

Acknowledging uncertainty in evolutionary reconstructions of ecological niches

Hannah Owens¹, Vivian Ribeiro², Erin Saupe³, Marlon E. Cobos⁴, Peter Hosner¹, Jacob Cooper⁵, Abdallah Samy⁶, Vijay Barve⁷, Narayani Barve⁷, Carlos Muñoz⁸, and A. Townsend Peterson⁴

¹University of Copenhagen

²Stockholm Environment Institute

³University of Oxford Department of Earth Sciences

⁴University of Kansas

⁵University of Chicago Committee on Evolutionary Biology

⁶Ain Shams University Faculty of Science

⁷University of Florida

⁸UNAM

April 23, 2020

Abstract

Reconstructing ecological niche evolution can provide insight into the biogeography and diversification of evolving lineages. However, comparative phylogenetic methods can infer the history of ecological niche evolution inaccurately because (1) species' niches are often poorly characterized; and (2) phylogenetic comparative methods rely on niche summary statistics rather than full estimates of species' environmental tolerances. Here we propose a new framework for coding ecological niches and reconstructing their evolution that explicitly acknowledges and incorporates the uncertainty introduced by incomplete niche characterization. Then, we modify existing ancestral state inference methods to leverage full estimates of environmental tolerances. We provide a worked empirical example of our method, investigating ecological niche evolution in the New World orioles (Aves: Passeriformes: *Icterus* spp.). Temperature and precipitation tolerances were generally broad and conserved among orioles, with niche reduction and specialization limited to a few terminal branches. Tools for performing these reconstructions are available in a new R package called *nichevol*.

Introduction

The environmental conditions under which species thrive (i.e. their fundamental ecological niches defined in abiotic environmental dimensions) evolve over time. The frequency and speed at which niches evolve in speciating lineages remains a key question in evolutionary biology (e.g., Peterson et al. 1999; Graham et al. 2004; Knouft et al. 2006; Losos 2008; Pearman et al. 2008; Evans et al. 2009; Vieites et al. 2009; Nyári and Reddy 2013; Owens et al. 2017; García-Navas and Rodríguez-Rey 2018). Methods for estimating fundamental ecological niches (Peterson et al. 2011; Hijmans and Elith 2015) and inferring macroevolutionary patterns from phylogenies (Swofford and Maddison 1987; Lanyon 1993; Freckleton et al. 2002; Pagel et al. 2004; O'Meara 2012; Revell 2012) have both advanced greatly in recent decades. These developments have facilitated a paradigm shift toward investigating biogeographic history in the context of reconstructed ancestral ecological niche characteristics (e.g., Rice et al. 2003; Graham et al. 2004; Knouft et

al. 2006; Pearman et al. 2008; Anciães and Peterson 2009; Evans et al. 2009; Vieites et al. 2009; Smith and Donoghue 2010; Nyári and Reddy 2013; Ribeiro et al. 2016; Owens et al. 2017). Still, modeling complex traits and their evolution remains a major challenge, and indeed reconstructing the evolution of abiotic ecological niches is particularly difficult.

Researchers have used several approaches to characterize ecological niches when attempting to reconstruct their evolutionary history. Studies have used means and standard errors of suitable abiotic niche characteristics (Rice et al. 2003; Anciães and Peterson 2009), minimum and maximum suitable abiotic niche values (Graham et al. 2004; Yesson and Culham 2006), central tendencies of suitable niche values (i.e. mean or median; Ackerly et al. 2006; Kozak and Wiens 2010; Cooper et al. 2011), and distributions of suitable niche values (Evans et al. 2009; Smith and Donoghue 2010). These data were derived either directly from the occurrence data (e.g. Ackerly et al. 2006; Kozak and Wiens 2010; Cooper et al. 2011) or from ecological niche model outputs (e.g. Rice et al. 2003; Smith and Donoghue 2010; Nyári and Reddy 2013). These approaches fit existing ancestral state reconstruction methodology relatively well, but at the cost of simplifying complex niches to summary statistics for each species.

Fundamental ecological niches, furthermore, are rarely characterized completely and unambiguously when they are estimated for real-world species on real landscapes, owing to biases and limitations in environmental conditions available across accessible areas of geographic space (Fig. 1; Saupe et al. 2012; Veloz et al. 2012; Owens et al. 2013; Guisan et al. 2014; Warren et al. 2014; Saupe et al. 2017). The fundamental ecological niche of a species is defined as the set of conditions under which it is able to maintain populations without immigrational input (Soberón 2007), and is the result of phenotypic traits subject to natural selection (Peterson 2011). However, the full suite of environmental conditions within a species’ fundamental niche is not necessarily represented on Earth, or across areas that are accessible to a species. This subset of the fundamental ecological niche that is present in geographic space at the time period of interest and is accessible to the species is referred to as the existing niche (Barve et al. 2011). A species’ realized ecological niche (i.e. environments where the species is observed) is determined by the further reduction of the existing niche by biotic factors such as competition, parasitism, etc. (Soberón 2007).

Figure 1. Uncertainty when characterizing fundamental niche based on realized niche. A. Map of simulation study area, distribution of suitable habitat for a simulated species (black), its accessible area (M; light gray), and a plot showing ranges of mean annual temperature present in each of these areas. Note that the species’ lower temperature limit is found within M, whereas its upper limit is not. B. Example of coding abiotic niche characteristics of species mapped in Fig. 1A, accounting for uncertainty. Bins with values suitable to the species are marked “1”. The lower limit is found within the species’ accessible area (M), so coding of bins below suitable range are coded without uncertainty (“0”). The upper limit is not found within the species’ accessible area (M), so coding of bins above known suitable range are coded as uncertain (“?”).

As such, any characterization of fundamental ecological niches that relies on inference from species’ geographic distributions (i.e. realized niche) will be incomplete (Saupe et al. 2012; Owens et al. 2013). A species’ fundamental niche becomes particularly difficult to approximate from its realized niche when its geographic range approaches the limits of the area to which it can disperse (as may be the case for many island endemic species; Saupe et al. 2012). Hence, although the estimated niche of a lineage through time may show variation in response to inherited adaptations that alters the lineage’s fundamental niche, that variation may also derive from changes in the set of environments accessible to that lineage, which do *not* represent a genetically-inherited set of adaptations or changes in the fundamental ecological niche (Araújo et al. 2013).

Methodologies that use estimates based on species’ realized niches to characterize ecological niches in phylogenetic analyses are known to overestimate true amounts of niche change (Saupe et al. 2017). Here, we

present a new framework to characterize species’ niches, which incorporates consideration of areas accessible to the species over relevant time periods (referred to as \mathbf{M} ; Soberón and Peterson, 2005; Phillips et al. 2009; VanDerWal et al. 2009; Barve et al. 2011;). Estimating and accounting for this accessible region has been recognized as important when generating niche or distribution models that use background or pseudo-absence data for calibration (Phillips et al. 2009; Elith et al. 2010; Barve et al. 2011). If regions accessible to a species are ignored when selecting the geographic extent for model calibration, fitted models may erroneously estimate suitable niche conditions. However, even niche estimates derived from presence data (i.e., without a modelling component) should consider \mathbf{M} , as doing so provides one of the only ways to assess in which cases niche estimates are likely to be truncated. Specifically, when environments across \mathbf{M} do not encompass conditions beyond those under which the species in question is observed, no evidence is available regarding the environmental limits of the species (Fig. 1b).

Our new binned-range (BR) character-coding method decomposes the broader environment occupied by and accessible to a clade into discrete bins, and scores each bin as suitable, unsuitable, or uncertain for a given species (Fig. 1b), thereby accounting for potential cases of incomplete niche characterization. We illustrate the utility of summarizing species’ niches in this way via simulation that compares ancestral niche reconstructions based on BR coding (Binned Ancestral Range; BAR) to those estimated using a more traditional analysis (generalized least-squares reconstructions based on the median suitable value of a variable for each species based on its realized niche). We demonstrate the utility of our approach with an empirical example, inferring patterns of ecological niche evolution in New World orioles (*Icterus* spp.; see, e.g., Fig. 2). This empirical example highlights the utility of BAR reconstructions in terms of incorporating uncertainty explicitly and considering species’ ecological niches as a ranged response instead of as a single value.

Figure 2. White-edged Oriole (*Icterus graceannae*) in Macara, Ecuador (ML202492981). This striking species is endemic to the Tumbesian dry forests of northwestern Peru and southwestern Ecuador, where its specialization on drier environments is example of niche reduction identified by our new methods. Photography credit: David M. Bell.

Methods

Coding ecological niches for analysis

Coding full niche ranges represents an initial challenge for ancestral niche reconstructions. We first determine relevant analytical limits with respect to single environmental dimensions as the minimum and maximum values within the union of all accessible areas polygons for the species comprising the clade of interest. Next, we parse this range of values into equal-width bins; for each species, bins with values falling within the range of environmental conditions occupied by the species are coded as “suitable” (1). Bins with values represented within the species’ \mathbf{M} but falling outside the range of environmental conditions occupied by the species are scored as “unsuitable” (0). In cases in which suitable niche conditions coincide with the limits of environmental conditions present in a species’ \mathbf{M} , all more extreme values—that is, values more extreme than those manifested within \mathbf{M} (e.g. Fig. 1)—are coded as unknown (?). This procedure allows explicit incorporation of uncertainty in our analyses. When suitability is unknown because the climatic values for the bin in question were not represented within the \mathbf{M} of the species, but the bin in question was flanked by two suitable bins, it is also scored as “suitable,” under an assumption of a unimodal response to environmental conditions (Maguire 1973). These steps can be achieved using *nichevol* v.0.1.17 (Cobos et al. 2020), an R (v.3.6.1; R Core Team 2019) package we created to facilitate studies of niche evolution, including many of the analyses presented in this paper. Package documentation includes a tutorial demonstrating an analytical workflow implementing *nichevol*.

Binned ancestral range reconstruction demonstration

As a simple illustration of how this analysis works, we simulated data for a scenario of a shift from an ancestral warm niche to a cool niche, to demonstrate the ability of our new method to identify instances of niche evolution. First, we simulated distributions and accessible areas for 1000 species across South America—500 with a fundamental mean annual temperature niche of 24–28°C (hereafter referred to as “cool niche species”) and 500 with a fundamental mean annual temperature niche of 25–29°C (hereafter referred to as “warm niche species”). For each set of species, accessible areas were generated using an initial population of 10,000 random polygons using *nichevol*. We assumed that each species could occupy all suitable cells within its corresponding **M** (i.e. we ignored biotic factors). Suitable cells were identified based on a 2.5’-resolution raster of annual mean temperature data (Bio 1) from WorldClim v.1.4 (Hijmans et al. 2005) and the *raster* package v.3.0.0 (Hijmans 2019) in R; simulated **M**s with no suitable conditions were removed ($n = 1$). The median suitable mean annual temperature value for each simulated species was calculated from suitable cell values within its **M**. We then determined the minimum and maximum mean annual temperature within the union of all simulated **M** polygons and parsed this range of values into equal-width, 1°C bins using *nichevol* in R. Raw and annotated R code the niche simulation and coding steps, as well as input and output data, can be found in supplementary materials provided via Dryad (Code Supplement 1, Annotated Code Supplement 1, Data Package 1; DOI: 10.5061/dryad.c866t1g3j).

We generated a single, 15-taxon stochastic birth-death tree (birth rate = 1, death rate = 0) using the R package *phytools* v.0.6-99 (Revell, 2012) and assigned simulated species from the cool niche group to a monophyletic clade of 7 taxa. The remaining tips in the tree were assigned simulated species from the warm niche simulation group. We then used *nichevol* tools to perform BAR reconstructions using maximum parsimony (as implemented in *castor* v.1.4.3; Louca and Doebeli 2018) and maximum likelihood (as implemented in *ape* v.5.3; Paradis and Schliep 2018). For both algorithms, ancestral state reconstructions were performed for each bin separately, treating bin scores (including “uncertain”) as discrete characters under an equal transition rate model of evolution. Results were smoothed such that reconstructed suitable ancestral niche bins at each node were not interrupted by unsuitable bins, following the assumption of a unimodal response to environmental conditions (Maguire, 1973), and accounting for evolutionary non-independence of bins. Raw and annotated R code for these steps, as well as input and output, can be found in supplementary materials provided via Dryad (Code Supplement 2, Annotated Code Supplement 2, Data Package 1; DOI: <https://doi.org/10.5061/dryad.c866t1g3j>).

We note that we have kept this initial example simple for the purpose of illustration—many improvements could be made to this methodology, such as implementation of different character evolution models, Bayesian approaches in inferring ancestral character states, stochastic character mapping, and consideration of joint effects of environmental dimensions (e.g. temperature, precipitation) that are here considered independently. Furthermore, phylogenetic comparative methods are notoriously “data-hungry,” and BAR reconstructions will benefit from further detailed simulation-based examinations in the future. Our purpose here is to illustrate the crucial importance of incorporating uncertainty explicitly in the inference of abiotic ecological niche evolution patterns.

Oriole analyses

We next used BAR reconstructions to infer patterns of niche evolution in 34 species of New World orioles (genus *Icterus*). We used the single best ultrametric maximum likelihood phylogeny from Powell et al. (2014; their Figure 4) inferred from mitochondrial and nuclear DNA sequences. Distributional data for each species were drawn from the Global Biodiversity Information Facility (GBIF, 2018), a large portion of which were derived from eBird (Supplemental Table 1, Data Package 2; via Dryad, DOI: <https://doi.org/10.5061/dryad.c866t1g3j>). We removed all records lacking geographic coordinates, and inspected those remaining with respect to known ranges of species based on expert assessment by four ornitho-

logists (authors Cooper, Hosner, and Peterson), removing records that reflected errors or outdated taxonomic arrangements. Species-specific hypotheses of areas accessible to the species (\mathbf{M}) were developed based on the biotic attributes and biogeographic history of the clade (Barve et al. 2011; Elith et al. 2010). That is, the ornithologists inspected patterns of occurrences for each species and outlined accessible-area hypotheses based on known barriers to dispersal (i.e., oceans, high mountain ranges, the Amazon River, deserts). While this step remains subjective, it is crucial to a realistic representation of the environments that should be considered within the species’ potential distribution (Phillips et al. 2009; Barve et al. 2011).

We then used BR to score species’ niches, explicitly scoring the parts of these profiles that were not observable (i.e. at the periphery of \mathbf{M}) as uncertain (see above). For mean annual temperature (Bio1 in WorldClim v.1.4; Hijmans et al. 2005), we used 32 equal-width, 1°C bins (3-4°, 4-5°, ... 34-35°) across the full range of temperature values represented in the union of all species’ \mathbf{M} areas. For annual precipitation (Bio12 in WorldClim v1.4; Hijmans et al. 2005), we used 80, 10 mm-width bins to cover the range of precipitation values from 0 to 800 mm across all species’ \mathbf{M} areas. For comparison to more traditional methods of coding species’ niches, we calculated median values for mean annual precipitation and temperature across species’ known occurrences. As with our simulated species, we characterized species’ niches using R; raw and annotated R code for analyses, as well as inputs and outputs, can be found in supplementary materials provided via Dryad (Code Supplement 3, Annotated Code Supplement 3, Data Package 2; DOI: <https://doi.org/10.5061/dryad.c866t1g3j>).

Finally, we inferred the evolutionary history of oriole temperature and precipitation niches using both BAR, as described above, and GLS reconstructions using the median temperature and precipitation values at species occurrences. For GLS reconstructions, we first examined the fits of Brownian motion, Ornstein-Uhlenbeck, early burst, and diffusion with linear trend models of evolution. We then performed ancestral state reconstructions using a continuous-value maximum likelihood algorithm (as implemented in *reconstruct* in *ape*) under the best-fit evolutionary model (Ornstein-Uhlenbeck) for both mean annual temperature and annual precipitation. Raw and annotated R code for analyses, as well as input and output can be found in supplementary materials provided via Dryad (Code Supplement 4, Annotated Code Supplement 4, Data Package 2; DOI: <https://doi.org/10.5061/dryad.c866t1g3j>).

Results

Binned ancestral range reconstruction demonstration

Our BAR reconstructions detected simulated niche shifts; maximum likelihood reconstructions performed more reliably than parsimony reconstructions. In our simulated example, using maximum likelihood, we were able to recover the expansion from a 25°C ancestral lower fundamental niche limit to a 24°C ancestral lower fundamental niche limit at the most recent common ancestor of the 7 cool-niche simulated species (Annotated Code Supplement 1). However, the parsimony-based reconstruction failed to recover this change fully, but did show an expansion to 24°C for simulated species “t1”, a species with a higher maximum known suitability than the other warm niche species (Annotated Code Supplement 1). By comparison, the GLS reconstruction performed qualitatively worse, reconstructing shifts to a *warmer* niche for the simulated cool-niche species and their ancestors. This is likely due to biased estimates of species’ realized niches based on their existing niches—“t3”, a cool niche species, had a median suitable temperature of 26.6 °C, tied for the highest temperature in the clade with “t5”, a simulated warm niche species. Interestingly, BAR using parsimony reconstruction tended to infer more uncertain character states at the cooler ends of ancestral niches, whereas BAR using ML reconstruction inferred more uncertain character states at the warmer ends of ancestral niches. See Annotated Code Supplement 1 for further detail.

Application to oriole niche evolution

Large numbers of occurrence points were available for this clade, thanks to recent advances in biodiversity informatics and community-science initiatives regarding bird distributions (Supplemental Table 1, Data Package 2). Niche estimates for some oriole species were completely characterized with respect to **M**, including the temperature and precipitation dimensions for *Icterus fuertesi*, and the precipitation dimensions for *I. graceannae* and *I. galbula* (Supplementary Figures 2 and 3). That is, estimated limits of suitable conditions were contained completely within the environments available in **M**, and did not appear to be truncated. The majority of species, however, were estimated to have niche ranges flanked by unknown maxima and/or minima.

In general, BAR reconstructions of species' ecological niches in *Icterus* did not recover reduction or gain in inferred suitable niche space, suggesting broad-scale evolutionary stability (Figs. 3 and 4; Supplementary Tables 2-5, Annotated Code Supplement 4). For temperature, both ML and parsimony BAR reconstructed a consistent range of mean annual temperature niche values across all ancestral nodes (hereafter referred to as a "core conserved niche"), although some individual nodes had lower minimum or higher maximum suitable temperatures; the estimated core conserved niche was much broader for maximum likelihood (21–26°C) than parsimony (24–25°C; Supplementary Tables 2 and 3). For precipitation, maximum likelihood reconstructed a core conserved annual precipitation niche range of 71–240 mm, whereas parsimony-based reconstructions recovered no clear core conserved niche for precipitation (Supplementary Tables 4 and 5).

Figure 3. New World oriole (*Icterus* spp.) mean annual temperature niche state inference, characterized using bins and reconstructed with maximum likelihood, showing general niche conservatism. Left panel shows bin-based characterization of niches at tips and reconstructed bin-based values at nodes; bars show bin coding from highest bin at top to lowest bin at bottom. Right panel shows bin-based characterization of niches at tips, comparison with ancestors at nodes. See Annotated Code Supplement 4 for maximum parsimony reconstructions (Dryad DOI: <https://doi.org/10.5061/dryad.c866t1g3j>).

Figure 4. New World oriole (*Icterus* spp.) annual precipitation niche state inference, characterized using bins and reconstructed with maximum likelihood, showing general niche conservatism. Left panel shows bin-based characterization of niches at tips and reconstructed bin-based values at nodes; bars show bin coding from highest bin at top to lowest bin at bottom. Right panel shows bin-based characterization of niches at tips, comparison with ancestors at nodes. See Annotated Code Supplement 4 for maximum parsimony reconstructions (Dryad DOI: <https://doi.org/10.5061/dryad.c866t1g3j>).

Because GLS and BAR reconstructions generate quite-different outputs, direct comparisons are not possible; however, qualitatively, single-value reconstructions appear to evolve much more quickly than bin-based reconstructions, especially near the base of the tree and in groups with a preponderance of island endemics. GLS reconstructions based on median environmental values fell within ranges of values inferred using the bin-based method at every node for both environmental variables (Supplementary Tables 3 and 5). However, these reconstructions were only congruent with maximum parsimony BAR reconstructions for 24 of 33 nodes for mean annual temperature, and for none of the nodes for annual precipitation (Supplementary Tables 2 and 4).

Discussion

New methodology

This contribution derives from careful examination and analysis of the growing suite of papers analyzing niche evolution across phylogenies (e.g., Peterson et al. 1999; Wiens and Graham 2005; Knouft et al. 2006; Losos 2008; Evans et al. 2009; Vieites et al. 2009; Nyári and Reddy 2013; Meseguer et al. 2015). It is likely that fundamental niches and realized niches are rarely equivalent, owing to constraints imposed by the set of environments that can be observed within areas accessible to a species (**M**) (Soberón and Peterson, 2011). The limited environments present in areas accessible to species typically will add variation to niche estimates that will bias analyses of niche evolution toward concluding increased niche lability (Ribeiro et al. 2016; Saupe et al. 2017). In addition, use of summary statistics to characterize species’ niches introduces further variation related to the environmental vagaries of sampling, which has its own intrinsic biases (Kadmon et al. 2004) that are—again—reflected in the environmental signature of the occurrence data that derive from the process (Saupe et al. 2017).

Analyzing ecological niche change on a phylogenetic tree without considering uncertainty produces more concise conclusions and is easier to implement (e.g., calculating the median or mean of environmental values across all occurrences for a species and performing a single reconstruction calculation). However, previous studies indicate that this approach comes with a cost: niche change may be over- or under-estimated, introducing biases in reconstructing evolutionary change in niches through time (Ribeiro et al. 2016; Saupe et al. 2017). Our empirical example using orioles shows patterns that are qualitatively consistent with these findings: GLS reconstructions of ancestral node characteristics varied more near the base of the tree and in clades dominated by narrow-range endemic species with incompletely-characterized abiotic ecological niches than in clades with fewer narrow-range endemics.

Admittedly, we currently lack a method for quantitative assessment of niche evolution rates estimated from BAR reconstructions that is comparable to rates calculated for single continuous value reconstructions (especially in light of differences in how evolutionary models are applied and how rates are estimated for continuous and discrete characters). However, our BAR reconstructions appear to be qualitatively more robust to noise in the data introduced by narrow-range endemics that are incompletely characterized, as it recovered conserved ranges of suitable habitat for all basal *Icterus* lineages (except for parsimony reconstructions of precipitation). Furthermore, BR coding less likely to be skewed by instances of biased sampling, (i.e. a greater frequency of occurrences within a particular environmental range can skew niche estimates based on summary statistics). Indeed, a greater abundance of occurrences in particular environments may not be due to those environments being more suitable to a species than another suite of environments, but merely that those environments are more common within a species’ **M** or more likely to be sampled by researchers. This is illustrated by our simulation reconstruction, in which the median ancestral temperature for the cool-niche simulated species was inferred to be *warmer* owing to biased tip state characterizations.

Oriole niche evolution

The genus *Icterus* exists in many different environments, which suggests that the niches of these species have diversified. Indeed, when we look at patterns of niche evolution inferred using GLS, we found frequent apparent niche shifts across the phylogeny, particularly within clades dominated by island endemics (Annotated Code Supplement 4, Supplementary Tables 2-5). However, BAR reconstructions found little evidence of change in the inferred fundamental ecological niche across the phylogenetic history of the genus, particularly when reconstructions were done using the maximum likelihood algorithm. This pattern is consistent with the fact that species of *Icterus* that “left” the Tropics (i.e. migratory species) move into northern areas of North America in the breeding season only—a special case of niche conservatism termed “niche following” in previous work (Joseph and Stockwell 2000; Nakazawa et al. 2004).

The overall tendency across the history of the genus *Icterus* was one of remarkable niche stability, notwith-

standing the GLS results. Particularly invariant was the upper end of the temperature tolerance spectrum (Fig. 3; Supplementary Fig. 2, Tables 2 and 3; Annotated Code Supplement 4). This observation coincides with recent results from Araújo et al. (2013), who presented a meta-analysis that concluded that heat tolerance was much more constrained over evolutionary history than cold tolerance. Importantly, though, our proposed framework for characterizing ecological niches and subsequent ancestral niche inference may underestimate true amounts of niche evolution because the method only concludes niche change when explicit evidence exists, which we consider to be a desirable quality. Still, further detailed simulation study is needed to examine fully the sensitivity of our proposed methods to true niche evolution in the face of various biasing or obfuscating factors.

Focusing on maximum likelihood BAR reconstructions, which showed clearer patterns with less uncertainty than parsimony BAR reconstructions, we identified niche reductions for species that are relative habitat specialists within *Icterus*. *Icterus* orioles are a predominately lowland group, although some species occur in foothills and low montane regions adjacent their core lowland ranges. We identified reductions in high temperature tolerance for two species that specialize in Mesoamerican montane habitats, *I. abeillei* and *I. maculialatus*, and reductions in low temperature tolerance for two strictly lowland tropical species, *I. fuertesi* and *I. chrysocephalus*. *Icterus* orioles occupy a variety of forest types across a variety of precipitation regimes. However, for two species that specialize in dry forest, *I. auratus* of the Yucatán Peninsula and *I. graceannae* of the Tuzexian region, we identified suitable niches corresponding to reduced precipitation.

Conclusion

The challenge of understanding change in species' ecological niches across evolutionary history lies in characterizing the entirety of a species' niche. We present a simple methodology that directly incorporates knowledge gaps based on incomplete niche characterization. We see a number of next steps in developing this methodology further—specifically, developing *nichevol* tools to encompass Bayesian estimation approaches and considering alternative evolutionary models. We would also take into account the frequency of occurrence of environmental conditions across the accessible area of each species in making conclusions about niche limitations (e.g. Meyer and Pie 2018)—that is, non-occurrence in relatively rare environments should perhaps not be taken as evidence of niche limitation. Finally, we plan to develop a method for estimating the likely range of niche evolution rates encompassing uncertainty using our bin-based method. We are exploring implementation of these next steps in coming applications of this methodology.

Data Accessibility Statement

Analysis scripts, annotated HTML script reports, M polygons for both the virtual species and empirical case, final oriole occurrence datasets, and results of niche characterization are accessible via Dryad (DOI: <https://doi.org/10.5061/dryad.c866t1g3j>).

Competing Interests Statement

The authors have no competing interests.

Author Contributions

All authors contributed to the conceptualization, theory, and design of the study, and discussion of results, led by HLO. HLO, EES, ATP and VR wrote the manuscript with input from all authors. PAH, JCC, HLO, EES, and ATP prepared data for the empirical section of the study. HLO, VR, MEC, VB, NB, and ATP performed elements of analysis. MEC aided HLO with re-analysis of the original study, and wrote the *nichevol* R package.

Acknowledgments

We thank the members of the University of Kansas Ecological Niche Modeling Group for their support, assistance, and advice, during the development of this contribution. Luis Escobar provided help with NicheA, which was central to early explorations during this project. We also thank Richard Glor, Mark Holder, Daniel Reuman, and Jorge Soberón (University of Kansas) for their time and advice regarding complex analytical challenges.

References

- Ackerly, D. D., D. W. Schilck, and C. O. Webb. 2006. Niche evolution and adaptive radiation: testing the order of trait divergence. *Ecology* 87:S50-S61.
- Anciães, M., and A. T. Peterson. 2009. Ecological niches and their evolution among Neotropical manakins (Aves: Pipridae). *Journal of Avian Biology* 40:591-604.
- Araújo, M. B., F. Ferri-Yáñez, F. Bozinovic, P. A. Marquet, F. Valladares, and S. L. Chown. 2013. Heat freezes niche evolution. *Ecology Letters* 16:1206-1219.
- Barve, N., V. Barve, A. Jimenez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J. Soberón, and F. Villalobos. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling* 222:1810-1819.
- Cobos, M. E., Owens, H. L., and A. T. Peterson. 2020. *nichevol*: tools for ecological niche evolution assessment considering uncertainty. R package version 0.1.19. <https://CRAN.R-project.org/package=nichevol>.
- Cooper, N., R. P. Freckleton, and W. Jetz. 2011. Phylogenetic conservatism of environmental niches in mammals. *Proceedings of the Royal Society B* 278:2384-2391.
- Elith, J., M. Kearney, and S. J. Phillips. 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution* 1: 330-342.
- Evans, M. E., S. A. Smith, R. S. Flynn, and M. J. Donoghue. 2009. Climate, niche evolution, and diversification of the “bird-cage” evening primroses (Oenothera, sections Anogra and Kleinia). *American Naturalist* 173:225-240.
- Freckleton, R. P., P. H. Harvey, and M. Pagel. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. *American Naturalist* 160:712-726.
- Garcia-Navas, V., and M. Rodriguez-Rey. 2018. The evolution of climatic niches and its role in shaping diversity patterns in diprotodontid marsupials. *Journal of Mammalian Evolution*: <https://doi.org/10.1007/s10914-10018-19435-z>.

- GBIF. 2018. GBIF Occurrence Download. <https://doi.org/10.15468/dl.x7ygeq>. Accessed 30 July 2018.
- Graham, C. H., S. R. Ron, J. C. Santos, C. J. Schneider, and C. Moritz. 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution* 58:1781-1793.
- Guisan, A., B. Petitpierre, O. Broennimann, C. Daehler, and C. Kueffer. 2014. Unifying niche shift studies: insights from biological invasions. *Trends in Ecology & Evolution* 29:260-269.
- Hijmans, R. J. 2019. raster: geographic data analysis and modeling. R package version 3.0-7. <https://CRAN.R-project.org/package=raster>
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965-1978.
- Hijmans, R. J., and J. Elith. 2015. Species distribution modeling with R. Available: <ftp://ctan.math.mun.ca/CRAN/web/packages/dismo/vignettes/sdm.pdf>.
- Joseph, L., and D. R. B. Stockwell. 2000. Temperature-based models of the migration of Swainson's Flycatcher (*Myiarchus swainsoni*) across South America: a new use for museum specimens of migratory birds. *Proceedings of the Academy of Natural Sciences of Philadelphia* 150:293-300.
- Kadmon, R., O. Farber, and A. Danin. 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications* 14:401-413.
- Knouft, J. H., J. B. Losos, R. E. Glor, and J. J. Kolbe. 2006. Phylogenetic analysis of the evolution of the niche in lizards of the *Anolis sagrei* group. *Ecology* 87:S29-S38.
- Kozak, K. H., and J. J. Wiens. 2010. Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecology Letters* 13:1378-1389.
- Lanyon, S. M. 1993. Phylogenetic frameworks—towards a firmer foundation for the comparative approach. *Biological Journal of the Linnean Society* 49:45-61.
- Losos, J. B. 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters* 11:995-1003.
- Louca, S., and M. Doebeli. 2018. Efficient comparative phylogenetics on large trees. *Bioinformatics* 34:1053-1055.
- Maguire, B. 1973. Niche response structure and the analytical potentials of its relationship to the habitat. *American Naturalist* 107:213-246.
- Meyer, A. L. S., and Pie, M. R. 2018. Environmental prevalence and the distribution of species richness across climatic niche space. *Journal of Biogeography* 45: 2348-2360. <https://doi.org/10.1111/jbi.13419>
- Meseguer, A. S., J. M. Lobo, R. Ree, D. J. Beerling, and I. Sanmartin. 2015. Integrating fossils, phylogenies, and niche models into biogeography to reveal ancient evolutionary history: the case of Hypericum (Hypericaceae). *Systematic Biology* 64:215-232.
- Nakazawa, Y., A. T. Peterson, E. Martinez-Meyer, and A. G. Navarro-Siguenza. 2004. Seasonal niches of Nearctic-Neotropical migratory birds: Implications for the evolution of migration. *Auk* 121:610-618.
- Nyari, A. S., and S. Reddy. 2013. Comparative phyoclimatic analysis and evolution of ecological niches in the scimitar babblers (Aves: Timaliidae: Pomatorhinus). *PLoS ONE* 8:e55629.
- O'Meara, B. C. 2012. Evolutionary inferences from phylogenies: A review of methods. *Annual Review of Ecology, Evolution, and Systematics* 43:267-285.
- Owens, H. L., L. P. Campbell, L. L. Dornak, E. E. Saupe, N. Barve, J. Soberon, K. Ingenloff, A. Lira-Noriega, C. M. Hensz, C. E. Myers, and A. T. Peterson. 2013. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling* 263:10-18.

- Owens, H. L., D. S. Lewis, J. R. Dupuis, A. L. Clamens, F. A. H. Sperling, A. Y. Kawahara, R. P. Guralnick, and F. L. Condamine. 2017. The latitudinal diversity gradient in New World swallowtail butterflies is caused by contrasting patterns of out-of and into-the-tropics dispersal. *Global Ecology and Biogeography* 26:1447-1458.
- Paradis E., and K. Schliep. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526-528.
- Pagel, M., A. Meade, and D. Barker. 2004. Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology* 53:673 - 684.
- Pearman, P. B., A. Guisan, O. Broennimann, and C. F. Randin. 2008. Niche dynamics in space and time. *Trends in Ecology and Evolution* 23:149-158.
- Peterson, A. T. 2011. Ecological niche conservatism: A time-structured review of evidence. *Journal of Biogeography* 38:817-827.
- Peterson, A. T., J. Soberon, R. G. Pearson, R. P. Anderson, E. Martinez-Meyer, M. Nakamura, and M. B. Araujo. 2011. *Ecological Niches and Geographic Distributions*. Princeton University Press, Princeton.
- Peterson, A. T., J. Soberon, and V. Sanchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. *Science* 285:1265-1267.
- Phillips, S. J., M. Dudik, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications* 19:181-197.
- Powell, A. F., F. K. Barker, S. M. Lanyon, K. J. Burns, J. Klicka, and I. J. Lovette. 2014. A comprehensive species-level molecular phylogeny of the New World blackbirds (Icteridae). *Molecular Phylogenetics and Evolution* 71:94-112.
- R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3:217-223.
- Revell, L. J., L. J. Harmon, and D. C. Collar. 2008. Phylogenetic signal, evolutionary process, and rate. *Systematic Biology* 57:591-601.
- Ribeiro, V., A. T. Peterson, F. P. Werneck, and R. B. Machado. 2016. Ecological and historical views of the diversification of Geositta miners (Aves: Furnariidae: Sclerurinae). *Journal of Ornithology* 158:15-23.
- Rice, N. H., E. Martinez-Meyer, and A. T. Peterson. 2003. Ecological niche differentiation in the Aphelocoma jays: A phylogenetic perspective. *Biological Journal of the Linnean Society* 80:369-383.
- Saupe, E. E., V. Barve, C. E. Myers, J. Soberon, N. Barve, C. M. Hensz, A. T. Peterson, H. L. Owens, and A. Lira-Noriega. 2012. Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling* 237:11-22.
- Saupe, E. E., N. Barve, H. L. Owens, J. C. Cooper, P. A. Hosner, and A. T. Peterson. 2017. Reconstructing ecological niche evolution when niches are incompletely characterized. *Systematic Biology* 67:428-438.
- Smith, S. A., and M. J. Donoghue. 2010. Combining historical biogeography with niche modeling in the Caprifolium clade of Lonicera (Caprifoliaceae, Dipsacales). *Systematic Biology* 59:322-341.
- Soberon, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters* 10:1115-1123.

- Soberon, J. and A. T. Peterson. 2011. Ecological niche shifts and environmental space anisotropy: A cautionary note. *Revista Mexicana de Biodiversidad* 82: 1348-1353.
- Swofford, D. L., and W. P. Maddison. 1987. Reconstructing ancestral character states under Wagner parsimony. *Mathematical Biosciences* 87:199-229.
- VanDerWal, J, L. P. Shoo, C. Graham, and S. E. Williams SE. 2009. Selecting pseudo- absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling* 220:589-594.
- Veloz, S. D., J. W. Williams, J. L. Blois, F. He, B. Otto-Bliesner, and Z. Liu. 2012. No-analog climates and shifting realized niches during the late quaternary: Implications for 21st-century predictions by species distribution models. *Global Change Biology* 18:1698-1713.
- Vieites, D. R., S. Nieto-Roman, and D. B. Wake. 2009. Reconstruction of the climate envelopes of salamanders and their evolution through time. *Proceedings of the National Academy of Sciences USA* 106:19715-19722.
- Warren, D. L., M. Cardillo, D. F. Rosauer, and D. I. Bolnick. 2014. Mistaking geography for biology: inferring processes from species distributions. *Trends in Ecology & Evolution* 29:572-580.
- Wiens, J. J., and C. H. Graham. 2005. Niche conservatism: Integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution and Systematics* 36:519-539.
- Yesson, C., and A. Culham. 2006. Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. *Systematic Biology* 55:785-802.