Different roads lead to Rome: Integrative taxonomic approaches lead to the discovery of two new lizard lineages in the *Liolaemus montanus* group (Squamata: Liolaemidae)

Cesar Aguilar  
*Brigham Young University - Provo*

Perry L. Wood Jr.  
*Brigham Young University - Provo*

Mark C. Belk  
*Brigham Young University - Provo*, mark_belk@byu.edu

Mike H. Duff  
*Brigham Young University - Provo*

Original Publication Citation  

BYU ScholarsArchive Citation  
Aguilar, Cesar; Wood, Perry L. Jr.; Belk, Mark C.; Duff, Mike H.; and Sites, Jack W. Jr., "Different roads lead to Rome: Integrative taxonomic approaches lead to the discovery of two new lizard lineages in the *Liolaemus montanus* group (Squamata: Liolaemidae)" (2016). *Faculty Publications*. 5413.  
https://scholarsarchive.byu.edu/facpub/5413

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Different roads lead to Rome: Integrative taxonomic approaches lead to the discovery of two new lizard lineages in the *Liolaemus montanus* group (Squamata: Liolaemidae)

CESAR AGUILAR¹,²,³*, PERRY L. WOOD JR.¹, MARK C. BELK¹, MIKE H. DUFF¹ and JACK W. SITES JR.¹

¹Department of Biology and M. L. Bean Life Science Museum, Brigham Young University (BYU), Provo, UT, 84602, USA
²Departamento de Herpetología, Museo de Historia Natural de San Marcos (MUSM), Av. Arenales 1256, Jesus Maria, Lima, Peru
³Facultad de Ciencias Biológicas, Instituto de Ciencias Biológicas Antonio Raimondi, Universidad Nacional Mayor de San Marcos, Lima, Peru

Received 21 February 2016; revised 30 July 2016; accepted for publication 30 July 2016

Integrative taxonomy (IT) is becoming a preferred approach to delimiting species boundaries by including different empirical criteria. IT methods can be divided into two types of procedures both of which use multiple kinds of evidence: step-by-step approaches test hypotheses by sequential evaluation in a hypothetic-deductive framework, while model-based procedures delimit groups based on statistical information criteria. In this study we used a step-by-step approach and a Gaussian clustering (GC) method to test species boundaries in the northernmost species of the *Liolaemus montanus* group. We used different methods based on mitochondrial and nuclear DNA sequence data, morphological measures and niche envelope variables. In contrast with GC, our step-by-step approach shows that one Andean population (Abra Apacheta) previously considered part of *L*. *melanogaster*, is actually nested within another clade; another Andean species, *L*. *thomasi*, is equivocally shown to be either a distinct species or conspecific with *L*. *ortizi*; and an additional Andean population (Abra Toccto) is delimited by concordance among most lines of evidence and different methods as a distinct lineage. However, one of the oldest and low-elevation populations (Nazca) is strongly delimited by all data sets and IT procedures as a new lineage distinct from any currently recognized species. © 2016 The Linnean Society of London, Biological Journal of the Linnean Society, 2017, **120**, 448–467.

KEYWORDS: lizards – Pacific lowland – Peruvian Andes – species boundaries.

**INTRODUCTION**

Integrative taxonomy (IT), the use of different kinds of data and methods for species discovery and hypothesis testing, is becoming a fundamental approach in species delimitation (Padial & De La Riva, 2010; Padial *et al*., 2010; Mckay *et al*., 2014; Pante, Schoelinck & Puillandre, 2015). This shift to IT as an alternative to species delimitation (SDL) studies based exclusively on molecular data is due to evidence that: (1) sequence data alone may not reflect accumulation of differences associated with reproductive isolation; (2) very young species or those that have diverged with ongoing gene flow in neutral regions of the genome may not be detected; (3) failures can occur when errors associated with initial assignment of individuals to species are not detected in upstream analyses; and (4) when molecular analyses, in general, are based on simplified assumptions about divergence processes (Camargo & Sites, 2013; Solís-Lemus, Knowles & Ané, 2014; Olave, Solà & Knowles, 2014a).

Integrative taxonomy approaches using different types of data should reveal cryptic diversity when divergence occurs (at least initially) along non-molecular axes of differentiation, or when divergence
occurs with gene flow (Solís-Lemus et al., 2014; Olave et al., 2014a). IT also exposes potential conflicts among the different kinds of data, and leads to more deeply informed and statistically rigorous assessments of biodiversity (Mckay et al., 2014). IT methods can be divided informally into two types of procedures: (1) step-by-step methods based on sequential analyses of independent data types, followed by a qualitative assessment of diversity in a hypothetico-deductive framework (Schlick-Steiner et al., 2010; Yeates et al., 2011; Andrújar et al., 2014); and (2) model-based methods that simultaneously evaluate multiple data types, followed by delimitation of species based on a statistical or information criterion (Guillot et al., 2012; Edwards & Knowles, 2014; Solís-Lemus et al., 2014). Both IT approaches can be used for the four focal areas of SDL: (1) validation of candidate species as evolutionary distinct lineages; (2) inferring species relationships; (3) detecting ‘cryptic diversity’; and (4) individual specimen assignment to a species group (Edwards & Knowles, 2014; Leavitt, Moreau & Lumbsch, 2015).

These SDL issues are highly relevant in the large and ecologically prominent temperate South American lizard genus Liolaemus (Aguilar et al., 2013), and in particular in the L. montanus group (Olave et al., 2014b). These are mainly viviparous lizards ranging from northern Argentina, Chile and Bolivia to central Peru, and from near sea level to more than 5000 m elevation (Aguilar et al., 2015). The group comprises 60 (24%) of the ∼250 known species in the genus (Uetz & Hosek, 2016). The northernmost Peruvian component of this group includes 12 recognized species (Fig. 1): L. annectens Boulenger, 1901, L. disjunctus Laurent, 1990; L. etheridgei Laurent, 1998; L. insolitus Cei & Pèfaur, 1982; L. melanogaster Laurent, 1998; L. ortizi Laurent, 1982; L. poconchilensis Valladares, 2004, L. polyxysticus Laurent, 1992; L. robustus Laurent, 1992; L. signifer (Dumeril and Bribon, 1837), L. thomasi Laurent, 1998 and L. williamsi Laurent, 1992. Most species descriptions in the northernmost species of this group have been based, at best, on only morphological data, and usually on a limited number of individuals from one or a few localities. In other cases, species descriptions were based on very small sample sizes or even a single specimen (e.g. L. ortizi and L. thomasi).

In addition to these issues, recent fieldwork and SDL studies have revealed examples of taxa representing a known species, but previously recognized as different based on a doubtful type locality (e.g. Liolaemus disjunctus; Aguilar et al., 2013). This kind of taxonomic error reflects the fact that new populations collected between type localities of known species are often difficult to identify based on the limited morphological characters of earlier studies. More complete geographic sampling and multiple lines of evidence often identify new lineages that were ‘hidden’ due to insufficiently informative phenotypic traits (Aguilar et al., 2013). Hypotheses of species limits based on adequate geographic sampling and multiple lines of evidence (molecular, ecological and morphological) are necessary for assigning populations to known species or for the discovery of new lineages as ‘candidate species’ requiring further study.

The Liolaemus montanus species group, like many others, exemplifies the need for a low-cost IT approach in megadiverse countries where research resources and infrastructure are limited, and immediate threats to biodiversity are an unfortunate reality. In the Peruvian Andes, habitat destruction and overexploitation are significant threats to some populations of the L. montanus species group, and some of these populations likely represent new species with restricted distributions. However, without formal descriptions and names, ‘cryptic diversity’, ‘candidate species’ and ‘distinct evolutionary lineages’ are not afforded legal protection or official recognition on species lists maintained by international conservation agencies (Pante et al., 2015). For instance, a recent Peruvian list of threatened species and IUCN evaluation of Andean squamates (lizard and snakes) shows an increase in the number species in the L. montanus group listed as either ‘threatened’ (L. insolitus and L. poconchilensis) or ‘near threatened’ (L. robustus and L. signifer) due to habitat destruction, pollution and overexploitation in their geographic ranges (Ministerio de Agricultura, 2014; IUCN unpubl. data). These same threats are likely present in areas inhabited by distinct lineages of unrecognized species, but without formal names and descriptions they cannot be included in current conservation planning.

Species descriptions based on IT analyses of multiple lines of evidence (molecular, morphological and bioclimatic data) can be implemented at minimal cost, and these descriptions are of higher quality than conventional descriptions based on a single line of evidence (e.g. morphology) that are sometimes without statistical support (Aguilar et al., 2013; Pante et al., 2015). The goal of this study is to delimit species boundaries in the northernmost taxa of the Liolaemus montanus group using IT step-by-step and model-based SDL procedures based on molecular, morphological and bioclimatic data. Specifically we would like to test if: (1) an Andean population identified as ‘Abra Apacheta’ and currently assigned to L. melanogaster, is in fact part of this lineage, or conspecific with its geographically closest species, L. polyxysticus; (2) L. ortizi and L. thomasi actually represent one or two lineages; (3) an Andean population...
called ‘Abra Toccto’ represents a distinct lineage; and (4) a low-elevation population from the Pacific Andean slopes (‘Nazca’) represents a new lineage. Formal taxonomic changes and species descriptions will be treated in separate papers.

MATERIAL AND METHODS

SAMPLING OF SPECIMENS
Specimens were collected from Liolaemus annectens, L. etheridgei, L. insolitus, melanogaster, L. ortizi, L. polystictus, L. robustus, L. signifer, L. thomasi and L. williamsi type localities, localities of paratypes if different from the type locality, and other locations which are represented by previous museum records (Fig. 1) or mentioned in taxonomic publications. Type specimens of L. ortizi, L. melanogaster, L. polystictus, L. robustus and L. williamsi and other museum specimens (Supporting Information, Appendix S1) were also examined and compared with collected specimens to propose initial species hypothesis and perform morphological analyses (see below).

DNA SAMPLING AND EXTRACTION
Lizards were collected by hand, photographed and euthanized with an injection of sodium pentobarbital. After liver and muscle tissues were collected for DNA samples, whole specimens were fixed in 10% formaldehyde, and transferred to 70% ethanol for permanent storage in museum collections. Tissue samples were collected in duplicate, stored in 96% ethanol and deposited at the M. L. Bean Life Science Museum at Brigham Young University (BYU) and Museo de Historia Natural de San Marcos (MUSM) in Lima, Peru, and voucher specimens were shared between these same institutions on a 50:50 basis. Total genomic DNA was extracted from liver/muscle tissue using the animal tissue extraction protocol in the Qiagen protocol (Qiagen Inc., Valencia, CA).

For in-groups and outgroups we used selected species of the subgenus Eulaemus that are assigned to different species groups and for which mtDNA sequences of cyt-b and 12S fragments are available from the GenBank database. Our ingroup included ten taxa that have been assigned to the Liolaemus montanus group: L. annectens, L. etheridgei, L. insolitus, L. melanogaster, L. ortizi, L. poconchilensis, L. robustus, L. polystictus, L. signifer, and L. williamsi (Lobo, Espinoza & Quinteros, 2010). To further resolve the relationships of the northernmost species of the L. montanus group, we sampled other species assigned to this species group (L. andinus Koslowsky, 1895, L. dorbignyi Koslowsky, 1898, L. famatinae Cei, 1980), the rothi complex (L. rothi Koslowsky, 1898), and the fitzingeri group (L. melanops Burmeister, 1888; Olave et al., 2014b). We used L. ornatus, a species belonging to the darwini group (Camargo et al., 2012) as the outgroup.
DNA amplification and sequencing

We sequenced part of the mitochondrial cyt-b gene (643 bp for 138 individuals from 31 localities; Supporting Information, Appendix S2). Redundant cyt-b haplotypes were identified using DnaSP v5 (Librado & Rozas, 2009), and individuals representing non-redundant cyt-b haplotypes were then sequenced for the mtDNA 12S region (~660 bp), and five nuclear gene regions, including: protein-coding KIF24 (440 bp), MAXRA5 (776 bp), EXPH5 (747 bp), and anonymous A12D (~580 bp) and A4B (~374 bp) DNA fragments. Individuals used for these fragments and sequencing primers are given in Supporting Information (Appendix S2) and Table 1, respectively. All new sequences are deposited in the GenBank database (accession numbers KX826506–KX826781; Supporting Information, Appendix S3) and alignments in Dryad. Double-stranded DNA polymerase chain reactions (PCR) amplified target regions under the conditions described in Aguilar et al. (2013) and Noonan & Yoder (2009), for mitochondrial and nuclear markers, respectively. PCR products were visualized on 10% agarose gels to ensure the targeted products were cleanly amplified, then purified using a MultiScreen PCR (μ) 96 (Millipore Corp., Billerica, MA), and directly sequenced using the BigDye Terminator v 3.1 Cycle Sequencing Ready Reaction (Applied Biosystems, Foster City, CA). The cycle sequencing reactions were purified using Sephadex G-50 Fine (GE Healthcare) and MultiScreen HV plates (Millipore Corp.). Samples were then analyzed on an ABI3730xl DNA Analyzer in the BYU DNA Sequencing Center.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Kind of marker</th>
<th>Substitution model</th>
<th>Primers</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYTB</td>
<td>mtDNA</td>
<td>HKY + G</td>
<td>IguA2F, IguA2R2</td>
<td>Corl et al. (2010)</td>
</tr>
<tr>
<td>12S</td>
<td>mtDNA</td>
<td>TrN + I + G</td>
<td>tphE, tphE</td>
<td>Wiens, Reeder &amp; De Oca (1999)</td>
</tr>
<tr>
<td>A4B</td>
<td>ANL</td>
<td>HKY</td>
<td>F, R</td>
<td>Camargo et al. (2012)</td>
</tr>
<tr>
<td>A12D</td>
<td>ANL</td>
<td>TPM2uf</td>
<td>F, R</td>
<td>Camargo et al. (2012)</td>
</tr>
<tr>
<td>EXPH5</td>
<td>Coding</td>
<td>HKY</td>
<td>F1, R1</td>
<td>Portik et al. (2012)</td>
</tr>
<tr>
<td>KIF24</td>
<td>Coding</td>
<td>HKY</td>
<td>F1, R2</td>
<td>Portik et al. (2012)</td>
</tr>
<tr>
<td>MXRA5</td>
<td>Coding</td>
<td>HKY</td>
<td>F, R</td>
<td>Portik et al. (2012)</td>
</tr>
</tbody>
</table>

Phylogenetic analyses

All sequences were aligned in the MUSCLE (Edgar, 2004) plug-in in GENEIOUS PRO v5.6.6 (Kearse et al., 2012), and protein-coding sequences were translated to check for premature stop codons. Bayesian Information Criteria in JMODELTEST v2.1.3 (Darriba et al., 2012) were used to identify the best-fit models of evolution. The concatenated mitochondrial fragments (cyt-b and 12S; 1298 nt, 63 individuals) were run in MRBAYES v3.2 (Ronquist et al., 2012). Two parallel runs were performed using four chains (one cold and three hot) for $1.1 \times 10^6$ generations with sampling every 200 generations from the Markov Chain Monte Carlo (MCMC) output. We determined stationarity by plotting the log likelihood scores of sample points against generation time; when the values reached a stable equilibrium and split frequencies fell below 0.01, stationarity was assumed. We discarded 100 000 samples and 10% of the trees as burn-in and a maximum clade credibility (MCC) tree was constructed using TREEANNOTATOR v2.1.2 (Bouckaert et al., 2014); we interpreted Bayesian posterior probabilities (PP) > 95% as evidence of significant support for a clade (Wilcox et al., 2002).

Multilocus concatenated and dating analysis

To estimate divergence times, we generated a concatenated tree that combined the mtDNA sequences and all nuclear region sequences using 117 terminals that include members of different species groups in the subgenus *Eulaemus*, and two species of the subgenus *Liolaemus* (Supporting Information, Appendix S3). Terminals include new sequences of 35 individuals of the northernmost species of the *L. montanus* group and sequences of 82 individuals downloaded from GenBank. Species of the subgenus *Liolaemus* were used as outgroups. We then calibrated the *Eulaemus* clade using a fossil (Albino, 2008) to date the divergence between *Liolaemus* (s.s.) and *Eulaemus* following Breitman et al. (2011) and Fontanella et al. (2012). This calibration prior was set to 20 Mya assuming a lognormal distribution and with a standard deviation of 0.13 (24.56–16.01), based on the recommendations of Ho (2007). This analysis was implemented in BEAST v1.8 (Drummond et al., 2012) and run for 100 million generations for each of ten independent runs. To check for convergence, we used Tracer v1.6 (Drummond et al., 2012) to ensure that all effective samples sizes (ESS) were greater than 200. We discarded 10% of the trees as burn-in and the remaining trees were combined using LogCombiner v1.8.0 and sampled at a lower frequency, resulting in 10 000 trees. An MCC
tree was then constructed using TreeAnnotator v1.8 (Drummond et al., 2012), and keeping mean heights.

Species tree analysis
A species tree analysis was also performed for mtDNA and all nuclear region sequences. We used 15 terminals representing taxa of the northernmost species of the Liolaemus montanus group and L. ornatus as the outgroup. Each nuclear DNA fragment was tested for presence of recombination using RDP v3.44 (Martin & Rybicki, 2000) and haplotypes of nuclear markers were phased using DnaSP v5 (Librado & Rozas, 2009). Each locus was included as a separate data partition (the two mitochondrial loci were linked) in an estimate of the species tree using *BEAST in BEAST v2.0 (Bouckaert et al., 2014). We used a relaxed log normal molecular clock model, a linear-with-constant-root model, and a Yule model for the species tree prior. Analyses were run for 100 million generations and samples taken every 4000 generations. We determined stationarity by plotting the log likelihood scores of sample points against generation time; when the values reached a stable equilibrium and split frequencies fell below 0.01, stationarity was assumed. We discarded 100 000 samples and 10% of the trees as burn-in, and constructed a MCC tree using TREEANNOTATOR v1.7.5 (Drummond et al., 2012). Analyses were run in the BYU Fulton Supercomputer Lab.

Tree distance and Rosenberg’s probability
We used the ‘species delimitation’ plug-in in the Genious software (Masters, Fan & Ross, 2011) as an exploratory tool to assess populations and ‘known’ species in our mitochondrial gene tree. This algorithm estimates average pairwise interspecific tree distance (ITD) and the Rosenberg probability, $\hat{P}_{ab}$, to test the null hypothesis that taxon A represented by a $a$ sequences is monophyletic, or in a clade of $a + b$ sequences the $a$ sequences will be reciprocally monophyletic with the remaining $b$ sequences, under a Yule model of random coalescence (Rosenberg, 2007). The rejection of the null hypothesis suggests that the random branching of the Yule model does not hold, perhaps because lineages were drawn from multiple genetically distinctive groups (Rosenberg, 2007). Specifically, we test whether monophyletic groups of populations might represent isolated lineages. We reject the null hypothesis of random branching when $P \leq 0.01$.

Morphological data and analyses
We collected three classes of morphological data from a total of 302 individuals (Supporting Information, Appendix S1). We scored the following 11 morphometric characters: (1; SVL) snout–vent length, (2; AGL) axilla–groin length (between the posterior insertion of forelimb and anterior insertion of thigh), (3; HL) head length (from snout to anterior border of auditory meatus), (4; HW) head width (at widest point), (5; FOL) forelimb length (distance from the attachment of the limb to the body to the terminus of the fourth digit), (6; HIL) hindlimb length (distance from the attachment of the limb to the body to the terminus of the fourth digit), (7; SL) snout length (from snout to anterior border of eye), (8; AMW) auditory meatus width, (9; AMH) auditory meatus height, (10; RW) rostral width, and (11; RL) rostral length. We also scored five meristic characters, including: (1; MBS) number of midbody scales (counted transversely at the middle of the body), (2; DTS) dorsal trunc scales (counted from the level of anterior border of the ears to anterior border of the thighs), (3; DHS) dorsal head scales (counted from the rostral scale to anterior border of ear), (4; VS) ventral scales (counted from the mental scales to the cloaca), and (5; SCI) number of scales in contact with the interparietal.

Measurements and counts were taken from the right side of the animal using a stereomicroscope. Morphometric data were only taken for adult males and females (adults were identified by size using the largest female and male for each species/population). We explored differences between sexes using Principal Component Analyses (PCA; Supporting Information, Appendix S4), and if sexes formed distinct clusters, we performed all subsequent analyses for males and females separately; otherwise data from both sexes were pooled. Correlation of morphometric characters was performed to avoid redundancy and variables with linear Pearson higher than 0.9 were discarded. Size correction was done using SVL as an independent variable and remaining morphometric characters as dependent variables in a multivariate linear model. We used unstandardized residuals of the linear model as variables. Correlation and linear model were performed in PAST v3.0 (Hammer, Harper & Ryan, 2001).

The third category of morphological data was head shape, as quantified using geometric morphometric methods. Ten landmarks on the dorsal head view of lizards (Supporting Information, Appendix S4) were set on digital pictures using tpsDig v1.4 (Rohlf, 2004), and shape analyses were performed using PCA after a Generalized Procrustes approach. Procrustes and PCA analyses were performed using MorphoJ v1.03d (Klingenberg, 2011), and PCA scores were extracted for further analyses (see below) using the Geomorph package (Adams & Otarola-Castillo, 2013) in R (R Core Team, 2014). We retained the
The first two principal components for all classes of data as they are used in the gap morphological analysis (see below). The morphological data are deposited in the MorphoBank database.

We inferred gaps in morphology for the three types of data (morphometric, meristic and head shape) as described in Zapata & Jimenez (2012). This method uses the multivariate morphological space derived by a PCA to estimate a ridgeline manifold, the corresponding probability density function (PDF), and ellipsoids of tolerance regions for each pair of samples to test for discontinuities in phenotypic values. We assumed normality for the three classes of morphological data, and used the principal components on correlation (for measurements and counts) and covariance (for head shape data) matrices as mentioned above. The ridgeline manifold is a surface image that includes the main characteristics (e.g. peaks and saddles) of a PDF in a mixed distribution and identifies the number of modes (Ray & Lindsay, 2005). A mixed distribution is used to model the multivariate data in a set of two components (two groups of samples) that might have more than one mode. If the ridgeline manifold of a PDF suggests that there is more than one peak for different values of a variable $\alpha$ (which varies from 0 at the multivariate mean of one component, to 1 at the multivariate mean of the other), then one can infer two modes and a gap in morphological space (Zapata & Jimenez, 2012).

When the PDF along the ridgeline manifold exhibits two modes, ellipsoids of tolerance regions for each component are estimated with different values of $\beta$ (a proportion of the multivariate distribution which varies from 0 to 1), and at fixed confidence level of 0.95 (Krishnamoorthy & Mondal, 2006; Krishnamoorthy & Mathew, 2009). Each tolerance region ellipsoid shares a single point along the ridgeline manifold (that corresponds to different values of $\alpha$) with another ellipsoid that defines a tolerance region for the other distribution (Zapata & Jimenez, 2012). Overlap of these ellipsoids for different proportions $\beta$ and values $\alpha$ along the ridgeline manifold can be visualized in a plot that shows the estimated phenotypic overlap between two samples. Following Wiens & Servedio (2000), we selected an a priori frequency cutoff of 10%, below which overlap of phenotypic values between samples indicates negligible gene flow. In other words, if the overlap in a plot is greater than $\beta = 0.9$, then the hypothesis that the sample of multivariate phenotypic values represents two taxa is supported.

Statistical analyses were performed using R packages ellipse (Murdoch & Chow, 2007), labdsv (Roberts, 2007), and mvtnorm (Genz et al., 2009). Although Liolaemus insolitus and L. poconchilensis are recognized as distinct species, they overlapped in most meristic data (see below) suggesting that pooling morphological data is justified. Additionally, sample sizes were small for Liolaemus insolitus and L. poconchilensis, hence data for these two species were pooled to compare with a similar taxon (Nazca).

**DISTRIBUTIONAL MODELS AND NICHE IDENTITY TESTS**

We used bioclimatic variables from the WorldClim v1.4 dataset with a resolution of 2.5 min (Hijmans et al., 2005) and to avoid over-parameterization of downstream analysis, we chose nine out of 19 variables that were not correlated with each other (Pearson coefficient $|r| < 0.7$). Bioclimatic variables were derived from monthly temperature and precipitation layers (Hijmans et al., 2005). Occurrence points without duplicates are: 13 for Nazca, ten for Abra Pampa, 12 for L. robustus, 22 for L. polystictus, 22 for Abra Tocoto, nine for (L. melanogaster + L. williamsi), nine for L. ortizi, and 11 for L. thomasi (Supporting Information, Appendix S5).

To visualize potential niche divergence between populations and species in the northernmost species of the Liolaemus montanus group we conducted a PCA using bioclimatic data derived from occurrence points. We then used the maximum entropy model implemented in the program MAXENT v3.3.3e (Phillips, Anderson & Schapire, 2006) to estimate potential distribution of lineages in the northernmost species of the Liolaemus montanus group. MAXENT generates distributional models (or ecological niche models; ENMs) using presence-only records, contrasting them with background/pseudoabsence data sampled from the remainder of the study area. We chose this approach because of its overall better performance with presence-only data and with small sample sizes (Elith et al., 2006). Because of small sample sizes some species occurrence points were pooled with closely related species (L. melanogaster + L. williamsi) enabling ENM development, but we were unable to develop ENMs for L. insolitus and L. poconchilensis because they are not hypothesized to be closely related.

Layers were trimmed to the areas surrounding each species or sample of populations that might represent candidate species, and then projected over a larger region that represents the whole geographic range of Peruvian species of the Liolaemus montanus group: $-10.793^\circ$ to $-18.543^\circ$ and $-75.423^\circ$ to $-70.009^\circ$.

For model calibration we used the default settings, but with a regularization multiplier of 2 to reduce overfitting (Radosavljevic & Anderson, 2014), with 1000 iterations, and the minimum training value averaged over the ten replicates as threshold with the default convergence threshold ($10^{-5}$). Due to our
small samples sizes, we used the cross-validation option with ten replicates for model calibration and evaluation, and averaged the results to estimate species niche and distributions. For model testing, we used occurrence points of closely related species (e.g. Liolaemus thomasi for L. ortizi and vice versa), or clades (e.g. L. melanogaster + L. williamsi for Abra Tocco and vice versa). We then used the area under the curve (AUC) to summarize the model’s ability to rank presence localities higher than a sample of random pixels (Peterson et al., 2011). AUC values ≤ 0.5 correspond to predictions that are equal or worse than random. AUC values > 0.5 are generally classed into: (1) poor predictors (0.5–0.7); (2) reasonable predictors (0.7–0.9); and (3) very good predictors (> 0.90; but see Peterson et al., 2011, for caveats on use of AUC in presence/background data). Model clamping (the process by which variables are constrained to remain within the range of values in the training data) was checked with the ‘fade by clamping’ option available in MAXENT v 3.3.3e.

Finally, the Schoener’s D metric was used as a measure of ‘niche similarity’ between pairs of populations (or species), and was estimated using the ENMTOOLS package (Warren, Glor & Turelli, 2010). We calculated these values by comparing the climatic suitability of each grid cell in the projected area obtained with MAXENT. This similarity measure ranges from 0 (niche ‘envelopes’ have no overlap) to 1 (niche ‘envelopes’ identical; Warren, Glor & Turelli, 2008). We estimated similarity measures and then tested whether the ENMs for two populations or species are ‘identical’ using the niche identity test in ENMTOOLS. One hundred randomly resampled pseudoreplicate data sets were generated to obtain a distribution of D scores under the null hypothesis that niche envelopes are random, and we reject the hypothesis of niche identity when the empirically observed value for D is significantly lower than the values expected from the pseudoreplicated data set (Warren et al., 2010).

**Gaussian Clustering**

For a small dataset of adults (N = 20 individuals) we used GC for our combined multilocus molecular, morphological (morphometric and meristic), and bioclimatic data. Most species and candidate species are known only from one or two localities, and bioclimatic data were redundant for most individuals within a locality, limiting the number of individuals that could be used for this method. We used the same individual for all datasets in most cases, but when this was not possible, we used another conspecific individual from the same locality. We used the same measurement and count variables as above, but two categorical variables were added: keeling in dorsal scales (absent/weak/strong), and enlarged ciliary scales (absent/present). We also used the 19 bioclimatic variables that were downloaded for each individual as mentioned above.

Euclidian and Gower distances were calculated for environmental and morphological data, respectively, using the cluster package in R (Maechler et al., 2015). Genetic distances were estimated using MEGA v. 6.06 (Tamura et al., 2013) with a Jukes-Cantor correction to account for multiple substitutions with substitution rates among sites following a Gamma distribution, and a Gamma parameter of 1. Genetic distances for individual loci were divided by mean pairwise distance to account for differences in substitution rates among loci, and individual distances were averaged across loci. Distance matrices for each data type were standardized using non-metric multidimensional scaling (NMDS) using the MASS package (Venables & Ripley, 2002). We followed the recommendations of Hausdorf & Hennig (2010) and chose four NMDS dimensions because these had stress values below 10% for each dataset, and were considered to be accurate estimates of clusters (but see Discussion). We concatenated the four NMDS dimensions of each dataset and estimated species groups using GC with the number of clusters determined by the Bayesian Information Criteria (BIC), using the mclust package (Fraley & Raftery, 2002). Noise (outliers) in the NMDS data was detected by the ‘noise’ estimator in prabclus (Hennig & Hausdorf, 2015), and for this we chose a tuning constant of 2 to detect clusters with few individuals.

**Integrative Taxonomic Procedure**

Assuming a General Lineage Concept (de Queiroz, 1998, 2007) and using molecular, morphological and niche envelope differences as criteria to delimit species, we implemented a step-by-step approach to evaluate four hypothesized alternatives of species limits in the northernmost taxa of the Liolaemus montanus group. We then used our time-calibrated concatenated and species tree analyses, as well as the model-based GC approach, to further evaluate our step-by-step results.

**Step-by-step approach**

First, field collected and museum specimens were initially identified and grouped based on type material and species descriptions. When a sample could not be assigned to any known species (e.g. Abra Tocco and Nazca), it was referred to by the name of the locality where it was first discovered. Nominal species and populations were then used as our primary species hypotheses. Second, we used the mtDNA
gene tree to identify the number of well-supported haploclades, and then used this topology to estimate interspecific tree distances between these groups. We tested for significant deviation of these groups under the null model of random coalescence using Rosenberg probabilities, and offered alternate species hypotheses from this test. Third, we used our morphological analyses to test the hypothesized species limits obtained in this second step, and last, we used the niche similarity test to evaluate the species hypotheses resolved in the second and third steps. Finally we integrated all evidence and designated candidate species.

**Species tree and dating analyses**

Relationships in our mitochondrial gene tree were evaluated using the time-calibrated concatenated tree and the multilocus species tree. These multilocus analyses were implemented to provide a plausible history of the group, and to incorporate this history as an integral part of the SDL approach.

**Comparison with Gaussian clustering and final delimitation**

We compare our previous results with those derived from the GC analyses, to further test the proposed candidate species based on the mitochondrial tree, and results from the time-calibrated concatenated and species tree analyses. This procedure leads to our best-supported species hypotheses, and also highlights the incongruence among evidence and methods (see Table 3 and Discussion).

**RESULTS**

**PRIMARY SPECIES HYPOTHESES**

In total 302 specimens were examined (Supporting Information, Appendix S3) and five primary species hypotheses are proposed. Three primary species hypotheses correspond to samples that could not be assigned to any known species (Abra Apacheta, Abra Toccto and Nazca; see Integrative Taxonomy, for further details). A fourth sample is slightly different from Liolaemus robustus and we call it L. robustus ‘Castrovirreyna’. One paratype of L. polystictus and our collected sample from the same locality are different from the holotype (and topotypes), and we call it L. polystictus ‘Castrovirreyna’. Species limits between these two last populations, L. robustus and L. polystictus will be treated in a separate paper.

**MITOCHONDRIAL TREE**

Our mitochondrial tree recovers all populations and all named species as monophyletic groups with high posterior probability (pp) support (= 1, Fig. 2) with the exception of Liolaemus polystictus and L. annectens; L. polystictus is well resolved as paraphyletic to Abra Apacheta with the structure: (L. polystictus ‘Castrovirreyna’ (L. polystictus + Abra Apacheta)) with nodal support value of pp = 0.98 (Fig. 2). This clade is the sister clade (pp = 1) to (L. robustus + L. robustus ‘Mina Martha’) (pp = 1). This larger clade is in turn the sister group (pp = 1) of a [Abra Toccto (L. melanogaster + L. williamsi)] clade (pp = 0.98), which we refer to as the L. robustus clade. This group is the sister clade to a (L. signifier (L. annectans ‘Lampa’ (L. annectans + L. etheridgei)))) clade (pp = 1 at the stem and all internal nodes). This large clade then forms an unresolved polytomy (pp < 0.9) with these other well supported clades: (L. ortizi + L. thomasi), L. dorbignyi, (L. andinus + L. famatinae), and (Nazca); pp = 1.0 for nodes of the three clades represented by two or more terminals. External to this larger clade is an unresolved polytomy with L. insolitus and two individuals of L. poconchilensis.

**DIVERGENCE ESTIMATES, CONCATENATION AND SPECIES TREE PHYLOGENIES**

Our concatenated (CT) and species tree (ST) analyses recovered topologies similar to the mtDNA gene tree, but with fewer strongly supported nodes and fewer paraphyletic terminals. Further, some relationships at deep nodes are more strongly supported in the ST relative to the CT (Fig. 3A, B; the complete tree of the dating analysis is shown in Supporting Information, Appendix S6). The CT (Fig. 3A) resolves the following clades with strong support: (L. poconchilensis), (Nazca), (L. ortizi + L. thomasi), (L. williamsi + L. melanogaster), (Abra Toccto), (L. robustus) and the L. robustus clade: ((L. williamsi + L. melanogaster) + Abra Toccto) + (Abra Apacheta (L. polystictus (L. polystictus ‘Castrovirreyna’))) + L. robustus).

In contrast, the ST recovers the two most deeply nested nodes with strong support, including (L. poconchilensis + (Nazca + L. insolitus + all Andean clades)) confirming paraphyly of the lowland groups. The Andean clade is not strongly supported, but well-supported nested clades include: the large (((L. robustus clade) + (L. signifier + (L. annectans + L. etheridgei)))) and external to this clade is a strongly supported (L. ortizi + L. thomasi) clade.

Our time-calibrated analysis corroborates this topology in suggesting that Andean taxa originated in the Pleistocene (< 3 Myr), and the older low-elevation lineages having a Pliocene (5–3 Mya) origin, albeit there is extensive overlap in the highest posterior density (HPD) error bars of these estimates (Fig. 3A).
INTERCLADE DISTANCE AND ROSENBERG’S PROBABILITY

All interclade tree distances (ITD) show values equal or higher than 0.03 with the exception of *Liolaemus ortizi* and *L. thomasi* (Table 2). ITD between all combinations of lowland taxa (Nazca, *L. insolitus*, *L. poconchilensis*) are equal to or higher than 0.09. The Abra Toccto clade has ITD values of 0.05 and 0.06 with *L. melanogaster* and *L. williamsi*, respectively. The ITD between Abra Apacheta and *L. polystictus* is 0.03, between (*L. polystictus ‘Castrovirreyna’* (Abra Apacheta + *L. polystictus*)) and *L. robustus* is 0.06. Rosenberg
probabilities are small (i.e. reject the null hypothesis of random monophyletic groups at \( P \leq 0.01 \)) for all pairwise comparisons except for \( L. ortizi \) vs. \( L. thomasi \), and \( Abra Apacheta \) vs. \( L. polystictus \) (Table 2).

### Table 2. Interspecific tree distances (ITD) and Rosenberg’s probabilities \( P(AB) \) based on mitochondrial markers between focal taxa and selected northernmost species of the \( L. montanus \) group

<table>
<thead>
<tr>
<th>Species</th>
<th>Closest species</th>
<th>ITD</th>
<th>Rosenberg’s ( P(AB) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L. thomasi )</td>
<td>( L. ortizi )</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Abra Apacheta</td>
<td>( L. polystictus )</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>( L. melanogaster )</td>
<td>( L. williamasi )</td>
<td>0.03</td>
<td>1.98E-03</td>
</tr>
<tr>
<td>( L. annectens )</td>
<td>( L. etheridgei )</td>
<td>0.03</td>
<td>1.00E-02</td>
</tr>
<tr>
<td>( L. melanogaster )</td>
<td>Abra Toccto</td>
<td>0.05</td>
<td>6.20E-07</td>
</tr>
<tr>
<td>( L. williamasi )</td>
<td>Abra Toccto</td>
<td>0.06</td>
<td>1.98E-03</td>
</tr>
<tr>
<td>( L. robustus )</td>
<td>( L. polystictus^{*} )</td>
<td>0.06</td>
<td>6.00E-04</td>
</tr>
<tr>
<td>( L. etheridgei )</td>
<td>( L. signifer )</td>
<td>0.08</td>
<td>5.10E-06</td>
</tr>
<tr>
<td>( L. insolitus )</td>
<td>( L. poconchilensis )</td>
<td>0.09</td>
<td>1.40E-05</td>
</tr>
<tr>
<td>Nazca</td>
<td>( L. insolitus )</td>
<td>0.11</td>
<td>5.00E-09</td>
</tr>
<tr>
<td>Nazca</td>
<td>( L. poconchilensis )</td>
<td>0.11</td>
<td>4.60E-04</td>
</tr>
</tbody>
</table>

The bold rows indicate ‘lowland’ from all Andean taxa. \( L. polystictus^{*} \) includes the clade (\( L. polystictus \) “Castrovirreyana” (Abra Apacheta + \( L. polystictus \))).

In gap analyses of morphometric and head shape data, \( L. orthiz\) and \( L. thomasi \) showed one mode (Supporting Information, Appendix S4), but two modes with the shape data (females only, Fig. 4D, E).

In gap analyses of morphometric, meristic and head shape data, Abra Toccto showed one mode with either \( L. melanogaster \) (Fig. 5A–F) or \( L. williamasi \) (Supporting Information, Appendix S4).

In gap analyses of meristic data, Nazca showed one mode with either \( L. insolitus \) or \( L. poconchilensis \), (Supporting Information, Appendix S4) but two modes with each species in our gap analyses of the morphometric (\( L. insolitus \) and \( L. poconchilensis \)) data were pooled; Supporting Information, Appendix S4) and head shape data (Fig. 6A–E).

In gap analyses of morphometric and head shape data, Nazca also showed one mode with either \( L. orthiz\) or \( L. thomasi \), but two clear modes with each species in our gap analyses of the meristic data (Supporting Information, Appendix S4).

In gap analyses of morphometric, meristic and head shape data, Nazca showed one mode with
Liolaemus robustus (Supporting Information, Appendix S4). In gap analyses of morphometric and meristic data, Nazca showed one mode with *L. polystictus*, but two clear modes with this species in our gap analyses of head shape data (Supporting Information, Appendix S4).

In gap analyses of morphometric and head shape data, Nazca showed one mode with Abra Apacheta, but two clear modes with this population in our gap analyses of meristic data (Supporting Information, Appendix S4).

In gap analyses of morphometric, meristic and head shape data, Nazca showed one mode with *Abra Toccto* (Supporting Information, Appendix S4). In gap analyses of head shape data, Nazca showed one mode with *L. williamsi*, but two modes with this species in our gap analyses of morphometric and meristic data (Supporting Information, Appendix S4).

In gap analyses of meristic data, Nazca showed one mode with *L. melanogaster*, but two modes with this species in our gap analyses of morphometric and head shape data (Supporting Information, Appendix S4).

DISTRIBUTIONAL MODELS AND NICHE IDENTITY TESTS

The first two principal components (PC) of the bioclimatic variables explained 99.7% of the variance in the data. The variables (Supporting Information, Appendix S5) contributing to most of the variation in both PCs are Temperature Seasonality (BIO4) and Annual Precipitation (BIO12). The PC plot (Supporting Information, Appendix S5) shows a clear break between lowland (Nazca, *Liolaemus poconchilensis* and *L. insolitus*) and Andean taxa (Abra Apacheta, Abra Toccto, *L. melanogaster*, *L. ortizi*, *L. polystictus*, *L. robustus*, *L. robustus* from Minas Martha, *L. thomasi* and *L. williamsi*).

All distributional models show AUC values > 0.90 with the exception of *Liolaemus ortizi* (AUC = 0.7284). Projections of niche models are shown in Supporting Information.
Information (Appendix S5). Niche identity tests show that the observed Schoener’s D metric of Abra Apacheta vs. L. polystictus (Fig. 7A), and of L. ortizi vs. L. thomasi (Fig. 7B) fall within the distribution of the pseudoreplicates, i.e. niche envelopes do not differ in either comparison. In contrast, niche identity tests show that the observed Schoener’s D metric of Abra Toccto vs. (L. melanogaster + L. williamsi) (Fig. 7C) fall outside the distribution of the pseudoreplicates, i.e. niche envelopes differ between these lineages.

**Gaussian Clustering**

NMDS stress values for each data type were below 5%. The best model (BIC = −246.5591) had six clusters: (1) Nazca; (2) Liolaemus poconchilensis; (3) L. ortizi, L. thomasi, and L. annectens; (4) L. melanogaster, L. williamsi, L. polystictus, Abra Apacheta, and L. signifer; (5) Abra Toccto and L. robustus ‘Mina Martha’; and (6) L. polystictus ‘Castrovirreyna’. Taxa and number of individuals identified as noise were L. insolitus (1), L. robustus (2), and L. etheridgei (1). Using the step-by-step approach as a benchmark, 69% of the individuals were correctly identified.

**Integrative Taxonomy**

Table 3 shows candidate species delimited through the step-by-step, concatenation, species tree analysis, and Gaussian clustering, and our consensus delimitation proposal. We summarize each final delimitation case below.

**Abra Apacheta**

Two individuals (one male adult and one juvenile) from Abra Apacheta were included as part of the paratype series in the species description of Liolaemus melanogaster (Laurent, 1998), but this locality is geographically closer to L. polystictus (Fig. 1). When these paratypes are compared with our collected samples, our primary hypothesis is that
**Liolaemus** individuals from this locality represent a population that cannot be assigned to *L. melanogaster*, *L. polystictus* or any known species. We identified this population ‘Abra Apacheta’ and the mtDNA gene tree recovers Abra Apacheta as the sister clade to *L. polystictus* (pp = 0.99), and distant from *L. melanogaster* by five strongly supported nodes (Fig. 2). Interclade tree distance (ITD) between Abra Apacheta and *L. polystictus* is similar to ITD between *L. melanogaster* and *L. williamsi* (0.03) from type localities, and lower than for all other pairs except for *L. ortizi* and *L. thomasi* (Table 2). Conversely, Rosenberg’s probability is not significant between Abra Apacheta and *L. polystictus* suggesting that separation of these taxa is random (Table 2).

Gap analyses of meristic data reveal separation in the multivariate space between Abra Apacheta and *Liolaemus polystictus*, but there is overlap in their tolerance regions (Fig. 4A–C), and niche identity tests give an Abra Apacheta–*L. polystictus* Schoener’s value within the distribution of the pseudoreplicate values (Fig. 7A). The CT and ST analyses recover Abra Apacheta grouped with *L. polystictus* ‘Castrovirreyna’, *L. polystictus* and *L. robustus*, but without significant support (Fig. 3A, B). Gaussian clustering of the concatenated four dimensions of each data set groups Abra Apacheta with *L. polystictus*, *L. williamsi* and *L. melanogaster*. Almost all available evidence suggests that Abra Apacheta should be a considered a distinct lineage related to (or conspecific with) *L. polystictus*, but not conspecific with *L. melanogaster* (Table 3).

**Liolaemus thomasi**

This species was described from a single specimen (Laurent, 1998), and is geographically close to *L. ortizi* (Fig. 1). However, our primary species hypothesis is that our collected topotypes should be considered *L. thomasi*. The mtDNA gene tree recovers *L. ortizi* and *L. thomasi* as distinct haploclades each
with strong support and also as sister groups, but with the lowest ITD and Rosenberg's probability was not significant (Fig. 2, Table 2). Morphological gap analyses of shape data separate *L. thomasi* and *L. ortizi* in multivariate space, but with overlap in their tolerance regions (Fig. 4D–F), and the niche identity test recovers a Schoener's value between these species within the distribution of the pseudoreplicate values (Fig. 7B). The CT and ST analyses recover *L. thomasi* and *L. ortizi* as sister clades with high support (Fig. 3). Gaussian clustering of the concatenated four dimensions of each data set groups *L. thomasi* with *L. ortizi*, but also with *L. annectens*. In summary, all available evidence suggests that *L. thomasi* should either be considered a distinct lineage or conspecific with *L. ortizi*.

**Abra Toccto**

Our primary species hypothesis is that field collected and museum specimens of this locality form a distinct population. The mtDNA gene tree recovers the Abra Toccto samples as a well supported haploclade (pp = 1.0), and sister group to a *L. melanogaster + L. williamsi* clade (pp = 0.98; Fig. 2). This lineage also has a larger ITD with both *L. melanogaster* and *L. williamsi* than the ITD between these two last taxa, and significant Rosenberg probabilities separating it from these two species (Table 2). Gap analyses of morphometric, meristic and head shape data does not reveal any separation in multivariate space between Abra Toccto and *L. melanogaster*, and Abra Toccto and *L. williamsi* (Fig. 5). Niche identity tests give a Schoener's value for Abra Toccto vs. (*L. melanogaster + L. williamsi*) which falls outside of the pseudoreplicate values (Fig. 7C).

CT analysis recovers all Abra Toccto individuals as a strongly supported clade, and these are also recovered in the ST analysis (Fig. 3B). However, in both analyses there is only weak support for Abra Toccto as the sister group to the (*L. melanogaster + L. williamsi*) clade (pp < 0.9 in both; Fig. 3). Gaussian clustering of the concatenated four dimensions of each data set shows that Abra Toccto forms a distinct group from *L. melanogaster* and *L. williamsi*, but it also grouped with *L. robustus* ‘Minas Martha’. In this example, niche identity tests and all phylogenetic and species tree analyses suggest that Abra Toccto is an independent lineage.

**Nazca**

Our primary species hypothesis is that field collected and museum specimens from Nazca form a distinct population. The mtDNA gene tree recovers all Nazca individuals as a clade with high support, and it falls outside of the well supported clade that includes all other taxa and populations of our ingroup (Fig. 2).
Further, CT and ST analyses recover this clade well outside of strongly supported more nested clades, although fewer of these nested nodes have significant support in the concatenated than in the species tree (Fig. 3A, B). This sample also has a larger ITD in comparison with other old lowland Liolaemus (L. insolitus and L. poconchilensis, Table 2), and significant Rosenberg probabilities with both L. insolitus and L. poconchilensis (Table 2). Morphological gap analyses show a separation in the multivariate head shape space between Nazca and L. insolitus, and Nazca and L. poconchilensis, but with overlap in the tolerance regions of both paired tests (Fig. 6). There are also two modes in multivariate meristic space between Nazca vs. L. ortizi, L. thomasi, Abra Apacheta and L. williamsi (Supporting Information, Appendix S4); two modes in multivariate head shape space between Nazca vs. L. polystictus and L. melanogaster (Supporting Information, Appendix S4); and two modes in multivariate morphometric space between Nazca vs. L. williamsi and L. melanogaster (Supporting Information, Appendix S4). However, overlap in tolerance regions is present in all paired tests (Supporting Information, Appendix S4). Gaussian clustering of the concatenated four dimensions of each data set shows that Nazca forms a distinct group from all other species and populations. Most available evidence and methods (mitochondrial, concatenated, species trees and Gaussian clustering) suggest that Nazca is an independent lineage.

DISCUSSION

Robust hypotheses of species boundaries come from the inference of using multiple operational (empirical) criteria (Leavitt et al., 2015). Different operational criteria emphasize the many contingent properties (monophyly, differences in morphological features, ecological niches, etc.) of diverging populations associated with the various evolutionary processes operating in various geographic contexts (de Queiroz, 2005a; Camargo & Sites, 2013). In contrast, using a single empirical criterion might artificially reduce the complexity of evolving lineages (de Queiroz, 2005b). In addition, the General Lineage Concept (de Queiroz, 1998, 2007) explicitly recognizes a ‘grey zone’ or fuzzy boundary where populations in various stages of divergence have not fully completed a speciation process, and under which all methods for delimiting species will occasionally fail or be discordant with each other (Sites & Marshall, 2003, 2004). However, an IT approach can provide evolutionary explanations for discordant species criteria and uncover complex evolutionary histories (Dejaco et al., 2016; Karanovic, Djurakic & Eberhard, 2016).
Here we used different delimitation criteria within an integrative taxonomy framework to test for species limits in the northernmost lineages of the Liolaemus montanus group. We have found that the Abra Apacheta population, previously recognized as conspecific with L. melanogaster, is part of a different clade. In this case, the mitochondrial tree, Rosenberg probability and niche identity tests recover Abra Apacheta as conspecific with L. polyistictus. In the same way, the mitochondrial tree distance, Rosenberg probability and niche identity tests recover L. thomasi as conspecific with L. ortizi. However, in both of these comparisons, relationships are not congruent (or not supported by) the concatenated and species tree analyses. This incongruence between data sets can reflect intricate evolutionary histories and can be explained considering Abra Apacheta as well as L. thomasi representing ‘grey zone’ lineages that have split from a common ancestor, but may not have fully diverged to the level of separate species. Although this is a limitation of IT approach to delimit species, it actually reflects the fact that speciation is a continuous process along different axes of divergence and for this reason some evolutionary entities will often be truly indistinct (Hey et al., 2003, 2004; Nosil, Harmon & Seehausen, 2009; Huang & Knowles, 2016).

In contrast, the Abra Toccto and Nazca samples represent lineages at more advanced stages of speciation. Most empirical criteria and methods, particularly for Nazca, support these populations as distinct lineages. The Nazca population was also recovered as a significantly distinct group in the model-based GC algorithm, while Abra Toccto was not distinguished using this method. With the exception in the Nazca case, GC did not recover any other candidate species or isolated populations as distinct, even when they were recovered as separate lineages in our step-by-step, concatenated, and species tree methods. This lack of correspondence between GC and other approaches might be due to the small sample sizes we used for the four NMDS dimensions (1–3 individuals per species or population), in contrast to what is advised for this method (a minimum of five individuals; Hausdorf & Hennig, 2010).

Further, some populations or species are known from only single localities, thereby compromising the collection of sufficient bioclimatic, morphological or genetic data for this method. Additionally, GC might be better applied to closely related species or populations (species complex) than to taxa belonging to different species complexes, as is the case in our study (Edwards & Knowles, 2014; Hausdorf & Hennig, 2010). Moreover our results could simply be an idiosyncratic limitation of GC for this particular group of lizards, yet we still rely upon qualitative judgements of species boundaries in a hypothetico-deductive framework, even when IT model-based approaches are used (Sites & Marshall, 2004; Yeates et al., 2011; Hausdorf & Hennig, 2010).

Our morphological gap analyses show different degrees of overlap using three different types of data [morphometric, meristic, and geometric morphometric (head shape)]. In these analyses Nazca, recovered as an older lineage in the time-calibrated concatenation analysis, shows the smallest degree of overlap among all of our paired gap analyses. Gap analyses between this taxon and lowland (L. insolitus, L. poonchilensis) and some Andean lineages, (L. ortizi, L. thomasi, L. polyistictus, Abra Apacheta, L. williamsi and L. melanogaster) show two modes (Fig. 6, Supporting Information, Appendix S4) in comparison with the same analyses for other Andean taxa where a single mode is recovered (Figs 4, 5, Supporting Information, Appendix S4). However, tolerance regions overlapped between Nazca and these taxa probably due to small sample sizes. The general pattern here is that delimitation between the oldest lineages (4–5 Myr) and youngest lineages (1–2.5 Myr) in our gap analyses also show that, despite other criteria which clearly separate these lineages, morphological features commonly used in Liolaemus taxonomy do not differentiate taxa even when species appear to have had sufficient time to acquire morphological discontinuities. Alternatively, many species of Liolaemus may be under selective constraint, which is a particularly important point for taxonomic studies of Liolaemus, one of the most remarkable species-rich temperate lizards genera on earth. Until recently most species descriptions have been based only on gross comparisons of morphological features, and lacking statistical rigor (e.g. Cei & Pèfaur, 1982; Laurent, 1982, 1990, 1992, 1998).

In contrast, the gap analyses we used here is probably one of the most comprehensive methods to detect species limits with morphological data and statistical rigor. It requires large samples sizes to infer no overlap (a gap) in the multivariate space of two taxa for 0.9 proportions of each statistical population (a surrogate for inferring limited gene flow) and at a confidence level of 95% (Zapata & Jimenez, 2012). Other univariate or multivariate methods that have been used to detect species limits with morphological data based on central tendencies (e.g. discriminant analyses, ANOVA) might have given statistically significant results, but they might not be appropriate as delimitation criteria for quantitative phenotypic characters. In an earlier study, we have shown that central tendency univariate methods for morphological characters might be misleading (even with statistically significant results) giving the
impression that a species have diverged to the point that diagnostic phenotypic characters have evolved (Aguilar et al., 2013).

Although morphological gap analyses were uninformative for all four test of species boundaries in this study, ENMs and niche identity tests were useful to test species boundaries based on other lines of evidence (Fig. 7, Supporting Information, Appendix S5). However, the use of ENMs as non-heritable surrogates for ecological phenotypes in SDL studies has been questioned. Some researchers interpret these data as having limited (Tocchio et al., 2014) or no relevance (Meik et al., 2015). Whereas others show that ENMs provide an independent line of evidence in SDL studies and support of hypotheses based on heritable traits (e.g. Pelletier et al., 2015; Huang & Knowles, 2016). We acknowledge the caution with which ENM data should be used in SDL studies, but these data might be useful as a proxy of some aspect of a species eco-physiological niche, and therefore relevant for studies of species boundaries (Warren, 2012).

We have shown, as have a growing number of previous studies in Liolaemus (Aguilar et al., 2013; Medina, Avila & Morando, 2013; Medina et al., 2014; Minoli, Morando & Avila, 2014; Breitman et al., 2015), that an IT approach (together with evolutionary explanations in cases where different lines of evidence disagree) provides an empirically richer means of delimiting species, thereby improving the quality of species hypotheses, and associated descriptions.

ACKNOWLEDGEMENTS

We thank J. Córdova, C. Torres (MUSM), A. Resetar (FMNH), J. Losos, J. Rosado (MCZ), F. Glaw (ZSM), L. Welton and R. Brown (KU) for loans and accessions of specimens under their care. We thank Dr. D. Edwards for advice on Gaussian clustering, and C. Ramirez, C. Salas, A. Guzman, A. Mendoza, V. Vargas, F. Huari and J. C. Cusi for assistance in the fieldwork. We also thank an anonymous reviewer, J. A. Allen and C. L. Malone for improving with their comments a previous version of this paper. Fieldwork was supported by the Wait Foundation-National Geographic Society (award W195-11 to CA and JWS), the BYU Bean Life Science Museum (JWS), and lab work by NSF-Emerging Frontiers award (EF 1241885 to JWS), Dr. C.G. Sites to JWS, and NSF-Doctoral Dissertation Improvement Grant (award #1501187 to JWS and CA). Permits (RD N’ 1280-2012-AG-DGFFS-DGEFFS, RD N’ 008-2014-MINAGRI-DGFSS-DGEFFS) were issued by the Ministerio de Agricultura, Lima, Peru, and the work was approved by the BYU Institutional Animal Care and Use Committee protocol number 12001 and in accordance with US law.

REFERENCES


Corl A, Davis AR, Kuchta SR, Comendant T, Sinervo B. 2010. Alternative mating strategies and the evolution of sexual size dimorphism in the side-blotched lizard, Uta


**SUPPORTING INFORMATION**

Additional Supporting Information may be found online in the supporting information tab for this article:

**Appendix S1.** Specimens used in this study for morphological analyses (if yes marked with an X) including museum number, types, sex, reproductive stage for each individual.

**Appendix S2.** Tissues used in this study including molecular markers (if yes marked by an X), museum numbers and locality for each individual.

**Appendix S3.** GenBank accession numbers used in this study.

**Appendix S4.** Landmarks used in shape analysis and other gap analyses.

**Appendix S5.** Geographic information, principal component analysis of bioclimatic data and Maxent projections of selected species.

**Appendix S6.** Extended time calibrated phylogenetic tree including 117 terminals.