Science Advances
AAAS

Supplementary Materials for

The poleward naturalization of intracontinental alien plants

Zhijie Zhang *et al.*

Corresponding author: Zhijie Zhang, zhijie.zhang@uni-konstanz.de

Sci. Adv. **9**, eadi1897 (2023)
DOI: 10.1126/sciadv.adi1897

This PDF file includes:

Figs. S1 to S13
Tables S1 and S2
Supplementary Methods

Figures S1-S8

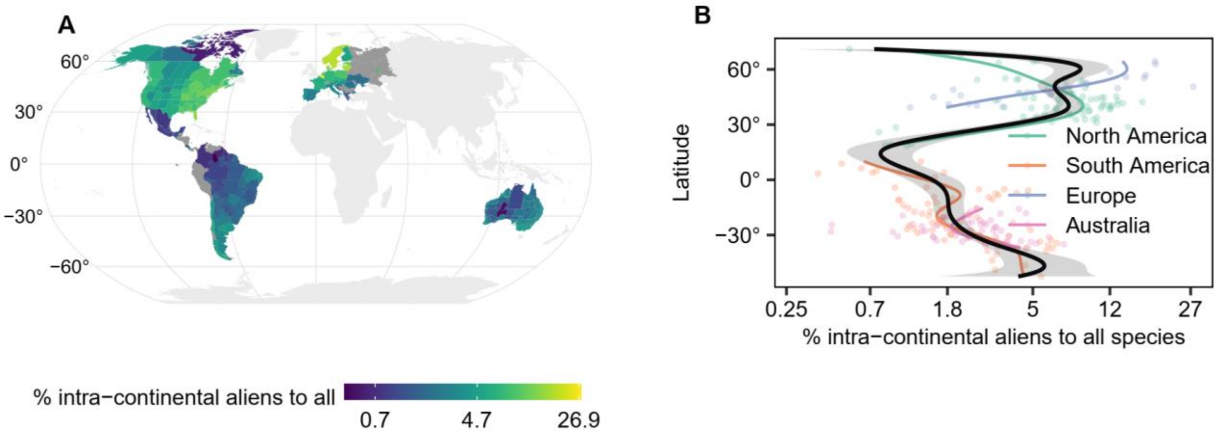


Figure S1 The proportion of intra-continental aliens to all plant species (native and alien) in the four focal continents. A, The proportion in the 243 regions. **B**, The relationship between latitude and the proportion of intra-continental aliens to all plants. The black line represents the general trend across continents, with the shaded area representing the 95% confidence interval. Colored lines represent trends at the continental level. GAMM showed that the peaks at 50 degrees in both hemisphere (edf = 8.33, F =37.1, P < 0.001). Proportion was logit-transformed in both panels.

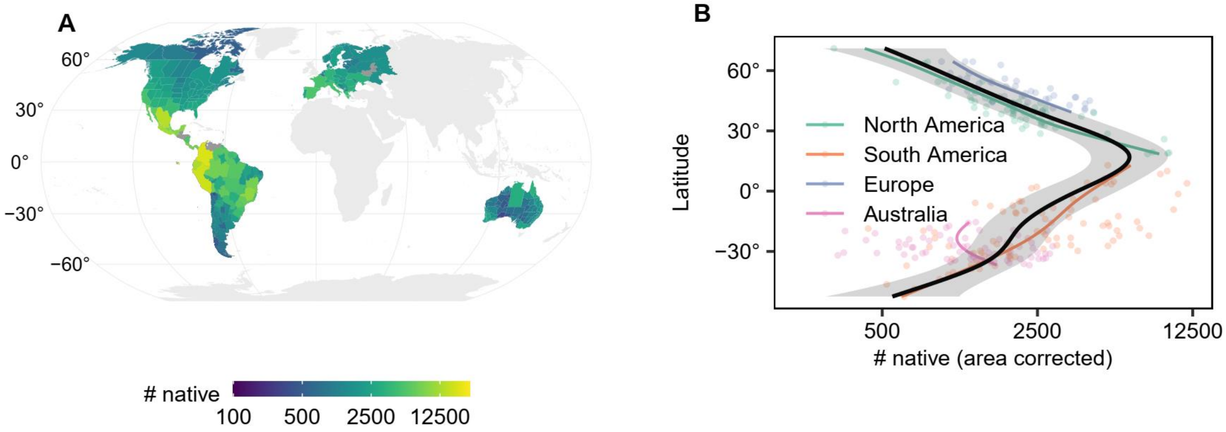


Figure S2 Native plant richness in the four focal continents The number of all native flowering plants in the 264 regions (**A**) and its relationship with latitude (**B**). In **B**, the black line represents the general trend across continents, with the shaded area representing the 95% confidence interval. Colored lines represent trends at the continental level. GAMM showed that the number of all native plants peaks at the equator (edf = 5.93, F = 30.8, $P < 0.001$).

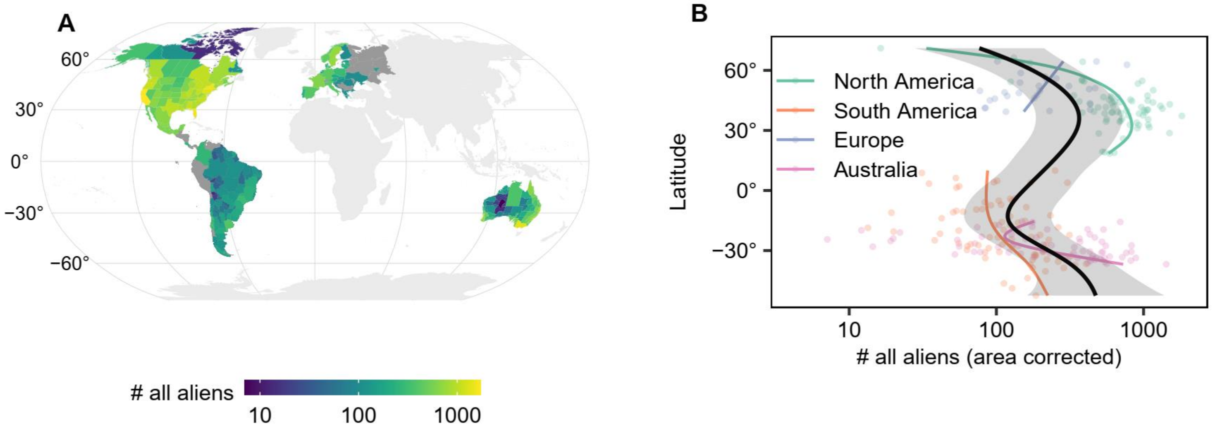


Figure S3 Naturalized alien plant richness in the four focal continents The number of all aliens (including both inter- and intra-continental aliens) in the 243 regions (**A**) and its relationship with latitude (**B**). In **B**, the black line represents the general trend across continents, with the shaded area representing the 95% confidence interval. Colored lines represent trends at the continental level. GAMM showed that the number of all naturalized aliens peaks at around 40 degrees in the Northern Hemisphere, and at around 50 degrees in the Southern Hemisphere (edf = 5.00, F = 8.89, P < 0.001).

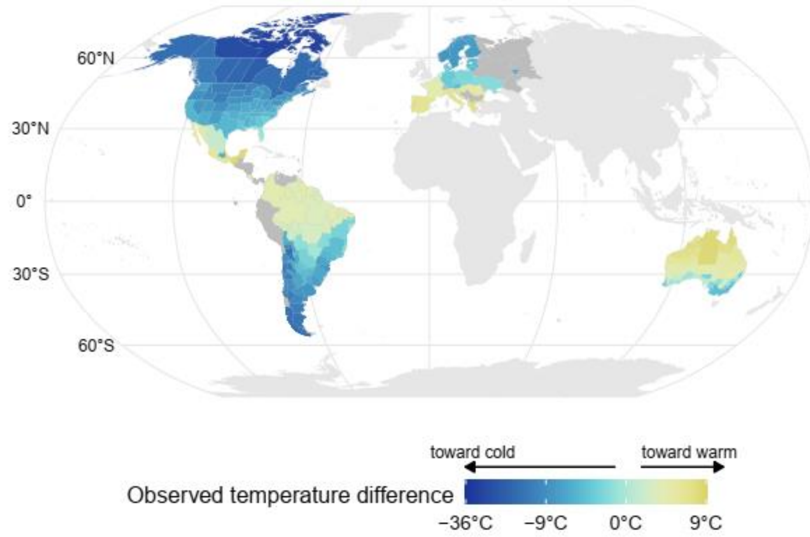


Figure S4 Observed difference in temperature between recipient regions and donor (i.e. native) regions of intra-continental alien plants. For each of the 243 regions, the median temperature difference between recipient and donor regions of naturalized intra-continental aliens was calculated. Blue shades indicate regions that receive mainly intra-continental aliens from warmer regions (i.e. intra-continental aliens naturalize to colder regions), while yellow shades indicate the opposite.

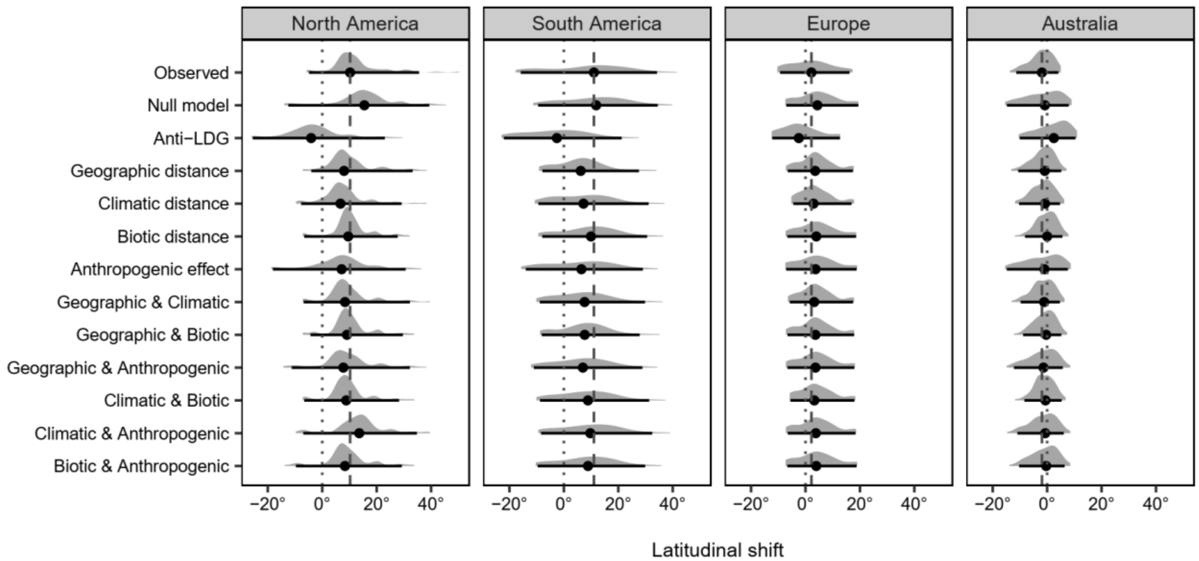


Figure S5 Observed and predicted latitudinal shifts of intra-continental alien plants. The null model assumed that the intra-continental aliens in a recipient region were randomly drawn (migrated) from the other regions. The 'anti-latitudinal-gradient' model offsets the latitudinal gradient in native species richness (e.g. species were less likely to be drawn from species rich latitudes, which usually are the low latitudes). Models informed by a single predictor assumed that the probability of a species migrating (with the help of humans) to a recipient region decreased with geographic, climatic or biotic distance between its native region(s) and the recipient region. Models informed by multiple drivers assumed that the probability of a species migrating to a recipient region was jointly determined by two or all of the three predictors (see Methods). Distributions of the latitudinal shifts are plotted with the median and 95% probability interval. Dotted lines indicate no latitudinal shift and dashed lines indicate the medians of the observed latitudinal shifts.

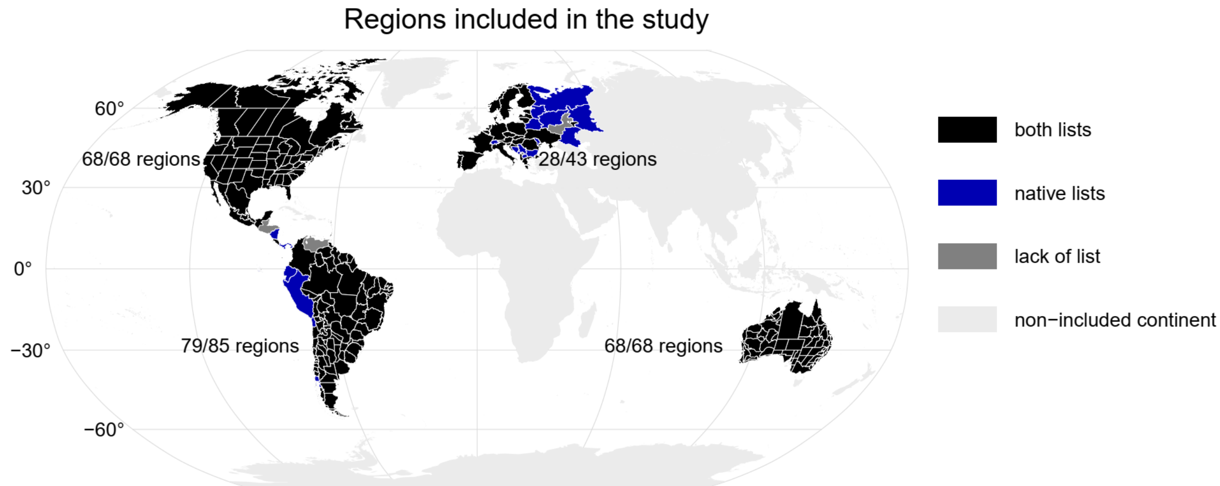


Figure S6 Regions included in the study. Black indicates regions that have both native and naturalized alien species lists. Blue indicates regions that have only native species lists and thus were used for identification of native ranges within the continents (e.g. Peru in South America). Dark gray color indicates regions lacking data (e.g. Honduras in South America) and light gray indicates continents or parts of a continent (i.e. the Northern Hemisphere regions of South America) not included in the present study (e.g. Africa). The numbers besides continents indicate the number of regions that have both native and alien lists and the number of regions (i.e. black) that have native lists (i.e. black + blue). South American regions that are located north of the equator were not included in the analysis on latitudinal shifts.

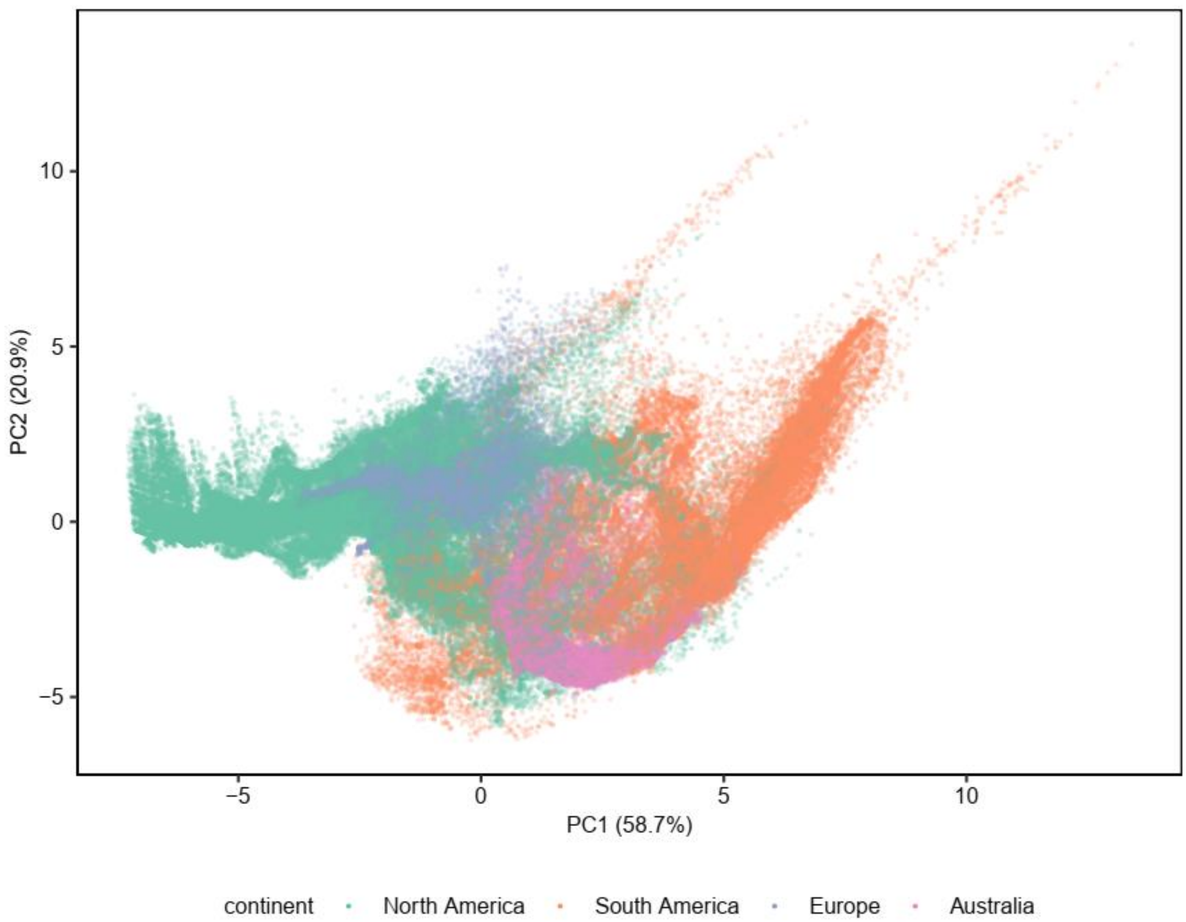


Figure S7 Principal component analysis (PCA) of the 19 bioclimatic variables. The raw data contains 3,672,571 locations (grid cells). To aid illustration, we randomly selected 5% of the data. Different colors indicate different continents.

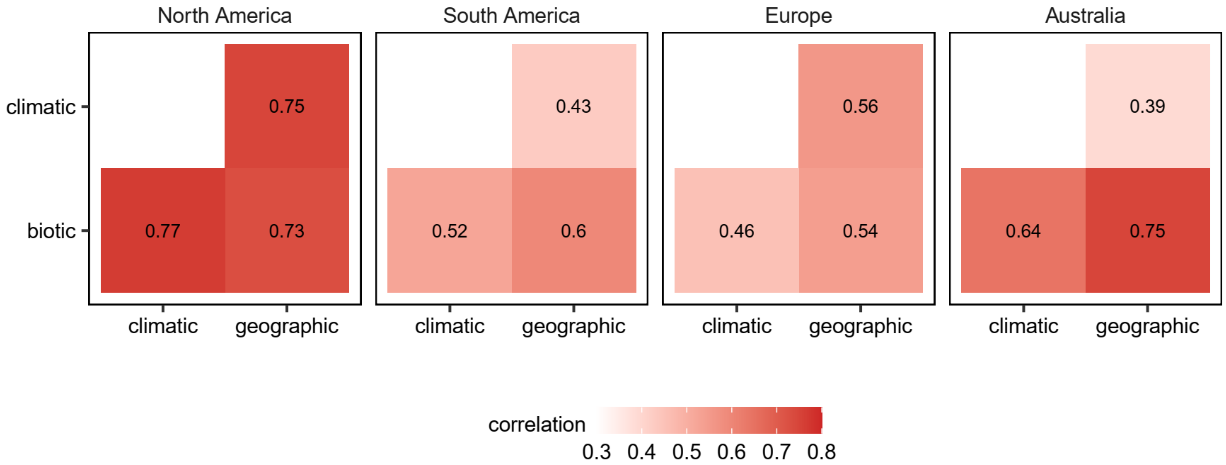


Figure S8 Pairwise correlations between geographic distance, climatic distance and biotic distance between regions in the four continents. The numbers are the Pearson correlation coefficients. In all cases, the correlations are significant (t > 10, P < 0.001)

Tables S1-S2

Table S1 Differences between observed and predicted latitudinal shifts[1].

	North America				South America				Europe				Australia			
	Estimate	Z	P		Estimate	Z	P		Estimate	Z	P		Estimate	Z	P	
Null model *vs.* Observed	3.303	7.5	<0.001	*	3.218	8.81	<0.001	*	2.57	12.7	<0.001	*	1.339	5.4	<0.001	*
Anti-LDG vs. Observed	-15.38	-34.9	<0.001	*	-10.996	-30.1	<0.001	*	-3.912	-19.3	<0.001	*	4.97	20.1	<0.001	*
Geographic distance vs. Observed	-2.099	-4.76	<0.001	*	-1.87	-5.12	<0.001	*	1.969	9.76	<0.001	*	1.246	5.03	<0.001	*
Climatic distance vs. Observed	-4.348	-9.87	<0.001	*	-0.841	-2.3	0.024	*	1.854	9.19	<0.001	*	1.062	4.28	<0.001	*
Anthropogenic effect vs Observed	-1.545	-3.51	<0.001	*	1.254	3.43	0.001	*	2.326	11.5	<0.001	*	2.475	9.99	<0.001	*
Biotic similarity vs. Observed	-5.208	-11.8	<0.001	*	-2.018	-5.53	<0.001	*	2.083	10.3	<0.001	*	1.17	4.72	<0.001	*
Geographic & Climate vs Observed	-2.624	-5.96	<0.001	*	-1.052	-2.88	0.005	*	1.909	9.46	<0.001	*	1.155	4.66	<0.001	*
Geographic & Biotic vs Observed	-1.759	-3.99	<0.001	*	-0.519	-1.42	0.17		2.049	10.2	<0.001	*	1.88	7.59	<0.001	*
Geographic & Anthroc vs Observed	-3.038	-6.89	<0.001	*	-1.267	-3.47	0.001	*	2.009	9.96	<0.001	*	1.109	4.47	<0.001	*
Climate & Biotic vs Observed	-2.632	-5.97	<0.001	*	0.176	0.48	0.629		2.09	10.4	<0.001	*	1.639	6.61	<0.001	*
Climate & Anthroc vs Observed	1.757	3.99	<0.001	*	1.506	4.12	<0.001	*	2.215	11.0	<0.001	*	1.26	5.09	<0.001	*
Biotic & Anthro vs Observed	-2.875	-6.53	<0.001	*	0.355	0.97	0.346		2.349	11.6	<0.001	*	1.84	7.42	<0.001	*
Anti-LDG vs Null	-18.683	-42.4	<0.001	*	-14.214	-38.9	<0.001	*	-6.482	-32.1	<0.001	*	3.631	14.7	<0.001	*
Geographic distance vs Null	-5.402	-12.3	<0.001	*	-5.088	-13.9	<0.001	*	-0.6	-2.98	0.004	*	-0.093	-0.38	0.74	
Climate distance vs Null	-7.651	-17.4	<0.001	*	-4.059	-11.1	<0.001	*	-0.716	-3.55	0.001	*	-0.277	-1.12	0.336	
Biotic distance vs Null	-4.848	-11	<0.001	*	-1.964	-5.38	<0.001	*	-0.244	-1.21	0.237		1.136	4.58	<0.001	*
Anthropogenic factor vs Null	-8.511	-19.3	<0.001	*	-5.236	-14.3	<0.001	*	-0.486	-2.41	0.019	*	-0.17	-0.68	0.541	
Geographic & Climate vs Null	-5.927	-13.5	<0.001	*	-4.271	-11.7	<0.001	*	-0.661	-3.28	0.002	*	-0.184	-0.74	0.526	
Geographic & Biotic vs Null	-5.062	-11.5	<0.001	*	-3.737	-10.2	<0.001	*	-0.52	-2.58	0.013	*	0.541	2.18	0.045	*
Geographic & Anthro vs Null	-6.34	-14.4	<0.001	*	-4.485	-12.3	<0.001	*	-0.56	-2.78	0.007	*	-0.23	-0.93	0.427	
Climate & Biotic vs Null	-5.935	-13.5	<0.001	*	-3.042	-8.33	<0.001	*	-0.48	-2.38	0.02	*	0.299	1.21	0.307	
Climate & Anthro vs Null	-1.546	-3.51	<0.001	*	-1.713	-4.69	<0.001	*	-0.355	-1.76	0.086	†	-0.079	-0.32	0.751	
Biotic & Anthro vs Null	-6.178	-14.0	<0.001	*	-2.864	-7.84	<0.001	*	-0.221	-1.09	0.274		0.501	2.02	0.062	†

We conducted LMMs for each continent. Like the analysis for predictive accuracy, the LMMs included latitudinal shift as response variable, type of approach (e.g. observed or prediction based on the null model) as the fixed effect, and identity of region as the random effect. Multiple comparisons were applied to test for differences between observed patterns vs. predicted patterns and between predictions based on the null model vs. another model.

Table S2 Differences in predictive accuracy among models.

	North America				South America				Europe				Australia			
	Estimate	Z	P		Estimate	Z	P		Estimate	Z	P		Estimate	Z	P	
Anti-LDG < Null model[1]	1.317	6.95	<0.001	*	1.153	6.45	<0.001	*	0.562	1.9	0.087	†	1.017	6.26	<0.001	*
Informed > Null model[2]																
Geographic distance > Null	0.906	4.78	<0.001	*	-0.006	-0.03	0.718		1.1	3.71	0.001	*	0.458	2.82	0.01	*
Climate distance > Null	0.714	3.77	<0.001	*	0.059	0.33	0.597		1.233	4.16	<0.001	*	0.689	4.24	<0.001	*
Biotic distance > Null	0.908	4.79	<0.001	*	0.302	1.69	0.191		0.11	0.37	0.574		-0.064	-0.39	0.857	
Anthropogenic factor > Null	0.332	1.75	0.069	†	0.315	1.76	0.191		0.311	1.05	0.28		0.054	0.33	0.517	
Geographic & Climate > Null	0.842	4.44	<0.001	*	0.071	0.4	0.597		0.943	3.18	0.004	*	0.568	3.5	0.001	*
Geographic & Biotic > Null	0.88	4.64	<0.001	*	0.063	0.35	0.597		0.455	1.54	0.131		0.15	0.93	0.31	
Geographic & Anthro > Null	0.538	2.84	0.005	*	0.325	1.81	0.191		0.458	1.54	0.131		0.228	1.41	0.168	
Climate & Biotic > Null	0.79	4.17	<0.001	*	0.14	0.78	0.457		0.504	1.7	0.117		0.336	2.07	0.051	†
Climate & Anthro > Null	0.259	1.37	0.129		0.046	0.26	0.597		0.224	0.76	0.393		0.43	2.64	0.014	*
Biotic & Anthro> Null	0.627	3.31	0.001	*	0.302	1.69	0.191		0.086	0.29	0.578		0.125	0.77	0.356	
Informed > Average of all[3]																
Geographic distance > Average	0.252	1.78	0.069	†	-0.186	-1.4	0.919		0.62	2.81	0.011	*	0.179	1.48	0.164	
Climate distance > Average	0.038	0.27	0.485		-0.114	-0.86	0.873		0.767	3.47	0.002	*	0.435	3.6	0.001	*
Biotic distance > Average	0.253	1.79	0.069	†	0.156	1.17	0.282		-0.481	-2.18	0.989		-0.402	-3.32	1	
Anthropogenic factor > Average	-0.386	-2.73	1		0.17	1.27	0.282		-0.257	-1.16	0.989		-0.27	-2.23	1	
Geographic & Climate > Average	0.18	1.28	0.141		-0.101	-0.75	0.873		0.445	2.02	0.077	†	0.301	2.48	0.02	*
Geographic & Biotic > Average	0.223	1.58	0.093	†	-0.11	-0.82	0.873		-0.097	-0.44	0.827		-0.163	-1.35	1	
Geographic & Anthro > Average	-0.157	-1.11	0.958		0.181	1.36	0.282		-0.094	-0.43	0.827		-0.077	-0.64	0.911	
Climate & Biotic > Average	0.123	0.87	0.252		-0.025	-0.18	0.752		-0.043	-0.19	0.807		0.043	0.35	0.517	
Climate & Anthro > Average	-0.468	-3.31	1		-0.128	-0.96	0.873		-0.354	-1.6	0.989		0.147	1.21	0.215	
Biotic & Anthro> Average	-0.059	-0.42	0.773		0.156	1.17	0.282		-0.507	-2.29	0.989		-0.191	-1.58	1	

[1]Whether the null model outperformed the anti-latitudinal-diversity-gradient model (one-tailed test)

[2]Whether the informed model outperformed the null model (one-tailed test).

[3]Whether the informed model outperformed the average of all other informed models (one-tailed test).

Supplementary Method 1 Species-level analyses

1.1 A GLMM on status ~ latitude

In addition to the region-level analyses reported in the main text, we also did an analysis at the species level. To test the naturalization of intra-continental aliens at the species level, we started with a binomial generalized linear-mixed effect model (GLMM) using *lme4* package. In the model, we included records of region status of intra-continental aliens (native or naturalized) as the response variable, absolute latitude of the region as the fixed effect, and species, genus and continent as the random effects. Significances of the fixed effects were assessed with likelihood-ratio tests (type II) with the *car* package.

Results

We found that most of the naturalized records of intra-continental aliens occur at higher latitudes than their native records ($\chi2 = 3654.2$, $P < 0.0001$), indicating that intra-continental aliens predominantly naturalized towards the poles.

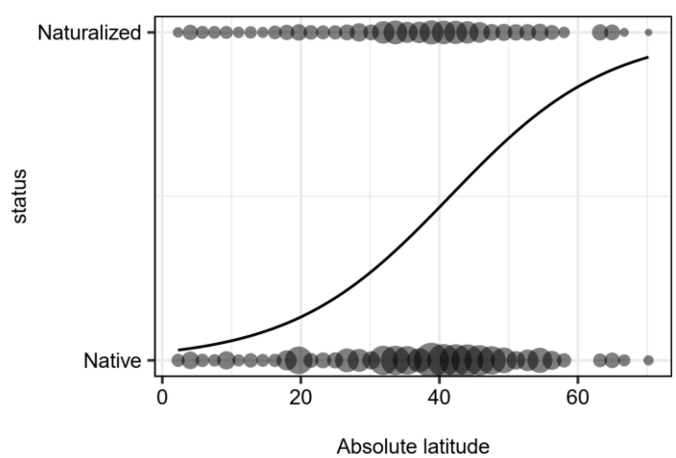


Figure S9 The relationship between region status and absolute latitude. The size of circles indicates the number of species at a certain latitude.

1.2 Latitudinal shifts at the species level

1.2.1 Quantifying median latitudinal shifts of intra-continental aliens

We quantified observed latitudinal shifts of intra-continental aliens as latitudinal changes from their native donor region(s) to the recipient region(s) at the species level. Specifically, for each intra-continental alien, we calculated the native latitude as the centroid latitude of its donor region, and the naturalized latitude as the centroid latitude of its recipient region. For species with multiple donor (or recipient) regions, we calculated the representative latitude as the median latitude of all regions. We used absolute values of latitude so that a high value indicates that the species is native to regions that are distant to the equator, irrespective of whether this is in the Southern or Northern Hemisphere. As in the main text, the 12 Northern-Hemisphere regions in South America were removed from the analyses.

Results

Across the four continents, intra-continental aliens naturalized towards higher latitudes by 5.2 latitude degrees. The same pattern was consistently found in each of the four continents, with the latitudinal shifts in Australia less apparent (Fig. S10).

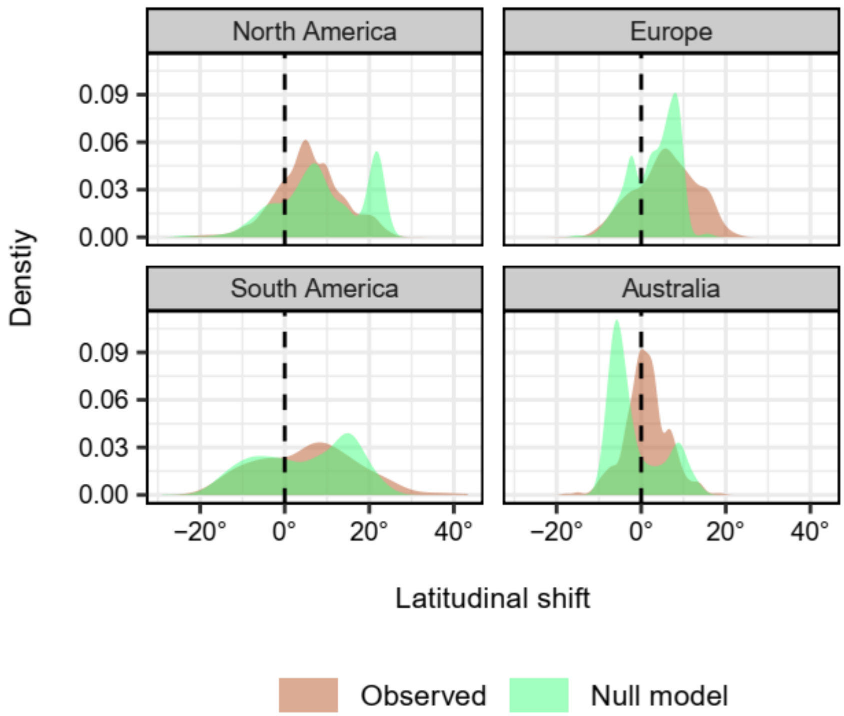


Figure S10 Observed and predicted latitudinal shifts of intra-continental alien plants in the four focal continents. For each of the naturalized intra-continental aliens, the median latitudinal shift was calculated. The prediction is based on a null model where each intra-continental alien was randomly pushed from its donor region(s) to recipient (nonnative) regions.

1.2.3 Modeling latitudinal shifts of intra-continental aliens and assessing the predictive accuracy

To predict median latitudinal shifts of intra-continental alien plants and assess their predictive accuracy at the species level, we ran models that are similar to those of the region-level analyses (in the main text). The difference is that, instead of drawing native species to each recipient region, here, we pushed each intra-continental alien from its donor region(s) to recipient regions either randomly or based on anthropogenic effect of the recipient region, geographic distance, climatic distance or biotic distance between the recipient and donor regions. Because this model focuses on individual species, the effect of the latitudinal diversity gradient cannot be tested.

For a given recipient species (species *j*), we calculated the predictive accuracy (acc_j) as the deviation between the observed ($shift\_obs_i$) and predicted latitudinal distance ($shift\_pred_i$), normalized to its largest latitudinal distance between nonnative regions.

$$acc_j = 1 - \frac{|shift\_obs_j - shift\_pred_j|}{|margin_{j,north} - margin_{j,south}|} \qquad (2)$$

Here $margin_{j,south}$ and $margin_{j,north}$ are the latitudes of the most southern and the most northern nonnative regions of species *j*. A high value of acc_j indicates a high predictive accuracy. For example, acc_j = 1 indicates that observed and predicted latitudinal shifts perfectly match for species *j*.

To assess model predictive accuracy and thus the potential determinants that best explain the latitudinal shifts of intra-continental aliens in each of the four continents, we conducted linear mixed-effect models (LMMs) for each continent with the *nlme* package. The LMMs included predictive accuracy as the response variable, type of modeling approach (i.e. null model and the informed models) as the fixed effect, and species and genus as the random effects (i.e. random intercepts). Because species that have more naturalized regions have larger sample sizes, we weighted the data according to the log-transformed (number of naturalization regions + 1). To improve the normal distribution of the residuals, we logit-transformed the predictive accuracy. We used multiple comparisons, with the *multcomp* package, to test 1) whether the informed models had higher predictive accuracies than the null model, and 2) which informed model(s) had higher predictive accuracies than the average of the other informed models. Significances of fixed effects were assessed with Z-tests, corrected with the Benjamini-Hochberg procedure to control for Type I errors in multiple comparisons.

Results

Consistent with the region-level analyses (in the main text), the null model, which randomly pushed intra-continentals to nonnative regions, predicted the observed latitudinal shifts reasonably well (Fig. S11). The models informed by climatic distance outperformed the null models in North America and Australia. Unlike the region-level analyses, the informed models did not outperform the null model in Europe. It is worth noting the species-level analyses focus on the variation among recipient regions. For example, in the geographic-distance informed model, an intra-continental alien is more likely to naturalize in a recipient region that is close to its donor regions than a recipient region that is far away. In contrast, the region-level analyses focus on variation among donor regions. Together, the two analyses indicate that in Europe, the variation among donor regions contributes more to naturalization of intra-continental alien than that among recipient regions.

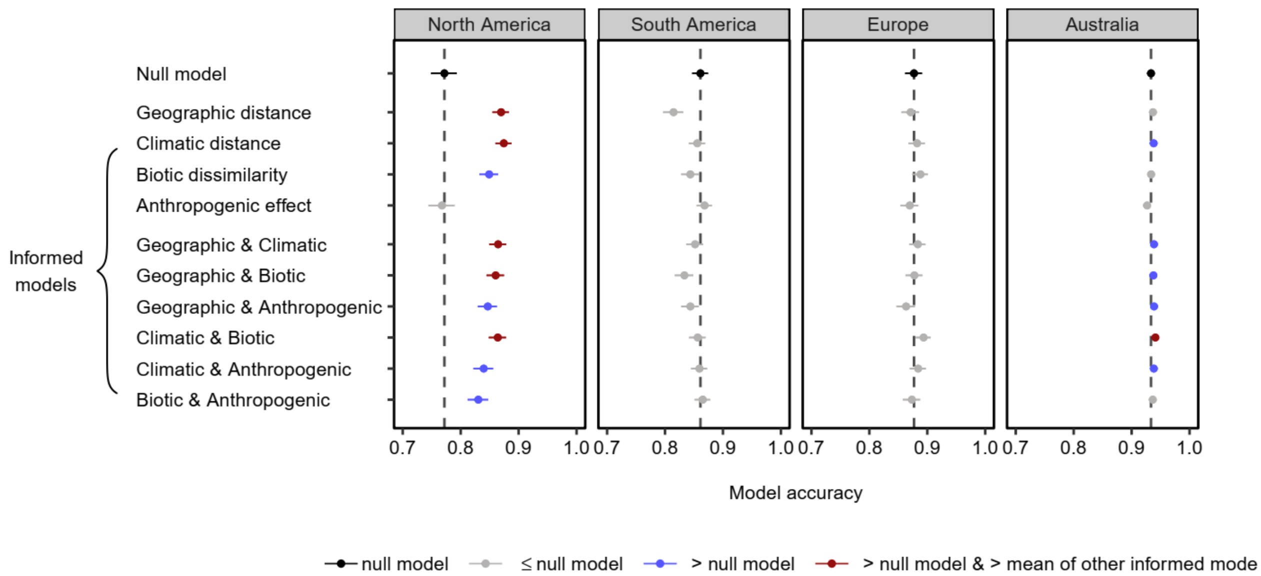


Figure S11 Accuracy of models used to predict latitudinal shifts of intra-continental alien plants. The null model assumed that the intra-continental aliens were randomly pushed (naturalized) from donor to recipient (nonnative) regions. Models informed by a single predictor assumed that the probability of a species naturalizing to a recipient region increased with the anthropogenic factor (GDP per capita) of the recipient region, or decreased with geographic, climatic or biotic distance between its donor and recipient regions. Models informed by multiple predictors assumed that the probability of a species naturalizing to a recipient region was jointly determined by two or all of the three predictors. Gray indicates models that did not outperform the null model (vertical dashed line), blue indicates models that outperformed the null model, and red indicates models that outperformed both the null model and the average of all other informed models. Error bars represent 95% confidence intervals.

1.2.3 Variation in the direction of latitudinal shifts

For a given species, even if its median latitudes show naturalization towards a certain direction, there could be variation in changes of its lowest and highest latitudes (i.e. range boundaries). Consequently, for each species, we quantified the shifts of its lowest and highest latitudes. Then, we quantified the proportion of species that expanded both range boundaries, one of the two, and neither of the two.

Results

Across the four continents, we found that while 49.8% of the species expanded only their poleward boundaries, 9.2%, 9.8 and 31.2% expanded only the equatorward boundaries, both boundaries and neither of the boundaries, respectively. (Fig. S12). This pattern is consistent across the four continents.

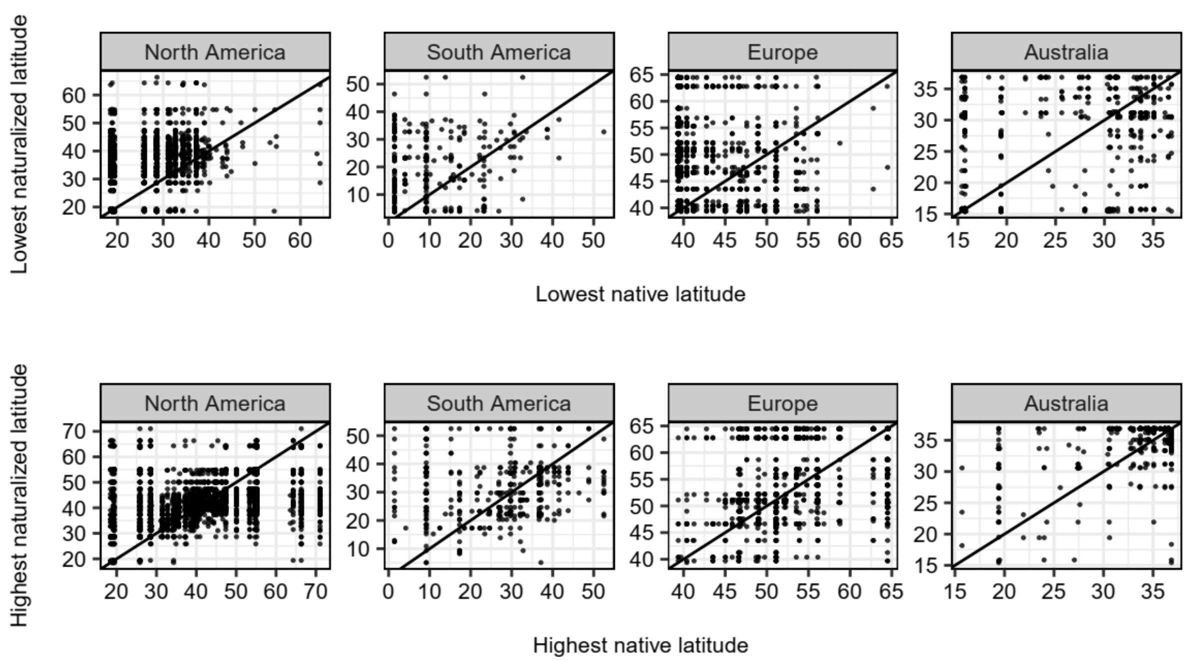


Figure S12 The relationship between boundaries of the native and non-native regions of intra-continental aliens. The upper panels show the equatorward boundaries (i.e. the lowest absolute latitudes), and the lower panels show the poleward boundaries. Each point represents an intra-continental alien species. One-one lines are added to aid visualization.

Supplementary Method 2 An alternative approach of combining multiple factors for predicting latitudinal shifts.

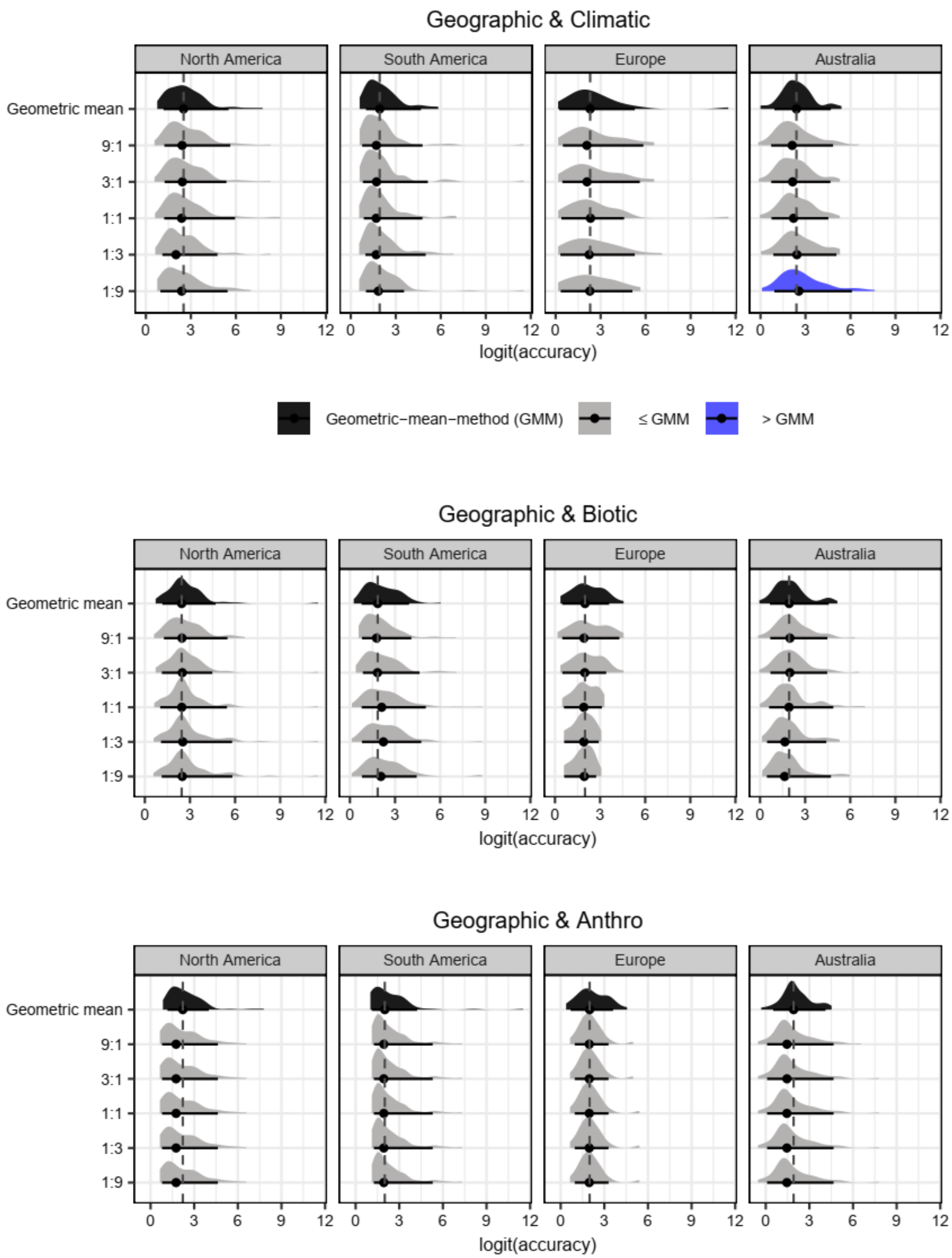
For the models that combine multiple factors, we assumed that P_{ij} (the probability of species *j* being drawn to region *i*) is proportional to the geometric mean of the factors of interest. For example, when combining geometric and climatic distance, P_{ij} is proportional to w_{ij}, where

$$w_{ij} = \sqrt{w_{ij,geometric\ distance} \times w_{ij,climatic\ distance}}$$

An alternative approach is assigning each factor a certain weight and calculating the sum (i.e. weighted sum), where

$$w_{ij} = c_{geo} \times w_{ij,geometric\ distance} + c_{clim} \times w_{ij,climatic\ distance}$$

We ran a series of models using different combinations of weights (9:1, 3:1, 1:1, 1:3 and 1:9), and tested whether these models had different predictive accuracies with the geometric-mean-approach. We found that only in 6 out of 120 cases, the alternative method outperformed the geometric-mean-method, revealing the reliable predictability of the latter approach. Furthermore, when the alternative method did outperform the geometric-mean-method, this was overall continent-dependent rather than method-dependent (Figure S6, indicated by blue). For example, when combining climatic and biotic distances (the upper panel in Figure S6) for regions in Australia, the models that assigned a relatively high weight to climatic distance always outperformed the geometric-mean approach. This is because climatic distance is a strong predictor of latitudinal shifts in Australia. Combining biotic dissimilarity, a poor predictor, with climatic distance will weaken the predictive accuracy.


Geographic & Climatic
North America
South America
Europe
Australia
Geometric mean
9:1
3:1
1:1
1:3
1:9
0
3
6
9
12
logit(accuracy)
Geometric−mean−method (GMM)
≤ GMM
> GMM
Geographic & Biotic
Geographic & Anthro

Continued

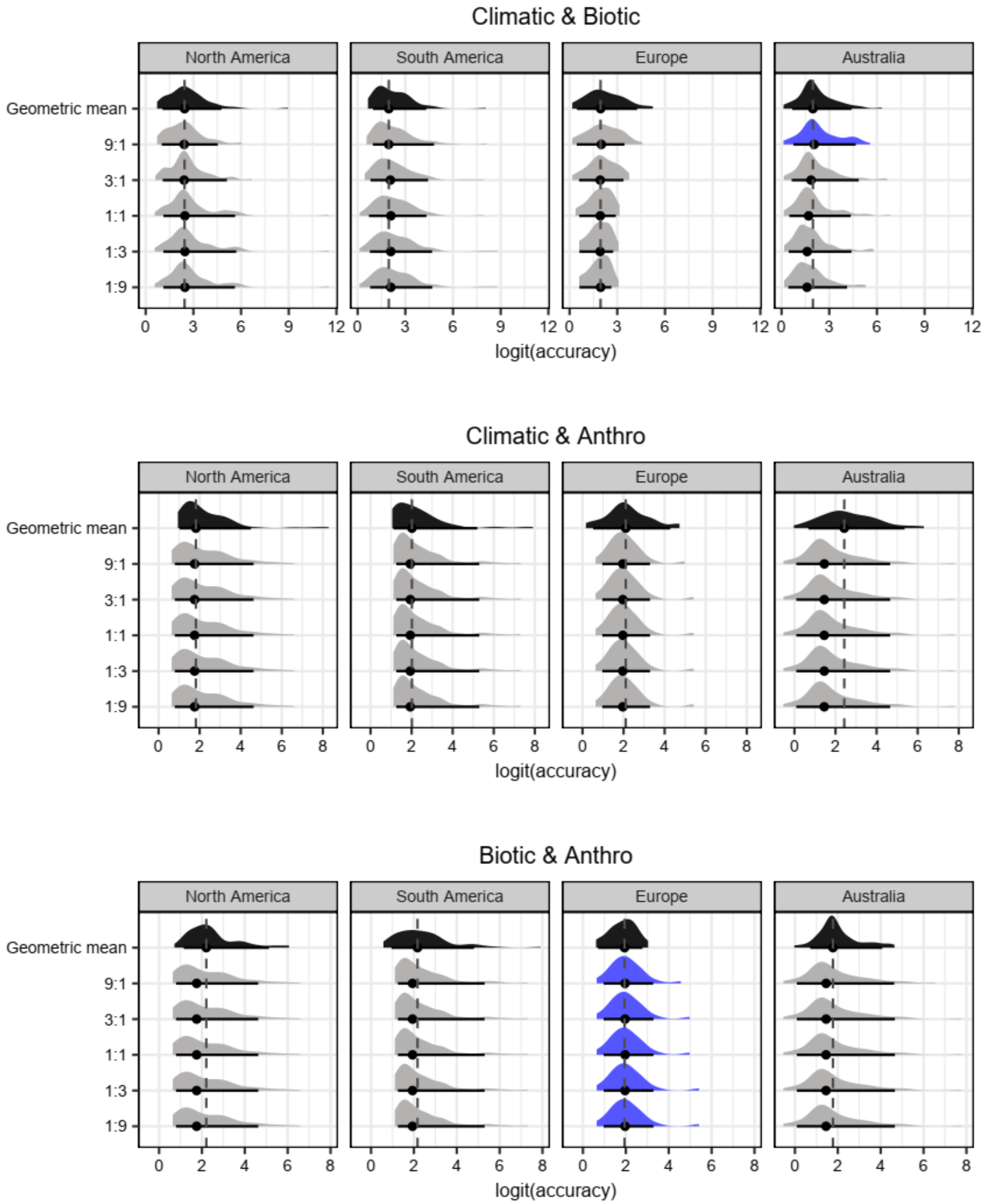


Figure S13 Accuracy of models that combined two factors as geometric mean or as weighted sum. The ratio along the y axis indicates the weight for the two factors of interest. For example, in the upper panel, '9:1' indicates that the weights of climatic and biotic distances are 9 and 1, respectively. Blue indicates models that outperformed the geometric-mean-approach.

Supplementary Method 3 The calculation of predicted latitudinal shifts

For a recipient region (region *i*), there are a total of *n* species that are native to its continent but not native to it. In our models, the probability (P_{ij}) of species *j* being drawn (i.e. to become naturalized to region *i*) is proportional to a weighting factor (w_{ij}). In a null model, w_{ij} is a constant. In an informed model, w_{ij} is the predictor (e.g. inverse of the climatic distance between region *i* and native region(s) of species *j*).

We calculated the predicted overall native latitudes of the *n* species ($native\_exp_i$) as the weighted median of the native latitude of the *n* species as follows:

1. We arranged the native latitude of the *n* species in ascending order: $native_{i1}$, $native_{i2}$, …, $native_{in}$.
2. We scaled the $w_{i1}, w_{i2}, \ldots, w_{in}$ such that $\sum_{j=1}^{n} w_{ij} = 1$.
3. We identified the $native\_exp_i$ (the weighted median) as the $native_{ik}$ satisfying the condition

$$\sum_{j=1}^{k-1} w_{ij} \leq \frac{1}{2} \text{ and } \sum_{j=k+1}^{n} w_{ij} \geq \frac{1}{2}.$$

In cases where two of the elements satisfy the *condition*, the weighted median is the median of the two elements.

We then calculated the expected overall latitudinal shift ($shift\_pred_i$) as the difference between the latitude of the region *i* and the predicted native latitude of the intra-continental aliens in the recipient region ($native\_pred_i$).

It can be seen from the calculation that, the calculated predicted latitudinal shift ($shift\_pred_i$) does not depend on the number of intra-continental alien species in the recipient region. In addition,

$shift\_pred_i$ will be similar to the result of a resampling test where a certain number of species are repeatedly drawn to the region *i*. With the resampling test increasing the number replicates, its median result will eventually approach to $shift\_pred_i$ (as mathematically it is supposed to be).