to make PD and FD independent from species richness. Specifically, we obtained 1000 null distributions randomizing the assemblage composition (presence/absence data) of each botanical country preserving marginal totals using ‘Curveball algorithm’ (Strona et al., 2014). We then calculated standardized effect size (SES) as

\[
\text{SES} = \frac{(\text{Metric}_{\text{obs}} - \text{mean} (\text{Metric}_{\text{null}}))}{\text{sd}(\text{Metric}_{\text{null}})}.
\]

which indicates if the observed value of the considered metric (i.e. sesFD or sesPD) is higher (or lower) than expected given the number of species present in the botanical country. Botanical countries hosting less than 10 species were removed from the dataset, leaving 364 botanical countries for further analyses (96% of the original dataset). Hereafter, when mentioning PD and FD, we always refer to sesPD and sesFD, which were used in the subsequent analyses. We then centred and scaled to unit variance the three diversity facets (i.e. log(SR), PD and FD). PD and FD were further re-scaled between 0 and 1 to obtain comparable ranges of variation and use them as weights for species richness (i.e. PD_{sc}, FD_{sc}). We did this re-scaling by comparing the PD and FD values with a cumulative normal distribution function with mean=0 and standard deviation=1. Re-scaling this way removes the effect that outliers (i.e. botanical countries with extremely high or low SES values) would have in other types of re-scaling like dividing by maximum SES values. We then calculated \( \mu \)-diversity for each botanical country as follows:

\[
\text{Ecological Distinctiveness (ED) = } \frac{\text{PD}_{sc} + \text{FD}_{sc}}{2},
\]

\[
\mu \text{-diversity = } S \times \text{ED},
\]

where ED reflects how diverse in functional and evolutionary terms the species in a given botanical country are, after accounting for species richness, and S is the number of species in a given botanical country. \( \mu \)-diversity reflects the taxonomic diversity weighted by its ecological distinctiveness. This mathematical formulation can be seen as in analogy with the concept of the effective number of species (ENS, MacArthur, 1965), whereby ENS is equal to species richness in the case of a perfectly even assemblage, otherwise is always smaller. Similarly, in the case of two assemblages with equal species richness, \( \mu \)-diversity will tend to richness value if ED = 1, otherwise it is always smaller because it is proportionally penalized as ED approaches 0. Our approach provides a new descriptive metric which is comprehensive since accounts for multiple diversity facets, standardized and scalable based on the resolution of the input data, which might be fine-tuned afterwards as soon as there is an increased data availability (e.g. more information on functional and phylogenetic aspects, or species spatial distribution at a finer scale). This metric is not intended to replace existing diversity metrics, but rather as a further step toward a more holistic comprehension of biodiversity.

2.5 | Environmental and anthropogenic data

The selected environmental data about climate and geological stability, water-energy dynamics (e.g. actual evapotranspiration) and anthropogenic factors (past and current) reflect the main
potential determinants influencing plant broad-scale patterns (Feng et al., 2019; Kreft & Jetz, 2007) (Table S2).

Climate variation since the Last Glacial Maximum (ca. 20ka) was estimated using the change of surface temperature as a proxy. In practice, we estimated the median rate of change of surface temperature during the time-series expressed in °C/century (Brown, Wigley, Otto-Bliesner, & Fordham, 2020) so that larger values indicate less climate stability and vice versa.

To obtain geological dynamics, we first extracted for each botanical country the number of active faults from the Global Earthquake Model Global Active Fault Database (Sytron & Pagani, 2020) and the number of active volcanoes since the Pleistocene from Global Volcanism Program (Global Volcanism Program, 2016). We then used PCA of these two variables and selected the score of each botanical country in the first PC (explaining 70% of total variation), which reflects a gradient of geological activity (Figure S1).

Land surface ruggedness (hereafter roughness, resolution 1 km²), considered as a proxy of spatial heterogeneity, was obtained from Amatulli et al. (2018).

We also retrieved actual evapotranspiration (AET), a good proxy of water-energy dynamics and plant productivity (Kreft & Jetz, 2007), as well as soil humidity, which reflects water availability. Gridded databases of AET and soil humidity were obtained from TerraClimate (Abatzoglou et al., 2018).

Global current human influence was retrieved from the 2009 Human Footprint index (HF) (Venter et al., 2016), a synthetic metric of the intensity of human pressure based on a weighted average of eight metrics (i.e. roads, railways, navigable waterways, built environments, crop land, human population density, pasture land, nighttime lights).

Past human influence was obtained from ArchaeoGLOBE Project (Stephens et al., 2019), which reports the year of onset of different land use types (e.g. agriculture, pastoralism, urbanism) in different regions of the globe.

To avoid unnecessary complexity as well as to prevent potential overfitting in the modelling procedure, we included only the variables showing a Pearson's correlation coefficient < 0.70 (Dormann et al., 2013) (Figure S2). To control for the well-known species-area relationships (Lomolino, 2000), we added the area (log10-transformed) of each botanical country, retrieved from Dawson et al. (2017), as a covariate in the model. The following variables were finally selected: AET, roughness, geological dynamics, Human Footprint index, year of onset of pastoralism and climate variation since LGM. Since this information were not available for all botanical countries, we retained only those botanical countries with complete data (247 botanical countries, 68% of the original dataset) for the subsequent modelling procedure. For each botanical country, data were extracted using the ‘sf’ (Pebesma, 2018) and ‘exactextract’ R packages (Daniel Baston, 2022). For gridded data, the values for each botanical country were obtained by calculating the average value of pixels that intersects a given polygon weighted by the cover fraction within the polygon. Maps’ colour schemes are based on the ‘viridis’ (Garnier et al., 2021) and ‘scico’ (Pedersen & Crameri, 2020) R packages.

2.6 | Statistical analyses

Spatial associations among original diversity facets (i.e. log(SR), sesPD, sesFD), and between these and μ-diversity were computed using a Pearson's correlation (r) corrected by spatial autocorrelation using a modified t-test (Clifford et al., 1989), as implemented in the ‘SpatialPack’ R package (Vallejos et al., 2020). The spatial overlap among the three dimensions of diversity was calculated as follows: for each diversity facet, we first ranked the areas based on diversity values and then extracted the top 10% and 50% of areas of global importance. For each group (i.e. 10% and 50%), we calculated the overlap as the intersection among the resulting vectors, namely by intersecting the vectors of the botanical countries for each diversity facet and calculating the size (i.e. the number of botanical countries) of individual partitions, which were finally relativized to the total to obtain percentage values. Spatial overlaps were calculated and visualized using Venn diagrams available through ‘eulerr’ and ‘VennDiagram’ R packages (Chen, 2022; Larsson, 2021). We also decomposed the variation of μ-diversity and extracted the unique and shared contribution of each diversity facet at the global scale and across continents following the level 1 of TDWG categorization through a variation partitioning approach using vegan R package (Borcard et al., 1992; Oksanen et al., 2022).

The relationship between plant μ-diversity and environmental, historical and anthropogenic data was assessed using random forests exploiting the framework provided in the ‘ml3’ (Lang et al., 2019) and ‘mlr3spatiotempcv’ (Schratz & Becker, 2021) R packages. The same modelling framework was also applied independently to each diversity facet. Model’s hyperparameters (i.e. number of variables randomly sampled at each split (mtry), sample fraction, number of trees, minimum size of terminal nodes) were optimized using five-folds spatial cross-validation, and 50 evaluations of the model’s setting using ‘paradox’ R package (Lang et al., 2021). For each evaluation, a random search within a user defined parameter space was performed, which root mean square error (RMSE) was later stored for comparison. The set of hyperparameters showing the lowest RMSE was finally selected to compute the model and related variable importance.

Variable importance was determined by measuring the mean change of RMSE after variable permutations (N = 500) using ‘DALEX’ R package (Biecek, 2018). This methodology assumes that if a given variable is important for the model, we expect to observe a decrease in the model’s performance after permuting its values (see Fisher et al., 2019). Marginal effects were visualized using partial dependence plots available in the ‘iml’ R package (Molnar et al., 2018). Accounting for spatial autocorrelation is a crucial step during the modelling exercise (Meyer & Pebesma, 2022; Ploton et al., 2020), therefore we evaluated modelling performance using an internal spatial cross-validation (Becker et al., 2021) by subsetting five spatially disjoint partitions that maximize the distance between training and validation set with respect to random partitioning (Love lace et al., 2019). Nested resampling was then used following Becker et al. (2021), which is a two-step resampling whereby inner resampling tunes model hyperparameters whereas the outer one evaluates the model’s performance. Due to
the computational effort, we used 50 evaluations of model settings coupled with 5-folds in inner resampling, and 5-folds with 25 repetitions each for outer resampling to reduce variance among replicates. We ran 4 independent models in total, one for μ-diversity and one for each diversity facet, after having scaled log(SR) as z-score. Model performances were evaluated using normalized root mean square error—NRMSE—whereby RMSE was normalized on the range of the response variable. Finally, we intersected μ-diversity with human footprint using bivariate maps, after classifying each variable in quartiles (i.e. 0–25%, 25–50%, 50–75%, 75–100%). The difference between HF for each quartile of μ-diversity was evaluated using a generalized least squares model (GLS) directly including the spatial covariance structure of the data to account for spatial autocorrelation; GLS was computed using ‘nlme’ R package (Pinheiro et al., 2022). Posthoc multiple comparisons were performed using Tukey’s HSD and Holm correction using ‘multcomp’ package (Hothorn et al., 2008). To further disentangle the effect between current and past human impacts, we first calculated ‘multcomp’ package (Hothorn et al., 2008). To further disentangle the effect between current and past human impacts, we first calculated the residuals of the relation between μ-diversity and HF, and then we related these residuals to the year of onset of pastoralism using the same modelling procedure described above. All statistical analyses were performed in R 4.2.1 (R Core Team, 2022).

2.7 | Sensitivity analysis

To evaluate the robustness of the random forest to sampling bias, we performed a sensitivity analysis across different quantiles of taxonomic completeness (i.e. the difference between the number of species for which we had functional and phylogenetic information compared to the total amount of species reported for that specific botanical country) from 0.01 (lower completeness, 3 botanical countries) to 0.25. For each quantile, we simulated the value of μ-diversity (n = 500) by sampling it from a uniform distribution without replacement based on the actual range of global μ-diversity but restricted only to values greater than the observed μ-diversity in that specific botanical country. This was done to increase the ecological realism of the sensitivity analysis, under the assumption that if we increased the amount of functional and phylogenetic information in a given botanical country, μ-diversity cannot decrease and remains stable only when redundant species are added to the present assemblage. After having randomized μ-diversity, we then merged this subset with the rest of the original observations and we ran the random forest. Finally, we computed the correlation coefficient (Pearson’s r) between the original predictions of the random forest model and the one with the randomized μ-diversity.

3 | RESULTS

3.1 | The geography of seed plant diversity

The spatial patterns of seed plant diversity strongly varied between the three diversity facets (Figures 1a–c and 2a–c). Species richness (Figures 1a and 2a) was higher (top 10%) in Central and South America (e.g. tropical Andes, Amazonia, Mexico and Costa Rica), Southern Africa (Cape region, Tanzania), South East Asia (e.g. Borneo, Southern China), Papua New Guinea and Australia. Likewise, PD (Figures 1b and 2b) was higher especially in the Caribbean and South-East Asia. The regions showing the highest FD are mostly in Papua such as Papua New Guinea, Northern Australia along with the Mekong region (i.e. Vietnam, Laos and Cambodia), and the Philippines (Figures 1c and 2c). It is worth noting that inventory completeness for PD and FD is higher in the northern hemisphere, while there is a higher uncertainty especially in some botanical countries of the Global South (Figure S3).

Combining the three diversity facets in a single synthetic index (μ-diversity, Figures 1d and 2d) revealed a clear distinction between hemispheres, with the botanical countries in the southern hemisphere and near the equator being generally more diverse than the ones in the northern hemisphere. In more detail, SE Asia and Papua (e.g. Philippines, Papua New Guinea, Vietnam and Cambodia; cf. Sundaland and Indo-Burma biodiversity hotspots in Myers et al., 2000), India, Eastern Australia, Mesoamerica and the Caribbean including Florida were the botanical countries with the highest seed plant μ-diversity. Furthermore, this index seemed more sensitive than species richness alone in describing the latitudinal gradient, showing higher sensitivity especially at higher latitudes (cf. Figure 2a.d). If we consider the spatial overlap of the top 10% areas of global importance for μ-diversity (Figure 3a), 37 botanical countries were included, but among them, only Papua New Guinea can be considered a hotspot for all the three individual dimensions of biodiversity representing 0.01% of all botanical countries. When considering 50% of the most diverse botanical countries, this percentage increased up to 19% (56 countries), further suggesting that different regions of the globe encompass diverse aspects of biodiversity. μ-diversity also proved to be differentially influenced by the three diversity facets showing a high contribution from SR in South America, tropical Asia, Australasia and Pacific, while temperate Asia and Europe were mostly determined by FD and PD, respectively (Figure 3b).

Considering the relationships between the different diversity facets, μ-diversity was correlated with FD (r = 0.79, p < 0.001) and PD (r = 0.75, p < 0.001) while the correlation was not significant with SR (r = 0.15, p = 0.16). Concerning raw diversity metrics, we detected a relatively high congruence between PD and FD (r = 0.69, p < 0.001), while SR seems to be more variable (Figure S2 and Table S3).

3.2 | Climate stability and water-energy dynamics determine global μ-diversity patterns

The selected variables were good predictors of broad-scale patterns of seed plant μ-diversity ($R^2 = 0.64 \pm 0.09$; normalized root mean square error—NRMSE = 0.24 ± 0.05; median ± SD. Table S4), and the model was relatively robust to taxonomic bias (Figure S4). Climate variation and AET were the most important
variables suggesting a prominent role of climate stability since LGM and water-energy dynamics followed by anthropogenic factors (Figure 4). Specifically, we detected a negative relationship with climate variation, meaning that more stable areas since LGM hosted higher $\mu$-diversity. We also observed a clear positive trend with AET suggesting that more productive areas of the planet host higher multifaceted diversity. Interestingly, we also detected an effect of the year of onset of pastoralism which had an effect stronger than recent human impacts, suggesting that $\mu$-diversity is lower in botanical countries where pastoralism started earlier in time. When looking at single diversity facets (i.e. SR, PD and FD; Figure 5), AET and climate variation overall are confirmed as the most important variables even though some differences can be observed among diversity facets. For instance, SR increased in relation to geological activities since the Pleistocene, whereas PD was mostly affected by climate variation and AET with a marked negative and positive trend, respectively. In contrast, FD showed a high importance of pastoralism onset (but a weak effect), which is comparable to the one of climate variation and much more important than current human impacts.

When comparing the geography of current anthropogenic disturbances with the spatial distribution of $\mu$-diversity, we observed that botanical countries mostly located in the Asian continent (e.g. India, the Philippines, Japan and the Korean peninsula) had a clear overlap suggesting that these regions hosting the highest $\mu$-diversity are disproportionally at risk, being under extreme anthropogenic pressure (Figure 6 and Table S5). Likewise, even though the human footprint is more nuanced in Tropical Andes and Atlantic Forest in South America as well as in South-Eastern Australia and part of New Zealand, these areas still host high levels of $\mu$-diversity, thus potentially resulting in immediate imperilment if anthropogenic disturbances further increase in these areas.

4 | DISCUSSION

Integrating multiple biodiversity dimensions into a new index ($\mu$-diversity) allowed us to highlight regions hosting high levels of multifaceted seed plant diversity. In particular, the regions hosting the highest plant $\mu$-diversity are mainly located in Papuasia, SE Asia,
the Caribbean, the Tropical Andes and some areas in Australia. Concerning the main determinants of $\mu$-diversity, we observed a primary role of historical factors (climate stability) and water-energy dynamics (AET), but we also detected a signal of past human impacts (i.e. pastoralism onset), which for FD was comparable to the effect of climate stability. We showed that a synergistic evaluation of biodiversity may disclose some hidden patterns which may be not seen when focusing only on species richness, but that can be truly...
disentangled only when accounting for multiple dimensions of biodiversity (cf. pastoral onset in Figures 4 and 5).

This can have important consequences for biodiversity conservation, as also recently outlined by other studies showing how biodiversity conservation would benefit from considering multiple biodiversity dimensions (Pollock et al., 2017; Tordoni et al., 2021), given also their small spatial overlap. Specifically, when considering the most diverse botanical countries (top 50%), we clearly detected that different areas on Earth can be considered of primary importance for different dimensions of biodiversity with relevant implications when planning biodiversity conservation actions and restoration efforts (Guo et al., 2022). Two crucial aspects of biodiversity, functional and phylogenetic diversity, were often decoupled from species richness. While PD can provide information about the evolutionary resilience of an assemblage in relation to extinction risk, given that extinction risk is non-randomly distributed across species (Purvis et al., 2000), FD is a good proxy of the range of ecophysiological strategies of the species in an assemblage.

In this sense, μ-diversity may help to disentangle the broad-scale patterns of diversity by accounting for the different facets of diversity in a balanced way. In turn, this can be exploited by policy makers grappling with conservation issues in the Anthropocene as a tool which is more actionable, although other aspects such as societal trends and nature contributions to people might be taken into account to have a more holistic perspective (Soto-Navarro et al., 2021). Even though we detected that the single biodiversity dimensions affected μ-diversity differentially across continents, we recognize that it can be difficult to effectively disentangle the contribution of each diversity facet in the current formulation of the metric. We also know that, especially for PD and FD, some regions still show a relatively low completeness (Figure S3) potentially underestimating how these diversity facets contribute to μ-diversity. Moreover, in recent times, there has been a proliferation of different methods to estimate functional and phylogenetic diversities (Mammola et al., 2021; Tucker et al., 2017), whereby it is important to make a careful selection of the metrics considering their purpose, and how these may ultimately affect the resulting diversity patterns, taking also into account their scale-dependency which can influence their degree of congruence. In this context, μ-diversity may serve as a first benchmark to evaluate how ecological distinctiveness affects species richness patterns. It also permits identifying the areas

![Figure 4](image-url) Variable importance ranked by the root mean square error (RMSE) loss after permutations (left panel) and their marginal effects (mean ± 2 SE) of the different predictors (right panel) of the random forest model using seed plant μ-diversity as response variable ($R^2 = 0.64±0.09$; NRMSE $= 0.24±0.05$, median ± SD). Boxplots represent the values of RMSE loss in each permutation (N = 500). Climate variation is the median rate of change of climate since the Last Glacial Maximum (expressed in °C/century). Pastoralism onset expresses the year of onset of these land uses sensu Stephens et al. (2019). Geological dynamics express a gradient of geological activity since the Pleistocene while roughness is an index of spatial heterogeneity. AET, actual evapotranspiration; HF, human footprint index. The lower inset represents the cumulative variable importance of anthropogenic (grey), environmental (green) and historical factors (orange). The log$_{10}$(Area) was used as a covariate to control the effect of the area of the botanical country.
where functional and phylogenetic information strongly influence the multifaceted diversity while potentially being underrepresented in the current conservation schemes, providing a generalized pattern that can be promptly delivered to policy makers and for science communication purposes. Nevertheless, in the case of low spatial congruence between functional and phylogenetic information, it is up to the conservation planners and managers to identify the single contribution of PD and FD to ecological distinctiveness and carefully plan how these facets might be prioritized.

4.1 | The biogeography of plant diversity

We identified higher species richness in Tropical Andes and SE Asia in agreement with previous works (Brummitt et al., 2021; Cai et al., 2023; Cazzolla Gatti et al., 2022; Kreft & Jetz, 2007). We detected also higher levels of phylogenetic and functional diversity in SE Asia, which is consistent with other studies which found extremely high levels of angiosperm’s PD in eastern Asia probably due to the combination of environmental and biogeographical factors along with the unique geological history of the area (Qian et al., 2017). Eastern Asia is a crossroad between temperate and tropical floras which could have contributed to increase diversity since the Tertiary (Qian & Ricklefs, 2000; Sodhi et al., 2004), making this area a cradle of vascular plant diversity (‘out of Asia’ hypothesis; Donoghue, 2008), hosting unique trait combinations and evolutionary history.

In their landmark study, Myers et al. (2000) identified biodiversity hotspots showing regions with higher species richness and irreplaceability under significant threats. Among the listed hotspots and also according to key biodiversity areas (https://www.keybiodive
Our study expands these findings for other dimensions of biodiversity, showing that Papuasia, Philippines, Sundaland and Indo-Burma regions along with the forests of northeastern Australia are not only rich in species but also characterized by unique evolutionary history and functional features. Notably, we detected a congruence between seed plant PD and FD (Table S3), which is not mirrored in tetrapod species (Tordoni et al., 2021). This might suggest that these centres of diversification promoted the evolution of unique features in plants, thanks also to the disproportionate accumulation of extremely rare species (Enquist et al., 2019), especially in tropical mountains (Testolin et al., 2021). It is worth noting the recent compelling evidence about the lower speciation rates in the tropics compared to temperate regions (Igea & Tanentzap, 2020; Tietje et al., 2022), suggesting that other mechanisms such as tropical conservatism (i.e. lower dispersal from the tropics to temperate zones; Wiens & Donoghue, 2004) or the higher ecological stability of these areas over evolutionary time-scales might have favoured the maintenance of older lineages which, in turn, may have promoted higher diversity (Trew & Maclean, 2021). Furthermore, we found an overlap between the areas of global importance for plants with the ones of tetrapod species, especially mammals and reptiles (Tordoni et al., 2021), further highlighting the importance of these regions in the conservation of global biodiversity. When considering the spatial overlap of global land area for the regions hosting the highest μ-diversity (top 10%), only Papua New Guinea stands out, suggesting that more efforts are needed to successfully capture different facets of biodiversity for prioritization analyses (Brum et al., 2017; Guo et al., 2022).

4.2 The role of climate stability, water-energy dynamics and anthropogenic impacts in shaping plant diversity

Our findings provide stronger support for climate stability and water-energy dynamics influencing both μ-diversity and the single diversity facets. The central role of climate stability for diversity has already been reported in several studies focusing on plants and other taxa (Feng et al., 2019; Gao & Liu, 2018; Morueta-Holme et al., 2013; Su et al., 2022). Indeed, lower extinction rates have usually been associated with more stable areas (Brown, Wigley, Otto-Bliesner, Rahbek, & Fordham, 2020), allowing a longer persistence of species and, consequently, increased time for species accumulation and diversification (Sandel et al., 2017). Water-energy dynamics are usually claimed as one of the main drivers of species richness broad-scale patterns (e.g. Field et al., 2009; Kreft & Jetz, 2007), since these are linked to productivity and can buffer against local extinctions (Evans et al., 2005). Regions with higher energy availability usually have assemblages characterized by a higher total number of individuals, which has been shown to be involved in diversity regulation (i.e. more-individuals hypothesis; Storch et al., 2018). The higher energy availability is also often associated with higher evolutionary speed and genetic diversity (Gillman & Wright, 2014), since it can lead to faster metabolism which in turn enhance mutation rates, ultimately fostering speciation when water availability is not a limiting factor. Within historical factors, geological dynamics also had a relatively important effect on species richness, while its effects on PD and FD seem less prominent. For instance, volcanism activities and sea-level
changes may have affected local extinctions, dispersal, and speciation events, especially on archipelagos close to mainland (e.g., Philippines or Sundaland) (Piombino, 2016; Sodhi et al., 2004; Weigelt et al., 2016), as well as strongly influencing local environmental conditions and environmental heterogeneity (Antonelli et al., 2018; Rahbek et al., 2019).

The output of our correlative models also suggests that the earlier onset of pastoralism might have influenced plant diversity globally, especially for FD and PD (see Figure S5), showing that botanical countries where pastoralism started earlier had lower diversity, and this pattern seems to be weakly linked to current human impacts (Figure S5). This is in agreement with recent findings highlighting that legacies of past anthropic disturbances might have strongly influenced current diversity patterns of both animals and plants (McMichael, 2021; McMichael et al., 2023; Polaina et al., 2019; Scerri et al., 2022). Recent mounting evidence has highlighted that extensive land use changes occurred much earlier in time than previously assessed (Mottl et al., 2021; Stephens et al., 2019), suggesting also that humans have been directly exploiting tropical forests since mid-Holocene (ca. 4 ka) (Scerri et al., 2022 and references therein). Specifically, the Neolithic revolution (i.e. the transition from a society of hunter-gatherers to the one of farmers) enhanced land use change through the spread of agriculture and pastoralism which could have influenced the taxonomic (and functional) composition of natural assemblages via forest clearances, plant domestication and widespread use of fire (Bodin et al., 2020; Bush et al., 2022; Nascimento et al., 2022; but see Karp et al., 2021), affecting local biodiversity (Hixon et al., 2021). In the last century, the increased extent and magnitude of land use change have even exacerbated these effects on global biodiversity posing several species with unprecedented imperils (Newbold et al., 2015; Powers & Jetz, 2019), given that the regions hosting the highest diversity are also the ones more exposed to high human pressures. Indeed, current anthropogenic disturbances are causing detrimental effects especially for small-ranged species (Gomes et al., 2019; Wright et al., 2009), with no signs of improvement in the near future (Williams et al., 2020).

Despite the wide coverage of the plant data considered in this study, we have to acknowledge the relatively coarse resolution of the data and the presence of artificial borders (i.e. administrative units), which might have influenced the observed patterns. Nonetheless, checklists still represent the most authoritative, trustworthy source of near-global characterization of plant distributions (Govaerts et al., 2021; Weigelt et al., 2020), and analyses using the same spatial resolution (i.e. botanical countries) already proved effective in identifying broad-scale patterns of plant diversity (Brummitt et al., 2021). Furthermore, the fact that global areas of importance for plant \( \mu \)-diversity tend to mirror the ones already described for species richness using georeferenced local plant assemblages (Sabatini et al., 2022), makes us confident enough that we were able to capture the main spatial patterns of seed plant variation. Future studies should examine if these findings remain stable across scales (Keil & Chase, 2019), since biodiversity patterns and processes are strongly scale-dependent. Further research is also needed to increase our knowledge about the spatial distribution of the species as well as the phylogenetic and functional features of natural assemblages, paying extra attention to data harmonization (Wüest et al., 2020). Especially from a functional point of view, better data coverage would have provided us with a more comprehensive picture of the functional structure of these assemblages, which would also help to anticipate the trajectories of change under the current global biodiversity crisis.

5 | CONCLUSIONS

This is one of the first attempts to disentangle patterns and drivers over multiple dimensions of seed plant diversity at the global scale (Cai et al., 2023; Guo et al., 2022). Our integrative approach resulted more sensitive in describing species diversity patterns and capturing latitudinal gradient than species richness alone, highlighting that Papua, SE Asia followed by Central and South America and parts of Australia are the most diverse regions globally when combining taxonomic, phylogenetic and functional information. We also detected a relatively small spatial overlap among diversity facets, suggesting that a few areas on Earth actually host high levels of all biodiversity dimensions. Regarding the main drivers, climate stability and water-energy dynamics were the most important factors which explained broad-scale patterns of plant multifaceted diversity. We also detected that some regions were under strong anthropogenic pressure, especially in Asia and Central and South America, and that the legacies of past human impacts (i.e. the onset of pastoralism) are at least as important as the current ones. Together with recent mounting evidence, this further corroborates the hypothesis that old civilizations might have shaped plant diversity more strongly than previously estimated.

AUTHOR CONTRIBUTIONS

Enrico Tordoni, Carlos Pérez Carmona, and Meelis Pärtel conceived the idea. Enrico Tordoni prepared the data and performed the statistical analysis. Carlos Pérez Carmona, Meelis Pärtel, Aurèle Toussaint and Riin Tamme contributed to explain and critically discuss the results. Enrico Tordoni led the writing of the first draft of the manuscript, which was edited by all co-authors.

ACKNOWLEDGMENTS

E.T., C.P.C., A.T., R.T. and M.P. were supported by the Estonian Research Council grants (MOBJD1030, PSG293, MOBERC40, PSG505, PRG609 and PRG2142) and by the European Regional Development Fund (Centre of Excellence EcolChange).

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The WCVP dataset is owned by “©Copyright Board of Trustees of the Royal Botanic Gardens, Kew” (https://www.kew.org/), which...


**BIOSKETCH**

Enrico Tordoni is a Research Fellow at University of Tartu. His current research focuses on disentangling global diversity spatio-temporal patterns of plants and vertebrates also accounting for functional and evolutionary aspects under the current scenario of accelerated global change.

**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.