Population genetic structure and species delimitation of a widespread, Neotropical dwarf gecko

Brendan J. Pinto a,⁎, Guarino R. Colli b, Timothy E. Higham c, Anthony P. Russell d, Daniel P. Scantlebury e, Laurie J. Vitt f, Tony Gamble e,⁎,⁎

a Department of Biological Sciences, Marquette University, Milwaukee, WI, USA
b Departamento de Zoologia, Universidade de Brasília, Brasília, DF, Brazil
c Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, Riverside, CA, USA
d Department of Biological Sciences, University of Calgary, Calgary, AB, Canada
e Resonate, Reston, VA, USA
f Sam Noble Museum and Biology Department, University of Oklahoma, Norman, OK, USA
g Bell Museum of Natural History, University of Minnesota, Saint Paul, MN, USA
h Milwaukee Public Museum, Milwaukee, WI, USA

ARTICLE INFO

Keywords:
Amazonia
RADseq
Species delimitation
Population genetics
Neotropical phylogeography

ABSTRACT

Amazonia harbors the greatest biological diversity on Earth. One trend that spans Amazonian taxa is that most taxonomic groups either exhibit broad geographic ranges or small restricted ranges. This is likely because many traits that determine a species range size, such as dispersal ability or body size, are autocorrelated. As such, it is rare to find groups that exhibit both large and small ranges. Once identified, however, these groups provide a powerful system for isolating specific traits that influence species distributions. One group of terrestrial vertebrates, gecko lizards, tends to exhibit small geographic ranges. Despite one exception, this applies to the Neotropical dwarf geckos of the genus Gonatodes. This exception, Gonatodes humeralis, has a geographic distribution almost 1,000,000 km² larger than the combined ranges of its 30 congeners. As the smallest member of its genus and a gecko lizard more generally, Gonatodes humeralis is an unlikely candidate to be a wide-ranged Amazonian taxon. To test whether or not Gonatodes humeralis is, in fact, a single species across its contiguous range in South America, we generated molecular genetic data using restriction-site associated sequencing (RADseq) and traditional Sanger methods for samples from across its range and conducted a phylogeographic study. We conclude that Gonatodes humeralis is, in fact, a single species across its contiguous range in South America. Thus, Gonatodes is a unique clade among Neotropical taxa, containing both wide-ranged and range-restricted taxa, which provides empiricists with a powerful model system to correlate complex species traits and distributions. Additionally, we provide evidence to support species-level divergence of the allopatric population from Trinidad and we resurrect the name Gonatodes ferrugineus from synonymy for this population.

1. Introduction

The use of genetic data to study variation among populations and delimit species has provided unprecedented insight into the patterns and processes of speciation (Castillas and Barbadilla, 2017; Domingos et al., 2017; Gratton et al., 2015; Harvey et al., 2017; Lemmon et al., 2012; McKay et al., 2013; Nazareno et al., 2017a, 2017b; Weir et al., 2015). Genetic data have been particularly useful in the investigation of poorly-studied taxa from Neotropical regions, such as Amazonia (Angulo and Icochea, 2010; Antonelli et al., 2011). Employing large genetic datasets to Neotropical biogeographic studies can vastly increase their accuracy and resolution relative to previous analyses. Most Neotropical work to date, however, has been conducted using a single type of data (largely mitochondrial data), and has likely led to the oversimplification in our understanding of this biogeographic system (Beheregaray, 2008; Turchetto-Zolet et al., 2013). Thus, in order to elucidate the complex historical scenarios across the Neotropics that have resulted in the immense biodiversity harbored there, studies utilizing larger datasets are needed for a diversity of animal groups.

Several hypotheses have been proposed to explain the historical and spatial patterns of range-limited Amazonian species (see Antonelli et al., 2011; Turchetto-Zolet et al., 2013 for thorough review), nearly all...
of which depend on the emergence of physical barriers to gene flow that result in allopatric speciation (Hafer, 1969, 1997; Endler, 1977; Vanzolini and Williams, 1981; Wallace, 1852). Although there has been considerable debate as to the timing of Amazonian speciation, it now seems clear that cladogenesis has been happening, continually, for tens of millions of years. For instance, many invertebrate, mammal, and bird groups display interspecific divergence between sister species during the Quaternary (< 2.6 million years ago [mya]), whereas many amphibians and reptiles exhibit earlier divergence times during the Neogene (> 2.6 mya) (Gamble et al., 2008; Antonelli et al., 2011; Fouquet et al., 2015; Turchetto-Zolet et al., 2013). Thus, determining the complex patterns that have generated Amazonian biodiversity may require testing several competing hypotheses and searching for patterns between large- and small-scale studies, across a variety of taxonomic groups. Indeed, and as datasets for Neotropical taxa increase in size, complex historical scenarios have been uncovered that were previously unidentifiable and/or untestable (Alexander et al., 2017; Avila-Pires et al., 2012; Fouquet et al., 2015; Lessa et al., 2003; Nazarenko et al., 2017a, 2017b; Prates et al., 2016; Werneck et al., 2012).

One trend that molecular genetic data have revealed is that many widely distributed tropical taxa are composed of multiple, often cryptic, species (Funk et al., 2012). These species are usually of smaller body size, with low vagility, and/or those that occupy narrow ecological niches (Camargo et al., 2006; Fouquet et al., 2007b; Wynn and Heyer, 2001). Indeed, even prior to the advent of molecular genetic data, it was predicted that very few widespread nominal taxa in the Neotropics would remain intact upon closer investigation (Lynch, 1979). Subsequently, phylogeographic studies of multiple populations have found that most widespread, non-volante, vertebrate taxa are in fact ‘species-complexes’ (i.e. composed of multiple undescribed and/or cryptic species). This pattern extends across many terrestrial vertebrate groups including, but not limited to: anole lizards (D’Angiolella et al., 2011; Glor et al., 2001), frogs (Camargo et al., 2006; Caminer et al., 2017; Chek et al., 2001; Fouquet et al., 2007a, 2014; Funk et al., 2012; Gehara et al., 2014; Guayasamin et al., 2017; Wynn and Heyer, 2001), gecko lizards (Bergmann and Russell, 2007; Gamble et al., 2011a; Geurges and Rodrigues, 2010; Kronauer et al., 2005), salamanders (Hervas et al., 2016), toads (Fouquet et al., 2007a; Funk et al., 2012; Murphy et al., 2017b), and other herpetofauna (Nunes et al., 2012; De Oliveira et al., 2016). Furthermore, identifying concordant patterns in species’ ranges is an important step in the testing of complex biogeographical scenarios that underpin the origins of biodiversity (Clarke et al., 2017a, 2017b; Costello et al., 2013; Da Silva and Patton, 1993; Díaz-Nieto et al., 2016; Ditchfield, 2000; Gazoni et al., 2018; Gehara et al., 2014; Miralles and Carranza, 2010; Stroud and Feely, 2017; Turchetto-Zolet et al., 2013).

Whereas many widespread Neotropical taxa appear to be composed of multiple, undescribed species, there are exceptions to this pattern and widely distributed Neotropical taxa do exist. However, these widespread taxa are less frequent than once thought and are typically species that exhibit traits that facilitate high vagility (e.g. being volant, having a large body size, and/or occupying broad ecological niches). Some notable examples of these widespread taxa include: the Amazon Tree Boa (Corallus hortulanus), Andersen’s Fruit-eating Bat (Artibeus anderseni), the Bushmaster (Lachesis muta), capybaras (Hydrochoerus hydrochaeris), jaguars (Panthera onca), the Green Anaconda (Eunectes murinus), the Green Iguana (Iguana iguana), and the Lesser Treefrog (Dendropsophus minutus) (Colston et al., 2013; Ditchfield, 2000; Eizirik et al., 2001; Gehara et al., 2014; Zamudio and Greene, 1997). These examples suggest that range size and abundance of Neotropical species are likely attributable to intrinsic factors such as body size, dispersal ability, and niche breadth, among other traits that have a strong phylogenetic component (Dexter and Chave, 2016; Meiri et al., 2017; Wynn and Heyer, 2001). Thus, some clades are composed mostly of wide-ranging species (large and dispersal-limited animals), while others are composed mainly of range-limited species (small and dispersal-limited animals). Studying differences in ecological traits and range distribution among these taxa can provide important insights into the patterns and processes responsible for Neotropical biodiversity. However, it is difficult to deduce the relative contribution of individual traits to range size disparities between species, because many traits are autocorrelated at the macroevolutionary scale (Beck and Kitching, 2007; Dexter and Chave, 2016; Hurlbert and White, 2007). Investigating clades that include both geographically widespread and restricted species may provide important insights into how phenotypic differences can influence species distributions (Gehara et al., 2014).

In line with these observations, most Neotropical lizard species have small distributions. However, there are a few notable exceptions, such as the dwarf gecko, Gonatodes humeralis, the geographic distribution of which (~7,600,000 km²) is larger than that of all of its congeners combined, by nearly 1,000,000 km² (~6,700,000 km²) (Roll et al., 2017). Gonatodes humeralis occurs across Amazonia and the Guiana Shield, as well as in forested enclaves and gallery forests in the adjacent Cerrado and Caatinga biomes, and on the island of Trinidad (Avila-Pires, 1995; Murphy, 1997; Ribeiro-Júnior, 2015; Roberto et al., 2014; Vanzolini, 1955). Overall, its current distribution occupies a geographic area marginally smaller than that of the continental United States and overlaps with 13 currently described congeneric species (Supplemental Fig. 1). Gonatodes humeralis also exhibits a broad niche breadth, occurring in a variety of habitat types including: primary and secondary forest, riparian forest, gallery forest, forest edges, bamboo forest, and human dwellings (Carvalho et al., 2008; Dixon and Soini, 1986; Higham et al., 2017; Hoogmoed, 1973; Vanzolini and Williams, 1981; Vitt and Zani, 1996; Vitt et al., 1997, 2000). Its massive distribution and extensive niche breadth contrast with those of its congeners, most of which occupy specialized niches with small, distributions in Central and South America and several islands of the Lesser Antilles (Supplemental Fig. 1). In the context of recent discoveries suggesting that widespread Neotropical taxa are uncommon, the diminutive G. humeralis (maximum 41.5 mm snout-vent length; Avila-Pires, 1995) is an unlikely candidate for being a single species. However, if G. humeralis is, in fact, one widespread species, then Gonatodes harbors both widespread and geographically restricted taxa, providing a powerful model system for identifying traits that may influence species distributions.

Previous investigations on G. humeralis have revealed evidence for genetic, ecological, and morphological variation between populations across its range (Avila-Pires, 1995; Avila-Pires et al., 2012; Rivero-Blanco, 1979; Vitt et al., 1997), and early hypotheses suggested that populations should exhibit relatively shallow divergence times, within the Pleistocene (Vanzolini and Williams, 1981; Vitt et al., 1997). Supporting this, the first multi-locus phylogenetic analysis of Gonatodes revealed that G. humeralis samples from eastern and western Amazonia likely shared a common ancestor in the late Pliocene or early/mid Pleistocene, approximately 1.9 (1.1–2.7) mya (Gamble et al., 2008). Later, the most comprehensive phylogeographic analysis to date investigated the history of G. humeralis populations in eastern Amazonia using two mitochondrial markers (Cytb & 16S) from 56 individuals (Avila-Pires et al., 2012). The authors found little phylogenetic resolution among sampled populations and no evidence that Amazonian rivers (namely, the Amazon and Tocantins) have acted as isolating mechanisms between sampled populations in eastern Amazonia. The authors concluded that range-wide sampling and the addition of nuclear markers would be necessary to obtain sufficient resolution of any phylogeographic hypothesis relating to this species.

We herein investigate the geographically widespread gecko, G. humeralis, across its range in northern South America and Trinidad. Specifically, we test two alternative hypotheses: (i) if G. humeralis is typical of most small, non-volante Neotropical vertebrates, we expect to uncover a species-complex composed of multiple cryptic, or morphologically similar, species; (ii) conversely, if G. humeralis is an atypical taxon, then we expect it to be a single, widespread species across its
contiguous Amazonian range, and potentially also on the island of Trinidad. To test this, we generated restriction-site associated DNA sequencing (RADseq) data, and a multi-locus Sanger-sequenced dataset using traditionally informative nuclear and mitochondrial markers. We began by investigating the population genetic structure of *G. humeralis* across its range, then reconstructed the relationships of those alleles between populations, and used these relationships to generate specific species delimitation hypotheses for further testing. Indeed, we predicted that *G. humeralis* would consist of multiple, cryptic species with distributions comparable in size to those of other species of *Gonatodes*. However, we found that *G. humeralis* is a single, widespread species across Amazonia, whereas the population on the island of Trinidad appears to be highly divergent and independently-evolving. We discuss these results in a comparative context with other Neotropical species and posit that this genus of geckos (*Gonatodes*) may yield unprecedented insights into the origins and maintenance of Neotropical biodiversity.

2. Materials and methods

2.1. Sampling

We sampled 31 individuals of *G. humeralis* from 13 localities across its range (Fig. 1). Three individuals of *G. antillensis* were included as an outgroup (Russell et al., 2015). We extracted genomic DNA from downstream genetic sequencing from tail clips or liver, using the Qiagen® DNeasy Blood and Tissue extraction kit.

2.2. RADseq data

We generated a reduced-representation genomic dataset for all *G. humeralis* individuals using restriction-site associated DNA sequencing (RADseq). RADseq libraries were constructed following a protocol modified from Etter et al. (2011), as described by Gamble et al. (2015). Briefly, genomic DNA was digested using high-fidelity *SbfI* restriction enzyme (New England Biolabs). We ligated individually barcoded P1 adapters onto the *SbfI* cut site for each sample. Samples were pooled into multiple libraries, sonicated, and size selected into 200- to 500-bpasepair (bp) fragments using magnetic beads in a PEG/NaCl buffer (Rohland and Reich, 2012). Libraries were blunt-end-repaired and dA-tailed. To each of the pooled libraries, we ligated a P2 adapter containing unique Illumina barcodes. Libraries were amplified using 16 PCR cycles with Phusion high-fidelity DNA polymerase (New England Biolabs), and were size-selected a second time into 250- to 600 bp fragments using magnetic beads in PEG/NaCl buffer. Libraries were sequenced using paired-end 125 bp reads on the Illumina® HiSeq2500 at the Institute for Integrative Genome Biology, University of California, Riverside. RADseq data for the 10 individuals from Trinidad were previously published (Gamble et al., 2018).

We trimmed and demultiplexed raw single-end Illumina sequencing reads by their individual-specific barcodes using the process_radtags command in STACKS [v1.23]; (Catchen et al., 2011). After the removal of low-quality reads, restriction site overhangs, and barcodes, the 3’ ends of the 125 bp reads were trimmed to 100 bp. Cleaned reads were imported into the PyRAD pipeline [v3.0.63] for de novo assembly [steps 2–7] (Eaton, 2014). One individual, CHUNB47049, was removed from the RADseq dataset prior to filtering, due to low-quality reads (adjusting our RADseq dataset, N = 30). This removed locality #8 (Fig. 1c) from all RADseq data analyses. We assayed various filtering criteria configurations, including varying the minimum read depth per locus from 4 to 12; maximum number of “N”s per locus from 4 to 6; within- and across-sample clustering threshold from 80 to 98%; and the minimum number of individuals with sequence data for a locus needed from 10 to 28. To obtain a dataset with > 10,000 and < 50,000-unlinked markers incorporating ≤10% missing data, we set the final filtering criteria for exclusion of any locus with a read depth of less than 8 reads, and missing data (“N” characters) to ≥5. We set the within- and across-sample clustering threshold to 95% sequence identity, and the minimum number of individuals required for data to be included in a final locus was set to 25 of the 30 individuals. All other PyRAD parameters used default settings. The final dataset consisted of 35,260 informative loci with 67,173 total single-nucleotide polymorphisms.

**Fig. 1.** (A) Maximum-likelihood tree computed using 22,486 unlinked SNPs executed in RAxML, bootstrap values ≥70 reported (bootstrap values of 100 = black circles, bootstrap values from 70 to 99 = gray circles). Bolded numbers correlate individual or clade with sampling locality depicted on map. (B) Distruct plot depicting proportions of shared alleles present in the *G. humeralis* lineage determined by STRUCTURE analysis, K = 3 (Supplemental Fig. 2). (C) Map indicating sampling localities, within the geographic range of *G. humeralis*, in relation to cluster assignments (Trinidad = circle, west Amazonia = solid square, east Amazonia = patterned square) in relation to their geographic locality (Supplemental Table 1). Further, locality 8 (represented by CHUNB47049) is absent from the RADseq tree in panel A (see Methods).
(SNPs), 26,486 of which were unlinked (sampling only one SNP per RAD locus). We subsampled and reformatted this final dataset for all downstream RADseq data analyses; further data specifics for each analysis are provided below.

2.3. Sanger sequence data

We also produced sequence data from fragments of six molecular markers using Sanger sequencing of PCR amplicons. This consisted of four nuclear genes: microtubule-associated protein 1b – exon 5 (MAP1b), recombination-activating gene 1 (RAG1), oocyte-maturation factor MOS (CMOS), and protein tyrosine phosphatase nonreceptor type 12 (PTPN12); and two mitochondrial genes: NADH dehydrogenase subunit 2 (ND2) and 16S ribosomal subunit (16S). PCR conditions and primer sequences are described elsewhere: MAP1b (Werneck et al., 2012), RAG1, CMOS, PTPN12 (Gamble et al., 2011b), ND2 (Jackman et al., 2008), and 16S (Gamble et al., 2008). We Sanger-sequenced PCR amplicons using GeneWiz® single-pass sequencing, then assembled and quality-trimmed raw sequences using Geneious® [v9.1.5] (Kearse et al., 2012). GenBank accession numbers for all sequences are listed in Supplemental Table 1. Sequences were aligned using MUSCLE [v3.8.425] (Edgar, 2004) and alignments refined by eye, if necessary. Models of molecular evolution were chosen based on AICc and BIC criteria, computed using MEGA7 (Kumar et al., 2016).

2.4. Population genetic analyses

We visualized the population-level genetic diversity within G. humeralis sensu lato and estimated the number of genetic populations in Hardy-Weinberg equilibrium present in our RADseq data using STRUCTURE [v2.3.4] (Pritchard et al., 2007). We investigated possible values of K (where K is equal to the number of populations of alleles) between 1 and 6 with a subset of the unlinked SNP data, using only the values of K (where K is equal to the number of populations of alleles) dataset using Arlequin [v3.5.2.2] (Excoffier et al., 2005). We investigated possible values of K (where K is equal to the number of populations of alleles) using the Geographic Distance Matrix Generator software [v2.3.4] (Pritchard et al., 2007). We investigated possible values of K (where K is equal to the number of populations of alleles) using the Mantel [v2.4.2] function, creating 999 bootstrap replicates.

We also produced a rooted mitochondrial gene tree in a Bayesian framework, to compare with the nuclear SNP data tree, using BEAST2 [v2.5.1] under a strict clock (Bouckaert et al., 2014) on the CIPRES cluster (Miller et al., 2010). The concatenated mitochondrial (mtDNA) data (ND2 and 16S) consisted of 34 samples, including three G. antillensis, for a total of 1484 bp. We used the GTR + I model and a Yule tree prior with $5 \times 10^9$ MCMC iterations with a 10% burnin. Bayesian analyses were replicated three times and examined by eye using Tracer [v1.6.1] to ensure convergence. Post-burnin trees from all three runs were combined to estimate final tree parameters using Log Combiner and Tree Annotator, respectively.

Next, we estimated divergence times among G. humeralis populations using the StarBEAST2 [v0.15.1] (Ogilvie et al., 2017) module of BEAST2 (Bouckaert et al., 2014). We used the multi-locus Sanger sequence data, sampling 15 Gonatodes species using a secondary calibration at the root following Higham et al. (2017) and individuals of three G. humeralis phylogeographic clusters identified by STRUCTURE: Trinidad, eastern, and western Amazonia, based on the (see Population Genetic Analyses). The final dataset used in this analysis included seven loci: ACM4, CMOS, mtDNA (ND2 + 16S), DCF (phosducin), PTPN12, RAG1, and RAG2; nuclear loci were phased using DNAsp [v5.0] (Librado and Rozas, 2009; Stephens et al., 2001). Loci used in this analysis were chosen specifically to minimize the amount of missing data per taxon while combining newly generated and previously published sequence data (Supplemental Table 2). Indeed, each locus was provided its own best-fit as calculated in MEGA7 (and has an available model in StarBEAST2), this was HKY + I for all nuclear loci and GTR + I for our concatenated mtDNA genes. We used an uncorrelated lognormal clock model, with secondary calibration from a previously published fossil-calibrated phylogenetic reconstruction, to provide a prior on the root age between Gonatodes and its sister clade Lepidoblepharis at approximately 72.5 (± 7.5) mya, with a uniform distribution to reflect confidence intervals (Gamble et al., 2015).

To corroborate these findings, we utilized the published rate of molecular evolution for the mitochondrial ND2 gene in geckos. We estimated the divergence time between the mainland and Trinidad using p-distances assuming a strict molecular clock. We calculated p-distances in MEGA7 (Kumar et al., 2016) and calculated the divergence time according to the previously published rate of molecular evolution for the ND2 locus in geckos, at 0.57% (per lineage rate) per million years (Macey et al., 1999), i.e. (p-distance/2 * 100) * 0.57 = lineage-divergence in millions of years.

2.6. Species delimitation

We assessed whether G. humeralis consists of one, two, or three
putative species using our phylogenetic and STRUCTURE results to guide assignment of individuals into putative species-level lineages using three species delimitation methods: Poisson Tree Processes (PTP), STACEY, and Bayes Factor Delimitation (BFD).

First, we analyzed species boundaries using our Bayesian mitochondrial gene tree with the single-rate PTP test, using the PTP web service (http://mppt.h-its.org/#/tree), with the p-value set at 0.001 (Kapli et al., 2017).

Second, we used STACEY [v1.2.4] (Jones, 2017) with our Sanger sequenced dataset (CMOS, MAP1b, PTPN12, RAG1, and mtDNA), including G. antillensis and G. concinnus as outgroups. In accordance with program documentation and additional specifications outlined by Barley et al. (2018), we provided an exponential distribution with a mean of 0.1 for the “popPriorScale” parameter, a lognormal distribution with a mean of 5 and a standard deviation of 2 to the “bdeGrowthRate” prior, and the “collapseWeight” was provided a uniform distribution with the lower and upper bounds set at 0 and 1, respectively (Barley et al., 2018). In addition, each gene partition was provided the best-fit model of molecular evolution used by the STACEY package (CMOS and PTPN12 – JC; MAP1b and RAG1 – HKY; mtDNA – TN93), an independent strict molecular clock, with rate priors calculated from a log-normal distribution that were given a mean of 0 and standard deviation of 1 (Barley et al., 2018). We ran three independent chains of 5.0 × 10^7 MCMC repetitions, sampling every 5000 trees, and compared trace files using Tracer [v1.7] (Rambaut et al., 2018). We combined tree files using LogCombiner, visualized them using Densi-Tree, and analyzed the resulting 30,000 trees using the SpeciesDelimitationAnalyzer [v1.8], herein STACEY and SpeciesDelimitationAnalyzer are referred to as SDSA. We used a burnin of 5000 trees and a collapse-height of 0.0001 to calculate our final species delimitation posterior.

Third, we compared two alternative species models, the 2-taxon (PTP: Trinidad/mainland) and 3-taxon models (SSDA: Trinidad/east Amazonia/west Amazonia), using BFD with the RADseq SNP dataset (Leaché et al., 2014). BFD utilizes the path-sampling analysis of the SNAPP package (Bryant et al., 2012) in BEAST2 (Bouckaert et al., 2014) to infer species boundaries directly from biallelic SNP data by comparing the likelihood of two differing species models using Bayes factors (Leaché et al., 2014). We used 48 path sampler steps with 100,000 MCMC repetitions and a 10% burnin to sample from 500,000 MCMC SNAPP replications. We systematically compared models using Bayes factors, calculated using BF = 2 * (|model 1| − |model 2|), where the “model” represents the marginal-likelihood estimate from the specific model being compared against (Ogilvie and Leaché, 2016). We ensured that each model was better than random by estimating the marginal-likelihood for a 3-taxon model, where all individuals were randomly assigned to a “species” to ascertain that both models were better than an unrealistic “null” model (Burbrink et al., 2011).

Lastly, we conducted topology tests to assess whether we could reject the hypothesis that eastern and western Amazonia were reciprocally monophyletic, potentially providing support for the hypothesis that each cluster is a distinct lineage. We constructed two sets of ML trees using RAxML (HPC2-v8.2.10) under a GTR + Γ model, with RAxML’s automatic bootstrapping function (Stamatakis, 2014), also implemented on the CIPRES cluster (Miller et al., 2010) for our RADseq SNP dataset (described above) and for our mtDNA (ND2 and 16S). We constructed an unconstrained tree and a tree for which we enforced a reciprocal monophyletic constraint between eastern and western Amazonia. We conducted topology tests between both trees using the Shimodaira-Hasegawa (SH) test (Shimodaira and Hasegawa, 1999) and Shimodaira’s Approximately Unbiased (AU) test (Shimodaira, 2002) in a likelihood framework under a GTR model with an estimated rate matrix. Topology tests were conducted in Phylogenetic Analysis Using Parsimony (PAUP*) [v 4.0a157] (Swoford, 2002). We calculated significance using 10,000 RELL bootstrap replications.

3. Results

3.1. Population genetic STRUCTURE

The best-fit model for the STRUCTURE analysis was for three populations of alleles in Hardy-Weinberg equilibrium (K = 3). These STRUCTURE results in light of phylogenetic reconstruction indicated that, Trinidadian individuals are distinct from the mainland, but most alleles are shared across the mainland. However, there is a small proportion of unique alleles specific to eastern Amazonia (Fig. 1, Supplementary Fig. 2), which could be due to a variety of factors (see Discussion). Alleles belonging to the allopatric Trinidad population were distinct from those of the mainland (‘orange’) (Fig. 1, Table 1), so we excluded Trinidadian individuals from certain subsequent population-level analyses (i.e. neutrality tests and testing for IBD). Further investigation into the population structure and demographic history of mainland G. humeralis involved three analyses. (i) we tested against a neutral model of molecular evolution for evidence of rapid population expansion across the mainland, and we looked for concordance between two test statistics, Tajima’s D and Fu’s Fs. Neither test showed a deviation from neutrality for either the mitochondrial or RADseq SNP data (Table 3). (ii) we tested for the presence of IBD across the mainland using Mantel’s test (Tables 1 and 2) by correlating a matrix of pairwise genetic distances and a matrix of geographic distances. This analysis revealed strong evidence for IBD across mainland South America (Table 2, R^2 = 0.637, p-value = 0.001). (iii) we estimated within-population genetic distance (p-distance) and within-population nucleotide diversity (π) for each population and the outgroup, G. antillensis, for mtDNA (Supplemental Table 3). These measurements showed that G. humeralis from eastern Amazonia exhibits more genetic diversity than western populations, and that Trinidadian G. humeralis display very little genetic diversity overall when compared to mainland populations.

3.2. Phylogenetic inference

Phylogenetic relationships at well-resolved nodes was largely concordant across the methodologies and data sets used (Figs. 1 and 2b). ML and Bayesian methods recovered reciprocally monophyletic Trinid-
dian and mainland populations using RADseq and Sanger sequenced mitochondrial and nuclear datasets (Figs. 1 and 2, Supplemental Figs. 3, 4, and 5). Indeed, overall relationships among mainland populations were concordant at well-supported nodes, with a broader Amazonian clade containing a nested monophyletic group from western Amazonia. Between-group mean genetic distances among G. humeralis phyloge
graphic clusters ranged from 0.05–0.1 and 0.012–0.023 for ND2 and 16S, respectively (Table 1). Divergence times between Trinidad and mainland G. humeralis lineages were estimated to occur in the early Pleistocene: 1.89 mya (0.90–2.42, 95% HPD) (Fig. 2; Supplemental Fig. 5) using a secondary calibration and 2.7 mya [2.45–2.91] assuming a strict clock using the published ND2 rate calibration in geckos [p-distance = 0.094 ± 0.008]. There was more consensus on the esti-
mated divergence time between populations in eastern and western Amazonia, where mean values varied from 1.59 [0.13–3.0]
3.3. Species delimitation

We utilized three well-documented statistical species delimitation methods (PTP, STACEY, BFD) to examine species limits between the three phylogeographic clusters previously identified by STRUCTURE (Fig. 1). Analysis of our mtDNA gene tree using PTP revealed significant species-level divergence between Trinidad and mainland clades (p-value = 0.001), but not between eastern and western Amazonia (Supplemental Fig. 3). Analysis of the multi-locus Sanger sequenced dataset with STACEY and SpeciesDA (SSDA) supported the Trinidad and mainland South American clades as being distinct, species-level lineages (pp = 0.999) (Table 2, Supplemental Fig. 4). SSDA analyses also yielded an additional species delimitation hypothesis within the mainland, identifying populations from eastern Amazonia and western Amazonia as separate species (Table 2, Supplemental Fig. 6). We used

<table>
<thead>
<tr>
<th>Geographic Cluster</th>
<th>Mantel's Test</th>
<th>Neutrality Test</th>
<th>Species Delimitation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R-square</td>
<td>P-value</td>
<td>Data</td>
</tr>
<tr>
<td>Trinidad</td>
<td>0.504</td>
<td>0.794</td>
<td>mtDNA</td>
</tr>
<tr>
<td>Mainland</td>
<td>0.637</td>
<td>0.001***</td>
<td>RADseq</td>
</tr>
<tr>
<td>Mainland (East)</td>
<td>0.045</td>
<td>0.386</td>
<td>RADseq</td>
</tr>
<tr>
<td>Mainland (West)</td>
<td>0.594</td>
<td>0.162</td>
<td>mtDNA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RADseq</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Geographic Cluster</th>
<th>Mantel's Test</th>
<th>Neutrality Test</th>
<th>Species Delimitation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R-square</td>
<td>P-value</td>
<td>Data</td>
</tr>
<tr>
<td>Trinidad</td>
<td>0.504</td>
<td>0.794</td>
<td>mtDNA</td>
</tr>
<tr>
<td>Mainland</td>
<td>0.637</td>
<td>0.001***</td>
<td>RADseq</td>
</tr>
<tr>
<td>Mainland (East)</td>
<td>0.045</td>
<td>0.386</td>
<td>RADseq</td>
</tr>
<tr>
<td>Mainland (West)</td>
<td>0.594</td>
<td>0.162</td>
<td>mtDNA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RADseq</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Geographic Cluster</th>
<th>Mantel's Test</th>
<th>Neutrality Test</th>
<th>Species Delimitation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R-square</td>
<td>P-value</td>
<td>Data</td>
</tr>
<tr>
<td>Trinidad</td>
<td>0.504</td>
<td>0.794</td>
<td>mtDNA</td>
</tr>
<tr>
<td>Mainland</td>
<td>0.637</td>
<td>0.001***</td>
<td>RADseq</td>
</tr>
<tr>
<td>Mainland (East)</td>
<td>0.045</td>
<td>0.386</td>
<td>RADseq</td>
</tr>
<tr>
<td>Mainland (West)</td>
<td>0.594</td>
<td>0.162</td>
<td>mtDNA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RADseq</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Geographic Cluster</th>
<th>Mantel's Test</th>
<th>Neutrality Test</th>
<th>Species Delimitation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R-square</td>
<td>P-value</td>
<td>Data</td>
</tr>
<tr>
<td>Trinidad</td>
<td>0.504</td>
<td>0.794</td>
<td>mtDNA</td>
</tr>
<tr>
<td>Mainland</td>
<td>0.637</td>
<td>0.001***</td>
<td>RADseq</td>
</tr>
<tr>
<td>Mainland (East)</td>
<td>0.045</td>
<td>0.386</td>
<td>RADseq</td>
</tr>
<tr>
<td>Mainland (West)</td>
<td>0.594</td>
<td>0.162</td>
<td>mtDNA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RADseq</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3

Species delimitation models compared using Bayes factors with BFD, ranked by marginal likelihood estimate (MLE). Bayes factors reported as pairwise comparisons of a randomized 3-taxon model versus being listed by each model (Bayes factor = \[2 * (\text{MLE model 1} - \text{MLE model 2})\]). Pairwise ln(BF) calculations select both the 2-taxon (10.4) and 3-taxon (10.8) models as being significantly better than random species assignments using the Kass and Raftery (1995) scale; where \(\text{ln(BF)} \geq 5\) there is strong support for the model with the higher MLE. Pairwise comparison between 2-taxon and 3-taxon models results in a ln(BF) = 9.5, providing decisive support in favor of the 3-taxon model.

<table>
<thead>
<tr>
<th>Taxon Statement Model Tested</th>
<th>MLE</th>
<th>Rank</th>
<th>Bayes Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized 3-taxon Statement</td>
<td>−104081.08</td>
<td>3</td>
<td>−</td>
</tr>
<tr>
<td>2-taxon (Trinidad &amp; Mainland)</td>
<td>−87439.54</td>
<td>2</td>
<td>16645.86</td>
</tr>
<tr>
<td>3-taxon (Trinidad, East, &amp; West)</td>
<td>−80789.26</td>
<td>1</td>
<td>6645.96</td>
</tr>
</tbody>
</table>

Fig. 2. Phylogenetic inference using two Bayesian inference methods. (A) Time-calibrated StarBEAST2 multi-locus phylogenetic inference (trimmed from Supplemental Fig. 5). Red dots at nodes indicate nodal support ≥0.95 posterior probability. Scale in millions of years before present (mya) and geological era indicated via shaded boxes (Plio = Pliocene, Ple = Pleistocene). (B) Mitochondrial gene tree generated with ND2 and 16S on zoomed in region from part A. Numeric values indicate posterior probability support for the adjacent node. Shallow, haplotype-level support values are removed for clarity. Precise posterior support for all nodes, however, are reported in cladogram format in Supplemental Fig. 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
BFD to compare the two-species (Trinidad + mainland) model, favored by PTP, and the three-species (Trinidad + eastern Amazonia + western Amazonia) model, favored by SSDA, using our RADseq data in a coalescent framework. Pairwise Bayes Factors (BF) calculations selected both the 2-taxa \( \text{ln}(\text{BF}) = 10.4\) and 3-taxa \( \text{ln}(\text{BF}) = 10.8\) models as being significantly better than random species assignments using the Kass and Raftery (1995) scale; if \( \text{ln}(\text{BF}) \geq 5\) there is strong support for the model with the higher MLE. The pairwise comparison between 2-taxon (PTP) and 3-taxon (SSDA) models provided stronger support for the 3-taxon model \( \text{ln}(\text{BF}) = 9.5\) (Table 3, Supplemental Fig. 4). To further examine the feasibility that \textit{G. humeralis} from eastern and western Amazonia belong to distinct species, we tested whether our data supported reciprocal monophyly between the populations using topology tests by generating constraint trees for each dataset (trees not shown). Indeed, both SH and AU tests rejected the hypothesis that eastern and western Amazonian populations are reciprocally monophyletic, using the RADseq SNP data (SH p-value < 0.0001, AU p-value ~ 0) and mtDNA data (SH p-value = 0.0055, AU p-value = 0.0006).

4. Discussion

Phylogenetic analyses recovered \textit{G. humeralis} populations from Trinidad as sister to mainland populations, with a western Amazonian clade nested within populations from eastern Amazonia (Figs. 1 and 2b). Furthermore, STRUCTURE analysis inferred three populations of alleles in Hardy-Weinberg equilibrium \( K = 3\), with no individuals belonging purely to the third “ghost” population (‘purple’). This STRUCTURE pattern can be the result from two scenarios (Lawson et al., 2018): (1) admixture with an extinct/unsampled population or (2) genetic diversity in eastern Amazonia that did not establish in western populations, potentially through isolation-by-distance (IBD) mediated gene flow or a population bottleneck during stepwise westward range expansion. Distinguishing between scenarios (1) and (2) is difficult and they are not mutually exclusive. At present, testing for admixture, scenario 1, is not possible with our current sampling as individuals from the putative “ghost” population are also needed. It’s possible that increased sampling across the Guiana Shield could identify \textit{G. humeralis} populations that harbor an increased frequency of these “ghost alleles”. Indeed, population differentiation in this region has been noted previously for other taxa (Noonan and Gaucher, 2005). However, we posit (2) is a more likely scenario, i.e. extensive genetic diversity specific to eastern Amazonian populations, for three reasons: (i) we found much greater genetic diversity in eastern Amazonia (Supplemental Table 3) and little evidence for shared mtDNA haplotypes between localities, as did Avila-Pires et al (2012), which would be expected under this scenario; (ii) we recovered a signal of IBD across the mainland, which could account for the eastern specificity of these alleles via dropout; and (iii) western Amazonian populations are monophyletic, which would be expected if there were a population bottleneck during westward colonization. However, apart from weighing these lines of evidence, the current state of knowledge and our current sampling provide no definitive way of differentiating them. Thus, future work may warrant further examination of these possibilities.

Our phylogenetic and STRUCTURE results informed the possibility that Trinidadian divergence from the mainland is sufficient to warrant taxonomic reevaluation. Examining species limits using multiple methods and data types consistently identified the Trinidad populations as distinct species from the mainland populations, while a subset of analyses (SSDA & BFD) further split populations from eastern and western Amazonia. We first address whether the Trinidad populations represent a distinct species from the mainland populations, and then discuss whether the South American populations consist of one or more species.

All species delimitation analyses recovered Trinidadian populations as being distinct from Amazonian \textit{G. humeralis} (Table 2; Supplemental Figs. 3 and 4). Additionally, uncorrected genetic distances in mitochondrial ND2 (10%) between Trinidad and eastern populations (Table 1) are comparable to mitochondrial genetic distances among other recognized sister species of geckos, which typically range from 4.1% to 35.5% (Botov et al., 2015; Grismer et al., 2014a, 2014b, 2017; Oliver et al., 2007; Pepper et al., 2006; Portik et al., 2013). Although species delimitation based solely on pre-determined sequence divergence values is difficult, if not impossible, to justify due to variations in effective population sizes and lineage-specific substitution rates (Barracough et al., 2009; Moritz and Cicero, 2004; Pons et al., 2006), genetic distances among putative taxa can highlight taxa that warrant closer examination using other species delimitation methodologies (Gamble et al., 2012a; Hickerson et al., 2006), e.g. PTP, SSDA, and BFD. Thus, the bulk of the evidence supports recognition of the Trinidadian population as an independently evolving metapopulation lineage, or species (de Queiroz, 2007), distinct from mainland \textit{G. humeralis}. Because the type locality of \textit{G. humeralis} is from Peru (Guichenot, 1855; Rivero-Blanco, 1979), mainland South American populations should retain that name. Geckos on Trinidad, however, were previously described as \textit{G. ferrugineus} (Cope, 1864) and we resurrect that name from synonymy for the Trinidadian population and briefly discuss its unusual nominal history.

\textit{Gonatodes ferrugineus} has a complex taxonomic history (see supplement for complete synonymy). Cope (1864) described \textit{G. ferrugineus} from material collected on Trinidad that Theodore Gill deposited in the Smithsonian. Although the original description was ambiguous, and the type presumably lost (Rivero-Blanco, 1979), Cope (1868) later identified a \textit{G. ferrugineus} specimen (presumably being unaware of \textit{G. humeralis}) among a collection of lizards from Peru and thus later naturalists assumed that \textit{G. ferrugineus} was morphologically similar-to, and perhaps a junior synonym of, \textit{G. humeralis} (Guichenot, 1855). \textit{Gonatodes ferrugineus} was eventually synonymized with \textit{G. humeralis}, although no justification was provided for the decision (Donoso-Barros, 1968). However, throughout the late 19th and most of the 20th centuries discrepancies in nomenclature were apparent. Some herpetologists appeared to be unaware of \textit{G. ferrugineus} and listed \textit{G. humeralis} as occurring on Trinidad, likely based on their own experiences with this species while working in South America (Parker, 1935; Roux, 1926). Others listed \textit{G. ferrugineus} as occurring on Trinidad and \textit{G. humeralis} on the mainland (Boulenger, 1885; Burt and Burt, 1933). Wermuth (1965) added to the confusion by indicating that both \textit{G. ferrugineus} and \textit{G. humeralis} co-occur on Trinidad. However, following the explicit synonymy of Donoso-Barros (1968) and Rivero-Blanco’s thorough scholarly review (1979), synonymy of \textit{G. ferrugineus} with \textit{G. humeralis} was unanimously accepted (Avila-Pires, 1995; Kluge, 1991, 1995, and 2001).

\textit{Gonatodes ferrugineus} is currently morphologically indistinguishable from \textit{G. humeralis} although there appear to be some qualitative differences in proportionality of the face, body size, and coloration in adult males that may, upon further investigation, diagnose this species (Authors’ pers. obs.; Rivero-Blanco, 1979). Coloration may be particularly useful as adult males from Trinidad are generally not as colorful as those from mainland South America (Supplemental Fig. 1). Trinidadian males lack red spots on the sides of the body and their heads tend to favor orange/yellow rather than red and white/blue, both of which are typical features of most South American populations (Authors’ pers. obs.; Rivero-Blanco, 1979). Similarly-colored males to those from Trinidad have also been observed in northern Venezuela (Rivero-Blanco, 1979), leading to the possibility that \textit{G. ferrugineus} occurs there as well (Supplemental Fig. 1). Indeed, several Trinidadian endemics exhibit distributions that extend into northern Venezuela, such as \textit{Gonatodes ceciae}, \textit{Gonatodes vittatus}, \textit{Polychrus auduboni}, and \textit{Flectonotus fitzgeraldi} (Murphy, 1997; Murphy et al., 2017a). Further, previous studies that have examined morphological variation within \textit{G. humeralis} have not included specimens from Trinidad (Avila-Pires, 1995; Avila-Pires et al., 2012; Vitt et al., 1997). Thus, future work should attempt to identify
diagnostic phenotypic differences to complement the identified genotypic characters between these two species and determine the geographical boundaries of these species (Supplemental Fig. 1). It is worth pointing out a gap in our sampling from the northern Guiana Shield to Trinidad. Indeed, having not sampled Venezuelan populations may confound species delimitation metrics. However, we find this unlikely as we see no evidence of gene flow between Trinidad and the mainland, even when K = 2 (Supplemental Fig. 2) and 9.4% pairwise divergence at the mitochondrial locus ND2 is considerable, and likely reflects substantial reproductive isolation.

Although G. ferrugineus was revealed to be unambiguously distinct from mainland populations in all analyses, the status among South American populations was less straightforward. SSDA and BFD both provided support for a species delimitation model that splits mainland G. humeralis into two species, occupying eastern and western Amazonia (Tables 2 and 3, Supplemental Fig. 4). This hypothesis was bolstered by the fact that western Amazonia did not possess a large proportion of eastern-specific alleles (Fig. 1b) and that western Amazonia is monophyletic, although not reciprocally monophyletic with relation to eastern populations (Fig. 1a). These data are also congruent with previous work showing that the western Amazonian populations exhibit ecological differences compared to eastern populations. Namely, eastern G. humeralis occurs in primary forest, whereas western G. humeralis occur frequently in clearings, secondary forests, and human dwellings (Vitt et al., 1997). Additionally, a model that supports a parapatric mode of speciation across Amazonia would support the gradient hypothesis of Amazonian biogeography (Endler, 1977). However, there is emerging evidence that intraspecific, population-level processes can confound assumptions made by coalescent species delimitation methods, such as SSDA and BFD (Ahrens et al., 2016; Barley et al., 2018; Gratton et al., 2015; Sukumaran and Knowles, 2017). This includes processes such as IBD, which we identified in our mainland samples, that can result in oversplitting species even in well-represented, continuously sampled populations. When considered in conjunction with our relatively sparse sampling, particularly in central Amazonia (Fig. 1), it is most likely that SSDA and BFD mis-interpreted this structure as speciation, and thus oversplit the mainland clade. Additionally, for both the mtDNA and RADseq data, eastern and western populations are not reciprocally monophyletic. While reciprocal monophyly at any specific locus is not a prerequisite for species delimitation (Hudson and Coyne, 2002; Palumbi, 2001), rapidly coalescing loci like mtDNA frequently form monophyletic sister species, reflecting their reproductive isolation (Wiens and Penkrot, 2002; Zink and Barrowclough, 2008). Thus, the failure to recover reciprocal monophyly, coupled with high proportions of shared alleles between eastern and western lineages, supports a single-species hypothesis for mainland, i.e. G. humeralis sensu stricto.

Our estimates of the divergence time between mainland Amazonia and Trinidad are moderately disparate (mean = 1.89 mya (secondary calibration) and 2.7 mya (ND2 rate)). This is as expected, because gene divergence occurs prior to species divergence (Edwards and Beerli, 2000). Thus, we err on the side of the more-recent species divergence estimate of 1.89 mya (Fig. 2), which then suggests that cladogenesis between G. ferrugineus and G. humeralis took place in the early- to mid-Pleistocene, coinciding with the published divergences separating sister taxa in other organisms distributed on Trinidad and South America, including: fishes (Jowers et al., 2008), frogs (Camargo et al., 2009), skinks (Hedges and Conn, 2012), and birds (Hunt et al., 2001). Concordance across animal clades is suggestive of a large-scale isolating event between groups of organisms on Trinidad and South America during this time-period due to Pleistocene glacial cycles. However, these divergences are ancient considering recent connections between the Paria peninsula of Venezuela and Trinidad as recently as 10,000 years ago (Comeau, 1991). This transient connector may have also provided G. ferrugineus with the means of re-colonizing the mainland in a similar manner to G. ceciliae and G. vitatus (Supplemental Fig. 1). This possibility presents an interesting testable hypothesis of testing co-divergence of these lineages. Nonetheless, testing this hypothesis using a model-based biogeographic analysis (such as Ree et al., 2005) is currently not possible, as we are still lacking a fully sampled Gonatodes phylogeny (Gamble et al., 2008; Schargel et al., 2010; Russell et al., 2015).

We are currently unable to devise definitive tests to differentiate between three competing phylogeographic scenarios: (1) Trinidad and mainland populations were isolated via vicariance during Pleistocene glacial cycles, (2) dispersal to Trinidad via river flotsam (from the Orinoco or other nearby river), or (3) the inverse scenario, dispersal to the mainland from Trinidad. Given the current data, we are unable to ascertain the approximate distribution of the most recent common ancestor to G. humeralis and G. ferrugineus. As discussed above, western Amazonian populations are nested within eastern populations of G. humeralis, excluding the possibility of an Andean origination (Fig. 1). G. humeralis possesses significantly greater genetic diversity in eastern Amazonia than G. ferrugineus, which suggests a founder effect bottleneck on Trinidad via (1) vicariance or (2) riverine dispersal and discourages (3) the inverse possibility of dispersal from Trinidad to South America (Supplemental Table 2). In many cases, high levels of genetic diversity correlate with a lineage’s point-of-origin as genetic diversity accumulates over time in stable populations (Ingman et al., 2000; Kimura, 1983). In addition, most Gonatodes species occur in South America, including a member of G. humeralis sensu lato’s sister group, G. connicinus, suggesting a continental origin, with Caribbean species resulting from subsequent dispersals from the mainland (Supplemental Fig. 1), unlike Anolis lizards (Glor et al., 2001). However, several species closely related to this clade, e.g. G. ocellatus, G. ceciliae, and G. antillensis (G. connicinus’ sister species), occur on islands north of South America, including Trinidad and Tobago (Supplemental Fig. 1). Thus, although the data are suggestive, these scenarios can, and should be, explicitly tested when sufficient data are available.

The identification of the recent radiation of G. humeralis across Amazonia provides a powerful framework for testing recent biogeographic theories using fine-scale sampling, given specific demographic and phyleogeographic predictions (Avila-Pires et al., 2012; Bush and Oliveira, 2006; Haffer, 1997; Prates et al., 2016; Wernicke et al., 2012). We found that G. humeralis does not diverge from a neutral model, suggesting a relatively constant population size over time. However, it is also known that small sample sizes (mainland N = 20) can confound true deviations from neutrality, although failure to diverge from a neutral model is also a common theme in Amazonian taxa and is not unique to G. humeralis (Lessa et al., 2003). This is still somewhat surprising since the divergence between G. humeralis in eastern and western Amazonia has occurred so recently (Fig. 2). This shallow time frame, however, provides the potential for Quaternary divergence hypotheses, namely the refuge (Haffer, 1969) and vanishing refuge (Vanzolini and Williams, 1981) hypotheses, to be tested by employing more fine-scale sampling than was available for this study. Thus, G. humeralis sensu stricto provides a model system for elucidating the recent history of Amazonia.

Within eastern Amazonia, our results are largely concordant with the findings of Avila-Pires et al. (2012), using mitochondrial data to infer high genetic diversity in eastern Amazonia (Supplemental Table 3). Along with the lack of genetic diversity in western Amazonia and on Trinidad, our data suggest the most recent common ancestor of G. humeralis sensu stricto occurred in eastern Amazonia, with subsequent westward expansion; as source populations typically have higher genetic diversity than their emigrated counterparts (Cann et al., 1987; Ingman et al., 2000). Previous investigations of geographic barriers that have affected G. humeralis have focused on riverine barriers (Avila-Pires et al., 2012). Rivers have played an important role in Amazonian biogeography by acting as barriers to gene flow in multiple taxa (Gracraft, 1985; Haffer, 1969; Oliveira et al., 2017; Wallace, 1852). However, there is little evidence that they have had much impact on the present-
day distribution of *G. humeralis*, as our time-calibrated phylogeny suggests that intraspecific divergence within *G. humeralis* took place < 2.4 mya (Fig. 2), which is more recent than the establishment of the present-day Amazon river (≥ 3.6 mya) or the paleo-Tocantins river (≈ 2.6 mya) (Figueiredo et al., 2009; Latrubesse et al., 2010). Future investigations, with more thorough geographic sampling, may be able to elucidate a role for riverine barriers in relation to migration and gene flow in *G. humeralis*. Furthermore, the adaptation(s) that have led to the unusually broad distribution of *G. humeralis* may be of greater macroevolutionary importance for further investigation. Here, we briefly discuss the current state of knowledge regarding *G. humeralis*’ lineage-specific adaptations.

### 4.1. Gonatodes as a phylogeographic model system

*Gonatodes humeralis* is distributed over a geographic range considerably larger than that of any of its congeners. Indeed, because most geckos exhibit small ranges, *G. humeralis* may possess one of the largest native ranges of any gecko species (Meiri et al., 2017; Roll et al., 2017). *Gonatodes humeralis* resembles its congeners in many respects, and there are several hypotheses to explain the large distribution of *G. humeralis*. The first involves increased thermal tolerance, which could allow *G. humeralis* to disperse across warm, open areas between forest fragments (Vanzolini and Williams, 1981). However, *G. humeralis* maintains the same body temperature as at least two congeners: *G. concinna* (Vitt and Zani, 1996); and *G. hasemani* (Vitt et al., 2000), and although it occupies slightly warmer microhabitats than *G. hasemani*, its thermal properties may be explained by differences in body size; as *G. humeralis* is the smallest member of its genus (Avila-Pires, 1995). To test this as a potential explanation for the relative success of *G. humeralis*, body and microhabitat temperatures for additional *Gonatodes* species will be needed (Hertz et al., 1995). Another hypothesis involves the presence of functionally adhesive digits in *G. humeralis*, and *G. ferrugineus*, a unique trait for these taxa (Higham et al., 2017; Russell et al., 2015). The gain and loss of adhesive toe pads in geckos has been hypothesized to represent a key innovation (Higham et al., 2017; Losos, 2011; Russell and Delaugerre, 2017). A key innovation is a behavioral or morphology adaptation that has the capacity to enhance competitive ability, relax adaptive trade-offs, or catalyze the exploitation of a novel resource, which, in turn enhances the number or longevity of a species (Hunter, 1998). Digital adhesion allows geckos to exploit vertical, low-friction surfaces and may have allowed *G. humeralis* to occupy habitats unavailable to its congeners, such as higher strata in the rainforest canopy or locomotion on a wide variety of substrates (Vitt et al., 1997; Russell et al., 2015). Although current genetic and fossil data are lacking to successfully correlate gain and loss of digital adhesion and diversification rates in geckos, it has been demonstrated that: (1) digital adhesion has been gained, and lost, multiple times throughout the evolutionary history of gecko lizards (Gekkota) (Gamble et al., 2012b), (2) under different environmental conditions, selection can favor the presence or absence of adhesive digits (Russell and Delaugerre, 2017), (3) the evolution of functional adhesion requires few morphological changes (Russell et al., 2015), and (4) small morphological changes can have marked impacts on function and the success of a lineage (Burggren, 1992; Higham et al., 2015, 2016; Hunter, 1998; Liem, 1973; Russell, 1979; Thomason and Russell, 1986; Webb, 1982). Although, key innovations are generally discussed in the context of adaptive radiations (Farrell, 1998; Stroud and Losos, 2016), it is evident that we witness evolutionary processes as a snapshot in time and, given a strong environmental impetus, a well-adapted (successful) lineage with a broad range may also be a lineage that is primed for subsequent diversification (Endler, 1977; Haffer, 1969). Thus, digital adhesion, which is absent from all other *Gonatodes* species, provides a putative mechanism for *G. humeralis sensu lato*, relative to other members of the genus, to have capitalized on available ecological opportunity across Amazonia and on Trinidad (see Wellborn and Langerhans, 2014) for a scholarly review of ecological opportunity).

### 5. Conclusion

We propose that *G. humeralis sensu lato* is composed of two species. (1) *G. humeralis sensu stricto* occupies mainland South America and (2) its sister species, *G. ferrugineus*, resides allopatrically on the island of Trinidad. However, we reject the hypothesis that *G. humeralis* is a species-complex made up of multiple species across Amazonia. More specifically, genetic analyses support the hypothesis that *G. humeralis sensu stricto* is a single species throughout its contiguous range across northern South America with substantial population structure (local diversity and IBD). This is extremely atypical for a small, non-volant Neotropical taxon, and this pattern contrasts with that of most Amazonian taxa, as well as other species of *Gonatodes*, which occupy small, disjunctive distributions, and this discrepancy in geographic range invites further investigation. Indeed, unlike many clades consisting of widespread Neotropical taxa, *Gonatodes* harbors both widespread and geographically restricted taxa, providing a powerful system for identifying traits that influence species distributions. Thus, future work should attempt to elucidate the evolutionary adaptations that have influenced the biogeography of *Gonatodes*.

### Author contributions

B.J.P. assisted in study design, performed lab work, analyzed data, and wrote the manuscript. G.R.C. conducted fieldwork. T.E.H. assisted in study design. A.P.R. assisted in study design and manuscript preparation. D.P.S. and L.J.V. conducted fieldwork. T.G. conducted fieldwork, assisted in study design, performed lab work, and manuscript preparation. In addition, all authors read and approved the final manuscript.

### Acknowledgements

Authors would like to thank S. Estrada-Villegas, S.V. Nielsen, and A.H. Griffin for helpful comments and insightful discussion on early drafts of the manuscript; anonymous reviewers for constructive and helpful comments on the manuscript; S. Meiri for providing data for our geographic area estimations; A. Schultz and A. M. Simons for lab help on an earlier iteration of this project; and Ambrose Monell Cryo Collection (AMCC) at the American Museum of Natural History, R. Brumfield and LSU Museum of Natural Science Collection of Genetic Resources, M. Forstner, R. M. Murphy (ROM), J. Simmons (RU), and G. Watkins–Colwell (YPM) provided tissues from material in their care. B. J. Pinto was supported in-part by NSF DEB1657662 (to T. Game); G. R. Colli thanks Coordenação de Apoio à Formação de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Apoio à Pesquisa do Distrito Federal (FAPDF), and the Partnerships for Enhanced Engagement in Research (PEER) program for financial support. This work was also funded by National Science Foundation (NSF) IOS grant (1147043) and start-up funds from UC Riverside to T. E. Higham, The Natural Sciences and Engineering Research Council of Canada Discovery Grant (9745-2008) to A. P. Russell, NSF DDIG (DEB-1110605) to D. P. Sacleburnie, the University of Oklahoma Research Council via a George Lynn Cross Research Professorship to L. J. Vitt, NSF IOS grant (1146820) to D. Zarkower, and start-up funds (Marquette University) and Dayton/Wilkie Fund (Bell Museum of Natural History) to T. Gamble.

### Data accessibility

GenBank and Short Read Archive (SRA) accession numbers provided in Supplemental Table 1.

62
B.J. Pinto et al.

Molecular Phylogenetics and Evolution 133 (2019) 54-66


Lerat, B.J. Pinto et al.
Molecular Phylogenetics and Evolution 133 (2019) 54–66


Rambaut, A., Drummond, A., Xie, D., Baele, G., Suchard, M., 2018. Posterior summar-


ing in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67 (5), 901–904.


