

A biome-dependent distribution gradient of tree species range edges is strongly dictated by climate spatial heterogeneity

Received: 24 April 2022

Accepted: 10 February 2023

Published online: 9 March 2023

 Check for updates

David Lerner¹✉, Marcos Fernández Martínez², Stav Livne-Luzon¹, Jonathan Belmaker^{3,4}, Josep Peñuelas^{2,5} & Tamir Klein¹

Understanding the causes of the arrest of species distributions has been a fundamental question in ecology and evolution. These questions are of particular interest for trees owing to their long lifespan and sessile nature. A surge in data availability evokes a macro-ecological analysis to determine the underlying forces limiting distributions. Here we analyse the spatial distribution of >3,600 major tree species to determine geographical areas of range-edge hotspots and find drivers for their arrest. We confirmed biome edges to be strong delineators of distributions. Importantly, we identified a stronger contribution of temperate than tropical biomes to range edges, adding strength to the notion that tropical areas are centres of radiation. We subsequently identified a strong association of range-edge hotspots with steep spatial climatic gradients. We linked spatial and temporal homogeneity and high potential evapotranspiration in the tropics as the strongest predictors of this phenomenon. We propose that the poleward migration of species in light of climate change might be hindered because of steep climatic gradients.

The geographical distributions of species are marked by their range limits. Understanding the causes of distribution arrest has been a fundamental question in ecology and evolution^{1–4}. Given the strong interplay among biotic, abiotic, demographic, physical and historical forces in predicting range edges (REs), it has been challenging to find underpinnings for their formation. Two main environmental forces seem to play a major role in the formation of REs: spatial environmental heterogeneity and habitat quality^{4–6}. Most models on the formation of species REs rely on the interplay between either one of these two forces and non-climatic pressures to explain their formation. For example, steep climate gradients combined with high dispersal and gene flow reduce species' fitness and genetically constrain their evolution into new environments^{7,8}. Likewise, low habitat quality reduces population size^{2,9}, increasing drift and migration load^{10–12}. Nevertheless, the significance of climate in the interplay between these two environmental

components in defining REs for an array of species or on a wide biogeographical scale remains elusive^{13,14}.

The field of biogeography has long sought an understanding of species ranges despite having limited tools^{3,15}. Given the surge in large-scale datasets, it is now possible to better identify the underpinnings of species distributions by studying the macro-eco-evolutionary processes involved in their formation^{12,13,16–18}. Although methods and results are disparate between studies, there is almost a consensus that the presence of large-scale biogeographical units confines species with climate as their primary predictor. For example, Bontager et al.¹² suggested distinct characteristics for RE populations dependent on their latitude. Likewise, niche conservatism and strong beta-diversity patterns seem to withhold at large macro-ecological scales^{17–19}. This is generally true for plant species, with biomes being the most consistent classifier on the basis of structural and functional similarity^{20,21}. Recent

¹Department of Plant & Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel. ²CREAF, Cerdanyola de Vallès, Spain. ³School of Zoology, Tel Aviv University, Tel Aviv, Israel. ⁴Steinhardt Museum of Natural History, Tel Aviv University, Tel Aviv, Israel. ⁵CSIC, Global Ecology Unit CREA-FCI-UAB, Bellaterra, Spain. ✉e-mail: david.lerner@weizmann.ac.il

efforts have been made to understand how accurate and substantial biome entities are at defining species distributions^{21,22}. Nevertheless, it remains an open question whether, and to what extent, the intersection between biomes is a source of species range-edge hotspots (REHs). Deciphering such patterns will enable a proper understanding of how communities of species redistribute and are structured geographically, and if similar biomes in distinct geographic areas have similar effects on the distribution of species.

Although there has been increased interest in defining the biogeographical underpinnings of species distributions, most techniques have used species relatedness and diversity metrics to test for the existence of shared niche space between species and communities. Direct analysis of the ecological and climatic limitations on geographical space, although trivial, remains elusive. Here, we look at the universal set of climatic factors and geographical patterns of species distributions by focusing directly on species' RE distributions.

We present the first-to-date global study of tree species REs, applying a novel, simple, yet effective method of delineating REs to (1) identify deterministic patterns of REs, as seen by RE-dense areas (REHs); (2) determine whether the classification of biomes as distinct community-level patterns of biodiversity properly delineates the niche of species; (3) identify global-scale REH patterns; and (4) discern the underlying niche factors responsible for RE formation. In particular, we investigate whether the spatial heterogeneity of abiotic factors or a universal predictor of habitat quality are determinants of RE formation. We focus in particular on tree species because they are an exemplary group in the study of the ecological changes predicted to occur at the peripheries of distributions, given their long-lived characteristics and fundamental role in many ecosystems^{23,24}, specifically biomes^{19,25}.

Although we do not study the interplay between climate and other ecological and evolutionary limiting factors to species distributions (for example, seed dispersal, plasticity and adaptation), discerning these patterns and the climatic components leading to such distributions will enable a better understanding of interplays between biotic and abiotic factors in future studies. A better understanding of the climatic factors affecting dispersal allows for better predictions of the success of species to track changing climates, and in turn, whether they will be subject to migration lags^{26–28}.

Results

Global dataset and REH distribution

We present the first report, to our knowledge, of the global distribution of tree REHs (Fig. 1), marked by hexagons with significant clustering of REs (Supplementary Fig. 1). We did not identify any significant coldspots, given the baseline presence of REs around the globe. The visual patterns emerging from these distributions indicated that distributions stopped disproportionately more often at the edges of biomes than within them. For example, northern REH occurred mostly at the intersection between a montane (Himalayas) and a desert biome (Gobi Desert) or at the edge of the tundra in North America. Southern REH in Africa and southern Eurasia tended to stop at the edges of desert biomes (for example, Sahara Desert) but mainly at the edges of temperate and montane grasslands in the Southern Hemisphere (pampas and Andes, respectively). Notably, eastern and western REH were mostly in similar geographical locations; for example, at the intersections between the Himalayas and central Asian deserts or at the edge of the Atacama Desert in South America.

To identify the underlying niche factors that define REs, we focused only on inland REs, because REs at the edge of a water source are probably due to an obvious geographical barrier rather than an ecological effect. The fraction of inland REs was globally similar between continents (between 65% and 75%), except for Australia (~45%; Supplementary Fig. 2a). The large fraction of inland REs (Supplementary Fig. 2b) was mainly associated with the edges of biomes (for

example, northern REs in Africa, Asia and North America were associated with the Sahara Desert, montane grasslands of the Himalayas and the tundra biome, respectively; Fig. 1). A significant fraction of eastern inland REs were in Europe, but almost no significant REHs were identified, implying sparse distribution of REs throughout the continent, or an effect of a smaller area for distribution compared with other continents.

Contribution of biome–biome intersections to REHs

We quantitatively identified the global patterns of arrest by analysing the fraction of REHs that stopped at biome–biome intersections—14 central global ecological regions best distinguished by their climate, fauna and flora obtained from the World Wildlife Fund (WWF, <http://www.worldwildlife.org/>) (Fig. 1)—and by delineating a buffer zone at points of intersection between two or more biomes. Although not significant, the number of REH was strongly associated with the intersection between biomes (arrow in Fig. 2a), indicating that climatic conditions were a probable reason for the REs of tree species at biome edges. Our results, however, identified an unequal contribution of the different biome edges to the fraction of REH (Fig. 2a). We then analysed the individual biome–biome intersections normalized by a global permutation (that is, from permutation of the global distribution of REH; Methods). Here, we identified a strong contribution of REH at the intersections between temperate and desert biomes (Fig. 2b(i)) in comparison with a weaker contribution at the intersections of tropical and subtropical biomes (between themselves and with temperate biomes). The significance of the contribution of REH ($P < 0.05$) (Methods) was attributed almost exclusively to the intersections within temperate biomes (asterisks in Fig. 2b). This unequal contribution of the temperate versus tropical and subtropical biomes was consistent globally. Nevertheless, under a per-biome permutation (that is, normalizing each intersect by a selective permutation from the respective biome combinations separately; Fig. 2b(ii)), we observed a much weaker contribution of REs to the formation of biomes. Only a selected number of biome–biome intersections made a significant contribution to the formation of REH, yet no specific pattern between temperate and tropical biomes was observed. Indeed, the strong positive correlation of the number of REs between biomes with the number of REs within biomes (Fig. 2c; $R = 0.9$, linear regression, $P < 0.001$) reflects the discrepancy between the panels in Fig. 2b, because the number of REH at the edge of biomes had a strong linear association with the number of REH within that biome. However, we identified five biomes outside the 95% confidence interval (CI) of the regression (Fig. 2c). The two biomes above the regression CI (desert and boreal) were ones in which the number of REH at the edges was larger than predicted (Fig. 2c). By contrast, temperate coniferous forest, mangrove and flooded grasslands fell below the CI of the regression, indicating a much larger number of REH within the biome compared with the edge.

A parallel analysis using the distribution of REs (rather than REH) was conducted to identify similarities and differences between the distributions. The results indicated a similar pattern of distribution (Supplementary Fig. 4a,b compared with Fig. 2b,c), suggesting global forces associated with REH.

Climatic predictors of RE formation

Intersections between biomes were a substantial cause of RE formation, therefore we also investigated the dependence of RE formation on climate. We tested both the 'absolute climate' (annual and seasonal average temperature and precipitation) and the 'spatial heterogeneity of climate' (SH; spatial variability of absolute climate) at each of the global hexagons using 19 bioclimatic variables obtained from WorldClim (Methods). Elevation (absolute and SH) and latitude were also accounted for at each hexagon resolution. Generalized linear models (GLMs) of regression between each climatic variable as an independent predictor indicated that all 40 climatic variables were significant

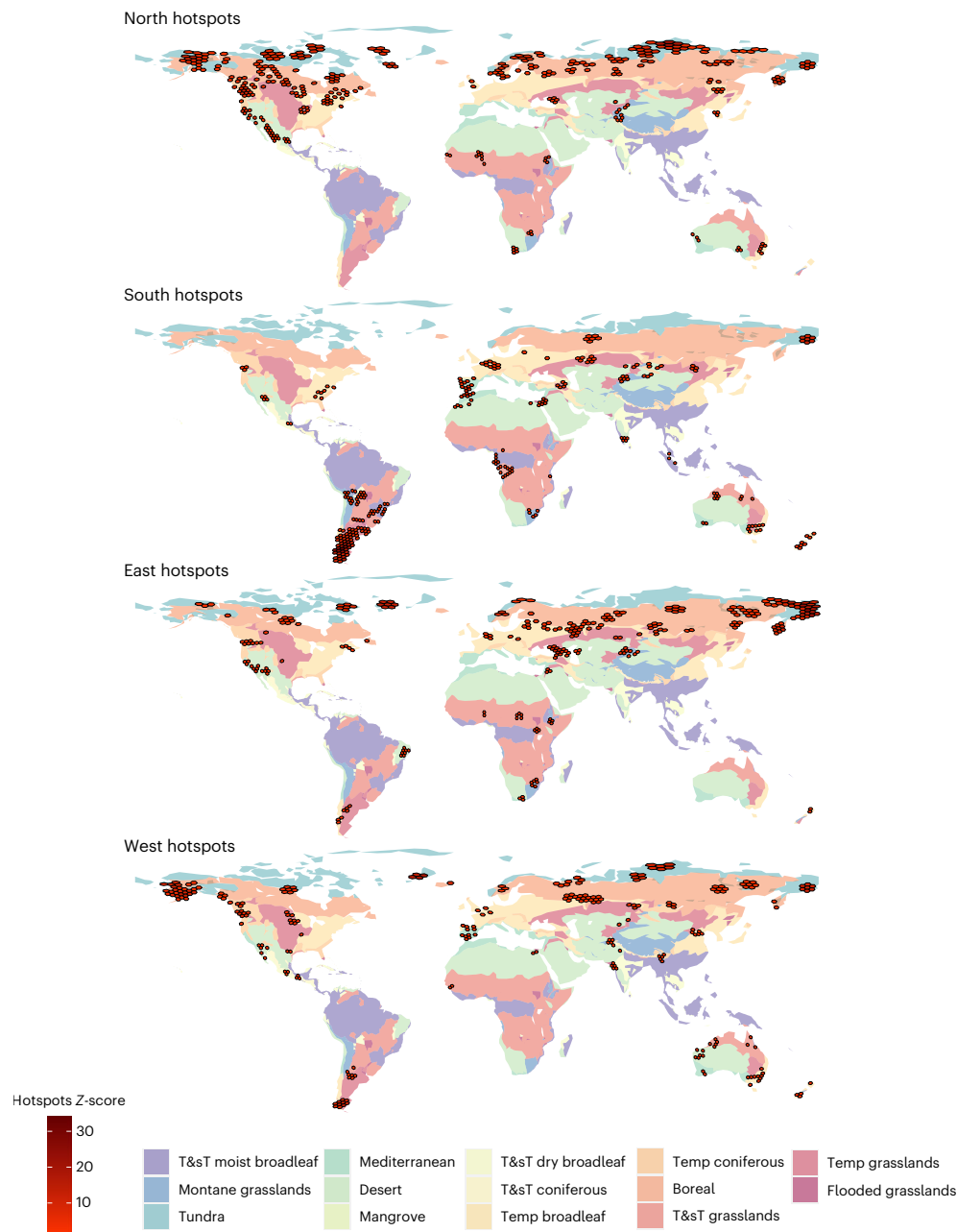


Fig. 1 | Range-edge hotspots. Hotspots were identified using Getis–Ord G_i^* analysis, which returns a Z-score for each hexagon in the world. Only hexagons with $P < 0.05$ are considered hotspots and subsequently shown here. The Z-score of each hexagon is represented by the colour gradient. Biomes (as defined by the WWF) are marked by colours. T&T, tropic and subtropic; Temp, temperate.

predictors of RE formation ($P < 0.05$) (Supplementary Fig. 5a). Interestingly, most absolute climatic variables were negatively associated with the global RE distribution (Fig. 3a,b), implying a general prediction of REs occurring in climates with low temperatures and low precipitation. A positive association was attributed to the four climatic variables that defined the temporal heterogeneity of temperature and precipitation (mean diurnal range, temperature seasonality, temperature annual range and precipitation seasonality; BioClim variables 2, 4, 7 and 15, respectively) (Fig. 3a and Supplementary Fig. 5) and all of the SH climatic variables. A mixed model with biomes and continents as random factors gave a reduced number of variables that significantly predicted the formation of RE (Fig. 3a). A crossed model with both continents and biomes consistently gave a stronger fit (Akaike information criterion (AIC) values) than only considering either of these categorical random

factors independently (Supplementary Fig. 6b). In this more stringent global analysis, latitude, absolute temperature and SH temperature, precipitation and elevation were significant predictors of REH. The goodness of fit (R^2) of each model, as a measure of predictive strength, indicated that spatial heterogeneity accounted for RE formation better than their absolute equivalents. A model selection was carried out to identify the most important factors associated with RE formation, followed by averaging of the models with a $\Delta AIC < 2$ (Fig. 3c). The SH climatic variables again defined RE better than the absolute climatic variables. Although absolute climatic variables such as isothermality (BioClim3), temperature of the wettest quarter (BioClim8) (both characteristic of tropical and subtropical climates) and annual precipitation (BioClim12) were strongly associated with REs in a GLM, temperature of the warmest quarter (BioClim10) was the only predictor strongly

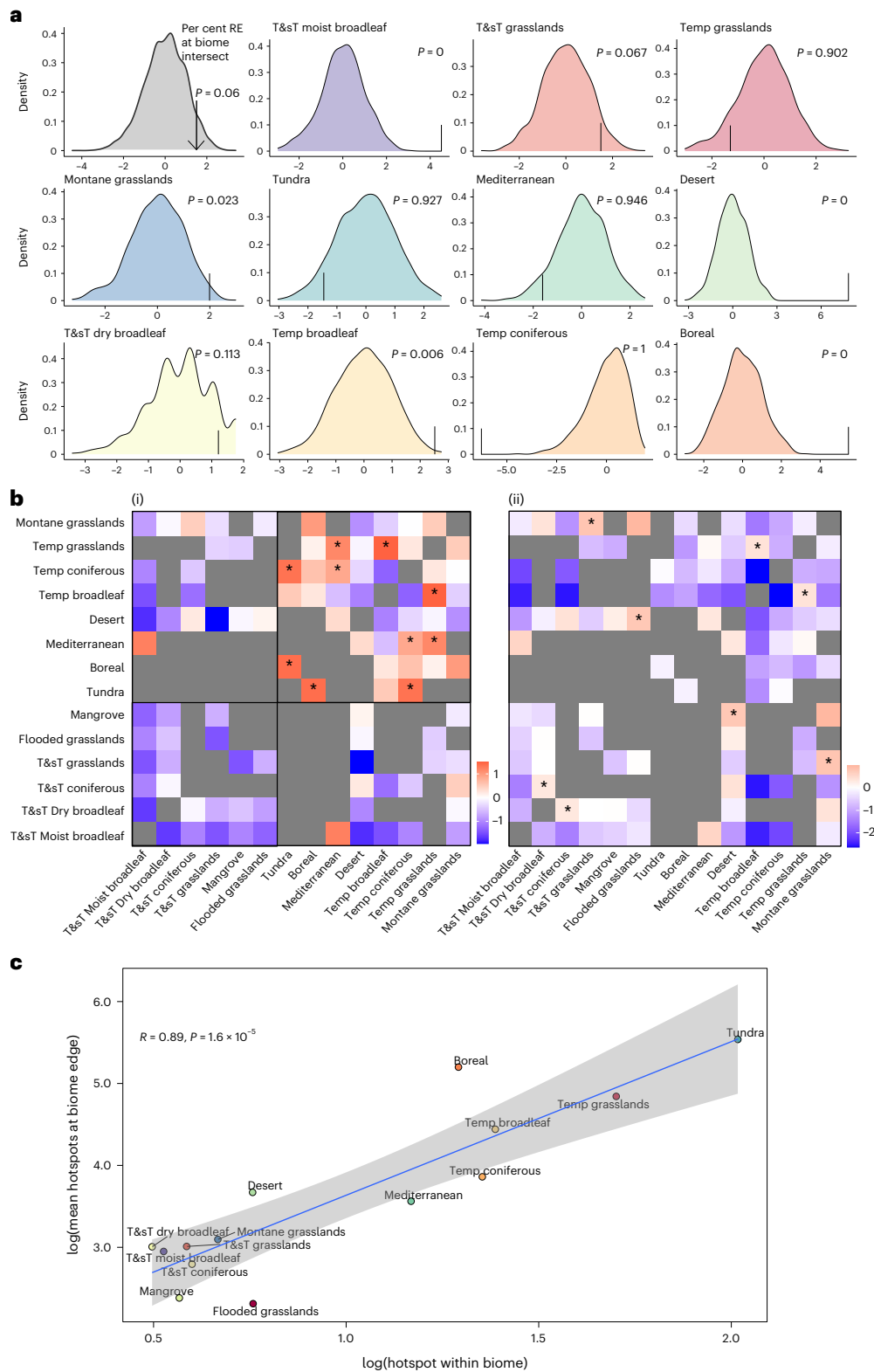


Fig. 2 | REH at intersections between biomes. a, Modelled distribution of the percentage of hotspots at the edges of biomes from a permuted (randomized) per-biome distribution of hotspots. The median value of all the biomes in the figure over a standardized Z-distribution of biomes is shown (first distribution), where the arrow marks the percentage of REHs at biome intersections in the dataset (one-sided P values). **b**, Heat maps of the percentage of REH (relative to the total number of hotspots) at the intersection between two biomes. A biome-pair intersection with a significant number of hotspots ($P < 0.08$) is marked with

an asterisk. Biome pairs that have no intersections are shown in grey. Normalized over (i) a global bootstrap and (ii) a per-biome bootstrap. **c**, Correlations (and regression lines) of the relationships between the number of REs at the edges of biomes and the number of REs within the biomes. The shaded area represents the 95% CI around the regression line. P values are calculated using a two-sided Student's t -test (degrees of freedom = 12). See Methods for further information on the methodology for obtaining significance levels (for **b**) and hotspot permutations. T&T, tropic and subtropic; Temp, temperate.

associated with RE under the mixed model (Fig. 3a and Supplementary Fig. 6). This difference in results can be seen visually when comparing between continents (Fig. 3b), for example, in Fig. 3b(iii), with temperature of the wettest quarter (BioClim8) having partial dissociations in Africa and South America; that is, the two continents with the most tropical biomes. Furthermore, SH isothermality (BioClim3), minimum temperature of the coldest month (SH BioClim6), precipitation (SH BioClim13 and BioClim16) and elevation change (SH elevation) were all predictors of REs. Interestingly, the absolute climatic predictors (BioClim8 and BioClim10) were strongly negatively correlated with spatial heterogeneity at the minimum temperature of the coldest month (SH BioClim6) (Supplementary Fig. 7b), indicating its representation of a tropical biome climate. The weaker relative importance of SH precipitation (Fig. 3c, left) is due to its dissociation with the continents that most strongly represent temperate regions: Europe and North America (Fig. 3b(iv)).

In parallel, we ran models with ENVIREM, a dataset of environmental variables complementary to WorldClim that are more ecophysiological meaningful for plant species²⁹. Most ENVIREM variables associated with REH are spatially heterogeneous (Fig. 3c and Supplementary Fig. 7). Absolute potential evapotranspiration (PET) of the coldest quarter was strongly correlated with both absolute temperature of the warmest quarter (BioClim10) and SH temperature of the coldest month (BioClim6); the two predictors are indicative of a transition between tropical and temperate biomes (Supplementary Fig. 7b). SH embergerQ and SH PET of the warmest quarter were strongly associated with SH precipitation (BioClim13 and BioClim16) and SH temperature of the warmest quarter (BioClim10), respectively. Despite the strong correlation of all absolute ENVIREM variables with absolute BioClim variables²⁹, we identified several SH ENVIREM variables to be weakly correlated with BioClim (for example, SH moisture index and PET of the driest month). All the results in Fig. 3 were robust against spatial autocorrelation (Supplementary Fig. 5b; see Methods for further information).

The REH distribution from our generated polygons was compared with a REH distribution from expert polygons (Methods) to test the accuracy of our generated dataset. Indeed, the results of GLMs converge, indicating the robustness of our generated dataset to describe REs (Supplementary Fig. 8).

Discussion

The results support our hypothesis of a non-random distribution of REs. We were able to confirm that many tree species REs were clustered rather than sparsely (stochastic) distributed by obtaining a larger number of significant REH. This finding suggests the underlying presence of ecological and evolutionary forces governing REH formation. Similarly, the matching results obtained from the biome analyses when accounting for REs (Supplementary Fig. 4) or REH (Fig. 2) were also indicative of the deterministic clustering of species REs at these specific ecological barriers (biome edges). Nevertheless, we identified

case-specific exceptions such as the scattered distribution of eastern REs throughout Europe, indicated by the strong identification of internal REs (Supplementary Fig. 2b), but not REH (Fig. 1c). This could potentially have occurred as a consequence of either intense anthropogenic activity throughout^{30,31} that prevented the distribution of tree species to reach their natural REs, or the effect of a smaller land area compared with other continents, altering the effect of how species interact with abiotic factors, thus affecting their distribution and adaptation.

Our results also indicated a strong dependence of REH on the edges of biomes, strongly supporting the many efforts to determine whether the division of the planet into discrete geographical units has been delineated appropriately^{22,32–35}. Our findings, however, were unexpectedly biome-specific, identifying the biome borders that most defined tree distributions (Fig. 2). There is no seemingly obvious pattern of differential contribution to REH when analysing each biome independently (per-biome bootstrap; Fig. 2a). However, a clear distinction between tropical/temperate distributions of REH is strongly observed under a biome–biome pairs analysis (global bootstrap; Fig. 2b). Specifically, REs were strongly dependent on desert, temperate and montane edges compared with their weak dependence on tropical and subtropical biome edges. Despite the marginally significant association of REs to biomes globally (Fig. 2a), a tropical–temperate REH distribution is not specifically dependent on biome edges, but rather represents a global biome trend (Fig. 2b(i) and 2c). In addition, the importance of biomes in describing climate's association to REs, as seen from a mixed effects model (Fig. 3a and Supplementary Fig. 6c), and the absence of latitude in predicting REs in a global model (Fig. 3c), is further indicative of the importance of biomes (rather than a latitudinal effect) in defining REs.

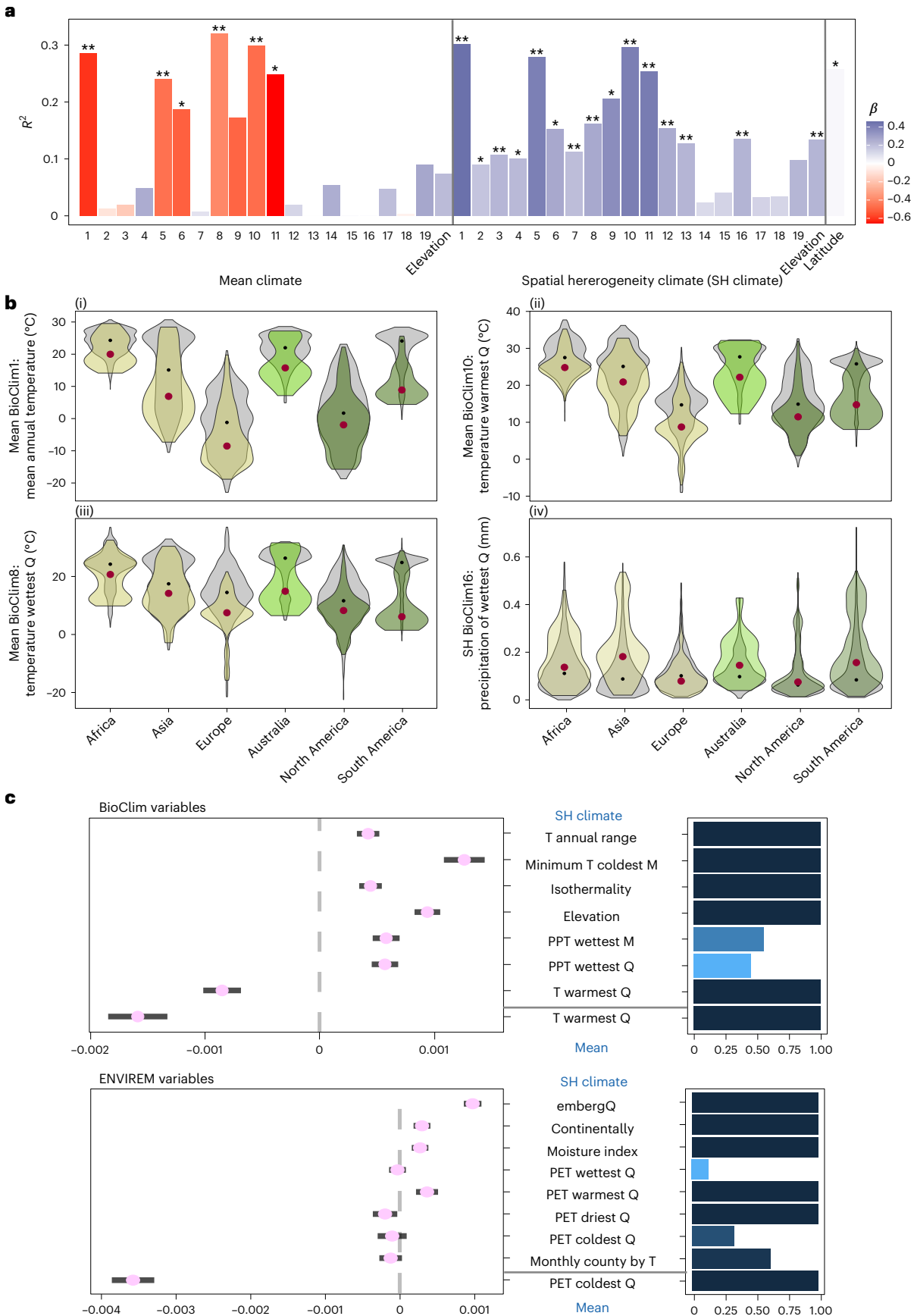
The large differences in the relative number of REs between the tropical and temperate biomes can indicate adaptive mechanisms between species residing in either of these two types of biomes. These non-trivial results may have been related to the latitude diversity gradient^{36–38}, a well-established pattern in which biodiversity is higher in the tropics than in temperate regions. The notion that these differences surged through the effective evolutionary time hypothesis^{37,39} suggests that genetic diversity is higher in the tropics owing to the possible longer times needed for adaptation and expansion. Similarly, long-term climatic oscillations have been suggested to reduce cladogenesis at higher latitudes, consistent with the observation that tropical areas are centres of evolutionary novelty^{37,40,41} and evolve faster than temperate regions³⁶, leading to the notion of tropics as centres for the radial expansion of clades and species. Recent studies have identified mechanisms for this 'out of the tropics expansion' model^{42,43} and have reported a higher fraction of bridge species (species violating niche conservatism) from the tropics compared with temperate regions. We conclude that the lower contribution to REs in tropical versus temperate regions is consistent with stronger radial expansions from the tropics than from temperate regions.

Fig. 3 | Climatic predictors of RE formation. Models of RE formation using the absolute climate (mean) and the spatial heterogeneity (PV index) of the 19 WorldClim variables and elevation. **a**, Individual (binomial) mixed regression models between each of the predictor variables and number of REs (1–19 are BioClim variables) accounting for continents and biomes as random effects. The estimated coefficients of the explanatory variables (β) are represented by the colour gradient. * $P < 0.05$, ** $P < 0.01$ (two-sided Student's t -test). BioClim1, annual mean temperature; BioClim2, mean diurnal range (mean of monthly (maximum temperature–minimum temperature)); BioClim3, isothermality; BioClim4, temperature seasonality; BioClim5, maximum temperature of the warmest month; BioClim6, minimum temperature of the coldest month; BioClim7, temperature annual range; BioClim8, mean temperature of the wettest quarter; BioClim9, mean temperature of the driest quarter; BioClim10, mean temperature of the warmest quarter; BioClim11, mean temperature of the coldest

quarter; BioClim12, annual precipitation; BioClim13, precipitation of the wettest month; BioClim14, precipitation of the driest month; BioClim15, precipitation seasonality (PV); BioClim16, precipitation of the wettest quarter; BioClim17, precipitation of the driest quarter; BioClim18, precipitation of the driest quarter; BioClim19, precipitation of the coldest quarter. **b**, Violin plots depicting the results from **a** for four predictor variables. The distribution of continental climates is shown in grey, contrasted with the climatic distribution specific to the REH (scaled to the intensity of the hotspot, that is, the Z -score). Q, quarter. **c**, Forest plot of a model average from the highest predicting LMM with both BioClim and ENVIREM variables (identified with a model selection). M, month; PPT, precipitation; T, temperature. Beta values (log-odds) are shown for each predictor. Absolute and SH climate was obtained from all of the inland global hexagonal units ($n = 5,851$). Error bars represent 95% CI around the average effect.

However, other mechanisms could explain the distinctive REH patterns between tropical and temperate regions. For example, based on several studies demonstrating the large physiological and

evolutionary effects of forest fragmentation^{30,44,45}, the increased long-term anthropogenic activity, and consequently excessive fragmentation, in forests in temperate regions compared with tropical



and subtropical regions may have hindered the adaptation of species to novel climates. The lack of association of REH specifically with temperate grasslands, temperate conifers and Mediterranean biomes (Fig. 2a) could be indicative of such an effect, because these biomes have been historically subject to strong anthropogenic activity.

Our analyses confirmed the strong dependence of RE on climate. Lower temperature, higher climatic heterogeneity (both temporal and spatial) and elevation changes were the strongest climatic predictors of REs. Analysis of the best predictive variables indicated a noticeable weak correlation between most of the leading factors determining REs (Supplementary Fig. 7b), suggesting that these factors were site-specific predictors of REs (for example, desert biome with low annual precipitation, montane grassland with high elevation, and tropical and subtropical biomes with temperature homogeneity (isothermality, temperature of the wettest quarter and spatial heterogeneity of low temperature—SH6)). We suggest that the strong negative correlation between these absolute climatic variables and the spatial heterogeneity of low temperature is an indication of different temperature patterns between tropical and temperate regions; that is, a buffered heterogeneous temperature in the tropics in contrast to the latitudinal effect of decreasing temperatures in the temperate biomes. There was an overall positive trend for the effect of precipitation (BioClim12–BioClim19) when controlling for biome because the signal appeared only when biome is included as a random factor (Supplementary Fig. 6). This means there must be other unaccounted for confounding factors altering the relationship when analysing all biomes together, suggesting that the effect is not biome dependent. Although this study does not make note of the complex interplay between biotic and abiotic forces or the ecological traits of tree species (for example, seed dispersal and phenology) in the formation of RE, we find a consistent global effect of temperature, and spatial heterogeneity of temperature and precipitation to predict their formation throughout the different models. This is indicative of their principal role as universal predictors of species distributions.

The strong prediction of REH formation from PET of the coldest quarter and its covariance to the tropical–temperate transition variables (SH BioClim6 and absolute BioClim10) could reflect a possible mechanistic evolutionary constraint for the distribution of woody species at temperate biomes. At higher latitudinal temperate biomes, where evapotranspiration is strongly reduced especially during the coldest quarter, there is a strong limitation on photosynthesis and growth caused by the significant time reduction in stomatal conductance^{46,47}. Similarly, given the strength of its prediction of REH, SH embergerQ (pluviothermic quotient) could indicate a more refined mechanism for RE formation than its covariates of precipitation (SH BioClim13 and BioClim16). This index describes mean annual precipitation in relation to annual changes in temperature. embergerQ thus increases the predictability of how precipitation also dictates the formation of REH in more temperate biomes (Fig. 3b(iv)). The consistency of our results indicates spatial and temporal heterogeneity of climate and topography as overwhelmingly stronger predictors of RE formation than their absolute climatic counterparts (Fig. 3) and, in turn, frail evidence for a universal poor habitat quality. Even in cases in which mean climatic variables strongly predict REH, these were strongly associated with this transition from spatial and temporal climatically homogenous tropical and subtropical biomes to more heterogeneous temperate biomes. Nevertheless, we note the importance of a lack of evapotranspiration, particularly in cold climates, as a main predictor of RE formation.

These observations have substantial implications for the effects of climate change on tree distributions and tree migration. Although prediction of a future steeper temperature gradient as a result of greenhouse gas emission and climate change is not trivial^{48–50}, such an increase in temperature gradients could greatly affect

the distribution of species. In particular, our results strengthen the growing understanding that the predicted poleward migration of tree species might not be as successful as previously predicted^{27,28}. The increase in stronger spatial gradients (especially at lower latitudes)^{28,49,51} or extreme and spontaneous events might all be causes of migration lags, despite the suitable temperatures at higher latitudes and altitudes. Likewise, the importance of PET from temperate biomes on the formation of REH presented here could also suggest a possible migration lag or loss of adaptation owing to the predicted reduction in PET at higher latitudes⁵². Our results thus highlight the importance of accounting for more precise spatial heterogeneity of climate as a critical feature in future models of species distribution and the development of more precise conservation efforts such as assisted migration.

Methods

Data acquisition and polygon formation

Supplementary Figure 9 visually summarizes the methodologies used to obtain global REH distributions. We downloaded a dataset of tree species from the open-source dataset Global Biodiversity Information Facility (GBIF; 5 July 2021, <https://doi.org/10.15468/dl.ajen6k>) using R packages ‘*rgbif*’ and ‘*taxize*’ and the Botanic Gardens Conservation list of 60,000 tree species. We downloaded occurrences with entries from 1980 onwards, removing any occurrence reported with a geo-spatial issue, species not belonging to the kingdom Plantae (in case of mismatched species names), and any occurrence marked as unlikely, mismatched or invalid. We removed occurrences that had reported uncertainties of >100 km and records based on fossils and unknown sources. We then used the R package ‘*CoordinateCleaner*’⁵³ to remove any occurrences with zero coordinates, equal *x* and *y* coordinates, duplicates, occurrences at sea, coordinates at capitals and centroids. To finalize, we again removed species with fewer than 300 occurrences. The data we used undoubtedly contained sampling bias⁵⁴, probably overrepresenting the number of REs in some regions with a reduced or negligible sampling effort. We tried to overcome this issue by basing our filtering steps on several previous studies^{6,53,54}. The strength of the critical filtering steps applied in our analysis resembled those presented previously⁵⁴.

We converted the georeferenced species occurrences (*x* and *y* coordinates) into distributional polygons in parallel using two independent techniques: through concave-hull (Supplementary Methods) and multivariate kernel-density estimation (described below). Given the strong similarity between the two methods (Supplementary Fig. 8), we discuss the methods and results in detail only for the kernel-density estimated polygons. We created polygons using two-dimensional kernel-density estimations. We first divided the extent of all the coordinates into 800 grid points in each dimension (longitude and latitude) to produce a matrix of 640,000 grid cells for each species. Subsequently, we selected for the grid cells with the highest 99% estimation of the species’ occurrence and rasterized these. Polygons were then delineated around the contour of the rasters.

Polygon groupings

All polygons belonging to the same species were grouped based on their absolute distance from one another. Polygons separated by ≤500 km were grouped together, with the assumption that fragmentation, gene flow and unreported data could all warrant two nearby populations being considered as one. We used the *hclust* function (package ‘*stats*’, agglomeration method: complete) to hierarchically cluster populations from a sequence of three or more populations by their distances, also using a cut-off of 500 km (*cutree* function, package ‘*stats*’) for determining the clusters. The final dataset comprised >3,600 tree species, ranging from one to nine populations (polygons) per species, for a total of 8,500 populations. All spatial data were analysed using R packages ‘*sf*’ and ‘*raster*’.

RE determination

RE-dense areas were determined by defining distinct global units, identifying the RE of each species and mapping species RE to the global units to calculate the density of REs per unit. In detail, first we rasterized the world map to spatially bin the density of REs. To overcome the problem of spatial distortion, we used hexagonal bins with the R package 'dgggridR', developed using the ISEA Discrete Global Grids system, a repetition of polygons on the surface of an icosahedron, allowing for the projection of equal-sized bins on a two-dimensional plane. We defined the size of each hexagon as ~23,000 km² (with an average spacing between centre nodes of 165 km). Second, we used coordinates of the cardinal directions (north, south, east and west) to represent species REs by subdividing each polygon cluster into four quartiles in the four cardinal directions (northeast, northwest, southeast and southwest). The REs for each quartile were determined as the two most-outward coordinates of the corresponding cardinal directions (for example, north and east cardinal coordinates for the northeast quartile). Eight REs were thus determined for each population (two for each cardinal direction). As a filtering step, we accounted for both REs from the same cardinal direction if they were >20 arc degrees apart, otherwise we accounted only for the farthest point from the centroid. Third, the total number of REs obtained using this method was normalized by the total number of species intersecting its respective hexagon. In parallel, we also defined REs by accounting for the perimeter of the polygon for each species (Supplementary Fig. 10). The 'perimeter' system may be a more realistic and complete system for identifying REs, but the 'cardinal coordinate' system, although more simplistic in nature, provides a clearer visual representation of the distribution of REs, and allows for the directionality of REs to be compared, essential farther along the pipeline by distinguishing between coastline and inland REs and identifying hotspots in the four cardinal directions.

We then classified coastline and inland REs, assuming that the arrest of species distribution at the edge of a water source was probably due to an obvious geographical barrier rather than an ecological effect. Coastline REs were determined by creating a semicircle (buffer of 3 arc degrees) around each RE in the direction of its cardinal coordinate and measuring the percentage overlap with water. Cardinal coordinates with >50% overlap were considered a 'coastal RE'.

To find the probability distribution of REs at the edge of a water source (Supplementary Fig. 2), we permuted the global population of REs (except for the Australian population) and used the mean of this permutation to compare with the number of REs in each of the continents.

Hotspot analysis

Hotspots were identified using Getis–Ord Gi* hotspot analysis^{55,56} to find spatial correlations between hexagon (inland and normalized) RE densities. We initially compiled a list of neighbours between all hexagons using the poly2nb function and then obtained the local G statistic using the weighted density (normalized number of REs) of the global hexagons and their relative distance from each other. The G statistic calculates a Z-score (measure of standard deviation) for each hexagon. P values were then determined using the critical Z-scores at 95% CI followed by a Bonferroni correction with the p.adjustSP function (using the number of neighbours between hexagons rather than the total number of hexagons). All analyses were carried out using the R package 'spdep'⁵⁷.

We compared the analysis from the linear models with expert-based polygons from three different sources: International Union for Conservation of Nature, Botanical Information and Ecological Network and European Forest Genetic Resources Programme. A randomized weighted sample of all of this dataset was used to generate a global distribution of REH by running this sample through our pipeline. GLMs were run on the global distribution of expert-based REH in the same way as with our generated polygons. Given the uneven

distribution of expert-based polygons globally (Supplementary Fig. 8a), we ran models excluding Asia and Africa to account for this bias. As seen by the strong similarity between the GLMs of expert polygons in a global and filtered model (Supplementary Fig. 8b,c(iii)), we observed an overrepresentation of the expert polygons for these continents. Likewise, the GLMs from our generated polygons are very similar to those obtained from the filtered model. In this case, practically all variables showed the same relationship with REs (either positive or negative β values) as well as similar magnitudes.

Statistical analyses

Contribution of biome edge to RE. We used the 14 biomes defined by the WWF for our analyses. The distributions were downloaded from the WWF webpage (<http://www.worldwildlife.org/>). To identify the intersection between biomes, we reduced the complexity of the polygon edge using the package 'rmapshaper', which can perform topologically aware polygon simplifications, thus maintaining the intersection between biomes upon reduction of the 'edginess' of the polygons. The intersections were delineated and subsequently enlarged (with a buffer distance of 0.1 arcminute, ~185 m at the equator). A global permutation assay was carried out by randomizing (1,000 iterations) the global distribution of hexagons with REH (absolute Z-score), and a per-biome stratified permutation was carried out by randomizing the distribution of hexagons with REH within each biome independently.

The averaged global bootstrap shown in Fig. 2a was calculated using a per-biome bootstrap to obtain the probability distribution of REH at biome edges (Supplementary Fig. 3a). Distributions that were not normally distributed as a result of their small size (flooded grassland, mangrove and tropical and subtropical coniferous forest; Supplementary Fig. 3b) were removed from the analysis. A general trend for the probability of REs falling at the intersection of biomes was therefore measured as a unified standardized Z-distribution, and compared with the median Z-score from the actual per cent overlap for each biome.

The density of hotspots at the intersection between biomes was calculated using the sum of Z-scores of the hotspots at that intersection, and the percentage contribution was then calculated using this value over the total Z-score at all biome intersections. The global or per-biome 1,000 permutation means were used as a normalizing denominator for the values obtained from our dataset (Fig. 2b and Supplementary Fig. 4a). The denominator could be either larger or smaller than the numerator, so we log-transformed the outcome to obtain a linear-like relationship. Contribution within a biome was calculated in the same way as for the contribution at the edge, using the mean from a permuted assay to normalize for the absolute value.

Climatic dependency of RE. We tested the relationship between RE density and climatic features by assigning a set of environmental variables to each hexagon. We used the bioclimatic attributes downloaded from WorldClim Global Climate Data⁵⁸ at a resolution of 5 arcminutes. The 19 BioClim variables and elevations for each hexagon were extracted using the R package 'raster'. We also used the 16 ENVIREM variables described by Title and Bemmels²⁹, downloaded from their website, at a resolution of 2.5 arcminutes. Absolute climate for each hexagon was obtained using the mean over all pixels. Climatic SH was calculated using the proportional variability (PV) index^{59,60} over all pixels in each hexagon.

We used an array of linear mixed models (LMM; Fig. 3 and Supplementary Figs. 5 and 6) to test for the dependence of REs on climate and the robustness of the results. LMM were carried out to account for biomes and continents. Both variables were introduced as random effects in random intercept models. A model selection analysis was used to determine the models that best predicted the formation of REs. Random intercept models using both continent and biome as random variables were run, and models with $\Delta AIC < 2$ were selected for.

Given the strong correlation between different predictor variables, we ran only models with variable combinations that had a Pearson's correlation value of $r < 0.7$. The relative contribution of the variable included in the model was calculated from the selected models. Analyses were run using R package 'MuMin'⁶¹.

All statistical analyses (individual GLMs and multiple-predictor GLMs) were tested for their robustness to spatial autocorrelation by creating a spatial autocovariate (function `autocov_dist` function, package 'spdep'), calculated as the distance-weighted average of neighbouring dependent variables⁶², so hexagons in proximity were averaged and those farther away received a lower weighting average. We set the predetermined distance to 200 km based on the average distance between cells. The spatial autocovariate was then included in the regression model as a dependent variable.

Reporting summary

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

Data availability

The occurrence points used from GBIF can be found on the GBIF webpage (<https://doi.org/10.15468/dl.ajen6k>). Polygons generated from occurrence points are provided in the public Zenodo repository 10.5281/zenodo.7613535. Biome polygons were obtained from the WWF webpage (<http://www.worldwildlife.org/>). Bioclimatic attributes were downloaded from WorldClim Global Climate Data⁵⁸. ENVIREM variables were downloaded from their webpage (<https://envirem.github.io/>).

Code availability

Custom codes related to this paper can be found in a GitHub repository at <https://github.com/dlerner/Global-Range-edges>

References

- Soule, M. The epistasis cycle: a theory of marginal populations. *Annu. Rev. Ecol. Syst.* **4**, 165–187 (1973).
- Brown, J. H. On the relationship between abundance and distribution of species. *Am. Nat.* **124**, 255–279 (1984).
- Gaston, K. J. *The Structure and Dynamics of Geographic Ranges* (Oxford Univ. Press, 2003).
- Sexton, J. P., McIntyre, P. J., Angert, A. L. & Rice, K. J. Evolution and ecology of species range limits. *Annu. Rev. Ecol. Syst.* **40**, 415–436 (2009).
- Gaston, K. J. Geographic range limits: achieving synthesis. *Proc. R. Soc. B* **276**, 1395–1406 (2009).
- Zizka, A. et al. No one-size-fits-all solution to clean GBIF. *PeerJ* **8**, e9916 (2020).
- Goldberg, E. E. & Lande, R. Species' borders and dispersal barriers. *Am. Nat.* **170**, 297–304 (2007).
- Bachmann, J. C., Rensburg, A. J. V., Cortazar-Chinarro, M., Laurila, A. & Buskirk, J. V. Gene flow limits adaptation along steep environmental gradients. *Am. Nat.* **195**, E67–E86 (2020).
- Hargreaves, A. L., Samis, K. E. & Eckert, C. G. Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. *Am. Nat.* **183**, 157–173 (2014).
- Henry, R. C., Bartoń, K. A. & Travis, J. M. J. Mutation accumulation and the formation of range limits. *Biol. Lett.* **11**, 20140871 (2015).
- Perrier, A., Sánchez-Castro, D. & Willi, Y. Environment dependence of the expression of mutational load and species' range limits. *J. Evol. Biol.* **35**, 731–741 (2022).
- Bontrager, M. et al. Adaptation across geographic ranges is consistent with strong selection in marginal climates and legacies of range expansion. *Evolution* **75**, 1316–1333 (2021).
- Santini, L., Pironon, S., Maiorano, L. & Thuiller, W. Addressing common pitfalls does not provide more support to geographical and ecological abundant-centre hypotheses. *Ecography* **42**, 696–705 (2019).
- Oldfather, M. F., Kling, M. M., Sheth, S. N., Emery, N. C. & Ackerly, D. D. Range edges in heterogeneous landscapes: integrating geographic scale and climate complexity into range dynamics. *Glob. Chang. Biol.* **26**, 1055–1067 (2020).
- Janzen, D. H. Why mountain passes are higher in the tropics. *Am. Nat.* **101**, 233–249 (1967).
- Maxwell, M. F., Leprieux, F., Quimbayo, J. P., Floeter, S. R. & Bender, M. G. Global patterns and drivers of beta diversity facets of reef fish faunas. *J. Biogeogr.* **49**, 954–967 (2022).
- Roy, K., Hunt, G., Jablonski, D., Krug, A. Z. & Valentine, J. W. A macroevolutionary perspective on species range limits. *Proc. R. Soc. B* **276**, 1485–1493 (2009).
- Loiseau, N. et al. Global distribution and conservation status of ecologically rare mammal and bird species. *Nat. Commun.* **11**, 5071 (2020).
- Kerkhoff, A. J., Moriarty, P. E. & Weiser, M. D. The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. *Proc. Natl Acad. Sci. USA* **111**, 8125–8130 (2014).
- Donoghue, M. J. & Edwards, E. J. Biome shifts and niche evolution in plants. *Annu. Rev. Ecol. Syst.* **45**, 547–572 (2014).
- Ringelberg, J. J., Zimmermann, N. E., Weeks, A., Lavin, M. & Hughes, C. E. Biomes as evolutionary arenas: convergence and conservatism in the trans-continental succulent biome. *Glob. Ecol. Biogeogr.* **29**, 1100–1113 (2020).
- Smith, J. R. et al. A global test of ecoregions. *Nat. Ecol. Evol.* **2**, 1889–1896 (2018).
- Paquette, A. & Messier, C. The effect of biodiversity on tree productivity: from temperate to boreal forests. *Glob. Ecol. Biogeogr.* **20**, 170–180 (2011).
- Pichancourt, J. B., Firn, J., Chadès, I. & Martin, T. G. Growing biodiverse carbon-rich forests. *Glob. Chang. Biol.* **20**, 382–393 (2014).
- Pennington, R. T., Lavin, M. & Oliveira-Filho, A. Woody plant diversity, evolution, and ecology in the tropics: perspectives from seasonally dry tropical forests. *Annu. Rev. Ecol. Syst.* **40**, 437–457 (2009).
- Zhu, K., Woodall, C. W. & Clark, J. S. Failure to migrate: lack of tree range expansion in response to climate change. *Glob. Chang. Biol.* **18**, 1042–1052 (2012).
- Corlett, R. T. & Westcott, D. A. Will plant movements keep up with climate change? *Trends Ecol. Evol.* **28**, 482–488 (2013).
- la Sorte, F. A., Butchart, S. H. M., Jetz, W. & Böhning-Gaese, K. Range-wide latitudinal and elevational temperature gradients for the world's terrestrial birds: implications under global climate change. *PLoS One* **9**, e98361 (2014).
- Title, P. O. & Bemmels, J. B. ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography* **41**, 291–307 (2018).
- Veresoglou, S. D. & Peñuelas, J. Variance in biomass-allocation fractions is explained by distribution in European trees. *New Phytol.* **222**, 1352–1363 (2019).
- Grantham, H. S. et al. Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. *Nat. Commun.* **11**, 5978 (2020).
- Holdridge, L. R. Determination of world plant formations from simple climatic data. *Science* **105**, 367–368 (1947).
- Whittaker, R. H. Classification of natural communities. *Bot. Rev.* **28**, 1–239 (1962).
- McDonald, R. et al. Species compositional similarity and ecoregions: do ecoregion boundaries represent zones of high species turnover? *Biol. Conserv.* **126**, 24–40 (2005).

35. von Humboldt, A. & Bonpland, A. *Essay on the Geography of Plants* (Univ. Chicago Press, 2013).
36. Cardillo, M. Latitude and rates of diversification in birds and butterflies. *Proc. R. Soc. Lond. B* **266**, 1221–1225 (1999).
37. Hillebrand, H. On the generality of the latitudinal diversity gradient. *Am. Nat.* **163**, 192–211 (2004).
38. Mittelbach, G. G. et al. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.* **10**, 315–331 (2007).
39. Hewitt, G. M. Genetic consequences of climatic oscillations in the Quaternary. *Phil. Trans. R. Soc. Lond. B* **359**, 183–195 (2004).
40. Crane, P. & Scott, L. Angiosperm diversification and paleolatitudinal gradients in Cretaceous floristic diversity. *Science* **246**, 675–678 (1989).
41. Jablonski, D. The tropics as a source of evolutionary novelty through geological time. *Nature* **364**, 142–144 (1993).
42. Jablonski, D. et al. Out of the tropics, but how? Fossils, bridge species, and thermal ranges in the dynamics of the marine latitudinal diversity gradient. *Proc. Natl Acad. Sci. USA* **110**, 10487–10494 (2013).
43. Antonelli, A. et al. An engine for global plant diversity: highest evolutionary turnover and emigration in the American tropics. *Front. Genet.* **6**, 130 (2015).
44. Jump, A. S. & Peñuelas, J. Running to stand still: adaptation and the response of plants to rapid climate change. *Ecol. Lett.* **8**, 1010–1020 (2005).
45. Morreale, L. L., Thompson, J. R., Tang, X., Reinmann, A. B. & Hutrya, L. R. Elevated growth and biomass along temperate forest edges. *Nat. Commun.* **12**, 7181 (2021).
46. Wilkinson, S., Clephan, A. L. & Davies, W. J. Rapid low temperature-induced stomatal closure occurs in cold-tolerant *Commelina communis* but not in cold-sensitive tobacco leaves, via a mechanism that involves apoplastic calcium but not abscisic acid. *Plant Physiol.* **126**, 1566–1578 (2001).
47. Brodribb, T. J. & Holbrook, N. M. Stomatal protection against hydraulic failure: a comparison of coexisting ferns and angiosperms. *New Phytol.* **162**, 663–670 (2004).
48. Davis, B. A. S. & Brewer, S. Orbital forcing and role of the latitudinal insolation/temperature gradient. *Clim. Dyn.* **32**, 143–165 (2009).
49. Seager, R. et al. Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. *Nat. Clim. Change* **9**, 517–522 (2019).
50. Xu, Y. & Ramanathan, V. Latitudinally asymmetric response of global surface temperature: implications for regional climate change. *Geophys. Res. Lett.* **39**, L13706 (2012).
51. Colwell, R. K., Brehm, G., Cardelús, C. L., Gilman, A. C. & Longino, J. T. Global warming, elevational range shifts, and lowland biotic attrition in the wet tropics. *Science* **322**, 258–261 (2008).
52. Basso, B., Martínez-Feria, R. A., Rill, L. & Ritchie, J. T. Contrasting long-term temperature trends reveal minor changes in projected potential evapotranspiration in the US Midwest. *Nat. Commun.* **12**, 1476 (2021).
53. Zizka, A. et al. CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods Ecol. Evol.* **10**, 744–751 (2019).
54. Serra-Diaz, J. M., Enquist, B. J., Maitner, B., Merow, C. & Svenning, J. Big data of tree species distributions: how big and how good? *For. Ecosyst.* **4**, 30 (2017).
55. Getis, A. & Ord, J. K. The analysis of spatial association by use of distance statistics. *Geogr. Anal.* **24**, 189–206 (1992).
56. Mendez, C. Spatial autocorrelation analysis in R. R Studio/RPubs. <https://rpubs.com/quarcs-lab/spatial-autocorrelation> (2020).
57. Bivand, R. S., Pebesma, E. & Gómez-Rubio, V. *Applied Spatial Data Analysis with R* (Springer, 2013).
58. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
59. Heath, J. P. Quantifying temporal variability in population abundances. *Oikos* **115**, 573–581 (2006).
60. Fernández-Martínez, M. et al. The consecutive disparity index, D: a measure of temporal variability in ecological studies. *Ecosphere* **9**, e02527 (2018).
61. Bartoń, K. MuMIn: multi-model inference. R package v.1.10.1. (2013).
62. F. Dormann, C. et al. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**, 609–628 (2007).

Acknowledgements

T.K. thanks the Edith and Nathan Goldenberg Career Development Chair; Mary and Tom Beck-Canadian Center for Alternative Energy Research; Larson Charitable Foundation New Scientist Fund; Yotam Project; Dana and Yossie Hollander; the Estate of Emile Mimran; and the Estate of Helen Nichunsky. D.L. was supported by the Sustainability and Energy Research Initiative PhD grant. M.F.M. and J.P. were supported by grants PID2019-110521GB-I00 and TED2021-132627B-I00 funded by Spain, Ministry for Science and Innovation (MCIN), AEI/10.13039/501100011033 and the NextGeneration EU/PRTR (Recovery, Transformation and Resilience Plan). M.F.M. was supported by a postdoctoral fellowship from 'la Caixa' Foundation (ID 100010434), code: LCF/BQ/PI21/11830010.

Author contributions

D.L., J.B. and T.K. designed the research. D.L. performed the research and analysed the data. M.F.M., J.P., T.K., J.B. and S.L.-L. provided scientific advice. M.F.M. and J.B. advised on statistical aspects. D.L. wrote the paper with special contributions from J.P., J.B., M.F.M. and T.K.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41477-023-01369-1>.

Correspondence and requests for materials should be addressed to David Lerner.

Peer review information *Nature Plants* thanks Giacomo Puglielli and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

Reporting Summary

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

Statistics

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

n/a | Confirmed

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

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Data collection was done through the free software R, version 4.1.2. Data on climate conditions was collected from high-resolution global raster layers using R, mainly with packages 'raster' (v3.5-2) and 'sf' (v1.0-3). Data on species occurrences was downloaded using 'rgbif' (v3.7.5) and 'taxize' (v0.9.100) packages. Complexity reduction of biomes was carried out using rmapshaper package (v0.4.5).

Data analysis Data analysis was done through the free software R, version 4.1.2. Data was cleaned using packages 'coordinate cleaner' (v2.0-2). All spatial data analyses were done using packages raster (v3.5-2), sf (v1.0-3), spdep (v1.1-12), ddgridR (v2.0.3). Statistical analyses were carried out using the packages MuMin (v1.43.17), stats (v4.1.2). Custom codes related to this paper can be found in a GitHub repository at <https://github.com/dlerner/Global-Range-edges>

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

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

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

The occurrence points used from GBIF can be found in the GBIF webpage (<https://doi.org/10.15468/dl.ajen6k>). Polygons generated from occurrence points are provided in the public Zenodo repository 10.5281/zenodo.7613535. Biome polygons were obtained from the WWF webpage (<http://www.worldwildlife.org/>). Bioclimatic attributes were downloaded from WorldClim Global Climate Data58. ENVIREM variables were downloaded from their webpage (<https://envirem.github.io/>).

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	<input type="text" value="N/A"/>
Population characteristics	<input type="text" value="N/A"/>
Recruitment	<input type="text" value="N/A"/>
Ethics oversight	<input type="text" value="N/A"/>

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

Field-specific reporting

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

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

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

Ecological, evolutionary & environmental sciences study design

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

Study description	This study generates the global map of tree range edge and tree range edge hotspots and describes the patters of distribution of such, as well as the main climatic attributes affecting their distribution using several statistical tests (LMM, model averaging, bootstrap). All models are generalized linear (mixed) models with fixed and random effects.
Research sample	Data collection was performed using all available spatial data from GBIF for woody tree species. DOI: https://doi.org/10.15468/dl.ajen6k
Sampling strategy	All tree species were collected. A final sample was used using a stringent filtering step described in the Methods section and Supplementary Materials
Data collection	Spatial data for the occurrence of tree species was downloaded from GBIF. The eco-climatic attributes from WorldClim and ENVIREM
Timing and spatial scale	Data from GBIF was downloaded on 05/072021. The original data was collected for occurrences from 1980 until present day at a global scale
Data exclusions	We omitted occurrence points as described in the Methods and Supplementary sections and omitted species that had >500 occurrence points
Reproducibility	This study uses open-source datasets. The initial raw data is easily accessible using the relevant DOI, as well as the final dataset produced (polygons). Together with the codes used for the data extraxtion and analysis, this study is therefore fully reproducible.
Randomization	Given the observational and global nature of the study, there is no randomization of data acquisition.
Blinding	Blinding is not relevant in this study

Did the study involve field work? Yes No

Reporting for specific materials, systems and methods

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

Materials & experimental systems

- | n/a | Included in the study |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

- | n/a | Included in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |