

# The nature of dispersal barriers and their impact on regional species pool richness and turnover

Jeffrey C. Nekola  | Jan Divíšek | Michal Horskák

Department of Botany and Zoology,  
Masaryk University, Brno, Czech Republic

## Correspondence

Jeffrey C. Nekola, Department of Botany  
and Zoology, Masaryk University,  
Kotlářská 2, CZ-611 37 Brno, Czech  
Republic.

Email: [nekola@sci.muni.cz](mailto:nekola@sci.muni.cz)

## Funding information

Grantová Agentura České Republiky,  
Grant/Award Number: 20-188275

## Abstract

**Aim:** We document realized and potential global species ranges based on empirically vetted species concepts in conjunction with global climate databases and climate suitability modelling. From this we investigate the nature of dispersal barriers and illustrate how they generate ecological uniqueness.

**Location:** Holarctic.

**Methods:** Fifty-two small body-size (i.e. < 5 mm) land snail taxa within the genera *Euconulus*, *Pupilla* and *Vertigo* were considered. These represent ~10% of all small body-size Holarctic land snails and are among the most proficient known passive dispersers. Their potential climatic ranges were determined using Maxent modelling based on 9205 occurrence records. From these we inferred the location, width and nature of isolating barriers and tested for their effects on regional species pool richness and turnover.

**Results:** Use of unvetted traditional taxonomic concepts and unverified occurrence records would have created up to threefold higher or lower estimates of species-specific climatic tolerances than the actual values. Modelling must thus only use high quality occurrence data. All but one taxon were shown at a global scale to possess multiple isolated areas of appropriate climate. While oceans represented the most common barrier (37%), intra-continental barriers were in total almost twice as frequent (inappropriate climate – 29%, habitat/history – 27% and the Greenland ice sheet – 7%). These barriers restricted taxa to only a subset of their potential range, with European taxa possessing approximately twice the global occupancy rates as North American ones (median scores of 62 vs. 34%). As a result, regional taxa pools were three times smaller than their potential sizes, with 50% change in composition occurring over ~2600-km distances.

**Main conclusions:** Even for these readily dispersing taxa, isolation barriers prevented species from saturating their potential global range, reduced the size of regional species pools by 2/3, and generated ecological uniqueness between them.

## KEYWORDS

climate suitability modelling, dispersal limitation, distinctiveness, Holarctic, isolation barriers, terrestrial gastropods

## 1 | INTRODUCTION

The fundamental role of dispersal barriers in stimulating biological diversification and biogeographic uniqueness has been suggested since the dawn of modern biology (e.g. Darwin, 1859; Humboldt & Bonpland, 1805). Evolutionary biologists have been early and strong advocates of isolation as a fundamental mechanistic driver (e.g. Jordan, 1905; Mayr, 1942). Biogeographers have identified ocean, mountain and adverse climate barriers as playing an essential role in the formation of global biogeographic regions (Lomolino et al., 2010; Riddle & Hafner, 2010). And, over the last few decades there has also been an increasing awareness of the role of barriers in generating ecological pattern from local (Hubbell, 2001) to regional (Hanski, 1999; Leibold & Chase, 2018), continental and global (Preston, 1960; Ricklefs & Schluter, 1993; Shmida & Ellner, 1984) scales.

In spite of this, the quantitative aspects of barriers and their influence on species pools remain relatively poorly investigated, especially at the global scale. This is likely due to inherent difficulties in their empirical documentation, which requires proof of two different forms of a negative (Lobo et al., 2010): (a) documentation of regions falling outside of a species' niche, and (b) identifying unoccupied areas that possess appropriate environmental conditions for establishment and persistence of a given species. It is thus one thing to say 'There is hardly a climate or condition in the Old World which cannot be paralleled in the New... [with] barriers of any kind, or obstacles to free migration, [being] related in a close and important manner [to their biological] differences...' (Darwin, 1859) and quite another to prove it.

However, our ability to quantitatively document, analyze and map species ranges – and thus barriers – across global extents has undergone a revolution over the last few decades. While little precise environmental data were once available, there are now multiple resources available providing – at least for climate – global coverage of dozens of variables at fine spatial resolution not only for current (e.g. Hijmans et al., 2005; Karger et al., 2017) but also past (ranging back to the Last Glacial Maximum – LGM; e.g. Brown et al., 2018; Tittle & Bemmels, 2018) and potential future (e.g. Navarro-Racines et al., 2020) conditions. Second, the empirical estimation of niche space has evolved greatly over the past few decades with the advent of increased computational power and development of nonparametric multivariate methods such as generalized additive (Hastie & Tibshirani, 1986) and Maxent (Phillips et al., 2006) modelling. Because of their accuracy in defining environmental niche space these approaches generate better estimates of potential global range (Cowell & Rangel, 2009), including the location and size of not only occupied but also unoccupied appropriate patches and the barriers separating them. Third, the development of DNA sequencing technology and molecular phylogenetics now allows for a priori determination of biologically valid groups and their correct identification features. This reduces environmental data errors related to misidentified, overlumped and/or oversplit occurrence data (Nekola & Horsák, 2022).

In spite of these advancements, determination of potential range saturation – and the factors influencing it – has been limited to regional or smaller scales: for example, European trees (Svenning & Skov, 2004, 2007) and carabid beetles (Calatayud et al., 2019) or Mexican mammals (Munguía et al., 2008). While such works have been crucial in documenting disequilibrium between species range and climate and a reduction of species pool size (Svenning et al., 2015), until now no one has considered the nature and impact of dispersal barriers at the global scales where such features have been classically invoked (Preston, 1960).

The principal goals for this paper are thus to estimate actual, realized and potential Holarctic-wide distributions for a suite of effectively dispersing land-snail taxa that have survived integrative empirical revision. From this we identify, measure and characterize the barriers impacting each. Our analyses document the statistical properties of species ranges and dispersal barriers, the mechanisms underlying barriers, and their impact on the size and composition of regional taxa pools. They also allow for consideration of how barriers impact: (a) realized range sizes, (b) global occupancy rates of appropriate habitats, (c) regional species pool richness, and (d) the strength of distance decay in similarity between species pools.

## 2 | METHODS

### 2.1 | Target group

We limit our analyses to small (shell diameter < 5 mm) Holarctic land snail species from three genera (*Euconulus*, *Pupilla* and *Vertigo*). All modelled taxa have survived empirical integrative taxonomic revision using multiple DNA and morphological signals (Horsáková et al., 2020; Nekola et al., 2015, 2018). We limited analysis to entities whose range centre is at least 40°N in eastern Eurasia and central/eastern North America and 50°N in central/western Eurasia with this difference being due to a shifting of boreal species ranges north due to warmer winter temperatures at similar latitudes in western Eurasia due to Gulf Stream maritime influence. One additional species (*Pupilla hebes*) was included even though its range centre lies farther south because it is limited to montane taiga. After some additional adjustments related to uncertain taxonomic status (see Appendix A), our sample system includes 5 *Euconulus*, 13 *Pupilla* and 34 *Vertigo* taxa (Appendix B).

These taxa represent an ideal system to analyse dispersal barriers because: (a) they have all undergone integrative taxonomic revision, occur throughout the entire region, and represent about 10% of the entire small-sized Holarctic fauna. They are also among the most frequent land snails within the Holarctic fauna, and occur throughout the region across all habitat types and trophic states (Barker, 2001; Némec et al., 2021; Pilsbry, 1948; Sysoev & Schileyko, 2009; Welter-Schultes, 2012); (b) post hoc analysis demonstrates that shell features allow for correct identification of all but one empirically validated taxon. It is thus possible to accurately determine occurrences from shells alone; (c) composition and abundance data are available

for over 5000 modern sites across much of the Holarctic (Horsák & Meng, 2018; Horsáková et al., 2018; Nekola, 2014); (d) while being poor active movers (e.g. Örstan et al., 2011), small land snails are among the most effective known passive dispersers, with multiple ocean crossings of 7500 km having been documented (Gittenberger et al., 2006). Rates of local endemism for small-sized species within the Holarctic fauna are thus quite low, with at least some taxa being thought to possess global ranges (Pilsbry, 1948). Focus on these taxa should thus identify the lower bound for the impact of barriers on ecological and biogeographic pattern.

## 2.2 | Input data compilation

Using genetically validated diagnostic shell traits, all target taxa occurrences from our 5000+ site-scale composition lists – as well material from the 15,000+ lot Brian Coles collection at the National Museum of Wales – were revised. We also reidentified all Holarctic *Pupilla* and *Vertigo* lots from the University of Michigan Museum of Zoology, Royal Ontario Museum and National Museum of Canada (Nekola & Coles, 2010). For species whose diagnostic features are easily observed and did not change, we also incorporated occurrence data from the UK/Ireland (Kerney, 1999) and Sweden (Waldén, 2007). Over 9200 individual occurrence records were accumulated (Figure 1). Each was reported in decimal degrees at 5 arc-min cell resolution, and all are provided in Supplementary File 'occurrences\_all.csv'. While these data obviously do not represent all known occurrences, accurate climatic envelope estimation is possible because the full known geographic, ecological and climatic ranges of each taxon have been included. Given that museum land snail records possess a non-trivial frequency of labelling errors and a 20% misidentification rate (Nekola et al., 2019) – which increases to more than 80% in online databases of some other invertebrate groups (Meier & Dikow, 2004) – we avoided their use as we had no way to independently verify identifications or generate simple a priori decision rules to exclude inaccurate reports. We thus opted to use a smaller but higher-quality dataset to ensure the most robust model outputs.

To avoid potential bias in climate suitability models caused by the uneven density of occurrence records, we resampled species

occurrences using an environmental filtering procedure (Castellanos et al., 2019; Varela et al., 2014). We assigned climatic variables obtained from the WorldClim v.1.4 (Hijmans et al., 2005) and ENVIREM (Title & Bemmels, 2018) databases at 5 arc-minute resolution to each record. For taxa with more than 80 occurrences, we placed each record into a multidimensional space generated through principal components analysis (PCA) that captured at least 90% of observed variation. We then successively removed occurrence records associated with the smallest average pairwise distance in reduced PCA space until all distances were 0.1 SDs or greater, or the number of remaining records reached 80, whichever came first. The number of records used to calibrate each climate suitability model is provided in Appendix B, and these are mapped in Appendix E.

## 2.3 | Climate suitability modelling

### 2.3.1 | Model parameterization

We used Maxent (version 3.4.1; Phillips & Dudík, 2008) to model suitable climates because it uses presence-only data and generally performs well at small sample sizes (Elith et al., 2006; Wisz et al., 2008). We excluded seven taxa primarily due to their very limited number of records (see Appendix A for details). Models were calibrated using resampled occurrence records (as described above) with 10,000 background points and taxon-specific subsets of best-performing climatic variables to avoid issues connected with model overfitting (for details on variable selection, see Appendix C). All suitability models and their resultant predictions were evaluated using fivefold cross-validation. Three metrics were utilized: (a) area under a receiver operating characteristic curve (AUC; the higher the score the better the model performance; Phillips et al., 2009); (b) difference between forecast and reference AUC (AucDiff or overfitting; the higher the score the poorer the model performance; Warren & Seifert, 2011); and (c) Boyce index (1 = models consistent with occurrence distribution, -1 = models which avoid occurrences; Boyce et al., 2002; Hirzel et al., 2006). Modelling was conducted in the dismo R package (Hijmans et al., 2017). All details on model settings are provided in the Overview, Data, Model, Assessment, Prediction (ODMAP) protocol found in Appendix C.

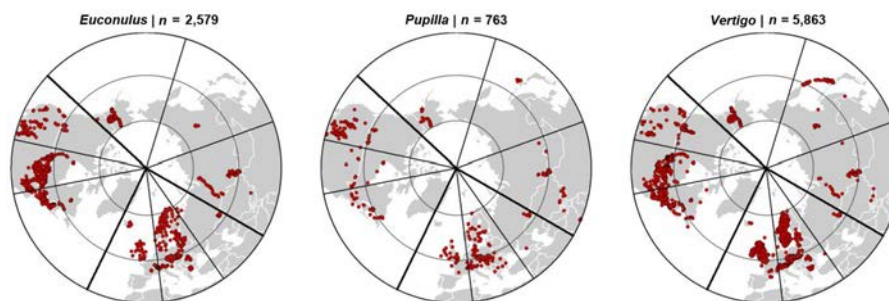


FIGURE 1 Location of the 9,205 occurrence records use to parameterize climate suitability models

### 2.3.2 | Potential range determination

The climatic suitability model for each taxon was projected across the Holarctic using WorldClim 1.4 and ENVIREM (Title & Bemmels, 2018) data at 5 arc-min resolution. Geographic coverage of suitable climates (hereafter referred to as 'potential range') was estimated by applying a Maxent routine that balances training omission, predicted area and threshold value (Phillips & Dudík, 2008). This procedure was used as it generates the most liberal estimates of potential range and thus the smallest possible barrier widths. For this reason, actual barrier widths are likely larger than our estimates.

## 2.4 | Range and barrier documentation

### 2.4.1 | Holarctic biogeographic regions

We divided the Holarctic into nine roughly equally dispersed biogeographic regions (Figure 1): western (Region 1: 26°W–7°E), central (Region 2: 7°E–35°E) and eastern (Region 3: 35°E–60°E) Europe; western (Region 4: 60°E–110°E), central (Region 5: 110°E–163°E) and eastern (Region 6: 163°E–133°W) Beringia; and western (Region 7: 133°W–102°W), central (Region 8: 102°W–79°W) and eastern (Region 9: 79°W–26°W) North America. Since genetically confirmed species occurrence data do not yet exist for Greenland, for the purposes of this work we defined Region 1 to begin just west of Iceland, with Region 9 ending along the western shore of the Labrador Sea.

### 2.4.2 | Actual range measurement

Taxon occupancy within each Holarctic region (see above) was recorded based on our data in addition to verifiable literature reports (e.g. Nekola & Coles, 2010; Pilsbry, 1948; Welter-Schultes, 2012). The known major and orthogonal minor range extents for each were measured using our data in combination with published literature. Based on its range centroid, each taxon was also assigned to one of three main biogeographic groups: European (26°W–60°E), Beringian (60°E–133°W), and North American (133°W–26°W). Range area was estimated by calculating the total area (km<sup>2</sup>) of appropriate climate found within all occupied regions. These represent the maximum possible because taxa almost never fully occupy their potential range (Hurlbert & White, 2005; Svenning & Skov, 2004). Recorded data for each species are presented in Appendix B.

### 2.4.3 | Potential range measurement

We measured major and orthogonal minor axes for the suitable climate patches associated with the areas where each species resides (Appendix B). These reflect the maximum potential range because taxa occurrences at range margins often represent microclimatically unique habitats – such as cooled air emanating from ice

caves (Nekola, 1999) or warm air associated with geothermal fields (Brunton, 1986; Carcaillet et al., 2018). The global climate suitability map for each species was also used to record potential occurrence within each biogeographic region. The simple presence of an appropriate 5 arc-min climate pixel within a region does not indicate that a given taxon should be expected there, however, with population colonization and persistence being unlikely when total appropriate area and largest patch size are too small (e.g. Gilpin & Soulé, 1986). Using actual modern Holarctic species occurrences as a guide, we set these lower thresholds at 90,000 and 45,000 km<sup>2</sup>, respectively. Thus, any taxa with appropriate climate area falling under both limits were coded as being absent from a given region even if some appropriate climate was actually present. The matrix of actual and potential regional occurrences is presented in Appendix D. Lastly, maximum potential global range area was determined by summing the total area of appropriate climate within regions that were recorded as potentially supporting populations.

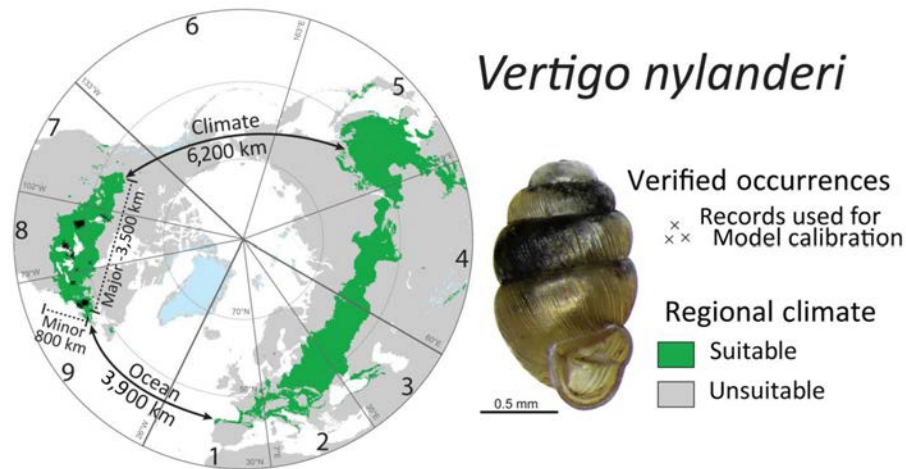
### 2.4.4 | Dispersal barrier measurement

For each modelled taxon, minimum barrier width separating appropriate climate areas was measured both to the east and west of the actual range. Distances were calculated from great-circle routes that did not exceed 70°N – for example, transpolar routes were disallowed. Patches within regions falling below the 90,000/45,000 km<sup>2</sup> thresholds were ignored. The underlying cause of the barrier (inappropriate climate, ocean or ice) was noted. If more than one appeared responsible, each was reported. An example is provided in Figure 2. When actual range was found to terminate significantly (e.g. 1500+ km) before the end of a contiguous predicted climate envelope, barrier width was recorded as missing with habitat/history being recorded as the barrier type. The resultant data matrix is provided in Appendix B.

## 2.5 | Statistical analyses

### 2.5.1 | Range size and occupancy

Cumulative Rank Frequency Distributions (CRFDs; Newman, 2005) for actual and potential range area and major axis extent were plotted. Because these (as well as barrier distributions) demonstrated POver-law/LOg-normal (POLO; Halloy 1998; Newman 2005) shapes (see below), and because normalizing them frequently leads to significant mathematical artifacts (Nekola et al., 2008), we relied on the Kruskal–Wallis test as it is the nonparametric equivalent of ANOVA, works well when groups contain at least five members (our data range from 9 to 66) and when distributions are not strongly skewed (none of ours are). In all cases where multiple comparisons were made on the same data, we adjusted the  $\alpha = .05$  significance threshold using a Bonferroni correction. The proportion of the potential global Holarctic range occupied by each taxon was also calculated. Significance of observed differences between European,



**FIGURE 2** Range and barrier analysis for *Vertigo nylanderi*. The appropriate climate envelope (marked in green) is based on WorldClim data gathered from 119 verified populations across the entire known range. Regions 8 and 9 are known to be occupied, while an additional six (1, 2, 3, 4, 5, 7) were deemed to be appropriate by possessing at least 90,000 km<sup>2</sup> of suitable climate with the largest patch size being at least 45,000 km<sup>2</sup>. We chose to consider Region 7 as being occupied because it supports a proper amount of climate (see Section 2.4.3) extending in a contiguous fashion from the current western range limit in western Manitoba to the base of the Rocky Mountains in Alberta. The reason that the species has yet to be reported from Region 7 is almost certainly due to lack of field surveys. Note that the small isolated appropriate climate zone in Newfoundland was considered unoccupied because no populations have ever been recorded from this area, which is separated from the nearest appropriate climate by 500 km of ocean and inappropriate climate. The actual major range axis is estimated to be 3,500 km with the minor axis to be 800 km. The distance of the potential major axis as indicated via climate niche modelling was 4,100 km with a potential minor axis of 1,300 km. The amount of appropriate climate within the occupied range is 3.4 million km<sup>2</sup>, the total amount of global modern appropriate climate within zones capable of supporting the species is 13.4 million km<sup>2</sup>, representing a global occupancy rate of 25.7%. The occupied range is separated by 3,900 km of ocean from the nearest suitable regional climate in the Cantabrian Mountains of western Spain, while the western range limit is separated by 6,200 km of inappropriate climate from the nearest suitable regional climate in the Amur Oblast of eastern Siberia

Beringian, and North American species was estimated using the Kruskal–Wallis test. All comparisons were graphically represented using box-plots. These and all subsequent statistical analyses were conducted in R version 3.4.1 (<https://www.r-project.org/>).

## 2.5.2 | Barrier size and type

CRFDs were generated for measured barrier widths. The Kruskal–Wallis test was used to identify the significance of differences between major versus minor and actual versus potential range extents and between major/minor range extents versus barrier widths. The number of instances in which climate, ocean, ice, and habitat/historical factors underlie barriers was calculated separately for European, Beringian, and North American taxa, with significance of deviation in cell counts from uniform in 2 × 3, 3 × 3, or 3 × 4 contingency tables being estimated using Fisher's exact test because it provides the most accurate results when sparse cells (e.g.  $n < 5$ ) are relatively frequent.

## 2.5.3 | Observed versus potential species pool turnover

The significance of differences between observed and potential regional richness was calculated using the Kruskal–Wallis test.

Observed versus potential distance decay was calculated for all pairwise comparisons using the Jaccard index (Nekola & White, 1999), with distance between regional centroids being estimated using great-circle routes that did not exceed 70°N or cross the Atlantic Ocean. Because distance decay may be modelled through either exponential or power law forms (Nekola & McGill, 2014), each was fit to the data with two free parameters using nonlinear regression in the NLS package of R with small sample-size corrected Akaike information criterion (AICc) being recorded for each along with optimum parameter values. Model  $p$  and assemblage 50% turnover distance (Nekola & White, 1999) were also noted. Relationships were plotted on untransformed distance and similarity axes.

## 3 | RESULTS

### 3.1 | Model evaluation

All climate suitability models accurately portrayed their respective occurrences (Table 1, Appendix E): AUC ranged from .716 to .996 (median = .953); overfitting from .0002 to .056 (median = .008); Boyce index from .285 to .937 (median = .792). In eight taxa (18% of total; *Euconulus alderi*, *Euconulus fresti*, *Vertigo alpestris*, *Vertigo cristata* agg., *Vertigo lilleborgi* agg., *Vertigo oughtoni*, *Vertigo pygmaea* and *Vertigo substriata*), populations were noted outside of the predicted

TABLE 1 Climate niche model evaluation. Density of blue coloration in the Boyce Index column cells indicates model efficacy, ranging from white (poor) to blue (excellent).

		Fivefold cross-validation					
		Evaluation based on 20% of data excluded from calibration dataset					
Taxon		Mean test AUC	±SD	Mean overfitting (AucDiff)	±SD	Mean Boyce index	±SD
<i>Euconulus</i>	<i>alderi</i>	.887	.021	.015	.026	.862	.153
<i>Euconulus</i>	<i>fresti</i>	.927	.023	.005	.028	.844	.154
<i>Euconulus</i>	<i>fulvus</i>	.716	.010	.027	.012	.913	.039
<i>Euconulus</i>	<i>polygyratus</i>	.956	.018	.008	.021	.813	.045
<i>Pupilla</i>	<i>alaskensis</i>	.986	.006	.002	.007	.515	.310
<i>Pupilla</i>	<i>alpicola</i>	.961	.033	.011	.039	.737	.081
<i>Pupilla</i>	<i>blandi</i>	.931	.079	.038	.089	.543	.179
<i>Pupilla</i>	<i>hebes</i>	.982	.006	.004	.007	.705	.130
<i>Pupilla</i>	<i>hokkaidoensis</i>	.993	.007	.000	.009	.502	.384
<i>Pupilla</i>	<i>hudsonianum</i>	.953	.048	.013	.051	.742	.083
<i>Pupilla</i>	<i>loessica</i>	.996	.002	.001	.002	.830	.064
<i>Pupilla</i>	<i>muscorum</i>	.890	.016	.014	.019	.792	.072
<i>Pupilla</i>	<i>sterrii</i>	.933	.026	.024	.028	.687	.281
<i>Pupilla</i>	<i>triplicata</i>	.912	.080	.038	.090	.758	.085
<i>Pupilla</i>	<i>turcmenia</i>	.983	.006	.008	.006	.718	.080
<i>Vertigo</i>	<i>alpestris</i>	.953	.008	.007	.009	.923	.035
<i>Vertigo</i>	<i>arctica</i>	.959	.010	.003	.012	.910	.042
<i>Vertigo</i>	<i>arthuri</i>	.887	.015	.020	.017	.832	.056
<i>Vertigo</i>	<i>beringiana</i>	.987	.003	.002	.004	.863	.065
<i>Vertigo</i>	<i>bollesiana</i>	.956	.018	.006	.021	.809	.117
<i>Vertigo</i>	<i>circumlabiata</i>	.978	.012	.008	.015	.778	.094
<i>Vertigo</i>	<i>columbiana</i>	.994	.004	.002	.005	.621	.176
<i>Vertigo</i>	<i>cristata</i> agg.	.855	.044	.028	.052	.935	.040
<i>Vertigo</i>	<i>extima</i>	.986	.004	.003	.005	.780	.060
<i>Vertigo</i>	<i>genesii</i>	.980	.009	.005	.010	.825	.071
<i>Vertigo</i>	<i>genesioides</i>	.932	.040	.028	.040	.523	.197
<i>Vertigo</i>	<i>geyeri</i>	.946	.008	.006	.010	.919	.031
<i>Vertigo</i>	<i>hannai</i>	.976	.015	.002	.017	.737	.155
<i>Vertigo</i>	<i>kurilensis</i>	.994	.005	.001	.006	.875	.136
<i>Vertigo</i>	<i>kushiroensis</i>	.984	.007	.008	.007	.724	.080
<i>Vertigo</i>	<i>lilljeborgi</i>	.930	.012	.007	.014	.916	.058
<i>Vertigo</i>	<i>microsphaera</i>	.975	.013	.012	.013	.697	.025
<i>Vertigo</i>	<i>modesta</i>	.874	.014	.019	.019	.852	.095
<i>Vertigo</i>	<i>morsei</i>	.940	.035	.016	.039	.791	.061
<i>Vertigo</i>	<i>nylanderi</i>	.952	.024	.008	.028	.694	.151
<i>Vertigo</i>	<i>oughtoni</i>	.986	.019	.002	.023	.445	.389
<i>Vertigo</i>	<i>parcedentata</i>	.977	.037	.010	.042	.808	.169
<i>Vertigo</i>	<i>perryi</i>	.938	.087	.056	.088	.285	.257
<i>Vertigo</i>	<i>pseudosubstriata</i>	.911	.132	.055	.140	.334	.484
<i>Vertigo</i>	<i>pygmaea</i>	.853	.012	.022	.014	.928	.033

TABLE 1 (Continued)

		Fivefold cross-validation					
		Evaluation based on 20% of data excluded from calibration dataset					
Taxon		Mean test AUC	$\pm$ SD	Mean overfitting (AucDiff)	$\pm$ SD	Mean Boyce index	$\pm$ SD
<i>Vertigo</i>	<i>ronnebyensis</i>	.925	.017	.008	.021	.881	.094
<i>Vertigo</i>	<i>substriata</i>	.920	.006	.011	.006	.937	.024
<i>Vertigo</i>	<i>ultima</i>	.942	.095	.042	.099	.569	.233
<i>Vertigo</i>	<i>ventricosa</i>	.887	.017	.025	.023	.899	.033
Mean		.941	.025	.014	.028	.751	.127

Abbreviation: AUC, area under a receiver operating characteristic curve; AucDiff, difference between forecast and reference AUC.

appropriate climatic range. These universally represented thermally distinct microclimates (e.g. fens, cold ocean shores, sky island forests, talus slopes with cold air drainage, irrigated yards/fields). We also note the *Pupilla alpicola* climate model accurately predicted its previously unreported presence on the north-west coast of Iceland (Horsák et al., 2022).

### 3.2 | Range and occupancy

CRFDs for actual and potential major range axes and areas all exhibit concave shapes characteristic of ecological systems (Figure 3), and likely represent underlying POLO distributions. Observed major axis extents (Appendix B) varied from c. 100 km (*Vertigo lilljeborgi* East Asian race) to 19,800 km (*Euconulus fulvus*) with a median value of 3500 km. Minor range axes ranged from c. 100 km (*Pupilla hokkaidoensis* and *V. lilljeborgi* East Asian race) to 4800 km (*Vertigo cristata* agg., *Vertigo modesta*) with a median value of 1000 km (Figure 4). While these minimum values are certainly impacted by underreporting (the actual major and minor range extents are likely at least 1000 km), regional endemics within these genera do rarely exist (Horsák et al., 2016). Median potential major axis extent was 7000 km and that for minor axes was 2800 km, ranging from 2200 to 19,800 and 200 to 5600 km, respectively. Following Bonferroni correction for eight comparisons ( $\alpha' = .05/8 = .00625$ ), significant differences ( $p < .001$ ) were noted between actual and potential major versus minor range extents.

The total area of appropriate climate within biogeographic regions supporting a given taxon (e.g. maximum actual range size) varied from  $3.3 \times 10^5$  (*Vertigo kurilensis*) to  $4.2 \times 10^7$  km<sup>2</sup> (*E. fulvus*) with a median value of  $5.6 \times 10^6$  km<sup>2</sup> (Figure 5). The total area of appropriate climate across the entire Holarctic from 30°N within regions supporting sufficient quantities to maintain a given taxon (e.g. maximum potential range size) varied from  $1.4 \times 10^6$  (*V. kurilensis*) to  $4.2 \times 10^7$  km<sup>2</sup> (*E. fulvus*) with a median value of  $1.4 \times 10^7$  km<sup>2</sup>. Kruskal–Wallis tests show that actual range sizes are

significantly smaller than their potential ( $p < .001$ ). Additionally, no significant differences were noted between European, Beringian, and North American taxa for actual ( $p = .861$ ) and potential ( $p = .106$ ) range sizes. Global occupancy rates ranged from 8.6% (*V. lilljeborgi* East Asian race) to 100% (*E. fulvus*), being highest for European (median = 61.7%) and lowest for North American (33.9%) taxa (Figure 6). While 71% of European taxa were estimated to occupy more than 50% of their potential global range, this was the case for only 50% of Beringian and 21% of North American taxa. The Kruskal–Wallis test found this difference to be significant ( $p = .028$ ).

### 3.3 | Barriers

Only one climate suitability map (*E. fulvus*) suggested existence of a single trans-Holarctic range. All remaining taxa possessed at least two (and often more) discrete areas of appropriate climate (Appendix E), with minimum barrier distances ranging from 200 km (*Pupilla alaskensis*, *Vertigo hannai*, *Vertigo modesta*, *Vertigo oughtoni*, *Vertigo ultima*) to 9200 km (*Pupilla hebes*) (Appendix B). The five smallest all represented Alaska taxa separated from appropriate climates in the Chukchi Peninsula of easternmost Siberia by the Bering Strait. It is also possible that their apparent absence in Siberia is due to inadequate field sampling. CRFD plots demonstrated that barrier widths possessed a concave shape (Figure 3), implying again an underlying POLO distribution. Median barrier width was 2450 km (Figure 4). Of the 66 measured barrier distances, 39 (~60%) were shorter than the ~3000 km ocean barrier separating eastern Newfoundland from western Ireland and the Hawaiian archipelago from the nearest continental landmass. Using Bonferroni corrected significance thresholds, actual barrier widths were found to be statistically similar to both actual major axis ( $p = .012$ ) and potential minor axis ( $p = .829$ ) distances. However, barrier widths were significantly larger ( $p < .001$ ) than actual minor axis and significantly smaller ( $p < .001$ ) than potential major axis distances.

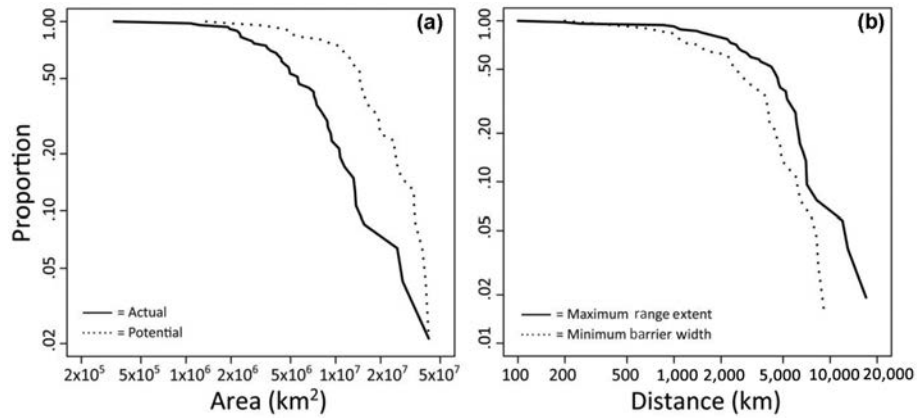


FIGURE 3 Cumulative rank frequency distribution of: (a) actual versus potential range area; (b) maximum range extent versus minimum barrier widths

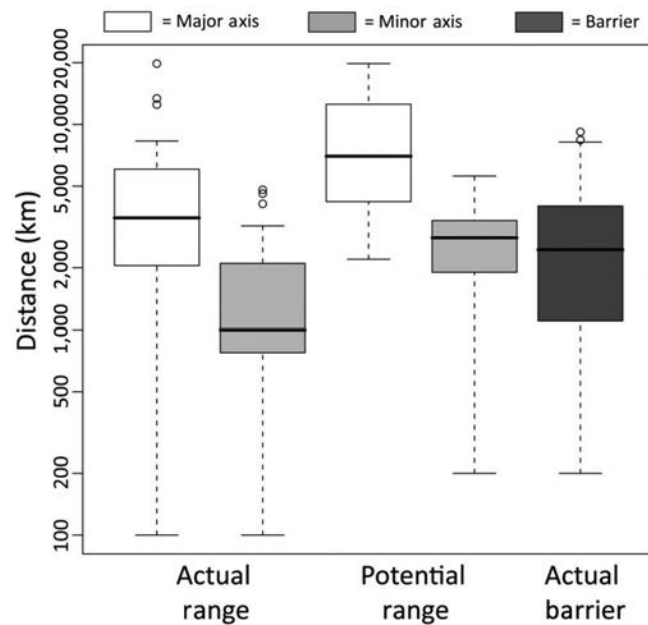


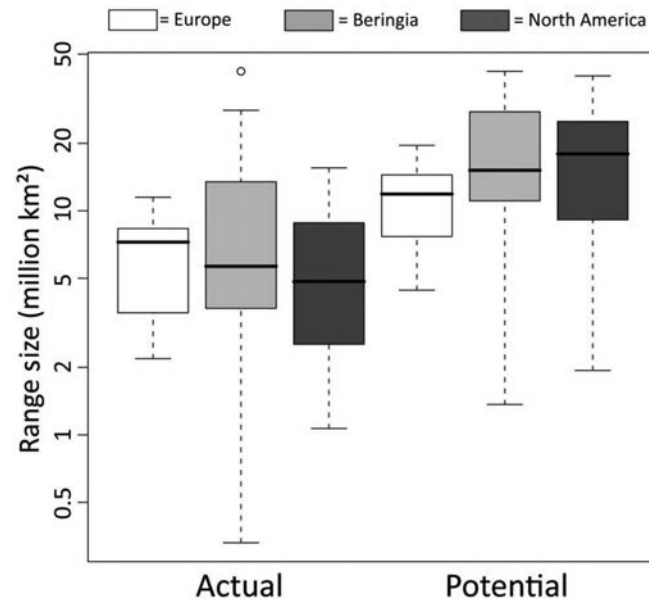
FIGURE 4 Major and minor range extent versus barrier width distances for actual and potential ranges. Kruskal–Wallis test statistics (all  $df = 1$ ) and  $p$ -values: actual major versus minor axis – 30.6518,  $p \ll .001$ ; potential major versus minor axis – 44.3291,  $p \ll .001$ ; actual major versus potential major axis 19.4006,  $p < .001$ ; actual minor versus potential minor axis – 21.4957,  $p \ll .001$ ; actual major axis versus barrier width – 6.2088,  $p = .0123$ ; actual minor axis versus barrier width – 13.8069,  $p < .001$ ; potential major axis versus barrier width 42.8913,  $p \ll .001$ ; potential minor axis versus barrier width 0.0465,  $p = .8292$ . Both actual range categories represent 52 observations; both potential range categories represent 47, while 66 barrier distances were considered. Bonferroni-adjusted .05 significance threshold for eight comparisons  $\alpha' = .00625$

Twenty-six range limits in 20 species were found to terminate at least 1500 km before cessation of appropriate climate (Table 2, Appendix A). While some of these cases are clearly related to disappearance of required habitat (e.g. for *Vertigo genesioides* the lack of peatlands between the Altai and Urals in Kazakhstan and south-western Siberia), others have no clear explanation and may be related to the vagaries of dispersal history. The frequency of habitat/historical barriers demonstrated a strong phylogenetic signal, being present in almost 60% of *Pupilla* range limits (10/11 taxa), but in only 20% of *Vertigo* (11/31 taxa) and 12.5% of *Euconulus*

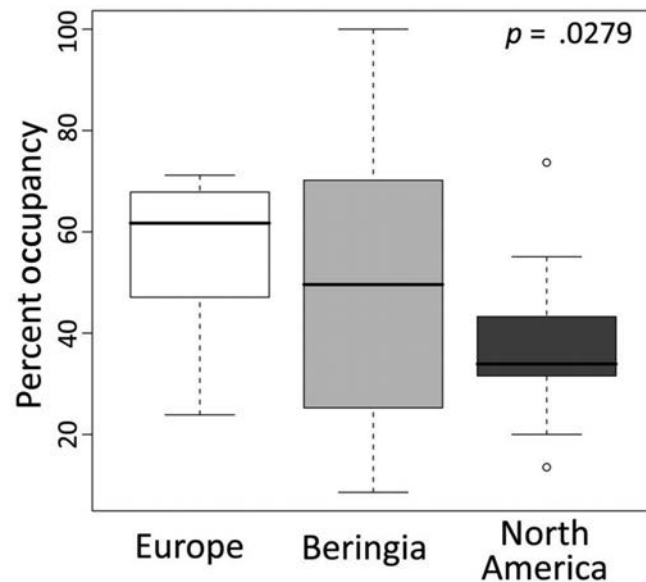
(1/5 taxa). Fisher’s exact test demonstrated this difference to be significant ( $p = .001$ ).

Of the 114 recorded barriers, oceans represented the most common underlying causative factor (37%) followed by inappropriate climate (29%), habitat/history (27%), and the Greenland ice sheet (7%). While frequencies significantly varied ( $p = .007$ ) across European, Beringian, and North American taxa, this result is completely driven by the three times higher level of habitat/history range limits in Beringia: there was no statistically significant difference among the other three barrier types across these three regions ( $p = .775$ ).





**FIGURE 5** Actual and potential range sizes by biogeographic affinity. Kruskal–Wallis test statistics and  $p$ -values: all actual versus all potential – 24.235,  $df = 1$ ,  $p < .001$ ; Europe versus Beringia versus North America differences in actual range – 0.8613,  $df = 2$ ,  $p = .6501$ , potential range – 4.4714,  $df = 2$ ,  $p = .1069$ . For both actual and potential categories, boxes represent data summarized across 14 European, 14 Beringian, and 19 North American species. Bonferroni-adjusted .05 significance threshold for three comparisons  $\alpha' = .0167$



**FIGURE 6** Percent global occupancy of potential modern range by biogeographic affinity. Kruskal–Wallis test statistic = 7.1598 with  $df = 2$ . Boxes represent data summarized across 14 European, 14 Beringian, and 19 North American species

### 3.4 | Taxa-pool richness and turnover

Observed actual regional taxa richness ranged from 20 (western Beringia) to 11 (central Beringia), with a median of 16 (Figure 7). Potential regional taxa richness ranged from 45 (eastern Europe, central Beringia, western North America) to 31 (western Europe), with a median of 43. This difference was highly significant (Kruskal–Wallis  $p < .001$ ).

Nonlinear least-squares fits of observed regional taxa pool similarity versus distance (Figure 8) demonstrated that AICc values were

virtually identical between the two-parameter exponential and power law forms (−49.5375 versus −49.1935). Because the exponential form is mathematically expected in situations where species pools vary across sample extent (Nekola & McGill, 2014), it was chosen for subsequent analysis. The exponential decay of regional composition was highly significant ( $p < .001$ ), explaining over 79% of observed variation. With a best-fit decay coefficient of −0.000263, a distance of ~2600 km is required for average regional taxa-pool similarity to fall by 50%. However, no significant compositional decay with distance was noted in potential regional species pools ( $p = .975$ , pseudo- $r^2 < .001$ ).

TABLE 2 Barrier type by biogeographic location. Bonferroni-adjusted .05 significance threshold for four comparisons  $\alpha' = .0125$

(a) By region				
	Climate	Ocean	Ice	Habitat/history
Region	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)	<i>n</i> (%)
Europe	11 (34.4)	12 (37.5)	4 (12.5)	5 (15.6)
Beringia	5 (17.9)	7 (25.0)	0 (0)	16 (57.1)
North America	17 (31.5)	23 (42.6)	4 (7.4)	10 (18.5)
Total	33 (28.9)	42 (36.8)	8 (7.0)	31 (27.2)

Fisher's exact test:

Climate/ocean/ice across Holarctic:  $p = .6071$

Climate/ocean/ice between regions:  $p = .7745$

All categories between regions:  $p = .0072$

(b) By genus			
Genus	Range limits corresponding to		
	Climate/Ocean/Ice	Habitat/History	(% Total)
<i>Euconulus</i>	7	1	12.5
<i>Pupilla</i>	10	14	58.3
<i>Vertigo</i>	66	16	19.5

Fisher's exact test:  $p = .0011$

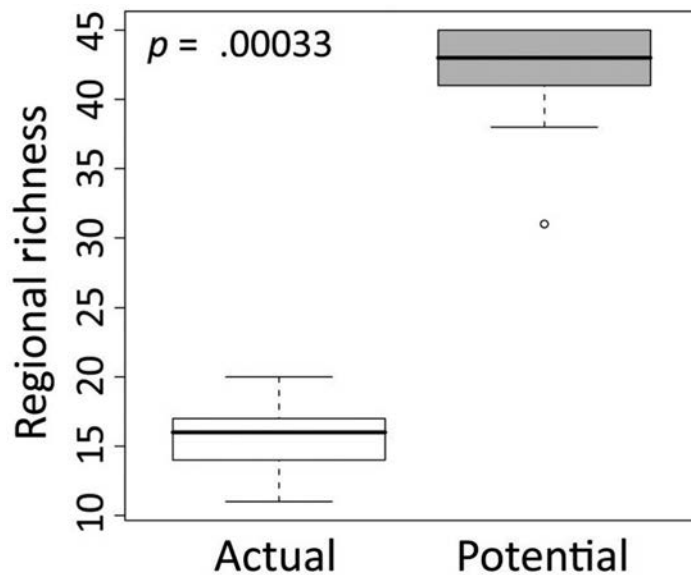


FIGURE 7 Actual versus potential regional taxa richness across all nine Holarctic biogeographic regions. Kruskal-Wallis test-statistic = 12.8959 at  $df = 1$

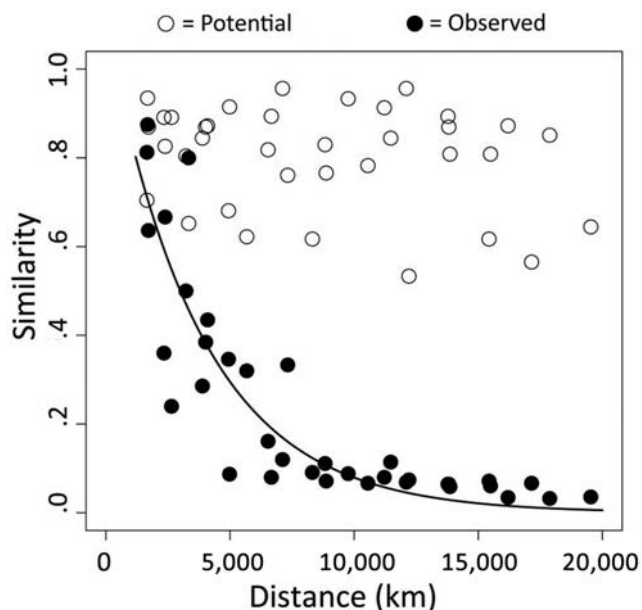
## 4 | DISCUSSION

This work illustrates that use of global climate databases in conjunction with modern niche modelling and empirically vetted taxonomic concepts provides important insights into the factors generating large-scale dispersal barriers and their impact on global range limits and regional taxa pools. It also helps document the importance of intra-continental barriers in isolating faunas, and suggests that

concern regarding exotic species should be expanded to include anthropogenic movements within continents.

### 4.1 | Protocol efficacy

Our protocol demonstrated excellent ability to accurately predict taxa occurrences with 33 taxa (75%) having Boyce index scores



**FIGURE 8** Observed and potential climate-driven taxa pool distance decay across all pairwise comparisons. Significance of nonlinear exponential fit for observed regional taxa pools  $p < .001$  (decay coefficient =  $-0.000263$ , pseudo- $r^2 = .79$ ); potential regional taxa pools based on climate envelopes  $p = .975$

exceeding .7. Only a single entity (*Vertigo perryi*) had a Boyce index value  $< .3$ , and even then its AUC score (.938) was high. The only occurrences outside of predicted appropriate climates were all associated with unique isolated microclimates. As a result, we are confident that our models accurately portray potentially appropriate climates and that they reliably estimate not only actual and potential maximum range, but also the position, width and nature of dispersal barriers.

As has been previously shown in invasive plant niche modeling (Ensing et al., 2013), perhaps the most important step in generating accurate projections was to limit model parameterization to high-quality occurrence data. Because more than half of the traditional taxonomic concepts considered here were initially in error, up to 90% of site composition lists would have also been initially incorrect (Nekola & Horsák, 2022). The use of such data would have created up to threefold larger or smaller estimates of species-specific climatic tolerances than the actual ones and lead to badly under- or over-estimated range sizes and dispersal barriers for the majority of species (Appendix A). We would thus treat as potentially suspect climate envelopes based solely on uncritical use of unvetted taxonomic concepts and unverified occurrence records. It is for this reason we only analysed data that we had empirically confirmed and avoided use of unverified records even though these are numerous in the literature and online databases. As long as they capture the full geographic and climatic range of a given taxon, such smaller sized, highly vetted input data likely represent the best way to accurately parameterizing climate niches.

## 4.2 | Statistical properties of range extents/sizes and barrier widths

The median observed Holarctic taxon range was  $3500 \times 1000$  km. Median predicted climate envelope range was  $7000 \times 2800$  km. Measured barriers were of the same general scale, with a median distance of 2450 km. Given the excellent passive dispersal abilities of these species, it is not surprising that some have been able to span these barriers and develop disjunct distributions: European and Altai *Pupilla alpicola* populations are separated by a 4000-km and southern Urals and European *Pupilla sterrii/Vertigo alpestris* by a 2000-km climate barrier. *Vertigo arctica* and *Vertigo genesii* both possess populations on either side of a ~600-km climate/water barrier between southern Scandinavia and the Alps. Icelandic *E. fulvus*, *P. alpicola*, *V. alpestris* and *V. arctica* populations are separated from those in Scandinavia/Great Britain by 1000 km of the north Atlantic. And, since European contact both *Pupilla muscorum* and *Vertigo pygmaea* have traversed the 3000+ km North Atlantic Ocean barrier between Europe and North America (Nekola et al., 2015, 2018). The origins of these disjunctions are almost certainly related to different mechanisms: consistent DNA sequence differences in *P. alpicola* populations between Europe and central Asia (Nekola et al., 2015) suggests old isolation age, perhaps related to splitting of once continuous Pleistocene distributions through Holocene climate change. The long branch for isolated Icelandic *V. alpestris* populations also suggests early- to mid-Holocene establishment. However, the lack of genetic deviance between *P. muscorum*, *V. arctica*, *V. genesii* and *V. pygmaea* populations (Nekola et al., 2015, 2018) suggest more recent barrier crossings.

One last important point is that both ranges and barriers demonstrate POLO statistical distributions. This could imply that they are driven by random multiplicative processes (McGill, 2003), or – within a complex systems perspective – that they are generated by a positive interaction correlation with agent size and an inverse correlation with increasing distance (Halloy, 1998). While data transformation could normalize such distributions, this process can also generate mathematical artifacts that complicate or inhibit the drawing of accurate conclusions (Nekola et al., 2008). Because of this, standard parametric statistics will be inappropriate for empirical confrontation of competing hypotheses (Hilborn & Mangel, 1997).

## 4.3 | Barrier configuration and causes

The most common factors associated with range limits/barriers were oceans (37%), followed by inappropriate climate (29%), habitat/history (27%) and ice sheets (7%). The importance of oceans and habitat/history in generating range limits is not surprising, with these having been given prominence since the start of biogeographic inquiry. However, intra-continental climate represents

20–35% of barriers across all three biogeographic affinities. These are especially important in separating European taxa from appropriate Beringian climate zones in the Himalayas and east of Lake Baikal, North American taxa from boreal eastern Eurasia, and eastern and western North American taxa from each other. Assuming that dispersal limitation is unimportant within continents (e.g. Krebs, 1985) must thus be seriously questioned.

The frequent existence of intra-continental barriers has significant conservation implications. For instance, the rapid human-driven range expansion of some eastern North American land snail taxa (e.g. *Triodopsis hopetonensis*, *Ventridens demissus*) has led to negative impacts on local species pools in newly colonized areas (Brian Coles, personal communication). And the ~120-km climate barrier associated with Puget Sound effectively isolated alpine habitats of the Cascade and Olympic ranges in Washington state. Movement of mountain goats (*Oreamnos americanus*) across this barrier by wildlife biologists in the 1920s led to subsequent serious degradation of alpine plant biodiversity within the Olympic range (Scheffer, 1993).

It is also interesting to note that even though in modern times continental ice serves as a potential ~1000-km barrier only in Greenland, roughly 10% of European and North American taxa may have range limits related to this feature. It thus seems likely that continental ice could have played a much more important role during cold phases of the Pleistocene when their coverage was more than an order of magnitude greater. Modern range disequilibrium related to position of the Last Glacial Maximum continental ice sheet has been previously suggested for both European forest plants (Svenning et al., 2008) and carabid beetles (Calatayud et al., 2019).

#### 4.4 | Impact of barriers on ecology, assemblage richness and uniqueness

Dispersal barriers have had a profound impact on species biogeography and regional taxa pools. Taxa frequently occupy less than half of their potential Northern Hemisphere climate range. These occupancy rates possess a significant biogeographic signal, with that of European taxa being roughly twice that of North American (median scores of 62 vs. 34%). This difference seems driven partly by the relative scarcity of the European boreal-maritime climate across the rest of the Holarctic. Range limitation from barriers has also led to regional taxa pools being three times smaller than they would be if ranges were in equilibrium with climate. This generates strong distance decay in regional taxa pools with observed overlap between regional pools falling by 50% after ~2600 km; in the absence of dispersal barriers there is no predicted change across the extent of the Holarctic. The age and stability of these barriers are also likely important, with abrupt but dynamic small-scale climatic transitions often leading to smaller faunal differences than longer-term and more stable ocean barriers related to plate tectonics (Ficetola et al., 2017).

Barriers and dispersal limitation thus appear to have profound implications upon ecological and biodiversity pattern and process.

First, these factors make it impossible for competitive sorting to solely explain composition because the bulk of species simply never have the opportunity to come into direct contact (Brown & Kurzius, 1987; Harmon & Harrison, 2015). Second, theoretical analyses have shown that typical species abundance distribution shapes are most likely related to incomplete dispersal and recruitment (Chave et al., 2002). As a result, dispersal barriers may well represent their underlying universal root cause (McGill, 2003; McGill & Nekola, 2010). The relaxation of dispersal limitation via anthropogenic activities thus has the potential to profoundly alter the shape of biodiversity statistical distributions by altering diversity generation and maintenance processes, especially for competitive co-equivalents, which require dispersal barriers to maintain co-existence (Shmida & Ellner, 1984).

#### ACKNOWLEDGMENTS

The study was primarily funded by the Czech Science Foundation (GA20-18827S).

#### DATA AVAILABILITY STATEMENT

All occurrence data used in this study are provided in the associated Supplementary File 'occurrences\_all.csv'.

#### ORCID

Jeffrey C. Nekola  <https://orcid.org/0000-0001-6073-0222>

#### REFERENCES

- Barker, G. M. (2001). Gastropods on land: Phylogeny, diversity and adaptive morphology. In G. M. Barker (Ed.), *The biology of terrestrial molluscs* (pp. 1–139). CABI Publishing.
- Boyce, M. S., Vernier, P. R., Nielsen, S. E., & Schmiegelow, F. K. A. (2002). Evaluating resource selection functions. *Ecological Modelling*, 157(2), 281–300. [https://doi.org/10.1016/S0304-3800\(02\)00200-4](https://doi.org/10.1016/S0304-3800(02)00200-4)
- Brown, J. L., Holl, D. J., Dolan, A. M., Carnaval, A. C., & Haywood, A. M. (2018). PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Scientific Data*, 5, 180254. <https://doi.org/10.1038/sdata.2018.254>
- Brown, J. H., & Kurzius, M. A. (1987). Composition of desert rodent faunas: Combinations of coexisting species. *Annales Zoologici Fennici*, 24, 227–237.
- Brunton, D. F. (1986). Status of the Southern Maidenhair Fern, *Adiantum capillus-veneris* (Adiantaceae), in Canada. *Canadian Field Naturalist*, 100, 404–408.
- Calatayud, J., Rodríguez, M. Á., Molina-Venegas, R., Leo, M., Horreo, J. L., & Hortal, J. (2019). Pleistocene climate change and the formation of regional species pools. *Proceedings of the Royal Society B: Biological Sciences*, 286, 2019021. <https://doi.org/10.1098/rspb.2019.0291>
- Carcaillet, C., Latil, J.-L., Abou, S., Ali, A., Ghaleb, B., Magnin, F., Roiron, P., & Aubert, S. (2018). Keep your feet warm? A cryptic refugium of trees linked to a geothermal spring in an ocean of glaciers. *Global Change Biology*, 24, 2476–2487. <https://doi.org/10.1111/gcb.14067>
- Castellanos, A. A., Huntley, J. W., Voelker, G., & Lawing, A. M. (2019). Environmental filtering improves ecological niche models across multiple scales. *Methods in Ecology and Evolution*, 10, 481–492. <https://doi.org/10.1111/2041-210X.13142>
- Chave, J., Muller-Landau, H. C., & Levin, S. A. (2002). Comparing classical community models: Theoretical consequences for patterns of diversity. *The American Naturalist*, 159, 1–23. <https://doi.org/10.1086/324112>

- Cowell, R. K., & Rangel, T. F. (2009). Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 19641–19658.
- Darwin, C. (1859). *On the origin of species*. John Murray.
- Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M. M., Townsend Peterson, A., ... Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>
- Ensing, D. J., Moffat, C. E., & Pither, J. (2013). Taxonomic identification errors generate misleading ecological niche model predictions of an invasive hawkweed. *Canadian Journal of Botany*, 91, 137–147. <https://doi.org/10.1139/cjb-2012-0205>
- Ficetola, G. F., Mazel, F., & Thuiller, W. (2017). Global determinants of zoogeographical boundaries. *Nature Ecology & Evolution*, 1, 0089. <https://doi.org/10.1038/s41559-017-0089>
- Gilpin, M. E., & Soulé, M. E. (1986). Minimum viable populations: Processes of species extinction. In M. E. Soulé (Ed.), *Conservation biology* (pp. 19–34). Sinauer Associates.
- Gittenberger, E., Groenenberg, D. S. J., Kokshoon, B., & Preece, R. C. (2006). Molecular trails from hitch-hiking snails. *Nature*, 439, 409. <https://doi.org/10.1038/439409a>
- Halloy, S. R. P. (1998). A theoretical framework for abundance distributions in complex systems. *Complexity International*, 6, 12.
- Hanski, I. (1999). Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes. *Oikos*, 87, 209–219. <https://doi.org/10.2307/3546736>
- Harmon, L. J., & Harrison, S. (2015). Species diversity if dynamic and unbounded at local and continental scales. *The American Naturalist*, 185, 584–593.
- Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical Science*, 3, 297–318.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978. <https://doi.org/10.1002/joc.1276>
- Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2017). *dismo: species distribution modeling*. R package version 1.1-4. <https://CRAN.R-project.org/package=dismo>
- Hilborn, R., & Mangel, M. (1997). *The ecological detective: Confronting models with data*. Princeton University Press.
- Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., & Guisan, A. (2006). Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, 199, 142–152. <https://doi.org/10.1016/j.ecolmodel.2006.05.017>
- Horsák, M., Horsáková, V., Divišek, J., & Nekola, J. C. (2022). Ecological niche divergence between extant and glacial land snail populations explained. *Scientific Reports*, 12, 806. <https://doi.org/10.1038/s41598-021-04645-2>
- Horsák, M., Juříčková, L., Škodová, J., & Ložek, V. (2016). *Pupilla alluvionica* Meng et Hoffmann, 2008: A land snail extant in the Altai refugium recognised for the first time in Central European early-middle Pleistocene glacials. *Malacologia*, 59, 223–230.
- Horsák, M., & Meng, S. (2018). *Punctum lozeki* n. sp. – a new minute land-snail species (Gastropoda: Punctidae) from Siberia and Alaska. *Malacologia*, 62, 11–20.
- Horsáková, V., Hájek, M., Hájková, P., Dítě, D., & Horsák, M. (2018). Principal factors controlling the species richness of European fens differ between habitat specialists and matrix-derived species. *Diversity and Distributions*, 24, 742–754. <https://doi.org/10.1111/ddi.12718>
- Horsáková, V., Nekola, J. C., & Horsák, M. (2020). Integrative taxonomic consideration of the Holarctic *Euconulus fulvus* group of land snails (Gastropoda, Stylommatophora). *Systematics and Biodiversity*, 18, 142–160.
- Hubbell, S. P. (2001). *A unified theory of biodiversity and biogeography*. Princeton University Press.
- Humboldt, A., & Bonpland, A. (1805). *Essai sur la géographie des plantes: accompagné d'un tableau physique des régions équinoxiales, fondé sur des mesures exécutées, depuis le dixième degré de latitude boréale jusqu'au dixième degré de latitude australe, pendant les années 1799, 1800, 1801, 1802 et 1803*. Chez Levrault, Schoell et compagnie, libraires.
- Hurlbert, A. H., & White, E. P. (2005). Disparity between range map- and survey-based analyses of species richness: Patterns and implications. *Ecology Letters*, 8, 319–327. <https://doi.org/10.1111/j.1461-0248.2005.00726.x>
- Jordan, D. S. (1905). The origin of species through isolation. *Science*, 22, 545–562. <https://doi.org/10.1126/science.22.566.545>
- Karger, D. N., Conrad, O., Böhrner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4, 170122. <https://doi.org/10.1038/sdata.2017.122>
- Kerney, M. (1999). *Atlas of the land and freshwater molluscs of Britain and Ireland*. Harley Books.
- Krebs, C. J. (1985). *Ecology: The experimental analysis of distribution and abundance*. Harper & Row.
- Leibold, M. A., & Chase, J. M. (2018). *Metacommunity ecology*. *Monographs in population biology* #59. Princeton University Press.
- Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, 33, 103–114. <https://doi.org/10.1111/j.1600-0587.2009.06039.x>
- Lomolino, M. V., Riddle, B. R., Whittaker, R. J., & Brown, J. H. (2010). *Biogeography* (4th ed.). Sinauer Associates.
- Mayr, E. (1942). *Systematics and the origin of species from the viewpoint of a zoologist*. Columbia University Press.
- McGill, B. (2003). Strong and weak tests of macroecology theory. *Oikos*, 102, 679–685.
- McGill, B., & Nekola, J. C. (2010). Mechanisms in macroecology: AWOL or purloined letter? Towards a pragmatic view of mechanism. *Oikos*, 119, 591–603. <https://doi.org/10.1111/j.1600-0706.2009.17771.x>
- Meier, R., & Dikow, T. (2004). Significance of specimen databases from taxonomic revisions for estimating and mapping the global species diversity of invertebrates and repatriating reliable specimen data. *Conservation Biology*, 18, 478–488. <https://doi.org/10.1111/j.1523-1739.2004.00233.x>
- Munguía, M., Peterson, A. T., & Sánchez-Cordero, V. (2008). Dispersal limitation and geographical distribution of mammal species. *Journal of Biogeography*, 35, 1879–1887.
- Navarro-Racines, C., Tarapues, J., Thornton, P., Jarvis, A., & Ramirez-Villegas, J. (2020). High-resolution and bias-corrected CMIP5 projections for climate change impact assessments. *Scientific Data*, 7, 7. <https://doi.org/10.1038/s41597-019-0343-8>
- Nekola, J. C. (1999). Paleoreugia and neoreugia: The influence of colonization history on community pattern and process. *Ecology*, 80, 2459–2473.
- Nekola, J. C. (2014). North American terrestrial gastropods through either end of a spyglass. *Journal of Molluscan Studies*, 80, 238–248.
- Nekola, J. C., Chiba, S., Coles, B. F., Drost, C. A., von Proschwitz, T., & Horsák, M. (2018). A phylogenetic overview of the genus *Vertigo* O. F. Müller, 1773 (Gastropoda: Pulmonata: Pupillidae: Vertigininae). *Malacologia*, 62, 21–161.
- Nekola, J. C., & Coles, B. F. (2010). Pupillid land snails of eastern North America. *American Malacological Bulletin*, 28, 29–57. <https://doi.org/10.4003/006.028.0221>
- Nekola, J. C., Coles, B. F., & Horsák, M. (2015). Species assignment in *Pupilla* (Gastropoda: Pulmonata: Pupillidae): Integration of

- DNA-sequence data and conchology. *Journal of Molluscan Studies*, 81, 196–216. <https://doi.org/10.1093/mollus/eyu083>
- Nekola, J. C., & Horsák, M. (2022). The impact of statistically unchallenged taxonomic concepts on ecological assemblages across multiple spatial scales. *Ecography*, 2022, e06063.
- Nekola, J. C., Hutchins, B. T., Schofield, A., Najev, B., & Perez, K. E. (2019). *Caveat Consumptor Notitia Museo*: Let the museum data user beware. *Global Ecology and Biogeography*, 28, 1722–1734.
- Nekola, J. C., & McGill, B. (2014). Scale dependency in the functional form of the distance decay relationship. *Ecography*, 37, 309–320. <https://doi.org/10.1111/j.1600-0587.2013.00407.x>
- Nekola, J. C., Šizling, A. L., Boyer, A. G., & Storch, D. (2008). Artifacts in the log-transformation of species abundance distributions. *Folia Geobotanica*, 43, 259–268. <https://doi.org/10.1007/s12224-008-9020-y>
- Nekola, J. C., & White, P. S. (1999). Distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, 26, 867–878.
- Němec, T., Liznarová, E., Birkhofer, K., & Horsák, M. (2021). Stable isotope analysis suggests low trophic niche partitioning among co-occurring land snail species in a floodplain forest. *Journal of Zoology*, 313, 297–306. <https://doi.org/10.1111/jzo.12859>
- Newman, M. E. J. (2005). Power laws, Pareto distributions, and Zipf's law. *Contemporary Physics*, 46, 323–351. <https://doi.org/10.1080/00107510500052444>
- Örstan, A., Sparks, J., & Pearce, T. (2011). Wayne Grimm's legacy: A 40-year experiment on the dispersal of *Cepaea nemoralis* in Frederick County, Maryland. *American Malacological Bulletin*, 29, 139–142.
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, 31, 161–175. <https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications*, 19, 181–197. <https://doi.org/10.1890/07-2153.1>
- Pilsbry, H. A. (1948). Land Mollusca of North America (North of Mexico). Monograph 1(1)-2(2), Academy of Natural Sciences of Philadelphia.
- Preston, F. W. (1960). Time and space and the variation of species. *Ecology*, 41, 612–627. <https://doi.org/10.2307/1931793>
- Ricklefs, R. E., & Schluter, D. (1993). Species diversity: Regional and historical influences. In R. E. Ricklefs, & D. Schluter (Eds.), *Species diversity in ecological communities* (pp. 350–363). University of Chicago Press.
- Riddle, B. R., & Hafner, D. J. (2010). Integrating pattern with process at biogeographic boundaries: The legacy of Wallace. *Ecography*, 33, 321–325. <https://doi.org/10.1111/j.1600-0587.2010.06544.x>
- Scheffer, V. B. (1993). The Olympic goat controversy: A perspective. *Conservation Biology*, 7, 916–920. <https://doi.org/10.1046/j.1523-1739.1993.740916.x>
- Schlick-Steiner, B. C., Steiner, F. M., Seifert, B., Stauffer, C., Christian, E., & Crozier, R. H. (2010). Integrative taxonomy: A multisource approach to exploring biodiversity. *Annual Review of Entomology*, 55, 421–438. <https://doi.org/10.1146/annurev-ento-112408-085432>
- Shmida, A., & Ellner, S. (1984). Coexistence of plant species with similar niches. *Vegetatio*, 8, 29–55. <https://doi.org/10.1007/BF00044894>
- Svenning, J.-C., Eiserhardt, W. L., Normand, S., Orodóñez, A., & Sandel, B. (2015). The influence of paleoclimate on present-day patterns in biodiversity and ecosystems. *Annual Review of Ecology, Evolution and Systematics*, 46, 551–572. <https://doi.org/10.1146/annurev-ecolsys-112414-054314>
- Svenning, J.-C., Normand, C., & Skov, F. (2008). Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography*, 31, 316–326.
- Svenning, J.-C., & Skov, F. (2004). Limited filling of the potential range in European tree species. *Ecology Letters*, 7, 565–573. <https://doi.org/10.1111/j.1461-0248.2004.00614.x>
- Svenning, J.-C., & Skov, F. (2007). Ice age legacies in the geographical distribution of tree species richness in Europe. *Global Ecology and Biogeography*, 16, 234–245.
- Sysoev, A., & Schileyko, A. (2009). *Land snails and slugs of Russia and adjacent countries*. Pensoft.
- Title, P. O., & Bemmels, J. B. (2018). ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*, 41, 291–307. <https://doi.org/10.1111/ecog.02880>
- Turgeon, D. D., Quinn, J. F. Jr, Bogan, A. E., Coan, E. V., Hochberg, F. G., Lyons, W. G., & Williams, J. D. (1998). *Common and scientific names of aquatic invertebrates from the United States and Canada: Mollusks* (2nd ed.). American Fisheries Society Special Publication 26.
- Varela, S., Anderson, R. P., García-Valdés, R., & Fernández-González, F. (2014). Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography*, 37, 1084–1091.
- Waldén, H. (2007). *Svensk landmolluskatlas*. Naturcentrum AB.
- Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21, 335–342. <https://doi.org/10.1890/10-1171.1>
- Welter-Schultes, F. W. (2012). *European non-marine molluscs, a guide for species identification*. Planet Poster Editions.
- Will, K. W., Mishler, B. D., & Wheeler, Q. D. (2005). The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology*, 54, 844–851. <https://doi.org/10.1080/10635150500354878>
- Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A., & NCEAS Predicting Species Distributions Working Group. (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, 14, 763–773. <https://doi.org/10.1111/j.1472-4642.2008.00482.x>
- Zurell, D., Franklin, J., König, C., Bouchet, P. J., Serra-Diaz, J. M., Dormann, C. F., Elith, J., Fandos Guzman, G., Feng, X., Guillera-Aroita, G., Guisan, A., Leitão, P. J., Lahoz-Monfort, J. J., Park, D. S., Peterson, A. T., Rapacciuolo, G., Schmatz, D. R., Schröder, B., Thuiller, W., ... Merow, C. (2020). A standard protocol for describing species distribution models. *Ecography*, 43, 1261–1277. <https://doi.org/10.1111/ecog.04960>

## BIOSKETCH

JEFFREY C. NEKOLA is an associate professor in the Department of Botany and Zoology at Masaryk University. His work spans the range of biodiversity studies, ranging from molecular genetics to macroecology and general theory.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Nekola, J. C., Divíšek J., & Horsák M. (2022). The nature of dispersal barriers and their impact on regional species pool richness and turnover. *Global Ecology and Biogeography*, 00, 1–31. <https://doi.org/10.1111/geb.13517>

## APPENDIX A

**Taxonomic control, taxa excluded from the modelling, and impact of limiting modelling on empirically vetted concepts**

We treated as distinct the western North American populations of *Euconulus alderi* and the eastern Eurasian populations of *Vertigo lilljeborgi* because of their unique sequence signatures in combination with geographic isolation. The only reason these entities have not yet been formally described is that there are not yet enough known sites and material to accurately demarcate their genetic/morphological variability and geographic range. We did not model *Vertigo* aff. *hoppi* because we are currently unable to distinguish it based on shell features and know of only three genetically confirmed sites. Additionally, because of uncertainty regarding taxonomic status (Nekola et al., 2018), we chose to lump *Vertigo ronneybyensis* + *Vertigo ultimathule* and *Vertigo coloradensis* + *Vertigo cristata* + *Vertigo pisewensis*.

Taxa with five or fewer modern occurrences (*Pupilla limata*, *Vertigo binneyana*, *Vertigo chytryi*) were not modelled. Even though falling under this threshold, we modelled western North American *E. alderi*, and Japanese *V. lilljeborgi* and *V. lilljeborgi vinlandica* by lumping their occurrences into remaining *E. alderi* and *V. lilljeborgi* records, respectively. This approach appeared justified as the resultant suitability models accurately portrayed not only the known ranges of these taxa but their better-known European relatives as well. *Pupilla alluvionica*, although having 14 known occurrences, was also not modelled

due to extreme spatial clustering and climatic homogeneity within these records (Horsák et al., 2016), which limits global interpolation due to large potential spatial autocorrelation effects.

The impact of using empirically vetted concepts from integrative taxonomic revision on analyses was estimated by calculating the overlap of climate polygons in 2-D principal components analysis (PCA) space for unvetted versus vetted taxonomic concepts in 10 taxa that required lumping, splitting, and/or altered diagnostic features (1 *Euconulus*, 4 *Pupilla*, 5 *Vertigo*). Traditional taxonomic concepts that required splitting into multiple distinct forms (e.g. *E. alderi*, *Pupilla muscorum*, *Vertigo genesii*) possessed initial PCA climate polygons up to 76% too large, while those that required other forms to be lumped into them (e.g. *Vertigo arthuri*, *Vertigo ventricosa*) possessed initial polygons up to 3.5 times too small. Species that had altered diagnostic characters (e.g. *Pupilla blandi*, *V. ronneybyensis*) expressed variable changes, ranging from a 50% decrease to threefold increase in polygon size (Figure A1). For instance, the continental range of *V. arthuri* would not have been apparent if the seven forms that had been incorrectly split from it had not been included. The limitation of *Vertigo alpestris* to western Eurasia would have been lost if *Vertigo beringiana* had been incorrectly lumped with it. And the low elevation eastern Rockies/Great Plains range of *P. blandi* would not have been generated if identification used traditional diagnostic features leading to conflation with either *P. muscorum* or the mid-high elevation forest *Pupilla hebes pithodes*.

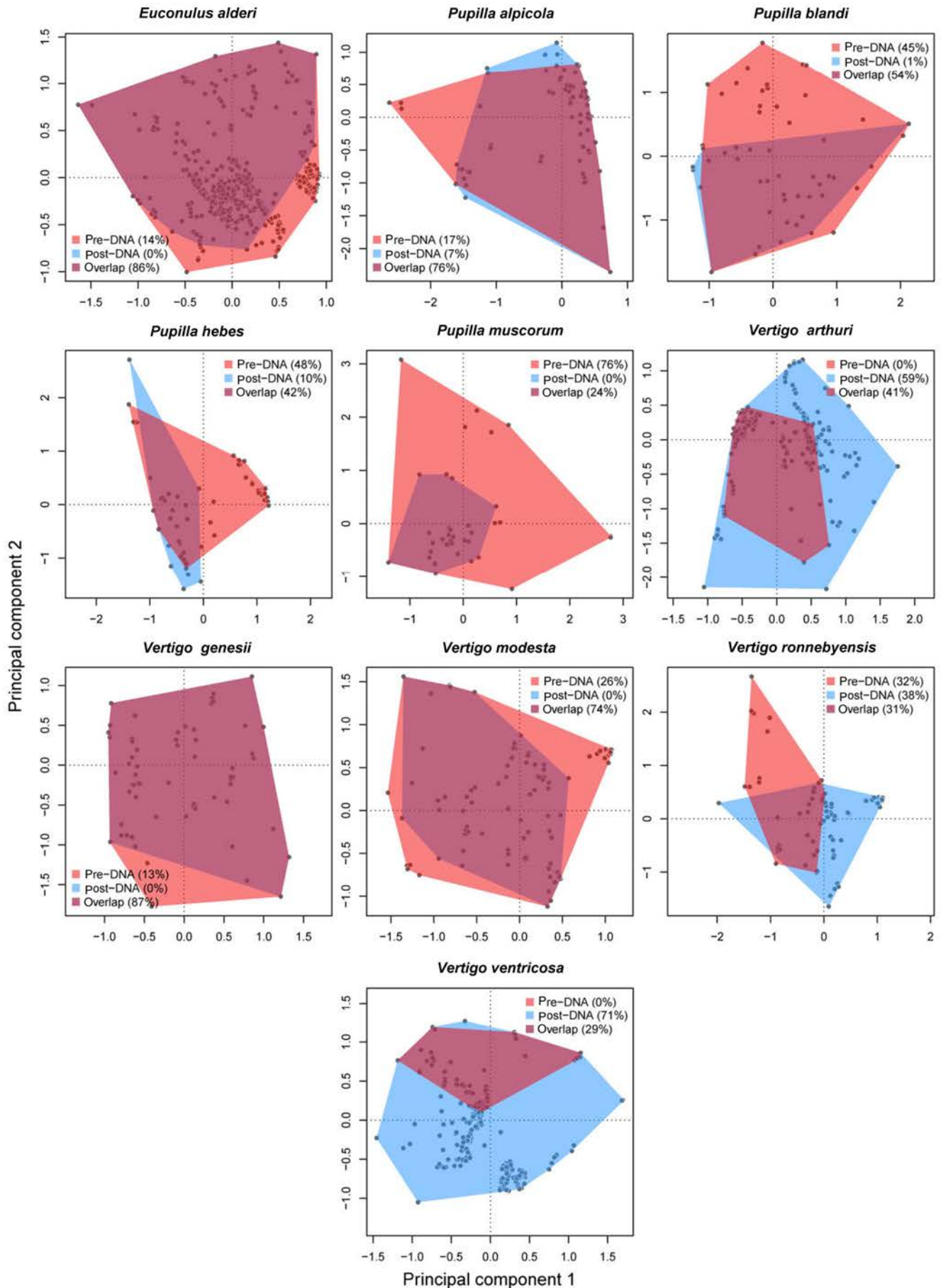


FIGURE A1 Comparison of overlap between niche polygons in 2-D principal components analysis (PCA) climate space before and after empirical integrative taxonomic revision



## APPENDIX B.

## Range and barrier statistics with the number of sites used to parameterize models

Species	Median longitude	Affinity	# Sites		Range extent (km)				Range size (km <sup>2</sup> )		Occupancy (%)		Barrier		
			Total	Filtered	Actual	Minor	Major	Potential	Actual	Potential	Type	Distance (km)	East	West	
					Major	Minor	Major	Minor			East	West	East	West	
<i>Euconulus alderi</i> (Gray, 1840)	39.4 E	E	471	205	6000	3200	7000	3500	11,496,357	19,571,937	58.7	C	O	2300	925
<i>Euconulus alderi</i> (NAM subpopulation)	134.5 W	N	-	-	4200	1000	4800	1000	4,554,682	19,571,937	23.3	H	C	-	2300
<i>Euconulus fresti</i> Horsáková et al., 2020	100.9 W	N	219	139	3900	2800	4200	2800	4,350,428	13,003,671	33.5	O	C	4800	4900
<i>Euconulus fulvus</i> (Müller, 1774)	138.8 E	B	1702	557	19,800	4600	19,800	4600	41,899,670	41,899,670	100.0	-	-	-	-
<i>Euconulus polygyratus</i> (Pilsbry, 1899)	80.1 W	N	187	106	2100	1000	3300	1000	2,767,890	5,227,391	53.0	O	C	4600	1500
<i>Pupilla alaskensis</i> Nekola et al., 2015	142.6 W	B	63	63	1300	800	12,000	3400	2,520,035	16,507,398	15.3	H	OH	-	200
<i>Pupilla alluvionica</i> Meng & Hoffman, 2008	87.4 E	B	-	-	200	200	-	-	-	-	-	-	-	-	-
<i>Pupilla alpicola</i> (Charpentier, 1837)	49.3 E	E	97	80	2900	1000	3000	2800	7,367,236	12,476,040	59.1	H	O	-	1000
<i>Pupilla blandi</i> Morse, 1865	104.5 W	N	24	24	1800	1000	5800	2000	4,836,699	14,771,582	32.7	CO	H	8200	-
<i>Pupilla hebes</i> (Ancey, 1881)	111.5 W	N	152	81	2100	800	2600	2500	2,306,398	4,963,607	46.5	H	C	-	9200
<i>Pupilla hokkaidoensis</i> Nekola et al., 2015	144.1 E	B	6	6	250	100	6300	1900	4,010,415	15,546,240	25.8	H	H	-	-
<i>Pupilla hudsonianum</i> Nekola et al., 2015	85.2 W	N	49	49	5600	1200	7000	2800	6,612,281	19,953,426	33.1	O	H	3800	-
<i>Pupilla limata</i> Schileyko, 1989	155.1 E	B	-	-	1000	500	-	-	-	-	-	-	-	-	-
<i>Pupilla loessica</i> Ložek, 1954	90.9 E	B	57	57	3000	500	5200	1200	4,911,392	8,633,772	56.9	H	C	-	500
<i>Pupilla muscorum</i> (Linnaeus, 1758)	16.4 E	E	163	134	4800	2300	4800	2800	7,905,603	11,716,070	67.5	C	O	1600	3000
<i>Pupilla sterrii</i> (Forster, 1840)	31.2 E	E	44	44	1000	1000	4200	2800	4,326,651	12,034,843	36.0	C	H	1000	-
<i>Pupilla triplicata</i> (Studer, 1820)	44.1 E	E	51	51	6100	1000	7300	2100	7,127,344	10,663,523	66.8	H	H	-	-
<i>Pupilla turcmenia</i> (O. Boettger, 1889)	77.9 E	B	57	57	2500	700	8100	3800	7,173,659	14,776,715	48.6	H	H	-	-
<i>Vertigo alpestris</i> Alder, 1838	17.4 E	E	239	200	2600	1500	3000	3700	3,505,027	5,140,840	68.2	C	OI	4400	2500

(Continues)

## APPENDIX B. (Continued)

Species	Median longitude	Affinity	# Sites		Range extent (km)			Range size (km <sup>2</sup> )			Occupancy (%)		Barrier			
			Total	Filtered	Actual	Minor	Major	Potential	Actual	Minor	Major	Potential	Type	Distance (km)		
					Major	Minor	Major	Minor	Major	Minor	Major	Minor	East	West	East	West
<i>Vertigo arctica</i> (Wallenberg, 1858)	3.8 E	E	379	199	2700	1900	5300	2000	2,189,285	2000	9,173,261	23.9	C	OI	2200	1200
<i>Vertigo arthuri</i> von Martens, 1882	102.5 W	N	455	171	6800	2500	7300	3400	8,796,787	3400	25,935,890	33.9	C	CO	4000	1400
<i>Vertigo beringiana</i> Nekola et al., 2018	155.0 E	B	57	57	7600	1000	12,100	3400	13,164,207	3400	18,755,093	70.2	H	H	-	-
<i>Vertigo binneyana</i> Sterki, 1890	104.5 W	N	-	-	1100	600	-	-	-	-	-	-	-	-	-	-
<i>Vertigo bollesiana</i> (Morse, 1865)	80.6 W	N	223	149	2000	1000	2600	1000	2,266,691	1000	3,076,278	73.7	OC	C	6500	1800
<i>Vertigo chyttryi</i> Nekola et al., 2018	88.2 E	B	-	-	1500	750	-	-	-	-	-	-	-	-	-	-
<i>Vertigo circumlabiata</i> Schileyko, 1984	176.7 E	B	44	44	4600	1000	13,400	2300	3,668,278	2300	14,525,437	25.3	H	C	-	2100
<i>Vertigo columbiana</i> Sterki, 1892	132.7 W	N	43	43	2700	800	6,000	1200	1,904,142	1200	3,789,168	50.3	CO	H	7600	-
<i>Vertigo cristata</i> Sterki, 1919 (aggregate)	98.9 W	N	636	190	7200	4800	16,500	4800	13,692,840	4800	34,183,042	40.1	O	CO	4000	700
<i>Vertigo extima</i> (Westerlund, 1877)	46.60 E	E	45	45	3100	700	4600	1200	2,320,676	1200	7,677,454	30.2	C	O	2400	1100
<i>Vertigo genesii</i> (Gredler, 1856)	13.67 E	E	120	80	2500	1400	3200	2800	2,845,534	2800	4,419,752	64.4	C	O	4500	1100
<i>Vertigo genesioides</i> Nekola et al., 2018	169.0 E	B	80	80	12,500	2000	15,200	5300	28,073,020	5300	33,320,590	84.3	O	H	2300	-
<i>Vertigo geyeri</i> Lindholm, 1925	11.9 E	E	473	204	2800	2200	2900	2800	4,140,154	2800	5,816,677	71.2	C	O	1,500	1,000
<i>Vertigo hannai</i> Pilsbry, 1919	118.8 W	N	78	78	3500	2000	14,300	5000	10,728,659	5000	33,446,437	32.1	H	OH	-	200
<i>Vertigo cf. hoppi</i> (Möller, 1842)	172.3 E	B	-	-	8200	1000	-	-	-	-	-	-	-	-	-	-
<i>Vertigo kurilensis</i> Nekola et al., 2018	149.9 E	B	14	14	1400	200	8000	200	329,983	200	1,366,601	24.2	O	C	700	8400
<i>Vertigo kushiroensis</i> Pilsbry & Hirase, 1905	115.5 E	B	65	65	4500	1500	8800	1500	5,606,816	1500	11,068,907	50.7	O	H	3800	-
<i>Vertigo liljeborgi</i> (Westerlund, 1871)	38.8 E	E	286	220	4900	2600	6400	3100	8,329,218	3100	14,452,455	57.6	C	OI	2500	2700

## APPENDIX B. (Continued)

Species	Median longitude	Affinity	# Sites		Range extent (km)			Range size (km <sup>2</sup> )		Occupancy (%)		Barrier			
			Total	Filtered	Actual	Minor	Major	Minor	Actual	Potential	East	West	East	West	
					Major	Minor	Major	Minor			East	West	East	West	
<i>Vertigo liljeborgi vinlandica</i> Nekola et al., 2018	67.4 W	B	-	-	1000	200	2200	2100	1,240,282	14,452,455	8.6	CO	C	4000	2500
<i>Vertigo liljeborgi</i> (East Asian)	144.2 E	N	-	-	100	100	8000	1600	1,956,887	14,452,455	13.5	OI	CO	2700	4000
<i>Vertigo microspiraera</i> Schileyko, 1984	148.2 E	B	51	51	8300	1201	13,000	1900	5,699,209	10,483,635	54.4	H	H	-	-
<i>Vertigo modesta</i> (Say, 1824)	102.9 W	N	310	107	6700	4800	17,000	5500	15,529,780	39,998,183	38.8	OI	OH	1400	200
<i>Vertigo morsei</i> Sterki, 1894	89.2 W	N	61	61	3600	1300	4000	2000	5,564,219	17,933,706	31.0	CO	C	5400	6100
<i>Vertigo nylanderi</i> Sterki, 1909	82.6 W	N	119	80	3500	800	4100	1300	3,431,968	13,379,671	25.7	O	C	3900	6200
<i>Vertigo oughtoni</i> Pilsbry, 1948	106.4 W	N	35	35	5200	1400	13,500	3000	9,418,521	25,302,356	37.2	OI	OH	3500	200
<i>Vertigo parcedentata</i> (Braun, 1847)	49.0 E	E	21	21	4700	1500	17,500	2000	7,480,927	15,886,549	47.1	H	OI	-	2900
<i>Vertigo perryi</i> Sterki, 1905	71.3 W	N	47	47	2800	600	2800	700	1,067,533	1,937,949	55.1	O	C	3900	2200
<i>Vertigo pseudosubstriata</i> Ložek, 1954	80.3 E	B	11	11	1500	500	16,000	5600	13,471,887	37,917,220	35.5	H	H	-	-
<i>Vertigo pygmaea</i> (Draparnaud, 1801)	39.1 E	E	667	337	6500	2600	7100	2600	10,570,127	15,580,638	67.8	C	O	4800	1000
<i>Vertigo ronneyensis</i> (Westerlund, 1871)	135.8 E	B	331	181	13,400	3000	16,200	3000	25,811,166	27,662,503	93.3	O	O	500	800
<i>Vertigo substriata</i> (Jeffreys, 1833)	37.8 E	E	522	295	6,200	3100	7000	3300	9,283,118	13,412,829	69.2	C	O	3400	1000
<i>Vertigo ultima</i> Pilsbry, 1919	106.6 W	N	26	26	6,000	1100	15,100	3000	4,946,733	24,729,534	20.0	OI	OH	2300	200
<i>Vertigo ventricosa</i> (Morse, 1865)	100.9 W	N	405	163	6,300	4100	6700	4100	8,881,393	24,675,136	36.0	O	CO	4100	3200

Biogeographic affinity codes: E, European; B, Beringian; N, North America.

## APPENDIX C.

Protocol generated by ODMAP 1.0 (Zurell et al., 2020)

## THE NATURE OF DISPERSAL BARRIERS AND THEIR IMPACT ON REGIONAL SPECIES POOL RICHNESS AND TURNOVER

2021-11-01

### OVERVIEW

#### Authorship

Nekola, J. C., J. Divisek & M. Horsak

#### Model objective

Model objective: Mapping and interpolation.

Target output: Estimation of species potential ranges.

#### Focal taxon

Focal taxon: Small land snails from genera *Euconulus*, *Pupilla* and *Vertigo*.

#### Location

Location: Holarctic region.

#### Scale of analysis

Spatial extent: -180, 180, 30, 90 (xmin, xmax, ymin, ymax).

Spatial resolution: 7681.097 m.

Temporal extent: 1996-2018.

Boundary: Rectangle.

#### Biodiversity data

Observation type: Field survey, museum collections.

Response data type: Presence-only.

#### Predictors

Predictor types: Climatic.

#### Hypotheses

Hypotheses: Species distributions at this spatial scale are mainly driven by suitable environment (climate) and dispersal barriers.

#### Assumptions

Model assumptions: Relevant ecological drivers (or proxies) of species distributions are included. Gathered occurrence data well represent the geographic and climatic range of each taxon (any biases are accounted for/corrected or discussed). Detectability does not change across habitat gradients. All occurrence records are taxonomically vetted.

#### Algorithms

Modelling techniques: Maxent.

Model complexity: To reduce model complexity, each model was calibrated using taxon-specific subset of best-performing climatic variables, which were selected based on fivefold cross-validation.

We used the default parameterization of the 'auto-feature' option. This uses all available mathematical transformations of predictor variables for sample sizes > 80, but only linear, quadratic and hinge features for sample sizes between 15 and 79.

Model averaging: No.

#### Workflow

Model workflow: (a) Taxonomic revision of occurrence data; (b) environmental filtering of occurrence data (sensu Castellanos et al., 2019; Varela et al., 2014) to reduce uneven sampling intensity; (c) selection of taxon-specific subsets of best-performing climatic variables; (d) Maxent modelling with constrained sampling of background points; (e) model evaluation using fivefold cross-validation and AUC, AUC.diff and Boyce index; (f) model projection and critical revision of resulting maps of potential ranges.

#### Software

Software: Maxent (version 3.4.1), R (version 4.0.2), dismo package (version 1.1-4).

Code availability: Available on request.

Data availability: Available in Supporting Information as 'occurrences\_all.csv'.

### DATA

#### Biodiversity data

Taxon names: Taxon names are available in [Appendices B-D](#).

Taxonomic reference system: Traditional species and subspecies-level taxonomic concepts were based on Welter-Schultes (2012; Europe), Sysoev and Schileyko (2009; central and east Asia) and Turgeon et al. (1998; North America).

Ecological level: Species, operational taxonomic units.

Data sources: Our own field collections; Brian Coles collection at the National Museum of Wales; the Royal Ontario Museum collection; the National Museum of Canada collection; species distribution atlases (Kerney, 1999 for UK and Waldén, 2007 for Sweden). We avoided the use of external museum and other remote data sources as we could not independently verify identifications or make a simple and consistent, non-arbitrary rule on which data can be used (because are likely to be correct) and which cannot.

Sampling design: Stratified to capture maximum environmental variability.

Sample size: The numbers of occurrence records for each taxon are provided in [Appendix A](#).

Clipping: Holarctic region.

Scaling: Because non-uniform occurrence distribution leads to climate reporting bias, accurate and robust models require input data rarefaction. We accomplished this for taxa with more than 80 occurrences by standardizing and ordinating occurrence climate data using principal components analysis (PCA), with the number of dimensions representing those needed to explain at least 90% of observed variation. We then removed the point associated with the smallest average pairwise distance in PCA space, and recalculated

the distance matrix. We repeated this process until all distances were 0.1 SDs or greater, or the number of remaining records reached 80, whichever came first. The number of records used to calibrate each climate suitability model is shown in [Appendix A](#).

**Cleaning:** We limited analyses to the three genera that have undergone integrative empirical revision (e.g. Schlick-Steiner et al., 2010; Will et al., 2005) across the Holarctic: *Euconulus*, *Pupilla* and *Vertigo* (see Horsáková et al., 2020; Nekola et al., 2015, 2018). We considered species-level concepts to be empirically validated when they were shown to be distinct from their nearest evolutionary neighbours across a consensus of data streams including some reasonable subset of mitochondrial DNA (mtDNA) sequence, nuclear DNA (nDNA) sequence, conchology, genitalic anatomy, behavior, ecological preference and/or biogeography. The actual conchological features distinguishing each taxon were then determined. We reserved subspecies-level assignments to those taxa strongly supported in mtDNA but unsupported in the more slowly evolving nDNA amplicons which we employ. Such taxa tend to also possess weakly differentiated shell characters and a unique biogeographic range from the nominate subspecies.

**Absence data:** Not available.

**Background data:** 10,000 background points randomly assigned within well-surveyed regions defined by a minimum convex envelope extending 100 km beyond the most marginal occurrences in our dataset

#### Data partitioning

**Training data:** Fivefold cross-validation with random assignment of species occurrences to folds; final models based on 100% of data.

**Validation data:** Fivefold cross-validation with random assignment of species occurrences to folds.

#### Predictor variables

**Predictor variables:** Climatic variables from the WorldClim and ENVIREM databases.

**Data sources:** WorldClim (version 1.4; <https://www.worldclim.org/data/v1.4/worldclim14.html>); ENVIREM (<https://envirem.github.io/>).

**Spatial extent:** -180, 180, 30, 90 (xmin, xmax, ymin, ymax).

**Spatial resolution:** 7681.097 m.

**Coordinate reference system:** North Pole Lambert Azimuthal Equal Area projection (+proj=laea +lat\_0=90 +lon\_0=0 +x\_0=0 +y\_0=0 +datum=WGS84 +units=m +no\_defs).

**Temporal extent:** 1960-1990.

#### Transfer data

**Data sources:** Climatic variables from the WorldClim and ENVIREM databases as described above.

**Spatial extent:** -180, 180, 30, 90 (xmin, xmax, ymin, ymax).

**Spatial resolution:** 7681.097 m.

**Temporal extent:** North Pole Lambert Azimuthal Equal Area projection (+proj=laea +lat\_0=90 +lon\_0=0 +x\_0=0 +y\_0=0 +datum=WGS84 +units=m +no\_defs)

**Models and scenarios:** N/A.

**Quantification of novelty:** N/A.

## MODEL

### Variable pre-selection

**Variable pre-selection:** We limited model calibration to a taxon-specific subset of the best-performing variables selected based on fivefold cross-validation (see below).

#### Multicollinearity

**Multicollinearity:** We limited calibration to a subset of the best-performing variables. These were determined by first calibrating a set of Maxent models for a given taxon containing only one climatic variable and a constant. The performance of each was documented using fivefold cross-validation. Resulting values of area under a ROC curve (AUC) were used to rank each climate variable from the best to worst in terms of predictive value. The two best-performing variables were selected, and if their variance inflation factor (VIF) was lower than 10, the third best-performing variable was added to the subset and collinearity again checked. This process was repeated until the VIF of any variable in the subset exceeded 10.

#### Model settings

**maxent:** betamultiplier (1), removeduplicates (true), autofeature (true), addsamplestobackground (true), extrapolate (true), doclamp (true), maximumiterations (500), convergencethreshold (0.00001), lq2lqptthreshold (80), l2lqthreshold (10), hingethreshold (15).

**Model settings (extrapolation):** Clamping applied.

#### Model selection – model averaging – ensembles

**Model averaging:** No model averaging, no ensemble.

#### Threshold selection

**Threshold selection:** Potential range of habitat suitability was estimated by applying a threshold that balances training omission, predicted area and threshold value provided by Maxent, with areas above this threshold being considered climatically suitable. This was used as it generates the most liberal estimates of potential range and thus the smallest possible barrier widths. Actual barrier widths are thus actually likely larger.

## ASSESSMENT

### Performance statistics

**Performance on training data:** AUC, Boyce index.

**Performance on validation data:** AUC, AUC difference between training and testing data, Boyce index.

#### Plausibility check

**Response shapes:** We used partial dependence plots to check the ecological plausibility of fitted relationships in Maxent models.

**Expert judgement:** We critically assessed resulting maps of species potential ranges.



## APPENDIX D. (Continued)

Species	Biogeographic zone occurrence																	
	Actual									Potential								
	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
<i>Vertigo lilljeborgi vinlandica</i>	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+
<i>Vertigo lilljeborgi</i> (East Asian)	-	-	-	-	+	-	-	-	-	+	+	+	+	+	+	+	+	+
<i>Vertigo microsphaera</i>	-	-	-	+	+	+	-	-	-	-	+	+	+	+	+	+	+	+
<i>Vertigo modesta</i>	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>Vertigo morsei</i>	-	-	-	-	-	-	+	+	+	-	+	+	+	+	-	+	+	+
<i>Vertigo nylanderi</i>	-	-	-	-	-	-	-	+	+	+	+	+	+	+	-	+	+	+
<i>Vertigo oughtoni</i>	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	+	+	+
<i>Vertigo parcedentata</i>	-	+	-	+	-	-	-	-	-	-	+	+	+	+	+	+	+	+
<i>Vertigo perryi</i>	-	-	-	-	-	-	-	+	+	+	+	+	-	-	-	-	+	+
<i>Vertigo pseudosubstriata</i>	-	-	-	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+
<i>Vertigo pygmaea</i>	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+
<i>Vertigo ronneyensis</i>	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+
<i>Vertigo substriata</i>	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	-	+
<i>Vertigo ultima</i>	-	-	-	-	-	+	-	-	+	+	+	+	+	+	+	+	+	+
<i>Vertigo ventricosa</i>	-	-	-	-	-	+	+	+	+	-	+	+	+	+	+	+	+	+

<sup>a</sup> Unmodelled species.

APPENDIX E CLIMATIC SUITABILITY MAPS

*Euconulus alderi* | n = 471/205



*Euconulus fresti* | n = 219/139



*Euconulus fulvus* | n = 1702/557



*Euconulus polygyratus* | n = 187/106






*Pupilla alaskensis* | n = 63/63



*Pupilla alpicola* | n = 97/80



 Suitable climate  
 Unsuitable climate

 Records used for model calibration



*Pupilla blandi* | n = 24/24



*Pupilla hebes* | n = 152/81



*Pupilla hokkaidoensis* | n = 6/6



*Pupilla hudsonianum* | n = 49/49






*Pupilla loessica* | n = 57/57

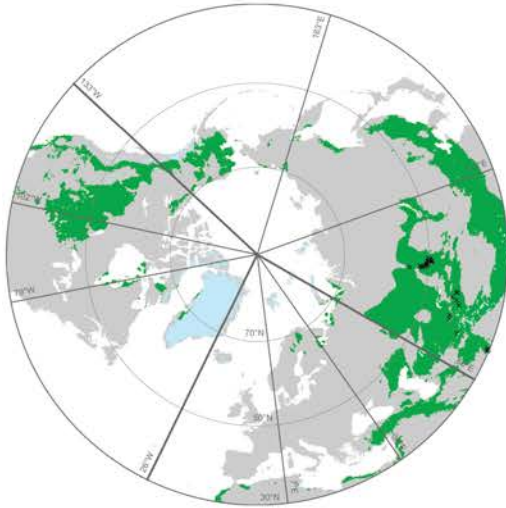
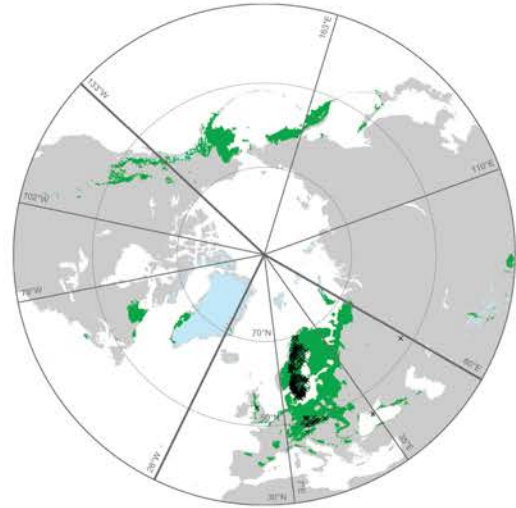
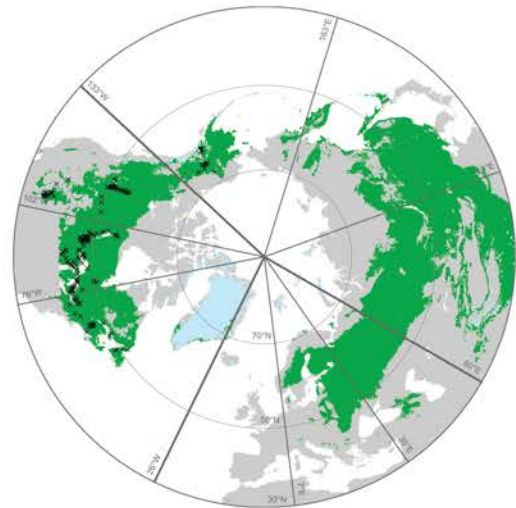




*Pupilla muscorum* | n = 163/134





 Suitable climate  
 Unsuitable climate

 Records used for model calibration

***Pupilla sterrii* | n = 44/44*****Pupilla triplicata* | n = 51/51*****Pupilla turkmenia* | n = 57/57*****Vertigo alpestris* | n = 239/200*****Vertigo arctica* | n = 379/199*****Vertigo arthuri* | n = 455/171**

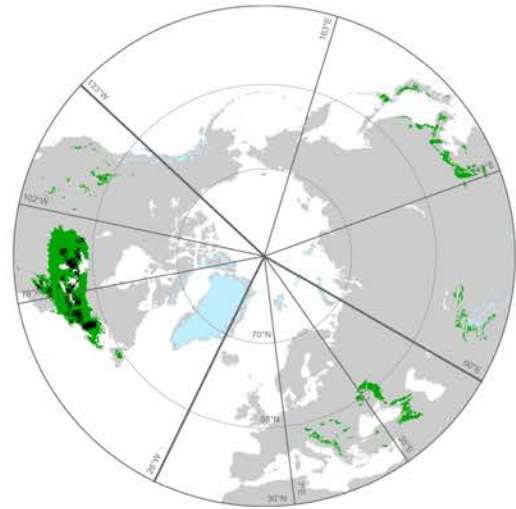
 Suitable climate  
 Unsuitable climate

  Records used for model calibration

*Vertigo beringiana* | n = 57/57



*Vertigo bollesiana* | n = 243/149



*Vertigo circumlabiata* | n = 44/44



*Vertigo columbiana* | n = 43/43





*Vertigo cristata* agg. | n = 636/190



*Vertigo extima* | n = 45/45



 Suitable climate  
 Unsuitable climate

x x  
x x Records used for  
model calibration

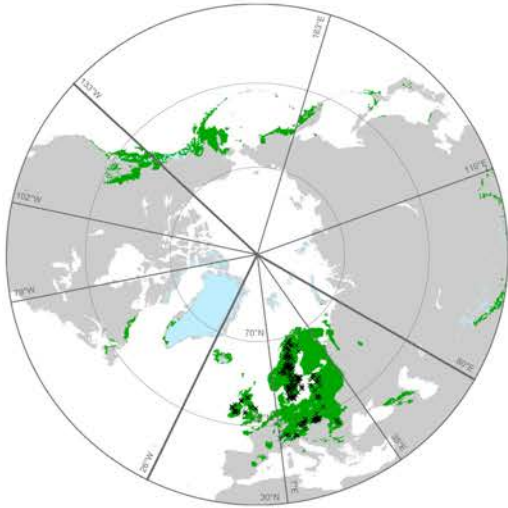
*Vertigo genesii* | n = 120/80



*Vertigo genesioides* | n = 80/80



*Vertigo geyeri* | n = 473/204



*Vertigo hannai* | n = 78/78







*Vertigo kurilensis* | n = 14/14



*Vertigo kushiroensis* | n = 65/65



 Suitable climate  
 Unsuitable climate

  Records used for model calibration

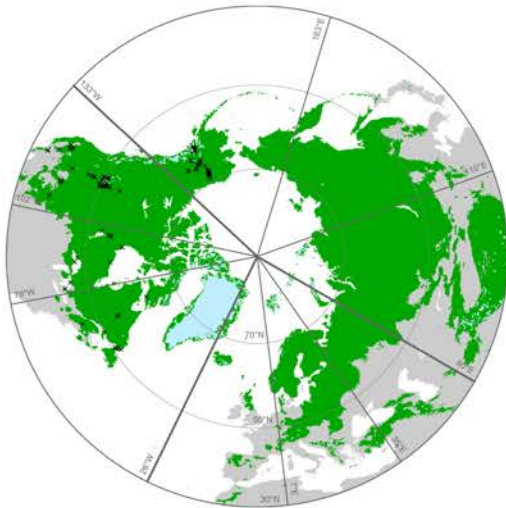
*Vertigo lilljeborgi* | n = 286/220



*Vertigo microsphaera* | n = 51/51



*Vertigo modesta* | n = 310/107



*Vertigo morsei* | n = 61/61






*Vertigo nylanderi* | n = 119/80

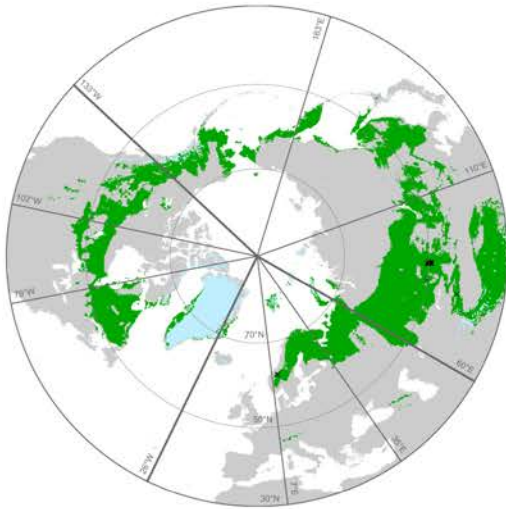
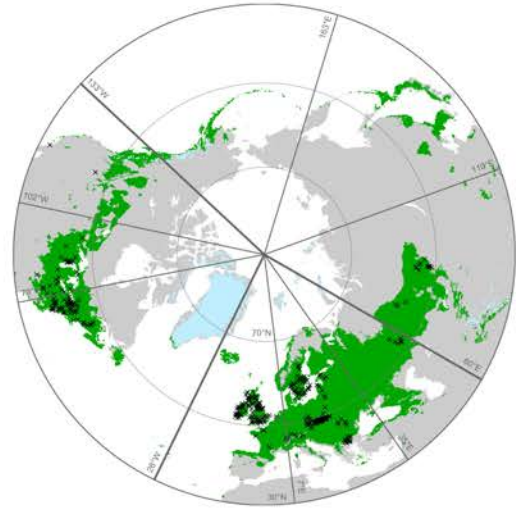
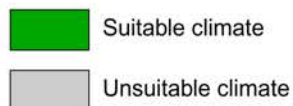


*Vertigo oughtoni* | n = 35/35



 Suitable climate  
 Unsuitable climate



 Records used for model calibration

***Vertigo parcedentata* | n = 21/21*****Vertigo perryi* | n = 47/47*****Vertigo pseudosubstriata* | n = 11/11*****Vertigo pygmaea* | n = 667/337*****Vertigo ronneyensis* | n = 331/181*****Vertigo substriata* | n = 522/295**

x x  
x x Records used for  
model calibration

*Vertigo ultima* | n = 26/26



 Suitable climate  
 Unsuitable climate

*Vertigo ventricosa* | n = 405/163



x x x  
Records used for  
model calibration