Biological Journal of the Linnean Society, 2018, **XX**, 1–15. With 4 figures.

Leapfrogging the Mexican highlands: influence of biogeographical and ecological factors on the diversification of highland species

ITZUE W. CAVIEDES-SOLIS* and ADAM D. LEACHÉ

Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Box 351800, Seattle, WA 98195, USA

Received 18 October 2017; revised 3 January 2018; accepted for publication 4 January 2018

In order to understand the processes that generate and maintain diversity, it is important to disentangle the roles of ecology and geography in speciation. We investigated the biogeographical and ecological factors that influenced the diversification of tree frogs (genus *Sarcohyla*) in the Mexican highlands, a region with high levels of endemism. Using single nucleotide polymorphism data for 58 samples, we found support for seven distinct genetic clusters within the *Sarcohyla bistincta* species complex, corresponding to *Sarcohyla calthula*, *Sarcohyla pentheter* and five populations within *S. bistincta*. A species tree analysis using the multispecies coalescent model did not support the monophyly of the five *S. bistincta* populations. We used niche modelling to calculate the ecological overlap among lineages; we found a degree of overlapping for most of the lineages, suggesting that ecological differentiation did not play a key role in their genetic divergence. Speciation and population structure in the complex have been shaped primarily by geological events, landscape modifications and climate changes during the Pleistocene. We discuss the relevance of genetic diversity for inferring the degree of species threats and recovery for conservation assessments.

ADDITIONAL KEYWORDS: divergence – Hylidae – niche modelling – phylogeography – population structure – *Sarcohyla*.

INTRODUCTION

The Neotropical highlands are among the most diverse and threatened ecosystems in the world (Myers *et al.*, 2000). The high levels of species richness and endemism in the Neotropics have typically been attributed to speciation via geographical isolation (Mayr, 1963; Barraclough & Vogler, 2000). Allopatric speciation is mainly related to geographical barriers that limit dispersal, e.g. orogenesis. After a vicariant event occurs, the resulting altitudinal gradient promotes parapatric speciation from the lowlands to the highlands (or vice versa; Weir, 2006). In addition, species isolation is reinforced by ecological divergence, related to ecological barriers that restrict gene flow, such as environmental conditions, microhabitat selection and species interactions (Schluter, 2000). Ecological and geographical factors influencing speciation are not independent or mutually exclusive; they are interconnected and affect species distinctively depending on the organism's intrinsic characteristics (Rundle & Nosil, 2005). For example, studies of alpine sedges and grasshoppers found that colonization and distributional shifts, caused by ecological factors, have a larger influence on speciation than geographical isolation (Massatti & Knowles, 2016; Knowles & Massatti, 2017). In order to understand the processes that have generated and maintain diversity in the Neotropical highlands, it is important to disentangle the roles of ecology and geography in speciation.

The factors affecting speciation in the highlands are relatively well understood when considered independently (Fine, 2015), but their interactions and consequences are species specific. Elucidating the influences of ecology and geography on speciation can benefit from inferring the population structure of species (Barraclough, Vogler & Harvey, 1998). The population structure of widespread species across the highlands can reflect patterns of gene flow, which can reveal geographical barriers (e.g mountains or rivers) with low dispersal capability or gene flow within specific environmental regions, independently of the

*Corresponding author. E-mail: itzuecs@uw.edu

barriers among them and their dispersal capability (Buckley *et al.*, 2010). Considering both geographical and ecological factors is important to make inferences about diversification processes, especially to avoid misleading interpretations when climatic events reduce the visible effects of geography (e.g. Pleistocene glaciations; Liu & Herbert, 2004). Identifying the geographical and ecological factors and their associated genetic variability is highly relevant for biological conservation. Maintaining high levels of genetic variability is the foundation for species persistence when facing environmental challenges (Vandergast *et al.*, 2008). However, the processes generating biogeographical patterns are rarely considered for conservation assessments (Moritz, 2002).

In the Mexican highlands, speciation events are affected by the geological history of the mountain systems, as well as ecological factors and climatic variables (McCormack *et al.*, 2008). Mountain forests (e.g. pine-oak forest and cloud forest) are distributed on a narrow elevation range and they share constant climatic conditions related to temperature, humidity and rainfall (Ornelas *et al.*, 2013). Consequently, their fragmentary distribution along the Mexican highlands has influenced the diversification of numerous groups of vertebrates (García-Moreno, *et al.*, 2004; León-Paniagua *et al.*, 2007; Navarro-Sigüenza *et al.*, 2008). The mountain forests in Mexico are considered to be the most diverse ecosystems in México and the second-richest ecosystem of vertebrates in Mesoamerica (Flores-Villea & Gerez, 1994). The biodiversity hot spots are mainly located within the cloud forests, despite the fact that cloud forest covers only 1% of the total country area, and there is only 50% of the original cloud forest coverage left (Ornelas *et al.*, 2013). Owing to the high rates of habitat loss (among other reasons), most of the species endemic to the mountain forests are considered endangered, and conservation policies do not account for species genetic diversity (IUCN, 2017). This is concerning, because although some species maintain similar biogeographical patterns, their specificity and levels of gene flow along the highlands are idiosyncratic.

In Mexico, amphibian diversity is closely associated with highland forest habitats with high humidity and rainfall (Pineda & Lobo, 2009; García, Ortega-Huerta & Martínez-Meyer. 2014). The highlands have high levels of endemism for tree frogs, and the species restricted to mountain forests are particularly sensitive to the climatic conditions (Smith *et al.*, 2007). Therefore, we selected a widespread tree frog species as our study system. The *Sarcohyla bistincta* (Cope, 1877) complex (Anura: Hylidae) is composed of three recognized species and five distinctive paraphyletic lineages (Caviedes-Solis & Nieto-Montes de Oca, 2018): *S. bistincta* (three lineages), *Sarcohyla pentheter* (Adler, 1965) and *Sarcohyla calthula* (Ustach *et al.*, 2000). These lineages are distributed along the main Mexican mountain systems, whose intricate topography combined with habitat shifts during the Pleistocene glaciation have previously been correlated with species diversification of tree frogs, e.g. *Hyla arenicolor* (Cope, 1866; Bryson *et al.*, 2010). In addition, tree frogs are prone to speciation related to habitat specificity (Hua & Wiens, 2010), making the *S. bistincta* complex an excellent study system to test the influence that geographical and ecological factors have in a mountain species complex.

The aim of the present study was to understand the patterns and processes influencing highland speciation. In particular, our goal was to disentangle how geographical (topographic) and ecological (climatic) factors have influenced species diversification by: (1) estimating the number of genetically distinctive populations within the *S. bistincta* complex across the Mexican highlands; (2) determining whether the mountain systems harbour independent and distinct genetic clusters, reflective of a biogeographical history of geographical isolation; and (3) estimating the ecological niche for each population and evaluating their degree of overlap across the geographical barriers. Our results also describe the genetic diversity across the highlands as a reference for future conservation efforts.

MATERIAL AND METHODS

Sampling

We sampled the clade that includes *S. pentheter*, *S. calthula* and *S. bistincta* (*sensu stricto*). We included four individuals for *S. pentheter*, three for *S. calthula* (both species with restricted distributions), and 51 individuals for the widespread species, *S. bistincta*. Samples correspond to 42 localities distributed in eight mountain systems: Cerro Piedra Larga (CPL), Sierra de Juarez (SJ), Sierra Mazateca (SMaz), Sierra Mixe (SMix), Sierra Madre Oriental (SMO), Sierra Madre Occidental (SMOcc), Sierra Madre del Sur (SMS) and Trans Volcanic Belt (TVB) (Table 1). We obtained tissue samples from two Herpetology collections: Museo de Zoología 'Alfonso L. Herrera' Facultad de Ciencias, Universidad Nacional Autónoma de México (MZFC) and University of Michigan Museum of Zoology (UMMZ).

Genetic data

We extracted total genomic DNA from liver or muscle, following the ammonium acetate procedure (Fetzner, 1999), and quantified the DNA using a Qubit Fluorometer. We sequenced Double Digest Restriction

Table 1. Distribution data for samples included in the study

Species Collector number	*N*	Mountain system	State: locality
Sarcohyla calthula (*N* = 4)			
AMH	1	SMix	OAX: Totontepec
JAC 21167	1	SMix	OAX: Santiago Zacatepec
JAC 22944	1	SMix	OAX: Santa Maria Guienagati
LCM	1	SMix	OAX: Santa Maria Alotepec
Sarcohyla pentheter (*N* = 3)			
ANMO 1057, 1058	2	SMS	OAX: San Gabriel Mixtepec
ISZ 497	1	SMS	OAX: Pluma Hidalgo
Sarcohyla bistincta (*N* = 51)			
ANMO 2728	1	CPL	OAX: Cerro Piedra Larga
ANMO 2352, 2353	2	SJ	OAX: Santiago Comaltepec
ANMO 2358	1	SJ	OAX: San Pedro Yolox
MK 716(1)	1	SJ	OAX: San Miguel Aloapan
MK 755(1)	1	SJ	OAX: Atepec Abejones
MK 705(2)	1	SMaz	OAX: Huahutla Oaxaca
RVG 199	1	SMaz	OAX: San Jeronimo Tecoatl
JLAL 174, UOGV 281	2	SMaz	OAX: Zoquiapan Sierra Monte Flor
UOGV 637, LCM 1202	2	SMaz	OAX: Tutepetongo
MK 748(2,4)	2	SMix	OAX: Ayutla Mixes
LCM 1140	1	SMix	OAX: Totontepec
CDR 241	1	SMO	PUE: San Bernardino
MK 700(2)	1	SMO	PUE: Zoquitlan Puebla
MK 697(3), 699(1)	2	SMO	VER: Xoxocotla
VHL 026(1–3)	3	SMOcc	NAY: Xalisco
ANMO 1898	1	SMS	OAX: Santa Maria Yucuhiti
ANMO 2640, 2644	2	SMS	OAX: San Vicente Laxichio
MK 721	1	SMS	OAX: Santa Maria Laxichio
JAC 22157	1	SMS	GRO: Carrizal de Bravo
JCBH 085	1	SMS	GRO: El Molote
MK 650(1–2)	2	SMS	GRO: Los Morros
MK 656(1)	1	SMS	GRO: Omiltemi
MK 662	1	SMS	GRO: Teitipac
MK 671(4), 672	2	SMS	GRO: Chichihualco
MK 685(2), 689 (2)	2	SMS	GRO: Puerto del Gallo
MK 691(4)	1	SMS	GRO: Yextla
MK 759, 760(1–2)	3	SMS	GRO: Ixtapa
UOGV 1834 -1836	3	SMS	GRO: Malinaltepec
MDL 083	1	TVB	JAL: Sierra de Quila
MK 618	1	TVB	MICH: Los Azufres
MK 627	1	TVB	MICH: Uruapan
MK 666	1	TVB	MICH: Zitacuaro
UM	1	TVB	MICH: Morelia
UOGV 1553	1	TVB	MICH: Indaparapeo
MK 600(1)	1	TVB	EDO MEX: Valle de Bravo
MK 645	1	TVB	MOR: Cuernavaca

Mountain systems are as follows: CP, Cerro Piedra Larga; SJ, Sierra de Juarez; SMix, Sierra Mixe; SMO, Sierra Madre Oriental; SMOcc, Sierra Madre Occidental; SMaz, Sierra Mazateca; SMS, Sierra Madre del Sur; TVB, Trans Volcanic Belt. State names are as follows: EDO MEX, Estado de Mexico; GRO, Guerrero; JAL, Jalisco; MICH, Michoacan; MOR, Morelos; OAX, Oaxaca; PUE, Puebla; VER, Veracruz.

Associated DNA (ddRAD), following the protocol described by Peterson *et al.* (2012). We double-digested samples with 500 µg for 2 h at 37 °C; we used a common cutter, *MspI* (restriction site 5′-CCGG-3′), and a rare cutter, *Sbf1* (restriction site 5′-CCTGCAGG-3′), using the manufacturer's recommended buffer

(New England Biolabs). We verified the digestion by running the samples on 2% agarose gel. Fragments were purified with AMPure XP beads before ligation of barcoded Illumina adaptors. Sequences for barcoding and Illumina indexes used in library preparations are provided by Peterson *et al.* (2012). After a second round of bead purification, we used a Pippin Prep size fractionator (Sage Science) to size select our libraries between 415 and 515 bp. For PCR amplification, we used proofreading Taq and Illumina's indexed primers. We estimated the size distribution and concentration of each pool on an Agilent 2200 TapeStation. Finally, the samples were pooled in equimolar rations and sequenced on a single Illumina HiSeq 2500 lane (50 bp, single-end reads) at the QB3 facility at UC Berkeley; ddRADseq data was deposited in the Dryad repository doi:10.5061/dryad.72r4b86

ddRADseq bioinformatics

We processed our raw Illumina reads with pyRAD v.3.0.64 (Eaton, 2014), which use a global alignment clustering that allows indel variation. We de-multiplexed the samples using the unique barcode and adapter sequences. We changed sites with Phred quality scores < 99.9% (Phred score = 33) to '*N*', and discarded reads containing > 10% *N*s (NQual = 4). We removed the 6 bp restriction site overhang (TGCAGG) and the 5 bp barcode, which reduced each locus from 50 to 39 bp. We assembled loci using a cluster threshold of 94% and we set the maximum proportion of shared polymorphic sites in a locus to 30% (MaxSH = 18). We selected the optimal threshold according to our data, following Ilut, Nydam & Hare (2014). We repeated the clustering process, incrementing sequentially the mismatches allowed between reads within a cluster (from 0.82 to 0.98 every 0.02). We calculated the proportion of clusters for homozygous and heterozygous (2 and > 2) and plotted the results. We selected 0.94 as the optimum, the point where there is an asymptote for all the haplotypes (Supporting Information, Appendix S1). As an additional filtering step, we discarded consensus sequences with low coverage (< 8), excessive heterozygous sites (> 5) or too many haplotypes (> 2 for diploids). To establish locus homology among species, we clustered consensus sequences across samples using identical thresholds used to cluster data within species. We aligned every locus using MUSCLE v 3.8.31 (Edgar, 2004). Finally, we assembled loci into three data matrices using different variables for the parameter minimal number of individuals (min. ind. = 52, 43 or 29). By increasing the number of individuals required to have data at each locus, the missing data allowed decreases, resulting in lower numbers of single nucleotide polymorphisms (SNPs). In this case, min. ind. = 52, 43 and 29 corresponds to 10, 25 and 50% missing data, respectively.

Population structure

To determine the factors influencing the population structure, we first estimated the number of independent genetic clusters within the *S. bistincta* complex. We used the clusters compositions and the patterns of gene flow among clusters to disentangle how geographical and environmental factors contributed to speciation. To identify the genetic clusters, we used the Adegenet software package (Jombart, Devillard & Balloux, 2010) in the R v3.3.2 (2016) software environment. Adegenet implements a discriminant analysis of principal components (DAPC), which estimates the differences between groups while minimizing variation within clusters. We performed independent DAPC on three pyRAD structure output files that contained unlinked SNPs corresponding to 10, 25 and 50% missing data, respectively. Jombart & Collins (2015) reported that unlike k-means, using fewer principal components (PCs) with DAPC has the advantage of reducing over-fitting to discriminant functions and instability of membership probabilities. Therefore, we used the first 19 PCs, corresponding to a third of the total samples (the maximum number suggested); we saved two discriminant functions. We also implemented the *find.clusters* function to identify the appropriate number of genetic clusters and to assign individuals to those clusters. We ran the k-means clustering algorithm successively for up to 15 groups using 58 PCs. We selected the appropriate number of clusters based on the value of k-means with the lowest Bayesian information criteria (BIC) score. Finally, we visualized the resultant clusters on DAPC scatterplots.

Biogeographical sample clustering

We partitioned samples into their corresponding mountain system to test whether clustering by biogeographical region produced genetically independent groups (Table 1). If each mountain system represents an independent genetic cluster, we would expect that their limits on the DAPC scatterplot would be well defined and distinctive. However, if there is overlap among the DAPC clusters, then the mountain systems might not represent biogeographical barriers between genetically distinctive populations. We performed a DAPC using the Adegenet package (Jombart *et al.*, 2010) in the R v3.3.2 (2016) software environment. For DAPC, we used the first 15 PCs, we saved two discriminant functions, and we visualized the resultant clusters on DAPC scatterplots for each group.

Phylogenetic analysis

We inferred the phylogeny for the group to determine the order of speciation events and the evolutionary relationships among lineages. Both are necessary for determining the biogeographical and ecological events related to speciation. We estimated the species tree with SNAPP (Bryant *et al.*, 2012) using BEAST2 v2.4.5 (Bouckaert *et al.*, 2014), which implements a multispecies coalescent model for unlinked bi-allelic markers (Bryant *et al.*, 2012). We coded the three SNP's matrices from pyRAD as bi-allelic markers using the R package Phrynomics (Leaché *et al.*, 2015). We assigned samples to populations according to the cluster results obtained from Adegenet. We used a gamma distribution with mean = 10 for the prior on the coalescence rate. We sampled mutation rates U and V, both with an upper value of 20, and we set their priors to Exponential (mean = 1). We performed two Markov chain Monte Carlo (MCMC) replicates of 500 000 generations each, and sampled every 100 steps. We used the program Tracer (Rambaut & Drummond, 2007) to assess stationarity and convergence of the two independent runs. We then combined the post-burn-in samples from the two individual replicates using LogCombiner v2.2.1. (Rambaut & Drummond, 2011), and finally, we used TreeAnnotator v2.2.1 (Rambaut & Drummond, 2012) to calculate the maximum-clade credibility (MCC) tree and to summarize the estimated posterior support for each branch.

Niche modelling

We modelled the ecological niche for all the independent genetic clusters of the *S. bistincta* complex. This group of frogs is highly sensitive to climatic conditions (e.g. rainfall and temperature). Therefore, we tested whether their distribution models along the highlands reflect their genetic identity (independently of the climatic variables) or if the distribution models reflect their ecology (independently of the genetic cluster to which they belong). Our total sampling for the modelling (after removal of duplicate coordinates) includes 97 individuals: 41 from the present study and 56 downloaded from GBIF (www.gbif.org; Supporting Information, Appendix S2). We estimated a minimum convex polygon for each of the seven genetically distinctive populations recovered, based on SNP clustering analyses, and then assigned GBIF occurrence data to each polygon (Supporting Information, Appendix S3). We used 19 bioclimatic variables from WorldClim (www.worldclim.org)) using the grid size of 2.5 arc-min (4.63 km at the equator; Hijmans *et al.*, 2005). We tested for collinearity among variables and selected a subset of the variables for ecological niche modelling (Supporting Information, Appendix S4). To estimate the ecological niche model for each population, we implemented the maximum entropy method in Maxent v3.4.1 (Phillips, Dudík & Schapire, 2017) using the R package 'Dismo' (Phillips, Anderson & Schapire, 2006; Hijmans *et al.*, 2011) in R v3.2.3 (2016). We assessed the significance of the model predictions by estimating the area under the curve (AUC; Phillips *et al.*, 2006). This method provides a measure of model accuracy with AUC values ranging from zero to one, where a score of one indicates high predictive performance, and ≤ 0.5 indicates model accuracy not better than random (Fielding & Bell, 1997).

To determine whether the ecological factors were related to speciation events, we tested whether the distribution models estimated for each lineage were independent. We calculate niche overlapping among models by using Schoener's D (Schoener, 1968) and I statistics; both measure and range the overlap between pairwise models from zero (no overlap) to one (identical) (Warren, Glor & Turelli, 2008). To evaluate niche overlap, we performed a 'background' or 'similarity test', which determines whether the observed niche overlap is significant compared with the overlapping expected by chance (Warren *et al.*, 2008). This test samples points randomly within the environmental conditions from the distribution models where both populations (combined) occur, and then it calculates the overlap between them. We performed 100 replicates for each of the possible combinations of lineages (N = 21). Niche overlap and the similarity test were calculated using the R package ENMTools (Warren, Glor & Turelli 2010). We plotted the results in a three-dimensional frame as a sidewise comparison between the lineages overlapping and the mountain systems; the x- and y-axes correspond to the longitude and latitude, respectively, and the z-axis corresponds to the elevation (in metres) or to the number of lineages per area. We considered a strong overlap when the value was > 75 and if it was significantly different than predicted by random according to the 'similarity test' (P-value > 0.5) (Swets, 1988; Elith, 2002; Kalkvik *et al.*, 2012).

Conservation mapping

Cloud forests distributed along the Mexican highlands are protected based on hierarchical priority areas (CONABIO, 2008). We evaluated whether the current designated priority areas correspond to the areas of the highest genetic diversity for the *S. bistincta* complex. Priority areas for cloud forest conservation (44 in total) were pre-assigned by CONABIO (2008) according to four categories: high (17), critical (15), medium (10) and pending (3). Each category was determined based on four main factors, namely: quality (e.g. species diversity, cover and connectivity); threat to

permanence (including multiple causes of habitat loss); opportunities for conservation (measures of previous conservation efforts); and social characteristics (e.g. marginalization). We calculated the number of lineages per raster cell for each priority area depending on their conservation category (minimum, mean and maximum), and the total area that lineages occupy outside of the conservation areas, both using the R package ENMTools (Warren *et al.*, 2008).

RESULTS

ddRADseq summary statistics

A summary of the ddRADseq data is provided in Table 2. On average, each sample had > 1 400 000 reads and nearly 10 000 loci (Table 2). The average number of raw reads and loci that passed quality filters was consistent among species; mean number of loci: *S. bistincta* = 11 005; *S. calthula* = 10 516; and *S. pentheter* = 13 572 (Table 2). The number of SNPs and loci in the final alignments were strongly influenced by the stringency that we assigned to the minimum number of individuals parameter. Close relatives share more SNPs than distant relatives. Therefore, when the number of individuals distantly related that are required to have data at each locus increases (less missing data allowed), the number of SNPs decline (Leaché & Oaks, 2017). This was the case for the present study; the number of ddRAD loci, SNPs and unlinked SNPs increased dramatically when larger amounts of missing data were allowed (Table 3). For example, the data matrix containing 10% missing data (min. ind. = 52) contains 721 loci and 2515 SNPs (unlinked SNPs = 666), whereas the matrix containing 50% missing data contains 3313 loci and 11 484 SNPs (unlinked SNPs = 3003) (Table 3).

Table 2. Characteristics of the ddRADseq data, sorted by species

Species	*N*	Reads*	Depth†	Loci‡
Sarcohyla bistincta	51	1 584 168	23.9	11 005
Sarcohyla calthula	4	1 461 547	24.7	10 516
Sarcohyla pentheter	3	1 484 718	20.6	13 572

N is the number of individuals.
*Mean number of raw read after de-multiplexing.
†Mean sequencing depth.
‡Mean number of loci passing quality filters.

Table 3. Summary of the single nucleotide polymorphism data matrices analysed for the *Sarcohyla bistincta* complex

Min. Ind.	Missing Data (%)	Loci	Total SNPs	Unlinked SNPs
52	10	721	2,515	666
43	25	1,613	5,784	1,493
29	50	3,313	11,484	3,003

Min. Ind., minimum number of individuals required to have data present at a locus for that locus to be included in the final matrix; SNP, single nucleotide polymorphism; Total SNPs, count of all variable sites; Unlinked SNPs, one SNP drawn from each locus.

Population structure

Adegenet clustering was mostly consistent across different percentages of missing data. The number of clusters (k) estimated using the BIC was k = 7 for all the levels of missing data (Supporting Information, Appendix S5). Adegenet clustering was mostly consistent across different percentages of missing data. The assignment of individuals to clusters was the same for 25 and 50% missing data; they differ from the 10% missing data results, where: (1) *S. bistincta* 3b is subdivided into two clusters (3b-1 and 3b-2); and (2) *S. bistincta* 1a and 1b are combined in the same cluster instead of subdivided in western TBV and SMOcc (1a) and eastern TVB (1b) (Supporting Information, Appendix S5). The Adegenet DAPC plot for the dataset corresponding to 50% missing data is shown in Fig. 1. Individuals are distributed among the seven clusters as follows: *S. calthula* (four from SMix); *S. pentheter* (three from SMS); *S. bistincta* 1a (one from TVB and three from SMOcc); *S. bistincta* 1b (six from TVB); *S. bistincta* 2 (seven from SMS and one from SMO); *S. bistincta* 3a (four from SMS, one from CP, three from SMix, six from SJ, six from SMaz, and three from SMO); and *S. bistincta* 3b (ten from SMS).

Biogeographical sample clustering

We partitioned the lineages into two groups according to the two main clades in the phylogeny, as follows: group 1, *S. calthula*, *S. pentheter*, *S. bistincta* 1a and 1b and *S. bistincta* 2; and group 2, *S. bistincta* 3a and 3b (Fig. 2B). Our findings suggest that the population clusters corresponding to *S. bistincta* 2 (SMS Oax and SMO), *S. calthula* (SMix) and *S. pentheter* (SMS Gro) are well-defined groups that do not overlap, and that these lineages are genetically distinct (Fig. 2C). *Sarcohyla bistincta* 1a (SMOcc + TVB) and 1b (TVB) overlap slightly, but are mainly defined according to their mountain system and represent distinctive genetic populations (Fig. 2C). For the populations belonging to *S. bistincta* 3 (3a and 3b), the most widespread lineages of the complex, most of the mountain systems have a genetic overlap according to their

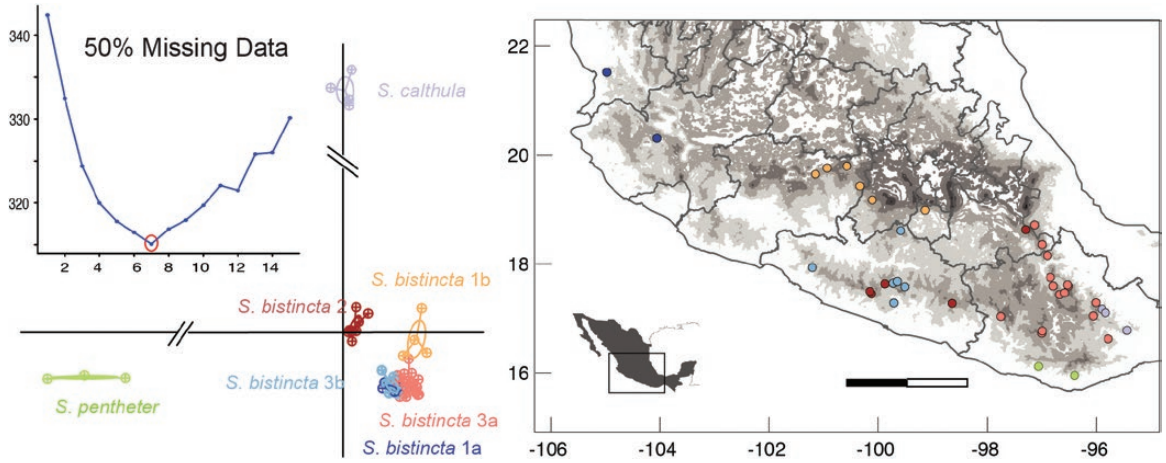

Figure 1. Discriminant analysis of principal components (DAPC) clustering results for the *Sarcohyla bistincta* complex. The analyses were produced using the data matrix containing 50% missing data.

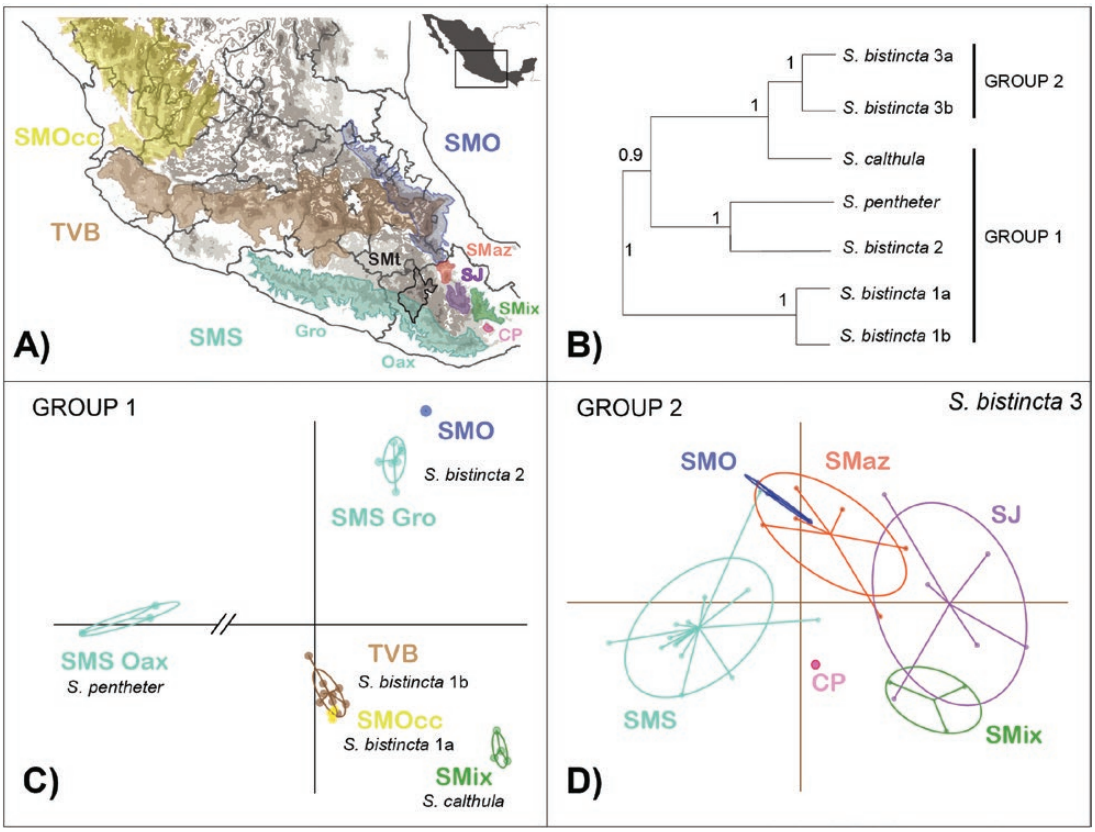

Figure 2. A, mountain systems distributions. B, species tree estimated with SNAPP. Numbers on nodes are posterior probability values. C, discriminant analysis of principal components (DAPC) clustering results for group 1 (*Sarcohyla bistincta* 1a, 1b and 2, *Sarcohyla calthula* and *Sarcohyla pentheter*), organized by mountain system. D, DAPC clustering results for group 2 (*S. bistincta* 3a and 3b), organized by mountain system.

geographical proximity to each other. The sequential order of overlapping is SMO-SMaz, SMaz-SJ and SJ; SMS overlaps with SMO as the result of the clustering of one individual (Fig. 2D). Finally, SMix and CP are isolated clusters, located in the eastern section of its distribution (Fig. 2D).

Phylogenetic analysis

The phylogeny with the highest support and larger number of loci (50% missing data) is shown in Fig. 2B. The five *S. bistincta* populations are not monophyletic and are divided into two groups. *Sarcohyla pentheter* is the sister taxon of *S. bistincta* 2, and *S. calthula* is sister to the clade composed by *S. bistincta* 3a and 3b. *Sarcohyla bistincta* 1 and *S. bistincta* 3 were subdivided into two independent lineages each (1a, 1b, 3a and 3b). The support for relationships was high, with posterior probabilities ≥ 0.9. Species trees estimated using fewer SNPs with fewer levels of missing data allowed (10 and 25% missing data) differ from the species tree shown in Fig. 2B with respect to the placement of *S. calthula* (Supporting Information, Appendix S6), which was placed as the sister taxon of *S. bistincta* 1a and 1b, but with low support (posterior probability < 0.6).

Niche modelling

Mexican highlands and their associated fragments of cloud forest in the southeast are the areas with the most niche overlap among populations. Figure 3 shows the comparison between the elevation gradient of the mountain systems (Fig. 3A) and the number of lineages per area (Fig. 3B).

Despite having a complex topography (Fig. 3A), this visualization allows us to remark that not all the mountain systems are equally diverse; instead, the higher overlapping of lineages is located in the south east (Fig. 3B). The high degree of overlapping among lineages also supports the niche conservatism hypothesis, which states that independent genetic clusters of the *S. bistincta* complex share similar ecological niches. We found that this prediction is supported by four pairwise comparisons estimated with *I* statistics: *S. bistincta* 3a with respect to *S. bistincta* 1a, *S. bistincta* 1b, *S. bistincta* 2 and *S. pentheter*. The lineage comparisons that did not fit this predication were *S. calthula* with respect to *S. bistincta* 1b, *S. bistincta* 2, *S. bistincta* 3a and *S. pentheter*. For the remaining 13 pairwise comparisons, the overlap values estimated with *I* statistics were not significant (Supporting Information, Appendix S7). The values estimated with *D* statistics were lower than those obtained with *I* statistics (Table 4). The patterns of overlap were not consistent with geographical proximity, not even for sympatric lineages.

The distribution model for each of the seven populations is shown in Supporting Information, Appendix S8. The performance of the models, as measured by the AUC, and the contribution of each bioclimatic variable to the models is shown in

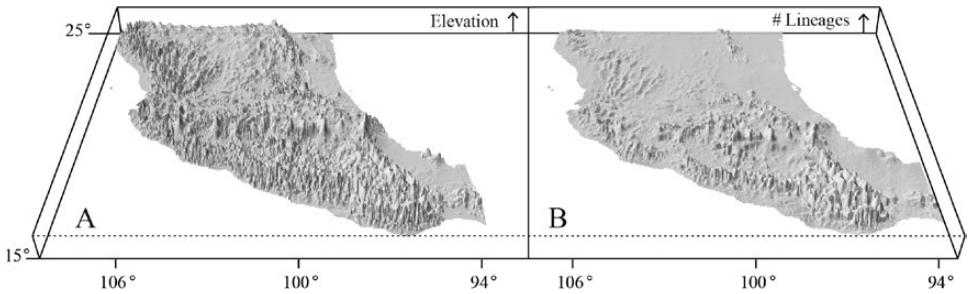

Figure 3. A, topography of mountain systems in Mexico. B, niche overlapping among lineages per area. Axes *x* and *y* correspond to the longitude and latitude, and the *z*-axis corresponds to the elevation (A) or the number of lineages per area (B).

Table 4. Background similarity test between populations. Schoener's *D* (lower diagonal) and *I* (upper diagonal) statistics

	Sb1a	Sb1b	Sb2	Sb 3a	Sb 3b	Sca	Spe
AUC	0.50	0.77	0.71	0.79	0.5	0.58	0.84
Sarcohyla bistincta 1a		0.88	0.71	0.86*	0.90	0.69	0.70
Sarcohyla bistincta 1b	0.63		0.67	0.80*	0.89	0.52*	0.57
Sarcohyla bistincta 2	0.39	0.33		0.83*	0.94	0.47*	0.95
Sarcohyla bistincta 3a	0.62*	0.52*	0.60		0.89	0.59*	0.75*
Sarcohyla bistincta 3b	0.69	0.59	0.71	0.64		0.67	0.92
Sarcohyla calthula	0.44	0.26*	0.25*	0.26*	0.37		0.60*
Sarcohyla pentheter	0.40	0.27	0.78	0.45*	0.69	0.35*	

Significant *P*-values are indicated with an asterisk. Area under the curve (AUC) values correspond to the performance of each model.

Table 4 and Supporting Information, Appendix S4 (respectively). Temperature variables contributed the most to the models, including temperature seasonality (*S. bistincta* 2, *S. bistincta* 3a, *S. bistincta* 3b and *S. pentheter*), maximum temperature of the warmest month (*S. bistincta* 1a), mean temperature of the wettest quarter, and temperature annual range (*S. calthula*).

Conservation mapping

We evaluated whether the current designated priority areas correspond to the areas of the highest genetic diversity for the *S. bistincta* complex by calculating the number of lineages per raster cell for each priority area (Fig. 4). The average number of lineages per raster cell was less than three, the minimum was zero, and maximum was five, for all priorities areas independently of the priority level (Supporting Information, Appendix S9).

DISCUSSION

Phylogeny of the group

Previous multilocus evidence suggested that the *S. bistincta* complex is composed of five main lineages, including three distinct lineages of *S. bistincta*, *S. calthula* and *S. pentheter* (Caviedes-Solis & Nieto-Montes de Oca, 2018). These are concordant with our findings based on SNP data; however, the present analysis subdivided *S. bistincta* further into five genetically independent groups instead of three (Fig. 1). A recent study recovered eight independent lineages within the *S. bistinca* complex using ultraconserved elements (UCE; Zarza *et al.*, 2017). They recovered similar patterns with two exceptions: (1) UCE data support three independent lineages along the TVB within *S. bistincta* 1a; and (2) the lineage *S. bistincta* 2 is composed of two independent lineages. *Sarcohyla calthula* and *S. pentheter* have always been recovered as sister taxa in previous phylogenetic inferences (Caldwell, 1974; Faivovich *et al.*, 2005; Pyron & Wiens, 2011; Duellman, Marion & Hedges, 2016; Caviedes-Solis & Nieto-Montes de Oca, 2018). Nonetheless, our species tree inferences based on SNP data suggest that *S. calthula* is an independent lineage, and *S. pentheter* is the sister taxon of *S. bistincta* 2 with high posterior probabilities (PP ≥ 95). This is concordant, in part, with the UCE analysis, which also recovered *S. bistincta* 2 as an independent lineage that is sister to *S. pentheter* (Zarza *et al.*, 2017).

Owing the lack of a fossil record for the genus *Sarcohyla*, divergence time estimates for this clade are controversial. Previous studies relied on secondary calibration points (Caviedes-Solis & Nieto-Montes de Oca, 2018), which are known to produce unreliable estimates (Schenk, 2016). Therefore, the goal of the present manuscript is not to associate divergence with

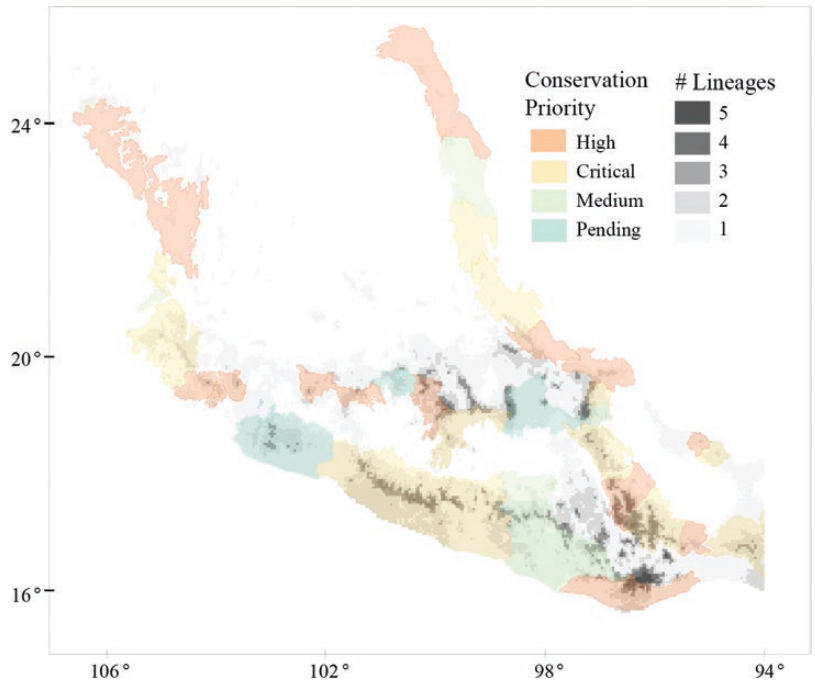

Figure 4. Niche overlap among populations and cloud forest distributions divided according to their conservation priority ranking (CONABIO, 2008).

geological events, but instead to determine whether the lineages distributed in different mountain systems are genetically independent.

Biogeographical factors influencing diversification

The geological history of the Mexican highlands has influenced the diversification of many vertebrates, and although many species maintain the same biogeographical patterns, their specificity and levels of gene flow are idiosyncratic (García-Moreno *et al.*, 2004; León-Paniagua *et al.*, 2007). Within frogs, population structure is highly determined by geological events and landscape modifications, especially in montane forest ecosystems, where habitat fragmentation represents a barrier to dispersal (Morrone & Crisci, 1995). Our study suggests that mountain systems in Mexico had a differential impact on the population structure of the *S. bistincta* complex. Species within this complex share similar morphologies, natural histories and dispersal capabilities, which minimize the intrinsic factors causing diversification.

In the *Sarcohyla bistinca* complex, lineages restricted to one mountain system include *S. bistincta* 1b (TVB), *S. bistincta* 3b (SMS), *S. pentheter* (SMS) and *S. calthula* (SMixe). Those three mountain systems (TVB, SMS and SMixe) are considered to be important biogeographical areas that have played a role in the diversification of multiple species, particularly those with low dispersal capabilities. The formation of the TVB (23–2.4 Mya) is considered to be one of the most important vicariant events in Mexico (Navarro-Sigüenza *et al.*, 2008; Ornelas *et al.*, 2013). The uplift of the TVB started in the west and continued to the east until it impacted the continuity of the SMO (Rosas-Elguera *et al.*, 2003; Becerra, 2005). Simultaneously, a secondary uplifting of the southern section of the SMOcc impacted the western and oldest section of the TVB (Riddle & Hafner, 2006). Splits between populations of the west TVB and southern SMOcc, like the ones recovered for *S. bistincta* 1b and 1a, have been reported previously in comparative phylogeographical analyses of cloud forest plants (Luna-Vega, Almeida-Leñero & Llorente-Bousquets, 1989; Luna-Vega *et al.*, 1999) and snakes in the genus *Pituophis* (Bryson, García-Vázquez & Riddle, 2011), all with different levels of specificity.

The SMS is one of the oldest mountain systems in México (Morán-Zenteno *et al.*, 1999). It extends across southern Mexico, along Guerrero (Gro) and Oaxaca (Oax) states, and is divided by the Sierra Mixteca (SMt). The SMt represents an important biogeographical barrier affecting the continuity of species distributions along the SMS (Nieto-Samaniego *et al.*, 2006). Within the *S. bistincta* species complex, the distributions of *S. pentheter* and *S. bistincta* 3 were considered to include the full extent of the SMS (both Gro and Oax). However, a recent molecular phylogeny revealed that *S. pentheter* is endemic only to the SMS (Oax) east of the SMt (Caviedes-Solis & Nieto-Montes de Oca, 2018). That result is also supported by the present study; in addition, we found that the SMt represents a barrier for *S. bistincta* 3, which is composed of two independent lineages (*S. bistincta* 3a and 3b). Finally, *S. calthula* (SMixe) is isolated by multiple geographical barriers, which include Oaxacan central valleys and multiple tributary rivers from the Atlantic (Martínez-Ramírez *et al.*, 2004).

The lineages *S. bistincta* 1a, *S. bistincta* 2 and *S. bistincta* 3a are each distributed in more than one mountain system. Despite the multiple geographical barriers that could potentially subdivide populations, the genetic clusters for those lineages are not independent. For example, we find no evidence using SNP data that the rise of the TVB and the secondary uplift of the SMOcc influenced the divergence of *S. bistincta* 1a, and the central valleys in Oaxaca did not contribute to the divergence of *S. bistincta* 2 distributed in the SMS and the SMO. The populations belonging to *S. bistincta* 3 are distributed on seven mountain systems across all the major biogeographical barriers previously mentioned, including those in Oaxaca known for their high rate of diversity and endemism (Parra-Olea, Flores-Villela & Mendoza-Almeralla, 2014). Previous biogeographical analyses recovered similar patterns for species of mammals, birds and reptiles distributed across multiple Mexican highlands (Sullivan, Markert & Kilpatrick, 1997; Navarro-Sigüenza *et al.*, 2008; Bryson *et al.*, 2011). Some of the connections are related to high dispersal capabilities of each species, but they can also be explained by expansions and contractions of pine-oak forest across the Mexican plateau during the Pleistocene glaciations (Bryson *et al.*, 2011).

Ecological factors influencing diversification

Models of speciation involving ecological niches can be categorized broadly as niche conservatism and niche divergence (Hua & Wiens, 2010). Our findings support the niche conservatism hypothesis within the *S. bistincta* complex. Their distribution models reflect that some lineages retained their niches across the geographical barriers and independently of the genetic cluster to which they belong. In addition, the niche models overlap independently of geographical or genetic distances. Therefore, we cannot infer that niche differentiation has an influence on lineage divergence. Is important to consider that most of the values recovered with *D* and *I* statistics show a pattern of niche overlap among lineages, and the lack

of significance for some lineages might be the effect of sample size. Some populations have only a few individuals, and the resultant overlap values are not significant when compared with the values estimated by random sampling (Warren *et al.*, 2008).

Ecological speciation along the Mexican highlands is taxon specific, similar to the differential influence that geographical factors have on species divergence. Ibarra-Cerdeña *et al.* (2014) also found evidence of niche conservatism for a species complex of Mexican insects. They recovered the same distribution models independently of the size of their area, the region where they are located or the degree of niche overlap. Some species, however, show evidence of niche divergence, including populations of angiosperms with different (non-overlapping) niches along the TVB (Ruiz-Sanchez & Specht, 2014). Niche modelling estimates the distributions based on recent climatic conditions, but is important to consider that climatic conditions constituting the ecological niches for each lineage vary through time. Past climatic variations and vicariant events interact to affect the connectivity among lineages. For example, secondary contacts among highland populations during the Pleistocene influenced the phylogeography of temperate species (Ornelas *et al.*, 2013).

Conservation areas and species diversity

Cloud forests are considered to be one of the most threatened ecosystems in the world (Toledo-Aceves, *et al.*, 2011). In Mexico, they are distributed along the highlands and represent the refuge for a large number of the endemic species. Mexican cloud forests are protected based on hierarchical priority areas (CONABIO, 2008). These conservation areas were ranked based on species richness and endemism (among other factors). However, previous studies have shown that genetic diversity is an important factor for conservation, because high genetic diversity has a positive effect to overcome species threats (Luck, Daily & Ehrlich, 2003; Hughes & Stachowicz, 2004). Our results suggest that conservation areas with the highest priority level do not correspond to the areas containing *S. bistincta* lineage diversity (e.g. areas in the mountain systems in Oaxaca). Instead, the average number of lineages protected by each conservation priority level appears random.

Genetic variation is one of the three measures of biodiversity recommended for conservation (McNeely *et al.*, 1990). However, genetic diversity and the factors influencing diversification are not usually determining factors for conservation initiatives. We found a lack of concordance between priority areas and *Sarcohyla* diversity, which is unfortunate given the fast rates of decline for tree frog species (Frías-Alvarez, Zúniga-Vega & Flores-Villela, 2010). One positive benefit of conserving genetic diversity is that the overall time required for species recovery decreases when genotypic diversity increases (Reed & Frankham, 2003; Hughes & Stachowicz, 2004). In addition, biogeographical and ecological factors influencing speciation can be combined to infer the degree of species threats and species recovery. For example, species with multiple populations that lack genetic or ecological distinctiveness, but are separated by geographical barriers, might be less sensitive than species with multiple populations that are genetically independent. Therefore, more detailed analyses including a larger fraction of tree frog diversity are needed to design conservation areas that successfully integrate cloud forest and frog conservation.

CONCLUSION

Our study contributes to the knowledge of Hylid frog biogeography and conservation in México. Tree frogs from the family Hylidae are an important group of species, which are particularly sensitive owing to their strict habitat requirements and low dispersal capabilities. The phylogeography of the *S. bistincta* complex has been influenced by the uplift of the Mexican highlands. Some lineages were more susceptible to geological events and Pleistocene climate changes than others; many lineages are microendemics, whereas others are widespread across the highlands. These mountain systems are considered to be biodiversity hotspots for amphibians and are responsible, in part, for the high rates of endemism. For lineages that lack niche differentiation, divergence is mostly correlated with geographical barriers. However, for the lineages in which neither niche differentiation nor geographical barriers represent vicariant events, other unknown factors might be playing a role. For example, why *S. bistincta* 3 was able to leapfrog over mountains whereas the rest of the lineages did not is a major question that remains to be answered. All the lineages share similar natural history, and the only morphological difference among the lineages of *S. bistincta* with respect to *S. pentheter* and *S. calthula* (besides some colour patterns) is the absence of the vocal slits (Duellman, 2001). In addition, all the lineages in the *S. bistincta* complex have similar ecological niches, and the extended overlap among them suggests that ecological differentiation did not play a role in their genetic divergence.

ACKNOWLEDGEMENTS

Itzue W. Caviedes-Solis was funded by a Consejo Nacional de Ciencia y Tecnologia, Mexico CONACyT

fellowship number 313501, 'Becas CONACyt al Extranjero' and the Burke Museum of Natural Hisory. We thank to O. A. Flores-Villela and A. Nieto-Montes de Oca from Museo de Zoología 'Alfonso L. Herrera' Facultad de Ciencias, Universidad Nacional Autónoma de México, M. Kaplan Zoology Museum, University of Michigan, M. Dominguez-Lasso, V. H. Luja, Jorge Orozco and J. Orozco for the donation of tissue samples. This work used the Vincent J. Coates Genomics Sequencing Laboratory at the University of California Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. We thank three anonymous reviewers for their helpful comments; your positive and constructive feedback improved the quality of this manuscript.

REFERENCES

Adler KA. 1965. Three new frogs of the genus Hyla from the Sierra Madre de Sur of Mexico. *Occasional Papers of the Museum of Zoology, University of Michigan* **642:** 1–18.

Barraclough TG, Vogler AP. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. *The American Naturalist* **155:** 419–434.

Barraclough TG, Vogler AP, Harvey PH. 1998. Revealing the factors that promote speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **353:** 241–249.

Becerra JX. 2005. Timing the origin and expansion of the Mexican tropical dry forest. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 10919–10923.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond A. 2014. BEAST2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10:** e1003537.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg N, RoyChoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution* **29:** 1917–1932.

Bryson RW, García-Vázquez UO, Riddle BR. 2011. Phylogeography of Middle American gophersnakes: mixed responses to biogeographical barriers across the Mexican Transition Zone. *Journal of Biogeography* **38:** 1570–1584.

Bryson RW Jr, Nieto-Montes de Oca A, Jaeger JR, Riddle BR. 2010. Elucidation of cryptic diversity in a widespread Nearctic treefrog reveals episodes of mitochondrial gene capture as frogs diversified across a dynamic landscape. *Evolution* **64:** 2315–2330.

Buckley LB, Davies TJ, Ackerly DD, Kraft NJ, Harrison SP, Anacker BL, Cornell HV, Damschen EI, Grytnes JA, Hawkins BA, McCain CM. 2010. Phylogeny, niche conservatism and the latitudinal diversity gradient in mammals. *Proceedings of the Royal Society of London B: Biological Sciences* **277:** 2131–2138.

Caldwell JP. 1974. A re-evaluation of the *Hyla bistincta* species group, with descriptions of three new species (Anura: Hylidae). *Occasional Papers of the Museum of Natural History, University of Kansas* **28:** 1–37.

Caviedes-Solis IW,Nieto-Montes de Oca A. 2018. A multilocus phylogeny of the genus *Sarcohyla* (Anura: Hylidae), and an investigation of species boundaries using statistical species delimitation. *Molecular Phylogenetics and Evolution* **118:** 184–193.

CONABIO 2008. *Categoría de prioridad para la conservación del bosque mesófilo de montaña en México.* Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) Metadatos.

Cope ED. 1866. On the structures and distribution of the genera of the arciferous Anura. *Journal of the Academy of Natural Sciences of Philadelphia* Series 2, **6:** 67–112.

Cope ED. 1877. Tenth contribution to the herpetology of tropical America. *Proceedings of the American Philosophical Society* **17:** 85–98.

Duellman WE. 2001. *The hylid frogs of Middle America*. Ithaca, NY, USA: Society for the Study of Amphibians and Reptiles, 1170.

Duellman WE, Marion AB, Hedges SB. 2016. Phylogenetics, classification, and biogeography of the treefrogs (Amphibia: Anura: Arboranae). *Zootaxa* **4104:** 1–109.

Eaton DAR. 2014. PyRAD: assembly of *de novo* RADseq loci for phylogenetic analyses. *Bioinformatics* **30:** 1844–1849.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32:** 1792–1797.

Elith J. 2002. Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In: Ferson S, Burgman M, eds. *Quantitative methods for conservation biology*. Springer, 39–58.

Faivovich J, Haddad CFB, Garcia PCA, Frost DR, Campbell JA, Wheeler WC. 2005. Systematic review of the frog family Hylidae, with special reference to Hylinae: phylogenetic analysis and taxonomic revision. *Bulletin of the American Museum of Natural History* **294:** 1–240.

Fetzner JW Jr. 1999. Extracting high-quality DNA from shed reptile skins: a simplified method. *BioTechniques* **26:** 1052–1054.

Fielding AH, Bell JF. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* **24:** 38–49.

Fine PV. 2015. Ecological and evolutionary drivers of geographic variation in species diversity. *Annual Review of Ecology, Evolution, and Systematics* **46:** 369–392.

Flores-Villela O, Gerez P. 1994. *Biodiversidad y Conservación en México: Vertebrados y uso de Suelo*. Comisión Nacional para el Uso y Conservación de la Biodiversidad, Universidad Nacional Autónoma de México.

Frías-Alvarez P, Zúniga-Vega JJ, Flores-Villela O. 2010. A general assessment of the conservation status and decline trends of Mexican amphibians. *Biodiversity and Conservation* **19:** 3699–3742.

García A, Ortega-Huerta MA, Martínez-Meyer E. 2014. Potential distributional changes and conservation priorities of endemic amphibians in western Mexico as a result of climate change. *Environmental Conservation* **41:** 1–12.

García-Moreno J, Navarro-Sigüenza AG, Peterson AT, Sánchez-González LA. 2004. Genetic variation coincides with geographic structure in the common bush-tanager (*Chlorospingus ophthalmicus*) complex from Mexico. *Molecular Phylogenetics and Evolution* **33:** 186–196.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25:** 1965–1978.

Hijmans RJ, Phillips S. Leathwick J, Elith J. 2011. *Package 'dismo'*. Available at: http://cran.r-project.org/web/packages/dismo/index.html

Hua X, Wiens J. 2010. Latitudinal variation in speciation mechanisms in frogs. *Evolution* **64:** 429–443.

Hughes AR, Stachowicz JJ. 2004. Genetic diversity enhances the resistance of a seagrass ecosystem to disturbance. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 8998–9002.

Ibarra-Cerdeña CN, Zaldívar-Riverón A, Peterson AT, Sánchez-Cordero V, Ramsey JM. 2014. Phylogeny and niche conservatism in North and Central American triatomine bugs (Hemiptera: Reduviidae: Triatominae), vectors of Chagas' disease. *PLoS Neglected Topical Diseases* **8:** 3266.

Ilut DC, Nydam ML, Hare MP. 2014. Defining loci in restriction-based reduced representation genomic data from nonmodel species: sources of bias and diagnostics for optimal clustering. *BioMed Research International* **2014:** 675158.

IUCN 2017. *The IUCN red list of threatened species, Version 2015–3*. Available at http://www.iucnredlist.org

Jombart T, Collins C. 2015. *A tutorial for discriminant analysis of principal components (DAPC) using* adegenet *2.0.0*. Avaliable at: http://cran.r-project.org/web/packages/adegenet/

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11:** 94.

Kalkvik HM, Stout IJ, Doonan J, Parkinson CL. 2012. Investigating niche and lineage diversification in widely distributed taxa: phylogeography and ecological niche modeling of the *Peromyscus maniculatus* species group. *Ecography* **35:** 54–64.

Knowles LL, Massatti R. 2017. Distributional shifts – not geographic isolation – as a probable driver of montane species divergence. *Ecography* **40:** 1475–1485.

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. *Systematic Biology* **64:** 1032–1047.

Leaché AD, Oaks JR. 2017. The utility of single nucleotide polymorphism (SNP) data in phylogenetics. *Annual Review of Ecology, Evolution, and Systematics* **48:** 69–84.

León-Paniagua L, Navarro-Sigüenza AG, Hernández-Baños BE, Morales JC. 2007. Diversification of the arboreal mice of the genus *Habromys* (Rodentia: Cricetidae: Neotominae) in the Mesoamerican highlands. *Molecular Phylogenetics and Evolution* **42:** 653–664.

Liu Z, Herbert TD. 2004. High-latitude influence on the eastern equatorial Pacific climate in the early Pleistocene epoch. *Nature* **427:** 720–723.

Luck GW, Daily GC, Ehrlich PR. 2003. Population diversity and ecosystem services. *Trends in Ecology & Evolution* **18:** 331–336.

Luna-Vega I, Alcántara-Ayala O, Espinosa-Organista D, Morrone JJ. 1999. Historical relationships of the Mexican cloud forests: a preliminary vicariance model applying parsimony analysis of endemicity to vascular plant taxa. *Journal of Biogeography* **26:** 1299–1305.

Luna-Vega I, Almeida-Leñero L, Llorente-Bousquets J. 1989. Florística y aspectos fitogeográficos del bosque mesófilo de montaña de las cañadas de Ocuilan, estados de Morelos y México. *Anales del Instituto de Biología, Universidad Nacional Autónoma de México, Serie Botánica* **59:** 63–87.

Martínez-Ramírez E, Doadrio-Villarejo I, De Sosota-Fernández A. 2004. Peces continentales. In: García-Mendoza AJ, Ordóñez MJ, Briones-Salas M, eds. *Biodiversidad de Oaxaca*. México: Instituto de Biología-UNAM/Fondo Oaxaqueño para la Conservación de la Naturaleza/World Wildlife Fund, 357–373.

Massatti R, Knowles LL. 2016. Contrasting support for alternative models of genomic variation based on microhabitat preference: species-specific effects of climate change in alpine sedges. *Molecular Ecology* **25:** 3974–3986.

Mayr E. 1963. *Animal species and evolution*. Harvard University Press.

McCormack JE, Peterson AT, Bonaccorso E, Smith TB. 2008. Speciation in the highlands of Mexico: genetic and phenotypic divergence in the Mexican jay (*Aphelocoma ultramarina*). *Molecular Ecology* **17:** 2505–2521.

McNeely JA, Miller KR, Reid WV, Mittermeier RA, Werner TB. 1990. *Conserving the world's biological diversity*. Washington, DC: World Conservation Union, World Resources Institute, Conservation International, World Wildlife Fund–US and the World Bank.

Morán-Zenteno DJ, Tolson G, Martınez-Serrano RG, Martiny B, Schaaf P, Silva-Romo G, Macıas-Romo C, Alba-Aldave L, Hernández-Bernal MS, Solıs-Pichardo GN. 1999. Tertiary arc-magmatism of the Sierra Madre del Sur, Mexico, and its transition to the volcanic activity of the Trans-Mexican Volcanic Belt. *Journal of South American Earth Sciences* **12:** 513–535.

Moritz C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology* **51:** 238–254.

Morrone JJ, Crisci JV. 1995. Historical biogeography: introduction to methods. *Annual Review of Ecology and Systematics* **26:** 373–401.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403:** 853–858.

Navarro-Sigüenza AG, Peterson AT, Nyari A, García-Deras GM, García-Moreno J. 2008. Phylogeography of the *Buarremon* brush-finch complex (Aves, Emberizidae) in Mesoamerica. *Molecular Phylogenetics and Evolution* **47:** 21–35.

Nieto-Samaniego AF, Alaniz-Álvarez SA, Silva-Romo G, Eguiza-Castro MH, Mendoza-Rosales CC. 2006. Latest Cretaceous to Miocene deformation events in the eastern Sierra Madre del Sur, Mexico, inferred from the geometry

and age of major structures. *Geological Society of America Bulletin* **118:** 238–252.

Ornelas JF, Sosa V, Soltis DE, Daza JM, Gonzalez C, Soltis PS, Gutierrez-Rodriguez C, De los Monteros AE, Castoe TA, Bell C, Ruiz-Sanchez E. 2013. Comparative phylogeographic analyses illustrate the complex evolutionary history of threatened cloud forests of northern Mesoamerica. *PLoS One* **8:** e56283.

Parra-Olea G, Flores-Villela O, Mendoza-Almeralla C. 2014. Biodiversidad de anfibios en México. *Revista Mexicana de Biodiversidad* **85:** 460–6.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for *de novo* SNP recovery and genotyping in model and non-model species. *PLoS One* **7:** e37135.

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

Phillips SJ, Dudík M, Schapire RE. 2017. *Maxent software for modeling species niches and distributions, Version 3.4.1*. Available at: http://biodiversityinformatics.amnh.org/open_source/maxent/

Pineda E, Lobo JM. 2009. Assessing the accuracy of species distribution models to predict amphibian species richness patterns. *Journal of Animal Ecology* **78:** 182–190.

Pyron RA, Wiens JJ. 2011. A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Molecular Phylogenetics and Evolution* **61:** 543–583.

R Development Core Team 2016. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, Available at: http://www.R-project.org

Rambaut A, Drummond AJ. 2007. *Tracer v.1.5.0*. Available at: http://beast.bio.ed.ac.uk/Tracer

Rambaut A, Drummond AJ. 2011. *LogCombiner v.2.2.1*. Available at: http://beast.bio.ed.ac.uk/

Rambaut A, Drummond AJ. 2012. *TreeAnnotator v.2.2.1*. Available at: http://beast.bio.ed.ac.uk/

Reed DH, Frankham R. 2003. Correlation between fitness and genetic diversity. *Conservation Biology* **17:** 230–237.

Riddle BR, Hafner DJ. 2006. A step-wise approach to integrating phylogeographic and phylogenetic biogeographic perspectives on the history of a core North American warm deserts biota. *Journal of Arid Environments* **66:** 435–461.

Rosas-Elguera J, Alva-Valdivia LM, Goguitchaichvili A, Urrutia-Fucugauchi J, Ortega-Rivera MA, Prieto JCS, Lee JK. 2003. Counterclockwise rotation of the Michoacan Block: implications for the tectonics of western Mexico. *International Geology Review* **45:** 814–826.

Ruiz-Sanchez E, Specht CD. 2014. Ecological speciation in *Nolina parviflora* (Asparagaceae): lacking spatial connectivity along of the Trans-Mexican Volcanic Belt. *PLoS One* **9:** e98754.

Rundle HD, Nosil P. 2005. Ecological speciation. *Ecology Letters* **8:** 336–352.

Schenk JJ. 2016. Consequences of secondary calibrations on divergence time estimates. *PLoS One* **11:** e0148228.

Schluter D. 2000. *The ecology of adaptive radiation*. Oxford: Oxford University Press.

Schoener TW. 1968. The Anolis lizards of Bimini: resource partitioning in a complex fauna. *Ecology* **49:** 704–726.

Smith SA, Nieto-Montes de Oca A, Reeder TW, Wiens JJ. 2007. A phylogenetic perspective on elevational species richness patterns in Middle American treefrogs: why so few species in lowland tropical rainforests? *Evolution* **61:** 1188–1207.

Sullivan J, Markert JA, Kilpatrick CW. 1997. Phylogeography and molecular systematics of the *Peromyscus aztecus* species group (Rodentia: Muridae) inferred using parsimony and likelihood. *Systematic Biology* **46:** 426–440.

Swets JA. 1988. Measuring the accuracy of diagnostic systems. *Science* **240:** 1285–1293.

Toledo-Aceves T, Meave JA, González-Espinosa M, Ramírez-Marcial N. 2011. Tropical montane cloud forests: current threats and opportunities for their conservation and sustainable management in Mexico. *Journal of Environmental Management* **92:** 974–981.

Ustach PC, Mendelson JR, McDiarmid RW, Campbell JA. 2000. A new species of Hyla (Anura: Hylidae) from the Sierra Mixes, Oaxaca, Mexico, with comments on ontogenetic variation in the tadpoles. *Herpetologica* **56:** 239–250.

Vandergast AG, Bohonak AJ, Hathaway SA, Boys J, Fisher RN. 2008. Are hotspots of evolutionary potential adequately protected in southern California? *Biological Conservation* **141:** 1648–1664.

Warren DL, Glor RE, Turelli M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* **62:** 2868–2883.

Warren DL, Glor RE, Turelli M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33:** 607–611.

Weir JT. 2006. Divergent timing and patterns of species accumulation in lowland and highland neotropical birds. *Evolution* **60:** 842–855.

Zarza E, Connors EM, Maley JM, Tsai WL, Heimes P, Kaplan M, McCormack JE. 2017. Bridging multilocus species delimitation and DNA barcoding through target enrichment of UCEs: A case study with Mexican highland frogs. *BioRxiv*. doi:10.1101/153601.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Optimal threshold for loci assemblage was estimated according to our data and following Ilut *et al.* (2014). This method repeats the clustering process, incrementing sequentially the mismatches allowed between reads within a cluster (from 0.82 to 0.98 every 0.02). It calculates the proportion of clusters for homozygous and heterozygous (2 and > 2) and plots the results. We selected 0.94 as the optimum, the point where there is an asymptote for all the haplotypes.
Appendix S2. Complete sampling for the niche modelling before removing duplicated data points. Includes 224 individuals (58 from the present study and 164 from GBIF).
Appendix S3. Assignment of GBIF data points to their corresponding lineage polygon.
Appendix S4. Bioclimatic variables and their contributions to each distribution model. Numbers in bold indicate the variable that contributed the most to each lineage. Sb1a = *S. bistincta* 1a; Sb1b = *S. bistincta* 1b; Sb2 = *S. bistincta* 2; Sb3a = *S. bistincta* 3a; Sb3b = *S. bistincta* 3b; Sca = *S. calthula*; Spe = *S. pentheter.* Hyphens correspond to variables that did not contribute to each distribution model.
Appendix S5. Adegenet analysis and phylogenetic inference with different amounts of missing data; comparison among 10, 25 and 50%.
Appendix S6. Species tree estimated with SNAPP. Numbers on nodes are posterior probability values. A, 50% missing data. B, 25% missing data. C, 10% missing data.
Appendix S7. Background similarity test distributions between populations. Schoener's *D* (pink) and *I* statistics (blue).
Appendix S8. Ecological niche model per population. Latitudes and longitudes are included in map G as a reference for the rest of the maps.
Appendix S9. Number of lineages distributed in each conservation priority area: minimum (blue), mean (green) and maximum (red). *N* = number of areas per priority level.