

ECOGRAPHY

Research

Postglacial recolonization of North America by spadefoot toads: integrating niche and corridor modeling to study species' range dynamics over geologic time

Iulian Gherghel and Ryan Andrew Martin

I. Gherghel (<https://orcid.org/0000-0001-7963-6498>) ✉ (iuliangherghel@case.edu) and R. A. Martin (<https://orcid.org/0000-0002-7448-5907>), Dept of Biology, Case Western Reserve Univ., Cleveland, OH, USA.

Ecography

43: 1499–1509, 2020

doi: 10.1111/ecog.04942

Subject Editor:

Jens-Christian Svenning

Editor-in-Chief: Miguel Araújo

Accepted 25 June 2020



Understanding the factors that shape species' distributions is a key topic in biogeography. As climates change, species can either cope with these changes through evolution, plasticity or by shifting their ranges to track the optimal climatic conditions. Ecological niche modeling (ENM) is a widespread technique in biogeography that estimates the niche of the organism by using occurrences and environmental data to estimate species' potential distributions. ENMs are often criticized for failing to take species' dispersal abilities into consideration. Here, we attempt to fill this gap by combining ENMs with dispersal and corridor modeling to study the range dynamics of North American spadefoot toads (Scaphiopodidae) over the Holocene. We first estimated the current and past distributions of spadefoot toads and then estimated their past distributions from the Last Glacial Maximum (LGM) to the present day. Then, we estimated how each taxon recolonized North America by using dispersal and corridor modeling. By combining these two modeling approaches we were able to 1) estimate the LGM refugia used by the North American spadefoot toads, 2) further refine these projections by estimating which of the putative LGM refugia contributed to the recolonization of North America via dispersal, and 3) estimate the relative influence of each LGM refugium to the current species' distributions. The models were tested using previously published phylogeographic data, revealing a high degree of congruence between our models and the genetic data. These results suggest that combining ENMs and dispersal modeling over time is a promising approach to investigate both historical and future species' range dynamics.

Keywords: dispersal, glacial refugia, last glacial maximum, spadefoot toads, species distribution modeling

Introduction

Understanding the factors that influence and shape species' distributions is a central topic in biogeography. Ecological niche modeling (ENM) estimates the niche of the organism by using occurrences and environmental data to estimate species' potential distributions (Franklin 2009, Peterson et al. 2011). ENMs are a commonly



www.ecography.org

© 2020 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

used technique to study a wide array of topics in conservation (Gómez-Ruiz et al. 2017, Taylor et al. 2018), climate change (Wauchope et al. 2017, Dyderski et al. 2018), evolution (Loera et al. 2012, Sillero et al. 2014), species invasion (Lantschner et al. 2017, Mutascio et al. 2017) and global health (Sehgal et al. 2011, Wang et al. 2018). ENMs are also used to study range shifts in response to both historical (Gray et al. 2017, Huang et al. 2017) and contemporary climate change (Veloz et al. 2012, Boria et al. 2017) by projecting the estimated niche on climate scenarios. Nonetheless, while ENMs contribute to our fundamental understanding of how species are distributed now and into the future, there is a need to further develop and improve these methods.

ENMs are often criticized for failing to take into consideration species' dispersal abilities. In a recent review, Miller and Holloway (2015) found just a handful of attempts that account for dispersal in ENMs (Engler and Guisan 2009, Dullinger et al. 2012, Sahlean et al. 2014). Failing to account for dispersal can lead to substantial errors in projected species' distribution estimates (Engler and Guisan 2009), especially if the models are being projected on areas that are inaccessible due to species' dispersal abilities (Dullinger et al. 2012, Sahlean et al. 2014). Engler and Guisan (2009) introduced the MigClim to run simulations of species' dispersal, similar to another approach proposed by Sahlean et al. (2014) which uses the landscape characters' resistance estimates to estimate dispersal. When species distribution estimates are not informed by species' dispersal abilities, the estimates often overestimate species' range shifts, and dispersal estimates using statistical methods are highly desired to study species' range shifts especially in the context of changing climates (Miller and Holloway 2015). To fill this gap, we combine ENMs with dispersal models over time to explore the dispersal patterns and corridors used by species to disperse from their glacial refugia. Combining the two methods (ENMs and dispersal modeling) offers the advantage of using these well-established methods to simulate how species shift their ranges across time.

Past climate change events, such as the cycles between glaciated (cold) and interglacial (warm) periods through the Pleistocene, offer opportunities to study species' range shifts as species track favorable climates. Postglacial migrations are usually sourced from a few glacial refugia (Svenning et al. 2008, Normand et al. 2011), and these refugia do not contribute equally to the postglacial colonization process (Tzedakis et al. 2013). Consequently, species do not often occupy their full potential distributions (Svenning et al. 2008, Normand et al. 2011, Dullinger et al. 2012, Bai et al. 2018). Moving forward, examining the dispersal patterns and migration routes from LGM refugia is important for understanding how current species ranges have been shaped by dispersal and niche-related processes. A common way of determining these dispersal patterns is through population and landscape genetic tools, for example via haplotype networks combined with environmental data. However, these approaches can be expensive in time and money because they require dense genetic sampling across the range (Stöck et al. 2012, Wielstra

and Arntzen 2012, Skourtanioti et al. 2016, Psonis et al. 2017, 2018). In our study, we explore whether combining ENMs with dispersal and corridor modeling can provide reliable insights on the dispersal patterns, range dynamics and corridors used by species to disperse from their putative LGM refugia at the end of last Ice Age. To model dispersal and corridors, we use least-cost paths, a tool developed in spatial analysis to find the path between two locations that travels thru the most cost-effective route between the locations (as a function of time or distance) also commonly used in population genetics (Chan et al. 2011). We first estimate the current distributions for North American spadefoot toads and then estimate their past distributions during the Last Glacial Maximum, 21 000 yr ago. We then estimate how each taxon recolonized North America by using dispersal and corridor modeling. By combining these two modeling approaches, we 1) project which LGM refugia were used by the North American spadefoot toads, and 2) further refine these projections by estimating which of the putative LGM refugia have contributed to the recolonization of North America via dispersal. Finally, we test the predictive ability of our models against previously published molecular phylogeography data (Neal et al. 2018). With this exercise, we aim to demonstrate a new approach for studying species' range dynamics by combining ENMs and dispersal modeling.

Methods

Study organism

To study whether combining ENMs with dispersal and corridor modeling provide a reliable method for studying species' range dynamics during the Holocene, we used the spadefoot toads (Scaphiropodidae) as a case study because their distributions were affected by the Ice Age (García-Paris et al. 2003). Scaphiropodidae is a group of spadefoot toads found only in North America, primarily in southwestern Canada, and throughout the United States and Mexico (Duellman and Trueb 1986). There are seven extant species of North American spadefoot toads, belonging to two genera, *Scaphiopus* (*S. couchii*, *S. holbrookii* and *S. hurterii*) and *Spea* (*S. bombifrons*, *S. hammondii*, *S. intermontana* and *S. multiplicata*) (García-Paris et al. 2003). These frogs are fossorial, spending most of the time burrowed in the sand, emerging en masse to reproduce on the first night of seasonal rains (Duellman and Trueb 1986). All species of both *Scaphiopus* and *Spea* have extremely rapid larval development, as fast as 7 d in *S. couchii*, 12–15 d in *S. holbrookii*, *S. hurterii*, *S. bombifrons* and *S. multiplicata* and under 30 d in *S. hammondii* and *S. intermontana* (Duellman and Trueb 1986, Zeng et al. 2014).

Species occurrence data

We collected occurrence data for the North American spadefoot toads from published literature and online

databases (HerpMapper 2019, Global Biodiversity Information Facility). Occurrence data without Latitude and Longitude were manually geo-referenced using Global Gazetteer ver. 2.1 (<www.fallingrain.com/world>). We then manually searched and georeferenced locations not found in the gazetteer using Google Earth (<https://earth.google.com/>) following the instructions on the museum collection label. In total, the dataset comprised over 8000 georeferenced records, the entire range of the studied taxa, providing a good representation of their distribution. We then tested for spatial bias in the occurrence dataset using a global Moran I test and found that the species occurrence data were heavily clumped (Moran I index = 0.38, z = 440, p ≤ 0.001). To mitigate this spatial bias of the data, we used the function ‘trim duplicate’ in the Pearl application ENMTools 1.3 (Warren et al. 2010), then we used SDMtoolbox (Brown and Anderson 2014) (<http://sdmtoolbox.org/>) to further rarefy the occurrence points based on the environmental datasets (see Environmental data section) and the distance between points (20 km). This allowed us to avoid spatial bias and resulted into 2741 unique, spatially and environmentally unbiased occurrence records for the studied taxa (*S. couchii* (703), *S. holbrookii* (369), *S. hurterii* (130), *S. bombifrons* (605), *S. hammondii* (138), *S. intermontana* (257), *S. multiplicata* (539)).

Climatic variables

We ran the ENMs using climatic data downloaded from WorldClim 1.4 database at 5 km resolution (Hijmans et al. 2005, <www.worldclim.org>). The WorldClim database is comprised of 19 bioclimatic variables (Margules and Austin 1990, sensu Nix and Busby 1986), which are a combination of seasonal and monthly temperature and precipitation variables that have proved to be good predictors for estimating species distributions (Nix and Busby 1986, Beaumont et al. 2005, Hijmans and Graham 2006). Variables with a high degree of multicollinearity (Pearson correlation, r > 0.75) were eliminated from the analysis (Dormann et al. 2013). This resulted in eight bioclimatic variables used to create the ENMs (mean diurnal temperature range, isothermality, minimum temperature of the coldest month, mean temperature of the wettest quarter, mean temperature of the driest quarter, mean temperature of the warmest quarter, annual precipitation, precipitation seasonality) (Table 1). For estimating the past species’ distributions of North American spadefoot toads, we used climatic reconstruction models corresponding to the Last Glacial Maximum (LGM) and mid-Holocene. The LGM data used in this study come from two global circulation model simulations, the Model for Interdisciplinary Research on Climate (MIROC) produced by the Center for Climate System Research (Hasumi 2007) and the Community Climate System Model (CCSM) produced by University Corporation for Atmospheric Research (Collins et al. 2006), <www.worldclim.org> (Hijmans et al. 2005).

Table 1. Summary of the average performance and variable contributions for the generated ENMs.

Species	AUC training	10% omission error (%)	Cohen's Kappa	Mean temperature of warmest quarter	Annual precipitation	Precipitation seasonality	Mean diurnal temperature range	Isothermality	Min temperature of coldest month	Mean temperature of wettest quarter	Mean temperature of driest quarter
<i>Scaphiopus couchii</i>	0.91	8	0.82	11.47	24.25	7.20	5.99	13.85	9.56	22.65	5.04
<i>Scaphiopus holbrookii</i>	0.89	8	0.75	10.56	14.62	25.13	0.67	4.88	5.84	16.35	21.96
<i>Scaphiopus hurterii</i>	0.92	8	0.75	25.85	13.74	11.24	5.70	15.79	7.21	14.32	6.15
<i>Spea intermontana</i>	0.91	9	0.79	4.54	39.41	21.35	0.76	2.86	3.51	24.55	3.02
<i>Spea bombifrons</i>	0.88	9	0.74	13.29	5.39	25.24	13.99	8.36	14.95	7.81	10.97
<i>Spea hammondii</i>	0.95	4	0.87	6.71	11.32	34.60	7.57	3.53	3.87	25.20	7.19
<i>Spea multiplicata</i>	0.88	10	0.76	6.57	13.54	4.80	23.84	11.56	7.22	28.02	4.46

Ecological niche modeling

To estimate the potential distributions of the North American spadefoot toads we used classification tree analysis, multivariate adaptive regression splines, generalized linear models, generalized additive models, artificial neural networks, Maxent and support vector machines using the 'SSDM' R package (Schmitt et al. 2017). To estimate the past species' distributions of the North American spadefoot toads, the niche models trained with present climate data were projected over the past climate reconstructions corresponding to the LGM (see Climatic variables section). Since significant inter-model variation is present when projecting models in time or space, we followed Araujo et al.' (2019) recommendations to run multiple algorithms, stacking all the results of the algorithms into one AUC weighted stacked projection for each species that were used in all of our analyses. Since accounting for dispersal is an important prior of ENMs (Barve et al. 2011), we used as the accessible area the current range of the North American spadefoot toads (IUCN 2016). Models were built using a random subset of occurrence points (75%, calibration dataset) and model performance was evaluated using the remaining occurrence points (25%, validation dataset) (Phillips et al. 2006, Peterson et al. 2011). Variable contribution to the model was calculated based on the Pearson correlation coefficient between the model with all variables and models where each variable was omitted in turn, using the SSDM package (Schmitt et al. 2017). The resulting models were evaluated using Receiver Operating Characteristic (ROC) Area Under the Curve (AUC) and Cohen's K test (Mouton et al. 2010). AUC's and Cohen's K are metrics that range from 0 (no model fit) to 1 (perfect model fit) (Hand and Till 2001) and are frequently used to evaluate ENMs (Phillips et al. 2006, Franklin 2009, Jiménez-Valverde 2012). In addition to the AUC and Cohen's K, we also used 10 percentile omission error as a measure of the models' power to discriminate between suitable and unsuitable climatic conditions for the North American spadefoot toads. This metric is more reliable for evaluating the fit of the models because it takes into consideration only the omission error, which is independently calculated from a validation data set corresponding to known species distributions (Mouton et al. 2010, Jiménez-Valverde 2012). To calculate the 10 percentile omission error, we took the thresholded binary present/absent prediction and calculated the ratio of occurrences predicted absent of the occurrences set aside for testing the models (the validation dataset, see above) (Franklin 2009, Peterson et al. 2011).

Estimating dispersal patterns and migration corridors

For estimating the dispersal patterns (here used to describe the trends observed in the models) and migration corridors (here used to discuss the estimated corridors resulted from the dispersal analysis) used by each North American spadefoot toad species from their estimated glacial refugia to their current known distribution (represented by the occurrence

points rarefied at 50 km to reduce spatial clumping and bias for each species), we first created a friction layer for each species. The friction layer assumes that movement can occur across a landscape in response to other landscape elements besides distance (the ease of movement is greatest where resistance/friction is least) (McRae et al. 2008). For creating the friction layer, we first summed the continuous SSDMs' suitability of the present and mid-Holocene conditions, and then inverted these summed values so that lower values in the final friction layer represent higher suitability SSDMs' values using SDMToolbox (Brown and Anderson 2014). We choose this method due to the overwhelming evidence that species distribution models' suitability scores correlate directly with functional connectivity and gene flow (Wang et al. 2008, Chan et al. 2011). Using this technique to estimate the friction layer in corridor and dispersal analyses has advantages over methods relying on expert opinion where values are chosen subjectively (Chan et al. 2011, Brown and Anderson 2014). Next, we used the friction layer to inform movement across the landscape from the estimated LGM refugia (used as sources) to present-day occurrence records using least-cost paths (LCPs) (McRae et al. 2008). This method allowed us to estimate the most likely migration corridors (LCPs) that spadefoot toads followed to spread across North America to result in their present-day distributions. For estimating the corridors, we combined all the LCPs and calculated the density of the LCPs across North America. We define corridors used by the species to migrate from their putative LGM refugia to inhabit the extant range as the areas where most LCPs are passing (high corridor=66–100% of the LCPs pass the area; intermediate corridor=33–65% of the LCPs pass the area; minimum corridor=5–32% of the LCPs pass the area, and the areas that are being passed by less than 10% of the LCPs were masked due to low congruence of LCPs). Next, to estimate how many of the putative LGM refugia contributed to current distributions, we examined the association of the LCP of the extant occurrence points to the putative LGM refugia. Following this methodology, if an extant point is directly linked with a putative LGM refugium by an LCP, then we assume that a) the population associated with the occurrence point has sourced from the putative LGM refugium linking the two by an LCP, and b) the population is within the influence of the putative glacial refugium. If a putative glacial refugium is linked to no extant occurrence points by an LCP, then we assume that it was not used. These analyses were repeated for each species using Spatial Analyst in ArcGIS 10.3.

Testing the dispersal and migration corridor estimates with independent data

Lastly, we tested our dispersal models by comparing our results to recent independent molecular phylogeographic data (Neal et al. 2018). This was done in ArcGIS 10.3, using the Spatial Analyst toolbox where we overlapped the previously published population genetic data (Neal et al. 2018) for one of the species (*S. hammondi*) used in our study with

the dispersal and corridor estimates to test their reliability. In this and previous studies, *S. hammondi* was found to have a high degree of genetic divergence between the southern and northern populations (Garcia-Paris et al. 2003, Neal et al. 2018). Therefore, we tested the degree of overlap between our estimated corridor model results and the results obtained from molecular phylogeography (Neal et al. 2018) as an independent method to test the predictions of our dispersal models. Population genetic data are currently insufficient to perform a similar test for the other North American spadefoot toad species, although we make more informal comparisons where possible.

Results

Performance of ecological niche models and estimated present distributions

All ENMs performed very well (train AUC: 0.91) and adequately predicted the current distribution of the species (the average 10% omission test: 8%) (Table 1). *Spea hammondi* had the best model fit (where only 4% of the occurrences left for testing the models were omitted under the threshold), whereas the model for *S. bombifrons* and *S. multiplicata* had a good AUC (train AUC: 0.88), but more occurrences used for testing were omitted under the threshold (Table 1). The most important predictor for estimating the distributions of North American spadefoot toads were the mean temperature of the wettest quarter (19.84%), precipitation seasonality (18.51%) and annual precipitation (17.46%). However, species differed in how important each variable was to their distributions (Table 1). For example, *S. holbrookii*'s distribution is highly affected by the mean temperature of the wettest quarter, whereas *S. multiplicata* is less affected by this variable (Table 1). Generally, the models identified the core of the known distributions of the North American spadefoot toads (Table 1, Fig. 1), with most predictions being within the known ranges of the studied species (Fig. 1). However, the models for *S. intermontana* and *S. multiplicata* identified suitable areas outside of their known present range, most notably in California and Mexico.

Past distributions and glacial refugia

The two LGM circulation models identified key areas for the past distributions of the North American spadefoot toads. The past distributions of *S. holbrookii*, *S. hammondi* and *S. multiplicata* highly overlap their current ranges (Fig. 1, 2). Most changes occurred in the northern part of the species' ranges. In contrast, *S. couchii*, *S. hurterii*, *S. intermontana* and *S. bombifrons* likely saw a drastic change in their ranges in the past compared with today (Fig. 1, 2). From the putative LGM refugia initially estimated, we have selected only the LGM refugia that were connected by a least cost path to a known presence occurrence of the spadefoot toad species (Fig. 3). Among all species, only 40% of all potential LGM

refugia were found to have contributed to the extant ranges of the species. Overall, we identified two different patterns during LGM: a) a continuous, large refugium, oftentimes used by many spadefoot toad species (*S. couchii*, *S. hurterii*, *S. bombifrons*, *S. multiplicata*), in the Rio Negro river basin, western coast of Mexico or by single species in the northern Rocky Mountains (in case of *S. intermontana*) or in Florida and south-eastern USA (in case of *S. holbrookii*) (Fig. 2); and b) refugia sparsely distributed in the Rocky Mountains, and along the California coast (in case of *S. intermontana* and *S. hammondi*) (Fig. 2).

Identifying the likely migration corridors used to migrate from LGM refugia to extant ranges

From the selected LGM refugia, we estimated the corridors used by each species to migrate to their extant ranges (Fig. 1–3). We found that each species has a unique number of high probability corridors and that few of these corridors are overlapping (Fig. 2). *Scaphiopus couchii*, *S. bombifrons* and *S. multiplicata* had overlapping corridors in western Texas on their way north into the Great Plains (Fig. 2). *Scaphiopus holbrookii* used major migration corridors northwards from the glacial refugium located in the southeastern USA along the Atlantic and Mexican Gulf Coast (Fig. 2). Other minor corridors used by the *S. holbrookii* to pass the Appalachian Mountains into the Midwest are also found along major rivers (Fig. 2). *Scaphiopus hurterii* used a few corridors from its refugium in eastern Texas to inhabit the rest of the range in the southern Great Plains (Fig. 2). *Spea intermontana* had multiple small corridors from multiple glacial refugia. *Spea hammondi* is clustered into two groups; the northern group likely never met the southern group as the two had separate glacial refugia and distinct corridors along the Pacific coast (Fig. 2). The northern group likely dispersed from one large refugia, whereas the southern group likely come from multiple small refugia (Fig. 2).

Testing dispersal patterns using genetic data: a case study in *Spea hammondi*

We found that our predictions (LGM potential distributions, corridor, and dispersal modeling and the influence of each LGM) agreed with the existing population genetic data for *S. hammondi* (Fig. 3). More specifically, our models predicted that *S. hammondi*'s distribution is split in two, generally characterized by a northern and a southern cluster that each used different glacial refugia and recolonized their range after glaciation using different corridors resulting in minimal overlap between the two groups, in agreement with genetic data (Fig. 2, 3).

Discussion

The role of dispersal is often overlooked in ecological niche models (Guisan and Zimmermann 2000, Waltari et al. 2007,

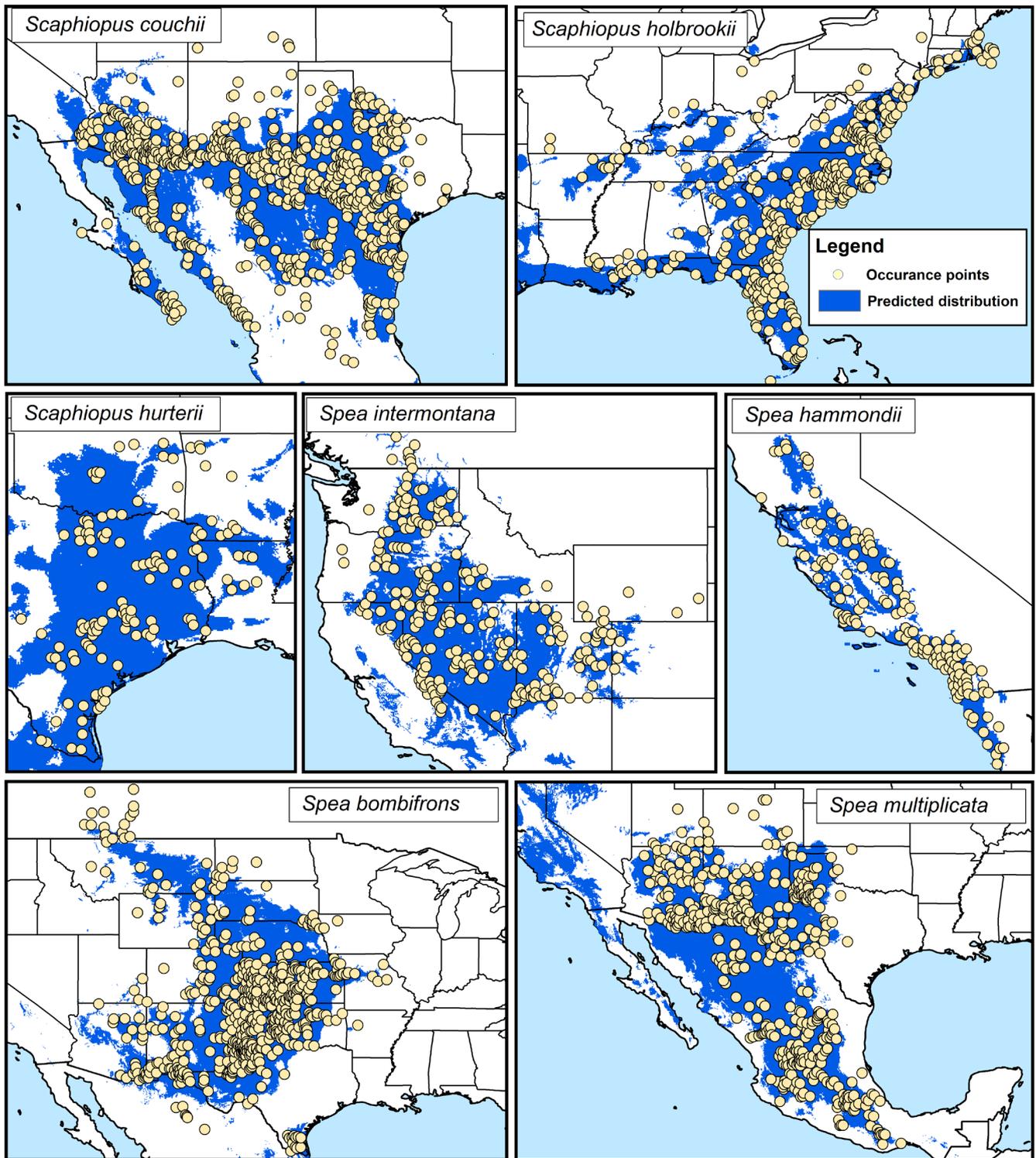


Figure 1. Current predicted distributions of the extant North American spadefoot toads.

Barve et al. 2011, Peterson et al. 2011, Holloway et al. 2016, Qiao et al. 2017). Currently, methods for estimating ENMs are rarely combined with dispersal analyses (Holloway et al. 2016), leading to an inaccurate estimation of the factors that shape species distributions. Hence, combining the

two techniques – ENMs and dispersal analyses – are crucial in developing more realistic models (Barve et al. 2011, Qiao et al. 2017). Here, we used ENMs and dispersal analyses to study the range dynamics of species following the Last Glacial Maximum by estimating the current and past

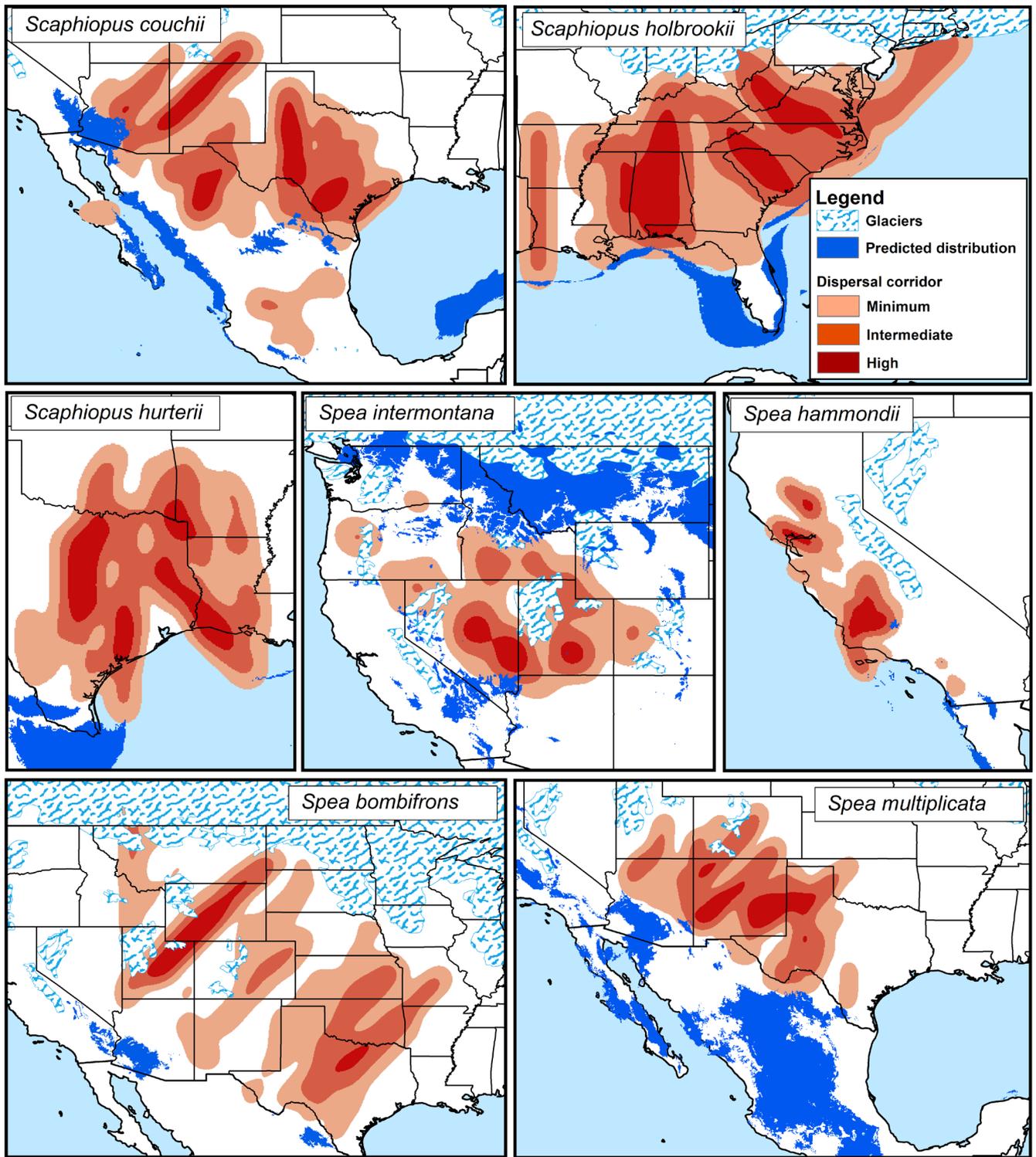


Figure 2. Past distributions of the extant North American spadefoot toads during the Last Glacial Maximum, and the corridors used by the species to expand their ranges during the Holocene. Dispersal corridors are mapped according to the density of least-cost paths in a given area. Intermediate and high corridors show that the majority of least-cost paths are passing a given area and were likely used as a corridor during the dispersal from the Last Glacial Maximum to their current extant range.

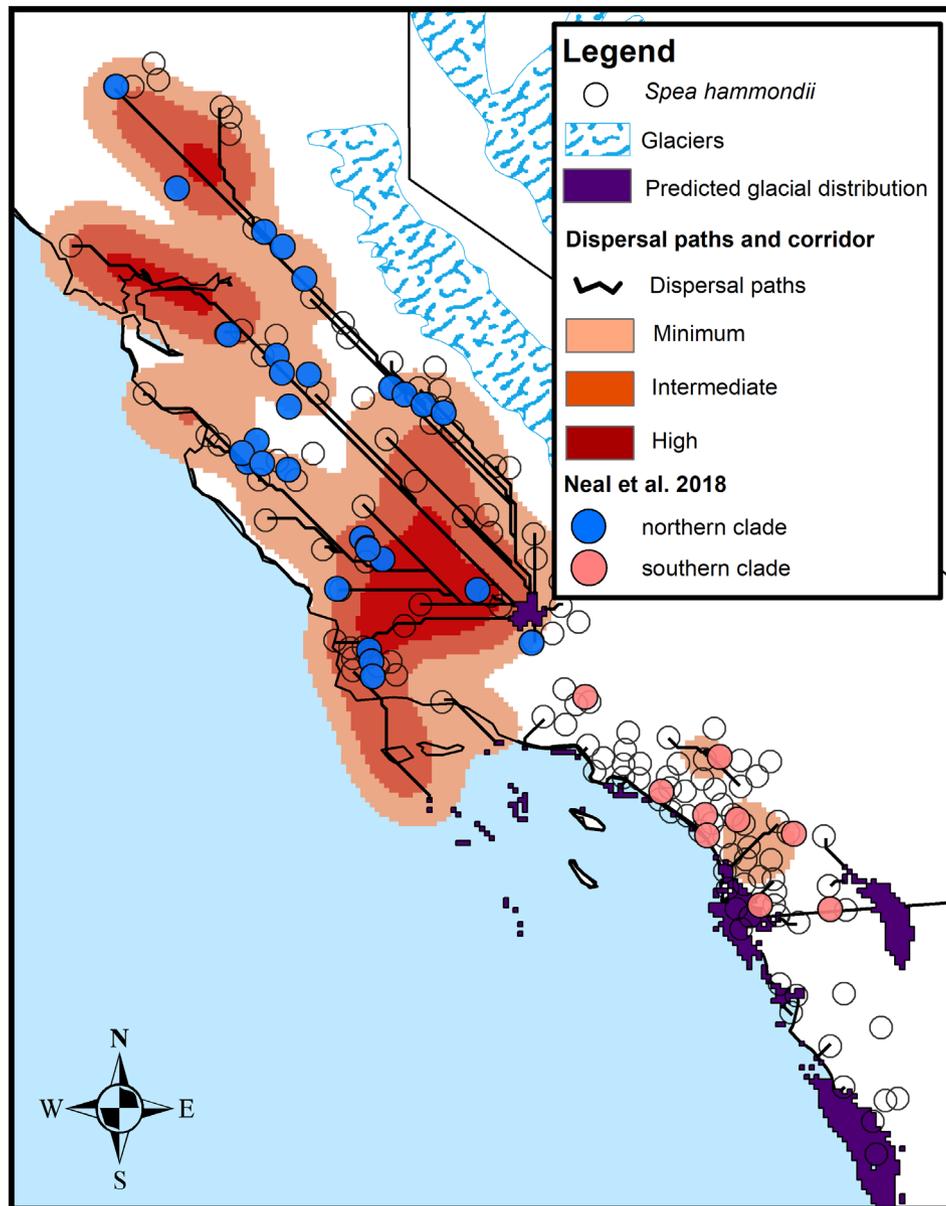


Figure 3. The congruence between current population genetic structure and predicted dispersal from refugia after the LGM for *S. hammondii*. The purple area is the predicted distribution during the last glacial maximum, the black polylines are the individual least-cost paths from the glacial refugia to the known present occurrences of the species. Filled yellow points are the present species occurrence points used for constructing the ENMs. Population genetic sampling data identifying a northern clade (blue filled points) and southern clade (pink filled points) are from Neal et al. (2018).

distributions for the North American spadefoot toads and the corridors used by these species to disperse and inhabit their current range (Fig. 2, 3). For one species (*S. hammondii*), we tested these predictions against previously published phylogeography data (Neal et al. 2018). This comparison revealed a high degree of agreement between the current population genetic structure of *S. hammondii* and the past range dynamics predicted by our models (Fig. 3). Therefore, we suggest that combining ENMs and dispersal modeling represents a promising approach for studying species' range dynamics and for generating hypotheses of the biogeographic past.

Species distribution models: past and present distributions

Our models identified the core of the current distributions of North American spadefoot toads (Fig. 1). Despite all of our models fitting the data well (AUC > 0.75), the model for *S. hurterii* and *S. intermontana* exhibited lower fit compared with other top-performing models (such as the model for *S. couchii*) (Table 1). This often happens when the occurrence data for the species is scarce in some parts of the species range (Franklin 2009, Peterson et al. 2011).

However, together these models recovered the general patterns of North American spadefoot species distributions. Interestingly, we found that, by the mid-Holocene, the climate was suitable for all of the North American spadefoot toads to an extent similar to that of the current species' ranges. For *S. holbrookii*, our models suggest the species inhabited a large Atlantic and Mexican Gulf coast refugium located in the southeastern USA and then expanded north along Atlantic Coast, the Mississippi river and other major rivers (Fig. 2). In the case of *S. hurterii*, the glacial refugium was probably in current day northern Mexico, and the species migrated north along the Great Plains. *Scaphiopus couchii* on the other hand likely had multiple glacial refugia, isolated by the Sierra Madre Occidental and Oriental Mountains during the LGM, and another glacial refugium in Baja California. This scenario suggests that the range of this species was fragmented by the LGM, leading to the isolation of these populations. Similarly, *S. hammondi* and *S. intermontana* had a series of multiple refugia, which should result in substantial genetic structure among populations (Wiens and Titus 1991, Neal et al. 2018) (Fig. 2, 3). Indeed, in the case of *S. hammondi*, our predictions of alternative refugia match well with the limited historic gene flow between the northern and southern populations, which have been recently proposed as distinct species (García-Paris et al. 2003, Neal et al. 2018). Similarly, we might predict that the multiple refugia identified in the Great Basin, as well as the Rocky Mountains, led to limited gene flow within *S. intermontana* (Fig. 2, 3). This finding is also supported by earlier work which found high genetic divergence between eastern and western *S. intermontana* populations (García-Paris et al. 2003). In the case of *S. bombifrons*, the predicted glacial refugium of the species was in Arizona. While previous molecular work suggests that the glacial refugium of *S. bombifrons* was in Kansas and Oklahoma (Rice and Pfennig 2008), these areas correspond with the most intensive sampling in their study, which could miss other centers of high diversity indicative of past refugia. Moreover, our models suggest that the isolated population of *S. bombifrons* from the southern part of Texas have been isolated from the rest of its range for at least 20 k yr (Fig. 2, 3). This suggests that those populations might have used a different glacial refugium compared with the rest of the range. Our models for *S. multiplicata* on the other hand, found that the species inhabited the same range during the LGM as the extant range of the species.

Using population genetic data to test our results

Incorporating multiple lines of evidence to test estimates from ENMs is highly desirable (Franklin 2009, Peterson et al. 2011). Our model estimates identified glacial refugia used by the North American spadefoot toads during the LGM (Fig. 2). By incorporating dispersal patterns and corridor modeling, we were able to create predictions of the paths these species used to colonize the continent after the climate became suitable (Fig. 2). We used previously published molecular

phylogeography results (Neal et al. 2018) to test our models using independent genetic data. Previously, *S. hammondi* has been found to have a high degree of genetic divergence between their southern and northern populations (García-Paris et al. 2003, Neal et al. 2018) (Fig. 3). Estimated corridor models confirm the presence of the two clusters and suggest that they used different migration corridors to disperse from their respective (distinct) glacial refugia. The comparison between our estimates and the published population genetic results showed a high degree of congruence, suggesting that our models are reliable and that our technique accurately predicts within-range species patterns associated with dispersal and niche dynamics (Fig. 3).

The role of dispersal

The distribution of any species is a complex interaction between its ecological niche (including physiological tolerances) and its evolutionary history (Brown et al. 1996). Dispersal also shapes species' distributions, as they must colonize suitable, available geographic space (MacArthur 1984, Pulliam 1988, Pulliam 2000). As a consequence, incorporating dispersal into the estimates of species distributions generated using ENMs projected into geographical space is important (Barve et al. 2011, Holloway et al. 2016). Nobis and Normand (2014) introduced a method to model accessible areas over time that further developed the field. However, estimating migration routes used by species during range shifts over time is still a topic that needs to be further studied in species distribution modeling. Another application performed by Gherghel and Papeş (2015) based solely on dispersal models (no ENMs or genetic data were considered) estimated the connectivity and migration corridors between all known populations of Danube crested newt *Triturus dobrogicus*. These models have since proven to be consistent with the genetic population structure of the species (Vörös et al. 2016, Wielstra et al. 2016). Based on our results and the lines of evidence from other studies, we can conclude that combining niche modeling with dispersal modeling can improve range dynamics that are influenced by climate and dispersal. When available, these models should be supported and tested against genetic and/or fossil data to avoid potential biases in predictions (Davis et al. 2014), and we encourage further research into combining ENMs and dispersal models.

Data availability statement

Data available from the Dryad Digital Repository: <<https://doi.org/10.5061/dryad.f4qrfj6td>> (Gherghel and Martin 2020).

Acknowledgements – We would like to thank S. Diamond, M. Dugas, L. Chick, A. Lenard, A. Perez, and M. Moore, J. C. Svenning, and two anonymous reviewers for helpful feedback that greatly improved the manuscript.

References

- Araujo, M. B. et al. 2019. Standards for distribution models in biodiversity assessments. – *Sci. Adv.* 5: eaat4858.
- Bai, Y. et al. 2018. Distributional dynamics of a vulnerable species in response to past and future climate change: a window for conservation prospects. – *PeerJ* 6: e4287.
- Barve, N. et al. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. – *Ecol. Model.* 222: 1810–1819.
- Beaumont, L. J. et al. 2005. Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. – *Ecol. Model.* 186: 251–270.
- Boria, R. A. et al. 2017. A single-algorithm ensemble approach to estimating suitability and uncertainty: cross-time projections for four Malagasy tenrecs. – *Divers. Distrib.* 23: 196–208.
- Brown, J. H. et al. 1996. The geographic range: size, shape, boundaries, and internal structure. – *Annu. Rev. Ecol. Syst.* 27: 597–623.
- Brown, J. L. and Anderson, B. 2014. SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. – *Methods Ecol. Evol.* 5: 694–700.
- Chan, L. M. et al. 2011. Integrating statistical genetic and geospatial methods brings new power to phylogeography. – *Mol. Phylogenet. Evol.* 59: 523–537.
- Collins, W. et al. 2006. The community climate system model version 3 (CCSM3). – *J. Clim.* 19: 2122–2143.
- Davis, E. B. et al. 2014. Ecological niche models of mammalian glacial refugia show consistent bias. – *Ecography* 37: 1133–1138.
- Dormann, C. F. et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. – *Ecography* 36: 27–46.
- Duellman, W. E. and Trueb, L. 1986. *Biology of amphibians*. – JHU Press.
- Dullinger, S. et al. 2012. Post-glacial migration lag restricts range filling of plants in the European Alps. – *Global Ecol. Biogeogr.* 21: 829–840.
- Dyderski, M. K. et al. 2018. How much does climate change threaten European forest tree species distributions? – *Global Change Biol.* 24: 1150–1163.
- Engler, R. and Guisan, A. 2009. MigClim: predicting plant distribution and dispersal in a changing climate. – *Divers. Distrib.* 15: 590–601.
- Franklin, J. 2009. *Mapping species distributions: spatial inference and prediction*. – Cambridge Univ. Press.
- García-Paris, M. et al. 2003. Phylogenetic relationships of Pelobatoidae re-examined using mtDNA. – *Mol. Phylogenet. Evol.* 28: 12–23.
- Gherghel, I. and Papeş, M. 2015. Landscape as a determinant of dispersal patterns and population connectivity in a newt species. – *Ecol. Inform.* 28: 1–6.
- Gherghel, I. and Martin, R. 2020. Data from: Postglacial recolonization of North America by spadefoot toads: integrating niche and corridor modeling to study species' range dynamics over geologic time. – Dryad Digital Repository, <<https://doi.org/10.5061/dryad.f4qrfj6td>>.
- Gómez-Ruiz, E. P. et al. 2017. Modelling the potential geographic distribution of an endangered pollination corridor in Mexico and the United States. – *Divers. Distrib.* 23: 67–78.
- Gray, L. K. et al. 2017. Potential effects of climate change on members of the palaeotropical pitcher plant family Nepenthaceae. – *PLoS One* 12: e0183132.
- Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – *Ecol. Model.* 135: 147–186.
- Hand, D. J. and Till, R. J. 2001. A simple generalisation of the area under the ROC curve for multiple class classification problems. – *Mach. Learn.* 45: 171–186.
- Hasumi, H. 2007. K-1 coupled GCM (MIROC) description. – Center for Climate System Research, Univ. of Tokyo.
- HerpMapper 2019. HerpMapper – a global herp atlas and data hub. – HerpMapper, <www.herpMapper.org>.
- Hijmans, R. J. and Graham, C. H. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. – *Global Change Biol.* 12: 2272–2281.
- Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – *Int. J. Climatol.* 25: 1965–1978.
- Holloway, P. et al. 2016. Incorporating movement in species distribution models: how do simulations of dispersal affect the accuracy and uncertainty of projections? – *Int. J. Geogr. Inf. Sci.* 30: 2050–2074.
- Huang, J. et al. 2017. Projecting the range shifts in climatically suitable habitat for Chinese Sea Buckthorn under climate change scenarios. – *Forests* 9: 9.
- IUCN 2016. The IUCN red list of threatened species. Version 2014.1. – IUCN, <www.iucnredlist.org>.
- Jiménez-Valverde, A. 2012. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. – *Global Ecol. Biogeogr.* 21: 498–507.
- Lantschner, M. V. et al. 2017. Predicting North American Scolytinae invasions in the southern hemisphere. – *Ecol. Appl.* 27: 66–77.
- Loera, I. et al. 2012. Diversification in North American arid lands: niche conservatism, divergence and expansion of habitat explain speciation in the genus *Ephedra*. – *Mol. Phylogenet. Evol.* 65: 437–450.
- MacArthur, R. H. 1984. *Geographical ecology: patterns in the distribution of species*. – Princeton Univ. Press.
- Margules, C. and Austin, M. 1990. *Nature conservation: cost effective biological surveys and data analysis*. – CSIRO Publishing.
- McRae, B. H. et al. 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. – *Ecology* 89: 2712–2724.
- Miller, J. A. and Holloway, P. 2015. Incorporating movement in species distribution models. – *Prog. Phys. Geogr.* 39: 837–849.
- Mouton, A. M. et al. 2010. Ecological relevance of performance criteria for species distribution models. – *Ecol. Model.* 221: 1995–2002.
- Mutascio, H. E. et al. 2017. Modeling relative habitat suitability of southern Florida for invasive Burmese pythons (*Python molurus bivittatus*). – *Landscape Ecol.* 33: 257–274.
- Neal, K. M. et al. 2018. Genetic structure and environmental niche modeling confirm two evolutionary and conservation units within the western spadefoot (*Spea hammondi*). – *Conserv. Genet.* 19: 937–946.
- Nix, H. and Busby, J. 1986. BIOCLIM, a bioclimatic analysis and prediction system. Annual report CSIRO. – CSIRO Division of Water and Land Resources, Canberra.
- Nobis, M. P. and Normand, S. 2014. KISSMig – a simple model for R to account for limited migration in analyses of species distributions. – *Ecography* 37: 1282–1287.
- Normand, S. et al. 2011. Postglacial migration supplements climate in determining plant species ranges in Europe. – *Proc. R. Soc. B* 278: 3644–3653.

- Peterson, A. T. et al. 2011. Ecological niches and geographic distributions. – Princeton Univ. Press.
- Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – *Ecol. Model.* 190: 231–259.
- Psonis, N. et al. 2017. Hidden diversity in the *Podarcis tauricus* (Sauria, Lacertidae) species subgroup in the light of multilocus phylogeny and species delimitation. – *Mol. Phylogenet. Evol.* 106: 6–17.
- Psonis, N. et al. 2018. Resolving complex phylogeographic patterns in the Balkan Peninsula using closely related wall-lizard species as a model system. – *Mol. Phylogenet. Evol.* 125: 100–115.
- Pulliam, H. R. 1988. Sources, sinks and population regulation. – *Am. Nat.* 132: 652–661.
- Pulliam, H. R. 2000. On the relationship between niche and distribution. – *Ecol. Lett.* 3: 349.
- Qiao, H. et al. 2017. Accessible areas in ecological niche comparisons of invasive species: recognized but still overlooked. – *Sci. Rep.* 7: 1213.
- Rice, A. and Pfennig, D. 2008. Analysis of range expansion in two species undergoing character displacement: why might invaders generally ‘win’ during character displacement? – *J. Evol. Biol.* 21: 696–704.
- Sahlean, T. C. et al. 2014. Refining climate change projections for organisms with low dispersal abilities: a case study of the Caspian whip snake. – *PLoS One* 9: e91994.
- Schmitt, S. et al. 2017. SSDM: an R package to predict distribution of species richness and composition based on stacked species distribution models. – *Methods Ecol. Evol.* 8: 1795–1803.
- Sehgal, R. N. et al. 2011. Spatially explicit predictions of blood parasites in a widely distributed African rainforest bird. – *Proc. R. Soc. B* 278: 1025–1033.
- Sillero, N. et al. 2014. Niche evolution and thermal adaptation in the temperate species *Drosophila americana*. – *J. Evol. Biol.* 27: 1549–1561.
- Skourtanioti, E. et al. 2016. A reinvestigation of phylogeny and divergence times of the *Ablepharus kitaibelii* species complex (Sauria, Scincidae) based on mtDNA and nuDNA genes. – *Mol. Phylogenet. Evol.* 103: 199–214.
- Stöck, M. et al. 2012. Cryptic diversity among Western Palearctic tree frogs: postglacial range expansion, range limits, and secondary contacts of three European tree frog lineages (*Hyla arborea* group). – *Mol. Phylogenet. Evol.* 65: 1–9.
- Svenning, J. C. et al. 2008. Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. – *Ecography* 31: 316–326.
- Taylor, A. T. et al. 2018. Incorporating fragmentation and non-native species into distribution models to inform fluvial fish conservation. – *Conserv. Biol.* 32: 171–182.
- Tzedakis, P. C. et al. 2013. Cryptic or mystic? Glacial tree refugia in northern Europe. – *Trends Ecol. Evol.* 28: 696–704.
- Veloz, S. D. et al. 2012. No-analog climates and shifting realized niches during the late quaternary: implications for 21st-century predictions by species distribution models. – *Global Change Biol.* 18: 1698–1713.
- Vörös, J. et al. 2016. Phylogeographic analysis reveals northerly refugia for the riverine amphibian *Triturus dobrogicus* (Caudata: Salamandridae). – *Biol. J. Linn. Soc.* 119: 974–991.
- Waltari, E. et al. 2007. Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. – *PLoS One* 2: e563.
- Wang, R. et al. 2018. Modeling and mapping the current and future distribution of *Pseudomonas syringae* pv. *actinidiae* under climate change in China. – *PLoS One* 13: e0192153.
- Wang, Y.-H. et al. 2008. Habitat suitability modelling to correlate gene flow with landscape connectivity. – *Landscape Ecol.* 23: 989–1000.
- Warren, D. L. et al. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. – *Ecography* 33: 607–611.
- Wauchope, H. S. et al. 2017. Rapid climate-driven loss of breeding habitat for Arctic migratory birds. – *Global Change Biol.* 23: 1085–1094.
- Wielstra, B. and Arntzen, J. W. 2012. Postglacial species displacement in *Triturus* newts deduced from asymmetrically introgressed mitochondrial DNA and ecological niche models. – *BMC Evol. Biol.* 12: 161.
- Wielstra, B. et al. 2016. Is the Danube crested newt *Triturus dobrogicus* polytypic? A review and new nuclear DNA data. – *Amphibia-Reptilia* 37: 167–177.
- Wiens, J. J. and Titus, T. A. 1991. A phylogenetic analysis of *Spea* (Anura, Pelobatidae). – *Herpetologica* 47: 21–28.
- Zeng, C. et al. 2014. Evolution of rapid development in spadefoot toads is unrelated to arid environments. – *PLoS One* 9: e96637.