Genome-scale data reveal deep lineage divergence and a complex demographic history in the Texas horned lizard (*Phrynosoma cornutum*) throughout the southwestern and central US

NICHOLAS FINGER[1], KEAKA FARLEIGH[2], JASON T. BRACKEN[2], ADAM D. LEACHÉ[3], OLIVIER FRANÇOIS[4], ZIHENG YANG[5], TOMAS FLOURI[5], TRISTAN CHARRAN[1], TEREZA JEZKOVA[2], DEAN A. WILLIAMS[6], CHRISTOPHER BLAIR[1,7]

[1]Department of Biological Sciences, New York City College of Technology, The City University of New York, 285 Jay Street, Brooklyn, NY 11201, USA

[2]Department of Biology, Miami University, 501 E High St, Oxford, OH 45056, USA

[3]Department of Biology & Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195, USA

[4]Faculty of Medicine, University Grenoble-Alpes, TIMC-IMAG UMR 5525, 38000 Grenoble, F38706 La Tronche, France

[5]Department of Genetics, Evolution and Environment, University College London, Darwin Building, Gower Street, London WC1E 6BT, UK

[6]Department of Biology, Texas Christian University, 2800 S University Dr, Fort Worth, TX 76129, USA

[7]Biology PhD Program, CUNY Graduate Center, 365 5th Ave, New York, NY 10016, USA

Correspondence: cblair@citytech.cuny.edu

RH: Phylogeography of the Texas horned lizard

© The Author(s) 2021. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Abstract

The southwestern and central US serve as an ideal region to test alternative hypotheses regarding biotic diversification. Genomic data can now be combined with sophisticated computational models to quantify the impacts of paleoclimate change, geographic features, and habitat heterogeneity on spatial patterns of genetic diversity. In this study we combine thousands of genotyping-by-sequencing (GBS) loci with mtDNA sequences (ND1) from the Texas Horned Lizard (*Phrynosoma cornutum*) to quantify relative support for different catalysts of diversification. Phylogenetic and clustering analyses of the GBS data indicate support for at least three primary populations. The spatial distribution of populations appears concordant with habitat type, with desert populations in Arizona and New Mexico showing the largest genetic divergence from the remaining populations. The mtDNA data also support a divergent desert population, but other relationships differ and suggest mtDNA introgression. Genotype-environment association with bioclimatic variables support divergence along precipitation gradients more than along temperature gradients. Demographic analyses support a complex history, with introgression and gene flow playing an important role during diversification. Bayesian multispecies coalescent analyses with introgression (MSci) analyses also suggest that gene flow occurred between populations. Paleo-species distribution models support two southern refugia that geographically correspond to contemporary lineages. We find that divergence times are underestimated and population sizes are over-estimated when introgression occurred and is ignored in coalescent analyses, and furthermore, inference of ancient introgression events and demographic history is sensitive to inclusion of a single recently admixed sample. Our analyses cannot refute the riverine barrier or glacial refugia hypotheses. Results also suggest that populations are continuing to diverge along habitat gradients. Finally, the strong evidence of admixture, gene flow, and mtDNA introgression among populations suggests that *P. cornutum* should be considered a single widespread species under

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

the General Lineage Species Concept.

Key words: demography, introgression, lizards, phylogeography, speciation

Statement of Significance

Many studies have documented cryptic diversity in diverse taxa inhabiting the arid regions of western North America, with divergence correlated with both Neogene vicariance and Pleistocene climate change. However, relatively few studies adopt a genomics approach and most implicitly assume that gene flow ceases once divergence begins. Using the Texas horned lizard (*Phrynosoma cornutum*) as a model, our results suggest a complex demographic history that includes episodes of gene flow. Results also suggest that divergence is continuing along environmental axes and that adequate model choice is imperative for demographic hypothesis testing. This study can serve as a model for how genomic data and new analytical tools can be used to test traditional evolutionary hypotheses throughout geologically and climatically diverse regions.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Introduction

Allopatric divergence has long been considered the most likely cause of speciation, and geographic barriers the primary hindrance to gene flow (Coyne & Orr, 2004). However, the origins of a particular diversification event can be both controversial and unclear, resulting in the various forces behind diversification becoming a current topic for discussion (Pyron & Burbrink, 2010; Butlin et al., 2008; Fitzpatrick et al., 2009; Nosil & Feder, 2012). Not only may the forces acting on species be disparate, but the diversification process can be episodic with periods of isolation interspersed with periods of gene flow leading to a history of reticulation (Blair & Ané 2020). As the climate changes, a population may fracture by seeking shrinking patches of ideal habitat, expand into newly habitable regions, or adapt, the latter of which can lead to niche divergence and ecological segregation (Castro-Insua et al., 2018; Jezkova et al., 2016; Wiens & Graham, 2005). As a species expands or contracts its range, it may encounter hard barriers to gene flow such as rivers, which have been shown to result in genetic divergence in multiple taxa (Pastorini et al. 2003; Nazereno et al. 2019). Populations and species likely to encounter disruptive barriers throughout their history tend to occupy a wide geographic range of varied habitat, yet possess low dispersal capabilities (Schield et al., 2018). Ectothermic species such as reptiles that exhibit these traits are also further influenced by climate differences (Huey & Kingsolver, 1993; Wogan & Richmond, 2015). Ultimately, understanding the evolutionary history of a species involves evaluating the geographic, genetic and climatic factors affecting divergence throughout its history (Fitzpatrick et al., 2009).

The Texas Horned Lizard (*Phrynosoma cornutum*) is spread across a diverse collection of ecological habitats making it an interesting candidate to examine adaptation and phylogeographic history. While its range does consist of many smaller environmental niches (Price, 1990), there exists a primary habitat divide that bisects the species' distribution providing an apparently stark environmental contrast through which to view its effects on the species. The

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

southwestern range inhabits the Chihuahuan desert of Arizona and New Mexico, whereas the northeastern range covers the Great Plains east of the Rocky Mountains throughout Texas, Oklahoma, and Kansas extending the furthest east of any horned lizard (Sherbrooke, 2003). As expansive as the range is, *P. cornutum* lives a sedentary life, maintaining fidelity to a home range with daily movement < 250 meters and limited long distance dispersal capabilities (Fair & Henke, 1999). The combined factors of the species' large geographic distribution, low dispersal ability and varied ecological niche (with respect to various environmental variables such as temperature and precipitation) across the range may increase the likelihood of regional adaptation (Lenormand, 2002; Newman & Austin, 2015). Of particular note is the broad range of annual precipitation values, from ~10 inches per year in the western deserts to ~50 inches per year in the Great Plains (Pittman et al., 2007). *Phrynosoma cornutum* has also developed mechanisms for water harvesting involving both behavioural and morphological adaptations (Sherbrooke, 1990). The lizard will adopt a rain-harvesting stance, spreading the dorsal surface so as to maximize retention of raindrops which are then carried through interscalar channels to the mouth (Sherbrooke, 1990). These behavioural and morphological adaptations are shared with other *Phrynosoma* (*P. modestum* and *P. platyrhinos*) inhabiting similar arid ecological niches (Sherbrooke, 1990; Sherbrook, 2003), and suggest that there may be clines in allele frequencies that are partially tied to temperature and/or precipitation.

With the uniqueness of these adaptations, along with their status in historical accounts and importance in use as symbols and mascots, *Phrynosoma* spp. have been the subject of interest in many evolutionary studies (Leaché & Linkem, 2015; Leaché & McGuire, 2006; Williams et al., 2019). The crown group of *Phrynosoma* diverged roughly 25 Ma and the genus now contains 17 species after the addition of three new species over the past decade. Recent studies focusing on the genetic structure and lineage divergence within the various species (Blair & Bryson, 2017; Bryson et al., 2012; Jezkova et al., 2016; Montanucci, 2015; Mulcahy et al., 2006) yielded the discovery of these three new additions, *P. cerroense, P. blainvilli and P.*

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

sherbrookei, to the taxonomy (de Oca et al., 2014; Leache et al., 2009). Previously, relationships both between and within species have been difficult to untangle due to hybridization, introgression, and incomplete lineage sorting (ILS) resulting in disagreement between concatenation vs coalescent-based methods, as well as discordance between trees inferred using mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) (de Oca et al., 2014; Leaché & McGuire, 2006). With the advent of reduced representation sequencing providing a random and more diverse view of the genome (Andrews et al., 2016), we are able to overcome these previous challenges in discerning phylogenetic and phylogeographic relationships caused by mtDNA introgression and gene tree/species tree discordance (Leaché & Linkem, 2015; Leaché et al., 2015). Given the comparatively large geographic range of *P. cornutum*, and the lack of genomic assessment across diverse habitats, the possibility of cryptic diversity is high.

A previous study of this species found strong divergence between the western desert and eastern plains populations using mtDNA data (Williams et al. 2019). It was hypothesized that the presence of an extensive late Pliocene pluvial lake, Lake Cabeza de Vaca, was the barrier that originally separated these two clades. Both clades gave a signal of population expansion in the Pleistocene. Nuclear microsatellite loci also revealed strong divergence between the western and eastern mitochondrial clades and found that the eastern plains were further subdivided into the South-Central Semi-Arid Prairies to the north of the Balcones Escarpment and the Southern Texas Plains south of the Escarpment (Williams et al. 2019). Although these results further advance our understanding of evolutionary pattern and process throughout the central-southern US, a genomic approach that takes advantage of sophisticated new analytical tools would provide additional power to disentangle competing hypotheses regarding historical and contemporary divergence.

In this study, we expand on previous results by including samples from more northern areas of the species range (Kansas and Oklahoma) and by examining the phylogeographic and demographic history of *P. cornutum* using both mtDNA sequences and thousands of nuclear

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

SNPs from a modified genotyping-by-sequencing (GBS) approach. We first use concatenated and coalescent-based phylogenetic analyses, species delimitation analyses, and clustering to test the hypothesis that the genomic and mtDNA data support the presence of cryptic diversity, which has been demonstrated in other species of *Phrynosoma* with large geographic distributions. Second, we use genotype-environment association analyses (GEA) to test the hypothesis that a proportion of SNPs are statistically correlated with bioclimatic variables and that the environmental gradient between the plains and desert habitat may be driving adaptation and furthering genetic divergence (McDonald, 1983; Wiens et al., 2013). We then adopt an explicit hypothesis testing framework to elucidate demographic history, testing three hypotheses of divergence likely important to the species. Specifically, we use our models to assess the relative importance of the Rio Grande as a hard allopatric barrier to gene flow between divergent lineages (Lanna et al., 2020), as compared to soft allopatric divergence due to cyclical paleoclimate change or ecological gradients. Both present day and historical species distribution models (SDMs) are used to further test the hypothesis that divergence was driven by Pleistocene climate fluctuations (Hewitt, 1996, 2000) as has been observed for other inhabitants in the region (Jezkova et al., 2016; Schield et al., 2015). Finally, we test the hypothesis that explicitly accommodating gene flow in Bayesian multispecies coalescent analyses (MSci; Flouri et al., 2020), leads to alternative estimates of demographic history (i.e. divergence times and effective population sizes).

Results

Data set characteristics

We obtained approximately 225 megabases of nGBS data from 75 *P. cornutum* samples and a single *P. solare* outgroup. After processing the data in ipyrad (Eaton & Overcast, 2020),

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

most individuals had *ca*. 30,000 loci (4,757–42,652; Supplementary Table S1). The full concatenated matrix consisted of 7,906,017 bp and 57,459 loci. The final mtDNA alignment consisted of 1,330 bp, 119 variable (but parsimony uninformative) sites, and 101 parsimony informative sites across 74 sequences including a single *P. solare* sequence used as the outgroup. Excluding the outgroup resulted in 27 variable (parsimony uninformative) sites and 100 parsimony informative characters.

Phylogenetic analysis

We used multiple phylogenetic analyses to test for the presence of cryptic lineages and elucidate the relationships among them. Concatenated ML analysis in RAxML-ng (Kozlov et al., 2019) yielded a topology consisting of three primary lineages (Figs. 1, 2). These lineages included a Desert clade (DST) consisting of samples from the Arizona and New Mexico portions of the Chihuahuan Desert (N. American Eco Region 10: North American desert), a Southern clade (STH) containing samples from the southern Texas plains (N. American Eco Region 9: Great Plains) and a Plains clade (PLN) of samples from Western Nevada, Northern Texas, Colorado, Kansas, and Oklahoma (N. American Eco Region 9: Great Plains; Fig. 1). The Desert lineage was supported by a bootstrap value of 100%, the Southern Lineage had a bootstrap value of 81% and the Plains lineage was also supported by 100% bootstrap value. The average relative Robinson-Foulds (RF) distance in this tree set was 0.079466 and the number of unique topologies in the tree set was 10. In all cases the three primary clades were recovered. Bayesian analysis in ExaBayes (Aberer et al., 2014) resulted in a nearly identical topology to the ML tree with 100% posterior probability for the three distinct lineages (Fig. 2). ESS values for all parameters indicated that the chain was run for an adequate duration (ESS > 200 for all parameters). Both the ML and Bayesian analyses provided some additional support for two lineages within the Plains clade. The bootstrap consensus tree from SVDQUARTETS (Chifman & Kubatko, 2014) yielded a topology consistent with the ML and Bayesian trees (Fig. 2).

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Bootstrap support for each clade was 100%. However, this topology did not support two distinct Plains lineages (Supplementary Fig. S1).

Bayesian analysis of the mtDNA data in BEAST (Bouckaert et al., 2019) yielded high ESS values for all parameters (>200). The coefficient of variation parameter under a relaxed clock model (which measures the extent of clock violation) had substantial posterior density near zero, indicating that a strict clock model was appropriate. The maximum clade credibility (MCC) tree showed a different tree topology compared to the three GBS based trees discussed above. The Desert clade was still present and strongly supported (minus sample KK104), but the remaining topology did not support a distinctive Southern or Plains population. Instead, individuals from the Southern and Plains populations were interspersed throughout two lineages that diverged approximate 1 Ma (assuming a substitution rate of 0.00805 substitutions per site per million years [Macey et al., 1999]. The mtDNA genealogy supported an initial divergence time of approximately 5 Ma for *P. cornutum* (Supplementary Fig. S2).

Population structure and GEA analysis

To complement the phylogenetic analyses, we performed genetic clustering using sNMF in the R package LEA (Frichot et al., 2014; Frichot & François, 2015). After filtering missing data and SNPs showing evidence of linkage disequilibrium from the initial matrix of 54,634 SNPs, population genomic analyses in sNMF provided support for K = 5 genetic groups (Fig. 2b,c; Supplementary Figs. S3, S4) based on the cross-entropy criterion. Results were similar to the phylogenetic analyses, showing strong evidence for the western Desert (DST) cluster with strong geographic structure, a small Southern (STH) population and a third larger Plains (PLN) population consisting of three subpopulations (Plains South, Plains Central, Plains North), with substantial shared ancestry amongst them (Fig. 2). We chose to treat the data as three populations for demographic modeling rather than five to focus on the deepest divergences from the phylogenetic analysis. Further, the additional structure detected with K = 5 likely

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

represented isolation by distance (IBD; see below). The major split between two groups separating the western (DST) and eastern (STH+PLN) populations (*K* = 2) was recovered in virtually all analyses, and runs with the lowest cross-entropy levels supported the partition shown in Fig. 2. For all demographic modeling (i.e. BPP, MOMENTS) we defined two sets of analyses on a reduced subset of individuals, one including sample KK104 (admixed) and one without (non-admixed). We focused on this individual for several reasons: (1) it was the only sample included in the analyses where <50% of its genome traced back to a single ancestral population (Fig. 2b); (2) the genomic background for the individual spanned two divergent lineages (Fig. 2); (3) this individual was placed in a mixed STH+PLN lineage based on the mtDNA data (Supplementary Fig. S2). These results were likely because the individual was captured near the boundary of two lineages (see Discussion for additional information). For all analyses, we compared models and parameter estimates to quantify the impact of this individual on the results.

Pairwise F_{st} and Nei's genetic distance estimates supported the split between the two groups inferred from the phylogenetic and sNMF analyses, separating the western (DST) and eastern populations (STH+PLN). Both F_{st} and genetic distance were higher between western and eastern populations than between the two eastern populations (Supplementary Table S2). Genetic distance within populations was higher among eastern populations than the western population (Supplementary Table S2). Analysis of spatial genetic structure revealed a significant pattern of isolation by distance ($p < 0.001$; Supplementary Fig. S5).

Our next objective was to test for a statistical association between SNPs and environmental gradients (genotype-environment association; GEA), which can provide evidence that these lizards may be adapting to divergent climatic conditions. Correlations between SNPs and environmental variables was first performed through redundancy analysis using the R package vegan. Our global model and first of two redundancy axes were significant ($P < 0.05$). The global model had an adjusted R^2 of 0.017. RDA identified 29 outlier SNPs based on locus

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

scores that were ± 2.5 SD, eight associated with mean temperature of the driest quarter and 21 associated with precipitation seasonality (Fig. 3a). Individuals from our Desert population showed a positive relationship with BIO15: precipitation seasonality, and individuals in our Central Plains subpopulation exhibited a negative relationship with BIO9: mean temperature of the driest quarter (Fig. 3b).

We also used LFMM (Frichot et al., 2013; Frichot & François, 2015; Caye et al., 2019) to statistically correlate SNPs among 5,560 loci with environmental gradients, after controlling for population structure (Supplementary Figure S6). The importance of bioclimatic gradients was evaluated by computing a multiple squared correlation between each variable and the SNPs detected by LFMM for that variable. The most important bioclimatic variables for association with allele frequencies were BIO9: mean temperature of driest quarter (correlated with 95 loci, R-squared = 0.78, P-value = 1.40e-09), BIO17: precipitation of driest quarter (correlated with 117 loci, R-squared = 0.82, P-value = 1.26e-05), BIO15: precipitation seasonality (correlated with 53 loci, R-squared = 0.86, P-value = 1.98e-17), BIO19: precipitation of coldest quarter (correlated with 54 loci, R-squared = 0.66, P-value = 5.23e-08), and BIO2: mean diurnal range (correlated with 10 loci, R-squared = 0.42, P-value = 3.7e-06, Fig. 3c). The high congruence between RDA and LFMM indicated that drought-related variables were important in shaping genomic variation in the species.

Historical demography under the MSC model

Bayesian Phylogenetics & Phylogeography (BPP; Yang, 2015; Flouri et al. 2018) was run for three purposes: to provide additional evidence for divergence among the three primary lineages (analysis A11), to estimate a species tree (analysis A01), and to estimate divergence times and effective population sizes (analysis A00). A11 analysis (species tree estimation and species delimitation) of both our admixed and non-admixed data resulted in posterior probabilities of ~1.0 for each of the three populations (DST, STH, PLN). All species tree

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

analyses placed the STH and PLN as sister with a posterior probability of 1.0. Effective population size (N_e) estimates from the A00 analysis showed signs of both population growth and decline following divergence (Supplementary Table S3). In comparing the N_e estimates for runs containing KK104 and runs without, six of the seven parameters overlapped within the 95% HPDs. The results differed most in their estimates for our DST (pop 1) population (admixed = 348,125 vs. non-admixed = 233,593) as well as the most recent common ancestor (MRCA) of our ingroup (admixed = 772,968 vs. non-admixed = 625,781). To minimize potential biases in parameter estimation, the following N_e values were from runs with KK104 removed. Our ingroup MRCA showed an N_e of *ca*. 625k with the descendant populations having N_e values of *ca*. 233k for DST and *ca*. 930k for the combined STH+PLN population. After the split of the STH+PLN populations, there was a reduction in N_e to STH (*ca*. 575k) and PLN (*ca*. 157k). These results are consistent with peripheral population expansion following divergence.

In addition to potential bias in N_e estimates due to admixture or mixed ancestry, we found evidence for biases in divergence times (Fig. 4). Including sample KK104 resulted in an older divergence time at the root while the divergence time of the ingroup was younger. Again, to minimize any biases regarding interpretation, we focused on the results with this sample removed. Assuming a divergence time of 20 Ma for *P. cornutum* and *P. solare* (Leaché & Linkem, 2015) resulted in an estimated substitution rate of 0.000535 substitutions per site per million years, similar to the previously estimated mean genome-wide rate for lizards of 0.0008 by Perry et al. (2018). Thus, independent data supported a relatively slow rate of substitution, compared with faster rates found in other studies (Green et al. 2014; Tollis et al. 2018). Basing our calibration on a rate of 0.0008 substitutions per site per million years, divergence times for both nodes fell clearly in the Quaternary (Supplementary Fig. S7). Combining these results with the divergence estimates from the mtDNA in BEAST (initial divergence of 5 Ma), a late Pliocene-early Pleistocene divergence event appears to be a likely scenario for the initial split.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Demographic models

Our MOMENTS (Jouganous et al., 2017) analyses were used to test three hypotheses regarding historical divergence: allopatric divergence due to the Rio Grande, divergence due to paleoclimate change, and divergence due to ecological gradients. Each hypothesis makes assumptions regarding the importance of gene flow during evolutionary history (Leaché et al., 2019). For consistency with the BPP analyses, we analyzed the same set of individuals. The top ranked models were similar across the two data sets (with and without the admixed sample KK104), consisting of an initial split between DST and the ancestral population of STH and PLN, followed by a period of no gene flow before final diversification between STH and PLN populations with gene flow (Fig. 5; Supplementary Table S4). The data set including KK104 suggested that gene flow only occurs between the STH and PLN populations. In contrast, the data set that did not include the admixed individual suggests that there was gene flow between DST and STH and between STH and PLN populations. We were unable to perform likelihood ratio tests for the data set without the admixed individual due to our top two models being unnested. Likelihood ratio tests for the data set including the admixed individual failed to reject the nested model suggesting a barrier to gene flow when compared to the model favored by the other data set, therefore it was considered the best model for the admixed data set (D_{adj} = -2515.84; p-value = 1). AIC weights for the admixed data set strongly supported the refugia_barrier model (0.9980), whereas the non-admixed data set favored the refugia_adj_2 model (wAIC = 0.7328; Table 1; Fig. 6). However, the refugia_barrier model was within the 95% confidence interval for the no admixture data set.

Accommodating gene flow under the MSci model

Although the MSC model can accommodate coalescent stochasticity due to ILS, it explicitly assumes no gene flow once populations diverge. This assumption is likely violated in many systems, particularly in analyses of closely related species or populations. Thus, we

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

performed a series of analyses under the MSC-with-introgression (MSci) model in BPP (Flouri et al., 2020) to compare demographic parameter estimates from the MSC analyses. We again analyzed both the admixed (with sample KK104) and non-admixed (without sample KK104) data sets (500 loci in each case). In each data set, there were two local peaks in the posterior distribution, which corresponded to two sets of parameter values and may be considered two demographic hypotheses (Figs. 7 and S8; Table 2). The two peaks fit the data nearly equally well because the species tree is close to a trichotomy with two divergence times close to each other. For the admixed data, the Markov chain Monte Carlo (MCMC) run often visited only one peak. For the non-admixed data, the MCMC run jumped between the peaks, with introgression probabilities φ_A and φ_B showing bimodal distributions. Note that the introgression probability φ_A is the proportion of population *A* composed of migrants from population *TB* while $1 - \varphi_A$ is the contribution from population *SA* (Fig. 7). In other words, when we trace the genealogical history of sequences sampled from modern species/populations backwards in time and reach node *A*, each sequence will take the two parental paths *BT* and *AS* with probabilities φ_A and $1 - \varphi_A$, respectively. We separated the samples for the two peaks depending on whether $\varphi_A > ½$. Peak 1 (with $\varphi_A > ½$) consisted of ~86% of the MCMC samples. The subsamples corresponding to the same peak were noted to be similar between runs and those from different runs were combined to produce the posterior summary for that peak (Table 2).

We discuss the genetic history implied by Peak 1 for the non-admixed data, and then examine the similarities and differences of Peak 2 and of the results from the admixed data. When we trace the history of the samples backwards in time, Peak 1 implies the following (Fig. 7a). The DST sequences mostly (with probability φ_A = 86.8%) trace back to node *B* (or branch *TB*), before taking the path *TSR* to the root of the tree. Sequences from STH will reach node *C* and then mostly (with probability $1 - \varphi_C$ = 93.5%) trace back to node *B*. Sequences from PLN will reach node *D* and mostly (with probability φ_D = 93.5%) take the *DCB* route to reach *B*. Thus,

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

most sequences from populations STH and PLN will be in the same ancestral population *C* by the time $\tau_C = \tau_D \approx 0.00017$, while most sequences from DST will meet those from STH or PLN in ancestral population *B* by time $\tau_A = \tau_B \approx 0.00141$. Note that in BPP, both divergence (or introgression) times (τs) and population sizes (θs) are measured in units of expected number of mutations per site.

Peak 2 for the non-admixed data is a minor peak in the posterior (Fig. 7b). It implies that most sequences from populations STH and PLN will be in the same ancestral population *C* at time $\tau_C = \tau_D \approx 0.00017$, while most sequences from DST will meet those from STH or PLN in ancestral population *A* by time $\tau_A = \tau_B \approx 0.0014$. Beyond nodes *AB*, the divergence times and population sizes on the paths to the root are similar between Peaks 1 and 2.

The two peaks for the admixed data are even more similar to each other because the inferred species tree has nearly a trichotomy with $\tau_S \approx \tau_T$, with near perfect matching of the parameters between the peaks: $\varphi'_A \approx 1 - \varphi_A$, $\varphi'_B \approx 1 - \varphi_B$, $\theta'_A \approx \theta_B$, and $\theta'_B \approx \theta_A$ (Supplementary Fig. S8, Table 2). Most sequences from populations STH and PLN meet in population *C* at time $\tau_C = \tau_D \approx 0.00013$, while most sequences from DST meet those from STH or PLN in population *T* at time $\tau_T = 0.00178$ according to Peak 1 or in population *S* at time $\tau_S = 0.00210$ according to Peak 2. Beyond nodes *S* or *T*, the divergence times and population sizes on the paths to the root are almost identical between Peaks 1 and 2. Thus, if we consider the expected coalescence times between sequences from the three populations, or if we consider similarly sequence distances between populations, the two peaks for each data set made very similar predictions.

Finally, we compared parameter estimates from the MSci model with those of the MSC model (Fig. 4). The MSci model simultaneously accommodates deep coalescence and gene flow when estimating common evolutionary parameters. In general, ignoring gene flow when it is present leads to underestimation of divergence times and overestimation of population sizes.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

There was a relatively large effect of including/excluding sample KK104 on divergence times. Assuming a mutation rate of 0.0008, calibrated divergence times under the MSci model were 4.83 Ma for node *T* and 7.68 Ma for node *S*. Introgression times were 1.78 Ma for $\tau_A = \tau_B$ and 213 Ka for $\tau_C = \tau_D$ (see Fig. 7 for node labels). We provide calibrated estimates for the non-admixed Peak 1 data set only, as that is our best estimate of the evolutionary history of these populations.

Species distribution modeling

We estimated SDMs to further test the hypothesis that lineage divergence was caused by paleoclimate change (Fig. 8a,b). The SDMs estimated from the Last Glacial Maximum (LGM) revealed niche space in northern Mexico and along the border in southern Texas and New Mexico. The eastern and central (near Big Bend) portion of this area held the highest probabilities of occurrence. The northern edge of the LGM niche space coincided with our current STH population in the east and the DST population in the central region. The models also revealed a potential disjunct niche space, albeit with lower probabilities of occurrence, between the western edge of the Chihuahuan desert to the east and the Sonoran Desert to the west (outside of the current range of the species). The current SDM shifted the suitable niche northward expanding across the plains of Texas, up into Colorado, Oklahoma and Kansas, and connecting with the expanding range in southern Arizona and New Mexico. The eastern and larger area of the current SDM occupies Level 1 Ecoregion 9 The Great Plains, whereas the western and smaller portion occurs over Ecoregion 10 North American Deserts. The PCA analysis of the climatic niche space occupied by our genetic clusters showed the greatest dissimilarity between the areas occupied by our DST and STH populations with no overlap on the PC1 axis (Fig. 8c). The climate niche space occupied by our three PLN subpopulations

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

showed the greatest similarity and considerable overlap on the PC1 axis. All PCAs indicate that the three main lineages/populations inhabit a substantially different niche space (Fig. 8d)

Discussion

Genetic structure and demography

Speciation occurs when barriers to gene flow arise and separate populations. Barriers can come in the form of hard geographical divides such as mountains and rivers, or soft divides where the barriers to gene flow are environmental factors. Recent studies have shown these soft ecological divides may have a greater impact on diversification and speciation than the traditional hard allopatric geographical divides (Castro-Insua et al., 2018; Moen & Wiens, 2017; Myers et al., 2019). The evolutionary history of *P. cornutum* appears to further the evidence for the importance of both hard and soft allopatry in shaping species and highlight the diverse history of populations across a species range.

We found similar population structure to Williams et al. (2019) with high divergence between a desert (DST), southern (STH) and plains (PLN) clade at nuclear SNPs that correspond respectively to the western, southern, and northern, populations in the earlier study. By incorporating analysis of SNP data in addition to mitochondrial data we were able to expand upon this earlier study by estimating divergence times between these groupings and elucidating the current and historic environmental factors that have influenced population structure. Divergence time estimates from both the mitochondrial and nuclear data (under the MSC model) suggest that *P. cornutum* populations initially diverged during the late Pliocene or early Pleistocene in the range of 2.5 - 3 Ma, supporting our hypothesis of cryptic diversity within the species. We arrive at this time interval based on multiple analyses of the nuclear data while taking into account the divergence estimates from our mtDNA analysis (~5 Ma). Given the

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

likelihood of over estimating divergence times from mtDNA due to substitution saturation owing to a quicker mtDNA mutation rate (Zheng et al., 2011), we focus predominantly on the nuclear estimates. However, we do recognize the present challenges of adopting nuclear genome-wide substitution rates. Importantly, our divergence times correlate with the onset of full scale North American glaciations (Zachos et al., 2001), which resulted in cooler and more arid conditions throughout much of the American Tropics and may also have facilitated the Great American Biotic Interchange in mammals (Bacon et al., 2016). However, our SDMs suggest that our study area in particular experienced cooler and wetter conditions, at least during the LGM.

The two primary lineages (DST, STH+PLN) may have roughly coincided geographically within refugial habitats that originated during the Pleistocene, in the Chihuahuan Desert to the east, and the Sonoran Desert to the west (Figs. 1 & 7). This deep divide may be the result of niche conservatism (Wiens & Graham, 2005), where these populations tracked habitats amidst a changing climate resulting in subsequent isolation, consistent with a refugial speciation model (Moritz et al., 2000). The finding of suitable habitat throughout the Sonoran Desert during the Pleistocene is noteworthy, as the current range of *P. cornutum* does not extend this far west. These historical patterns also appear congruent to those of other reptile taxa inhabiting the region, which also support a model of divergence in allopatry during the Pleistocene followed by secondary contact and gene flow (Schield et al., 2015,2018,2019). An alternative hypothesis for the initial split is that the Plio-Pleistocene Lake Cabeza de Vaca in the northern Chihuahuan Desert served as a biogeographic barrier leading to vicariance (Rosenthal & Forstner, 2014). Unfortunately, the results of our demographic modeling make it difficult to disentangle vicariance due to paleoclimate versus the lake, as both hypotheses predict initial divergence in allopatry followed by secondary contact and gene flow. From our nuclear data we show evidence of a second split occurring more recently in the eastern population as it expanded its range northward in response to a shifting climate opening up greater niche space as glaciation receded. It is these fluctuating Pleistocene climatic cycles driving habitat contraction and

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

expansion that are likely to have initially shaped the current population structure and set the groundwork for further divergence.

As a population expands its range through a series of founder events, the signatures of this expansion should be evident in a reduction of population size and genetic diversity in the populations occupying the new territory (Excoffier et al., 2009). This decrease in heterozygosity at the forefront of the expansion has been illustrated in many studies of wide-ranging species (Garcia-Elfring et al., 2017; Jezkova et al., 2016; Peter & Slatkin, 2013). This same signature of expansion is readily visible across our analyses. Consistent with this signature of expansion at nuclear loci, there is higher mtDNA haplotype diversity in the STH (south) population than the PLN (north) population which also suggests the expansion occurred from the south into more northern areas (Williams et al. 2019). Although our PLN population occupies by far the largest geographical area, stretching from Texas to Kansas, it appears to have the smallest population size. Our BPP analyses indicate a reduction in N_e after the STH and PLN populations diverged, furthering the evidence for this northward expansion originating from the south. Interestingly, evidence from our population structure analysis indicates that members of this expanding PLN population do share ancestral genetic variation with our DST population. The existence of some highly admixed individuals (KK104, 7R10L) support our demographic results and point towards secondary contact and gene flow post divergence. Taken together, these results suggest that climatic cycling during the Pleistocene was the most likely catalyst for range expansion and secondary contact. An alternative hypothesis for admixture may be due to human mediated movement of *P. cornutum* owing to its popularity as a pet and symbol of the American Southwest. Other studies have shown evidence of translocations with admixed individuals appearing far removed from boundary areas (Williams et al., 2019). This human mediated movement may play a role in the mitochondrial introgression. It may also provide the reason the Rio Grande does not appear to be an insurmountable barrier to gene flow between the populations. However, we note that signals of introgression and admixture are restricted to the

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

periphery of the range of each lineage. For example, sample KK104 was collected in Brewster Co., Texas, which is substantially farther east than other individuals in the clade and in close geographic proximity to samples encompassing our PLN population. This sample is also nested in the PLN+STH mtDNA lineage and not the DST lineage, indicating introgression. Similar geographic patterns are also found with sample 7R10L from Dimmit/La Salle County, Texas. A previous study with denser sampling in western Texas, found that the DST (western) population extended from El Paso Co. to Brewster Co. (Williams et al. 2019), on the opposite side of the Rio Grande. Admixture between the western and eastern groups was concentrated in Jeff Davis and Brewster Counties, although as previously mentioned, there were some admixed individuals that were far removed from this potential boundary area (Williams et al. 2019). More comprehensive sampling throughout Texas, particularly near contact zones, is required to determine the precise locations of lineage boundaries.

The Riverine Barrier hypothesis would suggest that the Rio Grande could act as a vicariant barrier to gene flow, isolating the groups on either side and shaping the population structure (Lanna et al., 2020; Pellegrino et al., 2005). Geographically, the river does appear to divide the populations (Fig. 1) with only three individuals from our DST population appearing on the eastern side of the river. It is possible that the river continues to serve as a moderate barrier to dispersal, and future studies should focus on obtaining samples from Mexico to test this hypothesis further. The demographic models we tested in MOMENTS supported different models depending on whether sample KK104 was included in the analysis. Models without KK104 (non-admixed data set) favored secondary contact with gene flow between the populations (i.e. the refugia_adj_2 model), though the model with an explicit barrier between populations (with no gene flow to/from DST) was within the 95% CI of AIC weights. The best demographic model that included KK104 (admixed data set) was the refugia_river_barrier model (wAIC = 1.0), that predicted gene flow only between the STH and PLN populations. These results highlight the importance of sampling scheme (even a single highly admixed

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

individual) for demographic inference, and further studies are needed to explore this phenomenon more closely. The presence of a heavily admixed specimen from the DST population (KK104) from the eastern side of the river, along with DST ancestral genetic variation appearing in individuals throughout the range suggests that the river is not an absolute barrier. The importance of rivers as vicariant barriers to gene flow has come under recent scrutiny with studies showing they may not provide the impasse once thought, with one study finding them non-effective in 99% of Amazonian species studied (Lanna et al., 2020; Nazareno et al., 2017; Santorelli et al., 2018). Again, it seems best to not approach this question as an all or nothing proposition as the river's width was correlated with the strength as a barrier to gene flow (Nazareno et al., 2017). Thus, it is possible that the reduced gene flow between these populations is at least partly due to the Rio Grande. Additional sampling throughout Mexico will likely result in more power to test the efficacy of the Rio Grande as a barrier to gene flow. We also note that the Sacramento Mountains in southern New Mexico may serve as a contemporary barrier to gene flow.

Niche divergence resulting from ecological gradients across the species' range may play a significant role in driving continued divergence in *P. cornutum*. Among ecological gradients, precipitation is considered a major factor in furthering diversity and determining a species' range (Hawkins et al., 2003; Wiens et al., 2013). The family Phrynosomatidae has historically existed in arid environments, with those currently occupying more mesic habitats being recently derived (Wiens et al., 2013). This historic trend highlights a family-wide pattern of migration (=recent colonization) towards areas of greater precipitation. Across the range of *P. cornutum* there exists a significant precipitation gradient, ranging from under 10 inches (25.4 cm) of average annual rainfall in the western desert to over 50 inches (127 cm) in the eastern reaches of the Great Plains (Pittman et al., 2007). Variables concerning precipitation account for our top three results from LFMM analysis. Further, 21 of 29 SNPs identified through redundancy analysis were associated with seasonal precipitation. Thus, we cannot refute the hypothesis that the

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

varied levels of precipitation from across the range of *P. cornutum* are causing adaptive divergence in this system. Because of the species' low vagility and extensive range, adaptations that prove advantageous may become fixed in the population with greater speed, compounding the effects of niche divergence (Ujvari et al., 2008). Considering morphological adaptations to arid environments are visible in the form of the interscalar channels *P. cornutum* uses to harvest rainwater (Sherbrooke, 1990), it would be interesting to see if morphological variation along precipitation gradients exists among the three populations.

There are additional populations of *P. cornutum* that reside in the southeastern United States, having been introduced in the 1920s as a form of pest control. These populations already show significant morphological differences from their west coast counterparts (Heuring et al., 2019) despite the short term of geographical separation. While it is not clear if the differences are the result of genetic drift or adaptation to unique environments, it does highlight the rapidity with which significant morphological changes can arise between populations. With the deep divergence between our DST and STH+PLN populations occurring > 3 Ma, not only does it vastly increase the time frame for adaptation and further divergence to occur, it places it amongst other speciation events seen in the genus. According to a recent time calibrated phylogeny of *Phrynosoma* (Leaché & Linkem, 2015), several species pairs diverged more recently than 5 Ma, with the *P. platyrhinos*- *P. goodei* split occurring concurrently with our DST and STH+PLNS divergence at ~3 Ma. Currently, *P. cornutum* is the second oldest lineage of the genus at 20 Ma, younger than only *P. asio*. In addition, recent genomic data (ddRADseq) suggest that *P. cornutum*, along with *P. asio* and *P. solare*, exhibits substantial genetic divergence among populations (Leaché et al., 2021). If there indeed does exist a cryptic species or subspecies within *P. cornutum*, further examination of both morphological and behavioural differences between the populations is necessary.

Accommodating gene flow in genomic studies

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

The recently developed MSci model was designed to explicitly accommodate both ILS and gene flow/introgression when estimating divergence times and effective population sizes (Flouri et al., 2020). Given the presumed ubiquity of inter- and intraspecific gene flow in natural populations, the model marks a significant advancement of the field. However, the current implementation of BPP assumes that the introgression model is specified *a priori*, and furthermore the program may not deal with recent hybrids when inferring ancient admixture events. Here, we compare and contrast results from several BPP analyses under the MSC and MSci models, both with and without sample KK104 (admixed and non-admixed data sets). For many parameters the 95% HPDs overlapped, though some interesting patterns emerged. Of particular note was the much older divergence times inferred from the non-admixed data set under the MSci model versus the other three analyses (MSC-admixed, MSC-non-admixed, MSci-peak1-admixed). The former analyses estimated divergence times of 7.68 Ma and 4.83 Ma, whereas in the remaining analyses divergence times occurred during the Pleistocene. The admixed data also produced a much smaller introgression time $\tau_A = \tau_B$ than the non-admixed data (posterior means 0.000178 vs. 0.00142; Table 2). Other divergence times (such as $\tau_C = \tau_D$, which is assumed to be smaller than $\tau_A = \tau_B$, and τ_T and τ_S) were also affected. Similar to MOMENTS, these differences can be explained by the impact of including a recent hybrid sample (KK104). Note that sequences sampled from two modern populations cannot coalesce until they are in the same ancestral population. Let $t_{1\text{-}23}$ be the smallest sequence divergence between 1.DST and 2.STH (or 3.PLN), minimized across all loci and all sequence pairs at each locus. Then $t_{1\text{-}23} > \tau_A$. As KK104 appears to be a recent hybrid, the divergence time at some loci can be very small, and those small distances will force τ_A to be very small. Note that under the coalescent model, species divergence times and introgression times are determined mostly by the minimum, rather than the average, sequence divergence between species. The result suggests that hybrid samples should be avoided when one aims to infer ancient introgression

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

history. Similarly, we suggest that the results from the non-admixed data may represent a more realistic description of the history of divergences and introgressions for those lineages. We leave it to future studies to more thoroughly assess the impact of admixed samples on the estimation of divergence times under the MSC and MSci models.

As discussed above, the peaks in the posterior of Figs. 7 and S8 are difficult to distinguish using genomic sequence data. According to the theory developed by Yang & Flouri (2021), bidirectional introgression (BDI) events generate unidentifiability issues of two types: *within-model* and *between-model*, depending on whether the species involved in the introgression are sister or non-sister species. The *within-model* unidentifiability is essentially a label switching issue as the MCMC samples parameters within a single model. We note that the two peaks in Figs. 7 and S8 represent alternative *within-model* hypotheses that are nearly equally supported by the data. The peaks are identifiable, but very hard to distinguish with genetic data because the two speciation events occurred in quick succession (with $\tau_S \approx \tau_T$ in Supplementary Fig. S8). The sequence data also provide equal support for multiple *between-model* hypotheses: the four alternative between-model hypotheses corresponding to Peak 1 for the non-admixed data set are shown in Supplementary Fig. S9. These models are unidentifiable, as they make exactly the same probabilistic predictions for the gene trees and thus the same predictions for the multilocus sequence data. It is then impossible to use genomic data to distinguish such models. Researchers will need to consider additional information (e.g. habitat requirements) to help elucidate the most likely history of the species/populations. To our knowledge, this study serves as the first empirical investigation of unidentifiability issues with BDI models in BPP, and we encourage researchers interested in these models to carefully examine both classes of unidentifiability issues that may confound analysis and interpretation (Yang & Flouri 2021).

Both the MSci and isolation-with-migration (IM) models can be used with genomic data to account for gene flow when estimating divergence times and population sizes (Gronau et al.,

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

2011; Flouri et al., 2020). The MSci model assumes periodic introgression events between species, whereas the IM model accommodates continuous migration rates every generation. Selecting the appropriate model for a given data set is not straightforward, and additional studies are needed to quantify the effect of model misspecification. However, our results indicate that ignoring gene flow when it is present can potentially bias parameter estimates. More specifically, divergence times are underestimated and population sizes are overestimated when gene flow is not explicitly accounted for. Interestingly, we find no effect on species delimitation or species trees. This result is most likely due to the small number of populations studied. Our results are remarkably similar to previous simulation studies that also demonstrated similar biases in parameter estimates (Leaché et al., 2014). Thus, we further advocate careful consideration of models, assumptions, and sampling regimens when estimating demographic histories from genomic data.

Conclusions

We investigated the history of diversification within *P. cornutum* throughout the southwestern and central US by using genomic data to examine the hard and soft allopatric forces that have shaped population genetic structure. We find evidence for an initial divergence during the Plio-Pleistocene (possibly the Miocene) that was likely driven by habitat fragmentation due to climate fluctuations, vicariance due to the Rio Grande, and potentially Lake Cabeza de Vaca, followed by a subsequent northward range expansion as the receding glaciation opened up novel habitats. This expansion facilitated divergence along sharp environmental clines and possible adaptation to a divergent niche space. Whether the population-level diversity uncovered through this study rises to the level of species will require further investigation (for example, estimation of hybridization rates in contact zones for comparison with the long-term introgression rate), additional data (i.e. morphology), and dense population sampling, especially throughout Mexico. The evolutionary history presented here

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

highlights the importance of both hard and soft allopatric forces in shaping a species through gene flow, as the lineage divergences appear at least partially influenced by a changing habitat and environmental niche. Finally, this study should serve as a foundation for the exploration of powerful new models of demographic inference that make use of genomic data sets.

Materials and Methods

Sampling and data collection

Tissue samples (75) of *P. cornutum* were obtained from both museum specimens and field samples collected from multiple sites throughout Kansas, Oklahoma, Colorado, New Mexico, Texas and Arizona (Fig. 1; Supplementary Table S1). A single *P. solare* individual from Pima County, Arizona was also included as an outgroup taxon. All new collections were approved by the IACUC Committee at Miami University (protocol number 992_2021_Apr).

Genomic DNA was extracted from liver or muscle tissue using the Qiagen DNeasy Blood & Tissue Kit (Hilden, Germany) following manufacturer protocols. DNA quantity and quality were measured on a NanoDrop spectrophotometer. Aliquots of DNA extracts were shipped to LGC Genomics (Berlin, Germany) for library prep and sequencing using a modified genotyping-by-sequencing (Arvidsson et al., 2016; Elshire et al., 2011) approach. The technique, termed normalized GBS (nGBS) digests genomic DNA using the *Ms*II restriction enzyme and utilizes a subsequent normalization step after adapter ligation to remove fragments with a high number of copies. The method is particularly suited for species lacking a reference genome. Size-selected fragments were QC-ed and sequenced on an Illumina NextSeq flow cell (150 bp PE). Data were demultiplexed using Illumina bcl2fastq v. 2.17.1.14. Two samples (FHSM16593, FHSM16898) were excluded from further analysis due to a low number of reads. All nGBS data were uploaded to the SRA (accessions provided upon acceptance).

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

The raw, demultiplexed data were processed using IPYRAD v. 0.7.30 (Eaton & Overcast, 2020). The demultiplexed data were first quality filtered to remove residual adapter sequences (using *cutadapt*) and low-quality bases. Reads were then clustered within and between individuals based on 85% similarity, which is the default value recommended by the program authors. A minimum of 30 individuals per locus (~39% of samples) was required to keep loci in the final assembly, resulting in a concatenated matrix of ~8 million base pairs and 57,459 retained loci. Default values were also used for the remaining parameters. We also performed additional assemblies using a clustering threshold of 90%, and the results were qualitatively similar.

We obtained new mtDNA sequences from all samples to compare with the GBS data. Approximately 1,400 bp of mtDNA were collected from each sample, encompassing the entire ND1 gene, tRNA leucine, tRNA isoleucine, tRNA glutamine, and portions of 16S and tRNA methionine. PCR amplification was performed using previously published primers (Leaché & McGuire, 2006) and the *Taq* PCR kit (New England Biolabs, Ipswich, MA). Reactions (25 µl) consisted of the following: 2.5 µl 10X reaction buffer, 0.5 µl 10 mM dNTPs, 0.5 µl 10 µM forward primer (16dR), 0.5 µl 10 µM reverse primer (tMet), 0.125 µl *taq* DNA polymerase, 19.875 µl ddH_2O, 1 µl template DNA. All PCRs were performed on a BIO-RAD T100 Thermal Cycler using the following cycling conditions: initial denaturation at 95 ºC (30 sec), 30 cycles of denaturation at 95 ºC (30 sec), annealing at 55 ºC (1 min), and extension at 72 ºC (1 min), followed by a final extension at 72 ºC for 5 min and samples held indefinitely at 4 ºC. Horizontal agarose gel electrophoresis (1%) was used to assess the success of reactions. Amplicons were enzymatically purified using ExoSAP-IT (ThermoFisher Scientific, Waltham, MA) following manufacturer's recommendations. Purified products were sent to GENEWIZ (South Plainfield, NJ) for Sanger sequencing. Due to the large fragment size, amplicons were sequenced in both directions. Raw sequence data were edited in FinchTV v. 1.5.0 (Geospiza, Inc.). Aliview v. 1.26 (Larsson, 2014) was used to form contigs and perform multiple sequence alignment using

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Muscle (Edgar, 2004). All new mtDNA sequences were deposited to GenBank (accessions provided upon acceptance).

Phylogenetic analysis

All phylogenetic analyses were implemented through the High-Performance Computing Center (HPCC) at The College of Staten Island (CUNY). We performed both concatenated and coalescent analyses on the genomic data, as both approaches have their strengths and weaknesses (Kubatko & Degnan, 2007; Chou et al., 2015; Edwards et al., 2016) and recent empirical studies show that performing both can potentially result in novel insights (Blair et al., 2019). Concatenated maximum likelihood (ML) phylogenetic analysis (unpartitioned) was implemented using the hybrid MPI/Pthreads version of RAxML-ng v. 0.8.1 (Kozlov et al., 2019). A standard non-parametric bootstrap (250 reps) and ML search was implemented under a GTRGAMMA model of nucleotide substitution. Trees were rooted using *P. solare*. We also performed 20 independent ML searches from 10 distinct maximum parsimony and 10 random starting trees to determine if multiple likelihood peaks were present in the data. Robinson-Foulds (RF) distances were calculated between the 20 unrooted trees. These analyses were performed using the full multi-locus data versus individual SNPs.

We also performed Bayesian phylogenetic analyses in ExaBayes v. 1.5 (Aberer et al., 2014). ExaBayes is explicitly geared towards Bayesian analysis of large phylogenomic data sets generated through next-generation sequencing, utilizing MPI parallelization to increase computational efficiency. Default priors were used for all parameters. Analyses were run for 50 million generations, sampling every 5000 generations. Mixing and effective sample sizes (target ESS >200) for all parameters was monitored in Tracer v.1.7.1 (Rambaut et al., 2018). A majority rule consensus tree was generated following a burnin of 25%. The unrooted topology was subsequently rooted using *P. solare*. Similar to the ML analyses, all ExaBayes runs used the full loci including invariable sites.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Coalescent-based phylogenetic analysis was performed using SVDQUARTETS (Chifman & Kubatko, 2014) in PAUP* v. 4.0a159 (Swofford, 2001). SVDQUARTETS is statistically consistent with the multispecies coalescent and first infers quartet relationships using site pattern frequencies and singular-value decomposition scores. The algorithm then uses QFM (Reaz et al., 2014) to assemble quartets into a full tree containing all taxa. Although SVDQUARTETS can be used with multi-locus sequence data, the method is particularly suited to large SNP data sets and has been recently used in other RADseq/GBS studies (Eaton et al., 2016; Leaché et al. 2015). We used the *.u.snps.phy* file from IPYRAD for all SVDQUARTETS analyses to minimize linkage of SNPs. All quartets were evaluated (1,150,626) and 100 nonparametric bootstrap replicates were used to assess nodal support. Trees were rooted using *P. solare*.

We used BEAST v. 2.6.3 (Bouckaert et al., 2019) to infer genealogical relationships and divergence times based on the mtDNA sequences. bModelTest v. 1.2.1 (Bouckaert & Drummond, 2017) was specified as the substitution model for all analyses, which uses reversible-jump MCMC to switch between models. A constant size coalescent tree prior was used, a relaxed log normal clock (Drummond et al., 2006), and all remaining priors were left as defaults. We also ran a strict clock analysis for comparison. Analyses were temporally calibrated using a mitochondrial substitution rate previously calculated for a similarly sized lizard (Macey et al., 1999) and used in other studies of both *Phrynosoma* and other lizards (Bryson et al., 2012; Jezkova et al., 2016). However, to accommodate uncertainty in the rate, we specified a normal prior with a mean of 0.00805 substitutions per site per million years and a sigma of 0.0005. Chains were run for 40 million generations, sampling every 4,000 for a total of 10,000 states over independent runs. Mixing, ESS values (target >200) and parameter estimates were monitored in Tracer. TreeAnnotator was used to construct a maximum clade credibility (MCC) tree annotating nodes using mean heights following a burnin of 10%.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Population structure and GEA analysis

Population structure was analyzed using the non-negative matrix factorization algorithm sNMF implemented in LEA v2.6.0, for which the number of genetic clusters, *K*, was evaluated from the cross-entropy criterion (Frichot et al., 2014; Frichot & François, 2015). This criterion measures the amount of statistical information conveyed by a model with *K* clusters by comparing predictions of masked alleles to their true value, and detects the most significant subdivisions in the data. Like STRUCTURE (Pritchard et al., 2000), sNMF is a descriptive method, and visual inspection of the clustering results was used to investigate finer population structure for *K* = 2-10. Before performing GEA analysis, SNPs were filtered out for loci with less than 50% missing data. The missing genotypes were then imputed using values predicted by the sNMF model (*K* = 5). SNPs with minor allele frequency lower than 5%, and SNPs in strong linkage disequilibrium (r2 > 0.96) were removed from the data set.

We calculated pairwise F_{st} values (Weir & Cockerham, 1984) between the three main populations inferred from both the sNMF and phylogenetic analyses using the R package hierfstat (Goudet, 2005). We made the decision to treat these as three populations rather than five to focus on the both the deepest divergences from the phylogenetic analysis and the geographic structure of the populations (see Results). The analysis was run for 1000 bootstraps using 95% confidence intervals to assess significance. Nei's genetic distances (Nei, 1978) were calculated using the R package StAMPP v 1.5.1 (Pembleton et al., 2013) to determine mean pairwise distances between populations and diversity within each population.

Spatial genetic structure was examined at an individual level using maximum likelihood population effects parametrization (MLPE, Clarke et al., 2002). We compared geographic distance and genetic distance to test for evidence of isolation by distance (IBD) throughout the sampled distribution. This was implemented using the R packages nlme (Pinheiro et al., 2012) and corMLPE (https://github.com/nspope/corML PE), with the correlation between population

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

pairs as covariates, and Akaike weights calculated using the MuMIn package (Bartoń, 2019). The outgroup taxon was excluded prior to performing these analyses.

Genome-wide associations with climatic gradients were investigated using latent factor mixed models (LFMM), as implemented in the R package lfmm (Frichot et al., 2013; Frichot & François, 2015; Caye et al., 2019). The number of factors in LFMMs were determined from the population structure analysis (K = 5). Climate data were obtained from the WorldClim v2 database at the 2.5 minute resolution (Fick & Hijmans, 2017). All 19 WorldClim bioclimatic variables were tested for association with SNPs and a joint correlation analysis for all bioclimatic variables was performed. Significance values were obtained after Bonferroni correction for multiple testing. The importance of bioclimatic variables was evaluated by computing the coefficient of determination for each variable and the SNPs detected by LFMM for that variable. Statistical significance of determination coefficients was evaluated using Fisher tests. R code and associated data files to reproduce sNMF and LFMM analyses are available on Dryad (doi provided upon acceptance).

We also implemented redundancy analyses (RDA) to assess correlation between SNPs and environmental variables using the R package *vegan* (Oksanen et al., 2016). RDA is a constrained ordination method that is a multivariate analog of linear regression and examines the amount of variation in one set of variables that explains variation in another set. In our case, how much genomic variation is explained by environmental predictors. RDA is a powerful method that can be used to infer selection, with low false positive and high true positive rates (Forester et al., 2018). The approach performs a PCA on the response variables (SNP matrix) while constraining the PCA axes as linear combinations of the predictor (environmental) variables. In our analyses, environmental variables were represented by two bioclimatic variables from WorldClim v2 (Fick & Hijmans, 2017): mean temperature of the driest quarter and precipitation seasonality. These variables were selected to account for major aspects of climate while avoiding autocorrelation among variables (Dormann

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

et al., 2013). The significance of the entire model and each axis was evaluated using an analysis of variance (ANOVA) with 999 permutations. Effects of collinearity between environmental predictors were assessed using the function *vif.cca* to evaluate variance inflation factors. We then identified candidate SNPs based on locus score that were ± 2.5 *SD* from the mean loading on all four constrained axes. We identified the environmental variables with the strongest associations with each candidate SNP using a Pearson's correlation coefficient (*r)*.

Species tree and historical demographic analysis

We used Bayesian Phylogenetics & Phylogeography (BPP) v4.1.3 (Yang, 2015; Flouri et al. 2018) to perform a series of coalescent-based analyses on reduced subsets of data (individuals and loci). This is a Bayesian MCMC implementation of the multispecies coalescent model with and without introgression. The full likelihood approach applied to multilocus sequence alignments makes full use of information contained in both gene tree topologies and branch lengths. Unlike concatenation, the approach accommodates the coalescent fluctuation in genealogical history across the genome. Unlike two-step approaches, the likelihood calculation in the MCMC algorithm averages over gene trees and branch lengths at individual loci, accommodating their uncertainties (Rannala & Yang 2003; Yang & Rannala, 2014; Rannala & Yang 2017; Flouri et al., 2020). Because our genetic clustering analyses indicated the possibility of admixture between some populations (see Results), one data set excluded a highly admixed individual with <50% of the genome originating from a single ancestral population (sample KK104) that was in an otherwise genetically distinct population while another included the individual. All other individuals used in analyses could trace >50% of their genome to a single cluster. Our goal was to test how inclusion of this sample might influence the estimation of common evolutionary parameters (e.g. species trees, divergence times, population sizes). Samples were assigned to one of three populations in *P. cornutum* (rooted with *P. solare*)

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

following the results of the concatenated analyses (i.e. RAxML-ng, ExaBayes, SVDquartets) and clustering in sNMF. We chose to analyze three populations/lineages to represent the deepest divergences in the genealogy. We did not divide the Plains lineage into two populations due to the results of SVDquartets (see Results). However, all BPP analyses used individuals from only one of the two Plains lineages inferred by RAxML-ng and ExaBayes. For computational reasons, all analyses were run using 500 loci.

We first performed a series of A11 analyses to provide additional support that the populations defined by previous analyses might represent distinct populations or species (Yang & Rannala, 2010; Yang & Rannala, 2014). This analysis compares MSC models that differ in the number of species and in the species phylogeny. Each MSC model involves two sets of parameters: the species divergence times (τs) and the population sizes (θs). Both parameters are measured in the expected number of mutations per site. Four independent A11 analyses were run (two using algorithm 0 and two algorithm 1). The species model prior assumed uniform rooted trees, and the starting tree topology was based off the concatenated analyses. We specified an inverse gamma (IG) prior of IG(3,0.004) for population sizes (θ) and IG(3,0.05) for the divergence time at the root of the species tree (τ_0). Runs were implemented using an initial burnin of 50,000 generations followed by sampling every 5 generations for 100,000 total samples. Convergence was assessed by examining consistency between runs. We then performed a series of species tree analyses in BPP (A01) using the same populations. Similar to previous analyses, runs were performed both with and without the admixed individual KK104 to quantify the potential impact of gene flow on species tree estimation. All A01 analyses used the same priors and sampling frequency as the A11 analyses. We compared the best tree and associated support values among runs. Finally, we performed multiple A00 analyses to estimate divergence times and effective population sizes (N_e) on the species tree inferred from the A01 analyses. Again, analyses included or excluded sample KK104 to determine how gene flow

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

might influence divergence times and population sizes. The parameter settings and priors were identical to the other BPP analyses, except that we used an initial burnin of 200,000 followed by sampling every 20 generations for 100,000 total samples. Mixing, convergence, and ESS values (target > 200) were assessed using Tracer v1.6.0 (Rambaut et al., 2018).

There is still no general consensus of accurate nuclear genome-wide substitution rates for lizards. Estimates from the literature suggest that lizard rates, on average, are slightly faster than snakes (0.00077 vs. 0.00074 substitutions per site per million years, respectively (Perry et al., 2018)). The assumptions and uncertainty about substitution rate directly translates to uncertainties about absolute divergence times, which can influence hypothesis testing. Thus, we used several sources of information to convert raw parameter estimates of θ and τ to units of effective number of individuals and millions of years, respectively. First, we used previous results for the divergence time (T) of *P. cornutum* and *P. solare* (~20 Ma; Leaché & Linkem, 2015) to obtain an empirical mutation rate (μ) estimate directly from the data ($\mu = \tau/T$) . This calculation provided additional evidence either supporting or refuting previous rate hypotheses. We then compared our rate estimate to independently estimated genome-wide neutral substitution rate for lizards and squamates (Green et al., 2014; Perry et al., 2018; Tollis et al., 2018). Our analysis provided support for slower substitution rates, supporting the recent estimates of Perry et al. (2018). Thus, our final calibrations were based on a rate of 0.0008 substitutions per site per million years (8×10^{-10} substitutions per site per year). To obtain estimates of N_e we assumed a generation time of 2 years (Jezkova et al., 2016).

Demographic model testing

To examine and compare the different models of the divergence of *P. cornutum* (riverine barrier, paleoclimate change, environmental gradients) we used MOMENTS (Jouganous et al., 2017) to simulate the three-dimensional joint site frequency spectrum (JSFS) of genetic

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

variation between the three populations based on results from our phylogenetic and population structure analysis. However, MOMENTS is based on the approximation of the discrete Wright-Fisher Model, meaning that it is not appropriate to pool populations that may be genetically distinct (e.g. Plains cluster). Therefore, we used the same individuals in MOMENTS as in BPP analysis. For each data set (with and without KK104), we tested 10 3D models that were based on various aspects of divergence previously hypothesized for species in the region ranging from simple models with no gene flow to more complex models involving multiple time periods and varying degrees of gene flow between populations (Fig. 6). We examined the possibility of river barriers preventing gene flow between adjacent populations, divergence in isolation with subsequent secondary contact, and various combinations involving models with allopatric and subsequent parapatric divergence along ecological clines (Jezkova et al., 2016; Myers et al., 2019; Schield et al., 2015).

The program easySFS (https://github.com/isaacovercast/easySFS) was used to determine the dimensions that would maximize segregating sites shared between samples when creating the folded JSFS; we also retained one SNP per locus to minimize linkage disequilibrium. MOMENTS is an efficient method of simulating the evolution of an allele frequency spectrum over time using differential equations. The basis of MOMENTS is similar to the diffusion approximation approach utilized in the program ∂a∂i and many of the models we tested were adapted from previously developed ∂a∂i and MOMENTS models (Gutenkunst et al., 2009; Leaché et al., 2019; Portik et al., 2017). For all models, we performed consecutive rounds of optimization with multiple replicates using the best scoring parameter (highest log-likelihood) estimates to base searches in the subsequent round (Portik et al., 2017; Leaché et al., 2019). Default settings in moments_pipeline (https://github.com/dportik/moments_pipeline) were used (replicates = 10, 20, 30, 40; maxiter = 3, 5, 10, 15; fold = 3, 2, 2, 1), and we optimized parameters using optimize_log_fmin, a simplex (a.k.a. amoeba) method in terms of log parameters. Optimized parameter sets of each replicate were used to simulate the 3D-JSFS,

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

and the multinomial approach was used to estimate the log-likelihood of the 3D-JSFS given the model. We ranked models according to AIC (lowest to highest) and estimated the standard deviation for each parameter using the Godambe Information Matrix with bootstrapped spectra. Finally, we determined the best model by comparing the two top ranked models for each data set using a likelihood-ratio test if they were nested. It should be noted that while that practices that we employed are common (e.g., selecting one SNP per locus, projecting down the JSFS), they can influence demographic inference. Projecting down the JSFS can result in composite likelihoods which can cause statistics such as AIC and BIC to favor more complex models (Coffman et al., 2016; Gao & Song, 2010).

Gene flow and the multispecies coalescent with introgression model

Because several of our analyses suggested that gene flow was important throughout the evolutionary history of *P. cornutum*, we utilized the multispecies coalescent model with introgression (MSci) in BPP (Flouri et al., 2020) to estimate introgression probabilities and reassess how divergence times and population sizes are affected when gene flow is explicitly modeled. Parameters and prior settings were virtually identical to the previous BPP analyses with a few exceptions. First, we used the best model from MOMENTS to specify a phylogenetic network (i.e. species tree with introgression events) for BPP to estimate parameters (i.e. θ, τ, and φ). This model included multiple reticulations in the species tree. For the introgression probability parameter (φ), we specified a beta prior of (1,1). We ran four independent analyses using a burnin of 200,000, followed by 500,000 samples that were taken every two generations. All BPP MSci analyses were run under a strict clock model (default) using BPP v. 4.3.0. Convergence was assessed by examining the trace plots in Tracer and checking for consistency between runs. All MSci analyses used the same 500 loci as the BPP MSC analyses. We performed analyses both with and without the admixed/outlier sample KK104.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

When included, KK104 was assigned to the Desert (DST) population following the results from the phylogenetic analyses.

Species distribution modeling

We reconstructed the suitable climatic niche of *P. cornutum* for current climatic conditions and those of the Last Glacial Maximum (LGM) across the range of the species using ecological niche modeling. This methodology uses environmental data associated with occurrence records to estimate habitat suitability across the landscape by means of various program-specific algorithms (Elith et al., 2006). For occurrence data, we used our sampling localities, supplemented by occurrence records from the Vertnet (vertnet.org; queried 1st May 2018) and iNaturalist (iNaturalist.org; queried 5th September 2021) databases. All records with the coordinate uncertainty of 5 km and temperature outliers were removed, as well as all localities outside the known native range of the species and non-research grade records. This yielded 1096 occurrence records. We then filtered the occurrence records using the R package spThin (Aiello-Lammens, Boria, Radosavljevic, Vilela, & Anderson, 2015) to only include one occurrence record per 120 km. This filtering alleviated potential bias caused by unequal sampling effort (Merow, Smith, & Silander, 2013) and differential coordinate access restrictions between states. This yielded 169 occurrence records used to inform the models.

We derived the current climatic niche of the species using 19 bioclimatic variables with resolution of 30 seconds (~1km) from the WorldClim dataset (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005). We derived the LGM climatic niche for *P. cornutum* using two simulation models of the LGM climate: community climate system model (CCSM ver. 3; Otto-Bliesner et al., 2006) with a resolution of 1°, and the model for interdisciplinary research on climate (MIROC ver. 3.2; (Sugiyama, Shiogama, & Emori, 2010)) with an original spatial resolution of 1.4° X 0.5° (Braconnot et al., 2007). These original climatic variables have been downscaled to the spatial resolution of 2.5 minutes (under the assumption of high spatial autocorrelation) and converted to

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

bioclimatic variables (Hijmans et al., 2005; Peterson & Nyári, 2008). These two models both indicate colder and wetter climate during the LGM. However, the CCSM model predicts lower values across temperature variables whereas the MIROC model predicts higher values across precipitation variables (see Jezkova et al., 2016). We constructed climatic niche models for each climatic data set in the program MAXENT v. 3.3.3k (Phillips, Anderson, & Schapire, 2006) using the R packages ENMeval (Muscarella et al., 2014) and dismo (Hijmans, Phillips, Leathwick, & Elith, 2015). MAXENT estimates relative probabilities of the presence of species within defined geographic spaces, with high probabilities indicating suitable environmental conditions (Phillips et al., 2006; Phillips & Dudík, 2008). We used 1000 background points randomly extracted from a polygon drawn around the occurrence records and expanded by 2 degrees in all directions. This selection of background points was chosen to exclude distant areas with very different environmental conditions, following recommendations by Merow *et al.* (2013). We explored values for the regularization multiplier (rm) between 0.5 and 4 (by increments of 0.5) and all combinations of available features (i.e., linear, quadratic, product, threshold, and hinge). We ran 3-fold cross-validation replicates to choose a model with the best fit, as assessed by the lowest AICc value. The best-fitting model for each climatic data set was visualized using logistic probability values (Merow et al., 2013). PCA analyses were also performed for current climate niche space occupied by the 3 and 5 genetic clusters derived from sNMF population structure and phylogenetic analyses and utilizing the 19 bioclimatic variables.

Data Accessibility

All raw GBS data have been deposited in the Sequence Read Archive (BioProjectID = PRJNA780191).

-All mtDNA sequences have been deposited in GenBank (OL549193 - OL549266).

-The following items are available on figshare (https://doi.org/10.6084/m9.figshare.c.5715668.v2):

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

1. R scripts to reproduce all sNMF and LFMM analyses.
2. All output from ipyrad.
3. The concatenated mtDNA alignment.

Full data citation:

Finger, Nicolas; Farleigh, Keaka; Bracken, Jason T; D. Leaché, Adam; François, Olivier; Yang, Ziheng; et al. (2021): Data and scripts used to investigate the phylogeography and demographic history of the Texas horned lizard (Phrynosoma cornutum). figshare. Collection. https://doi.org/10.6084/m9.figshare.c.5715668.v2

Acknowledgments

We wish to thank the curatorial staff at multiple museums for providing tissue loans. We specifically thank the staff at the Sternberg Museum of Natural History Herpetology Collection (Curtis J. Schmidt), The Yale Peabody Museum of Natural History (Gregory J. Watkins-Colwell, Jacques Gauthier), The Sam Noble Oklahoma Museum of Natural History (Brandi S. Coyner, Janet K. Braun), The Museum of Vertebrate Zoology at Berkeley (Carol L. Spencer, Jimmy A. McGuire), and The California Academy of Sciences (Lauren A. Scheinberg, Rayna Bell). We also thank Steven Hein for help with collecting tissue samples. This material is based on work supported by the National Science Foundation Grant No. DEB-1929679 issued to C. Blair, a PSC-CUNY Award, jointly funded by The Professional Staff Congress and The City University of New York, issued to C. Blair, and through startup funds provided by Miami University issued to T. Jezkova. K. Farleigh was supported by the National Science Foundation Graduate Research Fellowship Program (Award # 2037786). Texas Parks and Wildlife Department's Horned Lizard License Plate, Andrews Institute of Mathematics & Science Education at TCU, and TCU

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Research and Creative Activities Fund provided funding to D. Williams. We thank the HPCC at the College of Staten Island (CUNY) for providing computational resources. Finally. We thank three anonymous reviewers for their comments on the manuscript.

Author Contributions

Designed the study (NF, TJ, CB), collected data (NF, TC, TJ, DW), analyzed the data (NF, KF, JB, AL, OF, ZY, TF, TJ, CB,), wrote the initial manuscript (NF, CB). All authors were involved in contributing to the manuscript revisions.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

References

Aberer AJ, Kobert K, Stamatakis A. 2014. ExaBayes: massively parallel Bayesian tree inference for the whole-genome era. Mol Biol Evol. 31:2553–2556.

Aiello-Lammens ME, Boria RA, Radosavljevic A, Vilela B, Anderson RP. 2015. spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography 38:541–545.

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet. 17:81–92.

Arvidsson S, Fartmann B, Winkler S, Zimmermann W. 2016. Efficient high-throughput SNP discovery and genotyping using normalised Genotyping-by-Sequencing (nGBS). https://biosearch-cdn.azureedge.net/assetsv6/efficient-high-throughput-snp-discovery-genotyping-ngbs-app-note.pdf (19.09.2019)

Bacon CD, Molnar P, Antonelli A, Crawford AJ, Montes C, Vallejo-Pareja MC. 2016. Quaternary glaciation and the Great American Biotic Interchange. Geology 44:375–378.

Bartoń K. 2019. MuMIn: multi-model inference. R package version 1.43.6.

Blair C, Bryson RW. 2017. Cryptic diversity and discordance in single-locus species delimitation methods within horned lizards (Phrynosomatidae: *Phrynosoma*). Mol Ecol Resour. 17:1168–1182.

Blair C, Bryson RW, Linkem CM, Lazcano D, Klicka J, McCormack J. 2019. Cryptic diversity in the Mexican highlands: thousands of UCE loci help illuminate phylogenetic relationships, species limits and divergence times in montane rattlesnakes (Viperidae: *Crotalus*). Mol Ecol Resour. 19:349–365.

Blair C, Ané C. 2020. Phylogenetic trees and networks can serve as powerful and complementary approaches for analysis of genomic data. Syst Biol. 69:593-601.

Bouckaert R, Drummond AJ. 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. BMC Evol Biol. 17:42.

Bouckaert R, et al. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comp Biol. 15:e1006650.

Braconnot P, et al. 2007. Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum–Part 1: experiments and large-scale features. Clim Past. 3:261–277.

Bryson RW, García-Vázquez UO, Riddle BR. 2012. Diversification in the Mexican horned lizard *Phrynosoma orbiculare* across a dynamic landscape. Mol Phylogenet Evol. 62:87–96.

Butlin RK, Galindo J, Grahame JW. 2008. Sympatric, parapatric or allopatric: The most important way to classify speciation? Philos Trans R Soc B. 363:2997–3007.

Castro-Insua A, Gómez-Rodríguez C, Wiens JJ, Baselga A. 2018. Climatic niche divergence

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

drives patterns of diversification and richness among mammal families. Sci Rep. 8:8781.

Caye K, Jumentier B, Lepeule J, François O. 2019. LFMM 2: Fast and accurate inference of gene-environment associations in genome-wide studies. Mol Biol Evol. 36:852–860.

Chifman J, & Kubatko L. 2014. Quartet inference from SNP data under the coalescent model. Bioinformatics 30:3317–3324.

Chou J, Gupta A, Yaduvanshi S, Davidson R, Nute M, Mirarab S, Warnow T. 2015. A comparative study of SVDquartets and other coalescent-based species tree estimation methods. BMC Genomics 16:S2.

Clarke RT, et al. 2002. Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. J Agric Biol Environ Stat. 7:361–372.

Coffman AJ, et al. 2016. Computationally efficient composite likelihood statistics for demographic inference. Mol Biol Evol. 33:591–593.

Coyne JA, Orr HA. 2004. Speciation. Sinauer Associates, Sunderland, MA.

de Oca ANM, Arenas-Moreno D, Beltrán-Sánchez E, Leaché AD. 2014. A new species of horned lizard (Genus *Phrynosoma*) from Guerrero, México, with an updated multilocus phylogeny. Herpetologica 70:241–257.

Dormann CF, et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:027-046.

Dray S, Dufour AB. 2007. The ade4 package: implementing the duality diagram for ecologists. J Stat Softw. 22:1-20.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol 4e88.

Eaton DAR, Overcast I. 2020. ipyrad: Interactive assembly and analysis of RADseq datasets. Bioinformatics 36:2592-2594.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Edwards SV, Xi Z, Janke A, Faircloth BC, McCormack JE, Glenn TC, Zhong B, Wu S, Moriarty Lemmon E, Lemmon AR, Leaché AD, Liu L, Davis, CC. 2016. Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. Mol Phylogenet Evol. 94:447–462.

Elith J, et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

Elshire RJ, et al. 2011. A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. PLoS One 6:e19379.

Excoffier L, Foll M, Petit RJ. 2009. Genetic consequences of range expansions. Annu Rev Ecol

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Evol Syst. 40:481–501.

Fair WS, Henke SE. 1999). Movements, home ranges, and survival of Texas horned lizards (*Phrynosoma cornutum*). J Herpetol. 33:517–525.

Fick SE, Hijmans RJ. 2017. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas: New Climate Surfaces For Global Land Areas. Int J Climatol. 37:4302–4315.

[dataset] Finger N, et al. 2021. Supporting data for Texas horned lizard (*Phrynosoma cornutum*) phylogeography. To be submitted after manuscript acceptance.

Fitzpatrick BM, Fordyce JA, Gavrilets S. 2009. Pattern, process and geographic modes of speciation. J Evol Biol. 22:2342–2347.

Flouri T, Jiao X, Rannala B, Yang Z. 2018. Species tree inference with BPP using genomic sequences and the multispecies coalescent. Mol Biol Evol. 35:2585–2593.

Flouri T, Jiao X, Rannala B, Yang Z. 2020. A Bayesian implementation of the multispecies coalescent model with introgression for phylogenomic analysis. Mol Biol Evol. 37:1211–1223.

Forester BR, Lasky JR, Wagner HH, Urban DL. 2018. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. Mol Ecol. 27:2215–2233.

Frichot E, François O. 2015. LEA: An R package for landscape and ecological association studies. Methods Ecol Evol. 6:925–929.

Frichot E, Mathieu F, Trouillon T, Bouchard G, François O. 2014. Fast and efficient estimation of individual ancestry coefficients. Genetics 196:973–983.

Frichot E, Schoville SD, Bouchard G, François O. 2013. Testing for associations between loci and environmental gradients using latent factor mixed models. Mol Biol Evol. 30:1687–1699.

Gao X, Song PXK. 2010. Composite likelihood Bayesian information criteria for model selection in high-dimensional data. J Am Stat Assoc. 105:1531–1540.

Garcia-Elfring A, et al. 2017. Admixture on the northern front: Population genomics of range expansion in the white-footed mouse (*Peromyscus leucopus*) and secondary contact with the deer mouse (*Peromyscus maniculatus*). Heredity 119:447–458.

Green RE, et al. (2014). Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. Science 346:1254449.

Gottscho AD, et al. 2017. Lineage diversification of fringe-toed lizards (Phrynosomatidae: *Uma notata* complex) in the Colorado Desert: Delimiting species in the presence of gene flow. Mol Phylogenet Evol. 106:103–117.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Goudet J. 2005. Hierfstat, a package for R to compute and test hierarchical F-statistics. Mol Ecol Not. 5:184–186.

Gronau I, Hubisz MJ, Gulko B, Danko CG, Siepel A. 2011. Bayesian inference of ancient human demography from individual genome sequences. Nat Genet. 43:1031–1034.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genet. 5:e1000695.

Hamann A, Wang T, Spittlehouse DL, Murdock TQ. 2013. A comprehensive, high-resolution database of historical and projected climate surfaces for western North America. Bull Am Meteorol Soc. 94:1307–1309.

Harrington SM, Hollingsworth BD, Higham TE, Reeder TW. 2018. Pleistocene climatic fluctuations drive isolation and secondary contact in the red diamond rattlesnake (*Crotalus ruber*) in Baja California. J Biogeogr 45:64–75.

Hawkins BA, Porter EE, Felizola Diniz-Filho JA. 2003. Productivity and history as predictors of the latitudinal diversity gradient of terrestrial birds. Ecology 84:1608–1623.

Heuring C, et al. (2019). Genetics, morphology and diet of introduced populations of the ant-eating Texas Horned Lizard (*Phrynosoma cornutum*). Sci Rep. 9:11470.

Hewitt G. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linn Soc 58:247–276.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907–913.

Hey J, et al. 2018. Phylogeny estimation by integration over isolation with migration models. Mol Biol Evol. 35:2805–2818.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 25:1965–1978.

Hijmans RJ, Phillips SJ, Leathwick JR, Elith J. 2015. Species Distribution Modeling. Package "dismo." Dismo: Species Distribution Modeling. R Package Version 1.0-12. http://CRAN.R-Project.org/package=dismo, Version 1. Retrieved from http://cran.r-project.org/web/packages/dismo/index.html

Huey R, Kingsolver J. 1993. Evolution of resistance to high temperature in ectotherms. Am Nat. 142:S21-S46.

Jezkova T, et al. 2016. Range and niche shifts in response to past climate change in the desert horned lizard (*Phrynosoma platyrhinos*). Ecography, 39:437–448.

Jouganous J, Long W, Ragsdale AP, Gravel S. 2017. Inferring the joint demographic history of multiple populations: beyond the diffusion approximation. Genetics 206:1549–1567.

Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

35:4453–4455.

Kubatko LS, Degnan JH. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst Biol. 56:17–24.

Lanna FM, et al. 2020. Dwarf geckos and giant rivers: The role of the São Francisco River in the evolution of *Lygodactylus klugei* (Squamata: Gekkonidae) in the semi-arid Caatinga of north-eastern Brazil. Biol J Linn Soc. 129:88–98.

Larsson A. 2014. AliView: A fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30:3276–3278.

Leaché AD, Harris RB, Rannala B, Yang Z. 2014. The influence of gene flow on species tree estimation: a simulation study. Syst Biol. 63:17–30.

Leache AD, et al. (2009). Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). Proc Natl Acad Sci USA 106:12418–12423.

Leaché AD, Linkem CW. 2015. Phylogenomics of horned lizards (Genus: *Phrynosoma*) using targeted sequence capture data. Copeia 103:586–594.

Leaché AD, McGuire JA. 2006. Phylogenetic relationships of horned lizards (*Phrynosoma*) based on nuclear and mitochondrial data: Evidence for a misleading mitochondrial gene tree. Mol Phylogenet Evol. 39:628–644.

Leaché AD, et al. (2019). Exploring rain forest diversification using demographic model testing in the African foam-nest treefrog *Chiromantis rufescens*. J Biogeogr 46: 2706-2721.

Leaché AD, Banbury BL, Felsenstein J, Nieto-Montes de Oca A, Stamatakis, A. 2015. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. Syst Biol 64:1032–1047.

Leaché AD, Davis HR, Singhal S, Fujita MK, Lahti ME, Zamudio KR. 2021. Phylogenomic assessment of biodiversity using a reference-based taxonomy: an example with Horned Lizards (*Phrynosoma*). Front Ecol Evol. 437.

Lenormand T. 2002. Gene flow and the limits to natural selection. Trends Ecol Evol. 17:183–189.

Long C, Kubatko L. 2018. The effect of gene flow on coalescent-based species-tree inference. Syst Biol. 67:770–785.

Macey JR, Wang Y, Ananjeva NB, Larson A, Papenfuss TJ. 1999. Vicariant patterns of fragmentation among gekkonid lizards of the genus *Teratoscincus* produced by the Indian Collision: a molecular phylogenetic perspective and an area cladogram for Central Asia. Mol Phylogenet Evol. 12:320–332.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Mol Phylogenet Evol. 66:526–538.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Merow C, Smith MJ, Silander JA. 2013. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. Ecography 36:1058–1069.

Moen DS, Wiens JJ. 2017. Microhabitat and climatic niche change explain patterns of diversification among frog families. Am Nat. 190:29–44.

Montanucci RR. 2015. A taxonomic revision of the *Phrynosoma douglasii* species complex (Squamata: Phrynosomatidae). Zootaxa 4015:1-177

Moritz C, Patton JL, Schneider CJ, Smith TB. 2000. Diversification of rainforest faunas: an integrated molecular approach. Annu Rev Ecol Systemat 31:533–563.

Mueller NF, Ogilvie H, Zhang C, Drummond A, Stadler T. 2018. Inference of species histories in the presence of gene flow. bioRxiv:348391.

Mulcahy DG, Spaulding AW, Mendelson JR, Brodie ED. 2006. Phylogeography of the flat-tailed horned lizard (*Phrynosoma mcallii*) and systematics of the *P. mcallii-platyrhinos* mtDNA complex. Mol Ecol 15:1807–1826.

Muscarella R, et al. (2014). ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods Ecol Evol. 5:1198–1205.

Myers EA, et al. (2019). Environmental heterogeneity and not vicariant biogeographic barriers generate community-wide population structure in desert-adapted snakes. Mol Ecol. 28:4535–4548.

Nazareno AG, Dick CW, Lohmann LG. 2017. Wide but not impermeable: Testing the riverine barrier hypothesis for an Amazonian plant species. Mol Ecol. 26:3636–3648.

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590.

Newman CE, Austin CC. 2015. Thriving in the cold: Glacial expansion and post-glacial contraction of a temperate terrestrial salamander (*Plethodon serratus*). PLoS One 10:e0130131.

Nosil P, Feder JL. 2012. Genomic divergence during speciation: Causes and consequences. Philos Trans R Soc B. 367:332–342.

Oksanen O, et al. 2016. Vegan: Community Ecology Package. R Package Version 2.3-5. http://CRAN.R-project.org/package=vegan

Otto-Bliesner BL, et al. 2006. Last glacial maximum and Holocene climate in CCSM3. J Clim. 19:2526–2544.

Nazareno AG, Dick CW, Lohmann LG. 2019. A biogeographic barrier test reveals a strong genetic structure for a canopy-emergent Amazon tree species. Sci Rep. 9:18602.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Pastorini J, Thalmann U, Martin RD. 2003. A molecular approach to comparative phylogeography of extant Malagasy lemurs. Proc Natl Acad Sci USA. 100:5879-5884.

Pellegrino KCM, et al. 2005. Phylogeography and species limits in the *Gymnodactylus darwinii* complex (Gekkonidae, Squamata): Genetic structure coincides with river systems in the Brazilian Atlantic Forest (Gekkonidae, Squamata). Biol J Linn Soc. 85:13–26.

Pembleton LW, Cogan NOI, Forster JW. 2013. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Mol Ecol Res. 13:946-952.

Perry BW, et al. 2018. Molecular adaptations for sensing and securing prey and insight into amniote genome diversity from the garter snake genome. Genome Biol Evol. 10:2110–2129.

Peter BM, Slatkin M. 2013. Detecting range expansions from genetic data. Evolution 67:3274–3289.

Peterson AT, Nyári ÁS. 2008. Ecological niche conservatism and Pleistocene refugia in the thrush-like Mourner, *Schiffornis* sp., in the neotropics. Evolution 62:173–183.

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. Ecol Model. 190:231–259.

Phillips SJ, Dudík M. 2008. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31:161–175.

Pinheiro J, et al. 2012. nlme: linear and nonlinear mixed effects models. R package version 3.

Pittman EG, et al. 2007. Water for Texas 2007. Texas Water Development Board, Vol. II, Document No. GP-8-1.

Portik DM, et al. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. Mol Ecol. 26:5245–5263.

Prates I, Penna A, Rodrigues MT, Carnaval AC. 2018. Local adaptation in mainland anole lizards: Integrating population history and genome-environment associations. Ecol Evol. 8:11932–11944.

Price AH. 1990. *Phrynosoma cornutum* (Harlan): Texas Horned Lizard. Catalogue of American Amphibians and Reptiles 469:1-7

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Pyron R, Burbrink FT. 2010. Hard and soft allopatry: Physically and ecologically mediated modes of geographic speciation: Modes of allopatric speciation. J Biogeogr 37:2005-2015.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol. 67:901–904.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

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.

Rannala B, Yang Z. 2017. Efficient Bayesian species tree inference under the multispecies coalescent. Syst. Biol. 66:823–842.

Rosenthal J, Forstner MRJ. 2014. Effects of a Plio-Pleistocene barrier on Chihuahuan Desert herpetofauna. In: Proceedings of the Sixth Symposium on the Natural Resources of the Chihuahuan Desert Region, 269–282.

Santorelli S, Magnusson WE, Deus CP. 2018. Most species are not limited by an Amazonian river postulated to be a border between endemism areas. Sci Rep. 8:2294.

Schield DR, et al. 2018. Cryptic genetic diversity, population structure, and gene flow in the Mojave rattlesnake (*Crotalus scutulatus*). Mol Phylogenet Evol. 127:669–681.

Schield DR, et al. 2019. Allopatric divergence and secondary contact with gene flow – a recurring theme in rattlesnake speciation. Biol J Linn Soc. 128:149–169.

Schield DR, et al. (2015). Incipient speciation with biased gene flow between two lineages of the Western Diamondback Rattlesnake (*Crotalus atrox*). Mol Phylogenet Evol. 83:213–223.

Sherbrooke WC. 1990. Rain-harvesting in the lizard, *Phrynosoma cornutum*: Behavior and integumental morphology. J Herpetol 24:302–308.

Sherbrooke WC. 2003. Introduction to Horned Lizards of North America. University of California Press, Berkeley.

Solís-Lemus C, Ané C. 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. PLoS Genet. 12:e1005896.

Solís-Lemus C, Yang M, Ané C. 2016. Inconsistency of species tree methods under gene flow. Syst Biol. 65:843–851.

Sugiyama M, Shiogama H, Emori S. 2010. Precipitation extreme changes exceeding moisture content increases in MIROC and IPCC climate models. Proc Natl Acad Sci USA. 107:571–575.

Sukumaran J, Knowles LL. 2017. Multispecies coalescent delimits structure, not species. Proc Natl Acad Sci USA. 114:1607–1612.

Tollis M, et al. 2018. Comparative genomics reveals accelerated evolution in conserved pathways during the diversification of anole lizards. Genome Biol Evol. 10:489–506.

Ujvari B, Dowton M, Madsen T. 2008. Population genetic structure, gene flow and sex-biased dispersal in frillneck lizards (*Chlamydosaurus kingii*). Mol Ecol. 17: 3557-3564.

Wang T, Hamann A, Spittlehouse DL, Murdock TQ. 2012. Climate WNA - High-resolution spatial climate data for Western North America. J Appl Meteorol Climatol. 51:16–24.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Weir BS, Cockerham CC. 1984. Estimating F-Statistics for the analysis of population structure. Evolution 38:1358–1370.

Wiens JJ, Graham CH. 2005. Niche conservatism: Integrating evolution, ecology, and conservation biology. Annu Rev Ecol Evol Syst. 36:519–539.

Wiens JJ, Kozak KH, Silva N. 2013. Diversity and niche evolution along aridity gradients In North American lizards (Phrynosomatidae). Evolution 67:1715–1728.

Williams DA, Rains ND, Hale AM. 2019. Population genetic structure of Texas horned lizards: Implications for reintroduction and captive breeding. PeerJ 7:e7746.

Wogan GOU, Richmond JQ. 2015. Niche divergence builds the case for ecological speciation in skinks of the *Plestiodon skiltonianus* species complex. Ecol Evol. 5:4683–4695.

Yang Z. 2015. The BPP program for species tree estimation and species delimitation. Curr Zool. 61:854–865.

Yang Z, Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. USA 107:9264–9269.

Yang Z, Rannala B. 2014. Unguided species delimitation using DNA sequence data from multiple loci. Mol Biol Evol. 31:3125–3135.

Yang Z, Flouri T. 2021. Estimation of cross-species introgression rates using genomic data despite model unidentifiability. bioRxiv DOI: 10.1101/2021.08.14.456331.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292:686–693.

Zheng Y, Peng R, Kuro-o M, Zeng X. 2011. Exploring patterns and extent of bias in estimating divergence time from mitochondrial DNA sequence data in a particular lineage: A case study of salamanders (Order Caudata). Mol Biol Evol. 28:2521–2535.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Tables

Table 1. AIC, ΔAIC, relative likelihood, and weighted AIC (wAIC) calculations for each demographic model considered (see Fig. 6) for each data set (upper panel: non-admixed; lower panel: admixed) in the program MOMENTS. non-admixed = without KK104; admixed = with KK104.

non-admixed data set

Model	AIC	ΔAIC	relative L	wAIC
refugia_adj_2	514.28	0.00	1.00	0.73
refugia_asymmig_adjacent	517.90	3.62	0.16	0.12
refugia_barrier	518.48	4.20	0.12	0.09
split_nomig	519.38	5.10	0.08	0.06
refugia_adj_1	529.54	15.26	0.00	0.00
refugia_adj_3	553.04	38.76	0.00	0.00
split_asymmig_adjacent	559.26	44.98	0.00	0.00
split_sym_mig_all	600.70	86.42	0.00	0.00
split_symmig_adjacent	624.66	110.38	0.00	0.00
refugia_symmig_all	629.98	115.70	0.00	0.00

admixed data set

Model	AIC	ΔAIC	relative L	wAIC
refugia_barrier	738.38	0.00	1.00	1.00
refugia_adj_2	750.78	12.40	0.00	0.00
refugia_adj_1	778.82	40.44	0.00	0.00
refugia_adj_3	779.32	40.94	0.00	0.00
split_asymmig_adjacent	799.04	60.66	0.00	0.00
split_nomig	840.36	101.98	0.00	0.00
split_sym_mig_all	870.88	132.50	0.00	0.00
refugia_asymmig_adjacent	884.50	146.12	0.00	0.00
split_symmig_adjacent	992.34	253.96	0.00	0.00
refugia_symmig_all	1226.80	488.42	0.00	0.00

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Table 2. Posterior means and 95% HPD CIs (in parentheses) of parameters in the introgression (MSci) model of Fig. 7 obtained from BPP analyses of data that either include or exclude the admixed sample KK104. Estimates of θ and τ are ×1000.

	non-admixed data without KK104		admixed data with KK104	
Parameter	Peak 1 (φ_A > ½)	Peak 2 (φ_A < ½)	Peak 1	Peak 2
θ_{OG}	1.22 (0.49, 2.08)	1.22 (0.51, 2.08)	1.22 (0.50, 2.08)	1.23 (0.49, 2.09)
θ_{DST}	1.34 (1.10, 1.58)	1.34 (1.09, 1.59)	1.56 (0.89, 2.28)	1.52 (0.88, 2.25)
θ_{STH}	3.52 (2.02, 5.10)	3.54 (1.87, 5.25)	2.53 (1.22, 3.96)	2.50 (1.29, 3.93)
θ_{PLN}	0.93 (0.56, 1.33)	0.92 (0.51, 1.43)	0.79 (0.44, 1.19)	0.73 (0.41, 1.14)
θ_R	23.1 (12.3, 34.2)	24.1 (15.6, 33.5)	22.5 (14.2, 31.7)	22.6 (14.2, 31.9)
θ_S	2.69 (0.50, 5.37)	2.72 (0.75, 4.45)	3.82 (3.07, 4.55)	3.72 (3.01, 4.46)
θ_T	1.89 (0.40, 3.88)	3.51 (0.37, 10.2)	3.72 (0.46, 9.08)	3.74 (0.37, 11.3)
θ_A	**2.55 (0.34, 7.10)**	**2.90 (0.95, 4.33)**	**2.15 (1.46, 2.88)**	**3.80 (2.51, 5.15)**
θ_B	**2.95 (1.06, 4.63)**	**2.34 (0.38, 6.15)**	**3.49 (1.31, 5.12)**	**2.17 (1.51, 2.93)**
θ_C	4.27 (2.92, 5.65)	4.55 (3.11, 5.96)	2.10 (0.42, 4.98)	2.48 (0.45, 5.88)
θ_D	6.96 (0.34, 20.2)	6.37 (0.33, 19.5)	11.56 (1.00, 26.0)	11.27 (1.01, 25.4)
τ_R	9.75 (5.90, 15.1)	8.53 (5.65, 12.2)	10.11 (7.23, 13.5)	10.00 (7.21, 14.1)
τ_S	6.14 (2.72, 10.8)	4.17 (2.52, 5.86)	2.05 (1.70, 2.39)	2.10 (1.78, 2.43)
τ_T	3.86 (2.31, 5.60)	3.42 (1.48, 5.09)	1.78 (0.76, 2.34)	2.01 (1.62, 2.42)
$\tau_A = \tau_B$	1.42 (1.12, 1.70)	1.41 (1.10, 1.72)	0.18 (0.09, 0.26)	0.17 (0.10, 0.29)
$\tau_C = \tau_D$	0.17 (0.10, 0.24)	0.17 (0.08, 0.26)	0.13 (0.07, 0.20)	0.13 (0.07, 0.20)
φ_A	**0.868 (0.668, 0.998)**	**0.222 (0.034, 0.469)**	**0.129 (0.078, 0.180)**	**0.873 (0.817, 0.923)**
φ_B	**0.090 (0.004, 0.204)**	**0.871 (0.421, 1.000)**	**0.019 (0.000, 0.043)**	**0.985 (0.968, 1.000)**
φ_C	0.065 (0.016, 0.126)	0.055 (0.013, 0.110)	0.250 (0.079, 0.605)	0.165 (0.069, 0.263)
φ_D	0.935 (0.869, 0.991)	0.938 (0.877, 0.990)	0.817 (0.606, 0.981)	0.855 (0.752, 0.953)

Note.— There are two local peaks in the posterior under the model for both the non-admixed and admixed data, which differ mainly in four parameters, with $\varphi'_A \approx 1 - \varphi_A$, $\varphi'_B \approx 1 - \varphi_B$, $\theta'_A \approx \theta_B$, and $\theta'_B \approx \theta_A$ (highlighted in bold). MCMC samples around each peak are summarized separately. The introgression probability for any bidirectional introgression event is defined for the horizontal branch: for example, φ_A is for branch BA while the vertical branch SA has $1 - \varphi_A$ (Fig. 7). Divergence and introgression times (τ) are the ages of nodes on the tree. *Population sizes (θ) correspond to branches on the tree, identified by the daughter node of the branch* (e.g. θ_S is for branch *RS* and θ_A is for branch *SA*). Both τ and θ are measured in the expected number of mutations per site. *OG = outgroup; DST = Desert; STH = Southern; PLN = Plains.*

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Figure Legends

Fig. 1. Sample locations for all *Phrynosoma cornutum* used in this study within the EPA level I ecoregions. Population assignments are based on genotypes from the nGBS dataset using the program sNMF.

Fig. 2. (a) Maximum likelihood (ML) genealogy inferred using RAxML-ng on a concatenated nGBS matrix of 7,906,017 bp. Values at nodes (on top) represent ML bootstrap proportions/Bayesian posterior probabilities from ExaBayes (* = 1.0). Values at nodes (below) represent bootstrap support (100 replicates) from SVDquartets analyses on a matrix of 54,634 SNPs. The branch leading to the outgroup was pruned for clarity. (b) Population structure inferred using sNMF. (c) The cross entropy criterion supported 5 ancestral populations (K = 5).

Fig. 3. Results from the GEA analyses. Plots from the redundancy analyses for the first two constrained ordination axes. (a) Relationship between individuals from the sNMF population assignments (color-coded circles) and the tested environmental variables (arrows). (b) Outlier loci (color-coded to environmental variable) and directionality of the relationship between the climate variables (arrows). (c) Importance of environmental variables in LFMM analysis as indicated by p-values for multiple R-squared (F-tests, *** = p < 1e-04). bio1 = Annual Mean Temperature; bio2 = Mean Diurnal Range; bio3 = Isothermality; bio4 = Temperature Seasonality; bio5 = Max Temperature of Warmest Month; bio6 = Min Temperature of Coldest Month; bio7 = Temperature Annual Range; bio8 = Mean Temperature of Wettest Quarter; bio9 = Mean Temperature of Driest Quarter; bio10 = Mean Temperature of Warmest Quarter; bio11 = Mean Temperature of Coldest Quarter; bio12 = Annual Precipitation; bio13 = Precipitation of Wettest Month; bio14 = Precipitation of Driest Month; bio15 = Precipitation Seasonality; bio16 =

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Precipitation of Wettest Quarter; bio17 = Precipitation of Driest Quarter; bio18 = Precipitation of Warmest Quarter; bio19 = Precipitation of Coldest Quarter.

Fig. 4. Comparison of parameter estimates from multispecies coalescent (MSC) analysis in BPP (analysis A00) with (brown) and without (blue) the highly admixed/outlier individual from Pop1 (KK104). Purple bars depict parameter estimates based on the multispecies coalescent with introgression (MSci) model from the data including KK104, whereas orange bars represent MSci estimates without KK104. Error bars represent 95% HPDs. Pop1 = Desert (DST), Pop2 = Southern (STH), Pop3 = Plains (PLN). OG = outgroup (*Phrynosoma solare*).

Fig. 5. (a,b) The demographic model selected from the program MOMENTS for the *Phrynosoma cornutum* populations using the three-dimensional site frequency spectrum (3D-SFS) for the Admix (a) and NoAdmix (b) datasets. The reference population (N_{ref}) was calculated from estimates of theta produced during demographic modeling (theta = $4N_{ref}\mu$; see Supplementary Table S4) where μ is the substitution rate which was set to 0.0008 substitutions per site per million year. (c,d) The fits between the 3D-SFS model and the data with the resulting residuals (positive residuals indicate that the model predicted too many SNPs in that entry).

Fig. 6. Demographic models explored using the program MOMENTS. Analyses were performed with and without sample KK104 that had substantial mixed ancestry. The data set with KK104 favored the “refugia_barrier” model (blue), whereas the data set without KK104 supported the “refugia_adj_2” model (red).

Fig. 7. Two local peaks in the posterior for parameters in the MSci model in the BPP analysis of the data without the admixed sample KK104. The two peaks represent two hypotheses that

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

have nearly equal support from the data, due to the species tree being nearly a trichotomy. Posterior means of node ages (τs) are used to draw branches, and the node bars represent the 95% highest probability density (HPD) credibility intervals (CIs). Numbers next to branches are posterior means of population sizes (θs) (see Table 2); not all population sizes are shown. The model assumes two bidirectional introgression events (*A*↔*B* and *C*↔*D*), and the thickness of the horizontal branches indicates the estimated introgression probability (φ). According to the first peak (a), the lineage A-DST is comprised of φ_A = 86.8% of migrants from lineage *TB* and 1 – φ_A = 13.2% from lineage *SA*. In contrast, the second peak (b) suggests that the lineage *A*-DST is 22.2% from lineage *STB* and 77.8% from lineage *SA*. Estimates of φs at the other three nodes (*B*, *C*, and *D*; see Table 2) are interpreted in the same way. The phylogenetic network in the center represents the model specified in BPP.

Fig. 8. Climatic niche model for *Phrynosoma cornutum* built using the Wordclim bioclimatic variables with resolution of 2.5 minutes for the current climatic conditions (a) and projected on the MIROC and CCSM (b) of the Last Glacial Maximum climate (mean of models shown). The models were visualized using logistic probability values. Warmer colors indicate a higher probability for species presence. The outer blue line shows the known range of *P. cornutum.* Dots represent the spatially filtered occurrence records used to create models. Climatic niche space occupied by each of the 5 genetic clusters (color coded circles) identified in the sNMF analysis (c) and similar results for the primary three clusters/lineages used for demographic modeling (d). The first two principal components derived from 19 bioclimatic variables (arrows) of the WorldClim data set are shown.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

FIG. 1

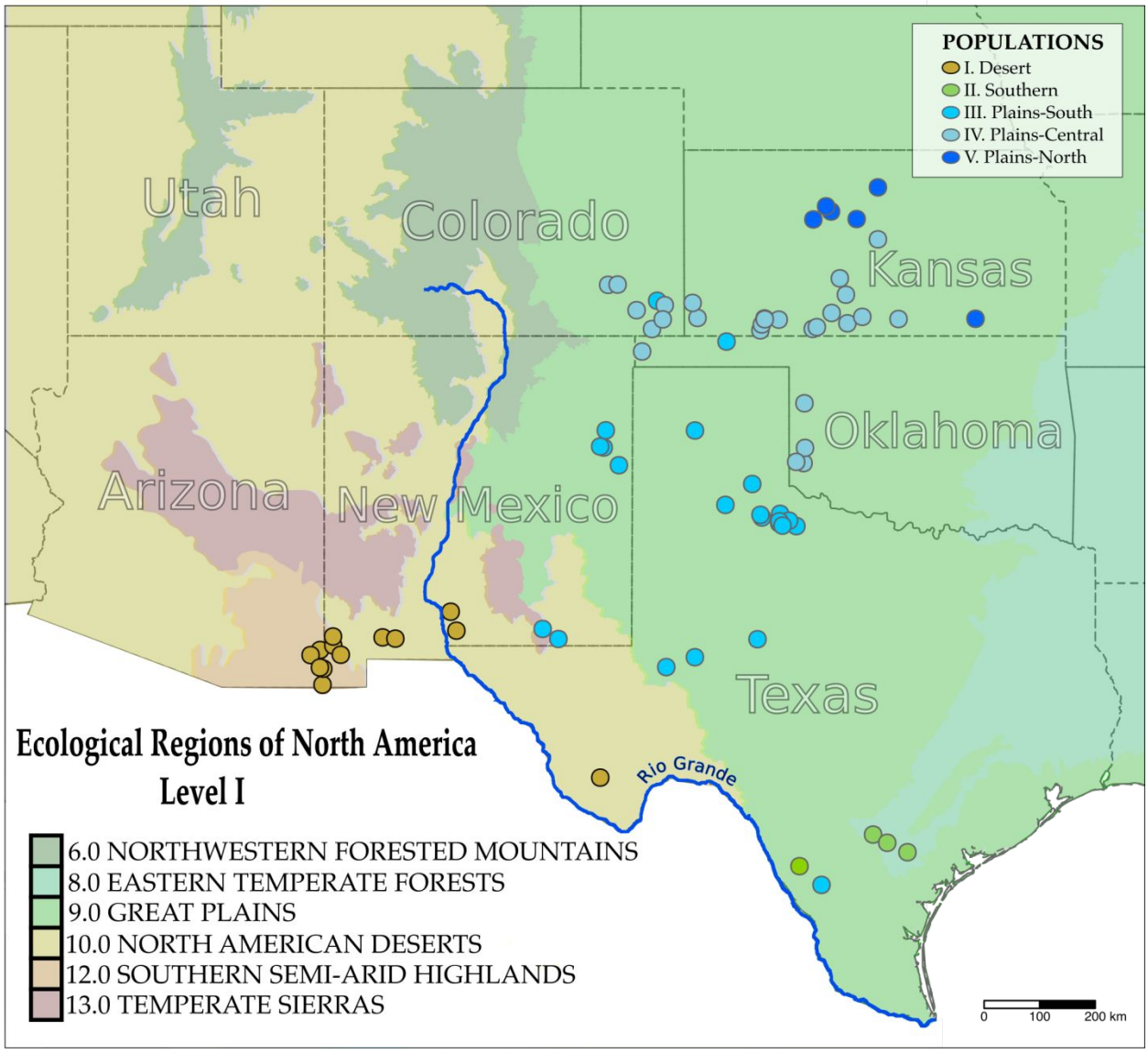

Fig. 1. Sample locations for all *Phrynosoma cornutum* used in this study within the EPA level I ecoregions. Population assignments are based on genotypes from the nGBS dataset using the program sNMF.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

FIG. 2

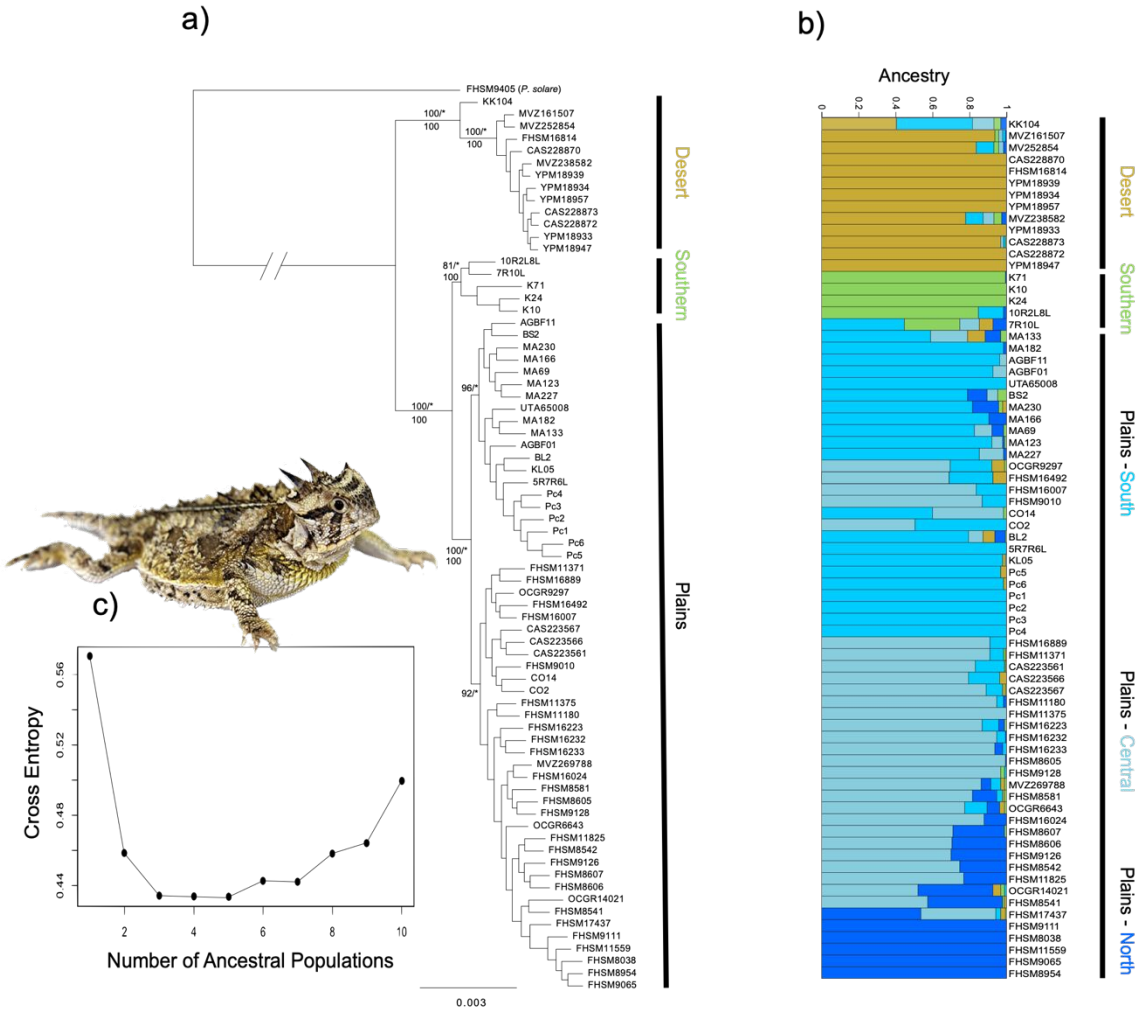

Fig. 2. (a) Maximum likelihood (ML) genealogy inferred using RAxML-ng on a concatenated nGBS matrix of 7,906,017 bp. Values at nodes (on top) represent ML bootstrap proportions/Bayesian posterior probabilities from ExaBayes (* = 1.0). Values at nodes (below) represent bootstrap support (100 replicates) from SVDquartets analyses on a matrix of 54,634 SNPs. The branch leading to the outgroup was pruned for clarity. (b) Population structure inferred using sNMF. (c) The cross entropy criterion supported 5 ancestral populations (K = 5).

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

FIG. 3

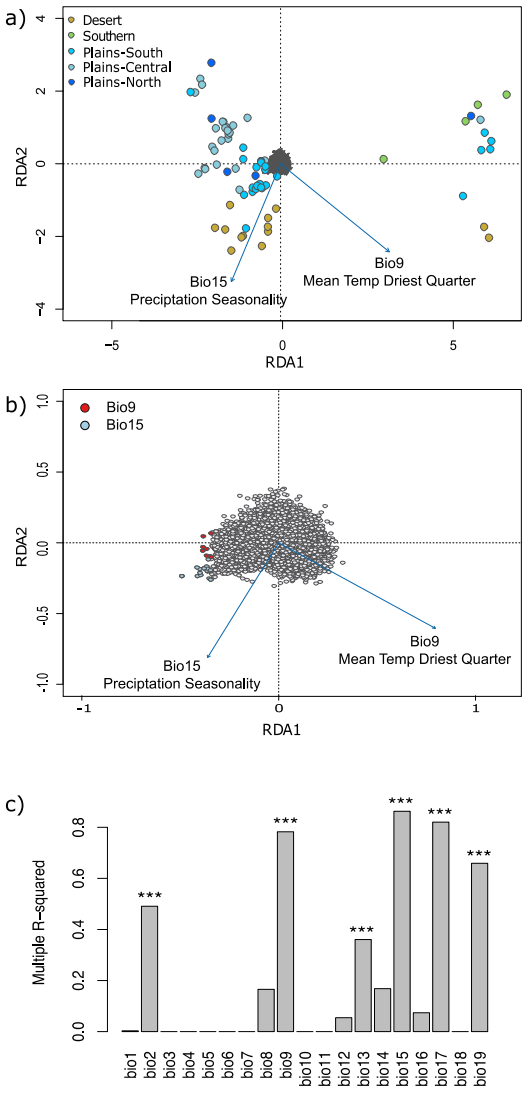

Fig. 3. Results from the GEA analyses. Plots from the redundancy analyses for the first two constrained ordination axes. (a) Relationship between individuals from the sNMF population assignments (color-coded circles) and the tested environmental variables (arrows). (b) Outlier loci (color-coded to environmental variable) and directionality of the relationship between the climate variables (arrows). (c) Importance of environmental variables in LFMM analysis as indicated by p-values for multiple R-squared (F-tests, *** = p < 1e-04). bio1 = Annual Mean Temperature; bio2 = Mean Diurnal Range; bio3 = Isothermality; bio4 = Temperature Seasonality; bio5 = Max Temperature of Warmest Month; bio6 = Min Temperature of Coldest Month; bio7 = Temperature Annual Range; bio8 = Mean Temperature of Wettest Quarter; bio9 = Mean Temperature of Driest Quarter; bio10 = Mean Temperature of Warmest Quarter; bio11 = Mean Temperature of Coldest Quarter; bio12 = Annual Precipitation; bio13 = Precipitation of Wettest Month; bio14 = Precipitation of Driest Month; bio15 = Precipitation Seasonality; bio16 = Precipitation of Wettest Quarter; bio17 = Precipitation of Driest Quarter; bio18 = Precipitation of Warmest Quarter; bio19 = Precipitation of Coldest Quarter.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

FIG. 4

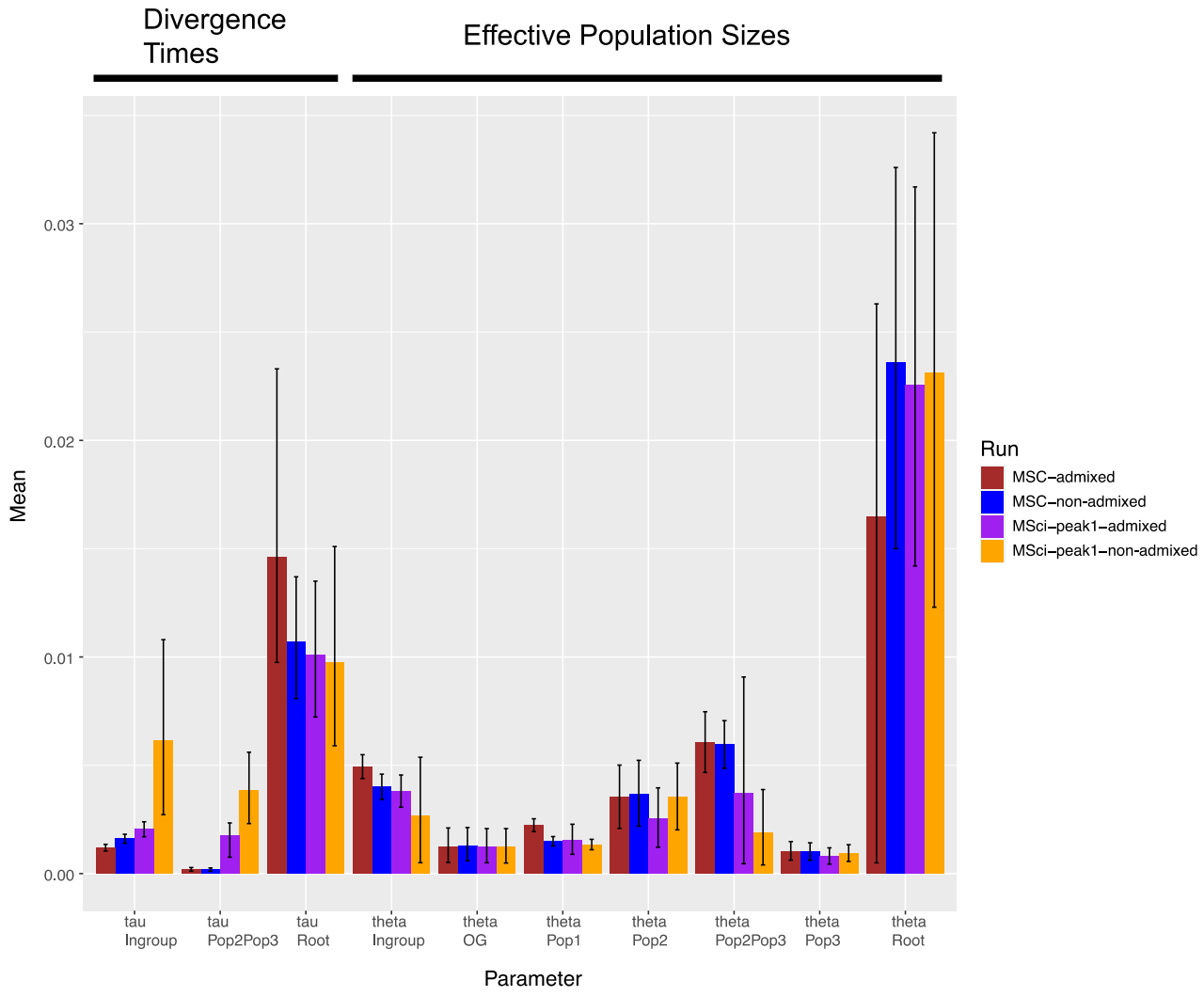

Fig. 4. Comparison of parameter estimates from multispecies coalescent (MSC) analysis in BPP (analysis A00) with (brown) and without (blue) the highly admixed/outlier individual from Pop1 (KK104). Purple bars depict parameter estimates based on the multispecies coalescent with introgression (MSci) model from the data including KK104, whereas orange bars represent MSci estimates without KK104. Error bars represent 95% HPDs. Pop1 = Desert (DST), Pop2 = Southern (STH), Pop3 = Plains (PLN). OG = outgroup (*Phrynosoma solare*).

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

FIG. 5

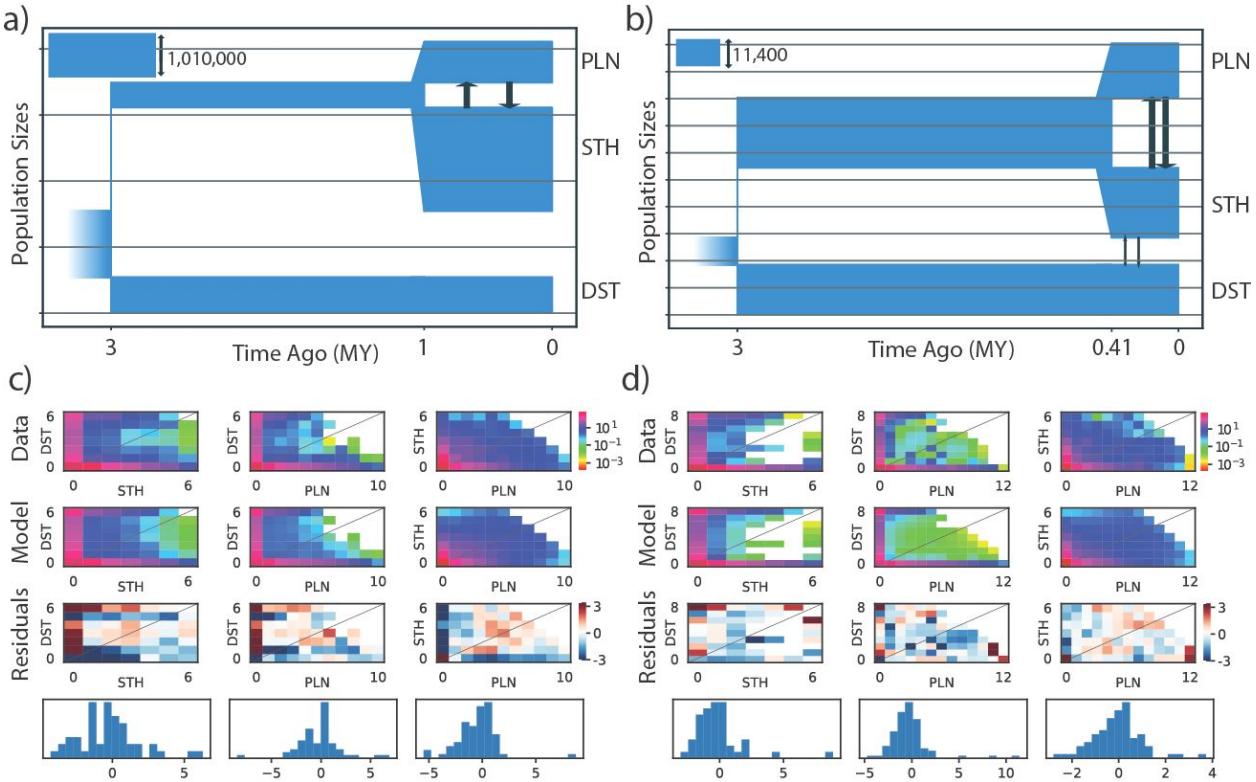

Fig. 5. (a,b) The demographic model selected from the program MOMENTS for the *Phrynosoma cornutum* populations using the three-dimensional site frequency spectrum (3D-SFS) for the admixed (a) and non-admixed (b) datasets. The reference population (N_{ref}) was calculated from estimates of theta produced during demographic modeling (theta = $4N_{ref}\mu$; see Supplementary Table S4) where μ is the substitution rate which was set to 0.0008 substitutions per site per million year. (c,d) The fits between the 3D-SFS model and the data with the resulting residuals (positive residuals indicate that the model predicted too many SNPs in that entry).

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

FIG. 6

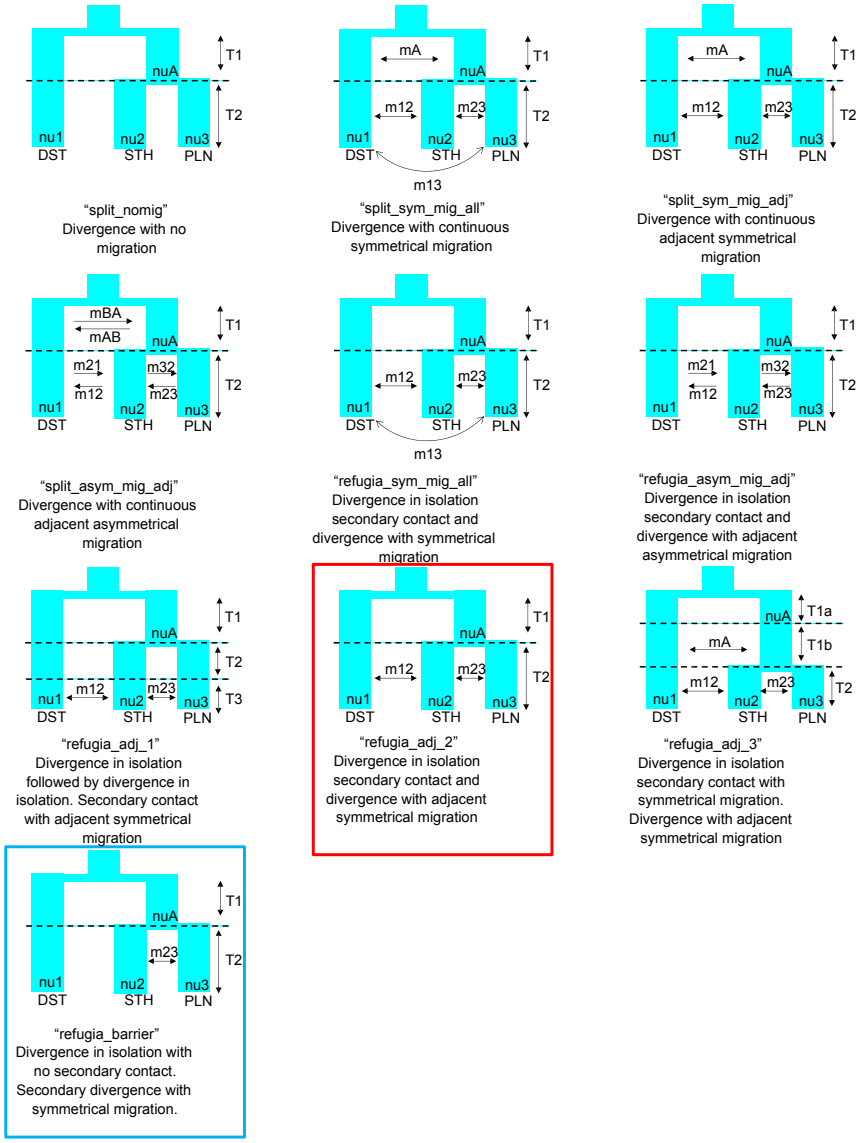

Fig. 6. Demographic models explored using the program MOMENTS. Analyses were performed with and without sample KK104 that had substantial mixed ancestry. The data set with KK104 favored the "refugia_barrier" model (blue), whereas the data set without KK104 supported the "refugia_adj_2" model (red).

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

FIG. 7

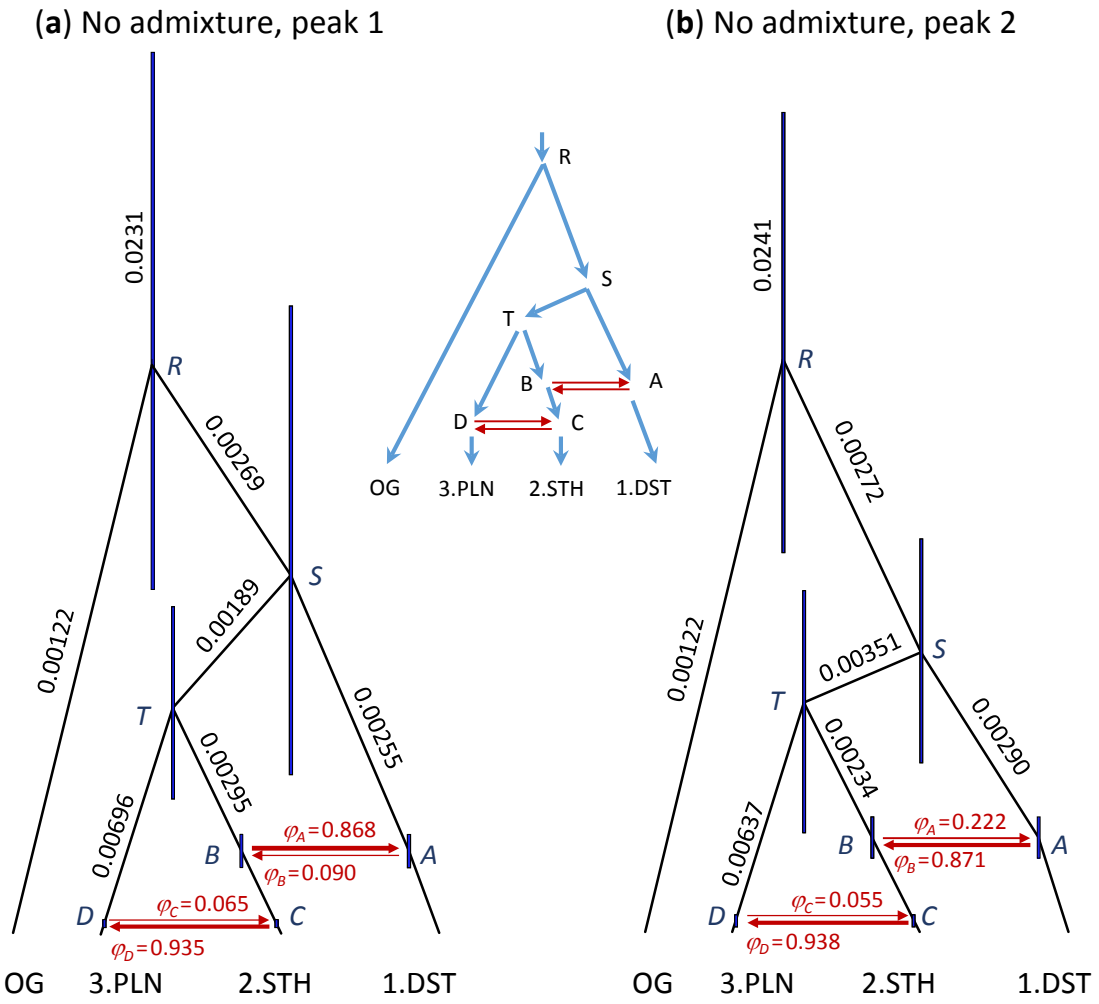

Fig. 7. Two local peaks in the posterior for parameters in the MSci model in the BPP analysis of the data without the admixed sample KK104. The two peaks represent two hypotheses that have nearly equal support from the data, due to the species tree being nearly a trichotomy. Posterior means of node ages (τs) are used to draw branches, and the node bars represent the 95% highest probability density (HPD) credibility intervals (CIs). Numbers next to branches are posterior means of population sizes (θs) (see Table 2); not all population sizes are shown. The model assumes two bidirectional introgression events ($A \leftrightarrow B$ and $C \leftrightarrow D$), and the thickness of the horizontal branches indicates the estimated introgression probability (φ). According to the first peak (a), the lineage A-DST is comprised of φ_A = 86.8% of migrants from lineage *TB* and 1 – φ_A = 13.2% from lineage *SA*. In contrast, the second peak (b) suggests that the lineage *A*-DST is 22.2% from lineage *STB* and 77.8% from lineage *SA*. Estimates of φs at the other three nodes (*B*, *C*, and *D*; see Table 2) are interpreted in the same way. The phylogenetic network in the center represents the model specified in BPP.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021

Fig. 8

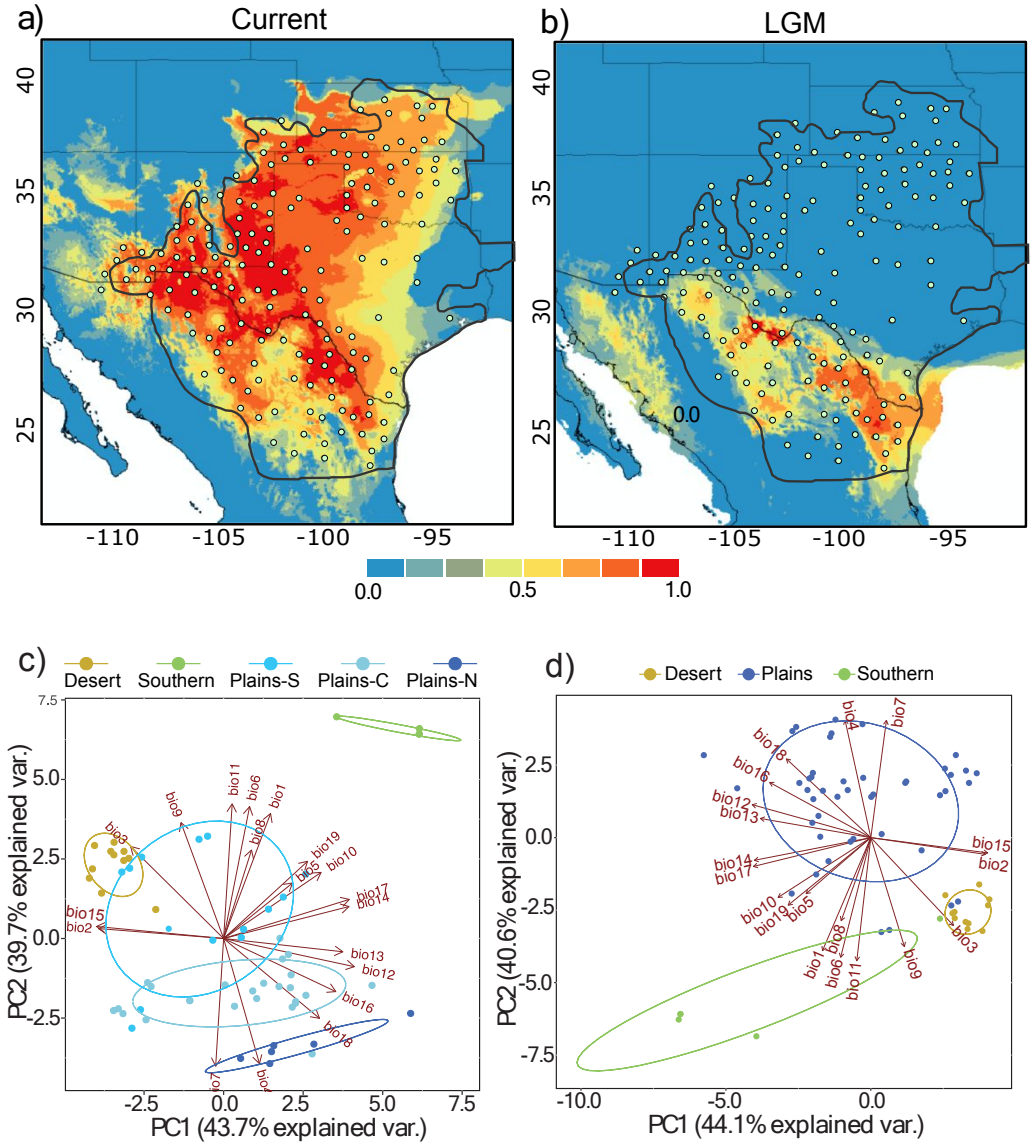

Fig. 8. Climatic niche model for *Phrynosoma cornutum* built using the Wordclim bioclimatic variables with resolution of 2.5 minutes for the current climatic conditions (a) and projected on the MIROC and CCSM (b) of the Last Glacial Maximum climate (mean of models shown). The models were visualized using logistic probability values. Warmer colors indicate a higher probability for species presence. The outer blue line shows the known range of *P. cornutum.* Dots represent the spatially filtered occurrence records used to create models. Climatic niche space occupied by each of the 5 genetic clusters (color coded circles) identified in the sNMF analysis (c) and similar results for the primary three clusters/lineages used for demographic modeling (d). The first two principal components derived from 19 bioclimatic variables (arrows) of the WorldClim data set are shown.

Downloaded from https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evab260/6443127 by University College London user on 03 December 2021