Plant taxonomic and phylogenetic turnover increases toward climatic extremes and depends on historical factors in European beech forests

Josep Padullés Cubino, Borja Jiménez-alfaro, Francesco Maria Sabatini, Wolfgang Willner, Zdeňka Lososová, Idoia Biurrun, Jörg Brunet, Juan Antonio Campos, Adrian Indreica, Florian Jansen, et al.

▶ **To cite this version:**

Josep Padullés Cubino, Borja Jiménez-alfaro, Francesco Maria Sabatini, Wolfgang Willner, Zdeňka Lososová, et al.. Plant taxonomic and phylogenetic turnover increases toward climatic extremes and depends on historical factors in European beech forests. Journal of Vegetation Science, 2021, 32 (1), pp.e12977. 10.1111/jvs.12977 . hal-03362995

HAL Id: hal-03362995
https://hal.science/hal-03362995

Submitted on 2 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Plant taxonomic and phylogenetic turnovers increase towards climatic extremes and depend on historical factors in European beech forests

Running title: Spatial turnover in European beech forests

Josep Padullés Cubino[1*], Borja Jiménez-Alfaro[2], Francesco Maria Sabatini[3,4], Wolfgang Willner[5], Zdeňka Lososová[1], Idoia Biurrun[6], Jörg Brunet[7], Juan Antonio Campos[6], Adrian Indreica[8], Florian Jansen[9], Jonathan Lenoir[10], Željko Škvorc[11], Kiril Vassilev[12], Milan Chytrý[1]

[1]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic

[2]Research Unit of Biodiversity (CSIC-UO-PA), University of Oviedo, Mieres, Principado de Asturias, Spain

[3]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

[4]Martin Luther University Halle-Wittenberg, Institute of Biology/Geobotany and Botanical Garden, Halle (Saale), Germany

[5]Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria

[6]Department of Plant Biology and Ecology, University of the Basque Country UPV/EHU, Bilbao, Spain

[7]Southern Swedish Forest Research Centre, Swedish University of Agricultural Sciences, Alnarp, Sweden

[8]Department of Silviculture, Transilvania University of Brasov, Brasov, Romania

[9]Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany

[10] UR « Ecologie et Dynamique des Systèmes Anthropisés » (EDYSAN, UMR 7058 CNRS), Université de Picardie Jules Verne, Amiens, France

[11]Faculty of Forestry, University of Zagreb, Zagreb, Croatia

[12]Department of Plant and Fungal Diversity and Resources, Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, Sofia, Bulgaria

*Corresponding author: Josep Padullés Cubino, Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, CZ-611 37 Brno, Czech Republic; padullesj@gmail.com

ORCID numbers

Josep Padullés Cubino: 0000-0002-2283-5004

Borja Jiménez-Alfaro: 0000-0001-6601-9597

Francesco Maria Sabatini: 0000-0002-7202-7697

Wolfgang Willner: 0000-0003-1591-8386

Zdeňka Lososová: 0000-0001-9152-7462

Idoia Biurrun: 0000-0002-1454-0433

Jörg Brunet: 0000-0003-2667-4575

Juan Antonio Campos: 0000-0001-5992-2753

Florian Jansen: 0000-0002-0331-5185

Jonathan Lenoir: 0000-0003-0638-9582

Željko Škvorc: 0000-0002-2848-1454

Kiril Vassilev: 0000-0003-4376-5575

Milan Chytrý: 0000-0002-8122-3075

Funding information

The study was supported by the Czech Science Foundation (project no.19-28491X). I.B. and J.A.C. were funded by the Basque Government (IT936-16).

Abstract

Aims: The effect of biogeographical processes on the spatial turnover component of beta-diversity over large spatial extents remains scarcely understood. Here, we aim at disentangling the roles of environmental and historical factors on taxonomic and phylogenetic turnover, while controlling for the effects of species richness and rarity.

Location: European beech (*Fagus sylvatica*) forests in Europe.

Methods: We aggregated plant species occurrences from vegetation plots in spatial grid cells of 0.25º × 0.25º to calculate the spatial turnover component of taxonomic (TBD_{turn}) and phylogenetic (PBD_{turn}) beta-diversity for each cell. We also calculated the deviation of PBD_{turn} given TBD_{turn} ($PBD_{dev\text{-}turn}$), which measures the importance of phylogenetic turnover after factoring out taxonomic turnover. Beta-diversity was calculated for each grid cell as the mean pairwise dissimilarity between the focal cell and all other cells. We used structural equation modeling (SEM) to examine the relationships between environmental (climate, soil pH, and distance from the geographical distribution limit of beech) and historical (distance from beech glacial refugia) predictors and beta-diversity metrics.

Results: We found a geographically consistent variation in taxonomic and phylogenetic turnover. Overall, TBD_{turn} and PBD_{turn} increased significantly towards more extreme climatic conditions, on more acidic soils, and towards the margins of beech distribution. The effects of environmental variables and the distance from glacial refugia on beta-diversity metrics were mediated by species richness and rarity. $PBD_{dev\text{-}turn}$ was higher in areas closer to glacial refugia.

Conclusions: Continental-scale patterns of beta-diversity in European beech forests are the result of complementary ecological and evolutionary processes. In general, beech forests are taxonomically and phylogenetically more distinct in climatically marginal areas of their European range. However, the spatial variation of beta-diversity in European beech forest flora is still strongly characterized by the distribution of groups of closely related species that evolved or survived in glacial refugia.

Keywords: beta-diversity; climatic gradient; community phylogenetics; deciduous forests; distribution range margin; European Vegetation Archive (EVA); *Fagus sylvatica*; plant diversity; post-glacial dispersal; species rarity; vegetation plots.

Introduction

A key issue in macroecology is understanding the role of biogeographical processes in determining spatial variation in biodiversity (Ricklefs, 2004). Classical biogeographical studies have mainly asked why some sites have more species than others, but much less attention has been paid to explaining spatial variation in the composition of communities, i.e., beta-diversity (Qian, 2009; Kraft *et al*., 2011; Myers *et al*., 2013). Exploring how beta-diversity varies among regions with contrasting ecological and evolutionary histories is fundamental for understanding the processes that structure natural communities (Leibold *et al*., 2004) and for guiding conservation efforts (Socolar *et al*., 2016).

Environmental, spatial, and historical factors can all drive patterns in beta-diversity. First, environmental filtering and species' ecological requirements (niche) determine species' presence in a given habitat (MacArthur, 1972; Tilman, 1988). Sites with more dissimilar environments should host more different sets of species and thus exhibit greater beta-diversity (Whittaker, 1975). Second, spatial processes can also influence species composition through stochastic events, ecological drift, and species' distinct dispersal abilities (Hubbell, 2001; Nekola and White, 1999; Leibold *et al*., 2004; Soininen *et al*., 2007). Third, historical processes, including speciation, extinction, range expansion from areas of origin, or recolonization from refugia, are also important drivers of species distributions (Svenning *et al*., 2015). Yet, it remains poorly understood how these three sets of drivers combine to shape beta-diversity across large spatial scales, and in specific habitat types.

The relationships between environmental, spatial, and historical processes and beta-diversity have often been examined using taxonomic information only (e.g., Qian, 2009; Keil *et al*., 2012; Sabatini *et al*., 2018). However, species are not evolutionarily independent, and historical processes coupled with species diversification are expected to leave significant

imprints on communities' phylogenetic structure (Emerson and Gillespie, 2008). Combining taxonomic and phylogenetic measures of beta-diversity can help understand the effect of ecological and evolutionary processes on biodiversity patterns (Graham and Fine, 2008; Swenson, 2011a). For example, in areas with a large fraction of narrow-ranging species with recent divergence time (i.e., neo-endemics), taxonomic beta-diversity should be high but phylogenetic beta-diversity low (Swenson, 2011a). In contrast, if the same narrow-ranging species have relatively old divergence time (i.e., belong to lineages with long-standing and disparate evolutionary histories), both taxonomic and phylogenetic beta-diversity should be high (Graham and Fine, 2008). Yet, taxonomic and phylogenetic beta-diversities have seldom been employed over large spatial extents to disentangle the relative importance of ecological and historical processes (Swenson, 2011b; Hardy *et al*., 2012; Kubota *et al*., 2014).

Several features make European beech forests a good model system to test the influence of environmental and historical factors on the patterns of taxonomic and phylogenetic beta-diversity. European beech (*Fagus sylvatica*) is one of the most abundant deciduous tree species in Europe, which typically forms monospecific, dense canopies, with a relatively homogeneous understory. It extends from the mountains of southern Europe (northern Spain, Italy, and Greece) to the lowlands in south-eastern England and southern Sweden (Bolte *et al*., 2007; Caudullo *et al*., 2017). Its distribution is mainly constrained by dry summers and frosty winters (Houston Durrant *et al*., 2016). Nonetheless, many of its understory species have different ecological and geographical requirements than beech itself and may be restricted to specific regions and ecological conditions (Willner *et al*., 2009). Furthermore, the geographical distribution of beech forest species has repeatedly been shown to be affected by dispersal limitations partially reflecting the history of post-glacial recolonization from refugia (Magri, 2008; Lehsten *et al*., 2014; Willner *et al*., 2009). Several studies have examined beech forest diversity (Willner *et al*., 2017) and its environmental and

historical drivers at the regional (Hrivnák *et al*., 2014; Ujházyová *et al*., 2016; Weigel *et al*., 2019) and continental scale (Jiménez-Alfaro *et al*., 2018; Willner *et al*., 2004, 2009). However, no studies have simultaneously tested for the effect of these drivers on both taxonomic and phylogenetic turnover (i.e., the beta-diversity component that reflects the change in species identities and phylogenetic relationships between two or more assemblages; Baselga, 2010) across the whole range of beech forests.

Here, we used plant occurrences from vegetation-plot records and aggregated them within grid cells of 0.25º × 0.25º spatial resolution across the range of European beech forests. Then we calculated the spatial turnover component of taxonomic and phylogenetic beta-diversity between a given cell and all the other cells. Specifically, we tested two hypotheses. First, (H1) taxonomic and phylogenetic spatial turnover increases in the cells with more distinct environmental conditions or closer to the margin of the geographical distribution of European beech (MacArthur, 1972; Whittaker, 1975). We expected this because beech forests closer to the edges of their environmental and geographical distribution can suffer from stronger climatic and environmental stress, and host unique sets of species, some of them potentially deriving from other habitat types. Second, (H2) we expected phylogenetic turnover to be low in relation to taxonomic turnover when moving closer to the putative glacial refugial areas of beech. These areas, mostly located on three southern European peninsulas (Iberian, Italian and Balkan) (Magri *et al*., 2006), can host a higher fraction of narrow-ranging and rare species within the understory of beech forests. In some cases, such species are closely related because they evolved as a consequence of geographical isolation during the cold stage (i.e., by vicariance) and did not disperse into newly available environments after glaciation (Nekola, 1999; Turner, 2004; Svenning and Skov, 2007).

Methods

Vegetation data

We compiled georeferenced vegetation-plot records from the European Vegetation Archive (EVA; Chytrý *et al*., 2016; accessed on June 28th, 2019). This database contains more than 1.5 million vegetation plots sampled across Europe (see Appendix S1 for an overview of contributing databases). European beech (*Fagus sylvatica*) forests are amongst the best-sampled vegetation types in EVA (~ 35,000 vegetation plots), providing a unique model to address macroecological questions (Jiménez-Alfaro *et al*., 2018). To select vegetation plots dominated by European beech forest from the EVA database, we used a classification expert system of European EUNIS habitats (Chytrý *et al*., 2020) run in the JUICE program (Tichý, 2002). This expert system contains definitions of individual EUNIS habitats based on their species composition and geographic location (Chytrý *et al*., 2020). We focused only on European beech forests to reduce the potential confounding effects of different evolutionary histories and species pools across habitat types.

We selected plots sampled after 1970 with areas ranging between 100 and 1,000 m^2, which were the most common plot sizes in the database, and avoided plots in which the location uncertainty of coordinates was larger than 10 km. However, in specific regions with lower sampling effort, we retained plots with unknown location uncertainty. For each plot, we obtained data on species occurrences and selected only angiosperms, thus excluding gymnosperms, pteridophytes, lycophytes, and non-vascular plants (~4% of total species). We focused on angiosperms to avoid inflated results due to a few taxa connected to deep phylogenetic nodes (Massante and Gerhold, 2020; Qian *et al*., 2020; see Appendix S2 for further justification and results including all vascular plants). We standardized the species

names in our dataset according to The Plant List using the R package 'Taxonstand' (Cayuela *et al*., 2017).

We aggregated vegetation plots in a raster of grid cells of 0.25º × 0.25º spatial resolution and derived the composition of the species pool of each cell (N = 2,035; total number of species = 2,213). The species pool of each cell was obtained by aggregating all plant occurrences, independent of the vegetation layer (understory and upper canopy layers), from the final set of vegetation plots classified as beech habitat and contained within the focal cell. Grid cells with less than five vegetation plots (1,049) were removed from subsequent analyses to guarantee a reliable identification of the pool of species found in the cell (N = 986; total number of species = 2,111). To avoid effects of a disproportionally high density of sample plots in particular grid cells, we conducted a heterogeneity-constrained random (HCR) resampling (Lengyel *et al*., 2011) with the R package 'vegclust' (De Cáceres *et al*., 2010). We did this separately for each grid cell and selected the same number (N = 5) of plots per grid cell. The HCR resampling strategy maximizes the mean and minimizes the variance of the compositional dissimilarity between pairs of plots. Consequently, it tends to select plots that are representative of compositional diversity of vegetation within each grid cell. We calculated dissimilarity between vegetation plots co-occurring in the same grid cell with the Simpson dissimilarity index (Baselga, 2010), computing 1,000 iterations within each grid cell. We also repeated the HCR resampling with 10 and 15 plots per grid cell, but these variants all yielded very similar results to those based on the threshold of five plots per grid cell (Appendix S3).

The total number of vegetation plots included in our study after the HCR resampling was 4,930. They were distributed across 986 grid cells and hosted a total of 1,739 angiosperm species.

Phylogenetic data

We generated a phylogeny by linking the species in our dataset (n = 1,739) with those in the mega-phylogeny implemented in the R package 'V.PhyloMaker' (Jin and Qian, 2019), which includes 74,533 species derived from two mega-trees (Zanne *et al*., 2014; Smith and Brown, 2018). We used the 'Scenario 3' approach implemented in the same R package to add missing species (~ 27%) to the phylogeny. In this scenario, the tip for a new genus is bound to 1/2 of the family branch, and the new tip of an existing genus is bound to the basal node of a given genus (for further details see Qian and Jin, 2016; Jin and Qian, 2019).

Environmental and historical data

We obtained current climatic data from the full set of 19 bioclimatic variables taken from the WorldClim v.2.1 database (Fick and Hijmans, 2017) at a resolution of 30 seconds (~1 km^2) (Appendix S4). We extracted climatic data for each cell with the R package 'raster' (Hijmans, 2019) and calculated the mean value of all raster pixels contained in the cell. We then used principal component analysis (PCA) to reduce the dimensionality and collinearity of the 19 bioclimatic variables. We standardized all variables to a mean of 0 and a standard deviation of 1 before implementing the PCA. We interpreted the first PCA component (37% of variation explained) as a gradient of aridity, as it distinguished between cells with low temperatures and high precipitation (mainly mountain areas of Central Europe) and cells with high temperatures and low precipitation (mainly Southern Europe; Appendix S4: Table S4.1 and Fig. S4.1). We only used this PCA axis in subsequent analyses for two reasons: (1) aridity has long been recognized as a major factor determining the distribution range of beech forests in Europe (Houston Durrant *et al*., 2016), and (2) it showed the strongest relationship

with our set of response variables (Appendix S4: Fig. S4.2). We thus refer to the first PCA component as an 'aridity' factor throughout the manuscript.

We additionally downloaded data for soil pH (at 15 cm depth) at a spatial resolution of ~250 m from the SoilGrids database (https://soilgrids.org/; Hengl *et al*., 2017) and calculated the mean value of all raster pixels contained in each cell. We also obtained a shapefile with the geographical distribution of European beech (*Fagus sylvatica*) from Caudullo *et al*. (2017), and used it to calculate the distance (i.e., great-circle distance) between each grid cell and the closest margin of the geographical distribution of European beech (Appendix S5: Fig. S5.1). We calculated distances with the R package 'geosphere' (Hijmans, 2019). We assigned zero values to cells located outside the distribution range (~ 9%). Grid cells located closer to the margin of the geographical distribution of European beech were associated with more extreme climatic conditions (Appendix S6: Fig. S6.2), supporting our assumption that geographically marginal areas can suffer from stronger climatic and environmental stress and present greater species turnover. For this reason, we refer to the distance from the distribution margin of beech as an 'environmental variable' througouth the study.

Finally, we also obtained tentative locations of the putative glacial refugia of beech from Magri *et al*. (2006) (see Fig. 9 therein and Appendix S5: Fig. S5.1) and calculated the minimum distance between each cell and these locations to get a measure of distance from refugial areas. Correlations among variables are shown in Appendix S6.

Species richness (SR) and Species rarity index (SRar)

Taxonomic and phylogenetic beta-diversity might be affected by both the number of species in a cell and their degree of uniqueness or rarity across grid cells. Therefore, we calculated

both the total species richness (SR), and a 'species rarity index' (SRar) for each focal grid cell, to explicitly control for these effects. SRar was calculated as follows:

$$SRar = 1 - median\left(\frac{x_i}{N-1}, \ldots, \frac{x_j}{N-1}\right) \quad \text{(Eq. 1)}$$

where *x* is the number of cells where species *i* to *j* that are present in the focal cell occur outside the focal cell, and *N* is the total number of cells. We subtracted median values from 1 to rank cells with rarer species high and cells with more common species low. Therefore, SRar is equal to 1 when at least half of the species are only present in the focal cell (i.e., a high degree of rarity), while it is equal to 0 when at least half of the species present in the focal cell also occur in all other cells (i.e., a low degree of rarity). We decided to use the median instead of the mean to reduce the potential effect of outliers, although both measures were highly correlated (Pearson's r = 0.97). The resulting metric was positively related to species richness (R^2 = 0.27; Appendix S7: Fig. S7.1A). We provide an R code to calculate SRar in Appendix S7.

Beta-diversity

We calculated the turnover component of spatial taxonomic (TBD_{turn}) and phylogenetic (PBD_{turn}) beta-diversity of each grid cell following Simpson's index of dissimilarity (Baselga, 2010; Leprieur *et al*., 2012):

$$TBD_{turn} = \frac{min(b,c)}{a+min(b,c)} \quad \text{(Eq. 2)}$$

where *a* is the number of species shared by both grid cells, *b* is the number of species exclusive to the focal grid cell, and *c* is the number of species exclusive to the compared grid cell. The equation for PBD_{turn} is equivalent to that for TBD_{turn}, although shared/exclusive 'species' are replaced by the length of shared/exclusive 'branches' in the phylogenetic tree (Leprieur *et al*., 2012). The TBD_{turn} and PBD_{turn} metrics reflect true substitution of species or lineages between cells, respectively.

Both the TBD_{turn} and PBD_{turn} metrics have values equal to 0 when two compared grid cells are identical in terms of species and branches, respectively, and values equal to 1 when they are entirely distinct. For each grid cell, turnover was calculated as the mean pairwise dissimilarity between a given cell and all other cells (Whittaker, 1972). This metric is equivalent to the 'local contribution to beta-diversity' proposed by Legendre and De Cáceres (2013). We preferred this approach over a moving window approach, which calculates beta-diversity among a group of adjacent cells (e.g., Peixoto *et al*., 2017; Qian *et al*., 2020), because we were interested in comparing areas based on their absolute levels of turnover rather than identifying areas where the local rate of change in turnover was maximum. This corresponds to the difference between the second and the third level of abstraction in beta-diversity analyses according to Tuomisto and Ruokolainen (2006).

We also calculated the deviations of PBD_{turn} given TBD_{turn} (i.e., $PBD_{dev\text{-}turn}$) to identify cells where the phylogenetic turnover was higher or lower than expected given taxonomic turnover (Graham and Fine, 2008; Peixoto *et al*., 2017):

$$PBD_{dev-turn} = TBD_{turn} - (PBD_{turn}/TBD_{turn}) \quad \text{(Eq. 3)}$$

High positive values indicate that phylogenetic turnover is low in relation to taxonomic turnover (i.e., limited lineage exchange), while high negative values indicate that

phylogenetic turnover is high in relation to taxonomic turnover (i.e., high lineage exchange). We calculated all beta-diversity measures using the R package 'betapart' (Baselga and Orme, 2012).

Statistical analyses

We used structural equation modeling (SEM) to evaluate the effects of environmental and historical variables on TBD_{turn}, PBD_{turn}, and $PBD_{dev\text{-}turn}$. The SEM approach allows causal relationships among variables of interest to be statistically defined and evaluated in the form of mutually interconnected equations (Grace, 2006). In our study system, SEM was deemed as appropriate because of the causal links between beta-diversity and environmental and historical drivers. Specifically, we assumed that these links are mediated by the degree of species uniqueness or rarity (measured here as 'SRar'). Also, we acknowledge that greater environmental heterogeneity, coupled with potential differences in the size of the sampled plots, can influence spatial turnover by affecting regional species richness (SR) (Nekola and White, 1999; Whittaker, 1975). Therefore, we introduced SR as a predictor of beta-diversity to control for the effect of this variable in our models. Because SR and SRar are non-independent components of diversity (Appendix S7: Fig. S7.1A), we included a non-causal correlation between SR and SRar in our SEM models.

Our data were spatially structured. We therefore implemented SEM using simultaneous autoregressive models (SAR; Kissling and Carl, 2007), which account for spatial autocorrelation by adding a spatial weight matrix that specifies the neighborhood of each cell and the relative weight of each neighbor. Spatial weight matrices were defined by successively fitting a SAR model and testing several distances between neighbors, ranging

from 100 to 500 km away from a given grid cell at intervals of 50 km. We ranked the resulting SAR models by their Akaike information criterion (AIC) values and selected the model with the lowest AIC value as our top SAR model. For all measures, the SAR model with the lowest AIC included neighbors located within 150 km from a given cell. Across all beta-diversity measures, the top SAR model also minimized residual spatial autocorrelation (Appendix S8).

We built a SAR model predicting each of the three beta-diversity metrics (i.e., TBD_{turn}, PBD_{turn}, and $PBD_{dev\text{-}turn}$), based on SR, SRar as well as on the selected environmental (i.e., climate, soil pH, geographical distance from the margin of beech distribution) and historical (i.e., distance from glacial refugia) variables. We modelled beta-diversity metrics as a function of the selected environmental and historical variables, rather than as a function of mean geographical and environmental distances because we were interested in exploring how absolute levels of turnover in grid cells varied along biogeographical gradients. We log-transformed the 'distance from the distribution margin', and the 'distance from glacial refugia' to adjust the data distribution closer to normality. We standardized all variables before running the models (Zuur *et al*., 2007). To account for the nonlinearity of our predictions, we included quadratic terms, modelled as $(x - mean(x))^2$ (Maureaud *et al*., 2019) for aridity, soil pH and SR. Relationships were included in the final SEM if they were statistically significant and improved the explained variance of each SAR model. The variables included in the models showed low levels of multicollinearity when measured using variance inflation factors (VIF < 2.5; Legendre and Legendre, 2012).

To reduce complexity and avoid model saturation of our SEM, we evaluated the goodness-of-fit with Fisher's test after removing non-significant pathways between predictors and beta-diversity metrics. We report Nagelkerke pseudo-R-squared (R^2) of final SAR models as a measure of the coefficient of determination. We ran SEM with the R package

'piecewiseSEM' (Lefcheck, 2016) and SAR models with the R package 'spdep' (Bivand *et al*., 2013). We provide an example of the code used for these analyses in Appendix S9. We established significance at $P < 0.01$ and performed all the analyses in R v. 3.5.3 (R Core Team, 2019).

Results

Effects of environmental and historical factors on taxonomic and phylogenetic turnover

We found the highest values of TBD_{turn} around mountain areas of northern Spain, southern France, southern Italy, Greece, and more locally in the Netherlands, Belgium, and western Germany (Fig. 1A). The lowest values of TBD_{turn} were found in the northern parts of central Pyrenees, the Jura mountains between France and Switzerland, and in Slovenia. PBD_{turn} showed more homogeneous patterns than TBD_{turn} across Europe, although it highlighted similar regions (Fig. 1B). Taxonomic (TBD_{turn}) and phylogenetic (PBD_{turn}) turnover were highly positively correlated (Pearson's r = 0.78; $P < 0.001$).

Our final SAR models explained 90% and 63% of the variation in TBD_{turn} and PBD_{turn}, respectively (Fig. 2A and B). The Fisher statistic showed an adequate goodness-of-fit for both TBD_{turn} (Fisher's C = 12.074; df = 10; $P = 0.280$) and PBD_{turn} (Fisher's C = 9.203; df = 12; $P = 0.685$). The TBD_{turn} and PBD_{turn} metrics were best predicted by the species rarity index (SRar) and species richness (SR), which showed a positive linear and a negative quadratic relationship with the two beta-diversity metrics, respectively (Fig. 2A and B; Fig. 3A and B; see also Appendix S10: Fig. S10.2). We also found a negative direct effect of the distance from the distribution margin of beech and the distance from the glacial refugia of beech on TBD_{turn} (Fig. 2B; Fig. 3B; Appendix S11). Soil pH only showed a positive direct

effect on PBD_{turn}, and aridity had no significant direct effect on either TBD_{turn} or PBD_{turn}. When SRar and SR were removed from our models, the explained variation decreased to 46% and 25% for TBD_{turn} and PBD_{turn}, respectively.

The SRar index was negatively related to the distance from the glacial refugia of beech and the distance from the distribution margin of beech but positively related to soil pH (Fig. 2; Appendix S12). The SRar index was lower at intermediate levels of aridity (Fig. 2; Appendix S12). Species richness was negatively related to aridity and the distance from the glacial refugia of beech but positively related to soil pH (Fig. 2; Appendix S12).

Overall, total effects indicated that both TBD_{turn} and PBD_{turn} were lowest at intermediate levels of aridity but highest towards hotter and drier areas (Fig. 3A and B). Furthermore, both TBD_{turn} and PBD_{turn} tended to decrease with an increase in soil pH and at a greater distance from the geographical distribution margin of beech. Only TBD_{turn} showed a tendency to decrease at a greater distance from the glacial refugia of beech.

Effects of environmental and historical factors on $PBD_{dev\text{-}turn}$

We observed a strong latitudinal pattern in $PBD_{dev\text{-}turn}$ across the geographic distribution of beech (Fig. 1C), as phylogenetic turnover in relation to taxonomic turnover increased towards higher and lower latitudes (see also Appendix S10: Fig. S10.1).

Our final SAR model explained 66% of the variation in $PBD_{dev\text{-}turn}$ (Fig. 2C). The Fisher statistic showed an adequate goodness-of-fit (Fisher's C = 7.447; df = 8; P = 0.489). The $PBD_{dev\text{-}turn}$ metric was best predicted by SRar (positively) and SR (negatively) (Fig. 2C; Fig. 3C; see also Appendix S10: Fig. S10.2). We also found a negative direct effect of the distance from the distribution margin of beech, the distance from the glacial refugia of beech, and soil pH on $PBD_{dev\text{-}turn}$ (Fig. 2C; Fig. 3C; Appendix S11). All variables had an indirect

effect on $PBD_{dev\text{-}turn}$ through SRar and SR, although the total effect of aridity was notably lower than for TBD_{turn} and PBD_{turn}. The total effect of the distance from the glacial refugia of beech on $PBD_{dev\text{-}turn}$ was notably higher than for TBD_{turn} and showed the opposite pattern than for PBD_{turn}.

Discussion

Our study provides evidence that environmental and historical processes simultaneously drive large-scale patterns of taxonomic and phylogenetic turnover in European beech forests, and that their effects are mediated mainly by the degree of species uniqueness or rarity in regional plant assemblages. Following our initial hypotheses, we demonstrate that (H1) taxonomic and phylogenetic spatial turnover increases in the cells with more distinct environmental conditions or closer to the margin of the geographical distribution of beech, and that (H2) phylogenetic turnover is low in relation to taxonomic turnover closer to the putative glacial refugia of beech, partially because these areas contain a higher fraction of closely related species.

Environmental drivers of taxonomic and phylogenetic turnover

Taxonomic and phylogenetic turnover increased towards both ends of the aridity gradient. Forests in these conditions are likely to host a relatively high proportion of angiosperm species that can frequently occur in adjacent forest types but only rarely in beech forests. This is in agreement with Hrivnák *et al*. (2014), who showed that the herb-layer of Carpathian beech forests is characterized by two species groups. First, shade-tolerant beech forest specialists are associated with this habitat especially at intermediate altitudes with mild

climatic conditions. Second, species typical of neighboring vegetation types are more frequent at both ends of the altitudinal gradient, where the climatic conditions are more extreme. Our study shows that a similar pattern exists at the continental scale and across the full range of European beech forests.

Overall, the highest taxonomic and phylogenetic dissimilarities occurred in hot and dry areas. This high species and lineage turnover might be explained by that fact that such areas generally contain a smaller fraction of the pool of species commonly found in beech forests than those in colder and wetter conditions. The absence of otherwise common species and the presence of uncommon ones, perhaps as a result of strong human influence, might also help explain the high spatial turnover in the relatively species-poor areas in the Netherlands and western Germany (Appendix S13). Our study also indicates that both taxonomic and phylogenetic turnovers increase at intermediate levels of species richness (Appendix S10: Fig. S10.2), suggesting that extremely species-poor beech forests usually host widespread species and lineages.

Beech forests on base-rich soils hosted, on average, more unique sets of species and had a relatively higher taxonomic and phylogenetic spatial turnover. However, the direct positive effect of soil pH on turnover was compensated by an indirect negative effect through species richness. In particular, beech forests on more acidic soils had a lower number of species (see also Ewald, 2003; Jiménez-Alfaro *et al*., 2018; Pärtel, 2002), which in turn promoted a more intense species and lineage turnover than on soils with a higher pH. The strength of these relationships, however, could have been limited by the scale mismatch between the fine-scale effect of soil factors on plants' distribution and the coarse-scale data we used (see also Bruelheide *et al*., 2018), as well as by the fact that we did not consider variability in soil pH within cells of high geodiversity. Yet, the contrasting effects of soil pH

on beta-diversity highlight the need to consider multiple interconnected causal relationships when examining processes structuring natural communities.

As we expected, spatial taxonomic and phylogenetic turnover increased towards the margin of the geographical distribution of beech, reflecting a change in plant composition with proximity to the edges of the habitat range (Ries *et al*., 2004). Distribution margins frequently represent physical transitions between habitat types and can be characterized by more unusual environmental conditions where edge specialists thrive (Laurance *et al*., 2007; Ries *et al*., 2004). However, distribution margins in different locations (e.g., polar, meridional, or continental) can differ fundamentally with respect to historical and other biogeographical factors (Bolte *et al*., 2007; Weigel *et al*., 2009). Thus, further research is needed to elucidate how beta-diversity varies among geographical distribution margins in different ecoregions, and whether there is a saturation effect of beta-diversity over distance from these margins.

Both taxonomic and phylogenetic turnovers were structured by similar processes, although with different relative importance. In particular, the explanatory power of our models was higher for taxonomic than for phylogenetic turnover even when we excluded species richness and rarity, indicating that the influence of current climatic conditions on beta-diversity weakens as we go deeper in evolutionary history. A better understanding of the influence of climatic oscillations over geological timescales and of other historical drivers on current patterns of phylogenetic beta-diversity would require complementary biodiversity measures that capture variation in species composition not only at a shallow but also at a deep phylogenetic level (Swenson, 2011a).

Even if based on the largest collection of vegetation-plot data available, our work does not come without uncertainties. First, there is a still a certain amount of unexplained variation in taxonomic and phylogenetic turnover, especially if we remove the effect of

species richness and rarity. We suspect this may be due to a combination of local stochastic processes, biotic interactions, distinct dispersal abilities of species and/or unmeasured environmental and spatial variables that ultimately determine the geographical distribution of beech forest species (Hubbell, 2001). Second, although our study covers almost the entire distribution range of European beech forests, further refinement of our conclusions would require additional data in underrepresented areas, such as Germany or western regions of the Balkans (compare Fig. 3 with Appendix S5: Fig. S5.1) and the inclusion of fine-scale ecological factors. Finally, more precise calculations of phylogenetic turnover are needed to challenge our interpretations as molecular data sources provide better and more complete phylogenies.

Historical drivers of taxonomic and phylogenetic turnover

Supporting our initial hypothesis, we found that phylogenetic turnover was less important than taxonomic turnover (i.e., high $PBD_{dev\text{-}turn}$) in the proximity of the glacial refugia of beech and that this effect was largely mediated by species rarity. This finding is consistent with the idea that glacial refugia in mountainous areas host a large fraction of narrow-ranged species (Magri 2008; Willner *et al*., 2009) and that the high replacement of these species is not followed by the replacement of complete phylogenetic lineages.

The observed patterns could likely be explained by the high speciation rate (i.e., vicariance) in mountain refugial areas (Graham and Fine, 2008; Turner, 2004). For example, repeated glacial isolations have led to recent speciation in some genera like *Cardamine* or *Pulmonaria* (Cesca and Peruzzi, 2002; Kirchner, 2004), most likely in isolated beech or other deciduous forest environments, resulting in differentiation of the original phylogenetic group into new varieties of species. Furthermore, in Europe, glaciation promoted relatively high

extinction rates at northern latitudes (Eiserhardt *et al*., 2015; Svenning and Skov, 2007) and allowed the accumulation of more species from ancient lineages in the base-rich mountainous refugial areas of the Iberian, Italian, and Balkan peninsulas (Magri, 2008; Willner *et al*., 2009). Dispersal limitations may have prevented some beech forest specialist species to recolonize areas affected by climate change during interglacial periods (Svenning and Skov, 2007; Willner *et al*., 2009). Altogether, these factors could explain why the putative glacial refugia of beech present a relatively lower variation in phylogenetic turnover in comparison with taxonomic turnover. Nevertheless, we admit that the locations of the glacial refugia of beech used in this study are tentative (Magri, 2008; Magri *et al*., 2006). We may be missing other potential refugia, such as cryptic micro-refugia in central Europe (Birks and Willis, 2008), which may have played a key role in driving current patterns of plant distribution. Further identification of areas with relatively high rates of speciation (potentially by comparing the coupling between phylogenetic and taxonomic turnover across regions) can help refine the localization of glacial refugia of temperate forests in Europe (see also de Lafontaine *et al*., 2014).

Phylogenetic turnover was also less important than taxonomic turnover closer to the glacial refugia of beech after controlling for the effect of species richness and the degree of species rarity in the cells (see also Appendix S11). This suggests that other mechanisms beyond those determining the geographical distribution of beech forest species underly the observed patterns. A plausible explanation is that, once species richness and rarity are held constant, plant assemblages located closer to glacial refugia are phylogenetically more similar to the 'average' assemblage across the range of beech forests. This interpretation is supported by the fact that phylogenetic turnover, unlike taxonomic turnover, decreased with the proximity to glacial refugia. This finding reinforces the idea that glacial refugial areas of

European beech forests harbor a large fraction of the plant lineages that can currently be found across beech forests in Europe (Petit *et al*., 2003).

Conclusions

Previous research on European beech forests has often focused on particular regions or examined the effects of major environmental and historical drivers on alpha- and gamma-diversity. Using a large-scale database of vegetation-plot records, we were able to test the effect of environmental and historical evolutionary processes on beta-diversity across the entire distribution range of European beech. We showed that forests located at both ends of the aridity gradient (particularly in warmer and drier areas) and forests located closer to the margins of the geographical distribution of beech hosted more taxonomically and phylogenetically distinct plant assemblages, possibly due to the presence of species associated with other forest types. Moreover, areas located closer to the putative glacial refugia of beech had a comparatively lower phylogenetic turnover in relation to taxonomic turnover. This pattern suggests that these areas still host groups of closely related species that evolved or survived in glacial refugia. We conclude that, despite the vastly coincident geographic patterns in taxonomic and phylogenetic turnover of European beech forests, these two metrics of turnover are distinctly affected by biogeographical processes. Taken together, these findings reinforce the importance of accounting for complementary ecological and evolutionary processes when examining the drivers of spatial beta-diversity at the continental scale.

Biodiversity conservation planning of particular habitats should consider protecting areas at the opposite ends of environmental gradients, as well as those closest to their geographical distribution margin. These areas host unique sets of species that promote greater

spatial taxonomic and phylogenetic turnover. Particular attention should be paid to the beech forests located in the warmest and driest areas. These forests have the highest beta-diversity, but their position makes them particularly sensitive to climate change. Further research is thus required to assess the susceptibility of these beech forest communities to climate change and the variation in their distribution, structure, and composition over time.

Acknowledgements

We thank Emiliano Agrillo, Henry Brisse, Laura Casella, Panayotis Dimopoulos, Xavier Font, Jean-Claude Gégout, Ute Jandt, Zygmunt Kącki, Remigiusz Pielech, Urban Šilc, Ioannis Tsiripidis, Milan Valachovič, Thomas Wohlgemuth, the Institut national de l'information géographique et forestière (https://inventaire-forestier.ign.fr), and all other database custodians and data contributors for vegetation-plot records; Stephan Hennekens for managing the EVA database; Ilona Knollová for preparing the EVA data for this project, and Irena Axmanová for providing a code to clean up the data.

Author contributions

JPC, BJ-A, FS, WW, and MC conceived the study; IB, JB, JAC, MC, AI, FJ, BJ-A, JL, ZS, KV, and WW collected data; JPC conducted the statistical analyses and wrote the paper; all authors discussed the results and commented on the manuscript.

Data availability statement

The data are available upon request from the European Vegetation Archive (EVA project no. 82).

References

Baselga, A. (2010) Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, 19, 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x

Baselga, A. and Orme, C.D.L. (2012) betapart: an R package for the study of beta diversity. *Methods in Ecology and Evolution*, 3, 808–812. https://doi.org/10.1111/j.2041-210X.2012.00224.x

Birks, H.J.B. and Willis, K.J. (2008) Alpines, trees, and refugia in Europe. *Plant Ecology & Diversity*, 1, 147–160. https://doi.org/10.1080/17550870802349146

Bivand, R.S., Pebesma, E. and Gomez-Rubio, V. (2013) *Applied spatial data analysis with R*, 2nd edition. New York: Springer.

Bolte, A., Czajkowski, T. and Kompa, T. (2007) The north-eastern distribution range of European beech – a review. *Forestry*, 80, 413–429. https://doi.org/10.1093/forestry/cpm028

Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S.M. *et al*. (2018) Global trait–environment relationships of plant communities. *Nature Ecology & Evolution*, 2, 1906–1917. https://doi.org/10.1038/s41559-018-0699-8

Caudullo, G., Welk, E. and San-Miguel-Ayanz, J. (2017) Chorological maps for the main European woody species. *Data in Brief*, 12, 662–666. https://doi.org/10.1016/j.dib.2017.05.007

Cayuela, L., Stein, A. and Oksanen, J. (2017) *Taxonstand: Taxonomic Standardization of Plant Species Names*. R package version 2.1. https://CRAN.R-project.org/package=Taxonstand

Cesca, G. and Peruzzi, L. (2002) A new species of *Cardamine* subgen. *Dentaria* (*Cruciferae*), apoendemic in Calabria (Southern Italy). *Plant Biosystems*, 136, 313–320. https://doi.org/10.1080/11263500212331351209

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F. *et al*. (2016) European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied Vegetation Science*, 19, 173–180. https://doi.org/10.1111/avsc.12191

Chytrý, M., Tichý, L., Hennekens, S.M., Knollová, I., Janssen, J.A.M., Rodwell, J.S. *et al*. (2020) EUNIS Habitat Classification: Expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science*, 23, 648–675. https://doi.org/10.1111/avsc.12519

De Cáceres, M., Font, X. and Oliva, F. (2010) The management of vegetation classifications with fuzzy clustering. *Journal of Vegetation Science*, 21, 1138–1151. https://doi.org/10.1111/j.1654-1103.2010.01211.x

Eiserhardt, W.L., Borchsenius, F., Plum, C.M., Ordonez, A. and Svenning, J.-C. (2015) Climate-driven extinctions shape the phylogenetic structure of temperate tree floras. *Ecology Letters*, 18, 263–272. https://doi.org/10.1111/ele.12409

Emerson, B.C. and Gillespie, R.G. (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution*, 23, 619–630. https://doi.org/10.1016/j.tree.2008.07.005

Ewald, J. (2003) The calcareous riddle: Why are there so many calciphilous species in the Central European flora? *Folia Geobotanica*, 38, 357–366. https://doi.org/10.1007/BF02803244

Fick, S.E. and Hijmans, R.J. (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302–4315. https://doi.org/10.1002/joc.5086

Grace, J.B. (2006) *Structural equation modeling and natural systems*. Cambridge: Cambridge University Press.

Graham, C.H. and Fine, P.V.A. (2008) Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology Letters*, 11, 1265–1277. https://doi.org/10.1111/j.1461-0248.2008.01256.x

Hardy, O.J., Couteron, P., Munoz, F., Ramesh, B.R. and Pélissier, R. (2012) Phylogenetic turnover in tropical tree communities: impact of environmental filtering, biogeography and mesoclimatic niche conservatism. *Global Ecology and Biogeography*, 21, 1007–1016. https://doi.org/10.1111/j.1466-8238.2011.00742.x

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A. *et al*. (2017) SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE*, 12(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Hijmans, R. (2019) *raster: Geographic Data Analysis and Modeling*. https://CRAN.R-project.org/package=geosphere

Houston Durrant, T., de Rigo, D. and Caudullo, G. (2016) *Fagus sylvatica* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T. and Mauri, A. (Eds), *European Atlas of Forest Tree Species*. Luxembourg: Publications Office of the EU, pp. 94-95

Hrivnák, R., Gömöry, D., Slezák, M., Ujházy, K., Hédl, R., Jarčuška, B. and Ujházyová, M. (2014) Species richness pattern along altitudinal gradient in central European beech forests. *Folia Geobotanica*, 49, 425–441. https://doi.org/10.1007/s12224-013-9174-0

Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton, NJ: Princeton University Press.

Jiménez-Alfaro, B., Girardello, M., Chytrý, M., Svenning, J.-C., Willner, W., Gégout, J.-C. *et al*. (2018) History and environment shape species pools and community diversity in European beech forests. *Nature Ecology & Evolution*, 2, 483–490. https://doi.org/10.1038/s41559-017-0462-6

Jin, Y. and Qian, H. (2019) V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography*, 42, 1353–1359. https://doi.org/10.1111/ecog.04434

Keil, P., Schweiger, O., Kühn, I., Kunin, W.E., Kuussaari, M., Settele, J. *et al*. (2012) Patterns of beta diversity in Europe: the role of climate, land cover and distance across scales. *Journal of Biogeography*, 39, 1473–1486. https://doi.org/10.1111/j.1365-2699.2012.02701.x

Kirchner, D.E. (2004) *Molekulare Phylogenie und Biogeographie der Gattung Pulmonaria L. (Boraginaceae)*. Aachen: Verlag Mainz.

Kissling, W.D. and Carl, G. (2007) Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*, 17, 59–71. https://doi.org/10.1111/j.1466-8238.2007.00334.x

Kraft, N.J.B., Comita, L.S., Chase, J.M., Sanders, N.J., Swenson, N.G., Crist, T.O. *et al*. (2011) Disentangling the drivers of diversity along latitudinal and elevational gradients. *Science*, 333, 1755–1758. https://doi.org/10.1126/science.1208584

Kubota, Y., Hirao, T., Fujii, S., Shiono, T. and Kusumoto, B. (2014) Beta diversity of woody plants in the Japanese archipelago: the roles of geohistorical and ecological processes. *Journal of Biogeography*, 41, 1267–1276. https://doi.org/10.1111/jbi.12290

de Lafontaine, G., Amasifuen Guerra, C.A., Ducousso, A., Sánchez-Goñi, M.-F. and Petit, R.J. (2014) Beyond skepticism: uncovering cryptic refugia using multiple lines of evidence. *New Phytologist*, 204, 450–454. https://doi.org/10.1111/nph.13089

Laurance, W.F., Nascimento, H.E.M., Laurance, S.G., Andrade, A., Ewers, R.M., Harms, K.E. *et al*. (2007) Habitat fragmentation, variable edge effects, and the landscape-divergence hypothesis. *PLoS ONE*, 2, e1017. https://doi.org/10.1371/journal.pone.0001017

Lefcheck, J.S. (2016) piecewiseSEM: Piecewise structural equation modeling in R for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, 7(5), 573-579. https://doi.org/10.1111/2041-210X.12512

Legendre, P. and De Cáceres, M. (2013) Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters*, 16, 951–963. https://doi.org/10.1111/ele.12141

Legendre, P. and Legendre, L. (2012) *Numerical Ecology*. 3rd edn. Amsterdam: Elsevier.

Lehsten, D., Dullinger, S., Hülber, K., Schurgers, G., Cheddadi, R., Laborde, H. *et al*. (2014) Modelling the Holocene migrational dynamics of *Fagus sylvatica* L. and *Picea abies* (L.) H. Karst: LPJ-CATS migration model. *Global Ecology and Biogeography*, 23, 658–668. https://doi.org/10.1111/geb.12145

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F. *et al*. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, 7, 601–613. https://doi.org/10.1111/j.1461-0248.2004.00608.x

Lengyel, A., Chytrý, M. and Tichý, L. (2011) Heterogeneity-constrained random resampling of phytosociological databases. *Journal of Vegetation Science*, 22, 175–183. https://doi.org/10.1111/j.1654-1103.2010.01225.x

Leprieur, F., Albouy, C., De Bortoli, J., Cowman, P.F., Bellwood, D.R. and Mouillot, D. (2012) Quantifying phylogenetic beta diversity: Distinguishing between 'true' turnover of lineages and phylogenetic diversity gradients. *PLoS ONE*, 7, e42760. https://doi.org/10.1371/journal.pone.0042760

MacArthur, R.H. (1972) *Geographical ecology*. New York, NY: Harper and Row.

Magri, D. (2008) Patterns of post-glacial spread and the extent of glacial refugia of European beech (*Fagus sylvatica*). *Journal of Biogeography*, 35, 450–463. https://doi.org/10.1111/j.1365-2699.2007.01803.x

Magri, D., Vendramin, G.G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D. *et al*. (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, 171, 199–221. https://doi.org/10.1111/j.1469-8137.2006.01740.x

Massante, J.C. and Gerhold, P. (2020) Environment and evolutionary history depict phylogenetic alpha and beta diversity in the Atlantic coastal white-sand woodlands. *Journal of Vegetation Science*, 31, 634–645. https://doi.org/10.1111/jvs.12900

Maureaud, A., Hodapp, D., van Denderen, P.D., Hillebrand, H., Gislason, H., Spaanheden Dencker, T. *et al*. (2019) Biodiversity–ecosystem functioning relationships in fish communities: biomass is related to evenness and the environment, not to species richness. *Proceedings of the Royal Society B: Biological Sciences*, 286, 20191189. https://doi.org/10.1098/rspb.2019.1189

Myers, J.A., Chase, J.M., Jiménez, I., Jørgensen, P.M., Araujo-Murakami, A., Paniagua-Zambrana, N. and Seidel, R. (2013) Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. *Ecology Letters*, 16, 151–157. https://doi.org/10.1111/ele.12021

Nekola, J.C. (1999) Paleorefugia and neorefugia: the influence of colonization history on community patterns and process. *Ecology*, 80, 2459–2473. https://doi.org/10.1890/0012-9658(1999)080[2459:PANTIO]2.0.CO;2

Nekola, J.C. and White, P.S. (1999) The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, 26, 867–878. https://doi.org/10.1046/j.1365-2699.1999.00305.x

Pärtel, M. (2002) Local plant diversity patterns and evolutionary history at the regional scale. *Ecology*, 83, 2361-2366. https://doi.org/10.1890/0012-9658(2002)083[2361:LPDPAE]2.0.CO;2

Peixoto, F.P., Villalobos, F., Melo, A.S., Diniz-Filho, J.A.F., Loyola, R., Rangel, T.F. and Cianciaruso, M.V. (2017) Geographical patterns of phylogenetic beta-diversity components in terrestrial mammals. *Global Ecology and Biogeography*, 26, 573–583. https://doi.org/10.1111/geb.12561

Petit, R.J., Aguinagalde, I., de Beaulieu, J.L., Bittkau, C., Brewer, S., Cheddadi, R. *et al*. (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, 300, 1563–1565. https://doi.org/10.1126/science.1083264

Qian, H. (2009) Beta diversity in relation to dispersal ability for vascular plants in North America. *Global Ecology and Biogeography*, 18, 327–332. https://doi.org/10.1111/j.1466-8238.2009.00450.x

Qian, H. and Jin, Y. (2016) An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *Journal of Plant Ecology*, 9, 233–239. https://doi.org/10.1093/jpe/rtv047

Qian, H., Jin, Y., Leprieur, F., Wang, X. and Deng, T. (2020) Geographic patterns and environmental correlates of taxonomic and phylogenetic beta diversity for large-scale angiosperm assemblages in China. *Ecography*. https://doi.org/10.1111/ecog.05190

R Core Team. (2019) *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Ricklefs, R.E. (2004) A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, 7, 1–15. https://doi.org/10.1046/j.1461-0248.2003.00554.x

Ries, L., Fletcher, R.J., Battin, J. and Sisk, T.D. (2004) Ecological responses to habitat edges: Mechanisms, models, and variability explained. *Annual Review of Ecology, Evolution, and Systematics*, 35, 491–522. https://doi.org/10.1146/annurev.ecolsys.35.112202.130148

Sabatini, F.M., Jiménez-Alfaro, B., Burrascano, S., Lora, A. and Chytrý, M. (2018) Beta-diversity of central European forests decreases along an elevational gradient due to the variation in local community assembly processes. *Ecography*, 41, 1038–1048. https://doi.org/10.1111/ecog.02809

Smith, S.A. and Brown, J.W. (2018) Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany*, 105, 302–314. https://doi.org/10.1002/ajb2.1019

Socolar, J.B., Gilroy, J.J., Kunin, W.E. and Edwards, D.P. (2016) How should beta-diversity inform biodiversity conservation? *Trends in Ecology & Evolution*, 31, 67–80. https://doi.org/10.1016/j.tree.2015.11.005

Soininen, J., McDonald, R. and Hillebrand, H. (2007) The distance decay of similarity in ecological communities. *Ecography*, 30, 3–12. https://doi.org/10.1111/j.0906-7590.2007.04817.x

Svenning, J.-C. and Skov, F. (2007) Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? *Ecology Letters*, 10, 453–460. https://doi.org/10.1111/j.1461-0248.2007.01038.x

Svenning, J.-C., Eiserhardt, W.L., Normand, S., Ordonez, A. and Sandel, B. (2015) The influence of paleoclimate on present-day patterns in biodiversity and ecosystems.

Annual Review of Ecology, Evolution, and Systematics, 46, 551–572. https://doi.org/10.1146/annurev-ecolsys-112414-054314

Swenson, N.G. (2011a) Phylogenetic beta diversity metrics, trait evolution and inferring the functional beta diversity of communities. *PLoS ONE*, 6, e21264. https://doi.org/10.1371/journal.pone.0021264

Swenson, N.G. (2011b) The role of evolutionary processes in producing biodiversity patterns, and the interrelationships between taxonomic, functional and phylogenetic biodiversity. *American Journal of Botany*, 98, 472–480. https://doi.org/10.3732/ajb.1000289

Tichý, L. (2002) JUICE, software for vegetation classification. *Journal of Vegetation Science*, 13, 451–453. https://doi.org/10.1111/j.1654-1103.2002.tb02069.x

Tilman, D. (1988) *Plant strategies and the dynamics and structure of plant communities*. Princeton, NJ: Princeton University Press.

Tuomisto, H. and Ruokolainen, K. (2006) Analyzing or explaining beta diversity? Understanding the targets of different methods of analysis. *Ecology*, 87, 2697-2708. https://doi.org/10.1890/0012-9658(2006)87[2697:AOEBDU]2.0.CO;2

Turner, J.R.G. (2004) Explaining the global biodiversity gradient: energy, area, history and natural selection. *Basic and Applied Ecology*, 5, 435–448. https://doi.org/10.1016/j.baae.2004.08.004

Ujházyová, M., Ujházy, K., Chytrý, M., Willner, W., Čiliak, M., Máliš, F. and Slezák, M. (2016) Diversity of beech forest vegetation in the Eastern Alps, Bohemian Massif and the Western Carpathians. *Preslia*, 88, 435–457.

Weigel, R., Gilles, J., Klisz, M., Manthey, M. and Kreyling, J. (2019) Forest understory vegetation is more related to soil than to climate towards the cold distribution margin

of European beech. *Journal of Vegetation Science*, 30, 746–755. https://doi.org/10.1111/jvs.12759

Whittaker, R.H. (1972) Evolution and measurement of species diversity. *Taxon*, 21, 213–251. https://doi.org/10.2307/1218190

Whittaker, R.H. (1975) *Communities and ecosystems*. New York, NY: MacMillan Publishing.

Willner, W., Moser, D. and Grabherr, G. (2004) Alpha and beta diversity in central European beech forests. *Fitosociologia*, 41, 15–20.

Willner, W., Di Pietro, R. and Bergmeier, E. (2009) Phytogeographical evidence for post-glacial dispersal limitation of European beech forest species. *Ecography*, 32, 1011–1018. https://doi.org/10.1111/j.1600-0587.2009.05957.x

Willner, W., Jiménez-Alfaro, B., Agrillo, E., Biurrun, I., Campos, J.A., Čarni, A. *et al*. (2017) Classification of European beech forests: a Gordian Knot? *Applied Vegetation Science*, 20, 494–512. https://doi.org/10.1111/avsc.12299

Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A., FitzJohn, R.G. *et al*. (2014) Three keys to the radiation of angiosperms into freezing environments. *Nature*, 506, 89–92. https://doi.org/10.1038/nature12872

Zuur, A.F., Ieno, E.N. and Smith, G.M. (2007) *Analysing Ecological Data*. New York, NY: Springer.

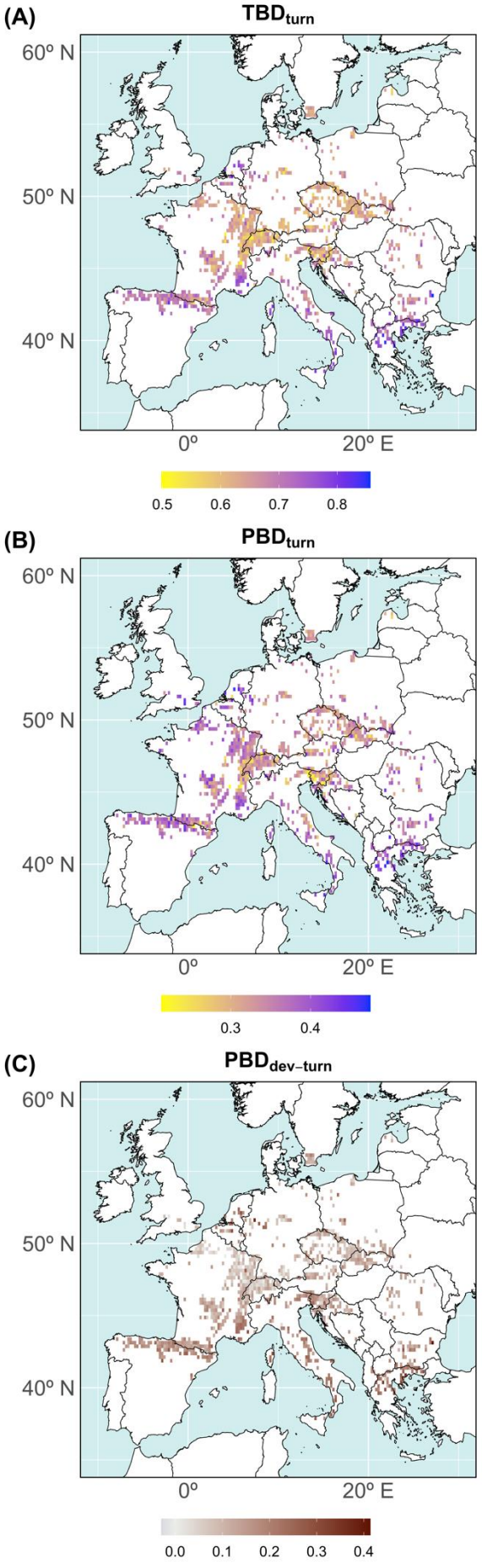
(A)
TBD turn
60° N
50° N
40° N
0°
20° E
0.5
0.6
0.7
0.8
(B)
PBD turn
60° N
50° N
40° N
0°
20° E
0.3
0.4
(C)
PBD dev−turn
60° N
50° N
40° N
0°
20° E
0.0
0.1
0.2
0.3
0.4

Fig. 1: Spatial patterns of plant beta-diversity in European beech forest vegetation. The maps show (A) the spatial turnover component of taxonomic beta-diversity (TBD_{turn}) and (B) phylogenetic beta-diversity (PBD_{turn}), and (C) the deviations of PBD_{turn} given TBD_{turn} ($PBD_{dev\text{-}turn}$).

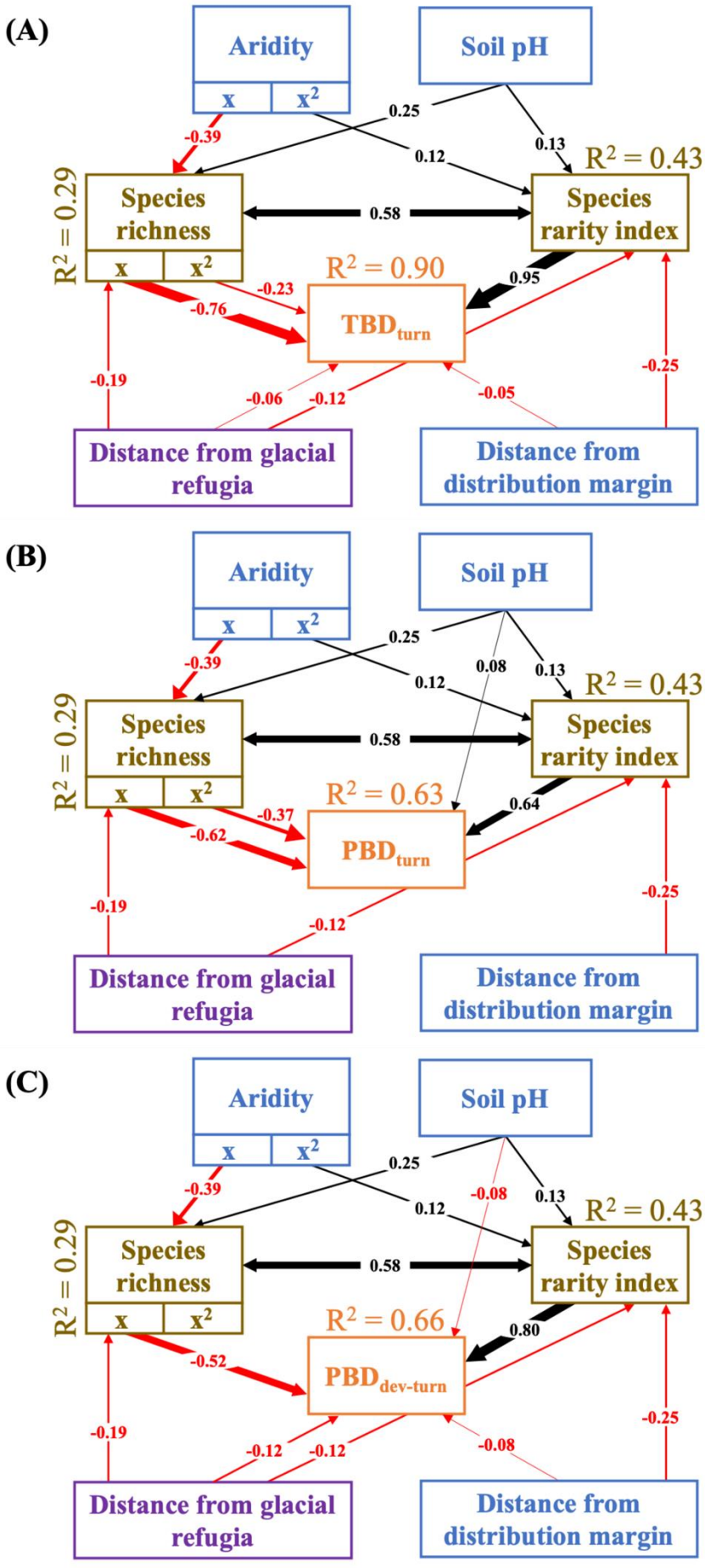
(A)
Aridity
x
x2
Soil pH
0.25
0.12
0.13
-0.39
R2 = 0.43
R2 = 0.29
Species richness
x
x2
0.58
Species rarity index
R2 = 0.90
0.95
-0.23
-0.76
TBDturn
-0.19
-0.06
-0.12
-0.05
-0.25
Distance from glacial refugia
Distance from distribution margin
(B)
Aridity
x
x2
Soil pH
0.25
0.12
0.08
0.13
-0.39
R2 = 0.43
R2 = 0.29
Species richness
x
x2
0.58
Species rarity index
R2 = 0.63
0.64
-0.37
-0.62
PBDturn
-0.19
-0.12
-0.25
Distance from glacial refugia
Distance from distribution margin
(C)
Aridity
x
x2
Soil pH
0.25
-0.08
0.12
0.13
-0.39
R2 = 0.43
R2 = 0.29
Species richness
x
x2
0.58
Species rarity index
R2 = 0.66
0.80
-0.52
PBDdev-turn
-0.19
-0.12
-0.12
-0.08
-0.25
Distance from glacial refugia
Distance from distribution margin

Fig. 2: Final structural equation models (SEM) explaining TBD_{turn} (A), PBD_{turn} (B), and $PBD_{dev\text{-}turn}$ (C). Environmental and historical variables are in blue and purple, respectively, while beta-diversity response variables are shown in orange. Intermediate response variables (species richness and species rarity index) are in brown. Nagelkerke pseudo-R-squared values (R^2) are shown for the dependent variables. 'x^2' indicates if the variable was transformed following: $(x - \text{mean}(x))^2$. Black and red directional arrows represent positive and negative causal links, respectively. Double-headed arrows represent non-causal correlations. Path thickness reflects the strength of the relationship (i.e., values of standardized β-coefficients). Standardized β-coefficients and correlation coefficients are also shown. 'Aridity' refers to the first PCA component obtained from a set of climatic variables and depicts an environmental gradient ranging from low temperatures and high precipitation to hot temperatures and low precipitation (Appendix S4). The variables 'distance from distribution margin' and the 'distance from glacial refugia' of beech were log-transformed.

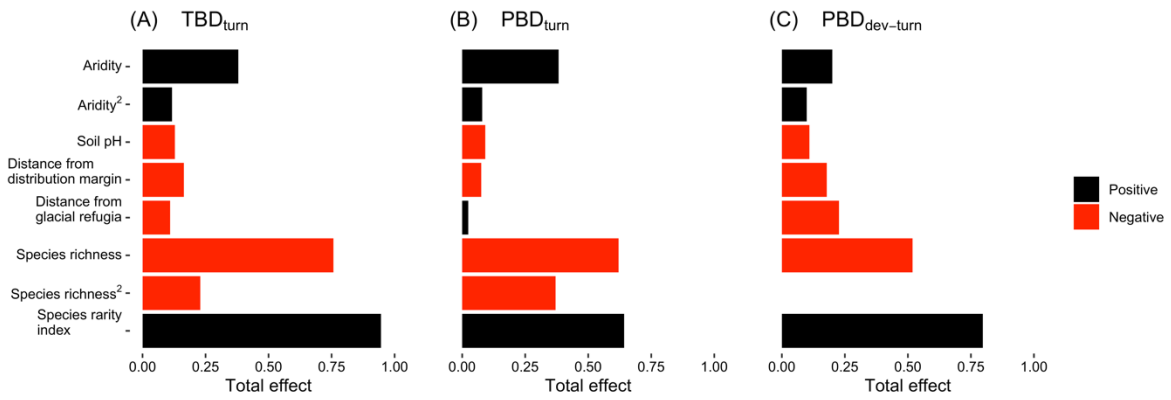

Fig. 3: Total effects of predictor variables on TBD_{turn} (A), PBD_{turn} (B), and $PBD_{dev\text{-}turn}$ (C). Total effects are defined as the sum of the direct and indirect paths between variables (excluding the paths denoted by double-headed arrows in Fig. 2). To allow for the comparison of total effects among variables, we placed both positive and negative effects on the same scale and colored them as black and red, respectively. 'Aridity' refers to the first PCA component obtained from the set of climatic variables and depicts an environmental gradient ranging from low temperatures and high precipitation to hot temperatures and low precipitation (Appendix S4). The variables 'distance from distribution margin' and the 'distance from glacial refugia' of beech were log-transformed.

SUPPORTING INFORMATION

Table of contents:

Appendix S1: Overview of the datasets included in this study.

Appendix S2: Justification for the removal of non-angiosperm species and results considering all vascular plants.

Appendix S3: General results considering a different number of plots per grid cell.

Appendix S4: Results from Principal Component Analysis (PCA) of environmental variables.

Appendix S5: Distribution of beech (*Fagus sylvatica*).

Appendix S6: Correlations between environmental and historical variables.

Appendix S7: Calculation of the 'species rarity index' (SRar) and relationship with species richness and the 'corrected weighted endemism' (CWE).

Appendix S8: Spatial correlograms.

Appendix S9: Example of the code used to implement SEM.

Appendix S10: Relationships between latitude, SR, and the 'species rarity index' (SRar) and TBD_{turn}, PBD_{turn}, and $PBD_{dev\text{-}turn}$ derived from ordinary least square (OLS) regressions.

Appendix S11: Comparison of the effects of predictor variables on response variables before and after controlling for the effect of other predictors.

Appendix S12: Total effects of environmental and historical variables on SR and SRar.

Appendix S13: Plant composition and diversity metrics in areas with consistently high and low taxonomic and phylogenetic turnover. [Spreadsheet]

SUPPORTING INFORMATION

Padullés Cubino et al. Plant taxonomic and phylogenetic turnovers increase towards climatic extremes and depend on historical factors in European beech forests. *Journal of Vegetation Science.*

Table of contents:

Appendix S1: Overview of the datasets included in this study.

Appendix S2: Justification for the removal of non-angiosperm species and results considering all vascular plants.

Appendix S3: General results considering a different number of plots per grid cell.

Appendix S4: Results from Principal Component Analysis (PCA) of environmental variables.

Appendix S5: Distribution of beech (*Fagus sylvatica*).

Appendix S6: Correlations between environmental and historical variables.

Appendix S7: Calculation of the 'species rarity index' (SRar) and relationship with species richness and the 'corrected weighted endemism' (CWE).

Appendix S8: Spatial correlograms.

Appendix S9: Example of the code used to implement SEM.

Appendix S10: Relationships between latitude, SR, and the 'species rarity index' (SRar) and TBD_{turn}, PBD_{turn}, and $PBD_{dev\text{-}turn}$ derived from ordinary least square (OLS) regressions.

Appendix S11: Comparison of the effects of predictor variables on response variables before and after controlling for the effect of other predictors.

Appendix S12: Total effects of environmental and historical variables on SR and SRar.

Appendix S13: Plant composition and diversity metrics in areas with consistently high and low taxonomic and phylogenetic turnover. [Spreadsheet]

Appendix S1: Overview of the datasets included in this study.

Table S1.1: Overview of the datasets included in this study. For each dataset, we provide the total number and the proportion of plots over the total after resampling procedure (*N* = 4,930). See the list of databases at http://euroveg.org/eva-database-participating-databases for details.

Database	GIVD code	Database custodian	Number of plots	% of plots
SOPHY	EU-FR-003	Henry Brisse	694	14.08
French National Forest Inventory			643	13.04
Czech National Phytosociological Database	EU-CZ-001	Milan Chytrý	375	7.61
Vegetation Plot Database - Sapienza University of Rome	EU-IT-011	Emiliano Agrillo	259	5.25
Swiss Forest Vegetation Database	EU-CH-005	Thomas Wohlgemuth	228	4.62
Iberian and Macaronesian Vegetation Information System (SIVIM)	EU-00-004	Xavier Font	218	4.42
Slovak Vegetation Database	EU-SK-001	Milan Valachovič	215	4.36
Austrian Vegetation Database	EU-AT-001	Wolfgang Willner	199	4.04
German Vegetation Reference Database (GVRD)	EU-DE-014	Ute Jandt	183	3.71
Romanian Forest Database	EU-RO-007	Adrian Indreica	174	3.53
Balkan Vegetation Database	EU-00-019	Kiril Vassilev	172	3.49
Hellenic Natura 2000 Vegetation Database (HelNatVeg)	EU-GR-005	Panayotis Dimopoulos	166	3.37

Polish Vegetation Database	EU-PL-001	Zygmunt Kącki	166	3.37
Croatian Vegetation Database	EU-HR-002	Željko Škvorc	157	3.18
Vegetation Database of Slovenia	EU-SI-001	Urban Šilc	142	2.88
Iberian and Macaronesian Vegetation Information System (SIVIM) – Deciduous Forests	EU-00-023	Juan Antonio Campos	140	2.84
Vegetation Database of Habitats in the Italian Alps - HabItAlp	EU-IT-010	Laura Casella	98	1.99
EcoPlant	EU-FR-005	Jean-Claude Gégout	97	1.97
Vegetation-Plot Database of the University of the Basque Country (BIOVEG)	EU-00-011	Idoia Biurrun	94	1.91
VegetWeb Germany	EU-DE-013	Florian Jansen	85	1.72
Forest Database of Southern Poland	EU-PL-003	Remigiusz Pielech	79	1.60
The Nordic Vegetation Database	EU-00-018	Jonathan Lenoir	60	1.22
Hellenic Woodland Database + Hellenic Beech Forests Database (Hell-Beech-DB)	EU-GR-006 + EU-GR-007	Ioannis Tsiripidis	50	1.01
INBOVEG	EU-BE-002	Els De Bie	47	0.95
UK National Vegetation Classification Database	EU-GB-001	John S. Rodwell	40	0.81
Dutch National Vegetation Database	EU-NL-001	Stephan Hennekens	35	0.71
VegMV	EU-DE-001	Florian Jansen	20	0.41
European Boreal Forest Vegetation Database	EU-00-027	Anni Kanerva Jašková	15	0.30

Vegetation Database of Ukraine and Adjacent Parts of Russia	EU-UA-006	Viktor Onyshchenko	15	0.30
Vegetation Database of Albania	EU-AL-001	Michele De Sanctis	15	0.30
Iberian and Macaronesian Vegetation Information System (SIVIM) – Sclerophyllous Forests	EU-00-004	Federico Fernández-González	15	0.30
SE Europe forest database	EU-00-021	Andraž Čarni	11	0.22
VegItaly	EU-IT-001	Roberto Venanzoni	8	0.16
Database Schleswig-Holstein (Northern Germany)	EU-DE-040	Joachim Schrautzer	5	0.10
Dutch Military Ranges Vegetation Database (DUMIRA)	EU-NL-003	Iris de Ronde	4	0.08
CircumMed Pine Forest database	EU-00-026	Gianmaria Bonari	3	0.06
Iberian and Macaronesian Vegetation Information System (SIVIM) – Floodplain forests	EU-00–024	Idoia Biurrun	2	0.04
Masaryk University's Gap-Filling Database of European Vegetation		Milan Chytrý	1	0.02

Appendix S2: Justification for the removal of non-angiosperm species and results considering all vascular plants.

In this study, we obtained, for each plot, data on species occurrences and selected only angiosperms, thus excluding gymnosperms, pteridophytes, lycophytes, and non-vascular plants (~ 4%) to avoid inflated results due to a few taxa connected to deep phylogenetic nodes (e.g., Massante and Gerhold, 2020; Qian et al., 2020). Although approximately 87% of the grid cells in our study had less than 15% of non-angiosperm vascular plant species in their species pool (Fig. S2.1), around 10% of the variation in the turnover component of phylogenetic beta-diversity (PBD_{turn}) was explained by the variation in the proportion of non-angiosperm species (log-transformed) in the grid cells (Fig. S2.2).

Our results showed that PBD_{turn} calculated using all vascular plants was less affected by species richness than PBD_{turn} calculated using only angiosperm species (see Fig. S2.5 vs. Fig. 3). Moreover, soil pH had the opposite effect on PBD_{turn} (i.e., a positive effect) which is largely attributed to the decrease in its indirect effect through species richness.

References:

Massante, J.C. and Gerhold, P. (2020) Environment and evolutionary history depict phylogenetic alpha and beta diversity in the Atlantic coastal white-sand woodlands. *Journal of Vegetation Science*, 31, 634–645. https://doi.org/10.1111/jvs.12900

Qian, H., Jin, Y., Leprieur, F., Wang, X. and Deng, T. (2020) Geographic patterns and environmental correlates of taxonomic and phylogenetic beta diversity for large-scale angiosperm assemblages in China. *Ecography*. https://doi.org/10.1111/ecog.05190

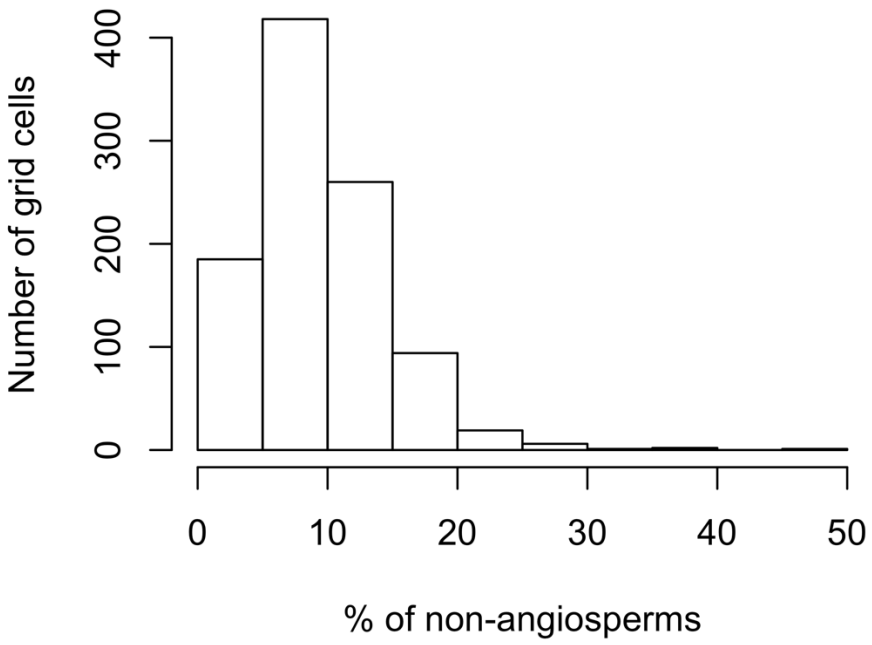

Fig. S2.1: Histogram of the proportion (%) of non-angiosperm species in the grid cells included in our study.

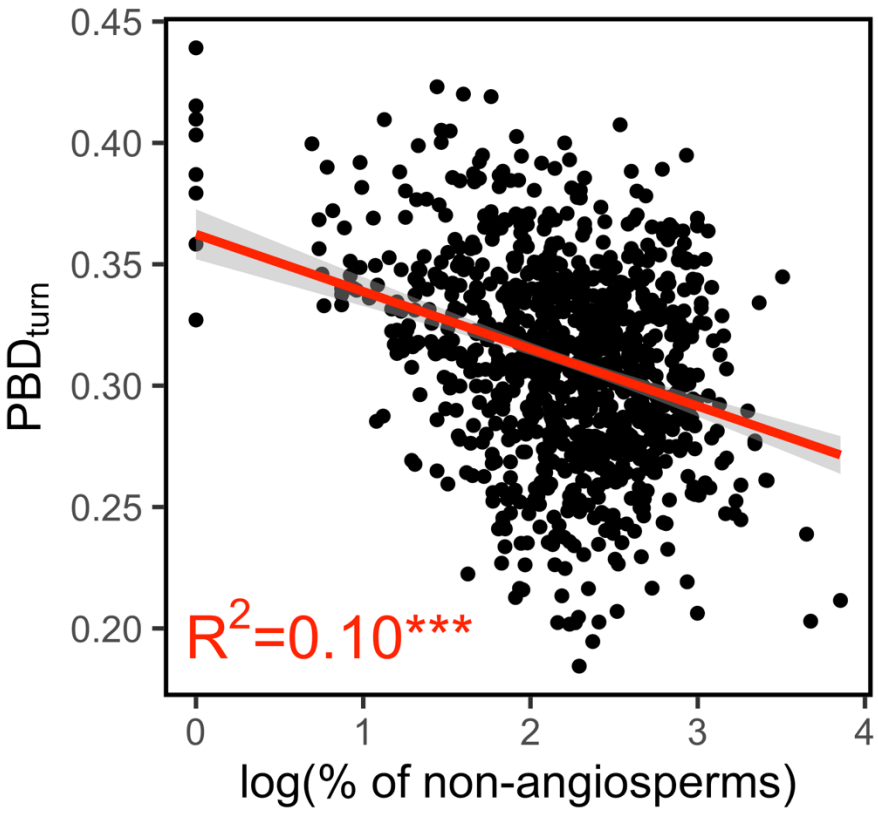

Fig. S2.2: Linear relationship between the proportion (%) of non-angiosperm species (log-transformed) and PBD_{turn}. Shaded grey areas represent 95% confidence intervals. R-squared is also provided. ***$P < 0.001$.

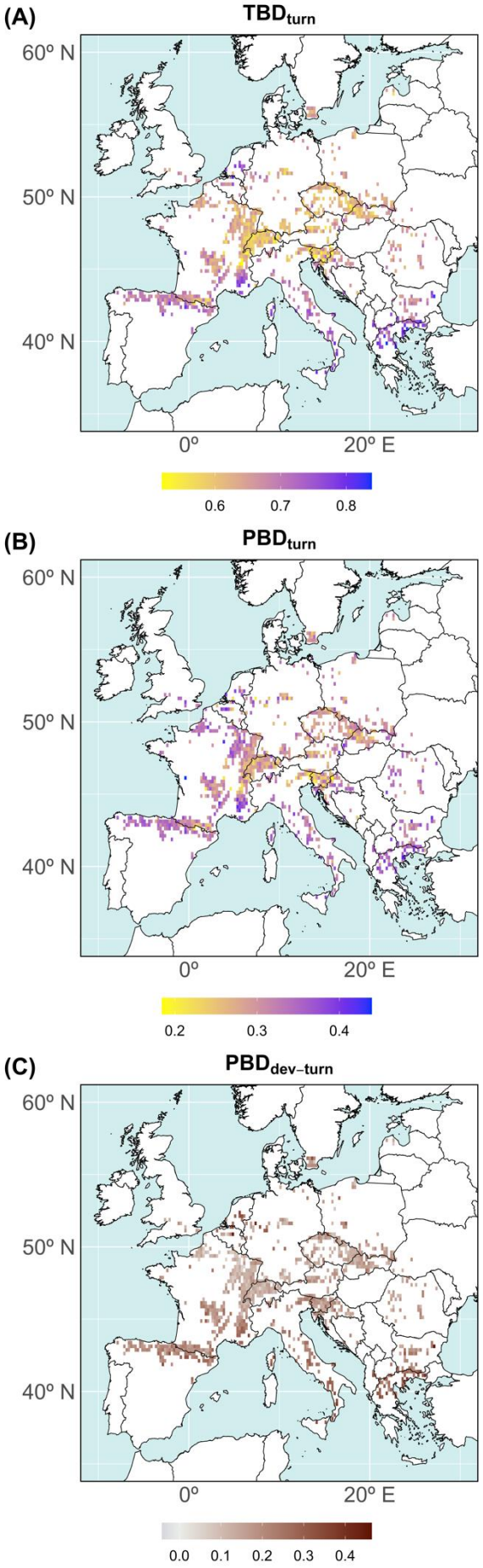
(A)
TBD turn
60° N
50° N
40° N
0°
20° E
0.6
0.7
0.8
(B)
PBD turn
60° N
50° N
40° N
0°
20° E
0.2
0.3
0.4
(C)
PBD dev−turn
60° N
50° N
40° N
0°
20° E
0.0
0.1
0.2
0.3
0.4

Fig. S2.3: Spatial patterns of beta-diversity of beech forest vegetation in Europe based on the data including all vascular plants. The maps show (A) the spatial turnover component of taxonomic beta-diversity (TBD_{turn}) and (B) phylogenetic beta-diversity (PBD_{turn}), and (C) the deviations of PBD_{turn} given TBD_{turn} ($PBD_{dev\text{-}turn}$).

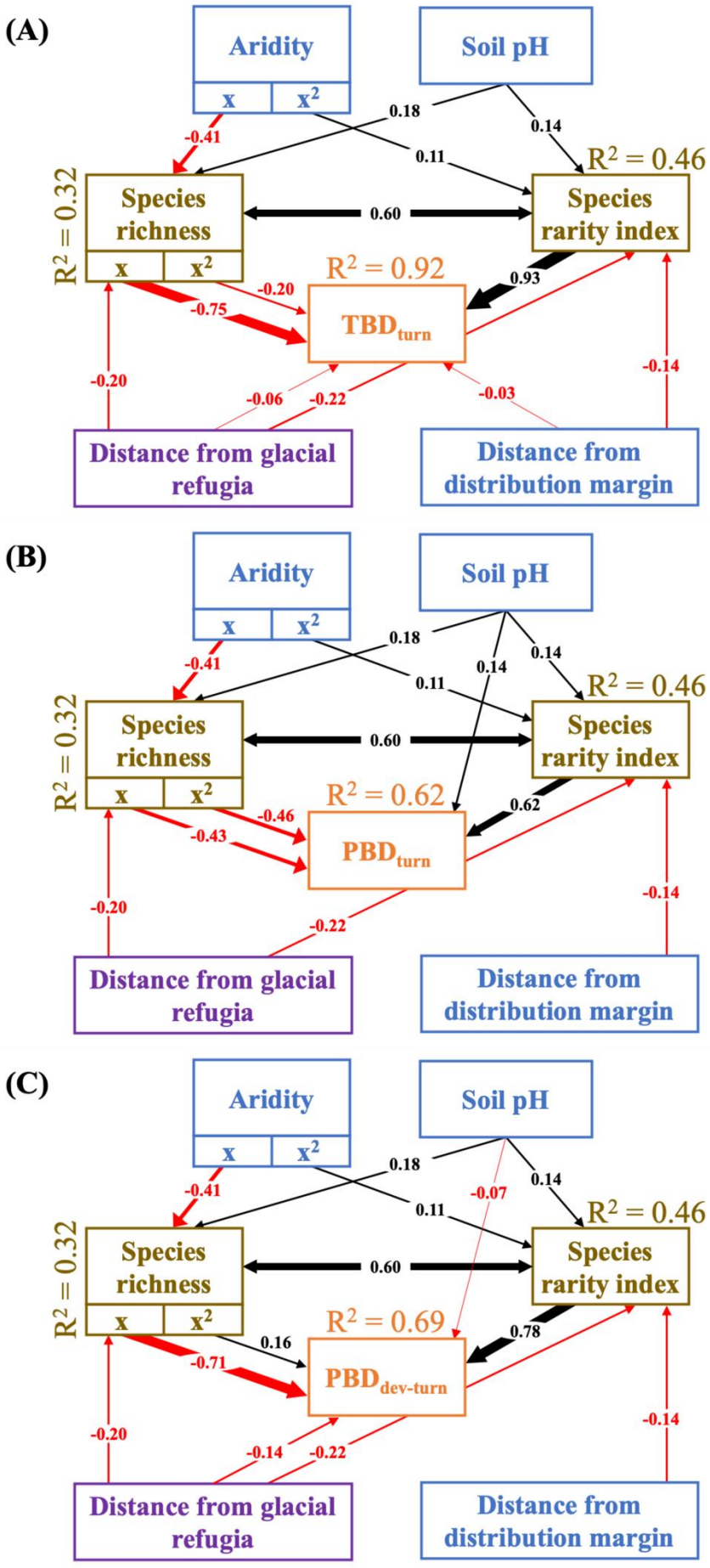
(A)
Aridity
x
x²
Soil pH
0.18
0.14
-0.41
0.11
R² = 0.46
Species richness
x
x²
0.60
Species rarity index
R² = 0.32
R² = 0.92
-0.20
0.93
-0.75
TBDturn
-0.20
-0.14
-0.06
-0.22
-0.03
Distance from glacial refugia
Distance from distribution margin
(B)
Aridity
x
x²
Soil pH
0.18
0.14
0.14
-0.41
0.11
R² = 0.46
Species richness
x
x²
0.60
Species rarity index
R² = 0.32
R² = 0.62
-0.46
0.62
-0.43
PBDturn
-0.20
-0.14
-0.22
Distance from glacial refugia
Distance from distribution margin
(C)
Aridity
x
x²
Soil pH
0.18
0.14
-0.41
-0.07
0.11
R² = 0.46
Species richness
x
x²
0.60
Species rarity index
R² = 0.32
R² = 0.69
0.16
0.78
-0.71
PBDdev-turn
-0.20
-0.14
-0.14
-0.22
Distance from glacial refugia
Distance from distribution margin

Fig. S2.4: Final structural equation models (SEM) explaining TBD_{turn} (A), PBD_{turn} (B), and $PBD_{dev\text{-}turn}$ (C) including all vascular plants. Environmental and historical variables are shown in blue and purple, respectively, while beta-diversity response variables are shown in orange. Intermediate response variables (species richness and species rarity index) are shown in brown. Nagelkerke pseudo-R-squared values (R^2) are shown for the dependent variables. 'x^2' indicates if the variable was transformed following: $(x - \text{mean}(x))^2$. Black and red directional arrows represent positive and negative causal links, respectively. Double-headed arrows represent non-causal correlations. Path thickness reflects the strength of the relationship (i.e., values of standardized β-coefficients). Standardized β-coefficients and correlation coefficients are also shown. 'Aridity' refers to the first PCA component obtained from the set of climatic variables and depicts an environmental gradient ranging from low temperatures and high precipitation to high temperatures and low precipitation (Appendix S4). The variables 'distance from distribution margin' and the 'distance from glacial refugia' of beech were log-transformed.

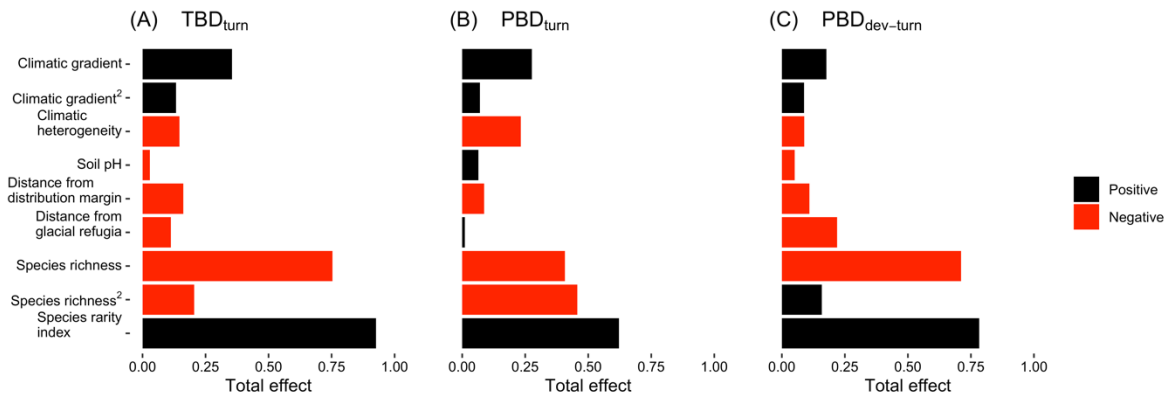

Fig. S2.5: Total effects of predictor variables on TBD_{turn} (A), PBD_{turn} (B), and $PBD_{dev\text{-}turn}$ (C) including all vascular plants. Total effects are defined as the sum of the direct and indirect paths between variables (excluding the paths denoted by double-headed arrows in Fig. S2.4). To allow for the comparison of total effects among variables, we placed both positive and negative effects on the same scale and colored them as black and red, respectively. 'Aridity' refers to the first PCA component obtained from the set of climatic variables and depicts an environmental gradient ranging from low temperatures and high precipitation to high temperatures and low precipitation (Appendix S4). The variables 'distance from distribution margin' and the 'distance from glacial refugia' of beech were log-transformed.

Appendix S3: General results considering a different number of plots per grid cell.

In this section, we present maps and results from SEM models considering thresholds of 10 and 15 plots per grid cell, respectively. Increasing the number of plots per grid cell reduced the total number of grid cells included in our study, but we found no substantial differences from the results produced with 5 per grid cell (see Fig. S3.1 vs. Fig. 1; Fig. S3.2 vs. Fig. 3).

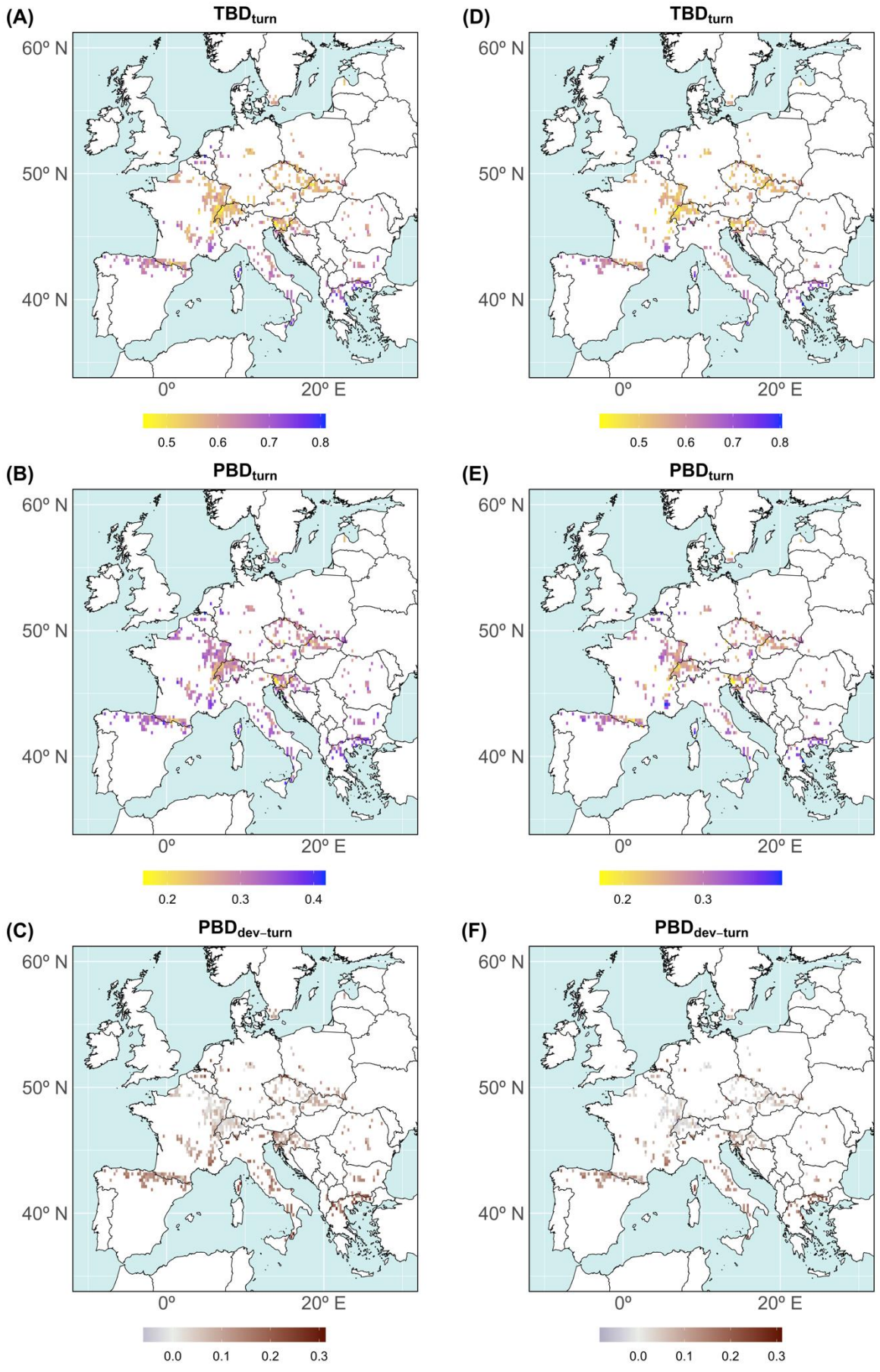
(A)
TBD turn
60° N
50° N
40° N
0°
20° E
0.5
0.6
0.7
0.8
(D)
TBD turn
60° N
50° N
40° N
0°
20° E
0.5
0.6
0.7
0.8
(B)
PBD turn
60° N
50° N
40° N
0°
20° E
0.2
0.3
0.4
(E)
PBD turn
60° N
50° N
40° N
0°
20° E
0.2
0.3
(C)
PBD dev−turn
60° N
50° N
40° N
0°
20° E
0.0
0.1
0.2
0.3
(F)
PBD dev−turn
60° N
50° N
40° N
0°
20° E
0.0
0.1
0.2
0.3

Fig. S3.1: Spatial patterns of beta-diversity of beech forest vegetation in Europe including 10 (A-C) and 15 (D-F) plots per grid cell after running heterogeneity-constrained random (HCR) resampling. The maps show (A & D) the spatial turnover component of taxonomic beta-diversity (TBD_{turn}) and (B & E) phylogenetic beta-diversity (PBD_{turn}), and (C & F) the deviations of PBD_{turn} given TBD_{turn} ($PBD_{dev\text{-}turn}$).

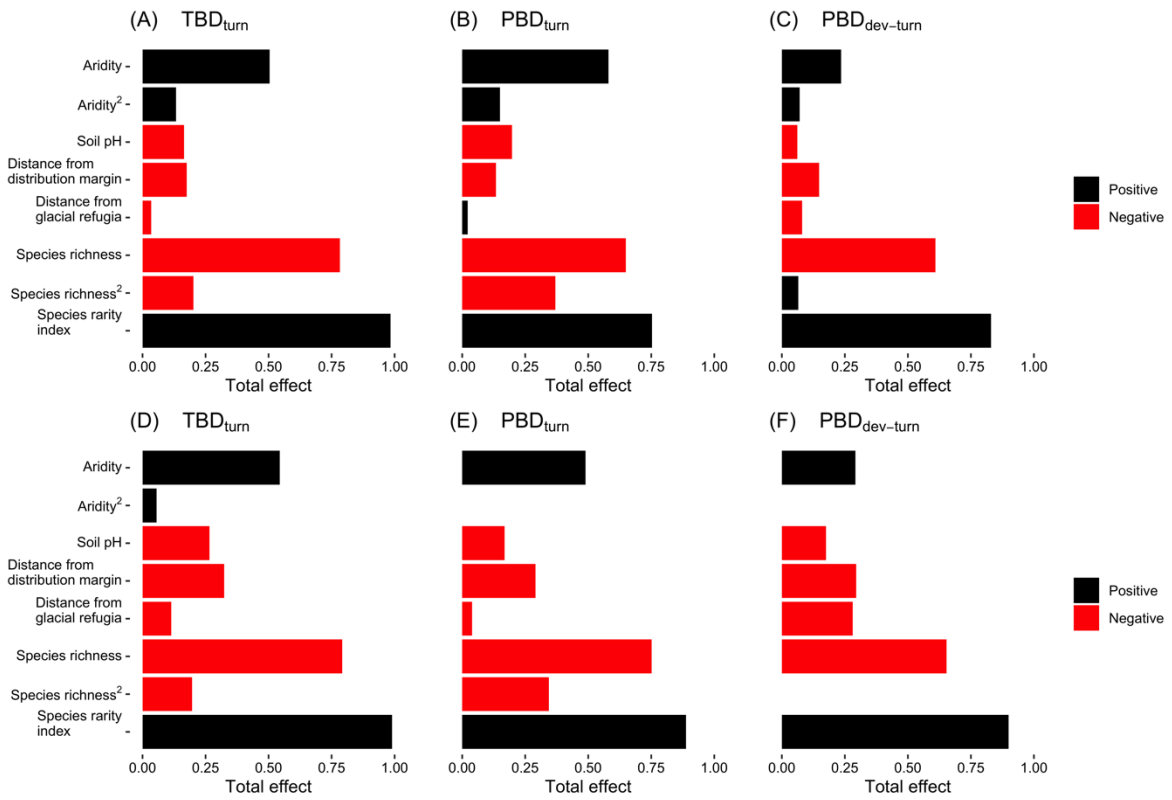

Fig. S3.2: Total effects of predictor variables on TBD_{turn} (A & D), PBD_{turn} (B & E), and $PBD_{dev\text{-}turn}$ (C & F) including 10 (A-C) and 15 (D-F) plots per grid cell after running heterogeneity-constrained random (HCR) resampling. Total effects are defined as the sum of the direct and indirect paths between variables in SEM models. To allow for the comparison of total effects among variables, we placed both positive and negative effects on the same scale and colored them as black and red, respectively. 'Aridity' refers to the first PCA component obtained from the set of climatic variables and depicts an environmental gradient ranging from low temperatures and high precipitation to high temperatures and low precipitation (Appendix S4). The variables 'distance from distribution margin' and the 'distance from glacial refugia' of beech were log-transformed.

Appendix S4: Results from Principal Component Analyses (PCA) of environmental variables.

Table S4.1: Loadings of climatic variables (**mean values**) on the first three PCA components. Loadings > 0.5 are shown in bold.

Code	Variable Name	PC1	PC2	PC3
BIO1	Annual Mean Temperature	**0.797**	0.420	0.272
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	0.138	-0.183	**0.814**
BIO3	Isothermality (BIO2/BIO7) (×100)	0.209	**0.587**	0.336
BIO4	Temperature Seasonality (standard deviation ×100)	-0.112	**-0.852**	0.379
BIO5	Max Temperature of Warmest Month	**0.768**	0.075	**0.596**
BIO6	Min Temperature of Coldest Month	**0.642**	**0.708**	-0.097
BIO7	Temperature Annual Range (BIO5-BIO6)	-0.015	**-0.716**	**0.648**
BIO8	Mean Temperature of Wettest Quarter	0.064	**-0.690**	0.168
BIO9	Mean Temperature of Driest Quarter	**0.547**	**0.638**	0.144
BIO10	Mean Temperature of Warmest Quarter	**0.791**	0.091	0.457
BIO11	Mean Temperature of Coldest Quarter	**0.670**	**0.699**	0.059
BIO12	Annual Precipitation	**-0.792**	0.495	0.301
BIO13	Precipitation of Wettest Month	**-0.792**	0.265	0.344
BIO14	Precipitation of Driest Month	**-0.746**	**0.518**	0.146
BIO15	Precipitation Seasonality (Coefficient of Variation)	0.089	**-0.594**	0.050
BIO16	Precipitation of Wettest Quarter	**-0.811**	0.275	0.314
BIO17	Precipitation of Driest Quarter	**-0.730**	0.570	0.202
BIO18	Precipitation of Warmest Quarter	**-0.881**	-0.166	0.165
BIO19	Precipitation of Coldest Quarter	-0.448	**0.810**	0.161
	Proportion of Variance	0.373	0.300	0.130
	Cumulative Proportion of Variance	0.373	0.673	0.802

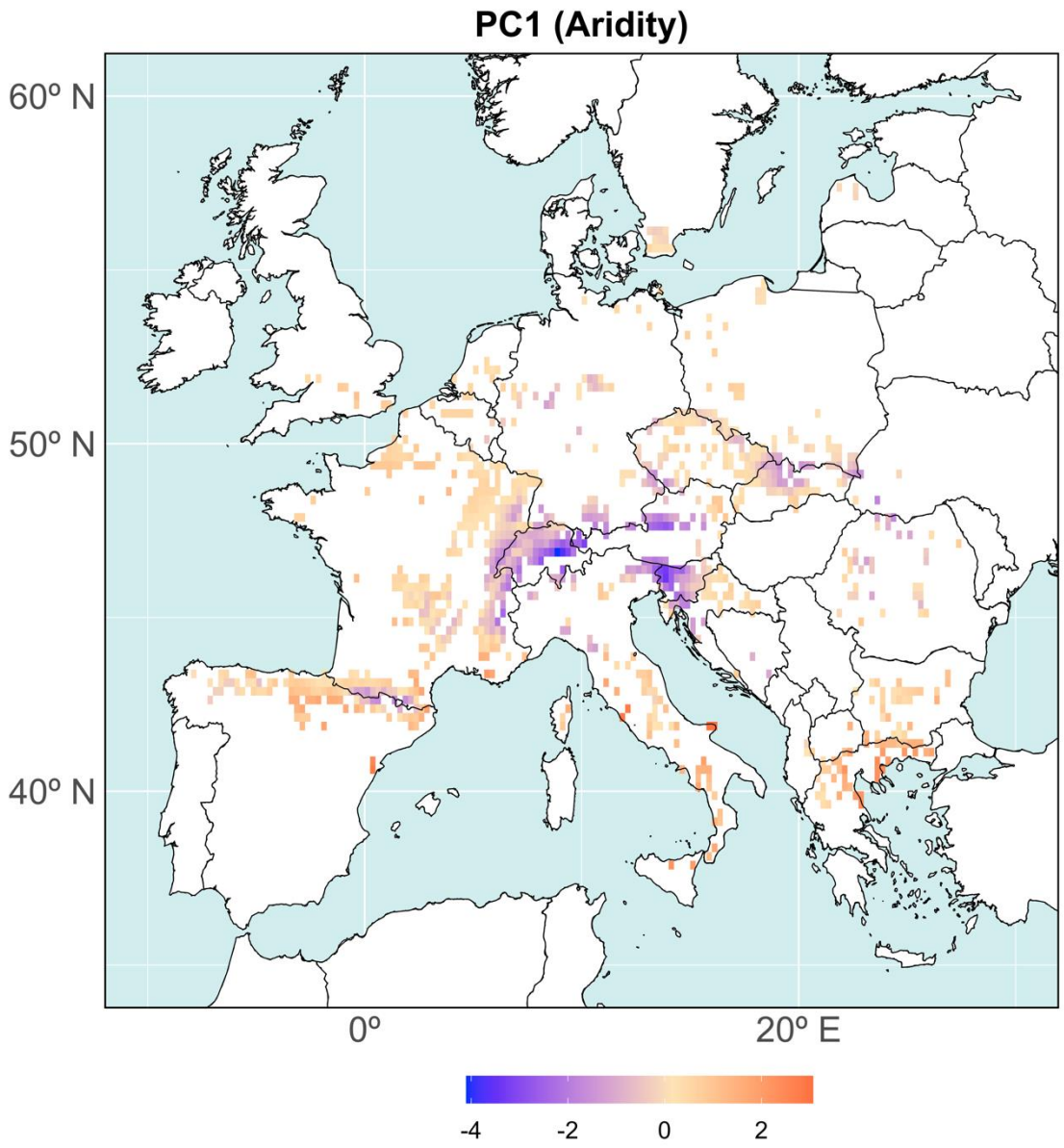

Fig. S4.1: Map of scores for the first PCA component across the studied grid cells. Areas with low values (blue tones) are characterized by low temperatures and high precipitation, while areas with high values (red tones) are characterized by high temperatures and low precipitation (see Table S4.1).

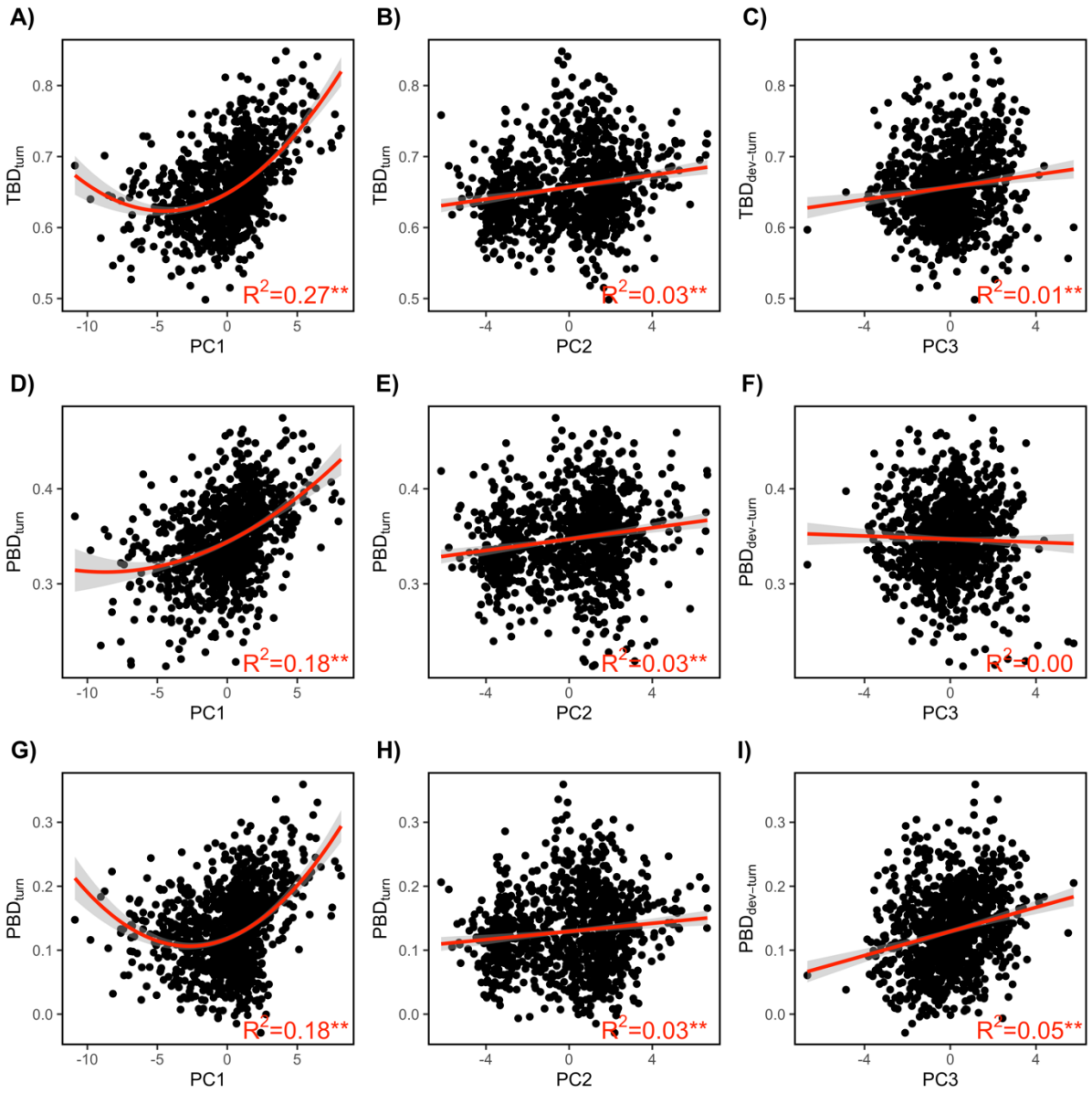

Fig. S4.2: Relationship between the three PCA components (PC1, PC2, and PC3) obtained with mean values of the 19 bioclimatic variables and TBD_{turn} (A-C), PBD_{turn} (D-F), and $PBD_{dev\text{-}turn}$ (G-I) derived from ordinary least square (OLS) regressions. All quadratic relationships had higher R-squared values than linear regressions. Shaded grey areas represent 95% confidence intervals. R-squared are also provided. $^{**}P < 0.001$.

Appendix S5: Distribution of beech (*Fagus sylvatica*).

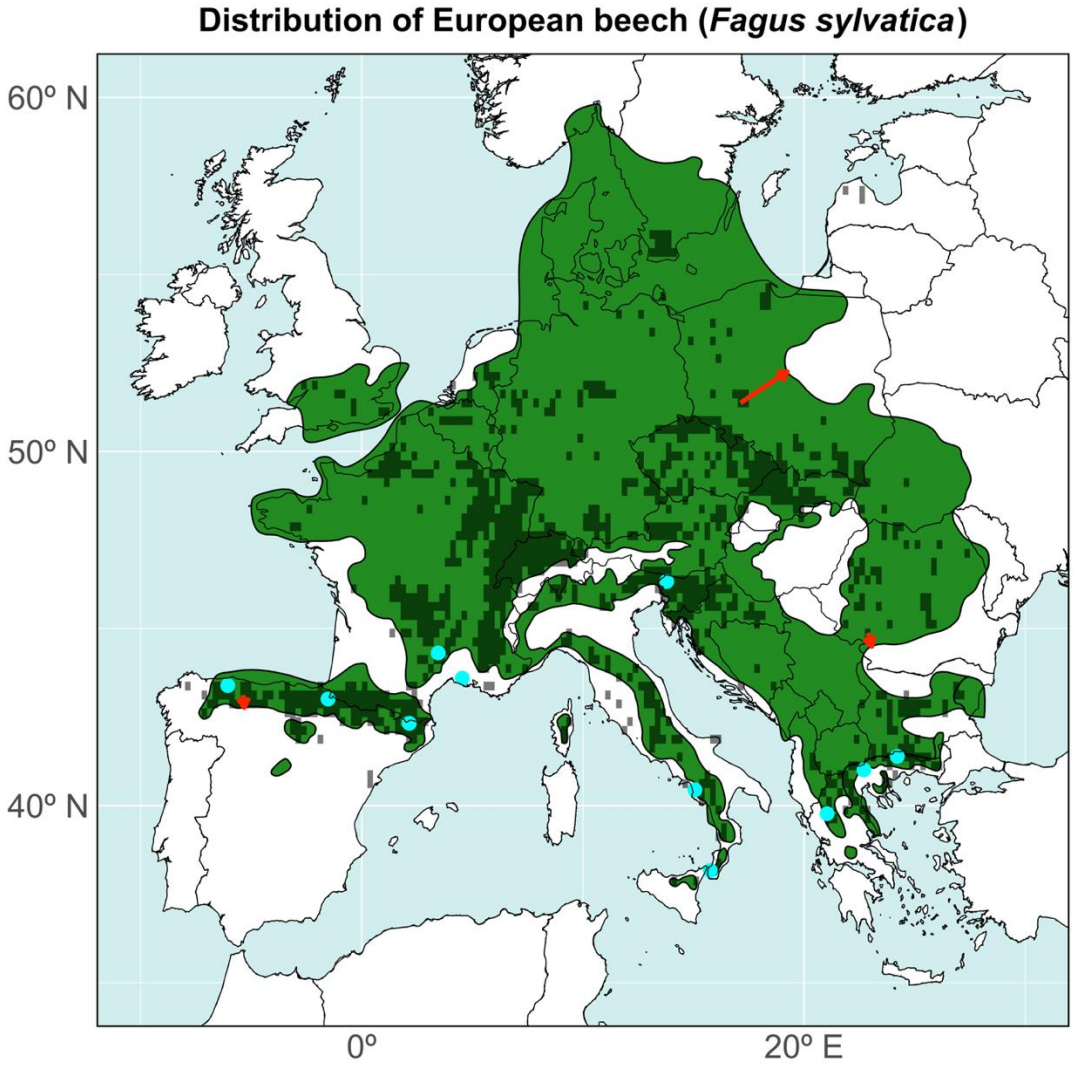

Fig. S5.1: Distribution of European beech (*Fagus sylvatica*) derived from the shapefile obtained from Caudullo et al. (2017) (green areas). Note the distribution of European beech shown in this map is broader than the distribution of forests dominated by European beech, which are the focus of this study. We have overlaid on top of the European beech distribution the grid cells used in the present study (dark green areas). Cyan dots are the locations of tentative glacial refugia used in the present study and derived from Magri et al. (2006) (see Methods). Red arrows connect three random grid cells to the closest point in the geographical distribution margin.

Appendix S6: Correlations between environmental and historical variables.

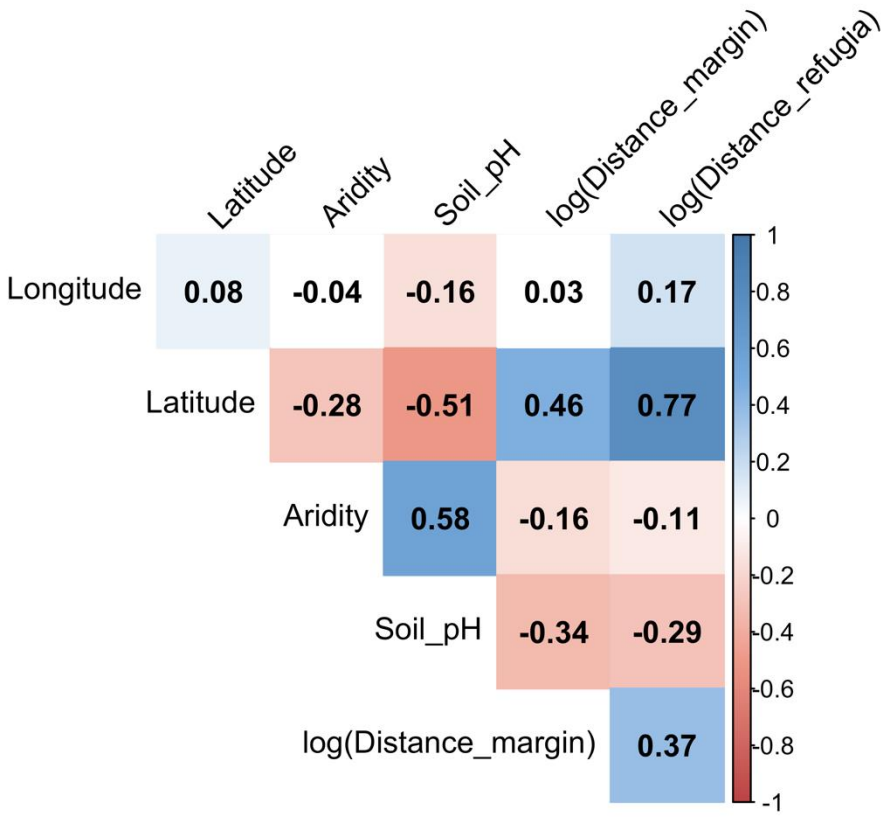

Fig. S6.1: Correlation matrix with Pearson's correlation coefficients between environmental and historical variables and latitude, longitude, species richness (SR) and the species rarity index (SRar). Only cells with significant correlations ($P < 0.01$) are colored.

log(Distance_margin) = Distance from the distribution margin (log-transformed);

log(Distance_refugia) = Distance from glacial refugia (log-transformed).

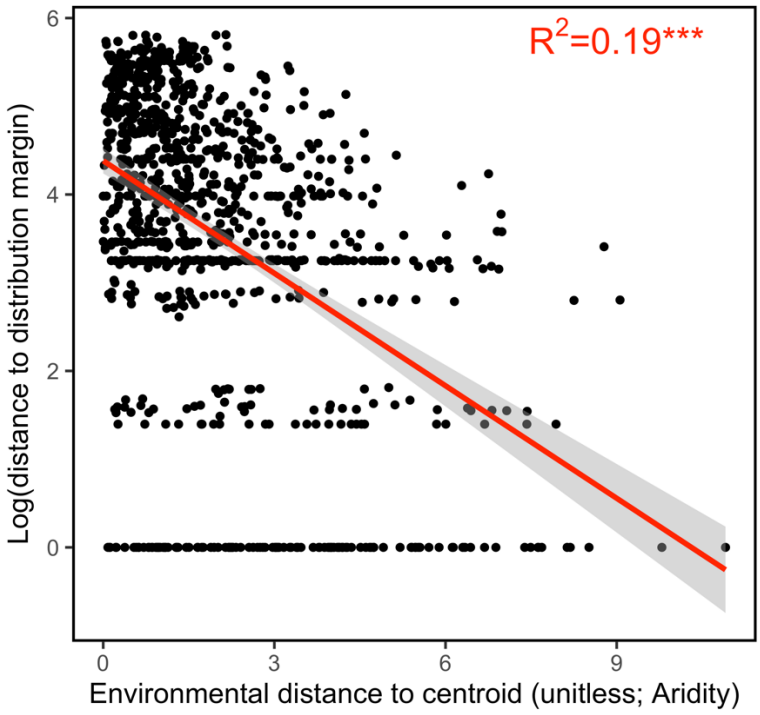

Fig. S6.2: Linear relationship between the environmental distance of each cell to its centroid (i.e., mean value of 'Aridity') and the log-transformed distance of each cell to the closest margin of the geographical distribution of European beech. Shaded grey areas represent 95% confidence intervals. R-squared is also provided. ***$P < 0.001$.

Appendix S7: Calculation of the 'species rarity index' (SRar) and relationship with species richness and the 'corrected weighted endemism' (CWE).

The 'species rarity index' (SRar) was positively related to species richness ($R^2 = 0.27$) and the log-transformed version of the 'corrected weighted endemism' (CWE) ($R^2 = 0.43$) proposed by Crisp et al. (2001). The CWE index, as SRar, is also a measure of endemism that is weakly related to species richness. However, CWE emphasizes rarer species over more common species, which is an undesirable property in our study. Therefore, we developed SRar that gives the same weight to all species. We present the correlations between SRar, SR, and the log-transformed version of CWE to clearly show the differences between these metrics.

References

Crisp, M.D., Laffan, S., Linder, H.P. and Monro, A. (2001) Endemism in the Australian flora: Endemism in the Australian flora. *Journal of Biogeography* 28, 183–198. https://doi.org/10.1046/j.1365-2699.2001.00524.x

R code function to calculate the 'species rarity index' (SRar).

data = matrix of species occurrences with sites as rows and species as columns.

```
srar<-function(data){
  disp.total.md<-numeric(length(data[,1]))
for(i in 1:nrow(data)){ # for each focal cell…
     disp<-numeric(length(data[1,]))
          # …calculate the relative frequency of all species found in it
     for(j in 1:ncol(data)){
       ifelse(data[i,j]==0, disp[j]<-NA , disp[j]<-((sum(data[,j])-1)/(nrow(data)-1)))
     }
     disp.total.md[i]<-1-median(disp, na.rm=T) # then, calculate median and reverse order
   }
  return(data.frame(cell=row.names(data), disp.total.m)) #store output
}
```

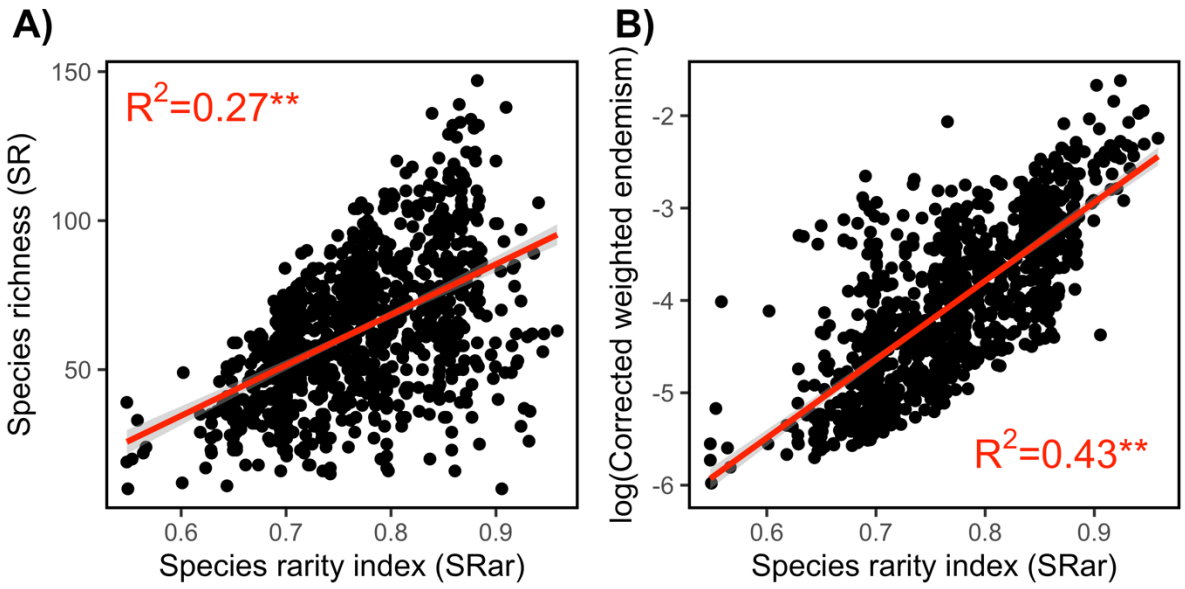

Fig. S7.1: Linear relationship between the 'species rarity index' (SRar) and (A) species richness (SR), and (B) the 'corrected weighted endemism' (log-transformed) proposed by Crisp et al. (2001). Shaded grey areas represent 95% confidence intervals. R-squared is also provided. $^{**}P < 0.01$.

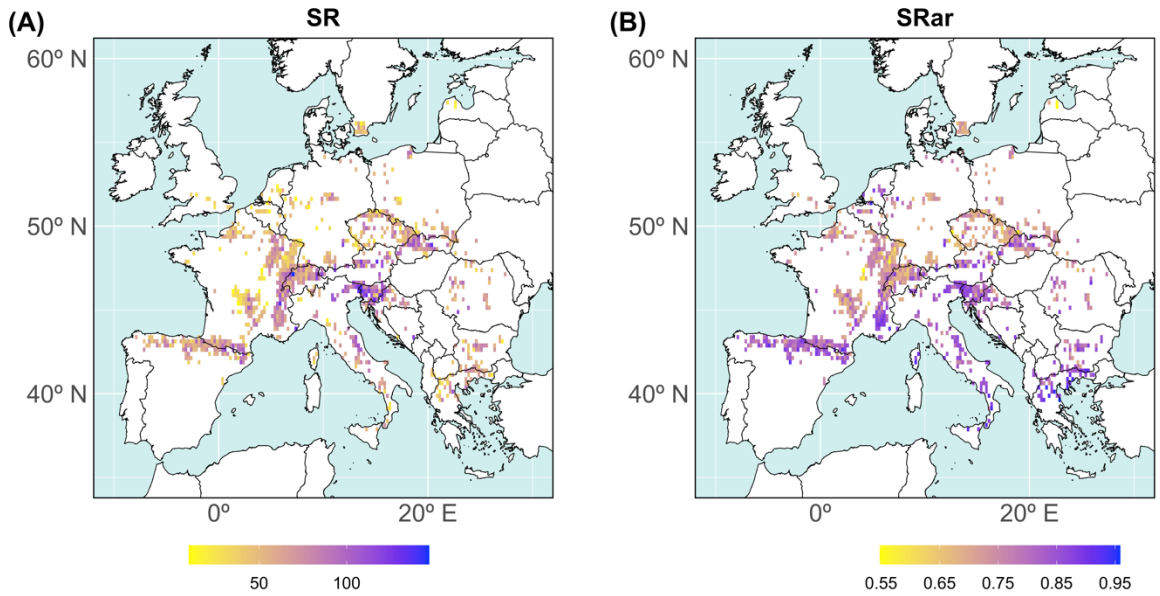

Fig. S7.2: Spatial distribution of (A) species richness and (B) the 'species rarity index' (SRar) of European beech forests.

Appendix S8: Spatial correlograms.

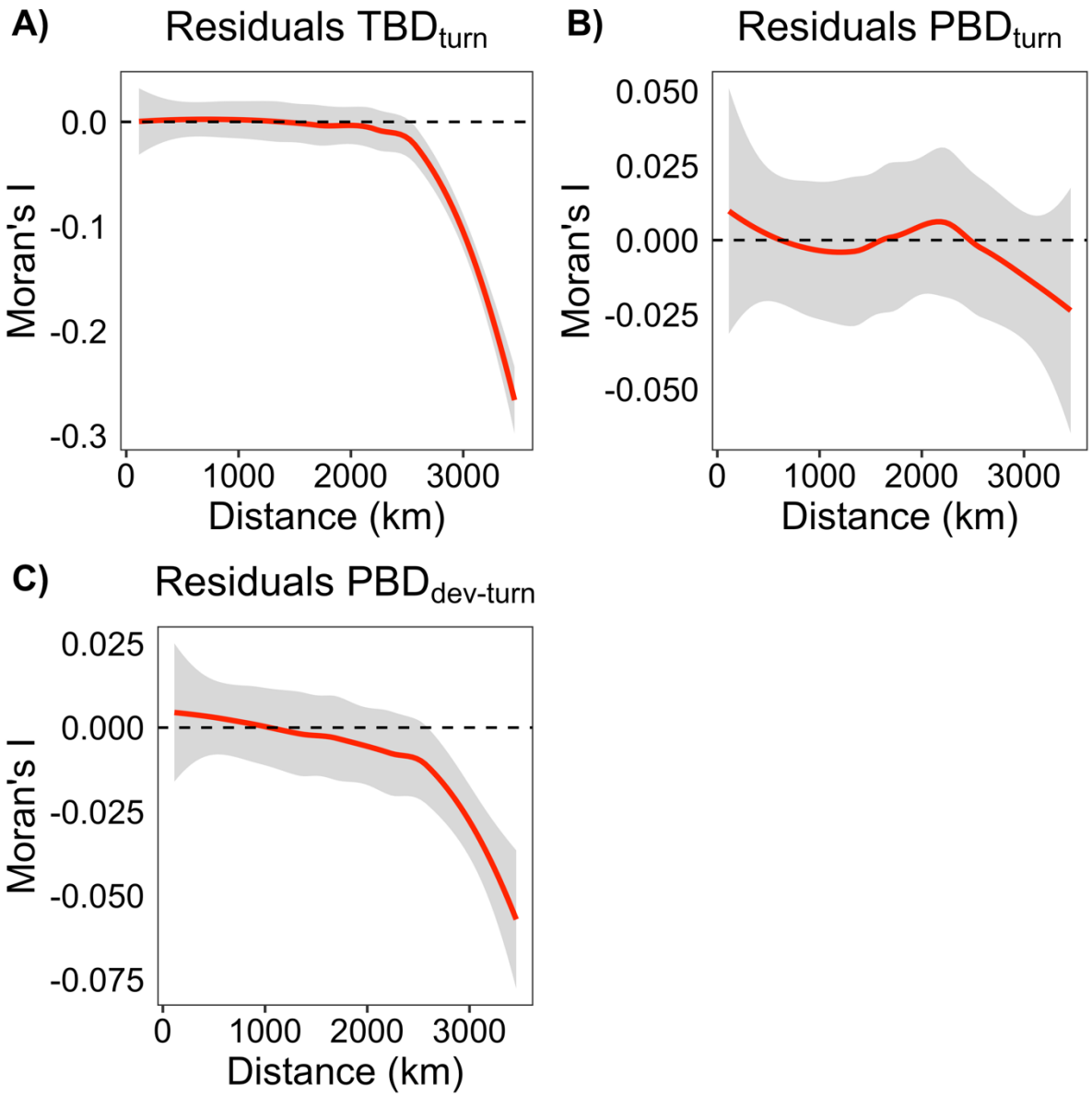

Fig. S8.1: Spatial correlograms of SEM residuals predicting TBD_{turn} (A), PBD_{turn} (B), and $PBD_{dev\text{-}turn}$ (C). Shaded grey areas represent 95% confidence intervals.

Appendix S9: Example of the R code used to implement SEM.

```
#############
# Example of R code used to implement Structural Equation Modelling (SEM)
# with Simultaneous Autoregressive Models (SAR)
#############

#Load libraries:
library(piecewiseSEM)
library(spdep)
library(spatialreg)

#Load data (rows are grid cells and columns are variables):
dat<-read.table("~/example.csv", sep=";")

#Create groups of neighboring cells:
dat2<-dat
coordinates(dat2)<- ~ longitude + latitude
W<-nb2listw(dnearneigh(coordinates(dat2), 0, 150, longlat = TRUE), style="W")
#neighbours within 150 km
plot(W, coordinates(dat2)) #visualize network of neighbors

#Calculate quadratic terms of variables to add to the models:
dat$climate.PC1.sq<- (dat$climate.PC1-mean(dat$climate.PC1))^2
dat$SR.sq<- (dat$SR-mean(dat$SR))^2

#Run SEM:
mod.tbd <- psem(
  errorsarlm(tbd.sim ~ SR + SR.sq + SRar + dist.margin + dist.refugia, data=dat, W),
  errorsarlm(SR ~ climate.PC1 + PH + dist.ref, data=dat, W),
  errorsarlm(SRar ~ dist.margin + PH + dist.refugia + climate.PC1 + climate.PC1.sq,
data=dat, W),
  SR %~~% disp.total.md,
```

```
  SR.sq %~~% disp.total.md,
  SR %~~% SR.sq,
  data=dat)

#Get model results:
summary(mod.tbd)
```

Appendix S10: Relationships between latitude, SR, and the 'species rarity index' (SRar) and TBD_{turn}, PBD_{turn}, and $PBD_{dev\text{-}turn}$ derived from ordinary least square (OLS) regressions.

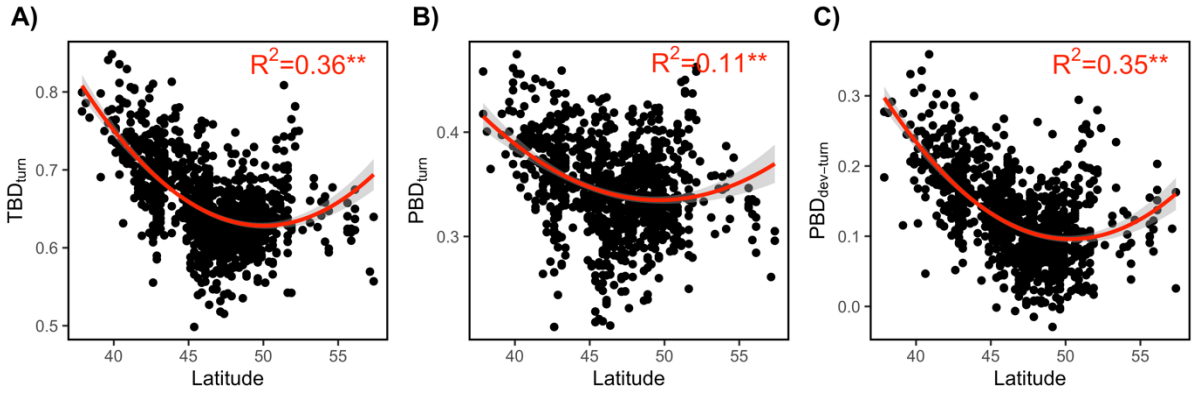

Fig. S10.1: Relationship between latitude and TBD_{turn} (A), PBD_{turn} (B), and $PBD_{dev\text{-}turn}$ (C) derived from ordinary least square (OLS) regressions. All quadratic relationships had higher R-squared values than linear regressions. Shaded grey areas represent 95% confidence intervals. R-squared are also provided. $^{**}P < 0.001$.

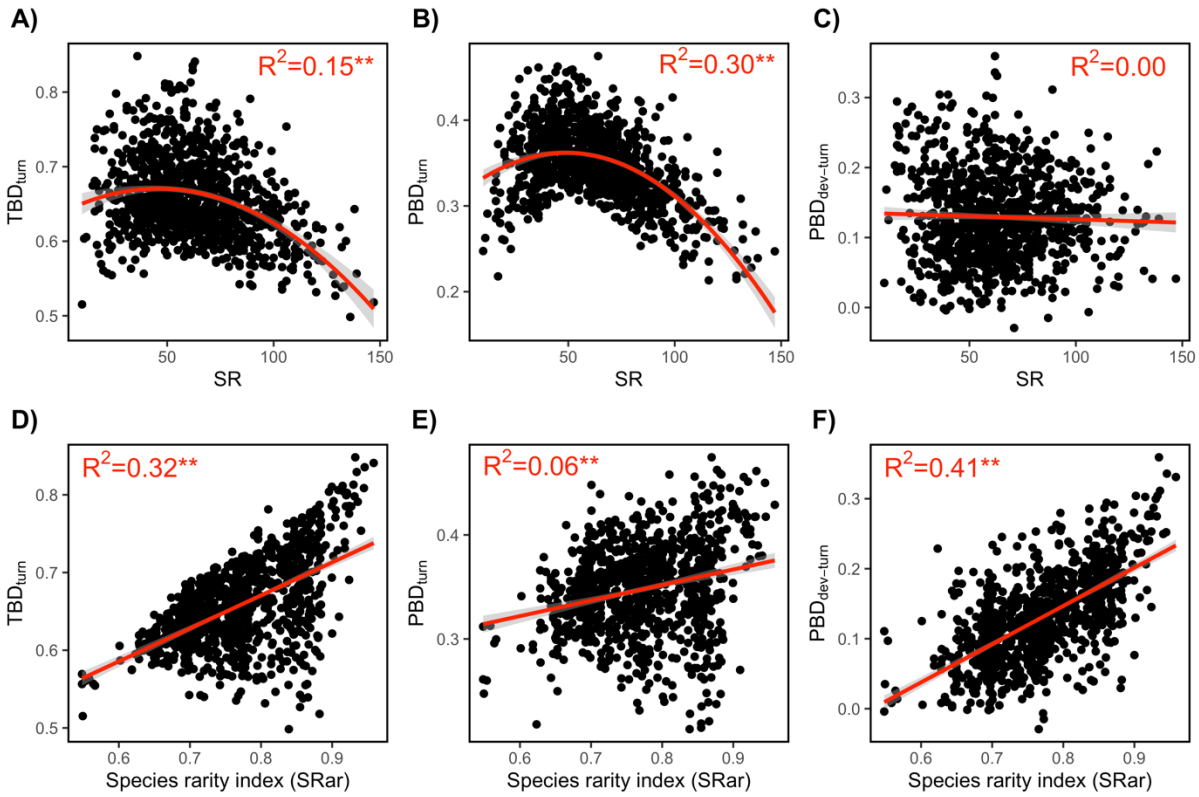

Fig. S10.2: Relationship between species richness (SR) (A-C) and the 'species rarity index' (SRar) (D-F) and TBD_{turn}, PBD_{turn}, and $PBD_{dev\text{-}turn}$ derived from ordinary least square (OLS) regressions. Quadratic relationships are shown when significant and when they had greater R-squared values than linear regressions. Shaded grey areas represent 95% confidence intervals. R-squared is also provided. $^{**}P < 0.001$.

Appendix S11: Comparison of the effects of predictor variables on response variables before and after controlling for the effect of other predictors.

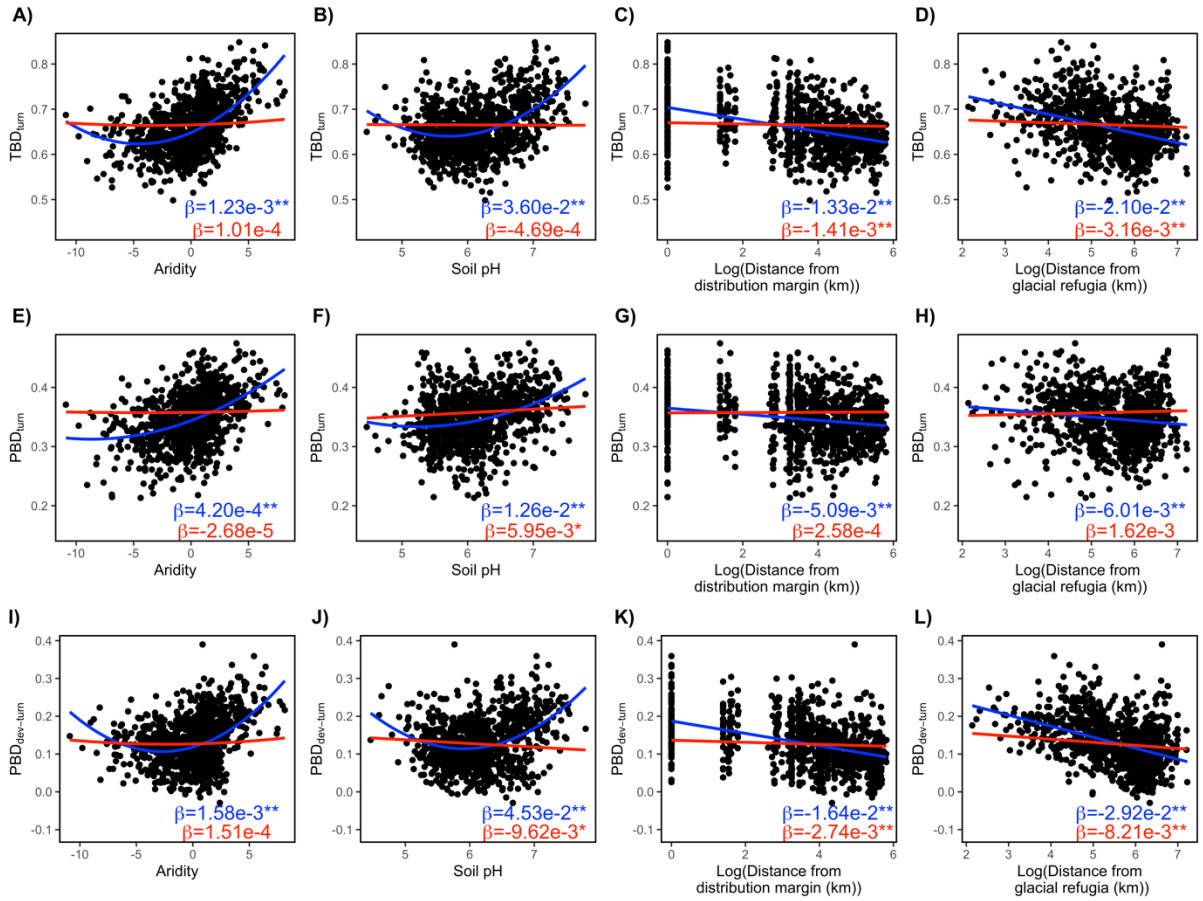

Fig. S11.1: Relationship between environmental and historical variables and TBD_{turn} (A-D), PBD_{turn} (E-H), and $PBD_{dev\text{-}turn}$ (I-L). The continuous red lines represent the effects of predictors after controlling for the effect of all other variables using SAR models. The blue lines are individual effects of predictors in ordinary least square (OLS) regressions. Standardized beta-coefficients are also provided. 'Aridity' refers to the first PCA component obtained from the set of climatic variables and depicts an environmental gradient ranging from low temperatures and high precipitation to high temperatures and low precipitation (Appendix S4). *P < 0.01; **P < 0.001.

Appendix S12: Total effects of environmental and historical variables on SR and SRar.

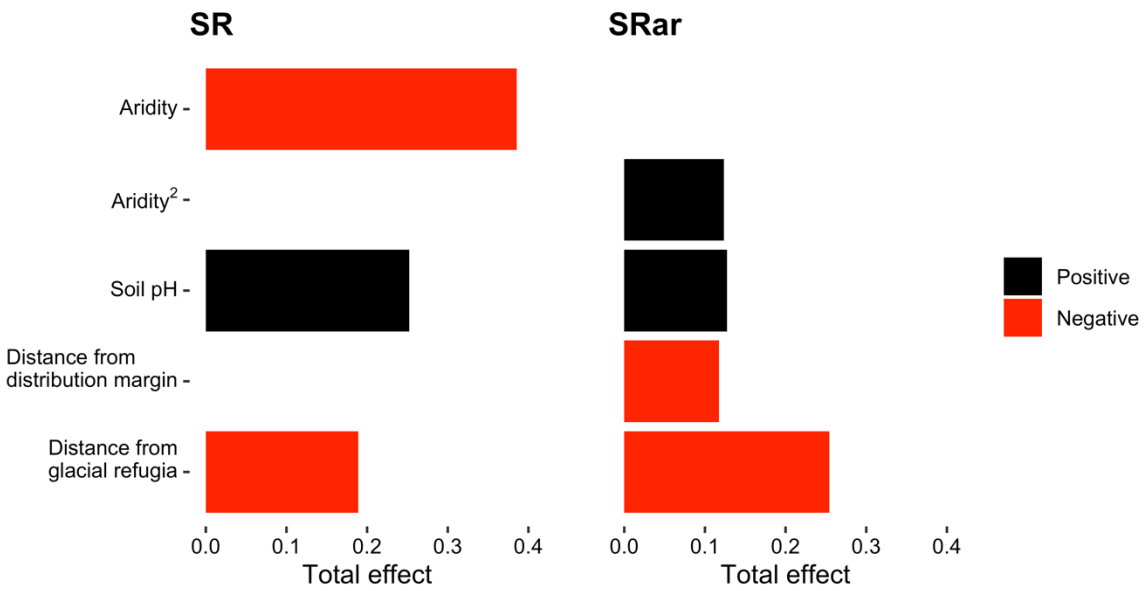

Fig. S12.1: Total effects of environmental and historical variables on species richness (SR) and the index of species rarity (SRar) derived from the final SEM (Fig. 2). Total effects are defined as the sum of the direct and indirect paths between variables (excluding the paths denoted by double-headed arrows in Fig. 2). To allow for the comparison of total effects among variables, we placed both positive and negative effects on the same scale and colored them as black and red, respectively. 'Aridity' refers to the first PCA component obtained from the set of climatic variables and depicts an environmental gradient ranging from low temperatures and high precipitation to high temperatures and low precipitation (Appendix S4). The variables 'distance from distribution margin' and the 'distance from glacial refugia' of beech were log-transformed.