Molecular Phylogenetics and Evolution 132 (2019) 284–295

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Complex patterns of hybridization and introgression across evolutionary timescales in Mexican whiptail lizards (*Aspidoscelis*)

Anthony J. Barley[a], Adrián Nieto-Montes de Oca[b,*], Tod W. Reeder[c], Norma L. Manríquez-Morán[d], José Carlos Arenas Monroy[b], Oswaldo Hernández-Gallegos[e], Robert C. Thomson[a]

[a] Department of Biology, University of Hawai'i, Honolulu, HI 96822, USA
[b] Laboratorio de Herpetología and Museo de Zoología Alfonso L. Herrera, Departamento de Biología Evolutiva, Facultad de Ciencias, Universidad Nacional Autónoma de México, Cd. Universitaria, Del. Coyoacán, Ciudad de México C.P. 04510, Mexico
[c] Department of Biology, San Diego State University, San Diego, CA 92182, USA
[d] Laboratorio de Sistemática Molecular, Centro de Investigaciones Biológicas, Universidad Autónoma del Estado de Hidalgo, Col. Carboneras, Mineral de la Reforma, Hidalgo, Mexico
[e] Laboratorio de Herpetología, Facultad de Ciencias, Universidad Autónoma del Estado de México, Toluca, Estado de México, Mexico

ARTICLE INFO

Keywords:
Admixture
Demographic modeling
D-statistics
RADseq
Parthenogenesis

ABSTRACT

Identifying patterns of introgression across the tree of life is foundational to understanding general mechanisms that govern the impacts of gene flow on the speciation process. There are few vertebrate groups in which hybridization is associated with as large a diversity of outcomes as in North American whiptail lizards (*Aspidoscelis*). Of particular interest is that hybridization among divergent whiptail species has repeatedly led to the formation of unisexual (parthenogenetic) lineages. Understanding the hybrid origin of these unisexual lineages requires an accurate understanding of species boundaries among gonochoristic whiptails. Doing so has historically been an extremely challenging problem which, in part, may be a consequence of widespread hybridization and incomplete reproductive isolation among lineages. The lack of a robust phylogenetic framework and uncertainty in species boundaries precludes studies of general patterns and mechanisms of introgression among whiptail species. Here, we use genomic data to reconstruct a robust estimate of evolutionary history in the largest clade of whiptail lizards (*A. sexlineatus* species group) and use it to identify patterns of introgression. Our results indicate substantial introgressive hybridization and admixture has occurred among multiple lineages of whiptails across diverse evolutionary time scales, and illustrate their impact on phylogenetic inference. Thus, hybridization among whiptail species appears to have been a prominent feature throughout their evolutionary history, which could, in part, explain why parthenogenesis has evolved so many times in whiptails in comparison to other vertebrate groups.

1. Introduction

The prominence of interspecific gene flow across the tree of life is increasingly being recognized (Nosil, 2008; Nosil and Feder, 2012; Abbott et al., 2013). However, scientists are only beginning to understand the impact of gene flow on the process of speciation in natural, empirical systems, and through evolutionary time. On one hand, gene flow has traditionally been considered a homogenizing force that counteracts species differentiation (Coyne and Orr, 2004). The historical basis for this perspective lies in foundational studies of speciation (e.g., Mayr, 1942), which led to the wide acceptance of the biological species concept. The view of gene flow as a more complex, unpredictable, and even potentially creative force during speciation has more recently become appreciated (Barton, 2001; Abbott et al., 2016; Elgvin et al., 2017). Studies that quantify gene flow among species in a broader diversity of empirical systems should allow scientists to understand general mechanisms that govern these patterns and make predictions about how gene flow will affect the evolutionary trajectory of species.

While the prevalence of gene flow among species in nature creates opportunities to study the process of evolutionary diversification, it also presents particular challenges to the field of systematic biology that have received increasing attention. For example, evolutionary biologists have begun to develop methods that account for gene flow in

* Corresponding author.
E-mail address: anietomontesdeoca@me.com (A. Nieto-Montes de Oca).

https://doi.org/10.1016/j.ympev.2018.12.016
Received 9 August 2018; Received in revised form 11 December 2018; Accepted 12 December 2018
Available online 15 December 2018

phylogenetic analyses, and to conceptualize the history of some groups of species as more closely resembling a phylogenetic network than a tree (Huson et al., 2010; Pease et al., 2016; Solís-Lemus and Ané, 2016; Wen et al., 2016). Gene flow among evolutionary lineages also complicates the recognition of species boundaries. From a methodological perspective, researchers increasingly rely on the use of phylogenetic trees and coalescent models for delimiting species, which frequently assume that no interspecific gene flow occurs. From a theoretical perspective, the conceptualization of species as general metapopulation lineages and the process of speciation as a continuum in evolutionary time (de Queiroz, 1998) is more compatible with gene flow during speciation than a more strict, categorical view of species using criteria such as reproductive isolation (e.g., Mayr, 1942). This conceptual synthesis initially helped resolve debates about species concepts and has more recently guided efforts to develop new methods that better account for some of the complexity associated with the speciation process (and that hold promise for identifying species boundaries in cases where gene flow complicates their recognition [e.g., Jackson et al., 2017]).

Whiptail lizards (=whiptails) of the genus *Aspidoscelis* are one of the most conspicuous and abundant groups of lizards in North America. They have long been of interest to biologists because they are one of the few known clades of vertebrates that include unisexual (parthenogenetic) species (Wright and Vitt, 1993). These unisexual lineages are derived from hybridization events between divergent, gonochoristic species (those composed of two sexes), when parthenogenetic reproduction is initiated in the hybrid offspring. Identifying species boundaries among gonochoristic whiptails has historically been considered one of the most challenging problems in herpetological systematics due to the morphological conservatism among whiptails as a whole, large phenotypic variation within and among populations, and phenotypic intergradation between many adjacent populations. An early discussion of this issue was provided by Gadow (1906) in his comprehensive treatment of Mexican whiptails where he remarked, "Most of the 'species' are so plastic, so variable, that they may well drive the systematist to despair. No two taxonomic authorities will, or can, possibly agree upon the number of admissible species."

It is known that some whiptail species hybridize where they come into contact (Dessauer et al., 2000; Manning et al., 2005), and this might, in part, have contributed to difficulties in identifying species boundaries. Whiptails are also one of the most diverse groups of lizards in North America, and thus, are an excellent system for understanding the impact of gene flow on diversification. A necessary first step, however, is to develop a phylogenetic framework based on a large sampling of genes and individuals to identify the major evolutionary lineages of whiptails, determine the extent to which introgression occurs among these lineages, and on what time scales it has occurred. Insights into speciation from novel systems across the tree of life will undoubtedly improve our general understanding of the process.

Here, we develop an evolutionary framework and quantify patterns of diversification and gene flow across the largest clade of whiptails: the *A. sexlineatus* species group. The taxonomic history of this group is chaotic. However, it currently consists of ~13 recognized gonochoristic species that are distributed from the southern U.S. south to Honduras and El Salvador (with the highest diversity occurring in Mexico; Reeder et al., 2002). Eight of the currently described species are polytypic and encompass multiple described subspecies or morphotypes, for a total of > 40 recognized subspecies (Table 1). At least one of the parental lineages of all of the known unisexual whiptail species are thought to be members of this species group. For six of the diploid and two of the triploid described unisexual taxa, all the parental lineages are thought to be members of the *A. sexlineatus* group. Thus, it is central to understanding the origin of parthenogenesis in whiptails. The most recent, comprehensive taxonomic treatment of the *A. sexlineatus* group, and an important landmark in the systematics of Mexican whiptails, was published by Duellman and Zweifel (1962). However, despite their comprehensive effort, many populations exhibited such a high degree of phenotypic variation that they precluded assignment to any described subspecies (e.g., populations on the Pacific Coast of Mexico related to *A. costatus*/*A. burti* and populations from the Mexican Plateau related to *A. gularis*). Additional research on the *A. sexlineatus* group has frequently resulted in the description of new species or subspecies (e.g., *A. alpinus*, a form thought to be related to *A. gularis* or *A. costatus*; Maslin and Walker, 1965), only to be followed by conflicting analyses or interpretations that call into question the validity of previous taxonomic decisions (e.g., Maslin and Secoy, 1986; Wright and Vitt, 1993;

Table 1
Summary of the taxonomic composition of the gonochoristic species in the *Aspidoscelis sexlineatus* species group, including descriptions of geographic ranges for each taxon. Here we primarily follow the taxonomy of Reeder et al. (2002) with some minor modifications. Several subspecies listed have been considered full species by some authors, but additional systematic work is needed to assess their distinctiveness.

Species	Range
A. a. angusticeps	Yucatán Peninsula in Mexico
A. a. petensis	Northern Guatemala, Belize
A. b. burti	Sonora (Mexico)
A. b. stictogrammus	Sonora (Mexico); Arizona and New Mexico (U.S.)
A. b. xanthonotus	Arizona (U.S.)
A. calidipes	Tepalcatepec and Balsas river basins (Michoacán, Mexico)
A. c. communis	Jalisco, Colima, Michoacán (Mexico)
A. c. mariarum	Tres Marías Islands (Nayarit, Mexico)
A. c. costatus	Upper Balsas Basin, Trans-Mexican Volcanic Belt (Puebla, Morelos, Tlaxcala; Mexico)
A. c. barrancorum	Barranca region of Chihuahua, Sonora (Mexico)
A. c. griseocephalus	Pacific Coast of Mexico in Sonora and Sinaloa
A. c. huico	Pacific Coast of Mexico in Sinaloa and Nayarit
A. c. mazatlanensis	Pacific Coast of Mexico near Mazatlán, Sinaloa
A. c. nigrigularis	Pacific Coast of Mexico in Sinaloa
A. c. occidentalis	Nayarit, Jalisco, Michoacán (Mexico)
A. c. zweifeli	Tepalcatepec and Balsas river basins in Michoacán, Guerrero (Mexico)
A. g. gularis	Southern U.S., Mexican Plateau
A. g. colossus	Querétaro, San Luis Potosí, Hidalgo, Mexico (Mexico)
A. g. pallidus	Near Cuatrociénegas, Coahuila (Mexico)
A. g. rauni	Nuevo León, Tamaulipas, San Luis Potosí, Aguascalientes, Jalisco (Mexico)
A. g. scalaris	Chihuahua, Coahuila, Durango (Mexico)
A. g. semiannulatus	Durango, Zacatecas, Aguascalientes, San Luis Potosí (Mexico)
A. g. semifasciatus	Coahuila, Nuevo León, Zacatecas, Durango (Mexico)
A. g. septemvittatus	Texas (U.S.); Chihuahua (Mexico)
A. i. inornatus	Nuevo León and Coahuila, (Mexico)
A. i. arizonae	Arizona (U.S.)
A. i. cienegae	Near Cuatrociénegas, Coahuila (Mexico)
A. i. chihuahuae	Chihuahua (Mexico)
A. i. gypsi	New Mexico (U.S.)
A. i. hepatogrammus	Chihuahua, Coahuila (Mexico); Texas, New Mexico (U.S.)
A. i. juniperus	New Mexico (U.S.)
A. i. llanuras	New Mexico (U.S.)
A. i. octolineatus	Nuevo León (Mexico)
A. i. pai	Arizona (U.S.)
A. i. paululus	Durango, Zacatecas, San Luis Potosí (Mexico)
A. labialis	Baja California (Mexico)
A. mexicanus	Semi-arid valleys of Central Oaxaca (Mexico)
A. motaguae	Central Oaxaca and Chiapas (Mexico); Guatemala, Honduras, El Salvador
A. parvisocius	Northern Oaxaca and southeastern Puebla (Mexico)
A. s. sackii	Oaxaca, Puebla (Mexico)
A. s. gigas	Puebla, Oaxaca, Guerrero, Morelos (Mexico)
A. s. sexlineatus	Central and Eastern U.S.
A. s. stephensae	Southern Texas (U.S.)
A. s. viridis	Texas, New Mexico, Oklahoma (U.S.)

Sullivan et al., 2014). Evolutionary relationships among species within the *A. sexlineatus* group are also largely unknown, as previous molecular phylogenetic studies have only included data for a small number of taxa (primarily those occurring in the U.S.).

Using the most comprehensive taxonomic, geographic and genetic sampling to date, we developed an evolutionary framework for the *A. sexlineatus* group and identified patterns of gene flow among lineages. Gene flow (i.e., the movement of genes between populations) can occur in a variety of evolutionary contexts. Here, we are primarily interested in gene flow between species or divergent populations that results from the process of hybridization (i.e., the interbreeding of individuals from two species or distinctive populations). Consistent with previous work, we use the term introgression (or introgressive hybridization) to refer to the incorporation of alleles from one evolutionarily divergent lineage into another (Harrison and Larson, 2014) via hybridization (in this case, 'divergent lineages' refer to non-sister taxa, and are largely synonymous with species, except where there is problematic taxonomy). We use admixture to refer to gene flow between closely related groups of populations, or lineages that have diverged much more recently (e.g., sister taxa from a phylogenetic perspective). Our analyses highlight many clear instances of introgressive hybridization between phenotypically and evolutionarily divergent taxa. Using phylogenomic and population genomic approaches we identify the major evolutionary lineages in the *A. sexlineatus* group and quantify divergence and gene flow among and within these lineages. By identifying how sources of discordance in our data arise from this extensive, but previously unknown, introgressive hybridization, our analyses highlight how whiptails can serve as a model system for studying hybridization and speciation in nature.

2. Methods

2.1. Data collection

We collected samples for all currently described gonochoristic species within the *A. sexlineatus* group. Because no extensive molecular studies of the group exist, we assembled two large datasets: one for the ND2 mitochondrial gene (n = 153 which included two sequences from GenBank–see Supporting Information for mitochondrial sequencing details) and one using reduced representation DNA sequencing (n = 188; RADseq). In total, 279 individuals were sampled for the research, with 62 being included in both the mitochondrial and RADseq datasets (Fig. 1; Supporting Information). DNA was extracted from each sample and quantified using a Qubit dsDNA BR assay kit (ThermoFisher, Waltham, MA). We sequenced these samples using a modified version of the original ddRADseq protocol (Peterson et al., 2012; see Supporting Information for details). The RADseq samples were sequenced in three separate batches. The first library was sequenced on a single lane of an Illumina HiSeq 2500 (Illumina, San Diego, CA) using a single end 100 bp protocol, whereas the second two libraries were each sequenced on half of a lane of an Illumina HiSeq 4000 with a single end 100 bp protocol.

2.2. Dataset assembly

We edited the mitochondrial DNA sequence data using Sequencher v4.1.4 (Gene Codes Corporation, Ann Arbor, MI) and aligned the sequences using MUSCLE (Edgar, 2004) in Geneious v7.1.5 (Kearse et al., 2012). We demultiplexed the RADseq data and assembled it *de novo* using pyRAD v3.0.66 (Eaton, 2014). We used default parameters in our analyses, except for the clustering threshold parameter which we selected following the Cluster Threshold Series approach of Ilut et al. (2014), the MaxSH parameter (see Nieto-Montes de Oca et al., 2017 for details), and using a minimum depth of coverage of 10x for genotype calls.

2.3. Phylogenetic analyses

We selected optimal partitioning strategies for the mitochondrial data and the RADseq data using PartitionFinder v2.1.1 (Lanfear et al., 2014). For the mitochondrial data, we selected among the potential partitioning strategies for each codon position and the adjacent tRNA sequence data using AICc, evaluating the 20 standard substitution models implemented in MrBayes. We selected the optimal partitioning strategy for the RAD loci using BIC and the 'rcluster' algorithm (allowing for each locus to potentially be grouped as a separate partition and selecting between the GTR and GTR + Γ models with the "RAxML" option). Each dataset was analyzed using Maximum Likelihood with RAxML v8.2.10 (Stamatakis, 2014), employing the selected partitioning scheme with the GTR + Γ substitution model, and the automatic bootstrapping option. The optimal partitioning scheme was also used to estimate phylogenetic relationships using Bayesian inference in MrBayes v3.2.6 (Ronquist et al., 2012). Analysis of the mitochondrial data consisted of two independent analyses, each with four chains, which were run for 10 million generations, sampling every thousand generations. Analyses of the RAD data were run for 20 million generations, sampling every 2,000 generations and consisted of two analyses with eight chains each. Convergence was assessed using the average standard deviation of split frequencies and potential scale reduction factor diagnostics in MrBayes, and by ensuring that all parameters had ESS values > 1000 using Tracer v1.6 (Rambaut et al., 2014). We also assessed topological convergence using the R package RWTY v1.0.1 (Warren et al., 2017), checking that each run had mixed well and converged on a stationary distribution, that independent runs were sampling from similar areas of tree space and that the posterior probabilities were correlated (correlation coefficient > 0.95), and that the topological ESS values were > 1000. We inferred the placement of the root of the phylogeny using a subset of the RADseq data in a partitioned MrBayes analysis that included an outgroup species (*A. deppii*), as well as with a strict-clock and no outgroup species in an unpartitioned phylogenetic analysis using BEAST v1.8.4, assessing convergence as above.

Previous studies have shown that hybridization occurs between some whiptail species and analyses of the mitochondrial data suggested that introgression is potentially common among whiptails when they come into contact. Therefore, we were concerned that reticulate evolution might be impacting our ability to accurately identify/discover lineages and reconstruct their general interrelationships in the phylogeny. To visualize uncertainty and phylogenetic conflict within our dataset that might result from reticulation, we used the NeighborNet algorithm in SplitsTree v4.14.4 (Huson, 1998) to examine genetic distances among individuals in our dataset. We also performed an additional Bayesian phylogenetic analysis on a subset of samples in the RAD dataset (after removing individuals that were suspected to have admixed ancestry based on the phylogenetic/population genetic analyses–see Results) to determine if this changed our estimate of evolutionary relationships, because introgression is not accounted for in standard phylogenetic models.

2.4. D-statistic tests for introgression

Preliminary phylogenetic analyses of the mitochondrial and RADseq data suggested that there was introgression occurring among multiple divergent lineages in the *A. sexlineatus* group. To identify the extent of genomic introgression among these lineages, we analyzed the RADseq data using D-statistics (Green et al., 2010; Eaton and Ree, 2013) to test several introgression hypotheses. D-statistic tests (also known as ABBA-BABA tests) evaluate frequencies of site patterns in alignments (among four taxa) and compare them to expectations under incomplete lineage sorting. Significantly skewed site patterns should indicate the extent of introgression between taxa. We assumed the RADseq topology from our phylogenetic analysis of the subset of individuals (see Results Fig. 5) in

Fig. 1. Map of Mexico showing location of samples genotyped using RADseq. Map is cropped to increase resolution of core study area (samples of *Aspidoscelis angusticeps* from the Yucatan and *A. gularis/A. sexlineatus* from Florida and Oklahoma not shown).

our D-statistic analyses. We tested for introgression between the following taxa: (1) *A. communis* and *A. costatus zweifeli*, (2) *A. motaguae* and *A. mexicanus*, (3) *A. mexicanus* and *A. sackii*, (4) *A. parvisocius* and *A. mexicanus*, *A. motaguae*, or *A. sackii*, (5) *A. costatus costatus* and *A. gularis* or *A. sackii,* and (6) *A. c. costatus* and several populations from northwestern Oaxaca. The first four hypotheses were generated from phylogenetic analyses of the mitochondrial data, the fifth was generated from the phylogenetic network analyses of the RAD data, and the sixth was generated from a combination of phylogenetic and population genetic analyses of the RAD data (see Results). All D-statistic analyses were run using ipyrad v0.7.15 and included all alignment sites with sufficient depth of coverage for genotype calls (see above) across each four taxon statement. For taxa that we had sampled many individuals, we only performed tests including the four individuals (per species) that had the least amount of missing data, because the number of possible combinations of tests become computationally prohibitive as the number of individuals increases.

2.5. Genetic diversity among closely related populations

Previous systematic studies of *Aspidoscelis* have identified extensive phenotypic variation among and within populations that is challenging to clearly partition into species level-diversity (Gadow, 1906; Duellman and Zweifel, 1962). Some of this uncertainty might be the result of lineage divergence followed by secondary contact (Wright and Vitt, 1993). An alternative is that this phenotypic variation has resulted from continuous gene flow among continuously distributed populations with few discrete boundaries separating them. After identifying the major genetic lineages within the *A. sexlineatus* group, we used a recently developed method (Bradburd et al., in press) to quantify continuous (isolation by distance) and discrete patterns of genetic structure within lineages that have also previously been identified as taxonomically problematic, polytypic species. Our goal here was to assess genetic diversity within each lineage that could guide future, more comprehensive studies aimed at conclusively resolving species boundaries. We performed the analyses using conStruct v0.0.0.9 for three different lineages identified in our phylogenetic analyses (see Results for details). Because this approach can be sensitive to missing data (Bradburd et al., in pres; a result we also noted in preliminary analyses of our dataset), the final datasets we used were filtered such that the samples and loci included in the analyses had < 15% missing data (and only used a single SNP per locus). We used the cross-validation approach to

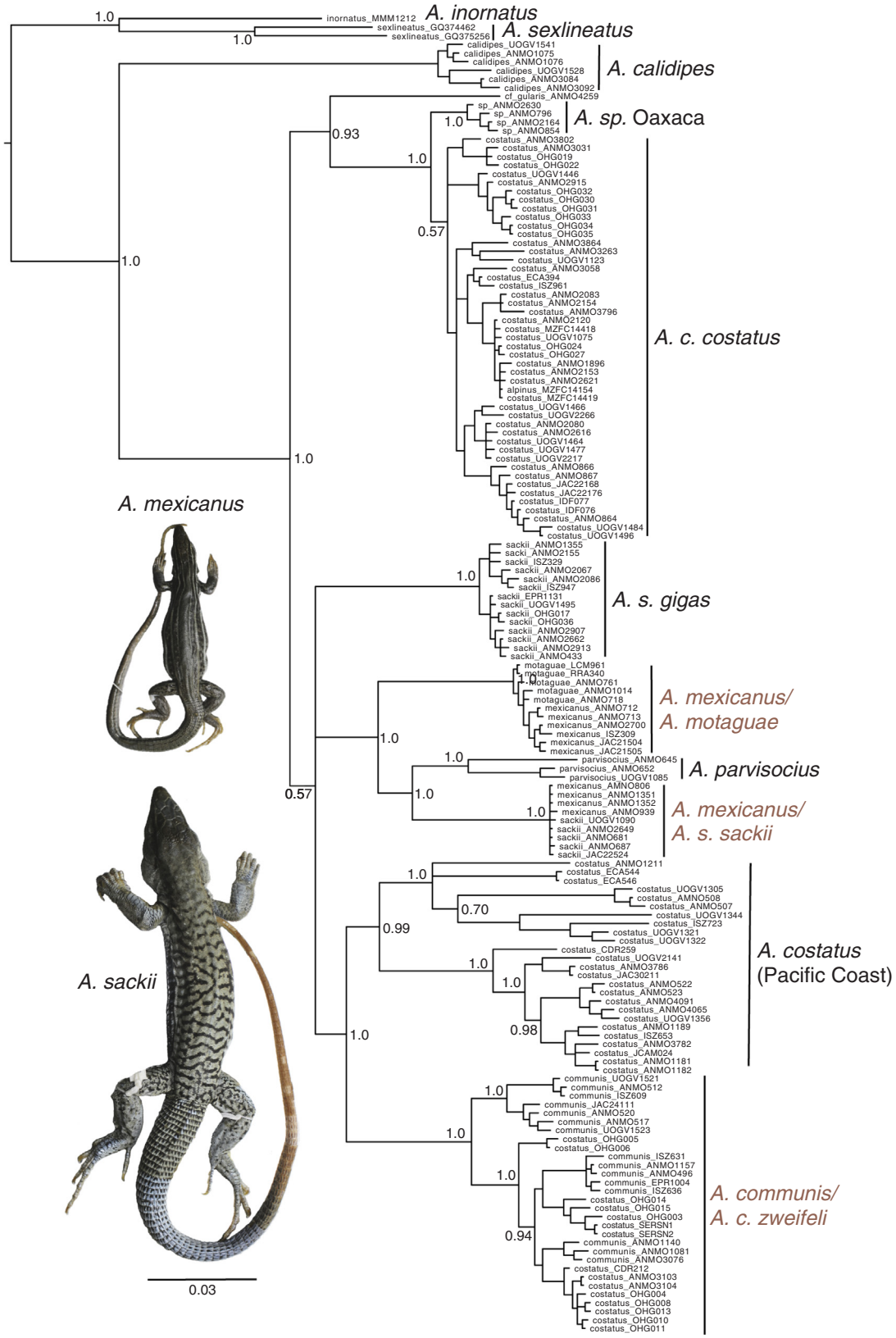

Fig. 2. Majority-rule consensus tree from Bayesian phylogenetic analysis of mitochondrial sequence data from the ND2 gene. Numbers at nodes indicate posterior probabilities. Clades with intermixed haplotypes of evolutionarily divergent species are highlighted in red. Exemplar specimens of phenotypically divergent species with introgressed haplotypes (*Aspidoscelis sackii* and *A. mexicanus*) shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

examine predictive accuracy across a broad range of values for the number of population groups ($K = 1–7$) with three replicates of both the spatial and non-spatial model. We ran all analyses for 10,000 iterations and visually checked for convergence using the trace plots as suggested in the documentation.

2.6. Demographic modelling

Finally, we sought to quantify divergence and gene flow between the groups of populations that we identified in the conStruct analyses that appear to be relatively recently diverged and may have had a complex admixture history. We did this by fitting coalescent models to the joint allele frequency spectrum in δaδi (Gutenkunst et al., 2009). Our motivation here was to identify the context and general time-scale on which admixture and divergence between these lineages has occurred. We selected among seven demographic models for each group using the Akaike information criterion that encompassed the range of scenarios we were interested in: (1) divergence with no gene flow, (2) divergence with constant, symmetrical gene flow, (3) divergence with constant, asymmetrical gene flow, (4) divergence with historical symmetrical migration, (5) divergence with historical asymmetrical migration, (6) divergence in isolation followed by secondary contact with symmetrical migration, (7) divergence in isolation followed by secondary contact with asymmetrical migration.

Our datasets consisted of a single SNP selected randomly from each locus, and in our analyses we assumed they were unlinked so that we could use the log-likelihood values as true likelihood values in our model comparisons. We analyzed the two-dimensional joint site frequency spectrum for each scenario and projected allele sample sizes down to account for missing data in our analyses by maximizing the number of segregating sites for each population (Gutenkunst et al., 2009). We performed a series of optimizations to identify the optimal parameter values and likelihood of each model. We performed initial optimizations by generating 50 sets of threefold randomly perturbed parameters, optimizing each using the Nelder-Mead method, and running each step for a maximum of 100 iterations. These optimized parameter sets were used to simulate the joint frequency spectrum for each model and estimate the likelihood of the spectrum given the model using "multinomial" optimization (Gutenkunst et al., 2009). We used the parameters from the replicate with the highest likelihood as starting values to run a second round of twofold perturbed parameter optimizations with 50 replicates. Finally, we used the optimal parameter values from this second round as starting parameters for a final onefold perturbed parameter optimization with 100 replicates to estimate the likelihood of the frequency spectrum given the model. We also generated 100 (nonparametric) bootstrap replicates and performed uncertainty analysis using the Godambe Information Matrix (Coffman et al., 2016) to obtain confidence intervals for the parameter estimates of each best-fit model for each dataset. Confidence intervals (CI) were calculated as: $CI = ML \pm 1.96 * sd$, where ML is the maximum likelihood parameter estimate and *sd* is the standard deviation from the uncertainty analysis.

3. Results

RAD sequencing resulted in a total of ~466 million sequence reads. Visual examination of sequence quality summaries in FastQC (Andrews, 2012) showed reads had generally high scores across the entire length of the sequences. Assembly of the data resulted in a total of ~60,000 loci, with coverage varying substantially across individuals and loci (15 individuals were removed from downstream analyses due to poor coverage). Because each of our analyses required datasets of different composition (i.e., in terms of individuals and total number of loci), we chose loci for each analysis that had the least amount of missing data. The composition of each dataset is detailed in the Results section for each analysis and the Supporting Information.

3.1. Phylogenetic analyses

Analysis of the mitochondrial DNA (1166 base pairs) resulted in a generally well-resolved phylogeny, however, many described species in our analysis were not monophyletic (Fig. 2). The reason for this appears to vary across taxa. In some cases, this is likely due to problematic/inaccurate taxonomy. Previous phylogenetic analyses (Reeder et al., 2002) found that some subspecies assigned to the taxon "*A. costatus*" are paraphyletic (a result confirmed in our RADseq data analyses). In others, it appears to be the result of mitochondrial introgression. For example, *A. mexicanus* and *A. sackii* are highly phenotypically distinct species (including having a nearly two-fold difference in adult snout-vent length), yet some individuals of each species have nearly identical ND2 haplotypes (Fig. 2). This scenario might be the result of introgression of *A. mexicanus* mitochondrial DNA into *A. sackii* individuals, as they have overlapping ranges in Oaxaca where these samples were collected. Additionally, samples of *A. sackii* from populations farther west in Oaxaca, Puebla, Guerrero, and Morelos form a distinct mitochondrial clade (the introgressed and non-introgressed mitochondrial clades appear to correspond to previously described subspecies of *A. sackii*–see Discussion below). It appears *A. mexicanus* also hybridizes with *A. motaguae* in Oaxaca where their ranges overlap, as the other *A. mexicanus* samples in our mitochondrial tree are nested within a clade of *A. motaguae* samples. Finally, mitochondrial haplotypes of *A. communis* and *A. costatus zweifeli* were extensively intermixed in our phylogenetic analysis, which could in theory be the result of mitochondrial introgression or incomplete lineage sorting between these more closely related lineages.

Phylogenetic analyses of the nuclear RADseq datasets were largely congruent, regardless of the number of taxa, loci, and amount of missing data that they were composed of (Fig. 3, Supporting Information). The same root for the phylogeny was inferred in both the MrBayes analysis with an outgroup, and the strict-clock analysis in BEAST, with *A. labialis* from Baja California being rather divergent from the rest of the *A. sexlineatus* group. The RADseq analyses also resolved much of the non-monophyly we observed in the mitochondrial tree, with *A. sackii*, *A. motaguae*, *A. communis*, and *A. c. zweifeli* all forming monophyletic groups (Fig. 3). The main exception to this was that the analyses strongly suggest the taxon *A. costatus* is not monophyletic and is composed of at least three distinct lineages. Additionally, we found several instances of paraphyly among more recently derived lineages which could reflect incomplete lineage sorting or gene flow between them (*A. costatus* populations on the Pacific Coast vs. *A. burti*, and *A. c. costatus* vs. *A. mexicanus* and several unassigned populations from northwestern Oaxaca). The placement of *A. parvisocius* in the RADseq tree also notably conflicts with its placement in the mtDNA tree, which we hypothesize might be the result of historical hybridization/introgression (see D-statistic results). Finally, the analyses suggested that *A. sexlineatus* is more closely related to several subspecies of *A. inornatus* from Mexico than several subspecies of *A. inornatus* from Arizona/New Mexico, however, it is worth noting that our sampling of this complex is highly incomplete.

In the phylogenetic network analysis of the RADseq data, most of the clades that were identified in the MrBayes analysis clustered into distinct groups, however, some individuals showed evidence of reticulation that could be the result of introgression (e.g., samples UOGV1477 & ANMO3864 of *A. costatus costatus* and UOGV1090 of *A. sackii*; Fig. 4). The phylogenetic analysis using a subset of the RADseq samples (removing individuals with evidence of admixed ancestry and larger amounts of missing data) were largely similar to the analyses of the full dataset, with several exceptions (Fig. 5). Most notably, is the placement of *A. calidipes* (see Discussion). Additionally (as might be expected), removing potentially admixed individuals cause individuals of the more recently diverged lineages to be recovered as monophyletic: (1) the unassigned populations from northwestern Oaxaca and *A. mexicanus* came out as sister to *A. c. costatus* rather than nested within

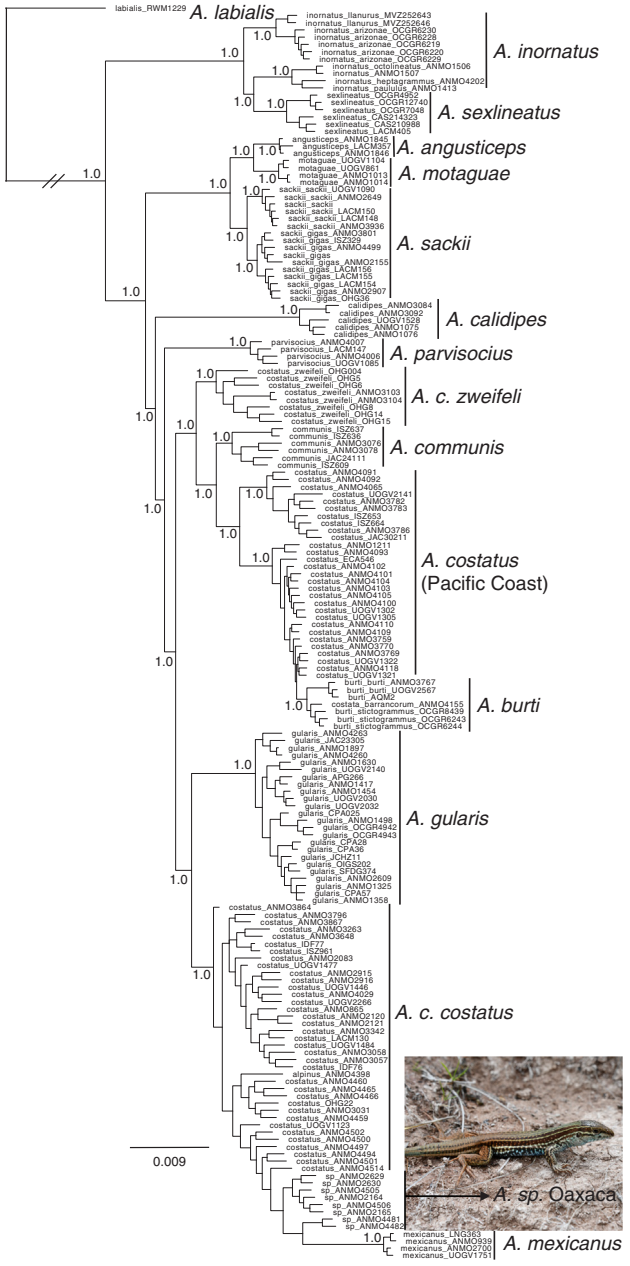

Fig. 3. Majority-rule consensus tree from partitioned, Bayesian phylogenetic analysis of the RADseq data. Numbers at nodes indicate posterior probabilities. Tree is based on an analysis of 183,107 bp for 173 individuals, from 1956 RAD loci.

it, and (2) *A. burti* samples were recovered as sister to one of the Pacific Coast lineages of *A. costatus*, rather than nested within it (see conStruct analyses below).

3.2. D-statistic tests for introgression

We formed several hypotheses about introgression which we examined using D-statistics calculated from the RADseq data (Table 2). These were primarily based on observed conflicts between the mitochondrial gene tree topology and specimen identifications based on phenotypic data. In a few cases, we performed tests to confirm a result suggested by other analyses of the RAD data, or to identify the most likely candidate for introgression from among several closely related species. These analyses suggested that the intermixing of mitochondrial haplotypes between *A. communis* and *A. c. zweifeli*, which occur sympatrically in Michoacán (Fig. 2), is the result of introgression (rather than incomplete lineage sorting). All of the D-statistic tests for this species pair had large test statistic values, which indicates that there has also been substantial nuclear introgression between them. We also found evidence of substantial nuclear introgression between *A. motaguae* and *A. mexicanus*, as well as between *A. sackii* and *A. mexicanus* (these taxa also had intermixed mitochondrial haplotypes; Fig. 2). Among *A. motaguae*, *A. sackii*, and *A. mexicanus* (whose mtDNA haplotypes were much more similar to those of *A. parvisocius* than would be expected based on the RADseq data; Figs. 2, 3), the best candidate for historical introgression based on D-statistics is *A. motaguae* (Table 2; though it is also possible that the majority of introgression occurred between *A. parvisocius* and the ancestor of the *A. sackii/A. angusticeps/A. motaguae* lineage). The placement of several samples of *A. c. costatus* in the phylogenetic network analyses of the RADseq data suggested there might be introgression between this species and an additional *A. sexlineatus* group lineage (such as *A. gularis* or *A. sackii*). The D-statistic analyses found strong evidence for introgression between *A. c. costatus* and *A. sackii*, but not *A. c. costatus* and *A. gularis* (which is congruent with the broadly overlapping distributions of the former pair). Finally, the D-statistic analyses supported the hypothesis that there has been introgression between the unassigned populations from northwestern Oaxaca and *A. c. costatus*.

3.3. Genetic diversity among closely related populations

Although the phylogenetic analyses of the RADseq data largely resolved evolutionary relationships among the major lineages within the *A. sexlineatus* group, uncertainty remains with respect to species boundaries within several clades that consist of closely related, phenotypically variable populations. We identified three lineages in our phylogenetic analyses within which we attempted to quantify the number of genetic groups using conStruct: (1) *A. sackii*, (2) the Pacific Coast lineage of *A. costatus/A. burti*, and (3) the Upper Balsas Basin/Trans-Mexican Volcanic Belt lineage of *A. c. costatus/A. mexicanus* (Fig. 3). Results of analyses using the spatial and non-spatial model were qualitatively similar. Fig. 6 shows the results of the spatial analyses for the preferred value of *K* for each lineage (results of the cross-validation analyses and layer contribution information can be found in the Supporting Information).

Analysis of the Pacific Coast lineage of *A. costatus/A. burti* dataset (34 individuals, 1153 SNPs) suggested that a model of three populations with limited amounts of shared ancestry between them was the best fit to the data (Fig. 6). This result is consistent with the phylogenetic analyses (Fig. 3). Populations allocated to *A. burti* form the northernmost population group. The southernmost group largely consists of populations that have been considered the subspecies *A. c. huico* and *A. c. occidentalis*, whereas the central group corresponds to populations that have been considered *A. c. griseocephalus*, *A. c. nigrigularis*, and *A. c. mazatlanensis* (Zweifel, 1959). The conStruct analyses indicated that samples from near Mazatlán have mixed ancestry from the southern and central populations. The analyses also suggest extensive shared ancestry between northern populations of *A. costatus* and *A. burti*, and the assignment of these populations has proven challenging historically (Duellman and Zweifel, 1962). Thus, it seems likely there is gene flow between each of these lineages where they come into contact.

We were initially uncertain about the assignment of some populations related to *A. c. costatus* and *A. mexicanus* from northwestern Oaxaca, and preliminary analyses conflicted in terms of which species they were most closely related to. Therefore, we included our *A. mexicanus* samples in the conStruct analyses of this entire clade (Fig. 6; 48 individuals, 2193 SNPs). The cross-validation analyses showed the best fit model for the dataset was K = 3. These population groups would correspond to (1) *A. mexicanus*, (2) the unassigned Oaxaca populations from near Santiago Yolomécatl, Villa Tejupam de La Unión, and Santo Domingo Yanhuitlán, and (3) populations of *A. c. costatus* from

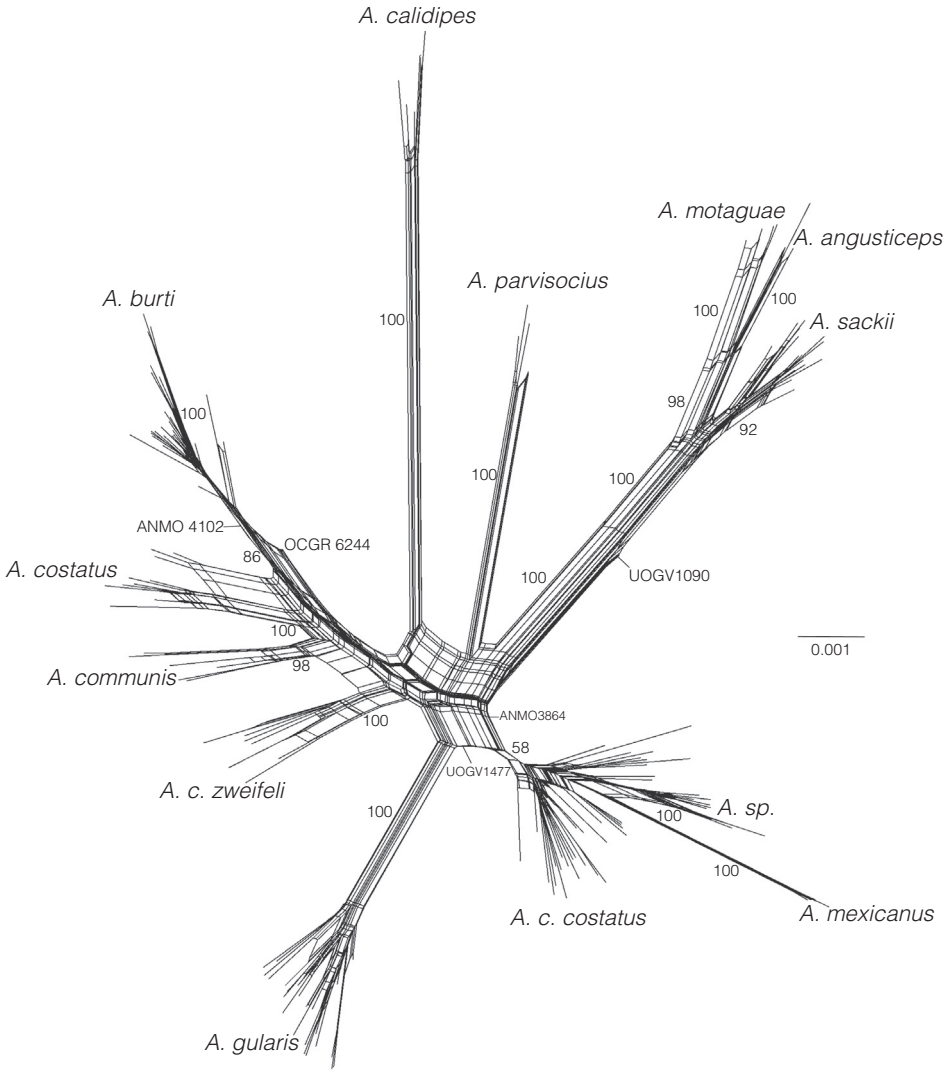

Fig. 4. Phylogenetic network from SplitsTree. Numbers indicate bootstrap proportions. Major lineages identified in MrBayes analysis are indicated, as are samples showing evidence of reticulation. Tree is based on an analysis of 183,107 bp for 173 individuals from 1956 RAD loci. Samples of *Aspidoscelis labialis*, *A. sexlineatus*, and *A. inornatus* were removed in order to increase resolution of focal species group (because they do not show evidence of reticulation and are highly divergent).

Tlaxcala, Morelos, Guerrero, Puebla, and northwestern Oaxaca. Populations of *A. c. costatus* from northwestern Oaxaca, central Puebla and eastern Guerrero have mixed ancestry. The conStruct analyses for *A. sackii* (16 samples, 1748 SNPs) suggested that two genetic clusters (corresponding to the two described subspecies within *A. sackii*) was the best fit to the data (with all samples having ancestry being predominantly derived from a single cluster; Fig. 6).

3.4. Demographic modeling

We performed the demographic analyses to elucidate the processes that may have led to the population structuring we identified using conStruct (Table 3). Coalescent analyses for the divergence between *A. burti* and the northern coastal lineage of *A. costatus* (19,499 SNPs, 24 individuals), between the northern and southern coastal lineages of *A. costatus* (19,618 SNPs, 27 individuals), between *A. c. costatus* and the northern Oaxaca populations (40,587 SNPs, 44 individuals) and between *A. s. sackii* and *A. s. gigas* (24,918 SNPs, 16 individuals), all strongly suggested that this genetic structuring is the result of initial divergence between these population groups (within the last several hundred thousand years) followed by secondary contact and symmetrical or asymmetrical gene flow (Table 3).

4. Discussion

Large genomic datasets have increasingly revealed discordance in phylogenetic history among genes, and that the process of evolutionary diversification is often more complex than a simple scenario of geographic isolation and subsequent speciation (Salichos and Rokas, 2013; Tigano and Friesen, 2016). Quantifying sources of phylogenetic discordance is a major focus in the field of systematic biology, and by identifying biological sources of discordance within evolutionary radiations, we stand to gain deeper insights into how the process of speciation occurs in nature. In developing an evolutionary framework for the largest clade of whiptails, we have shown that introgressive hybridization among species appears to be common in zones of geographic contact, even when lineages are phenotypically and evolutionarily divergent. We also found that admixture among more recently diverged groups of populations can impact phylogenetic inference, and may have complicated the resolution of species boundaries within whiptails. Although some of this complexity might be the result of secondary contact and hybridization following initial population divergence, work remains to identify how isolated these groups of populations are where they come into contact, and their likely evolutionary fate.

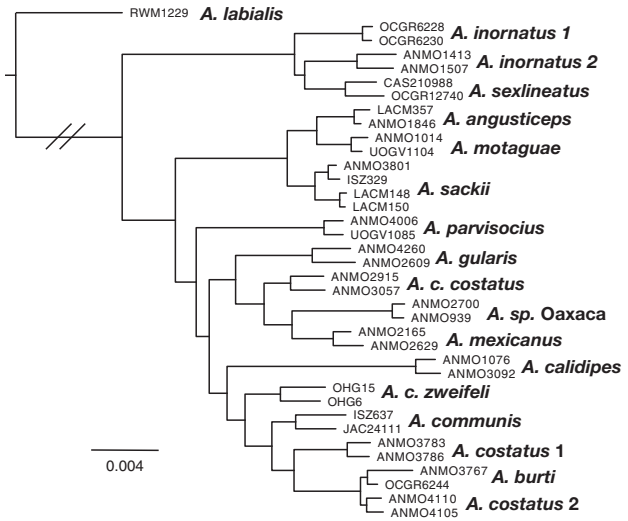

Fig. 5. Majority-rule consensus tree from partitioned, Bayesian phylogenetic analysis of a subsample of the RADseq data. All nodes have a posterior probability of 1.0. Tree is based on an analysis of 176,592 bp for 37 individuals from 1940 RAD loci.

4.1. Phylogenetics of the A. sexlineatus species group

Our phylogenetic analyses identify several new lineages within this species group, as well as confirm the status of many of the major lineages that were identified in previous morphological studies. The highest species diversity within the *A. sexlineatus* species group occurs in Oaxaca, and presumably the geographic complexity of the southern Sierra Madre del Sur region has been important in diversification in this group (see also Morrone, 2010). The complexity of the southern Sierra Madre del Sur region also appears to have shaped complex patterns of mitochondrial and nuclear introgression among species, as six different pairs of species appear to have hybridized in this region (Table 2; with some species showing evidence of introgression from multiple hybrid partners on different time scales). Populations or subspecies assigned to the taxon *A. costatus* correspond to three highly divergent lineages that are not monophyletic when considered in whole: (1) one distributed along the Pacific Coast of Northern Mexico, (2) one distributed across the Upper Balsas Basin and east-central portion of the Trans-Mexican Volcanic Belt (previously considered *A. c. costatus*), and (3) one in the Western Balsas Basin/Río Tepalcatepec Valley in Michoacán/Guerrero (previously considered *A. c. zweifeli*). Additionally, for the two former lineages, uncertainty remains with respect to diversity within them, and the nature of their relationship to other *Aspidoscelis* lineages (e.g., *A. burti*). Nodes corresponding to the earliest diverging lineages in the *A. sexlineatus* group have present day distributions in the north, with *A. labialis* in Baja California being highly divergent from the rest of the species complex (and warranting further study). *Aspidoscelis inornatus* and *A. sexlineatus* (which both occur in the southern U.S. and northern Mexico) appear to be closely related, and additional sampling is needed to resolve species boundaries in this species complex. Divergence of these lineages was followed by the divergence of lineages that ultimately have broad distributions in the south (*A. angusticeps*, *A. motaguae*, and *A. sackii*) and a narrowly distributed endemic species: *A. parvisocius*.

Finally, two sister lineages encompass species that are distributed (1) along the length of the Pacific Coast in northwestern Mexico and the Tepalcatepec River Basin (*A. burti*, *A. c. zweifeli*, *A. communis* and the Pacific lineages of *A. costatus*) and (2) the length of the Mexican Plateau and Trans-Mexican Volcanic Belt/Balsas River Basin (*A. gularis*, *A. mexicanus*, and *A. c. costatus*). Species from these two lineages likely have come into contact in various parts of Southern Mexico (Duellman and Zweifel, 1962), though we detected no signs of introgression between them. The placement of *A. calidipes* appears to be of some uncertainty. The phylogenetic analysis of the full RADseq dataset suggests it is highly divergent, being sister to *A. parvisocius* + the two lineages described above, whereas the analysis of a subset of the RAD data (removing individuals that show evidence of admixture or large amounts of missing data) suggests it is nested within the latter. Either placement seems plausible: morphologically it is most similar to *A. parvisocius* (Duellman, 1960), whereas its geographic range (between *A. communis* and *A. c. zweifeli*) suggest that its placement sister to a clade containing those lineages makes biogeographic sense. These alternative placements could be driven by taxon sampling differences, or genomic heterogeneity (not accounted for in either phylogenetic analysis, which assume a single evolutionary history across all genes in the alignment), for example, due to historical introgression. Future work employing model-based phylogenetic network approaches and analyses of individual gene trees will likely be valuable to generating new insights into the phylogenetic relationships among whiptails.

4.2. The biological nature of diverging whiptail lineages and species boundaries

The historically chaotic taxonomic history of whiptails appears to have been driven in part by complex dynamics of isolation and admixture among different groups of populations through evolutionary time. Despite being able to resolve species boundaries among some taxa within the *A. sexlineatus* group, several issues still remain, and determining which groups of populations should be recognized as full species will require a deeper biological understanding of the relationships among these lineages. Our population genetic analyses revealed distinct clusters of populations within each of the lineages we examined; however, our analyses also suggested that in many cases, individuals in these clusters are admixed, or share some genetic ancestry.

Table 2
Results of D-statistic tests of introgression hypotheses. Because we had multiple samples of each lineage, we performed multiple tests among the different combinations of samples. Table shows the range of Z-scores and D-statistics across all tests, the number of tests performed for each introgression hypothesis, and the average number of loci included across all tests. The D-statistic is the test statistic value based on the frequencies of site patterns in the alignment. Z-scores different from 0 indicate the presence of introgression between taxon P3 and taxon P2 (positive D-statistic) or taxon P1 (negative D-statistic). Lineages correspond to those indicated in Fig. 5 (which also shows the topology used in the analyses).

Test	P1	P2	P3	O	Z Range	D Range	*N* loci	*N* tests
A. c. zweifeli–A. communis	*A. costatus*	*A. communis*	*A. c. zweifeli*	*A. gularis*	(5.5, 19.5)	(0.25, 0.55)	5410	180
A. motaguae–A. mexicanus	*A. gularis*	*A. mexicanus*	*A. motaguae*	*A. inornatus*	(3.8, 11.3)	(0.34, 0.62)	2049	168
A. mexicanus–A. sackii	*A. gularis*	*A. mexicanus*	*A. sackii*	*A. inornatus*	(6.8, 10.4)	(0.26, 0.34)	5350	18
A. parvisocius–A. mexicanus	*A. c. costatus*	*A. mexicanus*	*A. parvisocius*	*A. inornatus*	(0.0, 1.4)	(−0.04, 0.07)	1410	54
A. parvisocius–A. motaguae	*A. gularis*	*A. parvisocius*	*A. motaguae*	*A. inornatus*	(3.8, 6.5)	(0.23, 0.34)	1627	45
A. parvisocius–A. sackii	*A. gularis*	*A. parvisocius*	*A. sackii*	*A. inornatus*	(0.29, 4.5)	(0.02, 0.23)	1663	45
A. c. costatus–A. gularis	*A. mexicanus*	*A. c. costatus*	*A. gularis*	*A. costatus*	(0.03, 1.5)	(−0.04, 0.01)	10,327	45
A. c. costatus–A. sackii	*A. gularis*	*A. c. costatus*	*A. sackii*	*A. inornatus*	(7.2, 15.8)	(0.26, 0.46)	4820	90
A. c. costatus–A. sp. Oax.	*A. mexicanus*	*A. sp.* Oax.	*A. c. costatus*	*A. gularis*	(4.4, 11.9)	(0.28, 0.37)	3537	90

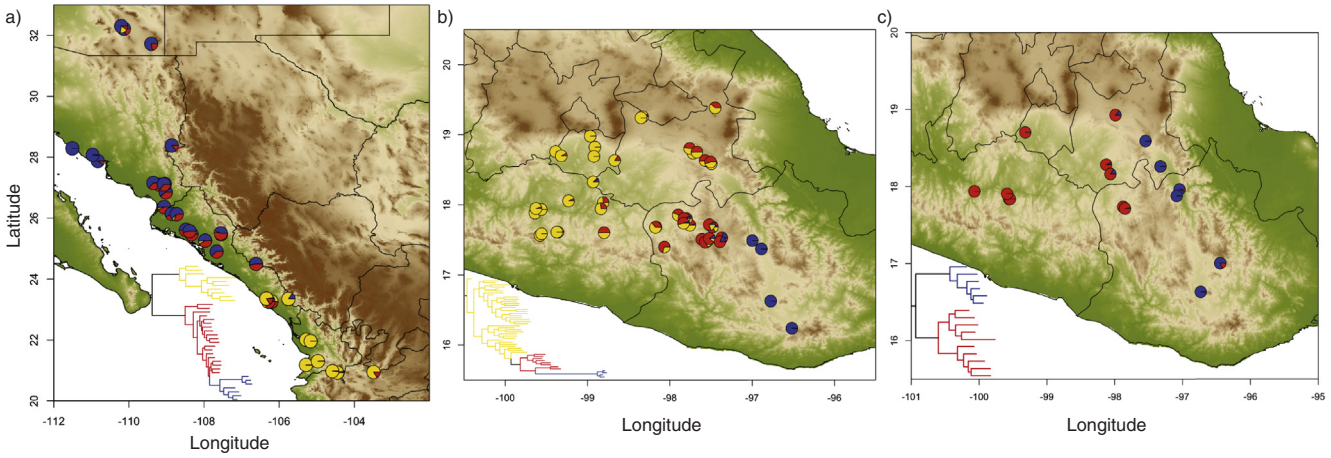

Fig. 6. Results from conStruct analyses using the optimal value of *K* for three different lineages within the *Aspidoscelis sexlineatus* group. (a) *A. burti* (blue) and *A. costatus* (yellow/red); (b) *A. c. costatus* (yellow), *A. mexicanus* (blue), and unassigned populations (red); (c) *A. sackii gigas* (red) and *A. s. sackii* (blue). Each pie diagram indicates the sampling location of an individual and the proportion of ancestry estimated to be derived from each population cluster. Inset phylogeny for each panel depicts the phylogenetic relationships among samples inferred from the RADseq dataset (Fig. 3), with tips colored according to the sample lineages/population clusters. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

This could theoretically result from gene flow following a relatively older initial divergence, or incomplete coalescence following a more recent divergence. Within the Pacific Coast lineage of *A. costatus/A. burti* we detected the presence of three lineages, which is somewhat consistent with previous taxonomic work based on morphology in that one corresponds to *A. burti*, and the contact zone for the two lineages within *A. costatus* (near Mazatalán) was previously shown to be an area of intergradation between the subspecies *A. c. mazatlanensis* and *A. c. huico* (Zweifel, 1959). Beyond that, however, analyses of the genetic data do not show support for the recognition of six subspecies within *A. costatus* (which were described by Zweifel [1959] based on dorsal color pattern and adult chin coloration). The analyses also suggest that populations assigned to *A. burti* share extensive ancestry with northern populations of *A. costatus* (in fact, one individual we sampled from the northernmost subspecies, *A. c. barrancorum*, appears to be more closely related to *A. burti*, than *A. costatus* (Fig. 3)). Duellman and Zweifel (1962) extensively discussed the morphological similarities between *A. burti* and *A. costatus*, particularly with regard to populations from

Table 3
Results of coalescent simulation analyses using δaδi. Bootstrapped confidence intervals of time and migration parameters for best fit model (and alternative models with similar AIC values) are shown. Parameter values were transformed into biologically meaningful values assuming a typical vertebrate mutation rate of 1×10^{-8} and an average generation time of 1.5 years. T1 refers to time between the initial population divergence and the present. T2 refers to the time between the present and initial secondary contact or the time migration stopped (for the ancestral migration model). Time is in units of years and migration rate is the fraction of individuals each generation in a population that are migrants.

Divergence	Model	Log-likelihood	AIC	ΔAIC	AIC ω	T1	T2	*m*12	*m*21
A. burti-costatus	No Mig.	−239.0	484.0	95.6	$1.1e^{-21}$	–	–	–	–
A. burti-costatus	Sym. Mig.	−200.7	409.5	21.1	$1.7e^{-5}$	–	–	–	–
A. burti-costatus	Asym. Mig.	−199.9	409.9	21.5	$1.4e^{-5}$	–	–	–	–
A. burti-costatus	Hist. Sym. Mig.	−239.0	488.1	99.7	$1.5e^{-22}$	–	–	–	–
A. burti-costatus	Hist. Asym. Mig.	−239.1	490.1	101.8	$5.2e^{-23}$	–	–	–	–
A. burti-costatus	Sec. Cont. Sym.	−189.8	389.6	1.2	0.354	150,031–187,208	4877–12,313	$5.0e^{-6}$–$9.2e^{-6}$	–
A. burti-costatus	Sec. Cont. Asym.	−188.2	388.4	–	0.646	152,009–183,740	554–4665	$2.1e^{-5}$–$2.3e^{-5}$	$7.1e^{-6}$–$1.3e^{-5}$
A. costatus N-S	No Mig.	−597.6	1201.2	198.5	$7.7e^{-44}$	–	–	–	–
A. costatus N-S	Sym. Mig.	−548.0	1104.0	101.3	$1.0e^{-22}$	–	–	–	–
A. costatus N-S	Asym. Mig.	−518.3	1046.7	44.0	$2.8e^{-10}$	–	–	–	–
A. costatus N-S	Hist. Sym. Mig.	−597.7	1205.3	202.7	$9.8e^{-45}$	–	–	–	–
A. costatus N-S	Hist. Asym. Mig.	−597.9	1207.8	205.2	$2.8e^{-45}$	–	–	–	–
A. costatus N-S	Sec. Cont. Sym.	−526.8	1063.5	60.9	$6.1e^{-14}$	–	–	–	–
A. costatus N-S	Sec. Cont. Asym.	−495.3	1002.7	–	1.00	245,653–273,284	20,677–33,759	0–$2.6e^{-5}$	0–$9.7e^{-6}$
A. c. costatus-Oax.	No Mig.	−474.0	954.0	470.2	$7.9e^{-103}$	–	–	–	–
A. c. costatus-Oax.	Sym. Mig.	−251.0	509.9	26.1	$2.1e^{-6}$	–	–	–	–
A. c. costatus-Oax.	Asym. Mig.	−255.9	521.8	38.0	$5.6e^{-9}$	–	–	–	–
A. c. costatus-Oax.	Hist. Sym. Mig.	−252.4	514.8	31.0	$1.9e^{-7}$	–	–	–	–
A. c. costatus-Oax.	Hist. Asym. Mig.	−250.0	512.0	28.2	$7.5e^{-7}$	–	–	–	–
A. c. costatus-Oax.	Sec. Cont. Sym.	−236.9	483.8	–	1.00	333,503–517,738	89,639–122,457	$6.0e^{-6}$–$7.8e^{-6}$	–
A. c. costatus-Oax.	Sec. Cont. Asym.	−245.6	503.2	19.4	$6.1e^{-5}$	–	–	–	–
A. sackii-gigas	No Mig.	−684.5	1375.0	638.5	$2.2e^{-139}$	–	–	–	–
A. sackii-gigas	Sym. Mig.	−456.2	920.3	183.8	$1.2e^{-40}$	–	–	–	–
A. sackii-gigas	Asym. Mig.	−442.8	895.6	159.1	$2.8e^{-35}$	–	–	–	–
A. sackii-gigas	Hist. Sym. Mig.	−524.0	1058.0	321.5	$1.5e^{-70}$	–	–	–	–
A. sackii-gigas	Hist. Asym. Mig.	−702.5	1417.0	680.5	$1.7e^{-148}$	–	–	–	–
A. sackii-gigas	Sec. Cont. Sym.	−375.3	760.5	24.0	$6.1e^{-6}$	–	–	–	–
A. sackii-gigas	Sec. Cont. Asym.	−361.7	735.5	–	1.00	157,378–358,914	0–56,428	0–$1.6e^{-5}$	0–$1.2e^{-5}$

central Sonora (for which we lack sampling) that were phenotypically intermediate between the two (but also highly phenotypically similar to *A. c. barrancorum*). Coalescent analyses suggest that the divergence between these groups of populations is likely the result of secondary contact following divergence, however, additional sampling is needed to understand the historical, and potentially contemporary interactions between these lineages.

Within the *A. c. costatus/A. mexicanus* species complex, our analyses also detected the presence of three lineages (corresponding to those two species plus distinctive populations in northwestern Oaxaca). The conflicting phylogenetic results regarding the relationships among these lineages could be explained by the Oaxacan populations diverging from *A. mexicanus* following the initial split of their ancestor with *A. c. costatus*, and subsequently experiencing secondary contact and introgression with *A. c. costatus* (with which they share genetic ancestry according to the conStruct analyses). This scenario is supported by D-statistic tests and demographic analyses, which suggest substantial nuclear introgression from the Oaxacan populations into *A. c. costatus* (Tables 2, 3). These distinctive populations from near Yanhuitlán Santo Domingo, as well as the populations with mixed ancestry from Puebla were also extremely challenging to identify for Duellman and Zweifel (1962) who first described their interesting morphology in terms of color pattern and scutelation. Clearly a re-examination of these traits in light of the genetic data is warranted. The phylogenetic analyses also indicate populations from Alchichica, Puebla (the type locality for *A. alpinus*) are closely related to other populations of *A. c. costatus* (with which previous workers synonymized them; Wright and Vitt, 1993).

In the most comprehensive previous taxonomic treatment of the *A. sexlineatus* group, Duellman and Zweifel (1962) split populations of Sack's Spotted Whiptail into two subspecies (*A. s. sackii* and *A. s. gigas*). They justified this based on differences in color pattern between the Oaxaca/E. Puebla populations (distinct dark crossbars on the dorsum; *A. s. sackii*) and the Guerrero/W. Puebla populations (reticulations and/or irregular spots on the dorsum; *A. s. gigas*), though there was a substantial gap in their sampling between those groups of populations. Interestingly, analyses of the mitochondrial data suggest that populations of *A. s. sackii* (but not *A. s. gigas*) have introgressed haplotypes from *A. mexicanus*, which makes biogeographic sense given the overlapping distribution of *A. s. sackii* and *A. mexicanus* in Oaxaca (Fig. 2). Phylogenetic analyses of the RADseq data found these two subspecies to be reciprocally monophyletic. The conStruct results suggested these two subspecies corresponded to distinct genetic clusters (with individuals sharing a very small amount of genetic ancestry). Our sampling helps clarify the biogeographic nature of the split between these two subspecies, with *A. s. gigas* occurring in the Balsas Basin in Puebla, Morelos, western Oaxaca, and Guerrero, and *A. s. sackii* occurring in the Tehuacán Valley in Puebla, south to the inland valleys of Central Oaxaca. The demographic analyses suggested the secondary contact model with little to no gene flow was the optimal divergence model.

Previous researchers as early as Gadow (1906) remarked on the value of whiptail lizards for studying both phenotypic variation and the process of speciation (which he suggested some lineages were in the process of undergoing). Further studies are needed to characterize phenotypic differences and the contact zones between many of these population groups. This would allow for several different potential scenarios to be distinguished: (1) if they correspond to narrow hybrid zones where gene flow is highly restricted, and potentially if reinforcement is contributing to additional reproductive isolation among these groups, or (2) if they are recent zones of contact where gene flow is largely unrestricted. As has been previously hypothesized (Wright and Vitt, 1993), our results suggest that contact zones may in some cases have resulted from secondary contact among previously isolated lineages (as opposed to divergence with migration due to ecological or geographic isolation among populations, for example). Studies of contact zones will likely lead to additional insights into how gene flow is shaping diversification in this unique group of lizards. Resolving these remaining issues will be important to future efforts to determine the ancestry of many of the unisexual whiptail species, as several of the *A. sexlineatus* lineages appear to have been involved in their formation (e.g., *A. costatus*, *A. gularis*, *A. inornatus*). Finally, nearly all species of whiptails include populations that have been documented to be phenotypically distinct (and have frequently been described as subspecies), and further sampling is needed to address the status of many of these groups (e.g., subspecies of *A. inornatus*, *A. gularis*, *A. burti*, *A. communis mariarum*, etc.).

4.3. Hybridization and whiptail evolution

Whiptails have long been of interest to evolutionary biologists because a significant proportion of species in *Aspidoscelis* (> 15) are parthenogenetic lineages that have been independently derived by hybridization between relatively divergent gonochoristic species. Analyses of genomic data suggest that hybridization among whiptails may be a recurrent process throughout their evolutionary history that also frequently leads to extensive introgression between distantly related lineages and admixture among diverging populations. In this case, none of the lineage pairings between which we detected evidence of introgression are known to produce parthenogenetic hybrid offspring. Thus, whiptails are an excellent study system for understanding the impact of gene flow on the evolutionary trajectory of species. For example, whiptails are a highly successful radiation of lizards: along with lizards of the genus *Sceloporus*, they are arguably the most conspicuous and abundant lizards found in most North American aridland ecosystems. Given that we know hybridization and admixture can sometimes facilitate evolution, it is possible that adaptive introgression has been important to the success of some lineages in new habitats (e.g., Kolbe et al., 2004; Lavergne and Molofsky, 2007). Several of the recently derived lineages appear to be directly associated with gene flow. Our analyses suggest that the undescribed lineage from northwestern Oaxaca that we identified here could plausibly have been derived from a hybrid background of *A. c. costatus* and *A. mexicanus*, while admixture appears to have been a prominent component of the history of *A. burti* and *A. costatus* on the northern Pacific Coast of Mexico. It appears that introgression from *A. mexicanus* into *A. sackii* may have primarily occurred into one of the phenotypically distinct subspecies (based on patterns of mitochondrial introgression). Our analyses also suggest that *A. c. zweifeli* and *A. communis* may have repeatedly hybridized following their divergence. Other species appear to have had multiple hybrid partners on recent time scales, including *A. mexicanus*, *A. motaguae*, and *A. sackii*. Understanding when and why hybridization facilitates diversification in different ways will require more in depth studies of interactions between species. Here we have developed a framework that can begin to help guide these efforts, however, it should be noted (given the prevalence of hybridization among whiptails) that extinction of one of the parental lineages could obscure signatures of past hybridization events.

5. Conclusions

We took advantage of extensive geographic sampling in Mexico, large genomic datasets, and recently developed statistical approaches to resolve several longstanding problems in the systematics of whiptails. Our results provide a framework for understanding evolutionary history in whiptails, suggest that many species have experienced introgressive hybridization across extensive evolutionary time, and demonstrate the ways in which RADseq data can help resolve complex evolutionary histories. The frequency of hybridization among gonochoristic whiptail species (as revealed by the results here) might shed some light on why clonal reproduction has evolved so many times in whiptails in comparison to other vertebrate groups. This should motivate future research investigating the demographic history among whiptail populations, and the mechanisms that contribute to reproductive isolation and

incompatibilities in this group.

Acknowledgments

We thank Chris Martin, Shana McDevitt, Jonathan Whitney, Jesse Weber, Amy Eggers, and Patrick Monnahan for advice and assistance with the RADseq protocol. Illumina sequencing was performed by the QB3 Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley. We used the University of Hawaii ITS HPC Cluster for performing many of the analyses in our study. Some of the analyses were also run using the CIPRES Science Gateway. Financial support was provided by a grant from DGAPA, UNAM (PAPIIT no. IN217818 to ANMO), a postdoctoral fellowship to AJB from the Arnold and Mabel Beckman Foundation, and grants from the National Science Foundation Division of Environmental Biology (DEB-1754350 and DEB-1754675). We thank Uri García-Vázquez, Levi Gray, and Israel Solano-Zavaleta for assistance with fieldwork and sample collection, and Sandra Elizalde-Rocha for her assistance in generating the mtDNA data. We thank Jonathan Campbell, Greg Pauly and the Los Angeles County Museum of Natural History, the Sam Noble Museum of Natural History, the Museum of Vertebrate Zoology, and the California Academy of Sciences for providing additional genetic samples for the research.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2018.12.016.

References

Abbott, R., et al., 2013. Hybridization and speciation. J. Evol. Biol. 26, 229–246.
Abbott, R.A., Barton, N.H., Good, J.M., 2016. Genomics of hybridization and its evolutionary consequences. Mol. Ecol. 25, 2325–2332.
Andrews, S., 2012. Fast QC Version 0.11.2. Available: < http://www.bioinformatics.babraham.ac.uk/projects/fastqc > .
Barton, N.H., 2001. The role of hybridization in evolution. Mol. Ecol. 10, 551–568.
Bradburd, G., Coop, G., Ralph, P.,. Inferring continuous and discrete population genetic structure across space. *Genetics* **210**, 2018, 33-52 .
Coffman, A.J., Hsieh, P., Gravel, S., Gutenkunst, R.N., 2016. Computationally efficient composite likelihood statistics for demographic inference. Mol. Biol. Evol. 33, 591–593.
Coyne, J.A., Orr, H.A., 2004. Speciation. Sinaur Associates, Inc., Sunderland, MA.
Dessauer, H.C., Cole, C.J., Townsend, C.R., 2000. Hybridization among western whiptail lizards (*Cnemidophorus tigris*) in southwestern New Mexico: population genetics, morphology, and ecology in three contact zones. Bull. Am. Mus. Nat. Hist. 246.
de Quieroz, K., 1998. The general lineage concept of species, species criteria, and the process of speciation. In: Howard, D.J., Berlocher, S.H. (Eds.), Endless Forms: Species and Speciation. Oxford University Press, New York, pp. 57–75.
Duellman, W.E., 1960. Variation, distribution, and ecology of the Mexican Teiid lizard *Cnemidophorus calidipes*. Copeia 1960, 97–101.
Duellman, W.E., Zweifel, R.G., 1962. A synopsis of the lizards of the sexlineatus species group (Genus *Cnemidophorus*). Bull. Am. Mus. Nat. Hist. 123.
Eaton, D.A.R., Ree, R.H., 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). Syst. Biol. 62, 689–706.
Eaton, D.A.R., 2014. PyRAD: assembly of *de novo* RADseq loci for phylogenetic analyses. Bioinformatics 30, 1844–1849.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and throughput. Nucleic Acids Res. 19, 1792–1797.
Elgvin, T.O., et al., 2017. The genomic mosaicism of hybrid speciation. Sci. Adv. 3, e1602996.
Gadow, H., 1906. A contribution to the study of evolution based upon the Mexican species of *Cnemidophorus*. J. Zool. 76, 277–375.
Green, R.E., et al., 2010. A draft sequence of the Neandertal genome. Science 328, 710–722.
Gutenkunst, R.N., Hernandez, R.D., Williamson, S.H., Bustamante, C.D., 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genet. 5, e1000695.
Harrison, R.G., Larson, E.L., 2014. Hybridization, introgression, and the nature of species boundaries. J. Hered. 105, 795–809.
Huson, D.H., 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14, 68–73.
Huson, D.H., Rupp, R., Scornavacca, C., 2010. Phylogenetic Networks: Concepts, Algorithms, and Applications. Cambridge University Press.
Ilut, D.C., Nydam, M.L., Hare, M.P., 2014. Defining loci in restriction-based reduced representation genomic data from non-model species: sources of bias and diagnostics for optimal clustering. BioMed. Res. Int. 2014, 675158.
Jackson, N.D., Carstens, B.C., Morales, A.E., O'Meara, B.C., 2017. Species delimitation with gene flow. Syst. Biol. 66, 799–812.
Kearse, M., et al., 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649.
Kolbe, J.J., Glor, R.E., Schettino, L.R., Lara, A.C., Larson, A., Losos, J.B., 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature 431, 177–181.
Lanfear, R., Calcott, B., Kainer, D., Mayer, C., Stamatakis, A., 2014. Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evol. Biol. 14, 82.
Lavergne, S., Molofsky, J., 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proc. Natl. Acad. Sci. 104, 3883–3888.
Manning, G.J., Cole, C.J., Dessauer, H.C., Walker, J.M., 2005. Hybridization between parthenogenetic lizards (*Aspidoscelis neomexicana*) and gonochoristic lizards (*Aspidoscelis sexlineata viridis*) in New Mexico: ecological, morphological, cytological, and molecular context. Am. Mus. Novit. 3492.
Maslin, T.P., Secoy, D.M., 1986. A checklist of the lizard genus *Cnemidophorus* (Teiidae). Univ. Color. Mus. Cont. Zool. 1, 1–60.
Maslin, T.P., Walker, J.M., 1965. *Cnemidophorus alpinus*: a new species of Teiid lizards from Puebla. Mexico. Univ. Color. Mus. Cont. Biol. 19, 1–8.
Mayr, E., 1942. Systematics and the Origin of Species from the Viewpoint of a Zoologist. Columbia University Press, New York, NY.
Morrone, J.J., 2010. Fundamental biogeographic patterns across the Mexican Transition Zone: an evolutionary approach. Ecography 33, 355–361.
Nieto-Montes de Oca, A., Barley, A.J., Meza-Lázaro, R.N., García-Vázquez, U.O., Zamora-Abrego, J.G., Thomson, R.C., Leaché, A.D., 2017. Phylogenomics and species delimitation in the knob-scaled lizards of the genus *Xenosaurus* (Squamata:Xenosauridae) using ddRADseq data reveal a substantial underestimation of diversity. Mol. Phylogenetics Evol. 106, 241–253.
Nosil, P., 2008. Speciation with gene flow could be common. Mol. Ecol. 17, 2103–2106.
Nosil, P., Feder, J.L., 2012. Genomic divergence during speciation: causes and consequences. Philos. Trans. R. Soc. 367, 332–342.
Pease, J.B., Haak, D.C., Hahn, M.W., Moyle, L.C., 2016. Phylogenomics reveals three sources of adaptive variation during a rapid radiation. PLoS Biol. 14, e1002379.
Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS ONE 7, e37135.
Rambaut, A., Suchard M.A., Xiew, W., Drummond, A.J., 2014. Tracer, version 1.6. Available at < http://beast.bio.ed.ac.uk/Tracer > .
Reeder, T.W., Cole, C.J., Dessauer, H.C., 2002. Phylogenetic relationships of whiptail lizards of the genus *Cnemidophorus* (Squamata: Teiidae): a test of monophyly, re-evaluation of karyotypic evolution, and review of hybrid origins. Am. Mus. Novit. 3365, 1–61.
Ronquist, F., Teslenko, M., van der Mark, P., Ayers, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Salichos, L., Rokas, A., 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature 497, 327–333.
Solís-Lemus, C., Ané, C., 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. PLoS Genet. 12, e1005896.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Sullivan, B.K., Douglas, M.R., Walker, J.M., Cordes, J.E., Davis, M.A., Anthonysamy, W.J.B., Sullivan, K.O., Douglas, M.E., 2014. Conservation and management of polytypic species: the little striped whiptail complex (*Aspidoscelis inornata*) as a case study. Copeia 2014, 519–529.
Tigano, A., Friesen, V.L., 2016. Genomics of local adaptation and gene flow. Mol. Ecol. 25, 2144–2164.
Warren, D.L., Geneva, A.J., Lanfear, R., 2017. RWTY (R We There Yet): an R package for examining convergence of Bayesian phylogenetic analyses. Mol. Biol. Evol. 34, 1016–1020.
Wen, D., Yu, Y., Hahn, M.W., Nakhleh, L., 2016. Reticulate evolutionary history and extensive introgression in mosquito species revealed by phylogenetic network analysis. Mol. Ecol. 25, 2361–2372.
Wright, J.W., Vitt, L.J., 1993. Biology of the Whiptail Lizards (Genus *Cnemidophorus*). Museum of Natural History, Oklahoma.
Zweifel, R.G., 1959. Variation in and distribution of lizards of western Mexico related to *Cnemidophorus sacki*. Bull. Am. Mus. Nat. Hist. 117, 2.