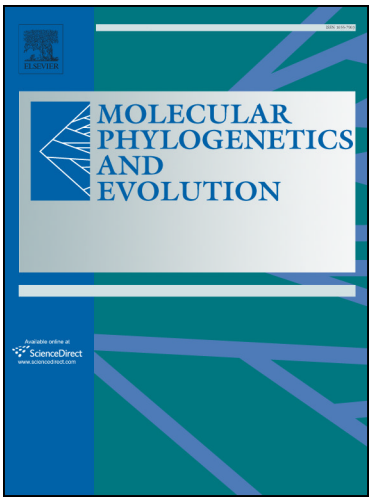


Phylogenomic evidence for a recent and rapid radiation of lizards in the Patagonian *Liolaemus fitzingerii* species group

Jared A. Grummer, Mariana M. Morando, Luciano J. Avila, Jack W. Sites Jr., Adam D. Leaché

PII: S1055-7903(17)30730-3
DOI: https://doi.org/10.1016/j.ympev.2018.03.023
Reference: YMPEV 6095

To appear in: *Molecular Phylogenetics and Evolution*

Received Date: 11 October 2017
Revised Date: 7 February 2018
Accepted Date: 15 March 2018

Please cite this article as: Grummer, J.A., Morando, M.M., Avila, L.J., Sites, J.W. Jr., Leaché, A.D., Phylogenomic evidence for a recent and rapid radiation of lizards in the Patagonian *Liolaemus fitzingerii* species group, *Molecular Phylogenetics and Evolution* (2018), doi: https://doi.org/10.1016/j.ympev.2018.03.023

Phylogenomic evidence for a recent and rapid radiation of lizards in the Patagonian *Liolaemus fitzingerii* species group

Jared A. Grummer[a,b,*], Mariana M. Morando[c], Luciano J. Avila[c], Jack W. Sites, Jr.[d], and Adam D. Leaché[a]

[a] *Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Box 351800, Seattle, WA 98195-1800, USA;*

[c] *Centro Nacional Patagónico – Consejo Nacional de Investigaciones Científicas y Técnicas (IPEEC-CONICET), Boulevard Almirante Brown 2915, ZC: U9120ACD, Puerto Madryn, Argentina;*

[d] *Department of Biology and M.L. Bean Life Science Museum, Brigham Young University, Provo, Utah 84602, USA;*

[b] *Current address: Biodiversity Research Centre and Beaty Biodiversity Museum, University of British Columbia, 6270 University Blvd., Vancouver, BC V6T 1Z4, Canada;*

[*] *Corresponding author: grummer@zoology.ubc.ca*

Abstract.— Rapid evolutionary radiations are difficult to resolve because divergence events are nearly synchronous and gene flow among nascent species can be high, resulting in a phylogenetic "bush". Large datasets composed of sequence loci from across the genome can potentially help resolve some of these difficult phylogenetic problems. A suitable test case is the *Liolaemus fitzingerii* species group of lizards, which includes twelve species that are broadly distributed in Argentinean Patagonia. The species in the group have had a complex evolutionary history that has led to high morphological variation and unstable taxonomy. We generated a sequence capture dataset for 28 ingroup individuals of 580 nuclear loci, alongside a mitogenomic dataset, to infer phylogenetic relationships among species in this group. Relationships among species were generally weakly supported with the nuclear data, and along with an inferred age of ∼2.6 million years old, indicate either rapid evolution, hybridization, incomplete lineage sorting, non-informative data, or a combination thereof. We inferred a signal of mito-nuclear discordance, indicating potential hybridization between *L. melanops* and *L. martorii*, and phylogenetic network analyses provided support for 5 reticulation events among species. Phasing the nuclear loci did not provide additional insight into relationships or suspected patterns of hybridization. Only one clade, composed of *L. camarones*, *L. fitzingerii*, and *L. xanthovirids* was recovered across all analyses. Genomic datasets provide molecular systematists with new opportunities to resolve difficult phylogenetic problems, yet the lack of phylogenetic resolution in Patagonian *Liolaemus* is biologically meaningful and indicative of a recent and rapid evolutionary radiation. The phylogenetic relationships of the *Liolaemus fitzingerii* group may be best be modeled as a reticulated network instead of a bifurcating phylogeny.
(Keywords: sequence capture, ultraconserved elements, coalescent, population, hybridization, Patagonia)

1.0 Introduction

Evolutionary radiations occur when one ancestral population diversifies into a variety of forms, typically over relatively short timescales, due to ecological opportunity or to evolutionary innovations (Schluter 2000; Glor 2010). However, non-adaptive radiations also occur, and these are also "evolutionary radiations". Rapid radiations are difficult to resolve because they are often characterized by incomplete lineage sorting (ILS), introgression, and few fixed differences between species (e.g., short internodes; Rokas and Carroll 2006, Patel et al. 2013). Resolving interspecific relationships in rapid radiations is important for accurate taxonomy, biogeography, trait evolution, and diversification studies.

Genomic scale datasets have become common for trying to resolve difficult phylogenetic problems because of reduced sequencing costs and recent developments in genome sequencing techniques (e.g. Baird et al. 2008; Faircloth et al. 2012; Lemmon et al. 2012; Peterson et al. 2012; Leaché et al. 2016). In addition to containing a large quantity of data for reconstructing phylogenies, genomic datasets also provide hundreds or thousands of independent estimates of the coalescent history across the genome, and therefore a better understanding of a group's evolutionary history. A common goal when trying to resolve rapid radiations is to collect and analyze more data (Rokas and Carroll 2006). However, more data will not help resolve "hard" polytomies, which result from near simultaneous divergence of many species; by definition, these cannot be resolved. Hard polytomies often characterize rapidly diversifying groups and can give the appearance of a bush rather than a tree. In contrast, "soft" polytomies are the result of analytical artifacts; these can be solved with the addition of more data or taxa, though this is not always successful (Maddison 1989; Olave et al. 2015). It is difficult to distinguish between hard and soft polytomies in rapid radiations because of the stochastic coalescent processes (e.g., incomplete lineage sorting) that cause a high degree of gene tree heterogeneity. In such cases, genomic datasets may not be able to resolve species-level relationships.

Sequence capture is a genomic data collection technique that targets specific

regions from across the genome, from tens to thousands of loci (McCormack et al. 2013). Because particular genomic regions are targeted, often something is known about the function or rate of evolution of those regions. Because the ability to sequence has proceeded faster than the ability to analyze large datasets, researchers are often faced with the challenge of finding an appropriate method for estimating a phylogeny from phylogenomic data. One common approach is to concatenate all loci together and analyze them together as one "supergene". However, simulation work has shown that concatenation can fail under certain circumstances and that it will provide increasing support for the wrong tree as more loci are added (Kubatko and Degnan 2007). Under certain demographic scenarios (e.g., population sizes and divergence times), the evolutionary history of some species is expected to be in the "anomaly zone", an area of tree space where the majority of gene tree topologies will not match the true species tree topology (e.g., Linkem et al. 2016). Multi-species coalescent methods attempt to model the independent coalescent histories among different loci, and therefore offer a more reliable alternative to concatenation (Yang and Rannala 2012; Edwards et al. 2016).

The impact of hybridization on species-level phylogenetic relationships under the multi-species coalescent model is in need of further exploration (but see Zhang et al. 2011, Leaché et al. 2013). Hybridization is common in nature with approximately 10% and 25% of animal and plant species known to hybridize, respectively (Mallet 2005). Whereas hybridization is often found to occur in limited geographic areas termed "contact" or "hybrid" zones (e.g. Barton and Hewitt 1985), hybridization is sometimes detected across broad areas of sympatry (e.g Martin et al. 2013). Nonetheless, it is difficult to document hybridization in remote geographic regions where the natural history of species is often understudied. Interspecific gene flow (e.g., hybridization) can result in the inferred phylogeny not matching the "true" phylogeny, but also distorts estimates of divergence times and population sizes (Leaché et al. 2013).

The genus *Liolaemus* (Squamata: Iguania: Liolaemidae) contains 250+ species distributed broadly across South America, and hybridization has been documented across several species including the *L. fitzingerii* species group (Morando et al. 2004;

Olave et al. 2011, 2017). The *L. fitzingerii* group is broadly distributed in coastal and Patagonian shrub-steppe habitats in central-southern Argentina (Fig. 1). This group is morphologically diverse, which has been the basis for many of the described species (e.g. Abdala et al. 2012b,a). Species range in maximum size (snout-vent length [SVL]) from 74.2 (*L. goetschi*) to 110mm (*L. fitzingerii*) (Abdala et al. 2012b,a), with sexual dichromatism absent in some species and evident in others. Unpublished morphological and molecular analyses have identified putative contact zones where individuals display intermediate patterning between parental species and mixing of mitochondrial parental haplotypes, both of which indicate localized hybridization.

Taxonomy of the *L. fitzingerii* group has been muddled since the 19th century when Charles Darwin incorrectly labeled the *L. fitzingerii* holotype as collected in "Chile", when in fact he collected this specimen in Puerto Deseado, Santa Cruz Province, Argentina (Cei 1980; Morando et al. 2004; Abdala 2007). Currently, twelve species are recognized in the *L. fitzingerii* group (Avila et al. 2006, 2008, 2010): five in the *fitzingerii* complex (*L. camarones*, *L. chehuachekenk*, *L. fitzingerii*, *L. shehuen*, and *L. xanthoviridis*), and 7 in the *melanops* complex (*L. casamiquelai*, *L. dumerili*, *L. goetschi*, *L. martorii*, *L. melanops*, *L. morenoi*, and *L. purul*). A fossil-calibrated analysis by Fontanella et al. (2012) determined the age of the *L. fitzingerii* species crown group to be 4.67 million years old. In slight contrast, unpublished analyses using a mutation rate of 0.019355 substitutions per site per million years calculated for the cytochrome B gene by (Olave et al. 2015) infer that the age of the *L. fitzingerii* group at ~2.6 million years old. A phylogeographic study performed by Avila et al. (2006) of the *L. fitzingerii* group recovered support for multiple range expansions, long-distance colonization events, secondary contact between described species in this group (*L. xanthoviridis* and *L. fitzingerii*), and species-level paraphyly within the larger *L. melanops* clade. Taken together, this information suggests a complex evolutionary history of range expansions, secondary contact, and possible hybridization, all of which occurred recently. To date, the *L. fitzingerii* group has not been the focus of an in-depth molecular-based phylogenetic study (but Olave et al. 2015 included

representatives of all species in the *L. fitzingerii* group in a sub-genus wide study).

In this study, we infer evolutionary relationships among species in the *L. fitzingerii* species group using a sequence capture dataset containing 585 loci and mitogenomic DNA. We sought to infer phylogenetic relationships to properly understand the evolutionary relationships among described species and candidate taxa in this group. To examine the impact of including putative hybrids on phylogenetic inference, we ran analyses with and without suspected hybrids. We analyze the data with multi-species coalescent approaches that account for ILS (e.g., BP&P [Yang 2015], SVDquartets [Chifman and Kubatko 2014]) in addition to a network approach that considers reticulate evolution (Than et al. 2008) to infer the evolutionary history of this group. Our results indicate that the *L. fitzingerii* species group evolved recently and then radiated rapidly. Furthermore, the inclusion of suspected hybrids did not affect the estimation of phylogenetic relationships.

2.0 Materials and Methods

2.1 Sampling

We performed sequence capture on all twelve species in the *L. fitzingerii* group (mentioned above) in addition to five individuals representing candidate species based on evidence for their potential status as distinct species (referred to as *Liolaemus* 16 – 19 and *L.* sp. Cona Niyeu; Olave et al. 2014), for a total of 28 ingroup individuals (1-4 individuals per species); sequence data from four ingroup samples were taken from a separate *Liolaemus*-wide phylogenetic study (Leaché et al., *in prep.*; Supplemental Table S1). Most individuals were assigned to species by geography (i.e., selecting individuals near type localities; Fig. 1). However, individuals collected further from type localities were assigned to species based on morphology. An additional five individuals were included because a study by Olave et al. (2014) provided evidence for their potential status as distinct species (referred to as *Liolaemus* 16 – 19 and *L.* sp. Cona Niyeu). Three geographically widespread species were represented by multiple

individuals (*L. fitzingerii*, *L. melanops*, and *L. xanthoviridis*), whereas all other lineages were represented by a single individual (Fig. 1; Supplemental Table S1). Four putative hybrid individuals were identified based on prior unpublished mtDNA and morphological analyses (*L. martorii* S, *L. melanops* C, S1, and S2; Fig. 1), and we performed all multi-species coalescent analyses with and without these suspected hybrids to examine how their inclusion affected results. All specimens were collected by hand in accordance with provincial permits from the Dirección de Fauna y Flora Slivestre and have been deposited into the LJAMM-CNP herpetology collection in the Centro Patagónico Nacional (IPEEC-CONICET), Puerto Madryn, Chubut, Argentina. Sequence data four other *Liolaemus* species (*L. bibronii*, *L. boulengeri*, *L. kingii*, and *L. rothi*) were used from Leaché et al. (*in prep.*) as outgroups for phylogenetic analyses (Supplemental Table S1). Sequence data from a single individual of *Liolaemus purul* were also included from Leaché et al. (*in prep.*) to test whether the placement of this recently described species in the *L. fitzingerii* species group based on morphological data (Abdala et al. 2012b) is also supported by the molecular phylogeny.

2.2 Sequence Capture Laboratory Protocol

We performed targeted sequence capture with a set of RNA probes specifically designed for Iguanian lizards (Leaché et al. 2015). We targeted 585 nuclear loci with a probe set that consisted of 1,170 RNA probes. Of the 585 targeted loci, 541 were from the Tetrapods-UCE-5Kv1 set (www.ultraconserved.org) and the remaining 44 were developed to capture loci from the Squamate Assembling the Tree of Life project (Wiens et al. 2012).

Genomic DNA was extracted from tissue (tail tips, liver) with either a Qiagen DNeasy blood and tissue extraction kit (Qiagen Inc., CA, USA) or NaCl extraction method (MacManes 2013). We used a Qubit fluorometer (Life Technologies, Carlsbad, CA) to measure DNA concentration of extracted samples and standardized to 400ng (nanograms) per sample. Genomic DNA was sheared to a target peak size of 400bp with a Bioruptor Pico (Diagenode Inc., Danville, NJ, USA). Library sequence

preparation was done with an Illumina TruSeq Nano kit (Illumina, San Diego, CA), and all cleanups in between steps were done with Ampure XP beads (Beckman Coulter Life Sciences, Indianapolis, IN). We first hybridized genomic DNA to the RNA probes, with a mixture of blocking probes consisting of TruSeq Nano forward and reverse complements, and then used chicken (Chicken Hybloc, Applied Genetics Lab Inc., Melbourne, FL) and salmon blockers to reduce the binding of repetitive DNA sequences; hybridization of RNA probes to genomic DNA lasted for 24 hours at 65°C. Following hybridization, libraries were enriched through 20 PCR cycles with TruSeq adapter primers and Phusion High-Fidelity DNA Taq polymerase (New England Biolabs Inc., Ipswich, MA). We quantified final libraries through quantitative PCR (qPCR) on an Applied Biosystems Step One Plus thermocycler (Applied Biosystems Inc., Foster City, CA) with probes that targeted five loci that are located on different chromosomes in the *Anolis carolinensis* genome. Final libraries were also quantified with an Agilent Tapestation 2200 (Agilent Technologies, Santa Clara, CA). All samples were pooled in equimolar ratios (based on qPCR results) and combined with 24 samples from other projects (a total of 48 individuals). Sequencing was performed on a single Illumina HiSeq 2500 lane (250bp paired-end, "Rapid run" mode) at the Vincent J. Coates QB3 Sequencing facility at UC Berkeley.

2.3 Bioinformatics and Dataset Assembly

We assembled a nuclear dataset consisting of phased alleles where each individual was represented by two alleles/haplotypes per locus. This dataset was assembled with a custom python pipeline (developed by Sonal Singhal, available at https://github.com/singhal/SqCL). We used Illumiprocessor and Trimmomatic (v0.36; Bolger et al. 2014) to remove adapters and barcodes, de-multiplex individuals, and remove low quality raw sequence reads (raw data stats can be found in Supplemental Table S1); clean reads were merged with PEAR (v0.9.10; Zhang et al. 2014). Reads were then assembled into contigs, per individual, in Trinity (v2.2.0; Grabherr et al. 2011). We then retained the assembled contigs that matched the 1170 probes (585 loci)

with BLAT (v36; Kent 2002). Next, we assembled pseudo-reference genomes (PRGs) for each species to be used in variant calling. If an individual's assignment to a species was ambiguous, we assigned that individual to its own "species". We then aligned the raw reads (for each individual) back to these PRGs to determine allelic variants with BWA (v0.7.12; Li and Durbin 2009), samtools (v1.3.1; Li et al. 2009), and Picard (v2.4.1; http://broadinstitute.github.io/picard/). GATK (v3.6; McKenna et al. 2010) was used to remove duplicates, identify SNPs and indels via standard hard filtering parameters and variant quality score recalibration according to best practices recommendations (Auwera et al. 2013). All bases, variant and invariant, were retained in the data matrix if they had ≥10x sequencing depth and a Phred quality score ≥20. SNPs were phased in relation to each other when paired reads spanned multiple variants, resulting in "blocks" of phased sequence that were hundreds of BPs long. With no good way to orient these phased blocks with respect to each other (e.g., long-range phasing), we oriented blocks randomly in relation to each other. Haplotypes were then combined by locus and then aligned in MAFFT (Katoh and Standley 2013). Resulting alignments were manually inspected one-by-one for poorly aligned ends and hand-edited as needed.

Mitochondrial ("mt") sequence data are often obtained as "by-catch", given that mitochondrial genomes are not targeted during library preparation, during sequence capture dataset sequencing. We used a pipeline developed by Alexander et al. (2017) and freely available on github (https://github.com/laninsky/Pulling-out-mitogenomes-from-UCE-data/) to assemble whole mitochondrial genomes for the individuals sequenced in this study. Briefly, we used NCBI BLAST (Altschul et al. 1990) and the mitochondrial genome of *Liolaemus chehuachekenk* (assembled into a single contig during *de novo* assembly and verified in NCBI BLAST) to serve as a reference library. We then performed a BLAST search of the Trinity contigs from each individual against the reference *L. chehuachekenk* genome at 75% similarity. The program seqtk (https://github.com/lh3/seqtk) was then used to extract the FASTA sequences of the contigs that matched the reference mt genome. A "sample-specific" mt genome was then generated for each individual, and contigs from

each individual were then searched against its own reference mt genome at 95% similarity to find any contigs we may have missed during the first search. We ran these last two steps iteratively (creating a sample-specific reference and BLASTing contigs to it) until no new contigs were found matching the reference genome. At that point, we used Geneious v10 (Biomatters; Auckland, New Zealand) to align these contigs to the reference *L. chehuachekenk* mt genome.

2.4 Phylogenetic Analyses

2.4.1 Multi-Species Coalescent Tree.— We inferred the species tree under the multi-species coalescent model (Rannala and Yang 2003; Yang and Rannala 2010) in the program BP&P v3.3 (Yang 2015). This Bayesian method does not account for gene flow and assumes gene tree discordance is due to ILS when estimating the species tree from sequence data. Individuals (and alleles) must be assigned to species before analysis, and we did so based on expert identification and the current taxonomy. Putative hybrids were conservatively identified (e.g., any suspected as hybrids based on previous morphological and mtDNA data), and assigned to their own lineage. Gene flow is a clear violation of the assumptions of many phylogenetic inference programs, so we ran two sets of analyses: one set including putative hybrids assigned to their own lineage, and the second set with putative hybrid individuals removed.

Two parameters must be specified by the user with priors in BP&P – θ and τ – which correspond to population sizes and divergence times, respectively. Note that to estimate θ, a minimum of two sequences per "species" is needed. We specified two different combinations of θ and τ priors to ensure results were stable, and conducted four replicates of each analysis. One set of analyses used a gamma prior G(5, 1000) on θ, giving a mean value of $5/1000 = 0.005$, with a gamma prior G(5, 2000) on τ, or a mean of 0.0025. These priors were based on the average pairwise sequence distances that we calculated across 40 loci with the highest variation in our dataset (e.g., ~1% sequence divergence within a locus). The second set used G(2, 200) for θ and G(2, 400)

for τ, representing larger population sizes and longer time between population divergences. We ran species tree analyses on two datasets, both with and without suspected hybrids, with a burn-in of 25,000 generations and post burn-in of 100,000 generations. Convergence was assessed by examining posterior estimates of θ, τ, and topological consistency across independent runs.

2.4.2 SVDquartets.— A new class of multi-species coalescent-based species tree estimation algorithms was recently designed, which do not utilize summary statistics nor gene trees, but rather infers a topology based on 4-taxon relationships inferred through site patterns (e.g., SNPs; Chifman and Kubatko 2014, Chifman and Kubatko 2015). The uncertainty in species-level relationships can then be quantified through non-parametric bootstrapping. This method is implemented in the program SVDquartets (through PAUP; Swofford 2003) and can be performed in seconds (inferring just the tree) or minutes (bootstrapping) on a standard desktop computer. Individuals/alleles were assigned to species as in the BP&P analyses. We inferred the species tree in SVDquartets with and without hybrids, evaluating all possible quartets with 100 bootstrap replicates to assess uncertainty in species-level relationships.

2.4.3 Concatenation.— We concatenated all nuclear loci and inferred a tree for this "super matrix" in RAxML v8.2 (Stamatakis 2014) with the GTR + Γ DNA substitution model with 100 bootstrap iterations. For each individual, all "1" alleles were concatenated together across loci, as were the "2" alleles, resulting in two "super alleles" per individual in the concatenated tree. We do not know the phase of each allele with respect to the alleles at the other loci, so the concatenation of alleles across loci is arbitrary.

2.4.4 Mitogenomic Tree.— We inferred the mitochondrial phylogeny from whole mitochondrial genomic alignments in BEAST v2.4.5 (Bouckaert et al. 2014). PartitionFinder2 (Lanfear et al. 2016) was used to determine the optimal partitioning scheme with a "greedy" search and BIC selection criterion. The analysis was run for 5 x 10^7 generations, with a burn-in of 10^7 generations. Stationarity was assessed in Tracer

v1.6 (Rambaut A and AJ 2014), where all parameters had effective sample size (ESS) values >200.

2.5 Testing for Hybridization

We used four methods to test for hybridization due to mito-nuclear discordance (see Results) and high morphological variation in restricted geographic areas. First, we used a network approach to infer the evolutionary history of this group with Phylonet (Than et al. 2008). This method requires gene trees for input, so we used jModelTest v2.1.7 (Guindon and Gascuel 2003; Darriba et al. 2012) on each alignment (including outgroup data) to infer the appropriate DNA substitution model based on the Bayesian Information Criterion. Gene trees were then inferred in RAxML v8.2 (Stamatakis 2014) with the top-ranking DNA substitution model and 100 bootstrap (BS) iterations for each locus, with sequence data for *Liolaemus rothi* rooting all gene trees. To mitigate alignment errors, we examined each gene tree for long branches and hand-checked dubious alignments. We also used these gene trees for detecting hybrids (see below). As in many "species tree" analysis programs, Phylonet requires that individuals must be assigned to species, so we based our assignments on current taxonomy and expert identification. Furthermore, the user specifies the number of reticulation events in the phylogeny to infer, which we explored for a range (0-5) of reticulation events. We were unable to explore >5 reticulation events because of exceeding computation wall time limits (40 days). Due to computational costs, we inferred each network under maximum pseudo-likelihood (MPL), with five replicates per analysis. We determined the best-fitting network through AIC model selection (Akaike 1998; Sullivan and Joyce 2005), where the number of free parameters (k) was the sum of internal branches, including the number of reticulations (Y. Yu, pers. comm.).

Secondly, we used a technique developed by Joly et al. (2015) that calculates genetic distances among individuals with SNPs. Using simulations, Joly et al. (2015) showed that these distances identify hybrids that are genetically intermediate between two parental species. The expectation is that a perfectly intermediate hybrid will have a

genetic distance ("I") of 0.5, where $I = \frac{D_{AX}}{(D_{AX}+D_{BX})}$; A and B are the parent species, X is the suspected hybrid, and D_{AX} is the genetic distance between parent A and the hybrid. To generate a random distribution of I values with which to compare the suspected hybrids, we assigned random trios of individuals as parents and hybrid. This distribution will generate an expectation of the average distance among any three individuals, thus providing a background set of I values with which to compare the suspected hybrids. We then compared I values of the suspected hybrids (3 *L. melanops* and 1 *L. martorii* individual) to this background "null" distribution. Joly et al. (2015) showed Nei's distance to be the most accurate at inferring hybrids, so we therefore calculated Nei's distance to infer hybrid individuals.

Third, we tested for putative hybrids through a discriminant analysis of principal components of genetic data in the R package Adegenet (Jombart et al. 2010; Jombart and Ahmed 2011). For this, we used all variable sites (12,651) and not just unlinked single nucleotide polymorphisms (SNPs). Hybrid individuals should fall outside the cluster (in PCA-space) of their parental species (when multiple individuals per species are available), and more specifically, in between (in PCA-space) parental species.

And finally, we used a qualitative approach via inspection of gene trees. With resolved and supported gene trees, putative hybrids can be identified based on distinct placement of their two alleles into divergent parental clades. We therefore searched all gene trees for divergent allelic placement of suspected hybrid individuals.

3.0 Results

3.1 Alignments

Alignment summaries (created by scripts from Portik et al. 2016), including the number of taxa, alignment lengths, number and percent of informative sites, and percent of gaps and missing data, were generated for datasets both with and without outgroup data and can be found in Table 1 and Supplemental Figures S1-2. Sequence data were poor for the outgroups *Liolaemus bibronii* and *L. kingii*, in addition to the

ingroup sample for *L. canqueli*, and therefore were not included in phylogenetic analyses (Supplemental Table S1). The final dataset therefore consisted of 27 ingroup individuals (including *L. purul*) and two outgroup individuals. We recovered 580 loci with > 75% taxon coverage per locus (Supplemental Table 1). On average, alignments are 510bp with 11.2 parsimony-informative sites per locus for the ingroup taxa (Fig. 2; Supplemental Fig. S2). The best-fit models of sequence evolution for each locus can be found in Supplemental Table S2.

3.2 Multi-Species Coalescent Tree

The monophyly of the *L. fitzingerii* species group is strongly supported with a posterior probability (pp) value of 1.0, with *L. purul* diverging first subsequent to outgroup taxa (Fig. 3; Supplemental Fig. S3). Nevertheless, relationships among species within this group are poorly supported. The τ prior had a noticeable impact on branch lengths, with shorter branches for trees estimated with larger prior mean values (Supplemental Fig. S3). However, inferred θ estimates were similar regardless of the prior values. One clade (*xanthoviridis*,(*fitzingerii*,*camarones*)) was consistently and strongly (pp≥0.95) recovered in both analyses. Also, *L. goetschi* and *L. martorii* are inferred as early diverging species with both datasets. Although placement for some taxa changed with the trees estimated with different priors (e.g., *L. dumerili* and *L.* sp. 19), none of the topological differences were strongly supported. Relationships did not significantly change when putative hybrid taxa were removed (Supplemental Fig. S4).

3.3 SVDquartets

In general, the trees inferred with SVDquartets are similar to those from BP&P, in terms of both support and topology (Fig. 3), and no significant topological differences resulted from including putative hybrids (Supplemental Fig. S5). Relationships among most species were poorly supported, with the northern species *L. goetschi*, *L.* sp. 17, and *L. martorii* diverging early from other species, and the southern (*xanthoviridis*,(*fitzingerii*,*camarones*)) clade strongly supported with both datasets.

3.4 Concatenation

The length of all loci combined was 297,000bp. *Liolaemus purul* was inferred to be sister to all other *L. fitzingerii* group species (Supplemental Fig. S6). Both "1" and "2" alleles within each individual were strongly supported as sister to each other, with the exception of *L. fitzingerii* N and *L. fitzingerii* Isla Leones; alleles from these individuals formed weakly supported relationships (BS <70) inter-digitated with each other (Supplemental Fig. S6). Individuals from the widespread species *L. melanops* form a strongly supported clade (BS = 100). The recently described *Liolaemus camarones* (Abdala et al. 2012a) was recovered within *L. fitzingerii*, rendering the latter taxon paraphyletic. The inclusion of putative hybrid individuals did not change overall support values (results not shown), maintaining generally low BS values across the tree; generally, suspected hybrids formed clades with geographically proximate individuals (except *L. martorii* S sister to *L. morenoi*).

3.5 mtDNA Phylogeny

The percent of the entire mt genome sequenced ranged across individuals from 38 to 89, or 6616 to 15379bp, with an average of 78% complete or 13,480bp (Supplemental Table S3). Seven partitions were selected, and their compositions and model choice can be found in Supplemental Table S4. Monophyly of the *L. fitzingerii* group is supported, with *L. purul* forming a clade with the outgroup taxa *L. boulengeri* and *L. rothi*. Within the *L. fitzingerii* species group, many relationships were supported with a posterior probability of 1.0, with only a single relationship receiving support >0.95 (Fig. 4; Supplemental Fig. S7). In general, clades were composed of geographically cohesive groups, with the exception of *L.* sp. 16 (sample #20) forming a clade with individuals much farther to the east. A clade of ((*L. fitzingerii*, *L. camarones*), *L. xanthoviridis*) was inferred with the mtDNA data, which matches the concatenated nDNA tree. However, some notable differences are evident between the mt- and nDNA concatenated phylogenies. First, *L. camarones* is sister to *L. fitzingerii* (based on a single *L. camarones* sample), vs. within *L. fitzingerii* as in the

concatenated nDNA tree. Second, the monophyly of *L. melanops* is not supported in the mtDNA tree. Interestingly, the individuals that have highly different placement between the mt- and nDNA trees map to phylogeographic clade boundaries of the mtDNA tree (Fig. 4). Similarly, the southern *L. martorii* sample is placed with *L. melanops* individuals, distant in the tree from the northern *L. martorii* individual.

3.6 Hybridization Detection

Via AIC model selection, the best-fitting network model included five reticulation events (Table 2; Fig. 5). However, many internodes between species were very short. Although the (*L. xanthoviridis*, (*L. fitzingerii*, *L. camarones*)) clade was not recovered in this network, those taxa were related by genomic inheritance from inferred ghost lineages. Two other reticulation events were inferred between *L. melanops* and suspected hybrids of *L. melanops* and *L. shehuen*. The final reticulation was inferred between *L. sp* 17 and the common ancestor of a large clade of many *L. fitzingerii* group species.

The background distribution of I calculations showed a somewhat bimodal distribution, with a large spike at $\sim$0.5 (Supplemental Fig. S8). The three suspected *L. melanops* hybrids had I values of 0.54 – 0.57, whereas the suspected *L. martorii* S hybrid had an I value of 0.38. Given that these values fall into the middle of the background distribution, this method did not detect hybrids with confidence.

Adegenet analyses provided evidence that the suspected *L. martorii* hybrid ("*L. martorii* S") is a hybrid. The specimen is inferred to be intermediate (in PCA-space) between its two suspected parental species (*L. martorii* and *L. melanops*; Supplemental Fig. S9). The three individuals sampled from a suspected hybrid zone between *L. melanops* and *L. shehuen* fall outside the space that encompasses the genetic diversity of *L. melanops* (Supplemental Fig. S9). However, these individuals do not lie between their suspected parental species. We took a conservative approach and treated these individuals as hybrids and performed all analyses both with and without them to ensure the stability of the phylogenetic results (which they were).

Regarding gene trees, the two most frequent models of DNA substitution were F81 and HKY85 (with or without I and/or Γ; Supplemental Tables S2). Resolution was low with very few well-supported clades within each gene tree, so we could not identify hybrids via placement of alleles in disparate clades.

4.0 Discussion

One might expect that morphologically divergent species would be genetically differentiated as well. However, in spite of the high level of morphological diversity seen in the *Liolaemus fitzingerii* group, this study showed that many of the relationships among species were poorly supported and that their history might best be modeled as a reticulated network. A comparison of n- and mtDNA phylogenies revealed strong discordance in terms of phylogenetic placement of certain individuals, and these individuals occur at phylogeographic clade boundaries (Fig. 4), suggesting introgression as the cause of this discordance (Funk and Omland 2003; Leaché 2009). However, two methods that we used specifically to detect hybrids lacked the power to support this hypothesis. These results suggest that the *L. fitzingerii* species group underwent a rapid radiation and that the lack of phylogenetic support is due to hybridization and/or insufficient information/variation present in the data to resolve phylogenetic relationships. The only clade consistently recovered was that of the southern-most species – *L. xanthoviridis*, *L. fitzingerii*, and *L. camarones*.

4.1 Resolving Rapid Evolutionary Radiations

Evolutionary radiations generally follow the evolution of morphological novelties or the availability of novel ecological niches in a particular environment, and are therefore inferred to be adaptive (Schluter 2000). Many radiations from an ancestral form are rapid. When this happens, the resulting phylogenetic pattern will approximate a "star" phylogeny, characterized by either short or non-existent internal nodes. For such radiations, estimating relationships among lineages is difficult at best. Many

simulation studies have shown that dozens or even thousands of loci are needed to obtain correct/accurate phylogenetic estimates (e.g. Liu et al. 2009). In this study, however, even a dataset of 580 loci cannot provide significant support for interspecific relationships in the *L. fitzingerii* species group.

One impediment to estimating a resolved phylogeny is homoplasy, which obscures the signal of ancient divergences that even model-based approaches fail to recover (e.g. Dopazo and Dopazo 2005). Rare genomic changes (RGCs), such as insertion-deletion events (particularly in coding regions), can be particularly informative for resolving ancient rapid radiations (e.g. Venkatesh et al. 2001; Murphy et al. 2007; King and Rokas 2017), but are more difficult to employ with younger radiations where these characters have either not evolved, or if they have, have not sorted by species. However, some research has shown that ultra-conserved elements are less prone to homoplasy than nuclear introns (and mitochondrial DNA; Meiklejohn et al. 2016). Homoplasy is not likely to be an issue for generating incongruent phylogenetic signals in a young radiation such as the *L. fitzingerii* group. A second factor responsible for failure to recover a well-supported phylogeny is the lack of phylogenetic signal in a dataset. Internal nodes exist because of shared nucleotide changes across descendent taxa, and in the case of a rapid radiation, little time exists for these stochastically evolved characters to sort to species (Rokas and Carroll 2006). Given the paucity of these changes, obtaining data from as much of the genome as possible will increase the odds of including the few characters that provide phylogenetic resolution.

It might be argued that using sequence capture datasets composed of "ultra-conserved elements" at shallow levels (e.g., population and inter-species studies) is ill-advised because these loci were developed to match genomic regions that have been conserved across deep evolutionary time (tens to hundreds of millions of years). However, some authors (e.g. Harvey et al. 2016) have shown that UCEs are useful in population-level studies. In addition, we included 44 loci that were developed for the Squamate Assembling the Tree of Life project (Wiens et al. 2012), which had higher levels of variation (Supplemental Table S5). The level of genetic variation and

informativeness of our dataset puts this species group in the realm of other study systems that did produce resolved phylogenies (Smith et al. 2014). Therefore, the incompletely resolved phylogeny of this group probably does not reflect limited genetic variation in the data. Nonetheless, an unresolved phylogeny based on a substantial dataset provides an important signal of evolutionary history of the focal group (Hoelzer and Meinick 1994; Rokas and Carroll 2006).

4.2 Detecting Hybridization with Sequence Data

Sequence data can effectively detect hybrids, particularly when viewed in a phylogenetic perspective. Based on unpublished morphological and mitochondrial analyses, we hypothesized that some individuals in this study were of hybrid origin. Because rapid radiations show short internodes, distinguishing between ILS and hybridization is difficult (Holder et al. 2001). Alternatively, when parent species are well-differentiated and belong to independent clades, the alleles of hybrid individuals are readily recovered in the two different clades (e.g. Leaché and McGuire 2006; Alexander et al. 2017). Furthermore, when an entire species/population is of hybrid origin, or when hybrid individuals are represented by a single consensus genotype (e.g., not phased alleles), phylogenetic support values will be reduced (due to the ambiguous placement of the admixed genotypes/individuals); this fact has been formalized into software that detects hybrids (Schneider et al. 2016). The placement of most suspected hybrids in the concatenated tree was strong with BS >60. We did not observe significant changes in bootstrap values when removing putative hybrid individuals from the dataset. In a related context, network approaches such as Phylonet seem promising for detecting hybridization events, because the majority of inferred reticulation events in the dataset corroborated independent hypotheses based on unpublished morphological and mtDNA analyses of hybridization in those individuals.

Another popular method for estimating gene flow with sequence data is via an isolation-migration model such as that implemented in IMa2 (Hey 2010). This method requires an input topology of species-level relationships, rendering it difficult to

implement when interspecific relationships are poorly supported, as is the case in the *L. fitzingerii* group. Thus, it was not possible to implement this method to test for gene flow with this method, so we sought to identify hybrids via variable sites alone – SNPs. The first approach we took calculated genetic distances among individuals based on phased SNPs; simulations showed that this approach can detect hybrids even with as few as tens of SNPs (Joly et al. 2015). However, these simulations were based on an allopolyploidization event between parental species that diverged 30,000 generations in the past (τ=0.003). The BP&P results indicate much shallower divergences for species in the *L. fitzingerii* group (τ<<0.001), providing little time for genetic drift or other evolutionary processes to generate differences between putative parental species. Morphologically, the parental *L. martorii* and *L. melanops* species differ in body size by ~15-20mm (*L. martorii* being smaller) as well as dorsal patterning (Abdala 2003). Putative *L. fitzingerii* group hybrids had I values in the 0.4 – 0.5 range (results not shown), which fell in the middle of the range of the randomized I distribution. This signifies that the genomes of many individuals/species in the *L. fitzingerii* group are equally/distantly divergent from one another, rendering hybrid detection difficult. It is possible, though not likely, that the *L. fitzingerii* group "species" actually represent a single, widespread panmictic species with a high level of phylogeographic structuring.

4.3 Systematics of the *Liolaemus fitzingerii* Species Group

The taxonomy of the *L. fitzingerii* group is particularly complex. Whereas some species have been described based on both molecular (generally mtDNA) and morphological characters (e.g., *L. chehuachekenk*, Avila et al. 2008; *L. casamiquelai*, Avila et al. 2010), other species have been described solely based on morphological characters (e.g., *L. dumerili* and *L. purul*, Abdala et al. 2012a; *L. camarones* and *L. shehuen*, Abdala et al. 2012b). Some of these characters are related to color patterning and melanism, the latter of which was shown to be uninformative for delimiting species in this group (Escudero et al. 2012). Relationships inferred from mtDNA and morphological characters are in stark contrast to one another (e.g., this study and Avila

et al. 2006; Abdala et al. 2012a and Abdala et al. 2012b). External morphological characters such as color and pattern are highly variable within species, and melanism, a character used in the diagnosis of many *L. fitzingerii* group species, varies ontogenetically between males and females (Escudero et al. 2016). An in-depth species delimitation analysis with finer-scale sampling would be necessary to fully test the species-level status of both described and undescribed taxa in the *Liolaemus fitzingerii* group.

Based on a fossil calibration applied to a combined n- and mtDNA dataset, Fontanella et al. (2012) inferred the date of the *L. fitzingerii* species crown group at 4.67 million years ago (mya). Based on a molecular clock rate of 1.9355% sequence divergence per million years for the *cyt B* locus that was calculated in Olave et al. (2015) (see their Table 2), we estimated an age of 2.55 million years (1.9 – 3.17mya 95% HPD) for the *L. fitzingerii* group (unpublished results). Despite the discrepancy in these estimates, both results confirm the young age of the *L. fitzingerii* group. The phylogenetic analyses showed *Liolaemus purul* as sister to the remaining *L. fitzingerii* group species (Fig. 3). Whether or not this species is a part of the *L. fitzingerii* group is ambiguous, as it could either be the earliest diverging member of the clade, or sister to the *L. fitzingerii* species group. Sampling other outgroup species that are close relatives of th *L. fitzingerii* group should provide more conclusive results in future studies of this group. Another consistent relationship inferred was the monophyly of the (*L. camarones* + *L. fitzingerii* + *L. xanthoviridis*) clade. These are the three southern-most taxa in the group and have low genetic diversity estimates, potentially indicative of post-glacial range expansions. This hypothesis is being tested through demographic analyses with SNP data (Grummer et al., *in prep.*).

A comparable amount of genetic variation seen in the *L. fitzingerii* species group has been found in other Squamate systems characterized by both multiple species with clear-cut boundaries as well as systems within which only a single species is recognized. For instance, the *Uma scoparia* and *Uma notata* complex had an average 11.2 segregating sites across 14 nuclear loci (Gottscho et al. 2014). Jackson and Austin

(2010) reported a similar diversity with an average of 14.1 parsimony-informative sites across seven nuclear loci (after removing the outlier locus "SELT") in the widespread and morphologically conserved eastern North American skink species *Scincella lateralis*. And lastly, more genetic variation exists across the *L. fitzingerii* species group than across 15 other *Liolaemus* species with the same loci (Panzera et al. 2017). The high phenotypic diversity seen in the *L. fitzingerii* group led to many species being described solely on external characteristics with little regard to molecular-based estimates of diversity and relationships. The level of molecular diversity we see in the *L. fitzingerii* species group is similar to other lizard species "complexes" where one to a few species are recognized. Thus, species in the *L. fitzingerii* group appear to be "over-split" in relation to other similar Squamate systems.

5.0 Conclusions

Our phylogenomic analyses support a rapid radiation in the *Liolaemus fitzingerii* species group. The conflicting set of relationships inferred between mt- and nDNA datasets, in particular with individuals at clade boundaries, strongly suggests a history of hybridization. The Patagonia region of South America that this group inhabits is characterized by a complex geologic and climatic history that has created many opportunities for range expansions and contractions that would facilitate hybridization (Sersic et al. 2011). Few phylogenetic relationships were well-supported, yet this information is important for understanding the evolutionary history of the *Liolaemus fitzingerii* species group. In fact, rapid radiations and hard polytomies may be common in the subgenus *Eulaemus* that the *L. fitzingerii* species group belongs to (Olave et al. 2015). Our results provide a phylogenetic hypothesis and historical context for understanding the evolutionary processes that gave rise to diversity in this species group.

Acknowledgments

This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. This research was funded in part by a National Science Foundation Doctoral Dissertation Improvement Grant (DEB-1500933) to JAG, and was facilitated though the use of advanced computational, storage, and networking infrastructure provided by the Hyak supercomputer system at the University of Washington. Special thanks to Sonal Singhal for help with the sequence capture bioinformatic pipeline, and to Yun Yu and Luay Nakhleh for assistance with running Phylonet.

Data Accessibility

Final aligned sequence data are available through GenBank, accession nos. #-#.

Conflict of Interest

We the authors declare no competing interests.

*

References

Abdala, C. S. 2003. Cuatro nuevas especies del género liolaemus (iguania: Liolaemidae), pertenecientes al grupo boulengeri, de la patagonia, argentina. Cuadernos de Herpetología 17.

Abdala, C. S. 2007. Phylogeny of the boulengeri group (iguania: Liolaemidae, liolaemus) based on morphological and molecular characters. Zootaxa Pages 1–84.

Abdala, C. S., J. Díaz Gómez, and V. Juarez Heredia. 2012a. From the far reaches of patagonia: new phylogenetic analyses and description of two new species of the *Liolaemus fitzingerii* clade (iguania: Liolaemidae). Zootaxa 3301:34–60.

Abdala, C. S., R. Semhan, D. M. Azocar, M. Bonino, M. Paz, and F. Cruz. 2012b. Taxonomic study and morphology based phylogeny of the patagonic clade *Liolaemus melanops* group (iguania: Liolaemidae), with the description of three new taxa. Zootaxa 3163:1–32.

Akaike, H. 1998. Information theory and an extension of the maximum likelihood principle. Pages 199–213 *in* Selected Papers of Hirotugu Akaike. Springer.

Alexander, A. M., Y.-C. Su, C. H. Oliveros, K. V. Olson, S. L. Travers, and R. M. Brown. 2017. Genomic data reveals potential for hybridization, introgression, and incomplete lineage sorting to confound phylogenetic relationships in an adaptive radiation of narrow-mouth frogs. Evolution 71:475–488.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. Journal of molecular biology 215:403–410.

Auwera, G. A., M. O. Carneiro, C. Hartl, R. Poplin, G. del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, et al. 2013. From fastq data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Current protocols in bioinformatics Pages 11–10.

Avila, L., M. Morando, and J. Sites. 2006. Congeneric phylogeography: hypothesizing species limits and evolutionary processes in patagonian lizards of the liolaemus boulengeri group (squamata: Liolaemini). Biological Journal of the Linnean Society 89:241–275.

Avila, L. J., M. Morando, and J. W. Sites Jr. 2008. New species of the iguanian lizard genus liolaemus (squamata, iguania, liolaemini) from central patagonia, argentina. Journal of Herpetology 42:186–196.

Avila, L. J., C. H. F. Pérez, M. Morando, and J. Sites Jr. 2010. A new species of liolaemus (reptilia: Squamata) from southwestern rio negro province, northern patagonia, argentina. Zootaxa 2434:47–59.

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, W. A. Cresko, and E. A. Johnson. 2008. Rapid snp discovery and genetic mapping using sequenced rad markers. PloS one 3:e3376.

Barton, N. H. and G. M. Hewitt. 1985. Analysis of hybrid zones. Annual review of Ecology and Systematics 16:113–148.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics Page btu170.

Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C.-H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond. 2014. Beast 2: A software platform for bayesian evolutionary analysis. PLOS Computational Biology 10:1–6.

Cei, J. M. 1980. L'identité des syntypes de proctotretus fitzingeri duméril et bibron. Bulletin Museum National d'Historie Naturelle (Paris) 4:317–320.

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

Chifman, J. and L. Kubatko. 2015. Identifiability of the unrooted species tree topology

under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. Journal of theoretical biology 374:35–47.

Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jmodeltest 2: more models, new heuristics and parallel computing. Nature methods 9:772–772.

Dopazo, H. and J. Dopazo. 2005. Genome-scale evidence of the nematode-arthropod clade. Genome biology 6:R41.

Edwards, S. V., Z. Xi, A. Janke, B. C. Faircloth, J. E. McCormack, T. C. Glenn, B. Zhong, S. Wu, E. M. Lemmon, A. R. Lemmon, et al. 2016. Implementing and testing the multispecies coalescent model: a valuable paradigm for phylogenomics. Molecular phylogenetics and evolution 94:447–462.

Escudero, P., I. Minoli, M. González Marín, M. Morando, and L. Avila. 2016. Melanism and ontogeny: a case study in lizards of the liolaemus fitzingerii group (squamata: Liolaemini). Canadian Journal of Zoology 94:199–206.

Escudero, P. C., I. Minoli, N. Frutos, L. J. Avila, and M. Morando. 2012. Estudio comparativo del melanismo en lagartijas del grupo liolaemus fitzingerii (liolaemini: Liolaemus). Cuadernos de herpetología 26:79–89.

Faircloth, B. C., J. E. McCormack, N. G. Crawford, M. G. Harvey, R. T. Brumfield, and T. C. Glenn. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic biology Page sys004.

Fontanella, F. M., M. Olave, L. J. Avila, M. MORANDO, et al. 2012. Molecular dating and diversification of the south american lizard genus liolaemus (subgenus eulaemus) based on nuclear and mitochondrial dna sequences. Zoological Journal of the Linnean Society 164:825–835.

Funk, D. J. and K. E. Omland. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial dna. Annual Review of Ecology, Evolution, and Systematics 34:397–423.

Glor, R. E. 2010. Phylogenetic insights on adaptive radiation. Annual Review of Ecology, Evolution, and Systematics 41:251–270.

Gottscho, A. D., S. B. Marks, and W. B. Jennings. 2014. Speciation, population structure, and demographic history of the mojave fringe-toed lizard (uma scoparia), a species of conservation concern. Ecology and evolution 4:2546–2562.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, et al. 2011. Trinity: reconstructing a full-length transcriptome without a genome from rna-seq data. Nature biotechnology 29:644.

Guindon, S. and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic biology 52:696–704.

Harvey, M. G., B. T. Smith, T. C. Glenn, B. C. Faircloth, and R. T. Brumfield. 2016. Sequence capture versus restriction site associated dna sequencing for shallow systematics. Systematic biology 65:910–924.

Hey, J. 2010. Isolation with migration models for more than two populations. Molecular biology and evolution 27:905–920.

Hoelzer, G. A. and D. J. Meinick. 1994. Patterns of speciation and limits to phylogenetic resolution. Trends in ecology & evolution 9:104–107.

Holder, M. T., J. A. Anderson, and A. K. Holloway. 2001. Difficulties in detecting hybridization. Systematic Biology 50:978–982.

Jackson, N. D. and C. C. Austin. 2010. The combined effects of rivers and refugia generate extreme cryptic fragmentation within the common ground skink (*Scincella lateralis*). Evolution 64:409–428.

Joly, S., D. Bryant, and P. J. Lockhart. 2015. Flexible methods for estimating genetic distances from single nucleotide polymorphisms. Methods in Ecology and Evolution 6:938–948.

Jombart, T. and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide snp data. Bioinformatics 27:3070–3071.

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

Katoh, K. and D. M. Standley. 2013. Mafft multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution 30:772–780.

Kent, W. J. 2002. Blat—the blast-like alignment tool. Genome research 12:656–664.

King, N. and A. Rokas. 2017. Embracing uncertainty in reconstructing early animal evolution. Current Biology 27:R1081–R1088.

Kubatko, L. S. and J. H. Degnan. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Systematic Biology 56:17–24.

Lanfear, R., P. B. Frandsen, A. M. Wright, T. Senfeld, and B. Calcott. 2016. Partitionfinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Molecular biology and evolution 34:772–773.

Leaché, A. D. 2009. Species tree discordance traces to phylogeographic clade boundaries in north american fence lizards (sceloporus). Systematic Biology 58:547–559.

Leaché, A. D., B. L. Banbury, C. W. Linkem, and A. N.-M. de Oca. 2016. Phylogenomics of a rapid radiation: is chromosomal evolution linked to increased diversification in north american spiny lizards (genus sceloporus)? BMC evolutionary biology 16:63.

Leaché, A. D., A. S. Chavez, L. N. Jones, J. A. Grummer, A. D. Gottscho, and C. W. Linkem. 2015. Phylogenomics of phrynosomatid lizards: conflicting signals from

sequence capture versus restriction site associated dna sequencing. Genome biology and evolution 7:706–719.

Leaché, A. D., R. B. Harris, B. Rannala, and Z. Yang. 2013. The influence of gene flow on species tree estimation: a simulation study. Systematic Biology 63:17–30.

Leaché, A. D. and J. A. McGuire. 2006. Phylogenetic relationships of horned lizards (phrynosoma) based on nuclear and mitochondrial data: evidence for a misleading mitochondrial gene tree. Molecular phylogenetics and evolution 39:628–644.

Lemmon, A. R., S. A. Emme, and E. M. Lemmon. 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Systematic biology Page sys049.

Li, H. and R. Durbin. 2009. Fast and accurate short read alignment with burrows–wheeler transform. Bioinformatics 25:1754–1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, et al. 2009. The sequence alignment/map format and samtools. Bioinformatics 25:2078–2079.

Linkem, C. W., V. N. Minin, and A. D. Leaché. 2016. Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (squamata: Scincidae). Systematic biology 65:465–477.

Liu, L., L. Yu, D. K. Pearl, and S. V. Edwards. 2009. Estimating species phylogenies using coalescence times among sequences. Systematic Biology 58:468–477.

MacManes, M. 2013. Available from: http://dx.doi.org/10.6084/m9.figshare.658946. Figshare 5.

Maddison, W. 1989. Reconstructing character evolution on polytomous cladograms. Cladistics 5:365–377.

Mallet, J. 2005. Hybridization as an invasion of the genome. Trends in ecology & evolution 20:229–237.

Martin, S. H., K. K. Dasmahapatra, N. J. Nadeau, C. Salazar, J. R. Walters, F. Simpson, M. Blaxter, A. Manica, J. Mallet, and C. D. Jiggins. 2013. Genome-wide evidence for speciation with gene flow in heliconius butterflies. Genome Research 23:1817–1828.

McCormack, J. E., S. M. Hird, A. J. Zellmer, B. C. Carstens, and R. T. Brumfield. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution 66:526–538.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, et al. 2010. The genome analysis toolkit: a mapreduce framework for analyzing next-generation dna sequencing data. Genome research 20:1297–1303.

Meiklejohn, K. A., B. C. Faircloth, T. C. Glenn, R. T. Kimball, and E. L. Braun. 2016. Analysis of a rapid evolutionary radiation using ultraconserved elements: evidence for a bias in some multispecies coalescent methods. Systematic biology 65:612–627.

Morando, M., L. J. Avila, J. Baker, J. W. Sites Jr, and M. Ashley. 2004. Phylogeny and phylogeography of the liolaemus darwinii complex (squamata: Liolaemidae): evidence for introgression and incomplete lineage sorting. Evolution 58:842–861.

Murphy, W. J., T. H. Pringle, T. A. Crider, M. S. Springer, and W. Miller. 2007. Using genomic data to unravel the root of the placental mammal phylogeny. Genome research 17:413–421.

Olave, M., L. J. Avila, J. W. Sites, and M. Morando. 2014. Multilocus phylogeny of the widely distributed south american lizard clade eulaemus (liolaemini, liolaemus). Zoologica Scripta 43:323–337.

Olave, M., L. J. Avila, J. W. Sites, and M. Morando. 2017. Hidden diversity within the lizard genus liolaemus: Genetic vs morphological divergence in the l. rothi complex (squamata: Liolaeminae). Molecular phylogenetics and evolution 107:56–63.

Olave, M., L. J. Avila, J. W. Sites Jr, and M. Morando. 2015. Model-based approach to test hard polytomies in the eulaemus clade of the most diverse south american lizard genus liolaemus (liolaemini, squamata). Zoological Journal of the Linnean Society 174:169–184.

Olave, M., L. E. Martinez, L. J. Avila, J. W. Sites, and M. Morando. 2011. Evidence of hybridization in the argentinean lizards liolaemus gracilis and liolaemus bibronii (iguania: Liolaemini): An integrative approach based on genes and morphology. Molecular Phylogenetics and Evolution 61:381–391.

Panzera, A., A. D. Leaché, G. D'Elía, and P. F. Victoriano. 2017. Phylogenomic analysis of the chilean clade of liolaemus lizards (squamata: Liolaemidae) based on sequence capture data. PeerJ 5:e3941.

Patel, S., R. T. Kimball, and E. L. Braun. 2013. Error in phylogenetic estimation for bushes in the tree of life. Journal of Phylogenetics & Evolutionary Biology .

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest radseq: an inexpensive method for *de novo* snp discovery and genotyping in model and non-model species. PloS one 7:e37135.

Portik, D. M., L. L. Smith, and K. Bi. 2016. An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (class: Amphibia, order: Anura). Molecular ecology resources 16:1069–1083.

Rambaut A, X. D., Suchard MA and D. AJ. 2014. Tracer version v1.6. http://tree.bio.ed.ac.uk/software/tracer/ .

Rannala, B. and Z. Yang. 2003. Bayes estimation of species divergence times and ancestral population sizes using dna sequences from multiple loci. Genetics 164:1645–1656.

Rokas, A. and S. B. Carroll. 2006. Bushes in the tree of life. PLoS Biol 4:e352.

Schluter, D. 2000. The ecology of adaptive radiation. OUP Oxford.

Schneider, K., S. Koblmüller, and K. M. Sefc. 2016. hext, a software supporting tree-based screens for hybrid taxa in multilocus data sets, and an evaluation of the homoplasy excess test. Methods in Ecology and Evolution 2016:358–368.

Sersic, A. N., A. Cosacov, A. A. Cocucci, L. A. Johnson, R. Pozner, L. J. Avila, J. W. Sites Jr, and M. Morando. 2011. Emerging phylogeographical patterns of plants and terrestrial vertebrates from patagonia. Biological Journal of the Linnean Society 103:475–494.

Smith, B. T., M. G. Harvey, B. C. Faircloth, T. C. Glenn, and R. T. Brumfield. 2014. Target capture and massively parallel sequencing of ultraconserved elements (uces) for comparative studies at shallow evolutionary time scales. Systematic biology 63:83–95.

Stamatakis, A. 2014. Raxml version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Sullivan, J. and P. Joyce. 2005. Model selection in phylogenetics. Annu. Rev. Ecol. Evol. Syst. 36:445–466.

Swofford, D. L. 2003. Paup*. phylogenetic analysis using parsimony (* and other methods). version 4. .

Than, C., D. Ruths, and L. Nakhleh. 2008. Phylonet: a software package for analyzing and reconstructing reticulate evolutionary relationships. BMC bioinformatics 9:322.

Venkatesh, B., M. V. Erdmann, and S. Brenner. 2001. Molecular synapomorphies resolve evolutionary relationships of extant jawed vertebrates. Proceedings of the National Academy of Sciences 98:11382–11387.

Wiens, J. J., C. R. Hutter, D. G. Mulcahy, B. P. Noonan, T. M. Townsend, J. W. Sites, and T. W. Reeder. 2012. Resolving the phylogeny of lizards and snakes (squamata) with extensive sampling of genes and species. Biology letters 8:1043–1046.

Yang, Z. 2015. The bpp program for species tree estimation and species delimitation. Current Zoology 61:854–865.

Yang, Z. and B. Rannala. 2010. Bayesian species delimitation using multilocus sequence data. Proceedings of the National Academy of Sciences 107:9264–9269.

Yang, Z. and B. Rannala. 2012. Molecular phylogenetics: principles and practice. Nature Reviews Genetics 13:303.

Zhang, C., D.-X. Zhang, T. Zhu, and Z. Yang. 2011. Evaluation of a bayesian coalescent method of species delimitation. Systematic Biology 60:747–761.

Zhang, J., K. Kobert, T. Flouri, and A. Stamatakis. 2014. Pear: a fast and accurate illumina paired-end read merger. Bioinformatics 30:614–620.

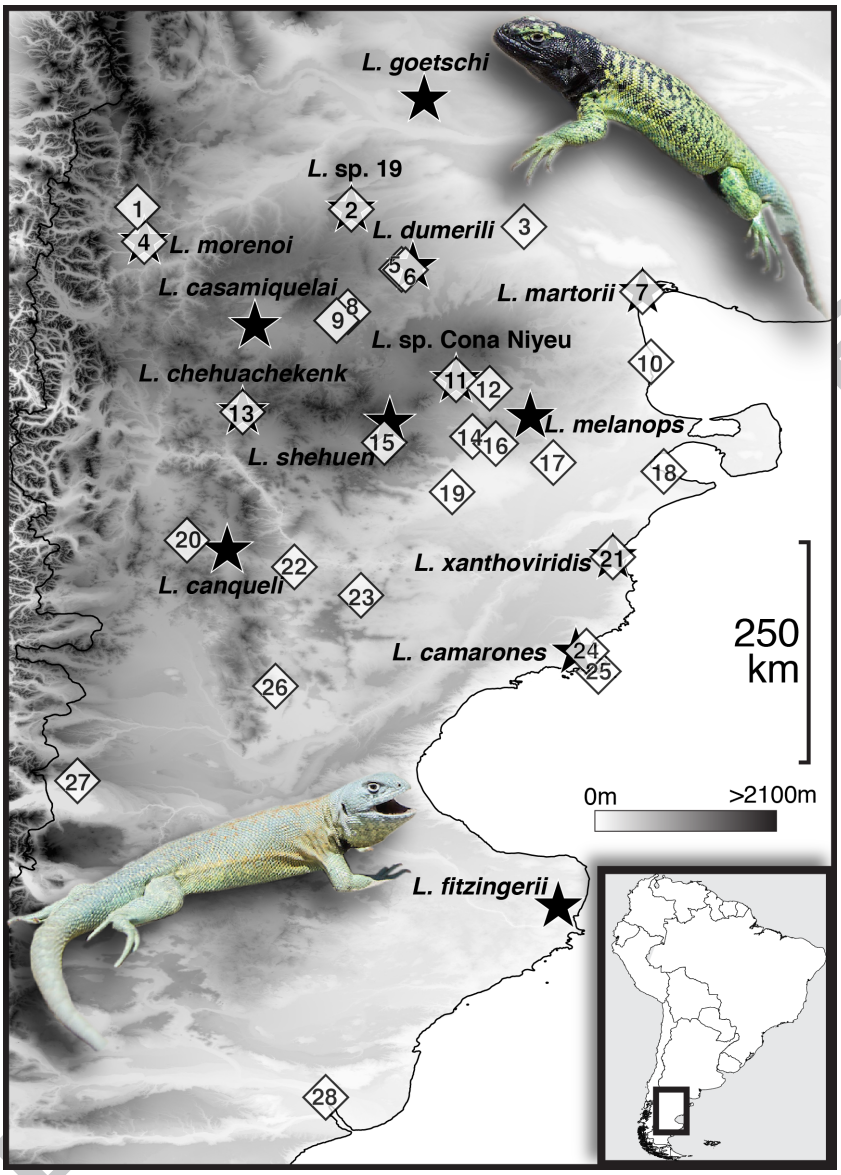


Figure 1: Sampling map of southern-central Argentina with type localities (stars) labeled by name for described and undescribed species in the *L. fitzingerii* species group, and locations where individuals were sampled (diamonds). Sampling numbers on the map correspond to the following individuals and their names used throughout this study: 1 - *Liolaemus purul*, 2 - *Liolaemus* sp. 19, 3 - *Liolaemus goetschi*, 4 - *Liolaemus morenoi*, 5 - *Liolaemus melanops* N1, 6 - *Liolaemus dumerili*, 7 - *Liolaemus martorii* N, 8 - *Liolaemus melanops* N2, 9 - *Liolaemus casamiquelai*, 10 - *Liolaemus martorii* S, 11 - *Liolaemus* sp. Cona Niyeu, 12 - *Liolaemus melanops* C, 13 - *Liolaemus chehuachekenk*, 14 - *Liolaemus* sp. 18, 15 - *Liolaemus shehuen*, 16 - *Liolaemus melanops* S1 (pictured, top-right), 17 - *Liolaemus melanops* S3, 18 - *Liolaemus* sp. 17, 19 - *Liolaemus melanops* S2, 20 - *Liolaemus* sp. 16, 21 - *Liolaemus xanthoviridis* E, 22 - *Liolaemus canqueli*, 23 - *Liolaemus xanthoviridis* W, 24 - *Liolaemus camarones*, 25 - *Liolaemus fitzingerii* Isla Leones, 26 - *Liolaemus fitzingerii* N (pictured, bottom-left), 27 - *Liolaemus fitzingerii* W, 28 - *Liolaemus fitzingerii* S.

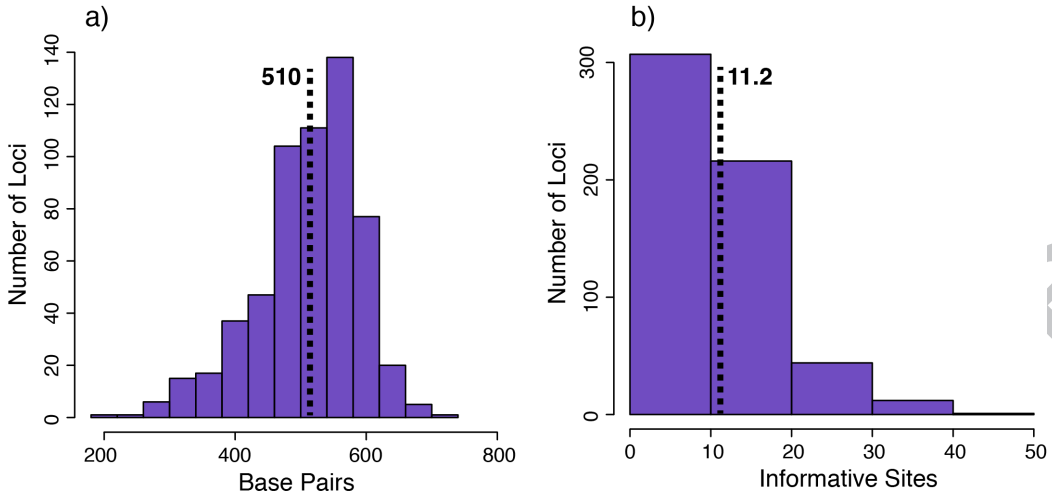


Figure 2: Sequence length (a) and number of informative sites (b) per nuclear locus for only ingroup individuals with means depicted with dashed lines. See Supplemental Figures S1-2 for further sequence statistics.

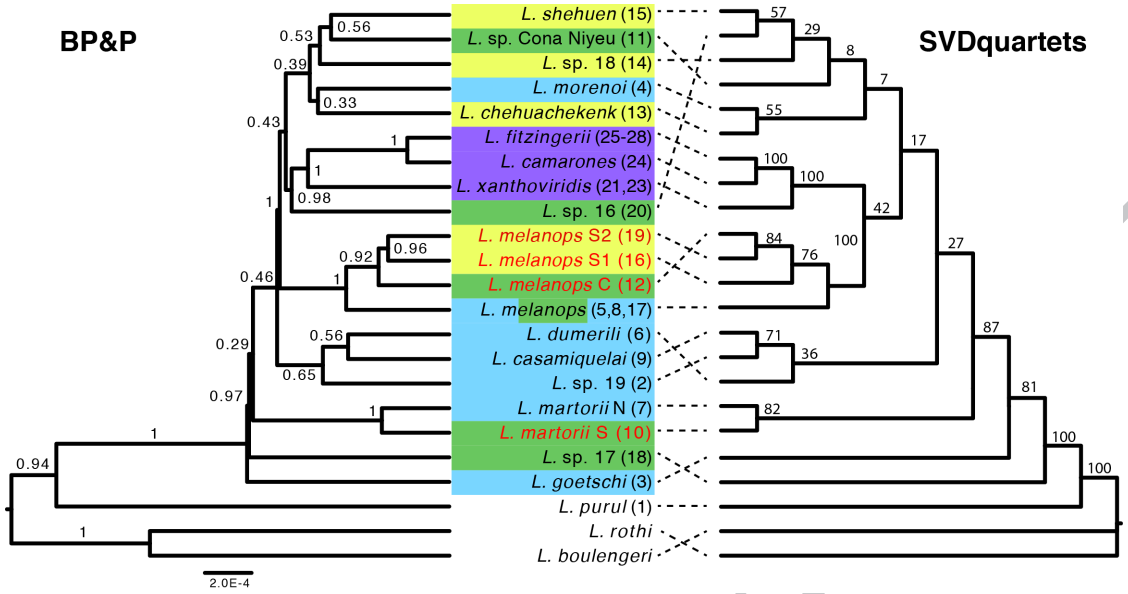


Figure 3: Multi-species coalescent phylogenies estimated with BP&P (G(2, 200) and G(2, 400) for the θ and τ priors, respectively) and SVDquartets (note the change in branch lengths for the BP&P analysis with smaller mean prior values in Supplemental Figure S3). Support values are posterior probabilities for the BP&P phylogeny and bootstraps for the SVDquartets phylogeny. Numbers following taxon names correspond to sample numbers in Figure 1, colors reflect mitochondrial clade memberships in Fig. 4, and branch lengths in the BP&P tree are in coalescent units. Tips labeled in red represent putative hybrid lineages (see Supplemental Figs. S4,5 for analyses without hybrids), and there are fewer tips than individuals because multiple individuals/alleles are assigned to each species in these trees.

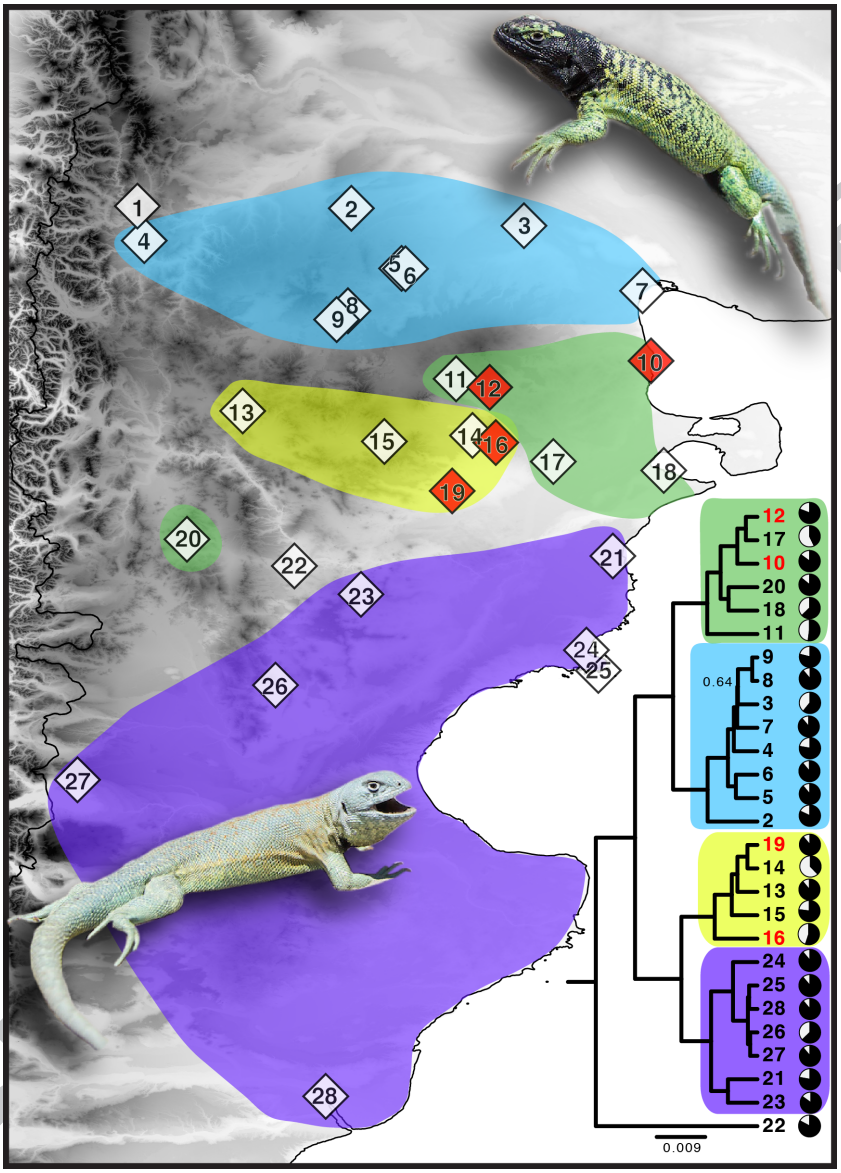


Figure 4: Phylogeny inferred from the mitogenomic dataset along with approximate geographic distributions of clades. Fraction of the mitogenome sequenced for each individual is shown in pie charts to the right (black = data present), branch lengths are in number of expected substitutions per site, and all nodes without support values shown received a posterior probability of 1.0. Sample numbering corresponds to the names given in Figure 1. Individuals labeled in red are suspected hybrids based on morphology and discordant placement in the nDNA tree. See Supplemental Figure S7 for the full mitochondrial genealogy including outgroup data.

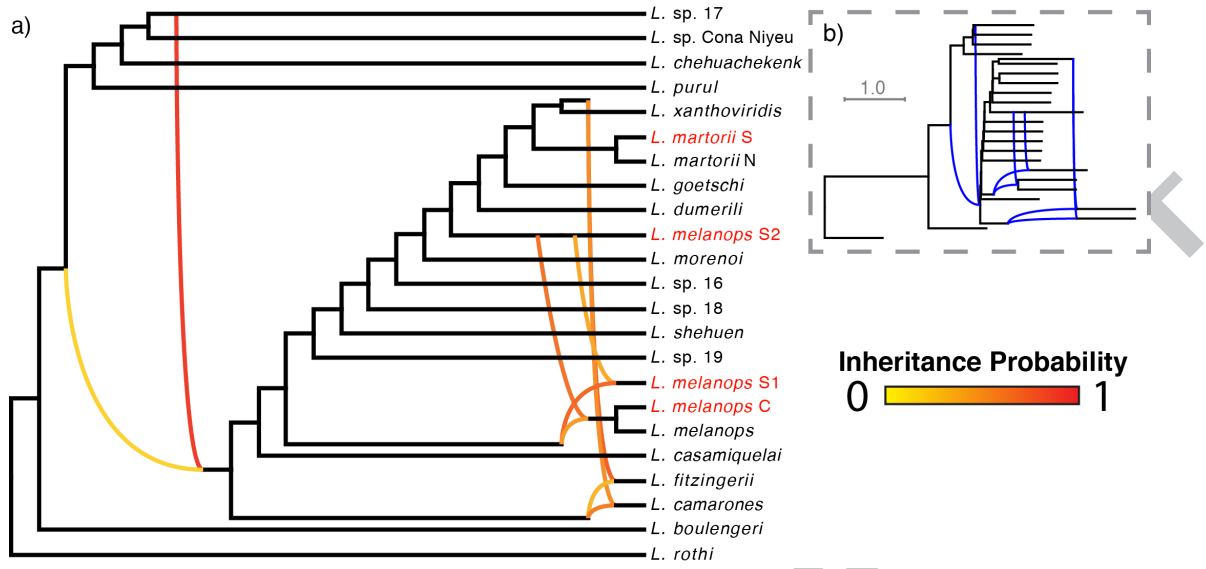


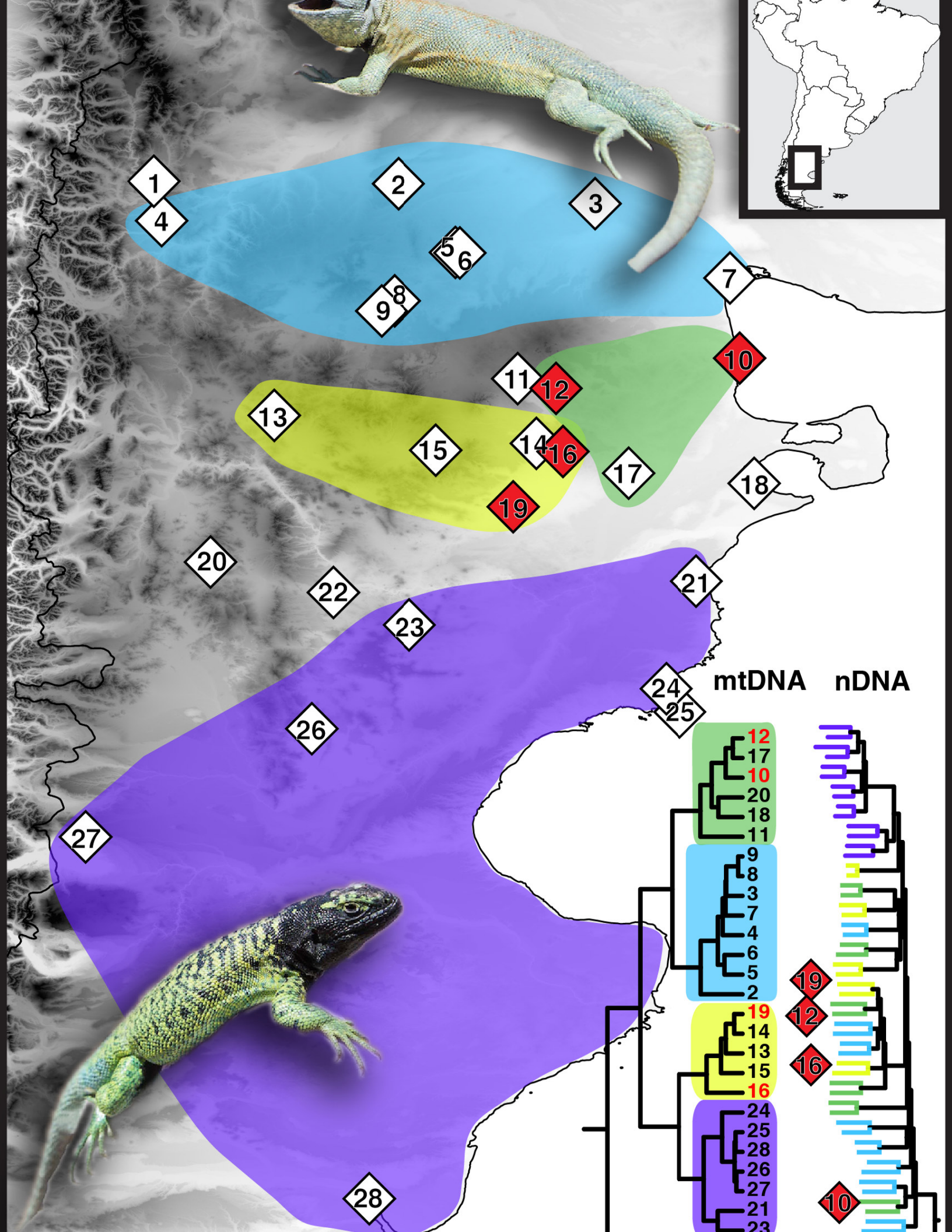
Figure 5: Phylonet network inferred showing the AIC-preferred five reticulations, with suspected hybrids in red. Reticulation events and relationships are shown in the larger network (a) and inferred branch lengths are shown in the (b) inset and represent coalescent units (number of generations divided by two times the effective population size). Note the inferred "ghost" lineage sister to *L. xanthoviridis* that is related to *L. fitzingerii* and *L. camarones*.

Table 1: Summary statistics for both nuclear (nDNA) and mitochondrial DNA (mtDNA) ingroup sequence data used in this study, with nuclear data shown by locus type and in aggregate. Averages for the nDNA and single values for the mtDNA are listed, whereas ranges are shown in parentheses. See Supplemental Figures S1,S2 and Supplemental Tables S3,S5 for further information.

	# Sequences	Length (bp)	# Inform. Sites	% Inform. Sites
Squamate TOL	50.4 (38-54)	428 (211-608)	16.78 (5-34)	4.22 (1-14)
UCE	50.99 (34-54)	518 (261-701)	10.82 (0-47)	2.09 (0-8.2)
nDNA Total	50.95	512	11.24	2.24
mtDNA	28	13,323 (6,616-15,370)	2,736	17.7

Table 2: Phylonet results and AIC phylogenetic network model selection, with the optimal network in bold. "BL" stands for number of branch lengths estimated, and k is the number of parameters used in the AIC calculation.

# Retics.	lnL	ΔlnL	# BLs	# Inferred Retics.	k	AIC	ΔAIC
0	-12015285		21	0	21	24030612	18821
1	-12011478	3807	22	1	23	24023002	11211
2	-12008493	2985	22	2	24	24017033	5242
3	-12007447	1046	23	3	26	24014945	3154
4	-12006527	920	22	4	26	24013105	1313
5	**-12005865**	**662**	**26**	**5**	**31**	**24011791**	**0**


1
4
2
3
5
6
9
8
7
10
11
12
13
15
14
16
17
18
19
20
22
23
21
24
25
26
27
28
mtDNA
nDNA
12
17
10
20
18
11
9
8
3
7
4
6
5
2
19
14
13
15
16
24
25
28
26
27
21
23
19
12
16
10

Research Highlights for *Phylogenomic evidence for a recent and rapid radiation in lizards of the Patagonian* Liolaemus fitzingerii *species group*

- Sequenced 580 nuclear loci and the mitogenome for 12 lizard species (n=28)
- Interspecific relationships based on nuclear DNA were weakly supported
- NDNA and mtDNA phylogenetic relationships were in conflict
- Formal tests provided ambiguous support for hybridization
- Species in the *Liolaemus fitzingerii* group constitute a recent and rapid radiation