Erschienen in: Nature Ecology & Evolution. Springer. 2023, 7, pp. 1633-1644
https://doi.org/10.1038/s41559-023-02172-z

Evolutionary imbalance, climate and human history jointly shape the global biogeography of alien plants

Trevor S. Fristoe[1]✉, Jonas Bleilevens[1,2], Nicole L. Kinlock[1], Qiang Yang[1,3], Zhijie Zhang[1], Wayne Dawson[4], Franz Essl[5], Holger Kreft[6,7], Jan Pergl[8], Petr Pyšek[8,9], Patrick Weigelt[6,7], Jean-Marc Dufour-Dror[14], Alexander N. Sennikov[10], Pawel Wasowicz[11], Kristine B. Westergaard[12] & Mark van Kleunen[1,13]

Human activities are causing global biotic redistribution, translocating species and providing them with opportunities to establish populations beyond their native ranges. Species originating from certain global regions, however, are disproportionately represented among naturalized aliens. The evolutionary imbalance hypothesis posits that differences in absolute fitness among biogeographic divisions determine outcomes when biotas mix. Here, we compile data from native and alien distributions for nearly the entire global seed plant flora and find that biogeographic conditions predicted to drive evolutionary imbalance act alongside climate and anthropogenic factors to shape flows of successful aliens among regional biotas. Successful aliens tend to originate from large, biodiverse regions that support abundant populations and where species evolve against a diverse backdrop of competitors and enemies. We also reveal that these same native distribution characteristics are shared among the plants that humans select for cultivation and economic use. In addition to influencing species' innate potentials as invaders, we therefore suggest that evolutionary imbalance shapes plants' relationships with humans, impacting which species are translocated beyond their native distributions.

Geographic imbalances in the origins of alien species that successfully establish populations beyond their native distributions (henceforth naturalize[1]) have long been recognized[2]. Asymmetric patterns of exchange are becoming increasingly clear[3–5] as species translocations continue unabated and information on alien distributions accumulates[6,7]. Despite such long-standing interest, the mechanisms driving the imbalanced biogeography of invasions have yet to be disentangled at the global scale. Recent work has shown that disparities in naturalization among plants originating from different geographic regions persist when accounting for introduction biases[4], suggesting a role of eco-evolutionary context in addition to anthropogenic factors. Indeed, similar imbalances have characterized many natural biotic interchanges throughout Earth's history[2,8,9]. To explain such patterns, Charles Darwin suggested that geographic barriers (for example, oceans for terrestrial species), in effect, divide the Earth's biota among various evolutionary arenas. The unique biogeographic conditions experienced by each regional biota influence the intensity of natural selection, resulting in differences in absolute fitness and, consequently, species' capacities to establish in new regions when barriers are removed[2].

A full list of affiliations appears at the end of the paper. ✉e-mail: trevor.fristoe@uni-konstanz.de

More recently, Darwin's ideas were formalized as the evolutionary imbalance hypothesis (EIH)[10]. The EIH posits that invasion potential should be highest for species originating from: (1) large regions able to support abundant populations with high genetic diversity that can facilitate effective selection and adaptation[11,12]; (2) regions of high biodiversity where lineages must compete and otherwise interact with many others to persist[13,14]; and (3) environmentally stable regions where selection is given ample time to optimize species for a given set of conditions[15,16]. The biogeographic characteristics proposed by these three mechanisms show considerable variation among global regions and have been invoked to explain patterns of species exchange during anthropogenic as well as natural episodes[2,8,9,17]. Despite their potential importance in mediating dynamics in biogeography, the tenets of the EIH have so far been assessed only for alien introductions among a limited set of regional biotas[10,18]. To fill this gap, we compiled data on native and alien distributions for nearly the entire global seed plant flora (336,245 species; ~99.5%; ref. 19) and quantified exchanges of naturalized species among biogeographic regions worldwide. Measuring conditions within species' native distributions, we then tested whether observed patterns of exchange are consistent with the three mechanisms proposed by the EIH.

Evolutionary imbalance is one among many potentially co-acting drivers of biotic invasions. Climate, for example, serves as a primary filter in determining which regions will be suitable for naturalization[20,21]. Relationships between plants and humans also shape patterns of exchange by determining which species are introduced into new regions[4,22,23]. It has also been suggested that species originating from regions with histories of intensive human land use are better adapted for persistence under anthropogenic disturbance or in semi-natural and artificial habitats[24], which may facilitate naturalization[25]. Accordingly, in our assessments of evolutionary imbalance mechanisms, we additionally accounted for the influence of climatic and anthropogenic factors on naturalization. Beyond providing a detailed characterization of the biogeography of alien plants and a global test of the EIH, our analyses therefore provide a globally comprehensive assessment of previously identified drivers of plant naturalization.

Results

Geographic imbalance in alien origins

For categorizing geographic origins, we used a simple yet novel approach to identify common patterns of plant distribution according to the continents where species occur as native. Specifically, we followed the continental definitions outlined by the World Geographic Scheme for Recording Plant Distributions[26] and classified species based on the combination of continents overlapped by their native distribution. Each unique distributional pattern, subsequently referred to as biogeographic syndrome, was therefore defined by a specific combination of native continents (Fig. 1a and Supplementary Fig. 1). For example, all species endemic to North America were assigned to a common biogeographic syndrome, while a separate biogeographic syndrome included all species with native distributions spanning both North and South America. We found that the 19 most common biogeographic syndromes described the distributions for over 99% of the species in our dataset (Supplementary Fig. 1; see also 'Defining regions of origin' in Methods) and focus on these in subsequent analyses. Using data from the Global Naturalized Alien Flora database (GloNAF)[7], we quantified naturalization success of the native flora classified to each biogeographic syndrome as the $\log_{10}$ ratio of the observed proportion of species that have naturalized outside their native distribution to the expected proportion of species. At the global scale, the expected value was based on a scenario where the probability of naturalization is equal for species regardless of biogeographic syndrome (that is, no biases in the geographic origins of naturalized aliens), and therefore equal to the proportion of naturalized species in the full global dataset (~4.3%; see 'Quantifying naturalization success' in Methods for further details on this metric). While the consequences of evolutionary imbalance are often framed in terms of fitness differences between native and potential alien species[10], we use our metric of naturalization success to compare evolutionary imbalance among potential donor regions. All else being equal, we expect aliens originating from large, biodiverse and stable native regions to be more likely to overcome biotic resistance and establish naturalized populations.

Global naturalization success varied greatly among biogeographic syndromes (Fig. 1a), with a clear tendency towards higher success for species from multi-continental syndromes (59% of species with native distributions spanning three or more continents were naturalized). Continental endemic syndromes, in contrast, have donated relatively few naturalized species globally (only 2% of species native to a single continent were naturalized). Our results also show that biogeographic syndromes defined by combinations of the 'Old World' continents of Africa, Europe, and temperate and tropical Asia have been particularly successful at donating naturalized alien species. For example, nearly 90% of the almost 500 species with distributions spanning all four of these continents have become naturalized. While previous studies have suggested that Europe and temperate Asia have been the largest donors of successful alien plants[3], our assessment of biogeographic syndromes clarifies that the apparent success of these continents is driven by species that span the interconnected Afro-Eurasian landmass.

Biogeographic drivers of naturalization

The greater naturalization success of species with multi-continental biogeographic syndromes suggests provisional support for the geographic area mechanism of the EIH. Nevertheless, to explicitly test whether the patterns in Fig. 1a are consistent with predictions based on evolutionary imbalance, we measured key characteristics of the native distributions of the species classified to each biogeographic syndrome. To correspond with our metrics of naturalization success, we measured distributional characteristics first at the species level and then aggregated to the scale of biogeographic syndromes (Methods). In line with long-standing interest in understanding geographic imbalance in the origins of naturalized plants dating back to Darwin[2], this approach allowed us to determine where successful aliens come from, as well as the biogeographic features of these regions. One limitation, however, is that aggregation ignores the sometimes substantial differences in distributions among species assigned to the same biogeographic syndrome. We therefore performed an additional set of complementary analyses of naturalization success at the species level (see 'Evolutionary imbalance and human use'). Data on the species' native distributions were compiled across 367 global regions that are commonly used to define plant distributions—the finest resolution available for such a complete set of species ($n = 336{,}245$)[26–30]. Native-range characteristics included estimates of the geographic area of all regions comprising the native range (EIH mechanism 1), as well as measures of the total and area-corrected (γ-diversity and α-diversity, respectively) richness and phylogenetic diversity of co-occurring species (EIH mechanism 2). To capture environmental stability within species' native distributions (EIH mechanism 3), we measured the variation of temperature and precipitation over the past 20,000 years[31], as well as the maximum extent of glaciation during the Last Glacial Maximum[32] (with higher values of each indicating lower stability).

In addition to variables associated with evolutionary imbalance mechanisms, we quantified species' global climatic suitability based on the climatic similarity of potential recipient regions (that is, those outside a species' native range) to the regions composing their native range, as well as numerous anthropogenic factors. We included information on the economic uses for each species and measured the number of territorial links (that is, colonies or dependencies) between their potential recipient regions and the countries overlapping their native range—both of which have been shown to be major pathways of plant naturalization[4,33,34]. In addition, we measured human population

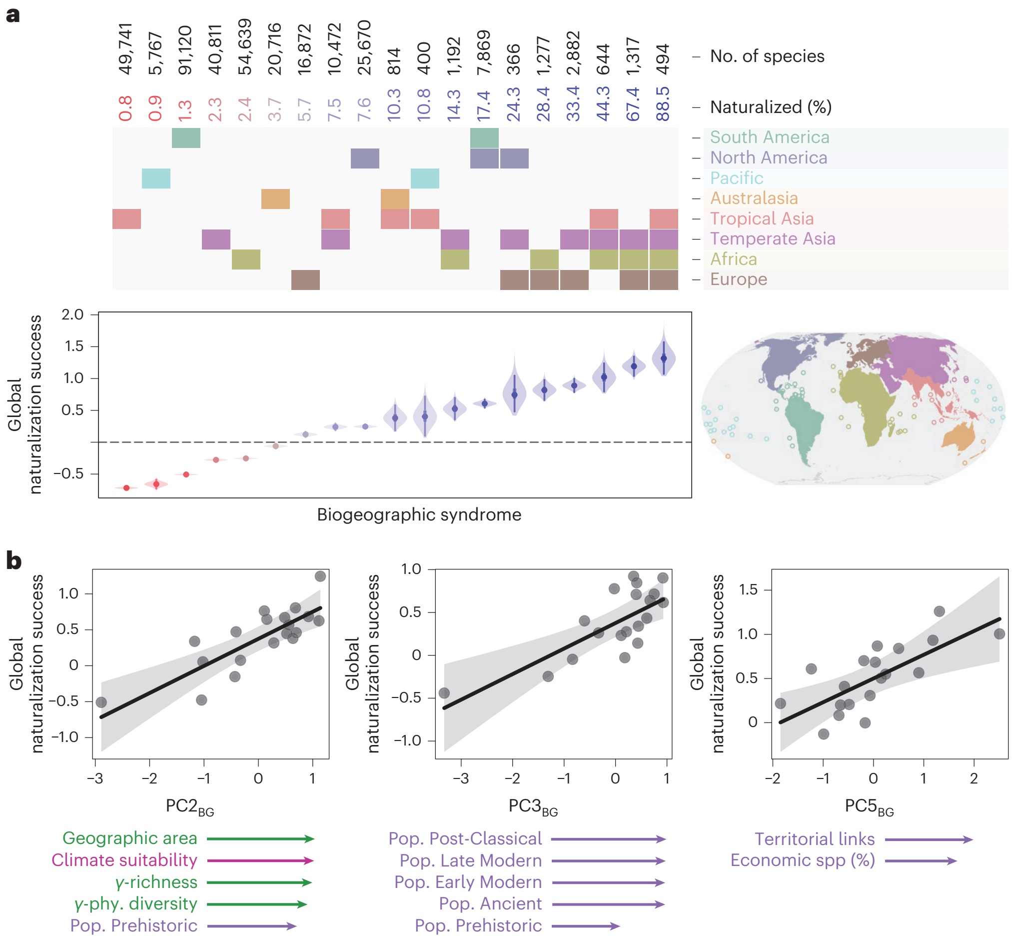


Fig. 1 | Variation and drivers of global naturalization success among plant origins. a,**b**, Global naturalization success across biogeographic syndromes (**a**; coloured from low success in red to high success in blue) and drivers of global naturalization success (**b**; with 'BG' indicating that principal components were derived from variables aggregated within biogeographic syndromes; $n = 19$; full results from linear regression in Extended Data Table 2). In the top panel of **a**, coloured boxes indicate the continents where the species of a given biogeographic syndrome are native; continent colours correspond to the inset map (islands smaller than 8,000 km^2 are marked with a circle). Violin plot of global naturalization success in **a** depicts uncertainty in estimated values due to differences in species numbers (distributions of values derived from $n = 10{,}000$ randomizations; see Methods). In the partial residual plots shown in **b**, main contributors to a given principal component (loading >0.50) are listed along the x axes with the length and direction of arrows indicating the value and sign of the variable loading (see Extended Data Table 1). Variables associated with the EIH are coloured green (phy. diversity, phylogenetic diversity), anthropogenic variables are in purple (pop., human population; economic spp (%), percentage of species with an economic use) and climatic suitability in pink. Shaded bands in **b** indicate 95% confidence bands.

density within the native range across five broad windows of human history[35]: Prehistoric (10000–5000 BCE), Ancient (5000 BCE–500 CE), Post-Classical (500–1500 CE), Early Modern (1500–1800 CE) and Late Modern (1800 CE–present). Human population is related to the frequency of potential encounters with and movements by people, but also to the degree of modification and disturbance of natural habitats within native distributions. Species' associations with high human population density in later periods may therefore indicate greater likelihood of transportation beyond the native range, while regions supporting large human populations through earlier periods may have higher prevalence of species adapted to anthropogenic habitats and disturbance[24].

Bivariate correlation tests showed that many of the distributional characteristics of biogeographic syndromes were associated with naturalization success (Supplementary Fig. 2). However, we also identified many correlations among potential drivers (Supplementary Fig. 3). We therefore focused on five composite variables derived using principal component analysis (PCA) to describe which features define the biogeographic syndromes that have been most successful in donating naturalized plants (denoted as $PC1_{BG}$–$PC5_{BG}$, with 'BG' specifying that variables are quantified at the biogeographic syndrome level; Extended Data Table 1). Biogeographic and anthropogenic factors together explained 73% of the variation in naturalization success (adjusted R^2; Extended Data Table 2). Consistent with the first two mechanisms of the EIH, the strongest predictor of naturalization success ($PC2_{BG}$, $P < 0.01$) indicated that successful syndromes included species with widespread distributions that overlap with a high diversity of co-occurring species (Fig. 1b). Species within these biogeographic syndromes also tended to have high climatic suitability globally and to occur in areas of high human density during the prehistoric period. We identified an additional positive effect of human population across time periods ($PC3_{BG}$, $P < 0.01$; temporal autocorrelation in population did not allow us to tease apart the relative importance of different periods), and for biogeographic syndromes with many economic plants and where species' native distributions include countries with many territorial links to potential recipient regions ($PC5_{BG}$, $P < 0.01$). Composite variables associated with measures of α-diversity ($PC1_{BG}$, $P = 0.06$) or environmental stability and glaciation disturbance ($PC1_{BG}$; $PC4_{BG}$, $P = 0.40$) were unrelated to naturalization success.

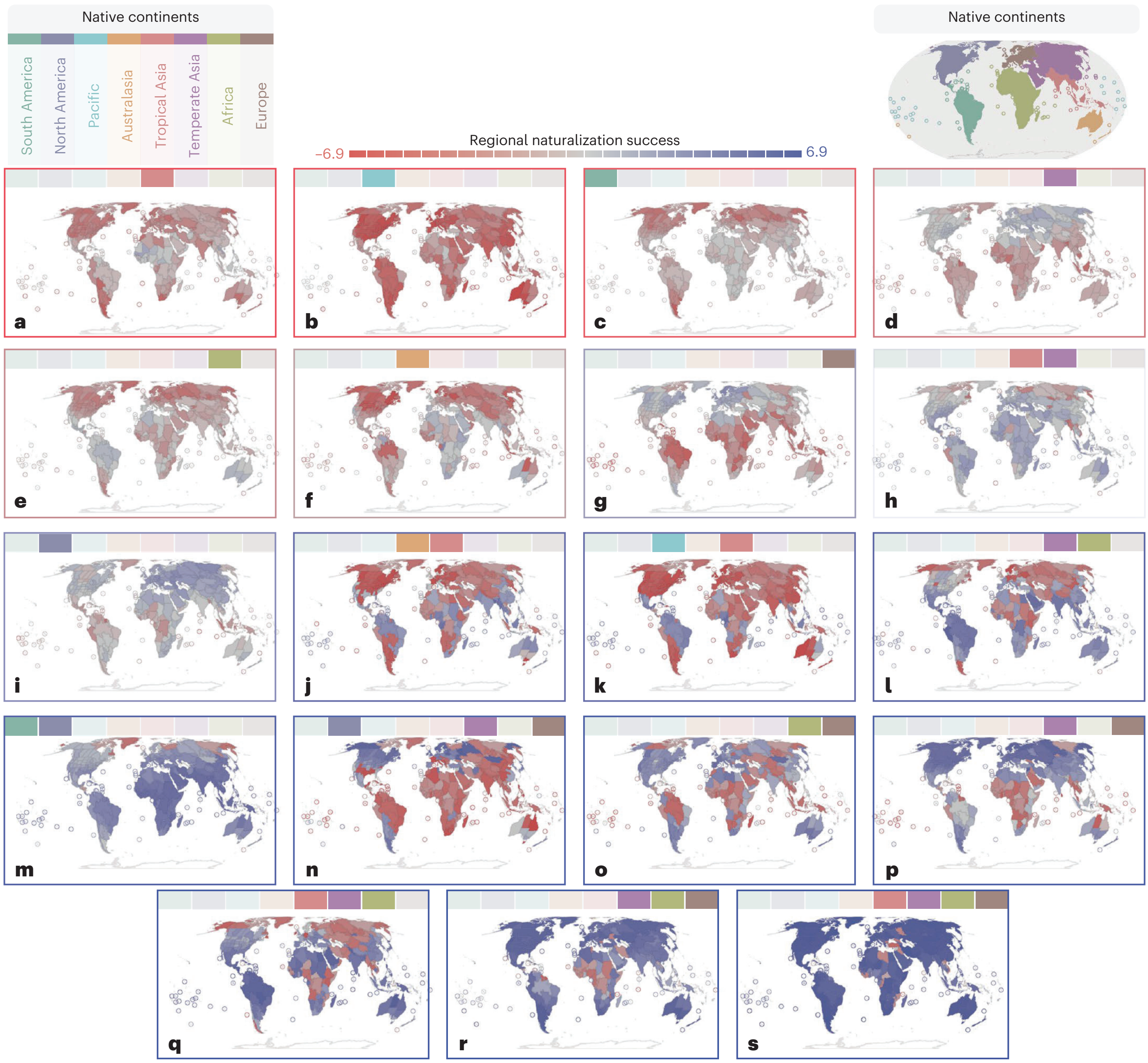


Fig. 2 | Geographic variation in naturalization success for species originating from each biogeographic syndrome. a–**s**, Naturalization success within each potential recipient region is depicted from low success in red (fewer species originating from the biogeographic syndrome have naturalized in the recipient region than expected by chance) to high success in blue (more naturalized species than expected by chance). Biogeographic syndromes are ordered from lowest to highest global naturalization success, with the colour of bounding boxes corresponding to the values in Fig. 1a. Bold coloured blocks at the top of each panel indicate the native continents composing each biogeographic syndrome (see key and map at top).

Climatic constraints on naturalization

Climatic suitability across potential recipient regions varied widely among biogeographic syndromes (Supplementary Fig. 4) and was a primary driver of global naturalization success (Fig. 1b and Supplementary Fig. 2). The importance of climate filtering in structuring the biogeography of alien plants was further illustrated by comparing geographic patterns of naturalization success for each biogeographic syndrome (Fig. 2). For these analyses, we calculated a regional naturalization success metric for the species native to each biogeographic syndrome within each potential recipient region (see ‘Quantifying naturalization success’ in Methods for further details on this metric). Very few biogeographic syndromes, primarily those that span temperate and tropical continents (for example, pan-American or Afro-Eurasian syndromes; Fig. 2m,s), have successfully donated naturalized species across the globe. For the species originating from most syndromes, including many that ranked highly in global naturalization success, broad climatic regions appear resistant to naturalization. For example, a high proportion of palaearctic species (that is, those native to Europe and temperate Asia) have naturalized somewhere beyond their native distributions (Fig. 1a), but relatively few have been successful in infiltrating tropical regions (Fig. 2p). Analysis of regional naturalization success confirmed climatic suitability as a powerful driver of these geographic patterns, but also detected a signal of territorial links between countries of the donor and recipient regions in shaping flows of aliens. Consistent with the analyses of global naturalization success, these results for regional naturalization success showed that successful aliens tend to have strong anthropogenic associations and originate from large biodiverse native distributions. In addition, however, we found a positive effect for biogeographic syndromes where glaciation was low, but precipitation variability high (Extended Data Fig. 1, and Extended Data Tables 3 and 4).

Evolutionary imbalance and human use

In addition to supporting the geographic area and diversity mechanisms of the EIH, our analyses of naturalization success across biogeographic syndromes demonstrate how human history has been a crucial factor in dictating the global biogeography of naturalized plants. As shown in previous studies[4], we found that plants cultivated and used by humans have been particularly successful in establishing outside their native distributions. But just as the course of human cultural evolution is shaping patterns of diversity and distributions in the Anthropocene, the ecology and evolutionary histories of plants form the basis of plant–human relationships, and in turn have the potential to influence trajectories of socio-economic and cultural systems[36–40]. Humans select species for characteristics that confer a particular societal or economic value, but the processes of cultural evolution should more generally favour associations with plants that possess the highest capacity for survival, growth and proliferation under a range of environmental conditions, and against a biotic background that includes potential competitors, herbivores and pathogens[41–43]. Therefore, we put forward that evolutionary imbalance may play an important role in determining the plants that humans form associations with, and ultimately transport beyond their native ranges.

We tested for this possible secondary effect of evolutionary imbalance in shaping naturalization patterns by applying species-level phylogenetic generalized linear models to assess the distributional characteristics associated with economic use and, subsequently, naturalization success (Fig. 3). As in our biogeographic level analysis, we derived composite variables ($PC1_{Spp}$–$PC7_{Spp}$, with 'Spp' specifying that variables are quantified at the species level) using PCA of predictor variables across all species (Extended Data Table 5). Our results (Extended Data Table 6) confirm anthropogenic factors that are expected to be important mediators of plant–human associations: plants with native distributions that have overlapped with centres of high human population from ancient to modern times ($PC2_{Spp}$, $P < 0.01$) and with countries that share many territorial links ($PC7_{Spp}$, $P < 0.01$) were more likely to be selected for economic use. Unexpectedly, however, we found a negative effect of prehistoric human population ($PC5_{Spp}$, $P < 0.01$). Consistent with evolutionary imbalance, economic plants tended to have large native distributions that encompass many other species ($PC3_{Spp}$, $P < 0.01$) and to persist in areas of high local diversity ($PC1_{Spp}$, $P < 0.01$; Fig. 3). The effects of environmental stability were more complex. While glacial disturbance and low temperature stability decreased the probability of economic use ($PC4_{Spp}$, $P < 0.01$), species originating from areas with low precipitation stability were more likely to have an economic use ($PC6_{Spp}$, $P < 0.01$). However, climatic suitability also increased along $PC6_{Spp}$, and was therefore generally higher for economic plants (see also results for $PC3_{Spp}$).

These results on the drivers of economic use, in combination with a species-level analysis of naturalization success (Fig. 3 and Extended Data Table 7), suggest that species with large distributions that overlap with biodiverse regions are more likely to form associations with humans, and secondarily will have a higher probability of naturalizing once transported outside their native distribution. Interestingly, however, while metrics of α-diversity were associated with increased naturalization success for economic species, the opposite was true for non-economic plants. This finding perhaps reflects that many of the species occupying the most biodiverse regions tend to be rare, with small, specialized distributions[44] located in landscapes where connectivity to trade and transportation networks is relatively low (for example, tropical forest or mountainous regions). These conditions probably mean decreased contact with humans and reduced likelihood of accidental translocation into new areas[45]. This possibility can partially be addressed by investigating the degree of naturalization success within the subset of naturalized plants, which, by definition, are known to have been moved by human activities. Indeed, among naturalized plants, we found that species with large, biodiverse native distributions have become naturalized in a greater number of regions worldwide (Fig. 4, and Extended Data Tables 8 and 9; note that for composite variables $PC1_{Nat}$–$PC6_{Nat}$, 'Nat' specifies that variables are quantified using data only from naturalized species); this effect was apparent in both economic and non-economic species.

Discussion

Our analyses of the distributional characteristics for over 99% of the known global seed plant flora reveal geographic signatures in the exchange of naturalized aliens that are largely consistent with expectations from the EIH. These patterns are generally consistent across scales of analyses (biogeographic syndromes or individual species) and across metrics of naturalization success (global or regional naturalization success; naturalization probability or number of naturalized regions at the species level). Our results suggest that the biogeographic conditions of certain global regions have preadapted species for success as invaders. In particular, species spanning vast, interconnected landmasses are overrepresented among naturalized plants. In addition, our results underscore the importance of climatic and anthropogenic factors in acting alongside evolutionary imbalance to shape the distribution of plants in the Anthropocene.

Species originating from large, biodiverse regions have been more successful at naturalizing beyond their native ranges[9]. This is in line with large areas facilitating effective selection, as proposed by the EIH, but wide-ranging distributions are also generally indicative of broad environmental tolerances as well as increased opportunities for interactions with humans[45–48]. Nevertheless, our finding of a positive association between native-range area and naturalization success was consistent when accounting for climate filtering directly (that is, analyses of regional naturalization success) and within the subset of plants with economic uses. Large geographic distributions also generally include greater numbers of co-occurring species, so the native ranges of successful aliens also tended to encompass diverse biotic environments (that is, high γ-diversity). Across biogeographic syndromes, however, the predicted link between native-range α-diversity and naturalization success was masked by confounding processes. We found that species originating from the biodiverse predominantly tropical continents were disproportionately rare among naturalized aliens, but this is probably because they tend to have less area of suitable climate globally (Supplementary Fig. 4) and to have less exposure to the global economic systems that drive anthropogenic movements of species[22,49]. Among economic plants, which have established histories of interacting with humans, species originating from regions of high local diversity were more likely to be naturalized; among all naturalized plants, native-range α-diversity was further associated with naturalization in more regions globally. Synthesizing results across scales of analyses was therefore key in disentangling the signal of evolutionary imbalance in global naturalization patterns.

The EIH proposes that stable conditions allow the time necessary for natural selection to work, but the predicted association between naturalization success and environmental stability received mixed support across our analyses. One possible explanation is that the regions that have experienced the greatest climatic shifts over the past 20,000 years are also the most variable over seasonal and interannual timescales[50]. Species that persist in these regions therefore must be adapted to a wider range of environmental conditions[51,52], which probably provides an advantage for establishing in a variety of new

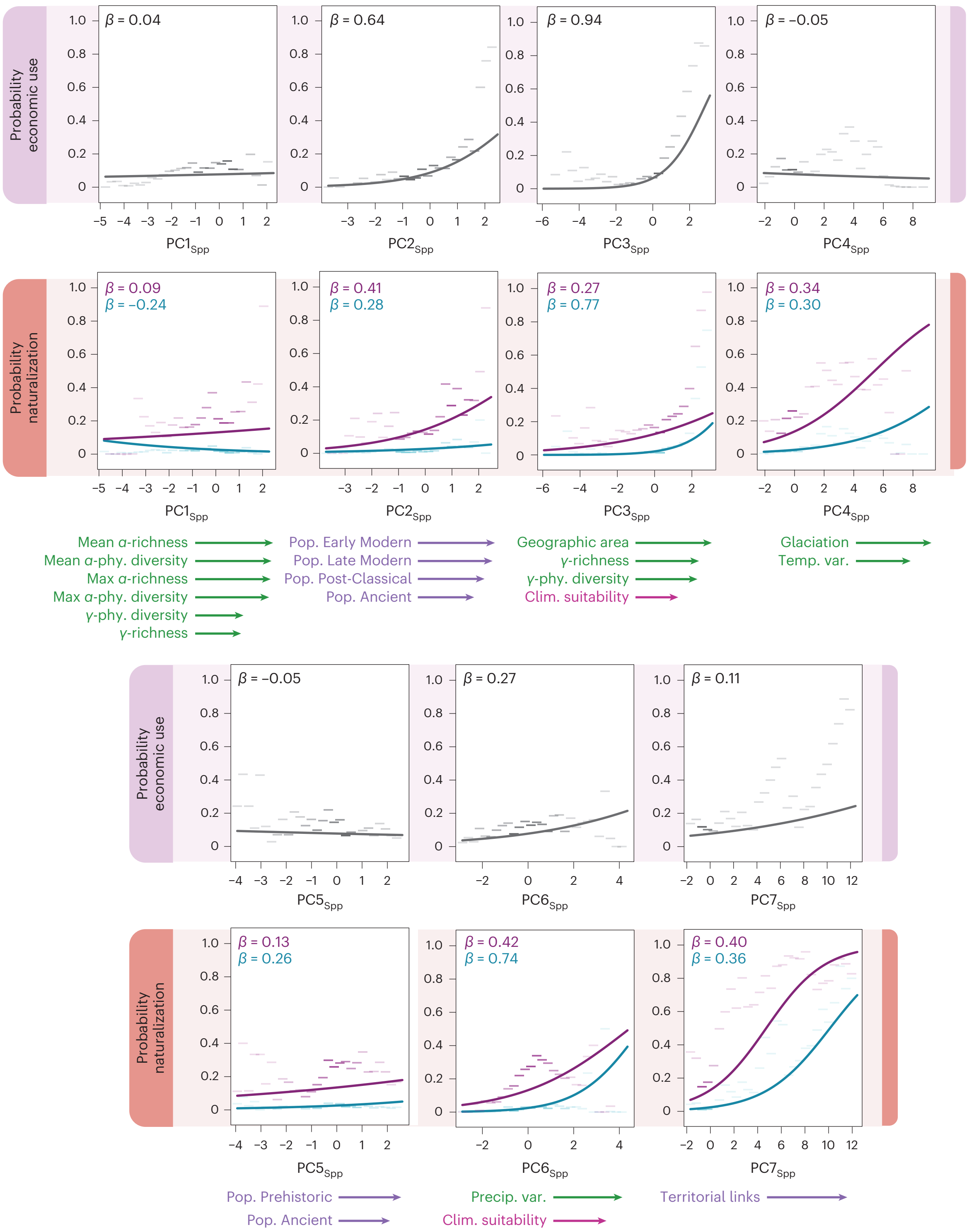


Fig. 3 | Drivers of economic use and naturalization across the global seed plant flora (n = 334,667). Panels illustrate the results of phylogenetic binomial regression predicting either plant economic use (top panels, outlined in purple) or naturalization (bottom panels, outlined in red; full results in Extended Data Tables 6 and 7). Lines show the estimated probability of economic use (top panels in black) or naturalization (bottom panels, with purple for economic species and turquoise for non-economic species) along a given species-level principal component ($PC1_{Spp}$–$PC7_{Spp}$, with 'Spp' indicating that PCs were derived using data for all species; see Extended Data Table 5), with all other principal components set to the median value. Bars within plots show the observed proportion of species with an economic use (top panels in black) or naturalized outside their native distributions (bottom panels, with purple for economic species and turquoise for non-economic species) within 30 even bins across values of a given principal component, with opacity indicating the relative number of species of a given group (for example, economic or non-economic) in the bin. Main contributors to a given principal component (loading >0.50) are listed along the *x* axes, with the length and direction of arrows indicating the value and sign of the variable loading. Variables associated with the EIH are coloured green (phy. diversity, phylogenetic diversity; temp. var., temperature variability; precip. var., precipitation variability), anthropogenic variables in purple (pop., human population) and climatic suitability (clim. suitability) in pink.

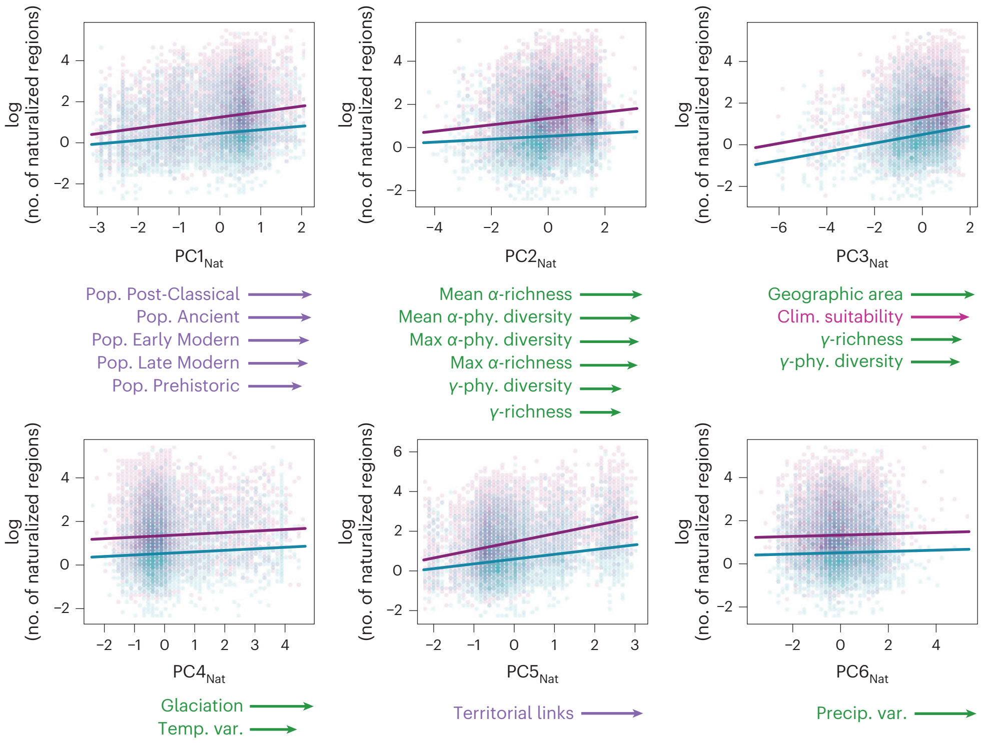


Fig. 4 | Drivers of naturalization extent among the naturalized alien seed plant flora (n = 13,280). Partial residual plots from phylogenetic linear regression predicting number of naturalized regions among the global naturalized flora (full results in Extended Data Table 9). Relationships for economic plants are shown in purple and non-economic plants in turquoise (darker coloured points indicate more species within a given interval of x and y values). Main contributors to a given principal component (loading >0.50) are listed along the x axes, with the length and direction of arrows indicating the value and sign of the variable loading (see Extended Data Table 8; 'Nat' indicates that PCs were derived using only data for naturalized species). Variables associated with the EIH are coloured green (phy. diversity, maximum area-corrected phylogenetic diversity; temp. var., temperature variability; precip. var., precipitation variability), anthropogenic variables in purple (pop., human population) and climatic suitability in pink.

regions. Additionally, we were only able to measure environmental stability for the areas within modern native distributions of the species. At continental scales, species have the potential to track favourable habitats to dampen the impacts of environmental change[53].

In quantifying global patterns in the exchange of naturalized plants, our results suggest an influence of evolutionary imbalance among biogeographic regions, but the movement of alien species is ultimately driven by human activities. Anthropogenic factors were consistently strong predictors of naturalization success across our analyses. Plants that have existed alongside large human populations, and by extension widespread anthropogenic disturbance, from prehistoric to modern time periods, especially in areas where cultures have established many territorial links (primarily via colonialism) across the world, have been particularly successful at naturalizing outside their native distributions. Proximity to humans and their global economic networks additionally increased the likelihood that species have been adopted for economic use, which our results reaffirm as a primary pathway of naturalization. However, we additionally revealed a previously unrecognized role of evolutionary imbalance in shaping plant–human relationships. The native distributions of economic plants tend to have characteristics consistent with those predicted by the EIH mechanisms; the same characteristics are further associated with naturalization success among these species. Interestingly, the global regions that have produced the most successful plant invaders have also given rise to human societies that have expanded into and colonized new lands—often beyond their continents of origin—to create vast empires[36,37]. While this association is partially explained by colonial networks providing opportunities for these plants to be transported beyond their native distributions[34], our analyses suggest that eco-evolutionary forces also play a role in the success of these species. Biogeographic factors have similarly been suggested to drive imbalances among human cultures that have been pivotal in shaping their expansions and contractions[36]. Our study therefore highlights intriguing parallels and potential interactions between evolutionary imbalance in biological and cultural systems.

When assessing evolutionary imbalance at the global scale, the resolution of available data necessarily imposes certain constraints. Measuring the success of aliens using data on naturalizations, for example, we are unable to capture the multidimensional and continuous nature of invasiveness[46]. Future studies evaluating consequences of evolutionary imbalance in invasion ecology will no doubt benefit from harnessing information on the abundances, impacts and extent of occurrence for aliens as it becomes available for more species in more regions. Owing to the context-dependent nature of biotic interactions, measuring native and alien distributions at the relatively coarse scales that are currently available may also overlook potentially meaningful details. For example, competition with natives, and by extension evolutionary imbalance, has been shown to be less important for mediating invasions in disturbed habitats[18]. While we considered broad differences in climate across potential recipient regions, our analyses were unable to account for the local habitats where species naturalize. However, considering the important role of disturbance in facilitating invasions globally, we believe that this speaks to the strength of the evolutionary imbalance signal present in our data. Ultimately, we emphasize that inferences from global analyses such as ours can be most impactful when considered alongside the results of studies conducted across scales, taxa, locations and diverse ecological contexts.

Understanding the characteristics and circumstances that allow alien species to establish is crucial for mediating the ongoing swell of biological invasions[54]. We identify evolutionary imbalance as a so-far poorly accounted driver of biological invasions. Focusing on the biogeographic conditions that select for invasion potential, our analyses cannot infer whether a common suite of traits or diverse ecological strategies characterize successful aliens. However, our results indicate that readily measured features of species' native distributions can be informative in forecasting potential invaders as well as potentially vulnerable regions. In addition, we identified eco-evolutionary processes that have shaped plant–human relationships and the global movements of alien species. While our insights into evolutionary imbalance have direct implications for invasion biology, they can also be applied more generally to understanding spatiotemporal dynamics in the distribution of biodiversity, such as the natural biotic exchanges that have occurred throughout Earth's history.

Methods

Distribution data

Native distributions for the global seed plant flora ($n = 420{,}748$ species and infraspecific taxa) were compiled from numerous sources. Most records were downloaded from Plants of the World Online ($n = 389{,}451$)[30], with additional taxa added from the Global Compositae Database ($n = 15{,}301$)[27], the United States Department of Agriculture Germplasm Resources Information Network ($n = 11{,}453$) and the International Union for Conservation of Nature Red List of Threatened Specied ($n = 4{,}543$)[28]. Native distributions were standardized to the 369 regions of level 3 (TDWG3) of the Biodiversity Information Standards World Geographical Scheme for Recording Plant Distributions[26]; in order to align with naturalized distribution data, the regions of Yemen and Socotra and the two portions of Turkey were each merged, respectively, resulting in a total of 367 regions. Distributions for infraspecific taxa were collapsed to the species level and matched to the taxonomy of the World Checklist of Vascular Plants following their provided list of synonyms[19]. Our final database included native distributions for 336,245 species. In the case of ancient introductions (that is, archaeophytes), determining the native status of a species in a particular region can be challenging, especially in areas with long histories of human disturbance[55]. While all four datasets that contributed native-range data make efforts to separate these non-native regions when reporting species' native distributions, some amount of uncertainty is unavoidable. However, archaeophytes should contribute few species to our overall dataset; in the case of Europe, a continent where the archaeophyte flora is relatively well documented, the richness of archaeophytes is only around 2% of the native flora richness[56]. In addition, because archaeophytes are typically defined as having been established within an area for at least five centuries, the regions where uncertainty in native status is highest are not necessarily irrelevant to a species' evolutionary history.

Naturalized distributions were taken from GloNAF[7]. Naturalized distributions were matched to TDWG3 regions as described in Supplementary Data 1, collapsed to the species level and harmonized with the taxonomy of the World Checklist of Vascular plants following their provided list of synonyms. In cases where a species was considered naturalized according to GloNAF and native according to the combined native distribution dataset, we classified the species as native and removed the naturalized record. In total, our dataset included naturalized distributions for 14,461 species (4.3% of the global seed plant flora). While data on naturalizations does not indicate a species' degree of invasiveness in a given region (that is, its ability to proliferate, spread or impact native ecosystems), we note that establishment is a crucial phase in the invasion process that can be strongly influenced by the biotic interactions that occur following introduction[57]. We therefore use global patterns in plant naturalizations to ask whether aliens originating from certain regions are more likely to overcome biotic resistance and establish populations than the aliens originating from other regions.

Defining regions of origin

One challenge in quantifying geographic biases in naturalization success is to define ecologically meaningful classifications for the geographic origins of species. Previous efforts to investigate geographic imbalances in species exchange have typically relied on broad classifications of species origins (for example, Northern versus Southern Hemispheres[2], temperate versus tropical zones[58] or Old versus New World[24,59]) or assigned species to individual continents of origin[3,4,60]. The former risks overlooking important differences within broad classifications, while the latter ignores the fact that many of the biogeographic features that define species distributions (for example, the neotropics, Mediterranean basin, or Eurasian taiga) span multiple continents—both are issues that may obscure important patterns. To overcome these limitations, we classified common patterns of plant distributions according to the combination of continents where each species occurs as native. Specifically, we followed the continental scheme outlined by the World Geographic Scheme for Recording Plant Distributions (that is, TDWG level 1 regions; TDWG1)[26] and refer to each unique pattern of native distribution across these nine continents as biogeographic syndromes. The TDWG1 regions generally follow geographic continents, with Asia divided into temperate and tropical Asia, and the Pacific Islands forming a single region (Supplementary Fig. 1). A species was considered native to a continent if it was recorded as native in at least one TDWG3 region within the continent. While 511 possible combinations of continents exist, we identified only 161 biogeographic syndromes among the global seed plant flora. Over 99% of species were classified to the 19 most common patterns; 104 (65%) of the patterns corresponded to fewer than 10 species, while 42 (26%) corresponded to a single species. Certain combinations of continents (biogeographic syndromes) were more common than others for a number of reasons: geographic diversity patterns (for example, primarily tropical continents have more species than primarily temperate); combinations with fewer continents are observed much more frequently because small range sizes are much more common across species than large range sizes; and because of geographic constraints on distributions such as oceans or other barriers (for example, thousands of species have native distributions that include North and South America, but exceedingly few occur in the combination of North America and Australasia).

Quantifying naturalization success

Naturalization success for the species within each biogeographic syndrome was measured as the $\log_{10}$ ratio of the observed proportion of species that have become naturalized compared to the expected proportion:

$$\text{Naturalization success} = \log_{10}\,(\text{observed naturalized}/\text{expected naturalized})$$

For each biogeographic syndrome, we calculated naturalization success globally, and within each potential TDWG3 recipient region. For global naturalization success, the observed naturalized value was the proportion of species with native distributions described by a given biogeographic syndrome that were recorded in the GloNAF dataset as naturalized outside of their native distribution. For regional naturalization success, the observed naturalized value was calculated as the proportion of species assigned to the given biogeographic syndrome that have naturalized in a focal recipient region (species native to the focal recipient region were excluded when calculating this proportion). For both metrics, the expected naturalized value is based on a scenario where the probability of naturalization is equal for species among biogeographic syndromes (that is, no biases in the geographic origins of successful aliens). For global naturalization success, the expected proportion was the proportion of naturalized species in the global flora (4.3%). For regional naturalization success, the expected proportion was the proportion of the global flora that have naturalized in the focal recipient region (species native to the focal recipient region were excluded when calculating this proportion). This metric is a convenient measure of naturalization success because it follows an asymptotic normal distribution and can be interpreted intuitively: values above or below zero indicate more or fewer, respectively, naturalized species than expected, and values are symmetric around zero (for example, 1 indicates 10 times more species than expected, −1 indicates 10 times fewer than expected). It also controls for the diversity of native species in each biogeographic syndrome (that is, naturalized species are more likely to originate from biodiverse regions by chance) and for the numbers of naturalized species recorded in potential recipient regions (that is, some regions are more invasible or have more complete naturalized species lists compared to others). In comparing naturalization success among biogeographic syndromes, we therefore did not need to include covariates associated with recipient region invasibility (for example, anthropogenic factors such as the degree of landscape modification) or quantify relative measures of evolutionary imbalance variables (for example, difference in biodiversity between donor and recipient regions). To explore uncertainty due to stochasticity in expected proportions, we additionally derived expected values through randomizations[4] when calculating global naturalization success. For each of 10,000 iterations, we randomly selected 14,461 species (the number of naturalized species in the GloNAF dataset). Uncertainty in expected values was higher for biogeographic syndromes with fewer species (Fig. 1a).

Evolutionary imbalance drivers

Variables associated with the three evolutionary imbalance mechanisms were measured first within the native distributions of species (based on the TDWG3 regions), and then aggregated to a single value for each biogeographic syndrome (the unique combination of TDWG1 continents) as the mean across species. Geographic area (mechanism 1) was measured as the combined area of all TDWG3 regions where a species was considered native. Species diversity (mechanism 2) was measured as species richness and, to account for the evolutionary distinctness of potentially interacting species, as Faith's phylogenetic diversity[61]. When measuring phylogenetic diversity, we used the global seed plant phylogeny from ref. 62 with missing species added to the root of their respective genera or family, following the methods in ref. 4. For each form of diversity, we calculated three metrics from lists of co-occurring native species for the TDWG3 regions of a given species' native distribution: the diversity of co-occurring species across the entire distribution (γ-diversity); the area-corrected diversity of the entire distribution (mean α-diversity); and the maximum area-corrected diversity of regions within the native distribution (max α-diversity). Area-corrected diversity was calculated as the residuals of the log–log relationship between geographic area and species richness or phylogenetic diversity across unique species distributions (mean α-diversity; richness: $\beta = 0.49$, $P < 0.01$, $R^2 = 0.44$; phylogenetic diversity: $\beta = 0.44$, $P < 0.01$, $R^2 = 0.39$) or regions (max α-diversity; richness: $\beta = 0.40$, $P < 0.01$, $R^2 = 0.63$; phylogenetic diversity: $\beta = 0.29$, $P < 0.01$, $R^2 = 0.58$)[63]. To capture environmental stability of species native distributions (mechanism 3), we measured the stability of temperature and precipitation from 21000 BCE to the present. Climatic data came from the StableClim database[31]; for both temperature and precipitation, we measured the average absolute change across 100-year intervals (henceforth temperature and precipitation variability). In the case of precipitation variability, changes were measured relative to the overall mean precipitation. Both sets of raster data were projected to Wagner IV equal-area projection and average values were extracted for species' native distributions. For precipitation variability, values were log-transformed following extraction. As the derived values provided a measure of change over time, lower values indicate more stable regions. We additionally measured the extent of glaciation during the Last Glacial Maximum, with higher glaciation corresponding to lower stability. The proportional glaciated area of species' native distributions was extracted from ref. 32.

Anthropogenic drivers

Data on plant economic uses came from the World Checklist of Useful Plants[64], with names standardized to the World Checklist of Vascular Plants taxonomy[19]. Following the results of ref. 4, we only considered plants with economic uses that have been shown to be associated with increased naturalization success: animal foods, environmental uses, gene sources, human foods, materials and medicines. For each biogeographic syndrome, we calculated the proportion of species used for an economic purpose.

Data on human population came from the HYDE 3.2 dataset[35]. The data and hindcasting methods used to produce these population estimates are accompanied by substantial uncertainties, especially for older time periods. However, for our purposes—relative comparisons of population across broad regions and coarse time periods—they should provide reasonable approximations. We calculated the average human population across geographic grid cells within five broad windows of human history: Prehistoric (10000–5000 BCE), Ancient (5000 BCE–500 CE), Post-Classical (500–1500 CE), Early Modern (1500–1800 CE) and Late Modern (1800 CE–present). Because human cultures have shown substantial variation in the direction and timing of their trajectories, it is impossible for broad time bins such as these to be globally representative. Acknowledging unavoidable cultural and geographic biases, we chose these windows because they roughly correspond to historical developments relevant to the biogeography of alien plants: the advent of cities, civilizations and their associated anthropogenic disturbances occurred during the Ancient period; the Post-Classical period saw the expansion of trade networks across the Afro-Eurasian landmass; colonial empires transported many plants between colonized and colonizing lands during the Early Modern period; and the Late Modern time period saw the advent of industrial systems and rapid acceleration in the speed and extent of transportation networks. Global rasters of population for each time period were projected to Wagner IV equal-area projection, and we calculated average values within each region. Then, we found for each species the maximum value of all TDWG3 regions within its native distribution. At the scale of biogeographic syndromes, we averaged values across species.

Information on territorial links came from the TRADHIST dataset[65], supplemented with additional information on colonies and dependent territories for the period starting in 1492 (that is, the year of Columbus' arrival in the Americas and the start of the period during which the vast majority of introductions and naturalizations have occurred), following the methods and sources in ref. 33. Because data on territorial links were provided at the country level, we matched TDWG3 regions to countries as described in Supplementary Data 2. For each species, we counted the number of potential recipient regions with territorial links to the countries that overlap with the native distribution. For analyses of regional naturalization success, we calculated the proportion of species from each biogeographic syndrome with territorial links between the countries in their native distribution and the given recipient region. When analysing global naturalization success, we calculated for each biogeographic syndrome the proportion of species with territorial links averaged across all potential recipient regions, weighted by the area of regions.

Climatic suitability

For estimating climatic suitability of species across potential recipient regions, we started by calculating the distance in climatic space defined by mean annual temperature (Bioclimatic variable 1 from the WorldClim version 2 database[66]) and square-root transformed annual precipitation (Bioclimatic variable 4) between species' native distributions and each potential recipient region. Climate raster data were originally downloaded at 10-min resolution and aggregated to 50-min resolution, in order to simplify downstream calculations, and projected to Wagner IV equal-area projection. For each cell within a given recipient region, we calculated the minimum Euclidean distance in two-dimensional climate space to the cells in a focal species' native distribution and then averaged these minimum values across the region. Values of zero for this metric of climatic distance indicate a recipient region with climatic conditions completely contained within the species' native distribution; increasing values indicate increasing climatic distance, similar to the metric proposed by ref. 67. Our measure of climatic suitability was then the additive inverse of this value, such that lower (that is, more negative) values indicate lower climatic suitability. The climatic suitability within each potential recipient region for a given biogeographic syndrome was the average across species, excluding those native to the recipient region. Global climatic suitability for a biogeographic syndrome was the average of these values, weighted by the geographic area of regions. Global climatic suitability at the species level was also calculated as the mean across regions weighted by area.

Data for all variables for biogeographic syndromes can be found in Supplementary Data 3, data for analysis of regional naturalization success in Supplementary Data 4 and species-level data in Supplementary Data 5.

Statistical analyses

We performed five complementary regression analyses to test for drivers of (1) global and (2) regional naturalization success among biogeographic syndromes, (3) probability of economic use and (4) naturalization among species, and, for species that have naturalized beyond their native distribution, (5) the number of naturalized regions. All analyses were performed in R version 4.2.0 (ref. 68). For each analysis, we performed PCA with varimax rotation on predictor variables, maintaining rotated components with cumulative variance explained of at least 90% (Extended Data Tables 1, 3, 5 and 8). Climatic suitability and territorial links were not included in the PCA for regional naturalization success because these values varied for every region by biogeographic syndrome comparison. While there are various methods for dealing with collinearity between predictor variables, we chose the PCA approach because it is transparent in presenting correlations and rightly calls attention to the inherent uncertainty in which (combinations of) correlated variables may be driving a given relationship[69]. Removing redundancy by retaining only the best predicting variable to represent each cluster of correlated variables is an alternative suggestion for easing model interpretability, but this can be misleading in hiding correlated variables that may be ecologically relevant. However, in line with this option and to show more completely the relationships among our data, we include the results from single predictor models in Supplementary Data 6. Within these results, we highlight which variable is the best predictor among the variables clustering with a given PC.

For analysis of naturalization success among biogeographic syndromes (analyses 1 and 2), we performed linear regression with the log of native species number for each biogeographic syndrome as weight. Our regional analysis (analysis 2) included metrics of naturalization success for species from each biogeographic syndrome within each recipient region; to account for each biogeographic syndrome contributing multiple measures of regional naturalization success, we used a linear mixed model (in the nlme package[70]) with the biogeographic syndrome included as a random effect. To account for the possibility of spatial autocorrelation across recipient regions, we compared models incorporating various types of spatial correlation structure (exponential, gaussian, linear, rational quadratic, spherical, no correlation structure) and selected exponential based on comparison of the Akaike information criterion.

Probability of economic use and naturalization among species (analysis 3 and 4) were analysed using phylogenetic binomial regression using the phyloglm function from the phylolm package[71]. The phylogenetic tree was the same as used for calculating phylogenetic diversity (Supplementary Data 7), with tree height scaled to 1 (refs. 33,62). Finally, to analyse the number of naturalized regions for naturalized species (analysis 5), we used phylogenetic linear regression[71] with log region number as the response. The use of phylogenetically informed analyses reduced the possibility that any geographic patterns are driven by the imbalanced representation of certain taxa or (assuming phylogenetic trait conservatism) traits across regional native floras. For analyses of naturalization probability and number of naturalized regions (analyses 4 and 5), we included interactions between economic use and other predictor variables. Non-significant interactions were removed during model selection. To assure that there was no implicit dispersion[72] in our binary regression models (analyses 3 and 4), we followed advice in the DHARMa R package[73]; specifically, we calculated simulation-based standardized residuals, grouped them by each predictor variable or by biogeographic syndrome ID, and in each case calculated dispersion for the grouped residuals. R code for reproducing analyses can be found in Supplementary Code 1.

References

1. Richardson, D. M. et al. Naturalization and invasion of alien plants: concepts and definitions. *Divers. Distrib.* **6**, 93–107 (2000).
2. Darwin, C. *On the Origin of Species* (John Murray, 1859).
3. van Kleunen, M. et al. Global exchange and accumulation of non-native plants. *Nature* **525**, 100–103 (2015).
4. van Kleunen, M. et al. Economic use of plants is key to their naturalization success. *Nat. Commun.* **11**, 3201 (2020).
5. Dyer, E. E. et al. The global distribution and drivers of alien bird species richness. *PLoS Biol.* **15**, e2000942 (2017).
6. Dyer, E. E., Redding, D. W. & Blackburn, T. M. The global avian invasions atlas, a database of alien bird distributions worldwide. *Sci. Data* **4**, 170041 (2017).
7. van Kleunen, M. et al. The Global Naturalized Alien Flora (GloNAF) database. *Ecology* **100**, e02542 (2019).
8. Vermeij, G. J. When biotas meet: understanding biotic interchange. *Science* **253**, 1099–1104 (1991).
9. Vermeij, G. in *Species Invasions: Insights into Ecology, Evolution, and Biogeography* (eds Sax, D. F. et al.) 315–340 (Sinauer, 2005).
10. Fridley, J. D. & Sax, D. F. The imbalance of nature: revisiting a Darwinian framework for invasion biology. *Glob. Ecol. Biogeogr.* **23**, 1157–1166 (2014).
11. Leimu, R., Mutikainen, P., Koricheva, J. & Fischer, M. How general are positive relationships between plant population size, fitness and genetic variation? *J. Ecol.* **94**, 942–952 (2006).
12. Tilman, D. Diversification, biotic interchange, and the universal trade-off hypothesis. *Am. Nat.* **178**, 355–371 (2011).
13. Dobzhansky, T. Evolution in the tropics. *Am. Sci.* **38**, 209–221 (1950).

14. MacArthur, R. H. *Geographical Ecology: Patterns in the Distribution of Species* (Princeton Univ. Press, 1972).
15. Cody, M. L. & Mooney, H. A. Convergence versus nonconvergence in Mediterranean-climate ecosystems. *Annu. Rev. Ecol. Syst.* **9**, 265–321 (1978).
16. Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am. Nat.* **138**, 1315–1341 (1991).
17. Leigh, E. G. Jr, Vermeij, G. J. & Wikelski, M. What do human economies, large islands and forest fragments reveal about the factors limiting ecosystem evolution? *J. Evolut. Biol.* **22**, 1–12 (2009).
18. Fridley, J. D., Jo, I., Hulme, P. E. & Duncan, R. P. A habitat-based assessment of the role of competition in plant invasions. *J. Ecol.* **109**, 1263–1274 (2021).
19. *World Checklist of Vascular Plants, Version 2.0* (Royal Botanic Gardens Kew, 2022); http://wcvp.science.kew.org/
20. Haeuser, E. et al. European ornamental garden flora as an invasion debt under climate change. *J. Appl. Ecol.* **55**, 2386–2395 (2018).
21. Liu, C., Wolter, C., Xian, W. & Jeschke, J. M. Most invasive species largely conserve their climatic niche. *Proc. Natl Acad. Sci. USA* **117**, 23643–23651 (2020).
22. Seebens, H. et al. Global trade will accelerate plant invasions in emerging economies under climate change. *Glob. Change Biol.* **21**, 4128–4140 (2015).
23. Bertelsmeier, C., Ollier, S., Liebhold, A. & Keller, L. Recent human history governs global ant invasion dynamics. *Nat. Ecol. Evol.* **1**, 0184 (2017).
24. di Castri, F. in *Biological Invasions: A Global Perspective* (eds Drake, J.A. et al.) 1–30 (Wiley, 1989).
25. MacDougall, A. S. & Turkington, R. Are invasive species the drivers or passengers of change in degraded ecosystems? *Ecology* **86**, 42–55 (2005).
26. Brummitt, R. *World Geographical Scheme for Recording Plant Distributions, Edition 2* (Hunt Institute for Botanical Documentation, Carnegie Mellon University, 2001); http://www.tdwg.org/standards/109
27. *Global Compositae Database* (Compositae Working Group, 2022); https://www.compositae.org/
28. *The IUCN Red List of Threatened Species* (IUCN, 2022); https://www.iucnredlist.org
29. *USDA-ARS Germplasm Resources Information Network (GRIN)* (United States Department of Agriculture, 2022); https://www.ars-grin.gov
30. *Plants of the World Online* (Royal Botanic Gardens Kew, 2022); http://www.plantsoftheworldonline.org/
31. Brown, S. C., Wigley, T. M. L., Otto-Bliesner, B. L. & Fordham, D. A. StableClim, continuous projections of climate stability from 21000 BP to 2100 CE at multiple spatial scales. *Sci. Data* **7**, 335 (2020).
32. Ehlers, J., Gibbard, P. L. & Hughes, P. D. *Quaternary Glaciations - Extent and Chronology: A Closer Look* (Elsevier, 2011).
33. Yang, Q. et al. The global loss of floristic uniqueness. *Nat. Commun.* **12**, 7290 (2021).
34. Lenzner, B. et al. Naturalized alien floras still carry the legacy of European colonialism. *Nat. Ecol. Evol.* **6**, 1723–1732 (2022).
35. Klein Goldewijk, K., Beusen, A., Doelman, J. & Stehfest, E. Anthropogenic land use estimates for the Holocene – HYDE 3.2. *Earth Syst. Sci. Data* **9**, 927–953 (2017).
36. Diamond, J. M. *Guns, Germs, and Steel: The Fates of Human Societies* (W.W. Norton, 1997).
37. Diamond, J. & Bellwood, P. Farmers and their languages: the first expansions. *Science* **300**, 597–603 (2003).
38. Vilela, B. et al. Cultural transmission and ecological opportunity jointly shaped the spread of human agriculture. *Evol. Hum. Sci.* **2**, E53 (2020).
39. Balick, M. J. & Cox, P. A. *Plants, People, and Culture: The Science of Ethnobotany* (Garland Science, 2020).
40. Vavilov, N. I., Vavylov, M. I. & Dorofeev, V. F. *Origin and Geography of Cultivated Plants* (Cambridge Univ. Press, 1992).
41. Phillips, O. & Gentry, A. H. The useful plants of Tambopata, Peru: II. Additional hypothesis testing in quantitative ethnobotany. *Econ. Bot.* **47**, 33–43 (1993).
42. Gaoue, O. G. et al. Theories and major hypotheses in ethnobotany. *Econ. Bot.* **71**, 269–287 (2017).
43. Milla, R. et al. Phylogenetic patterns and phenotypic profiles of the species of plants and mammals farmed for food. *Nat. Ecol. Evol.* **2**, 1808–1817 (2018).
44. Enquist, B. J. et al. The commonness of rarity: global and future distribution of rarity across land plants. *Sci. Adv.* **5**, eaaz0414 (2019).
45. Pyšek, P. et al. The global invasion success of central European plants is related to distribution characteristics in their native range and species traits. *Divers. Distrib.* **15**, 891–903 (2009).
46. Fristoe, T. S. et al. Dimensions of invasiveness: links between local abundance, geographic range size, and habitat breadth in Europe's alien and native floras. *Proc. Natl Acad. Sci USA* **118**, e2021173118 (2021).
47. Sheth, S. N. & Angert, A. L. The evolution of environmental tolerance and range size: a comparison of geographically restricted and widespread *Mimulus*. *Evolution* **68**, 2917–2931 (2014).
48. Pyšek, P. et al. Naturalization of central European plants in North America: species traits, habitats, propagule pressure, residence time. *Ecology* **96**, 762–774 (2015).
49. Hulme, P. E. Trade, transport and trouble: managing invasive species pathways in an era of globalization. *J. Appl. Ecol.* **46**, 10–18 (2009).
50. Fristoe, T. S., Iwaniuk, A. N. & Botero, C. A. Big brains stabilize populations and facilitate colonization of variable habitats in birds. *Nat. Ecol. Evol.* **1**, 1706–1715 (2017).
51. Sunday, J. M., Bates, A. E. & Dulvy, N. K. Global analysis of thermal tolerance and latitude in ectotherms. *Proc. R. Soc. B* **278**, 1823–1830 (2011).
52. Khaliq, I., Hof, C., Prinzinger, R., Böhning-Gaese, K. & Pfenninger, M. Global variation in thermal tolerances and vulnerability of endotherms to climate change. *Proc. R. Soc. B* **281**, 20141097 (2014).
53. Svenning, J.-C., Eiserhardt, W. L., Normand, S., Ordonez, A. & Sandel, B. The influence of paleoclimate on present-day patterns in biodiversity and ecosystems. *Annu. Rev. Ecol. Evol. Syst.* **46**, 551–572 (2015).
54. Seebens, H. et al. Projecting the continental accumulation of alien species through to 2050. *Glob. Change Biol.* **27**, 970–982 (2021).
55. Preston, C. D., Pearman, D. A. & Hall, A. R. Archaeophytes in Britain. *Bot. J. Linn. Soc.* **145**, 257–294 (2004).
56. Ecseri, K. & Honfi, P. Comparison of European archaeophyte lists in the light of distribution data. *Not. Bot. Horti Agrobot. Cluj Napoca* **48**, 480–491 (2020).
57. van Kleunen, M., Bossdorf, O. & Dawson, W. The ecology and evolution of alien plants. *Annu. Rev. Ecol. Evol. Syst.* **49**, 25–47 (2018).
58. Lenzner, B. et al. Role of diversification rates and evolutionary history as a driver of plant naturalization success. *N. Phytol.* **229**, 2998–3008 (2021).
59. Pyšek, P. et al. Naturalized alien flora of the world. *Preslia* **89**, 203–274 (2017).

60. Lonsdale, W. M. Global patterns of plant invasions and the concept of invasibility. *Ecology* **80**, 1522–1536 (1999).
61. Faith, D. P. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **61**, 1–10 (1992).
62. Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105**, 302–314 (2018).
63. Dengler, J. Which function describes the species–area relationship best? A review and empirical evaluation. *J. Biogeogr.* **36**, 728–744 (2009).
64. Diazgranados, M. et al. *World Checklist of Useful Plant Species* (Knowledge Network for Biocomplexity, 2020); https://doi.org/10.5063/F1CV4G34
65. Fouquin, M. & Hugot, J. *Two Centuries of Bilateral Trade and Gravity Data: 1827–2014* (CEPII, 2016).
66. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
67. Broennimann, O. et al. Distance to native climatic niche margins explains establishment success of alien mammals. *Nat. Commun.* **12**, 2353 (2021).
68. *R: A Language and Environment for Statistical Computing* (R Development Core Team, 2022).
69. Dormann, C. F. et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46 (2013).
70. Pinheiro, J., Bates, D. & R Core Team nlme: Linear and Nonlinear Mixed Effects Models http://CRAN.R-project.org/package=nlme (2023).
71. Tung Ho, L. S. & Ané, C. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol.* **63**, 397–408 (2014).
72. Hilbe, J. M. *Logistic Regression Models* (CRC Press, 2009).
73. Hartig, F. DHARMa: Residual Diagnostics for HierARchical Models http://florianhartig.github.io/DHARMa/ (2022).
74. Fristoe, T. S. et al. Evolutionary imbalance, human history, and the global biogeography of alien plants. *Figshare* https://doi.org/10.6084/m9.figshare.21512145 (2023).

Acknowledgements

M.v.K. and N.L.K. thank the German Research Foundation DFG for funding (grants 264740629 and 432253815 to M.v.K.). F.E. appreciates funding by the Austrian Science Foundation FWF (grant no. I 5825-B). P.P. and J.P. were supported by EXPRO grant no. 19-28807X (Czech Science Foundation) and long-term research development project RVO 67985939 (Czech Academy of Sciences). J.-M.D.-D. is an independent ecologist.

Author contributions

T.S.F. and J.B. conceived and designed the study with input from N.L.K., Q.Y., Z.Z. and M.v.K. T.S.F. and J.B. analysed the data. T.S.F. and J.B. wrote the initial draft with input from M.v.K. T.S.F., J.B., N.L.K., Q.Y., Z.Z., W.D., F.E., H.K., J.P., P.P., P. Weigelt, J.-M.D.-D., A.N.S., P. Wasowicz, K.B.W. and M.v.K. contributed data and to subsequent manuscript revisions.

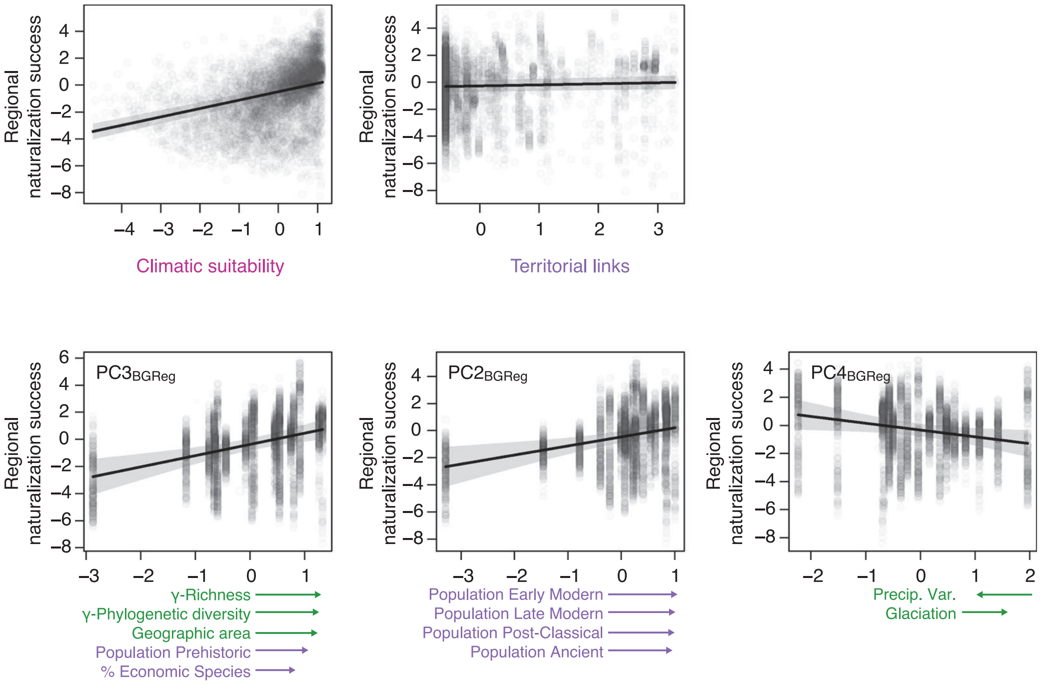
Competing interests

The authors declare no competing interests.

Correspondence and requests for materials should be addressed to Trevor S. Fristoe.

---

[1]Ecology, Department of Biology, University of Konstanz, Konstanz, Germany. [2]Centre for Organismal Studies (COS) Heidelberg, Biodiversity and Plant Systematics, Heidelberg University, Heidelberg, Germany. [3]The German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. [4]Department of Biosciences, Durham University, Durham, UK. [5]BioInvasions, Global Change, Macroecology Group, Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria. [6]Biodiversity, Macroecology & Biogeography, University of Goettingen, Göttingen, Germany. [7]Centre of Biodiversity and Sustainable Land Use (CBL), University of Goettingen, Göttingen, Germany. [8]Department of Invasion Ecology, Institute of Botany, Czech Academy of Sciences, Průhonice, Czech Republic. [9]Department of Ecology, Faculty of Science, Charles University, Prague, Czech Republic. [10]Botanical Museum, Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland. [11]Icelandic Institute of Natural History, Borgir vid Nordurslod, Akureyri, Iceland. [12]Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology, Trondheim, Norway. [13]Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation, Taizhou University, Taizhou, China. [14]Unaffiliated ✉e-mail: trevor.fristoe@uni-konstanz.de



Extended Data Fig. 1 | Drivers of regional naturalization success among biogeographic syndromes (n = 6061 biogeographic syndrome x recipient region comparisons). Partial residual plots from linear mixed model predicting regional naturalization success among biogeographic syndromes (full results Extended Data Table 4). In the bottom panels, main contributors to a given principal component (loading > 0.50) are listed along the x-axes with the length and direction of arrows indicating the value and sign of the variable loading (see Extended Data Table 3; 'BGReg' indicates that PCs for the regional analysis were derived using data for biogeographic syndromes, but not including climate suitability or territorial links). Variables associated with the evolutionary imbalance hypothesis are colored green (Precip. Var. = precipitation variability), anthropogenic variables purple, and climatic suitability in pink. Shaded bands indicate 95% confidence bands.

Extended Data Table 1 | Loadings for varimax rotated principal components of predictors of global naturalization success measured for biogeographic syndromes ('BG')

	PC1$_{BG}$	**PC2$_{BG}$**	**PC3$_{BG}$**	**PC4$_{BG}$**	**PC5$_{BG}$**
Log(γ-Richness)	0.27	**0.86**	0.31	−0.04	0.25
Max α-Richness	**0.8**	0.38	0.35	0.04	0
Mean α-Richness	**0.96**	−0.09	0.05	0.01	−0.08
Log(γ- Phy. Diversity)	0.35	**0.82**	0.29	−0.05	0.3
Max α-Phy. Diversity	**0.92**	0.32	0.14	−0.01	−0.04
Mean α- Phy. Diversity	**0.97**	−0.16	0.01	−0.03	0.02
log(Geographic area)	−0.19	**0.88**	0.26	−0.02	0.24
Temperature variability	**−0.7**	0.36	0.22	0.47	0.08
log(Precipitation variability)	−0.39	0.48	−0.09	**−0.73**	0.03
Glaciation	**−0.54**	0.1	0.19	**0.75**	0.04
Climate Suitability	−0.3	**0.87**	0.23	−0.09	0.03
log(Population Prehistoric)	0.11	**0.73**	**0.56**	0.12	−0.12
log(Population Ancient)	0.03	0.37	**0.92**	0.05	0.08
log(Population Post-Classical)	0.13	0.29	**0.94**	0.06	0.06
log(Population Early Modern)	0.01	0.11	**0.92**	0.06	0.31
log(Population Late Modern)	0.08	0.27	**0.93**	0.11	0.17
Territorial Links	−0.26	0.29	0.36	0.14	**0.72**
% Economic Species	0.14	0.48	0.41	−0.34	**0.59**
Cumulative variance	0.26	0.53	0.78	0.87	0.94

Loadings with an absolute value greater than 0.50 are in bold; less than 0.10 are in grey.

Extended Data Table 2 | Coefficients from linear regression predicting global naturalization success across biogeographic syndromes (n=19)

	Estimate	Std. Error	t-value	p
Intercept	0.30	0.07	4.13	0.001
$PC1_{BG}$	−0.15	0.07	−2.06	0.060
$PC2_{BG}$	0.38	0.07	5.07	< 0.001
$PC3_{BG}$	0.30	0.07	4.28	< 0.001
$PC4_{BG}$	−0.07	0.08	−0.85	0.406
$PC5_{BG}$	0.27	0.08	3.45	0.004
Adjusted $R^2 = 0.74$				

Significance values based on two-sided tests.

Extended Data Table 3 | Loadings for rotated principal components of predictors of regional naturalization success measured for biogeographic syndromes (‘BGReg’ for biogeographic syndrome by region analysis)

	PC1$_{\text{BGReg}}$	PC2$_{\text{BGReg}}$	PC3$_{\text{BGReg}}$	PC4$_{\text{BGReg}}$
γ-Diversity	0.17	0.33	**0.91**	−0.12
Max α-Diversity	**0.76**	0.32	0.46	0.1
Mean α-Diversity	**0 97**	0.04	0.01	0.13
Log(γ- Phy. Diversity)	0.24	0.32	**0.88**	−0.14
Max α-Phy. Diversity	**0.88**	0.13	0.42	0.07
Mean α- Phy. Diversity	**0 97**	0.02	−0.05	0.07
log(Geographic area)	−0.29	0.28	**0.88**	−0.15
Temperature variability	**−0.78**	0.23	0.3	0.36
log(Precipitation variability)	−0.35	−0.1	0.34	**−0.78**
Glaciation	**−0.64**	0.2	0.12	**0.65**
log(Population Prehistoric)	0.03	**0.5**	**0.74**	0.13
log(Population Ancient)	−0.01	**0.9**	0.39	0.04
log(Population Post-Classical)	0.1	**0.93**	0.32	0.07
log(Population Early Modern)	−0.02	**0.97**	0.15	0.01
log(Population Late Modern)	0.03	**0.94**	0.31	0.09
% Economic Species	0.09	**0.5**	**0.54**	−0.48
Cumulative variance	0.29	0.56	0.83	0.93

Climate suitability and territorial links are not included because they were measured for every biogeographic syndrome by region pair. Loadings with an absolute value greater than 0.50 are in bold; less than 0.10 are in grey.

Extended Data Table 4 | Coefficients from linear mixed model predicting regional naturalization success across biogeographic syndromes (n = 6061 biogeographic syndrome x recipient region comparisons)

	Estimate	Std. Error	t-value	p
Intercept	−0.71	0.21	−3.37	< 0.001
$PC1_{BGReg}$	0.28	0.22	1.28	0.222
$PC2_{BGReg}$	0.67	0.22	3.08	0.008
$PC3_{BGReg}$	0.83	0.22	3.80	0.002
$PC4_{BGReg}$	−0.48	0.22	−2.22	0.043
Climate Suitability	0.62	0.04	15.52	< 0.001
Territorial links	0.08	0.04	2.06	0.040
Marginal R^2 = 0.42				
Conditional R^2 = 0.56				

Biogeographic syndrome was included as a random effect (sd = 0.89). The range parameter, approximating the distance beyond which spatial autocorrelation is no longer detectable, was estimated as 819.69 km. Significance values based on two-sided tests.

Extended Data Table 5 | Loadings for varimax rotated principal components of predictors of probability of economic use and probability of naturalization among species ('Spp')

	PC1$_{\mathrm{Spp}}$	**PC2$_{\mathrm{Spp}}$**	**PC3$_{\mathrm{Spp}}$**	**PC4$_{\mathrm{Spp}}$**	**PC5$_{\mathrm{Spp}}$**	**PC6$_{\mathrm{Spp}}$**	**PC7$_{\mathrm{Spp}}$**
Log(γ-Richness)	**0.57**	0.08	**0.78**	−0.06	0.15	0.13	0.04
Max α-Richness	**0.94**	0.15	0.19	−0.07	0.12	0	−0.03
Mean α-Richness	**0.97**	0.09	−0.07	−0.13	0.05	−0.05	−0.04
Log(γ- Phy. Diversity)	**0.59**	0.09	**0.77**	−0.11	0.12	0.05	0.05
Max α-Phy. Diversity	**0.92**	0.11	0.27	−0.13	0.09	−0.1	−0.07
Mean α- Phy. Diversity	**0.95**	0.11	−0.06	−0.18	0.02	−0.15	−0.03
log(Geographic area)	−0.14	0.01	**0.95**	0.04	0.12	0.19	0.09
Temperature variability	−0.4	0.07	0.03	**0.68**	0.3	0.25	0.28
log(Precipitation variability)	−0.14	−0.31	0.13	−0.06	0.02	**0.86**	0.07
Glaciation	−0.18	0.04	−0.02	**0.94**	−0.07	−0.05	0.03
Climate Suitability	−0.06	0.16	**0.53**	0.24	0.11	**0.66**	0
log(Population Prehistoric)	0.23	0.42	0.26	0.05	**0.78**	0.12	0
log(Population Ancient)	0.12	**0.7**	0.17	0.02	**0.64**	0.01	0.08
log(Population Post-Classical)	0.17	**0.83**	0.14	−0.01	0.45	−0.07	0.04
log(Population Early Modern)	0.05	**0.96**	0.01	0.02	0.06	−0.1	0.1
log(Population Late Modern)	0.15	**0.93**	−0.01	0.07	0.06	−0.06	0.08
Territorial Links	−0.09	0.18	0.1	0.13	0.02	0.05	**0.96**
Cumulative variance	0.27	0.35	0.51	0.60	0.66	0.74	0.94

Loadings with an absolute value greater than 0.50 are in bold; less than 0.10 are in grey.

Extended Data Table 6 | Coefficients from phylogenetic binomial regression predicting probability of economic use among species (n = 334,667)

	Estimate	Std. Error	t-value	p
Intercept	−2.61	0.02	−157.53	< 0.001
$PC1_{Spp}$	0.04	0.01	5.97	< 0.001
$PC2_{Spp}$	0.64	0.01	80.35	< 0.001
$PC3_{Spp}$	0.94	0.01	106.71	< 0.001
$PC4_{Spp}$	−0.05	0.01	−5.62	< 0.001
$PC5_{Spp}$	−0.05	0.01	−6.37	< 0.001
$PC6_{Spp}$	0.27	0.01	32.08	< 0.001
$PC7_{Spp}$	0.11	0.00	25.95	< 0.001
				α = 29.79
				R^2_{Lik} = 0.29

The α parameter estimates phylogenetic signal, with low values indicating strong signal and high values low; for a tree scaled to a height of one, values above 50 correspond to negligible phylogenetic signal. Goodness of fit is measured using R^2_{Lik}[75]. Significance values based on two-sided tests.

Extended Data Table 7 | Coefficients from phylogenetic binomial regression predicting probability of naturalization among species (n = 334,667)

	Estimate	Std. Error	t-value	p
Intercept	−3.71	0.04	−84.25	< 0.001
Economic Use	1.83	0.03	57.39	< 0.001
$PC1_{Spp}$	−0.24	0.01	−22.53	< 0.001
$PC2_{Spp}$	0.28	0.01	25.11	< 0.001
$PC3_{Spp}$	0.77	0.01	55.22	< 0.001
$PC4_{Spp}$	0.30	0.01	31.91	< 0.001
$PC5_{Spp}$	0.26	0.01	20.17	< 0.001
$PC6_{Spp}$	0.74	0.01	52.30	< 0.001
$PC7_{Spp}$	0.36	0.01	51.01	< 0.001
Economic Use* $PC1_{Spp}$	0.33	0.02	18.63	< 0.001
Economic Use* $PC2_{Spp}$	0.16	0.02	9.46	< 0.001
Economic Use* $PC3_{Spp}$	−0.50	0.02	−26.02	< 0.001
Economic Use* $PC4_{Spp}$	0.04	0.02	2.84	0.005
Economic Use* $PC5_{Spp}$	−0.13	0.02	−7.16	< 0.001
Economic Use* $PC6_{Spp}$	−0.32	0.02	−15.88	< 0.001
Economic Use* $PC7_{Spp}$	0.04	0.01	3.43	< 0.001
				$\alpha = 12.63$
				$R^2_{Lik} = 0.38$

The α parameter estimates phylogenetic signal, with low values indicating strong signal and high values low; for a tree scaled to a height of one, values above 50 correspond to negligible phylogenetic signal. Goodness of fit is measured using R^2_{Lik}[75]. Significance values based on two-sided tests.

Extended Data Table 8 | Loadings for rotated principal components of predictors of number of naturalized regions among species that have naturalized beyond their native distributions ('Nat')

	PC$_{\text{Nat}}$	**PC2$_{\text{Nat}}$**	**PC3$_{\text{Nat}}$**	**PC4$_{\text{Nat}}$**	**PC5$_{\text{Nat}}$**	**PC6$_{\text{Nat}}$**
log(γ-Richness)	0.26	**0.58**	**0.73**	−0.17	0.12	0.02
Max α-Richness	0.40	**0.83**	0.20	0.08	−0.14	0.12
Mean α-Richness	0.09	**0.90**	0.08	−0.35	0.05	−0.07
log(γ-Phy. Diversity)	0.27	**0.59**	**0.70**	−0.21	0.14	−0.02
Max α-Phy. Diversity	0.29	**0.85**	0.27	−0.03	−0.23	0.02
Mean α-Phy. Diversity	0.12	**0.87**	0.11	−0.38	0.10	−0.12
log(Geographic area)	0.28	0.05	**0.92**	0.05	0.12	0.08
Temperature variability	0.14	−0.44	0.04	**0.67**	0.46	−0.02
log(Precipitation variability)	−0.24	−0.05	0.25	−0.16	0.08	**0.89**
Glaciation	−0.01	−0.24	0.01	**0 92**	0.01	−0.15
Climate Suitability	0.22	0.15	**0.84**	0.11	0.03	0.23
log(Population Prehistoric)	**0.76**	0.26	0.34	0.16	−0.15	0.17
log(Population Ancient)	**0.91**	0.18	0.22	−0.03	0.15	0.02
log(Population Post-Classical)	**0.92**	0.22	0.20	−0.07	0.10	−0.09
log(Population Early Modern)	**0.88**	0.10	0.13	−0.10	0.26	−0.21
log(Population Late Modern)	**0.86**	0.14	0.15	0.14	0.08	−0.18
Territorial Links	0.27	−0.03	0.18	**0.10**	**0.89**	0.08
Cumulative variance	0.26	0.51	0.69	0.79	0.86	0.92

Loadings with an absolute value greater than 0.50 are in bold; less than 0.10 are in grey.

Extended Data Table 9 | Coefficients from phylogenetic linear regression predicting number naturalized regions among species that have naturalized beyond their native distributions (n = 13,280)

	Estimate	Std. Error	t-value	p
Intercept	0.51	0.43	1.19	0.234
Economic Use	0.83	0.02	36.07	< 0.001
$PC1_{Nat}$	0.17	0.02	10.95	< 0.001
$PC2_{Nat}$	0.07	0.02	4.12	< 0.001
$PC3_{Nat}$	0.21	0.01	18.47	< 0.001
$PC4_{Nat}$	0.07	0.01	6.17	< 0.001
$PC5_{Nat}$	0.24	0.02	13.36	< 0.001
$PC6_{Nat}$	0.03	0.01	2.70	0.007
Economic Use* $PC1_{Nat}$	0.10	0.02	4.66	< 0.001
Economic Use* $PC2_{Nat}$	0.08	0.02	3.79	< 0.001
Economic Use* $PC3_{Nat}$	(−0.02)	(0.02)	(−1.14)	(0.255)
Economic Use* $PC4_{Nat}$	(0.02)	(0.02)	(1.10)	(0.270)
Economic Use* $PC5_{Nat}$	0.17	0.02	7.86	< 0.001
Economic Use* $PC6_{Nat}$	(0.02)	(0.02)	(0.87)	(0.384)
$\lambda = 0.38$				
adjusted $R^2 = 0.24$				

Values for interactions removed during model selection are shown in parentheses from the last model before they were removed. Phylogenetic signal is measured as λ, with 0 indicating no phylogenetic signal and values of 1 corresponding to expectations of Brownian motion. Significance values based on two-sided tests.