Firneno, T. J., Roelke, C. E., Leaché, A. D., Rains, N. D., Hibbitts, T. J., Ryberg, W. A., Laduc, T. J., Singhal, S., & Fujita, M. K. (2024). Genome-wide data reinforces the evolutionary relationships of previously problematic earless lizards (Phyrnosomatidae: *Holbrookia*). *Bulletin of the Society of Systematic Biologists*. https://doi.org/10.18061/bssb.v3i2.9459

Investigations

Genome-wide Data Reinforces the Evolutionary Relationships of Previously Problematic Earless Lizards (Phyrnosomatidae: *Holbrookia*)

Thomas J. Firneno Jr.[1], Corey E. Roelke[1], Adam D. Leaché[2], Nathan D. Rains[3], Toby J. Hibbitts[4], Wade A. Ryberg[5], Travis J. Laduc[6], Sonal Singhal[7], Matthew K. Fujita[1]

[1] Department of Biology, University of Texas at Arlington, [2] Department of Biology, University of Washington, [3] Texas Parks and Wildlife Department, [4] Department of Wildlife and Fisheries Sciences, Texas A&M University, [5] Natural Resources Institute, Texas A&M University, [6] Department of Integrative Biology, University of Texas at Austin, [7] Department of Biology, California State University, Dominguez Hills

Keywords: Holbrookia, species delimitation, biodiversity

https://doi.org/10.18061/bssb.v3i2.9459

Bulletin of the Society of Systematic Biologists

Abstract

In the face of anthropogenic change and the potential loss of species, documenting biodiversity – including accurately delimiting species complexes – is of tantamount importance. Genome-wide data are powerful for investigating lineage divergence, though deciding if this divergence represents species-level differentiation remains challenging. Here, we use genome-wide data to investigate species limits in four currently recognized species of Earless Lizards (Phrynosomatidae: *Holbrookia*), with a focus on *H. lacerata* and *H. subcaudalis*, the latter having potentially imperiled populations. This group's taxonomy has been repeatedly revised; most recently, *H. lacerata* and *H. subcaudalis* were elevated to species status using conserved morphological data and a few molecular markers. In this study, we used double-digest restriction-site associated DNA sequencing to delineate species limits for our focal taxa. We recovered five populations that corresponded to five well-supported lineages with very little gene flow among them. Our results support the recognition of *H. lacerata* and *H. subcaudalis* as two separate species, based on strong phylogenetic support for these lineages and genetic divergence measures that exceed those of currently recognized species within *Holbrookia*. Genomic methods for species delimitation offer a promising approach to assess biodiversity in taxonomically confounded taxa or organisms of conservation priority.

1 Introduction

The delimitation and description of new species is a fundamental and critical pursuit in biology, particularly with respect to understanding and inventorying the breadth of biodiversity and the application of conservation management. This is especially true at the present, where anthropogenic change is leading to the incredible loss of biodiversity worldwide (Ceballos et al., 2015; Hooper et al., 2012; Wake & Vredenburg, 2008). However, while great strides have been made in developing ways to rapidly assess biodiversity and species boundaries (e.g. DNA barcoding, robust quantitative techniques to carry out species delimitation), debates over the methodological and philosophical approaches to species delimitation persist (Burbrink & Ruane, 2021; Carstens et al., 2013; Hillis et al., 2021; Ratnasingham & Hebert, 2007; Sites & Marshall, 2003). Integrative taxonomic approaches have mitigated many of these issues (Dayrat, 2005; de Queiroz, 2007; Fujita et al., 2012; Padial et al., 2010; Padial & de la Riva, 2009; Pante et al., 2015; Will et al., 2005). Yet, these approaches are time consuming, and increasingly, people are not trained in implementing formal taxonomy and/or identifying fine scale morphological features used to define taxonomies (Engel et al., 2021; Riedel et al., 2013).

Emerging methods that use genomic data to delimit species have also mitigated many of these issues (Jones et al., 2015; Kapli et al., 2017; Smith & Carstens, 2020; Sukumaran et al., 2021; Yang & Rannala, 2010). However, these approaches do not always address whether the delimited units should be recognized as populations or species (Fujita et al., 2012; Leaché et al., 2019), and can potentially lead to an over- or under-estimation of species, causing estimates of biodiversity to be skewed (Campillo et al., 2020; Carstens et al., 2013; Rannala, 2015; Sukumaran & Knowles, 2017). One solution to these challenges is to apply a reference-based approach, in which divergence among "good" species

is used to benchmark and define the boundary between population and species for putative taxa. Such reference-based approaches have been used for some time (e.g., heuristic cutoffs in DNA barcoding; Hebert et al., 2004; Hebert & Gregory, 2005). They are being modernized to leverage genomic data and to provide a more empirical perspective on how genetic divergence relates to the "species continuum" (Chan & Grismer, 2019; Leaché et al., 2021; Poelstra et al., 2021). Reference-based approaches based on genomic data can also serve as a starting point that can be followed by other integrated datatypes for a more holistic and robust taxonomy. More robust approaches to species delimitation will allow us to increase taxonomic stability, to provide a framework for comparing the results of different species delimitation methods, and to implement more robust conservation management plans.

Here, we use genomic data to rapidly assess species boundaries in a taxonomically complex group of lizards (Phrynosomatidae: *Holbrookia*). While the monophyly of the lizard genus *Holbrookia* is well-supported within the family Phyrnosomatidae, the taxonomy of lineages within this genus has a turbulent history and several of the species still contain contentious subspecies designations (de Queiroz, 1992; Schulte & de Queiroz, 2008; Wilgenbusch & de Queiroz, 2000). Until recently, the Northern (Plateau) Spot-tailed Earless Lizard (*H. lacerata*) and Southern (Tamaulipan) Spot-tailed Earless Lizard (*H. subcaudalis*) were recognized as subspecies of *H. lacerata* (Hibbitts et al., 2019). These taxa are separated by ~100 kilometers across the Balcones Escarpment (Hibbitts et al., 2019). The Balcones Escarpment consists of sloping, limestone rocks that likely act as a barrier for these lizards, which are more typically found on flat, gravel soils. These two taxa are morphologically distinct (Axtell, 1956, 1958), and genetic data also supported the elevation of *H. subcaudalis* to species status (Hibbitts et al., 2019; Roelke et al., 2018). Further, these studies recovered two distinct mitochondrial lineages (eastern and western) within *H. subcaudalis*; these lineages are also geographically isolated by both historical and contemporary unsuitable habitat (Fig. 1). However, the study by Roelke et al. (2018) was based on only mitochondrial data, and Hibbitts et al. (2019) was based on one mitochondrial and nuclear gene each. Single locus methods, particularly those reliant on just mitochondrial data, can often be misrepresentative. The mitochondrial genome sorts faster and has a higher mutation rate than nuclear loci, leading to patterns of more rapid divergence than we would expect otherwise (Hudson & Turelli, 2003), and many species delimitation methods can be misled by patterns of introgression and incomplete lineage sorting when relying on topologies derived from a single locus (Dupuis et al., 2012; Knowles & Carstens, 2007). Thus, this revised taxonomy should ideally be confirmed using a multi-locus or genomic dataset.

An accurate species taxonomy is crucial for both *Holbrookia lacerata* and *H. subcaudalis* because both species are facing anthropogenic threat. Both *Holbrookia lacerata* and *H. subcaudalis* are early successional species that prefer patches of grassland, where disturbances produce a mosaic of sparse open patches of bare ground mixed with less-dense vegetation structure and occasional shrubs or trees (Hibbitts et al., 2021). Before the human colonization of North America, *H. lacerata* and *H. subcaudalis* were likely most abundant in grasslands heavily impacted by the effects of fire and grazing by large herbivores. Today, both species of *Holbrookia* are often found in high abundance in plowed agricultural fields or overgrazed pastures, as the microhabitat conditions in these landscapes are favorable to foraging and reproductive behaviors in these taxa (Hibbitts et al., 2021; Roelke et al., 2018). Recent evidence suggests that populations of both *H. lacerata* and. *H. subcaudalis* are in decline throughout their distributions in Texas in the southwestern United States, likely due to widespread use of pesticides and agricultural practices, urbanization, and invasive flora and fauna (Duran & Axtell, 2010; Wolaver et al., 2018). These declines have made them the focus for investigating land management practices (Duran & Axtell, 2010; Hibbitts et al., 2021). *Holbrookia lacerata* is currently listed as near threatened by the International Union for Conservation of Nature organization (IUCN). Additionally, due to perceived population declines or extirpations across much of their range in Texas (Duran & Axtell, 2010), both species of spot-tailed earless lizards (*H. lacerata* and *H. subcaudalis*) are currently undergoing a status review for federal listing in the U.S. (USFWS, 2011).

Here, we used a multi-locus, reference-based approach (mitochondrial DNA and ddRADseq) to determine the species status for the threatened species, *H. lacerata* and *H. subcaudalis*. We then compared our estimated species statuses to our results from four different approaches to species delimitation to assess the ability of these techniques to discern populations from species.

2 Methods

2.1 Taxon Sampling and DNA Extraction

We collected tissues from 75 individuals across the U.S. distribution of the focal taxa (*H. lacerata*, *H. subcaudalis*, and close sister species *H. propinqua*, and *H. maculata*). The large gap in sampling in *H. subcaudalis* falls along the break between the eastern and western lineages, and it spans a geographic region where no known intermediate populations are known to exist (Fig. 1). We extracted DNA from muscle or liver tissue stored in RNALater or 70% ethanol using a standard phenol-chloroform protocol (Sambrook & Russell, 2006). DNA extraction quality was checked using a 1% agarose gel, and DNA concentration was quantified using a QUBIT 2.0 Fluorometer (Life Technologies, Carlsbad, CA).

2.2 Molecular data generation and processing

We compiled a mitochondrial DNA dataset of the NADH dehydrogenase subunit 2 (*NAD2*) gene using sequences available from Hibbitts et al. (2019) and Roelke et al. (2018). We generated a multiple sequence alignment in MEGA7 (Kumar et al., 2016) using the MUSCLE algorithm (Edgar, 2004) with default parameters. The final alignment con-

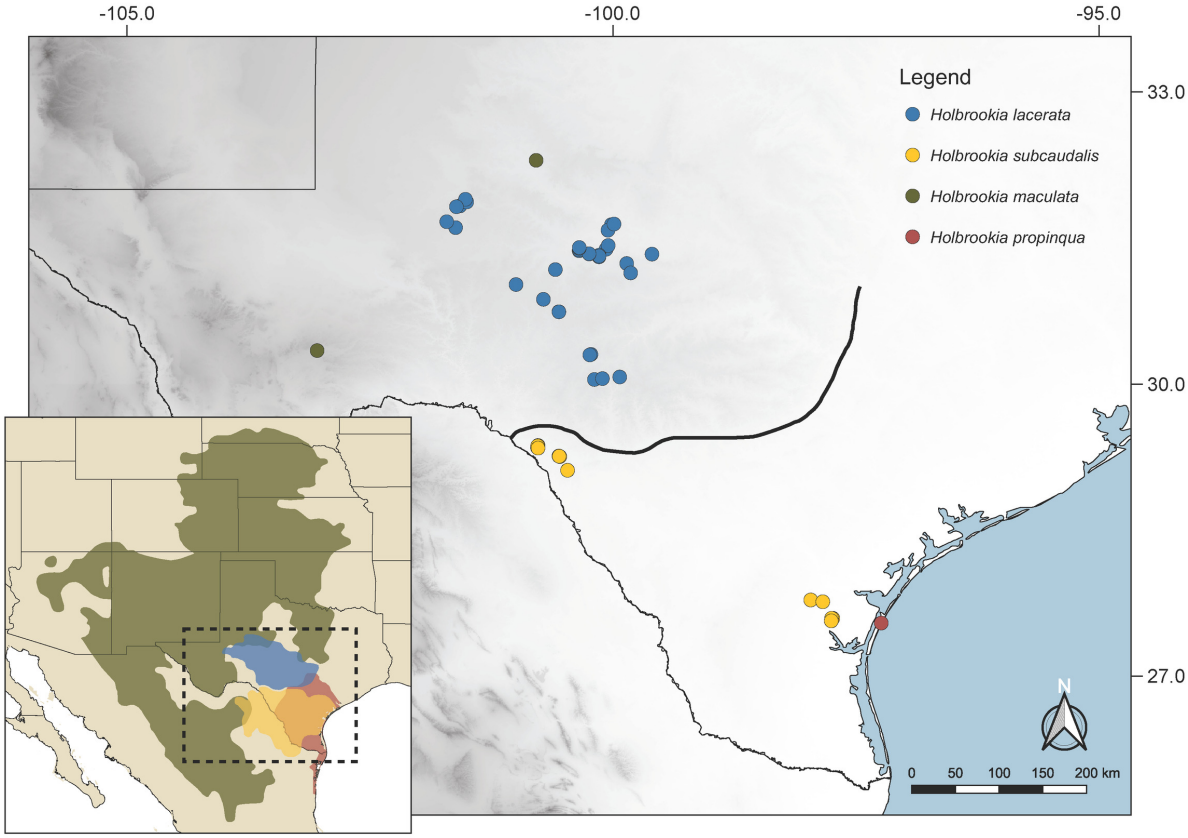

Figure 1. Sampling localities of the four focal taxa (*Holbrookia lacerata, H. maculata, H. propinqua,* and *H. subcaudalis*) included in this study, with approximate geographic distributions in southwestern North America shown in the inset. The solid black line represents the Balcones Fault/Escarpment, the natural biogeographic barrier between *H. lacerata* and *H. subcaudalis*. Gray shading is SRTM 1 Arc-Second Global elevation data (www.usgs.gov).

tained 1029 bp for 57 samples, which include 49 samples of our four focal taxa plus 8 other closely related species, including *Holbrookia elegans*, *Cophosaurus texanus*, *Callisaurus draconoides*, *Uma exsul*, *U. paraphygas*, *U. inornata*, *U. notata*, and *U. scoparia* used as outgroups (Table S1).

We collected ddRADseq data for 75 individuals from our four focal taxa following the protocol described in Peterson et al. (2012) and following parameters specified in Streicher et al. (2014). Our final library was sequenced on one Illumina HiSeq2500 lane (150 bp single end reads) at the Genomic Sequencing and Analysis Facility (GSAF) at The University of Texas (https://www.wikis.utexas.edu/display/GSAF). The workflow for data processing, filtering, and formatting was automated using scripts available from Portik et al., 2017 (https://github.com/dportik/Stacks_pipeline). In brief, the raw Illumina reads were demultiplexed using stacks v2.53 (Rochette et al., 2019), the restriction site overhangs were removed using the fastx_trimmer module of the fastx-toolkit (www.hannonlab.cshl.edu/fastx_toolkit), and the sequencing quality was examined on a per sample basis using fastqc v0.10.1 (www.bioinformatics.babraham.ac.uk/projects/fastqc). Loci were created, catalogued, and identified using ustacks, cstacks, and sstacks, respectively. The "populations" option was then used to export loci present in 80% of all individuals, which resulted in 1,872 loci. Subsequent filtering removed invariant loci (n=59), "blank" loci (n=11), non-biallelic loci (n=0), and loci containing at least one individual with more than two alleles (n=0). For loci containing multiple SNPs, we randomly chose a single SNP to be used for subsequent analyses. Any samples missing data for more than 60% of loci were removed (n=9). After completing the above filtering steps, our final SNP dataset consisted of 66 samples and 1,802 SNPs.

2.3 Population Structure, Gene Flow, and Isolation-by-Distance

We determined the number of discrete populations present across the sampled ranges of our focal lineages using Bayesian and likelihood clustering analyses and multivariate methods. We used STRUCTURE v2.3.4 (Falush et al., 2003; Pritchard et al., 2000) to examine the number of population clusters and potential admixture between populations. Hierarchical analyses were performed for 10 runs per K, up to a maximum of 6 populations, using the admixture model with a burn-in of 10,000 steps followed by 100,000 steps. We summarized our results using structureharvester (Earl & vonHoldt, 2012) and evaluated the number of populations based on inspection of likelihood plots and following Evanno et al (2005). To complement our structure analysis, we used a maximum likelihood approach with ADMIXTURE (Alexander et al., 2009). We performed ten repli-

cate analyses to evaluate up to 6 populations. To assess the best K value, we performed 10-fold cross-validation and selected the K value with the lowest cross-validation error. We also evaluated the number of discrete populations using a discriminant analysis of principal components (DAPC) with adegenet v2.0.0 (Jombart, 2008; Jombart & Ahmed, 2011). A maximum of 10 clusters were investigated using the k-means algorithm. The preferred number of clusters was evaluated using BIC scores. To minimize overfitting, an initial DAPC was used to find the a-score for each set of clusters, and this value was used to select the number of principal components to retain for subsequent analyses (Jombart, 2008; Jombart & Ahmed, 2011). To independently assess the validity of population differentiation and assignment, we used the fineRADstructure software package (Malinsky et al., 2018) to construct a co-ancestry matrix from the RADseq data. We used 100,000 burn-in followed by 100,000 MCMC steps sampling every 1,000 steps and the tree was constructed with 10,000 hill-climbing iterations. The results were visualized using the fineRADstructureplot.R and fineRADstructurelibrary.R scripts (included in the fineRADstructure package file).

To test whether there is introgression occurring between any of our focal populations (*H. lacerata*, *H. subcaudalis* "East", and *H. subcaudalis* "West") we analyzed a reduced individual dataset (individuals being assigned based on population clustering results) including the entire RAD locus for all sites (variable and invariant sites) under the multispecies coalescent with introgression (MSCi) model as implemented in BPP v4.3.0 (Flouri et al., 2020). We tested for introgression and estimated the strength of introgression using the introgression probability or admixture proportion (φ) by analyzing eight models that each included one unidirectional introgression event on the species tree (Table 1; Fig. 4). Diffuse inverse-gamma (IG) priors were designated as follows: θ ~ IG(3, 0.002) with mean 0.002/(3 – 1) = 0.001, and τ ~ IG(3, 0.01) with mean 0.01/(3 – 1) = 0.005. The prior for introgression was beta (1,1). Each analysis was run three times for 100,000 generations with the first 10,000 discarded as burn-in. Convergence was assessed by ensuring the stability of parameter estimates. We tested the significance of introgression using the Savage-Dickey density ratio and Bayes factors (Ji et al., 2023). Models with a BF ⩾ 20 were considered significant, and two introgression thresholds (the percent of the genome introgressed) were tested, 0.01 and 0.05 (Ji et al., 2023; Table 1).

Divergence between two geographically proximate species can result from a discrete break in gene flow due to geographic and/or reproductive isolation or the accrual of continuous divergence over space (isolation-by-distance; IBD). To test whether IBD explains divergence between populations of *H. lacerata* and *H. subcaudalis*, we used a Mantel test to compare pairwise genetic distances (d_{xy}) and pairwise geographic distances. Specifically, we tested for IBD between *H. lacerata* and the "East" and "West" lineages of *H. subcaudalis*. We tested for significant relationships between genetic distances (d_{xy}) and pairwise geographic distances using the mantel.randtest function in the R package 'adegenet' (Jombart, 2008). We performed 99,999 permutations to generate a random distribution of values to compare to our empirical pairwise values. We plotted genetic and geographical distances to visually inspect for evidence of a continuous cline of genetic distance versus distinctive genetic breaks between populations.

2.4 Phylogenetic relationships and divergence dating

We estimated a gene tree for the mtDNA data using the Bayesian software package BEAST v2.6.6 (R. Bouckaert et al., 2014). The stem group of the clade of *Cophosaurus* and *Holbrookia* is at least 16.0 Myr old, given the fossil *Holbrookia antiqua* from the Marsland Quarry from the Hemingfordian Formation (16.0-20.4 Myr) of Nebraska (Yatkola, 1976). Therefore, we used 17.0 Ma as a calibration point mean (offset = 1; sigma = 0) under a normal distribution for the divergence between *Cophosaurus* and *Holbrookia* (Wiens et al., 2013). We performed the analysis using an HKY+G model of nucleotide substitution with four gamma categories, relaxed clock model, and a Yule tree prior. We ran the analysis of three independent runs using a random starting tree with a Markov Chain Monte Carlo (MCMC) run for 20,000,000 generations, sampling every 1,000 generations producing a total of 20,000 trees. We examined the evidence for the lack of convergence using Tracer v1.6 (Rambaut & Drummond, 2009), and runs were combined in LogCombiner v2.6.6. A burn-in of 10% was discarded, and a maximum clade credibility (MCC) tree with median heights was created from the remaining trees.

We additionally estimated a species tree for our SNP dataset using SNAPP v1.3.0 (Bryant et al., 2012) implemented in BEAST2 (R. Bouckaert et al., 2014). The SNAPP model is based on the coalescent process and can accommodate ILS but assumes no gene flow and can be misguided by gene flow. To reduce run times, we randomly subsampled 2 – 8 representatives from each of five populations, for a total of 19 individuals. We defined populations based on the consensus found across clustering analyses (see Fig. 2). The mutation rates *u* and *v* were set to 1.0. We set the prior for the expected genetic divergence (θ) using a gamma distribution θ ~ G(2.5, 250) with a mean of alpha/beta = 0.01. We assigned a gamma hyperprior for the speciation rate parameter λ ~ G(2,200) with a mean alpha x beta = 400. We performed two independent runs with a chain length of 1,000,000 generations, sampling every 1000 generations. Runs were assessed using Tracer v1.6 to examine convergence (Rambaut & Drummond, 2009), and tree topologies and node heights were visualized using DENSITREE (R. R. Bouckaert, 2010).

2.5 Reference-based Taxonomy and Species Delimitation

To define the reference-based taxonomy for *Holbrookia*, we calculated the genetic divergence among recognized species and putative new taxa identified in our analysis. We measured genetic divergence by calculating mean d_{xy} and F_{ST} among all population pairs using the populations func-

tion in STACKS (Meirmans, 2006; Nei, 1987; Reich et al., 2009).

We then evaluated species boundaries using four different approaches for comparative molecular species delimitation: two coalescent-based approaches (Bayes Factor Delimitation*, BPP), a heuristic-based approach (*gdi*), and a protracted speciation approach (DELINEATE). Because each species delimitation approach makes different assumptions, we compared across multiple methods to identify robust delimitations (Smith & Carstens, 2022).

We used two methods that implement the multispecies coalescent model (MSC; Rannala & Yang, 2003). First, we conducted Bayes Factor Delimitation (BFD*) implemented in SNAPP v1.3.0 (Bryant et al., 2012) in BEAST2 v2.5.4 (R. Bouckaert et al., 2014) with our SNP dataset following Leaché et al. (2014). We used a dataset for all models that only included SNPs that were present across all considered taxa (subsampled 2 – 8 representatives from each of five populations), since differences in the number of SNPs can make the comparison of models inaccurate (Leaché & Oaks, 2017). We performed our analyses testing three models, including the current taxonomy (4 species – *H. lacerata*, *H. subcaudalis*, *H. maculata*, and *H. propinqua*), a taxonomy that lumped *H. lacerata* and *H. subcaudalis* (3 species – *H. lacerata* + *H. subcaudalis*, *H. maculata*, and *H. propinqua*), and a taxonomy that split populations of *H. subcaudalis* (5 species – *H. lacerata*, *H. subcaudalis* "East", *H. subcaudalis* "West", *H. maculata*, and *H. propinqua*). Using BFD*, we ranked all the models and selected the model with the highest maximum likelihood estimate (MLE) score. We calculated the Bayes Factor (BF) support for the best model using the equation BF = 2(MLE model 1 – MLE model 2) (Kass and Raftery 1995). We utilized the same parameters as our species tree estimation, but we ran the analyses for 100,000 generations, sampling every 1000 generations. We estimated marginal likelihood values using the PathSampleAnalyzer (alpha = 0.3) with 48 steps (10000 iterations, 1000 pre-burnin).

Second, we used Bayesian species delimitation as implemented in Bayesian Phylogenetics and Phylogeography (BPP) v4.3.0 (Flouri et al., 2020). The entire RAD locus for all sites (variable and invariant sites) was used for this analysis and were phased within the BPP analysis (J. Huang et al., 2022). We implemented the A10 (species delimitation using a fixed guide tree) analysis under the multispecies coalescent model. We used a consensus of relationships derived from population and phylogenetic (specifically, SNAPP) analyses allowing five populations (based on consensus of population analyses) as a guide tree (see Fig. 4). We used the priors used for θ and τ as they are described above for our MSCi analysis in BPP. Each analysis was run three times for 100,000 generations with the first 10,000 discarded as burnin. Convergence was assessed by ensuring the stability of parameter estimates.

Third, we used heuristic species delimitation which estimates the genealogical divergence index (*gdi*) to measure the overall genetic divergence between two populations based on the combined effects of genetic isolation and gene flow. This approach is therefore useful for studying structured populations (Jackson et al., 2017; Leaché et al., 2019). This method assumes that species should show greater evolutionary independence from one another compared to populations (Hey & Pinho, 2012). The *gdi* is continuous between 0 (panmixia) and 1 (strong divergence), which can be indicative of where a population lies on the path to speciation. Though there is no fixed delimitation cutoff between populations and species, it has been suggested that there are distinct species when *gdi* >0.7, the same species when *gdi* <0.2, and/or ambiguous delimitation status if 0.2> *gdi* <0.7 (Jackson et al., 2017; Pinho & Hey, 2010). We used *gdi* to compare the observed divergence between *H. lacerata* and *H. subcaudalis* with respect to species-level divergences between other lineages of *Holbrookia* (Chan & Grismer, 2019; Leaché et al., 2018). To estimate the *gdi*, we used the priors used for θ and τ as they are described above for our MSCi analysis in BPP. Four separate runs were performed for 100,000 generations with the first 10,000 discarded as burn-in and converged runs were combined to generate posterior distributions for the multispecies coalescent parameters that were then used to calculate *gdi* following the equation: $gdi = 1\text{-}e^{-2\tau/\theta}$ (Jackson et al., 2017; Leaché et al., 2018). When calculating the *gdi*, Population A was distinguished from Population B using the equation $2t_{AB}/\theta_A$, where θ_A is the θ of the focal population in this comparison. Conversely, $2\tau_{AB}/\theta_B$ was used to differentiate between Population B from Population A, where θ_B is the θ of the focal population in this comparison (Jackson et al., 2017; Leaché et al., 2018).

Finally, to address shortcomings of MSC methods (e.g., overestimating species based on population boundaries), we implemented the protracted speciation model through the program DELINEATE (Sukumaran et al., 2021). DELINEATE assesses potential species boundaries of observed genetic clusters by combining elements of tree structure and branch length, while modeling an extended time period after population splitting during which populations can go extinct, further split, or evolve into species (initiation and completion rate of speciation). We assigned candidate genetic clusters based on population clustering, phylogenetic results, and geography. We used the species tree from our SNAPP analysis as a guide tree. Within our control file, we distinguished our assigned genetic clusters as constrained (i.e., well-established species) or unconstrained lineages (those that remain the focus of species delimitation). We constrained the *H. maculata* and *H. propinqua* as species-level lineages but left the *H. lacerata*, *H. subcaudalis* "West", and *H. subcaudalis* "East" lineages as unconstrained to determine if they are inferred to be separate species.

3 Results

3.1 Population Inference, Isolation-by-Distance, and Gene Flow

Both hierarchical Bayesian (STRUCTURE) and maximum likelihood (ADMIXTURE) population clustering analyses based on 1,802 SNPs resulted in the detection of four primary populations that followed the nominal species designations (K=4; Fig. 2a,b). Further, we found limited evidence

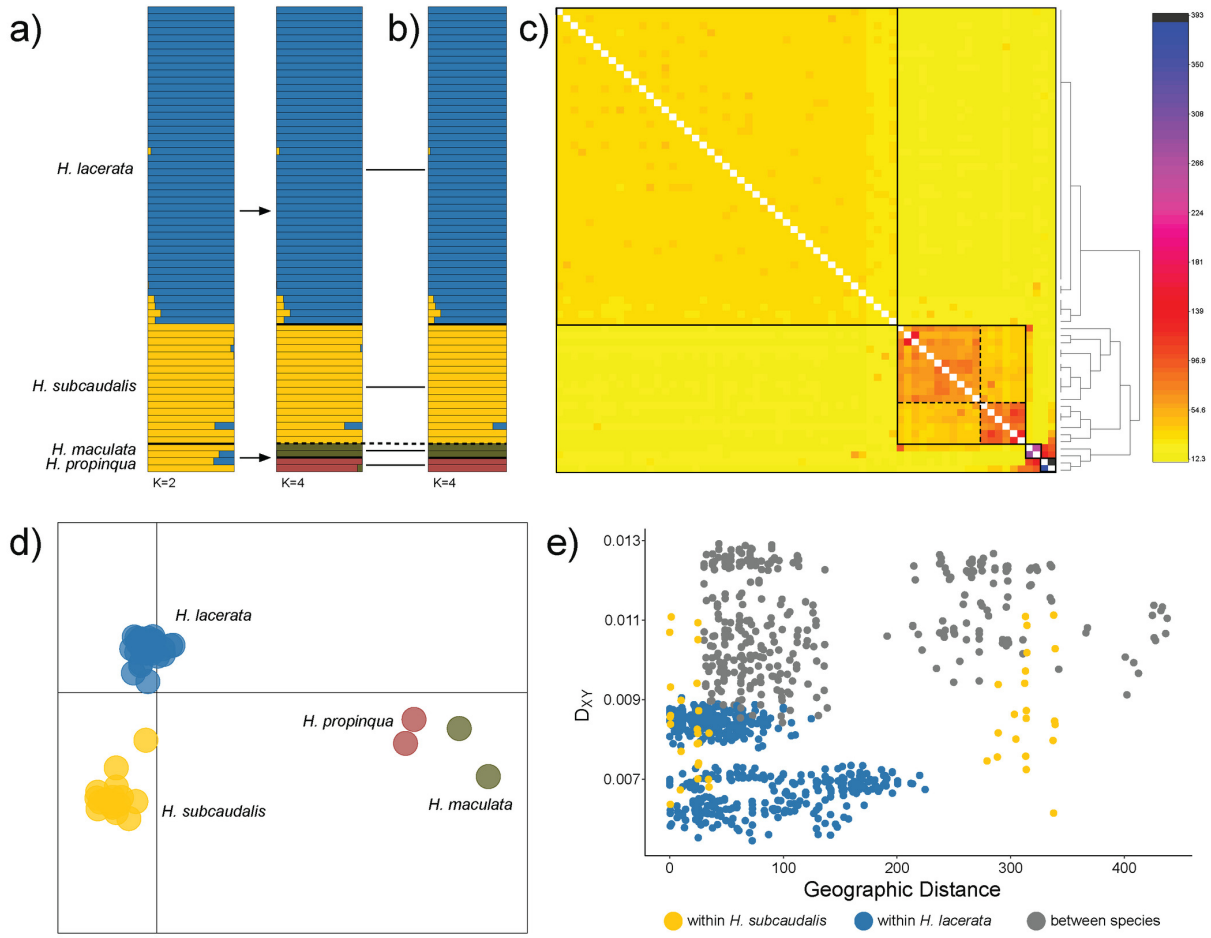

Figure 2. Population assignment for 66 individuals of species *Holbrookia lacerata, H. maculata, H. propinqua,* and *H. subcaudalis* from SNP data based on a) hierarchical Bayesian population clustering using STRUCTURE, b) maximum-likelihood population clustering using ADMIXTURE (dashed line indicates where populations were analyzed separately for hierarchical analysis), c) co-ancestry matrix from fineRADstructure (coefficients of ancestry are colored from low (yellow) to high (black), the dendrogram depicts a clustering of individual samples based on the pairwise matrix of co-ancestry coefficients, and boxes indicate population structuring), d) discriminant analysis of principal components. Across all clustering approaches, the best-fit model identified four clusters corresponding to the four nominal taxa. e) The relationship between pairwise geographic and genetic distances within and between populations *subcaudalis*).

for admixture between all populations, only identifying a few putatively admixed individuals in *H. lacerata*. The resulting dendrogram and coancestry matrix from our fineRADstructure analysis confirmed the results of our hierarchical population analyses, while also identifying substructuring within *Holbrookia subcaudalis* ("East" and "West" populations; dashed boxes in Fig. 2c). The DAPC results identified the same populations as the STRUCTURE and ADMIXTURE clustering analyses (Fig. 2d).

The results of our introgression analyses indicated low levels of introgression between all populations (ϕ = 0.044-0.326), with the highest estimates of introgression being estimated between "East" and "West" lineages of *H. subcaudalis*, specifically from the "West" lineage into the "East" lineage (ϕ = 0.326) (Table 1, Fig. 4). Most of the introgression estimates were significant using a Bayes factor test with a 1% threshold, but only three introgression events were significant with a 5% threshold (including *subcaudalis* W → *subcaudalis* E, *lacerata* → *subcaudalis* ancestor, and *subcaudalis* ancestor → *lacerata*; Table 1).

The Mantel test comparison between *H. lacerata* and populations of *H. subcaudalis* was significant (p = 0.0203) showing a positive correlation between geographic and genetic distance. However, individual patterns of between and within species divergence do not follow similar patterns and do not seem to be positively correlated (Fig. 2e).

3.2 Phylogenetic Relationships and Divergence Dating

The divergence dating analysis of mtDNA using a fossil calibration at the node of *Cophosaurus* and *Holbrookia* supported Miocene divergences for a majority of our major clades within *Holbrookia* (Fig. 3). We estimated the divergence for *H. lacerata* + *H. subcaudalis* and *H. elegans* + *H. propinqua* + *H. maculata* occurring in the mid-Miocene (13.84 Ma [95% HPD 9.79-17.2], the divergence for *H. lac-*

Table 1. Introgression events tested in *Holbrookia*. The posterior mean and 95% highest posterior distribution confidence interval (HPD CIs; in parentheses) for introgression probability (ϕ) are listed with Bayes factor test results. Introgression tests are shown for 1% and 5% introgression thresholds (percent of the genome introgressed). Introgression events with Bayes factors ⩾ 20 were considered significant. Bayes factor = ∞ if all ϕ values in the posterior distribution exceeded the threshold.

Introgression Events	ϕ estimate	Bayes factor for 1% introgression	Bayes factor for 5% introgression
lacerata → subcaudalis E	0.044 (0.034 – 0.052)	∞	0.01
subcaudalis E → lacerata	0.061 (0.052 – 0.071)	∞	2.35
lacerata → subcaudalis W	0.046 (0.035 – 0.057)	∞	0.02
subcaudalis W → lacerata	0.058 (0.049 – 0.068)	∞	2.25
subcaudalis E → subcaudalis W	0.057 (0.001 – 0.129)	0.3	0.05
subcaudalis W → subcaudalis E	0.326 (0.196 – 0.489)	∞	28.17
lacerata → subcaudalis Anc.	0.138 (0.117 – 0.159)	∞	318.93
subcaudalis Anc. → lacerata	0.140 (0.113 – 0.167)	∞	∞

erata and *H. subcaudalis* occurring in the mid- to late-Miocene (10.58 Ma [95% HPD 6.48-14.7]), the divergence for *H. elegans* and *H. propinqua*+*H. maculata* occurring in the late-Miocene (7.54 Ma [95% HPD 2.49-13.1]), and the divergence for *H. propinqua* and *H. maculata* occurring in the mid-Pliocene (3.97 Ma [95% HPD 0.56-8.60]).

In agreement with previous studies (Hibbitts et al., 2019; Roelke et al., 2018), we see four distinct and well-supported (pp ⩾ 0.9) lineages representing our four focal taxa for both our mtDNA (Fig. 3) and SNP (Fig. 4) datasets, along with two well-supported (pp ⩾ 0.9) lineages within *H. subcaudalis* reflecting the split of eastern and western populations as seen in our fineRADstructure inference.

3.3 Reference-based Taxonomy and Species Delimitation

Absolute and relative divergence between our currently recognized species (*H. propinqua* and *H. maculata)* were modest (d_{xy}: 1.1%; F_{ST}: 0.36; Table 2). Divergence between *H. lacerata* and *H. subcaudalis* were comparable to the divergence between these recognized species (d_{xy}: 0.8 – 0.9%; F_{ST}: 0.3 – 0.34). However, relative divergence between the two *subcaudalis* lineages "East" and "West" was noticeably lower compared to other lineage comparisons (d_{xy}: 0.7%; F_{ST}: 0.11).

The species delimitation methods yielded similar results to each other (Fig. 4). The MSC delimitation methods estimated five species units, in which all nominal species were delimited along with the "East" and "West" lineages of *H. subcaudalis*. BFD* favored the five species model (MLE =-5936.34) over the three and four species models (MLE =-6728.83 and MLE=-6008.20, respectively) with a BF = -143.72 (Table S2). BPP supported five species units with a posterior probability of 1.0. Heuristic species delimitation using the *gdi* values provided moderate to high values for *H. lacerata* (0.84) and the two populations of *H. subcaudalis* ("East" and "West", 0.78 and 0.71, respectively) (Fig. 4, S1). These values exceeded values compared to *H. maculata* (0.62) and *H. propinqua* (0.68). However, it should be noted that the *gdi* values for *H. maculata* and *H. propinqua* may be inflated due to only being calculated from two specimens each and also coming from only one or two sampling sites that cover a very small portion of the overall range of the two species. Finally, the protracted speciation model delimitation method (DELINEATE) supported five species units.

4 Discussion

Delimiting species that are morphologically conserved, inhabit similar niches, or are understudied and data deficient can be very difficult. These challenges can be particularly pressing for taxa that may be of conservation concern, like this study's focal species *H. lacerata* and *H. subcaudalis*. Here, we integrate population genomic, phylogenetic, and species delimitation to examine species limits between these two species and between the "East" and "West" lineages of *H. subcaudalis*.

Overall, our analyses support the recognition of *H. lacerata* and *H. subcaudalis* as separate species, validating the single locus (*ND2* and whole mitochondrial) genetic estimates of previous studies (Hibbitts et al., 2019; Roelke et al., 2018), as well as morphological and ecological assessments (Hibbitts et al., 2019). Our clustering analyses identify these two species as separate genetic populations, and our estimates of introgression using the MSCi indicate low to non-existent levels of admixture and gene flow between all populations of *H. lacerata* and *H. subcaudalis*. While overall patterns of IBD were significant, between-species divergence was higher than within-species divergence, suggesting that IBD was not the cause of structure among populations of *H. lacerata* and *H. subcaudalis*. Our divergence dating analysis with mtDNA data estimates a mid- to late-Miocene (10.58 Ma [95% HPD 6.48-14.7]) divergence between *H. lacerata* and *H. subcaudalis*, which coincides with the formation of the Balcones Escarpment (see Fig. 1; Hill & Vaughan, 1898; Weeks, 1945). The Bal-

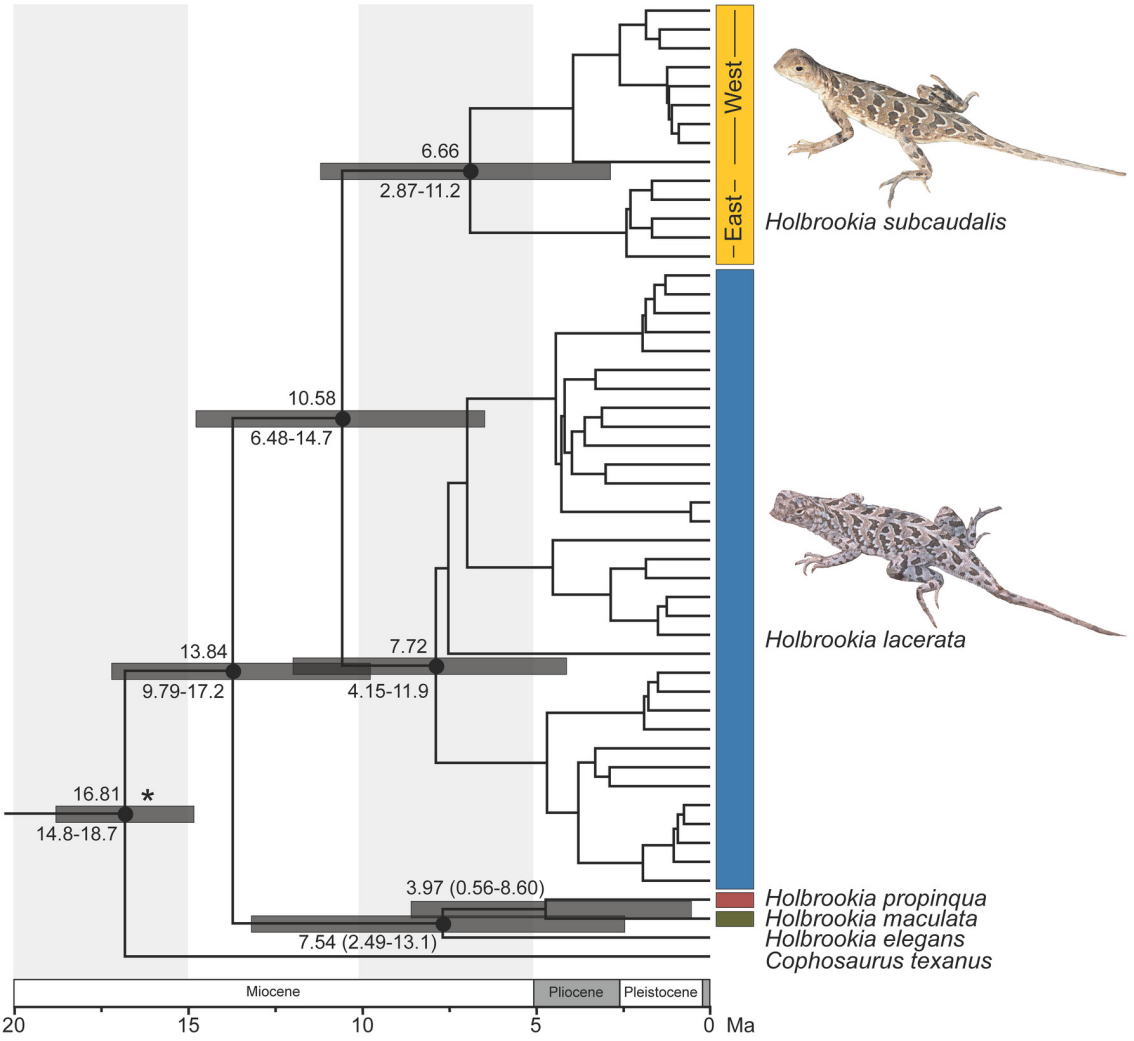

Figure 3. Chronogram of focal *Holbrookia* taxa from a BEAST analysis of ND2 mtDNA calibrated using a fossil of *H. antiqua* (indicated by *) of 17.0 Ma. Nodes with high support (posterior probability > 0.9) are indicated by black dots. Pictured are median ages; error bars at nodes represent the 95% highest posterior densities for node age (HPD). Photos by Troy J. Hibbitts.

Table 2. Comparisons of mean between d_{xy} (above gray diagonal), π (on gray diagonal), and F_{ST} (below gray diagonal) between our focal lineages (*Holbrookia lacerata, H. maculata, H. propinqua,* and *H. subcaudalis* "East" and "West" populations).

	H. lacerata (*n* = 45)	*H. subcaudalis* "East" (*n* = 6)	*H. subcaudalis* "West" (*n* = 11)	*H. propinqua* (*n* = 2)	*H. maculata* (*n* = 2)
H. lacerata	0.0043	0.0089	0.0080	0.0101	0.0112
H. subcaudalis "East"	0.3470	0.0061	0.0068	0.0116	0.0127
H. subcaudalis "West"	0.2998	0.1154	0.0067	0.0109	0.0114
H. propinqua	0.4221	0.4328	0.4180	0.0085	0.0113
H. maculata	0.4306	0.4443	0.3881	0.3668	0.0106

cones Escarpment has been shown to be the primary biogeographic barrier separating *H. lacerata* and *H. subcaudalis* (Axtell, 1958; Hibbitts et al., 2019), as well as other reptile taxa (e.g., *Masticophis taeniatus* and *M. schotti*; Camper & Dixon, 1994). Migration between the two species is thus likely to be limited given their border spans a major biogeographic break. Together, these results suggest that these two lineages are likely to continue diverging in isolation.

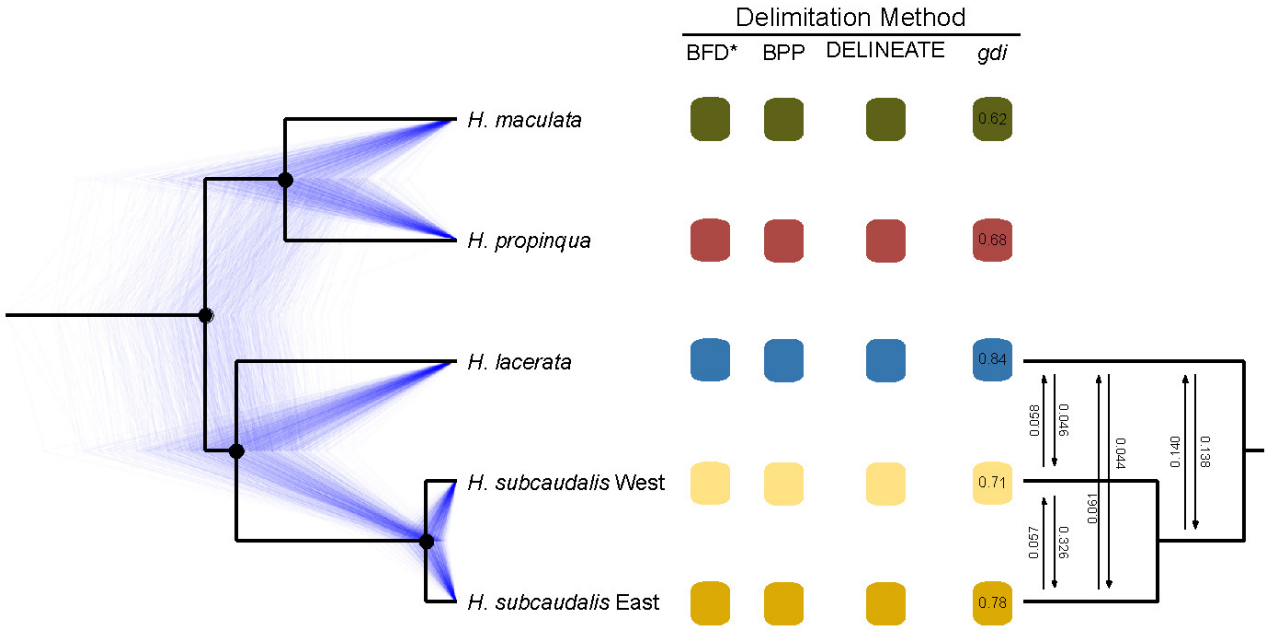

Figure 4. Summary of the species delimitation and introgression analyses for four species within *Holbrookia,* with a species tree of SNP dataset generated using SNAPP (left). Nodes with high support (posterior probability ⩾ 0.9) are indicated by black dots. Blocks represent species units that have been estimated by each species delimitation analysis. Mean gdi values are shown within blocks that represent species units tested by that method. The joint introgression model (right) constructed in this study with eight unidirectional introgression events showing introgression probability (phi) estimates (Table 1).

Further, genetic divergence estimates (d_{xy}) between *H. lacerata* and populations of *H. subcaudalis* indicate relatively high estimates of divergence (~0.8%; Table 2) for nuclear data (Maldonado et al., 2020; Sullivan et al., 2006), which are nearly as high as estimates seen between other nominal taxa (~1%; Table 2), and our species delimitation analyses show a consensus that *H. lacerata* and *H. subcaudalis* are separate species units (Fig. 4). Overall, our genomic analyses, considered along with current and previous analyses (Hibbitts et al., 2019; Roelke et al., 2018), and estimates of divergence based on morphology and ecology, support the distinctiveness of *H. lacerata* and *H. subcaudalis* as separate species.

Both Hibbitts et al., (2019) and Roelke et al., (2018) found distinct "East" and "West" lineages for *H. subcaudalis* (as also indicated in our *ND2* analysis in Fig. 3). The sampling gap observed here and in other studies (Hibbitts et al., 2019; Roelke et al., 2018) between these two lineages is not due to sampling bias, but an actual gap in the distribution where populations of *H. subcaudalis* do not exist. While there has always been a distributional gap within the eastern part of the range of *H. subcaudalis* due to the South Texas Sand Sheet (Axtell, 1956, 1958), there are historical records that extend into the central portion of their range (see Hibbitts et al., 2019) but that are no longer found there most likely due to invasive grass species highly altering the habitat in south Texas (Archer, 1989; Scott, 1996). The genomic evidence for these lineages as unique and independently evolving is mixed. While all species delimitation analyses identified these lineages as species-level taxa (Fig. 4), our population clustering approaches best supported a model in which the "West" and "East" lineages were lumped into one cluster (Fig. 2). Further, genetic divergence between these two lineages is significantly lower (F_{ST} = 0.12) relative to divergences seen between nominal species (F_{ST} = 0.29 – 0.44; Table 2). This lower divergence could partially result from historical or on-going introgression between these populations. We infer slightly elevated estimates of introgression from the "West" lineage into "East" lineage of *H. subcaudalis*. While these lineages are currently allopatric, this could suggest historical introgression if these populations were once connected (Fig. 4). Overall, the "East" and "West" lineages of *H. subcaudalis* are relatively less distinct and exhibit much higher rates of gene flow than other currently recognized species in the genus. Thus, while we found support for two distinct populations of *H. subcaudalis*, we find no compelling evidence that these populations should be elevated to species. Still, these populations could be earmarked for conservation by evaluating and managing Distinct Population Segments under the Endangered Species Act.

We also compared four multilocus species delimitation methods that utilize different frameworks and models. Each of these frameworks has been noted as having benefits and drawbacks. The MSC is one of the most used frameworks for molecular species delimitation and can overcome issues with incomplete lineage sorting and low levels of introgression (Fujita et al., 2012; Zhang et al., 2011). However, methods implementing the MSC (e.g., BFD* and BPP) can tend to delimit species along population boundaries, there-

fore overestimating the number of species within a focal group (Sukumaran & Knowles, 2017). To correct for this, some studies have used a heuristic framework (e.g., *gdi*; Jackson et al., 2017; Leaché et al., 2019, 2021) or the protracted speciation model (e.g., DELINEATE; Sukumaran et al., 2021). The *gdi* has been noted as having several weaknesses, including: (1) that ambiguity may arise when populations of different sizes are being compared because the criterion depends on the population divergence time relative to the population size (small bottlenecked populations and shallow divergence times may skew results), (2) that the metric has a rather large zone in which lineages are not clearly delineated as populations or species ($0.2 < gdi < 0.7$) and thus remains relatively subjective, and (3) the utility of the method may be limited if not used in a comparative manner that includes "good" species (Jackson et al., 2017; Leaché et al., 2019). Likewise, it has been argued that the protracted speciation model is unrealistic in delimiting species because it models instantaneous speciation in a single generation, which does not reflect how natural species convert from being incipient species to a true species (Leaché et al., 2019; Sukumaran & Knowles, 2017). The *gdi* and DELINEATE have been shown to provide more conservative results (fewer species), even when other informative criteria (morphology, ecology, behavior) support more species (e.g., Firneno et al., 2021; J. P. Huang, 2021). Despite the different assumptions made by these methods, all the species delimitation methods found support for five *Holbrookia* species within our dataset (Fig. 4), including support for the relatively less distinct "East" and "West" lineages of *H. subcaudalis*.

Why did these comparative species delimitation methods find such unequivocal support for the species-level status of the "East" and "West" lineages of *H. subcaudalis*, when all other data was relatively mixed? First, these two lineages are distinct, they are just not nearly as distinct as other species-level taxa within *Holbrookia*. These delimitation methods might be identifying – and perhaps overinterpreting – this fine-scale structuring as species-level divergence. Further, while we observed *gdi* between populations of *H. subcaudalis* that exceeded those of other species of *Holbrookia*, this could be due to the relatively small sample sizes used for the reference populations (*H. maculata* and *H. propinqua*) within our analyses causing populations of *H. subcaudalis* to appear more divergent. However, this potential issue should be alleviated using thousands of SNPs to evaluate genetic diversity between these lineages and the even levels of genetic diversity within populations of *H. subcaudalis*, which could influence and bias coalescent-based estimates of genetic divergence. These results do suggest, however, that species delimitation methods should not be used in isolation and should rather be used as part of a broader suite of analyses investigating the distinctiveness of putative species-level taxa.

The issue with many species delimitation approaches is the speed and efficiency at which these methods can be executed (Engel et al., 2021; Riedel et al., 2013). The genomic, reference-based approach we used here is not necessarily speedier or more efficient than other delimitation approaches. This approach requires researchers to collect comparative population genomic datasets across the reference species set, which can be non-trivial for rare or elusive taxa and can be costly as well. For example, in our own study, our own reference set consisted of only two nominal species (*H. maculata* and *H. propinqua*) due to the focus on populations/species in Texas, and because adding more divergent species to our ddRAD dataset would have reduced the number of homologous loci we would have been able to assemble and add to our dataset. For future studies, a sequencing method (e.g., target capture, low-coverage whole genome sequencing) that allows for the reference taxon set to be more comprehensive would be ideal (Leaché et al., 2021; Tobias et al., 2010). Regardless of the sequencing method, these approaches require specialized bioinformatics knowledge, and processing data can sometimes be slow. But, as pipelines for molecular delimitation methods are honed and made more publicly available, as these molecular delimitation methods are improved to be more biologically relevant (e.g., including gene flow or other evolutionary processes into models) in how they estimate species boundaries, and as methods are made more computationally efficient in handling larger genomic datasets, a reference-based approach has the potential to become more efficient. Further, future studies of putative species in *Holbrookia* and other closely related genera can build on this existing dataset and thus be completed more efficiently.

Further, in order to ensure a robust and holistic taxonomy, these genomic delimitation approaches should be combined with integrative data types (e.g., analyses of morphological, behavioral, and mating patterns across putative taxonomic units). Such data can help determine if this genomic divergence correlates to phenotypic changes that are likely to isolate taxonomic units. Although implementing these approaches does not require the specialized skills of trained taxonomists (e.g., morphological and diagnostic characterizations), translating the results from these approaches into updated taxonomies does. Ultimately, we must train more taxonomists who can formally diagnose new species based on several traits (genetics, morphology, behavior, ecology), especially if we want our species delimitation results to inform conservation decisions.

Acknowledgements

We would like to thank biologists from the Texas Parks and Wildlife Department, as well as the private and public landowners and land managers who granted us access to their land for lizard surveys. This work was funded by a Texas Parks and Wildlife Section Six Grant awarded to CER and MKF (TPWD 474241), National Science Foundation grants awarded to MKF (DEB-SBS-2024014), ADL (DEB-SBS-2023723), and SS (DEB-SBS-2023979), and a contract funded by the State Comptroller of Public Accounts (IAC-14000679). Some of this material is based upon work while MKF served at the National Science Foundation. Animal research on this project was approved by the University of Texas at Arlington's IACUC (#A16.010) and collec-

tion of animals in Texas was carried out under Scientific Permit Number SPR-0814-159.

Data Accessibility

All raw, unprocessed sequences were deposited in NCBI Sequence Read Archive (BioProject PRJNA894589; BioSample accession numbers SAMN31467436-SAMN3147501) for ddRADseq data. We have included a large package on DRYAD (https://datadryad.org/stash/share/aaCMVureaq1rYnbrVrjIKC5noFIM4DKBg9-_Pvx5jtY) that includes ddRADseq filtered "haplotypes" files, *ND2* alignment files, and resulting input files for several analysis programs (STRUCTURE, ADMIXTURE, DAPC, fineRADstructure, BFD*, BPP, *gdi*, DELINEATE, BEAST, and SNAPP).

Author Contributions

M.K.F., C.E.R., T.J.F., S.S., and A.D.L. conceptualized and designed the research project. M.K.F., C.E.R., N.D.R., W.A.R., T.J.H., and T.J.L. performed fieldwork to obtain samples. T.J.F. collected, processed, and analyzed all of the molecular data. T.J.F. wrote the manuscript with contributions from M.K.F., S.S., A.D.L., and C.E.R., and all authors approved the final manuscript.

Submitted: March 15, 2023 EDT, Accepted: April 02, 2024 EDT

References

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109

Archer, S. (1989). Have southern Texas savannas been converted to woodlands in recent history? *American Naturalist*, *134*(4), 545–561. https://doi.org/10.1086/284996

Axtell, R. A. (1956). A solution to the long neglected Holbrookia lacerata problem, and the description of two new subspecies of Holbrookia. *Bulletin of the Chicago Academy of Science*, *10*, 163–181. https://webapps.fhsu.edu/ksHerp/bibFiles/854.pdf

Axtell, R. A. (1958). *A monographic revision of the Iguanid genus Holbrookia* [Ph.D Dissertation]. The University of Texas.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., & Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*, 1–6. https://doi.org/10.1371/journal.pcbi.1003537

Bouckaert, R. R. (2010). DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, *26*, 1372–1373. https://doi.org/10.1093/bioinformatics/btq110

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, *29*, 1917–1932. https://doi.org/10.1093/molbev/mss086

Burbrink, F. T., & Ruane, S. (2021). Contemporary philosophy and methods for studying speciation and delimiting species. *Ichthyology and Herpetology*, *109*(3), 874–894. https://doi.org/10.1643/h2020073

Camper, J. D., & Dixon, J. R. (1994). Geographic variation and systematics of the striped whipsnakes (Masticophis taeniatus complex; Reptilia: Serpentes: Colubirdae). *Annals of the Carnegie Museum*, *63*(1), 1–48. https://doi.org/10.5962/p.226641

Campillo, L. C., Barley, A. J., & Thomson, R. C. (2020). Model-based species delimitation: are coalescent species reproductively isolated? *Systematic Biology*, *69*, 708–721. https://doi.org/10.1093/sysbio/syz072

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology*, *22*, 4369–4383. https://doi.org/10.1111/mec.12413

Ceballos, G., Ehrlich, P. R., & Rodolfo, D. (2015). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences USA*, *114*(30), 6089–6096. https://doi.org/10.1073/pnas.1704949114

Chan, K. O., & Grismer, L. L. (2019). To split or not to split? Multilocus phylogeny and molecular species delimitation of southeast Asian toads (family: Bufonidae). *BMC Evolutionary Biology*, *95*, 1–12. https://doi.org/10.1186/s12862-019-1422-3

Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society*, *85*(3), 407–417. https://doi.org/10.1111/j.1095-8312.2005.00503.x

de Queiroz, K. (1992). Phylogenetic relationships and rates of allozyme evolution among the lineages of sceloporine sand lizards. *Biological Journal of the Linnean Society*, *45*(4), 333–362. https://doi.org/10.1111/j.1095-8312.1992.tb00648.x

de Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*(6), 879–886. https://doi.org/10.1080/10635150701701083

Dupuis, J. R., Roe, A. D., & Sperling, F. A. H. (2012). Multi-locus species delimitation in closely-related animals and fungi: one marker is not enough. *Molecular Ecology*, *21*(18), 4422–4436. https://doi.org/10.1111/j.1365-294X.2012.05642.x

Duran, M., & Axtell, R. W. (2010). *A Rangewide Inventory and Habitat Model for the Spot-tailed Earless Lizard, Holbrookia lacerata*. Report submitted to Texas Parks and Wildlife Department.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTUREHARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340

Engel, M. S., Ceríaco, L. M., Daniel, G. M., Dellapé, P. M., Löbl, I., Marinov, M., ... Zacharie, C. K. (2021). The taxonomic impediment: a shortage of taxonomists, not the lack of technical approaches. *Zoological Journal of the Linnean Society*, *193*(2), 381–387. https://doi.org/10.1093/zoolinnean/zlab072

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, *164*(4), 1567–1587. https://doi.org/10.1093/genetics/164.4.1567

Firneno, T. J., Jr., O'Neill, J. R., Itgen, M. W., Kihneman, T. A., Townsend, J. H., & Fujita, M. K. (2021). Delimitation despite discordance: Evaluating the species limits of a confounding species complex in the face of mitonuclear discordance. *Ecology and Evolution*, *11*(18), 12739–12753. https://doi.org/10.1002/ece3.8018

Flouri, T., Jiao, X., Rannala, B., & Yang, Z. (2020). A Bayesian implementation of the multispecies coalescent model with introgression for phylogenomic analysis. *Molecular Biology and Evolution*, *37*, 1211–1223. https://doi.org/10.1093/molbev/msz296

Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology and Evolution*, *27*, 480–488. https://doi.org/10.1016/j.tree.2012.04.012

Hebert, P. D., & Gregory, T. R. (2005). The promise of DNA barcoding for taxonomy. *Systematic Biology*, *54*(5), 852–859. https://doi.org/10.1080/10635150500354886

Hebert, P. D., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *Proceedings of the National Academy of Sciences*, *101*(41), 14812–14817. https://doi.org/10.1073/pnas.0406166101

Hey, J., & Pinho, C. (2012). Population genetics and objectivity in species diagnosis. *Evolution*, *66*(5), 1413–1429. https://doi.org/10.1111/j.1558-5646.2011.01542.x

Hibbitts, T. J., Ryberg, W. A., Harvey, J. A., Voelker, G., Lawing, A. M., Adams, C. S., Neuharth, D. B., Dittmer, D. E., Duran, C. M., Wolaver, B. D., & Pierre, J. P. (2019). Phylogenetic structure of Holbrookia lacerata (Cope 1880) (Squamata: Phrynosomatidae): one species or two. *Zootaxa*, *4619*(1), 139–154. https://doi.org/10.11646/zootaxa.4619.1.6

Hibbitts, T. J., Walkup, D. K., LaDuc, T. J., Wolaver, B. D., Pierre, J. P., Duran, M., Neuharth, D., Frizzell, S., Adams, C. S., Johnson, T. E., & Yandell, D. (2021). Natural history of the spot-tailed earless lizards (Holbrookia lacerata and H. subcaudalis). *Journal of Natural History*, *55*(7–8), 495–514. https://doi.org/10.1080/00222933.2021.1907469

Hill, R. T., & Vaughan, T. W. (1898). Geology of the Edwards Plateau and Rio Grande Plain Adjacent to Austin and San Antonio, Texas. In *U.S. Geological Survey 18th Annual Report* (pp. 103–144).

Hillis, D. M., Chambers, E. A., & Devitt, T. J. (2021). Contemporary Methods and Evidence for Species Delimitation. *Ichthyology and Herpetology*, *109*(3), 895–903. https://doi.org/10.1643/h2021082

Hooper, D. U., Adair, E. C., Cardinale, B. J., Byrnes, J. E., Hungate, B. A., Matulich, K. L., Gonzalez, A., Duffy, J. E., Gamfeldt, L., & O'Connor, M. I. (2012). A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature*, *486*(7401), 105–108. https://doi.org/10.1038/nature11118

Huang, J., Bennett, J., Flouri, T., Leaché, A. D., & Yang, Z. (2022). Phase Resolution of Heterozygous Sites in Diploid Genomes is Important to Phylogenomic Analysis under the Multispecies Coalescent Model. *Systematic Biology*, *71*(2), 334–352. https://doi.org/10.1093/sysbio/syab047

Huang, J. P. (2021). The Genealogical Divergence Index Across a Speciation Continuum in Hercules Beetles. *Insect Systematics and Diversity*, *5*(6), 1. https://doi.org/10.1093/isd/ixab021

Hudson, R. R., & Turelli, M. (2003). Stochasticity overrules the "three-times rule": genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution*, *57*(1), 182–190. https://doi.org/10.1111/j.0014-3820.2003.tb00229.x

Jackson, N. D., Morales, A. E., Carstens, B. C., & O'Meara, B. C. (2017). PHRAPL: Phylogenetic inference using approximate likelihoods. *Systematic Biology*, *66*, 1045–1053. https://doi.org/10.1093/sysbio/syx001

Ji, J., Jackson, D. J., Leaché, A. D., & Yang, Z. (2023). Power of Bayesian and Heuristic Tests to Detect Cross-Species Introgression with Reference to Gene Flow in the Tamias quadrivittatus Group of North American Chipmunks. *Systematic Biology*, *72*, 446–465. https://doi.org/10.1093/sysbio/syac077

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jones, G., Aydin, Z., & Oxelman, B. (2015). DISSECT: an assignment-free Bayesian discovery method for species delimitation under the multispecies coalescent. *Bioinformatics*, *31*, 991–998. https://doi.org/10.1093/bioinformatics/btu770

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., & Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics*, *33*(11), 1630–1638. https://doi.org/10.1093/bioinformatics/btx025

Knowles, L. L., & Carstens, B. C. (2007). Delimiting Species without Monophyletic Gene Trees. *Systematic Biology*, *56*(6), 887–895. https://doi.org/10.1080/10635150701701091

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*, 1870–1874. https://doi.org/10.1093/molbev/msw054

Leaché, A. D., Davis, H. R., Singhal, S., Fujita, M. K., Lahti, M. E., & Zamudio, K. R. (2021). Phylogenomic Assessment of Biodiversity Using a Reference-Based Taxonomy: An Example With Horned Lizards (Phrynosoma). *Frontiers in Ecology and Evolution*, *2021*, 437. https://doi.org/10.3389/fevo.2021.678110

Leaché, A. D., Fujita, M. K., Minin, V. N., & Bouckaert, R. R. (2014). Species delimitation using genome-wide SNP data. *Systematic Biology*, *63*, 534–542. https://doi.org/10.1093/sysbio/syu018

Leaché, A. D., McElroy, M. T., & Trinh, A. (2018). A genomic evaluation of taxonomic trends through time in coast horned lizards (genus Phrynosoma). *Molecular Ecology*, *27*(13), 2884–2895. https://doi.org/10.1111/mec.14715

Leaché, A. D., & Oaks, J. R. (2017). The utility of single-nucleotide polymorphism (SNP) data in phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, *48*, 69–84. https://doi.org/10.1146/annurev-ecolsys-110316-022645

Leaché, A. D., Zhu, T., Rannala, B., & Yang, Z. (2019). The spectre of too many species. *Systematic Biology*, *68*, 168–181. https://doi.org/10.1093/sysbio/syy051

Maldonado, J. A., Firneno, T. J., Jr., Roelke, C. E., Rains, N. D., Mwgiri, J., & Fujita, M. K. (2020). Transcriptome sequencing reveals signatures of positive selection in the Spot-Tailed Earless Lizard. *PloS ONE*, *15*(6), e0234504. https://doi.org/10.1371/journal.pone.0234504

Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population inference for RADseq data. *Molecular Biology and Evolution*, *35*(5), 1284–1290. https://doi.org/10.1093/molbev/msy023

Meirmans, P. G. (2006). Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*, *60*(11), 2399–2402. https://doi.org/10.1111/j.0014-3820.2006.tb01874.x

Nei, M. (1987). *Molecular evolutionary genetics*. Columbia University Press. https://doi.org/10.7312/nei-92038

Padial, J. M., & de la Riva, I. (2009). Integrative taxonomy reveals cryptic Amazonian species of Pristimantis (Anura: Strabomantidae). *Zoological Journal of the Linnean Society*, *155*, 97–122. https://doi.org/10.1111/j.1096-3642.2008.00424.x

Padial, J. M., Miralles, A., de la Riva, I., & Vences, M. (2010). The integrative future of taxonomy. *Frontiers in Zoology*, *7*(16), 2010. https://doi.org/10.1186/1742-9994-7-16

Pante, E., Schoelinck, C., & Puillandre, N. (2015). From integrative taxonomy to species description: one step beyond. *Systematic Biology*, *64*(1), 152–160. https://doi.org/10.1093/sysbio/syu083

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Pinho, C., & Hey, J. (2010). Divergence with gene flow: models and data. *Annual Reviews in Ecology, Evolution, and Systematics*, *41*, 215–230. https://doi.org/10.1146/annurev-ecolsys-102209-144644

Poelstra, J. W., Salmona, J., Tiley, G. P., Schübler, D., Blanco, M. B., Andriambeloson, J. B., ... Yoder, A. D. (2021). Crytpic patterns of speciation in cryptic primates: microendemic mouse lemurs and the multispecies coalescent. *Systematic Biology*, *70*(2), 203–218. https://doi.org/10.1093/sysbio/syaa053

Portik, D. M., Leaché, A. D., Rivera, D., Barej, M. F., Burger, M., Hirschfeld, M., Rödel, M. O., Blackburn, D. C., & Fujita, M. K. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Molecular Ecology*, *26*, 5245–5263. https://doi.org/10.1111/mec.14266

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959. https://doi.org/10.1093/genetics/155.2.945

Rambaut, A., & Drummond, A. (2009). *Tracer v1.5.0*. http://beast.bio.ed.ac.uk/

Rannala, B. (2015). The art and science of species delimitation. *Current Zoology*, *61*(5), 846–853. https://doi.org/10.1093/czoolo/61.5.846

Rannala, B., & Yang, Z. (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, *164*, 1645–1656. https://doi.org/10.1093/genetics/164.4.1645

Ratnasingham, S., & Hebert, P. D. (2007). BOLD: The Barcode of Life Data System (http://www.Barcodinglife.Org). *Molecular Ecology Notes*, *7*(3), 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Reich, D., Thangaraj, K., Patterson, N., Price, A. L., & Singh, L. (2009). Reconstructing Indian population history. *Nature*, *461*, 489–494. https://doi.org/10.1038/nature08365

Riedel, A., Sagata, K., Suhardjono, Y. R., Tänzler, R., & Balke, M. (2013). Integrative taxonomy on the fast track-towards more sustainability in biodiversity research. *Frontiers in Zoology*, *10*(1), 1–9. https://doi.org/10.1186/1742-9994-10-15

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737–4754. https://doi.org/10.1111/mec.15253

Roelke, C. E., Maldonado, J. A., Pope, B. W., Firneno, T. J., Jr., Laduc, T. J., Hibbitts, T. J., Ryberg, W. A., Rains, N. D., & Fujita, M. K. (2018). Mitochondrial genetic variation within and between Holbrookia lacerata lacerata and Holbrookia lacerata subcaudalis, the spot-tailed earless lizards of Texas. *Journal of Natural History*, *52*(13–16), 1017–1027. https://doi.org/10.1080/00222933.2018.1436726

Sambrook, J., & Russell, D. W. (2006). Purification of nucleic acids by extraction with phenol:chloroform. In J. Sambrook & D. W. Russell (Eds.), *Commonly Used Techniques in Molecular Cloning* (p. Appendix8). Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/pdb.prot4045

Schulte, J. A., II, & de Queiroz, K. (2008). Phylogenetic relationships and heterogeneous evolutionary processes among phrynosomatine sand lizards (Squamata, Iguanidae) revisited. *Molecular Phylogenetics and Evolution*, *47*(2), 700–716. https://doi.org/10.1016/j.ympev.2008.01.010

Scott, N. J. (1996). Evolution and management of the North American grassland herpetofauna. In D. M. Finch (Ed.), *Ecosystem disturbance and wildlife conservation in western grasslands—A symposium proceedings* (pp. 40–53). USDA Forest Service, Rocky Mountain Forest and Range Experiment Station.

Sites, J. W., & Marshall, J. C. (2003). Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology and Evolution*, *18*(9), 462–470. https://doi.org/10.1016/S0169-5347(03)00184-8

Smith, M. L., & Carstens, B. C. (2020). Process-based species delimitation leads to identification of more biologically relevant species. *Evolution*, *74*(2), 216–229. https://doi.org/10.1111/evo.13878

Smith, M. L., & Carstens, B. C. (2022). Species Delimitation Using Molecular Data. In J. S. Wilkins, F. E. Zachos, & I. Y. Pavlinov (Eds.), *Species Problems and Beyond* (pp. 145–160). CRC Press. https://doi.org/10.1201/9780367855604-9

Streicher, J. W., Devitt, T. J., Golberg, C. S., Malone, J. H., Blackmon, H., & Fujita, M. K. (2014). Diversification and asymmetrical gene flow across time and space: Lineage sorting and hybridization in polytypic barking frogs. *Molecular Ecology*, *23*(13), 3273–3291. https://doi.org/10.1111/mec.12814

Sukumaran, J., Holder, M. T., & Knowles, L. L. (2021). Incorporating the speciation process into species delimitation. *PLoS Computational Biology*, *17*(5), e1008924. https://doi.org/10.1371/journal.pcbi.1008924

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Science USA*, *114*, 1607–1612. https://doi.org/10.1073/pnas.1607921114

Sullivan, J. P., Lundberg, J. G., & Hardman, M. (2006). A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear gene sequences. *Molecular Phylogenetics and Evolution*, *41*(3), 636–662. https://doi.org/10.1016/j.ympev.2006.05.044

Tobias, J. A., Seddon, N., Spottiswoode, C. N., Pilgrim, J. D., Fishpool, L. D., & Collar, N. J. (2010). Quantitative criteria for species delimitation. *Ibis*, *152*(4), 724–746. https://doi.org/10.1111/j.1474-919X.2010.01051.x

USFWS. (2011). Endangered and threatened wildlife and plants; 90-day finding on petition to list the Spot-tailed Earless Lizard as endangered or threatened. *Federal Register*, *76*(100), 30082–30087.

Wake, D. B., & Vredenburg, V. T. (2008). Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences USA*, *105*, 11466–11473. https://doi.org/10.1073/pnas.0801921105

Weeks, A. W. (1945). Balcones, Luling, and Mexia fault zones in Texas. *American Association of Petroleum Geologists Bulletin*, *29*, 1733–1737. https://doi.org/10.1306/3D9337B0-16B1-11D7-8645000102C1865D

Wiens, J. J., Kozak, K. H., & Silva, N. (2013). Diversity and niche evolution along aridity gradients in North American lizards (Phyrnosomatidae). *Evolution*, *67*(6), 1715–1728. https://doi.org/10.1111/evo.12053

Wilgenbusch, J., & de Queiroz, K. (2000). Phylogenetic relationships among the phrynosomatid sand lizards inferred from mitochondrial DNA sequences generated by heterogeneous evolutionary processes. *Systematic Biology*, *49*(3), 592–612. https://doi.org/10.1080/10635159950127411

Will, K. W., Mishler, B. D., & Wheeler, Q. D. (2005). The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology*, *54*(5), 844–851. https://doi.org/10.1080/10635150500354878

Wolaver, B. D., Pierre, J. P., Labay, B. J., LaDuc, T. J., Duran, C. M., Ryberg, W. A., & Hibbitts, T. J. (2018). An approach for evaluating changes in land-use from energy sprawl and other anthropogenic activities with implications for biotic resource management. *Environmental Earth Sciences*, *77*(5), 1–4. https://doi.org/10.1007/s12665-018-7323-8

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences USA*, *107*(20), 9264–9269. https://doi.org/10.1073/pnas.0913022107

Yatkola, D. A. (1976). Mid-Miocene lizards from western Nebraska. *Copeia*, *1974*(4), 645–654. https://doi.org/10.2307/1443444

Zhang, C., Zhang, D. X., Zhu, T., & Yang, Z. (2011). Evaluation of a Bayesian coalescent model of species delimitation. *Systematic Biology*, *60*, 747–761. https://doi.org/10.1093/sysbio/syr071