Molecular Ecology

WILEY

MOLECULAR ECOLOGY

ORIGINAL ARTICLE

Elevational Range Impacts Connectivity and Predicted Deme Sizes From Models of Habitat Suitability

Connor M. French[1,2] | Roberta P. Damasceno[3] | Mariana M. Vasconcellos[3,4] | Miguel T. Rodrigues[3] | Ana C. Carnaval[1,2] | Michael J. Hickerson[1,2,5]

[1]Biology Department, City College of New York, New York, USA | [2]Biology Ph.D. Program, Graduate Center, City University of New York, New York, USA | [3]Departamento de Zoologia, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil | [4]Department of Biology, Evolutionary Ecology Group, University of Antwerp, Antwerp, Belgium | [5]Division of Invertebrate Zoology, American Museum of Natural History, New York, New York, USA

Correspondence: Connor M. French (french.connor.m@gmail.com)

Received: 18 July 2024 | **Revised:** 4 November 2024 | **Accepted:** 6 November 2024

Handling Editor: Mitchell Cruzan

Funding: C.M.F., A.C.C. and M.J.H. thank NSF DBI 2104147 and NSF DEB 1343578. C.M.F. thanks the Doctoral Student Research Grant from the Graduate Center, City University of New York. M.T.R. thanks Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP # 2003/10335-8 and 2011/50146-6), CNPq, Furnas Centrais Elétricas, Hussam Zaher, Marcio Borges Martins, Marco Freitas, and all members of his lab for help in field. R.P.D. thanks Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP # 2013/22477-3 and 2015/11498-5). M.M.V. is supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP #2019/08308-0 and #2023/16814-9).

Keywords: landscape genetics | lizards | modeling | Niche modelling | phylogeography | population genetics-empirical

ABSTRACT

In integrative distributional, demographic and coalescent (iDDC) modelling, a critical component is the statistical relationship between habitat suitability and local population sizes. This study explores this relationship in two *Enyalius* lizard species from the Brazilian Atlantic Forest: the high-elevation *E. iheringii* and low-elevation *E. catenatus* and how this transformation affects spatiotemporal demographic inference. Most previous iDDC studies assumed a linear relationship, but this study hypothesises that the relationship may be nonlinear, especially for high-elevation species with broader environmental tolerances. We test two key hypotheses: (1) The habitat suitability to population size relationship is nonlinear for *E. iheringii* (high-elevation) and linear for *E. catenatus* (low-elevation); and (2) *E. iheringii* exhibits higher effective migration across populations than *E. catenatus*. Our findings provide clear support for hypothesis (2), but mixed support for hypothesis (1), with strong model support for a nonlinear transformation in the high-elevation *E. iheringii* and some (albeit weak) support for a nonlinear transformation also in *E. catenatus*. The iDDC models allow us to generate landscape-wide maps of predicted genetic diversity for both species, revealing that genetic diversity predictions for the high-elevation *E. iheringii* align with estimated patterns of historical range stability, whereas predictions for low-elevation *E. catenatus* are distinct from range-wide stability predictions. This research highlights the importance of accurately modelling the habitat suitability to population size relationship in iDDC studies, contributing to our understanding of species' demographic responses to environmental changes.

1 | Introduction

How species respond to their environments is a complex, multidimensional phenomenon (Fenderson, Kovach, and Llamas 2020). In particular, the dynamic and complex history underlying changes in species abundance, spatial distribution and connectivity is difficult to detect without detailed models and data from multiple sources (Fenderson, Kovach, and Llamas 2020; Lee-Yaw et al. 2022). Species distribution models (SDMs) can provide one direct way to explore how species'

© 2024 John Wiley & Sons Ltd.

current and historic range dynamics are influenced by the environment (Elith and Leathwick 2009). Another approach complimenting SDMs is geographically explicit population genomic coalescent models that use geo-referenced genetic data to infer processes underlying complex and dynamic spatiotemporal demographic histories (Bradburd and Ralph 2019; Liu et al. 2019; Kamm et al. 2020; Excoffier et al. 2021; Baumdicker et al. 2022).

Integrative distributional, demographic and coalescent modelling (iDDC) provides a powerful framework that combines SDMs and population genomic modelling to understand how past environmental change has shaped species' spatial demographic and evolutionary histories (Brown and Knowles 2012; He, Edwards, and Knowles 2013; Brown et al. 2016; Prates et al. 2016). iDDC models use habitat suitability values extracted from SDMs to inform spatially explicit demographic models of local population sizes and migration. Parameters from these demographic models are then applied to a coalescent model that simulates population genomic data that can be directly compared with the observed data (Alvarado-Serrano and Knowles 2014). In addition to hypothesis testing and parameter inference, these models can generate maps of predicted present-day spatial patterns of genetic diversity across both sampled and unsampled areas within a species' range. This summary statistic of genetic diversity (pi) is a key biodiversity metric for assessing species' responses to environmental change over recent and historical time scales (Exposito-Alonso et al. 2022; French et al. 2023; Theodoridis et al. 2020).

A critical element of iDDC models is the assumed relationship between habitat suitability and local population sizes (also referred to as deme sizes). How this relationship is parameterised can dramatically impact the inferences drawn from iDDC models (Brown and Knowles 2012), yet explicit consideration of this impact is relatively scarce. Typically, iDDC studies (Knowles and Massatti 2017; Pan et al. 2020; Castilla et al. 2024), assume that local deme sizes increase by one unit for each unit increase in habitat suitability, implying a linear relationship. However, this assumption may not hold across several ecological conditions (VanDerWal et al. 2009; Weber et al. 2017). For instance, the relationship could instead follow a wedge-shaped curve, where local deme sizes rise rapidly to a maximum at a certain SDM suitability threshold, or it might take another intermediate nonlinear form (VanDerWal et al. 2009). Furthermore, at the coarse spatial resolution typical of SDM-based inferences, access to diverse microhabitats—often not captured by coarse climatic variables—may lead to a decoupling between SDM-determined suitability and the actual local deme density, resulting in a nonlinear rather than a linear relationship (Lu and Jetz 2023; VanDerWal et al. 2009).

This nonlinear relationship may be particularly relevant in montane habitats, where the elevational gradient can enable species to more readily climb up or down in altitude and hence access suitable environments more easily than those species in lowland habitats (Dobrowski 2011). If the steep topography enables high-elevation species to more easily buffer environmental changes that reduce climate suitability, allowing them to maintain high deme sizes despite those conditions, we expect that iDDC models with a nonlinear relationship between local SDM suitability based on macroclimatic data and local deme sizes to be a better fit for high-elevation species than models assuming a linear relationship. In this case, local deme sizes are expected to reach a high value even under relatively low suitability scores. Conversely, a linear relationship can be expected in low-elevation species, in which local deme sizes are limited by the suitability of the landscape. If this is the case, then the spatial resolution of the SDMs could be aligned more closely with the granularity of habitat turnover in low- rather than in high-elevation species.

In addition, emerging evidence has shown that high-elevation ectotherm species tend to show greater resilience to climate fluctuations, potentially due to broader environmental tolerances relative to lowland animals (Janzen 1967; von May et al. 2017; Strangas et al. 2019; Bovo et al. 2023). This resilience may result from highland species experiencing a wider range of climate variability, driving physiological adaptations or plasticity (Brett 1956; Janzen 1967; Bozinovic, Calosi, and Spicer 2011), or behavioural buffering (Sunday et al. 2014; Muñoz 2022). Consequently, these broader environmental tolerances may confer greater metapopulation connectivity (Qiao et al. 2016; Carscadden et al. 2020). We therefore expect iDDC models of high-elevation species to infer higher connectivity (*Nm*) compared to low-elevation species, as indicated by higher estimated migration rates.

To test if these expectations hold, we develop iDDC models for two lizard species endemic to the Brazilian Atlantic Forest (BAF), which features a pronounced elevational gradient whereby low-elevation species in the north are expected to have experienced greater range stability than those in the south, which in turn experience more climate variability due to higher elevations and a more temperate climate (Carnaval and Moritz 2008). We specifically focus on two *Enyalius* lizard species, *E. iheringii* and *E. catenatus*, whose contrasting elevation ranges yield different predictions regarding how climate change may have influenced their habitat suitability and local density over time (Figure 1). Both species inhabit forest and forest-edge habitats with allopatric distributions (Rodrigues et al. 2014). *Enyalius iheringii* occurs in the southeastern BAF, a region farther from the equator and more environmentally variable at both current and historical time scales (Carnaval et al. 2014). In the northern part of its distribution, this species primarily inhabits regions around 700 m above sea level (asl), reaching lower elevations at higher latitudes (Figure 1). *Enyalius catenatus*, on the other hand, occurs in the northeastern BAF from sea level to around 200 m asl, becoming rare above this elevation (Figure 1) (Rodrigues et al. 2014). Both species are semi-arboreal, diurnal sit-and-wait foragers, and similar in size averaging around 90 mm (Jackson 1978; Liou 2008; Rautenberg and Laps 2010; Rodrigues et al. 2014; Bruscagin et al. 2017). They live up to 6 m above the ground, between the base of tree trunks and shrubs with leaves (Jackson 1978; Rodrigues et al. 2014), descending to the ground only to feed or reproduce.

In this study, we implement a novel iDDC method that integrates supervised machine learning (SML) to (1) test two alternative statistical relationships (linear vs. nonlinear) between local SDM suitability and local deme sizes across space and time, predicting a linear relationship for the low-elevation *E. catenatus* and a nonlinear relationship for the high-elevation

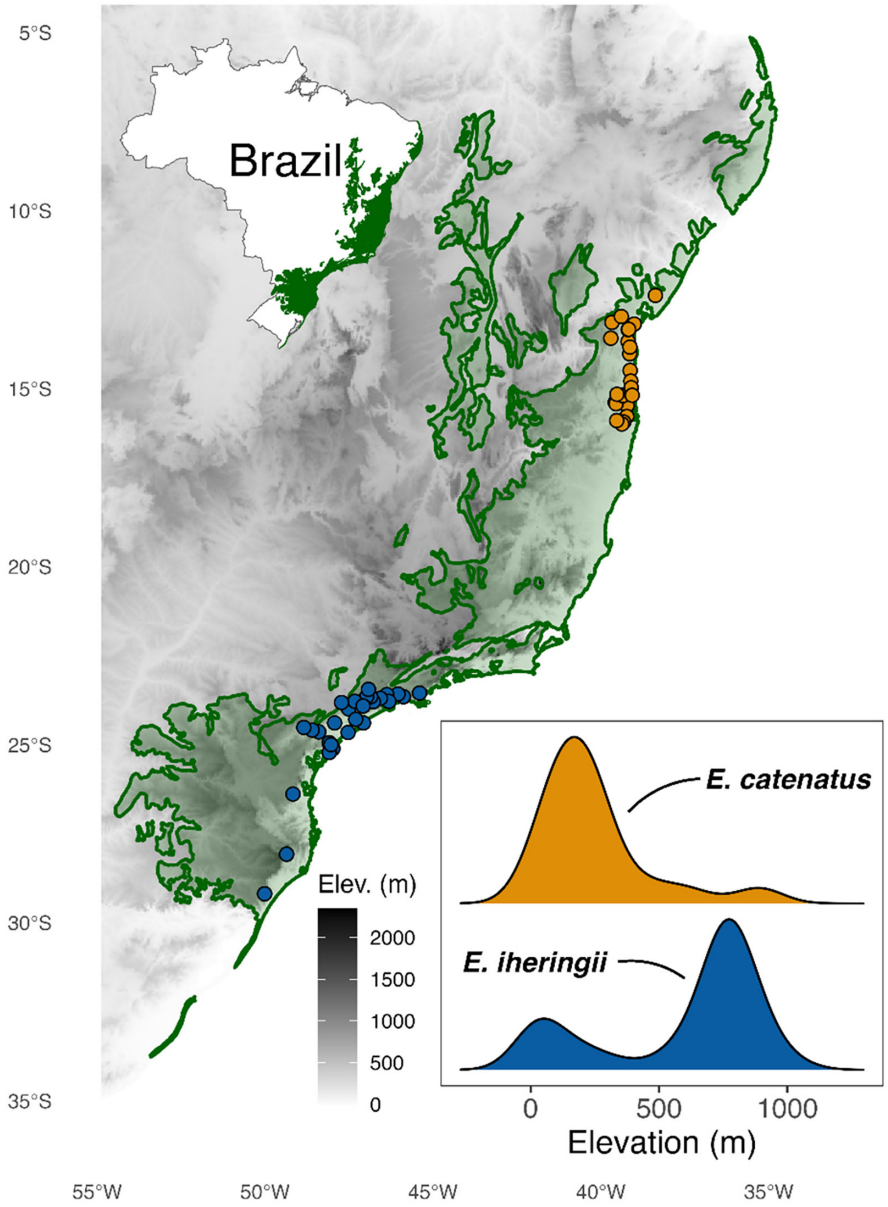

FIGURE 1 | A map of *Enyalius catenatus* (orange) and *Enyalius iheringii* (blue) sampling localities and their elevational distributions. *Enyalius catenatus* is distributed across lower elevations in the northern portion of the BAF, while *E. iheringii* is distributed across higher elevations at the southern portion of the BAF. The northern BAF is closer to the equator and experiences less annual climate fluctuation than the southern BAF.

E. iheringii; (2) evaluate the relationship between elevation and connectivity (by estimating local deme sizes (*N*) and effective migration rates (*Nm*) for each species), expecting lower *Nm* for *E. catenatus* compared to *E. iheringii*; and (3) use our estimates of local *N* and *Nm*, along with the best-fit transformation model, to generate landscape-wide maps of current predicted genetic diversity encompassing unsampled areas within each species' predicted range. For that, we combine SDMs with genomic data in a coalescent simulation framework across a grid-like landscape, where each grid cell represents a local population (deme). By approximating a spatially explicit coalescent model, we are able to efficiently generate simulated genotypes that match the spatial sampling configuration of the observed data (Figure 2). To assess the sensitivity of the assumption of a linear relationship between habitat suitability values and local deme sizes, we compare the effect of linear versus nonlinear transformations of SDM-generated suitabilities on local deme sizes in our iDDC model, using geo-referenced genome-wide SNP data from the two *Enyalius* species differing in elevation and latitude. We then incorporate this uncertainty in transformation probabilities into our iDDC-based analysis producing estimates of historical demographic parameters and predictions of landscape-wide genetic diversities.

2 | Methods

2.1 | Spatially Explicit Coalescent iDDC Model

To test the relative fit of the linear versus threshold transformation of SDM suitabilities into local effective population sizes, estimate demographic parameters under the best

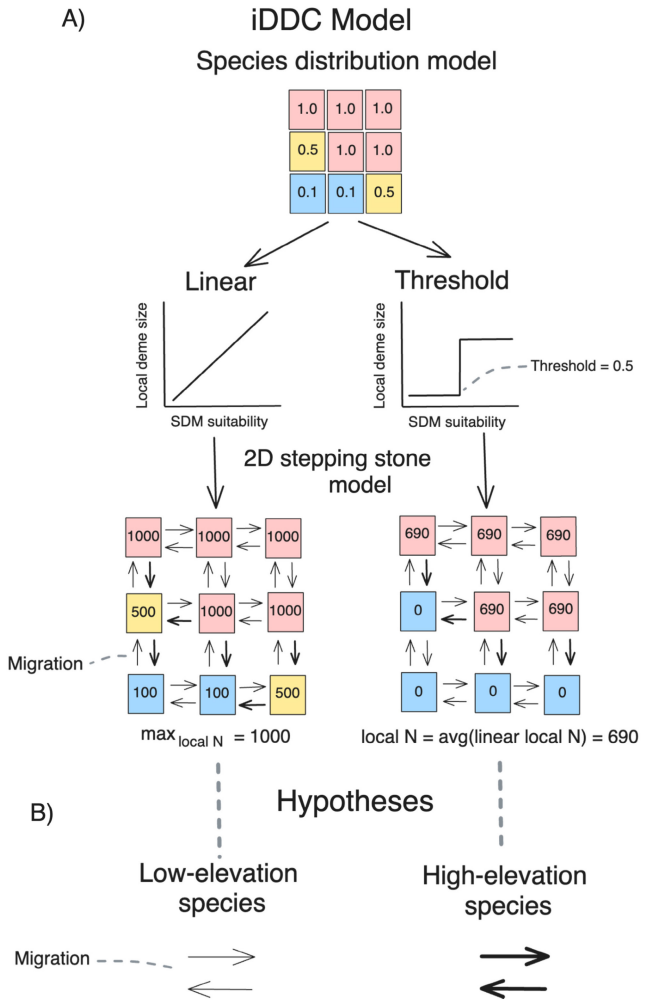

FIGURE 2 | Conceptual diagram of the iDDC model with hypotheses defined by SDM and local deme size transformation functions. (A) The transformation of SDM-derived habitat suitability values (ranging from 0 to 1) into effective population sizes of local demes within the 2D stepping-stone (2DSS) model using a transformation function. Two transformation functions are considered: Linear (left) and thresholded (right). For the linear transformation, each SDM suitability value is multiplied by a maximum local deme size ($max_{local\ N}$) making local deme sizes directly proportional to the SDM suitability. For the thresholded transformation, only SDM cells with suitability values above a specified threshold are occupied, while those below or equal to the threshold remain unoccupied. To ensure comparability of total population size across the landscape between linear and thresholded models, the constant local deme size in a threshold transformation is calculated as the average of all occupied demes in the corresponding linear transformation with the same $max_{local\ N}$. Migration between neighbouring demes is scaled by neighbouring population sizes with the formula $m_{N1-N2}=(N_{N1}/N_{N2}) \times M$, where M is a global migration rate, and N is the local deme size. (B) Predictions for elevation-based transformation hypotheses where a linear transformation is predicted for low-elevation species and a threshold transformation is predicted for high-elevation species. Additionally, higher migration rates (indicated by thicker lines) are expected for high-elevation species relative to low-elevation species. A full conceptual diagram that illustrates demographic change over time and the non-spatial portion of simulations is available in Figure S1.

transformation, and create predicted contemporary genetic diversity maps, we develop a new iDDC model that explicitly integrates SDMs with a coalescent two-dimensional stepping-stone model (2DSS) that simulates spatial patterns of genomic diversity for direct comparison with observed samples. By calculating the same array of summary statistics in the simulated and observed datasets, we use SML to probabilistically discriminate SDM transformation models, estimate demographic history model parameters and generate the genetic diversity maps (Figure 3).

SDMs use rasterised data to create probability maps of species occurrence, which can be used to infer ecological parameters like potential local abundance and resistance to dispersal, following the same precautions outlined in the Introduction (Guisan and Thuiller 2005; VanDerWal et al. 2009; Weber et al. 2017). We translate these ecological parameters into the population genetic parameters used by a 2DSS model (Figure 2). Specifically, the 2DSS model consists of a grid of rectangular demes connected by migration along their edges with effective population sizes scaled by the predicted SDM suitabilities. The grid is arranged in a rectangle that matches the extent of the SDM, where each cell with a predicted suitability in the SDM corresponds to a deme in the 2DSS, and cells without predicted suitability correspond to empty demes. It is a simplification of a continuous geographical space, where demes located nearer to each other are more likely to exchange alleles, generating patterns of isolation-by-distance (Kimura, Weiss, and Weiss 1964). Additionally, deme sizes and migration can change through time, where SDMs projected to environmental conditions in the past are used to parameterise deme size changes and subsequent migration rate changes. This implementation of the 2DSS model simulates local deme sizes and migration backwards-in-time in a fully coalescent framework, similar to Szép, Trubenová, and Csilléry (2022), as opposed to a paired forward-time demography, backward-time coalescent simulation approach like SPLATCHE3 and similar methods that implement the iDDC approach (Currat, Ray, and Excoffier 2004; Currat et al. 2019). Because backwards simulation under a coalescent model is far more efficient than forward time simulation, this iDDC approach directly links SDM suitability values with population genetic parameters in a computationally efficient manner.

To test the hypothesis of the influence of linear versus nonlinear transformations on the inference of local deme sizes, we directly translate SDM suitability values into 2DSS local deme sizes using either a linear transformation function, where the SDM suitability values, which range from zero to one, are multiplied by a maximum local deme size, or a thresholded transformation function. While there are a variety of nonlinear shapes the relationship between SDM suitability values and local deme sizes could take, we opted to consider a thresholded binary distribution, the most extreme nonlinear realisation of this transformation function, which has been shown to influence the predicted genetic diversity of iDDC models (Brown et al. 2016). The threshold transformation uses a minimum SDM suitability value as a threshold, where SDM suitability values above the threshold are considered occupied

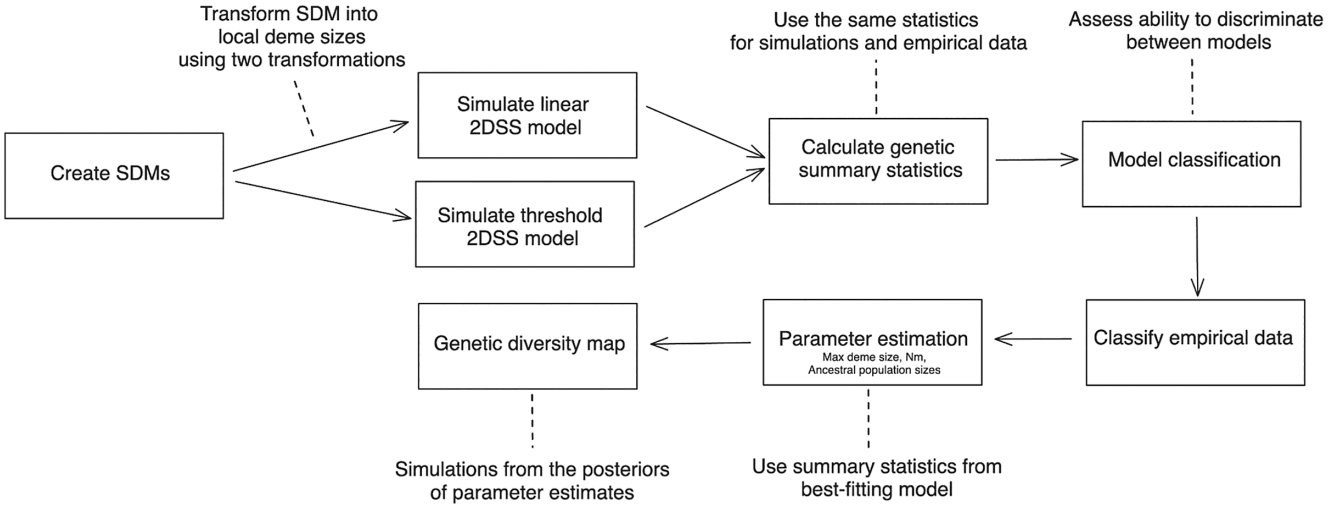

FIGURE 3 | Full analytical workflow for each species, from creating species distribution models (SDMs) to generating genetic diversity maps. SDMs are developed using weather station-derived bioclimatic variables (temperature and precipitation), resulting in habitat suitability ranging from 0 to 1. The SDMs are projected to palaeoclimates back to the last glacial maximum (LGM, 22 kya) in 1000-year time steps. The suitability values for each projection are translated into local deme sizes in a two-dimensional stepping-stone (2DSS) model, where demes are arranged in a grid that matches the dimensions of the SDM and each 2DSS deme corresponds to an SDM suitability pixel. A full description of the model is available in the Methods section 'Spatially explicit coalescent iDDC model'. The model outputs a genotype matrix for individuals matching empirical sampling localities, upon which a suite of genetic summary statistics are calculated. The same set of summary statistics is also calculated for empirical data. After running each model- linear and threshold- for approximately 50,000 replicates and aggregating the summary statistics, a supervised machine learning (SML) classification approach is employed to evaluate the model's ability to discriminate between the two transformation methods. After training the classifier with these simulations, the empirical data are classified into the best-fit transformation for the species. Following classification, demographic parameters are estimated using an SML regression approach with the summary statistics simulated from the best-fit transformation model. In addition to point estimates, 95% prediction intervals are inferred for each demographic parameter. Finally, to generate genetic diversity maps for each species, 100 random samples from the 95% prediction interval for each demographic parameter are used as inputs for 100 additional landscape-wide simulations. Genetic diversity is estimated for each deme in the current landscape containing at least two individuals based on the transformation from the SDM. The mean across these 100 simulations is used to derive an average genetic diversity estimate for each deme.

and SDM suitability values below the threshold are considered empty (Figure 2). The maximum local deme size is assigned to all occupied demes.

2.2 | Species Distribution Modelling

2.2.1 | Occurrence Data

We obtained occurrence data from across the species' ranges through field sampling and museum records, validating them with expert review and field notes (Table S1, Figure 1). Additionally, all localities for *Enyalius catenatus* and nearly all localities ($N-2$) for *E. iheringii* that were used for species distribution modelling have at least one individual who was sequenced for genomic analysis. However, four inland localities were discarded for *E. catenatus* after modelling and before thresholding the model output, due to their presence in very low suitability habitat and isolated relict forest fragments (Table S1). Including marginal habitat in SDMs is known to reduce the ability to detect environmental barriers to the species' occupation, so using them to threshold model output would result in overestimating the species' potential range (Soley-Guardia et al. 2016). These individuals were also not used in genetic simulations. We spatially thinned the occurrence data using a buffer distance of 20 km, which is twice the resolution of the environmental data, to reduce the impact of clustered sampling bias and the resultant artefactual spatial autocorrelation on the models using the spThin R package (Boria et al. 2014; Aiello-Lammens et al. 2015).

2.2.2 | Environmental Data

The environmental data used for modelling were the 19 bioclimatic variables from the CHELSA v2.1 database, which provide descriptions of temperature and precipitation based on statistical interpolation of weather station data (Karger et al. 2017, 2018). These data were downloaded at 30 arc-second (approximately 1 km^2 at the equator) resolution. We aggregated the layers to five arc-minute (approximately 10 km^2 at the equator) resolution using bilinear interpolation before modelling to allow for the use of lower-resolution observation points and to match the resolution we chose for spatial demographic modelling.

For model calibration, we delineated the study area for each species using a minimum convex polygon around the species' localities, with a 0.5° buffer. This approach captures habitat the species can reasonably disperse to without encroaching deep into the Cerrado. This reduces the chance of bias imposed by sampling suitable environments that are unreachable due to the species' dispersal abilities (Anderson and Raza 2010). The area within the polygon was used as the background environment for

SDM model building. All spatial processing was done using the *terra* v1.6 and *sf* v1.0 R packages (Pebesma 2018; Hijmans 2022; Pebesma and Bivand 2023).

2.2.3 | SDM Model Building

We modelled the potential distribution of each species using *Maxent* v3.4.4 (Phillips, Anderson, and Schapire 2006), a supervised machine learning technique and software package designed for using presence-background data, implemented in the *dismo* v1.3 R package (Hijmans et al. 2022). The maximum number of available background points were selected for each species' study area (considering a single point per pixel): 2355 points for *E. catenatus* and 6515 points for *E. iheringii*. To reduce the impact of multicollinearity on model interpretability and reduce the potential of projecting to non-analogue environments in historical time periods (Fitzpatrick and Hargrove 2009), we filtered predictor variables for each species and retained only those with a variance inflation factor less than 10 using the *usdm* v2.1 R package (Dormann et al. 2013; Naimi et al. 2014).

We considered a suite of candidate models under varying complexity, preferring simple over complex models to prevent model overfitting. In *Maxent*, the linear and quadratic feature classes, along with regularisation multipliers from 0.5 to 5.0 in 0.5 increments, were considered. Models were constructed using the ENMeval v2.0.4 R package (Kass et al. 2021). To further prevent model extrapolation, we clamped predictors to the range of the training data.

2.2.4 | SDM Model Evaluation

Due to the low number of occurrence records after spatial thinning, we used leave-one-out cross validation for model selection to maximise the amount of information available (Shcheglovitova and Anderson 2013). We used two criteria for model selection, the Akaike Information Criterion corrected for low sample sizes (AICc) and the minimum training presence (MTP) omission error rate, which is the lowest occurrence suitability value used to train the model. We first selected models with the lowest AICc values, then for all models within <2 AICc of each other, selected the model with the lowest MTP omission error rate. If there were multiple models with the same MTP omission error rate, we then considered the model with the highest average area under the curve (AUC) score performance on validation data. This approach for model selection prioritises simple models that reliably predict training data and are able to distinguish occurrences from background environment (Radosavljevic and Anderson 2014).

2.2.5 | Projecting Models to Historical Time Periods

To estimate distribution pattern shifts since the last glacial maximum (LGM, approximately 21 kya), we projected each species' SDM to environmental reconstructions in 22 sequential 1000-year timesteps from the present going back to the LGM, using Maxent's cloglog output. We obtained 30 arc-second resolution historical bioclimate variables from the CHELSA-TraCE21k dataset and aggregated them to five arc-minutes (Karger et al. 2023). To represent areas the lizards may reasonably disperse to and between, we limited the study area for projections to a 2.0° buffer around a minimum convex hull polygon of the observed localities. To facilitate visual evaluation of range stability through time, we summarised the change in per-pixel suitability through time as stability maps, where MTP thresholded projections were summed across time. Pixels were considered more stable if they had a higher number of timesteps with predicted presence, and less stable if they had a lower number.

2.3 | Genomic Data

Samples for sequencing the two *Enyalius* species were obtained from 89 individuals collected from 66 sampling locations across the Brazilian Atlantic Forest, along with three outgroup individuals (Figure 1 and Table S1). These samples were obtained under ICMBIO permits #10126, #30309, and #10754. Genomic DNA was isolated from liver or muscle tissues preserved in 100% ethanol using the Qiagen DNeasy blood and tissues kit extraction method. DNA quality was assessed by visualising high molecular weight bands on a 1% agarose gel, and DNA concentrations quantified using a Qubit 2.0 fluorometer (Life Technologies). All DNA extractions were normalised at a concentration of 10 ng/μL in a total volume of 50 μL for reduced-representation genomic library preparation.

A RAD-seq library (Baird et al. 2008) was prepared by Floragenex Inc. (http://floragenex.com) generating restriction-site associated DNA of thousands of short fragments distributed across the entire genome. Each individual's genomic DNA was digested with the enzyme SbfI (5′ CCTGCAGG 3′), then ligated to customised Illumina adapters containing 10-bp individual-specific barcodes and the enzyme overhang (6 bp). After ligation, all samples were pooled into a single sequencing library followed by sonication and end repair. Fragments ranging from 300 to 500 bp in size were selected. The fragments were then PCR-amplified, and finally sequenced across multiple sequencing lanes of the Hiseq2500 System Illumina sequencer, generating single-end reads of 100 bp.

2.3.1 | Read Processing and Assembly

Raw reads were demultiplexed to individuals based on unique barcode sequences. Demultiplexed sequencing reads were then assembled across species with three outgroups using *ipyrad* v0.7.22 (Eaton and Overcast 2020). A *de novo* reference was created, since no reference genome was available. Prior to assembly, adapters were trimmed from reads based on quality scores, and low-quality reads were filtered using a PhredQ score cutoff of less than 33, with a maximum of five low- quality bases per read. We determined the optimal clustering threshold following McCartney-Melstad, Gidiş, and Shaffer (2019) (Methods S1 and Figures S6–S9). A minimum of four samples per locus were retained, with further missingness thresholds explored on a per-species basis. All remaining parameters were set to default values for the *ipyrad* assembly pipeline. Following assembly, we assessed the presence of batch effects, visualising potential differences in expected heterozygosity, missing data biases, and

genetic structure among sequencing runs and found no presence of a significant effect (Methods S1 and Figures S10–S12). A single outlier individual was removed prior to applying SNP filters for each species due to a combination of high levels of allele sharing across species and a high proportion of missing data (Table S1).

2.3.2 | SNP Calling

We performed additional filtering on single nucleotide polymorphisms (SNPs) after quality filtering the reads, using vcftools v0.1.16 (Danecek et al. 2011) and scikit-allel v1.3.7 (Miles et al. 2023). Only biallelic sites were retained. We assessed the impact of missing data using four missingness thresholds: 50%, 40%, 30% and 20% of sites missing across individuals per species. We additionally filtered for unlinked SNPs using a linkage disequilibrium (r^2) threshold of 0.10. Finally, we removed invariant sites and singletons, which are likely overrepresented due to sequencing errors common in RADseq data.

2.3.3 | Population Structure

We determined the number of non-spatial ancestral populations of each *Enyalius* species using the program sNMF (Frichot et al. 2014). The sNMF algorithm uses sparse non-negative matrix factorisation to estimate ancestry coefficients without relying on population genetic model assumptions and is optimised for large genomic datasets (Frichot et al. 2014). We ran the algorithm with empirical samples and explored 1–5 ancestral populations (K), with 20 replicates per K value. Multiple alpha regularisation parameter values (1, 10, 100) were explored to assess the robustness of results to differing levels of model complexity. We selected the best-fit K for each species according to the cross-entropy criterion.

2.4 | 2D Stepping-Stone Coalescent Model

We used a fully coalescent implementation of the 2DSS, described above. Parameterised 2DSS models were simulated using the *msprime* coalescent simulator (v1.3.0, Baumdicker et al. 2022).

We discretised the landscape of predicted suitabilities derived from each species' SDM into a matrix of demes connected by migration (Figure 2A). Each deme has a local deme size, where N is the number of diploid individuals (i.e., the local effective population size), connected by continuous migration rates (m), which is the fraction of the recipient deme replaced by the donor deme. At discrete timesteps in the past, N and m are updated to reflect population size changes. The discrete timesteps correspond with SDM past projections (see *Projecting models to historical time periods*). Given the coalescent framework, population size changes and migration rate changes are instantaneous (Figure S1B).

For the demographic history before the LGM, the demes then collapse into two to three non-spatial ancestral populations (Figure S1C) whose number and sample assignment correspond to the number of ancestry clusters (K) determined from sNMF (Frichot et al. 2014). This non-spatial approximation of the coalescent process of the remaining uncoalesced sample gene lineages allows for simulating the part of the data that is affected by older time periods for which SDMs are not available and are expected to be well approximated by the collecting phase of the coalescent under a variety of meta-population models in which the number of demes scales with a non-spatial panmictic effective population size parameter (Wakeley 2000, 2004).

The simulated demes corresponding to empirical sampling localities are sampled at the current time step identically to the observed samples with regards to numbers of individuals. The output from a simulation is a tree sequence representing the genealogical history of all sampled individuals under the standard coalescent with recombination (Baumdicker et al. 2022), with mutations overlaid onto this tree sequence according to a Poisson distribution with a mean following a mutation rate μ (mutation/base pair/generation). Each tree sequence is then converted to a genotype matrix whereby a fraction of the matrix is masked to correspond with missing data in the empirical genotype matrix, and genotypes were filtered according to linkage disequilibrium similar to the filter applied to empirical data ($r < 0.1$). Finally, the array of chosen summary statistics are calculated from this simulated and filtered genotype matrix.

2.4.1 | Alternative Models of SDM to Deme Size Transformation

The two competing models of SDM suitability to 2DSS deme size transformation were the linear model (where the SDM suitability is multiplied by a maximum deme size to give each deme their N), and the threshold model (where an SDM suitability threshold determines whether a deme is occupied or not and the deme size is set at a fixed value for N). We determined the threshold as the minimum SDM suitability value that an empirical locality is found when projecting the SDM across the present-day landscape. To ensure that the total N across the total landscape (i.e., the sum across all demes) was comparable between linear and threshold models, the constant local deme size for threshold models were set as the average of all non-zero deme sizes of corresponding linear models with a set maximum deme size. Under both models, migration is scaled by the donor and recipient deme size, where m between a donor and recipient is $(N_{\text{donor}}/N_{\text{recipient}}) \times m_{\text{global}}$, where m_{global} is a globally assigned migration rate. This ensures that a proportional number of individuals are exchanged between demes relative to their deme sizes. All empty demes were given a very small deme size (1e-10) to comply with *msprime*. Any deme that was preceded by an empty deme migrated to its neighbours backwards-in-time in proportion to the neighbour's sizes to mimic short-distance colonisation.

A 1-year generation time was assumed for all *Enyalius* species, comparable with ecologically similar *Anolis* species (Jezkova, Leal, and Rodríguez-Robles 2009; Prates et al. 2016). Given that SDMs were projected to 1000-year intervals into the past, local deme sizes and continuous migration rates were updated every 1000 generations back to the LGM. All demes collapsed into K

ancestral populations at 22,000 generations in the past according to results from sNMF. Ancestral populations merged into a single population 1,000,000 generations into the past, a parameter whose choice had minimal impact on genetic summary statistics in exploratory simulations.

2.4.2 | 2DSS Model Parameters

Parameter space was explored through a grid search procedure, with 10 values selected for each parameter. Maximum local deme size was evenly spaced from 100 to 5000, *Nm* was geometrically spaced from 0.1 to 5, and the ancestral deme size priors varied among species, with a full range from 5000 to 3,000,000 (Table S3). To account for stochastic coalescent and mutational variation, each parameter combination was simulated five times. For both species, a mutation rate of 1e-9 mutations per base pair per generation and a recombination rate of 1e-9 were assumed (Perry et al. 2018).

After obtaining the better fit of the two alternate models, we use the best-fit model to estimate three classes of parameters by way of simulation-based SML techniques: (1) the maximum local deme size; (2) the migration rate (*Nm*, expressed as the number of individuals); and (3) the ancestral deme sizes for each ancestral population, where all species had K ancestral populations. The same set of simulations was used for model fitting and parameter estimation.

2.4.3 | Genetic Summary Statistics

From the observed and simulated sample of each of the two species, we calculated the same suite of seven spatial and nonspatial summary statistics. All statistics were calculated from each set of samples that correspond to each of the K ancestral populations according to sNMF results. The full array of statistics includes the first two Hill numbers (q = 1 and q = 2) of the site-frequency spectrum (SFS) corrected for the number of allele frequency classes (to remove the potential bias introduced by sampling effort). The q = 1 Hill number is a measure of allelic diversity and corresponds with Shannon's entropy of allele frequencies, while the q = 2 Hill number corresponds with heterozygosity (Sherwin et al. 2017; Gaggiotti et al. 2018; Overcast et al. 2021). Additional nonspatial summary statistics include nucleotide diversity (π) and Tajima's D. Two spatial summary statistics included the slope and R^2 of an ordinary least squares regression between pairwise geographic distances and genetic distances among sampled demes, where geographic distances are the great-circle distances among the centroids of the sampled demes and genetic distances are the average number of nucleotide differences between two sets of sampled individuals (d_{xy}). Prior to the regression, the geographic distances were $\log_{10}$ transformed and the genetic distances were corrected with the function $(d_{xy}/(1-d_{xy})$ according to Rousset (1997)). Finally, we included raw pairwise d_{xy} values calculated between sets of sampled demes. Given the varying number of K ancestral populations determined by sNMF, the total number of summary statistics varies per species. Summary statistics were calculated using the Python packages *scikit-allel* v1.3.7 (Miles et al. 2023), *numpy* v1.26.4 (Harris et al. 2020) and *scikit-learn* 1.3.0 (Pedregosa et al. 2011).

2.4.4 | Model Selection With SML

We used gradient boosting to classify the two competing models of SDM suitability to 2DSS deme size transformation, implemented in the Python packages *xgboost* v2.0.4 (Chen and Guestrin 2016) and *scikit-learn* v1.3.7. Gradient boosting is a SML approach that builds an ensemble of simple decision tree models that in composite form a strong predictive model. We centred and scaled all predictor data prior to training the algorithm. Simulated data, comprising summary statistics calculated from simulated genetic data along with the associated model indices and parameter values, were split into 80% for training and 20% for testing. The model was tuned using five-fold cross-validation paired with Bayesian optimisation, an efficient algorithm for selecting an optimal set of hyperparameters (Table S4) (Wu et al. 2019). Model performance was assessed by the area under the receiver operating curve (AUC) and prediction accuracy. The best performing model was chosen according to its AUC and variable importance was determined using the total gain metric, which measures the reduction in training error for each decision tree node the variable is present in, summed across nodes. The best performing model was then fit to the full training set and predicted to the withheld test data to assess the generalisation capabilities of the model. The model is not overfit if the model performance is similar on the test data compared to the training data. Finally, the SDM suitability to 2DSS deme size transformation was estimated from the empirical data using the best performing predictive model.

Additionally, we assessed the goodness of fit of the models by comparing the empirical summary statistics with those from the simulation output of both transformation models using a principal component analysis (PCA). To this end, we projected the empirical data into the PC space of the simulated dataset replicates, visualising them along the first two PC axes.

2.4.5 | Parameter Estimation With SML

We estimated the maximum local deme size, effective migration rate (*Nm*), and ancestral population sizes for each species. The summary statistics from the best-fit simulation model (linear or threshold) were used as the training data. We used a multi-output gradient boosted regression approach, which functions identically to the gradient boosting approach used for model classification, except outputs are transformed for continuous, rather than categorical data. The multi-output functionality also allows for efficient estimation of the parameters given that each parameter is dependent on the same set of training data. For each model, we centred and scaled all predictor data prior to training the algorithm. Simulated data was split into 80% training and 20% testing. The model was tuned over a suite of seven hyperparameters (Table S4) using five-fold cross-validation paired with Bayesian optimisation (Wu et al. 2019). Model performance was assessed using the mean absolute error (MAE) and variable importance was determined using the total gain metric. The best performing model was chosen according to its MAE. The best performing model was then fit to the full training set and predicted to the withheld test data to assess the generalisation capabilities of the model. The model is not overfit if the model performance

is similar on the test data compared to the training data. Finally, the empirical data were predicted using the best performing model. We used the CV+ approach to estimate prediction error (95% prediction interval), which uses the error from five-fold cross-validation to calibrate prediction intervals (Barber et al. 2021). All machine learning analyses were performed with Python v3.10.9.

2.4.6 | Maps of Predicted Genetic Diversity

One hundred additional simulations were conducted to create landscape-wide maps of predicted genetic diversity for each species using the parameter estimates from the best model given the observed data for each species. We sampled from the 95% prediction intervals of each estimated parameter to generate an additional 100 simulations for each species. For each simulation, we sampled two individuals from demes with $N \geq 2$ and calculated per-deme π using *tskit* v0.5.6. To efficiently estimate diversity across the landscape, we calculated π directly on tree sequences, returning π values that diverge from those calculated as simulation summary statistics, but follow the same spatial patterns (Baumdicker et al. 2022). To create the maps, we took the average π across simulations for each deme.

3 | Results

3.1 | Sequencing and SNP Calling

An average of 3,311,792 reads were sequenced per individual, with an average of 3,306,666 reads retained after quality control filtering. A clustering threshold of 0.92 was determined to be the most appropriate for this data set (Methods S1 and Figures S6–S9). An average of 62,541 high-depth clusters per sample were recovered with an average read coverage of 42.37. After filtering, 14,916 SNPs with 40% missing data were ascertained for *E. catenatus* and 12,722 unlinked SNPs with 40% missing data for *E. iheringii*.

3.2 | Species Distribution Modelling

The species distribution models (SDMs) chosen for each species were of simple to moderate complexity, with moderate performance (Figure 4). The optimal models for *E. iheringii* and *E. catenatus* had the linear feature class and 1.5 and 0.5 regularisation multipliers, respectively. Three predictors were used for each species (Figure S2). The test AUC for each species was 0.768 for *E. iheringii* and 0.813 for *E. catenatus*. Highest suitability was along the Brazilian coast for *E. catenatus* and *E. iheringii*.

3.3 | Spatial Coalescent Simulations

The 2DSS matrices used for simulations had dimensions of 92 × 65 demes (2824 occupied at present day) for *E. catenatus* and 118 × 110 demes (4900 occupied) for *E. iheringii*. A total of 88,000 simulations were completed for *E. catenatus* and 100,000 simulations were completed for *E. iheringii*. Descriptive statistics of the simulations indicate comparable simulated demographic histories given the linear and threshold transformations (Table S2).

3.4 | Population Structure

The sNMF analysis revealed different levels of population structure for the two species (Figure 4 and Figure S4). The best-fit number of populations (K), assessed by minimising the cross-entropy score, was three for *E. catenatus* and two for *E. iheringii*. The population splits followed a latitudinal gradient for both species (Figure 4). These population assignments were used as ancestral populations for 2DSS coalescent simulations.

3.5 | Genetic Summary Statistics

Given the different number of sampled localities and ancestral populations for each species, the number of genetic summary statistics varied per species. *Enyalius catenatus* had 102 summary statistics and *E. iheringii* had 279, the majority being raw pairwise d_{xy} values. Raw summary statistics for empirical samples and raw summary statistics for simulation output are available on https://github.com/connor-french/enyalius_project.

3.6 | Model Classification

PCAs indicate that the empirical data falls within the range of the simulation data (Figure 5). Classification of the linear and threshold iDDC models was more confident for the higher elevation *E. iheringii* compared to lower elevation *E. catenatus* (Figure 6). The best model, chosen based on the highest AUC score, predicted a threshold transformation for *E. iheringii* with a 99% probability and a threshold transformation for *E. catenatus* with a 60% probability (Figure 7).

The best model chosen through five-fold cross-validation for *E. catenatus* had an AUC of 0.74 on withheld test data and an average prediction accuracy of 67.09%, while *E. iheringii* had an AUC of 0.86 and an accuracy of 77.45% (Figure 6). A mix of population-wide summary statistics (e.g., π, Tajima's D, and Hill numbers of the site-frequency spectrum) and between-deme d_{xy} values were present in the top ten most important variable scores, with a majority of d_{xy} statistics (Figure S5).

3.7 | Parameter Estimation

According to performance on withheld test data, parameter estimation accuracy (represented by observed versus estimated plots) was higher in the higher elevation *E. iheringii* compared to lower elevation *E. catenatus* for all parameters (Figure 6 and Figure S3). Mean absolute error (MAE) for max deme size was 520.31 for *E. iheringii* and 880.47 for *E. catenatus*. MAE for *Nm* was 0.33 for *E. iheringii* and 0.57 for *E. catenatus*. Predictions were also less biased for *E. iheringii*, with lower over- or under-prediction made at the ends of the distributions of observed values. Maximum deme size was estimated to be higher in *E. catenatus* (point estimate = 3083, 95% prediction interval = [1162, 4970]) versus *E. iheringii* (1588, [275, 2987]), but

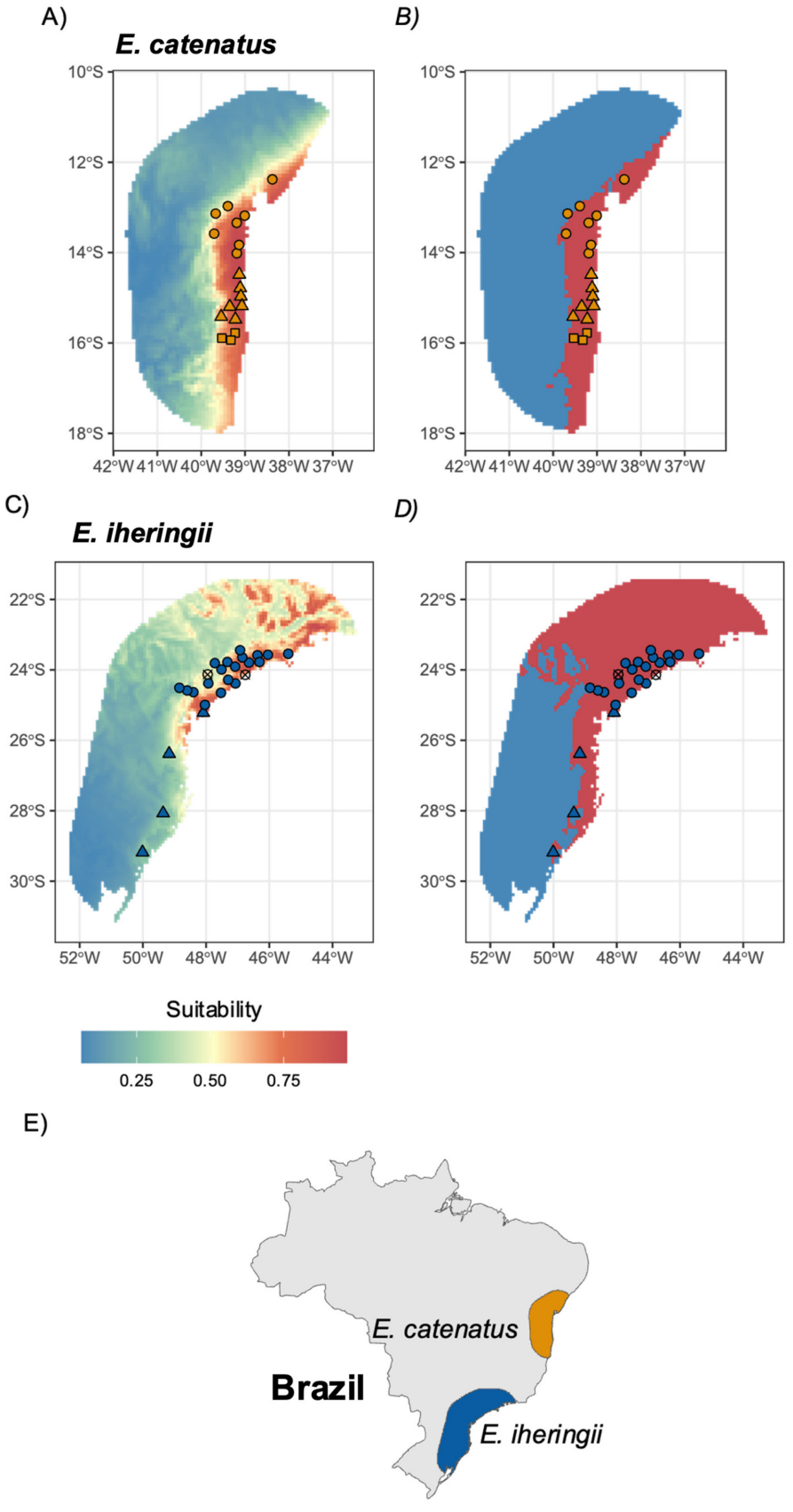

FIGURE 4 | Legend on next page.

FIGURE 4 | Raw (A, C) and thresholded (B, D) projections of SDMs for *Enyalius catenatus* (A, B) and *Enyalius iheringii* (C, D) under current climates. The minimum training presence threshold was applied to the thresholded projections, aligning with the threshold used in iDDC modelling. In the raw projections, cooler colours indicate lower suitability, while warmer colours indicate higher suitability. In the thresholded models, blue indicates non-suitable areas and red indicates suitable areas. Shapes in (A–D) correspond to sNMF admixture population assignments, where the best *K* for *E. catenatus* (A, B) was three and the best *K* for *E. iheringii* (C, D) was two. Circles, triangles, and squares represent assigned populations, while open, crossed circles indicate localities used for SDM modelling without genetic samples. Further details on sNMF modelling are available in the Methods and Figure S4.

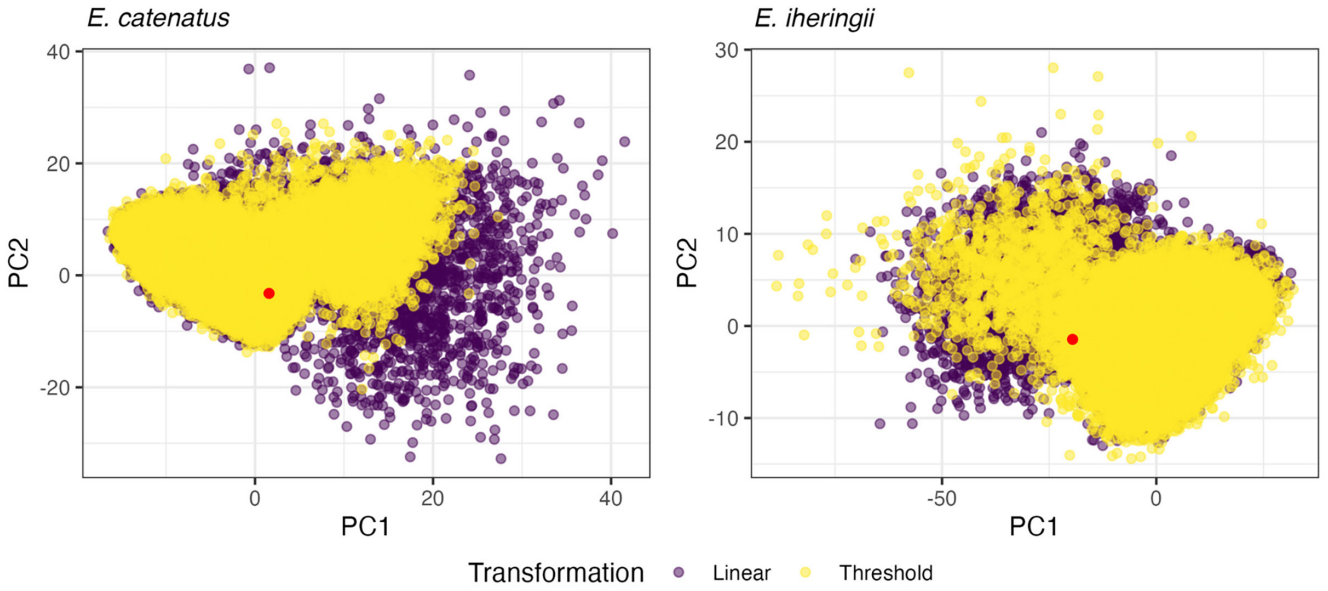

FIGURE 5 | The first two principal components (PCs) of a principal component analysis (PCA) conducted on the full summary statistic simulation output for *Enyalius catenatus* and *Enyalius iheringii*. Yellow and purple dots represent individual simulations, while red dots indicate empirical data projected into the simulation PC space.

migration (*Nm*) was inferred to be lower in *E. catenatus* (0.59, [−0.84, 2.04]) than *E. iheringii* (1.47, [0.30, 2.68]) (Figure 7). Ancestral population size estimates varied widely and are reported in Figure S3. A mix of population-wide summary statistics and between-deme d_{xy} values were present in the top 10 most important variable scores, with a majority of d_{xy} statistics (Figure S5).

3.8 | Genetic Diversity Maps

Genetic diversity is predicted to be higher at the core of *E. catenatus*'s range compared to the edges, while genetic diversity is predicted to be high through most of *E. iheringii*'s range except towards the southern portion (Figure 8A,B). Genetic diversity roughly corresponds with range stability since the last glacial maximum for *E. iheringii*, but diverges for *E. catenatus* (Figure 8C,D). Genetic diversity is predicted to be higher in the centre of *E. catenatus*'s range and lower towards the edges of the range, while stability is predicted through the extent of their range. For *E. iheringii*, high genetic diversity is predicted in areas surrounded by more stable environments, while low genetic diversity is predicted in areas surrounded by less stable environments. Additionally, the total number of individuals across the landscape are expected to have fluctuated more widely in *E. catenatus* compared with *E. iheringii*, with an overall decrease in size for *E. catenatus* and no change for *E. iheringii* (Figure 8E,F). This value is different from the effective population size (*Ne*). It is calculated as the sum of the local deme sizes across the entire simulated landscape, without taking into account the impact of migration on *Ne*.

4 | Discussion

4.1 | Using iDDC to Test the Fit of Alternative Transformation Functions in Montane and Lowland Enyalius

Comparing the fit of alternative iDCC models with different SDM suitability-deme size transformation functions provided mixed support for our hypothesised relationship between the transformation function and elevational range. The iDDC model with a thresholded transformation function had a higher probability given the data from both *Enyalius* species, yet the classification probability and AUC score were higher for the high elevation *E. iheringii*. This suggests that while the nonlinear, that is threshold, transformations are the best fit for both species, the lower-elevation *E. catenatus* may achieve maximum densities in areas where they are predicted to be present, even in regions where the SDM predicts intermediate levels of habitat suitability. These findings agree with previous studies on the relationship between SDM suitabilities and ecological abundance, which suggest that a wedge-shaped,

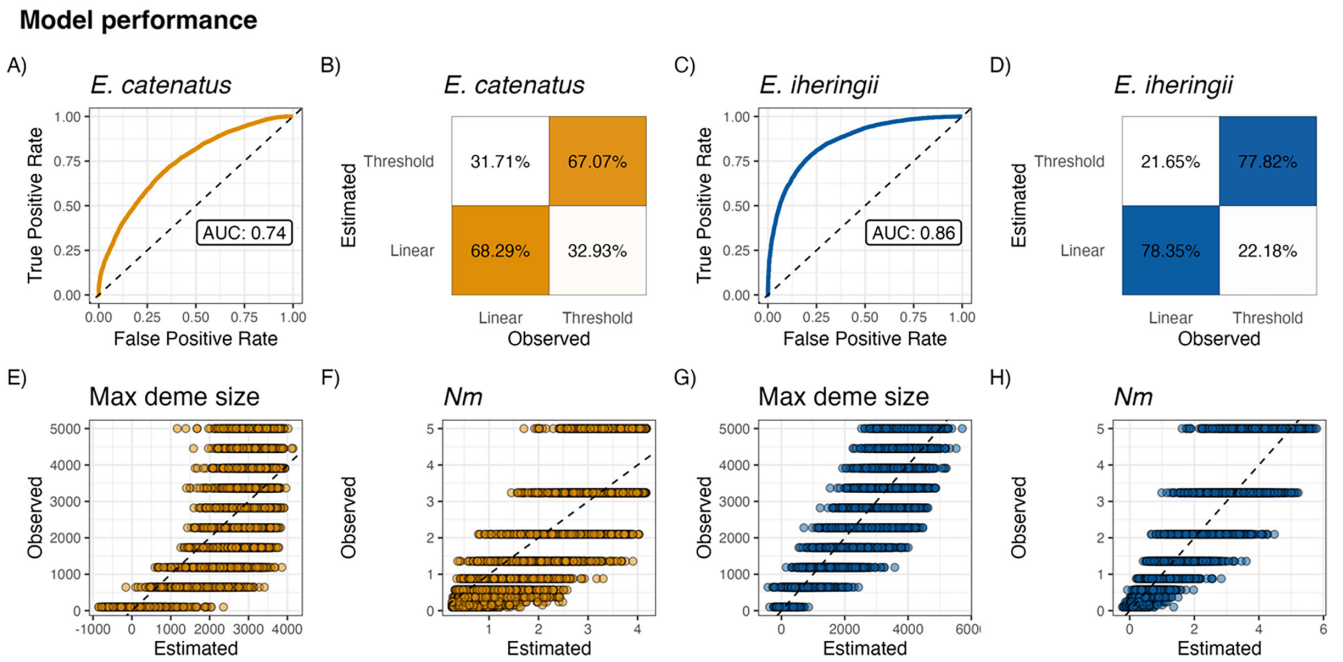

FIGURE 6 | Model performance plots, assessed on withheld test data, coloured by species (*Enyalius catenatus* in orange, *Enyalius iheringii* in blue). (A) and (C) are receiver operator curves displaying the true positive rate relative to the false positive rate for classification predictions. (B) and (D) confusion matrices showing the average prediction accuracy, where coloured squares indicate correct classifications and blank squares indicate incorrect classifications. (E–H) observed versus estimated plots for the two parameters of interest: Maximum deme size and *Nm*. Estimated values are derived from the best-fit SML model, while observed values are those used to generate simulations. The dashed line indicates a perfect relationship between observed and estimated values.

rather than linear, relationship best explains the connections between SDM suitabilities and empirical abundance estimates (VanDerWal et al. 2009; Weber et al. 2017). The fact that thermal physiological adaptation and plasticity are more prominent than previously thought (Muñoz et al. 2014; Strangas et al. 2019; Bodensteiner et al. 2021) may also be contributing to the maintenance of higher population densities in areas predicted to be poor habitats by SDM. Since adaptation to novel thermal environments may be more pronounced in high-elevation species compared to low-elevation counterparts (von May et al. 2017), *E. iheringii* could indeed exhibit wider environmental tolerances, allowing them to achieve high densities even in habitats classified as sub-optimal. Yet, beyond the influence of macroclimate, both biotic interactions (Wisz et al. 2013) and microclimate variation (Stark and Fridley 2022) may also mediate the relationship between SDM suitabilities and local deme sizes.

A fundamental factor contributing to the probabilistic ambiguity in discriminating between the two transformations in *E. catenatus* may be the nearly bimodal distribution of the underlying SDM suitabilities prior to the linear transformation. This resulted in a demographic model closely resembling the binary, thresholded transformation model (Figure 4). Our inability to distinguish a linear vs. thresholded transformation function in *E. catenatus* could also be due to modelling choices. For instance, removing inland *E. catenatus* localities likely raised the SDM values used as a threshold by limiting the marginal habitats considered in determining the minimum training presence threshold. This may have contributed to only highly suitable habitats being considered occupied, excluding potentially habitable areas from the model. Likewise, the presence of uniformly suitable habitat, regardless of the SDM transformation applied, may have caused local deme size to be inferred from sampling density, thereby decoupling the relationship between local deme size and genetic diversity. In general, larger local deme sizes correlate with larger neighbourhood sizes and consequently higher genetic diversity (Nunney 2016). An alternative approach could involve considering various thresholds, which would result in different levels of habitat patchiness and total landscape population size potentially allowing for better discrimination between linear and threshold models (Brown and Knowles 2012). When the goal is to identify the best-fitting relationship between SDM suitabilities and local deme sizes, a more flexible sigmoid transformation may be warranted. However, this approach comes with a trade-off of requiring additional parameters to be inferred, increasing model complexity and demanding additional computational resources, which may not be necessary depending on the study's objectives.

4.2 | Testing Expectations of Migration Rates in High- and Low-Elevation Enyalius

Our finding that estimates of effective migration are higher in the high-elevation *E. iheringii* aligns with Janzen's mountain pass hypothesis, which predicts increased migration through lowland mountain passes for species exposed to greater seasonality, such as high-elevation species or those from more temperate regions. Although Janzen's original predictions compared temperate versus tropical mountains (Janzen 1967), *E. iheringii* experiences greater seasonality due to its occurrence in higher

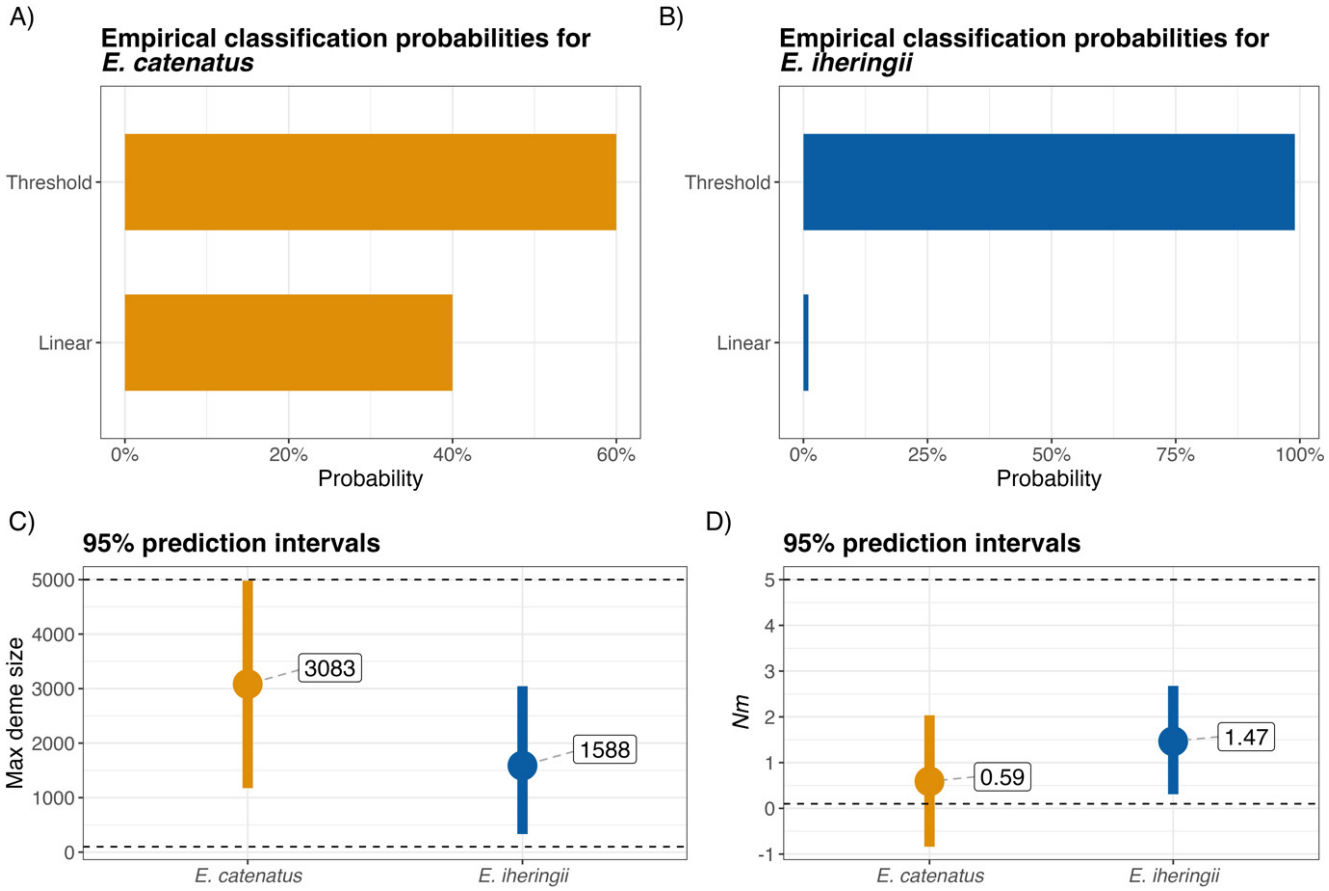

FIGURE 7 | Classification probabilities for empirical data (A, B) and 95% prediction intervals for empirical estimates of maximum deme size (C) and effective migration (*Nm*) (D). In the prediction interval plots (C, D), point estimates are displayed with 95% prediction intervals. Horizontal dashed lines indicate the full range of values used as inputs in the simulations.

elevations and its occurrence farther away from the equator compared to *E. catenatus*. While greater habitat heterogeneity across the range of *E. iheringii* might imply increased genetic structure (as suggested by the isolation-by-environment hypothesis, Wang and Summers 2010), our estimates of higher effective migration indicate that adaptation to substantial temporal environmental fluctuation may buffer the effects of spatial environmental variation. Conversely, the lower effective migration observed in *E. catenatus* also highlights a potential vulnerability to future climate change. Reduced connectivity can lead to range gaps, as the species struggles to track temperature isotherms upslope due to decreased dispersal ability (Colwell et al. 2008).

The higher estimates of maximum deme size in *E. catenatus* suggests that this low-elevation taxon may exist at higher local densities than *E. iheringii* populations. Although published natural history information on the two species is scarce, field collections and observations indicate that *E. catenatus* occurs at higher local densities than *E. iheringii* (Rodrigues, personal obs.). Other aspects of their natural history are similar, including comparable clutch sizes (*E. catenatus* = 5–17, *E. iheringii* = 4–17) and body sizes (mean snout-vent-length, *E. catenatus* = 92.9 mm, *E. iheringii* = 92.0 mm) (Rand 1982; Liou 2008; Rautenberg and Laps 2010; Carilo Filho et al. 2017). Home range size information is only available for *E. iheringii*, which is estimated from two individuals (88.05 and 154.44 m^2) (Liou 2008). One potential explanation for the higher effective deme size estimates in *E. catenatus* could be a greater degree of geographic genetic structure, which may contribute to higher levels of genetic diversity within the 10 km^2 grid cells used in this study (Laporte and Charlesworth 2002). This stronger genetic structure is consistent with the lower migration rate estimates observed between demes, likely due to limited dispersal (Prosser, Gibbs, and Weatherhead 1999; Walters, Trujillo, and Berg 2022). However, other forces that typically impact genetic structure acting at small spatial scales, like biotic interactions, social behaviour, and environmental variation may also be at play (Costa 1998; Adams et al. 2016).

4.3 | Mapping Genetic Diversity Across the Range of the Target Species

The predicted spatial distribution of genetic diversity across the putative ranges of both species show distinct patterns: genetic diversity in *E. catenatus* is predicted to be highest at the range core, while in *E. iheringii* it is predicted to be highest across the northern two-thirds of its range (Figure 8). For *E. iheringii*, this aligns with climate stability maps (Figure 8), where higher genetic diversity is observed in the northern, more equatorial portion of its range, which has remained relatively stable since the LGM. In contrast, the southern portion of *E. iheringii*'s range is within a narrow strip of historically stable habitat surrounded by relatively unstable habitat since the LGM.

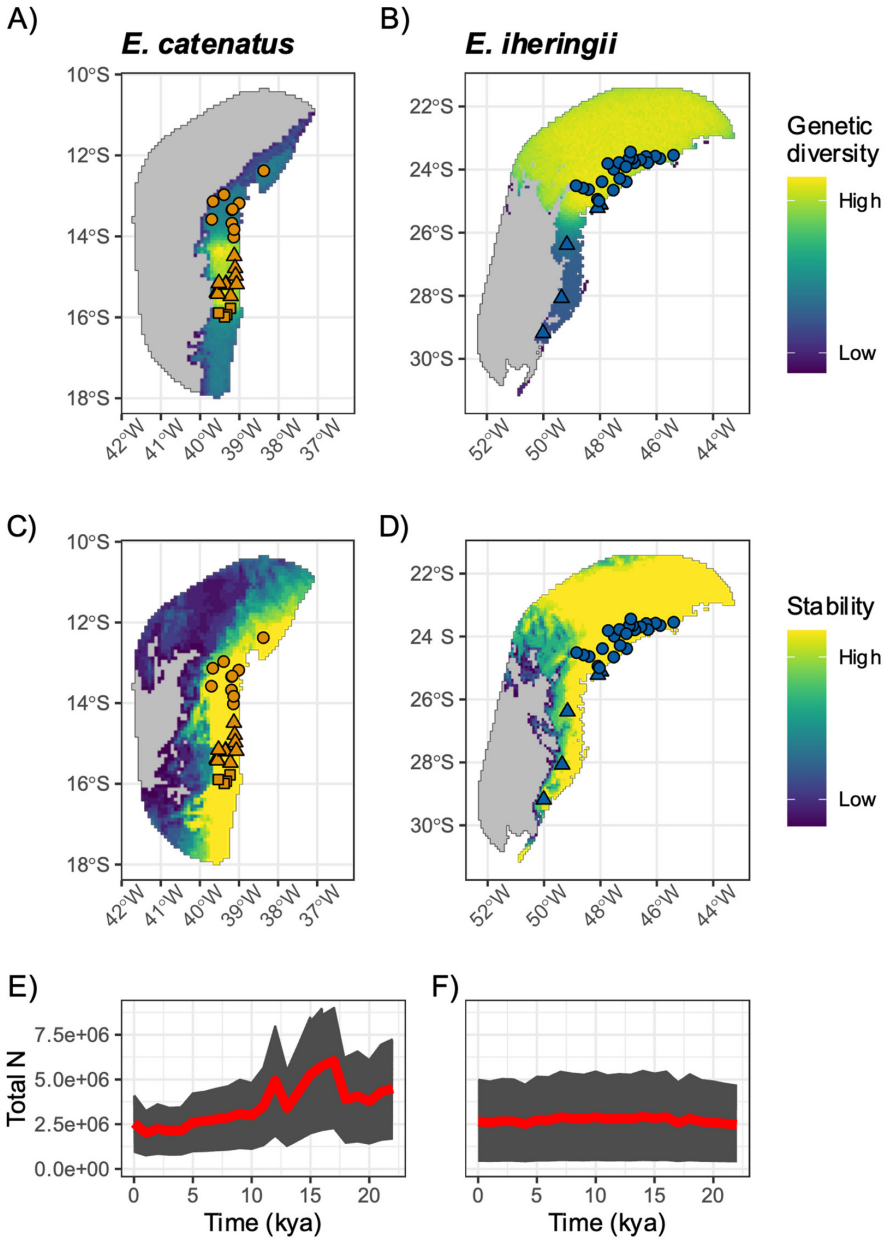

FIGURE 8 | Maps of genetic diversity (A, B) and range stability since the last glacial maximum (LGM) (C, D) for the low-elevation *Enyalius catenatus* (in orange) (A, C) and the high-elevation *Enyalius iheringii* (in blue) (B, D). Grey areas in the maps indicate regions with no predicted genetic diversity or no predicted suitability since the LGM. Shapes in (A–D) represent sNMF admixture population assignments, where the best $K=3$ for *E. catenatus* (A, B) and the best $K=2$ for *E. iheringii* (C, D). Panels E and F depict line plots showing the changes in the total number of individuals (Total *N*) simulated across the landscape since the LGM. Red lines indicate the Total *N* for landscapes calculated using the point estimate of maximum deme size for each species, while grey shading indicates 95% prediction intervals of the estimates.

For *E. catenatus*, higher levels of genetic diversity predicted to be found in the core may result from asymmetric, increased migration from the range edges as deme sizes have reduced over time (Figure 8E). However, the region with highest stability extends beyond the region with highest genetic diversity (Figure 8). This inconsistency may stem from greater uncertainty in demographic parameter inferences from the iDDC model for *E. catenatus*, suggesting that key demographic processes occur at different spatiotemporal scales, potentially biasing genetic diversity predictions (Cushman and Landguth 2010; Mascarenhas et al. 2019). This pattern, in addition to *E. catenatus*'s lower effective migration and potentially higher fragmentation, may render it vulnerable to additional factors like anthropogenic activity impacting its survival (Schlaepfer et al. 2018). Further conservation research for the species may be warranted.

Predictions on withheld test data indicate that the focal parameter estimates from the iDDC model (maximum deme size, effective migration, and ancestral population sizes) for *E. catenatus* were underestimated at high values and overestimated at low values (Figure 6 and Figure S3). This is likely due to the model overfitting the mean due to limited signal in the summary statistics to precisely estimate these parameters. High uncertainty in parameter estimates from

likelihood-free simulation based methods such as the SML approach used here and approximate Bayesian computation can arise when parameter values are not well correlated with summary statistics generated from the simulation model (Aeschbacher, Beaumont, and Futschik 2012). Despite this, the cross-validation analysis did recover some correlation between parameter values and estimates given that both focal iDDC models (linear and thresholded transformation) were successfully distinguished by the SML algorithm with moderate accuracy for *E. catenatus* (AUC: 0.74) and higher accuracy for *E. iheringii* (AUC: 0.86) (Figure 6).

4.4 | iDDC Models and Their Challenges

A general challenge for iDDC models is addressing uncertainties inherent to both approaches used (SDMs and spatial demographic models) within a unified framework (Brown and Knowles 2012; Alvarado-Serrano and Knowles 2014; Brown et al. 2016). The SDMs are influenced by different decisions concerning modelling and data (Radosavljevic and Anderson 2014; Soley-Guardia, Alvarado-Serrano, and Anderson 2024), while the population genetic simulations incorporate the uncertainty from coalescent and mutational variance (Edwards Scott and Beerli 2000) and the necessary modelling simplifications to simulate spatial data (Dabi and Schrider 2024). The suitabilities inferred from SDM can vary with modelling choices, potentially leading to simulated demographic scenarios that may not accurately reflect the species' ecological tolerances (Lee-Yaw et al. 2022). A potential solution is to include SDM uncertainty in the demographic model selection, where alternative models reflect different SDM modelling choices (Castilla et al. 2024). Simplifying assumptions in our demographic models—such as instantaneous population size changes and continuous migration over 1000-generation intervals—may not fully capture the demographic dynamics of the species, where population sizes and migration might fluctuate over shorter intervals and in a more heterogeneous way. Furthermore, biased migration or deep fine-scale fragmentation below the five arc-minute resolution used could inflate estimates of local deme size, while recurrent extinction events might reduce them (Ryman, Laikre, and Hössjer 2019). Still, population genetic models are inherently oversimplifications of reality aimed at balancing tractability with capturing relevant parameters in complex natural systems. By directly parameterising coalescent models from SDMs, without intermediate forward-in-time steps, we significantly reduce simulation time without sacrificing accuracy. Goodness of fit comparisons between simulated and observed summary statistics demonstrate that this approach effectively captures key features for both species. In addition, our efficient simulation model enables future studies to rigorously explore modelling choices such as SDM grain size and demographic parameter space moving forward iDDC-based approaches.

5 | Conclusions

In summary, our results show that iDDC model discrimination led to a better model fit for the thresholded transformation in both species, with higher confidence in model discrimination in the high-elevation *Enyalius iheringii* compared to the low-elevation *E. catenatus*. This supports our prediction that high elevation species, exposed to greater seasonality, exhibit wider environmental tolerances and a corresponding nonlinear relationship between SDM habitat suitability and local deme size. Moreover, consistent with Janzen's hypothesis that high-latitude and high-elevation species have better dispersal potential, we estimated lower local deme sizes and higher effective migration rates in the high-elevation species. These insights underscore the importance of carefully selecting SDM suitability-to-deme transformations in iDDC modelling frameworks, advancing our understanding of both species distribution and demography while providing a robust methodological approach for testing ecological hypotheses.

Author Contributions

C.M.F., A.C.C. and M.J.H. conceived of and framed the study. R.P.D., M.M.V. and M.T.R. performed fieldwork and collected sequencing data. C.M.F. and M.J.H. carried out the analyses. C.M.F. and M.J.H. led the writing. C.M.F., R.P.D., M.M.V., M.T.R., A.C.C. and M.J.H. contributed to interpretation of the results and to the writing, and all have approved the submission.

Acknowledgements

We thank the Hickerson, Carnaval, Anderson, and Lohman labs, along with Dr. Katharine Marske, Dr. Peter Ralph, and Dr. Phillip Staniczenko for valuable feedback on this manuscript. C.M.F., A.C.C. and M.J.H. thank NSF DBI 2104147 and NSF DEB 1343578. C.M.F. thanks the Doctoral Student Research Grant from the Graduate Center, City University of New York. M.T.R. thanks Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP # 2003/10335-8 and 2011/50146-6), CNPq, Furnas Centrais Elétricas, Hussam Zaher, Marcio Borges Martins, Marco Freitas, and all members of his lab for help in field. R.P.D. thanks Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP # 2013/22477-3 and 2015/11498-5). M.M.V. is supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP #2019/08308-0 and #2023/16814-9).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Demultiplexed RADseq reads and simulation output are available at https://doi.org/10.6084/m9.figshare.c.7522053. Code to reproduce all analyses is available at https://doi.org/10.5281/zenodo.14029198.

References

Adams, R. V., S. E. Lazerte, K. A. Otter, and T. M. Burg. 2016. "Influence of Landscape Features on the Microgeographic Genetic Structure of a Resident Songbird." *Heredity* 117: 63–72.

Aeschbacher, S., M. A. Beaumont, and A. Futschik. 2012. "A Novel Approach for Choosing Summary Statistics in Approximate Bayesian Computation." *Genetics* 192: 1027–1047.

Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela, and R. P. Anderson. 2015. "spThin: An R Package for Spatial Thinning of Species Occurrence Records for Use in Ecological Niche Models." *Ecography* 38: 541–545.

Alvarado-Serrano, D. F., and L. L. Knowles. 2014. "Ecological Niche Models in Phylogeographic Studies: Applications, Advances and Precautions." *Molecular Ecology Resources* 14: 233–248.

Anderson, R. P., and A. Raza. 2010. "The Effect of the Extent of the Study Region on GIS Models of Species Geographic Distributions and Estimates of Niche Evolution: Preliminary Tests With Montane Rodents (Genus Nephelomys) in Venezuela." *Journal of Biogeography* 37: 1378–1393.

Baird, N. A., P. D. Etter, T. S. Atwood, et al. 2008. "Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers." *PLoS One* 3: e3376.

Barber, R. F., E. J. Candès, A. Ramdas, and R. J. Tibshirani. 2021. "Predictive Inference With the Jackknife+." *Annals of Statistics* 49: 486–507.

Baumdicker, F., G. Bisschop, D. Goldstein, et al. 2022. "Efficient Ancestry and Mutation Simulation With Msprime 1.0." *Genetics* 220: iyab229.

Bodensteiner, B. L., G. A. Agudelo-Cantero, A. Z. A. Arietta, et al. 2021. "Thermal Adaptation Revisited: How Conserved Are Thermal Traits of Reptiles and Amphibians?" *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology* 335: 173–194.

Boria, R. A., L. E. Olson, S. M. Goodman, and R. P. Anderson. 2014. "Spatial Filtering to Reduce Sampling Bias Can Improve the Performance of Ecological Niche Models." *Ecological Modelling* 275: 73–77.

Bovo, R. P., M. N. Simon, D. B. Provete, M. Lyra, C. A. Navas, and D. V. Andrade. 2023. "Beyond Janzen's Hypothesis: How Amphibians That Climb Tropical Mountains Respond to Climate Variation." *Integrative Organismal Biology* 5: obad009.

Bozinovic, F., P. Calosi, and J. I. Spicer. 2011. "Physiological Correlates of Geographic Range in Animals." *Annual Review of Ecology, Evolution, and Systematics* 42: 155–179.

Bradburd, G. S., and P. L. Ralph. 2019. "Spatial Population Genetics: It's About Time." *Annual Review of Ecology, Evolution, and Systematics* 50: 427–449.

Brett, J. R. 1956. "Some Principles in the Thermal Requirements of Fishes." *Quarterly Review of Biology* 31: 75–87.

Brown, J. L., and L. L. Knowles. 2012. "Spatially Explicit Models of Dynamic Histories: Examination of the Genetic Consequences of Pleistocene Glaciation and Recent Climate Change on the American Pika." *Molecular Ecology* 21: 3757–3775.

Brown, J. L., J. J. Weber, D. F. Alvarado-Serrano, M. J. Hickerson, S. J. Franks, and A. C. Carnaval. 2016. "Predicting the Genetic Consequences of Future Climate Change: The Power of Coupling Spatial Demography, the Coalescent, and Historical Landscape Changes." *American Journal of Botany* 103: 153–163.

Bruscagin, R. T., M. Dixo, S. Famelli, and J. Bertoluci. 2017. "Patch Size Effects on Richness, Abundance, and Diversity of Leaf-Litter Lizards From Atlantic Rainforest Fragments." *Salamandra* 53: 59–65.

Carilo Filho, L. M., T. Medeiros, M. Sena, and A. Argolo. 2017. "*Enyalius catenatus* (Wied's Fat-Headed Anole). Clutch Size and Hatchling Size." *Herpetological Review* 48: 188–189.

Carnaval, A. C., and C. Moritz. 2008. "Historical Climate Modelling Predicts Patterns of Current Biodiversity in the Brazilian Atlantic Forest." *Journal of Biogeography* 35: 1187–1201.

Carnaval, A. C., E. Waltari, M. T. Rodrigues, et al. 2014. "Prediction of Phylogeographic Endemism in an Environmentally Complex Biome." *Proceedings of the Royal Society B: Biological Sciences* 281: 20141461.

Carscadden, K. A., N. C. Emery, C. A. Arnillas, et al. 2020. "Niche Breadth: Causes and Consequences for Ecology, Evolution, and Conservation." *Quarterly Review of Biology* 95: 179–214.

Castilla, A. R., A. Brown, S. Hoban, et al. 2024. "Integrative Demographic Modelling Reduces Uncertainty in Estimated Rates of Species' Historical Range Shifts." *Journal of Biogeography* 51: 325–336.

Chen, T., and C. Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16*, 785–794. New York: Association for Computing Machinery. https://doi.org/10.1145/2939672.2939785.

Colwell, R. K., G. Brehm, C. L. Cardelús, A. C. Gilman, and J. T. Longino. 2008. "Global Warming, Elevational Range Shifts, and Lowland Biotic Attrition in the Wet Tropics." *Science* 322: 258–261.

Costa, J. T. 1998. "Social Behavior and Its Effects on Colony- and Microgeographic Genetic Structure in Phytophagous Insect Populations." In *Genetic Structure and Local Adaptation in Natural Insect Populations: Effects of Ecology, Life History, and Behavior*, edited by S. Mopper and S. Y. Strauss, 205–238. Boston, MA: Springer US. https://doi.org/10.1007/978-1-4757-0902-5_10.

Currat, M., M. Arenas, C. S. Quilodràn, L. Excoffier, and N. Ray. 2019. "SPLATCHE3: Simulation of Serial Genetic Data Under Spatially Explicit Evolutionary Scenarios Including Long-Distance Dispersal." *Bioinformatics* 35: 4480–4483.

Currat, M., N. Ray, and L. Excoffier. 2004. "Splatche: A Program to Simulate Genetic Diversity Taking Into Account Environmental Heterogeneity." *Molecular Ecology Notes* 4: 139–142.

Cushman, S. A., and E. L. Landguth. 2010. "Scale Dependent Inference in Landscape Genetics." *Landscape Ecology* 25: 967–979.

Dabi, A., and D. Schrider. 2024. "Population Size Rescaling Significantly Biases Outcomes of Forward-In-Time Population Genetic Simulations." 2024.04.07.588318. https://www.biorxiv.org/content/10.1101/2024.04.07.588318v2.

Danecek, P., A. Auton, G. Abecasis, et al. 2011. "The Variant Call Format and VCFtools." *Bioinformatics* 27: 2156–2158.

Dobrowski, S. Z. 2011. "A Climatic Basis for Microrefugia: The Influence of Terrain on Climate." *Global Change Biology* 17: 1022–1035.

Dormann, C. F., J. Elith, S. Bacher, et al. 2013. "Collinearity: A Review of Methods to Deal With It and a Simulation Study Evaluating Their Performance." *Ecography* 36: 27–46.

Eaton, D. A. R., and I. Overcast. 2020. "Ipyrad: Interactive Assembly and Analysis of RADseq Datasets." *Bioinformatics* 36: 2592–2594.

Edwards Scott, V., and P. Beerli. 2000. "Perspective: Gene Divergence, Population Divergence, and the Variance in Coalescence Time in Phylogeographic Studies." *Evolution* 54: 1839–1854.

Elith, J., and J. R. Leathwick. 2009. "Species Distribution Models: Ecological Explanation and Prediction Across Space and Time." *Annual Review of Ecology, Evolution, and Systematics* 40: 677–697.

Excoffier, L., N. Marchi, D. A. Marques, R. Matthey-Doret, A. Gouy, and V. C. Sousa. 2021. "fastsimcoal2: Demographic Inference Under Complex Evolutionary Scenarios." *Bioinformatics* 37: 4882–4885.

Exposito-Alonso, M., T. R. Booker, L. Czech, et al. 2022. "Genetic Diversity Loss in the Anthropocene." *Science* 377: 1431–1435.

Fenderson, L. E., A. I. Kovach, and B. Llamas. 2020. "Spatiotemporal Landscape Genetics: Investigating Ecology and Evolution Through Space and Time." *Molecular Ecology* 29: 218–246.

Fitzpatrick, M. C., and W. W. Hargrove. 2009. "The Projection of Species Distribution Models and the Problem of Non-Analog Climate." *Biodiversity and Conservation* 18: 2255–2261.

French, C. M., L. D. Bertola, A. C. Carnaval, et al. 2023. "Global Determinants of Insect Mitochondrial Genetic Diversity." *Nature Communications* 14: 5276.

Frichot, E., F. Mathieu, T. Trouillon, G. Bouchard, and O. François. 2014. "Fast and Efficient Estimation of Individual Ancestry Coefficients." *Genetics* 196: 973–983.

Gaggiotti, O. E., A. Chao, P. Peres-Neto, et al. 2018. "Diversity From Genes to Ecosystems: A Unifying Framework to Study Variation

Across Biological Metrics and Scales." *Evolutionary Applications* 11: 1176–1193.

Guisan, A., and W. Thuiller. 2005. "Predicting Species Distribution: Offering More Than Simple Habitat Models." *Ecology Letters* 8: 993–1009.

Harris, C. R., K. J. Millman, S. J. van der Walt, et al. 2020. "Array Programming With NumPy." *Nature* 585: 357–362.

He, Q., D. L. Edwards, and L. L. Knowles. 2013. "Integrative Testing of How Environments From the Past to the Present Shape Genetic Structure Across Landscapes." *Evolution* 67: 3386–3402.

Hijmans, R. J. 2022. "terra: Spatial Data Analysis." https://CRAN.R-project.org/package=terra.

Hijmans, R. J., S. Phillips, J. Leathwick, and J. Elith. 2022. "dismo: Species Distribution Modeling." https://CRAN.R-project.org/package=dismo.

Jackson, J. F. 1978. "Differentiation in the Genera Enyalius and Strobilurus (Iguanidae): Implications for Pleistocene Climatic Changes in Eastern Brazil." *Arquivos de Zoologia* 30: 1–79.

Janzen, D. H. 1967. "Why Mountain Passes Are Higher in the Tropics." *American Naturalist* 101: 233–249.

Jezkova, T., M. Leal, and J. A. Rodríguez-Robles. 2009. "Living Together but Remaining Apart: Comparative Phylogeography of *Anolis poncensis* and *A. cooki*, Two Lizards Endemic to the Aridlands of Puerto Rico." *Biological Journal of the Linnean Society* 96: 617–634.

Kamm, J., J. Terhorst, R. Durbin, and Y. S. Song. 2020. "Efficiently Inferring the Demographic History of Many Populations With Allele Count Data." *Journal of the American Statistical Association* 115: 1472–1487.

Karger, D. N., O. Conrad, J. Böhner, et al. 2017. "Climatologies at High Resolution for the Earth's Land Surface Areas." *Scientific Data* 4: 170122.

Karger, D. N., O. Conrad, J. Böhner, et al. 2018. "Data From: Climatologies at High Resolution for the Earth's Land Surface Areas. 7266827510 Bytes." https://datadryad.org/stash/dataset/doi:10.5061/dryad.kd1d4.

Karger, D. N., M. P. Nobis, S. Normand, C. H. Graham, and N. E. Zimmermann. 2023. "CHELSA-TraCE21k – High-Resolution (1 km) Downscaled Transient Temperature and Precipitation Data Since the Last Glacial Maximum." *Climate of the Past* 19: 439–456.

Kass, J. M., R. Muscarella, P. J. Galante, et al. 2021. "ENMeval 2.0: Redesigned for Customizable and Reproducible Modeling of Species' Niches and Distributions." *Methods in Ecology and Evolution* 12: 1602–1608.

Kimura, M., G. H. Weiss, and G. H. Weiss. 1964. "The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation With Distance." *Genetics* 49: 561–576.

Knowles, L. L., and R. Massatti. 2017. "Distributional Shifts – Not Geographic Isolation – As a Probable Driver of Montane Species Divergence." *Ecography* 40: 1475–1485.

Laporte, V., and B. Charlesworth. 2002. "Effective Population Size and Population Subdivision in Demographically Structured Populations." *Genetics* 162: 501–519.

Lee-Yaw, J. A., J. L. McCune, S. Pironon, and S. N. Sheth. 2022. "Species Distribution Models Rarely Predict the Biology of Real Populations." *Ecography* 2022: e05877.

Liou, N. S. 2008. "História Natural de Duas Espécies Simpátricas de Enyalius (Squamata, Leiosauridae) na Mata Atlântica do Sudeste Brasileiro." *Zoology* PhD. https://teses.usp.br/teses/disponiveis/41/41133/tde-26022009-163611/publico/Noraly.pdf.

Liu, J., X. Guo, D. Chen, J. Li, B. Yue, and X. Zeng. 2019. "Diversification and Historical Demography of the Rapid Racerunner (*Eremias velox*) in Relation to Geological History and Pleistocene Climatic Oscillations in Arid Central Asia." *Molecular Phylogenetics and Evolution* 130: 244–258.

Lu, M., and W. Jetz. 2023. "Scale-Sensitivity in the Measurement and Interpretation of Environmental Niches." *Trends in Ecology & Evolution* 38: 554–567.

Mascarenhas, R., C. Y. Miyaki, R. Dobrovolski, and H. Batalha-Filho. 2019. "Late Pleistocene Climate Change Shapes Population Divergence of an Atlantic Forest Passerine: A Model-Based Phylogeographic Hypothesis Test." *Journal für Ornithologie* 160: 733–748.

McCartney-Melstad, E., M. Gidiş, and H. B. Shaffer. 2019. "An Empirical Pipeline for Choosing the Optimal Clustering Threshold in RADseq Studies." *Molecular Ecology Resources* 19: 1195–1204.

Miles, A., pyup.io bot, R. M, P. Ralph, et al. 2023. "cggh/scikit-allel: v1.3.7." https://doi.org/10.5281/zenodo.7945985.

Muñoz, M. M. 2022. "The Bogert Effect, a Factor in Evolution." *Evolution* 76: 49–66.

Muñoz, M. M., M. A. Stimola, A. C. Algar, et al. 2014. "Evolutionary Stasis and Lability in Thermal Physiology in a Group of Tropical Lizards." *Proceedings of the Biological Sciences* 281: 20132433.

Naimi, B., N. Hamm, T. A. Groen, A. K. Skidmore, and A. G. Toxopeus. 2014. "Where is Positional Uncertainty a Problem for Species Distribution Modelling." *Ecography* 37: 191–203.

Nunney, L. 2016. "The Effect of Neighborhood Size on Effective Population Size in Theory and in Practice." *Heredity* 117: 224–232.

Overcast, I., M. Ruffley, J. Rosindell, et al. 2021. "A Unified Model of Species Abundance, Genetic Diversity, and Functional Diversity Reveals the Mechanisms Structuring Ecological Communities." *Molecular Ecology Resources* 21: 2782–2800.

Pan, D., K. Hülber, W. Willner, and G. M. Schneeweiss. 2020. "An Explicit Test of Pleistocene Survival in Peripheral Versus Nunatak Refugia in Two High Mountain Plant Species." *Molecular Ecology* 29: 172–183.

Pebesma, E. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *R Journal* 10: 439–446.

Pebesma, E., and R. Bivand. 2023. *Spatial Data Science: With Applications in R*. London, UK: Chapman and Hall/CRC. https://r-spatial.org/book/.

Pedregosa, F., G. Varoquaux, A. Gramfort, et al. 2011. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12: 2825–2830.

Perry, B. W., D. C. Card, J. W. McGlothlin, et al. 2018. "Molecular Adaptations for Sensing and Securing Prey and Insight Into Amniote Genome Diversity From the Garter Snake Genome." *Genome Biology and Evolution* 10: 2110–2129.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. "Maximum Entropy Modeling of Species Geographic Distributions." *Ecological Modelling* 190: 231–259.

Prates, I., A. T. Xue, J. L. Brown, et al. 2016. "Inferring Responses to Climate Dynamics From Historical Demography in Neotropical Forest Lizards." *PNAS* 113: 7978–7985.

Prosser, M. R., H. L. Gibbs, and P. J. Weatherhead. 1999. "Microgeographic Population Genetic Structure in the Northern Water Snake, *Nerodia sipedon sipedon* Detected Using Microsatellite DNA Loci." *Molecular Ecology* 8: 329–333.

Qiao, H., E. E. Saupe, J. Soberón, A. T. Peterson, and C. E. Myers. 2016. "Impacts of Niche Breadth and Dispersal Ability on Macroevolutionary Patterns." *American Naturalist* 188: 149–162.

Radosavljevic, A., and R. P. Anderson. 2014. "Making Better Maxent Models of Species Distributions: Complexity, Overfitting and Evaluation." *Journal of Biogeography* 41: 629–643.

Rand, A. S. 1982. "Clutch and Egg Size in Brazilian Iguanid Lizards." *Herpetologica* 38: 171–178.

Rautenberg, R., and R. R. Laps. 2010. "Natural History of the Lizard *Enyalius iheringii* (Squamata, Leiosauridae) in Southern Brazilian Atlantic Forest." *Iheringia Serie Zoologia* 100: 287–290.

Rodrigues, M. T., C. E. V. Bertolotto, R. C. Amaro, Y. Yonenaga-Yassuda, E. M. X. Freire, and K. C. M. Pellegrino. 2014. "Molecular Phylogeny, Species Limits, and Biogeography of the Brazilian Endemic Lizard Genus Enyalius (Squamata: Leiosauridae): An Example of the Historical Relationship Between Atlantic Forests and Amazonia." *Molecular Phylogenetics and Evolution* 81: 137–146.

Rousset, F. 1997. "Genetic Differentiation and Estimation of Gene Flow From F-Statistics Under Isolation by Distance." *Genetics* 145: 1219–1228.

Ryman, N., L. Laikre, and O. Hössjer. 2019. "Do Estimates of Contemporary Effective Population Size Tell Us What We Want to Know?" *Molecular Ecology* 28: 1904–1918.

Schlaepfer, D. R., B. Braschler, H.-P. Rusterholz, and B. Baur. 2018. "Genetic Effects of Anthropogenic Habitat Fragmentation on Remnant Animal and Plant Populations: A Meta-Analysis." *Ecosphere* 9: e02488.

Shcheglovitova, M., and R. P. Anderson. 2013. "Estimating Optimal Complexity for Ecological Niche Models: A Jackknife Approach for Species With Small Sample Sizes." *Ecological Modelling* 269: 9–17.

Sherwin, W. B., A. Chao, L. Jost, and P. E. Smouse. 2017. "Information Theory Broadens the Spectrum of Molecular Ecology and Evolution." *Trends in Ecology & Evolution* 32: 948–963.

Soley-Guardia, M., D. F. Alvarado-Serrano, and R. P. Anderson. 2024. "Top Ten Hazards to Avoid When Modeling Species Distributions: A Didactic Guide of Assumptions, Problems, and Recommendations." *Ecography* 2024: e06852. https://onlinelibrary.wiley.com/doi/abs/10.1111/ecog.06852.

Soley-Guardia, M., E. E. Gutiérrez, D. M. Thomas, J. Ochoa-G, M. Aguilera, and R. P. Anderson. 2016. "Are We Overestimating the Niche? Removing Marginal Localities Helps Ecological Niche Models Detect Environmental Barriers." *Ecology and Evolution* 6: 1267–1279.

Stark, J. R., and J. D. Fridley. 2022. "Microclimate-Based Species Distribution Models in Complex Forested Terrain Indicate Widespread Cryptic Refugia Under Climate Change." *Global Ecology and Biogeography* 31: 562–575.

Strangas, M. L., C. A. Navas, M. T. Rodrigues, and A. C. Carnaval. 2019. "Thermophysiology, Microclimates, and Species Distributions of Lizards in the Mountains of the Brazilian Atlantic Forest." *Ecography* 42: 354–364.

Sunday, J. M., A. E. Bates, M. R. Kearney, et al. 2014. "Thermal-Safety Margins and the Necessity of Thermoregulatory Behavior Across Latitude and Elevation." *Proceedings of the National Academy of Sciences of the United States of America* 111: 5610–5615.

Szép, E., B. Trubenová, and K. Csilléry. 2022. "Using gridCoal to Assess Whether Standard Population Genetic Theory Holds in the Presence of Spatio-Temporal Heterogeneity in Population Size." *Molecular Ecology Resources* 22: 2941–2955.

Theodoridis, S., D. A. Fordham, S. C. Brown, S. Li, C. Rahbek, and D. Nogues-Bravo. 2020. "Evolutionary History and Past Climate Change Shape the Distribution of Genetic Diversity in Terrestrial Mammals." *Nature Communications* 11: 2557.

VanDerWal, J., L. P. Shoo, C. N. Johnson, and S. E. Williams. 2009. "Abundance and the Environmental Niche: Environmental Suitability Estimated From Niche Models Predicts the Upper Limit of Local Abundance." *American Naturalist* 174: 282–291.

von May, R., A. Catenazzi, A. Corl, R. Santa-Cruz, A. C. Carnaval, and C. Moritz. 2017. "Divergence of Thermal Physiological Traits in Terrestrial Breeding Frogs Along a Tropical Elevational Gradient." *Ecology and Evolution* 7: 3257–3267.

Wakeley, J. 2000. "The Effects of Subdivision on the Genetic Divergence of Populations and Species." *Evolution* 54: 1092–1101.

Wakeley, J. 2004. "Metapopulation Models for Historical Inference." *Molecular Ecology* 13: 865–875.

Walters, A. D., D. A. Trujillo, and D. J. Berg. 2022. "Micro-Endemic Species of Snails and Amphipods Show Population Genetic Structure Across Very Small Geographic Ranges." *Heredity* 128: 325–337.

Wang, I. J., and K. Summers. 2010. "Genetic Structure Is Correlated With Phenotypic Divergence Rather Than Geographic Isolation in the Highly Polymorphic Strawberry Poison-Dart Frog." *Molecular Ecology* 19: 447–458.

Weber, M. M., R. D. Stevens, J. A. F. Diniz-Filho, and C. E. V. Grelle. 2017. "Is There a Correlation Between Abundance and Environmental Suitability Derived From Ecological Niche Modelling? A Meta-Analysis." *Ecography* 40: 817–828.

Wisz, M. S., J. Pottier, W. D. Kissling, et al. 2013. "The Role of Biotic Interactions in Shaping Distributions and Realised Assemblages of Species: Implications for Species Distribution Modelling." *Biological Reviews* 88: 15–30.

Wu, J., X.-Y. Chen, H. Zhang, L.-D. Xiong, H. Lei, and S.-H. Deng. 2019. "Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimizationb." *Journal of Electronic Science and Technology* 17: 26–40.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.