Species Trees and Species Delimitation with Multilocus Data and Coalescent-based Methods: Resolving the Speciation History of the *Liolaemus darwinii* Group (Squamata, Tropiduridae)

Arley Camargo Bentaberry

*Brigham Young University - Provo*

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Biology Commons

BYU ScholarsArchive Citation


https://scholarsarchive.byu.edu/etd/2649

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen amatangelo@byu.edu.
Species Trees and Species Delimitation with Multilocus Data and Coalescent-based Methods: Resolving the Speciation History of the *Liolaemus darwinii* group (Squamata: Tropiduridae)

Arley Camargo Bentaberry

A dissertation submitted to the faculty of Brigham Young University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Jack W. Sites, Jr., Chair
Mark C. Belk
Keith Crandall
Jerald B. Johnson
Leigh A. Johnson

Department of Biology
Brigham Young University

April 2011

Copyright © 2011 Arley Camargo Bentaberry
All Rights Reserved
ABSTRACT

Species Trees and Species Delimitation with Multilocus Data and Coalescent-based Methods: Resolving the Speciation History of the *Liolaemus darwinii* group (Squamata: Tropiduridae)

Arley Camargo Bentaberry
Department of Biology, BYU
Doctor of Philosophy

The inference of species boundaries and phylogenetic relationships are fundamental for evolutionary, ecological, and conservation studies. The resolution of species boundaries and the inference of phylogenetic relationships among species are required to define the units of analysis and to find the most closely related units for evaluating alternative models of speciation. I highlight lizards as model organisms for ecological and evolutionary studies, emphasizing their contributions to advances in understanding linkages between phylogeography and speciation. In this dissertation, I focus on the phylogenetic relationships of the lizards in the *Liolaemus darwinii* group, and the species boundaries of a nested clade within the group, the *L. darwinii* complex, because of several advantages that make these taxa ideal for phylogeographic studies of speciation. I infer a phylogeny for the *L. darwinii* group based on DNA sequences of 20 loci (19 nuclear and 1 mitochondrial) using species trees methods that take into account the incongruence among gene trees. I found the minimum number of loci, number of sequences per species, and number of base pairs per locus that should be included in an analysis for an accurate and precise estimate of the species tree. The species tree based on all available data support a clade of closely related species (*L. darwinii, L. grosseorum*, and *L. laurenti*) known as the *L. darwinii* complex. A new method for species delimitation using Approximate Bayesian Computation is introduced and is shown to accurately delimit species given that limited or no gene flow has occurred after divergence and despite biased estimates of demographic parameters. ABC analyses supported the distinctness of two lineages within *L. darwinii* under a model of speciation with gene flow. Based on the species tree and the species limits obtained in this dissertation, phylogenetic comparative methods can be carried out to address the morphological and ecological evolution in the *L. darwinii* group and several sister species can be used for testing the alternative speciation models via correlation analyses of genetic, morphological, and ecological datasets. Future studies should assess the role speciation due to adaptive processes and its association the species' ecological niches and life histories.

Keywords: species trees, species delimitation, coalescent model, speciation, *Liolaemus*
ACKNOWLEDGMENTS

I would like to thank my graduate committee for their support and advice, and I wish to express my gratitude to the faculty and staff of the Department of Biology for their feedback and encouragement throughout the PhD program. I greatly appreciate financial support from the Department of Biology, the College of Life Sciences, the Bean Life Science Museum, and BYU Graduate Studies for a graduate research fellowship, a graduate mentoring fellowship, and several travel grants, which made possible the completion of this dissertation. I acknowledge also external financial support from a NSF-PIRE award, the Society of Systematic Biologists, and the Society for the Study of Amphibians and Reptiles.

I should thank undergraduate students for their assistance in the lab, especially Greg Walker for his dedication and commitment. Many graduate students and postdocs in the Sites Lab here in the US, the Morando and Avila Lab in Argentina, and the Victoriano Lab in Chile were great companions during field trips and during time spent in the lab, meetings, and socials.

Old friends back at home and new friends here in the US made the time spent during this PhD much happier; I deeply enjoyed their friendship and affection. For being always there, in very special moments, I thank Edgar, Fernanda, Rafael, Eduardo, Raúl, and Andrés.

My advisor, Jack, and his wife Joanne, were invaluable through their guidance and support during difficult times. Thank you both for caring and for helping me, and my family, to grow professionally and personally.

Thanks to my family, especially my parents and brother, for their unconditional encouragement and affection.
This dissertation is dedicated to my wife, Melody, and my daughter, Violeta. Their infinite patience, tolerance, and consolation were essential during this journey. I am so happy that both of you are part of my life; I will always love you.
TABLE OF CONTENTS

1. General Introduction . . . . . . . . . . . . . 1
   1.1. Speciation models . . . . . . . . . . . . . 1
   1.2. Species trees . . . . . . . . . . . . . . . 3
   1.3. Species delimitation . . . . . . . . . . . . 6
   1.4. Conclusions and future directions . . . . . . 8
   1.5. References . . . . . . . . . . . . . . . . . 11

2. Lizards as Model Organisms for Linking Phylogeographic and Speciation Studies 15
   2.1. Introduction . . . . . . . . . . . . . . . . . 17
   2.2. Lizards as models for evolutionary studies . . . . . . . . . 18
   2.3. Emerging themes in phylogeographic research . . . . . . . . . 20
   2.4. Lizard phylogeography: patterns and trends . . . . . . . . . 21
   2.5. Conceptual contributions to phylogeography from lizards . . . . 24
      2.5.1. Ecotones and hybrid zones . . . . . . . . 24
      2.5.2. Species delimitation . . . . . . . . . . 25
      2.5.3. Novel single species studies . . . . . . . 27
      2.5.4. Novel multi-species studies . . . . . . . . 29
   2.6. Liking phylogeography to population divergence and speciation . . . . 31
      2.6.1. Speciation modes and patterns . . . . . . . 31
      2.6.2. What modes can phylogeographic patterns distinguish? . . . 32
      2.6.3. What kinds of data are needed? . . . . . . . 35
2.6.4. What kinds of analyses are appropriate? 37

2.6.5. Integrating genetic, phenotypic and environmental data into tests of alternative speciation modes 39

2.7. Synthesis and future directions 43

2.8. References 47

2.9. Tables. 78

2.10. Figures. 81

2.11. Appendices 88

3. Accuracy and Precision of Species Trees: An Empirical Evaluation of Performance in Lizards of the Liolaemus darwinii Group (Squamata: Tropiduridae) under Varying Sub-sampling Designs 130

3.1. Introduction 132

3.2. Methods 136

3.3. Results 142

3.4. Discussion 146

3.4.1. Performance 147

3.4.2. Systematics 154

3.5. Conclusions 156

3.6. References 160

3.7. Tables. 169

3.8. Figures 171

3.9. Appendices 181
4. Species Delimitation using ABC: Accounting for Speciation with Gene Flow in Lizards of the *Liolaemus darwinii* Complex (Squamata: Tropiduridae) 201

4.1. Introduction 203

4.2. Methods 209

4.2.1. Simulation testing 209

4.2.2. Empirical analyses 211

4.3. Results 215

4.3.1. Simulation testing 215

4.3.2. Empirical analyses 216

4.4. Discussion 218

4.4.1. Species delimitation using ABC and other methods 218

4.4.2. Species limits in the *L. darwinii* complex 223

4.5. Conclusions 225

4.6. References 227

4.7. Tables 236

4.8. Figures 241

4.9. Appendices 245
GENERAL INTRODUCTION

SPECIATION MODELS

The delimitation of species, the units of biodiversity, and the resolution of their phylogenetic relationships is fundamental for evolutionary, ecological, and conservation studies (Sites and Marshall 2003; 2004). For example, the study of the population-level processes during the early stages of divergence is necessary to assess the ultimate causal factors of speciation, and phylogenetic (species) trees are required to address higher-level patterns of diversification (Barraclough 2010). After resolution of species limits and species trees, patterns of divergence in morphological, ecological, physiological, and other organismal traits, can be used to evaluate alternative models of speciation based on the role of stochastic and deterministic evolutionary forces. Allopatric speciation models represent the scenario where disjunct geographic distributions produces isolation and therefore leads to genetic divergence due to random genetic drift within populations. With sufficient time, this process can drive differentiation in compatibility or recognition systems generating reproductive isolation during an eventual geographic contact between former isolated lineages (Mayr 1963). An alternative consists of ecological speciation where disruptive selection across the geographic range and subsequent adaptive change may similarly promote differentiation between populations and reproductive isolation, even under parapatric or sympatric conditions (Schluter 2009). Still other models, like the morphic speciation, combine processes of natural and sexual selection to explain patterns of divergence in morphological, physiological, behavioral, and other life-history traits associated with social morphs (Corl et al. 2010). Regardless of the speciation model, a clear understanding
of species boundaries and species relationships is a prerequisite for addressing the processes and mechanisms underlying patterns of trait divergence during speciation in any taxonomic group.

In Chapter 2, I highlight lizards as model organisms for ecological and evolutionary studies, review the published population genetics/phylogeography literature to summarize general patterns and trends, and describe some studies that have contributed to conceptual advances. My review of 452 studies of lizard phylogeography and population genetics reflects a general trend of exponential growth associated with the theoretical and empirical progress of the discipline. I highlight several studies that have contributed to advances in understanding linkages between phylogeography and speciation, and suggest ways to expand phylogeographic studies to test alternative pattern-based models of speciation.

I develop a conceptual, pattern-based framework for predicting trait correlations for alternative speciation models. The application of this historical perspective for the study of speciation requires both: (1) the resolution of species boundaries, in order to define the units of analysis, and (2) the estimation of phylogenetic relationships among species, to find the most closely related units for assessing trait correlation and distinguishing between alternative speciation models. I propose an expanded approach to compare patterns of variation in phylogeographic data sets with morphological and environmental data to discriminate among alternative speciation models based on the role of deterministic forces in driving divergence between populations, including: (i) passive divergence by genetic drift; (ii) adaptive divergence by natural selection (niche conservatism or ecological speciation); and (iii) socially-mediated morphic speciation. This Chapter has been published as an invited review in the journal *Molecular Ecology* (Camargo et al. 2010).
In order to apply this pattern-based study of speciation, I estimated a phylogeny and assessed species boundaries in a clade of South American lizards of the genus *Liolaemus*. The focal taxa of this dissertation are the *Liolaemus darwinii* group, and a nested clade within the group, the *L. darwinii* complex. The *L. darwinii* group contains 18 species inhabiting the arid lands of the Monte Desert region of central and northwestern Argentina, while the *L. darwinii* complex includes *L. darwinii*, *L. grosseorum*, and *L. laurenti* distributed in sandy shrub lands of the central and southern Monte Desert. This *L. darwinii* group offers a number of advantages for the goals of this dissertation: (1) most species were sampled, (2) the *L. darwinii* complex was densely sampled, and (3) the latter shows potential instances of speciation in isolation and speciation with gene flow. In the lab, these biological advantages were coupled with methodological and analytical advances including (1) development of multiple nuclear loci for the *L. darwinii* group, (2) the application of new coalescent-based approaches for analyzing multilocus datasets, and (3) the use of parallel processors in a supercomputer (cluster marylou5 in Fulton Supercomputing Lab, BYU) for increased efficiency of data analysis. Chapter 3 of this dissertation deals with species trees of the *L. darwinii* group and it has been submitted to the journal *Systematic Biology*. The Chapter 4 consists of species delimitation analyses of the *L. darwinii* complex and is planned for submission to the journal *Evolution*.

**Species trees**

Until recently, standard phylogenetic analysis of multiple loci consisted of concatenating sequences into a single 'super-gene' implicitly assuming that all loci shared the same gene tree topology and matched the underlying species tree. However, gene trees from different loci can
be heterogeneous and be discordant from the species tree due to a number of processes including estimation error, incomplete lineage sorting, horizontal gene transfer, introgression, and gene duplication/loss (Pamilo and Nei 1988; Avise 1989; Maddison 1997). As an alternative to concatenation, Maddison (1997) introduced the idea of estimating species trees via summary-statistics (e.g., deep coalescences) that minimize the discordance between gene trees and the species trees.

Recently, molecular phylogenetics entered a new era in which species trees are estimated from a collection of gene trees by accommodating heterogeneity due to incomplete lineage sorting (Edwards 2009; Knowles and Kubatko 2010) based on the multispecies coalescent model (Rannala and Yang 2003; Degnan and Rosenberg 2009). These novel, model-based frameworks have led to the development of species trees methods based on maximum likelihood and Bayesian inference (BEST, Liu 2008; STEM, Kubatko et al. 2009; Liu et al. 2009; *BEAST, Heled and Drummond). The performance of these methods is beginning to be investigated with simulations to assess the impact of sampling strategies including: number of loci, number of individuals, and sequence length (McCormack et al. 2009; Huang et al. 2010). These simulation-based evaluations have used only accuracy measures, but Liu et al. (2009) suggested using also precision estimators based on the variance of species trees estimates (e.g., posterior distribution of trees from Bayesian analyses). In addition, the variation in locus 'informativeness' or phylogenetic signal (Knowles 2009) and taxon sampling typical of empirical datasets, have not been explored previously in the context of species trees. In addition, differences in species trees have considered topology only (e.g., Robinson-Foulds distance), but there are alternatives to quantify tree similarity that accomodate both topology and branch lengths (i.e., K tree score, Soria-Carrasco et al. 2007).
In Chapter 3, I estimate a phylogeny for the *L. darwinii* group using a species tree method (*BEAST*) based on DNA sequences of 20 loci. I used an empirical approach consisting of subsampling of species, loci, sequences, and base pairs to evaluate patterns of accuracy and precision of species trees with different amounts of information. Based on previous studies, I expected that with fewer data for analysis, the accuracy of the species tree estimate relative to best estimate based on all available data, would be lower. In addition, the precision of the estimate, which represents the variation in topology and branch lengths of the species tree estimate based on the posterior distribution obtained with a Bayesian method (*BEAST*) will also be lower (= higher variance). I found minimum sample sizes for number of sequences per species and number of base pairs per locus that should be included in an analysis in order to preserve accuracy and precision for a given number of loci. Fewer loci were necessary to estimate species trees with fewer species. Patterns of accuracy and precision in relation to the number of loci suggest that accuracy has been maximized but that additional loci will be needed to increase precision. The best estimate of the species tree recovered two major clades as in earlier studies, but with different species compositions. One clade includes all species closely related and formerly considered to be *L. ornatus*, with the addition of other species not previously included in this clade. The other clade includes species formerly considered to be *L. darwinii*, and confirmed the existence of a clade of closely related species (*L. darwinii*, *L. grosseorum*, and *L. laurenti*) that became the focus of the last chapter, due to high sampling density of these species, and the potential occurrence of additional species within the nominal *L. darwinii*. 
Species delimitation is a major research focus in evolutionary biology because the accurate assessment of species boundaries is a prerequisite for the study of speciation, the ultimate process responsible for biodiversity. The practice of species delimitation with molecular data is expanding rapidly due in part to the extension of coalescent models to the interspecific level with the multispecies or 'censored' coalescent (Rannala and Yang 2003; Degnan and Rosenberg 2009). For example, SpeDeSTEM 0.1.1 (Ence and Carstens 2010) finds the maximum likelihood estimate of the species tree for different models of species boundaries, which are compared with Akaike information criteria (Carstens and Dewey 2010). An alternative Bayesian approach consists of sampling from the posterior distribution of models of species limits, while assuming a fixed species tree, using reversible-jump Markov chain Monte Carlo as implemented in the program BPP 2.0 (Yang and Rannala 2010).

Coalescent methods used in current species delimitation approaches accommodate gene tree discordance as a result of incomplete lineage sorting (Knowles and Carstens 2007) but do not explicitly accommodate gene flow after divergence (Ence and Carstens 2010; Yang and Rannala 2010). One way of incorporating gene flow into species delimitation is via the likelihood-free method known as Approximate Bayesian Computation (ABC). ABC calculates summary statistics from observed and simulated genetic data based on coalescent models, and estimates parameters via an algorithm that retains those simulations that are more similar to the observed data (Lopes and Beaumont 2010). In addition to parameters, different demographic models can be compared statistically to select models based on posterior probabilities and/or Bayes factors.
In Chapter 4, I use SDMs in lizards of the *Liolaemus darwinii* complex that includes *L. darwinii*, *L. grosseorum*, and *L. laurenti*, which form a clade within the more inclusive *L. darwinii* group (Chapter 3). The populations of *L. darwinii* have been partitioned into northern (*L. darwinii*-N) and southern (*L. darwinii*-S) groups based on geographic distributions, morphological distinctness, and genetic differentiation (Etheridge 2001; Morando et al. 2004; Abdala 2007). While the taxon pair *L. laurenti* vs. *L. grosseorum* is likely associated with a speciation-in-isolation model, the lineages *L. darwinii*-N vs. -S are possibly the result of speciation with gene flow. I evaluated the performance of an ABC approach for delimiting species using simulated data, applied the method to empirical data of the *L. darwinii* complex, and compared the results with those obtained with other likelihood-based methods.

I found that the ABC method for species delimitation can infer the speciation model accurately given that no gene flow has occurred after divergence or when migration rates are low. In addition, the accuracy in model choice is high despite biased estimates of demographic parameters. There are several ways to improve the performance of the ABC method including the use of: different summary statistics and prior distributions, more complex models, and simultaneous delimitation of multiple lineages. Both likelihood-based methods (SpeDeSTEM and BPP) and ABC consistently supported the distinctness of southern and northern lineages within *L. darwinii*. I conclude that further simulation studies are necessary to evaluate the performance of ABC, likelihood-based, and other SDMs using genetic data derived from speciation models that differ in a number of demographic parameters, especially migration. The ABC framework represents an appropriate solution to the problem of species delimitation, especially in the face of speciation with gene flow, and contributes toward a unified approach that can simultaneously estimate species limits, species trees, and demographic parameters.
CONCLUSIONS AND FUTURE DIRECTIONS

Modern coalescent-based approaches offer a unique opportunity to assess species boundaries and phylogenetic (species) trees of model taxonomic groups for the study of speciation. In addition, new marker development coupled with increased computational power of parallel supercomputers ('clusters') also allows sampling and analyzing multiple loci from the nuclear genome, which are required for parameter estimation of coalescent models used in species delimitation and species trees inference. These methodological advances are maximized in those focal taxa that display diversification patterns representing multiple models of speciation including allopatric vs. parapatric/sympatric divergence with gene flow. Lizards in the *Liolaemus darwinii* complex and in the more inclusive *L. darwinii* group are amenable of these kinds of studies. Alternative speciation patterns, sexual dimorphism, and abundance and ease of observation/collection in the field makes these ideal clades for phylogenetic and phylogeographic studies of speciation. I combined this potential with the development of multiple anonymous nuclear loci (ANL) to produce sequence data for robust estimation of species trees, and for testing species delimitation hypotheses with coalescent-based methods that require intensive computation, which was provided by the MaryLou5 cluster in the Fulton Lab at BYU.

Not surprisingly, I found that multiple loci were necessary to estimate a species tree, and that additional loci were needed to increase nodal support and when including more species (16 of 18 species were sampled for this dissertation). Several ANL that were discarded from the analysis due to time constraints and extreme polymorphism that could not be resolved with
analytical approaches alone (i.e., cloning techniques are required to resolve these heterozygote individuals), but these loci could be added to the multilocus dataset for a re-estimation of the species tree in the future. Despite a relatively deep diversification history of the group (e.g., 13 million years), short internal branches in the species tree suggest rapid diversification during some periods of time that could have produced the high heterogeneity in gene trees observed across loci. The best estimate of the species tree shows a split of the group into two clades associated with different ecological niches, life histories, and reproductive strategies: one clade includes species distributed at higher altitudes including the Pre-Puna and Puna region in Argentina; these species are herbivorous and viviparous. The other clade is restricted to lower altitudes of the Monte region and includes insectivorous and oviparous species. Within the latter clade, the species tree confirmed the close relationships among the species of the *L. darwinii* complex that became the focus of my chapter on species delimitation.

I demonstrate that Approximate Bayesian Computation (ABC) methods can be used for species delimitation in a model selection framework while also producing estimates of demographic parameters relevant for the speciation process (i.e., population sizes, divergence times, and migration rates). Using simulations and empirical data, I also show that ABC correctly inferred species limits and successfully delimited species in the *L. darwinii* complex, consistent with results from likelihood-based methods. More importantly, the flexibility and efficiency of the ABC approach can incorporate gene flow into the speciation model, which represents an improvement over other methods that accommodate only account incomplete lineage sorting. Based on multilocus coalescent-based SDMs, I demonstrate that there are two sister evolutionary lineages within *L. darwinii* (North and South), which can be described as distinct species under the general lineage concept of species. Further, I also show that two very different
speciation patterns have occurred within the *L. darwinii* complex: an example of speciation in isolation between *L. grosseorum* and *L. laurenti*, and a case of speciation with gene flow between the *L. darwinii* lineages, which represent two independent comparisons for evaluation of alternative speciation models.

I anticipate that phylogenetic comparative methods can be used to study the morphological and ecological evolution of the group based on the phylogeny obtained in this dissertation. For example, the correlation between morphology and the evolution of climatic/habitat niche can shed light on the role of adaptive processes in the diversification of the group. The reconstruction of the evolution of discrete morphs associated with alternative life history strategies and the evolution of sexual dimorphism can also be used to evaluate predictions from socially-mediated morphic speciation, where sexual selection is involved in the origin and maintenance of reproductive isolation among divergent morphs. More generally, phylogenetic comparative approaches can be used in a model choice framework to assess the fit of alternative evolutionary models to organismal traits that have evolved in a stochastic vs. deterministic fashion.

Finally, based on the conceptual framework developed in Chapter 2, several sister species can be selected for testing the alternative speciation models discussed in this dissertation via correlation analyses among genetic, morphological, and ecological datasets. As discussed at the beginning of this chapter, alternative speciation models predict different correlation patterns of trait divergence between lineages. Careful selection of the species pairs for these analyses, based on the species limits and species relationships described in these dissertation manuscripts, could be carried out in the future to assess: (1) the relative importance of stochastic factors vs. adaptive process in the diversification of the group; and (2) the association of alternative speciation
models with species' attributes such as ecological niches and life-histories. For example, this dissertation demonstrated that the two lineages within *L. darwinii* have diverged with gene flow and therefore, they could be used to test the hypothesis of speciation due to adaptive processes since even limited gene flow can prevent population divergence unless local adaptation is strong enough to maintain reproductive isolation between the incipient species.

REFERENCES


LIZARDS AS MODEL ORGANISMS FOR LINKING PHYLOGEOGRAPHIC AND SPECIATION STUDIES

Arley Camargo\textsuperscript{1}, Barry Sinervo\textsuperscript{2}, and Jack W. Sites, Jr.\textsuperscript{1}

\textsuperscript{1}Department of Biology and Bean Life Science Museum, Brigham Young University, Provo, UT 84602, USA

\textsuperscript{2}Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95064, USA

Key words: lizards, phylogeography, adaptation, divergence, speciation, molecular markers

Corresponding author: Arley Camargo, Department of Biology, Brigham Young University, Provo, Utah, USA 84602; phone: 801/422-2203; fax: 801/422-0090; email: arley.camargo@gmail.com
Abstract.—Lizards have been model organisms for ecological and evolutionary studies from individual to community levels at multiple spatial and temporal scales. Here we highlight lizards as models for phylogeographic studies, review the published population genetics/phylogeography literature to summarize general patterns and trends, and describe some studies that have contributed to conceptual advances. Our review includes 426 references and 452 case studies: this literature reflects a general trend of exponential growth associated with the theoretical and empirical expansions of the discipline. We describe recent lizard studies that have contributed to advances in understanding of several aspects of phylogeography, emphasize some linkages between phylogeography and speciation, and suggest ways to expand phylogeographic studies to test alternative pattern-based modes of speciation. Allopatric speciation patterns can be tested by phylogeographic approaches if these are designed to discriminate among four alternatives based on the role of selection in driving divergence between populations, including:

(a) passive divergence by genetic drift, (b) adaptive divergence by natural selection (niche conservatism or ecological speciation), and (c) socially-mediated speciation. Here we propose an expanded approach to compare patterns of variation in phylogeographic data sets that, when coupled with morphological and environmental data, can be used to to discriminate among these alternative speciation patterns.
“Lizards” are a paraphyletic group of non-avian reptiles that, together with the odd “worm lizards” (Amphisbaenia) and the snakes (Serpentes), comprise the clade Squamata (Lee et al. 2004; Pough et al. 2004). The well-supported clades Amphisbaenia and Serpentes are unambiguously nested within the Squamata but for simplicity we refer to lizards as all squamates that do not belong to these other clades. This group includes at least 5,354 species (The Reptile Database: http://www.reptile-database.org/, accessed on 15 February 2010) in about 25–26 crown clades usually recognized as families (Pough et al. 2004). Figure 1 depicts a phylogenetic hypothesis from Townsend et al. (2004) that summarizes relationships and distributions in these major groups, and while this arrangement has been challenged (Lee et al. 2004; Vidal & Hedges 2005; Conrad 2008), we present the hypothesis merely to acquaint readers with some aspects of the evolutionary history of the group. Lizards are widely distributed geographically, occupy a wide range of habitats, and are characterized by a striking range of morphologies, ecologies, and body sizes (Pianka & Vitt 2003; Vitt & Caldwell 2009). As a group, lizards show about 150 independent origins of lateral toe fringes (for sand running; Luke 1986), about 100 independent origins of viviparity (lizards + snakes; Blackburn 2006), multiple transitions to a snakelike body form with limb reduction (Greer 1991; Wiens et al. 2006; Brandley et al. 2008), and multiple origins of obligate parthenogenesis (Kearney et al. 2009).

Numerous symposia (Milstead 1967; Huey et al. 1983; Vitt & Pianka 1994; Fox et al. 2003; Reilly et al. 2007) and texts (Roughgarden 1995; Pianka & Vitt 2003; Losos 2009) have focused on lizards as model organisms, because they share attributes relevant to the study of many biological processes, and they are often abundant and easy to manipulate. In this review,
we: (1) summarize some important aspects of lizard diversity and evolution, (2) describe some advantages lizards offer as models for phylogeographic studies, (3) identify some emerging themes and review available lizard phylogeographic studies to summarize trends and patterns, (4) describe case studies which have yielded important insights into broader aspects of phylogeography, (5) emphasize some explicit linkages between phylogeography and patterns of speciation and alternative speciation models, (6) synthesize the current state of knowledge and suggest ways to capitalize on attributes of lizards to improve resolution of phylogeographic studies capable of discriminating among alternative speciation patterns. Such questions can be framed in several alternative speciation contexts, and we suggest that multi-disciplinary studies can highlight linkages of phylogeographic patterns to divergence processes, and integrate some aspects of both phylogeographic and pattern-based speciation studies to allow deeper and more synthetic levels of inquiry.

LIZARDS AS MODELS FOR EVOLUTIONARY STUDIES

Lizards have become model organisms for evolutionary studies due to the accumulated knowledge of long-term demographics, life history strategies, and adaptive ecomorphology and ecophysiology, which together provide an ideal framework for phylogeographic and speciation studies. Lizards are easy to find, approach, and capture in the field for mark-release-recapture methods. They tolerate experimental manipulation, such as in vivo ablation of egg yolk mass (Sinervo & Huey 1990), alteration of body mass (Olsson et al. 2009), and removal of energy reserves by caudal autotomy (Naya et al. 2007).
Long-term demographic studies of several species (Bull 2000; Sinervo & McAdam 2008; Vercken et al. 2008) have provided deep pedigrees leading to novel insights into microevolutionary processes (Sinervo et al. 2007, 2008), patterns of heritable variation and covariation (see Pemberton 2008, for general issues), and patterns of natural selection (Sinervo & McAdam 2008). As an example, *Uta stansburiana* has one of the deepest vertebrate pedigrees in existence that covers both sexes; the pedigree for long-term demographic studies at Los Baños, California, currently spans 21 generations and 7,464 individuals (1988-2008; 400 years in the human sense of time; Sinervo & McAdam 2008).

Some of the earliest tests of the theory of density-dependent natural selection, based on r- vs K-selected life history strategies (MacArthur & Wilson 1967; Pianka 1970), were carried out on lizards (Tinkle et al. 1970; Pianka & Parker 1975). Thermal biology and biophysical ecology models in lizards that emerged from early physiological studies (Huey 1982; Tracy 1982; Porter et al. 2000) are now being applied to estimate geographic distributions based on thermal requirements and climate (Kearney & Porter 2004, 2009; Buckley 2008; Sinervo et al. 2010), whereas comparative ecophysiological methods have been used to explain multi-species distributional patterns (Navas 2002). Lizards have been ideal for investigating the mechanisms and targets of selection based on locomotor performance (Irschick 2000; Van Damme et al. 2008), showing that morphological variation is functionally and ecologically relevant because it translates into performance differences (Goodman 2007). These and related population processes have been emphasized in frameworks for studying the biological basis for allopatric speciation (Wiens 2004a), and we return to this point at the end of this review.
EMERGING THEMES IN PHYLOGEOGRAPHIC RESEARCH

The term “phylogeography” originally described analyses of gene genealogies within species or among closely-related species in explicit geographic contexts (Avise et al. 1987). Hewitt (2001) expanded phylogeography to also include considerations of hybrid zone dynamics and speciation patterns, especially in the context of Quaternary & Holocene histories of regional biotas. Moreover, because the original goal of phylogeography was, and still is, to bridge population genetics with phylogenetics, the analysis of genotypic and allele frequency data for phylogeographic inference of very recent events or ongoing processes has been incorporated into a more inclusive discipline (Garrick et al. 2010). As recent reviews attest, the field is experiencing rapid growth in many directions (Beherenayar 2008; Riddle et al. 2008; Avise 2009; Brito & Edwards 2009; Edwards 2009; Knowles 2009; Nielsen & Beaumont 2009). The availability of nuclear genetic markers, advances in coalescent theory, and new GIS tools for generating ecological niche and paleoclimate models, are rapidly increasing the scope of phylogeographic studies (Swenson 2008; Buckley 2009; Hickerson et al. 2010; Sinervo et al. 2010). For example, some studies incorporate external climatic and geologic data to generate \textit{a priori} predictions that can then be tested with molecular phylogographic approaches (e.g., Richards et al. 2007; Knowles et al. 2007; Knowles & Carstens 2007a; Moriarty-Lemmon et al. 2007; Carnaval & Moritz 2008; Werneck et al. in review), but statistical methods can also estimate phylogeographic history without \textit{a priori} hypotheses (Templeton 2004, 2010a,b; Lemmon & Moriarty-Lemmon 2008). The application of multi-locus coalescence methods to link phylogeography to species delimitation issues is growing rapidly (Carstens & Knowles 2007; Liu & Pearl 2007; Brumfield et al. 2008; Liu et al. 2008; Degnan & Rosenberg 2009; Yang & Rannala 2010), as are multi-species assessments of the role of gene flow vs. natural
selection across environmental gradients (Rosenblum 2006), biodiversity patterns and processes in regional landscapes (Leaché et al. 2007; Victoriano et al. 2008; Carnaval et al. 2009; Hurt et al. 2009; Moritz et al. 2009), and the incorporation of these data into conservation planning (Davis et al. 2008). Statistical and computational issues remain challenging (Knowles 2008; Nielsen & Beaumont 2009; Templeton 2009a,b, 2010a,b; Beaumont et al. 2010), but phylogeography will continue to expand and incorporate other disciplines (Beheregaray 2008; Avise 2009; Knowles 2009).

**Lizard Phylogeography: Patterns and Trends**

Similar to other organisms (Beheregaray 2008), population genetic and phylogeographic studies of lizards have grown rapidly due to the refinement of data collection and analytical techniques, including the use of molecular markers with finer resolving power (Avise 2000; Garrick et al. 2010), coupled with increasingly powerful analytical methods (Hickerson et al. 2006; Richards et al. 2007; Nielsen & Beaumont 2009; Templeton 2009a). Here we review the primary literature published on population genetics and phylogeography of lizards. While phylogeography originally referred to molecular studies linking the geographic distribution and genealogical relationships among intraspecific evolutionary lineages (Avise et al. 1987), and population genetics did not incorporate a genealogical component, we review both for two reasons. First, early population genetic studies often established a basis for subsequent phylogeographic studies, and second, the biology of population isolation and divergence, which ultimately drives speciation (Wiens 2004a, b), requires the integration of multiple approaches based on different data sets relevant to different time scales (Hewitt 2001; Templeton 2001;
The same is true for the issue of species delimitation, and both tree-based (coalescent) and non-genealogical gene flow methods are relevant to these interrelated issues (Knowles & Carstens 2007b; Petit & Excoffier 2009; Carstens & Dewey in press).

The search for published studies of lizards and the information extracted from these publications to analyze temporal, geographical, taxonomical, and methodological trends are described in Appendix I. We found 426 references representing 452 study cases (some studies included multiple taxa) (Table 1, Appendix I). Seventeen families and 117 genera were included in these studies, with the European lacertid Podarcis being the most commonly studied genus (40 references), followed by the North American phrynosomatid Sceloporus (35), and the Caribbean Anolis (35) (Appendix II). The first references appeared in 1980 (population genetic studies), while the first phylogeographic study (sensu Avise et al. 1987) appeared in 1989 (Sites & Davis 1989). Numbers of papers remained relatively stable with slight increases through 1996, and in 1997 the increase began a trajectory of nearly exponential growth (Fig. 2). These studies have been published in a total of 87 different journals with the most frequent being Molecular Ecology (63), Molecular Phylogenetics and Evolution (56), and Evolution (33).

Most studies were based in North America (24%), followed by Europe, Australia, and Asia (14–18%), the West Indies and Africa (~10% each), then South America with Atlantic and Pacific oceanic islands (3–6%); other regions include less than 1% of the studies (some studies cover more than one region; Appendix II). Coverage is taxonomically biased; the Lacertidae and Phrynosomatidae (the dominant clades in Europe and North America, respectively) are over-represented relative to their species diversity, while the species-rich clades Scincidae and Gekkonidae are under-represented. At the level of generic diversity, these families plus the
Agamidae are well represented, but other families remain poorly studied (e.g., Gymnophthalmidae; Appendix II). Lizards are better studied in the Southern Hemisphere proportional to their diversity; 33% of all studies were based in Africa, Australia, or South America, a much higher proportion relative to their diversity than for other groups of vertebrates (Beheregaray 2008).

Mitochondrial DNA has been the most frequently used (51%) marker, followed by allozymes (25%), mini- and microsatellites (16%), and AFLP/RFLP/RAPD markers (10%), chromosomes (8%), and nuclear sequences (5%). Allozymes were the earliest used, then mtDNA (restriction sites [1989] and [1993]), microsatellites in 1997, and since 1998 both have been preferred for different temporal scales, with nuclear sequences incorporated since 2004 (Fig. 3). The reconstruction of intraspecific (or congeneric) phylogenies and networks (category E; Table 2) has been the most common method employed (63%), usually together with estimates of differentiation and gene flow (A, B, and D; 58%; Appendix II). Tests of population structure (C) have been common (28%), but new phylogeographic/population genetic methods (F, G, and H) are becoming popular (16%), while classification and correlation methods (I and J) have infrequently been applied (16%). From 1999–2009, new methods (Nested Clade Phylogeographic Analysis = NCPA, coalescent, assignment/clustering algorithms; F, G, and H, respectively) were applied in ~21% of all studies. Over this same time frame, mtDNA use has remained about the same (53%), while nDNA and microsatellites usage almost doubled (11% and 28%, respectively), and allozymes and RFLP/AFLP/RAPD were rarely used (5% and 2%, respectively). The average sampling design was similar for studies using the newer methods (12.4 individuals per locality, standard deviation [s.d.] = 13.9, N = 70) relative to that used across all studies in 1999–2009 (12.7, s.d. = 36.3, N = 329).
CONCEPTUAL CONTRIBUTIONS TO PHYLOGEOGRAPHY FROM LIZARDS

Ecotones and hybrid zones

Studies of ecological gradients and parapatric hybrid zones can suggest what evolutionary forces may have contributed to divergence/adaptation to different habitats (Ogden & Thorpe 2002; Rosenblum 2006), while studies of narrow contact zones can reveal patterns such as linkage disequilibrium, heterozygote deficits, and coincident clines suggestive of post-zygotic selection against hybrids (Phillips et al. 2004). Novel investigations of divergence across ecotones include the Schneider et al. (1999) study of the skink *Carlia rubrigularis* in the Australian Wet Tropics. Morphology (body size, limb length, and head shape) and life history (age at maturity) in this species shift abruptly across a sharp ecotone between forest types, and avian predation (as estimated from beak marks on plasticine models) was one likely driver of this divergence despite of gene flow. Rosenblum (2006) studied phenotypic transitions of three lizard species (*Holbrookia maculata, Sceloporus undulatus* [Phrynosomatidae], and *Aspidoscelis inornata* [Teiidae]) characterized by “blanched” color morphs on the gypsum dunes, wild type morphs on brown soils, and intermediate colors in the narrow ecotones (color morphs have a genetic basis [Rosenblum 2005]). Neutral processes could not explain color variation but natural selection was sufficiently strong to produce divergent phenotypic responses despite species-specific differences in population structure, demographic history, and ecology (Rosenblum 2006). Rosenblum et al. (2010) have now shown that different molecular mechanisms in the same gene have produced these blanched phenotypes.
In Mexico, two chromosomal races of *Sceloporus grammicus* (Phrynosomatidae) form a hybrid zone in a pine-oak/chaparral ecotone characterized by steep concordant clines in three diagnostic autosomal markers. Sites *et al.* (1995) used tension zone theory (Barton & Hewitt 1985; Barton & Gale 1993) to calculate a cline width (~830 m), and inferred that this zone is likely maintained by both endogenous (genetic) and exogenous (environmental) selection adapting the races to different habitats. Studies of (1) fitness correlates in hybrid/back-cross males heterozygous for different autosomal rearrangements (Reed *et al.* 1995a,b), (2) female fecundity among parental, F₁, and back-cross genotypes (Reed & Sites 1995), (3) cyto-nuclear structure (Sites *et al.* 1996), and (4) modeling of cline shapes for multiple unlinked markers (Marshall & Sites 2001) confirmed earlier findings, and provided among the first multi-faceted studies of the dynamics of a vertebrate mosaic hybrid zone. Recent detailed studies of a contact zone between two mtDNA haploclades of *Lacerta schreiberi* in the Iberian Peninsula found a steep cline (Godinho *et al.* 2008) with asymmetric gene flow unrelated to female mating preferences (Stuart-Fox *et al.* 2009a), but consistent with body condition (based on parasite load) or male color (associated with aggressiveness) (Stuart-Fox *et al.* 2009b).

*Species delimitation*

Species delimitation has become inter-connected with phylogeography because (1) phylogeography deals with patterns and processes occurring at the intra/interspecific boundary, and (2) coalescent methods are relevant to both topics (Knowles 2009; Knowles & Carstens 2007b; O’Meara 2010; Carstens & Dewey in press). Methods of species delimitation have
already been the topic of several recent reviews (Sites & Marshall 2003, 2004; Padial & de la Riva 2006; Wiens 2007), and lizards have served as models for new approaches.

Wiens & Penkrot (2002) described tree-based species delimitation methods using molecular and morphological data, to test species boundaries in the Mexican *Sceloporus jarrovi* complex. Gene trees were constructed, and an inference key used to assess species boundaries, which identified a total of five species in this group but only two of these were identical across different data and criteria. This was among the earliest studies to test performance of clearly articulated methods for concordance in the species recovered. Morando *et al.* (2003) sequenced multiple mtDNA regions in the Patagonian *Liolaemus elongatus-kriegi* complex (Tropiduridae) and described an efficient hierarchical sampling design to simultaneously evaluate intra- and interspecific variation in poorly known clades, and to identify “candidate species” for further study. Raxworthy *et al.* (2007) used GIS-based ecological niche modeling (ENM; Phillips *et al.* 2006) to delimit species in two complexes of *Phelsuma* geckos endemic to Madagascar based on niche overlap predictions. The ENM of all named taxa combined was expected to overpredict niche space if the taxa occupied different niches, and results revealed that both species complexes included taxa that occupied divergent niche space. In some cases morphological data corroborated ENM inferences for species, despite low levels of molecular divergence (0.47% uncorrected mtDNA p-distance).

Marshall *et al.* (2006) compared the performance of several species delimitation methods in the *Sceloporus grammicus* complex by designating four “hypothesized evolutionary species” (HES) from molecular data, and then evaluating the accuracy of five methods in recovering four HES units. No single method strongly delimited all of these, but two showed some support of all four, revealing that co-dominant markers are likely to be successful at delimiting species by any
number of methods, given their success in this complex characterized by recent race
diversification and multiple hybrid zones.

Leaché et al. (2009) integrated mitochondrial and nuclear gene sequences, niche
envelopes, and morphometric assessments of horn shape, to delimit species in the Phrynosoma
coronatum complex. The mtDNA gene tree recovered five haploclades distributed linearly from
central California south through the Baja Peninsula. The other data sets were largely congruent
with each other and the mtDNA haploclades at the deepest divergence levels, but at recent levels
of divergence the other data sets were discordant, and nuclear gene flow between these could not
be rejected. The authors recognized three species concordant with the deepest mtDNA
haploclades, all of which were ecologically and morphologically diagnosable.

**Novel single species studies**

GIS-based ENM is now routine in many phylogeographic studies, but biophysical niche
modeling methods (Kearney 2006) can decipher functional links between organismal physiology
and predictor variables that may limit species distributions (Kearney & Porter 2004, 2009).
Strasburg et al. (2007) integrated this approach in a phylogeographic reconstruction of
Pleistocene range expansions in two parthenogenetic forms of the Australian gecko Heteronotia
binoei. Both have had relatively recent origins and subsequently expanded at different times; the
3N1 race at ~24,000 yr ago, and the 3N2 race at ~7,000 yr (estimates from NCPA and mismatch
distribution analyses). ENM and biophysical modeling (Kearney & Porter 2004) supported these
conclusions, and showed that the southern range limit of one bisexual race coincides closely with
the thermal limit for successful egg development, an inference that could not have been made from correlational modeling alone.

Rosenblum et al. (2007) studied colonization histories of Sceloporus undulatus (= S. cowlesi, in Leaché & Reeder 2002) in novel habitats in the Tularosa Basin of New Mexico. Geologically recent “islands” of white sand dunes (dated to ~10,000 ybp) and black rocks (Carrizozo lava flow; ~5,000 ybp) provide independent but analogous experiments in selection. Multiple loci used in an Approximate Bayesian Computation (ABC) framework (Hickerson et al. 2006) revealed that: (1) population reductions were associated with initial colonization of both habitats; and (2) these were more severe during colonization of the black lava habitat. Reductions in inbreeding effective population size ($N_e$) may be more dramatic when colonization is accompanied by a change in selection regime, an idea consistent with a demographic cost of adaptation to novel environments (Haldane 1957; Lande & Shannon 1996).

Gifford & Larson (2008) used multiple loci to infer two fragmentation events concordant with Pliocene and Pleistocene marine transgressions in Ameiva chrysolaema (Teiidae) on the island of Hispanola. Multi-locus NCPA (Templeton 2010a,b) and Bayesian coalescent analyses recovered signatures of population expansion and asymmetric migration consistent with the relative magnitude and duration of inundations for each region. In the Iberian Peninsula, Godinho et al. (2008) described a hybrid zone between two lineages of the lacertid Lacerta schreiberi, delineated by a combination of slow- and fast-evolving markers; a sharp transition between mtDNA clades but smooth clines in the nuclear data suggested a chronology of historical events including a late Pliocene fragmentation, recontact during glacial cycles with formation of a hybrid zone and recent population expansions.
Novel multi-species studies

McGuire et al. (2007) investigated patterns of mtDNA paraphyly in 12 species of the family Crotaphytidae (Crotaphytus and Gambelia) in southwestern North America, coupled with GIS-modeling of current and “last glacial maximum” (LGM, ~21,000 ybp) distributions to identify contact areas today or in the recent past. This revealed a unique pattern of mtDNA variation that suggested repeated cycles of introgression of *C. collaris* mtDNA haplotypes into *C. bicinctores*. The authors hypothesized an “introgression conveyer” model with three phases of unidirectional introgression, followed by substantial mtDNA divergence between each of the three events.

Dolman & Moritz (2006) estimated isolation and divergence in the Australian skink genus *Carlia*, using three well-defined mtDNA clades representing the sister species *Carlia rubrigularis* and *C. rhomboidalis* to assess interaction between geographic isolation, genetic drift, introgression, and divergent selection, on speciation and divergence processes in rainforest faunas (Moritz et al. 2000). A coalescent method (IM; Hey & Nielsen 2004) applied to sequence data from seven nuclear genes revealed large $N_e$ in *C. rhomboidalis*, suggesting that drift did not likely contribute to its divergence from *C. rubrigularis*, while the processes that maintained phenotypic stasis within *C. rubrigularis* and drove divergence between the two species were not clear. Recent studies on co-distributed lizard species in this same region, such as Moussalli et al. (2009) evaluation of climatic niche specialization in *Saproscincus* skinks, focused on responses Quaternary cycles of forest contraction/expansion. Current climate preferences of these species extrapolated to past climates were concordant with geographic patterns of mtDNA genetic diversity, suggesting that all have maintained their respective climate preferences at least through
the late Pleistocene, and that niche conservatism (Wiens & Graham 2005) contributed to genetic diversification within this system.

Victoriano et al. (2008) implemented a super-trees approach (Lapointe & Rissler [2005]), to estimate co-divergence in three species of *Liolaemus* (Tropiduridae) with partially overlapping distributions in the Chilean Andes. Concordance of area trees was tested by treating co-occurring taxa as host-parasite associations, and congruent patterns were inferred when tests were significant. Environments from three *a priori* recognized climate zones were quantified by six variables and tested against a null model (no difference) by permutation. Significant spatial co-divergence between *L. tenuis* and *L. pictus* and between *L. tenuis* and *L. lemniscatus*, and significant positive correlations between the supertree distance and the climate matrixes, suggest that, in sympatry, these species have responded in parallel to shared historical events.

Leaché et al. (2007) tested for simultaneous divergence across a shared phylogeographic break (the mid-peninsular seaway in 12 species co-distributed along the Baja Peninsula (mtDNA from four lizards, two snakes, and six rodents). A hierarchical Approximate Bayesian Computation (hABC) analysis suggested two temporally disjunct divergence events; seven taxon pairs diverged ~2.3–15.3 MYA, while five diverged ~0.6–3.4 MYA. In the Australian Wet Tropics, Moritz et al. (2009) included nine lizard species in a comparative phylogeographic analysis of a “suture zone” (a region where an assemblage of species establish secondary contact [Remington 1968; Swenson & Howard 2004]), and showed that individual hybrid zones were significantly clustered in a region between two major Quaternary refugia, and most of these occurred in areas of low environmental suitability relative to the adjacent refugia. MtDNA sequence divergences varied between sister lineages (2–15%), as did the extent of reproductive isolation (random admixture to speciation by reinforcement [Hoskin et al. 2005]). Moritz et al.
(2009) suggested was that suture zones are better defined by shared expansion times to contact, rather than common divergence times.

**LINKING PHYLOGEOGRAPHY TO POPULATION DIVERGENCE AND SPECIATION**

*Speciation modes and patterns*

An integrated phylogeographic perspective can provide important insights into three components of speciation research: (1) the geographic context of speciation, (2) the processes driving divergence, and (3) the origin of reproductive isolation (Nosil 2008, 2009; Nosil *et al.* 2009; Sobel *et al.* 2010). Speciation patterns have been categorized by geographic modes since the modern synthesis (Mayr 1942); the classical allopatric (Mayr 1963) and more recently peripatric, parapatric, and sympatric modes (Coyne & Orr 2004). This classification spans the continuum of geographic modes and clarifies some key questions in speciation research (Butlin *et al.* 2008), but despite theoretical (Gavrilets 2004) and empirical treatments (Coyne & Orr 2004; Futuyma 2005; Price 2008) on the frequency of these modes, the allopatric model remains widely corroborated (Barraclough & Vogler 2000; Coyne & Orr 2004; Phillimore *et al.* 2008; Price 2008).

A classification of speciation modes by evolutionary processes was presented by Losos (2009), and recognized ‘adaptive’ vs ‘non-adaptive’ patterns. In a theoretical context, Gavrilets (2004) has recognized stochastic vs deterministic factors responsible for the origin of reproductive isolation, which are responsible for ‘non-adaptive’ and ‘adaptive’ patterns of speciation. In practice, Futuyma (2005) suggested formulating and testing a null hypothesis of
speciation due to stochastic forces, which in the simplest case is the “passive divergence” or “drift-only” paradigm, because rejecting this hypothesis is probably easier than demonstrating the action of other evolutionary forces.

Stochastic processes alone are considered unlikely to drive speciation because drift is relatively inefficient in producing reproductive isolation (Sobel et al. 2010), but if species are independent evolutionary lineages (de Queiroz 1998) that can be detected using neutral genetic markers and coalescent-based methods (O’Meara 2010), then a phylogeographic approach can distinguish among some modes of speciation. Simulation studies show that, given enough loci, coalescent methods can delimit species at shallow levels of divergence (~0.3 N_e) when they still display considerable incomplete lineage sorting (Knowles & Carstens 2007b); these stochastic forces therefore have a role in generating independent lineages. Gene trees and geographic distributions alone are insufficient to distinguish among geographic modes of origin because assumptions about the distributional ranges of the populations/species may not be met (Losos & Glor 2003). Here we suggest that a phylogeographic focus on population/species divergence in terms of the spatio-temporal isolation of lineages, combined with environmental and phenotypic data, are sufficient to discriminate “drift only” vs “selection-driven” divergence, and then among some three classes of the latter. Establishment of a strongly corroborated pattern would then require follow-up studies to explicitly link lineage divergence to the origin of reproductive isolation (Sobel et al. 2010; Wiens 2004a, b).

What modes can phylogeographic patterns distinguish?
Ecologically-based adaptive processes can produce selectively-driven departures from a neutral divergence pattern in two different ways. First, niche conservatism can limit gene flow and therefore promote divergence between allopatric sister lineages by constraining adaptation at the geographic barrier separating them (Wiens 2004a; Kozak & Wiens 2006). Alternatively, adaptation to different ecological niches can also limit gene flow between allopatric or parapatric lineages and lead to ‘ecological’ speciation when these changes result in reproductive isolation (Rundle & Nosil 2005; Nosil et al. 2009; Schluter 2009). The niche conservatism scenario is predicted to produce more similar ENMs between species’ ranges relative to the unoccupied region separating them, whereas the adaptive divergence model predicts more different ENMs relative to the barrier region (Hua & Wiens 2010). Although the environmental factors used in ENM can distinguish between divergent ecological niches, it is important to evaluate if these factors have diverged beyond expectations due to geographic distance. In this vein, McCormack et al. (2010) developed null expectations for differentiation in ENM to distinguish between adaptive patterns (causing niche divergence or conservatism) and differentiation due geographic separation only. This approach can be strengthened by quantification of morphological divergence (or absence thereof) in isolated populations, because this information will capture some niche dimensions not included in climate modeling. For example, if the two lineages have similar ENM, we would expect them to also display similar phenotypes as a result of adaptation to “identical” ecological niches. Alternatively, if ENM are different, divergent ecological selection is expected to drive some phenotypic divergence between the lineages.

A third alternative is the hypothesis of socially mediated speciation, in which hybrid unfitness is due to alternative local mating systems, rather than divergent ecological forces underlying adaptive traits (Sinervo & Svensson 2002; Hochberg et al. 2003). The well-studied
“rock-paper-scissors” (RPS) mating system in the lizard *Uta stansburiana*, in which three male throat color morphs fluctuate via frequency-dependent selection within local populations, has been suggested as an example of speciation by this mode (Corl *et al.* 2010b; the genetic basis for these phenotypes, details of the model for divergence and possibly reproductive isolation, are described in Appendix III). Corl *et al.* (2010b) suggested that geographic variation in the RPS polymorphism among *Uta* populations offers the opportunity for speciation when ancestral tri-morphic systems collapse to di- or mono-morphic systems in isolated populations, and reproductive isolation then evolves upon secondary contact due to an interaction of natural and sexual selection forces. Loss of one or two male color morphs in novel environments alters the 3-morph RPS equilibrium, and this is followed by “character release” and rapid phenotypic evolution of the remaining color morphs in body size, sexual dimorphism, and probably other life history traits such as clutch size (Corl *et al.* 2010a). Body size and clutch size are likely not the only traits that have diverged with morph loss, because these morphs also differ in other heritable traits including behavior, hormone levels, clutch size, egg mass, and immunocompetence (see Appendix III), and strong correlational selection on the color locus and other trait loci (e.g., multi-trait selection) generates the highest standing levels of linkage disequilibrium observed within a species (Sinervo *et al.* 2006). Divergence in color traits involved in male signaling and female choice can then promote reproductive isolation through assortative mating by color (Bley & Sinervo 2007), multi-trait female preference (Lancaster *et al.* 2009) and/or operation of reinforcement processes on post-zygotic (Dobzhansky-Mueller) differences between populations (see details in Appendix III). The rapid phenotypic divergence of isolated di- or monomorphic populations (some recognized as different species or subspecies) suggests that a ‘morphic’ speciation process has been operating in *Uta stansburiana* (Corl *et al.*
Because color polymorphism may be the first step in speciation in this mode (Levene 1953; Maynard-Smith 1966), and reproductive isolation may arise from social competition rather than ecology *per se* (West-Eherhard 1983, 1986, 2003), we view this as a distinct alternative to the ecological modes described above. Table 2 summarizes patterns of ecological and morphological divergence expected among these speciation modes.

*What kinds of data are needed?*

**Geographic sampling.**—Dense geographic sampling and geo-referencing of specimens are necessary to document distributions of species accurately (Buckley 2009), which can then be used to address a range of ecological and evolutionary questions (Wiens & Graham 2005; Kozak *et al.* 2008). Ideally, sampling design should be based on *a priori* knowledge of population structure and life history, guided by the assumptions of the analytical methods to be used, and included as another parameter in analyses to evaluate the impact of sampling on the inferences (Buckley 2009). Sampling impact can be minimized by correcting for ascertainment bias due to poor sampling of rare polymorphisms (Rosenblum & Novembre 2007), assessing sampling completeness (Dixon 2006), and evaluating limitations of inferences made from finite samples (Templeton 2009a).

**Genetic data.**—While mtDNA will likely remain the preferred phylogeographic ‘first pass’ marker for lizards, as in most other organisms (Beheregaray 2008; Zink & Barraclough 2008; Avise 2009; Barrowclough & Zink 2009), the continued development of new markers and analytical methods will expedite the incorporation of multiple loci into phylogeographic studies.
Multiple loci increase the accuracy of estimation of parameters such as population sizes and divergence times (Edwards & Beerli 2000; Felsenstein 2006; Heled & Drummond 2008; Kuhner 2008), and the resolving power of coalescent-based species delimitation methods even at shallow time horizons (Knowles & Carstens 2007b). While the discovery of nuclear markers with sufficient variability for phylogeographic analysis is still challenging, anonymous loci seem to be promising for these purposes and even for phylogenetic inference (Brito & Edwards 2009). Complementary studies of hybrid zones and gene flow will likely rely on microsatellite markers (Petit & Excoffier 2009).

*Morphological data:*—Many studies have identified morphological traits relevant to functional performance with fitness consequences, including body size, limb proportions, and head size/shape, as these relate to locomotion, microhabitat use, niche convergence, anti-predator behavior, and social interactions (Harmon et al. 2005, 2007, 2008; Losos et al. 2006; Calsbeek 2008; Vervust et al. 2007; Herrel et al. 2008; Losos 2009). Frequency of tail autotomy provides information on predation efficiency/intensity (Medel et al. 1988; Cooper et al. 2004; Pafilis et al. 2009), and/or intrasexual competition (Hofmann & Henle 2006; Corl et al. 2010a). Body color patterns display adaptive variation in association with habitat (Thorpe 2002; Rosenblum 2006; Schneider 2008), and in cases of sexual dimorphism, color patterns or morphs usually act as signaling traits carrying relevant information for social interactions and mate recognition (Lancaster et al. 2009). If sexual dimorphism in color is relevant, then phenotypes can be scored by eye from photographs (Sinervo et al. 2006, 2007) or spectrophotometry if colors are beyond the visible range (e.g., UV; Côte et al. 2008; Vercken et al. 2008), and some pigmentation
patterns can be quantified from museum vouchers (Leaché & Cole 2007), especially melanism (Camargo et al. unpub. data).

**Environmental layers/niche modelling.**—Climate and topographic data with global coverage at several levels of geographic resolution are available from public databases (i.e., www.worldclim.org), and vegetation and soil properties can be derived from remote-sensing data (Zimmermann et al. 2007). GIS software enables preparation of these environmental layers for subsequent analyses to ensure identical resolution and area coverage, and to extract data associated with point localities. These layers represent the input data, together with georeferenced localities from field-collected specimens for estimating ENM under alternative scenarios. While most studies use the maximum entropy approach (Phillips et al. 2006; Phillips & Dudík 2008), mechanistic methods are emerging (Kearney & Porter 2009; Monahan 2009), and recent reviews have highlighted the utility of ENM for addressing questions in speciation research (Wiens & Graham 2005; Rissler & Apodaca 2007; Kozak et al. 2008).

*What kinds of analyses are appropriate?*

**Genetic data.**—DNA sequences can be analyzed by a number of methods for testing species boundaries, thus delimiting the units to be compared with phenotypic and environmental data. Coalescent-based approaches statistically test the fit of gene genealogies to alternative hypotheses of species boundaries, including likelihood methods that accommodate incomplete lineage sorting (Knowles & Carstens 2007b), and extensions (O’Meara 2010) to search simultaneously for both the optimal delimitation of multiple species and the species tree. We
anticipate that the current surge of coalescent methods, including likelihood (STEM, Kubatko et al. 2009) and Bayesian approaches (BEST, Liu & Pearl 2007; *BEAST, Heled & Drummond 2010; GLASS, Mossel & Roch 2010), will continue to improve by accommodating other processes (i.e., gene flow) that may occur between recently diverged species. Genetic data can also be used to calculate distance matrices based on pairwise comparisons between population mean values or individual values (FST for example).

**Morphological data.**—Differentiation based on morphological data can be summarized with the PST statistic, a surrogate for the quantitative genetic differentiation (QST) under some assumptions (Gay et al. 2009). Multivariate analyses can be applied to extract a few major axes that account for most of the variation in phenotypic variables (e.g., morphometric or meristic data, coordinates from geometric morphometrics, etc.), and multidimensional Euclidean distances between individuals can be obtained from the data space defined by these axes. Obviously neither the environmental nor the morphological data may fully capture some variables linked directly to adaptive change, and they may fail to detect adaptive divergence, but a focus on traits for which links between direct fitness and adaptive processes have been documented should minimize this problem. In lizards, differences in body size, shape, coloration, or limb and head proportions (Harmon & Gibson 2006; Harmon et al. 2007) are most often associated with adaptive responses to shifts in habitat (Kearney & Porter 2004, 2009; Calsbeek & Sinervo 2007; Calsbeek & Smith 2007; Losos et al. 2000, 2001; Ogden & Thorpe 2002), climate (Guillette 1993; de Fraitpont et al. 1996), or biotic interactions (e.g., competitors, parasites, prey, or predators; Kearney 2006; Buckley 2008; Stuart-Fox et al. 2009b).
Environmental data.—ENMs for species can be tested for significant between-species differences with new methods (ENM Tools, Warren et al. 2008; McCormack et al. in press). Further, environmental data extracted from georeferenced localities can be used to obtain environmental distances between localities based on Euclidean distances derived from multivariate approaches such as Principal Components Analysis. In addition to the usual climatic and topographic layers, remote-sensing data that are strongly associated with habitat differences and have low spatial autocorrelation can increase the resolving power of ENM analyses (see McCormack et al. 2010).

Integrating Genetic, Phenotypic and Environmental Data into Tests of Alternative Speciation Modes

Neutral genetic markers can be used to assess the significance of ecological/phenotypic divergence between populations, since they represent the predicted levels of differentiation due to drift only (within populations) or drift and gene flow (between populations; Nosil et al. 2008; Gay et al. 2009). Phenotypic divergence exceeding or discordant with neutral expectations suggests the influence of selection in driving divergence, and if also correlated with environmental divergence, a causal mechanism can be hypothesized for observed patterns. ENM can be used to test for significant niche divergence by accommodating the effect of geographic distance separating species/populations, and distinguishing between selection-driven ecological divergence vs. niche differentiation due strictly to geographic separation (McCormack et al. 2010). Because tests of niche differentiation may be insufficient to distinguish between adaptive and non-adaptive speciation modes, we suggest that the inclusion of neutral genetic and
morphological data provide a more inclusive context for ENM differentiation. We propose that by looking at both (1) the divergence between lineages in each dataset, and (2) the correlations between these three datasets, we can distinguish among the three alternative speciation patterns addressed above.

The first step consists of obtaining distance matrices for the three genetic, morphological (or other phenotypic), and environmental data sets. As explained above, $F_{ST}$ distances based on genetic variation and Euclidean distances from multivariate analyses of morphological and environmental data can provide these matrices. The next step uses a partial matrix correspondence (Mantel) test (MCT) to evaluate possible correlations between environmental and phenotypic matrices, while accommodating neutral genetic variation based on the residuals of the pairwise correlations with the genetic distance matrix (Smouse et al. 1986; Thorpe et al. 1996; Thorpe 2002); this is required in order to hypothesize a potential adaptive phenotypic response to the environmental conditions. A partial MCT is then used to test the null hypothesis that drift alone explains the phenotypic differentiation, by evaluating the correlation between phenotypic and environmental distance matrixes, after controlling for the genetic differentiation in neutral markers (Thorpe & Stenson 2003; Rosenblum 2006; Rosenblum et al. 2007; Richards & Knowles 2007). There are two alternative outcomes of this test (Fig. 4). If the partial MCT is not significant, we can infer that drift explains the observed phenotypic differentiation between lineages, as expected under non-adaptive divergence. Alternatively, if the partial MCT (between ENM and multivariate summary of morphological data) shows significant differentiation between lineages, we can hypothesize a process of adaptive divergence. Another possible outcome would be significantly conserved niches and morphologies, which would support an hypothesis of niche conservatism and phenotypic stasis (Fig. 4).
Additional evidence of demographic history can reveal important perspectives that might clarify results of the MCT and the evolutionary forces involved in divergence of lineages. For example, support for both an adaptive divergence model in MCT analyses and a model of isolation-with-migration in IM analyses, suggests that selective forces have maintained differentiation in spite of gene flow (Nosil 2008; Hey 2009). Multi-locus coalescent-based analyses may also detect population bottlenecks, providing evidence for selection associated with adaptive differentiation to the new habitat (Rosenblum et al. 2007). Alternatively, phenotypic divergence explained by neutral genetic divergence (non-significant MCT) coupled with a strict isolation model without significant gene flow, is more consistent with non-adaptive divergence. In summary, the interplay of evolutionary forces (drift, gene flow, and selection) during population divergence can result in two distinct divergence patterns (adaptive divergence with gene flow or neutral divergence in isolation) that are being currently examined with empirical data (usually in single species pairs; Nosil et al. 2008; Nosil 2009; Gay et al. 2009).

Under socially-mediated speciation model, divergence leading to possible reproductive isolation can result from a build-up or loss of color morphs (see Sinervo et al. 2008), which Corl et al. (2010b) refer to as morphic speciation (a sub-category of SMS, Appendix III). One-strategy (monomorphic) systems can be invaded and converted to two-strategy systems (dimorphic), which can be invaded by a third to generate a trimorphic RPS dynamic; this system can then collapse back to a two or a one-morph system, and these can be reconstructed in a well-designed phylogeographic study (Fig. 5). Lineage diversification is expected to be non-random under a SMS mode (Corl et al., 2010b), and closely related tip lineages should be more dissimilar in social systems, presumably for the same reasons as noted above for ecologically-driven speciation based on divergent selection. If social systems contribute to lineage divergence
and the buildup of reproductive isolation, then upon secondary contact they should begin to limit
gene flow via pre-reproductive isolating mechanisms or by reinforcement processes if hybrid
fitness is reduced by Dobzhansky-Muller incompatibilities (Corl et al. 2010b). If morphs evolve
in both sexes (or in females only) the expected phylogeographic patterns become more complex,
and will likely relate to social strategies of density regulation (Sinervo et al. 2000, 2007; Corl et
al. 2010b). Because completion of SMS may require evolution of a reinforcement mechanism
(Butlin 1989), it does not fit the strict “origin of allopatry” research paradigm described by
Wiens (2004a). Further, lineage-based phylogeographic predictions derived from this mode of
speciation [any number of social forces can drive this system besides morphs per se (Hochberg
et al. 2003; Sinervo & Clobert 2008; Sinervo et al. 2008)] are distinctly different from both
passive and adaptive divergence expectations (Table 2).

Potential limitations of this approach should be considered to reduce over-confidence in
the interpretations. First, the approach represents a minimum test to reject non-adaptive
explanations for observed between-lineage patterns of variation, but alone it does not imply that
divergence in phenotypic traits is directly responsible for speciation (if this has indeed occurred)
because correlations alone do not link the diverged phenotypes to the origin of reproductive
isolation. For example, Rundell & Price (2009) pointed out that ecological and morphological
differentiation between species can either occur together with reproductive isolation (‘ecological’
speciation) or evolve after non-adaptive speciation. In this case, coalescent-based methods can
clarify whether divergence occurred in complete isolation (e.g., no gene flow), supporting the
hypothesis of ‘non-adaptive’ speciation (Nosil 2008) even if species show different niches and
phenotypes today. Second, there is the possibility that phenotypic similarity between lineages is
due to “counter-gradient selection” in which genetic and environmental factors compensate each
other across an ecological gradient, with a net outcome of apparently conserved phenotypes (Conover & Schultz 1995; Conover et al. 2009). Third, these correlation tests cannot point unambiguously to a single process underlying observed patterns because different processes can be responsible for similar patterns (Revell et al. 2008). For example, phenotypic divergence following a neutral Brownian-like pattern can be due to either pure genetic drift or randomly fluctuating selection over time (Losos 2008).

Lastly, other kinds of analyses can also be used to evaluate association between genotype and phenotype or ecology; matrix correlation approaches are not the only options. For example, a nested random permutation procedure has been developed to test for ‘cohesion’ species based on significant associations between genetic lineages (inferred from past fragmentation events in a NCPA) and reproduction-related phenotypic and/or ecological traits (Templeton et al. 2000; Templeton 2001).

SYNTHESIS AND FUTURE DIRECTIONS

The lizard studies of ecotones and hybrid zones described provide strong evidence for the influence of natural selection in promoting population divergence across narrow environmental gradients with ongoing gene flow, in a variety of taxa and ecological contexts. Rapid ecomorphological divergence documented within some species as a result of adaptive change to different environments (Losos et al. 1997, 2006; Rosenblum 2006; Vervust et al. 2007; Herrel et al. 2008) also suggests that phenotypic divergence can occur quickly given strong selection pressures, even in the presence of gene flow. Future studies of lizard hybrid zones will continue to rely on dense sampling, multiple genetic markers, and the use of well-developed cline theory
(Sites et al. 1995; Phillips et al. 2004), but there is room for a greater focus on the role of color and other cues as these relate to mating preferences and the fitness consequences of mate selection (Stuart-Fox et al. 2009b; Corl et al. 2010b). Further, quantification of color traits coupled with an emerging understanding of the genetic basis of coloration (Morrison et al. 1995; Sinervo et al. 2001, 2006; Rosenblum et al. 2004, 2010), will permit advancement of mechanistic hypotheses about the role of color signals in promoting reproductive isolation under ecological and/or social modes of speciation. Genomics data sets will also expedite searches of “outlier loci” possibly linked to “speciation genes” under divergent selection (Hendry 2009; Nosil et al. 2009; Schluter 2009).

Lizards have figured prominently in species-delimitation studies, including the development of various approaches, comparisons of performance of different methods, and syntheses of multiple data sets. The increasing availability of nuclear markers will enhance coalescent frameworks for estimating species trees and parameters such as ancestral population sizes, divergence times, and demographic histories (Butlin et al. 2009). We suggest here also a place for non-model based approaches to the same issue, for example the two-stage hypothesis testing protocol outlined by Templeton (2001). This approach is based on the use of NCPA to:
(1) test for presence of separate evolutionary lineages as indicated by an inference for historical fragmentation at some clade level, and upon rejection of the null (panmixia), this is followed by
(2) a test that these separate lineages constitute different cohesion species. This second test evaluates whether these different lineages are genetically exchangeable and/or ecologically interchangeable within themselves, but not across lineage boundaries (Templeton 2001).
Comparative evaluations of the performance of model vs non-model approaches would be instructive in a few well-studied systems, and should different approaches prove discordant with
respect to species numbers and boundaries, the reasons would likely be apparent and informative about the validity of assumptions made by each approach (Wiens & Penkrot 2002; Marshall et al. 2006).

Bioclimatic (ENM) contributions will likely remain important for species delimitation, testing speciation hypotheses (Graham et al. 2004), predicting extinctions (Sinervo et al. 2010), and generating a priori phylogeographic predictions (Richards et al. 2007; Carnaval & Moritz 2008; Werneck et al. in review). Detailed biophysical data will not be available for the majority of species, and while limitations of ENM data as assessments of ‘niche’ are widely appreciated (omission of soils, vegetation, etc.), we suggest that morphological and other kinds of “niche” data available from museum specimens have yet to be fully explored. For example, morphological and morphometric data can be used to assess features of niche divergence not included in ENM (sexual dimorphism, differences in trophic structure, etc.; Fontanella et al. in review), and this approach could easily be expanded to assess potential competitors (congeneric species in sympatry or allopatry, for example). Other data available from vouchers include parasites and diets (food items are retained longer in poikilotherms in general, Vitt & Pianka 2003), both of which are likely to be informative about similarities or differences in niches.

The issue of sampling effects on phylogeographic inference has received little attention (but see Maddison & Knowles 2006) and future studies should distinguish among four aspects of sampling: number of individuals, number of loci, sequence length (Brito & Edwards 2009), and the number and distribution of sampling localities relative to the geographical range (Templeton 2009a; Leaché 2009). The relative impact of these different levels of sampling could be assessed by sub-sampling, which will help to incorporate ascertainment bias (Rosenblum & Novembre 2007) and to design better sampling strategies targeted to specific research questions.
Relative to other vertebrate taxa (Beheregaray 2008), fewer comparative studies of co-distributed taxa have been carried out in lizards (<4% of the studies summarized here), and we predict an increase in these kinds of studies, given rapid developments of methods for testing across multiple taxa for spatial and temporal co-divergence (see in examples in Leaché et al. 2007; Victoriano et al. 2008; Moussali et al. 2009), and other shared historical events (Moritz et al. 2009). Well-designed comparative lizard studies are also likely to contribute to biodiversity conservation via continued discovery of cryptic species, identification of regions of high diversity and endemism, and regions where evolutionary processes are likely to continue to operate (Davis et al. 2008).

We envision future phylogeographic and speciation research based on more explicit integration of multiple kinds of data from several disciplines, especially from earth sciences (Beheregaray 2008) and geography (Kidd & Ritchie 2006), and expanded assessments of phylogeographic patterns based on phenotypic and ecological data. Here we have outlined one correlative approach to compare genetic, morphological, and ecological divergence patterns in a framework easily applied to phylogeographic studies, and suggest that such data sets are capable of discriminating among alternative speciation patterns.

ACKNOWLEDGMENTS

For feedback of various drafts of this manuscript we thank Jonathan Losos, the Jim McGuire lab, the Erica Rosenblum – Luke Harmon labs, Fernanda Werneck, Mariana Morando, Michael Kearney, and Frank Fontanella. AC and JWS were supported by a NSF grant from the Office of International Science and Engineering (Collaborative Research: Establishing Sustainable
International Collaborations in Evolution, Ecology, and Conservation Biology; OISE 0530267) made to the following institutions (in alphabetical order): Brigham Young University, Centro Nacional Patagónico (Argentina), Dalhousie University, Instituto Botánico Darwinion (Argentina), Universidad Austral de Chile, Universidad Nacional de Comahue (Argentina), Universidad Nacional de Cordoba (Argentina), Universidad de Concepción (Chile), and University of Nebraska. BS was supported by an LTREB DEB 051597 (awarded to B. Sinervo and A. G. McAdam).

REFERENCES


Conover DO, Schultz ET (1995) Phenotypic similarity and the evolutionary significance of


Heled J, Drummond AJ (2008) Bayesian inference of population size history from multiple loci. BMC Evolutionary Biology, 8, 289.


phylogeography fulfilling original promise by integrating with emerging fields. Molecular Phylogenetics and Evolution, 54, 291–301.


Levene H (1953) Genetic equilibrium when more than one ecological niche is available. American Naturalist, 87, 331-335.


mutualism and transient altruism. Proceedings of the National Academy of Sciences of the USA, 102, 7372–7377.


Sinervo B, Méndez-de-la-Cruz F, Miles DB, Heulin B, Bastiaans E, Villagran-Santa Cruz M, Lara-Resendiz R, Martínez-Méndez N, Calderón-Espinosa ML, Meza-Lázaro RN, Gadsden H.


Templeton AR (2009b) Why does a method that fails continue to be used? The answer. Evolution, 63, 807–812.


Wiens JJ (2004a) What is speciation and how should we study it? American Naturalist, 163, 914–923.


Table 1. Summary of phylogeographic studies of lizards published by family (Fig. 1) through December 31, 2009. All details for each study are given in Appendix I. The second column shows the number of genera and species in each family based on the Reptile Database (www.reptile-database.org). The third column shows the number of genera sampled and the number of studies reviewed in each family. The fourth column is the mean number of localities and the fifth column represents the mean number of individuals sampled per locality (range in parentheses). Geographic region: AS = Asia, AF = Africa, AU = Australia, NA = North America, EU = Europe, SA = South America, WI = West Indies, PI = Pacific Ocean islands, AI = Atlantic Ocean islands, IO = Indic Ocean islands. Genetic marker: MT = mitochondrial DNA, NU = nuclear DNA, AZ = allozymes, MS = microsatellites, AFLP = amplified fragment length polymorphism, RFLP = restriction fragment length polymorphism, RAPD = random amplification of Polymorphic DNA, CS = chromosomes. Analytical method: (A) within-population differentiation, (B) between-population differentiation, (C) tests of population structure, (D) gene flow estimates, (E) tree-based methods, (F) nested clade phylogeographic analysis, (G) coalescent-based methods, (H) clustering/assignment/non-coalescent methods, (I) ordination and classification methods, (J) correlation analyses, (K) neutrality and equilibrium tests, (L) mating system/parentage/relatedness, and (M) cline analysis and hybrid indices.

<table>
<thead>
<tr>
<th>Family</th>
<th>Genera / Species</th>
<th>Genera / Studies</th>
<th>Geographic region</th>
<th>Number localities</th>
<th>Sample size</th>
<th>Genetic marker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agamidae</td>
<td>55/424</td>
<td>11/24</td>
<td>AS/AF/AU</td>
<td>16.2</td>
<td>11.5 (1-80)</td>
<td>MT/AZ/MS/RAPD/AFLP</td>
</tr>
<tr>
<td>Anguidae</td>
<td>12/115</td>
<td>2/3</td>
<td>NA</td>
<td>22.7</td>
<td>1.3 (1-3)</td>
<td>MT</td>
</tr>
<tr>
<td>Anniellidae</td>
<td>1/2</td>
<td>1/2</td>
<td>NA</td>
<td>26.5</td>
<td>2.0 (1-6)</td>
<td>MT/NU</td>
</tr>
<tr>
<td>Chamaeleonidae</td>
<td>9/183</td>
<td>4/6</td>
<td>AF/EU</td>
<td>32.4</td>
<td>1.5 (1-7)</td>
<td>MT/RAPD</td>
</tr>
<tr>
<td>Cordylidae</td>
<td>3/55</td>
<td>2/2</td>
<td>AF</td>
<td>9.0</td>
<td>5.2 (1-10)</td>
<td>MT</td>
</tr>
<tr>
<td>Crotaphytidae</td>
<td>2/12</td>
<td>3/7</td>
<td>NA</td>
<td>47.4</td>
<td>9.3 (1-189)</td>
<td>RFLP/AZ/MS/MT</td>
</tr>
<tr>
<td>Gekkonidae</td>
<td>101/1283</td>
<td>21/47</td>
<td>AU/EU/AS/PI/AF/NA/WI/AI</td>
<td>25.9</td>
<td>6.8 (1-87)</td>
<td>MT/AZ/NU/CS/RAPD/RFLP</td>
</tr>
<tr>
<td>Gymnophthalmidae</td>
<td>41/221</td>
<td>2/4</td>
<td>SA/WI</td>
<td>11.6</td>
<td>2.9 (1-16)</td>
<td>MT/CS/AZ</td>
</tr>
<tr>
<td>Iguanidae</td>
<td>8/39</td>
<td>9/14</td>
<td>PI/NA/CA/WI</td>
<td>17.2</td>
<td>16.3 (1-186)</td>
<td>MS/MT/AZ/RFLP/NU</td>
</tr>
<tr>
<td>Lacertidae</td>
<td>32/303</td>
<td>16/122</td>
<td>EU/AF/AS/AI</td>
<td>18.0</td>
<td>16.0 (1-542)</td>
<td>MT/AZ/MS/RAPD/RFLP/NU/CS</td>
</tr>
<tr>
<td>Phrynosomatidae</td>
<td>10/136</td>
<td>8/54</td>
<td>NA</td>
<td>24.9</td>
<td>9.5 (1-131)</td>
<td>MT/NU/AS/CS/RFLP/AFLP/MS</td>
</tr>
<tr>
<td>Polychrotidae</td>
<td>9/409</td>
<td>1/35</td>
<td>WI/NA</td>
<td>29.8</td>
<td>12.5 (1-57)</td>
<td>MT/NU/AS/MS/AFLP</td>
</tr>
<tr>
<td>Scincidae</td>
<td>133/1445</td>
<td>29/79</td>
<td>EU/AF/AS/PI/AI/NA/IO/SA</td>
<td>21.0</td>
<td>16.9 (1-202)</td>
<td>MT/MS/RFLP/NU/AZ/CS</td>
</tr>
<tr>
<td>Teiidae</td>
<td>9/127</td>
<td>5/27</td>
<td>SA/WI/NA</td>
<td>11.7</td>
<td>35.9 (1-88)</td>
<td>AZ/MT/RAPD/NU/CS/RFLP/MS</td>
</tr>
<tr>
<td>Tropiduridae</td>
<td>11/367</td>
<td>4/20</td>
<td>SA/PI</td>
<td>24.8</td>
<td>10.6 (1-64)</td>
<td>MT/CS/AZ/MS/RFLP</td>
</tr>
<tr>
<td>Varanidae</td>
<td>1/69</td>
<td>1/3</td>
<td>AS</td>
<td>10.7</td>
<td>10.7 (6-27)</td>
<td>MS/MT</td>
</tr>
<tr>
<td>Xantusiidae</td>
<td>3/30</td>
<td>1/4</td>
<td>NA</td>
<td>66.8</td>
<td>2.2 (1-18)</td>
<td>MT/AZ/NU</td>
</tr>
<tr>
<td>Family</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
</tr>
<tr>
<td>-------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Agamidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Anguidae</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Anniellidae</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chamaeleonidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cordylidae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crotaphytidae</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gekkonidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gymnophthalmidae</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iguanidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lacertidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Phrynosomatidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Polychrotidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Scincidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Teiidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Tropiduridae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Varanidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Xantusiidae</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Table 2. Expected patterns of divergence between sister species (or phylogroups) in allopatry, under alternative speciation scenarios.

<table>
<thead>
<tr>
<th>Pattern-based divergence</th>
<th>Non-adaptive (passive) divergence</th>
<th>Adaptive divergence</th>
<th>Phenotypic stasis</th>
<th>Socially-mediated speciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological variation (size or shape)(^1)</td>
<td>Variance and co-variance of traits within lineages proportional to those between lineages</td>
<td>Divergence due to adaptation to a novel environment reflected in divergent ecomorphological traits</td>
<td>No divergence due to failure to adapt to novel environment reflected in conserved ecomorphological traits</td>
<td>If male based polymorphism, then color genes will have effect on male size and accentuate sexual dimorphism</td>
</tr>
<tr>
<td>Sexual dimorphism(^2)</td>
<td>Divergence as above</td>
<td>Divergence in allopatry (or no divergence)</td>
<td>Divergence in allopatry (or no divergence)</td>
<td>Female only morphs – females larger than males; male only morphs, males larger than females Correlated to shifts in mating strategies (1→2→3, in phylogenetic sequence) in males OR in females (r)- vs. (K)-strategies</td>
</tr>
<tr>
<td>Color pattern polymorphism(^3)</td>
<td>Divergence as above</td>
<td>Divergence in allopatry (or no divergence)</td>
<td>Divergence in allopatry (or no divergence)</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)These are characters thought to be influenced by natural selection favoring adaptation to niche dimensions such as crypsis, microhabitat, thermoregulation, or interactions related to competition, predation, or parasitism.

\(^2\)Characters such as color and/or body size differences usually attributed to the influence of sexual selection or natural selection on female fecundity or male resource defense.

\(^3\)Characters such as color polymorphisms segregating within a single breeding group, and attributed to frequency–dependent selection on local mating dynamics.
List of Figures

Figure 1. Schematic phylogeny of squamate reptiles showing relationships of major lizard clades with non-lizard taxa (Amphisbaenia and Serpentes), modified from Townsend et al. (2004). The geographic distributions of lizard lineages are identified as follows: NW = New World, NA = North America, SA = South America, CA = Central America, WP = West Pacific, MA = Madagascar, OW = Old World, COSM = Cosmopolitan, AS = Asia, EUAS = Eurasia, and AF = Africa. The relationships within the clade Iguania were taken from Schulte et al. (2003).

Figure 2. Annual number of phylogeographic studies published between 1980-2009. The curve represents the best-fit exponential function for the period 1980-2008 ($R^2 = 0.92$, $P < 0.01$). Studies from 2009 were not included in the regression analysis because the number of references from last year found in the internet databases is probably an underestimate of the real number of publications.

Figure 3. Annual distribution of study cases between 1980–2009 for four classes of genetic markers, allozymes, RFLP/AFLP/RAPD, mtDNA, nuclear DNA, and microsatellites.

Figure 4. Diagram of inferential steps for evaluating three common speciation modes. First, a partial matrix correlation test (MCT) between phenotype and environment controlled by genotype (based on neutral markers) is used to test the null expectation: random drift accounts for the observed phenotypic divergence. Absence of significant correlation is consistent with the null model. Significant MCT between-lineage differences in phenotype and environment (based
on niche models) support an adaptive divergence scenario, and indistinguishable phenotypes and environments are consistent with a niche conservatism/phenotypic stasis model. The bottom panel summarizes patterns of phenotypic divergence under the three speciation modes: (A) non-adaptive divergence of phenotypes in isolated lineages due to random drift, (B) retention of conserved phenotypes more similar to each other than expected due to phenotypic stasis, and (C) phenotypes are more different from each other than expected due to adaptive divergence. Solid lines indicate the realized phenotypic divergence and dotted lines represent the expected pattern due to non-adaptive divergence. Socially-mediated speciation can potentially occur under all three scenarios since sexual selection can produce divergence in phenotypic traits (e.g., colour morphs) linked to mating systems and RPS dynamics (see text).

Figure 5. Hypothetical phylogeny illustrating taxa with the predicted signature of socially mediated speciation based on a RPS set of color morphs. Monomorphic outgroups suggest a blue ancestral color/mating system, which was subsequently, the blue is invaded by orange to create a dimorphic mating system, which is then invaded by yellow to create the full complement of three colors. This RPS mating system can also generate new mono- and dimorphic taxa by subsequent loss of morphs. Notice that the three monomorphic descendants of an RPS ancestor exhibit only one morph, but each descendant species is a different color. We label terminals as species here for convenience, but these can also be divergent intra-specific clades (Corl et al. 2010b).
Figure 2
Figure 3
Figure 4
Figure 5
Appendix I

We searched for references published in English using CrossSearch in the ISI Web of Knowledge website using the following key words: lizard, squamata (excluding snakes), phylogeography, speciation, population, structure, gene, cline, contact zone, hybridization, hybrid zone, and gene flow. Searches were performed on three main databases: ISI Web of Science (1984-present), BIOSIS (1980-present), and Zoological Record (1978-present). Articles published until December 31, 2009 have been retrieved using saved searches and e-mail alert service from the Web of Science, but some references were also obtained from cited references of articles and from authors’ internet websites, that were not recovered in the literature searches. Articles that focused strictly on the reconstruction of phylogenetic relationships were excluded, and only those that addressed the phylogeography of a single species or a complex of closely related species were included for this review. This review also included references that sampled a few or one locality when estimated demographic parameters were based on a large number of specimens.

The following information was extracted from each reference: species name, family name, geographic region, number of sampled localities, mean and range of sample size per locality, genetic marker(s), number of loci, analytical methods, author(s), and publication year. The current nomenclature and taxonomic diversity are based on the JCVI-TIGR Reptile Database (http://www.reptile-database.org/). If papers focused on two or more species, each species was considered as a separate study case. Studies were assigned to one or more of the following geographic regions: North America (NA), South America (SA), Africa (AF), Europe (EU), Asia (AS), Central America (CA), West Indies (WI), Australia (AU), Atlantic Ocean islands (AI), and Pacific Ocean islands (PI). Genetic markers were grouped according to the targeted genetic element and/or technique used to quantify genetic variation: chromosomes, allozymes, mitochondrial DNA sequences (mtDNA), nuclear DNA sequence (nuDNA), microsatellites/minisatellites, and RFLP/AFLP/RAPD fragments. Analytical methods were grouped into the following categories: (A) within-population differentiation: karyotypes, nucleotide and haplotype diversity, heterozygosity, allelic richness, allele number / size range per locus; (B) between-population differentiation: similarity indexes, Nei’s & Roger’s distances, uncorrected divergence; (C) tests of population structure: Hardy-Weinberg tests, AMOVA/SAMOVA, BARRIER, hierarchical structure using nucleotide diversity
and tree nodes, Chi-square, G and Fisher’s exact tests of frequencies, randomization, Wilcoxon-sign rank test; (D) gene flow estimates: Fst, Gst, Rst, Nst, Cockerham & Weirs θ, rare alleles method of Slatkin (1985); (E) tree-based methods: phylogenetic trees and networks; (F) nested clade phylogeographic analysis; (G) coalescent-based methods: programs FLUCTUATE, BOTTLENECK, IM/IMa, MIGRATE, MCMCcoal, GENETREE, MS-BAYES, MESQUITE; (H) clustering/assignment/non-coalescent methods: programs STRUCTURE, STRUCTURAMA, TESS, BAPS, GENECLASS, GENELAND, NEWHYBRIDS, BAYALLELE, FSTAT (assignment index); (I) ordination and classification methods: PCA, DA, MDS, cluster analysis (UPGMA), factorial correspondence analysis; (J) correlation analyses: Fst vs. geographic distance isolation, Mantel’s test, spatial autocorrelation, matrix correspondence test, genetic landscape surface; (K) neutrality and equilibrium tests: Fu’s and Tajima’s tests, mismatch analysis, Hudson-Kreitman-Aguadé test; (L) mating system/parentage/relatedness; and (M) cline analysis and hybrid indices.

In the table below we list all references about lizard phylogeography reviewed in this study indicating the family and species studied, the geographic region, the number of localities sampled, the sample size representing the mean and range of individuals sampled per locality, the genetic markers used, the methods of analyses applied, the citation, and the journal where the study was published. References included in this review are listed below the summary table. Abbreviations of column headers: GR=Geographic Region, NL=Number of Localities, SS=Sample Size, GM=Genetic Marker(s), and AM=Analytical Method(s).

<table>
<thead>
<tr>
<th>Family/Species</th>
<th>GR</th>
<th>NL</th>
<th>SS</th>
<th>GM</th>
<th>AM</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGAMIDAE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acanthosaura spp.</td>
<td>AS</td>
<td>27</td>
<td>2.3 (1-12)</td>
<td>MT</td>
<td>E</td>
<td>Kalyabina-Hauf et al. (2004a)</td>
</tr>
<tr>
<td>Agama atra</td>
<td>AF</td>
<td>13</td>
<td>3.0 (1-5)</td>
<td>MT</td>
<td>E</td>
<td>Matthee &amp; Flemming (2002)</td>
</tr>
<tr>
<td>Agama atra</td>
<td>AF</td>
<td>46</td>
<td>2.5 (2-16)</td>
<td>MT</td>
<td>ACDEJ</td>
<td>Swart et al. (2009)</td>
</tr>
<tr>
<td>Agama atra</td>
<td>AF</td>
<td>47</td>
<td>2.5 (2-16)</td>
<td>MT</td>
<td>EGJ</td>
<td>Tolley et al. (2009)</td>
</tr>
<tr>
<td>Agama impalearis</td>
<td>AF</td>
<td>17</td>
<td>5.2 (1-10)</td>
<td>MT</td>
<td>EF</td>
<td>Brown et al. (2002)</td>
</tr>
<tr>
<td>Amphibolorus nobbi</td>
<td>AU</td>
<td>14</td>
<td>6.6 (3-14)</td>
<td>AZ/MT</td>
<td>DEF</td>
<td>Driscoll &amp; Hardy (2005)</td>
</tr>
<tr>
<td>Chlamydosaurus kingii</td>
<td>AU</td>
<td>3</td>
<td>24.7 (39-14)</td>
<td>MT</td>
<td>BE</td>
<td>Ujvari et al. (2007)</td>
</tr>
<tr>
<td>Chlamydosaurus kingii</td>
<td>AU</td>
<td>3</td>
<td>24.7 (39-14)</td>
<td>MT/AFLP</td>
<td>EHK</td>
<td>Ujvari et al. (2008)</td>
</tr>
<tr>
<td>Ctenophorus ornatus</td>
<td>AU</td>
<td>1</td>
<td>80</td>
<td>MS</td>
<td>L</td>
<td>Lebas (2001)</td>
</tr>
<tr>
<td>Laudakia caucasia species group</td>
<td>AS</td>
<td>10</td>
<td>1.0 (1)</td>
<td>MT</td>
<td>E</td>
<td>Macey et al. (1998)</td>
</tr>
<tr>
<td>Laudakia caucasia species group</td>
<td>AS</td>
<td>12</td>
<td>1.0 (1)</td>
<td>MT</td>
<td>E</td>
<td>Macey et al. (2000)</td>
</tr>
<tr>
<td>Leiolepis reevesii/guentherpetersi</td>
<td>AS</td>
<td>2</td>
<td>4.5 (3-6)</td>
<td>MS</td>
<td>B</td>
<td>Malsheva et al. (2006)</td>
</tr>
<tr>
<td>Species</td>
<td>Geographic Region</td>
<td>Size Range</td>
<td>Analysis Method</td>
<td>Authors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-------------------</td>
<td>------------</td>
<td>-----------------</td>
<td>----------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phrynocephalus helioscopus</td>
<td>AS</td>
<td>2.1 (1-3)</td>
<td>MT</td>
<td>Pang et al. (2003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phrynocephalus przewalskii</td>
<td>AS</td>
<td>2.7 (?)</td>
<td>MS</td>
<td>Urruchart et al. (2005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phrynocephalus versicolor complex</td>
<td>AS</td>
<td>1.9 (1-4)</td>
<td>MT</td>
<td>Wang &amp; Fu (2004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phrynocephalus vlanguilii</td>
<td>AS</td>
<td>9.2 (1-25)</td>
<td>MT</td>
<td>ABCEFK J. et al. (2008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phrynocephalus vlanguilii</td>
<td>AS</td>
<td>58.5 (44-72)</td>
<td>MS/M T</td>
<td>Wang et al. (2009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trapelus aegilis</td>
<td>AS</td>
<td>7.0 (1-11)</td>
<td>AZ</td>
<td>Macey &amp; Ananjeva (2004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tymanocryptis pinguicolla</td>
<td>AU</td>
<td>3.0 (1-8)</td>
<td>MT</td>
<td>Melville et al. (2007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tymanocryptis pinguicolla</td>
<td>AU</td>
<td>7.0 (4-13)</td>
<td>MT</td>
<td>Scott &amp; Keogh (2000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tymanocryptis spp</td>
<td>AU</td>
<td>1.3 (1-2)</td>
<td>AZ</td>
<td>Smith et al. (1999)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uromastyx spp.</td>
<td>AF/AS</td>
<td>1.3 (1-2)</td>
<td>MT</td>
<td>Amer &amp; Kumazawa (2005)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ANGUIDAE**

<table>
<thead>
<tr>
<th>Species</th>
<th>Geographic Region</th>
<th>Size Range</th>
<th>Analysis Method</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barisia imbricata</td>
<td>NA</td>
<td>1.0 (1)</td>
<td>MT</td>
<td>Zaldivar-Riveron et al. (2005)</td>
</tr>
<tr>
<td>Elgaria multicarinata</td>
<td>NA</td>
<td>1.1 (1-2)</td>
<td>MT</td>
<td>Feldman &amp; Spicer (2006)</td>
</tr>
<tr>
<td>Elgaria multicarinata</td>
<td>NA</td>
<td>1.7 (1-3)</td>
<td>MT</td>
<td>Mahoney et al. (2003)</td>
</tr>
</tbody>
</table>

**ANNIELLIDAE**

<table>
<thead>
<tr>
<th>Species</th>
<th>Geographic Region</th>
<th>Size Range</th>
<th>Analysis Method</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anniella pulchra</td>
<td>NA</td>
<td>2.4 (1-4)</td>
<td>MT</td>
<td>Pearse &amp; Pogson (2000)</td>
</tr>
<tr>
<td>Anniella pulchra</td>
<td>NA</td>
<td>1.5 (1-6)</td>
<td>MT/NU</td>
<td>Parham &amp; Papenfuss (2009)</td>
</tr>
</tbody>
</table>

**CHAMAEOLEONIDAE**

<table>
<thead>
<tr>
<th>Species</th>
<th>Geographic Region</th>
<th>Size Range</th>
<th>Analysis Method</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brachypodion spp.</td>
<td>AF</td>
<td>?</td>
<td>?</td>
<td>Tolley et al. (2006)</td>
</tr>
<tr>
<td>Brachypodion spp.</td>
<td>AF</td>
<td>1.6 (1-7)</td>
<td>MT</td>
<td>Tolley et al. (2009)</td>
</tr>
<tr>
<td>Calumma brevicorne complex</td>
<td>AF</td>
<td>2.1 (1-5)</td>
<td>MT</td>
<td>Boumans et al. (2007)</td>
</tr>
<tr>
<td>Chamaeleo chamaeleon</td>
<td>EU/A F</td>
<td>1.0 (1)</td>
<td>RAPD</td>
<td>Padilla et al. (2004)</td>
</tr>
<tr>
<td>Furcifer lateralis</td>
<td>AF</td>
<td>1.7 (1-5)</td>
<td>MT</td>
<td>Boumans et al. (2007)</td>
</tr>
<tr>
<td>Furcifer polleni</td>
<td>AF</td>
<td>1.3 (1-2)</td>
<td>MT</td>
<td>Rocha et al. (2005a)</td>
</tr>
</tbody>
</table>

**CORDYLIDAE**

<table>
<thead>
<tr>
<th>Species</th>
<th>Geographic Region</th>
<th>Size Range</th>
<th>Analysis Method</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cordylus cordylus/oelofseni</td>
<td>AF</td>
<td>8.5 (7-10)</td>
<td>MT</td>
<td>Daniels et al. (2004)</td>
</tr>
<tr>
<td>Platysaurus spp.</td>
<td>AF</td>
<td>1.8 (1-4)</td>
<td>MT</td>
<td>Scott et al. (2004)</td>
</tr>
</tbody>
</table>

**CROTAPHYTIDAE**

<table>
<thead>
<tr>
<th>Species</th>
<th>Geographic Region</th>
<th>Size Range</th>
<th>Analysis Method</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crotaphytus collaris</td>
<td>NA</td>
<td>6.2 (2-15)</td>
<td>RFLP</td>
<td>Campbell &amp; McCoy (2002)</td>
</tr>
<tr>
<td>Crotaphytus collaris</td>
<td>NA</td>
<td>31.5 (2-189)</td>
<td>AZ</td>
<td>Hranitz &amp; Baird (2000)</td>
</tr>
<tr>
<td>Crotaphytus collaris</td>
<td>NA</td>
<td>12.1 (10-21)</td>
<td>MS</td>
<td>Hutchison &amp; Templeton (1999)</td>
</tr>
<tr>
<td>Crotaphytus collaris</td>
<td>NA</td>
<td>11.0 (?)</td>
<td>MS</td>
<td>Hutchison (2003)</td>
</tr>
<tr>
<td>Crotaphytus collaris</td>
<td>NA</td>
<td>1.0 (1)</td>
<td>MT</td>
<td>Hutchison et al. (1999)</td>
</tr>
<tr>
<td>Gambelia wislizenii</td>
<td>NA</td>
<td>1.4 (1-3)</td>
<td>MT</td>
<td>Orange et al. (1999)</td>
</tr>
<tr>
<td>Crotaphytus spp./Gambelia spp.</td>
<td>NA</td>
<td>2.2 (1-?)</td>
<td>MT</td>
<td>McGuire et al. (2007)</td>
</tr>
</tbody>
</table>

**GEKKONIDAE**

<table>
<thead>
<tr>
<th>Species</th>
<th>Geographic Region</th>
<th>Size Range</th>
<th>Analysis Method</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carphodactylus laevis</td>
<td>AU</td>
<td>3.4 (?)</td>
<td>MT</td>
<td>Schneider et al. (1998)</td>
</tr>
<tr>
<td>Christinus marmoratus</td>
<td>AU</td>
<td>3.3 (1-41)</td>
<td>AZ</td>
<td>Donnellan et al. (2000)</td>
</tr>
<tr>
<td>Cyrtodactylus kotschyi</td>
<td>EU/AS</td>
<td>2.3 (1-5)</td>
<td>MT</td>
<td>Kasapidis et al. (2005)</td>
</tr>
<tr>
<td>Species/Complex</td>
<td>Distribution</td>
<td>Mitochondrial locus</td>
<td>Nucleolar locus</td>
<td>Notes</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>--------------</td>
<td>---------------------</td>
<td>----------------</td>
<td>-------</td>
</tr>
<tr>
<td>Diplodactylus pulcher/stenodactylus group</td>
<td>AU</td>
<td>MT/NU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diplodactylus viattus complex</td>
<td>AU</td>
<td>MT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gekhyra oceanica/matifatina</td>
<td>PI/AS</td>
<td>CS/AZ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gekhyra variegata/nana</td>
<td>AU</td>
<td>7.4 (4-10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gekhyra variegata/punctata group</td>
<td>AU</td>
<td>5.6 (1-87)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gekko gecko</td>
<td>AS</td>
<td>?</td>
<td>RAPD</td>
<td>Qin et al. (2005)</td>
</tr>
<tr>
<td>Gekko hokouensis</td>
<td>AS</td>
<td>18.9 (14-22)</td>
<td>AZ</td>
<td>ABDI</td>
</tr>
<tr>
<td>Gekko tawaensis/japonicus</td>
<td>AS</td>
<td>27.1 (6-40)</td>
<td>AZ</td>
<td>BCI</td>
</tr>
<tr>
<td>Gekko tawaensis/japonicus</td>
<td>AS</td>
<td>17.4 (4-38)</td>
<td>AZ</td>
<td>ABDEJ</td>
</tr>
<tr>
<td>Gekko yakenensis/hokouensis</td>
<td>AS</td>
<td>23.1 (?)</td>
<td>AZ</td>
<td>AHM</td>
</tr>
<tr>
<td>Gymnodactylus darwinii complex</td>
<td>SA</td>
<td>1.9 (1-4)</td>
<td>MT</td>
<td>EJK</td>
</tr>
<tr>
<td>Hemidactylus mercatorius</td>
<td>AF</td>
<td>1.6 (1-3)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td>Hemidactylus spp.</td>
<td>AF</td>
<td>1.7 (1-4)</td>
<td>MT/NU</td>
<td>E</td>
</tr>
<tr>
<td>Heteronotia binoei complex</td>
<td>AU</td>
<td>2.8 (1-8)</td>
<td>RFLP</td>
<td>AE</td>
</tr>
<tr>
<td>Heteronotia binoei complex</td>
<td>AU</td>
<td>3.1 (1-9)</td>
<td>RFLP</td>
<td>AE</td>
</tr>
<tr>
<td>Heteronotia binoei complex</td>
<td>AU</td>
<td>3.0 (1-13)</td>
<td>RFLP</td>
<td>E</td>
</tr>
<tr>
<td>Heteronotia binoei complex</td>
<td>AU</td>
<td>3.3 (1-25)</td>
<td>AZ</td>
<td>A</td>
</tr>
<tr>
<td>Heteronotia binoei complex</td>
<td>AU</td>
<td>9.2 (4-36)</td>
<td>CS/AZ</td>
<td>ABCDEI</td>
</tr>
<tr>
<td>Lepidodactylus lugubris</td>
<td>PI</td>
<td>4.0 (1-10)</td>
<td>AZ</td>
<td>A</td>
</tr>
<tr>
<td>Lepidodactylus lugubris</td>
<td>PI</td>
<td>2.7 (2-4)</td>
<td>CS</td>
<td>A</td>
</tr>
<tr>
<td>Lepidodactylus lugubris/moestus/spp.</td>
<td>PI</td>
<td>6.0 (1-4)</td>
<td>CS/AZ/MT</td>
<td>AE</td>
</tr>
<tr>
<td>Lepidodactylus spp.</td>
<td>PI</td>
<td>?</td>
<td>RFLP/NU</td>
<td>E</td>
</tr>
<tr>
<td>Lucasium stenodactylum</td>
<td>AU</td>
<td>2.2 (1-7)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td>Nactus arnouxii</td>
<td>AU/PI</td>
<td>5.5 (2-25)</td>
<td>CS/AZ</td>
<td>EBE</td>
</tr>
<tr>
<td>Nactus pelagicus</td>
<td>AU/AS/PI</td>
<td>4.9 (3-1)</td>
<td>CS/AZ/RFLP</td>
<td>ABDE</td>
</tr>
<tr>
<td>Oedura leseurii</td>
<td>AU</td>
<td>7.1 (5-10)</td>
<td>AZ</td>
<td>BE</td>
</tr>
<tr>
<td>Oedura leseurii</td>
<td>AU</td>
<td>4.3 (1-8)</td>
<td>MT/NU</td>
<td>AE</td>
</tr>
<tr>
<td>Phelsuma spp.</td>
<td>AF</td>
<td>3.0 (1-8)</td>
<td>MT/NU</td>
<td>AE</td>
</tr>
<tr>
<td>Phelsuma spp.</td>
<td>AF</td>
<td>1.95 (1-9)</td>
<td>MT</td>
<td>BE</td>
</tr>
<tr>
<td>Phelsuma lineata</td>
<td>AF</td>
<td>1.4 (1-4)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td>Phyllodactylus lineatus complex</td>
<td>AF</td>
<td>4.8 (1-9)</td>
<td>AZ</td>
<td>BE</td>
</tr>
<tr>
<td>Phyllodactylus unctus</td>
<td>NA</td>
<td>6.5 (5-8)</td>
<td>AZ</td>
<td>A</td>
</tr>
<tr>
<td>Phyllurus ossa</td>
<td>AU</td>
<td>6.8 (3-11)</td>
<td>MT</td>
<td>ACE</td>
</tr>
<tr>
<td>Rhacodactylus sp.</td>
<td>PI</td>
<td>2.3 (1-7)</td>
<td>AZ</td>
<td>BE</td>
</tr>
<tr>
<td>Rhinochordura ornata</td>
<td>AU</td>
<td>1.3 (1-2)</td>
<td>MT/NU</td>
<td>E</td>
</tr>
<tr>
<td>Saltuarius cornutus</td>
<td>AU</td>
<td>2.7 (?)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td>Sphaerodactylus fantasticus</td>
<td>WI</td>
<td>10 (10)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td>Sphaerodactylus nicholsi</td>
<td>WI</td>
<td>15.3 (5-26)</td>
<td>AZ</td>
<td>ABD</td>
</tr>
<tr>
<td>Tarentola delalandii</td>
<td>AI</td>
<td>2.5 (1-10)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td>Tarentola mauritanica</td>
<td>EU/AU</td>
<td>1.3 (1-3)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td>Tarentola mauritanica</td>
<td>EU</td>
<td>2.3 (?)</td>
<td>MT</td>
<td>E</td>
</tr>
</tbody>
</table>
**GYMNOPHTHALMIDAE**

<table>
<thead>
<tr>
<th>Species</th>
<th>Code</th>
<th>Length (mm)</th>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gymnophthalmus underwoodi/speciosus</td>
<td>SA/WI</td>
<td>9</td>
<td>MT</td>
<td>E Kizirian et al. (1999)</td>
</tr>
<tr>
<td>G. underwoodi/speciosus/pleei</td>
<td>SA/WI</td>
<td>5.5</td>
<td>CS/AZ</td>
<td>A Cole et al. (1990)</td>
</tr>
<tr>
<td>Protoporus pachyrurus group</td>
<td>SA</td>
<td>20</td>
<td>MT</td>
<td>E Doan et al. (2005)</td>
</tr>
</tbody>
</table>

**IGUANIDAE**

<table>
<thead>
<tr>
<th>Species</th>
<th>Code</th>
<th>Length (mm)</th>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amblyrhynchus cristatus</td>
<td>PI</td>
<td>16.7</td>
<td>MS/MT</td>
<td>ACDEJ Rassmann et al. (1997a)</td>
</tr>
<tr>
<td>Amblyrhynchus cristatus</td>
<td>PI</td>
<td>23</td>
<td>MS/MT</td>
<td>CD Rassmann et al. (2004)</td>
</tr>
<tr>
<td>Brachylophus fasciatus / B. vitiensis</td>
<td>PI</td>
<td>20</td>
<td>MS/MT</td>
<td>ABCDEIK Burns et al. (2006)</td>
</tr>
<tr>
<td>Calisaurus draconoides</td>
<td>NA</td>
<td>12</td>
<td>AZ</td>
<td>B Adest (1987)</td>
</tr>
<tr>
<td>Conolophus cristatus / A. cristatus</td>
<td>PI</td>
<td>4.5</td>
<td>MT/RFLP</td>
<td>A Rassmann et al. (1997b)</td>
</tr>
<tr>
<td>Conolophus spp.</td>
<td>PI</td>
<td>9</td>
<td>MT</td>
<td>C Rassmann et al. (2004)</td>
</tr>
<tr>
<td>Conolophus spp.</td>
<td>PI</td>
<td>16</td>
<td>MS</td>
<td>ABCGH Tzika et al. (2008)</td>
</tr>
<tr>
<td>Ctenosaura quinquecarinata</td>
<td>CA</td>
<td>21</td>
<td>MT</td>
<td>EF Hashbun et al. (2005)</td>
</tr>
<tr>
<td>Ctenosaura pectinata</td>
<td>NA</td>
<td>44</td>
<td>MT/NU</td>
<td>EFK Zarza et al. (2008)</td>
</tr>
<tr>
<td>Cyclura carinata carinata</td>
<td>WI</td>
<td>29</td>
<td>MS</td>
<td>CDJ Welch et al. (2004)</td>
</tr>
<tr>
<td>Dipscosaursaurus dorsalis</td>
<td>NA</td>
<td>13</td>
<td>RFLP</td>
<td>E Lamb et al. (1992)</td>
</tr>
<tr>
<td>Iguana iguana</td>
<td>CA</td>
<td>5</td>
<td>AZ</td>
<td>CD Bock &amp; McCracken (1988)</td>
</tr>
<tr>
<td>Sauromalus obesus</td>
<td>NA</td>
<td>25</td>
<td>RFLP</td>
<td>E Lamb et al. (1992)</td>
</tr>
</tbody>
</table>

**LACERTIDAE**

<table>
<thead>
<tr>
<th>Species</th>
<th>Code</th>
<th>Length (mm)</th>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acanthodactylus erythrurus</td>
<td>EU/AF</td>
<td>14</td>
<td>MT</td>
<td>E Harris et al. (2004a)</td>
</tr>
<tr>
<td>Acanthodactylus spp.</td>
<td>AF</td>
<td>15</td>
<td>AZ</td>
<td>ABEI Blanc &amp; Cariou (1987)</td>
</tr>
<tr>
<td>Acanthodactylus erythrurus</td>
<td>EU/AF</td>
<td>22</td>
<td>MT/NU</td>
<td>E Fonseca et al. (2009)</td>
</tr>
<tr>
<td>Archaeolacerta bedriagae</td>
<td>EU</td>
<td>13</td>
<td>AZ</td>
<td>ABCEJ Salvi et al. (2009a)</td>
</tr>
<tr>
<td>Archaeolacerta bedriagae</td>
<td>EU</td>
<td>5</td>
<td>AZ</td>
<td>ABCDI Salvi et al. (2009b)</td>
</tr>
<tr>
<td>Darevskia armeniaca</td>
<td>AS</td>
<td>4</td>
<td>MS</td>
<td>A Malysheva et al. (2007)</td>
</tr>
<tr>
<td>Darevskia armeniaca</td>
<td>AS</td>
<td>3</td>
<td>MS</td>
<td>B Martirosyan et al. (2003)</td>
</tr>
<tr>
<td>Darevskia armeniaca</td>
<td>AS</td>
<td>3</td>
<td>AZ</td>
<td>Fu et al. (2000b)</td>
</tr>
<tr>
<td>Darevskia armeniaca</td>
<td>AS</td>
<td>7</td>
<td>AZ</td>
<td>A MacCulloch et al. (1995b)</td>
</tr>
<tr>
<td>Darevskia armeniaca/dahli/unisexualis</td>
<td>AS</td>
<td>8</td>
<td>MS/ABD</td>
<td>Petrosian et al. (2003)</td>
</tr>
<tr>
<td>D. bendimahienis/sapphirina/tuzzelli</td>
<td>AS</td>
<td>3</td>
<td>AZ</td>
<td>A Fu et al. (2000c)</td>
</tr>
<tr>
<td>Darevskia caucasia complex</td>
<td>AS</td>
<td>4</td>
<td>AZ</td>
<td>ABCDI Fu et al. (1995)</td>
</tr>
<tr>
<td>Darevskia dahli</td>
<td>AS</td>
<td>2</td>
<td>MS</td>
<td>A Kan et al. (1998)</td>
</tr>
<tr>
<td>Darevskia dahli</td>
<td>AS</td>
<td>6</td>
<td>AZ</td>
<td>A Murphy et al. (1997)</td>
</tr>
<tr>
<td>Darevskia dahli/armeniaca/unisexualis</td>
<td>AS</td>
<td>9</td>
<td>MS</td>
<td>A Tokarskaya et al. (2001)</td>
</tr>
<tr>
<td>Darevskia derjugini/praticola</td>
<td>AS</td>
<td>5</td>
<td>AZ</td>
<td>ACD MacCulloch et al. (1997c)</td>
</tr>
<tr>
<td>Darevskia mixta/armeniaca/dahli</td>
<td>AS</td>
<td>12</td>
<td>MT</td>
<td>E Fu et al. (1999)</td>
</tr>
<tr>
<td>Darevskia portschinskii</td>
<td>AS</td>
<td>4</td>
<td>AZ/ACD</td>
<td>E MacCulloch et al. (1997a)</td>
</tr>
<tr>
<td>Darevskia praticola/derjugini/rudis</td>
<td>AS</td>
<td>10</td>
<td>RAPD</td>
<td>B Ryabinina et al. (2002)</td>
</tr>
<tr>
<td>Darevskia radei</td>
<td>AS</td>
<td>17</td>
<td>IS-PCR/RAPD</td>
<td>AB Grechko et al. (2007)</td>
</tr>
<tr>
<td>Darevskia raddaei complex</td>
<td>AS</td>
<td>21</td>
<td>MT</td>
<td>E Fu et al. (2000a)</td>
</tr>
<tr>
<td>Species/Mixture</td>
<td>Country</td>
<td>Reference</td>
<td>Method</td>
<td>Genotypes</td>
</tr>
<tr>
<td>-----------------------------------------------------</td>
<td>---------</td>
<td>-----------</td>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td><em>Mesalina guttulata</em></td>
<td>AS</td>
<td>Mayer &amp; Tiedemann (1985)</td>
<td>RFLP</td>
<td>B</td>
</tr>
<tr>
<td><em>Darevskia raddei/nairensis</em></td>
<td>AS</td>
<td>Bobyn et al. (1996)</td>
<td>RFLP/MT/RAPD</td>
<td>E</td>
</tr>
<tr>
<td><em>Darevskia rostombekovi</em></td>
<td>AS</td>
<td>Martirosyan et al. (2002)</td>
<td>MT</td>
<td>B</td>
</tr>
<tr>
<td><em>Darevskia rostombekovi</em></td>
<td>AS</td>
<td>MacCulloch et al. (1997b)</td>
<td>MT</td>
<td>B</td>
</tr>
<tr>
<td><em>Darevskia saxicola group</em></td>
<td>AS</td>
<td>MacCulloch et al. (2000)</td>
<td>MT</td>
<td>D</td>
</tr>
<tr>
<td><em>Darevskia saxicola complex</em></td>
<td>AS</td>
<td>Ryabina et al. (1998)</td>
<td>MT</td>
<td>A</td>
</tr>
<tr>
<td><em>Darevskia unisexualis</em></td>
<td>AS</td>
<td>Korchagin et al. (2004)</td>
<td>MT</td>
<td>A</td>
</tr>
<tr>
<td><em>Darevskia unisexualis</em></td>
<td>AS</td>
<td>Ryskov et al. (2003)</td>
<td>MT</td>
<td>B</td>
</tr>
<tr>
<td><em>Darevskia unisexualis</em></td>
<td>AS</td>
<td>Tokarskaya et al. (2004)</td>
<td>MT</td>
<td>A</td>
</tr>
<tr>
<td><em>Darevskia unisexualis</em></td>
<td>AS</td>
<td>Fu et al. (1998)</td>
<td>MT</td>
<td>B</td>
</tr>
<tr>
<td><em>Darevskia unisexualis</em></td>
<td>AS</td>
<td>Ryskov et al. (2000)</td>
<td>MT</td>
<td>A</td>
</tr>
<tr>
<td><em>Darevskia unisexualis</em></td>
<td>AS</td>
<td>Tokarskaya et al. (2000)</td>
<td>MT</td>
<td>A</td>
</tr>
<tr>
<td><em>Darevskia valentini/portschinskii/rudis</em></td>
<td>AS</td>
<td>Ljubisavljevic et al. (2007)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Gallotia galloti</em></td>
<td>AS</td>
<td>Bloor et al. (2008)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Gallotia galloti</em></td>
<td>AS</td>
<td>Brown et al. (2006)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Gallotia galloti</em></td>
<td>AS</td>
<td>Thorpe &amp; Richards (2001)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Gallotia galloti</em></td>
<td>AS</td>
<td>Thorpe (1996)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Gallotia galloti</em></td>
<td>AS</td>
<td>Thorpe et al. (1994)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Gallotia galloti</em></td>
<td>AS</td>
<td>Thorpe et al. (1993b)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Gallotia galloti</em></td>
<td>AS</td>
<td>Thorpe et al. (1993a)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Gallotia galloti</em></td>
<td>AS</td>
<td>Almeida et al. (2002)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Iberolacerta monticolica/cyreni</em></td>
<td>AS</td>
<td>Grechko et al. (2006)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>I. monticolica/cyreni/martinezricai</em></td>
<td>AS</td>
<td>Gullberg et al. (1998)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Iberolacerta spp.</em></td>
<td>AS</td>
<td>Gullberg et al. (1999)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Iberolacerta spp.</em></td>
<td>AS</td>
<td>Gullberg et al. (1999)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta agilis</em></td>
<td>AS</td>
<td>Beebee &amp; Rowe (2001)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta agilis</em></td>
<td>AS</td>
<td>Grechko et al. (2006)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta agilis</em></td>
<td>AS</td>
<td>Gullberg et al. (1998)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta agilis</em></td>
<td>AS</td>
<td>Gullberg et al. (1999)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta agilis</em></td>
<td>AS</td>
<td>Gullberg et al. (1999)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta agilis</em></td>
<td>AS</td>
<td>Kalyabina-Hauf et al. (2004b)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta agilis</em></td>
<td>AS</td>
<td>Madsen et al. (2000)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta schreiberi</em></td>
<td>AS</td>
<td>Godinho et al. (2001)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta schreiberi</em></td>
<td>AS</td>
<td>Godinho et al. (2003)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta schreiberi</em></td>
<td>AS</td>
<td>Godinho et al. (2006a)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta schreiberi</em></td>
<td>AS</td>
<td>Godinho et al. (2006b)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta schreiberi</em></td>
<td>AS</td>
<td>Paulo et al. (2001)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta schreiberi</em></td>
<td>AS</td>
<td>Paulo et al. (2002)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta schreiberi</em></td>
<td>AS</td>
<td>Godinho et al. (2008)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta spp.</em></td>
<td>AS</td>
<td>Godinho et al. (2005)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta spp.</em></td>
<td>AS</td>
<td>Moritz et al. (1992)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta trilineata / viridis</em></td>
<td>AS</td>
<td>Mayer &amp; Tiedemann (1985)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta viridis viridis</em></td>
<td>AS</td>
<td>Böhme et al. (2006)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta viridis viridis</em></td>
<td>AS</td>
<td>Böhme et al. (2007a)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Lacerta viridis/bilineata complex</em></td>
<td>AS</td>
<td>Böhme et al. (2007b)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Mesalina guttulata</em></td>
<td>AS</td>
<td>Kapli et al. (2008)</td>
<td>MT</td>
<td>E</td>
</tr>
</tbody>
</table>
**Scelarcis perspicillata**

**Psammodromus algirus**

**Podarcis**

**Podarcis vaucheri**

**Podarcis spp.**

**Podarcis sicula/muralis**

**Podarcis sicula**

**Podarcis hispanicus**

**Podarcis hispanicus/bocagei**

**Podarcis hispanicus complex**

**Podarcis lilfordi**

**Podarcis lilfordi**

**Podarcis melisellensis**

**Podarcis muralis**

**Podarcis pityusensis/lilfordi**

**Podarcis raffoni**

**Podarcis sicula**

**Podarcis sicula**

**Podarcis sicula /tiliguerta**

**Podarcis sicula / wagleriana**

**Podarcis sicula/muralis**

**Podarcis spp.**

**Podarcis spp.**

**Podarcis spp.**

**Podarcis spp.**

**Podarcis spp.**

**Podarcis spp.**

**Podarcis taurica**

**Podarcis tiliguerta**

**Podarcis tiliguerta**

**Podarcis vaucheri species complex**

**Podarcis vaucheri**

**Podarcis vaucheri**

**Podarcis wagleriana/sicula**

**Psammodromus algirus**

**Psammodromus algirus**

**Scelarcis perspicillata**

<table>
<thead>
<tr>
<th>Species</th>
<th>Code</th>
<th>Value</th>
<th>Location 1</th>
<th>Location 2</th>
<th>Location 3</th>
<th>Location 4</th>
<th>Location 5</th>
<th>Location 6</th>
<th>Location 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedio planus burchelli</td>
<td>AF</td>
<td>49</td>
<td>2.1 (1-10)</td>
<td>MT</td>
<td>E</td>
<td>Tolley et al. (2009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phoenicola ceria kulzeri complex</td>
<td>AS</td>
<td>9</td>
<td>2.4 (2-4)</td>
<td>CS/MT</td>
<td>E</td>
<td>in den Bosch et al. (2003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis atrata</td>
<td>EU</td>
<td>4</td>
<td>5.0 (4-6)</td>
<td>MT</td>
<td>CDE</td>
<td>Castilla et al. (1998)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis bocagei</td>
<td>EU</td>
<td>4</td>
<td>41.8 (36-55)</td>
<td>AZ</td>
<td>ABE</td>
<td>Sa-Sousa et al. (2000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis bocagei/carbonell/vaucher</td>
<td>EU/AF</td>
<td>57</td>
<td>4.3 (1-10)</td>
<td>MT</td>
<td>ACEGK</td>
<td>Pinho et al. (2007a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis bocagei/cardenoielli</td>
<td>EU</td>
<td>5</td>
<td>29.2 (17-59)</td>
<td>AZ/MT</td>
<td>ABCHI</td>
<td>Pinho et al. (2009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis erhardii</td>
<td>EU</td>
<td>40</td>
<td>1.1 (1-2)</td>
<td>MT</td>
<td>E</td>
<td>Poulakakis et al. (2003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis erhardii / taurica subgroup</td>
<td>EU</td>
<td>125</td>
<td>2.0 (2-4)</td>
<td>MT</td>
<td>E</td>
<td>Poulakakis et al. (2005c)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis erhardii</td>
<td>EU</td>
<td>20</td>
<td>19 (1-30)</td>
<td>MS/MT</td>
<td>ACEJ</td>
<td>Hurston et al. (2009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis filfolensis</td>
<td>EU</td>
<td>3</td>
<td>7.4 (1-19)</td>
<td>AZ/MT</td>
<td>BCE</td>
<td>Saclea et al. (2004)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis hispanicus</td>
<td>EU/AF</td>
<td>25</td>
<td>1.24 (1-3)</td>
<td>MT</td>
<td>E</td>
<td>Harris &amp; Sa-Sousa (2002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis hispanicus</td>
<td>EU/AF</td>
<td>28</td>
<td>1.25 (1-2)</td>
<td>MT</td>
<td>E</td>
<td>Harris et al. (2002a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis hispanicus/bocagei</td>
<td>EU/AF</td>
<td>19</td>
<td>1.5 (1-4)</td>
<td>MT</td>
<td>E</td>
<td>Harris &amp; Sa-Sousa (2001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis hispanicus complex</td>
<td>EU</td>
<td>83</td>
<td>1.0 (?)</td>
<td>MS/MT/NU</td>
<td>ABCEGKH</td>
<td>Renoult et al. (2009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis lilfordi</td>
<td>EU</td>
<td>14</td>
<td>20.5 (6-38)</td>
<td>AZ</td>
<td>BI</td>
<td>Ramon et al. (1991)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis lilfordi</td>
<td>EU</td>
<td>43</td>
<td>2.7 (1 - 5)</td>
<td>MT</td>
<td>ACEGK</td>
<td>Terrasa et al. (2009a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis lilfordi / hispanica</td>
<td>EU</td>
<td>5</td>
<td>18.0 (13-20)</td>
<td>AZ</td>
<td>ABD</td>
<td>Pettipierre et al. (1987)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis lilfordi</td>
<td>EU</td>
<td>43</td>
<td>2.7 (1.5)</td>
<td>MT</td>
<td>F</td>
<td>Terrasa et al. (2009b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis melisellensis</td>
<td>EU</td>
<td>52</td>
<td>1.4 (?)</td>
<td>MT</td>
<td>E</td>
<td>Podnar et al. (2004)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis muralis</td>
<td>EU</td>
<td>11</td>
<td>10.8 (4-17)</td>
<td>AZ</td>
<td>BC</td>
<td>Capula (1997)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis pityusensis/lilfordi</td>
<td>EU</td>
<td>12</td>
<td>3.9 (2-5)</td>
<td>MT</td>
<td>E</td>
<td>Terrasa et al. (2004b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis sicula</td>
<td>EU</td>
<td>14</td>
<td>6.0 (2-16)</td>
<td>AZ</td>
<td>D</td>
<td>Capula &amp; Ceccarelli (2003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis sicula</td>
<td>EU</td>
<td>86</td>
<td>1.1 (1-2)</td>
<td>MT</td>
<td>E</td>
<td>Podnar et al. (2005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis sicula</td>
<td>EU</td>
<td>4</td>
<td>23.5 (20-34)</td>
<td>MS/MT</td>
<td>AEGHJK</td>
<td>Fuglione et al. (2008)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis sicula /tiliguerta</td>
<td>EU</td>
<td>6</td>
<td>13.3 (7-19)</td>
<td>AZ</td>
<td>B</td>
<td>Capula (2002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis sicula / wagleriana</td>
<td>EU</td>
<td>10</td>
<td>30.2 (8-101)</td>
<td>AZ</td>
<td>B</td>
<td>Capula (1993)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis sicula/muralis</td>
<td>EU</td>
<td>38</td>
<td>1.0 (1-1)</td>
<td>MT/NU</td>
<td>AE</td>
<td>Podnar et al. (2007)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis spp.</td>
<td>EU</td>
<td>?</td>
<td>?</td>
<td>AZ/MT</td>
<td>?</td>
<td>Terrasa et al. (2004a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis spp.</td>
<td>EU/AF</td>
<td>11</td>
<td>21.7 (14-34)</td>
<td>AZ</td>
<td>BCE</td>
<td>Pinho et al. (2003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis spp.</td>
<td>EU/AF</td>
<td>28</td>
<td>1.0 (1-2)</td>
<td>MT</td>
<td>BE</td>
<td>Pinho et al. (2006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis spp.</td>
<td>EU/AF</td>
<td>32</td>
<td>17.8 (8-34)</td>
<td>AZ</td>
<td>CDEH</td>
<td>Pinho et al. (2007b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis spp.</td>
<td>EU/AF</td>
<td>50</td>
<td>2.8 (1-7)</td>
<td>MT/NU</td>
<td>ABDEGK</td>
<td>Pinho et al. (2008)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis spp.</td>
<td>EU</td>
<td>?</td>
<td>?</td>
<td>AZ/MT</td>
<td>?</td>
<td>Terrasa et al. (2004a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis taurica</td>
<td>EU</td>
<td>6</td>
<td>3.5 (1-6)</td>
<td>MT</td>
<td>E</td>
<td>Poulakakis et al. (2005b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis tiliguerta</td>
<td>EU</td>
<td>14</td>
<td>14.0 (4-19)</td>
<td>AZ</td>
<td>D</td>
<td>Capula (1996)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis tiliguerta</td>
<td>EU</td>
<td>13</td>
<td>1.4 (1-2)</td>
<td>MT</td>
<td>E</td>
<td>Harris et al. (2005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis vaucheri species complex</td>
<td>EU/AF</td>
<td>10</td>
<td>2.4 (1-5)</td>
<td>AZ/MT</td>
<td>BEI</td>
<td>Busack et al. (2005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis vaucheri</td>
<td>EU/AF</td>
<td>35</td>
<td>1.4 (1-2)</td>
<td>MT</td>
<td>E</td>
<td>Lima et al. (2009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Podarcis wagleriana/sicula</td>
<td>EU</td>
<td>14</td>
<td>8.6 (4-32)</td>
<td>AZ</td>
<td>BDEI</td>
<td>Capula (1994)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psammodromus algirus</td>
<td>EU/AF</td>
<td>8</td>
<td>2.7 (1-5)</td>
<td>AZ/MT</td>
<td>BEIJ</td>
<td>Busack &amp; Lawson (2006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psammodromus algirus</td>
<td>EU/AF</td>
<td>69</td>
<td>1.3 (1-5)</td>
<td>MT</td>
<td>E</td>
<td>Carranza et al. (2006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scelarcis perspicillata</td>
<td>AF</td>
<td>9</td>
<td>16.7 (1-29)</td>
<td>MT</td>
<td>E</td>
<td>Perera et al. (2007)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Method</td>
<td>Length</td>
<td>Technique</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------------</td>
<td>--------</td>
<td>---------------</td>
<td>----------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sceloporus arenicola</td>
<td>Alkali</td>
<td>5</td>
<td>6.2 (2-9)</td>
<td>A Jesus et al. (2005b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sceloporus jarrovi</td>
<td>Alkali</td>
<td>9</td>
<td>45.6 (6-65)</td>
<td>CD Brehm et al. (2001b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sceloporus jarrovi</td>
<td>Alkali</td>
<td>10</td>
<td>9.6 (9-10)</td>
<td>MT EK Brehm et al. (2003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timon lepidus/tangitanus</td>
<td>EU/AF</td>
<td>?</td>
<td>7</td>
<td>AB Busack &amp; Lawson (2008)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timon lepidus/pater</td>
<td>EU/AF</td>
<td>18</td>
<td>3.9 (2-9)</td>
<td>BI Mateo et al. (1996)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zootoca vivipara</td>
<td>EU</td>
<td>3</td>
<td>6.3 (3-11)</td>
<td>A Kupriyanova (1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zootoca vivipara</td>
<td>EU</td>
<td>14</td>
<td>8.9 (1-19)</td>
<td>AE Surget-Gropa et al. (2002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zootoca vivipara</td>
<td>EU</td>
<td>18</td>
<td>1.2 (1-4)</td>
<td>E Heulin et al. (1999)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zootoca vivipara</td>
<td>EU</td>
<td>1</td>
<td>542</td>
<td>L Hoffman &amp; Henle (2006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zootoca vivipara</td>
<td>EU</td>
<td>57</td>
<td>2.5 (1-7)</td>
<td>E Surget-Gropa et al. (2001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### PHRYNOSOMATIDAE

<table>
<thead>
<tr>
<th>Species</th>
<th>Method</th>
<th>Length</th>
<th>Technique</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Callisaurus draconoides</td>
<td>Alkali</td>
<td>41</td>
<td>1.0 (1-2)</td>
<td>E Lindell et al. (2005)</td>
</tr>
<tr>
<td>Callisaurus draconoides</td>
<td>Alkali</td>
<td>7</td>
<td>1.0 (1)</td>
<td>E Wilgenbusch &amp; de Queiroz (2000)</td>
</tr>
<tr>
<td>Holbrookia maculata</td>
<td>Alkali</td>
<td>8</td>
<td>8.1 (2-12)</td>
<td>CEG Rosembhlm (2006)</td>
</tr>
<tr>
<td>Holbrookia maculata</td>
<td>Alkali</td>
<td>2</td>
<td>19.0 (15-23)</td>
<td>ADK Rosenbhill et al. (2004)</td>
</tr>
<tr>
<td>Holbrookia spp</td>
<td>Alkali</td>
<td>10</td>
<td>1.0 (1)</td>
<td>E Wilgenbusch &amp; de Queiroz (2000)</td>
</tr>
<tr>
<td>Petroscus thalassinus</td>
<td>Alkali</td>
<td>6</td>
<td>17.8 (4-27)</td>
<td>D Aguilars et al. (1988)</td>
</tr>
<tr>
<td>Phrynosoma coronatum</td>
<td>Alkali</td>
<td>96</td>
<td>1.1 (1-2)</td>
<td>BEG Leaché et al. (2009)</td>
</tr>
<tr>
<td>Phrynosoma douglasi</td>
<td>Alkali</td>
<td>38</td>
<td>1.7 (1-5)</td>
<td>E Zamudio et al. (1997)</td>
</tr>
<tr>
<td>Phrynosoma mcallii-platyrhinos complex</td>
<td>Alkali</td>
<td>32</td>
<td>3.2 (1-14)</td>
<td>MT EF Mulcahy et al. (2006)</td>
</tr>
</tbody>
</table>

### SCELOPORUS ARENICOLAS

<table>
<thead>
<tr>
<th>Species</th>
<th>Method</th>
<th>Length</th>
<th>Technique</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. arenicola</td>
<td>Alkali</td>
<td>25/7</td>
<td>2.2 (?)/ 32.9 (?)</td>
<td>ACDEGHJ AK Chan et al. (2009)</td>
</tr>
<tr>
<td>S. cowlesi/tristichus</td>
<td>Alkali</td>
<td>93</td>
<td>6.9 (1-80)</td>
<td>CS/MT AE Leache &amp; Cole (2007)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>93</td>
<td>14.9 (1-73)</td>
<td>CS/MT AE Arevalo et al. (1991)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>6</td>
<td>14.5 (3-26)</td>
<td>CS/AZ/RFLP BCE Arevalo et al. (1993)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>9</td>
<td>1.0 (1)</td>
<td>MT E Arevalo et al. (1994)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>28</td>
<td>15.5 (10-131)</td>
<td>CS A Dosselman et al. (1998)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>2</td>
<td>18.0 (8-28)</td>
<td>CS C Goyenechea et al. (1996)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>28</td>
<td>17.9 (?)</td>
<td>CS/AFLP/RFLP M Marshall &amp; Sites (2001)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>42</td>
<td>14.2 (1-68)</td>
<td>CS ABCI Porter &amp; Sites (1986)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>13</td>
<td>24.2 (7-35)</td>
<td>AZ B Sites &amp; Boyce (1985)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>19</td>
<td>3.8 (2-4)</td>
<td>CS/AZ/RFLP E Sites &amp; Davies (1989)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>13</td>
<td>24.2 (7-35)</td>
<td>AZ ABD Sites &amp; Greenbaum (1983)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>89</td>
<td>2.2 (1-10)</td>
<td>CS A Sites (1983)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>8</td>
<td>10.2 (1-32)</td>
<td>CS/AZ/RFLP BCDI Sites et al. (1993)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>33</td>
<td>16.8 (3-37)</td>
<td>CS ADM Sites et al. (1995)</td>
</tr>
<tr>
<td>S. grammicus</td>
<td>Alkali</td>
<td>23</td>
<td>14.9 (1-36)</td>
<td>CS/MT ACM Sites et al. (1996)</td>
</tr>
<tr>
<td>S. grammicus complex</td>
<td>Alkali</td>
<td>1</td>
<td>22.0 (10-35)</td>
<td>AZ ABD Sites et al. (1988)</td>
</tr>
<tr>
<td>S. grammicus/grammaticus</td>
<td>Alkali</td>
<td>1</td>
<td>22.0 (10-35)</td>
<td>CS/AZ ABCDI Thompson &amp; Sites (1986)</td>
</tr>
<tr>
<td>S. jarrovi</td>
<td>Alkali</td>
<td>30</td>
<td>1.5 (1-2)</td>
<td>MT EF Wiens et al. (1999)</td>
</tr>
<tr>
<td>S. jarrovi</td>
<td>Alkali</td>
<td>30</td>
<td>1.5 (1-2)</td>
<td>MT E Wiens et al. (1999)</td>
</tr>
<tr>
<td>S. jarrovi</td>
<td>Alkali</td>
<td>2</td>
<td>21 (?)</td>
<td>MS A Zamudio &amp; Wieczorek (2000)</td>
</tr>
<tr>
<td>Species</td>
<td>State</td>
<td>MA/MS</td>
<td>Method</td>
<td>Species</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Anolis allisoni/porcatus</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis allisoni/porcatus</td>
</tr>
<tr>
<td>Anolis bimaculatus group</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis bimaculatus group</td>
</tr>
<tr>
<td>Anolis carolinensis subgroup</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis carolinensis subgroup</td>
</tr>
<tr>
<td>Anolis chlorocyamus</td>
<td>NA/WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis chlorocyamus</td>
</tr>
<tr>
<td>Anolis cooki</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis cooki</td>
</tr>
<tr>
<td>Anolis cristatellus</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis cristatellus</td>
</tr>
<tr>
<td>Anolis cybotes</td>
<td>NA/WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis cybotes</td>
</tr>
<tr>
<td>Anolis cybotes/whitemani</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis cybotes/whitemani</td>
</tr>
<tr>
<td>Anolis desechensis/menensis/cooki/cristatellus</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis desechensis/menensis/cooki/cristatellus</td>
</tr>
<tr>
<td>Anolis distinctus</td>
<td>NA/WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis distinctus</td>
</tr>
<tr>
<td>Anolis denticristic</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis denticristic</td>
</tr>
<tr>
<td>Anolis equestris</td>
<td>NA/WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis equestris</td>
</tr>
<tr>
<td>Anolis extremus</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis extremus</td>
</tr>
<tr>
<td>Anolis garmani</td>
<td>NA/WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis garmani</td>
</tr>
<tr>
<td>Anolis grahami series</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis grahami series</td>
</tr>
<tr>
<td>Anolis marmoratus</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis marmoratus</td>
</tr>
<tr>
<td>Anolis marmoratus complex</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis marmoratus complex</td>
</tr>
<tr>
<td>Anolis oculatus</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis oculatus</td>
</tr>
<tr>
<td>Anolis oculatus</td>
<td>WI</td>
<td>35.0</td>
<td>RFLP</td>
<td>Anolis oculatus</td>
</tr>
</tbody>
</table>

**POLYCHORIDAE**

- Anolis allisoni/porcatus
- Anolis bimaculatus group
- Anolis carolinensis subgroup
- Anolis chlorocyamus
- Anolis cooki
- Anolis cristatellus
- Anolis cybotes
- Anolis cybotes/whitemani
- Anolis desechensis/menensis/cooki/cristatellus
- Anolis distinctus
- Anolis denticristic
- Anolis equestris
- Anolis extremus
- Anolis garmani
- Anolis grahami series
- Anolis marmoratus
- Anolis marmoratus complex
- Anolis oculatus
- Anolis oculatus

**References**

- Stenson et al. (2002)
- Malhotra & Thorpe (2000)
- Malhotra & Thorpe (1994)
- Schneider et al. (2001)
- Schneider (1996)
- Kolbe et al. (2007b)
- Thorpe et al. (2005)
- Miyamoto et al. (1986)
- Glor et al. (2003)
- Kolbe et al. (2007b)
- Eales et al. (2008)
- Rodríguez & Fuentes (2007)
- Murphy et al. (2005)
- Upton & Murphy (1997)
- Mahoney et al. (2003)
- E Upton & Murphy (1997)

**Methods**

- RFLP
- MS
- MT
- NU
Anolis oculatus/roquet  WI  10  13.5 (? )  AFLP  I  Ogden & Thorpe (2002a)
Anolis porcatus  NA/WI  33  3.8 (1-14)  MT  CE  Kolbe et al. (2007b)
Anolis roquet  WI  24  ? (15-20)  MS  CD  Ogden & Thorpe (2002b)
Anolis roquet  WI  63  1.2 (1-6)  MT  EJ  Thorpe & Stenson (2003)
Anolis roquet  WI  17  38.4 (28-48)  MS  ABH  Johansson et al. (2008a)
Anolis sagrei  WI  16  48 (48)  MT  M  Johansson et al. (2008b)
Anolis sagrei  WI  5  51.0 (45-52)  MS  ACDHJ  Calsbeek et al. (2007)
Anolis sagrei  NA/WI  130  4.7 (?)  MT  CE  Kolbe et al. (2004)
Anolis sagrei  NA  17  16.1 (?-20)  MT  E  Kolbe et al. (2007a)
Anolis sagrei  NA/WI  18  16.1 (?-20)  MT  CE  Kolbe et al. (2007b)
Anolis sagrei  NA/WI  18  14.0 (?-20)  MS/MT  ABJ  Kolbe et al. (2008)
Anolis sagrei  WI  104  2.8 (1-36)  MT  AEK  Eales & Thorpe (2009)
Anolis trinitatis  WI  11  2.0 (2)  MT  E  Thorpe (2002)

SCINCIDAE

Ablepharus budaki / kitaibelii  EU/AS  37  2.0 (2)  MT  E  Poulakakis et al. (2005a)
Acontias meleagris meleagris  AF  15  4.7 (1-10)  MT  E  Daniels et al. (2005)
Acontias meleagris  AF  24  5.0 (1-30)  MT/NU  CEK  Daniels et al. (2009)
Carlia rhomboidalis  AU  30  ?  MS/MT  BDE  Smith et al. (2001b)
Carlia rhomboidalis  AU  8  6.3 (3-13)  MT  ACE  Stuart-Fox et al. (2001)
Carlia rubrigularis  AU  9  18.0 (?-20)  MS/RFLP  CDHM  Phillips et al. (2004)
Carlia rubrigularis / rhomboidalis  AU  7  4.3 (2-7)  MT/NU  ?  Dolman & Moritz (2006)

Carlia fasca complex  AU  29/12  4  CS/AZ/MT  AE  Donnellan et al. (2009)
Chalcides sexlineatus  AI  24  4.0 (3-7)  MT  DEK  Pestano & Brown (1999)
Chalcides spp.  AI  21  2.2 (1-4)  MT  E  Brown & Pestano (1998)
Chalcides spp.  EU/AF  121  1.2 (1-4)  MT  E  Carranza et al. (2008)
Chalcides viridanus  AI  17  5.0 (1-8)  MT  CDEJK  Brown et al. (2000)
Cyclodina aenea species complex  PI  16  1.1 (1-2)  MT  E  Chapple et al. (2008a)
Cyclodina spp.  PI  72  1.3 (?)  MT  E  Chapple et al. (2008b)

Cyclodomorphus praealtus  AU  4  28 (5 - 41)  MS/MT  ACDEGH  K  Koumoundouros et al. (2009)
Cryptoblepharbus boutonii  AF  24  2.0 (1-5)  MT  E  Roha et al. (2006)
Cryptoblepharbus nigropunctatus  AS  11  32.7 (1-202)  MT  ABCE  Hayashi et al. (2009)
Ctenotus brooksi complex  AU  37  1.5 (1-2)  AZ  BCI  Hutchinson et al. (2006)
Ctenotus spp.  AU  25  1.96 (1-5)  AZ  BCEI  Hutchinson & Donnellan (1999)
Ctenotus robustus  AU  8  2.0 (1-6)  MT  AE  Colgan et al. (2009)
Ctenotus taeniolatus  AU  7  3.4 (1-8)  MT  AE  Colgan et al. (2009)
Ctenotus leonhardii/quattuordecimlineatus  AU  31  1.4 (?)  MT/NU  EG  Rabosky et al. (2009)
Egernia cunninghami  AU  2  153.0 (141-165)  MS  C  Stow & Briscoe (2005)
Egernia cunninghami  AU  1  1.2 (1-3)  MS  AHJL  Stow et al. (2001)
E. guthega/margaretae/modesta/montana  AU  85  1.2 (1-3)  MT  E  Chapple et al. (2005)
Egernia inornata/multiscutata/strigata  AU  93  1.0 (1-2)  MT  E  Chapple et al. (2004)

Anolis oculatus/roquet  WI  10  13.5 (?)  AFLP  I  Ogden & Thorpe (2002a)
| Species                                      | AU | UK | Canada | MT | US | Canada | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | Mexico | M
<table>
<thead>
<tr>
<th>Species</th>
<th>Continent</th>
<th>Species Code</th>
<th>Geographic Range</th>
<th>Genotype/Method</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Sphenomorphus jobiensis</em></td>
<td>AS</td>
<td>11</td>
<td>2.6 (1-7)</td>
<td>AZ</td>
<td>E</td>
</tr>
<tr>
<td><em>Tiliqua rugosa</em></td>
<td>AU</td>
<td>1</td>
<td>60</td>
<td>MS</td>
<td>L</td>
</tr>
<tr>
<td><em>Tiliqua rugosa</em></td>
<td>AU</td>
<td>1</td>
<td>86</td>
<td>MS</td>
<td>L</td>
</tr>
<tr>
<td><em>Tiliqua adelaidensis</em></td>
<td>AU</td>
<td>6</td>
<td>38.2 (5-116)</td>
<td>MS</td>
<td>ABCGHIJ</td>
</tr>
<tr>
<td><em>Trachydosaurus rugosus</em></td>
<td>AU</td>
<td>10</td>
<td>27.2 (9-53)</td>
<td>AZ</td>
<td>ABCDE</td>
</tr>
<tr>
<td><em>Trachylepis elegans</em></td>
<td>AF</td>
<td>12</td>
<td>1.4 (1-4)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>Trachylepis gravenhorstii</em></td>
<td>AF</td>
<td>24</td>
<td>1.2 (1-4)</td>
<td>MT</td>
<td>E</td>
</tr>
<tr>
<td><em>TEIIDAE</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Ameiva ameiva</em></td>
<td>SA</td>
<td>1</td>
<td>25</td>
<td>AZ</td>
<td>C</td>
</tr>
<tr>
<td><em>Ameiva chrysolaema</em></td>
<td>WI</td>
<td>14</td>
<td>3.9 (1-7)</td>
<td>MT</td>
<td>EFK</td>
</tr>
<tr>
<td><em>Ameiva chrysolaema</em></td>
<td>WI</td>
<td>9</td>
<td>13 (5-26)</td>
<td>MT/NU</td>
<td>M</td>
</tr>
<tr>
<td><em>Ameiva chrysolaema</em></td>
<td>WI</td>
<td>37</td>
<td>4.6 (1-10)</td>
<td>MT/NU</td>
<td>CEFKG</td>
</tr>
<tr>
<td><em>Ameiva exsul</em></td>
<td>WI</td>
<td>2</td>
<td>6</td>
<td>RAPD</td>
<td>AL</td>
</tr>
<tr>
<td><em>Aspidoscelis inornata</em></td>
<td>NA</td>
<td>1</td>
<td>25</td>
<td>AZ/mtDNA (ASO dot-blot)</td>
<td>CS/AZ</td>
</tr>
<tr>
<td><em>Aspidoscelis inornata</em></td>
<td>NA</td>
<td>2</td>
<td>19.5 (15-24)</td>
<td>MT/NU</td>
<td>ADK</td>
</tr>
<tr>
<td><em>Aspidoscelis neomexicana</em></td>
<td>NA</td>
<td>10</td>
<td>10.0 (9-11)</td>
<td>AZ</td>
<td>A</td>
</tr>
<tr>
<td><em>Aspidoscelis sexlineata</em></td>
<td>NA</td>
<td>7</td>
<td>7.0 (2-22)</td>
<td>AZ</td>
<td>A</td>
</tr>
<tr>
<td><em>C. gramivagus/lemniscatus/ocellifer/parecis</em></td>
<td>SA</td>
<td>5</td>
<td>31.2 (12-44)</td>
<td>AZ</td>
<td>C</td>
</tr>
<tr>
<td><em>A. laredoensis/ gularis/ sexlineatus</em></td>
<td>NA</td>
<td>11</td>
<td>2.6 (1-5)</td>
<td>RFLP</td>
<td>A</td>
</tr>
<tr>
<td><em>Cnemidophorus lemniscatus</em></td>
<td>SA</td>
<td>8</td>
<td>8.0 (3-20)</td>
<td>RFLP</td>
<td>E</td>
</tr>
<tr>
<td><em>Cnemidophorus longicaudatus</em></td>
<td>NA</td>
<td>36</td>
<td>2.0 (1-5)</td>
<td>MT</td>
<td>ACEFK</td>
</tr>
<tr>
<td><em>A. neomexicanus/ inornatus/tigris</em></td>
<td>NA</td>
<td>7</td>
<td>19.3 (1-35)</td>
<td>AZ</td>
<td>A</td>
</tr>
<tr>
<td><em>Aspidoscelis sexlineatus group</em></td>
<td>NA</td>
<td>19</td>
<td>4.5 (1-33)</td>
<td>RFLP</td>
<td>E</td>
</tr>
<tr>
<td><em>Aspidoscelis tesselata/tigris</em></td>
<td>NA</td>
<td>?</td>
<td>34</td>
<td>CS/AZ</td>
<td>A</td>
</tr>
<tr>
<td><em>Aspidoscelis tigris</em></td>
<td>NA</td>
<td>10</td>
<td>9.9 (9-10)</td>
<td>AZ</td>
<td>D</td>
</tr>
<tr>
<td><em>Aspidoscelis tigris</em></td>
<td>NA</td>
<td>48</td>
<td>12.7 (1-88)</td>
<td>AZ/MT</td>
<td>D</td>
</tr>
<tr>
<td><em>Cnemidophorus vanzoi</em></td>
<td>WI</td>
<td>2</td>
<td>? (14-42)</td>
<td>MS/MT</td>
<td>ACDHI</td>
</tr>
<tr>
<td><em>Aspidoscelis velox/sexanguis</em></td>
<td>NA</td>
<td>10</td>
<td>2.4 (1-5)</td>
<td>RFLP</td>
<td>AE</td>
</tr>
<tr>
<td><em>A. neomexicanus/tesselatus complex</em></td>
<td>NA</td>
<td>24</td>
<td>4.0 (1-17)</td>
<td>RFLP</td>
<td>E</td>
</tr>
<tr>
<td><em>Kentropyx altamazonica</em></td>
<td>SA</td>
<td>1</td>
<td>25</td>
<td>AZ</td>
<td>C</td>
</tr>
<tr>
<td><em>Tupinambis merianae</em></td>
<td>SA</td>
<td>1</td>
<td>13</td>
<td>AZ</td>
<td>C</td>
</tr>
</tbody>
</table>

**TROPIDURIDAE**

*Eurolophosaurus*

*divaricatus/amathites/manucae* | SA | 10 | 2.2 (?) | MT | E | Passoni et al. (2008) |

*Liolaemus bibronii / gracilis* | SA | 41 | 3.4 (1-6) | MT | CEFGK | Morando et al. (2006) |

*Liolaemus boulangerti group* | SA | 100 | 2.8 (1-5) | MT | ACEFK | Avila et al. (2006) |
<table>
<thead>
<tr>
<th>Species</th>
<th>Location</th>
<th>mtDNA</th>
<th>Mitochondrial</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liolaemus darwinii complex</td>
<td>SA</td>
<td>45</td>
<td>2.6 (1-3)</td>
<td>MT</td>
</tr>
<tr>
<td>Liolaemus elongatus-kriegi complex</td>
<td>SA</td>
<td>49</td>
<td>4.0 (1-11)</td>
<td>MT</td>
</tr>
<tr>
<td>Liolaemus koslowskyi</td>
<td>SA</td>
<td>43</td>
<td>3.4 (1-8)</td>
<td>MT</td>
</tr>
<tr>
<td>Liolaemus lenniscatus</td>
<td>SA</td>
<td>14</td>
<td>3.7 (1-10)</td>
<td>MT</td>
</tr>
<tr>
<td>Liolaemus monticola</td>
<td>SA</td>
<td>5</td>
<td>27.8 (2-64)</td>
<td>CS</td>
</tr>
<tr>
<td>Liolaemus monticola</td>
<td>SA</td>
<td>17</td>
<td>12.3 (2-32)</td>
<td>CS</td>
</tr>
<tr>
<td>Liolaemus monticola</td>
<td>SA</td>
<td>21</td>
<td>17.1 (2-32)</td>
<td>CS</td>
</tr>
<tr>
<td>Liolaemus monticola</td>
<td>SA</td>
<td>28</td>
<td>1.5 (1-5)</td>
<td>MT</td>
</tr>
<tr>
<td>Liolaemus monticola</td>
<td>SA</td>
<td>3</td>
<td>46.7 (26-52)</td>
<td>AZ/MT</td>
</tr>
<tr>
<td>Liolaemus pictus</td>
<td>SA</td>
<td>16</td>
<td>18.9 (2-7)</td>
<td>CS/AZ</td>
</tr>
<tr>
<td>Liolaemus tenius</td>
<td>SA</td>
<td>15</td>
<td>5.5 (1-11)</td>
<td>MT</td>
</tr>
<tr>
<td>Liolaemus tenius</td>
<td>SA</td>
<td>12</td>
<td>19.4 (8-50)</td>
<td>AZ</td>
</tr>
<tr>
<td>Liolaemus tenius</td>
<td>SA</td>
<td>41</td>
<td>3.5 (1-8)</td>
<td>MT</td>
</tr>
<tr>
<td>Microlophus albemarlensis</td>
<td>PI</td>
<td>1</td>
<td>17.4 (4-38)</td>
<td>MS</td>
</tr>
<tr>
<td>Microlophus albomarlenensis</td>
<td>PI</td>
<td>10</td>
<td>1.3 (1-2)</td>
<td>MT</td>
</tr>
<tr>
<td>Tropidurus nanucae group</td>
<td>SA</td>
<td>6</td>
<td>7.2 (4-16)</td>
<td>RFLP</td>
</tr>
<tr>
<td>Varanus komodoensis</td>
<td>AS</td>
<td>6</td>
<td>14.6 (6-27)</td>
<td>MS</td>
</tr>
<tr>
<td>Varanus komodoensis</td>
<td>AS</td>
<td>7</td>
<td>15.9 (8-27)</td>
<td>MS</td>
</tr>
<tr>
<td>Varanus rosenbergi</td>
<td>AU</td>
<td>19</td>
<td>1.6 (1-4)</td>
<td>MT</td>
</tr>
<tr>
<td>XANTUSIIDAE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xantusia henshawi</td>
<td>NA</td>
<td>13</td>
<td>1.4 (1-4)</td>
<td>MT</td>
</tr>
<tr>
<td>Xantusia vigilis</td>
<td>NA</td>
<td>87</td>
<td>1.4 (1-5)</td>
<td>MT</td>
</tr>
<tr>
<td>Xantusia vigilis species complex</td>
<td>NA</td>
<td>156</td>
<td>3.2 (1-18)</td>
<td>MT/NU</td>
</tr>
<tr>
<td>Xantusia vigilis/henshawi/riversiana</td>
<td>NA</td>
<td>11</td>
<td>2.6 (1-4)</td>
<td>AZ</td>
</tr>
</tbody>
</table>


Grechko VV, Kataev MV, Mel'nikova MN, Darevskii IS (1993) The DNA relationships of the parthenogenetic forms of the Lacerta lizards species and supposed parental bisexual species as it may be revealed by polymerase chain reaction with arbitrary single primer (AP-RAPD). Molekulyarnaya Biologiya, 27, 6, 1415-1424.

eastern Australia.


Kupriyanova L (1997) Is the Baltic Sea basin a zone of secondary contact between different chromosomal forms of Zootoca vivipara? Memoranda Societatis pro Fauna et Flora Fennica, 73, 3-4, 115-117 ER.


Ablepharus kitaibelii (Sauria : Scincidae).


Figure IIa. Percentage of phylogeographic studies in lizard families.
Figure IIb. Number of phylogeographic studies in lizard genera.
Figure IIc. Number of phylogeographic studies in lizards from several geographic regions.
Figure IId. Comparison of the percentage of studies published in lizard families (open bars) with the taxonomic richness (solid bars). (A) Species richness and (B) generic richness.
Figure IIe. Frequency of phylogeographic methods based on the number of study cases. Categories are non-exclusive because many studies used more than one kind of phylogeographic method.
Appendix III

A model for socially mediated speciation (Hochberg et al. 2003) hypothesizes that reproductive isolation can evolve within a species with a locus for distinct signals, given two other loci: one for mate choice and one mediating social interactions like altruistic donation or social competition. For example, the color trimorphic *Uta stansburiana* exhibits all three loci required for socially mediated speciation, including (1) signal, (2) mate choice, and (3) donation loci (Sinervo et al. 2006; Bleay et al. 2007). Genes that control such social interactions are referred to as “greenbeards”. The side-blotched lizard system represents the first example of greenbeard dynamics in vertebrates, but insects (Keller et al. 1997; Keller & Ross 1998), protists (Queller et al. 2003), and prokaryotes like *E. coli* exhibit similar greenbeard and/or RPS dynamics (Kerr et al. 2002, 2004). Obvious parallels among the RPS system in *Uta stansburiana*, the RPS in the lizard *Lacerta vivipara* of Europe (Sinervo et al. 2007), and those of protists (Queller et al. 2003), and prokaryotes (Kerr et al. 2002), suggest that many more examples will be uncovered (reviewed in Sinervo et al. 2007, 2008).

Based on gene mapping studies in the field pedigree and mate choice studies, the side-blotched lizard exhibits all 4 required loci (loci i-iii: color signals that elicit recognition, and social acts of altruism or competition [Sinervo & Clobert 2003, Sinervo et al. 2006] and locus iv: females exhibit self-color preferences for these signals [O for O and Y for Y: Bleay & Sinervo 2007; B for B Sinervo et al. 2006]. The number of loci that are required for the functioning of the social systems (minimum of 4 loci, above), and their distribution across the genome on separate linkage groups will generate unfitness between hybrids produced from trimorphic lineages crossed to di- or mono-morphic lineages (different socially “coadapted genes” in each lineage), and ensuing evolution of a RIM.

To generate hybrid unfitness, socially mediated speciation (Hochberg et al. 2003) requires separate and unlinked loci for: i) a signal locus (e.g., color), ii) recognition loci for (self) signals, iii) signal recognition elicits social acts of altruism (to self) or
competition (to non-self), and iv) females exhibit (loci) preferences for self (color signals). Even though these loci are unlinked, in spatially prescribed neighborhoods, the model reproducibly generates species with differently colored signals eliciting social acts in one species versus the other. Reduced hybrid fitness at junction between evolving species leads to the selection of mating preferences, which then spread to the core areas of the respective species (Hochberg et al. 2003). The number of loci that are required for the functioning of the social systems (minimum of 4 loci, above), and their distribution across the genome on separate linkage groups will generate unfitness between hybrids produced from trimorphic lineages crossed to di- or mono-morphic lineages (different socially “coadapted genes” in each lineage), and ensuing evolution of a RIM.

The process of coadapted mate preferences (e.g., preference for the same color morph: Bleay & Sinervo (2007) is amplified by multi-trait preferences involving color and many other loci (Lancaster et al. 2009). Therefore, the levels of linkage disequilibrium due to correlation selection (Sinervo & Svensson 2002), mate preference (Bleay & Sinervo 2007; Lancaster et al. 2009) within an RPS mating system are extremely high (Sinervo et al. 2006), rivaling the linkage disequilibrium found at a contact zone between species. Fixation and loss of color morphs is expected to result in extremely rapid genetic change (Corl et al. 2010a, b) because many other alleles at loci besides the master morph locus will be lost as the alternative optima for the now missing color alleles disappear. Thus, the genes present in a reduced mating system with fewer morphs will be dramatically different than the ancestral RPS system, perhaps generating Dobzhansky-Muller incompatibilities. The formation of Dobzhansky-Muller incompatibilities increases with the number of genes that diverge and with the number of associations among genes (Coyne & Orr 2004), so the loss of a morph could potentially generate a large number of incompatibilities (given the number of autosomal strategic loci with which the OBY locus interacts, Sinervo et al. 2006).

References


ACCURACY AND PRECISION OF SPECIES TREES: AN EMPIRICAL EVALUATION OF PERFORMANCE IN LIZARDS OF THE LIOLAEMUS DARWINII GROUP (SQUAMATA, TROPIDURIDAE) UNDER VARYING SUB-SAMPLING DESIGNS

Arley Camargo¹*, Mariana Morando², Luciano J. Avila² & Jack W. Sites, Jr.¹

¹Department of Biology and Bean Life Science Museum, Brigham Young University, Provo, UT 84602, USA
²CONICET-CENPAT, Boulevard Almirante Brown 2825, U9120ACF, Puerto Madryn, Chubut, Argentina

Key words: lizards, phylogeography, adaptation, divergence, speciation, molecular markers

Corresponding author: Arley Camargo, Department of Biology, Brigham Young University, Provo, Utah, USA 84602; phone: 801/422-2203; fax: 801/422-0090; email: arley.camargo@gmail.com
Abstract.—Molecular phylogenetics has entered a new era in which species trees are estimated from a collection of gene trees using methods that accommodate their heterogeneity and discordance with the species tree due to incomplete lineage sorting. Empirical evaluation of species trees are necessary to assess the performance of these methods with real data, which consists of gene genealogies likely shaped by different historical and demographic processes. We analyzed 20 loci for 16 species of the South American lizards of the *Liolaemus darwinii* species group and reconstructed a species tree with *BEAST*, then compared the performance of this method under different sampling strategies of loci, individuals, sequence lengths, and species. We found an asymptotic increase in the accuracy and precision of species trees with the number of loci and improvements in accuracy when using >1 individual per species and >1/4 of the original datasets for any number of loci. In addition, locus 'informativeness' was an important factor when using a few loci, but it became increasingly irrelevant with additional loci. Results show that more loci should be used when analyzing a larger number of species, suggesting that the addition of unsampled species to our phylogeny will require additional sequencing effort. Our results provide some guidance for empirical phylogeneticists constrained by time and/or technical difficulties, by showing that there is an optimal range of sampling effort of loci, individuals, and sequence length for a given speciation history and data type. Future studies should be directed towards further assessment of other factors that can impact performance of species trees, including gene flow, data 'informativeness', tree shape, missing data, and uncertain species boundaries.
INTRODUCTION

Molecular phylogenetics has entered a new era in which species trees are estimated from a collection of gene trees by accommodating their heterogeneity and discordance with the species tree due to incomplete lineage sorting (Edwards 2009; Knowles and Kubatko 2010). Over two decades ago it was realized that gene trees could be highly heterogeneous and be discordant with the species tree due to a variety of processes including estimation error, incomplete lineage sorting, horizontal gene transfer, and gene duplication/loss (Pamilo and Nei 1988; Avise 1989; Maddison 1997). Until recently, standard approaches assumed that all gene trees matched the underlying species tree and relied on sequence concatenation, which was shown to be more accurate than consensus methods (Gadagkar et al. 2005). However, simulation studies have found that concatenation is inconsistent in an “anomaly zone” (Kubatko and Degnan 2007) in which the most frequent gene trees do not match the species tree (anomalous gene trees, AGT; Degnan and Rosenberg 2006). As an alternative to concatenation, a gene tree parsimony approach based on reconciliation of the gene trees with the species tree was proposed over a decade ago (Page 1998; Slowinski and Page 1999), but concatenation remained the preferred choice in practice. Subsequently, Maddison (1997) introduced the idea of a summary-statistic approach based on minimizing deep coalescences across multiple gene trees, and more recently, a variety of other approaches have been proposed that use summary statistics (STAR and STEAC, Liu et al. 2009a; GLASS tree, Mossel and Roch 2010), consensus/supertree methods (Degnan et al. 2009), and Bayesian concordance factors (Ané et al. 2007).
A new generation of methods has explicitly incorporated gene tree heterogeneity due to incomplete lineage sorting into species trees estimation, based on the multispecies coalescent model (Rannala and Yang 2003; Degnan and Rosenberg 2009). These novel, model-based frameworks have led to the development of maximum likelihood and Bayesian inference approaches (Liu et al. 2009a). The ML approach is implemented in the program STEM (Kubatko et al. 2009), which combines user-provided constant population size, estimated gene trees, and relative rates among loci to obtain the ML species tree with branch lengths that accommodate rate variation and ploidy level across loci (Kubatko et al. 2009). A Bayesian approach has been implemented in two programs, including BEST (Liu 2008) and BEAST (*BEAST, Drummond and Rambaut 2007; Heled and Drummond 2010). BEST applies a hierarchical design to estimate the joint posterior distribution of species trees and gene trees, conditional on the observed sequence data with the restriction that species divergence times cannot predate the coalescence times of alleles. In addition, BEST estimates gene trees without assuming a molecular clock and then ultrametricizes branch lengths (Castillo-Ramírez et al. 2010).

On the other hand, *BEAST relaxes the molecular clock for estimating gene trees and accommodates for changing population sizes across the species tree (Heled and Drummond 2010). These methods further assume that loci are unlinked (free recombination between loci), with no intra-locus recombination, and that gene tree heterogeneity is due to incomplete lineage sorting only. Newer approaches that incorporate hybridization to the coalescent-based species trees are in active phase of development and first empirical results are promising (Kubatko 2009; Kubatko and Meng...
while a recently proposed summary-statistic method appears to be robust to horizontal gene transfer (Liu et al. 2009a).

The performance of multilocus species tree methods is beginning to be investigated with simulations to assess the impact of sampling strategies (McCormack et al. 2009; Castillo-Ramírez et al. 2010) and to disentangle the relative influence of coalescent vs. mutational variance (Huang et al. 2010). The performance of some of these new methods (STEM and BEST) has also been evaluated in the context of species delimitation (Carstens and Dewey 2010). There are at least three dimensions in the size of datasets that can be subsampled to evaluate their impact on performance: number of loci, number of individuals, and sequence length (Brito and Edwards 2009). Another dimension of sampling is the variation in locus 'informativeness' or phylogenetic signal (Knowles 2009) which, although a dominant factor in empirical studies, has been rarely explored in simulation studies because all loci are virtually identical (e.g. same substitution model, same sequence length, etc). In addition, taxon sampling in the context of species trees has not been explored previously, even though this factor has received substantial attention in traditional phylogenetic inference (Pollock et al. 2002; Zwickl and Hillis 2002; Hillis et al. 2003; Heath et al. 2008).

Recently, Liu et al. (2009a) suggested using accuracy and precision estimators to compare the performance of species trees methods under varying conditions and sampling effort. In an empirical context, performance can be seen as the ability of the method to estimate the species tree with limited information, and those methods that resolve the species tree accurately with fewer data are judged to perform better. Accuracy can be assessed as the distance between the best estimate of the species tree with all the
available data, and the best estimate obtained with a subsample of the data. In addition, Bayesian methods allow the evaluation of precision estimates via the variation in tree distances among the set of trees in the posterior distribution. One such measure that can be applied to compare tree similarity is the $K$ tree score (Soria-Carrasco et al. 2007), a measure of the difference in both topology and branch lengths between two phylogenetic trees.

Herein, we produced multi-locus sequence data for the *Liolaemus darwinii* group of South American lizards (Squamata, Tropiduridae) to reconstruct a species tree for the group, and to compare the performance of one species tree method under different sampling strategies. Lizards of the *L. darwinii* group inhabit the arid lands of the Monte Desert region of central and northwestern Argentina (Fig. 1). Several morphological characters support the monophyly of the group, and it currently contains 18 recognized species, many of which have been described in the last two decades after the most recent taxonomic review (Etheridge 1993). A recent combined molecular/morphological study recovered this group as a strongly supported clade nested within the more inclusive *L. boulengeri* clade (Abdala 2007). In this study, we sampled multiple loci from most described species in the *L. darwinii* group to re-assess relationships within this clade, and to present a working hypothesis that will serve as a framework for ongoing phylogeographic, species delimitation, and speciation studies of these lizards.

The conceptual focus of this study is to empirically evaluate the performance of a Bayesian method of species tree reconstruction based on our use of multiple loci under varying sampling designs. Specifically, we evaluated the convergence of the alternative models with increasing numbers of species, numbers of loci, numbers of intra-locus base
pairs, and the effect of intraspecific sampling. We used the tree estimated from all available data as our best estimate of the true species tree to quantify accuracy and precision, assuming that the species tree method is statistically consistent, meaning that the estimate is more accurate when more data are added to the analysis. Maximum accuracy occurs when the mean parameter estimate is equal to the true value of the parameter (topology and branch lengths), while precision quantifies the uncertainty in the parameter estimate (see definitions in Hillis and Bull 1993). To compare our best tree estimate with those based on subsampling of individuals, loci, and base pairs, we used the K tree score, which takes into account topology and branch lengths (Soria-Carrasco et al. 2007). This approach allowed us to evaluate optimal sampling strategies that converged on the best-supported species tree with fewer data (fewer loci, individuals per species, base-pairs, and species).

**Methods**

*Taxon sampling.*—We included all described species in the *Liolaemus darwinii* group except for *L. montanezi* and *L. cinereus* (Appendix I). For all species except those in the *L. darwinii* complex (*L. darwinii*, *L. grosseorum*, *L. laurenti*, and *L. olongasta*), which were sampled more widely, we collected at or near type localities to include only described lizard lineages. Taxonomic knowledge of the *L. darwinii* group is still incomplete and there are several potential new species awaiting further study (Avila, unpub. data). Based on the cyt *b* topology (see below), three individuals with divergent cyt *b* haplotypes were sampled for each species and for all nuclear loci (Appendix I).
Gene flow likely occurs between some species in the *L. darwinii* complex (Morando et al. 2004), which might affect species tree methods that typically do not accommodate for this source of gene tree discordance (Leaché 2009). Therefore, we excluded individuals from phylogeographic borders or contact zones with other species in the group to minimize the potential impact of intermixed/migrant individuals. Because potential outgroups did not consistently amplify across all loci, *L. boulengeri* was used in half of the loci (LJAMM 2187 and 3476) but other species were used for other loci: *L. cf. montanus* LJAMM 12020 (ACM4, B1D, B8H, BA3, EXPH5, KIF24, MXRA5), *L. telsen* LJAMM 5530 (B3F), *L. capillitas* LJAMM 2788 (B5B), and *L. rothi* LJAMM 2163 (B9G). We therefore made a 'composite' outgroup that was used to test the placement of the root in the species tree (see below).

*Sequence data.*—Genomic DNA was extracted with the DNAeasy Qiagen kit (Qiagen). We used the Green Go Taq PCR kit (Promega) for all PCR reactions in PTC-200 DNA Engine (MJ Research) or GeneAmp PCR 9700 thermal cyclers (Applied Biosystems, Inc.). Sequencing reactions used the Big-Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Inc.) in a GeneAmp PCR 9700 thermal cycler (Applied Biosystems, Inc.). Sequencing products were cleaned with Sephadex G-50 Fine (GE Healthcare Bio-Sciences AB) and sequenced in an automated sequencer ABI 3730xl DNA Analyzer (Applied Biosystems, Inc.). The *cyt b* mtDNA gene was sequenced for all individuals of the *L. darwinii* group (~900 sequences) following methods in Morando et al. (2004). Anonymous nuclear loci (ANL) developed from an individual of *L. darwinii* (LJAMM 7097) were screened for all species included in this
study based on the protocols of Noonan and Yoder (2009). From 30 ANL tested in PCR reactions across the sampled individuals, 20 produced positive PCR reactions for most samples, and the 12 most informative (B9G, A8F, A4B, B3F, A1D, A6D, A9C, B5B, A12D, B8H, B1D, and A9E) were selected for subsequent analyses (Camargo et al. unpub. data). One highly variable protein-coding gene (PRLR from Townsend et al. 2008), five additional protein-coding genes (CMOS, Saint et al. 1998; ACM4, Gamble et al. 2008; MXRA5, Portik et al. unpub. data; EXPH5 and KIF24, Portik et al. 2010), and one intron (BA3, Waltari and Edwards 2002) were also included; this provided a total of 20 loci for analysis (Table 1).

Individuals heterozygous for indels were analyzed with CodonCode Aligner (CodonCode Corp.) to resolve position and length of indels in one of the alleles. Ambiguity codes were used to represent polymorphisms in heterozygous individuals in which gap polymorphisms were coded as 'N'. We did not phase alleles for these heterozygotes in order to reduce computation time in analyses, and even though one previous study did not find differences in species trees estimated with phased vs. unphased data (Kubatko and Gibbs 2010), this issue still deserves detailed simulation testing in future studies. Each locus was analyzed with RDP3 beta35 (Martin et al. 2005) to test for recombination signal and alignments were also examined to check for fixed heterozygote positions that might suggest the occurrence of multiple-copy loci (Thomson et al. 2010). PCR reactions for all nuclear loci were performed with the temperature profile of Noonan and Yoder (2009), but for PRLR we used the touchdown cycling protocol for nuclear genes described in (Reyes-Velasco and Mulcahy 2010). Sequences were aligned in Clustal X 2.0 (Larkin et al. 2007). Best-fit substitution models were
obtained in jModeltest 0.1.1 (Posada 2008) with a BIC criterion for model choice (Table 1). We calculated the correlation between the locus variation (proportion of variable sites) and the proportion of informative sites for each locus, and also calculated the correlation between locus variation and the support index for the corresponding gene tree (proportion of highly supported nodes with posterior probability > 0.95).

Species tree.—Each locus was included as a separate dataset in an estimate of the species tree using a Bayesian method in *BEAST. We chose this method because it has been shown to outperform other methods under relaxed molecular clock assumptions (Heled and Drummond 2010). In addition, this approach provides an easy and appropriate way to obtain variance estimates from the posterior distribution of trees. First, we ran analyses with all the available data including 20 loci and a maximum of 3 individuals per species per gene. In addition to the species tree, we also estimated the gene tree from each locus and evaluated their relative discordance with the species tree using the number of deep coalescences (DC) in Mesquite 2.73 (Maddison and Maddison 2010). We calculated the correlation between locus variation and a standardized measure of gene tree discordance consisting of the ratio between DC and the number of gene copies (GC).

We prepared two new reduced datasets for each locus by randomly removing one and then two individuals per species. To sub-sample sequence length for each locus, we prepared new datasets for each locus by removing 25%, 50%, and 75% of the sites from all sequences. We also made new datasets with fewer species but always including species from each of the two basal clades in the group (see below). Datasets with 6 species included: *L. abauca*o, *L. crepuscularis, L. irregularis, L. laurenti, L. quilmes*, and
L. ornatus; for 10 species, we added L. albiceps, L. darwinii, L. grosseorum, and L. lavillai; and we then added L. chacoensis, L. espinozai, L. koslowskyi, and L. olongasta for the 14-species datasets. Based on these new datasets with fewer individuals, base-pairs, and species, we randomly subsampled loci to run analyses with 4, 8, 12, and 16 loci (Fig. 2). To minimize the influence of locus-specific effects, subsampling was done in a nested fashion in which 12 loci were sampled from the pool of 16 loci, 8 loci from the pool of 12 loci, and 4 loci from the pool of 8. Five replicate sampling trials were done to evaluate the effect of different locus combinations within each of these 4 subsets. In Beauti (Drummond and Rambaut 2007), we made all possible combinations of number of loci (4, 8, 12, 16, and 20) with number of individuals (1, 2, and 3), number of loci with proportion of sites (25%, 50%, 75%), and number of loci with number of species (6, 10, and 14).

In addition, because the locus 'informativeness' may also have an impact on species tree performance, we analyzed three groups of loci based on their variability (Table 1): (1) most variable (MV) loci, (2) most conserved (MC) loci, and (3) a mix of variable and conserved loci (VC). For example, we selected the two MV and the two MC loci for four loci, the four MV and four MC for 8 loci, and so on. Moreover, we also ranked loci based on gene tree discordance (DC/GC) with the best species tree (Table 1) in three classes: (1) most discordant (MD) loci, (2) least discordant (LD) loci, and (3) a mix of most and least discordant loci (ML). We used a separate relaxed molecular clock model for each gene with estimation of relative clock rates. We used random starting gene trees under the coalescent model, a Yule process and gamma-distributed population sizes for the species tree prior, and a continuous population size model with a constant
Analyses were run for 100 million generations and samples taken every 4,000 generations with default prior distributions and operator settings. Log files were inspected in Tracer (Rambaut and Drummond 2007) to determine an appropriate burn-in sample for obtaining point estimates and credible intervals of species trees.

**Performance.**—Liu et al. (2009a) suggested that the sampling impacts on species tree estimation could be assessed with a measure of the variance in the tree estimate. More informative datasets are expected to produce more precise estimates coupled with a lower variance. For example, a metric such as the branch length distance (BLD) (Kuhner and Felsenstein 1994), which takes into account topology and branch lengths, could be used to measure distance between trees in the posterior distribution. A modified version of this metric, the minimum BLD or K tree score, measures differences in tree topology and *relative* branch lengths and consequently, the absolute differences in tree depth are scaled to be the same (Soria-Carrasco et al. 2007). This metric is not symmetric and therefore is appropriate when one single reference tree (the target 'true' tree in our case) is compared to estimates of the reference tree when evaluating the performance of phylogenetic methods (Soria-Carrasco et al. 2007). We assumed in our empirical study that the species tree estimated with all data represents our best estimate, which is a reasonable assumption if the method is statistically consistent. In this context, we considered 'accuracy' not as an estimate of the true species tree but as a measure of convergence towards our best tree based on all the data. In order to calculate accuracy, we summarized the posterior species tree distributions to obtain the maximum clade credibility tree, using TreeAnnotator (Drummond and Rambaut 2007). Our best estimate
was used as a reference tree, whereas the best trees obtained with different combinations of loci, sites, and individuals were used as comparison trees to calculate K scores with Ktreedist v.1 (http://molevol.cmima.csic.es/castresana/Ktreedist.html, Soria-Carrasco et al. 2007). Lower K tree scores were considered as more accurate estimates. To estimate the precision of the species tree, we subsampled 100 trees from the posterior distribution using LogCombiner (Drummond and Rambaut 2007), and used them as comparison trees for calculating K tree scores. Based on these 100 scores, we calculated their variance as an estimator of precision since a lower variance in K tree scores represents a more precise estimate of the species tree. In addition, we calculated the correlation between the support index and the precision of species trees estimated with varying number of loci and individuals.

**RESULTS**

*Sequence data.*—All sequences are available in Genbank (accession numbers XXXX, to be provided upon publication). Sequence length varied between 291 bp (B1D) and 867 bp (MXRA5). Percentage of variable sites across the 20 loci ranged between 3% (A9E) and 35% (cyt b) (Table 1). Single-bp indels were rare in ANL sequences, there were a few multiple-bp indels in some ANL, and alignments were unambiguous except for a 16-bp segment in the A9C locus that was excluded from analyses. No signal of recombination was detected in any of the datasets. Species represented by only one individual occurred in only ~ 6% of all cases (21 of 320 species/gene combinations [= 20 loci x 16 species]) whereas species represented by 2-3 individuals were common across
all loci (94%), and only 7% of sequences sampled from within species were identical. The proportion of variable sites was significantly correlated with their proportion of informative sites ($R^2 = 0.75$, $F_{1,17} = 51.6$, $P < 0.01$), and with clade support in their corresponding gene trees ($R^2 = 0.50$, $F_{1,18} = 17.9$, $P < 0.01$) (Appendix II).

**Gene trees and species tree:**—Burn-in plots suggested discarding 25% of samples (5,000) for estimating posterior distributions of gene and species trees with the remaining 20,000 samples. In spite of intraspecific variation, gene trees show little paraphyly within species except for some interdigitation of samples between *L. irregularis* and *L. albiceps*, *L. espinozai* and *L. quilmes*, and *L. laurenti* and *L. grosseorum* (see Appendix III). In addition, lineages representing species pairs in the best-supported species tree were divergent in all cases except for the (*L. irregularis* + *L. albiceps*) clade, suggesting good support for species-level differentiation. The best-supported species tree derived from complete datasets for 20 loci recovers 8 of 15 nodes with a posterior probability $\geq 0.95$ (Fig. 3). Two well supported basal clades are present: clade A, which groups *L. albiceps*, *L. irregularis*, *L. ornatus*, *L. lavillai*, *L. crepuscularis*, *L. calchaqui*, *L. espinozai*, *L. quilmes*, *L. chacoensis* and *L. uspallatensis*, and clade B, including *L. grosseorum*, *L. laurenti*, *L. darwinii*, *L. koslowskyi*, *L. abaucan* and *L. olongasta* (Fig. 3). The A9C locus was the most discordant with the species tree topology whereas cyt *b* was the least discordant, based on the ratio between deep coalescences and number of gene copies (Table 1, Appendix III). The placement of the root was identical when the outgroup was excluded from the analysis but there were fewer well supported branches, especially at basal nodes (Appendix II). The proportion of variable sites was not correlated with the
discordance of gene trees ($R^2 = 0.02$, $F_{1,17} = 0.30$, $P = 0.59$), but the precision of species trees and the clade support index were significantly correlated ($R^2 = 0.84$, $F_{1,13} = 66.52$, $P < 0.01$, Appendix II).

**Sampling of loci and individuals.**—There was an increase in the accuracy with the number of loci for all numbers of individuals sampled per species but accuracy was consistently higher when sampling 2-3 individuals (Fig. 4a). Precision also shows a steep increase between 4 and 12 loci, and again sampling of 2-3 individuals per species is always better than sampling only one individual (Fig. 4b). The magnitude of the differences in K score between the best supported tree and trees estimated with fewer loci involved differences in topology: trees based on 16 and 12 loci had different relationships within clade B, and the 8-locus and 4-locus trees grouped *L. uspallatensis* and *L. chacoensis* within clade A (Appendix IV). In all cases, branch support, and consequently precision of species trees, declined with fewer loci.

**Sampling of base pairs.**—The accuracy of species trees estimated from datasets with variable number of sites show a similar pattern of improvement with increases in the number of loci and reached a plateau between 12 and 16 loci (Fig. 5a). Accuracy was almost indistinguishable for datasets with full sequence length, 75% (~440 bp/locus), and 50% (~295 bp/locus) of the sites, but analyses run with only 25% of the sites (~147 bp/locus) had lower accuracy for any given number of loci. Precision showed a similar pattern with a large increase between 4 and 12 loci and an apparent plateau after 16 loci.
whereas sequences with only 25% of the sites usually led to lower precision values (Fig. 5b).

**Sampling of species.**—Both accuracy and precision increased with the number of loci for all sampling levels of species (6, 10, and 14 species) and as expected, species trees with fewer species required fewer loci, i.e. increase in accuracy is faster for 6 species, intermediate for 10 species, and slower for 14 species (Fig. 6a, b). As seen above for analyses with all 16 species, there was a sustained increase in accuracy with number of loci (Fig. 6a) but precision stabilized after ~12 loci (Fig. 6b). The species tree topology for 6 species was identical for all number of sampled loci, but convergence in topology was achieved with 16 loci for 10 species, and with 12 loci for 14 species (not shown).

**Locus 'informativeness'.**—Sets of loci with different levels of variation generated species trees with roughly the same accuracy when 12 and 16 loci were included in the analyses, but performance was better with the most variable loci when only 4 and 8 loci were analyzed (Fig. 7a). An increase of precision with number of loci was clear for all three sampling designs of loci and consistently, the set of most variable loci gave more precise estimates, the set of most conserved loci gave the least precise estimates and the mixed set of loci produce an expected intermediate pattern (Fig. 7b). When the discordance of gene trees was used as an indicator of locus 'informativeness', the least discordant loci had the highest accuracy for all numbers of loci followed by the most discordant loci, and the mix of least and most discordant loci had the lowest accuracy.
However, precision was similar among these groups of loci with different levels of discordance with the species tree (Fig. 8b).

**DISCUSSION**

In empirical phylogenetics, researchers have finite data at hand for estimating species trees because time and/or technical constraints make it difficult to obtain sequence data from a large number of loci. Limited budgets also usually forces researchers to sample either more loci or more individuals for a given effort to minimize sequencing costs (Felsenstein 2006). Moreover, many potentially useful markers with variability sufficient for resolution of recent species radiations such as ANL might be poorly informative if very short sequences free of recombination are used in analyses. In this study we assembled 20 loci for 16 species, which represents a large dataset derived from direct sequencing techniques compared to sample sizes often used in empirical studies of species trees (Belfiore et al. 2008; Liu et al. 2008; Espregueira-Themudo et al. 2009; Leaché 2009; Kubatko and Gibbs 2010). This data coverage allowed us to evaluate the performance of one species tree method with decreasing number of loci, individuals, sequence length, and species. We measured performance in terms of both accuracy and precision because higher precision is correlated with higher overall support of the species tree, which is of interest for evaluating confidence of inferred trees in empirical phylogenetics. We found an asymptotic increase in accuracy and precision with the number of loci and an improvement in accuracy when using >1 individual/species and/or >1/4 of original datasets for any number of loci. Locus 'informativeness' was an
important factor when using a limited number of loci but it was increasingly irrelevant with more loci and accuracy decreased with higher gene tree heterogeneity. Not surprisingly, we found that the number of loci required for resolving species trees increases with the number of species, thus more loci will have to be sequenced to increase nodal support for future studies of this clade that include the remaining species.

**Performance**

Rather than compare performance of different species tree methods, here we focused on one method that relaxes many assumptions of other methods (Heled and Drummond 2010). Our choice for a fully-probabilistic method was based on findings that these methods usually outperform summary-statistic methods (e.g. Liu et al. 2009b). Further, *BEAST has been shown to outperform BEST (Heled and Drummond 2010), which in turn is more accurate than STEM and concatenation over a range of divergence times, population sizes, and species tree topologies (Leaché and Rannala accepted). However, summary-statistic methods are more efficient when dealing with genomic datasets that are beyond computational capabilities for highly parametric, coalescent methods (Liu et al. 2009b; Knowles and Kubatko 2010). One advantage of *BEAST over other methods with practical implications is the possibility of estimating the root position without using outgroups, which sometimes are difficult to sample with sequence-based markers due to the annealing specificity of PCR primers (as in this study). In fact, we found that *BEAST correctly inferred the root position using our dataset without an outgroup, as suggested by previous analyses (Heled and Drummond 2010).
Evaluation of species tree methods requires simulation studies but empirical studies are also necessary to assess performance with real data, which include gene genealogies shaped by different but unknown historical and demographic processes. Empirical analyses are based on sample sizes (individuals, loci, and base pairs) limited by budgetary and time constraints. On the other hand, simulation studies use a common, simple evolutionary model for all loci and sometimes dataset sizes are unrealistically large at least for direct sequencing techniques (i.e., 100 loci in Leaché and Rannala accepted; 6,400 bp/locus in Heled and Drummond 2010). In addition, more empirical studies based on realistic data and divergence scenarios are needed to apply recommendations about sampling strategies derived from theoretical and simulation studies (Castillo-Ramírez et al. 2010; Knowles 2010). Another important contribution of this study was our measurement of performance based on topology and also branch lengths using the K scores; most studies to date have quantified accuracy in terms of topology only (Liu and Edwards 2009; McCormack et al. 2009; Leaché and Rannala accepted).

Our evaluation of species tree inference using a coalescent-based method is valid as long as gene tree heterogeneity was a result of incomplete lineage sorting only. When including multiple individuals for reconstruction of species trees, sampling from phylogeographic lineage boundaries might obfuscate phylogenetic signal due to gene flow between species (Leaché 2009). Our geographic sampling strategy probably minimized the impact of gene flow as a source of discordance because we avoided regions of known or suspected contact zones. However, the distributions of several of the species in the *L. darwinii* group are not fully known, and hybridization/introgression of
mtDNA is known in some closely-related species within the *L. darwinii* complex (Morando et al. 2004). If this has occurred in our samples, it was likely restricted to closely related or sister species, which should not lead to major phylogenetic estimation error (Brumfield et al. 2008; Eckert and Carstens 2008). Future studies of this and other groups are planned which will incorporate newer methods being developed to include introgression as another source of gene tree heterogeneity (Kubatko and Meng 2010).

*Number of loci and individuals.*—Performance patterns found with subsampling of individuals and loci in our study mimic simulation results obtained for deep divergence histories. In our analysis, larger gains in performance were obtained with an increase in number of loci (up to 12 loci when convergence was reached in topology). Besides providing more accurate estimates of species trees (by the criterion applied in this study), increasing the number of loci seems to be related to the benefits of using more loci for robust estimates of population sizes (Felsenstein 2006), which is a critical parameter for accommodating gene tree discordance with the species tree in the multispecies coalescent model (Castillo-Ramírez et al. 2010). In contrast, there was an increase in accuracy when more than 1 individual per species was sampled, but no further gain between 2 and 3 individuals. In a simulation study, a larger gain in accuracy was obtained with more loci instead of more individuals for deep divergences and a given species sampling effort, but the biggest gains in accuracy for shallow histories were obtained when more individuals were sampled (McCormack et al. 2009). Gains in performance with more sampled loci, but not with more sampled individuals, suggest that our species tree represents a relatively deep diversification history within the *L. darwinii*
Multiple individuals are preferred over multiple loci in recent radiations because more gene copies are likely to cross the species boundary and coalesce in ancestral populations (Heled and Drummond 2010; Knowles 2010). Consequently, more individuals should be sampled when gene lineages within species have not yet sorted to monophyly, but when species are recovered as well-supported clades in older divergences (as in our gene trees, Appendix III) only additional loci contain phylogenetic signal about species relationships (Knowles 2010).

In addition to very low performance, sampling of only one individual/species caused poor convergence and mixing of MCMC chains in *BEAST, suggesting low information content in the data. The impact of using single individuals/species could be more serious for *BEAST because this method also estimates population sizes of extant lineages (Heled and Drummond 2010), while methods such as BEST only estimate population sizes for ancestral lineages where multiple alleles can coexist (Castillo-Ramírez et al. 2010). Even though there is an increase in performance with more loci (especially in deep divergences) and individuals (at shallow divergences) because of reduced coalescence variance, there is a concomitant increase of mutational variance with trade-offs for the relative gains of increased sampling effort (Huang et al. 2010). For example, when sampling additional individuals without increasing sequence variation, the search through a tree space with more alternative topologies becomes more difficult and leads to more uncertainty in gene and species trees (Huang et al. 2010).

Sequence length.—Our subsampling of base-pairs to assess the effect of sequence length also yielded similar results to those from simulation studies. Simulations suggest
that increasing locus length is a better strategy than sequencing additional loci for a given total number of base-pairs using BEST (Castillo-Ramírez et al. 2010). For a given topology and number of species, these authors found larger gains in accuracy with loci of 500 vs. 250 bp, but similar accuracy with loci of 500 or 1,000 bp. In addition, Heled & Drummond (2010) found that precision (using credible sets of trees) was doubled when sequence length increased from 200 to 800 bp in *BEAST. These simulations are consistent with the substantial drop in accuracy (and precision) that we found when using 25% of original sequences (~147 bp), but convergence in accuracy when analyzing 50% of sites or more (> 295 bp). These comparisons suggest that, for a given speciation history, there may exist a minimum threshold in sequence length below which the mutational variation is too low for robust estimation of gene trees, and resulting in inaccurate and poorly supported species trees. The impact of mutational variance could potentially be reduced by methods that incorporate gene tree uncertainty (i.e., BEST and *BEAST), although these highly parametric methods might have poor performance with limited genetic variation (Huang et al. 2010). However, recent simulations have shown that BEST is more accurate than other methods based on point estimates of gene trees when there is low genetic variation (Leaché and Rannala accepted).

*Number of species.*—We subsampled species from each of the two most inclusive clades in order to preserve the same tree depth among analyses with varying number of loci. Tree depth influences the performance of species trees because deep divergences reduce the extent of incomplete lineage sorting and produce less gene tree discordance (Maddison and Knowles 2006). Therefore, our sampling strategy attempted to control
this factor to separately assess the effect of species number on performance. Performance was higher for trees with fewer species at any given number of sampled loci, and the difference was greater with fewer loci, suggesting that less information is required to estimate species trees with fewer species. Although, the parameter space might be easier to explore for trees with fewer species (i.e., fewer internal branches and topologies), these results might still apply to our specific speciation history only, so simulations will be required to assess this issue in a more general context. Subsampling of species and loci showed that only 4 loci were necessary to recover a 6 species tree, and between 12-16 loci were adequate for 10-14 species trees. We did not attain convergence for the full 16 species tree, but precision (i.e., branch support) increased consistently with the addition of more loci. Therefore, 20 loci might be enough for the 16 species included in our study, but to increase confidence in some poorly supported branches (i.e., higher precision) additional loci should be sequenced. More importantly, future sampling will add the remaining species (e.g., *L. montanezi* and *L. cinereus*), and/or several undescribed forms (Etheridge 1993; Morando et al. 2004; Abdala 2007) that would require more loci to resolve an accurate species tree for the group.

*Locus 'informativeness'.—*Another sampling dimension relevant in empirical phylogenetics but that has been poorly explored in simulation studies is the relative information content of loci (Knowles 2009; 2010). In simulations, all loci usually have the same length and a common, sometimes simplified, substitution model is employed which reduces the rate variation across loci. Here, we evaluated performance with loci that differed in variability, length, substitution model, and discordance with the species
tree (Table 1), and our results show reduced performance with conserved loci, but this improved with more variable loci, and was highest when the most variable loci were used. These results are intuitive and agree with simulations demonstrating that the low number of informative sites is the most relevant factor decreasing accuracy under some simulation conditions (Castillo-Ramírez et al. 2010), probably as a result of limited phylogenetic signal for estimating well supported gene trees (McCormack et al. 2009; Huang et al. 2010). Because both higher number and 'informativeness' of loci increased performance, inclusion (to increase quantity) vs. exclusion (to increase 'informativeness') of a locus can be justified depending on which strategy provides a larger gain in performance. Although excluding conserved loci to increase performance is an appealing option, this is not an advisable strategy because all loci contain information about the coalescent process at some level of divergence, and arbitrary exclusion might introduce ascertainment biases in parameter estimates (Knowles 2010).

The next most relevant factor impacting performance was gene tree heterogeneity (Castillo-Ramírez et al. 2010). When gene tree heterogeneity increases, more loci have to be analyzed to estimate species trees (Edwards et al. 2007). We tested this prediction by analyzing loci with varying amounts of discordance with the species tree and found precisely that the most heterogeneous mix of discordant and concordant loci had the lowest accuracy. High levels of gene tree heterogeneity are common, not only in recent species radiations, but also when short branches in the species tree generate frequent AGTs and consequently, more loci are required to estimate species trees accurately (Knowles 2010). Even though our species tree seems to reflect a deep speciation history,
short internodes might be responsible for the high degree of heterogeneity observed in our gene trees since none of them matches the species tree topology (Appendix III).

**Systematics**

Previous phylogenetic studies of the *L. darwinii* group also recovered a basal split into two major clades but with some differences in their species composition compared to our species tree. All previous studies found support for the 'ornatus' clade (nested within clade A in Fig. 3) that includes *L. albiceps*, *L. irregularis*, *L. ornatus*, *L. lavillai*, *L. calchaqui*, and *L. crepuscularis*. In addition, these studies also grouped *L. darwinii*, *L. laurenti*, *L. grosseorum*, *L. chacoensis* and *L. olongasta* into the 'grosseorum' clade (Abdala 2007). These studies differed in the kind of data used to infer the phylogeny and their species sampling of the *L. darwinii* group: Etheridge (2000) used morphological and behavioral characters of 10 species, Schulte et al. (2000) sequenced 3 mtDNA genes plus several tRNAs of 11 species, Morando (2004) analyzed 3 mtDNA and 2 nuclear genes of 11 species, and Abdala (2007) inferred a morphological + molecular phylogeny for 16 species (the same to those used in this study). Our species tree placed the well-supported (*L. quilmes*, *L. espinozai*) clade as sister to the 'ornatus' clade within clade A and the well-supported (*L. koslowskyi*, *L. abaucan*) clade within clade B (Fig. 3).

The major difference is the placement of *L. uspallatensis* and *L. chacoensis* within clade A in our species tree, which were usually assigned to the 'grosseorum' clade in previous studies [except for *L. uspallatensis* which was basal in Abdala (2007)]. These conflicting results might be a result of shared ancestral polymorphisms of *L. uspallatensis*.
and *L. grosseorum* with clade B, which would bias the concatenated analyses used in previous studies that do not account for the process of incomplete lineage sorting. On the other hand, *L. chacoensis* is morphologically more similar to members of the *L. wiegmannii* group, an outgroup of the *L. darwinii* group (Abdala 2007), which could also have impacted the morphology-based, or combined morphological-molecular analyses.

From a biogeographic perspective, our species tree implies a clear transition between ecoregions because clade B occurs mostly in the Monte Desert of south-central and west-central Argentina at lower altitudes (Etheridge 1993). Within clade A, the 'ornatus' clade occupies the Puna and Prepuna of northwestern Argentina at higher altitudes (Abdala 2007). Sister to the 'ornatus' clade, the (*L. espinozai, L. quilmes*) clade inhabits the Prepuna-Monte ecotone and the northernmost region of the Monte Desert, respectively (see maps in Roig-Juñent et al. 2001; Abdala 2005). The species external to these clades occur in the isolated Uspallata-Calingasta valley in west-central Argentina (*L. uspallatensis*) and the Chaco lowlands of central and northern Argentina (*L. chacoensis*). In addition, the topology of our species tree is also consistent with the natural history of the group (Abdala and Díaz Gómez 2006) with viviparity evolving twice, once in *L. espinozai* and once in the 'ornatus' clade (assuming no reversals from viviparity to oviparity). Although our goal was not to address species limits in the group, the virtual lack of genetic divergence and slight morphological differentiation between *L. albiceps* and *L. irregularis* (Lobo and Laurent 1995) warrant further sampling and phylogeographic analysis to elucidate their potential conspecificity.
CONCLUSIONS

Diversification in the *Liolaemus darwinii* group appears to be old but with episodes of rapid speciation that appear to have resulted in short internal branches in the species tree and high gene tree heterogeneity. In fact, relaxed clock estimates with BEAST based on a mean rate of 0.65% substitutions/site/million years for cyt *b* (Morando et al. 2004) results in a mean age of ~13 Ma for the ancestral basal node of the *L. darwinii* group. The diversification of this clade represents a unique and specific speciation history, but our results are consistent with simulation studies that investigated a range of speciation histories with a variety of sampling designs. Results of all of these studies consistently suggest minimal sequence lengths, numbers of individuals, and numbers of loci for species trees inference that depends on speciation history, kind of data, and the specific inference method used. However, there also appears to be upper limits for sampling effort since increasing number of individuals and loci increases mutational variance (Huang et al. 2010), and increasing sequence length increases the probability of intra-locus recombination, which can mislead species trees methods due to the violation of model assumptions (Castillo-Ramírez et al. 2010). Therefore, there appears to be an optimal range of sampling effort for a given level of information content in the data set which can be increased by including loci with more informative sites since this is the most relevant factor impacting accuracy of species tree inference (Castillo-Ramírez et al. 2010). In this vein, new genome sequencing technologies and genome-enabled markers (Thomson et al. 2010) will probably allow the discovery of highly informative loci for more accurate and precise estimation of species trees. Until these
more informative markers become widely available, research should probably focus on the limits of species tree inference when there are mixed levels of nucleotide variation.

The choice of an appropriate species tree method and sampling design will depend on the data available and an unknown speciation history. In general, more data (especially more individuals and longer sequences with high number of informative sites) should be analyzed when speciation has been recent (shallow divergences) and rapid (short branch lengths). However, development of coalescent-based species tree methods is accelerating (Knowles and Kubatko 2010), as are sequencing technologies, so more studies are necessary to further assess optimal sampling designs for clades with varying numbers of species, tree topologies, branch lengths, population sizes, and information content of data, using simulation approaches (Knowles 2010). While generalizations about minimum and optimal sampling designs will be difficult, a potential approach could consist of the subsampling strategy used in this study, to assess convergence of species tree estimates (topology and branch lengths) and branch support (precision). In this study our subsampling explorations of the data suggested that more individuals per species would not be necessary, but that additional loci might be required given the high levels of gene tree heterogeneity, and when the unsampled species are included.

Several aspects of the speciation history, data, and inference methods that should be further investigated via theoretical, simulation, and empirical studies include tree shape (symmetric vs. pectinate), and variance of branch lengths across the species tree; both could have substantial effects on species trees accuracy (McCormack et al. 2009), particularly in the AGT parameter space (Degnan and Rosenberg 2006). While Huang and Knowles (2009) found that in practice, AGTs are unlikely to be problematic because
estimated gene trees from the anomaly zone are often unresolved, more research is needed to predict how the pattern and timing of diversification can confound the reconstruction of species trees. Second, the impact of gene flow has not been adequately explored but it probably is very influential on species tree methods that only model incomplete lineage sorting, especially if gene flow has occurred deep in the species tree and involved lineages that are not closely related (Eckert and Carstens 2008). In addition, the impact of missing data on species tree estimation has not been assessed although this is of substantial interest because complete datasets containing many loci, individuals, and species are difficult to obtain in practice. Based on ever increasing computational power and more efficient inference approaches, new coalescent-based methods dealing with massive genomic datasets (Liu et al. 2010) and new methods that can account for uncertainty in species limits (O'Meara 2010; Zhang and Cui 2010), are promising developments that will continue to move molecular systematics towards a new research paradigm.

**FUNDING**

AC acknowledges financial support from the BYU Department of Biology and Graduate Studies Office through research, mentoring, and travel scholarships. He also thanks support from the Society for Systematic Biologists and the Society for the Study of Amphibians and Reptiles for research and travel grants. Funding for this study was provided by the NSF “Partnership for International Research and Education” award (OISE 0530267) for collaborative research on Patagonian biodiversity, granted to the
following institutions (listed alphabetically): BYU, Centro Nacional Patagónico, Dalhousie University, Darwinion Botanical Institute, George Washington University, Universidad Nacional de Cordoba, Universidad Austral de Chile, Universidad Nacional del Comahue, and the Universidad de Concepción; and also by the NSF “Assembling the Tree-of-Life – Deep Scaly” project on squamate reptiles (subaward EF 0334966) to JWS, Jr. LJA and MM acknowledge several grants under programs PIP & IBol (CONICET) and FONCYT PICT (ANPCYT) (Argentina).

ACKNOWLEDGMENTS

We thank students and posdocs of Sites, Morando, and Avila labs for their assistance and support throughout the completion of this study. A.C. thanks his graduate committee for feedback on his dissertation research/manuscripts. We thank also valuable comments and suggestions of A. Leaché, S. V. Edwards, E. Benavides, and F. Werneck that greatly improved earlier drafts of the manuscript. Specimens were collected following local regulations and permits issued by Administración de Parques Nacionales, Direcciones Provinciales de Fauna de Río Negro, Chubut, Neuquen, La Pampa, Mendoza, Catamarca, and Administración de Áreas Protegidas de Neuquen y Mendoza (Argentina).
REFERENCES


Table 1. List of molecular markers ranked by percentage of variable sites. DC = deep coalescences, GC = gene copies.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Kind of marker</th>
<th>Substitution model</th>
<th>Length (bp)</th>
<th>Variation (%)</th>
<th>DC</th>
<th>GC</th>
<th>DC / GC</th>
<th>Primers</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYTB</td>
<td>mtDNA</td>
<td>TPM2uf+G</td>
<td>710</td>
<td>35.5%</td>
<td>10</td>
<td>48</td>
<td>0.21</td>
<td>Glu, CB3, F1, C2</td>
<td>Morando et al. 2004</td>
</tr>
<tr>
<td>B9G</td>
<td>ANL</td>
<td>K80+G</td>
<td>513</td>
<td>13.7%</td>
<td>62</td>
<td>39</td>
<td>1.59</td>
<td></td>
<td>This study</td>
</tr>
<tr>
<td>A8F</td>
<td>ANL</td>
<td>HKY</td>
<td>645</td>
<td>12.3%</td>
<td>56</td>
<td>39</td>
<td>1.44</td>
<td>F 5' - AGAGAGGGGAAAGGGGTTTG-3' R 5' - TCCCTTGTATATTCACAGACTTAAC-3'</td>
<td>This study</td>
</tr>
<tr>
<td>A4B</td>
<td>ANL</td>
<td>HKY+G</td>
<td>517</td>
<td>12.0%</td>
<td>31</td>
<td>45</td>
<td>0.69</td>
<td>F 5' - ACTCCCGTGAGATCTCTGTTG-3' R 5' - GGGAAAGGGGTGGCTTTAG-3'</td>
<td>This study</td>
</tr>
<tr>
<td>B3F</td>
<td>ANL</td>
<td>K80+I</td>
<td>505</td>
<td>11.9%</td>
<td>23</td>
<td>46</td>
<td>0.50</td>
<td>F 5' - CAATTCCTGCAAATCCACCCTA-3' R 5' - CCCTCCTCATTTACCTTACATGC-3'</td>
<td>This study</td>
</tr>
<tr>
<td>A1D</td>
<td>ANL</td>
<td>HKY</td>
<td>783</td>
<td>10.6%</td>
<td>51</td>
<td>42</td>
<td>1.21</td>
<td>F 5' - CTCAGCGCACAGCAAGAA-3' R 5' - TTAGAGAAAATCTCGAGTAGCA-3'</td>
<td>This study</td>
</tr>
<tr>
<td>A6D</td>
<td>ANL</td>
<td>HKY+I</td>
<td>476</td>
<td>9.9%</td>
<td>27</td>
<td>47</td>
<td>0.58</td>
<td>F 5' - CAATTCCTGCAAATCCACCCTA-3' R 5' - TTTCAGAAAGATGCTGCAAAT-3'</td>
<td>This study</td>
</tr>
<tr>
<td>A9C</td>
<td>ANL</td>
<td>K80+G</td>
<td>745</td>
<td>9.8%</td>
<td>100</td>
<td>38</td>
<td>2.63</td>
<td>F 5' - GATAGTTACTGCAAGGGTGCCA-3' R 5' - GGGCTATTGGTGAGCGGTCTC-3'</td>
<td>This study</td>
</tr>
<tr>
<td>B5B</td>
<td>ANL</td>
<td>HKY</td>
<td>623</td>
<td>9.8%</td>
<td>13</td>
<td>44</td>
<td>0.30</td>
<td>F 5' - GCAGAGCCAAGCCATGT-3' R 5' - GTAGGCTTTGAGCTTAGCTC-3'</td>
<td>This study</td>
</tr>
<tr>
<td>A12D</td>
<td>ANL</td>
<td>HKY</td>
<td>632</td>
<td>9.0%</td>
<td>61</td>
<td>37</td>
<td>1.65</td>
<td>F 5' - GTTTCCTGGAAAGGCTATGAA-3' R 5' - TCCAAAATGCTACACTGAGG-3'</td>
<td>This study</td>
</tr>
<tr>
<td>B8H</td>
<td>ANL</td>
<td>HKY</td>
<td>625</td>
<td>8.8%</td>
<td>35</td>
<td>38</td>
<td>0.92</td>
<td>F 5' - TTCTTTAAAAGCCAGCCAGACA-3' R 5' - GTTCCTACAGGCTGTGG-3'</td>
<td>This study</td>
</tr>
<tr>
<td>EXPH5</td>
<td>Coding</td>
<td>HKY+I</td>
<td>790</td>
<td>7.6%</td>
<td>73</td>
<td>43</td>
<td>1.70</td>
<td>F1, R1</td>
<td>Portik et al. in press</td>
</tr>
<tr>
<td>PRLR</td>
<td>Coding</td>
<td>HKY+I</td>
<td>552</td>
<td>7.6%</td>
<td>18</td>
<td>46</td>
<td>0.40</td>
<td>F1, R3</td>
<td>Townsend et al. 2008</td>
</tr>
<tr>
<td>B1D</td>
<td>ANL</td>
<td>HKY</td>
<td>291</td>
<td>7.2%</td>
<td>66</td>
<td>44</td>
<td>1.50</td>
<td>F 5' - GATATCGAGGGATTTTCGTTTCC-3' R 5' - CCAGTGTTTATGAGCAACTGAGTA-3'</td>
<td>This study</td>
</tr>
<tr>
<td>KIF24</td>
<td>Coding</td>
<td>HKY+I</td>
<td>551</td>
<td>7.1%</td>
<td>15</td>
<td>48</td>
<td>0.31</td>
<td>F1, R1</td>
<td>Portik et al. in press</td>
</tr>
<tr>
<td>BA3</td>
<td>Intron</td>
<td>K80</td>
<td>347</td>
<td>6.9%</td>
<td>65</td>
<td>41</td>
<td>1.59</td>
<td>F, R</td>
<td>Waltari &amp; Edwards 2002</td>
</tr>
<tr>
<td>MXRA5</td>
<td>Intron</td>
<td>HKY</td>
<td>867</td>
<td>5.4%</td>
<td>27</td>
<td>42</td>
<td>0.64</td>
<td>F, R</td>
<td>Portik et al. unpub. data</td>
</tr>
<tr>
<td>Reference</td>
<td>Mutations</td>
<td>Model</td>
<td>Length</td>
<td>% Insertions</td>
<td>% Deletions</td>
<td>% Substitutions</td>
<td>F/R Primers</td>
<td>Sequences</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-------</td>
<td>--------</td>
<td>--------------</td>
<td>-------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>ACM4</td>
<td>Coding</td>
<td>TPM1uf+I</td>
<td>453</td>
<td>4.9%</td>
<td>22</td>
<td>42</td>
<td>F, R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A9E</td>
<td>ANL</td>
<td>HKY+I</td>
<td>615</td>
<td>4.2%</td>
<td>73</td>
<td>38</td>
<td>F 5'-TGAACATGCCAGACAGAAACA-3' R 5'-TCCCGTAGTCCACAAACTGG-3'</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>CMOS</td>
<td>Coding</td>
<td>K80+I</td>
<td>523</td>
<td>4.2%</td>
<td>13</td>
<td>48</td>
<td>G73, G78</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Gamble et al. 2008

Saint et al. 1998
List of Figures

Figure 1. Map of Argentina showing sampled localities for species of the *Liolaemus darwinii* group. Numbers refer to localities listed in Appendix I. The inset shows the map of South America with the sampled region in gray. Pictures are not scaled. Photos by L. J. Avila (*L. darwinii, L. laurenti, L. calchaqui, L. quilmes, L. olongasta, L. chacoensis, L. ornatus, L. albiceps*, and *L. abaucan*) and C. Pérez (*L. lavillai, L. uspallatensis, L. grosseorum, L. irregularis*, *L. crepuscularis, L. koslowskyi*, and *L. espinozai*).

Figure 2. Diagram showing the strategy used in this study to sample loci, individuals, base-pairs, and species from 20 loci. MCT = maximum credibility tree; HPD = highest posterior density.

Figure 3. Species tree of the *Liolaemus darwinii* group based on 20 loci. Numbers on branches represent posterior probabilities. Transitions from oviparity to viviparity are marked with a transversal bar on branches.

Figure 4. Accuracy (a) and precision (b) of species trees estimated from different number of loci (4, 8, 12, 16, and 20) and different number of individuals (1, 2, and 3).

Figure 5. Accuracy (a) and precision (b) of species trees estimated from different number of loci (4, 8, 12, 16, and 20) and different number of base-pairs (100%, 75%, 50%, and 25% of full sequences).
Figure 6. Accuracy (a) and precision (b) of species trees estimated from different numbers of loci (4, 8, 12, 16, and 20) and different numbers of species (6, 10, and 14) representing the two most inclusive clades of the species tree.

Figure 7. Accuracy (a) and precision (b) of species trees estimated from different number of loci (4, 8, 12, 16, and 20) and different loci combinations based on their proportion of variable sites (see Table 1).

Figure 8. Accuracy (a) and precision (b) of species trees estimated from different number of loci (4, 8, 12, 16, and 20) and different loci combinations based on their discordance with the species tree (see Table 1).
Figure 1
Figure 2

- **20 loci, 16 spp, 3 ind, 100% bp**
  - **Species tree (reference tree)**
    - 14 spp
      - 10 spp
        - 6 spp
    - 3 → 2 ind
    - 25% bp
  - 2 → 1 ind
  - 50% bp
  - 75% bp

- **16 loci, 12 loci, 8 loci, 4 loci**

- **K tree score**
  - Accuracy (MCT)
  - Precision (95%-HPD)

Figure 2
Figure 3
Figure 4

(a) Accuracy (Mean K scores)

(b) Precision (Variance of K scores)
Figure 5
Figure 6
Figure 7

(a) Accuracy (Mean K scores)

(b) Precision (Variance of K scores)
Figure 8
### Appendix I. List of specimens sequenced for this study.

<table>
<thead>
<tr>
<th>Specimen</th>
<th>Number</th>
<th>Coordinates</th>
<th>Province</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>L. abaucan</em></td>
<td>LJAMM 2359</td>
<td>27º 26' 50&quot; S, 67º 40' 44&quot; W</td>
<td>Ruta Provincial 36, 16 Km S Palo Blanco, Tinogasta, Catamarca</td>
</tr>
<tr>
<td></td>
<td>LJAMM 2362</td>
<td>27º 26' 50&quot; S, 67º 40' 44&quot; W</td>
<td>Ruta Provincial 36, 16 Km S Palo Blanco, Tinogasta, Catamarca</td>
</tr>
<tr>
<td></td>
<td>LJAMM 2371</td>
<td>27º 26' 50&quot; S, 67º 40' 44&quot; W</td>
<td>Ruta Provincial 36, 16 Km S Palo Blanco, Tinogasta, Catamarca</td>
</tr>
<tr>
<td><em>L. albiceps</em></td>
<td>LJAMM 12040</td>
<td>24º 59' 57&quot; S, 66º 09' 15&quot; W</td>
<td>Road towards Nevado del Acay, 5 km S Estacion Muñano, from Ruta Nacional 51, Rosario de Lerma, Salta</td>
</tr>
<tr>
<td></td>
<td>LJAMM 12046</td>
<td>24º 59' 57&quot; S, 66º 09' 15&quot; W</td>
<td>Road towards Nevado del Acay, 5 km S Estacion Muñano, from Ruta Nacional 51, Rosario de Lerma, Salta</td>
</tr>
<tr>
<td></td>
<td>2646</td>
<td>24° 27' 02&quot; S, 65° 57' 05&quot; W</td>
<td>Santa Rosa de Tastil, Rosario de Lerma, Salta</td>
</tr>
<tr>
<td><em>L. calchaqui</em></td>
<td>LJAMM 12834</td>
<td>26º 22' 45&quot; S, 65º 43' 54&quot; W</td>
<td>Ruta Provincial 352, 38.3 km W Hualinchay, Cumbres Calchaquies, Trancas, Tucumán</td>
</tr>
<tr>
<td></td>
<td>LJAMM 12837</td>
<td>27º 22' 45&quot; S, 65º 43' 54&quot; W</td>
<td>Ruta Provincial 352, 38.3 km W Hualinchay, Cumbres Calchaquies, Trancas, Tucumán</td>
</tr>
<tr>
<td></td>
<td>12842</td>
<td>28º 22' 45&quot; S, 65º 43' 54&quot; W</td>
<td>Ruta Provincial 352, 38.3 km W Hualinchay, Cumbres Calchaquies, Trancas, Tucumán</td>
</tr>
<tr>
<td><em>L. chacoensis</em></td>
<td>LJAMM 10649</td>
<td>32º 53' 19&quot; S, 66º 49' 04&quot; W</td>
<td>1 km NE La Calera, 8.7 km SW Ruta Nacional 147, Sierra del Gigante, Belgrano, San Luis</td>
</tr>
<tr>
<td></td>
<td>LJAMM 10854</td>
<td>30º 38' 20&quot; S, 67º 22' 38&quot; W</td>
<td>Ruta Provincial 511, 7.6 km E San Agustin del Valle Fertil, Valle Fertil, San Juan</td>
</tr>
<tr>
<td></td>
<td>5042</td>
<td>30º 32' 57&quot; S, 66º 55' 40&quot; W</td>
<td>Ruta Provincial 27, 41.6 Km N El Portezuelo; 6.6 Km N San Ramon, Gral. Juan F. Quiroga, La Rioja</td>
</tr>
<tr>
<td><em>L. crepuscularis</em></td>
<td>LJAMM 12635</td>
<td>27º 21' 58&quot; S, 66º 22' 25&quot; W</td>
<td>Puesto Lopez, Ruta Provincial 47, 2 km S Mina Capillitas, Andalgalá, Catamarca</td>
</tr>
<tr>
<td></td>
<td>LJAMM 12642</td>
<td>27º 21' 58&quot; S, 66º 22' 25&quot; W</td>
<td>Puesto Lopez, Ruta Provincial 47, 2 km S Mina Capillitas, Andalgalá, Catamarca</td>
</tr>
<tr>
<td></td>
<td>12644</td>
<td>28º 21' 58&quot; S, 66º 22' 25&quot; W</td>
<td>Puesto Lopez, Ruta Provincial 47, 2 km S Mina Capillitas, Andalgalá, Catamarca</td>
</tr>
<tr>
<td><em>L. darwinii</em></td>
<td>LJAMM 10582</td>
<td>37º 04' 29&quot; S, 67º 47' 07&quot; W</td>
<td>Ruta Provincial 16, 23.6 km W junction with Ruta Nacional 151, Puelén, La Pampa</td>
</tr>
<tr>
<td></td>
<td>LJAMM 11022</td>
<td>42º 46' 00&quot; S, 65º 03' 00&quot; W</td>
<td>Puerto Madryn, Biedma, Chubut</td>
</tr>
<tr>
<td></td>
<td>5104</td>
<td>36º 08' 19&quot; S, 68º 17' 23&quot; W</td>
<td>Ruta Provincial 190, 2 Km N Agua Escondida, Malargüe, Mendoza</td>
</tr>
<tr>
<td><em>L. espinozai</em></td>
<td>LJAMM 12666</td>
<td>27º 03' 19&quot; S, 66º 11' 51&quot; W</td>
<td>Ruta Provincial 47, 21 km S El Arenal, Santa Maria, Catamarca</td>
</tr>
<tr>
<td></td>
<td>LJAMM 12668</td>
<td>28º 03' 19&quot; S, 66º 11' 51&quot; W</td>
<td>Ruta Provincial 47, 21 km S El Arenal, Santa Maria, Catamarca</td>
</tr>
<tr>
<td></td>
<td>4338</td>
<td>27º 07' 30&quot; S, 66º 13' 03&quot; W</td>
<td>Ruta Provincial 47, 20 Km S Punta de Balasto, Campo Arenal, Santa María, Catamarca</td>
</tr>
<tr>
<td><em>L. grosseorum</em></td>
<td>LJAMM 4019</td>
<td>35º 17' 09&quot; S, 68º 41' 52&quot; W</td>
<td>Ruta Provincial 180, 30 Km S El Nihuil, San Rafael, Mendoza</td>
</tr>
<tr>
<td></td>
<td>LJAMM 4046</td>
<td>36º 37' 17&quot; S, 68º 36' 38&quot; W</td>
<td>Ruta Provincial 180, 28.1 Km S south entrance to La Matancilla, Malargüe, Mendoza</td>
</tr>
<tr>
<td></td>
<td>7825</td>
<td>38º 13' 49&quot; S, 68º 57' 36&quot; W</td>
<td>Ruta Provincial 8, 23 km N Añelo, Añelo, Neuquén</td>
</tr>
<tr>
<td><em>L. irregularis</em></td>
<td>LJAMM 12795</td>
<td>24º 13' 13&quot; S, 66º 16' 59&quot; W</td>
<td>Ruta Nacional 51, 6 km E San Antonio de los Cobres, San Antonio de los Cobres, Salta</td>
</tr>
<tr>
<td></td>
<td>LJAMM 12798</td>
<td>25º 13' 13&quot; S, 66º 16' 59&quot; W</td>
<td>Ruta Nacional 51, 6 km E San Antonio de los Cobres, San Antonio de los Cobres, Salta</td>
</tr>
<tr>
<td>Collection Code</td>
<td>Number</td>
<td>Latitude</td>
<td>Longitude</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------</td>
<td>------------</td>
<td>-----------</td>
</tr>
<tr>
<td>LJAMM 2629</td>
<td>18</td>
<td>24º 12’ 02” S, 66º 24’ 04” W</td>
<td></td>
</tr>
<tr>
<td><em>L. koslowskyi</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJAMM 4159</td>
<td>19</td>
<td>28º 32’ 09” S, 67º 22’ 21” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 4206</td>
<td>20</td>
<td>28º 50’ 19” S, 67º 24’ 47” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 5011</td>
<td>21</td>
<td>28º 11’ 35” S, 67º 08’ 03” W</td>
<td></td>
</tr>
<tr>
<td><em>L. laurenti</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJAMM 12735</td>
<td>23</td>
<td>24º 36’ 25” S, 66º 11’ 41” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 12812</td>
<td>24</td>
<td>25º 14’ 14” S, 65º 54’ 02” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 12815</td>
<td>24</td>
<td>26º 14’ 14” S, 65º 54’ 02” W</td>
<td></td>
</tr>
<tr>
<td><em>L. lavillai</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJAMM 10751</td>
<td>25</td>
<td>31º 14’ 20” S, 68º 39’ 04” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 10783</td>
<td>26</td>
<td>30º 13’ 43” S, 68º 19’ 35” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 10821</td>
<td>27</td>
<td>29º 41’ 17” S, 68º 01’ 41” W</td>
<td></td>
</tr>
<tr>
<td><em>L. ornatus</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJAMM 12019</td>
<td>28</td>
<td>23º 16’ 38” S, 65º 49’ 09” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 12021</td>
<td>28</td>
<td>23º 16’ 38” S, 65º 49’ 09” W</td>
<td></td>
</tr>
<tr>
<td><em>L. quilmes</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJAMM 12713</td>
<td>29</td>
<td>25º 36’ 53” S, 66º 11’ 43” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 4346</td>
<td>30</td>
<td>25º 42’ 19” S, 66º 07’ 42” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 4404</td>
<td>31</td>
<td>25º 52’ 44” S, 65º 56’ 38” W</td>
<td></td>
</tr>
<tr>
<td><em>L. uspallatensis</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJAMM 12500</td>
<td>32</td>
<td>32º 23’ 26” S, 69º 23’ 20” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 12506</td>
<td>33</td>
<td>31º 25’ 06” S, 69º 26’ 31” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 12630</td>
<td>34</td>
<td>30º 52’ 43” S, 69º 25’ 12” W</td>
<td></td>
</tr>
<tr>
<td><em>L. boulengeri</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJAMM 2187</td>
<td></td>
<td>45º 27’ 22” S, 69º 45’ 56” W</td>
<td></td>
</tr>
<tr>
<td><em>L. capillitas</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJAMM 2788</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>L. cf. montaus</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJAMM 12020</td>
<td>23</td>
<td>23º 16’ 38” S, 65º 49’ 09” W</td>
<td></td>
</tr>
<tr>
<td>LJAMM 2163</td>
<td></td>
<td>47º 46’ 45” S, 70º 37’ 43” W</td>
<td></td>
</tr>
<tr>
<td><em>L. telsen</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJAMM 5530</td>
<td>42° 22' 04'' S, 67° 39' 22'' W</td>
<td>Ruta Provincial 4, 65.5 Km W Telsen, Telsen, Chubut</td>
<td></td>
</tr>
</tbody>
</table>
Appendix II

Figure IIa. Regression between % of variable sites and % of informative sites for all 19 nuclear loci. The mtDNA cyt b gene was excluded because it is an outlier.

Figure IIb. Regression between % of variable sites and resolution (clade support index) for all 19 nuclear loci. The mtDNA cyt b gene was excluded because it is an outlier.
Figure IIc. Species tree of the *Liolaemus darwinii* group estimated without an outgroup. Numbers on branches represent posterior probabilities. The scale bar is in units of substitutions per site.

Figure IId. Regression between % of variable sites and discordance (number of deep coalescences / number of gene copies) for all 19 nuclear loci. The mtDNA cyt b gene was excluded because it is an outlier.
Figure IIe. Regression between precision of the posterior distribution of species trees and the clade support index (proportion of highly supported clades) for analyses with different number of loci.
Appendix III. Gene trees of all 20 loci included in species trees analyses estimated with *BEAST. Loci are described in Table 1.
Modified from SpeciesTree-True

A9C

A9E
Figure Va. Species tree of the *Liolaemus darwinii* group based on 16 loci.
Figure Vb. Species tree of the *Liolaemus darwinii* group based on 12 loci.
Figure Vc. Species tree of the *Liolaemus darwinii* group based on 8 loci.
Figure Vd Species tree of the *Liolaemus darwinii* group based on 4 loci.
SPECIES DELIMITATION USING ABC: ACCOUNTING FOR SPECIATION WITH GENE FLOW IN LIZARDS OF THE *LIOLAEMUS DARWINII* COMPLEX (SQUAMATA: TROPIDURIDAE)

Arley Camargo\textsuperscript{1*}, Mariana Morando\textsuperscript{2}, Luciano J. Avila\textsuperscript{2} & Jack W. Sites, Jr.\textsuperscript{1}

\textsuperscript{1}Department of Biology and Bean Life Science Museum, Brigham Young University, Provo, UT 84602, USA

\textsuperscript{2}CONICET-CENPAT, Boulevard Almirante Brown 2825, U9120ACF, Puerto Madryn, Chubut, Argentina

Key words: species delimitation, Approximate Bayesian Computation, coalescent model, speciation, gene flow

Corresponding author: Arley Camargo, Department of Biology, Brigham Young University, Provo, Utah, USA 84602; phone: 801/422-2203; fax: 801/422-0090; email: arley.camargo@gmail.com
Abstract.—The practice of species delimitation is a major research focus in evolutionary biology because the accurate assessment of species boundaries is a prerequisite for the study of speciation, the ultimate process responsible for biodiversity. New species delimitation methods (SDMs) can accommodate non-monophyletic species and gene tree discordance as a result of incomplete lineage sorting via the coalescent model, but do not explicitly accommodate for gene flow after divergence. One way of incorporating gene flow into species delimitation is via Approximate Bayesian Computation (ABC) to estimate the relevant parameters of the speciation process and to test alternative species delimitation hypotheses. We evaluate the performance of an ABC approach for delimiting species using simulated data, apply the method to lizards of the L. darwinii complex, and compare the results with those obtained with likelihood-based methods. We find that while gene flow impacted the accuracy of the ABC method, the speciation model can still be correctly inferred when migration rates are low and despite biased estimates of demographic parameters. Both likelihood-based methods and ABC consistently supported the distinctness of southern and northern lineages within L. darwinii. Further simulation studies are necessary to evaluate the performance of ABC, likelihood-based, and other SDMs using genetic data derived from speciation models that differ in a number of demographic parameters. The ABC framework represents an appropriate solution to the problem of species delimitation, especially in the face of speciation with gene flow, and contributes toward a unified approach that can simultaneously estimate species limits, species trees, and demographic parameters.
The practice of species delimitation is a major research focus in evolutionary biology because the accurate assessment of species boundaries is a prerequisite for the study of speciation, the ultimate process responsible for biodiversity. While progress has been made in separating the species concepts (de Queiroz 2007) from the criteria used to delimit species boundaries in nature (Sites and Marshall 2003, 2004), the issue of delimiting species in practice has in comparison received little attention (Wiens 2007). However, species delimitation has been a growing topic in the literature especially during the last five years based on the results of database searches (Appendix I). During this period, several new criteria have been introduced based on gene trees that apply: inference keys with assessments of gene flow (Wiens and Penkrot 2002), measures of lineage exclusivity (Cummings et al. 2008), a statistical fit of the threshold between inter- and intra-specific divergence (Pons et al. 2006), and more recently, optimization approaches for minimizing gene tree discordance across species limits (O'Meara 2010). The practice of species delimitation with molecular data is expanding rapidly due in part to the extension of coalescent models to the interspecific level with the multispecies or 'censored' coalescent (Rannala and Yang 2003; Degnan and Rosenberg 2009), and the recent application of these new conceptual and methodological frameworks in phylogenetic systematics for inference of species relationships (Edwards 2009). Many species trees approaches are based on the a priori assignment of individual gene copies to population-level lineages assuming that boundaries between species are known with certainty, but in reality, estimation of species trees should ideally be performed simultaneously with the estimation of species limits (Carstens and Dewey 2010).
Most coalescent-based species delimitation methods (SDMs) can accommodate non-monophyletic species and gene tree discordance as a result of incomplete lineage sorting (ILS) via the coalescent model (Knowles and Carstens 2007). In this context, large ancestral population sizes and shallow divergence times are expected to increase levels of ILS and consequently the paraphyly of species and the topological discordance of gene trees (Funk and Omland 2003). One approach consists of hypothesizing a species tree (topology and branch lengths) and species boundaries (i.e., assignment of individual samples to species), and then calculating the probability of all gene trees under that species history. Subsequently, likelihoods are calculated for nested species histories (as when two species are collapsed into one) and a likelihood ratio statistic is used to test if the alternative, simpler species history fits the data better than the null hypothesis (Knowles and Carstens 2007). However, when models of species limits are non-nested, a more appropriate approach uses Akaike information criteria to test alternatives (Carstens and Dewey 2010). This method is implemented in the java pipeline SpeDeSTEM 0.1.1 (Ence and Carstens 2010) that is based in the program STEM 1.0 for estimating a species tree from a collection of gene trees known without error (point estimates) and assuming that population sizes have remained constant along the species tree (Kubatko et al. 2009). In this approach, a molecular clock is enforced on estimated gene trees, the mutation rate parameter \( \theta \) is estimated from the data, and species trees are re-estimated for each species delimitation hypothesis.

An alternative Bayesian approach consists of sampling from the posterior distribution of models of species limits using reversible-jump Markov chain Monte Carlo (rjMCMC) in order to move across models of different dimensionality (i.e., number of parameters) as implemented in the program BPP 2.0 (Yang and Rannala 2010). This approach uses a fixed, fully resolved guide
tree of lineages hypothesized to represent separate species, which is used to construct models of species limits by sequentially collapsing internal nodes until all nodes are collapsed into a single species. Priors should be given for population sizes and divergence times of the species tree in addition to the priors and proposal mechanisms of the regular MCMC chains for estimating the gene trees. Therefore, in this approach the uncertainty in the gene trees is explicitly incorporated in the models, theta is estimated and allowed to vary along the species tree, but the species tree is provided with a fixed topology.

Coalescent methods used in current species delimitation approaches do not explicitly accommodate for gene flow after divergence (Ence and Carstens 2010; Yang and Rannala 2010). However, speciation with limited gene flow appears to be common in nature (Nosil 2008; Pinho and Hey 2010) and therefore species delimitation should also take into account the process of divergence with gene flow in addition to complete isolation associated with allopatric models (Hey 2009, 2010). Most often, disruptive selection is the main cause of divergence in spite of gene flow (Pinho and Hey 2010), but also demographic history, in particular some levels of intraspecific gene flow, might also play a role (Zhou et al. 2010). If two species have diverged with occasional gene flow, SDMs accomodating only ILS are expected to collapse these species into a single lineage due to the homogenizing influence of gene flow. Alternatively, if species are not collapsed, the effect of gene flow could instead lead to underestimates of divergence times between species (Nielsen and Wakeley 2001). However, it is possible that even if gene flow is not accounted for, some SDMs may still be robust to the impact of gene flow and correctly separate species (Ence and Carstens 2010). Regardless, when multiple species are involved, an ideal approach to species delimitation would be to jointly estimate both the parameters that influence lineage divergence (population size, divergence time, gene flow, etc.) and the topology.
of the species trees, which are known to affect the outcome of species delimitation (Leaché and Fujita 2010).

One way of incorporating gene flow into species delimitation is via Approximate Bayesian Computation (ABC) methods. The use of ABC techniques started in the field of population genetics in 1997 but they have recently become a popular analytical tool for model choice and parameter estimation also in phylogeography, ecology, epidemiology, and phylogenetics (Beaumont 2010; Bertorelle et al. 2010; Csilléry et al. 2010a). A search in ISI Web of Science databases shows that the topic accumulated 112 references until 2008, but this accelerated rapidly since 2009 and exceeded 200 references in 2010 (Appendix I). ABC actually represents a group of likelihood-free algorithms that in its most basic formulation consists of: (1) sampling parameter values from prior distributions to generate simulated data; (2) calculating summary statistics (SuSt) from simulated and observed data and the Euclidean distance between them; and (3) approximating the posterior distribution of parameters with a rejection algorithm that retains those simulations that have an Euclidean distance smaller than a pre-specified threshold or tolerance (Lopes and Beaumont 2010). This procedure represents the original ABC formulation known as rejection-ABC, but step (3) has been modified to include a weighted local linear regression to correct the discrepancy between observed and simulated SuSt for increased performance (regression-ABC)(Beaumont et al. 2002).

In addition to parameters, different demographic models combining migration rates, divergence times, and population sizes can be compared statistically to select models based on posterior probabilities and/or Bayes factors. For example, in the context of rejection-ABC, the frequency of retained simulations generated under one model relative to all retained simulations represents its posterior probability, given that all models have the same prior number of
simulations (Pritchard et al. 1999). An improved and more accurate estimator of model probabilities includes an adjustment using weighted multinomial logistic regression (Beaumont 2008). Recently, a machine learning approach based on non-linear neural networks regression has been introduced for parameter estimation and model choice that relaxes assumptions and outperforms linear regression-ABC (Blum and François 2010). A complete ABC analysis requires not only simulating and estimating parameters of models, but also validating and testing the performance of the selected model using prior and posterior predictive tests (Bertorelle et al. 2010; Csilléry et al. 2010a). Despite some recent criticisms about incoherent and illogical model inference (Templeton 2009, 2010a, b), the statistical foundation of the ABC within a Bayesian framework has been shown to be solid (Beaumont et al. 2010; Berger et al. 2010; Csilléry et al. 2010b). Therefore, ABC seems an appropriate technique for estimation of the relevant parameters of the speciation process and for testing alternative models of species delimitation, while simultaneously handling large multilocus datasets due to the use of summary statistics and the likelihood-free nature of the approach (Beaumont 2010).

We use methods of species delimitation in lizards of the *Liolaemus darwinii* complex (Squamata, Tropiduridae) that occupy sandy habitats in the southern and central portions of the Monte Desert in Argentina (Etheridge 1993). The complex includes *L. darwinii*, *L. grosseorum*, and *L. laurentii*, which form a clade within the more inclusive *L. darwinii* group (Camargo et al. submitted). Several species in the group were formerly included within a single widespread species, *L. darwinii*, ranging across the whole Monte Desert from the Atlantic coast of Chubut Province to central Salta Province (Etheridge 1993). Detailed morphological studies revealed that *L. laurentii* from La Rioja and Catamarca Provinces, and *L. grosseorum* from Neuquen and Mendoza Provinces, were distinct species from *L. darwinii* based on diagnostic meristic
characters (scale and precloacal pores), male color patterns, and tail and body proportions (Etheridge 1992, 2001). The remaining geographic distribution of *L. darwinii* has been partitioned into northern (*L. darwinii*-N) and southern (*L. darwinii*-S) populations based on an apparent distributional gap in south-central Mendoza Province, scale count/color variation, and genetic differentiation (Etheridge 2001; Morando et al. 2004; Abdala 2007). For example, Etheridge (2001) found larger, darker dorsal spots and ventral patches in *L. darwinii*-N samples, and a mtDNA study (Morando et al. 2004), recovered a single southern (= *L. darwinii*-S) and several northern (N1 and N2 = *L. darwinii*-N) clades, which were interpreted as candidate species. Based on paraphyletic patterns in the mtDNA gene tree, geographic distributions, and coalescent expectations, incomplete lineage sorting and/or gene flow with introgression was inferred to occur between *L. darwinii*-N vs. *L. laurenti*, and between *L. darwinii*-S and *L. grosseorum* (Morando et al. 2004). Subsequently, Abdala (2007) assigned *L. darwinii*-N and two lineages of *L. darwinii*-S to different terminals in his phylogenetic analyses, and concluded that they probably represent different species pending more detailed analyses. These studies also found that a morphologically distinct and probably parthenogenetic form appears nested among *L. darwinii*-S and *L. darwinii*-N (Morando et al. 2004; Abdala 2007).

While mtDNA gene trees have suggested a closer relationship between *L. laurenti* and *L. darwinii*-N, species trees based on multi-locus datasets recovered a sister relationship between *L. laurenti* and *L. grosseorum* (Camargo et al. submitted). Based on these relationships, two different diversification patterns appear to have occurred in this complex: (1) a morphologically-divergent species pair (*L. laurenti* vs. *L. grosseorum*) with fully allopatric distributions and likely associated with a speciation-in-isolation model; and (2) a morphologically-more conserved pair of lineages (*L. darwinii*-N vs. -S) with nearly parapatric distributions (maybe in contact in the
past), possibly a result of divergence with gene flow, that have been identified as candidate species in previous studies. Therefore, these taxon-pairs represent appropriate empirical examples of two different evolutionary histories for analysis with species delimitation approaches. Herein, we introduce and evaluate the performance of an ABC approach for delimiting species using simulated data, apply the method to empirical data of the *L. darwinii* complex, and compare the results with those obtained with other likelihood-based methods using the same datasets. We excluded from these analyses the parthenogen form because its possible hybrid parentage (M. Morando, pers. com.) cannot be handled by the SDMs used in this study that assume strictly bifurcating species trees.

**Methods**

*Simulation testing*

*Model parameterization.*—We simulated data for the no-speciation (model A) and the speciation scenarios (model B). Model A included the per locus mutation parameter \( \theta (= 4N_o \mu) \) for a single lineage while model B included two additional parameters: the divergence time \( \tau \) (in units of \( 4N_o \) generations) and the migration rate between the divergent lineages \( m \) (the proportion of gene copies replaced by immigrant gene copies each generation). For each model, we simulated 100,000 coalescent genealogies with the program msABC (Pavlidis et al. 2010), a modification of the ms program (Hudson 2002) for simulating multiple loci, including prior distributions of parameters, and calculating a large array of SuSt for ABC analysis. Number of individuals, number of loci, sequence length, and prior distributions were chosen as to include
the observed parameter values in the empirical data sets (see below). We simulated 4 loci for 40 individuals (20 for each population in model B) with uniform priors of \( \theta (0.1–20) \), \( \tau (0.01–10) \), and \( m (0–m_{\text{max}}) \). Because increasing migration is expected to make more difficult the distinction between models A and B, we compared the performance of the ABC method for four different values of \( m_{\text{max}} \): 0 (no migration, divergence in isolation), 0.25 (low migration), 0.50 (moderate migration), and 1 (high migration). All 16 global SuSt were included in the simulations (Table 1), which represents a reasonable number of SuSt in ABC studies (Bertorelle et al. 2010) but we excluded the population-specific SuSt because they cannot be compared between models with different numbers of populations. Simulations for each model were concatenated into a single prior file including a binary index parameter that identifies the model under which the simulations were generated (e.g., 0 = model A, 1 = model B).

Performance.—To evaluate the performance of ABC to distinguish between speciation vs. no-speciation models, we input the prior file into the R package 'abc' (Csilléry et al. 2010c) (http://cran.r-project.org/web/packages/abc/index.html) to run the model selection function 'postpr'. In R we performed the leave-one-out approach that sequentially uses one simulation as a pseudo-observed dataset (pods) and the rest of simulations as the prior distribution of parameters and SuSt. This procedure was repeated with 100 pods to evaluate error rates in model selection using three different ABC algorithms available in 'abc': simple rejection, multinomial logistic, and neural-network. We followed a similar testing procedure to evaluate accuracy of parameter estimates with the function 'abc'. We calculated the bias in parameter estimates as the difference between the real parameter value and the mean estimated value based on the posterior distribution. We also calculated the relative bias with respect to the range used in the prior
distributions and parameter coverage, which is the percentage of simulations where the true value falls within the 95%-highest posterior density (HPD). Tolerance was set to 0.0005 to retain 100 accepted simulations for estimating posterior distributions. Logit regression was used to transform and to insure that parameter estimates were within the prior bounds used in simulations.

To compare the performance of the ABC method to those of likelihood-based methods (SpeDeSTEM and BPP), we also simulated model A and B for 5 loci and 5 sequences per lineage (10 individuals in model A) and used 100 pods to estimate model probabilities for three different conditions: (1) \( m = 0, \tau = 0.2N_o \) generations; (2) \( m = 0, \tau = 0.5N_o \) generations; and (3) \( m = 0.1 \) and \( \tau = 0.5N_o \) generations. The parameter values in (1) represent the conditions under which BPP has an accuracy of ~60% (see Fig. 2b in Yang and Rannala 2010), while (2) and (3) are two conditions under which SpeDeSTEM has been shown to successfully detect speciation (Ence and Carstens 2010) given that 5 loci and 5 individuals per population are sampled.

**Empirical analyses**

*Field sampling.*—We sampled 398 individuals of *L. darwinii* from 134 localities including the distribution of southern and northern lineages. We also sampled 69 individuals of *L. grosseorum* from 20 localities and 38 individuals of *L. laurenti* from 8 localities (Fig. 1, Appendix II). Tissue samples from liver and/or tail muscle were preserved in absolute ethanol and stored at -20 °C. Specimens were fixed in 10-20% formalin, later transferred to 70% ethanol, and deposited in the herpetological collection of L. J. Avila and M. Morando (LJAMM) of the
Centro Nacional Patagónico (CENPAT–CONICET) and the herpetological collections of Monte L. Bean Museum, Brigham Young University (BYU).

**Sequence data.**—Genomic DNA was extracted with the DNAeasy Qiagen kit (Qiagen). We used the Green Go Taq PCR kit (Promega) for all PCR reactions in PTC-200 DNA Engine (MJ Research) or GeneAmp PCR 9700 thermal cyclers (Applied Biosystems, Inc.). Sequencing reactions used the Big-Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Inc.) in a GeneAmp PCR 9700 thermal cycler (Applied Biosystems, Inc.). Sequencing products were cleaned with Sephadex G-50 Fine (GE Healthcare Bio-Sciences AB) and sequenced in an ABI 3730xl DNA Analyzer (Applied Biosystems, Inc.). We sequenced the cytochrome b (cyt b) mtDNA gene for all available individuals of *Liolaemus darwini* (including north and south lineages), *L. grosseorum*, and *L. laurenti* (~500 individuals) following protocols in Morando et al. (2004). We subsampled 10 individuals across the geographic distribution of each lineage representing the variation found in mtDNA haplotypes for screening three Anonymous Nuclear Loci (ANL) (Appendix II). We sequenced 3 ANL (A1D, A9C, and B6B) that were developed from the genomic DNA of a *L. darwinii* individual (LJAMM 7097) following protocols in Noonan & Yoder (2009). We generated ~200 random fragments, cloned, sequenced, and BLAST searched these to confirm they were anonymous. We then designed primers for fragments with confirmed anonymity and used the PCR temperature profile of Noonan & Yoder (2009) to amplify ANL in all sampled individuals.

Chromatograms were checked by eye and ambiguity codes were used to represent polymorphisms of heterozygote individuals in Sequencher 4.7 (Gene Codes Corporation). Gametic phase of heterozygote individuals were resolved with the program Phase 2.1.1.
(Stephens et al. 2001). Sequences were aligned with ClustalX 2.0.10 (Larkin et al. 2007). Aligned ANL sequences were inspected by eye to check that there were no fixed heterozygotes at any polymorphic site to insure we were not using multiple-copy markers (Thomson et al. 2010). Alignments were tested for recombination with the program RDP3 beta35 (Martin et al. 2005). For estimation of a species tree, we selected 3 individuals to represent each lineage from localities distant from haploclade boundaries or contact zones to minimize the potential impact of intermixed/migrant individuals (Leaché 2009). We included also L. olongasta, a closely related species to the focal clade in this study, and L. boulengeri, an outgroup of the L. darwinii group (Appendix II, Camargo et al. submitted).

Species delimitation with ABC.—Based on the aligned sequence data, we calculated the same 16 SuSt for all 4 loci using first, the 'fas2ms' perl script to transform base pairs to polymorphic sites, and then msABC with the option '--obs'. Observed SuSt were calculated for two paired combinations of lineages: L. darwinii-S vs. L. darwinii-N (dS-dN) and L. laurenti vs. L. grosseorum (L1-Lg). We assessed how well the simulated models fit the observed data via a Principal Components Analysis (PCA) of simulated prior SuSt and observed SuSt. A good fit in this prior predictive plot was interpreted when the observed SuSt occurred within the cloud of simulated SuSt. The observed SS and simulated prior SuSt were analyzed with 'postpr' to estimate posterior probabilities and statistical support of speciation vs. no-speciation models, and also with 'abc' to estimate posterior distributions of θ (both models), τ, and m (speciation model only). We also evaluated how well the selected model and estimated parameters predicted the observed data via simulating data with the estimated 95%-HPD of parameters and subsequently, running a PCA analysis of the simulated posterior SuSt and observed SuSt. As above, we
concluded that the selected model was a good fit of the data when the observed SuSt in this posterior predictive plot was found within the cloud of simulated SuSt defined by the PCA axes.

Species delimitation with other methods.—We analyzed the empirical dataset with SpeDeSTEM 0.9.4 (Ence and Carstens 2010) and BPP 2.0 (Yang and Rannala) for comparison with the ABC results. SpeDeSTEM is a java-based pipeline that uses gene trees obtained with PAUP* to estimate maximum likelihood species trees with the program STEM (Kubatko et al. 2009) for alternative models of species limits that are evaluated with Akaike's information criteria (Ence and Carstens 2010). Best-fit substitution models for the species delimitation dataset were estimated with jModelTest 0.1.1 from the pool of 88 competing models using the Bayesian information criterion (Posada 2008). Relative mutation rates of loci and average theta across lineages were calculated with Migrate-n 3.2.1 (Beerli and Palczewski 2010), using two independent maximum-likelihood runs each consisting of 10 short chains and 5 long chains sampled every 50 steps for a total of 2,000 and 30,000 generations, respectively, and a burn-in period of 20,000 steps. The relative mutation rate of cyt b was divided by two to account for the haploid status of this locus (Kubatko et al. 2009). In SpeDeSTEM, we randomly subsampled 5 sequences from each lineage in 50 replicates following the manual's recommendations (Ence and Carstens 2010) and tested species limits for the sister-lineage pairs *L. darwinii*-N vs. *L. darwinii*-S and *L. laurenti* vs. *L. grosseorum*.

We estimated a species tree with BEAST v1.6.1 (Drummond and Rambaut 2007; *BEAST, Heled and Drummond 2010) to be used as a guide tree in BPP analyses. Analyses were run for 40 million generations and samples taken every 4,000 generations with the same prior distributions and model settings used in a recent study of the complete *L. darwinii* group.
(Camargo et al. submitted). Log files were inspected in Tracer v1.5 (Rambaut and Drummond 2007) to determine an appropriate burn-in sample to estimate the posterior distribution of species trees. In BPP, we analyzed 20 sequences per lineage and per locus (except L. laurenti, 16 sequences for cyt b) with both algorithm 0 (ε = 5 and 10) and 1 (α = 2, m = 1 and α = 1, m = 2) used in the rjMCMC moves between alternative models of species delimitation. In both cases, we varied the parameters α and β of the gamma-distributed priors for θ and τ to take into account a range of speciation histories: large population size/deep divergence (both priors with α = 2 and β = 2000), small population size/shallow divergence (both priors with α = 1 and β = 10), and large population size/shallow divergence (α = 2 and β = 2000 for θ prior; α = 1 and β = 10 for τ prior) (Leaché and Fujita 2010). Fine-tuned parameters for the regular MCMC chains were adjusted in preliminary runs to achieve acceptance rates between 0.3–0.4 as recommended in the program's manual. We used the same relative rates per locus as specified in the SpeDeSTEM analyses. All runs consisted in 50,000 samples taken every 5 steps with a burn-in period of 10,000 steps.

RESULTS

Simulation testing

The performance of the ABC method was very high (> 95% accuracy) when the single species model was compared with the speciation-in-isolation model (m = 0) for all three ABC algorithms, but performance decreased with increasing migration in speciation-with-gene flow models (Fig. 2). The neural network outperformed other algorithms when the single-species
model was correct and accuracy ranged between 89% ($m_{\text{max}} = 0.25$) and 76% ($m_{\text{max}} = 1.00$) (Fig. 2a). In general, performance was lower when the speciation model was correct and both neural network and multinomial logistic outperformed the simple rejection algorithm with accuracy ranging between 82–83% ($m_{\text{max}} = 0.25$) and 75–76% ($m_{\text{max}} = 1$), respectively (Fig. 2b). Accuracy in parameter estimates measured as relative bias also decreased with increasing levels of migration with the highest accuracy for $\theta$ and the lowest for migration; when $m = 0$, $\tau$ was estimated more accurately than $\theta$ (Table 2). Likewise, accuracy measured as 95%-HPD coverage of the true parameter value also decreased with increasing migration especially in the case of $\tau$ that went from 100% to 20% when migration was included in the analysis (Table 2).

Accuracy of ABC species delimitation, measured as mean model probabilities for the correct model, when simulating 5 loci and 10 sequences total (5 per lineage) was 35% when $m = 0$, $\tau = 0.2N_o$ (compared to ~60% for BPP). Accuracy was 50% for $m = 0$ and $\tau = 0.5N_o$, and 44% when $m = 0.1$ and $\tau = 0.5N_o$, suggesting that the ABC method (for the specified priors and number of simulations) cannot detect speciation at these shallow levels of divergence time with migration between species while SpeDeSTEM successfully delimited species under these conditions.

**Empirical analyses**

We sequenced 713 bp for cyt $b$, 692 for A1D, 481 for A9C, and 415 bp for B6B. Up to 20 sequences were sampled per lineage and per locus, but 16 cyt $b$ sequences were obtained for *L. laurenti*, and 12 A1D and 14 B6B sequences were included for *L. darwinii*-S. No gaps were found in the multiple sequence alignments in any locus. We did not find any fixed heterozygotes
at any polymorphic site suggesting that ANL were single-copy markers (Thomson et al. 2010).

We did not find evidence of recombination in any of the loci analyzed in this study.

Species delimitation with ABC.—Based on the higher performance of the neural-net algorithm in the simulation testing, we used this procedure to analyze the empirical data. The ABC model selection for the pair *L. darwinii*-N vs. *L. darwinii*-S resulted in high support for the speciation model with migration ($m_{\text{max}} = 0.25$) with posterior probability of 0.967 (Bayes factor ~ 29). In addition, a similar procedure for the pair *L. laurenti* vs. *L. grosseorum* yielded high support for the speciation model with isolation with a posterior probability of 0.914 (Bayes factor ~ 11). The maximum number of iterations was set to 5,000 to reach convergence in both analyses. Parameter estimates for selected models in ABC analyses are given in Table 3. PCA analyses show that observed SS are within the bounds of the prior sample of SS, suggesting good model fitting (Appendix III). However, observed SS were outliers beyond the bounds of the simulated SS using the posterior parameter estimates (Appendix III).

Species delimitation with other methods.—The relative mutation rates for loci were: 1.75 (cyt b), 0.83 (A1D), 1.00 (A9C), and 0.42 (B6B). Mean $\theta$ per site across the four lineages analyzed was 0.0072. Based on the substitution models for each locus, relative mutation rates, and mean $\theta$ per site, SpeDeSTEM analyses selected the model with all four lineages as separate species with strong AIC support (Table 4). This model had an Akaike weight of ~0.99, indicating that it has a 99% chance of being the best model among the four alternatives analyzed (Table 4). The model collapsing *L. darwinii*-N and -S had very low support, while other models had insignificant model likelihoods (Table 4).
The species tree recovered the pair *L. darwinii*-N and *L. darwinii*-S as a strongly supported clade, the pair *L. laurenti* and *L. grosseorum* received moderate support, and these two species pairs were recovered as sister clades with strong support (Fig. 3). Based on this guide tree and the estimated relative mutation rates, BPP consistently found very high speciation probabilities (1.0) for all internal nodes across multiple analyses with different algorithms and prior distributions.

**DISCUSSION**

*Species delimitation with ABC and other methods*

ABC is a powerful and flexible approach for model choice and estimating model parameters including for example the number of populations (Bertorelle et al. 2010). Based on the capabilities of ABC and the recent availability of software for model simulation and model choice with novel summary statistics (i.e., msABC), and more sophisticated algorithms (non-linear regression-ABC), we applied this approach to species delimitation. Specifically, we tested the sensitivity of this ABC method to the influence of altered parameter values, especially the migration rate parameter. We find that while migration impacted the accuracy of the ABC method, its performance is still high when migration rates varied between 0 and 0.25, with the upper bound representing the case when one-fourth of the individuals in one population are replaced with migrants from another population. Further, the speciation model is correctly inferred based on posterior probabilities and Bayes factors even when estimates of demographic parameters are biased by increasing migration rates. This result is congruent with a recent ABC-
based study, which was able to distinguish statistically between two alternative demographic models in spite of imprecise parameter estimates (Carstens and Knowles 2010; Peter et al. 2010).

The differences in performance with different ABC algorithms using the simulation testing procedure are similar to those found previously in the context of model choice. In particular, logistic regression has been found to outperform simple rejection and non-linear regression in turns outperforms linear regression (Beaumont 2008; Bertorelle et al. 2010; Blum and François 2010). We found this same pattern in our analyses: logistic and neural networks always outperformed simple rejection, and neural network outperformed logistic regression when the no-speciation model was true (Fig. 2a). While ABC approaches are demonstrating the benefits of simultaneous model choice and parameter estimation in phylogeographic inference (Carstens and Knowles 2010), our results also suggest that the model-based ABC framework represents an appropriate solution to the problem of species delimitation, especially in the face of divergence with gene flow.

**Improving ABC.**—The flexibility of the ABC framework provides many opportunities to improve the accuracy of the method for both detecting the correct species delimitation model, and for parameter estimation conditional on the selected model. Although the prior predictive test indicates that the simulated models are a plausible explanation of the observed SuSt, their poor fit with the posterior predictive test suggests that parameter estimates are biased or that the selected model is inappropriate (Bertorelle et al. 2010). Biased parameter estimates can be improved with increased number of simulations since ABC studies suggest that more simulations are required for accurate parameter estimation than for model choice (Carstens and Knowles 2010). In addition, a reduced number of SuSt could be selected to improve accuracy and at the
same time reduce the dimensionality of SuSt, which might impact the efficiency of ABC techniques (Beaumont et al. 2002; Csilléry et al. 2010a).

ABC methods can accommodate highly parameterized models and to exploit this ability to its maximum potential, more realistic models can be conceived and compared with simpler versions (Bertorelle et al. 2010). For example, population size could be allowed to vary between and within lineages to account for population expansion or contraction, while models including population structure can also be formulated for correctly testing between stable or varying population sizes (Peter et al. 2010). The inclusion of population substructure within lineages in these models will also help to evaluate the impact of intraspecific gene flow in maintaining species distinctness in spite of interspecific gene flow (Zhou et al. 2010). In addition, while we only compared taxon pairs, the ABC models can be extended to perform species delimitation and parameter estimation for multiple lineages in a single analysis including the topology of the species tree as an additional parameter (Fan & Kubatko, submitted) (Beaumont 2008). However, recent simulations have demonstrated that model-based species delimitation for more than two lineages requires more data (almost the double of the number of loci) to achieve the same levels of accuracy as when delimiting only two species (Ence and Carstens 2010).

Improvements of the ABC procedure might also include using priors different from the uniform distributions used in this study. Based on empirical patterns of variation in nature, the migration rate and the divergence time parameters are frequently sampled from exponential distributions while mutation rate parameters are often drawn from log-uniform distributions (Bertorelle et al. 2010). In addition, other kinds of SuSt can be used that are more sensitive to detecting the relative contributions of migration vs. isolation in generating the observed genetic data, such as the variance of pairwise sequence differences (Wakeley 1996; Nielsen and Wakeley
2001). This SuSt might capture enough signal in the distribution of genetic variation between lineages to discriminate between speciation in isolation vs. with gene flow, and to obtain more accurate estimates of the migration rate parameter. This latter parameter was the most biased in our study, and may have contributed to the poor fit between the selected speciation model and the data. Our primary goal was to show the ability of ABC to distinguish between speciation and no-speciation models with the by-product of parameter estimates. However, after selecting the best model, parameters can be re-estimated with other available methods and software in ABC (DIY-ABC, Cornuet et al. 2008; popABC, Lopes et al. 2009) and likelihood-frameworks (IMa, Hey and Nielsen 2007; 3s, Yang 2010) to test between speciation models with and without the gene flow.

Comparison with other methods.—In spite of the assumptions of the likelihood-based (constant and known $\theta$) and Bayesian (known species tree) methods, both of these and the ABC method consistently detected separate species within both taxon pairs analyzed. While it is expected that full-likelihood methods could have difficulty in separating species in the case of speciation with gene flow (LdS vs. LdN taxon pair), they detected distinct species with high posterior probability (BPP) and/or model support (SpeDeSTEM). These results suggest that these methods might be robust to the effects of limited gene flow, that our study system represents an easy delimitation problem, or both. For example, long divergence times combined with little or no gene flow and small population sizes lead to fully sorted ancestral polymorphisms within populations and congruent gene trees, facilitating species delimitation (Ence and Carstens 2010). In contrast, shallow divergence times, large population sizes and post-divergence gene flow make species delimitation more challenging due to extensive gene tree conflict (incomplete
lineage sorting) and poorly resolved gene trees (reduced genetic variation) (Yang and Rannala 2010).

Previous simulations with the likelihood-based SpeDeSTEM method show that two species that have diverged as recently as 0.5\(N_o\) generations ago can be distinguished as separate lineages, even under moderate gene flow \((m = 0.1)\), when 5 loci and 5 sequences per species are sampled (Ence and Carstens 2010). Simulations with BPP found that the method cannot distinguish separate species when they diverged up to 0.2\(N_o\) generations ago without gene flow for the same sampling conditions (mean speciation probability was \(\sim 0.6\)) (Yang and Rannala 2010). Our simulations under similar sampling designs revealed a poorer performance of ABC compared to these likelihood-based methods but these are preliminary comparisons because SpeDeSTEM calculates model likelihoods (instead of model probabilities as BPP and ABC do) and for this method, simulations included coalescent and mutational variance [only coalescent variance for BPP in Yang and Rannala (2010) and ABC in this study].

Further evaluation is required to examine a range of parameter values for divergence times, migration rates, and sampling designs, including models with varying population size (which is estimated by BPP and ABC but assumed to remain constant in SpeDeSTEM) for an adequate evaluation of performance across these methods. It should be expected that with increasing post-divergence gene flow and shallower divergence times, these methods would tend to perform poorer but they would be able to detect separate species under these strict conditions when more loci (and/or possibly more sequences) are sampled (Ence and Carstens 2010). In addition to these coalescent-based methods, the evaluation can also include other methods that have been introduced and used empirically with genetic data including: clustering methods (Pritchard et al. 2000; Huelsenbeck and Andolfatto 2007; Hausdorf and Hennig 2010), networks.
(Chen et al. 2010; Flot et al. 2010), a mixed Yule-coalescent model (Pons et al. 2006), and non- and semi-parametric approaches (O'Meara 2010).

Species limits in the L. darwinii complex

Species delimitation analyses with three different coalescent-based methods support the distinctness of the northern and southern lineages of L. darwinii. Previously, Etheridge (2001) found morphological differences in male color patterns between these same populations suggesting species-level distinctness and a geographic barrier between these forms in south-central Mendoza Province. Based on a phylogeographic analysis of the cyt b gene, Morando et al. (2004) found that southern populations form a single, well supported clade with shallow divergences, while northern populations did not comprise a single clade, because the morphologically distinct L. laurenti was nested among the L. darwinii-N terminals. If true, then the northern and southern lineages of L. darwinii are not sister taxa. More recently, Abdala (2007) presented a combined, molecular and morphological phylogenetic analysis of the L. boulengeri group, and recovered L. darwinii-N and L. darwinii-S as sister taxa relative to L. laurenti and L. grosseorum. In this study, a coalescent-based, multi-locus analysis for estimating species trees demonstrated that L. darwinii-N and -S are sister lineages and L. laurenti and L. grosseorum also form a clade (see also Camargo et al. submitted). The differences in phylogenetic relationships recovered in this study relative to previous analyses could be a result of gene tree incongruence as a result of incomplete lineage sorting [which could bias the single-locus study of Morando et al. (2004)] and/or gene tree heterogeneity (Degnan and Rosenberg 2009; Edwards 2009; Liu et al. 2009), which would not be accommodated by the concatenation
approach of Abdala (2007). The parthenogen form, that appears nested between *L. darwinii*-N and -S, was not included in the delimitation analysis because of its possible hybrid origin (M. Morando, pers. com.). Although this parthenogen is probably reproductively isolated from other lineages and therefore a valid new species, SDMs can still be applied to confirm the genetic distinctness of this lineage within an ABC framework via an admixture model and/or with new likelihood-based approaches that can test for hybrid speciation (Kubatko 2009).

This study supported the distinctness of *L. darwinii*-N and *L. darwinii*-S using model-based, multi-locus ABC analyses that accommodate non-monophyletic lineages and discordance among gene trees due to incomplete lineage sorting and gene flow after speciation. These factors are especially relevant since levels of gene tree discordance can be quite high for large population sizes due to extensive incomplete lineage sorting. Both new forms, but especially *L. darwinii*-S, occupy large geographic areas and were found at high densities during the field trips, which allowed sampling of multiple individuals per locality. Moreover, taking into account gene flow is relevant because the distributions of these forms approach each other in central Mendoza, suggesting parapatric distributions that could have facilitated past and/or ongoing gene flow after speciation. Finally, while coalescent-models could also be biased because they cannot account for all the details of empirical demographic and speciation histories, the cross-validation of our ABC results with two likelihood-based methods minimizes the impact of each method’s assumptions and provides further strength to our conclusions about species limits. New morphological analyses including for example geometric morphometrics and quantification of color patterns will be required to establish diagnostic criteria for formal description of the northern form since the type locality of *L. darwinii* occurs in the southern region of the distribution (Etheridge 1993). Although new methods for species delimitation enable the
discovery of new, usually cryptic evolutionary independent lineages, traditional taxonomic practices still demand formal morphological descriptions to apply valid names to new forms (Bauer et al. 2010). However, the incorporation of a genealogical perspective and a rigorous statistical framework for species discovery and description will be required to achieve the ultimate goal of a phylogenetically-informed and stable taxonomy (Fujita and Leaché 2010) consistent with current views about species as independently evolving, segments of population-level lineages (de Queiroz 2007).

CONCLUSIONS

The ABC method introduced here represents a contribution toward improving model-based statistical species delimitation with genetic data and developing a unified approach that can simultaneously address species limits, infer species relationships (species trees), and estimate multiple demographic parameters. Several recently proposed methods are either capable of estimating two of these components, for example species trees and species limits (SpeDeSTEM with STEM), species limits and parameters (BPP), or species tree and limits (Brownie, O'Meara 2010). The ABC approach used in this study can delimit species while incorporating one critical parameter, the migration rate, which can potentially erase species divergence and hamper the ability to detect separate species. Ideally, ABC-based delimitation can be generalized for multiple lineages including the possibility of also estimating species trees to get closer to an integrative, unified approach. In addition to methods for genetic data, other kinds of data can be used to aid in species delimitation when genetic data are not available or informative. For example, new criteria have been introduced to delimit species with ecological (Rissler and
and morphological data (Ezard et al. 2010). Finally, a fully integrated approach with different types of data will consist of a single, joint analysis instead of evaluating concordance between several independent analyses. For example, Bayesian inference based on genetic data can incorporate other sources of evidence into the analysis via prior distributions that weights SDMs differentially based on non-genetic information (Yang and Rannala 2010). Moreover, this approach could also incorporate other processes generating gene tree discordance (i.e., gene flow, etc) and take into account uncertainty in the species assignments (O'Meara 2010), which would allow inference of species trees and species boundaries using individual-based data with ambiguous species affinities.

**FUNDING**

AC acknowledges financial support from the BYU Department of Biology and Graduate Studies Office through research, mentoring, and travel scholarships. He also thanks support from the Society for Systematic Biologists and the Society for the Study of Amphibians and Reptiles for research and travel grants. Funding for this study was provided by the NSF “Partnership for International Research and Education” award (OISE 0530267) for collaborative research on Patagonian biodiversity, granted to the following institutions (listed alphabetically): BYU, Centro Nacional Patagónico, Dalhousie University, Darwinion Botanical Institute, George Washington University, Universidad Nacional de Cordoba, Universidad Austral de Chile, Universidad Nacional del Comahue, and the Universidad de Concepción; and also by the NSF “Assembling the Tree-of-Life – Deep Scaly” project on squamate reptiles (subaward EF
0334966) to JWS, Jr. LJA and MM acknowledge several grants under programs PIP & IBol (CONICET) and FONCYT PICT (ANPCYT) (Argentina).

ACKNOWLEDGMENTS

We thank students and posdocs of Sites, Morando, and Avila labs for their assistance and support throughout the completion of this study. A.C. thanks his graduate committee for feedback on his dissertation research/manuscripts. Specimens were collected following local regulations and permits issued by Administración de Parques Nacionales, Direcciones Provinciales de Fauna de Río Negro, Chubut, Neuquen, La Pampa, Mendoza, Catamarca, and Administración de Áreas Protegidas de Neuquen y Mendoza (Argentina).

REFERENCES


Simulation, Genetics and Human Prehistory. McDonald Institute for Archaeological Research, Cambridge.


Csilléry, K., O. François, and M. G. B. Blum. 2010c. abc: estimation and model selection with Approximate Bayesian Computation (ABC). 1.


http://beast.bio.ed.ac.uk/Tracer

Rissler, L. J., and J. J. Apodaca. 2007. Adding more ecology into species delimitation:
Ecological niche models and phylogeography help define cryptic species in the black


Templeton, A. R. 2009. Statistical hypothesis testing in intraspecific phylogeography: nested
clad phylogeographical analysis vs. approximate Bayesian computation. Mol Ecol 18:319-331.


25:488-489.

Thomson, R. C., I. J. Wang, and J. R. Johnson. 2010. Genome-enabled development of DNA

Wakeley, J. 1996. The variance of pairwise nucleotide differences in two populations with

56:875-878.


Table 1. Range of summary statistics obtained for the three simulated models, the two posterior predictive distributions, and the observed values for the two taxon-pairs analyzed. Prior single species corresponds to the no-speciation model. Prior dN-dS is the two-species, speciation with gene flow model used for analyzing the pair *L. darwinii*-N and *L. darwinii*-S. Prior Ll-Lg is the two-species, speciation in isolation model used for analyzing the pair *L. laurenti* and *L. grosseorum*. Post dN-dS and Ll-Lg are the posterior predictive distributions obtained from the simulation of speciation models using the estimated parameter values. Obs dN-dS and Ll-Lg are the observed summary statistics obtained from the sequence data of four loci for each taxon-pair. Summary statistics are: ss (segregating sites), \( \theta_s \) (theta pi), \( \theta_w \) (Watterson's theta), D (Tajima 1983), ZnS (Kelly 1997), H (Fay & Wu XXXX), dvk (number of haplotypes, Depaulis & Veuille 1998), and dvh (haplotype diversity, Depaulis & Veuille 1998). Mean and variance values of summary statistics were calculated across four simulated or observed loci.

<table>
<thead>
<tr>
<th>Sum. Stats.</th>
<th>Prior Single Sp Min–Max</th>
<th>Prior dN-dS Min–Max</th>
<th>Prior Ll-Lg Min–Max</th>
<th>Post dN-dS Min–Max</th>
<th>Post Ll-Lg Min–Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ss</td>
<td>1–313</td>
<td>1–1033</td>
<td>4–5478</td>
<td>37.25-167.5</td>
<td>168.5-387.5</td>
</tr>
<tr>
<td>Variance ss</td>
<td>0–50500</td>
<td>0–184281</td>
<td>0.25–1029877</td>
<td>0-9534.917</td>
<td>0.6667-42172.25</td>
</tr>
<tr>
<td>Mean ( \theta_s )</td>
<td>0.0619–95.4</td>
<td>0.13526–438.4</td>
<td>0.7625–2315</td>
<td>6.23109-46.125</td>
<td>40.5997-133.56</td>
</tr>
<tr>
<td>Variance ( \theta_s )</td>
<td>0.00001–10500</td>
<td>0.00159–50173.8</td>
<td>0.07096–231364</td>
<td>0.00544-2251.596</td>
<td>0.04-10570.93</td>
</tr>
<tr>
<td>Mean ( \theta_w )</td>
<td>0.235–73.6</td>
<td>0.2351–242.9</td>
<td>0.94039–1288</td>
<td>8.75741-39.379</td>
<td>39.614-91.1</td>
</tr>
<tr>
<td>Variance ( \theta_w )</td>
<td>0–2790</td>
<td>0–10185.4</td>
<td>0.01382–56922</td>
<td>0.527.006</td>
<td>0.0368-2330.91</td>
</tr>
<tr>
<td>Mean D</td>
<td>-1.83–2.16</td>
<td>-1.66498–3.4</td>
<td>-1.55493–3</td>
<td>-1.63199-1.455</td>
<td>-0.6274-1.98</td>
</tr>
<tr>
<td>Variance D</td>
<td>0.00007–5.29</td>
<td>0.00001–8.3</td>
<td>0–4</td>
<td>0.00016-4.128</td>
<td>0.0001-2.55</td>
</tr>
<tr>
<td>Mean ZnS</td>
<td>0.00066–1</td>
<td>0.00104–0.9</td>
<td>0.04399–1</td>
<td>0.05253-0.34</td>
<td>0.107-0.39</td>
</tr>
<tr>
<td>Variance ZnS</td>
<td>0–0.5</td>
<td>0–0.5</td>
<td>0–0</td>
<td>0–0.069</td>
<td>0–0.05</td>
</tr>
<tr>
<td>Mean H</td>
<td>-3.73–0.99</td>
<td>-2.56382–1</td>
<td>-1.4093–1</td>
<td>-1.29213-1.015</td>
<td>-0.3927-1.08</td>
</tr>
<tr>
<td>Variance H</td>
<td>0.00002–15.8</td>
<td>0.00001–13.1</td>
<td>0–5</td>
<td>0.00001-7.098</td>
<td>0–4.53</td>
</tr>
<tr>
<td>Mean dvk</td>
<td>1.75–32.8</td>
<td>2–36.8</td>
<td>4.25–40</td>
<td>15.5-31.5</td>
<td>24.75-36</td>
</tr>
<tr>
<td>Variance dvk</td>
<td>0–65</td>
<td>0–64.3</td>
<td>0–61</td>
<td>0–74.25</td>
<td>0–58.25</td>
</tr>
<tr>
<td>Mean dvh</td>
<td>0.0594–0.97</td>
<td>0.11125–1</td>
<td>0.47188–1</td>
<td>0.87656-0.962</td>
<td>0.9334-0.97</td>
</tr>
<tr>
<td>Variance dvh</td>
<td>0–0.18</td>
<td>0–0.2</td>
<td>0–0</td>
<td>0–0.009</td>
<td>0–0</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Sum. Stats.</th>
<th>Obs LdN - LdS</th>
<th>Obs Ll - Lg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ss</td>
<td>38.0000</td>
<td>21.5000</td>
</tr>
<tr>
<td>Variance ss</td>
<td>298.0000</td>
<td>220.3333</td>
</tr>
<tr>
<td>Mean $\theta_x$</td>
<td>9.8263</td>
<td>2.7740</td>
</tr>
<tr>
<td>Variance $\theta_x$</td>
<td>31.0697</td>
<td>3.2813</td>
</tr>
<tr>
<td>Mean $\theta_w$</td>
<td>10.7110</td>
<td>5.0546</td>
</tr>
<tr>
<td>Variance $\theta_w$</td>
<td>23.6763</td>
<td>12.1781</td>
</tr>
<tr>
<td>Mean D</td>
<td>-0.4089</td>
<td>-1.3828</td>
</tr>
<tr>
<td>Variance D</td>
<td>0.3196</td>
<td>0.3726</td>
</tr>
<tr>
<td>Mean ZnS</td>
<td>0.2212</td>
<td>0.0939</td>
</tr>
<tr>
<td>Variance ZnS</td>
<td>0.0151</td>
<td>0.0012</td>
</tr>
<tr>
<td>Mean H</td>
<td>0.9676</td>
<td>0.5247</td>
</tr>
<tr>
<td>Variance H</td>
<td>0.0511</td>
<td>0.0127</td>
</tr>
<tr>
<td>Mean dvk</td>
<td>13.5000</td>
<td>14.5000</td>
</tr>
<tr>
<td>Variance dvk</td>
<td>1.0000</td>
<td>67.0000</td>
</tr>
<tr>
<td>Mean dvh</td>
<td>0.8963</td>
<td>0.7528</td>
</tr>
<tr>
<td>Variance dvh</td>
<td>0.0005</td>
<td>0.0218</td>
</tr>
</tbody>
</table>
Table 2. Relative bias and 95%-HPD coverage for mean parameter estimates of speciation models with varying levels of migration ($m_{\text{max}}$) used in ABC analyses. Relative bias is the percentage of the absolute bias relative to the range of the parameter sampled from the prior distributions. Coverage is the % of simulations where the true parameter value falls within the 95%-HPD of the parameter estimate.

<table>
<thead>
<tr>
<th>$m_{\text{max}}$</th>
<th>$\theta$</th>
<th>$\tau$</th>
<th>$m$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>Coverage</td>
<td>Bias</td>
</tr>
<tr>
<td>0</td>
<td>3.5%</td>
<td>65%</td>
<td>0.4%</td>
</tr>
<tr>
<td>0.25</td>
<td>7.8%</td>
<td>22%</td>
<td>10.7%</td>
</tr>
<tr>
<td>0.5</td>
<td>10.4%</td>
<td>11%</td>
<td>20.2%</td>
</tr>
<tr>
<td>1</td>
<td>10.1%</td>
<td>16%</td>
<td>30.1%</td>
</tr>
</tbody>
</table>
Table 3. Posterior probabilities and parameter estimates of speciation vs. no-speciation models for two taxon-pairs analyzed with ABC

<table>
<thead>
<tr>
<th></th>
<th>L. darwinii-N vs. L. darwinii-S</th>
<th>L. laurenti vs. L. grosseorum</th>
</tr>
</thead>
<tbody>
<tr>
<td>No-speciation</td>
<td>0.033</td>
<td>0.086</td>
</tr>
<tr>
<td>Speciation with gene flow</td>
<td>0.967</td>
<td>-</td>
</tr>
<tr>
<td>Speciation in isolation</td>
<td>-</td>
<td>0.914</td>
</tr>
<tr>
<td>Bayes Factor</td>
<td>29.14</td>
<td>10.57</td>
</tr>
<tr>
<td>$\theta$</td>
<td>1.17(0.90-1.45)</td>
<td>2.88(2.76-3.00)</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.49(0.31-0.71)</td>
<td>1.47(1.17-1.70)</td>
</tr>
<tr>
<td>$m$</td>
<td>0.16(0.06-0.22)</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 4. Model selection with AIC criteria of alternative species delimitation profiles estimated with SpeDeSTEM based on 50 subsampling replicates. Collapsed lineages are shown between parentheses. AIC = Akaike information criterion, K = number of parameters, $\Delta_i = \text{AIC} - \text{AIC}_{\text{min}}$, $w_i = \text{Akaike weight}$

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean AIC</th>
<th>K</th>
<th>$\Delta_i$</th>
<th>$w_i$</th>
<th>Model-likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>dN, dS, Ll, Lg</td>
<td>715.0534</td>
<td>3</td>
<td>0</td>
<td>0.986736</td>
<td>1</td>
</tr>
<tr>
<td>(dN-dS), Ll, Lg</td>
<td>723.6726</td>
<td>2</td>
<td>8.61920</td>
<td>0.013261</td>
<td>0.013439</td>
</tr>
<tr>
<td>dN, dS, (Ll-Lg)</td>
<td>739.9219</td>
<td>2</td>
<td>24.86847</td>
<td>3.93E-06</td>
<td>3.98E-06</td>
</tr>
<tr>
<td>(dN-dS), (Ll-Lg)</td>
<td>750.2925</td>
<td>1</td>
<td>35.23914</td>
<td>2.20E-08</td>
<td>2.23E-08</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1. Map of Argentina showing sampled localities for species of the *Liolaemus darwinii* complex. The inset shows the map of South America with the shaded sampled region. Red circles = *L. darwinii*-S; green circles = *L. darwinii*-N; blue circles = *L. laurenti*; yellow circles = *L. grosseorum*; white circles = *L. olongasta* (used in species tree only). Symbols with a white ring indicates localities sampled for the nuclear loci.

Figure 2. Accuracy of the ABC method (mean posterior probability of the correct model) for species delimitation relative to the amount of gene flow (migration rate) based on three different algorithms (simple rejection, multinomial logistic, and neural networks). (a) the no-speciation model is true; (b) the speciation model is true. Vertical bars represent standard errors based on 100 simulations.

Figure 3. Species tree of the *L. darwinii* complex based on a Bayesian analysis in *BEAST*. Numbers above branches are posterior probabilities of clades. The scale bar is in units of substitutions per site.
Figure 1
Figure 2
Figure 3
Figure Ia. Number of references published between 1981 and 2010 retrieved from the ISI Web of Science that contained the keyword "species delimitation".

Figure Ib. Number of references published between 1992 and 2010 retrieved from the ISI Web of Science that contained the keyword "approximate bayesian computation".
Appendix II. List of specimens sequenced for this study. Loci: 1 = cyt *b*, 2 = A1D, 3 = A9C, and 4 = B6B. * = sequence used in species tree analysis. LJAMM = L. J. Avila & M. Morando herpetological collection (CONICET-CENPAT). Code represents the numbers used to identify sampled localities in Fig. 1.

<table>
<thead>
<tr>
<th>Lineage/Specimen</th>
<th>Code</th>
<th>Loci</th>
<th>Coordinates</th>
<th>Locality</th>
</tr>
</thead>
<tbody>
<tr>
<td>L. darwinii-N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMCN 8LA</td>
<td>1</td>
<td>1, 3, 4</td>
<td>31°40'0&quot; 68°16'0&quot; W</td>
<td>Medanos Grandes, near Caucete, Caucete, San Juan</td>
</tr>
<tr>
<td>IMCN 19LA</td>
<td>2</td>
<td>1, 2, 3, 4</td>
<td>31°33'0&quot; 68°00'0&quot; W</td>
<td>W slope Sierra de Pie de Palo, Caucete, San Juan</td>
</tr>
<tr>
<td>LJAMM 1970</td>
<td>3</td>
<td>1</td>
<td>30°11'21&quot; S, 67°41'33&quot; W</td>
<td>Ruta Provincial 26, Km 90, Independencia, La Rioja.</td>
</tr>
<tr>
<td>LJAMM 1983</td>
<td>3</td>
<td>1, 2, 3, 4</td>
<td>30°11'21&quot; S, 67°41'33&quot; W</td>
<td>Parque Provincial Talampaya, Ruta Provincial 26, Km 92, Independencia, La Rioja</td>
</tr>
<tr>
<td>LJAMM 1984</td>
<td>3</td>
<td>2, 3, 4</td>
<td>30°11'21&quot; S, 67°41'33&quot; W</td>
<td>Parque Provincial Talampaya, Ruta Provincial 26, Km 92, Independencia, La Rioja</td>
</tr>
<tr>
<td>LJAMM 2276</td>
<td>4</td>
<td>1, 2</td>
<td>31°41'38&quot; S, 68°09'55&quot; W</td>
<td>Ruta Nacional 141, 15 Km E Bermejo, Caucete, San Juan</td>
</tr>
<tr>
<td>LJAMM 2287</td>
<td>5</td>
<td>1</td>
<td>31°35'53&quot; S, 67°41'33&quot; W</td>
<td>Parque Nacional Talampaya, Ruta Provincial 26, Km 92, Independencia, La Rioja</td>
</tr>
<tr>
<td>LJAMM 2295</td>
<td>6</td>
<td>1*, 2*, 3*, 4*</td>
<td>31°36'25&quot; S, 67°42'47&quot; W</td>
<td>Ruta Nacional 141, 21 Km E Bermejo, Caucete, San Juan</td>
</tr>
<tr>
<td>LJAMM 4038</td>
<td>7</td>
<td>2, 3, 4</td>
<td>33°19'33&quot; S, 68°04'03&quot; W</td>
<td>Parque Provincial 153, 2 Km S Las Catitas, Santa Rosa, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4083</td>
<td>8</td>
<td>2, 3, 4</td>
<td>30°12'21&quot; S, 67°40'26&quot; W</td>
<td>Ruta Provincial 510, 2 Km E Baldecitos, Km 88, Valle Fértil, San Juan</td>
</tr>
<tr>
<td>LJAMM 4078</td>
<td>8</td>
<td>2, 3, 4</td>
<td>30°12'21&quot; S, 67°40'26&quot; W</td>
<td>Ruta Provincial 510, 2 Km E Baldecitos, Km 88, Valle Fértil, San Juan</td>
</tr>
<tr>
<td>LJAMM 4081</td>
<td>8</td>
<td>1</td>
<td>30°12'21&quot; S, 67°40'26&quot; W</td>
<td>Ruta Provincial 510, 2 Km E Baldecitos, Km 88, Valle Fértil, San Juan</td>
</tr>
<tr>
<td>LJAMM 4086</td>
<td>8</td>
<td>1</td>
<td>30°12'21&quot; S, 67°40'26&quot; W</td>
<td>Ruta Provincial 510, 2 Km E Baldecitos, Km 88, Valle Fértil, San Juan</td>
</tr>
<tr>
<td>LJAMM 4089</td>
<td>8</td>
<td>1, 2, 3, 4</td>
<td>30°12'21&quot; S, 67°40'26&quot; W</td>
<td>Ruta Provincial 510, 2 Km E Baldecitos, Km 88, Valle Fértil, San Juan</td>
</tr>
<tr>
<td>LJAMM 4090</td>
<td>8</td>
<td>1</td>
<td>30°12'21&quot; S, 67°40'26&quot; W</td>
<td>Ruta Provincial 510, 2 Km E Baldecitos, Km 88, Valle Fértil, San Juan</td>
</tr>
<tr>
<td>LJAMM 4091</td>
<td>9</td>
<td>1</td>
<td>32°11'39&quot; S, 67°41'22&quot; W</td>
<td>Ruta Nacional 141, 28 km W Encon, 2 km W limita San Juan-San Luis, 25 de Mayo, San Juan</td>
</tr>
<tr>
<td>LJAMM 4092</td>
<td>9</td>
<td>1*, 2*, 3*, 4*</td>
<td>32°11'39&quot; S, 67°41'22&quot; W</td>
<td>Ruta Nacional 141, 28 km W Encon, 2 km W limita San Juan-San Luis, 25 de Mayo, San Juan</td>
</tr>
<tr>
<td>LJAMM 4130</td>
<td>10</td>
<td>1, 2, 3, 4</td>
<td>32°36'57&quot; S, 68°02'47&quot; W</td>
<td>Ruta Nacional 141, 18.2 km N Costa de Araujo, Lavalle, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4185</td>
<td>11</td>
<td>1, 2, 3, 4</td>
<td>29°04'09&quot; 67°22'48&quot; W</td>
<td>Catinzaco, Chilecito, La Rioja</td>
</tr>
<tr>
<td>LJAMM 4186</td>
<td>11</td>
<td>2, 3, 4</td>
<td>29°04'09&quot; 67°22'48&quot; W</td>
<td>Catinzaco, Chilecito, La Rioja</td>
</tr>
<tr>
<td>LJAMM 4187</td>
<td>11</td>
<td>1, 2, 3, 4</td>
<td>29°04'09&quot; 67°22'48&quot; W</td>
<td>Catinzaco, Chilecito, La Rioja</td>
</tr>
<tr>
<td>LJAMM 5780</td>
<td>12</td>
<td>1, 2, 3, 4</td>
<td>35°00'51&quot; S, 68°02'47&quot; W</td>
<td>Ruta Provincial 141, 35.7 km SW Soitue, San Rafael, Mendoza</td>
</tr>
<tr>
<td>LJAMM 10671</td>
<td>13</td>
<td>1</td>
<td>32°19'49&quot; 67°17'42&quot; W</td>
<td>Ruta Nacional 20, 29 km W La Chañarienta, 2 km W limita San Juan-San Luis, 25 de Mayo, San Juan</td>
</tr>
<tr>
<td>LJAMM 10686</td>
<td>14</td>
<td>1</td>
<td>31°45'12&quot; 67°55'05&quot; W</td>
<td>Ruta Nacional 141, 12.7 km E Malvín, Caucete, San Juan</td>
</tr>
<tr>
<td>LJAMM 10703</td>
<td>15</td>
<td>2, 3, 4</td>
<td>30°35'01&quot; 67°33'16&quot; W</td>
<td>Ruta Nacional 141, 10.1 km E acceso E a Bermejo, Caucete, San Juan</td>
</tr>
<tr>
<td>LJAMM 10720</td>
<td>15</td>
<td>2, 3, 4</td>
<td>30°35'01&quot; 67°33'16&quot; W</td>
<td>Ruta Nacional 141, 10.1 km E acceso E a Bermejo, Caucete, San Juan</td>
</tr>
<tr>
<td>LJAMM 10722</td>
<td>15</td>
<td>2, 3, 4</td>
<td>30°35'01&quot; 67°33'16&quot; W</td>
<td>Ruta Nacional 141, 10.1 km E acceso E a Bermejo, Caucete, San Juan</td>
</tr>
<tr>
<td>LJAMM 10728</td>
<td>16</td>
<td>1</td>
<td>31°25'02&quot; 67°20'66&quot; W</td>
<td>Ruta Nacional 141, 5 km E border San Juan-La Rioja, Rosario Vera Peñaloza, San Juan</td>
</tr>
<tr>
<td>LJAMM 10734</td>
<td>16</td>
<td>1</td>
<td>31°25'02&quot; 67°20'66&quot; W</td>
<td>Ruta Nacional 141, 5 km E border San Juan-La Rioja, Rosario Vera Peñaloza, San Juan</td>
</tr>
<tr>
<td>LJAMM 10735</td>
<td>16</td>
<td>2*, 3*, 4*</td>
<td>31°25'02&quot; 67°20'66&quot; W</td>
<td>Ruta Nacional 141, 5 km E border San Juan-La Rioja, Rosario Vera Peñaloza, San Juan</td>
</tr>
<tr>
<td>LJAMM 10880</td>
<td>17</td>
<td>3, 4</td>
<td>38 41° S 65 20° W</td>
<td>Ruta Nacional 150, 3 km W Patquiá, Independencia, La Rioja</td>
</tr>
</tbody>
</table>

| L. darwinii-S    |      |      |             |          |
| LJAMM fn5       | 18   | 1     | 38 41° S 65 20° W | Ruta Provincial 34, 0.5 km W junction Ruta Provincial 13 to Lihué Cale, Lihué Cale, La Pampa |
LJAMM fn9 19 1 38° 43' S 65° 59' W  
Ruta Provincial 34, 0.5 km W Estancia San Eduardo, Curacó, La Pampa

LJAMM fn18 20 1 40° 01' S 66° 00' W  
Ruta Provincial 4, 84 km S junction Ruta Nacional 250, Valcheta, Río Negro

LJAMM fn19 20 1 40° 01' S 66° 00' W  
Ruta Provincial 4, 84 km S junction Ruta Nacional 250, Valcheta, Río Negro

LJAMM fn63 21 1 40° 56' S 66° 38' W  
Ruta Provincial 60, 1 km N Chipauquil, Valcheta, Río Negro

LJAMM fn64 21 1 40° 56' S 66° 38' W  
Ruta Provincial 60, 1 km N Chipauquil, Valcheta, Río Negro

LJAMM fn65 21 1 40° 56' S 66° 38' W  
Ruta Provincial 60, 1 km N Chipauquil, Valcheta, Río Negro

LJAMM 2993 22 1 37°52'17.0", 67°06'35.1"  
1 km N Ruta Provincial 23 y Ruta a 25 de Mayo, Curacó, La Pampa

LJAMM 2994 22 1 37°52'17.0", 67°06'35.1"  
1 km N Ruta Provincial 23 y Ruta a 25 de Mayo, Curacó, La Pampa

LJAMM 2995 22 1 37°52'17.0", 67°06'35.1"  
1 km N Ruta Provincial 23 y Ruta a 25 de Mayo, Curacó, La Pampa

LJAMM 3014 23 1 40°30'16.8", 66°32'47.1"  
Ruta Nacional 23, Estacion Nahuel Niyeu, Valcheta, Río Negro

LJAMM 3016 23 1 40°30'16.8", 66°32'47.1"  
Ruta Nacional 23, Estacion Nahuel Niyeu, Valcheta, Río Negro

LJAMM 3022 18 1 38°41'18.8", 65°20'49.8"  
Ruta Provincial 34, 0.5 km W desvío Ruta Provincial 13 a Lihue Calel, Lihue Calel, La Pampa

LJAMM 3023 18 1 38°41'18.8", 65°20'49.8"  
Ruta Provincial 34, 0.5 km W desvío Ruta Provincial 13 a Lihue Calel, Lihue Calel, La Pampa

LJAMM 3026 21 1 40°24'08.5", 66°02'44.9"  
Ruta Provincial 4, Laguna del Indio Muerto, Valcheta, Río Negro

LJAMM 3027 21 1 40°24'08.5", 66°02'44.9"  
Ruta Provincial 4, Laguna del Indio Muerto, Valcheta, Río Negro

LJAMM 3028 21 1 40°24'08.5", 66°02'44.9"  
Ruta Provincial 4, Laguna del Indio Muerto, Valcheta, Río Negro

LJAMM 3030 20 1 40°06'20.2", 66°00'26.5"  
Ruta Provincial 4, 84 km S junction Ruta Nacional 250, Valcheta, Río Negro

LJAMM 3035 19 1 38°43'19.7", 65°59'03.0"  
Ruta Provincial 34, 0.5 km W Estancia San Eduardo, Curacó, La Pampa

LJAMM 4037 25 1 *360959.7, 681458.0  
Ruta Provincial 10, 4.6 Km E Agua Escondida, Chical Co, La Pampa

LJAMM 5104 26 1*, 2*, 3*, 4* 36°08'19", 68°17'23"  
Ruta Provincial 190. 2 Km N Agua Escondida, Malargüe, Mendoza

LJAMM 5180 27 2, 4 25 Km N Naval. Roca, General Roca, Río Negro

LJAMM 5374 28 2 38°18'22.5", 70°02'55.3"  
3 Km S Camino Va. Del Agrio, Zapala, Neuquén

LJAMM 5755 29 2 36°38'24.5", 69°49'55.4"  
Ruta Nacional 40, 3.2 km N Ranquil Norte, Malargüe, Mendoza

LJAMM 7796 30 2 37°40'54.1"S, 69°06'55.8"W  
Ruta Provincial 5, 37.6 km N empalme Ruta Provincial 7, Pehuenches, Neuquén

LJAMM 8313 31 2* 40°28'31.5"S 65°23'54.5"W  
Camino vecinal a 52,4 km NW San Antonio Oeste, bajo del Gualicho, San Antonio, Río Negro

LJAMM 8355 32 3 37°25'34.7"S 67°28'07.7"W  
35.8 km N Ruta Nacional 152, camino vecinal 17 km NE Puelén, Puelén, La Pampa

LJAMM 8340 33 3  
Ruta provincial 62, 50 km N Nahuel Niyeu, Valcheta, Río Negro

LJAMM 8346 32 2 37°25'34.7"S 67°28'07.7"W  
35.8 km N Ruta Nacional 152, camino vecinal 17 km NE Puelén, Puelén, Río Negro

LJAMM 8361 34 3 39°58'11.1"S 66°34'13.5"W  
Camino vecinal a bajo Santa Rosa, entre Ruta provincial 62 y 63, 3 km SE Santa Rosa, Valcheta, Río Negro

LJAMM 8364 34 3 39°58'11.1"S 66°34'13.5"W  
Camino vecinal a bajo Santa Rosa, entre Ruta provincial 62 y 63, 3 km SE Santa Rosa, Valcheta, Río Negro

LJAMM 8369 35 3 39°05'35.6"S 65°28'38.9"W  
Ruta provincial 56, 24.5 km empalme Ruta Nacional 22, 29,5 km NE Choele Choel, Avellaneda, Río Negro

LJAMM 8389 36 3 39°41'43.3"S 66°06'27"W  
Ruta provincial 63, 37,3 km E empalme Ruta provincial 62, camino a Lamarque, Avellaneda, Río Negro

LJAMM 8393 36 3 39°41'43.3"S 66°06'27"W  
Ruta provincial 63, 37,3 km E empalme Ruta provincial 62, camino a Lamarque, Avellaneda, Río Negro

LJAMM 8402 37 3 37°34'22.2"S 67°23'32.4"W  
camino vecinal 36.9 km E Puelén, Puelén, La Pampa

LJAMM 8409 38 3 37°56'23"S 67°06'32.1"W  
Ruta provincial 23, 6,6 km S Ruta provincial 26, Puelén, La Pampa

LJAMM 10344 39 4 37°39'16.1"S 68°46'33.5"W  
Auca Mahuida, Cantera 1, Neuquén

LJAMM 10391 40 4* 40°20'55.8"S 65°02'59.4"W  
Gran Bajo del Gualicho, 42,4 Km NW San Antonio Oeste, por Ruta Provincial 2, San Antonio, Río Negro

LJAMM 10518 41 4* 37°04'29.8"S 67°47'07.6"W  
Ruta Provincial 16, 23,6 km W empalme Ruta Nacional 151, Puelén, La Pampa

LJAMM 10582 41 1* 37°04'29.8"S 67°47'07.6"W  
Ruta Provincial 16, 23,6 km W empalme Ruta Nacional 151, Puelén, La Pampa

LJAMM 11022 42 1*, 2*, 3* -64.97092,-42.79344  
Puerto Madryn, Biedma, Chubut

LJAMM 11331 43 4 364750.6, 685639.4  
Ruta Provincial 180, 17.5 km NE junction with road El Clavado, 5 km SE La Matancilla, Malargüe, Mendoza
<table>
<thead>
<tr>
<th>L. laurenti</th>
<th>43</th>
<th>364750.6, 685639.4</th>
<th>Ruta Provincial 180, 17.5 km NE junction with road El Clavado, 5 km SE La Matancilla, Malargüe, Mendoza</th>
</tr>
</thead>
<tbody>
<tr>
<td>LJAMM 2334</td>
<td>44</td>
<td></td>
<td>'Ruta Nacional 40 y Río La Puerta. Km 1298, Tinogasta, Catamarca</td>
</tr>
<tr>
<td>LJAMM 4093</td>
<td>9</td>
<td></td>
<td>'Ruta Nacional 20. 2.8 Km W Encon, 25 de Mayo, San Juan</td>
</tr>
<tr>
<td>LJAMM 4115</td>
<td>8</td>
<td></td>
<td>'Ruta Provincial 510, Km 88, 2 Km E Baldecitos, Valle Fértel, San Juan</td>
</tr>
<tr>
<td>LJAMM 4116</td>
<td>8</td>
<td></td>
<td>'Ruta Provincial 510, Km 88, 2 Km E Baldecitos, Valle Fértel, San Juan</td>
</tr>
<tr>
<td>LJAMM 4117</td>
<td>8</td>
<td></td>
<td>'Ruta Provincial 510, Km 88, 2 Km E Baldecitos, Valle Fértel, San Juan</td>
</tr>
<tr>
<td>LJAMM 4118</td>
<td>8</td>
<td></td>
<td>'Ruta Provincial 510, Km 88, 2 Km E Baldecitos, Valle Fértel, San Juan</td>
</tr>
<tr>
<td>LJAMM 4160</td>
<td>45</td>
<td></td>
<td>'Ruta Nacional 40, Km 657, 9 Km E Pituí, Fatamanta, La Rioja</td>
</tr>
<tr>
<td>LJAMM 4161</td>
<td>45</td>
<td></td>
<td>'Ruta Nacional 40, Km 657, 9 Km E Pituí, Fatamanta, La Rioja</td>
</tr>
<tr>
<td>LJAMM 4162</td>
<td>45</td>
<td></td>
<td>'Ruta Nacional 40, Km 657, 9 Km E Pituí, Fatamanta, La Rioja</td>
</tr>
<tr>
<td>LJAMM 4164</td>
<td>45</td>
<td></td>
<td>'Ruta Nacional 40, Km 657, 9 Km E Pituí, Fatamanta, La Rioja</td>
</tr>
<tr>
<td>LJAMM 4165</td>
<td>45</td>
<td></td>
<td>'Ruta Nacional 40, Km 657, 9 Km E Pituí, Fatamanta, La Rioja</td>
</tr>
<tr>
<td>LJAMM 4166</td>
<td>45</td>
<td></td>
<td>'Ruta Nacional 40, Km 657, 9 Km E Pituí, Fatamanta, La Rioja</td>
</tr>
<tr>
<td>LJAMM 4208</td>
<td>46</td>
<td></td>
<td>'Entrada a Antinaco, 3.8 Km E Ruta Nacional 40, Fatamanta, La Rioja</td>
</tr>
<tr>
<td>LJAMM 4209</td>
<td>46</td>
<td></td>
<td>'Entrada a Antinaco, 3.8 Km E Ruta Nacional 40, Fatamanta, La Rioja</td>
</tr>
<tr>
<td>LJAMM 5745</td>
<td>47</td>
<td>31º25'25,1&quot;, 67º02'50,9&quot;</td>
<td>Ruta Nacional 141, 7.6 km W Mascalín, Salinas de Mascalín, Rosario Vera Peñaloza, La Rioja</td>
</tr>
<tr>
<td>LJAMM 5835</td>
<td>46</td>
<td>28º50'20,1&quot;, 67º24'47,7&quot;</td>
<td>Entrada a Antinaco, 2.4 km E Ruta Nacional 40, Fatamanta, La Rioja</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>L. grosseorum</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>LJAMM fn416</td>
<td>48</td>
<td></td>
<td>El Nihuil, San Rafael, Mendoza</td>
</tr>
<tr>
<td>LJAMM fn417</td>
<td>48</td>
<td></td>
<td>El Nihuil, San Rafael, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4019</td>
<td>49</td>
<td></td>
<td>'Ruta Provincial 180, 30 Km S El Nihuil, San Rafael, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4020</td>
<td>49</td>
<td></td>
<td>'Ruta Provincial 180, 30 Km S El Nihuil, San Rafael, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4021</td>
<td>49</td>
<td></td>
<td>'Ruta Provincial 180, 30 Km S El Nihuil, San Rafael, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4022</td>
<td>49</td>
<td></td>
<td>'Ruta Provincial 180, 30 Km S El Nihuil, San Rafael, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4026</td>
<td>50</td>
<td></td>
<td>'Ruta Provincial 10, 5.1 Km E La Humada, Chical Co, La Pampa</td>
</tr>
<tr>
<td>LJAMM 4027</td>
<td>50</td>
<td></td>
<td>'Ruta Provincial 10, 5.1 Km E La Humada, Chical Co, La Pampa</td>
</tr>
<tr>
<td>LJAMM 4029</td>
<td>51</td>
<td></td>
<td>4 Km E Agua del Toro, Malargüe, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4030</td>
<td>51</td>
<td></td>
<td>4 Km E Agua del Toro, Malargüe, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4045</td>
<td>52</td>
<td></td>
<td>'Ruta Provincial 180, 28.1 Km N entrada sur a La Matancilla, Malargüe, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4046</td>
<td>52</td>
<td></td>
<td>'Ruta Provincial 180, 28.1 Km N entrada sur a La Matancilla, Malargüe, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4047</td>
<td>52</td>
<td></td>
<td>'Ruta Provincial 180, 28.1 Km N entrada sur a La Matancilla, Malargüe, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4048</td>
<td>52</td>
<td></td>
<td>'Ruta Provincial 180, 28.1 Km N entrada sur a La Matancilla, Malargüe, Mendoza</td>
</tr>
<tr>
<td>LJAMM 4105</td>
<td>53</td>
<td></td>
<td>Camino a Chos Malal (huella), 54 Km W enpalme Ruta Nacional 151, Chical Co, La Pampa</td>
</tr>
<tr>
<td>LJAMM 4107</td>
<td>53</td>
<td></td>
<td>Camino a Chos Malal (huella), 54 Km W enpalme Ruta Nacional 151, Chical Co, La Pampa</td>
</tr>
<tr>
<td>LJAMM 4108</td>
<td>53</td>
<td></td>
<td>Camino a Chos Malal (huella), 54 Km W enpalme Ruta Nacional 151, Chical Co, La Pampa</td>
</tr>
<tr>
<td>LJAMM 4109</td>
<td>53</td>
<td></td>
<td>Camino a Chos Malal (huella), 54 Km W enpalme Ruta Nacional 151, Chical Co, La Pampa</td>
</tr>
<tr>
<td>LJAMM 4112</td>
<td>53</td>
<td></td>
<td>Camino a Chos Malal (huella), 54 Km W enpalme Ruta Nacional 151, Chical Co, La Pampa</td>
</tr>
<tr>
<td>LJAMM 4113</td>
<td>53</td>
<td></td>
<td>Camino a Chos Malal (huella), 54 Km W enpalme Ruta Nacional 151, Chical Co, La Pampa</td>
</tr>
<tr>
<td>Ref.</td>
<td>Row</td>
<td>Col</td>
<td>Localización</td>
</tr>
<tr>
<td>-----------</td>
<td>-----</td>
<td>-----</td>
<td>------------------------------</td>
</tr>
<tr>
<td>LJAMM 5040</td>
<td>54</td>
<td>2, 3, 4</td>
<td>35°26'12&quot;, 68°04'06&quot;</td>
</tr>
<tr>
<td>LJAMM 7707</td>
<td>55</td>
<td>2</td>
<td>37°42'22.6&quot;S, 68°27'12.4&quot;W</td>
</tr>
<tr>
<td>LJAMM 7708</td>
<td>55</td>
<td>2, 3</td>
<td>37°42'22.6&quot;S, 68°27'12.4&quot;W</td>
</tr>
<tr>
<td>LJAMM 7709</td>
<td>55</td>
<td>2, 3</td>
<td>37°42'22.6&quot;S, 68°27'12.4&quot;W</td>
</tr>
<tr>
<td>LJAMM 7710</td>
<td>55</td>
<td>2</td>
<td>37°42'22.6&quot;S, 68°27'12.4&quot;W</td>
</tr>
<tr>
<td>LJAMM 7713</td>
<td>55</td>
<td>2, 3</td>
<td>37°42'22.6&quot;S, 68°27'12.4&quot;W</td>
</tr>
<tr>
<td>LJAMM 7716</td>
<td>55</td>
<td>3</td>
<td>37°42'22.6&quot;S, 68°27'12.4&quot;W</td>
</tr>
<tr>
<td>LJAMM 7730</td>
<td>56</td>
<td>3</td>
<td>37°37'56.1S, 69°35'45.4&quot;W</td>
</tr>
<tr>
<td>LJAMM 7742</td>
<td>57</td>
<td>2</td>
<td>37°38'41.2&quot;S, 69°32'5.8&quot;W</td>
</tr>
<tr>
<td>LJAMM 7825</td>
<td>58</td>
<td>1*, 2*</td>
<td>381349, 685736.2</td>
</tr>
<tr>
<td>LJAMM 7827</td>
<td>58</td>
<td>2, 3*</td>
<td>381349, 685736.2</td>
</tr>
<tr>
<td>LJAMM 8753</td>
<td>59</td>
<td>3</td>
<td>37°37'36.3&quot;S, 68°04'35.2&quot;W</td>
</tr>
</tbody>
</table>

*L. olongasta*

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Row</th>
<th>Col</th>
<th>Localización</th>
</tr>
</thead>
<tbody>
<tr>
<td>LJAMM 10750</td>
<td>60</td>
<td>4*</td>
<td>311420.3, 683904.4</td>
</tr>
<tr>
<td>LJAMM 10751</td>
<td>60</td>
<td>1*</td>
<td>311420.3, 683904.4</td>
</tr>
<tr>
<td>LJAMM 10783</td>
<td>61</td>
<td>1*, 2*, 3*</td>
<td>301343, 681935.3</td>
</tr>
<tr>
<td>LJAMM 10785</td>
<td>61</td>
<td>4*</td>
<td>301343, 681935.3</td>
</tr>
<tr>
<td>LJAMM 10820</td>
<td>62</td>
<td>4*</td>
<td>294117.3, 680141.9</td>
</tr>
<tr>
<td>LJAMM 10821</td>
<td>62</td>
<td>1*, 2*, 3*</td>
<td>294117.3, 680141.9</td>
</tr>
</tbody>
</table>

*L. boletengeri*

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Row</th>
<th>Col</th>
<th>Localización</th>
</tr>
</thead>
<tbody>
<tr>
<td>LJAMM 2187</td>
<td>1*, 2*, 3*, 4*</td>
<td>45°27'22&quot;, 69°45'56&quot;</td>
<td>Ruta Provincial 20, 23 Km W Los Manantiales, 576 msnm, Río Senguer, Chubut</td>
</tr>
</tbody>
</table>
Appendix III

Figure IIIa. Plot of PC1 and PC2 for the prior predictive distribution of summary statistics for the no-speciation (green triangles) and the speciation-with-migration (red diamonds) models. The observed summary statistics from the taxon-pair *L. darwinii*-N vs. *L. darwinii*-S are shown with a pink dot.

Figure IIIb. Plot of PC1 and PC3 for the prior predictive distribution of summary statistics for the no-speciation (green triangles) and the speciation-with-migration (red diamonds) models. The observed summary statistics from the taxon-pair *L. darwinii*-N vs. *L. darwinii*-S are shown with a pink dot.
Figure IIIc. Plot of PC2 and PC3 for the prior predictive distribution of summary statistics for the no-speciation (green triangles) and the speciation-with-migration (red diamonds) models. The observed summary statistics from the taxon-pair *L. darwinii*-N vs. *L. darwinii*-S are shown with a pink dot.

Figure IIIId. Plot of PC1 and PC2 for the prior predictive distribution of summary statistics for the no-speciation (green triangles) and the speciation-in-isolation (red diamonds) models. The observed summary statistics from the taxon-pair *L. laurenti* vs. *L. grosseorum* are shown with a pink dot.
Figure IIIe. Plot of PC1 and PC3 for the prior predictive distribution of summary statistics for the no-speciation (green triangles) and the speciation-in-isolation (red diamonds) models. The observed summary statistics from the taxon-pair *L. laurenti* vs. *L. grosseorum* are shown with a pink dot.

Figure IIIf. Plot of PC2 and PC3 for the prior predictive distribution of summary statistics for the no-speciation (green triangles) and the speciation-in-isolation (red diamonds) models. The observed summary statistics from the taxon-pair *L. laurenti* vs. *L. grosseorum* are shown with a pink dot.
Figure IIIg. Plot of PC1 and PC2 for the posterior predictive distribution of summary statistics for the speciation-with-migration model. The observed summary statistics from the taxon-pair *L. darwinii*-N vs. *L. darwinii*-S are shown with a green triangle.

Figure IIIh. Plot of PC1 and PC3 for the posterior predictive distribution of summary statistics for the speciation-with-migration model. The observed summary statistics from the taxon-pair *L. darwinii*-N vs. *L. darwinii*-S are shown with a green triangle.
Figure IIIi. Plot of PC2 and PC3 for the posterior predictive distribution of summary statistics for the speciation-with-migration model. The observed summary statistics from the taxon-pair *L. darwini*-N vs. *L. darwini*-S are shown with a green triangle.

Figure IIIj. Plot of PC1 and PC2 for the posterior predictive distribution of summary statistics for the speciation-in-isolation model. The observed summary statistics from the taxon-pair *L. laurenti* vs. *L. grosseorum* are shown with a green triangle.
Figure IIIk. Plot of PC1 and PC3 for the posterior predictive distribution of summary statistics for the speciation-in-isolation model. The observed summary statistics from the taxon-pair *L. laurenti* vs. *L. grosseorum* are shown with a green triangle.

Figure IIIl. Plot of PC2 and PC3 for the posterior predictive distribution of summary statistics for the speciation-in-isolation model. The observed summary statistics from the taxon-pair *L. laurenti* vs. *L. grosseorum* are shown with green triangle.