MOLECULAR ECOLOGY

Molecular Ecology (2014) **23**, 3273–3291 doi: 10.1111/mec.12814

Diversification and asymmetrical gene flow across time and space: lineage sorting and hybridization in polytypic barking frogs

JEFFREY W. STREICHER,*[1] THOMAS J. DEVITT,† CAREN S. GOLDBERG,‡[2] JOHN H. MALONE,§ HEATH BLACKMON* and MATTHEW K. FUJITA*

*Department of Biology, The University of Texas at Arlington, Arlington, TX 76010, USA, †Department of Integrative Biology, The University of Texas at Austin, Austin, TX 78712, USA, ‡Fish and Wildlife Sciences, University of Idaho, Moscow, ID 83844, USA, §Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269, USA

Abstract

Young species complexes that are widespread across ecologically disparate regions offer important insights into the process of speciation because of their relevance to how local adaptation and gene flow influence diversification. We used mitochondrial DNA and up to 28 152 genomewide single nucleotide polymorphisms from polytypic barking frogs (*Craugastor augusti* complex) to infer phylogenetic relationships and test for the signature of introgressive hybridization among diverging lineages. Our phylogenetic reconstructions suggest (i) a rapid Pliocene–Pleistocene radiation that produced at least nine distinct lineages and (ii) that geographic features of the arid Central Mexican Plateau contributed to two independent northward expansions. Despite clear lineage differentiation (many private alleles and high between-lineage F_{ST} scores), *D*-statistic tests, which differentiate introgression from ancestral polymorphism, allowed us to identify two putative instances of reticulate gene flow. Partitioned *D*-statistics provided evidence that these events occurred in the same direction between clades but at different points in time. After correcting for geographic distance, we found that lineages involved in hybrid gene flow interactions had higher levels of genetic variation than independently evolving lineages. These findings suggest that the nature of hybrid compatibility can be conserved overlong periods of evolutionary time and that hybridization between diverging lineages may contribute to standing levels of genetic variation.

Keywords: ABBA-BABA tests, *Eleutherodactylus*, phylogeography, RADseq, SNAPP, species trees

Received 27 November 2013; revision received 20 May 2014; accepted 23 May 2014

Introduction

Gene flow can act as a creative force by introducing novel genetic variation into populations or as a homogenizing force by decreasing genetic differences between populations (Slatkin 1985). For closely related organisms, the extent to which gene flow affects genetic variation among populations varies greatly owing to different natural histories, geographic distributions, selective regimes and mutational processes (Endler 1973; Slatkin 1985, 1987). The 'age of genomics' (sensu Walsh 2001) has enhanced our understanding of these influences on gene flow by providing evidence for the frequent occurrence of introgressive hybridization across multiple evolutionary tiers (e.g. Rheindt & Edwards 2011; Cui *et al.* 2013) and between diverging species (e.g. Martin *et al.* 2013). In plants, hybridization has been implicated as an important source of genetic variation that can increase the potential for adaptive responses to environmental change (Anderson 1949;

Correspondence: Jeffrey W. Streicher, Fax: 001 520 621 9190; E-mail: streicher@email.arizona.edu; Matthew K. Fujita, Fax: 001 817 272 2855; E-mail: mkfujita@uta.edu
[1]Present address: Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA
[2]Present address: School of the Environment, Washington State University, Pullman, WA 99164, USA

Stebbins 1950; Lavergne & Molofsky 2007). However, the potential for hybrid gene flow to increase genetic variation in animal species is less explored (Dowling & Secor 1997; Hedrick 2013). Unlike some plants, opportunities for introgressive gene flow in animals may be restricted to closely related species or young species complexes (sensu Shaffer & McKnight 1996; Mullen *et al.* 2008) because of genic incompatibilities that evolve as species diverge (Dobzhansky 1936; Muller 1939). If hybridization can occur, interspecific gene flow may directly introduce novel alleles (or new allelic combinations) that increase adaptive potential in the short term. However, if hybridizing lineages are closely related, it may also broaden existing variation in a neutral/nearly neutral manner that increases adaptive potential in the long term. Thus, an important step towards understanding the role hybrid gene flow has in animal evolution is to ask: Do diverging lineages that have exchanged genes in the past have more genetic diversity than those that evolved in relative isolation? In other words, does hybridization actually 'broaden' standing variation during the course of speciation?

In vertebrates, several studies support the tenet that hybrid gene flow can increase adaptive potential (e.g. lizards, Kolbe *et al.* 2004; fish, Lucek *et al.* 2010). However, to the best of our knowledge, these studies have all examined hybridization events that occurred in the recent past (<300 years ago). While the evolutionary implications of these short timescale studies are clear, identifying whether hybrid gene flow occurred during (and perhaps facilitated) the evolution of older clades (>1 million years old) is challenging for several reasons (also see review in Rheindt & Edwards 2011). First, characterizing gene flow dynamics across closely related species or populations involves differentiating introgression from incomplete lineage sorting (ILS; i.e. ancestral polymorphism; Maddison 1997). Second, analytical tools for hypothesizing the directionality (asymmetric vs. symmetric) and evolutionary timing (ancient vs. recent) of introgression were not available until recently (Eaton & Ree 2013). Finally, using empirical data sets to examine gene flow often requires an extensive knowledge of the study organism so that patterns of genetic diversity associated with natural history and demography can be accounted for (e.g. sex-biased dispersal).

How can these challenges be met? Differentiating introgression and ancestral polymorphism is not an insurmountable task, but it requires a large number of unlinked loci (Kronforst 2008; Yu *et al.* 2011) or substantial mitochondrial (mtDNA) introgression (McGuire *et al.* 2007; Sequeira *et al.* 2011). Increasingly affordable next-generation DNA sequencing methods such as restriction-site-associated DNA sequencing (RADseq; Davey & Blaxter 2011), which can identify thousands of single nucleotide polymorphisms (SNPs) in nonmodel organisms, enable high-resolution examination of lineage sorting dynamics (Davey *et al.* 2011). Perhaps one of the more appealing attributes of RADseq data sets is that they allow ancient gene flow to be distinguished from ancestral polymorphism via the 'ABBA-BABA' test (Green *et al.* 2010; Durand *et al.* 2011; Skoflund & Jakobsson 2011; Meyer *et al.* 2012; Heliconius Genome Consortium 2012). This test uses Patterson's *D*-statistic (Durand *et al.* 2011) to compare the frequencies of discordant SNP genealogies in a four-taxon pectinate tree. Assuming equal substitution rates among lineages and unlinked loci, asymmetrical distributions of genealogical patterns suggest that introgression has occurred. Eaton & Ree (2013) extended the original ABBA-BABA test to include a fifth taxon, developing the 'partitioned *D*-statistic' that can identify the source population of introgression events when more than one extant lineage may have hybridized with the in-group. Thus, RADseq data and *D*-statistic tests offer a robust analytical framework for differentiating introgression from ILS and identifying the timing and directionality of gene flow events.

What animal systems are most ideal to explore patterns of introgression? Two important criteria for selecting candidates are (i) groups with demonstrable interlineage gene flow that contain multiple statistically supported genomic lineages (i.e. biological replicates) and (ii) groups with natural history that do not greatly bias the spatial distribution of genetic markers. Polytypic amphibian complexes often posses these characteristics (e.g. Hillis 1988; Robertson & Zamudio 2009; O'Neill *et al.* 2013). Hybridization is common between amphibian species (Malone & Fontenot 2008; Smith *et al.* 2013) and can occur between distantly related taxa (up to 21 million years divergent; Prager & Wilson 1975). Genetic patterns consistent with hybridization are also observed within young polytypic amphibian complexes (Devitt *et al.* 2011). Furthermore, relative to other tetrapods amphibians display high levels of philopatry and limited dispersal capability (Beebee 1996), which is often reflected in clear phylogeographic structure (Stuart *et al.* 2006; Zeisset & Beebee 2008). Thus, we believe polytypic amphibians represent ideal study systems for investigating introgressive interactions during the early stages of speciation (Wang 2009; Smith *et al.* 2013).

Here, we used recently developed statistical tools on mitochondrial (mtDNA; 12S gene) and nuclear DNA (nucDNA; up to 28 152 SNPs) from a Pliocene–Pleistocene radiation of barking frogs (*Craugastor augusti* complex) to (i) infer phylogenetic relationships, (ii) identify the timing and directionality of introgression events across distinct lineages and (iii) test whether more genetic variation is observed in lineages that have experienced

introgression. After identifying putative introgression events, we examined levels of genetic variation in those lineages that experienced introgressive gene flow (recipients), those that contributed gene flow (donors), and those that appear to have evolved in relative isolation (loners) to ask whether recipient lineages possess higher levels of genetic variation than loner and/or donor lineages.

Materials and methods

Terminology

In sexually reproducing species, a semantic distinction is often made between genetic material that is exchanged (i) between divergent populations of the same species, termed 'gene flow' and (ii) between different species or subspecies, termed 'introgression' (Hedrick 2013). While this distinction is helpful for understanding many systems, for young species complexes, this terminology is less useful. Because diverging lineages within young species complexes often possess a spectrum of reproductive compatibilities, population densities, and geographic distributions, the terminological distinction between gene flow and introgression is often trivial. Thus, to describe genetic exchanges in barking frogs, we use the terms 'hybrid gene flow', 'gene flow between lineages', 'introgression', 'introgressive hybridization' and 'reticulate evolution/gene flow' interchangeably to describe any transfer of genetic material between distinct genomic clusters (sensu Cui *et al.* 2013; Rheindt *et al.* 2014).

Study system

The genus *Craugastor* (Anura: Craugastoridae) has diversified extensively in the last 30–40 million years (Crawford & Smith 2005; Heinicke *et al.* 2007). There are at least 115 species (Frost 2013) and all members of the group are direct-developing (i.e. they lack an aquatic larval phase). While most *Craugastor* diversity is confined to the New World tropics, a small number of species have geographic distributions that extend into subtropical and temperate biomes of North America. Almost all species with subtropical and temperate distributions belong to the *Craugastor augusti* Species Series (sensu Hedges *et al.* 2008). This group contains two species, *C. augusti* and *C. tarahumaraensis*. These species are endemic to North America (Zweifel 1956) with *C. tarahumaraensis* restricted to high elevation pine-oak forests of the Sierra Madre Occidental in Mexico (Fig. 1G) and *C. augusti* occurring across Texas, New Mexico, Arizona, and throughout Mexico (Fig. 1A–F, H–I). Goldberg *et al.* (2004) examined mitochondrial sequence divergence and male advertisement calls among populations of *C. augusti* and discovered that populations from the United States and one locality in Mexico possessed regionally circumscribed and divergent haplotypes consistent with a polytypic species complex. With the exception of populations from western Texas and eastern New Mexico (Fig. 1A), which inhabit rodent burrows in creosote bush flats, barking frogs are always found near cliffs, caves and limestone outcrops (Schwalbe & Goldberg 2005). This likely indicates a conserved microhabitat preference across the group. Barking frogs have a surprisingly rich fossil record for an amphibian (Holman 2003). Fossils have been found in Pleistocene deposits in all three of their northernmost populations including Eddy County, New Mexico, USA (Applegarth 1980); Bexar County Texas, USA (Mecham 1959); and Rancho la Brisca, Sonora, Mexico (Van Devender *et al.* 1985). All of these fossils originate from deposits that have been dated to range from 11 000 to 240 000 years ago (Holman 2003). This may indicate that barking frogs had acquired their contemporary geographic distribution by the close of the Pleistocene at the latest.

Geographic and taxonomic sampling

To better understand the geographic distribution of barking frogs, we geo-referenced locality records of *C. augusti* and *C. tarahumaraensis* from natural history collections in the United States using data from HERPNET 2 (Fig. 1, Appendix S1, Supporting information; www.herpnet2.org). DNA was extracted from buccal swabs, muscle tissue or liver tissue of 68 barking frogs (Fig. 2; Appendix S2, Supporting information). This genetic sampling contained individuals from across the southwestern United States and Mexico including representatives of *C. tarahumaraensis*, and three of the four subspecies of *C. augusti* (all except *C. a. fuscofemora*) described by Zweifel (1956). We used *C. uno*, a member of the closely related *Craugastor bocourti* Species Series (sensu Hedges *et al.* 2008), as an out-group taxon.

DNA extraction and mitochondrial analysis

Genomic DNA was isolated using either a commercial spin-column kit (Qiagen®, Valencia, CA, USA) or a salt extraction protocol (sensu Rovito *et al.* 2012). We examined the quality of our isolates using visualization on a 1–2% agarose gel and a QUIBIT® 2.0 fluorometer (Life Technologies, Grand Island, NY, USA).

We sequenced a 461 base pair fragment of the small subunit ribosomal RNA gene (12S) for 68 individuals (Appendix S2, Supporting information). We amplified this fragment using the primers 12SF (5′ AAA CTG GGA TTA GAT ACC CCA CTA 3′) and 12SR (5′ GTR

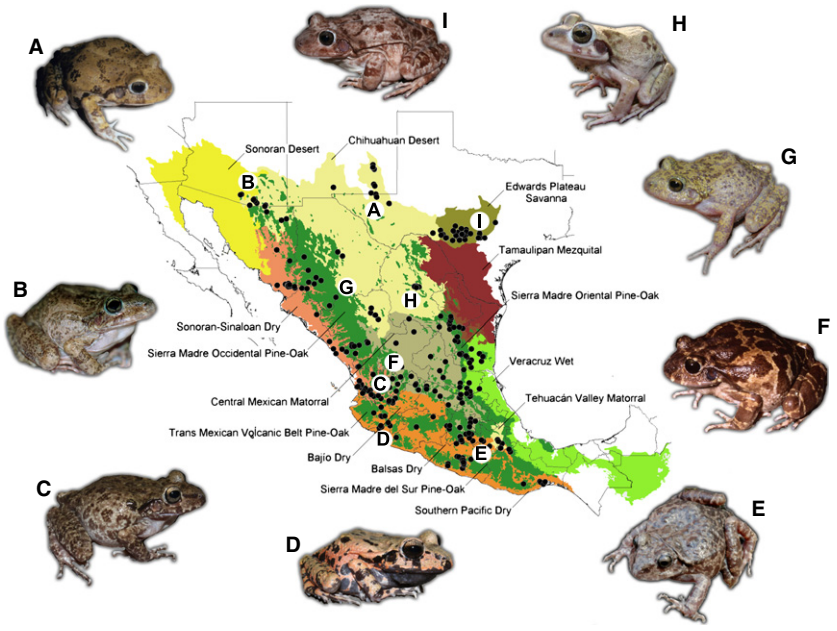

Fig. 1 Geographic distribution of the *Craugastor augusti* Species Series (*Craugastor augusti* and *C. tarahumaraensis*) inferred from 608 geo-referenced museum locality records. Across this ecologically dynamic range these species vary in colour pattern and adult body size (A–I). Terrestrial ecoregions follow Olson *et al.* (2001).

CGC TTA CCW TGT TAC GAC TT 3′). Each PCR amplification was performed using either GoTaq® Green Master Mix (Promega, Madison, WI, USA) or individual reagents from New England Biolabs (NEB), Ipswich, MA, USA (M0320L, *Taq* DNA polymerase; N0447S, dNTP solution). See Streicher *et al.* (2009) for the thermal cycling profile used to amplify 12S. PCR purifications were performed using AMPure XP (Beckman Coulter, Brea, CA, USA), Sera-Mag Speedbeads (Fisher Scientific, Pittsburgh, PA, USA; Rohland & Reich 2012) or ExoSAP-IT (Affymetrix, Santa Clara, CA, USA; according to the manufacturer's instructions). Cycle sequencing reactions and DNA sequencing were conducted using BigDye® terminator (Life technologies) chemistry and standard cycling profiles by either SeqWright (www.seqwright.com; Houston, TX, USA) or the UTA genomics core facility (gcf.uta.edu; Arlington, TX, USA). We sequenced the forward and reverse complement for each 12S fragment. Chromatograms were aligned and cleaned using the programs SEQUENCHER 5.0 (GeneCodes Corp., Ann Arbor, MI, USA) and Geneious R6 (Biomatters; www.geneious.com). All 12S sequences were deposited in GenBank (Appendix S2, Supporting information). We also downloaded 12S sequences for two *C. augusti* (Darst & Cannatella 2004; Frost *et al.* 2006) and one *C. uno* (Streicher *et al.* 2011) from GenBank.

Alignments of 12S sequences were made using the programs MEGA 5.1 (Tamura *et al.* 2011) and SEQUENCHER 5.0. We identified the best-fit model of nucleotide evolution using Bayesian and Akaike information criteria in MEGA 5.1. We used the programs MEGA 5.1 and MRBAYES 3.1.2 (Ronquist & Huelsenbeck 2003) to generate hypotheses of mitochondrial evolution using maximum-likelihood (ML) and Bayesian Markov chain Monte Carlo (BAYES) algorithms. In ML analyses, we used 2000 bootstrap pseudoreplicates to assess nodal support (Hedges 1992). In BAYES analysis, we ran parallel searches for 10 million generations using four chains (one hot, three cold; default parameters) with sampling occurring every 1000 generations to avoid autocorrelation. We examined convergence of likelihood scores and topologies with the online software 'Are We There Yet?' (AWTY, Wilgenbusch *et al.* 2004). To assess the levels of sequence divergence among mtDNA haplogroups, we constructed a pairwise matrix of between group mean divergence (i.e. Dxy values; Nei 1987) using uncorrected *p* distances in MEGA 5.1 (Table 1).

RADseq SNP library generation and analysis

In addition to the potential for examining introgression, we used RADseq because several studies have demonstrated that this approach is effective at resolving phylogenetic and phylogeographic structure among closely related populations and species (e.g. Emerson *et al.* 2010; Hohenlohe *et al.* 2010; Rubin *et al.* 2012; Jones *et al.* 2013; Stölting *et al.* 2013; Wagner *et al.* 2013). Although the RADseq method produces genealogical biases that may preclude the use of some population genetic statistics, general measures of population differentiation appear robust to these issues (Arnold *et al.* 2013).

We constructed genomic libraries for 46 individuals of the *C. augusti* Species Series (45 *C. augusti* and one *C. tarahumaraensis*) and one *C. uno* via double-digest RADseq, a method that employs two restriction enzymes

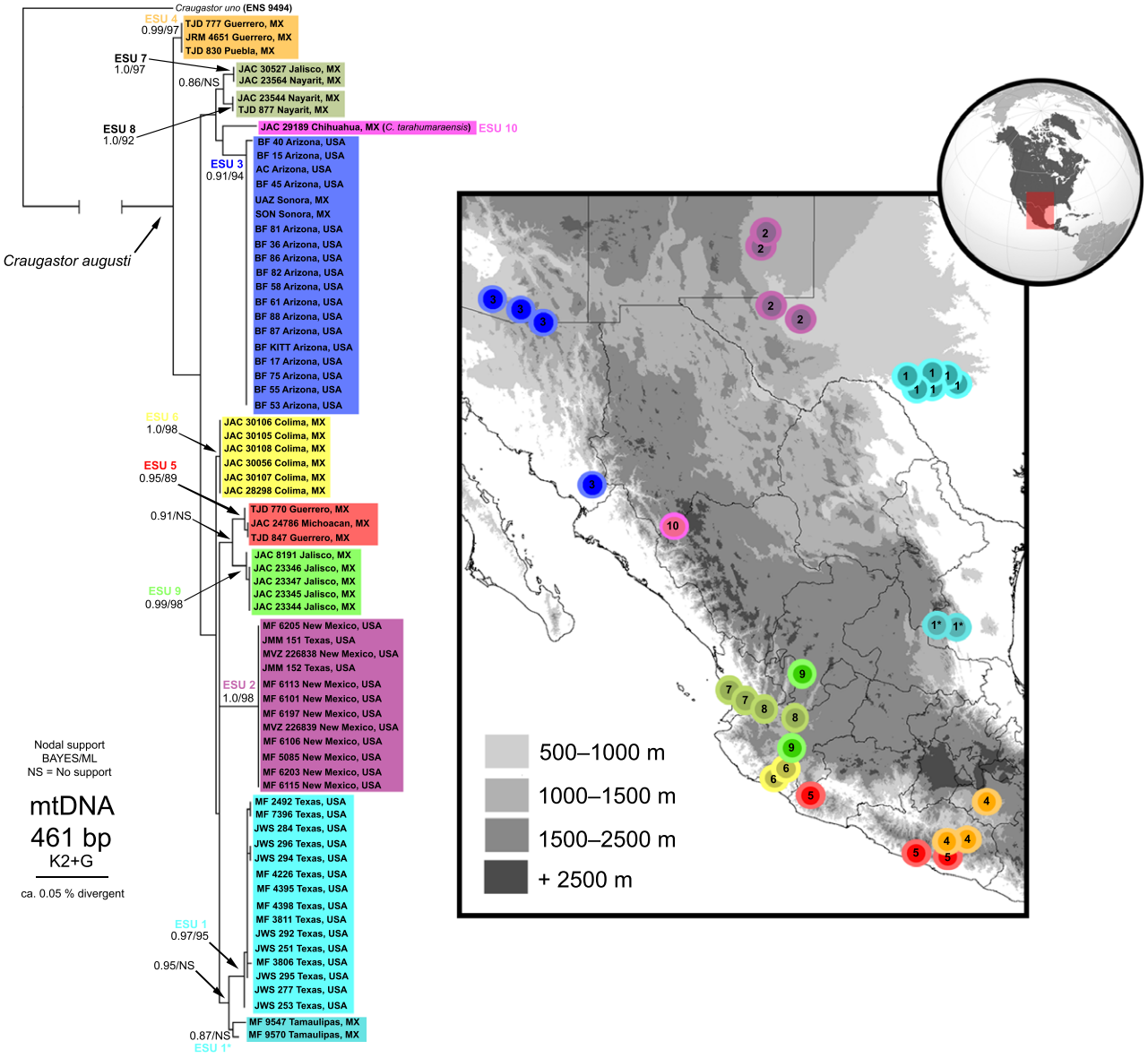

Fig. 2 Maximum-likelihood (ML) phylogram constructed from a fragment of mitochondrial DNA (Left) and the corresponding geographic distribution of *Craugastor augusti* Species Series samples used in this study (Right). Nodal support is derived from ML bootstrap analysis and Bayesian MCMC (BAYES MCMC) posterior probabilities, respectively. ESUs are described in Table 3.

and precise size selection to produce highly correlated read counts across individuals (Peterson *et al*. 2012). Collectively, this sampling included representatives of each major lineage, we identified using mtDNA. After analysing DNA isolates on a Qubit fluorometer using a dsDNA HS assay kit (Life technologies), we digested approximately 200 ng of DNA for each individual with the enzymes *Sbf*I and *Msp*I (NEB). See Appendix S3 (Supporting information) for our specific adaptor/sample design. Following adaptor ligation, cleaned samples were pooled by Illumina® index group (Appendix S3, Supporting information) and size selected for fragments ranging between 450 and 550 bp using the Blue Pippin electrophoresis platform (Sage Science, Beverly, MA, USA). RAD libraries were amplified via PCR with a Phusion® polymerase kit (NEB#M0530L). Successful library generation was confirmed by examining samples with a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) using a DNA 7500 chip. Our five Illumina® index group libraries were then pooled and sequenced using a single lane of an Illumina® HISEQ PE100 run at the University of Texas Southwestern Medical Center Genomics Core Facility (genomics.swmed.edu).

Table 1 Mean between group divergences (Dxy) generated from uncorrected *p* distances among 12S mitochondrial haplogroups in the *Craugastor augusti* species series and *Craugastor uno*

	ESU 1	ESU 2	ESU 3	ESU 4	ESU 5	ESU 6	ESU 7	ESU 8	ESU 9	ESU 10
ESU 1 Texas, USA										
ESU 2 New Mexico, USA	0.039									
ESU 3 Arizona, USA	0.047	0.044								
ESU 4 Eastern MX	0.051	0.048	0.043							
ESU 5 Southern MX	0.033	0.043	0.055	0.047						
ESU 6 Colima, MX	0.024	0.031	0.036	0.038	0.026					
ESU 7 Western MX	0.046	0.040	0.031	0.043	0.036	0.035				
ESU 8 Western MX	0.045	0.043	0.033	0.046	0.043	0.033	0.014			
ESU 9 Jalisco MX	0.035	0.042	0.051	0.041	0.023	0.024	0.033	0.043		
ESU 10 *C. tarahumaraensis*	0.054	0.058	0.036	0.056	0.063	0.044	0.035	0.033	0.050	
C. uno	0.162	0.159	0.155	0.163	0.156	0.159	0.162	0.166	0.163	0.175

We used STACKS (Catchen *et al.* 2011, 2013) to process our restriction-site-associated DNA markers (RAD-tags) and to produce SNP data sets. We used the recommended workflow described by Catchen *et al.* (2011) to identify RAD-tags containing SNPs (i.e. process_radtags, ustacks, cstacks and sstacks modules). We used default settings in (i) process_radtags to remove low quality reads (−*q* flag; if the average quality score of a read was below 90%, the read was discarded), (ii) ustacks to set a maximum distance allowed between individual stacks (−*M* flag; no more than two differences) and (iii) cstacks to set a distance between catalogue loci (−*n* flag; setting of 0). We believe the default setting of the distance between catalogue loci parameter is a conservative approach for inferring orthology among RAD-tags because although it can lead to a small number of haplotypes of the same locus being represented independently in the catalogue, it reduces the risk of assembling erroneous or paralogous RAD-tags. While we constructed our individual ustacks libraries using a minimum stack depth of three (−*m* 3; sensu Jones *et al.* 2013), we used only those RAD-tags that possessed a stack depth of five (−*m* 5) when compiling our final SNP alignments for clustering analyses. Furthermore, we used a custom Perl script to remove any RAD-tag that matched two or more stacks.

To estimate relationships among barking frogs from RAD-tags, we used two approaches: (i) a Bayesian clustering method that used bi-allelic SNP data from each frog (STRUCTURE, Pritchard *et al.* 2000) and (ii) neighbour joining (NJ) and ML phylogenetic analyses in MEGA 5.1 and RAXML 8.0 (Stamatakis 2014) that only used SNPs identified as fixed within individuals (i.e. homozygous) but polymorphic among individuals. We selected this sampling strategy because it is how the STACKS 'populations' module sorts homozygous/heterozygous SNP data for STRUCTURE and PHYLIP files, respectively. Because some of our analyses assumed unlinked loci (e.g. STRUCTURE), we only analysed 'read 1' data from our PE100 sequences to reduce the number of SNPs occurring in close genomic proximity to one another. In STRUCTURE analyses, we iteratively conducted runs of $K = 1$–13 with a burn-in of 10 000 generations and each analysis sampling every 100 generations for 100 000 generations. Each K value was run for five iterations. We used the online software STRUCTURE HARVESTER (Earl & vonHoldt 2012) to implement the ΔK statistic of Evanno *et al.* (2005) to identify an appropriate number of clusters. We also identified the appropriate number of genetic clusters using *k*-means clustering as implemented in the software ADEGENET 1.3-7 (Jombart 2008). This method ranks clustering patterns using BIC scores from axes derived from a Principle Components Analysis (PCA; run using ade4 package in R version 2.15.2, R Core Team 2012). We evaluated *k*-means clustering schemes up to $K = 40$. In MEGA 5.1 analyses, we performed 2000 bootstrap replicates to assess nodal support for SNP trees using ML criteria. Model-based (ML)

analyses were run using the GTR+G model (sensu Cariou *et al.* 2013). A more detailed account of the methods we used to generate and analyse RADs is provided in Appendix S4 (Supporting information).

Rubin *et al.* (2012) and Cariou *et al.* (2013) found that they were able to accurately estimate known phylogenies using RADseq data matrices with large amounts of missing data and suggested that RADseq based phylogenies may be *per se* robust to missing data issues. To examine the effect of missing data in our SNP analyses, we constructed three separate data matrices: (i) ≤50% missing data/site, (ii) ≤25% missing data/site and (iii) ≤10–20% missing data/site (10% in NJ/ML, 20% in STRUCTURE analyses). We generated these matrices by modifying the output commands in the 'populations' module of STACKS. All six of our data matrices and the commands used in our STACKS pipeline are available through the Dryad website (www.datadryad.org). The presence of F_{ST} outlier loci can lead to erroneous conclusions about phylogenetic relationships inferred from SNP data (Luikart *et al.* 2003). To address this issue, we used the program BAYESCAN to identify F_{ST} outliers (Foll & Gaggiotti 2008) on our 20% missing data matrix. This programme has been demonstrated via simulation to recover fewer false positives than alternative methods for detecting outlier loci (Pérez-Figueroa *et al.* 2010). We used two definitions of population structure to test for outlier loci at different levels of relatedness. We used PDGSPIDER 2.0.5.1 (Lischer & Excoffier 2012) to prepare files for BAYESCAN analysis and the R statistical package to summarize and plot results.

Inferring species trees

We employed a multispecies coalescent model (Rannala & Yang 2003) to estimate species trees (i.e. phylogenetic relationships among well-supported terminal clades). We considered a group of individuals to be a well-supported lineage if they possessed matched mtDNA and nucDNA clustering patterns. These lineages were used as a priori designated 'species' in our species tree analysis, and we hereafter refer to them as evolutionarily significant units (ESUs, Moritz 1994; Table 3). We used the program SNAPP 1.1.1, to generate species tree support measures (Bryant *et al.* 2012). At the time of our analysis, SNAPP 1.1.1 did not incorporate missing data, so we used a subset of our taxonomic sampling that maximized the number of SNPs available for analysis while retaining representatives of each ESU. In one instance, we combined two sister ESUs (supported in mtDNA and some nucDNA analyses) so that each 'species' contained data from at least two individuals. We used default settings and ran the analysis for 1 million generations sampling every 1000 generations. We confirmed convergence of our runs by examining log files for evidence of convergence (i.e. little variation in $-lnL$ score, ESS > 100) in the program TRACER 1.5 (Rambaut & Drummond 2007). We analysed tree files with the TreeSetAnalyser program that is distributed with SNAPP 1.1.1 to identify species trees that were contained in the 95% highest posterior density (HPD) set. Resulting tree files were visualized using the program DENSITREE (Bouckaert 2010).

Tests for introgression

All *D*-statistic tests are designed to test if the number of derived alleles (designated 'B' in ABBA-BABA patterns) shared by reciprocally monophyletic taxa is greater than would be expected under the null model of ILS (Durand *et al.* 2011). Derived alleles are identified in the context of an out-group taxon that possesses a different (i.e. ancestral) allelic state (designated 'A' in ABBA-BABA patterns). To describe taxon placement for these tests, we used the terminology of Durand *et al.* (2011) for four-taxon tests ((P1, P2), P3), O) and Eaton & Ree (2013) for five taxon tests ((P1, P2), ($P3_1$, $P3_2$)), O). Each Patterson's *D*-statistic test contained two ESUs (P1 & P2) recovered as sister taxa in our phylogenetic reconstructions, a divergent ESU that may have experienced admixture with one or both of the sister taxa (see Appendix S4 for details on how we selected P3 taxa, Supporting information) and an out-group (O). Each partitioned *D*-statistic test used a similar sampling strategy except we used two P3 taxa ($P3_1$ and $P3_2$) to infer the putative source of introgression. To summarize the number of sites with ABBA and BABA (Patterson's *D*-statistic) and ABBBA, BABBA, ABBBA, BABAA, ABABA and BAABA (partitioned *D*-statistic) patterns, we used custom R scripts based on the equations presented in Durand *et al.* (2011) and Eaton & Ree (2013). Our FASTA formatted alignments of four and five taxa are available on Data Dryad. We assessed confidence in *D*-statistics using a similar bootstrapping method to Eaton & Ree (2013) by employing 1000 bootstrap pseudoreplicates in all tests and using *Z*-test *P*-values to assess if allelic ratios were significantly different from zero. Scripts were executed in R version 2.15.2 and are available as part of the evobiR package (http://cran.r-project.org/web/packages/evobiR/index.html).

We interpreted *D*-statistic results in the following way: In four-taxon tests, a significant *D*-statistic indicated more derived SNPs were shared between P3 and P1 (negative value) or P3 and P2 (positive value) than expected from ILS. In five taxon tests, three *D*-statistics were calculated: D_1, D_2 and D_{12} (Eaton & Ree 2013). A significant D_1 statistic indicated introgression-like patterns between P1 and $P3_1$ (negative value) or P2 and

$P3_1$ (positive value). A significant D_2 statistic indicated introgression-like patterns between P2 and $P3_1$ (negative value) or P2 and $P3_2$ (positive value). A significant D_{12} statistic indicated introgression-like patterns between P2 and $P3_1 + P3_2$ (negative value) or P1 and $P3_1 + P3_2$ (positive value). We inferred the directionality of gene flow via partitioned *D*-statistic tests by using inverse taxonomic combinations where the positions of P1 and P2 are switched with $P3_1$ and $P3_2$. If we observed significant *D*-statistics in only one direction of an inverted comparison, we interpreted this as evidence for unidirectional gene flow (Eaton & Ree 2013).

Isolation-by-distance tests and recipient/donor/loner comparisons

We tested for isolation by distance (IBD; Wright 1943) using (i) 12S mtDNA sequences and (ii) the 10% missing data SNP alignment to address two issues. First, we wanted to compare spatial nucDNA and mtDNA patterns for evidence of sex-biased dispersal. While not a direct test, given the unidirectional inheritance (matrilineal) of mtDNA, matched mtDNA-nucDNA IBD patterns are consistent with sexes possessing broadly similar dispersal capabilities (Prugnolle & de Meeus 2002). Second, we wanted to understand the extent of spatial autocorrelation within our nucDNA data set so that we could correct for this common effect when comparing within lineage genetic variation values. Geographic distances were calculated using the program GENALEX 6.5 (Peakall & Smouse 2012). Statistical significance of the correlation between geographic and genetic distance was calculated using a Mantel test (Mantel 1967) with 999 randomizations in GENALEX 6.5.

Following introgression tests we categorized ESUs into three groups: recipients, donors, and loners (see Introduction for definitions) and generated mean within group genetic distance values (generated using the 20% missing SNP data set in MEGA) to compare these groups. We acknowledge that these categorical groupings are not mutually exclusive (i.e. if symmetrical introgressive gene flow had occurred, donor lineages would also have to be considered recipient lineages); however, we believe they are a conceptually useful analogy to describe the three patterns of introgressive gene flow we observed in barking frogs.

Results

An ecologically diverse geographic distribution

In total, we collected 608 locality records for barking frogs in the United States and Mexico (Fig. 1; Appendix S1, Supporting information). We mapped point locality records onto the terrestrial ecoregions and Ecozones of Olson *et al.* (2001). This examination confirmed previous reports of the broad distribution of barking frogs (Zweifel 1956), and suggests that barking frogs are absent or rarely encountered across most of the Chihuahuan Desert in north central Mexico and the Tamaulipan Mezquital region of northeastern Mexico.

Mitochondrial divergence suggests a rapid Pleistocene radiation

The K2 + G model was selected as the most appropriate for our 12S data set (*ln*L = −1398.866). In our BAYES MCMC analysis, the average standard deviation of split frequencies was 0.007 following 10 million generations. AWTY analysis indicated that topological convergence occurred around 5 million generations. Thus, we discarded the first 50% of sampled trees as burn-in. As noted by Goldberg *et al.* (2004), sequence divergence levels between the in-group taxa and their closest living relatives (members of the *C. bocourti* Species Series [*C. uno* in this study]) are large. Within the *C. augusti* species series, no pairwise comparison of mtDNA divergence levels exceeds 6% with most comparisons ranging between 3% and 5% (Table 1). We recovered nine well-supported mtDNA clades (>92 ML; 0.95 BAYES) of *C. augusti* (Fig. 2). These haplogroups are more or less geographically circumscribed and typically restricted to particular ecogeographic zones (Figs 1 and 2). We recovered only limited statistical support for relationships among haplogroups. We also recovered *C. tarahumaraensis* as nested within *C. augusti* (Fig. 2). The 12S locus evolves at a rate of about 1–1.5% per million years in amphibians (Mueller 2006). Based on this generalized rate, the oldest barking frog populations would have started diverging from one another at the onset of the Pliocene (Table 1). Additionally, the shape of the mtDNA topology, a deep root (*C. uno*) and short ancestral internodes, may suggest that a rapid radiation across Mexico occurred during the Pleistocene evolution of barking frogs (sensu Rothfels *et al.* 2012).

Single nucleotide polymorphism data provide evidence of extensive phylogeographic structure

Our STACKS workflow generated SNP data sets that contained between 28 152 and 271 SNPs depending on the amount of missing data (Table 2). We found that the amount of missing data did not substantially affect topology or clustering (Appendix S4, Supporting information). As in the mtDNA analysis, branch lengths between the in-group and *C. uno* were large (Fig. S3, Supporting information). Additionally, while the individual of *C. uno* had many high quality reads and a

Table 2 Number of orthologous single nucleotide polymorphisms (SNPs) between *Craugastor augusti* and the two out-group taxa used in this study (*C. tarahumaraensis* and *C. uno*; top and bottom, respectively). SNP counts were obtained using STACKS output files with variable amounts of missing data. See text for descriptions of bi-allelic and fixed loci

	50% missing	25% missing	20–10% missing
C. tarahumaraensis			
Bi-allelic	4160/8842	919/1455	590/892
Fixed alleles	14 472/28 152	2740/4319	394/580
C. uno			
Bi-allelic	142/5016	47/665	31/362
Fixed alleles	649/26 272	184/3873	24/271

catalogue of over 5500 unique alleles, depending on the amount of missing data we allowed, only between 24 and 649 SNPs were recovered for this taxon when comparing it with orthologs from *C. augusti* and *C. tarahumaraensis* (Table 2). This discrepancy may be related to nonrandom sampling of haplotypes (i.e. Arnold *et al.* 2013) because *C. uno* and the in-group shared a common ancestor at least 25 million years ago (Crawford & Smith 2005). Based on the relatively small number of orthologous SNPs we recovered from *C. uno* and a putative sister relationship between *C. augusti* and *C. tarahumaraensis* (Fig. S3, Supporting information), we removed *C. uno* from subsequent analyses and used *C. tarahumaraensis* as an out-group taxon. Across all three SNP data sets (50%, 25%, and 20% missing data), ΔK statistics indicated that a K of 2–3 was most appropriate (Fig. S1, Supporting information). However, unlike the analyses implementing 2–3 clusters, K values of 9–12, typically assigned individuals to clusters with high posterior assignment (i.e. >0.95; see Fig. 3 where $K = 11$). Furthermore, these clusters largely matched the mtDNA clades recovered in our analysis of 12S (Fig. 2). This discrepancy may be explained by ΔK values being highest at the uppermost clustering level (which often necessitates removing taxa to detect substructure within discrete units; for example Devitt *et al.* 2013), and known caveats of analysing data sets with sample size

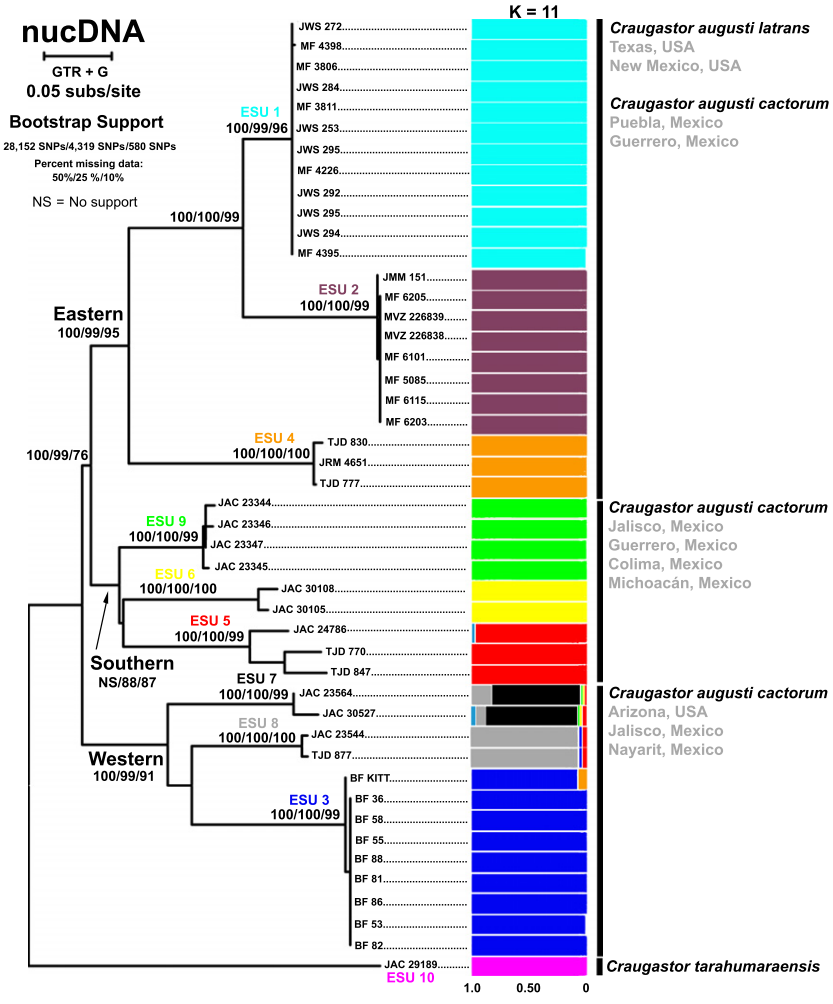

Fig. 3 Maximum-likelihood phylogram generated from 580 single nucleotide polymorphisms (SNPs; Left) and STRUCTURE plot ($K = 11$) generated from 892 SNPs (Right). Nodal support is derived from 2000 bootstrap pseudoreplicates conducted using 28 152 (50% missing), 4319 (25% missing) and 580 (10% missing) SNP data sets, respectively. ESUs are described in Table 3.

variation in STRUCTURE (Kalinowski 2011). In contrast, ADEGENET *k*-means clustering was consistent across varying levels of missing data with increasing BIC scores occurring around $K = 10$. This may indicate that for large and variable SNP data sets (i.e. those that include variation beyond population levels) *k*-means clustering of principal components is a more robust 'single analysis' discovery procedure for determining the approximate number of natural groups.

Our SNP trees and STRUCTURE analyses recovered all nine mtDNA haplogroups of *C. augusti* as well-supported clusters (Fig. 3). Based on these matched patterns of mtDNA and nucDNA clustering, we designated nine ESUs within *C. augusti* and a tenth ESU corresponding to *C. tarahumaraensis* (Table 3). Phylogenetic reconstructions of SNP data revealed statistical support for three geographically circumscribed groups of *C. augusti*: (i) a western group containing ESUs from Arizona, USA and the Mexican states of Nayarit and Jalisco, (ii) a southern group containing ESUs from the southern Mexican states of Jalisco, Nayarit, Colima, Michoacán, and Guerrero and (iii) an eastern group containing ESUs from Texas and New Mexico, USA as well as the Mexican states of Puebla and Guerrero. Statistical support was also recovered for the reciprocal monophyly of the eastern and southern groups relative to the western group. BAYESCAN analyses were performed using two 'population' schemes: (i) a nine population scheme where we treated each ESU (Table 3) as a population and (ii) a three population scheme where we treated the geographically circumscribed groups as populations. Using a false discovery rate of 0.05 we found zero outlier loci. However, mean F_{ST} values as inferred from the posterior distributions were high for both schemes (0.90 and 0.70, nine and three population, respectively; Fig. S2, Supporting information), which indicates clearly differentiated populations. Thus, the detection of outlier loci may have been limited by the observed F_{ST} distributions.

Species tree analysis was performed using 353 SNPs and 39 samples for nine species that, with one exception, correspond to the ESUs listed in Table 3. Excluding 10% of topologies as burn-in, 33 topologies (out of 70) were in the 95% HPD. The majority (*ca*. 71%) of 95% HPD species trees placed *C. tarahumaraensis* as the most basal member of the *C. augusti* Species Series, and recovered the major clades of *C. augusti* identified by the ML analyses: (i) ESUs 1, 2 and 4 (the eastern group), (ii) ESUs 3, 8 and 9 (the western group) and (iii) ESUs 5, 6 and 7 (the southern group). Most statistical uncertainty occurred among the three major lineages of *C. augusti* and within the southern group (Fig. 4).

Table 3 Barking frog evolutionarily significant units (ESUs) inferred from matched clustering patterns of mtDNA and genome wide single nucleotide polymorphism (SNP) data. Relevant taxonomic, geographic and distributional information are also listed

Unit	Geographic distribution	Taxonomic components	Eco-region
ESU 1	Edwards Plateau regions of Central Texas extending into northeastern Mexico in Tamaulipas state	*C. augusti latrans* *C. augusti augusti**	Savanna/Shrubland; post-oak, mesquite
ESU 2	Chihuahuan desert region including parts of west Texas and southwestern New Mexico	*C. augusti latrans*	Desert/Arid Scrub; Chihuahuan Desert
ESU 3	Arid regions of southern Arizona extending into northwestern Mexico in Sonora state	*C. augusti cactorum*	Sky islands of the Sonoran desert; extending into Dry Tropical Forests in western Mexico.
ESU 4	High elevation regions of the Sierra Madre del Sur	*C. augusti cactorum*	Tropical pine-oak forests
ESU 5	Mid-elevation regions of the Sierra Madre del Sur in southern Mexico	*C. augusti cactorum*	Transitional tropical dry to pine-oak forest
ESU 6	Mid-elevational regions southwestern Mexico	*C. augusti cactorum*	Transitional tropical dry to pine-oak forest
ESU 7	Mid-elevation regions of Jalisco and Nayarit in western Mexico	*C. augusti cactorum*	Transitional tropical dry to pine-oak forest
ESU 8	High elevation regions of Jalisco and Nayarit in western Mexico	*C. augusti cactorum*	Transitional tropical dry to pine-oak forest
ESU 9	High elevation regions of Jalisco	*C. augusti cactorum*	Tropical pine-oak forests
ESU 10	Sierra Madre Occidental	*C. tarahumaraensis*	Pine-oak forest

*Based on mtDNA data only.

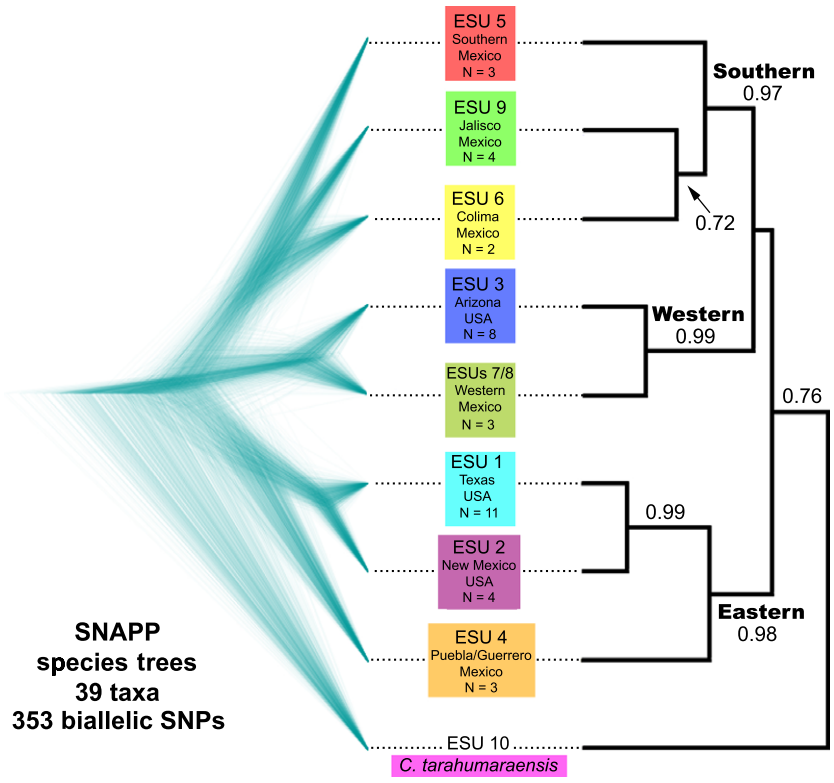

Fig. 4 Figure 4 Densitree diagram depicting topologies of 9999 species trees obtained from an analysis of 353 single nucleotide polymorphism loci from 39 barking frogs using the program SNAPP (Left), and associated root canal depicting a consensus topology from this analysis (Right). Nodal support values on the root canal are posterior probabilities that correspond to strongly supported nodes not designated a priori in the species tree analysis. ESUs are described in Table 3.

Patterns of introgression may indicate a long history of asymmetrical gene flow in western Mexico

To examine the nine ESUs of *C. augusti* for evidence of introgression, we reduced STACKS population maps to four or five individuals that each represented an ESU of interest (see Appendix S4 for sampling design, Supporting information). In all scenarios, we used *C. tarahumaraensis* (ESU 10) as the out-group taxon. To increase the number of fixed SNPs available for analysis, we lowered the *−m* parameter of the STACKS 'populations' module to 3 (sensu Jones *et al.* 2013). For Patterson's *D*-statistic tests, we tested SNP alignments ranging in size from 4032 to 7520 SNPs (Table 4). While some of these tests revealed no evidence of introgression (T1–T4; Table 4), several identified introgression-like genealogical patterns (T5–T13; Table 4). To further investigate these patterns, we used partitioned *D*-statistics (Table 5). These five taxon tests were conducted using SNP alignments ranging in size from 2408 to 4413 SNPs. Similar to Eaton & Ree (2013) we found that there were relatively few sites possessing the patterns used by the test (usually <50/genealogical comparison; Table 5). Thus, although we present all of the resulting partitioned *D*-statistics for transparency, to be conservative we only interpreted *D*-statistics derived from comparisons of more than 50 genealogical patterns as meaningful. Partitioned *D*-statistics revealed two interesting patterns: (i) relative to other nonsister ESU parings, ESUs 8 and 9 had more shared SNPs than expected from ILS (D_2 statistics, T17, T22, T24–25; Table 5) and (ii) the western ESUs possessed more shared SNPs with the southern ESUs than expected from ILS (D_{12} statistics, T14–T18; Table 5) whereas the southern ESUs did not display this pattern when compared the western ESUs (D_{12} statistics, T19–T23;

Table 4 Tests for introgression as inferred from four-taxon set *D*-statistics. ESU designations correspond to Table 3

Four-taxon ESU tests (((P1, P2), P3), O)	SNPs	ABBA	BABA	*D*	*P*
T1: (((1, 2), 3), 10)	4870	41	39	0.03	0.4555
T2: (((1, 2), 4), 10)	5433	78	75	0.02	0.6614
T3: (((1, 4), 5), 10)	5819	167	167	0	0.5218
T4: (((6, 9), 7), 10)	4946	113	99	0.07	0.1860
T5: (((7, 3), 6), 10)	5810	66	146	−0.38	**0.0000**
T6: (((7, 3), 9), 10)	5771	64	144	−0.38	**0.0000**
T7: (((7, 3), 5), 10)	4690	60	98	−0.24	**0.0016**
T8: (((7, 8), 5), 10)	6295	57	290	−0.67	**0.0000**
T9: (((7, 8), 6), 10)	5935	63	156	−0.43	**0.0000**
T10: (((7, 8), 9), 10)	4517	45	103	−0.39	**0.0000**
T11: (((8, 3), 5), 10)	6527	85	264	−0.51	**0.0000**
T12: (((8, 3), 6), 10)	4034	64	156	−0.42	**0.0000**
T13: (((8, 3), 9), 10)	7520	83	471	−0.70	**0.0000**

Significant deviations ($P > 0.005$) from symmetrical ancestral allele patterns are bolded.

Table 5). We interpret the first pattern as evidence that introgressive gene flow occurred between ESUs 8 and 9. Given the significant D_{12} statistic from test T24 and the nonsignificant D_{12} statistic from T25 (Table 5), the directionality of this putative introgression was inferred to be asymmetric (Fig. 5; event 1.1). The second pattern is also consistent with asymmetrical gene flow. Ignoring comparisons involving ESUs 8 and 9 (where southern clade ancestral polymorphisms that ESU 8 would have acquired from ESU 9 could create false positive SNP patterns when compared with ESUs 5 and 6), we recovered SNP patterns consistent with unidirectional introgression from ESUs 5 and 6 into ESUs 3 and 7 (D_{12} statistics T14–15, T18-T20, T23; Table 5). Given the geographic distribution of these groups (Fig. 2), it is unlikely that they have had recent opportunities to hybridize. Thus, we tentatively interpret these statistics as evidence for introgressive gene flow from the ancestral stock of the southern clade into the ancestral stock of the western clade (Fig. 5; event 1.2.); however, we acknowledge the evidence is not as compelling as for event 1.1 (Fig. 5). Based on the putative patterns of introgression observed among ESUs, we considered ESUs 1, 2 and 4 as loners, ESUs 5, 6 and 9 as donors and ESUs 3, 7 and 8 as recipients.

Detecting introgression events caused us to be concerned about the accuracy of our 'total-evidence' SNP-based topologies (Figs 3 and 4) given that introgression can obscure phylogenetic inference (Rheindt & Edwards 2011; Eaton & Ree 2013). To test the influence of introgressed alleles on our phylogenetic reconstructions, we conducted a series of NJ reconstructions using our 10% missing data alignment where we removed ESUs that were identified as sources of introgression. Regardless of which ESUs were removed, the major groups (eastern, western, and southern) and branching patterns among ESUs did not vary topologically.

Matched IBD patterns in mitochondrial and nuclear genomes

Mantel tests revealed that levels of both mtDNA 12S and SNP divergence had a positive and significant ($P < 0.0001$) relationship with geographic distance. Interestingly, geographic distance explained similar amounts of variation in the mtDNA ($r^2 = 0.41$) and genome wide SNP ($r^2 = 0.43$) alignments. Additionally, both comparisons featured several cases of highly divergent genetic pairs occurring in close geographic proximity (Fig. S4, Supporting information). Despite this finding, where redundant locality sampling was performed (ESUs 1, 2, 3, 6, and 9), we found no evidence for the syntopic occurrence of divergent genetic lineages.

Table 5 Five taxon partitioned *D*-statistic tests for ESUs belonging to the western and southern clades of *Craugastor augusti*

Five taxon ESU tests (((P1, P2), ($P3_1$ $P3_2$)) O)	Sites	Derived only $P3_1$				Derived only $P3_2$				Derived P3 shared				
		ABBAA	BABAA	D_1	P	ABABA	BAABA	D_2	P	ABBBA	BABBA	D_{12}	P	
T14: (((3, 7), (6, 5)), 10)	3478	8	18	−0.38	0.0001	9	12	−0.14	0.1000	**21**	**52**	**−0.42**	**0.001**	1.1
T15: (((3, 7), (6, 9)), 10)	3915	13	24	−0.29	0.0060	12	27	−0.38	0.0001	**15**	**54**	**−0.57**	**0.000**	1.1
T16: (((7, 8), (6, 5)), 10)	3390	30	14	0.36	0.0090	22	18	0.00	0.6600	**42**	**16**	**0.45**	**0.000**	1.1
T17: (((7, 8), (5, 9)), 10)	3449	16	19	−0.09	0.4700	**61**	**11**	**0.69**	**0.0000**	**49**	**10**	**0.66**	**0.000**	1.1/1.2
T18: (((3, 7), (5, 9)), 10)	2472	11	15	−0.15	0.2604	24	9	0.45	0.0000	25	16	0.22	0.042	1.1/1.2
T19: (((6, 5), (3, 7)), 10)	2920	13	4	0.53	0.0000	7	14	−0.33	0.0066	*36*	*37*	*−0.01*	*0.890*	1.1
T20: (((6, 9), (3, 7)), 10)	2408	9	10	−0.05	0.7200	17	15	0.06	0.4140	*39*	*29*	*0.15*	*0.288*	1.1
T21: (((6, 5), (7, 8)), 10)	3390	18	14	0.125	0.7300	*22*	*30*	*−0.15*	*0.160*	*48*	*56*	*−0.08*	*0.520*	1.1
T22: (((5, 9), (7, 8)), 10)	3452	11	19	−0.27	0.0175	**61**	**16**	**0.58**	**0.0000**	*65*	*45*	*0.18*	*0.892*	1.1/1.2
T23: (((5, 9), (3, 7)), 10)	2472	9	15	−0.25	0.0420	24	11	0.37	0.0030	*36*	*27*	*0.14*	*0.740*	1.1/1.2
T24: (((7, 8), (6, 9)), 10)	4413	21	22	−0.02	0.8000	**80**	**22**	**0.57**	**0.0000**	**77**	**15**	**0.67**	**0.000**	1.2
T25: (((6, 9), (7, 8)), 10)	4413	22	22	0.00	0.9200	**80**	**21**	**0.58**	**0.0000**	*67*	*56*	*0.08*	*0.660*	1.2

Figure 5

Significant tests ($P < 0.005$) where more than 50 sites were used to compare genealogical patterns are bolded. Nonsignificant tests where more than 50 sites were used are italicized. See Appendix S4 (Supporting information) for additional information. Alternating shading groups together values corresponding to the three test statistics generated by the method; D_1 , D_2, and D_{12}.

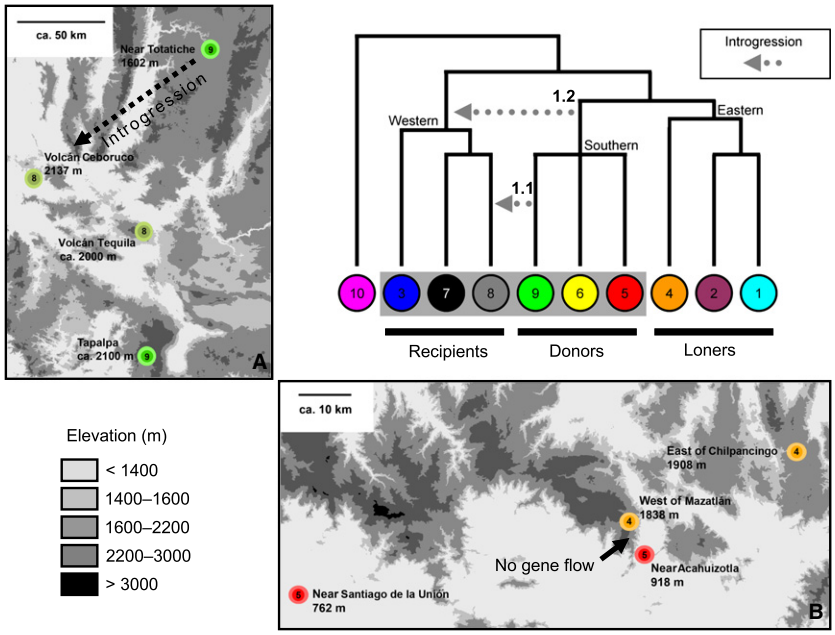

Fig. 5 Diagram depicting the directionality of introgression in the *Craugastor augusti* complex as inferred from Patterson's *D*-statistic tests (shaded grey bar; Table 4) and partitioned *D*-statistic tests (events 1.1 and 1.2, loner, donor, recipient lineage designations; Table 5). Inset maps depict geographic proximity of ESUs 8 and 9 where introgression was detected (A) and ESUs 4 and 5 where introgression was not detected (B). ESUs are described in Table 3.

Recipient and donor lineages possess higher levels of genetic variation than loner lineages

To correct for the effect of geographic distances (Fig. S4, Supporting information), we performed a regression analysis using average geographic distance (in km) between individuals in each ESU as an independent variable and mean within ESU genetic variation values as a dependent variable. The regression was significant ($n = 9$, $r^2 = 0.86$, $P = 0.003$), so we used residuals as a proxy for distance-corrected within-ESU genetic variation. This residual analysis revealed that donor and recipient lineages had higher levels of genetic variation than loner lineages (Fig. 6).

Discussion

We used mitochondrial DNA and genome wide SNP data to construct phylogenetic hypotheses, analyse spatial patterns of genetic variation, and provide evidence for unidirectional nucDNA introgression among ESUs of the *C. augusti* complex. Additionally, we found that those ESUs involved in putative hybridization events possessed more average genetic variation than ESUs with no evidence of historic introgression (Fig. 6). As with previous studies that have detected introgression using *D*-statistics (Durand *et al.* 2011; Eaton & Ree 2013), our findings have important implications for interpreting DNA sequence variation in well-differentiated populations and species. Specifically, assumptions of how gene flow occurs in (i) coalescent-based methods for inferring phylogenies (Leaché 2009; Liu *et al.* 2009; Fujita *et al.* 2012) and (ii) historical demographic inferences (see Paz-Vinas *et al.* 2013) may lead to data misinterpretation if a significant portion of standing genetic variation is explained by introgressive gene flow and not time since initial divergence. For example, we note that if we had not used *D*-statistic tests, our

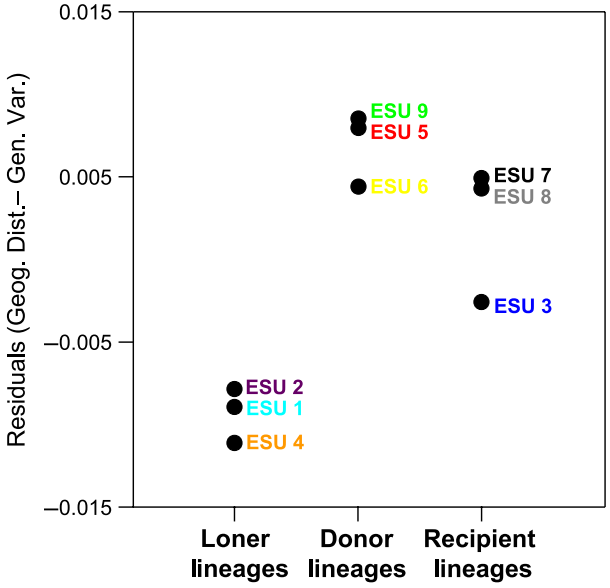

Fig. 6 Residual scores from a regression analysis of mean geographic distance between samples (in km) and mean genetic variation between samples (uncorrected '*p*' distances) for nine ESUs of *Craugastor augusti* grouped by loner, donor and recipient categories. ESUs that we found to have allelic patterns consistent with introgressive hybridization (donors and recipients; Tables 4 and 5) had higher residual scores. ESUs are described in Table 3.

findings would not have included the identification of hybrid gene flow in western Mexico. In this scenario, we would have concluded that barking frogs have a clear genetic structure with no evidence of post-diversification admixture (STRUCTURE plot; Fig. 3). This underscores not only the importance of testing for introgression in groups with seemingly well-resolved (i.e. statistically supported) genetic structure, but also adds to the growing body of evidence that interspecific nucDNA gene flow is a widespread agent of evolutionary diversification in animals (Rheindt & Edwards 2011; Hedrick 2013).

Barking frog phylogenetics, phylogeography and taxonomy

In general, widespread amphibian complexes possess high levels of genetic variation which is often geographically circumscribed (Stuart *et al.* 2006). We found that barking frogs are consistent with this trend in that they have extensive mtDNA and nucDNA phylogeographic structure. While our nucDNA phylogenies were more or less well resolved, we recovered little statistical support for higher-level relationships in the mtDNA phylogeny. The lack of nodal support for relationships among ESUs in the mtDNA phylogeny may be explained by mtDNA introgression from *C. augusti* into *C. tarahumaraensis*. Alternatively, if barking frogs radiated rapidly early in their evolution, then the 12S locus may not be evolving quickly enough to resolve relationships and a more rapidly evolving mtDNA marker may be better suited to the task of resolving phylogenetic relationships. In some previous comparisons between matched RADseq and mtDNA data, contrasting phylogenetic signal has been reported (Jones *et al.* 2013). Although we recovered congruent nodal support for terminal clades (i.e. individuals we designated as ESUs), there were topological differences between our barking frog mtDNA and SNP phylogenies. One notable difference was the recovery of *C. tarahumaraensis* as nested among *C. augusti* haplogroups in the mtDNA whereas our SNP phylogenetic trees placed *C. tarahumaraensis* as the most basal *C. augusti* Species Series (Fig. 3 and Fig. S3, Supporting information). Although the phylogenies produced using the SNP data sets suggests our sample of *C. tarahumaraensis* is sister to a monophyletic *C. augusti* complex, the posterior probability of this relationship was low in our species tree analysis (0.76, Fig. 4).

Interestingly, the branching patterns that unite ESUs of *C. augusti* in our SNP trees are consistent with a narrative of parallel northern range expansion events along the Sierra Madre Occidental and Sierra Madre Oriental to explain the origin of *C. augusti* in Arizona and New Mexico/Texas, respectively (Fig. 3). This is consistent with the Goldberg *et al.* (2004) conclusion that barking frogs in New Mexico and Texas (ESUs 1 and 2) are more closely related to one another than either is to the Arizona population (ESU 3). Our phylogenomic results have several taxonomic implications for the *C. augusti* Species Series. While ESUs referable to *C. augusti latrans*, the Balcones Barking Frog, form a monophyletic group in our SNP trees (Fig. 3), populations from central Texas (ESU 1) and western Texas and New Mexico (ESU 2) form a polytomy with several Mexican groups in the mtDNA analysis. Furthermore, our 12S analysis recovered the central Texas population of *C. augusti latrans* as more closely related to *C. augusti augusti* from Tamaulipas, Mexico than *C. augusti latrans* from New Mexico (although only with nodal support from the BAYES MCMC analysis; Fig. 2). Both 12S mtDNA and genome wide SNP analyses indicate that *C. augusti cactorum*, the Western Barking Frog, is nonmonophyletic (Figs 2 and 3; Table 3). Given our findings, following additional geographic sampling and species delimitation analyses (e.g. Fujita *et al.* 2012) taxonomic revision will be necessary. Specifically, samples should be included that represent (i) the subspecies *C. augusti fuscofemora* (known only from the Cuatro Ciénegas region of Coahuila, Mexico; Fig. 1I) and (ii) additional samples of *C. augusti augusti* and *C. augusti cactorum* from the Trans Mexican Volcanic Belt and Central Mexican Plateau.

The role of parapatry, elevation and time in determining patterns of hybridization and genetic variation

Prior to examining our SNP data for the signature of introgression, we expected that any detectable hybridization events would be limited to lineages with contemporary parapatric distributions. Our findings, while in part consistent with this tenet (Fig. 5A), also revealed at least two deviations from our null expectation.

First, the putative signature of ancient introgression between the ancestors of the southern and western clades (Fig. 5; event 1.2) coupled with no evidence of recent introgressive gene flow in all but two ESUs (Fig. 5; event 1.1) may indicate that barking frogs from Mexico experienced complex shifts in geographic distribution leading to variable episodes of secondary contact. This narrative is consistent with evidence that hybrid zones shift spatio-temporally (Eaton & Ree 2013; Smith *et al.* 2013). Using the evolutionary rates from Mueller (2006) and divergence levels from Table 1, the ancient introgression (Fig. 5; event 1.1) would have occurred sometime between 3 and 5 Ma, whereas the more recent introgression (Fig. 5; event 1.2) would have

occurred between 0.75 and 1 Ma. An interesting implication of this result is that the directionality of hybrid interactions may be conserved over long periods of time (see also Eaton & Ree 2013), which is consistent with how hybrid compatibility is thought to evolve (Orr 1995). Second, we did not detect gene flow between several ESUs that are putatively parapatric (e.g. Fig. 5B). This pattern could be related to recent range expansion (Spaulding *et al.* 2006), ecological speciation (Nosil *et al.* 2012), or partitioned elevational dispersal corridors (Cushman *et al.* 2006). We suspect that elevation may be the most likely explanation for this pattern as (i) we found evidence of hybridization in ESUs occupying adjacent high elevational strata (ESUs 8 and 9; Fig. 5A), yet not between ESUs occupying adjacent intermediate and high elevational strata (e.g. ESUs 4 and 5; Table 4; Fig. 5B) and (ii) previous work suggests that elevation strongly influences genetic structure in frogs (Monson & Blouin 2004; Gonzalez-Voyer *et al.* 2011).

The result of our loner, donor and recipient genetic variation analysis is consistent with a history of hybridization leading to higher levels of genetic variation (Fig. 6); however, our findings did not meet the entirety of our predications. Specifically, if hybridization explained the majority of genetic variation, we expected that recipient lineages would possess the highest levels of genetic variation. Instead, we observed that donor and recipient lineages both had high levels of genetic variation relative to loner lineages (Figs 5 and 6), a pattern expected based on our Patterson's *D*-statistic test results (Table 4), but not our partitioned *D*-statistic test results (Table 5). There are several putative explanations for this pattern. First, it is almost certain that introgressive gene flow and geographic distance are not the sole determinants of genetic variation in barking frogs; variation in historical demography could easily explain the higher levels of genetic variation observed in donor lineages (Fig. 6). Second, the number of sites with the specific genealogical patterns used in our partitioned *D*-statistic tests was relatively low (4–80 sites/pattern; Table 5) compared the Patterson's *D*-statistic sites (39–471 sites/pattern; Table 4). Thus, we cannot rule out that ascertainment bias may have misled or limited our inference of gene flow directionality. Third, the appropriateness of our out-group choice for *D*-statistic tests warrants some discussion. Because *C. tarahumaraensis* is (i) closely related to the *C. augusti* complex and (ii) there may be a history of mtDNA introgression with ESU 3 (Fig. 2), the possibility of undetected ancestral or introgressed loci between *C. augusti* and *C. tarahumaraensis* should be considered when interpreting evidence for early introgression events (event 1.2.; Fig. 5). Finally, we note that the spatial distribution of our sampling was limited and that future examinations of how genetic variation may relate to introgression, demography and phylogenetic descent should include increased geographic sampling.

Barking frogs are the only member of the genus *Craugastor* to have successfully colonized or persisted in arid regions of the northern subtropics (Fig. 1). As such, the northernmost lineages of barking frog (ESUs 1–3) represent an ideal opportunity to investigate the processes that allow tropical lineages to expand their climatic tolerances.

Acknowledgements

We thank C. Liang at the UT Southwestern genomics core and J. Castoe and M. Muenzier at the UTA genomics core facility. We sincerely thank L. Canseco-Márquez and L. Ochoa-Ochoa for assistance with georeferencing localities. We are indebted to J. Campbell and O. Flores-Villela for logistical assistance with fieldwork. The Demuth laboratory at UTA provided helpful discussion on *D*-statistic tests. We thank the Wiens laboratory (University of Arizona) for discussion that greatly improved the quality of the manuscript. DNA sequencing was performed using start-up funds awarded to MKF by UTA. G. Weatherman, E. Smith, U. Garcia-Vázquez and J. Meik provided photographs. J. Meik, J. Mendelson III, C. Sheehy III, C. Cox, B. Fontenot, T. Eimermacher, J. Reyes-Velasco, E. Smith, E. Enderson, J. Rorabough, and a large number of Arizona volunteers assisted with field work. Fieldwork performed by TJD was funded in part by a UC MEXUS-CONACyT postdoctoral fellowship. Fieldwork in the US was partially supported by an AMNH Theodore Roosevelt grant to JWS. Fieldwork performed by JHM was funded by the Society for the Study of Amphibians and Reptiles, Texas Herpetological Society, East Texas Herpetological Society, the Share with Wildlife Research Grant from New Mexico Game and Fish and was permitted by the Texas Parks and Wildlife Natural Resources Program (Permit Nos. 32–98, 60–98, and 25–00), New Mexico Game and Fish Collecting Permit No. 3090, and USFW Special Use Permit No. 89794. Mexican fieldwork was supported by NSF grants DEB-0102383 and DEB-0613802 to J. Campbell. Additional fieldwork was conducted using the following scientific collection permits: Mexico (FAUT-0015 to O. Flores-Villela), Texas (SPR-0707-1387 to C. Franklin), New Mexico (No. 3557 to MKF) and Arizona (SP849158 to C. Schwalbe).

References

Anderson E (1949) *Introgressive Hybridization*. John Wiley and Sons, New York, New York.

Applegarth JS (1980) Herpetofauna (anurans and lizards) of Eddy County, New Mexico: quaternary changes and environmental implications. *Abstracts International, Section B: The Sciences and Engineering*, **40**, 3613.

Arnold B, Corbett-Detig RB, Hartl D, Bomblies K (2013) RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology*, **22**, 3179–3190.

Beebee TJC (1996) *Ecology and Conservation of Amphibians*, pp. 214. Chapman and Hall, London.

Bouckaert RR (2010) DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, **26**, 1372–1373.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, 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.

Cariou M, Duret L, Charlat S (2013) Is RAD-seq suitable for phylogenetic inference? *Ecology and Evolution*, **3**, 846–852.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) *Stacks*: building and genotyping loci *de novo* from short-read sequences. *Genes Genomes Genetics*, **1**, 171–182.

Catchen JM, Hohenlohe P, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.

Crawford AJ, Smith EN (2005) Cenozoic biogeography and evolution in direct-developing frogs of Central America (Leptodactylidae: *Eleutherodactylus*) as inferred from a phylogenetic analysis of nuclear and mitochondrial genes. *Molecular Phylogenetics and Evolution*, **35**, 536–555.

Cui R, Schumer M, Kruesi K, Walter R, Andolfatto P, Rosenthal G (2013) Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. *Evolution*, **67**, 2166–2179.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *The American Naturalist*, **168**, 486–499.

Darst CR, Cannatella DC (2004) Novel relationships among hyloid frogs inferred from 12S and 16S mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **31**, 462–475.

Davey JW, Blaxter ML (2011) RADseq: next-generation population genetics. *Briefings in Functional Genomics*, **9**, 416–423.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499–510.

Devitt TJ, Baird SJE, Moritz C (2011) Asymmetric reproductive isolation between terminal forms of the salamander ring species *Ensatina escholtzii* revealed by fine-scale genetic analysis of a hybrid zone. *BMC Evolutionary Biology*, **11**, 245.

Devitt TJ, Cameron Devitt SE, Hollingsworth BD, McGuire JA, Moritz C (2013) Montane refugia predict population genetic structure in the Large-blotched Ensatina salamander. *Molecular Ecology*, **22**, 1650–1665.

Dobzhansky T (1936) Studies on hybrid sterility. II. Localization of sterility factors in Drosophila pseudoobscura hybrids. *Genetics*, **21**, 113–135.

Dowling TE, Secor CL (1997) The role of hybridization and introgression in the diversification of animals. *Annual Review of Ecology and Systematics*, **28**, 593–619.

Durand EY, Patterson N, Reich D, Slatkin M (2011) Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution*, **28**, 2239–2252.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Eaton DAR, Ree RH (2013) Inferring phylogeny and introgression using RADseq data: an example from flowering plants (Pedicularis: Orobanchaceae). *Systematic Biology*, **62**, 689–706.

Emerson KJ, Merz CR, Catchen JM, et al. (2010) Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences, USA*, **107**, 16196–16200.

Endler JA (1973) Gene flow and population differentiation. *Science*, **179**, 243–250.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Foll M, Gaggiotti O (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–993.

Frost DR (2013) *Amphibian Species of the World: An Online Reference. Version 5.6*. American Museum of Natural History, New York, New York. Available from http://research.amnh.org/herpetology/amphibia/index.html (accessed on 9 January 2013).

Frost DR, Grant T, Favovich J *et al.* (2006) The amphibian tree of life. *Bulletin of the American Museum of Natural History*, **297**, 370.

Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology and Evolution*, **27**, 480–488.

Goldberg CS, Sullivan BK, Malone JH, Schwalbe CR (2004) Divergence among barking frogs (*Eleutherodactylus augusti*) in the Southwestern United States. *Herpetologica*, **60**, 312–320.

Gonzalez-Voyer A, Padial JM, Castroviejo-Fisher S, De La Riva I, Vilà C (2011) Correlated of species richness in the largest Neotropical amphibian radiation. *Journal of Evolutionary Biology*, **24**, 931–942.

Green RE, Krause J, Briggs AW *et al.* (2010) A draft sequence of the Neanderthal genome. *Science*, **328**, 710–722.

Hedges SB (1992) The number of replications needed for accurate estimation of the bootstrap p value in phylogenetic studies. *Molecular Biology and Evolution*, **9**, 366–369.

Hedges SB, Duellman WE, Heinicke MP (2008) New World direct-developing frogs (Anura: Terrarana): molecular phylogeny, classification, biogeography, and conservation. *Zootaxa*, **1737**, 182.

Hedrick PW (2013) Adaptive introgression in animals: examples and comparison to new mutation using standing variation as sources of adaptive variation. *Molecular Ecology*, **22**, 4606–4618.

Heinicke MP, Duellman WE, Hedges SB (2007) Major Caribbean and Central American frog faunas originated by ancient oceanic dispersal. *Proceedings of the National Academy of Sciences USA*, **104**, 10092–10097.

Heliconius Genome Consortium (2012) Butterfly genome reveal promiscuous exchange of mimicry adaptations among species. *Nature*, **487**, 94–98.

Hillis DM (1988) Systematics of the *Rana pipiens* complex: puzzle and paradigm. *Annual Review of Ecological, Evolution, and Systematics*, **19**, 39–63.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6**, e10000862.

Holman JA (2003) *Fossil Frogs and Toads of North America*. Indiana University Press, Bloomington, Indiana.

Jombart T (2008) ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jones JC, Fan S, Franchini P *et al.* (2013) The evolutionary history of *Xiphophorus* fish and their sexually selected sword: a

genome-wide approach using restriction site-associated DNA sequencing. *Molecular Ecology*, **22**, 2986–3001.

Kalinowski ST (2011) The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for humans population structure. *Heredity*, **106**, 625–632.

Kolbe JJ, Glor RE, Schettino LRG, Lara AC, Larson A, Losos JB (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.

Kronforst MR (2008) Gene flow persists millions of years after speciation in *Heliconius* butterflies. *BMC Evolutionary Biology*, **8**, 98.

Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences, USA*, **104**, 3883–3888.

Leaché AD (2009) Species tree discordance traces to phylogeographic clade boundaries in North American fence lizards (*Sceloporus*). *Systematic Biology*, **58**, 547–559.

Lischer HEL, Excoffier L (2012) PDGSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, **28**, 298–299.

Liu L, Yu L, Kubatko L, Pearl DK, Edwards SV (2009) Coalescent methods for estimating phylogenetic trees. *Molecular Phylogenetics and Evolution*, **53**, 320–328.

Lucek K, Roy D, Bezault E, Sivasundar A, Seehausen O (2010) Hybridization between distant lineages increases adaptive variation during a biological invasion: stickleback in Switzerland. *Molecular Ecology*, **19**, 3995–4011.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

Maddison WP (1997) Gene trees in species trees. *Systematic Biology*, **46**, 523–536.

Malone JH, Fontenot BE (2008) Patterns of reproductive isolation in toads. *PLoS One*, **3**, e3900.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

Martin SH, Dasmahapatra KK, Nadeau NJ *et al.* (2013) Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research*, **23**, 1817–1828.

McGuire JA, Linkem CW, Koo MS *et al.* (2007) Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of crotaphytid lizards. *Evolution*, **61**, 2879–2897.

Mecham JS (1959) Some Pleistocene amphibians and reptiles from Friesenhahn cave, Texas. *Southwestern Naturalist*, **3**, 17–27.

Meyer M, Kircher M, Gansauge MT *et al.* (2012) A high-coverage genome sequence from an archaic denisovan individual. *Science*, **338**, 222–226.

Monson KJ, Blouin MS (2004) Extreme isolation by distance in a montane frog *Rana cascadae*. *Conservation Genetics*, **5**, 827–835.

Moritz C (1994) Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology and Evolution*, **9**, 373–375.

Mueller RL (2006) Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. *Systematic Biology*, **55**, 289–300.

Mullen SP, Dopman EB, Harrison RG (2008) Hybrid zone origins, species boundaries, and the evolution of wing-pattern diversity in polytypic species complex of North American butterflies (Nymphalidae: *Limenitis*). *Evolution*, **2008**, 1–18.

Muller HJ (1939) Reversibility in evolution considered from the standpoint of genetics. *Biological Reviews of the Cambridge Philosophical Society*, **14**, 261–280.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, New York.

Nosil P, Gompert Z, Farkas TE *et al.* (2012) Genomic consequences of multiple speciation processes in a stick insect. *Proceeding of the Royal Society B*, **279**, 5058–5065.

Olson DM, Dinerstein E, Wikramanayake ED *et al.* (2001) Terrestrial ecoregions of the world: a new map for life on earth. *BioScience*, **51**, 933–938.

O'Neill EM, Schwartz R, Bullock CT *et al.* (2013) Parallel tagged amplicon sequencing reveals major lineages and phylogenetic structure in the North American tiger salamander (*Ambystoma tigrinum)* species complex. *Molecular Ecology*, **22**, 111–129.

Orr HA (1995) The population genetics of speciation: the evolution of hybrid compatabilities. *Genetics*, **159**, 1805–1813.

Paz-Vinas I, Quéméré E, Chikhi L, Blanchet S (2013) The demographic history of populations experiencing asymmetric gene flow: combining simulated and empirical data. *Molecular Ecology*, **22**, 3279–3291.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, **28**, 2537–2539.

Pérez-Figueroa A, García-Pereira MJ, Saura M, Rolán-Alvarez E, Canallero A (2010) Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology*, **23**, 2267–2276.

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

Prager EM, Wilson AC (1975) Slow evolutionary loss of the potential for interspecific hybridization in brids: a manifestation of slow regulatory evolution. *Proceedings of the National Academy of Sciences USA*, **72**, 200–204.

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

Prugnolle F, de Meeus T (2002) Inferring sex-biased dispersal from population genetic tools: a review. *Heredity*, **88**, 161–165.

R Core Team (2012) *R: A Language for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rambaut A, Drummond AJ (2007) *Tracer v1.5*. Available from http://beast.bio.ac.uk/Tracer.

Rannala B, Yang ZH (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, **164**, 1645–1656.

Rheindt FE, Edwards SV (2011) Genetic introgression: an integral but neglected component of speciation in birds. *The Auk*, **128**, 620–632.

Rheindt FE, Fujita MK, Wilton PR, Edwards SV (2014) Introgression and phenotypic assimilation in *Zimmerius* flycatchers (Tyrannidae): population genetics and phylogenetic inferences from genome-wide SNPs. *Systematic Biology*, **63**, 134–152.

Robertson JM, Zamudio KR (2009) Genetic differentiation, vicariance, and selection in a polytypic frog. *Journal of Heredity*, **100**, 715–731.

Rohland N, Reich D (2012) Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research*, **22**, 939–946.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **20**, 407–415.

Rothfels CJ, Larsson A, Kuo LY, Korall P, Chiou WL, Pryer KM (2012) Overcoming deep roots, fast rates, and short internodes to resolve the ancient rapid radiation of eupolypod II ferns. *Systematic Biology*, **61**, 490–509.

Rovito SM, Parra-Olea G, Vásquez-Almazán CR, Luna-Reyes R, Wake DB (2012) Deep divergence and extensive phylogeographic structure in a clade of lowland tropical salamanders. *BMC Evolutionary Biology*, **12**, 255.

Rubin BER, Ree RH, Moreau CS (2012) Inferring phylogenies from RAD sequence data. *PLoS One*, **7**, e33394.

Schwalbe CR, Goldberg CS (2005) Species account for *Eleutherodactylus augusti* (Dugés, 1879). In: *Amphibian Declines: The Conservation Status of United States Species* (ed. Lannoo M). University of California Press, Berkeley, California.

Sequeira F, Sodré D, Ferrand N *et al.* (2011) Hybridization and massive mtDNA unidirectional introgression between the closely related Neotropical toads *Rhinella marina* and *R. schneideri* inferred from mtDNA and nuclear markers. *BMC Evolutionary Biology*, **11**, 264.

Shaffer HB, McKnight ML (1996) The polytypic species revisited: Phylogenetics of the tiger salamander *Ambystoma tigrinum* (Amphibia: Caudata) complex. *Evolution*, **50**, 417–433.

Skoflund P, Jakobsson M (2011) Archaic human ancestry in East Africa. *Proceedings of the National Academy of Sciences, USA*, **108**, 18301–18306.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecological, Evolution, and Systematics*, **16**, 393–430.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.

Smith KL, Hale JM, Gay L *et al.* (2013) Spatio-temporal changes in the structure of an Australian frog hybrid zone: a 40-year perspective. *Evolution*, **67**, 3442–3454.

Spaulding AW, Mock KE, Schroeder MA, Warheit KI (2006) Recency, range expansion, and unsorted lineages: implications for interpreting neutral genetic variation in the sharp-tailed grouse (*Tympanuchus phasianellus*). *Molecular Ecology*, **15**, 2317–2332.

Stamatakis A (2014) RAxML version 8.0: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.

Stebbins GL (1950) *Variation and Evolution in Plants*. Columbia University Press, New York, New York.

Stölting KN, Nipper R, Lindtke D *et al.* (2013) Genomic scan for single nucleotide polymorphism reveals patterns of divergence and gene flow between ecologically divergent species. *Molecular Ecology*, **22**, 842–855.

Streicher JW, Crawford AJ, Edwards CW (2009) Multilocus molecular phylogenetic anlaysis of the montane *Craugastor podiciferus* species complex (Anura: Craugastoridae) in Isthmian Central America. *Molecular Phylogenetics and Evolution*, **53**, 620–630.

Streicher JW, Meik JM, Smith EN, Campbell JA (2011) Low levels of genetic diversity among morphologically distinct populations of an enigmatic montane frog from Mexico (*Craugastor uno*: Craugastoridae). *Amphibia-Reptilia*, **32**, 125–131.

Stuart BL, Inger RF, Voris HK (2006) High levels of cryptic species diversity revealed by sympatric lineages of Southeast Asian forest frogs. *Biology Letters*, **2**, 470–474.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary analysis using maximum likelihood, evolutionary distance, and maximum parsimony. *Molecular Biology and Evolution*, **28**, 2731–2739.

Van Devender TR, Rea AM, Smith ML (1985) The Sangamon interglacial vertebrate fauna from Rancho La Brisca, Sonora, Mexico. *Transactions of the San Diego Society of Natural History*, **21**, 23–55.

Wagner C, Keller J, Wittwer S *et al.* (2013) Genome wide RAD sequencing data provides unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, **22**, 787–798.

Walsh B (2001) Quantitative genetics in the age of genomics. *Theoretical Population Biology*, **59**, 175–184.

Wang IJ (2009) Fine-scale population structure in a desert amphibian: landscape genetics of the black toad (*Bufo exsul*). *Molecular Ecology*, **18**, 3847–3856.

Wilgenbusch JC, Warren DL, Swofford DL (2004) *AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference*. Available from http://ceb.csit.fsu.edu/awty.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

Yu H, Than C, Degnan JH, Nakhleh L (2011) Coalescent histories on phylogenetic networks and detection of hybridization despite incomplete lineage sorting. *Systematic Biology*, **60**, 138–149.

Zeisset I, Beebee TJ (2008) Amphibian phylogeography: a model for understanding historical aspects of species distributions. *Heredity*, **101**, 109–119.

Zweifel RG (1956) A survey of the frogs of the *augusti* group, genus *Eleutherodactylus*. *American Museum Novitates*, **1813**, 1–35.

J.W.S. and M.K.F. conceived and designed the project. Fieldwork was conducted by T.J.D., C.S.G., J.H.M., J.W.S. and M.K.F. Laboratory work was performed by J.W.S. and M.K.F. Analyses were performed by J.W.S., M.K.F., and H.B. The manuscript was written by J.W.S. with input from T.J.D., C.S.G., J.H.M., H.B., and M.K.F.

Data accessibility

Mitochondrial sequences have been deposited in GenBank (JX001714–48 and KF739349–85). A file containing the 12S mitochondrial alignment, SNP data matrices and STACKS commands has been deposited in Dryad: doi:10.5061/dryad.n420r.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Statistical procedures (log-likelihood plateaus, ΔK scores and *k*-means clustering) used to identify how many clusters are useful in explaining SNP variation across 20%, 25% and 50% missing data SNP matrices.

Fig. S2 F_{ST} score posterior distributions resulting from BAYESCAN analysis.

Fig. S3 Maximum-likelihood phylogram generated from RADtag data for the *Craugastor augusti* Species Series and a distantly related out-group, *Craugastor uno*.

Fig. S4 Isolation-by-distance plots for genome wide SNPs (Top) and mitochondrial DNA (Bottom) in *Craugastor augusti*.

Appendix S1 Voucher information for barking frog museum specimens used to construct map in Fig. 1.

Appendix S2 Voucher information for barking frogs from the United States (USA) and Mexico (MX) used in mitochondrial and SNP analyses.

Appendix S3 Barcode and index primers used to make adaptors and track chain of custody for ddRADseq *Craugastor* libraries sequenced for this study. Bolded text indicates hexamer barcode and index sequences.

Appendix S4 Additional laboratory notes and results.