nature ecology & evolution

Article

https://doi.org/10.1038/s41559-024-02589-0

Global decoupling of functional and phylogenetic diversity in plant communities

Received: 9 January 2024

Accepted: 24 October 2024

Published online: 03 December 2024

Check for updates

A list of authors and their affiliations appears at the end of the paper

Plant communities are composed of species that differ both in functional traits and evolutionary histories. As species' functional traits partly result from their individual evolutionary history, we expect the functional diversity of communities to increase with increasing phylogenetic diversity. This expectation has only been tested at local scales and generally for specific growth forms or specific habitat types, for example, grasslands. Here we compare standardized effect sizes for functional and phylogenetic diversity among 1,781,836 vegetation plots using the global sPlot database. In contrast to expectations, we find functional diversity and phylogenetic diversity to be only weakly and negatively correlated, implying a decoupling between these two facets of diversity. While phylogenetic diversity is higher in forests and reflects recent climatic conditions (1981 to 2010), functional diversity tends to reflect recent and past climatic conditions (21,000 years ago). The independent nature of functional and phylogenetic diversity makes it crucial to consider both aspects of diversity when analysing ecosystem functioning and prioritizing conservation efforts.

Climate change and biodiversity loss are pressing environmental issues, with rising temperatures and shifting precipitation patterns increasingly driving plant species extinctions[1]. These changes have substantial implications for ecosystems and human societies alike, with impacts ranging from altered agricultural yields to increased risk of natural disasters[2–4]. To understand and mitigate the effects of climate change and biodiversity loss, it is crucial to determine how plant species assemble into communities and how these communities respond to changing environmental and climatic conditions[5,6]. To do this, we need to understand the underlying mechanisms of plant community assembly and how environmental conditions, species' functional traits and evolutionary histories interact to mediate these mechanisms[7].

Community assembly reflects several processes that can reinforce or oppose each other[8]. On the one hand, environmental filters tend to favour similar phenotypic traits generating clustering within a community[9,10]. On the other hand, biotic interactions such as competitive exclusion often limit how similar phenotypes can be as species with different traits coexist more readily, fostering trait divergence[11,12]. Attributing convergence or divergence to specific mechanisms is difficult; however, competitive exclusion can also generate convergence when other traits are associated with low competitive abilities[8]. Likewise, divergence can stem from habitat filtering when traits become correlated with distinct sets of environmental controls[13] or when interacting environmental factors select for resident species[14]. Whatever the underlying mechanism, the functional traits of species play an important role in community assembly while also reflecting how species evolved within specific environments. In other words, functional traits reflect past selection and are often conserved within phylogenetic lineages. Species closely related on the evolutionary tree are thus more likely to share similar traits compared to less closely related species. Depending on the pace of evolution, specific traits can be more or less conserved on the phylogenetic tree[15,16]. Indices based on Brownian motion models of trait evolution such as Blomberg's K and Pagel's λ (refs. 17,18) allow us to test whether traits are phylogenetically conserved. These indices are based on correlations between species' distances in trait values and distances along their shared phylogeny[7,19,20].

When species within a community share similar traits, the community is said to show phenotypic clustering, reducing functional diversity (FD). Phenotypic clustering can be associated with two patterns, either a combination of phylogenetic clustering with trait conservatism (Fig. 1, bottom left) or a combination of phylogenetic dispersion with

e-mail: georg.haehn@idiv.de

https://doi.org/10.1038/s41559-024-02589-0

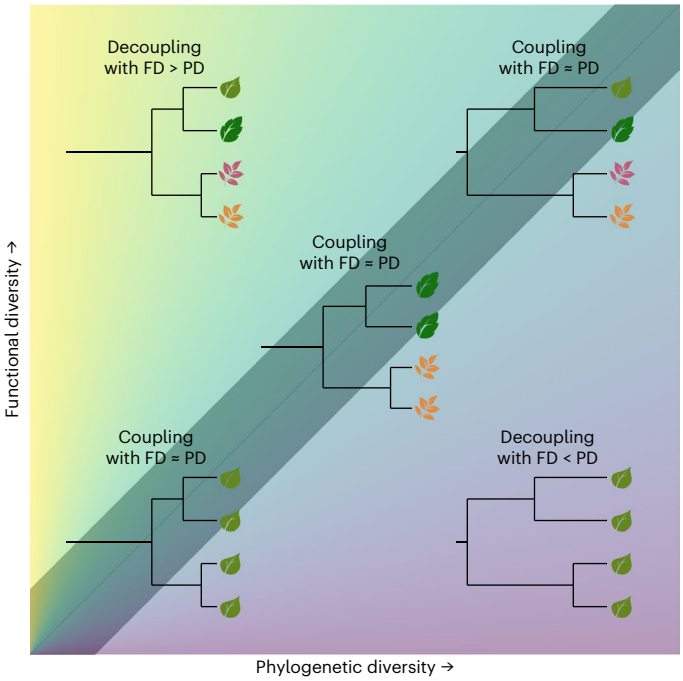

Fig. 1 | Conceptual figure of the relationship between functional and phylogenetic diversity. The figure is conceived after refs. 20,21. If FD is proportional to community PD, we consider the community to be coupled (diagonal). The extremes are the results of either phylogenetic clustering in combination with trait convergence (bottom left) or phylogenetic overdispersion in combination with trait divergence (top right). Decoupled communities can be observed if a community shows phylogenetic overdispersion in combination with trait convergence (bottom right) or if it shows phylogenetic clustering with trait divergence (top left).

trait convergence (Fig. 1, bottom right)[7,15,21]. In the former case, there is a positive covariation between phylogenetic and functional distances, which is why we call the resulting diversity metrics coupled. In the latter case, the phylogenetic and functional distances are inversely related, and thus, we call the resulting diversity metrics decoupled.

In contrast, if species in a community have dissimilar traits, the community has a high phenotypic variation, which is equivalent to a high FD. High FD can either happen in combination with high phylogenetic variation (Fig. 1, top right) or phylogenetic clustering (Fig. 1, top left). Again, in the former case, phylogenetic diversity (PD) and FD are coupled, while being inversely related, and therefore decoupled, in the latter case[21,22]. Many local studies found a prevalence of coupled communities with positive covariation of FD and PD[23–25], but negative covariations[26,27] and unclear patterns[28] have also been encountered. However, it is not yet known under which conditions communities express coupled or decoupled FD and PD.

By calculating FD and PD for 1,781,836 vegetation plots from sPlot[29], the global vegetation plot database, we tested whether patterns of coupling or decoupling (1) dominate at the global level, (2) show regional patterns, (3) differ between forest and non-forest ecosystems and (4) correlate with recent and past climatic gradients. We hypothesized an overall coupled pattern of FD and PD, since PD has often been found to reflect functional trait diversity, especially for those phylogenetically conserved traits that are not easily measurable in plants, such as herbivore and pathogen resistance[15,20,30]. We expected higher PD in forests than in non-forest ecosystems due to the co-occurrence of woody and non-woody plant species, given that the herbaceous habit has evolved from the ancestral woody state multiple times and in different lineages[31–34]. Since PD and FD metrics are correlated with species richness, we used null models to calculate standardized effect sizes (SES) and quantify how much PD and FD differed from random expectations before comparing them[35].

Results

The relationship of functional and phylogenetic diversity

We modelled the relationship between functional and phylogenetic diversity indices expressed as SES of Rao's quadratic entropy based on functional traits (standardized effect size of functional diversity ($SES.FD_Q$)) and phylogenetic distances (standardized effect size of phylogenetic diversity ($SES.PD_Q$)). We considered three functional traits representing the main dimensions of the global spectrum of plant form and function, namely the leaf economics spectrum (specific leaf area), the size-seed mass dimension (plant height) and the root collaboration gradient (specific root length)[36,37]. Both diversity indices were calculated as SES, on the basis of biome-specific null models that account for the varying species richness across plots and use the relative frequencies of species occurrences within each biome to weight species resampling probabilities. This was done

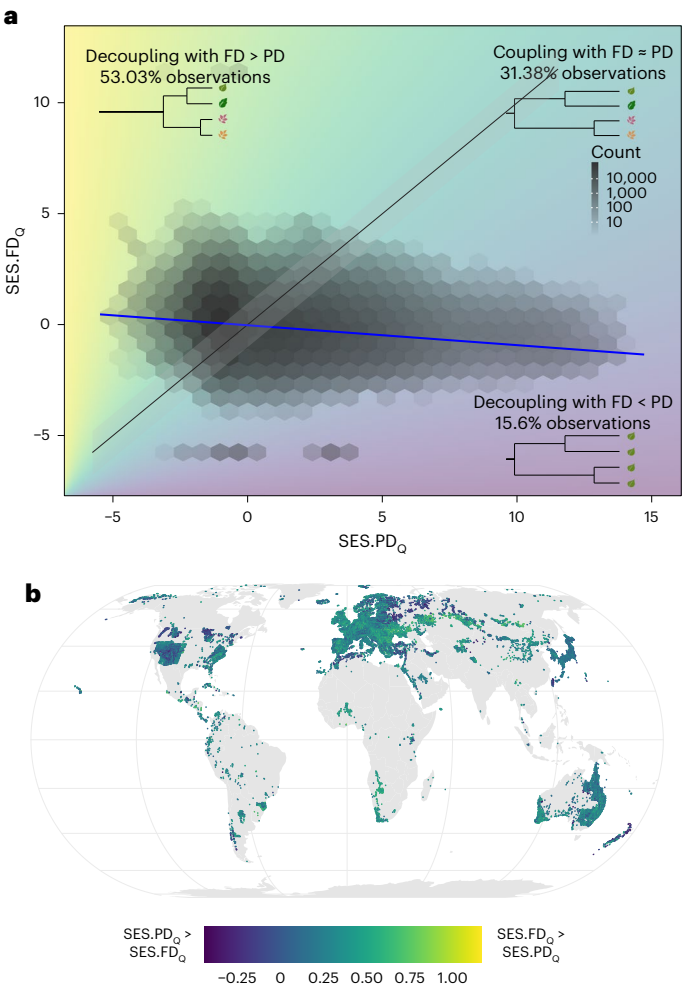

Fig. 2 | The relationship of $SES.FD_Q$ and $SES.PD_Q$. $SES.FD_Q$ is based on three functional traits: specific leaf area, plant height and specific root length. **a**, $SES.FD_Q$ as a function of $SES.PD_Q$ with the linear regression slope (blue) after accounting for spatial autocorrelation within a GAM (7.8% explained deviance). Additionally, the line of coupling with the 1:1 relationship (black) and the confidence interval (grey; Methods), with 31.38% of the observations lying within the confidence interval and 53.03% and 15.6% show decoupling, with either FD > PD or FD < PD, respectively. **b**, Mean log ratio of $SES.FD_Q$ and $SES.PD_Q$ per raster cell (863.8 km^2). Negative values indicate higher observed $SES.PD_Q$ than $SES.FD_Q$ and vice versa. The extracted values from the spatial smoothing spline from the GAM can be found in Supplementary Fig. 2d.

https://doi.org/10.1038/s41559-024-02589-0

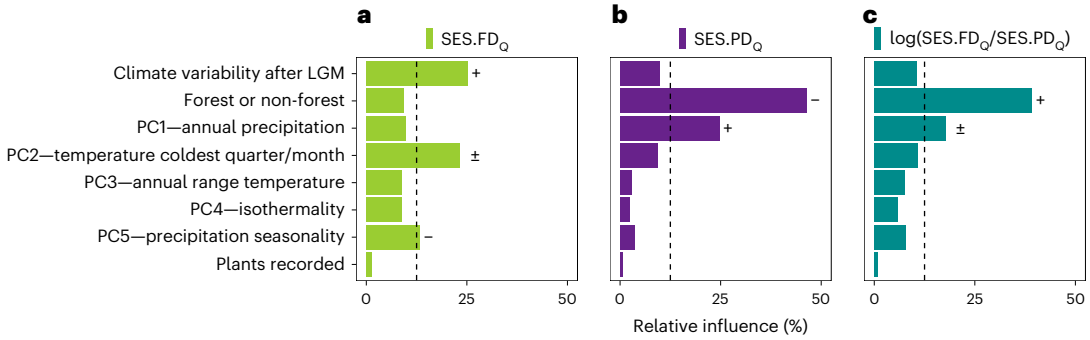

Fig. 3 | Relative influence of environmental variables on functional and phylogenetic diversity. a–**c**, Results of the BRT for functional diversity (SES.FD$_Q$) (**a**), phylogenetic diversity (SES.PD$_Q$) (**b**) and the logarithm of the ratio between SES.FD$_Q$ and SES.PD$_Q$ (**c**). An explanatory variable was considered relevant in the model when its relative influence was >12.5%, indicated by the dashed line, which is the expected influence of a variable if all eight predictors had the same relative importance. The signs indicate the direction of the significant effects based on the partial dependence models (Supplementary Figs. 3 and 4).

because both FD and PD are tightly related to species richness. Out of 1,781,836 vegetation plots, 31.38% showed trait and phylogenetic coupling as SES.FD$_Q$ and SES.PD$_Q$ were simultaneously high or low; 53.03% of the vegetation plots had higher SES.FD$_Q$ than SES.PD$_Q$ and 15.6% had higher SES.PD$_Q$ than SES.FD$_Q$, suggesting that decoupled plant communities are twice as common as coupled ones and that, on average, global communities are more functionally than phylogenetically diverse (Fig. 2a). These results did not change after removing non-significant standardized effect values, that is, values between −1.96 and 1.96 s.d. from the mean (6.9% coupled communities, 45.8% decoupled with high FD values and 17.3% decoupled with high PD values).

We did not find any clear geographical pattern at the global scale (Fig. 2b). Decoupled communities with high SES.FD$_Q$ and low SES.PD$_Q$ (see Methods for definition of high and low values of SES.FD$_Q$ and SES.PD$_Q$) occurred in the western United States and locally across Europe, while communities with low SES.FD$_Q$ and high SES.PD$_Q$ were found close to the Arctic Circle in Scandinavia and Siberia and in New Zealand and Japan. Coupled communities with high values of both diversity indices were encountered in the eastern United States and Central Europe as well as in New Zealand and Japan.

Overall, we found a negative relationship between SES.FD$_Q$ and SES.PD$_Q$. Accounting for the spatial structure of the data by adding a smoothing spline, our general additive model (GAM) explained 7.8% of the deviance in SES.FD$_Q$ (Fig. 2a). Modelling the raw values of FD$_Q$ against the raw values of PD$_Q$, hence not accounting for the effect of species richness, also returned a negative relationship with 18.5% of deviance explained (Supplementary Fig. 1a). The explained deviance increased to 36.2% when the distance matrix of phylogenetic distances was square root-transformed, accounting for the nonlinearity of trait evolution (Supplementary Fig. 1b).

The negative relationship between SES.FD$_Q$ and SES.PD$_Q$ was robust to the use of alternative null models, diversity indices, selections of functional traits and subsets of vegetation plot data (Methods). Using a null model based on a global species pool, SES.PD$_Q$ together with the spatial smoothing spline explained 5.8% of the deviance in SES.FD$_Q$, which increased to 6.2% when the phylogenetic distances were square root-transformed (Supplementary Fig. 1c,d). On the basis of a biome-specific but unweighted species pool, the explained deviance was 6.8% (Supplementary Fig. 1f). When null models were constrained on the basis of a phytogeographic[38] species pool, the explained deviance was 7.8% (Supplementary Fig. 1g). The same negative relationship was found when using alternative indices of FD and PD; that is, when modelling SES of functional dispersion (FDis) against mean pairwise distance (MPD). The explained deviance in this case was 7.1% (Supplementary Fig. 1e). Considering each trait individually, or including additional traits (eight; Methods) but only for an environmentally balanced subset of vegetation plot data (sPlotOpen[39]), also returned a negative relationship between FD$_Q$ and PD$_Q$ (Supplementary Fig. 7 and Supplementary Table 1).

The environmental predictors

We used boosted regression trees (BRT) to select the environmental variables that best explain either SES.FD$_Q$ or SES.PD$_Q$. The BRTs suggested climatic variables to be most relevant for shaping patterns of SES.FD$_Q$ (Fig. 3a). Temperature of the coldest quarter and coldest month (both reflected by PC2 in a principal component analysis (PCA) based on 19 bioclimatic variables) had the highest relative influence on SES.FD$_Q$, followed by the climate variability after the last glacial maximum (LGM) and precipitation seasonality (PC5). Partial dependence plots suggested a predominantly positive relationship between SES.FD$_Q$ and climate variability after the LGM and a negative one with precipitation seasonality (PC5; Supplementary Fig. 3). SES.FD$_Q$ first increased and then decreased with increasing temperatures of the coldest quarter and coldest month (PC2).

Regarding PD, SES.PD$_Q$ was especially related to the vegetation formation type (forest versus non-forest, classified on the basis of the cover of the tree layer and species traits, such as growth form and height; Methods), being higher in forest compared to non-forest ecosystems and tended to increase with annual precipitation (PC1; Fig. 3b and Supplementary Fig. 4a).

When modelling the log ratio of SES.FD$_Q$ to SES.PD$_Q$, BRTs showed that the classification of forest or non-forest and annual precipitation (PC1) had the highest relative influence, resembling what we observed for SES.PD$_Q$ (Fig. 3c and Supplementary Fig. 4b).

From the BRTs, we chose variables with a relative influence >12.5% (the relative influence expected by chance given by 100% or eight explanatory variables) to use in GAMs predicting SES.FD$_Q$ or SES.PD$_Q$ after accounting for spatial autocorrelation. The model for SES.FD$_Q$ explained 4.6% of the deviance and suggested that FD increases with increased climate variability after the LGM and temperatures of the coldest quarter or month (PC2; Fig. 4) and decreases with precipitation seasonality (PC5).

In contrast, the model for PD showed higher explanatory power (37.3% of the deviance) with annual precipitation (PC1), vegetation type and the spatial spline all affecting SES.PD$_Q$. Forests and sites with increased precipitation had higher SES.PD$_Q$ (Fig. 5). Modelling the log ratio between SES.FD$_Q$ and SES.PD$_Q$ confirmed that effects of SES.PD$_Q$ dominate, accounting for 30.8% of the deviance (Fig. 6).

To explore effects of environmental predictors on overall patterns of coupling and decoupling, we modelled the relationship between SES.FD$_Q$ and SES.PD$_Q$ as an ordered categorical variable with three states.

https://doi.org/10.1038/s41559-024-02589-0

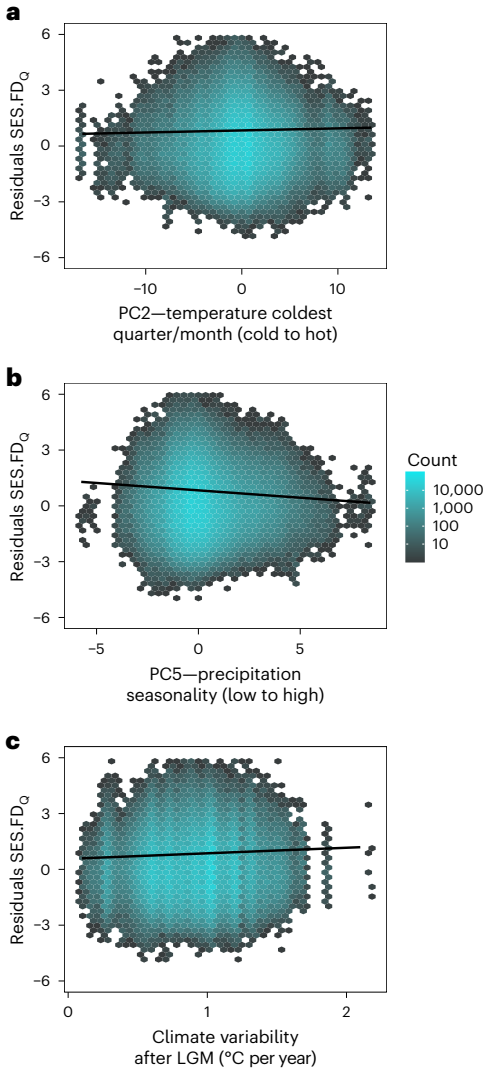

Fig. 4 | Environmental predictors of SES.FD$_Q$. a–**c**, Residuals of SES.FD$_Q$ as a function of temperature of the coldest quarter and month (PC2) (**a**), precipitation seasonality (PC5) (**b**) and climate variability after the LGM (**c**). The GAM explained 4.6% of the deviance. The solid line shows the regression obtained from the GAM. The density hexagons show the distribution of the residuals of the model without the explanatory variable of interest. The smooth term of SES.FD$_Q$ can be found in Supplementary Fig. 6a.

This acknowledges that, while there is only one way for communities to be coupled, decoupling can occur with either PD > FD or FD > PD. Doing this resulted in a model that explained 10.2% of the deviance (Supplementary Fig. 5). Annual precipitation (PC1), precipitation seasonality (PC5) and forest or non-forest had the most power to discriminate the three categories.

Discussion

Plant communities differ in their functional and phylogenetic composition. Here we modelled relationships between FD and PD in plant communities across the globe to infer which factors best predict these separate facets of diversity. Values of FD and PD tend to be decoupled, suggesting that global patterns of community assembly are primarily driven by either FD or PD rather than the two being integrated. Recent climatic conditions and past climatic conditions tended to drive differences in FD. As predicted, we found higher PD in forest versus non-forest communities. The log ratio of FD and PD varied with vegetation type (forest versus non-forest) and recent climatic conditions, in line with what we observed for PD.

Contrary to our hypothesis, we found a negative but weak relationship between FD and PD at the global scale (Fig. 2a). As PD is often considered to be a proxy for capturing unmeasured patterns of species functional traits, we expected a positive relationship between FD and PD[40], as postulated also by theoretical studies[25]. The negative correlation observed at the global scale shows that FD and PD are more often decoupled than coupled in plant communities, with communities either having high PD or FD, which is in line with recent results in grassland communities[26]. Additionally, distribution of traits across phylogenies can vary at small spatial scales, leading to both trait clustering and overdispersion[15,20]. This indicates that, contrary to the expected coupling of FD and PD, closely related species often exhibit considerable differences in trait values, while phylogenetically distant species can often share similar trait values. It is possible that co-occurring species with similar traits differ in other, not easily measurable traits; for example, herbivory resistance, which are captured by phylogeny but less so by other functional traits. Functional clustering could reflect equalizing competitive dynamics in neutrally assembled communities[41] or broader-scale environmental filters. Additionally, when considering the biogeographic histories of lineages, phylogenetic clustering could arise due to recent stochastic extinctions or limited dispersal following allopatric speciation[42].

The observed negative covariation between PD and FD might also be explained by the different impacts of biotic interactions and environmental filtering across communities[41,43,44]. In phylogenetically clustered communities, competitive exclusion may act as a primary mechanism, favouring the co-existence of species with dissimilar phenotypes and thus higher FD. In contrast, environmental filtering seems to be the driving process in communities with low FD and high PD. Here only species with specific phenotypes are admitted to the community[45], but if these come from different clades, the community will exhibit functional convergence but phylogenetic variation. This pattern also suggests that species can differ in features not captured by the traits we use to calculate FD[46]. Since most communities show decoupling with high FD (53%), competition may drive global plant community assembly processes most strongly. However, we must consider that trait divergence can also arise from environmental factors that are spatially nested and interact with each other in filtering species within a community. That is, trait divergence is generated within the studied community units when the filtering effects of fine-scale environmental factors, such as those related to soil and herbivory, interact with and are nested within coarse-scale factors, such as climate[14]. In communities with intermediate values of PD, environmental filtering and competitive exclusion appear to be equally important, resulting in coupled communities. However, the relative importance of these mechanisms is difficult to test as we do not know whether species are excluded from any given community as a result of the environmental conditions, biotic interactions, dispersal limitation or interactions among multiple factors[14,47]. FD and PD could then be decoupled in communities where geographical and local drivers differentially combine with biotic interactions to affect the functional and phylogenetic relationships of species.

We observed no clear spatial patterns relating FD to PD. Plots with coupled and decoupled FD and PD often occurred in geographical proximity, suggesting that local factors can dominate community assembly within regions (Fig. 2b). Previous studies reported geographical patterns of FD based on climatic conditions, such as precipitation gradients[48]. Similarly, PD tends to decrease polewards[49,50]. Studies on the global distribution of PD showed striking differences across ecoregions or biomes[51,52]. Such regional diversity patterns rarely translate into global patterns as broad-scale environmental conditions rarely correspond to local ecological conditions. Nevertheless, treating relationships between FD and PD as a three-level categorical variable (decoupling with higher PD, coupling and decoupling with higher FD) allowed us to demonstrate that coarse-scale environmental factors do

https://doi.org/10.1038/s41559-024-02589-0

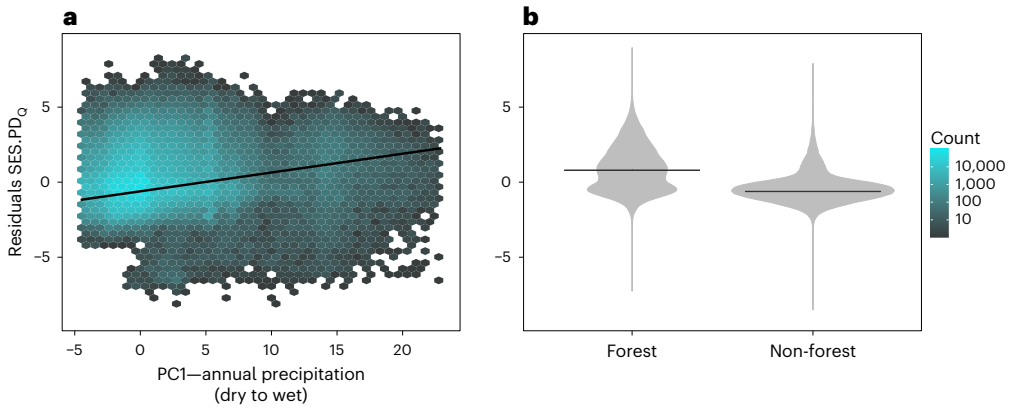

Fig. 5 | Environmental predictors of SES.PD$_Q$. a,**b**, Residuals of SES.PD$_Q$ as a function of annual precipitation (PC1) (**a**) and vegetation type (**b**). The GAM explained 37.3% of the deviance. The solid line shows the regression obtained from the GAM. The density hexagons show the distribution of the residuals of the model without the explanatory variable of interest. The smooth term of SES.PD$_Q$ can be found in Supplementary Fig. 6b.

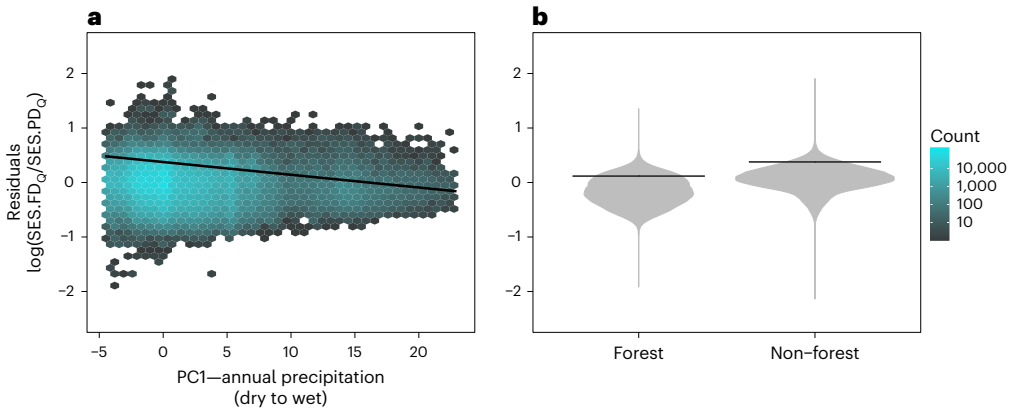

Fig. 6 | Environmental predictors of the log ratio between the SES.FD$_Q$ and SES.PD$_Q$. a,**b**, Residuals of log(SES.FD$_Q$/SES.PD$_Q$) as a function of annual precipitation (PC1) (**a**) and vegetation type (**b**). The GAM explained 30.8% of the deviance. The solid line shows the regression obtained from the GAM. The density hexagons show the distribution of the residuals of the model without the explanatory variable of interest. The smooth term of log(SES.FD$_Q$/SES.PD$_Q$) can be found in Supplementary Fig. 6c.

play a role (Supplementary Fig. 5). This suggests that even though we could not explain the full range of possible combinations of FD and PD, broader biogeographical patterns emerge.

Although SES.FD$_Q$ and environmental conditions sometimes covary, we failed to show that SES.FD$_Q$ is strongly driven by those conditions at the global scale (Fig. 4). In particular, FD was not well explained by recent climatic conditions and climate variability after the LGM. This is in line with studies suggesting that the functional composition of local communities depends mostly on local factors, such as land-use history, soil properties and microclimatic conditions[24,53]. However, a fine classification of biomes as functional units or vegetation types, as was done in a recent Europe-wide analysis on climate–trait relationship[54], might increase the explanatory power of our model.

SES.PD$_Q$ was consistently higher in forests compared to non-forest ecosystems, suggesting that different layers within forest communities support diverse evolutionary histories (Fig. 5). Most tree species belong to predominantly woody families, many of which are phylogenetically distant from other plant families, augmenting the PD found in forest ecosystems[31–33]. This is particularly true for conifers which represent a clade of woody species that separated from the angiosperms of today as early as 300 million years ago (ref. 19). Many forest understories also support cryptogams (including vascular ferns and lycopods) with distinct evolutionary histories relative to trees, further increasing PD in forests[55,56]. These taxa also occur as epiphytes in tropical forests, contributing to their increased PD. Stable microclimatic conditions under a closed canopy could also create conditions favouring species from distinct families[57,58]. Although stratification appeared to increase PD, it did not increase FD.

Overall, our findings suggest that, while forest ecosystems display high PD, the FD of plant species in forests may be limited by convergence in functional traits across different layers. These analyses represent an attempt to understand global relationships between FD and PD but come with limitations. Although sPlot represents a global harmonized database of vegetation plots, its coverage is uneven across biomes and vegetation types, potentially biasing our results. We attempted to correct for this by down-sampled data from the temperate zone in favour of data from the tropics to an environmentally balanced subset. However, we observed an even stronger negative relationship between FD and PD. This suggests that tropical plant communities contribute disproportionately to this pattern. In addition, data in sPlot were collected using various sampling protocols and approaches, sometimes including only woody species and using plots of different shapes and sizes. We sought to partially overcome this problem by including predictors related to plot record characteristics (Methods) and by calculating standardized effect sizes. Still, we do not know how these biases may have affected correlations between FD and PD. We also lacked information on the successional state of the vegetation plots, potentially influencing our results if early successional stages

are lower in FD and PD compared to later successional communities. Because species abundance data are not well standardized in sPlot, it was more robust to use presence–absence data, but this might limit comparisons with other studies. It is also possible that the functional traits we selected might affect the relationships between FD and PD that we observed, especially given that we used only three traits to calculate FD. We note, however, that our results were robust to which traits were selected, individually or jointly, for calculating FD, with these not affecting the relationship between FD and PD (Supplementary Fig. 7 and Supplementary Table 1).

Polytomies included in constructing the phylogeny might have led us to underestimate PD[59], which is why we used SES for PD. Additionally, we found the same negative pattern when we considered FDis and MPD (Supplementary Fig. 1e) as proxy for FD and PD, where the latter is known to show different dispersion patterns than PD_Q (ref. 60). However, when including PD as an explanatory variable in future studies, it is important to consider the relationship between traits and phylogeny and the potential nonlinearity of trait evolution. Additionally, our analysis revealed that none of the potential traits exhibited a strong phylogenetic signal in all families considered in this study (Supplementary Fig. 7b). Moreover, it appeared that certain families tend to possess more conserved traits compared to others. This is in line with other findings that evolutionary conservation can be associated with specific traits and lineages[16], but this is not a common pattern. Consequently, depending on the sampled community and plant species, different patterns may emerge in the relationship between FD and PD. While both plant characteristics and evolutionary history play crucial roles in community assembly processes, just which interacting mechanisms operate on which underlying biotic and abiotic factors remains unclear.

Our findings on the relationship of $SES.FD_Q$ and $SES.PD_Q$ imply that ecological communities can exhibit many combinations of FD and PD. The decoupling of FD and PD found here plus the overall slightly negative correlation imply that competitive exclusion may commonly occur in plant communities. Our results also highlight the need to conserve both FD and PD if we are to safeguard biodiversity. Both FD and PD play key roles in community assembly and probably affect how species and their interactions within communities will respond to changing climates and other drivers of global change. Future research may reveal which regional conditions contribute to hotspots of FD and PD and why. Understanding the diverse and context-dependent nature of FD and PD will shed light on the complex dynamics of ecological communities and help us to design schemes to better protect the diversity they support.

Methods

Species community data

The vegetation plot database sPlot[29] (www.idiv.de/splot) is a harmonized collection of national- and regional-scale vegetation plot datasets. sPlot provides georeferenced information on the presence and abundance of all vascular plants co-occurring in a sampling area, that is, vegetation plot. The database sPlot v.3.0 holds a total number of 1,977,637 vegetation plot records from 160 datasets collected between 1873 and 2019, across six continents and most biomes, including 76,912 vascular plant species (for v.2.1; ref. 29). The size of a plot varies according to the type of vegetation being sampled, from 1 m^2 in grasslands to 250,000 m^2 in forest ecosystems. The vegetation type of a plot was classified as forest and non-forest on the basis of tree layer cover and the growth form of dominant species[29]. Vegetation plot records were included in the study if the cumulative coverage of species for which both trait and phylogenetic information was available accounted for at least 50% of the relative vegetation cover in that plot (see below).

In addition, we used sPlotOpen[39], which is an environmentally balanced, open-access subset of sPlot, as a benchmark of our results, both when testing for the effect of trait selection when calculating FD and for the effect of uneven coverage of sPlot data across the Earth's biomes.

Functional diversity

Plant functional traits were available from the gap-filled TRY v.5.0 database[61–64]. We calculated FD as Rao's quadratic entropy (FD_Q) as well as FDis for all vegetation plots in sPlot 3.0. The calculation of Rao's quadratic entropy[65] is based on a Gower distance matrix calculated for the species present in each vegetation plot. FDis was computed from the uncorrected species–species distance matrix with the function dbFD from the R package FD[66,67]. We based this calculation on three functional traits selected to cover most of the variation within plant traits and to represent different axes in the plant economic spectrum, that is, belowground and resource strategy of acquisition or conservation (specific root length and specific leaf area) and reproduction strategy of quality or quantity (plant height)[37,68]. To evaluate the influence of trait selection on the relationship of FD and PD, we calculated FD_Q on eight functional traits (specific leaf area, specific root length, seed mass, plant height, leaf phosphorus and nitrogen content, leaf dry matter content, and chromosome number), both taken individually and jointly. We did this additional analysis based on the sPlotOpen subset only, since calculating SES (see below) of FD calculated on eight traits in all plots was computationally unfeasible, even using a high-performance cluster. Additionally, considering all eight traits for the complete dataset would have led to a loss of ~2,000 species (~10% of species considered in this study, see below) due to missing data in the TRY database.

Functional traits can be conserved in the phylogeny. This was tested with two evolutionary models (Blomberg's K and Pagel's λ), where the latter is known to be more robust against incomplete resolved phylogenies or suboptimal branch lengths[17,18]. K and λ were calculated using the function phylosig from the R package picante[69]. In contrast to other tests for phylogenetic signals, both models can be used to compare phylogenetic signals across different phylogenies[17], which needs to be done as a global plant phylogeny is simply too large for an appropriate calculation of phylogenetic signals. Therefore, the phylogenetic signal for each trait was calculated within each family. All eight functional traits showed either no or low phylogenetic signals for λ and K (Supplementary Fig. 7b,c). Therefore, we assume that there is also no phylogenetic signal across vascular plants for the considered traits.

Phylogenetic diversity

For all species present in sPlot, a phylogenetic tree was built using the function phylo.maker from the R package V.PhyloMaker[70]. The phylogenetic backbone of the package is the combination of GenBank taxa with a backbone provided by the Open Tree of Life v.9.1, for seed plants[71] and the clade of pteridophytes[72]. Missing genera were inserted to the half point of the family tree. This approach was evaluated by ref. 73, who showed that phylogenetic indices based on the calculated tree were highly correlated with indices based on the 'PhytoPhylo megaphylogeny' (updated phylogenetic tree from ref. 72). Species that could not be inserted by the phylo.maker were bound to the half of the terminal level of a sister species if only one species was available in this genus or to the most recent ancestor (MRCA) if the genus included more than one species. This additional binding was done with the bind.node function from the R package phytools[74].

The computed phylogenetic tree for sPlot contained 160 families with 68,052 of 76,912 species (88%) present within the database. Additional 3,802 species were included, with 3,348 being bound to the node of the MRCA of already-present sister species and 454 species to the half of the terminal level on the family node. The final phylogenetic tree contained 71,854 species on 32,395 nodes. A total of 31,727 species in the phylogeny also had traits in the TRY database. Of this subset, 322 species (~1%) were bound to the half of the terminal level on the family node and 2,766 (~9%) to the MRCA. Vegetation plot records were only included in the analysis if both trait and phylogenetic information was available for at least 50% of the total relative

cover of the species in that plot. In total, 1,781,836 out of 1,977,637 plot records remained.

PD was calculated as Rao's quadratic entropy (PD_Q) which amounts to the mean nearest taxon distance for presence–absence data. We used the function raoD from the R package picante[69], which is based on the cophenetic distance of all *n* species in the phylogeny, pruned to contain only the species in that plot. To account for the nonlinearity of evolutionary histories, we also calculated PD_Q on the basis of the square root-transformed cophenetic distance[75]. Additionally we calculated MPD, to be compared with FDis, as MPD could show opposite dispersion patterns compared with PD_Q (ref. 60). Only species with both trait information and known phylogeny were used to calculate FD and PD.

Standardized effect size

The species richness of the vegetation plot records ranged from 1 to 412 species (Supplementary Fig. 8). FD and PD indices are known to depend on species richness[76–78]. Especially for FD, a higher number of species in a community is more likely to return higher FD values than are communities with fewer species[77]. We controlled for species richness by calculating the SES of each diversity index for every vegetation plot record[79], fixing the number of species of the plot record and drawing species randomly, which is equivalent to shuffling traits across species. As species do not equally occur across the globe, we calculated our null expectations on the basis of biome-specific species pools accounting for the frequency of species in the plot records in each biome. However, to see if the patterns also hold true for broader species pools we used the following hierarchical approach with four stages of defined species pools. For the simplest species pool, we calculated our null expectations based on all species present in the whole sPlot database, so we allowed each species to occur everywhere in the world. For a more geographically constrained approach we calculated the null expectations based on species pools within 16 phytogeographical units[38] (stage 2) and 10 predefined biomes (stage 3) in response to global climate variation[29,80], namely: alpine, boreal zone, dry midlatitudes, dry tropics and subtropics, polar and subpolar zone, subtropics with winter rain, subtropics with year-round rain, temperate midlatitudes, tropics with summer rain and tropics with year-round rain. The fourth and most complex null model was based on the species pool within each biome, additionally sampling the species weighted by their frequency in the plot records within each biome. This means a species that occurred more frequently within a biome was randomly drawn more often to recalculate the null diversity index, compared to a species occurring less often. For each of the four null models, we calculated the mean and standard deviation of the distribution of null functional and phylogenetic indices across 499 draws. Vegetation plots only containing one species, or for which trait and phylogenetic information was not available, were excluded from functional or phylogenetic diversity calculations. SES were obtained by subtracting the mean index of the randomized data from the observed index and dividing the result by the standard deviation of the index of the randomized data.

Definition of coupling and decoupling

To measure the percentage of coupled and decoupled communities, a confidence interval was defined. We randomly drew one million values from a uniform distribution, defined between the minimum and maximum of observed standardized effect sizes of Rao's quadratic entropy based on functional traits ($SES.FD_Q$) as explanatory variable. We created a correlated response variable by adding an error from a normal distribution, obtained from the mean and the standard deviation of the observed $SES.FD_Q$. We fitted a linear model and extracted the intercept and the confidence interval. Communities with an observed value of $SES.FD_Q$ were considered coupled if the standardized effect sizes of Rao's quadratic entropy based on phylogenetic distance ($SES.PD_Q$) fell within this interval. On the basis of this, we defined three categories of community patterns; that is, decoupling with higher FD than PD, coupling and decoupling with lower FD than PD. This variable was later used as an ordered categorical response. Additionally, we calculated the log ratio between $SES.FD_Q$ and $SES.PD_Q$ as $\log(SES.FD_Q/SES.PD_Q)$ after scaling the values between 0.001 and 1. Positive and negative values define the deviation with higher and lower $SES.FD_Q$ than $SES.PD_Q$, respectively, from a perfect coupled community.

Explanatory variables

Recent climatic conditions (1981–2010) were represented by the 19 bioclimatic variables from CHELSA v.2.1 (refs. 81,82). A PCA was performed to reduce data dimensionality. In the following analyses, we only used the first five PCA axes, collectively accounting for 92.3% of the explained variation. We interpreted the axes on the basis of the highest loadings of the corresponding climatic variable as follows: annual precipitation for PC1; mean daily air temperature of the coldest quarter and mean daily minimum air temperature of the coldest month for PC2; annual air temperature range for PC3; isothermality for PC4; and precipitation seasonality for PC5 (Supplementary Table 2 and Supplementary Fig. 9).

Mean air temperature variability after the LGM was derived from the open-access StableClim v.1.1. dataset, containing estimates from 21,000 years ago at 2.5° spatial resolution[83]. Climatic variability represents rapid global warming during the last deglaciation during the Bølling–Allerød transition[84] on land and sea. The mean temperature variability between 21,000 BP and AD 100 was used as index for the climatic variability after the LGM.

All climatic variables were extracted for each plot with the extract function from the R package raster[85].

Not all vegetation plot records were complete in terms of the sampled functional groups. Records from tropical forest plots often contained either only tree data or tree and shrub data. As the exclusion of those plots would have substantially reduced the spatial coverage of our model, we added the nominal predictor variable called 'plants recorded' to our models to partially control for this source of bias. The variable 'plants recorded' has four values: all vascular plants, only dominant species, all woody plants and only trees. Additionally, we used the vegetation type (forest versus non-forest) from the vegetation plot database sPlot as predictor variable.

In total, we prepared eight explanatory variables, five related to the recent climatic conditions, one to past climatic variability and two to plot record characteristics.

Statistical modelling

A GAM was used to model the relationship between FD and PD, either expressed as observed Rao's quadratic entropy (for PD also after a square root transformation of the distance matrix) or as standardized effect size of Rao's quadratic entropy, FDis and MPD. In a GAM, the linear response can depend on unknown smooth functions of the explanatory variables. To account for the spatial structure of the data, the spatial coordinates were included as smooth spherical splines. All GAMs included a basis penalty smoother spline on the sphere (bs = "sos"), applied to the geographic coordinates of every plot, thus taking spatial autocorrelation into account. The explanatory variable was included as linear predictors without any smooth function. The model was performed using the function gam from the R package mgcv[86–91], defined as follows:

$$gam(SES.FD_Q \sim SES.PD_Q + s(Longitude, Latitude, bs = \text{"sos"}), family = \text{"gaussian"}, method = \text{"REML"})$$

$SES.FD_Q$ is the standardized effect size of Rao's quadratic entropy based on the three selected functional plant traits and $SES.PD_Q$ is the standardized effect size of Rao's quadratic entropy based on the phylogenetic distances of species present in the community. This step was done for the complete dataset and for the sPlotOpen subset, for

which we considered the eight traits, both individually and jointly, for calculating standardized effect size of FD.

To model the relationship between either FD or PD and the set of the eight explanatory variables described above, we used a two-step approach. In the first step, we used BRTs to select relevant explanatory variables and quantify their relative influence. In the second step, we fitted GAMs using functional, PD or their log ratio as response variables and the predictors selected in the first step as explanatory variables. We did this because fitting a full GAM algorithm with all predictors would lead to convergence issues, due to the huge number of data points.

BRTs are a machine learning technique used in regression and classification having few prior assumptions and being robust against overfitting and collinearity. They are known to uncover nonlinear relationships as well as interactions among predictors. The parameters of the BRT were set as follows: a tree complexity of 5 and a bag fraction of 0.5. The learning rate was set to 0.01 with a maximum number of 20,000 trees. The BRTs were calculated using the gbm.step routine from the dismo package[92]. An explanatory variable was considered relevant in the model if its relative influence was >12.5%, which is the expected influence of a variable if all the eight predictors had an equal relative importance.

The variables that were considered as relevant from the BRTs were then used in a second set of GAMs, having as response variable SES.FD_Q, SES.PD_Q or their log ratio; and as explanatory variables those that turned out to be relevant in the corresponding BRT. Additionally, we fitted a GAM with the ordered categorical response of coupling and decoupling against the environmental predictors, which were selected by the BRTs for FD and PD. As the three categories were not equally represented, we sampled 10,000 communities for each category and repeated the GAM 100 times, besides running the same model on the complete (unbalanced) dataset. The spatial coordinates were included as smooth spherical splines in all models as explained above. As not all vegetation plot entries in sPlot are classified as forest/non-forest the number of observations for the environmental models was 1,497,238. The prediction of each explanatory variable was performed using the prediction function from the R package marginaleffects[93] by predicting the explanatory variable based on the sequence between the minimum and maximum of the variable in the original data and the GAM model. The plotted regressions were obtained by extracting the residuals from a GAM without the explanatory variable of interest.

For plotting, functional and phylogenetic variables were averaged for each grid cell with a size of 863.8 km^2. The spatial smoother within the GAM was plotted at the same resolution based on the following model (example based on SES.FD_Q):

gam(SES.FD_Q ~ 1 + s(Longitude, Latitude, bs = "sos"), family = "gaussian", method = "REML").

All analyses were performed in R v.4.1.3 (ref. 94).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All calculated biodiversity indices necessary to reproduce the results of this paper are available at https://doi.org/10.25829/idiv.3574-mpmk21 (ref. 95). The vegetation plot raw data for sPlotOpen are available at https://www.idiv.de/de/splot/splotopen.html. The vegetation plot raw data contained in the sPlot database are available upon request by submitting a project proposal to the sPlot Steering Committee. The proposals should follow the Governance and Data Property Rules of the sPlot Working Group available on the sPlot website (www.idiv.de/splot). Source data are provided with this paper.

Code availability

All R scripts used for this study can be found in our GitHub repository at https://github.com/georghaehn/Haehn-et-al-2024-FD-PD-coupling.

References

1. O'Connor, B., Bojinski, S., Röösli, C. & Schaepman, M. E. Monitoring global changes in biodiversity and climate essential as ecological crisis intensifies. *Ecol. Inform.* **55**, 101033 (2020).
2. Anwar, M. R., Liu, D. L., Macadam, I. & Kelly, G. Adapting agriculture to climate change: a review. *Theor. Appl. Climatol.* **113**, 225–245 (2013).
3. Benevolenza, M. A. & DeRigne, L. The impact of climate change and natural disasters on vulnerable populations: a systematic review of literature. *J. Hum. Behav. Soc. Environ.* **29**, 266–281 (2019).
4. IPCC *Climate Change 2023: Synthesis Report* (eds Core Writing Team, Lee, H. & Romero, J.) (IPCC, 2023).
5. Fahad, S. et al. *Climate Change and Plants: Biodiversity, Growth and Interactions* (CRC, 2021).
6. Corlett, R. T. & Westcott, D. A. Will plant movements keep up with climate change? *Trends Ecol. Evol.* **28**, 482–488 (2013).
7. Cavender-Bares, J., Kozak, K. H., Fine, P. V. A. & Kembel, S. W. The merging of community ecology and phylogenetic biology. *Ecol. Lett.* **12**, 693–715 (2009).
8. Götzenberger, L. et al. Ecological assembly rules in plant communities—approaches, patterns and prospects. *Biol. Rev.* **87**, 111–127 (2012).
9. Rieseberg, L. H., Wood, T. E. & Baack, E. J. The nature of plant species. *Nature* **440**, 524–527 (2006).
10. Verdú, M. & Pausas, J. G. Fire drives phylogenetic clustering in Mediterranean Basin woody plant communities. *J. Ecol.* **95**, 1316–1323 (2007).
11. Ackerly, D. D., Schwilk, D. W. & Webb, C. O. Niche evolution and adaptive radiation: testing the order of trait divergence. *Ecology* **87**, 50–61 (2006).
12. Pillar, V. D., Duarte, L. d. S., Sosinski, E. E. & Joner, F. Discriminating trait-convergence and trait-divergence assembly patterns in ecological community gradients. *J. Veg. Sci.* **20**, 334–348 (2009).
13. Pillar, V. D., Sabatini, F. M., Jandt, U., Camiz, S. & Bruelheide, H. Revealing the functional traits linked to hidden environmental factors in community assembly. *J. Veg. Sci.* **32**, e12976 (2021).
14. Pillar, V. D. Trait divergence in plant community assembly is generated by environmental factor interactions. *J. Veg. Sci.* **35**, e13259 (2024).
15. Ackerly, D. Conservatism and diversification of plant functional traits: evolutionary rates versus phylogenetic signal. *Proc. Natl Acad. Sci. USA* **106**, 19699–19706 (2009).
16. Ávila-Lovera, E., Winter, K. & Goldsmith, G. R. Evidence for phylogenetic signal and correlated evolution in plant–water relation traits. *New Phytol.* **237**, 392–407 (2023).
17. Münkemüller, T. et al. How to measure and test phylogenetic signal. *Methods Ecol. Evol.* **3**, 743–756 (2012).
18. Molina-Venegas, R. & Rodríguez, M. Á. Revisiting phylogenetic signal; strong or negligible impacts of polytomies and branch length information? *BMC Evol. Biol.* **17**, 53 (2017).
19. Melzer, R., Wang, Y.-Q. & Theißen, G. The naked and the dead: the ABCs of gymnosperm reproduction and the origin of the angiosperm flower. *Semin. Cell Dev. Biol.* **21**, 118–128 (2010).
20. Cavender-Bares, J., Ackerly, D. D., Baum, D. A. & Bazzaz, F. A. Phylogenetic overdispersion in Floridian oak communities. *Am. Nat.* **163**, 823–843 (2004).
21. Cadotte, M., Albert, C. H. & Walker, S. C. The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecol. Lett.* **16**, 1234–1244 (2013).
22. Srivastava, D. S., Cadotte, M. W., MacDonald, A. A. M., Marushia, R. G. & Mirotchnick, N. Phylogenetic diversity and the functioning of ecosystems. *Ecol. Lett.* **15**, 637–648 (2012).

23. Webb, C. O. Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *Am. Nat.* **156**, 145–155 (2000).
24. Flynn, D. F. B., Mirotchnick, N., Jain, M., Palmer, M. I. & Naeem, S. Functional and phylogenetic diversity as predictors of biodiversity—ecosystem–function relationships. *Ecology* **92**, 1573–1581 (2011).
25. Tucker, C. M., Davies, T. J., Cadotte, M. W. & Pearse, W. D. On the relationship between phylogenetic diversity and trait diversity. *Ecology* **99**, 1473–1479 (2018).
26. Večeřa, M. et al. Decoupled phylogenetic and functional diversity in European grasslands. *Preslia* **95**, 413–445 (2023).
27. Prinzing, A. et al. Less lineages—more trait variation: phylogenetically clustered plant communities are functionally more diverse. *Ecol. Lett.* **11**, 809–819 (2008).
28. Kluge, J. & Kessler, M. Phylogenetic diversity, trait diversity and niches: species assembly of ferns along a tropical elevational gradient. *J. Biogeogr.* **38**, 394–405 (2011).
29. Bruelheide, H. et al. sPlot—a new tool for global vegetation analyses. *J. Veg. Sci.* **30**, 161–186 (2019).
30. Castagneyrol, B., Jactel, H., Vacher, C., Brockerhoff, E. G. & Koricheva, J. Effects of plant phylogenetic diversity on herbivory depend on herbivore specialization. *J. Appl. Ecol.* **51**, 134–141 (2014).
31. Qian, H., Hao, Z. & Zhang, J. Phylogenetic structure and phylogenetic diversity of angiosperm assemblages in forests along an elevational gradient in Changbaishan, China. *J. Plant Ecol.* **7**, 154–165 (2014).
32. Honorio Coronado, E. N. et al. Phylogenetic diversity of Amazonian tree communities. *Divers. Distrib.* **21**, 1295–1307 (2015).
33. Mastrogianni, A., Kallimanis, A. S., Chytrý, M. & Tsiripidis, I. Phylogenetic diversity patterns in forests of a putative refugial area in Greece: a community level analysis. *For. Ecol. Manag.* **446**, 226–237 (2019).
34. Klimeš, A., Šímová, I., Zizka, A., Antonelli, A. & Herben, T. The ecological drivers of growth form evolution in flowering plants. *J. Ecol.* **110**, 1525–1536 (2022).
35. Chai, Y. et al. Patterns of taxonomic, phylogenetic diversity during a long-term succession of forest on the Loess Plateau, China: insights into assembly process. *Sci. Rep.* **6**, 27087 (2016).
36. Díaz, S. et al. The global spectrum of plant form and function. *Nature* **529**, 167–171 (2016).
37. Weigelt, A. et al. An integrated framework of plant form and function: the belowground perspective. *New Phytol.* **232**, 42–59 (2021).
38. Carta, A., Peruzzi, L. & Ramírez-Barahona, S. A global phylogenetic regionalization of vascular plants reveals a deep split between Gondwanan and Laurasian biotas. *New Phytol.* **233**, 1494–1504 (2022).
39. Sabatini, F. M. et al. sPlotOpen—an environmentally balanced, open-access, global dataset of vegetation plots. *Glob. Ecol. Biogeogr.* **30**, 1740–1764 (2021).
40. Reich, P. B. et al. The evolution of plant functional variation: traits, spectra, and strategies. *Int. J. Plant Sci.* **164**, 143–164 (2003).
41. Mayfield, M. M. & Levine, J. M. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecol. Lett.* **13**, 1085–1093 (2010).
42. Pigot, A. L. & Etienne, R. S. A new dynamic null model for phylogenetic community structure. *Ecol. Lett.* **18**, 153–163 (2015).
43. Godoy, O., Kraft, N. J. B. & Levine, J. M. Phylogenetic relatedness and the determinants of competitive outcomes. *Ecol. Lett.* **17**, 836–844 (2014).
44. Kraft, N. J. B., Godoy, O. & Levine, J. M. Plant functional traits and the multidimensional nature of species coexistence. *Proc. Natl Acad. Sci. USA* **112**, 797–802 (2015).
45. de Bello, F. et al. *Handbook of Trait-Based Ecology: From Theory to R Tools* (Cambridge Univ. Press, 2021).
46. Owen, N. R., Gumbs, R., Gray, C. L. & Faith, D. P. Global conservation of phylogenetic diversity captures more than just functional diversity. *Nat. Commun.* **10**, 859 (2019).
47. Kraft, N. J. B. et al. Community assembly, coexistence and the environmental filtering metaphor. *Funct. Ecol.* **29**, 592–599 (2015).
48. Zuo, X. et al. Functional diversity response to geographic and experimental precipitation gradients varies with plant community type. *Funct. Ecol.* **35**, 2119–2132 (2021).
49. Massante, J. C. et al. Contrasting latitudinal patterns in phylogenetic diversity between woody and herbaceous communities. *Sci. Rep.* **9**, 6443 (2019).
50. Cai, H. et al. Geographical patterns in phylogenetic diversity of Chinese woody plants and its application for conservation planning. *Divers. Distrib.* **27**, 179–194 (2021).
51. Tietje, M. et al. Global hotspots of plant phylogenetic diversity. *New Phytol.* **240**, 1636–1646 (2023).
52. Qian, H., Zhang, J. & Jiang, M. Global patterns of taxonomic and phylogenetic diversity of flowering plants: biodiversity hotspots and coldspots. *Plant Divers.* **45**, 265–271 (2023).
53. De Pauw, K. et al. Taxonomic, phylogenetic and functional diversity of understorey plants respond differently to environmental conditions in European forest edges. *J. Ecol.* **109**, 2629–2648 (2021).
54. Kambach, S. et al. Climate–trait relationships exhibit strong habitat specificity in plant communities across Europe. *Nat. Commun.* **14**, 712 (2023).
55. Pryer, K. M. et al. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. *Nature* **409**, 618–622 (2001).
56. Rothfels, C. J. et al. The evolutionary history of ferns inferred from 25 low-copy nuclear genes. *Am. J. Bot.* **102**, 1089–1107 (2015).
57. De Frenne, P. et al. Forest microclimates and climate change: importance, drivers and future research agenda. *Glob. Change Biol.* **27**, 2279–2297 (2021).
58. Kovács, B., Tinya, F. & Ódor, P. Stand structural drivers of microclimate in mature temperate mixed forests. *Agric. For. Meteorol.* **234–235**, 11–21 (2017).
59. Swenson, N. Phylogenetic resolution and quantifying the phylogenetic diversity and dispersion of communities. *PLoS ONE* **4**, e4390 (2009).
60. Sessa, E. B. et al. Community assembly of the ferns of Florida. *Am. J. Bot.* **105**, 549–564 (2018).
61. Kattge, J. et al. TRY plant trait database—enhanced coverage and open access. *Glob. Change Biol.* **26**, 119–188 (2020).
62. Shan, H. et al. Gap filling in the plant kingdom: trait prediction using hierarchical probabilistic matrix factorization. Preprint at https://arxiv.org/abs/1206.6439 (2012).
63. Fazayeli, F., Banerjee, A., Kattge, J., Schrodt, F. & Reich, P. B. Uncertainty quantified matrix completion using bayesian hierarchical matrix factorization. In *2014 13th International Conference on Machine Learning and Applications* (eds Chen, X.-w. et al.) 312–317 (IEEE, 2014).
64. Schrodt, F. et al. BHPMF—a hierarchical Bayesian approach to gap-filling and trait prediction for macroecology and functional biogeography. *Glob. Ecol. Biogeogr.* **24**, 1510–1521 (2015).
65. Rao, C. R. Diversity and dissimilarity coefficients: a unified approach. *Theor. Popul. Biol.* **21**, 24–43 (1982).
66. Laliberté, E. & Legendre, P. A distance-based framework for measuring functional diversity from multiple traits. *Ecology* **91**, 299–305 (2010).
67. Laliberté, E., Legendre, P. & Shipley, B. FD: measuring functional diversity from multiple traits, and other tools for functional ecology. R package version 1.0-12 (2014).

68. Walker, A. P., McCormack, M. L., Messier, J., Myers-Smith, I. H. & Wullschleger, S. D. Trait covariance: the functional warp of plant diversity? *New Phytol.* **216**, 976–980 (2017).
69. Kembel, S. W. et al. picante: integrating phylogenies and ecology. R package version 1.8.2 (2020).
70. Jin, Y. & Qian, H. V. PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography* **42**, 1353–1359 (2019).
71. Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105**, 302–314 (2018).
72. Zanne, A. E. et al. Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).
73. Qian, H. & Jin, Y. An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *J. Plant Ecol.* **9**, 233–239 (2016).
74. Revell, L. J. phytools: phylogenetic tools for comparative biology (and other things). R package version 2.1-1 (2023).
75. Letten, A. D. & Cornwell, W. K. Trees, branches and (square) roots: why evolutionary relatedness is not linearly related to functional distance. *Methods Ecol. Evol.* **6**, 439–444 (2015).
76. de Bello, F., Carmona, C. P., Lepš, J., Szava-Kovats, R. & Pärtel, M. Functional diversity through the mean trait dissimilarity: resolving shortcomings with existing paradigms and algorithms. *Oecologia* **180**, 933–940 (2016).
77. Petchey, O. L. & Gaston, K. J. Extinction and the loss of functional diversity. *Proc. R. Soc. Lond. B* **269**, 1721–1727 (2002).
78. Cadotte, M. W. et al. Phylogenetic diversity metrics for ecological communities: integrating species richness, abundance and evolutionary history. *Ecol. Lett.* **13**, 96–105 (2010).
79. Gotelli, N. J. & McCabe, D. J. Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. *Ecology* **83**, 2091–2096 (2002).
80. Schultz, J. *The Ecozones of the World: The Ecological Division of the Geosphere* (Springer Nature, 2005).
81. Karger, D. N. et al. Climatologies at high resolution for the Earth's land surface areas. *Sci. Data* **4**, 170122 (2017).
82. Karger, D. N. et al. Data from: Climatologies at high resolution for the Earth's land surface areas. *Dryad* https://doi.org/10.5061/DRYAD.KD1D4 (2018).
83. Brown, S. C., Wigley, T. M. L., Otto-Bliesner, B. L. & Fordham, D. A. StableClim, continuous projections of climate stability from 21000 BP to 2100 CE at multiple spatial scales. *Sci. Data* **7**, 335 (2020).
84. Renssen, H. & Isarin, R. F. B. The two major warming phases of the last deglaciation at ~14.7 and ~11.5 ka cal BP in Europe: climate reconstructions and AGCM experiments. *Glob. Planet. Change* **30**, 117–153 (2001).
85. Hijmans, R. J. Raster: Geographic data analysis and modeling. R package version 3.6-30 (2023).
86. Wood, S. mgcv: Mixed GAM computation vehicle with automatic smoothness estimation. R package version 1.9-1 (2023).
87. Wood, S. N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. B* **73**, 3–36 (2011).
88. Wood, S. N. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *J. Am. Stat. Assoc.* **99**, 673–686 (2004).
89. Wood, S. N. *Generalized Additive Models: An Introduction with R* (Chapman and Hall/CRC, 2017).
90. Wood, S. N. Thin-plate regression splines. *J. R. Stat. Soc. B* **65**, 95–114 (2003).
91. Wood, S. N., Pya, N. & Säfken, B. Smoothing parameter and model selection for general smooth models. *J. Am. Stat. Assoc.* **111**, 1548–1563 (2016).
92. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. dismo: species distribution modeling. R package version 1.3-14 (2022).
93. Arel-Bundock, V. marginaleffects: predictions, comparisons, slopes, marginal means, and hypothesis tests. R package version 0.18.0 (2023).
94. R Core Team *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2022).
95. Hähn, G. J. A., Damasceno, G., Sabatini, F. M. & Bruelheide, H. Global decoupling of functional and phylogenetic diversity in plant communities (version 1.0). *iDiv* https://doi.org/10.25829/idiv.3574-mpmk21 (2024).

Acknowledgements

We are thankful for the efforts of thousands of vegetation scientists sampling and digitalizing vegetation data and making them available in regional, national or international databases. We appreciate the support of the German Research Foundation for funding sPlot as one of the iDiv research platforms (DFG FZT 118, 202548816). The scientific results have (in part) been computed at the High-Performance Computing Cluster EVE, a joint effort of both the Helmholtz Centre for Environmental Research—UFZ (http://www.ufz.de/) and the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig (http://www.idiv-biodiversity.de/). We thank the administration and support staff of EVE who keep the system running and support us with our scientific computing needs: T. Schnicke, B. Langenberg, G. Schramm, T. Harzendorf, T. Strempel and L. Schurack from the UFZ and C. Krause from iDiv. We thank the iDiv Data & Code Unit for assistance with curation (done by L. Figueiredo) and archiving of the dataset. F.M.S. gratefully acknowledges financial support from the Rita Levi Montalcini (2019) programme, funded by the Italian Ministry of University and Research (MUR). J.-C.S. considers this work a contribution to Center for Ecological Dynamics in a Novel Biosphere (ECONOVO), funded by Danish National Research Foundation (grant no. DNRF173) and his VILLUM Investigator project 'Biodiversity Dynamics in a Changing World', funded by VILLUM FONDEN (grant no. 16549). V.D.P. received support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil, grant no. 313315/2022-1). I.B. and J.A.C. were funded by the Basque Government (IT1487-22). A.D.B was supported by the Knut and Alice Wallenberg Foundation (WAF KAW 2019.0202) and the Swedish Foundation for Strategic Research (FFL21-0194). A.G.-d.-M. has been supported by National Forestry and Wildlife Service (SERFOR) of Peru (AUT-IFL-2023-017) and Fundación Universitaria San Pablo-CEU, grant nos. GNRI 2023 and GNRI 2024. A.Č., F.K. and U. Šilc were supported by the Slovenian Research and Innovation Agency (P1-0236).

Author contributions

G.J.A.H., F.M.S. and H.B. conceived the idea. G.J.A.H. performed the analysis with substantial input from F.M.S., G.D. and H.B. G.J.A.H. drafted the first version of the paper with support by F.M.S., G.D., M. Sporbert and H.B. E.A.-D., I.A., M.B., E.B., I.B., A.D.B., G.B., Z.B.-D., J.A.C., A.Č., M.C., R.Ć., A.L.d.G, M.D.S., J. Dengler, J. Dolezal, M.A.E.-S., M.F., A.G.-d.-M., E.G., H.G., V.G., S.H., M.Z.H., B.H., J.H., U.J., F.J., A.J., J.K., M.K., L.K., H.K., F.K., J.L., J.E.M., L.M., A.N., J.N., A.P.-H., O.L.P., V.D.P., G.R.-T., E.R., B. Sandel, M. Schmidt, U.S., S.S., F.S., U.Š., B. Sparrow, M. Sporbert, Z.S., B. Strohbach, J.-C.S., C.Q.T., Z.T., A.C.V., C.V., D. Waller, D. Wana, H.-F.W., T.W. and G.Z. provided parts of the data. All co-authors edited the paper and provided suggestions on how to improve the analyses.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-024-02589-0.

Correspondence and requests for materials should be addressed to Georg J. A. Hähn.

Peer review information *Nature Ecology & Evolution* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Georg J. A. Hähn[1,2,3] ✉, Gabriella Damasceno[1,2], Esteban Alvarez-Davila[4], Isabelle Aubin[5], Marijn Bauters[6], Erwin Bergmeier[7], Idoia Biurrun[8], Anne D. Bjorkman[9,10], Gianmaria Bonari[11], Zoltán Botta-Dukát[12], Juan A. Campos[8], Andraž Čarni[13,14], Milan Chytrý[15], Renata Ćušterevska[16], André Luís de Gasper[17], Michele De Sanctis[18], Jürgen Dengler[19], Jiri Dolezal[20], Mohamed A. El-Sheikh[21], Manfred Finckh[22], Antonio Galán-de-Mera[23], Emmanuel Garbolino[24], Hamid Gholizadeh[11], Valentin Golub[25], Sylvia Haider[26], Mohamed Z. Hatim[27], Bruno Hérault[28,29], Jürgen Homeier[30], Ute Jandt[1,2], Florian Jansen[31], Anke Jentsch[32], Jens Kattge[2,33], Michael Kessler[34], Larisa Khanina[35], Holger Kreft[36], Filip Küzmič[13], Jonathan Lenoir[37], Jesper Erenskjold Moeslund[38], Ladislav Mucina[39,40], Alireza Naqinezhad[41], Jalil Noroozi[42], Aaron Pérez-Haase[43], Oliver L. Phillips[44], Valério D. Pillar[45], Gonzalo Rivas-Torres[46], Eszter Ruprecht[47], Brody Sandel[48], Marco Schmidt[49], Ute Schmiedel[22], Stefan Schnitzer[50], Franziska Schrodt[51], Urban Šilc[13], Ben Sparrow[52], Maria Sporbert[1,2], Zvjezdana Stančić[53], Ben Strohbach[54], Jens-Christian Svenning[55], Cindy Q. Tang[56], Zhiyao Tang[57], Alexander Christian Vibrans[58], Cyrille Violle[59], Donald Waller[60], Desalegn Wana[61], Hua-Feng Wang[62], Timothy Whitfeld[63], Georg Zizka[64], Francesco Maria Sabatini[3,65,66] & Helge Bruelheide[1,2,66]

[1]Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Halle, Germany. [2]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. [3]Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy. [4]Escuela ECAPMA, Universidad Nacional Abierta y a Distancia (Colombia), Bogotá, Colombia. [5]Great Lakes Forestry Centre, Canadian Forest Service, Sault Ste Marie, Ontario, Canada. [6]Department of Environment, Ghent University, Gent, Belgium. [7]Department of Vegetation and Phytodiversity Analysis, University of Göttingen, Göttingen, Germany. [8]Department of Plant Biology and Ecology, University of the Basque Country UPV/EHU, Bilbao, Spain. [9]Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden. [10]Gothenburg Global Biodiversity Centre, Gothenburg, Sweden. [11]Biological and Environmental Sciences, University of Siena, Siena, Italy. [12]Institute of Ecology and Botany, Centre for Ecological Research, Vácrátót, Hungary. [13]Jovan Hadži Institute of Biology, Research Centre of the Slovenian Academy of Sciences and Arts, Ljubljana, Slovenia. [14]School for Viticulture and Enology, University of Nova Gorica, Nova Gorica, Slovenia. [15]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic. [16]Institute of Biology, Faculty of Natural Sciences and Mathematics, University of Ss. Cyril and Methodius, Skopje, North Macedonia. [17]Universidade Regional de Blumenau, Blumenau, Brazil. [18]Department of Environmental Biology, Sapienza University of Rome, Rome, Italy. [19]Zurich University of Applied Sciences (ZHAW), Wädenswil, Switzerland. [20]Institute of Botany, Czech Academy of Science, Trebon, Czechia. [21]Botany and Microbiology Department, College of Science, King Saud University, Riyadh, Saudi Arabia. [22]Institute of Plant Science and Microbiology, University of Hamburg, Hamburg, Germany. [23]Botany Lab, Universidad San Pablo-CEU, CEU Universities, Madrid, Spain. [24]MINES Paris PSL—ISIGE, Fontainebleau, France. [25]Institute of Ecology of the Volga River Basin, Samara Federal Research Scientific Center, Togliatti, Russia. [26]Institute of Ecology, School of Sustainability, Leuphana University of Lüneburg, Lüneburg, Germany. [27]Plant Ecology and Nature Conservation Group, Environmental Sciences Department, Wageningen University, Wageningen, the Netherlands. [28]CIRAD, UPR Forêts et Sociétés, Campus de Baillarguet, Montpellier, France. [29]University Montpellier, Montpellier, France. [30]Resource Management, HAWK Goettingen, Goettingen, Germany. [31]University of Rostock, Rostock, Germany. [32]Bayreuth Center of Ecology and Environmental Research, Department of Disturbance Ecology, University of Bayreuth, Bayreuth, Germany. [33]Max Planck Institute for Biogeochemistry, Jena, Germany. [34]Systematic and Evolutionary Botany, University of Zurich, Zurich, Switzerland. [35]Branch of the M.V. Keldysh IAM RAS, IMPB RAS, Pushchino, Russia. [36]Department of Biodiversity, Macroecology and Biogeography, University of Göttingen, Göttingen, Germany. [37]UMR CNRS 7058 Ecologie et Dynamique des Systèmes Anthropisés (EDYSAN), Université de Picardie Jules Verne, Amiens, France. [38]Department of Ecoscience, Aarhus University, Aarhus C, Denmark. [39]Harry Butler Institute, Perth, Western Australia, Australia. [40]Department of Geography and Environmental Studies, Stellenbosch University, Matieland, South Africa. [41]Department of Environmental Sciences, College of Science and Engineering, University of Derby, Derby, UK. [42]Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria. [43]Institut de Recerca de la Biodiversitat (IRBio), Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Universitat de Barcelona, Barcelona, Spain. [44]University of Leeds, Leeds, UK. [45]Department of Ecology, Universidade Federal do Rio Grande do Sul, Porto Alegro, Brazil. [46]Colegio de Ciencias Biológicas y Ambientales, Estación de Biodiversidad Tiputini, Universidad San Francisco de Quito USFQ, Quito, Ecuador. [47]Hungarian Department of Biology and Ecology, Faculty of Biology and Geology, Babeș-Bolyai University, Cluj-Napoca, Romania. [48]Department of Biology, Santa Clara University, Santa Clara, CA, USA. [49]Palmengarten der Stadt, Frankfurt, Germany. [50]Marquette University, Milwaukee, WI, USA. [51]University of Nottingham, Nottingham, UK. [52]The School of Biological Sciences, University of Adelaide, Glen Osmond, South Australia, Australia. [53]Faculty of Geotechnical Engineering, University of Zagreb, Varaždin, Croatia. [54]Biodiversity Research Center, Faculty of Health, Natural Resources and Applied Sciences, Namibia University of Science and Technology, Windhoek, Namibia. [55]Center for Ecological Dynamics in a Novel Biosphere (ECONOVO), Department of Biology, Aarhus University, Aarhus C, Denmark. [56]College of Ecology and Environmental Science, Institute of Ecology and Geobotany, Yunnan University, University Town, China. [57]College of Urban and Environmental Sciences, Department of Ecology,

Peking University, Beijing, China. [58]Universidade Regional de Blumenau (FURB), Blumenau, Brazil. [59]CEFE, CNRS, EPHE, IRD, University Montpellier, Montpellier, France. [60]Department of Botany, University of Wisconsin-Madison, Madison, WI, USA. [61]Department of Geography and Environmental Studies, Addis Ababa University, Addis Ababa, Ethiopia. [62]Sanya Nanfan Research Institute, Hainan University, Sanya, China. [63]Bell Musuem, University of Minnesota, St. Paul, MN, USA. [64]Senckenberg Research Institute and Natural History Museum Frankfurt and Department Botany and Molecular Evolution, Goethe University, Frankfurt, Germany. [65]Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Prague, Czech Republic. [66]These authors jointly supervised this work: Francesco Maria Sabatini, Helge Bruelheide. ✉e-mail: georg.haehn@idiv.de

nature portfolio

Corresponding author(s): Georg J. A. Hähn

Last updated by author(s): Oct 3, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	All analyses were performed with R version 4.1.3. All scripts are available at https://github.com/georghaehn/Haehn-et-al-2024-FD-PD-coupling.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The vegetation-plot raw data contained in the sPlot database are available upon request by submitting a project proposal to sPlot's Steering Committee. The proposals should follow the Governance and Data Property Rules of the sPlot Working Group available on the sPlot website (www.idiv.de/splot). The data including the coordinates and all calculated diversity indices are available at the iData portal.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status). Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.) Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Analyses of the relationship between functional and phylogentic diversity within plant communities.
Research sample	Global distributed plant communities.
Sampling strategy	The sPlot database is a collection of 116 datasets from all over the World, each being collected according to a proper sampling scheme.
Data collection	Each vegetation plot in each of the 116 datasets was surveyed in the field by trained botanists, who recorded the complete list of vascular plants occurring in the vegetation plot
Timing and spatial scale	Data in sPlot was sampled between 1873 and 2019.
Data exclusions	Data was excluded if the number of species in the community with known traits and phylogeny was below two.
Reproducibility	All Code to reproduce the analyses is available at https://github.com/georghaehn/Haehn-et-al-2024-FD-PD-coupling.
Randomization	The selection of vegetation data within sPlot was not random as this would not represent the non-random global distribution of species and plant communities across environmental conditions.
Blinding	Being a correlative analysis, no blinding was needed when performing the statistical analyses.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A