Received: 15 November 2017 | Revised: 9 April 2018 | Accepted: 17 April 2018

DOI: 10.1111/mec.14715

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

A genomic evaluation of taxonomic trends through time in coast horned lizards (genus *Phrynosoma*)

Adam D. Leaché | Matthew T. McElroy | Anna Trinh

Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Seattle, Washington

Correspondence
Adam D. Leaché, Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195-1800, USA.
Email: leache@uw.edu

Funding information
National Science Foundation, Grant/Award Number: DBI-1144630

Abstract

Determining the boundaries between species and deciding when to describe new species are challenging practices that are particularly difficult in groups with high levels of geographic variation. The coast horned lizards (*Phrynosoma blainvillii*, *Phrynosoma cerroense* and *P. coronatum*) have an extensive geographic distribution spanning many distinctive ecological regions ranging from northern California to the Cape Region of Baja California, Mexico, and populations differ substantially with respect to external morphology across much of this range. The number of taxa recognized in the group has been reevaluated by herpetologists over 20 times during the last 180 years, and typically without the aid of explicit species delimitation methods, resulting in a turbulent taxonomy containing anywhere from one to seven taxa. In this study, we evaluate taxonomic trends through time by ranking 15 of these species delimitation models (SDMs) using coalescent analyses of nuclear loci and SNPs in a Bayesian model comparison framework. Species delimitation models containing more species were generally favoured by Bayesian model selection; however, several three-species models outperformed some four- and five-species SDMs, and the top-ranked model, which contained five species, outperformed all SDMs containing six species. Model performance peaked in the 1950s based on marginal likelihoods estimated from nuclear loci and SNPs. Not surprisingly, SDMs based on genetic data outperformed morphological taxonomies when using genetic data alone to evaluate models. The de novo estimation of population structure favours a three-population model that matches the currently recognized integrative taxonomy containing three species. We discuss why Bayesian model selection might favour models containing more species, and why recognizing more than three species might be warranted.

KEYWORDS
Bayes factor, ddRADseq, *Phrynosoma coronatum*, phylogeography, SNP, species delimitation, taxonomy

1 | INTRODUCTION

Species delimitation is the practice of determining the number of species and their boundaries and is a critical step for identifying the fundamental biological units of the Genealogy of Life. Genetic data play a crucial role in species delimitation, and a large number of methods are available that delimit species using genetic distances (Sites & Marshall, 2004), phylogeny (Pons et al., 2006) or coalescent models (Fujita, Leaché, Burbrink, McGuire, & Moritz, 2012). In general, genetic-based species delimitation methods can be grouped into two categories: species discovery vs. species validation (Carstens, Pelletier, Reid, & Satler, 2013). Species discovery methods do not require any assumptions regarding species membership and are therefore useful for initial assessments of biodiversity and discovering

new species (Kapli et al., 2017). Species validation aims to estimate the probability that predefined species represent independent evolutionary lineages (Hotaling et al., 2016). Taxonomic groups with long and detailed histories of systematic study provide a rich set of models for species validation model testing and investigating trends in the performance of species delimitation models (SDMs) through time.

Bayesian model comparison is a statistical framework for evaluating predefined SDMs and ranking competing taxonomies (Grummer, Bryson, & Reeder, 2013). Differences among the competing SDMs may include alternative partitioning schemes for the number of species, different assignments of populations into species or both. After defining SDMs that represent each taxonomic model, the marginal likelihood for the model can be estimated from multilocus genetic data. The marginal likelihood is the fit of a model to a data set that is weighted with respect to the prior, over the entire parameter space, and can therefore naturally compare models with different numbers of parameters (Oaks, Cobb, Minin, & Leaché 2018). The SDMs are ranked by their marginal likelihoods, and then, Bayes factors (Jeffreys, 1935) are used to assess support for model rankings (Kass & Raftery, 1995). The aim of this study was to evaluate existing taxonomic models using a Bayesian model comparison framework to investigate trends in taxonomy through time in a geographically variable and widespread terrestrial vertebrate species group with a long history of taxonomic instability.

The number of taxa recognized within the coast horned lizards (*Phrynosoma coronatum*, *Phrynosoma cerroense* and *Phrynosoma blainvillii*) has been reevaluated by herpetologists over 20 times since the original description of *P. coronatum* by De Blainville, 1835 (Table 1). The species complex has an extensive distribution spanning many distinctive phytogeographic regions from northern California to the Cape Region of Baja California, Mexico, and populations differ with respect to external morphology across much of this range (Grismer, 2002). An increased understanding of geographic variation in external morphology underlies most of the taxonomic changes made from the 1800s to the early 1900s, and these reappraisals resulted in seven new taxa, most with type localities separated by approximately 200 kilometres. By the 1930s, most of the extensive sampling gaps in the geographic distribution were filled, and areas of gradual variation were characterized (Klauber, 1936). However, the number of recognized taxa continued to fluctuate between two and six with the accumulation of new specimens and data, and differing views on whether taxonomic units should occupy the level of species, subspecies or pattern class contributed to further taxonomic revisions (Grismer & Mellink, 1994).

Using genomic data to rank taxonomic hypotheses in a group with a long history of study by many of the distinguished naturalists of the 19th and 20th centuries presents an opportunity to investigate trends in taxonomy through time. The first rigorous statistical examination of morphological and colour pattern characters within *P. coronatum* provided evidence for four distinct species, one of which, *Phrynosoma wigginsi*, was described as new (Montanucci, 2004). The first phylogeographic study of the group used mitochondrial DNA (mtDNA) and identified five major clades (Figure 1); however, an integrative taxonomy recognizing three species was advocated based on a combination of evidence from phylogeny, morphology, ecology and gene flow (Leaché et al., 2009). A more recent species delimitation analysis using two nuclear loci provided evidence for recognizing the five mtDNA clades as separate species (Yang & Rannala, 2014), although no taxonomic changes were made.

TABLE 1 Taxonomic hypotheses evaluated for coast horned lizards using Bayesian model comparison

Model[a]	Taxa[b]
De Blainville (1835)[c]	1
Stejneger (1893)	3
Van Denburgh (1894)	3
Bryant (1911)	4
Schmidt (1922)	5
Van Denburgh (1922)	4
Linsdale (1932)	5
Klauber (1936)[d]	5
Smith and Taylor (1950)	5
Reeve (1952)	6
Presch (1969)	2
Jennings (1988)	6
Montanucci (2004)	4
Leaché et al. (2009)	3
Yang and Rannala (2014)	5

Notes. [a]Additional models that could not be included are described in Materials and Methods. [b]Species or subspecies. [c]The one-species model was proposed again by Grismer and Mellink (1994) and Brattstrom (1997). [d]Tevis (1944) proposed a similar model.

We hypothesize that taxonomic accuracy should increase through time as more data and specimens become available for species delimitation analysis. If true, there should be a positive relationship between SDM performance (as measured by marginal likelihood scores) and publication date. Here, we evaluate 15 SDMs for coast horned lizards using coalescent species tree analyses within a Bayesian model comparison framework. For each SDM, we partition the samples to reflect the taxonomic model, and then rank SDMs by their marginal likelihood scores and test their support using Bayes factors. No distinction is made between the species and subspecies category, because species tree inference using the multispecies coalescent (MSC) model does not consider gene flow, and therefore treats both as independent and genetically isolated lineages. In addition to conducting species validation analyses by ranking preexisting SDMs, we also use species discovery approaches for the de novo estimation of species diversity and population structure.

2 | MATERIALS AND METHODS

2.1 | Sampling

We included representatives from all taxa described within the *Phrynosoma coronatum* complex, which facilitates an evaluation of

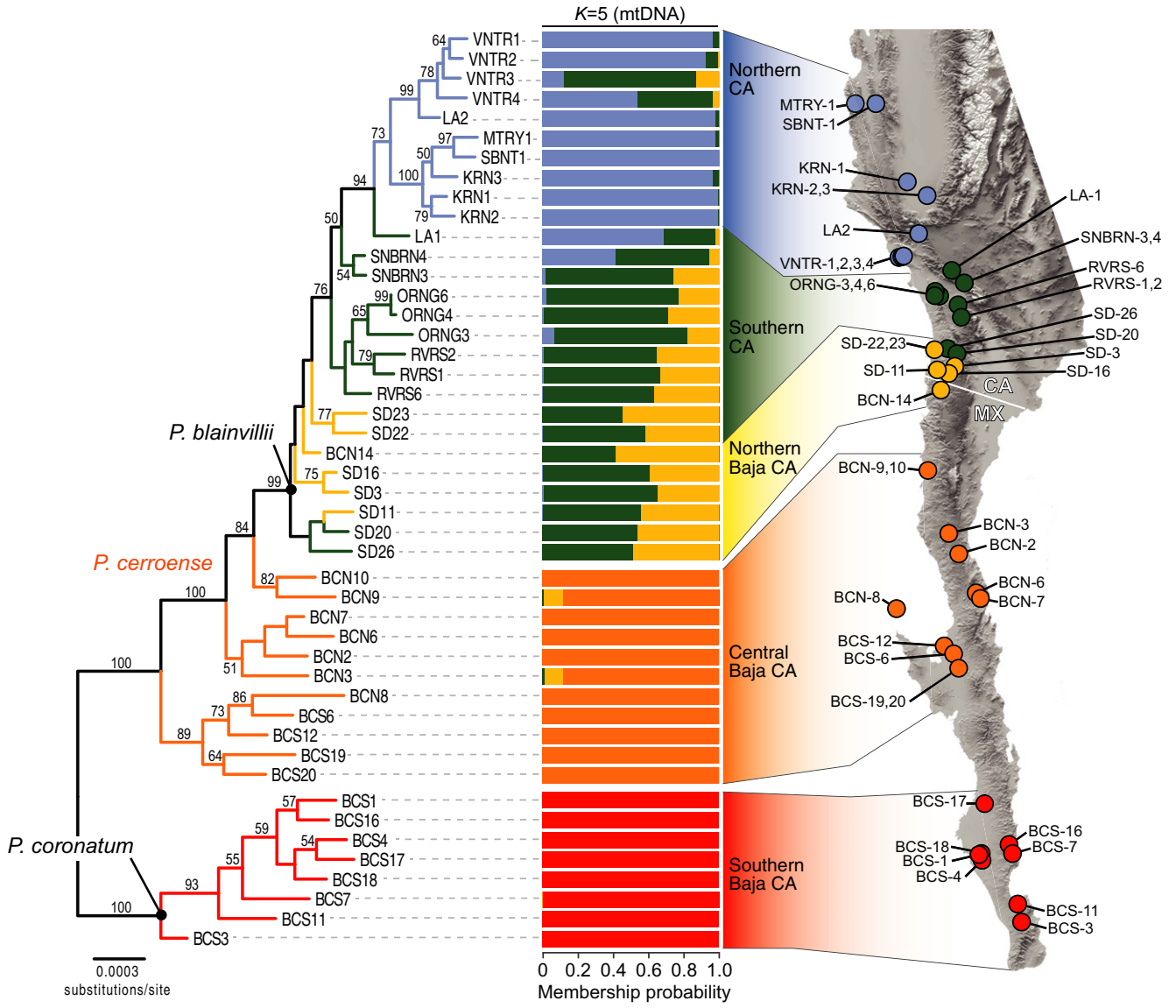

FIGURE 1 Phylogeographic structure of coast horned lizards. The maximum-likelihood phylogeny was estimated using a concatenated SNP matrix in RAXML v.8.2.0 (Stamatakis, 2014) with an HKY85 model (Hasegawa et al., 1985) and Felsenstein acquisition bias correction (Leaché, Banbury, et al., 2015). The phylogeny and probability of assignment of individuals to different clusters (estimated using the *compoplot* function in *adegenet*) are colour-coded to reflect the five mtDNA groups from Leaché et al. (2009). Bootstrap values calculated from 1,000 replicates ≥50% are shown. The SNP data suggest extensive admixture among the Californian populations [Colour figure can be viewed at wileyonlinelibrary.com]

most taxonomic proposals. For each SDM evaluated, we partitioned the samples to reflect the appropriate number of taxa, and the samples were assigned to their relevant partitions. Some published studies could not be included, because it was not possible to provide accurate species assignments. For example, Tevis (1944) did not discuss or illustrate the geographic limits of "*blainvillii*" and did not mention the status of "*cerroense*." Similarly, Cuesta Terron (1932) and Martín del Campo (1934) lacked the information required to establish accurate boundaries between taxa. In addition, the SDM proposed by Schmidt (1922) includes two sympatric species on Cedros Island (*Phrynosoma cerroense* and *Phrynosoma jamesi*); Cedros Island is represented by only one sample in our study, and it could only be assigned to one species. Later studies by Linsdale (1932), Cuesta Terron (1932), and Klauber (1936) recognized only one taxon on Cedros Island ("*cerroense*") and applied "*jamesi*" to population on the Baja California Peninsula. Supporting Information Figures S1 and S2 illustrate the geographic distributions assumed for taxa under each SDM.

2.2 | Nuclear genes

We used eight nuclear genes to compare the 15 SDMs (Table 1, Supporting Information Figure S1), including six loci sequenced for this study (*BACH1*, *EXPH5*, *NKTR*, *NOS1*, *PNN and R35*), and two

TABLE 2 Characteristics of the eight nuclear loci used for species delimitation

Gene	Sequences	Base pairs	Variable sites	Per cent variable
BACH1	38	939	19	2.0
BDNF	66	529	14	2.6
EXPH5	40	790	38	4.8
NKTR	41	550	19	3.5
NOS1	39	636	35	5.5
PNN	42	860	21	2.4
R35	39	701	21	3.0
RAG1	68	1054	44	4.2

additional loci downloaded from GenBank (*BDNF* PopSet: 255983692 and *RAG1* PopSet: 255983556). The number of samples sequenced for each locus ranged from 38 to 68 (Table 2; Supporting Information Table S1). One outgroup species, *Phrynosoma solare*, which is supported as one of the closest extant relatives of the *P. coronatum* complex (Leaché & Linkem, 2015), was included to permit an evaluation of a one-species model for the entire *P. coronatum* complex. Laboratory protocols for extraction, PCR and Sanger sequencing followed Nieto-Montes de Oca, Arenas-Moreno, Beltrán-Sánchez, and Leaché (2014). We aligned the nuclear loci using MUSCLE v3.8 (Edgar, 2004) and phased each locus computationally using the program PHASE v2.1.1 (Stephens & Donnelly, 2003; Stephens, Smith, & Donnelly, 2001). We selected nucleotide substitution model for each locus from among a set of 24 candidate models using the Bayesian information criterion (BIC) implemented in JMODELTEST v2.1.3 (Darriba, Taboada, Doallo, & Posada, 2012). We conducted species delimitation analyses using genotype alignments with heterozygotes coded using standard ambiguity codes, as well as with haplotype alignments using phased alleles. The phased data analyses included all alleles regardless of the level of phasing confidence. Removal of the low phase probability alleles removed (probability < 0.9) did not result in appreciable differences in SDM rankings (results not shown).

2.3 | SNP data

We collected new SNP data for 36 samples to evaluate SDMs (Table 3; Supporting Information Table S2) using double-digestion restriction site-associated DNA sequencing (ddRADseq; Peterson, Weber, Kay, Fisher, & Hoekstra, 2012). We downloaded data for ten additional samples used in a phrynosomatid lizard phylogeny study from the Sequence Read Archive database (SRA PRJNA294316; Leaché, Banbury, Felsenstein, Nieto-Montes de Oca, & Stamatakis, 2015). The laboratory procedure for library preparation follows Leaché, Banbury, et al. (2015). Briefly, we double-digested DNA using the restriction enzymes SbfI and MspI (New England Biolabs), followed by bead purification and ligation of barcoded Illumina adaptors, size selection and library quantification. Samples were sequenced on a single Illumina HiSeq 2500 lane (50-bp, single-end reads) shared with samples for other projects at the QB3 facility at UC Berkeley.

We processed raw Illumina reads using the program IPYRAD v.0.7.13 (Eaton, 2014). We clustered the filtered reads for each sample using the program VSEARCH v.1.1.0 (Edgar, 2010), using a clustering threshold of 88% and using previous ddRADseq assemblies of phrynosomatid lizards as a guide (Leaché, Banbury, et al., 2015; Leaché, Chavez, et al. 2015). We removed consensus sequences that had low coverage (<6 reads), excessive undetermined or heterozygous sites (>5), too many haplotypes (>2 for diploids) and an excess of shared heterozygosity among samples (paralog filter = 0.5). We used two different levels of missing data for the final SNP alignments. For population structure analyses, we excluded loci containing ≥50% missing data, and for coalescent species delimitation, we only included loci with no missing data. As a further filtering step, we excluded singleton SNPs. Allelic dropout (Arnold, Corbett-Detig, Hartl, & Bomblies, 2013) caused a drastic reduction in the number of shared SNPs when including an outgroup species, and this prevented the inclusion of the one-species SDM.

2.4 | Bayesian model comparison

We conducted Bayesian model comparison with the nuclear gene data using Bayes factor delimitation (BFD; Grummer et al., 2013). For each SDM, we conducted species tree estimation and calculated the marginal likelihoods necessary for model comparison using BEAST v.1.8 (Drummond, Suchard, Xie, & Rambaut, 2012). The species tree analysis uses the MSC to estimate posterior probability distributions for the species tree topology, divergence times and population sizes (Heled & Drummond, 2010). The species tree is estimated during marginal likelihood estimation, and therefore, the tree topology does not have to be assumed prior to analysis. The site models, clock models and gene trees were unlinked across loci, and the strict clock was applied to each locus. We applied the HKY model (Hasegawa, Kishino, & Yano, 1985) to each locus, because this model was included in the 95% credible set of models computed using the BIC (results not shown). Six replicate analyses were conducted with random starting seeds and chain lengths of 250 million generations with

TABLE 3 Summary of the ddRADseq data. Values are averages for all samples within a species

Species	*N*	Reads[a]	Passed[b]	Depth[c]	Loci[d]	Polymorphic[e]
P. blainvillii	27	1,752,384	1,751,647	11.5	4,271	0.00396
P. cerroense	11	1,562,597	1,562,154	14.7	3,820	0.00452
P. coronatum	8	1,269,113	1,268,575	11.9	3,467	0.00458

Notes. [a]Raw read count after sample demultiplexing. [b]Reads passing quality filters. [c]Mean sequencing depth per cluster. [d]Loci passing quality filters. [e]Frequency of polymorphic sites.

parameters sampled every 100,000 steps. The analysis with the highest marginal likelihood was used for model comparisons. The analyses achieved high effective sample sizes (≥200) for parameters, and this was used as a proxy for convergence. To estimate the marginal likelihood for each SDM, we used stepping-stone sampling (Baele, Li, Drummond, Suchard, & Lemey, 2013; Baele et al., 2012) with 100 steps (one million iterations each, 10% burn-in), with each step differing in the contribution of the data to the overall marginal likelihood.

We conducted Bayesian model comparison with the SNP data following the BFD* protocol (Leaché, Fujita, Minin, & Bouckaert, 2014). For each SDM, we conducted species tree estimation and calculated marginal likelihoods using SNAPP v1.3.0 (Bryant, Bouckaert, Felsenstein, Rosenberg, & RoyChoudhury, 2012) in BEAST v.2.4.8 (Bouckaert et al., 2014). We implemented a normal prior on the root of the species tree (mean = 8.4; sigma = 1.5) to obtain divergence times that match estimates from a phylogenomic study (6–11 million years; Leaché & Linkem, 2015). The mutation rates u and v were set to 1.0. We set the prior for the expected genetic divergence (θ) using a gamma distribution $\theta \sim G(2, 400)$ with a mean of alpha/beta = 0.005. We assigned a gamma hyperprior for the speciation rate parameter lambda (λ) $\sim$ G (2, 200) with a mean alpha × beta = 400. To estimate the marginal likelihood for each SDM, we performed a stepping-stone analysis using the PathSampleAnalyser (alpha = 0.3) with 40 steps (50,000 iterations, 12,000 pre-burn-in). We repeated each analysis twice using random starting seeds to ensure stable marginal likelihood estimation.

2.5 | Taxonomic trends

We investigated taxonomic trends by performing curve fitting and model selection in R v.3.4.0 (R Core Team 2017). We used the marginal likelihood score for each SDM as a measure of taxonomic performance. Models with higher marginal likelihood scores were considered to have outperformed those with lower scores. Several empirical studies have demonstrated that models containing more species tend to receive higher marginal likelihoods (Bryson et al., 2014; Nieto-Montes de Oca et al., 2017). Therefore, we tested for a positive relationship between SDM performance and the number of taxa in the model. We tested for a relationship between publication date and marginal likelihood score to investigate trends in the performance of SDMs through time. For each relationship, three linear models were fit to the data (first-, second- and third-degree polynomials), and their likelihood scores were compared using their Akaike information criterion (AIC) scores (Akaike, 1973). We ranked the regression models by their weighted AIC scores, and the best model was selected for regression analysis.

2.6 | Population structure

Given that many if not most of the taxonomic changes in this system reflect differing opinions on how to partition geographic variation into discrete units, we were interested in estimating population structure de novo using the multilocus data. First, we used the eight nuclear loci to estimate a spatially explicit population structure model using GENELAND v.4.0 (Guillot, Estoup, Mortier, & Cosson, 2005; Guillot, Santos, & Estoup, 2008). The method estimates posterior probabilities for the number of populations and sample assignments, and produces a graphical output of the spatial distribution of populations (Guillot et al., 2005). We conducted an initial analysis to identify the most probable number of population clusters (K = 1 to 6), as inferred by the posterior density of population clusters visited during the MCMC analysis (one million iterations, 10% burn-in). Next, we conducted an analysis under the K value that received the highest posterior density from step one. To check for convergence, we replicated the MCMC analyses twice.

We estimated population structure with the SNP data using the R package *adegenet* v.2.0.1 (Jombart, 2008; Jombart & Ahmed, 2011). We explored population cluster models using the *find.clust* function using 40 principal components (PCs) to maximize our ability to find groups. We conducted de novo population clustering for K = 1 to 10. To select the optimal population cluster model, we used the BIC. We also conducted discriminant analysis of principal components (DAPC, Jombart, Devillard, & Balloux, 2010) using the mitochondrial DNA (mtDNA) clade assignments (Leaché et al., 2009). Finally, we used the *compoplot* function to generate a barplot showing the probabilities for assignments of individuals to the five different mtDNA clades (Figure 1).

2.7 | Single-locus species discovery

We compared distance-based and tree-based single-locus species discovery methods using mtDNA data. These analyses used previously published mtDNA sequence data from three gene regions, *12S*, *ND1* and *ND2* (Leaché et al., 2009). The Automatic Barcode Gap Discovery (ABGD) method is a distance-based method that calculates the location of the barcode gap separating intraspecific and interspecific distances (Puillandre, Lambert, Brouillet, & Achaz, 2012). We used the default prior range for maximum intraspecific divergence (0.001 to 0.1) and a minimum slope increase (X) of 1.5. The mPTP model is a phylogeny-aware Poisson tree process that accommodates different rates of coalescence within clades and uses a phylogenetic tree (nonultrametric) to determine the transition points from speciation to coalescent processes (Kapli et al., 2017). The gene trees (one for each gene) were estimated using maximum-likelihood (ML) with the GTRGAMMA model in RAXML v.8.2.0 (Stamatakis, 2014). Redundant sequences were removed from the alignments to avoid unnecessary computations during the mPTP analysis. We used both ML and MCMC analyses with the multioption to enable variable coalescence rates and the default minimum branch length of 0.0001. The MCMC simulations used ten replicate MCMC runs (steps = 10M, sample interval = 1M, burn-in = 1M), each starting from a random delimitation.

3 | RESULTS

3.1 | Nuclear genes and SNP data

Table 2 provides a summary of the genetic diversity of the eight nuclear loci. The variability among the loci ranged from 2.0%

(*BACH1*) to 5.5% (*NOS1*), and these calculations are based on coast horned lizard genetic variation (the outgroup species *P. solare* is excluded). Illumina sequencing provided an average of 1.2 to 1.7 million raw sequence reads and an average sequencing depth of 12.7 × (Table 3). The final SNP assemblies contained 4,960 SNPs for population structure analysis (sampling one random SNP per locus and excluded loci containing ≥50% missing data), and 243 SNPs for species delimitation (allowing no missing data and after removing singletons).

3.2 | Model rankings

Table 4 provides the rankings of SDMs based on Bayesian analyses of the nuclear loci and the SNP data. A five-species SDM (Yang & Rannala, 2014) is the top-ranked model. The Bayes factors separating the top model from the others are decisive (≥12). The marginal likelihood estimates obtained using nuclear loci (genotypes or phased alleles) and SNPs place the same SDMs in the top five (Yang & Rannala, Jennings, Linsdale, Reeve, Smith & Taylor), but their rankings differ among the data sets (Table 4).

3.3 | Taxonomic trends

Curve fitting supported a parabolic relationship between publication year and the number of taxa, with the number of taxa increasing prior to the 1960s and decreasing after (adjusted $R^2 = 0.47$; $P = 0.0222$; Figure 2a). We inferred a plateauing positive relationship between the number of taxa in an SDM and the marginal likelihood score for nuclear genes (adjusted $R^2 = 0.8$; $P < 0.001$) and a positive linear regression between the SNP model performance and number of taxa (adjusted $R^2 = 0.277$; $P = 0.046$; Figure 2b). The number of species in an SDM does not necessarily predict model rankings. For example, several three-species models outperformed some four- and five-species SDMs, and the top-ranked five-species model outperformed the SDMs containing six species (Figure 2b). The relationship between publication year and model performance reaches a plateau in the 1950s, and this relationship is significant using nuclear genes (adjusted $R^2 = 0.4$; $P < 0.046$) and SNPs (adjusted $R^2 = 0.515$; $P = 0.016$; Figure 2c). Supporting Information Table S3 provides AIC results comparing different curve fitting models.

3.4 | Population structure

Analyses of the nuclear loci using a spatially explicit population structure model suggest that a three-population model best describes the data (posterior probability = 0.48; Figure 3a). The geographic distributions for populations under the $K = 3$ model coincide with the currently recognized distributions of *Phrynosoma blainvillii*, *P. cerroense* and *P. coronatum* (Figure 3a). Sample clustering using the SNP data with a de novo approach also suggests that $K = 3$ provides the optimal population model (Figure 3b), and the geographic distributions under that model match those estimated with the nuclear loci (Figure 3). Population models higher than $K = 3$ are nearly equivalent in BIC score and further divide samples from Northern California (results not shown). We also used DAPC analysis of the SNP data to calculate assignment probabilities of samples to mtDNA clade assignments; these assignments are

TABLE 4 Statistical ranks for 15 SDMs, ordered chronologically, including the marginal likelihood estimate (MLE) and Bayes factor (BF)

		Nuclear genes (alleles)			Nuclear genes (genotypes)			SNPs		
Model	Taxa	Rank	MLE	2lnBF	Rank	MLE	2lnBF	Rank	MLE	2lnBF
De Blainville (1835)[a]	1	15	−12204.0	+622	14	−9835.0	+208	–	–	–
Stejneger (1893)[b]	3	13	−12122.2	+458	13	−9815.2	+169	–	–	–
Van Denburgh (1894)	3	6	−11958.6	+131	9	−9756.0	+50	11	−6304.4	+1444
Bryant, (1911)	4	11	−11988.5	+191	12	−9768.2	+75	12	−6359.4	+1554
Schmidt, (1922)	5	7	−11961.0	+136	10	−9759.0	+56	9	−6152.9	+1141
Van Denburgh (1922)	4	10	−11985.4	+185	11	−9766.0	+70	10	−6215.0	+1265
Linsdale, (1932)	5	3	−11923.1	+60	2	−9737.1	+12	2	−5592.9	+21
Klauber, (1936)	5	8	−11962.6	+139	6	−9743.9	+26	7	−5893.3	+622
Smith & Taylor (1950)	5	5	−11938.6	+91	3	−9737.7	+14	5	−5820.3	+476
Reeve (1952)	6	4	−11925.3	+65	5	−9743.5	+25	4	−5820.0	+455
Presch (1969)[b]	2	14	−12172.6	+559	15	−9836.0	+210	–	–	–
Jennings, (1988)	6	2	−11915.8	+46	4	−9739.5	+17	3	−5600.1	+35
Montanucci, (2004)	4	12	−11994.2	+202	7	−9747.7	+34	8	−5984.8	+805
Leaché et al., (2009)	3	9	−11967.3	+149	8	−9748.8	+36	6	−5839.0	+513
Yang & Rannala (2014)	5	1	−11893.0	0	1	−9730.9	0	1	−5582.4	0

Notes. [a]The one-species model could not be tested with the SNP data due to a lack of shared data with the outgroup, *P. solare.* [b]The SNP analysis failed to convergence. The sample partitions are highly imbalanced, and the posterior estimates for population size and clock rate become trapped on extreme values.

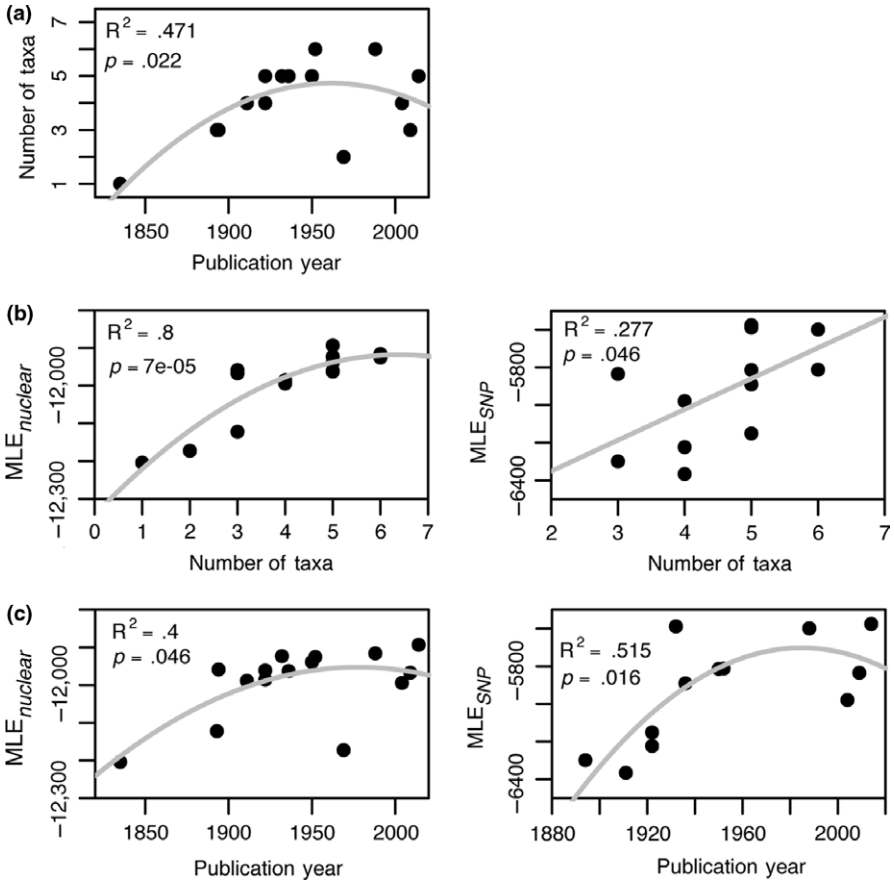

FIGURE 2 Taxonomic trends through time in coast horned lizards. (a) There is a nonlinear relationship between publication year and number of taxa with a peak in the mid-1900s. (b) SDMs containing more species tend to receive higher MLE scores. (c) The relationship between publication year and model performance based on nuclear genes ($MLE_{nuclear}$) and SNPs (MLE_{SNP}). The AIC scores for curve fitting models are available in Supporting Information Table S3

equivalent to the groupings used in the top-ranked five-species SDM (Table 3). The SNP data are able to distinguish the southern and central Baja California groups, but they cannot distinguish the Southern California populations where their genotypes appear admixed (Figure 1).

3.5 | Single-locus species discovery

Distance-based and tree-based species discovery using mtDNA data supported as few as four (*12S*) and as many as eight (*ND2*) putative species (Supporting Information Table S4). The ABGD and mPTP methods support the same number of species using the *12S* gene, although ABGD discovers fewer species than mPTP using both *ND1* (five vs. six) and *ND2* (five vs. eight). This high level of uncertainty in single-locus species discovery using mtDNA is similar to results reported for other species of *Phrynosoma* (Blair & Bryson, 2017).

4 | DISCUSSION

4.1 | Taxonomic trends

We investigated trends in the performance of SDMs through time in coast horned lizards, a taxonomic group with a long and detailed history of systematic study. First attempts at delimiting species in understudied groups are often limited by data availability, either in the form of few specimens, few data or both. Ideally, taxonomic hypotheses should become more robust through time as new data and specimens accumulate, and this should result in an increase in the performance of SDMs through time. However, differing philosophies regarding systematic practices can result in conflicting taxonomic hypotheses that are independent of the analysis of empirical data (Vane-Wright, 2003).

Using genetic data and marginal likelihood estimation as the arbiters of performance, we found that the accuracy of taxonomic hypotheses cannot necessarily be judged according to the decade or century when they were proposed. The taxonomic trend in coast horned lizards is for sharp fluctuations in the number of taxa recognized through time (Table 1; Figure 2a). Taxonomic trends in mammals support an increase in the number of taxa through time not through the discovery of new species, but largely due to taxonomic inflation stemming from reclassifications (Isaac, Mallet, & Mace, 2004). Regardless of what causes increasing numbers of taxa through time (species discovery or splitting), our evaluation of SDMs for coast horned lizards using a Bayesian model comparison framework does not support a positive linear relationship between SDM performance and publication date. Instead, both nuclear genes and SNPs suggest that SDM performance reached a plateau in the 1950s (Figure 2c). For coast horned lizards, taxonomic models have not necessarily improved through time. The main factors underlying fluctuations in the number of species recognized within coast horned lizards stem from different philosophies regarding lumping and splitting, and opinions on how to partition geographic variation into discrete units.

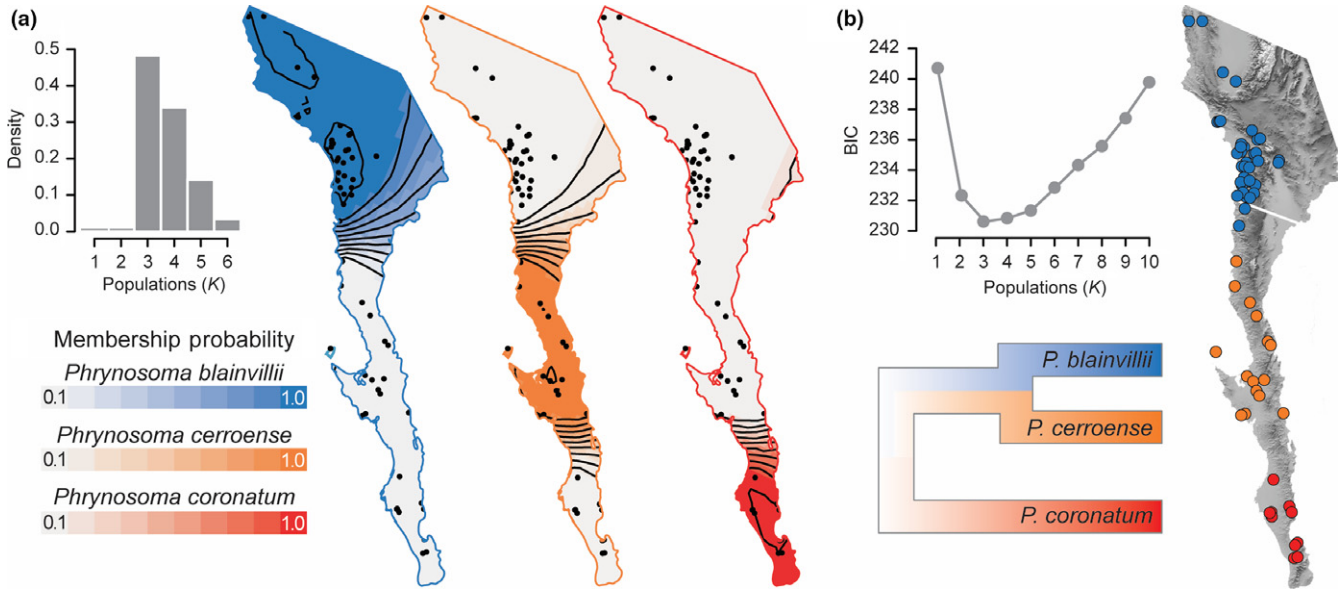

FIGURE 3 Three species of coast horned lizards are supported by de novo analyses of the nuclear genes and SNP data. (a) Spatially explicit estimation of population structure using the eight nuclear loci in Geneland. The posterior distribution for (*K*) and the geographic distributions of populations under the $K = 3$ model. Geographic areas with low population assignment probability are in white (0.1), and the contour lines depict the spatial change into regions of high assignment probability (1.0). (b) SNP-based species discovery; DAPC clustering inference using the BIC criterion in *adegenet* and the geographic distributions under the $K = 3$ model. The species tree topology was estimated with SNAPP. The *P. blainvillii* + *P. cerroense* clade has a posterior probability of 1.0 [Colour figure can be viewed at wileyonlinelibrary.com]

When conducting coalescent-based species delimitation, failing to partition samples into the correct groups can have a greater negative consequence on the performance of a model than lumping species. Several empirical studies have demonstrated that models containing more species tend to receive higher marginal likelihoods (Bryson et al., 2014; Nieto-Montes de Oca et al., 2017), raising the general concern that coalescent methods may result in oversplitting (Sukumaran & Knowles, 2017). We found that SDM performance varied considerably depending on the number of species in the model (Table 4; Figure 2b). Several three-species models outperformed some four- and five-species SDMs, and the top-ranked five-species model outperformed the six-species SDMs (Figure 2b). An important factor determining model performance is the number of species in a model, and this is often the focus of species validation model testing. However, for models containing the same number of species, accurately partitioning samples into species is critical. As a consequence, we predict that different taxonomic philosophies regarding lumping or splitting species could have less of an influence on model performance than does the correct identification of species boundaries.

4.2 | Bayesian model comparison

Despite being in the early stages of development, Bayesian model comparison has some advantages that make it particularly useful for comparing taxonomic hypotheses. First, taxonomic hypotheses often contain specific details on how to partition samples into the appropriate number of species, and how to assign samples to those species, resulting in explicit SDMs. Second, the SDMs can be nonnested, providing the flexibility to compare taxonomic hypotheses containing different numbers of species or different species assignments. Third, the methods integrate over gene trees, species trees and other model parameters, allowing SDMs to be compared and ranked without conditioning on any parameters being known, and without regard to the number of parameters in the models. Finally, using Bayesian model comparison approaches to evaluate and rank taxonomic hypotheses is a step towards helping taxonomy become a twenty-first-century information science (Adams, 2001; Godfray, 2002).

There are several critical limitations of Bayesian model comparison for species delimitation. Requiring predefined species assignments restricts the utility of the method to comparing existing models as opposed to searching among all possible species assignments, and this is an obvious disadvantage for studies aiming to discover cryptic diversity (Carstens et al., 2013). In the context of species validation, simulation studies have found that SDMs that lump species together are easier to distinguish from the true model vs. SDMs that oversplit samples into multiple species (Grummer et al., 2013; Leaché, Fujita, et al., 2014), suggesting that the method may be prone to oversplitting, especially if the SDMs are designed and tested using the same data. However, we note that the issue of oversplitting is not unique to Bayesian model comparison or even to molecular systematics; many of the coast horned lizard SDMs based on morphology have at least as many proposed taxa as those based on genetics.

Missing data can cause a potential problem during Bayesian model comparison where SDM rankings are influenced by unequal numbers of SNPs as well as by sample partitioning. RADseq

methods are prone to missing data (Wagner et al., 2013), and SDMs containing more species might contain fewer SNPs simply because SNAPP removes loci that do not contain alleles for each species. Conversely, SDMs containing few species will tend to have more SNPs. As a consequence, the model with the fewest SNPs (and most species) will have a higher marginal likelihood score compared to a model with more SNPs (and fewer species), and therefore, the model with the most species will rank highest (Noguerales, Cordero, & Ortego, 2018). We compared SDMs using equal numbers of the same SNPs, which should provide more accurate results that are reflective of differences in sample assignments and not levels of missing data.

A major limitation of current implementations of the MSC for species delimitation is that they do not consider gene flow. Incongruence among loci is assumed to be the result of incomplete lineage sorting by most MSC methods (including those used here), and distinguishing this process from gene flow is difficult (Leaché, Harris, Rannala, & Yang, 2014). Even when populations are admixed, as is the case with the California populations of *Phrynosoma blainvillii* (Figure 1), Bayesian model comparison using the MSC supports them as distinct. This overconfidence in selecting a model containing more species seems to be related to the more general problem in Bayesian analysis of choosing among incorrect models (Yang & Zhu, 2018). In terms of species delimitation with gene flow, if the true model is a two-species isolation–migration model, then an MSC model with no migration will view a two-species isolation model as less wrong than a one-species model, even if the migration rates are high enough to consider the populations a single species (Leaché, Zhu, Rannala, & Yang, 2018). In the context of coast horned lizards, the taxonomic implication of this problem is that Bayesian model selection favours the structured populations within *Phrynosoma blainvillii* as three separate species despite being connected by gene flow. The species validation tests are trying to select among models that are all incorrect because they ignore gene flow, and therefore, a model containing more species is favoured. There is a real need for the continued development of methods that can deal directly with gene flow during species delimitation (Jackson, Carstens, Morales, & O'Meara, 2017).

4.3 | Coast horned lizard systematics

Why has it been so difficult to identify species boundaries in the *Phrynosoma coronatum* species complex for the last 180 years? The root of the problem is that geographic variation in phenotypes and genotypes is often continuous and therefore difficult to bin into discrete taxonomic units (Vane-Wright, 2003; Wake, 2006). Early attempts at delimiting species in this group were limited by a lack of specimens from remote parts of Baja California, Mexico, which resulted in large sampling gaps. Klauber (1936) was the first to declare that the last sampling gap was filled and that specimens were finally available for a rigorous morphological study. However, more extensive geographic sampling did not help stabilize morphology-based taxonomy and instead resulted in the description of another new species (Montanucci, 2004). With respect to genetic data, large sampling gaps remain, and specimens from these regions are needed to refine population boundaries and to test for gene flow in areas of apparent phenotypic intergradation. Furthermore, multilocus genetic data can help identify populations that are divergent on deep timescales, enabling biologists to better interpret widespread geographic variation within species. If in the future additional genetic data support a model of isolation by distance or broad areas of genetic intergradation between species, then coast horned lizard taxonomy may continue to fluctuate.

The nuclear loci and SNP data analysed in this study both support $K = 3$ population models with similar population compositions (Figure 3). The de novo approach to population estimation using nuclear loci and SNPs provides additional evidence supporting the integrative taxonomy model for the group that already draws on information from morphology, ecology, genetics (mtDNA) and reproductive isolation (Leaché et al., 2009). The three species recognized under this model are *Phrynosoma blainvillii*, *Phrynosoma cerroense* and *P. coronatum*, with geographic distributions in northern Baja California and California, central Baja California (including Cedros Island) and southern Baja California, respectively (Figure 3). There is a general consensus among biologists that an integrative approach that combines diverse types of data from ecology, morphology and genetics will help increase taxonomic accuracy (Dayrat, 2005; Noguerales et al., 2018; Solís-Lemus, Knowles, & Ané, 2015).

The integrative taxonomy model containing three species is at odds with the outcome of Bayesian model selection, which favours five species. Integrative taxonomy is generally the preferred choice, as the consideration of multiple lines of evidence is more holistic and can strengthen taxonomic hypotheses (De Queiroz, 2007). The Bayesian model selection approach might be overconfident in its support for five species, because it is being forced to select from among misspecified models that do not account for gene flow. However, there could be some merit in recognizing additional species beyond just *P. blainvillii*, *P. cerroense* and *P. coronatum*. There is evidence for structured populations within *P. blainvillii* in California, including mtDNA support for three clades with limited sympatry, narrow geographic zones of admixture based on SNPs and the $K = 4$ population structure model (not shown) supports the Northern California population (*P. "frontale"*) as distinct. Fine-scale analyses of population structure within *P. blainvillii* across California are needed to determine the geographic extent of gene flow among these structured populations, which will assist in establishing the persistence and potential durability of these populations (Singhal, Hoskin, Couper, Potter, & Mortiz, 2018). *Phrynosoma blainvillii* is a Species of Special Concern in California (Thomson, 2016), and evidence for distinct genetic lineages in any region of California would have important implications for management and conservation.

ACKNOWLEDGEMENTS

We thank J. Grummer and K. Epperly for help with ddRADseq data collection. We thank R. Bouckaert for help with SNAPP analyses.

We thank B. Banbury, N. Bouzid, R. Bryson, I. Caviedes-Solis, J. Grummer, L. Jones, P. Kornilios, C. Linkem, M. Morando, J. Oaks, J. Richmond and Z. Yang for helpful discussions and comments on the manuscript. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. The research was funded by a National Science Foundation grant to A.D.L. (DBI-1144630).

DATA ACCESSIBILITY

Nuclear gene sequences: GenBank accessions GQ279665-GQ279730 and KU744987-KU745180. Demultiplexed ddRADseq data: NCBI SRA accession SRP070683 and SRP063316. BEAST XML files containing the final nuclear gene alignments and SNP assemblies for each SDM are available at the Dryad Digital Repository: https://doi.org/10.5061/dryad.k7k4m.

AUTHOR CONTRIBUTIONS

The study was designed by A.D.L. The ddRADseq data were collected by M.T.M. Nuclear gene sequences were collected by A.T. All authors conducted analyses, contributed to the text and approved the final manuscript.

ORCID

Adam D. Leaché http://orcid.org/0000-0001-8929-6300

REFERENCES

Adams, B. J. (2001). The species delimitation uncertainty principle. *Journal of Nematology*, *33*, 153–160.

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In B. N. Petrov, & F. Csaki (Eds.), *Second international symposium on information theory* (pp. 267–281). Budapest, Hungary: Akademiai Kiado.

Arnold, B., Corbett-Detig, R. B., Hartl, D., & Bomblies, K. (2013). RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology*, *22*, 3179–3190. https://doi.org/10.1111/mec.12276

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M. A., & Alekseyenko, A. V. (2012). Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, *29*, 2157–2167. https://doi.org/10.1093/molbev/mss084

Baele, G., Li, W. L. S., Drummond, A. J., Suchard, M. A., & Lemey, P. (2013). Accurate model selection of relaxed molecular clocks in Bayesian phylogenetics. *Molecular Biology and Evolution*, *30*, 239–243.

Blair, C., & Bryson, R. W. (2017). Cryptic diversity and discordance in single-locus species delimitation methods within horned lizards (Phrynosomatidae: Phrynosoma). *Molecular Ecology Resources*, *17*, 1168–1182. https://doi.org/10.1111/1755-0998.12658

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., … Drummond, A. J. (2014). EAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Brattstrom, B. H. (1997). Status of the subspecies of the coast horned lizard, *Phrynosoma coronatum*. *Journal of Herpetology*, *31*, 434–436. https://doi.org/10.2307/1565675

Bryant, H. C. (1911). The horned lizards of California and Nevada of the genera *Phrynosoma* and *Anota*. *University of California Publications in Zoology*, *9*, 1–84.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, *29*, 1917–1932. https://doi.org/10.1093/molbev/mss086

Bryson, R. W. Jr, Linkem, C. W., Dorcas, M. E., Lathrop, A., Jones, J. M., Alvarado-Diaz, J., … Murphy, R. W. (2014). Multilocus species delimitation in the *Crotalus triseriatus* species group (Serpentes: Viperidae: Crotalinae), with the description of two new species. *Zootaxa*, *3826*, 475–496. https://doi.org/10.11646/zootaxa.3826.3.3

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology*, *22*, 4369–4383. https://doi.org/10.1111/mec.12413

Cuesta Terron, C. (1932). Los camaleones mexicanos. *Anales del Instituto de Biologia, Universidad Nacional de México*, *3*, 95–121.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, *9*, 722.

Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society*, *85*, 407–415. https://doi.org/10.1111/j.1095-8312.2005.00503.x

De Blainville, H. M. D. (1835). Description de quelques especes de reptiles de la Californie precedee de l'analyse d'un systeme general d'herpetologie et d'amphibiologie. *Nouvelles annales du Museum d'histoire naturelle, Paris, France*, *3*, 232–296.

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*, 879–886. https://doi.org/10.1080/10635150701701083

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*, 1969–1973. https://doi.org/10.1093/molbev/mss075

Eaton, D. A. R. (2014). PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*, 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*, 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, *27*, 480–488. https://doi.org/10.1016/j.tree.2012.04.012

Godfray, H. C. J. (2002). Challenges for taxonomy. *Nature*, *417*, 17–19. https://doi.org/10.1038/417017a

Grismer, L. L. (2002). *Amphibians and Reptiles of Baja California, including its pacific islands, and the islands in the sea of Cortés*. Berkeley, CA, University of California Press.

Grismer, L. L., & Mellink, E. (1994). The addition of *Sceloporus occidentalis* to the herpetofauna of Isla de Cedros, Baja California, Mexico and its historical and taxonomic implications. *Journal of Herpetology*, *28*, 120–126. https://doi.org/10.2307/1564694

Grummer, J. A., Bryson, R. W., & Reeder, T. W. (2013). Species delimitation using Bayes factors: Simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). *Systematic Biology*, *63*, 119–133.

Guillot, G., Estoup, A., Mortier, F., & Cosson, J. F. (2005). A spatial statistical model for landscape genetics. *Genetics*, *170*, 1261–1280. https://doi.org/10.1534/genetics.104.033803

Guillot, G., Santos, F., & Estoup, A. (2008). Analysing georeferenced population genetics data with Geneland: A new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics*, *24*, 1406–1407. https://doi.org/10.1093/bioinformatics/btn136

Hasegawa, M., Kishino, H., & Yano, T. A. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, *22*, 160–174. https://doi.org/10.1007/BF02101694

Heled, J., & Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, *27*, 570–580. https://doi.org/10.1093/molbev/msp274

Hotaling, S., Foley, M. E., Lawrence, N. M., Bocanegra, J., Blanco, M. B., Rasoloarison, R., … Weisrock, D. W. (2016). Species discovery and validation in a cryptic radiation of endangered primates: Coalescent-based species delimitation in Madagascar's mouse lemurs. *Molecular Ecology*, *25*, 2029–2045. https://doi.org/10.1111/mec.13604

Isaac, N. J., Mallet, J., & Mace, G. M. (2004). Taxonomic inflation: Its influence on macroecology and conservation. *Trends in Ecology & Evolution*, *19*, 464–469. https://doi.org/10.1016/j.tree.2004.06.004

Jackson, N. D., Carstens, B. C., Morales, A. E., & O'Meara, B. C. (2017). Species delimitation with gene flow. *Systematic Biology*, *66*, 799–812.

Jeffreys, H. (1935). Some tests of significance, treated by the theory of probability. *Proceedings of the Cambridge Philosophical Society*, *31*, 203–222. https://doi.org/10.1017/S030500410001330X

Jennings, M. R. (1988). *Phrynosoma coronatum*. *Catalogue of American Amphibians and Reptiles*, *428*, 1–5.

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., & Ahmed, I. (2011). Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 1.

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., & Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics*, *33*, 1630–1638.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795. https://doi.org/10.1080/01621459.1995.10476572

Klauber, L. M. (1936). The horned toads of the *coronatum* group. *Copeia*, *1936*, 103–110. https://doi.org/10.2307/1436652

Leaché, A. D., Banbury, B. L., Felsenstein, J., Nieto-Montes de Oca, A., & Stamatakis, A. (2015). Short tree, long tree, right tree, wrong tree: New acquisition bias corrections for inferring SNP phylogenies. *Systematic Biology*, *64*, 1032–1047. https://doi.org/10.1093/sysbio/syv053

Leaché, A. D., Chavez, A. S., Jones, L. N., Grummer, J. A., Gottscho, A. D., & Linkem, C. W. (2015). Phylogenomics of Phrynosomatid lizards: Conflicting signals from sequence capture versus restriction site associated DNA sequencing. *Genome Biology and Evolution*, *7*, 706–719. https://doi.org/10.1093/gbe/evv026

Leaché, A. D., Fujita, M. K., Minin, V. N., & Bouckaert, R. R. (2014). Species delimitation using genome-wide SNP data. *Systematic Biology*, *63*, 534–542. https://doi.org/10.1093/sysbio/syu018

Leaché, A. D., Harris, R. B., Rannala, B., & Yang, Z. (2014). The influence of gene flow on species tree estimation: A simulation study. *Systematic Biology*, *63*, 17–30. https://doi.org/10.1093/sysbio/syt049

Leaché, A. D., Koo, M. S., Spencer, C. L., Papenfuss, T. J., Fisher, R. N., & McGuire, J. A. (2009). Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 12418–12423. https://doi.org/10.1073/pnas.0906380106

Leaché, A. D., & Linkem, C. W. (2015). Phylogenomics of horned lizards (Genus: *Phrynosoma*) using targeted sequence capture data. *Copeia*, *103*, 586–594. https://doi.org/10.1643/CH-15-248

Leaché, A. D., Zhu, T., Rannala, B., & Yang, Z. (2018). The spectre of too many species. *Systematic Biology*, *bioRxiv*, 263996. https://doi.org/10.1101/263996

Linsdale, J. M. (1932). Amphibians and reptiles from Lower California. *University of California Publications in Zoology*, *38*, 345–386.

Martín del Campo, R. (1934). Adición a los camaleones mexicanos. *Anales del Instituto de Biologia, Universidad Nacional de México*, *5*, 327–331.

Montanucci, R. R. (2004). Geographic variation in *Phrynosoma coronatum* (Lacertilia, Phrynosomatidae): Further evidence for a peninsular archipelago. *Herpetologica*, *60*, 117–139. https://doi.org/10.1655/02-100

Nieto-Montes de Oca, A., Arenas-Moreno, D., Beltrán-Sánchez, E., & Leaché, A. D. (2014). A new species of horned lizard (Genus *Phrynosoma*) from Guerrero, México, with an updated multilocus phylogeny. *Herpetologica*, *70*, 241–257. https://doi.org/10.1655/HERPETOLOGICA-D-13-00077

Nieto-Montes de Oca, A., Barley, A. J., Meza-Lázaro, R. N., García-Vázquez, U. O., Zamora-Abrego, J. G., Thomson, R. C., & Leaché, A. D. (2017). Phylogenomics and species delimitation in the knob-scaled lizards of the genus *Xenosaurus* (Squamata: Xenosauridae) using ddRADseq data reveal a substantial underestimation of diversity. *Molecular Phylogenetics and Evolution*, *106*, 241–253. https://doi.org/10.1016/j.ympev.2016.09.001

Noguerales, V., Cordero, P. J., & Ortego, J. (2018). Integrating genomic and phenotypic data to evaluate alternative phylogenetic and species delimitation hypotheses in a recent evolutionary radiation of grasshoppers. *Molecular Ecology*, *2018*, 1–16.

Oaks, J. R., Cobb, K. A., Minin, V. N., & Leaché, A. D. (2018). Marginal likelihoods in phylogenetics: a review of methods and applications. *arXiv*, 1805.04072.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non- model species. *PLoS ONE*, *7*, e37135. https://doi.org/10.1371/journal.pone.0037135

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., … Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*, 595–609. https://doi.org/10.1080/10635150600852011

Presch, W. (1969). Evolutionary osteology and relationships of the horned lizard genus *Phrynosoma* (family Iguanidae). *Copeia*, *1969*, 250–275. https://doi.org/10.2307/1442073

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, *21*, 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org/

Reeve, W. L. (1952). Taxonomy and distribution of the horned lizards genus *Phrynosoma*. *University of Kansas Science Bulletin*, *34*, 817–960.

Schmidt, K. P. (1922). The amphibians and reptiles of Lower California and the neighboring islands. *Bulletin of the American Museum of Natural History*, *46*, 607–707.

Singhal, S., Hoskin, C. J., Couper, P., Potter, S., & Mortiz, C. (2018). A framework for resolving cryptic species: A case study from the lizards of the Australian Wet Tropics. *Systematic Biology*, in press.

Sites, J. W. Jr, & Marshall, J. C. (2004). Operational criteria for delimiting species. *Annual Review of Ecology, Evolution, and Systematics*, *35*, 199–227. https://doi.org/10.1146/annurev.ecolsys.35.112202.130128

Smith, H. M., & Taylor, E. H. (1950). An annotated checklist and key to the reptiles of Mexico exclusive of the snakes. *Smithsonian Institution, United States National Museum Bulletin*, *199*, 1–253.

Solís-Lemus, C., Knowles, L. L., & Ané, C. (2015). Bayesian species delimitation combining multiple genes and traits in a unified framework. *Evolution*, *69*, 492–507. https://doi.org/10.1111/evo.12582

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stejneger, L. (1893). Annotated list of the reptiles and batrachians collected by the Death Valley expedition in 1891, with descriptions of new species. *North American Fauna*, *7*, 159–228. https://doi.org/10.3996/nafa.7.0002

Stephens, M., & Donnelly, P. (2003). A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *The American Journal of Human Genetics*, *73*, 1162–1169. https://doi.org/10.1086/379378

Stephens, M., Smith, N. J., & Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *The American Journal of Human Genetics*, *68*, 978–989. https://doi.org/10.1086/319501

Sukumaran, J., & Knowles, L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 1607–1612. https://doi.org/10.1073/pnas.1607921114

Tevis, L. (1944). Herpetological notes from Lower California. *Copeia*, *1944*, 6–18. https://doi.org/10.2307/1438242

Thomson, R. C. (2016). *California Amphibian and Reptile Species of Special Concern*. Berkeley, CA: University of California Press.

Van Denburgh, J. (1894). Descriptions of three new lizards from California and Lower California, with a note on *Phrynosoma blainvillii*. *Proceedings of the California Academy of Sciences*, 2nd Series, *4*, 296–301.

Van Denburgh, J. (1922). The reptiles of western North America, Volume 1. Lizards. *Occasional Papers of the California Academy of Sciences*, *10*, 1–611.

Vane-Wright, R. I. (2003). Indifferent philosophy versus almighty authority: On consistency, consensus and unitary taxonomy. *Systematics and Biodiversity*, *1*, 3–11. https://doi.org/10.1017/S1477200003001063

Wagner, C. E., Keller, I., Wittwer, S., Selz, O. M., Mwaiko, S., Greuter, L., … Seehausen, O. (2013). Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, *22*, 787–798. https://doi.org/10.1111/mec.12023

Wake, D. B. (2006). Problems with species: Patterns and processes of species formation in salamanders. *Annals of the Missouri Botanical Garden*, *93*, 8–23. https://doi.org/10.3417/0026-6493(2006)93[8:PWSPAP]2.0.CO;2

Yang, Z., & Rannala, B. (2014). Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution*, *31*, 3125–3135. https://doi.org/10.1093/molbev/msu279

Yang, Z., & Zhu, T. (2018). Bayesian selection of misspecified models is overconfident and may cause spurious posterior probabilities for phylogenetic trees. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 1854–1859. https://doi.org/10.1073/pnas.1712673115

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Leaché AD, McElroy MT, Trinh A. A genomic evaluation of taxonomic trends through time in coast horned lizards (genus *Phrynosoma*). *Mol Ecol*. 2018;27:2884–2895. https://doi.org/10.1111/mec.14715