

Sampling Strategies for Delimiting Species: Genes, Individuals, and Populations in the *Liolaemus elongatus-kriegi* Complex (Squamata: Liolaemidae) in Andean–Patagonian South America

MARIANA MORANDO,^{1,2} LUCIANO J. AVILA,^{1,2} AND JACK W. SITES JR.¹

¹Department of Integrative Biology and M. L. Bean Life Science Museum, Brigham Young University, Provo, Utah 84602, USA;
E-mail: mariana@email.byu.edu (M. M.)

²CONICET, Mendoza y Entre Rios s/n 5301, Anillaco, La Rioja, Argentina

Abstract.—Recovery of evolutionary history and delimiting species boundaries in widely distributed, poorly known groups requires extensive geographic sampling, but sampling regimes are difficult to design a priori because evolutionary diversity is often “hidden” by inadequate taxonomy. Large data sets are needed, and these provide unique challenges for analysis when they span intra- and interspecific levels of divergence. However, protocols have been designed to combine methods of analysis for DNA sequences that exhibit both very shallow and relatively deeper divergences. In this study, we combined several tree-based phylogeny reconstruction methods with nested-clade analysis to extract maximum historical signal at various levels in the poorly known *Liolaemus elongatus-kriegi* lizard complex in temperate South America. We implemented a recently described tree-based protocol for DNA sequences to test for species boundaries, and we propose modifications to accommodate large data sets and gene regions with heterogeneous substitution rates. Combining haplotype trees with nested-clade analyses allowed testing of species boundaries on the basis of a priori defined criteria. The results obtained suggest that the number of putative species in the *L. elongatus-kriegi* complex could be doubled. We discuss these findings in the context of the advantages and limitations of a combined approach for retrieval of maximum historical information in large data sets and with reference to the yet formidable unresolved issues of sampling strategies. [*Liolaemus*; lizards; mitochondrial DNA; nested-clade analysis; phylogeny; sampling design; species boundaries.]

Two frequently stated goals of systematic biology are to discover relationships among species and to discover, delimit, and describe species (Wheeler and Meier, 2000). A large body of theory and many methods are available to estimate phylogenetic relationships and test specific evolutionary hypotheses; however, empirical delimitation of species remains a focus of discussion and controversy, and few explicit discovery methods have been proposed (Avice and Ball, 1990; Davis and Nixon, 1992; Baum and Shaw, 1995; Mallet, 1995; Brower, 1999; Wiens, 1999; Wiens and Servedio, 2000; Puerto et al., 2001; Templeton, 2001; Wiens and Penkrot, 2002). Evolutionary relationships above and below the species level are different in nature, yet simultaneous resolution of both within a single clade may be necessary to formulate robust hypotheses about issues ranging from mechanisms of speciation to genotype-phenotype associations (Posada and Crandall, 2001). Traditional methods for inferring interspecific relationships make assumptions that are violated at the population level and therefore lose statistical power as terminals approach intraspecific levels of divergence (Crandall and Fitzpatrick, 1996; Clement et al., 2000; Templeton, 2001).

In poorly known groups, initial inferences about species boundaries are almost always based on a subset of characters and methods of analysis that by contemporary standards would be judged inadequate, and this problem is often compounded by inadequate sampling of specimens and localities. Species described on this basis fit what Good (1994:194) has termed the “inertial species concept” whereby “species limits are set solely by historical precedence.” Given that dense population sampling is required to provide statistical power for resolution of species boundaries and other evolutionary questions (Templeton, 1998, 2001), designing an a priori sampling protocol on the basis of a dated alpha tax-

onomy becomes an immense challenge. These issues are further complicated in groups that are widely distributed across topographically diverse landscapes in which limited access precludes randomized sampling of taxa. For these groups, one may have little choice but to begin with the existing alpha taxonomy and work incrementally toward better resolution (Hedin and Wood, 2002).

From a molecular perspective, geographic sampling of poorly known groups may include haplotypes spanning the range of genetic divergence both within and between species. This sampling design vastly increases the number of possible tree topologies with even modest sequence divergence within and/or between localities and makes phylogenies computationally difficult to estimate (Crandall and Fitzpatrick, 1996; Swofford et al., 1996). Crandall and Fitzpatrick (1996; hereinafter CF) recently showed that using a combination of networks for population-level phylogenies with either maximum parsimony (MP) or maximum likelihood (ML) tree-reconstruction methods improves DNA sequence-based phylogenetic estimates when haplotypes span a range of divergences.

The CF strategy can now be analytically extended above and below the species boundary by algorithms that were unavailable to those authors. Bayesian methods (Rannala and Yang, 1996; Yang and Rannala, 1997; Larget and Simon, 1999), for example now enable efficient phylogenetic analysis of large data sets by incorporation of Markov Chain Monte Carlo (MCMC) algorithms within an explicit likelihood framework (Lewis, 2001), which appear to provide good approximations of joint probability distributions of nodal support (Huelsenbeck and Bollback, 2001; Huelsenbeck et al., 2001; Leaché and Reeder, 2002) conditional on the ML model. For population level genealogies, several algorithms for inferring demographic processes are

available (Posada and Crandall, 2001), and at least one of these, the network algorithm of Templeton et al. (1992), can be incorporated into a nested-clade analysis (NCA; Templeton et al., 1995) to provide a framework for distinguishing historical from ongoing demographic processes at the intraspecific–interspecific interface.

Wiens and Penkrot (2002) recently combined tree estimation and NCA of mitochondrial DNA (mtDNA) haplotypes to assess species boundaries in lizards of the poorly known *Sceloporus jarrovi* complex, (see Wiens et al., 1999, for details). These authors were to our knowledge the first to emphasize the advantage for species delimitation of rapid sorting to monophyly of the mtDNA locus (but see their caveats) versus its use in phylogeny estimation (Moore, 1995). Their method (referred to hereinafter as WP), however, describes a general tree-based inference protocol estimated from any nonrecombining sequence (Wiens and Penkrot, 2002: fig. 1).

Here, we combined methods for efficient model-based tree estimation with NCA to retrieve deep and shallow phylogenetic histories and to explore species boundaries in South American lizards of the *Liolaemus elongatus-kriegi* complex (*sensu* Cei, 1979; Ortiz, 1981). We implemented the basic protocol of Wiens and Penkrot (2002) but with two modifications. First, because our molecular data set was larger than that of Wiens and Penkrot, we devised a protocol to sample sequence variation in multiple mtDNA gene regions (from all localities) to determine the region with greatest haplotype variability. The advantage here is that because the genes are linked, identification of haplotypes that are exclusive for one gene will be exclusive for all (Hudson and Coyne, 2002). However, use of several gene regions varying in substitution rates permits estimation of different parts of the gene genealogy, and slower regions need not be sequenced for all individuals. The most variable region is then sequenced in almost all individuals representing all focal and nonfocal taxa in the study and is used in a first-round assessment of clade exclusivity across all hierarchical levels.

The second step is to use the exclusive groups estimated from a phylogeny of the most variable gene region to define a subset of samples from which the more conserved gene regions are sequenced and then to combine these data sets to formulate initial hypotheses of species boundaries. Our hypotheses of species limits are provisional because of our focus primarily on mtDNA sequences, but we emphasize our agreement with many others (Wiens and Penkrot, 2002, and references therein) that the strongest evidence for delimiting species should be based on concordance of independent data sets (we have advocated and implemented this view elsewhere; Sites and Crandall, 1997; Benavides et al., 2002). Here, we follow the protocols of Wiens and Penkrot (2002) and Templeton (2001) to hypothesize species boundaries because both protocols are based on a degree of statistical rigor that readily accommodates subsequent testing with independent data sets. The “gene” sampling design described here should be useful with increasingly large molecular data sets, a point important in the context of

the often neglected issues of individual and population sampling for testing species boundaries.

THE GENUS *LIOLAEMUS* AS A MODEL SYSTEM

The genus *Liolaemus* is the second most species-rich genus of New World lizards. It includes more than 160 described species, of which at least 36 have been described in the last decade. The genus is distributed over a wide geographic region and occupies a large range of latitudinal (14°30′–52°30′S), altitudinal (0–4,500 m), and climatic regimes, from the extremely arid Atacama desert to temperate *Nothofagus* rainforest (Hellmich, 1951; Donoso-Barros, 1966; Cei, 1986, 1993; Etheridge and de Queiroz, 1988; Frost and Etheridge, 1989; Etheridge and Espinoza, 2000; Lobo, 2001). The wide geographic range of many species groups in combination with extreme morphological polytypy makes the delimitation of species boundaries extraordinarily difficult.

The first comprehensive taxonomic studies of *Liolaemus* were performed by Ortiz (1981) and Laurent (1983, 1984, 1985), but intrageneric relationships remain poorly known (Etheridge, 1995; Schulte et al., 2000; Lobo, 2001). The *L. elongatus-kriegi* complex, defined by Cei (1979) on the basis of several diagnostic morphological characters, is nested within the “*chiliensis* group” (Etheridge, 1995) and includes widespread polytypic species for which species recognition rests on an inadequate taxonomy. Different taxonomic groupings for this complex have been proposed by Cei (1974, 1979, but also see 1975, 1986, 1993), Ortiz (1981), Etheridge (1995), Espinoza et al. (2000), Schulte et al. (2000), and Lobo (2001) and are summarized in Table 1.

The *elongatus-kriegi* complex is composed of medium-sized (up to 112 mm snout–vent length), saxicolous, viviparous, and insectivorous or omnivorous lizards found in rock outcrops between 350 and 3,900 m above sea level. Latitudinally, the complex extends from 27°S (near the San Francisco Pass in Catamarca Province; Avila and Lobo, 1999; Avila, unpubl. data) south along a narrow strip with an east–west width of ≤ 200 km (Cei, 1974, 1986; Avila, unpubl. data), and includes a small distribution in central Chile (Donoso Barros, 1966; Nuñez, 1992; Nuñez and Torres-Mura, 1992) but with a broader east–west spread in the Payunia and Patagonian steppes (Cei, 1986; Scolaro, 1993; Avila and Morando, unpubl. data). The southern distributional limit extends to the northern edge of Santa Cruz province at 46°S (Avila, unpubl. data), and the complex reaches the extra-Andean Sierras del Aconquija and Cumbres Calchaquíes in Catamarca and Tucumán Province (Espinoza et al., 2000) and the Sierra de Famatina and Sierra del Velazco ranges in La Rioja Province (Cei, 1993, Avila and Morando, unpubl. data; Fig. 1).

MATERIALS AND METHODS

Taxon Sampling and Outgroup Choice

Mitochondrial DNA sequence data were collected from a total of 207 lizards, of which 198 samples from

TABLE 1. Simplified history of the proposed taxonomic groupings for the *Liolaemus* species included in the *elongatus-kriegi* complex. With each author is the original name used for the groupings, and the species in each group are indicated (X).

Study	<i>Liolaemus</i> species										
	<i>elongatus</i> ^a	<i>petrophilus</i> ^a	<i>austromendocinus</i> ^a	<i>capillitas</i> ^a	<i>thermarum</i>	<i>heliodermis</i>	<i>kriegi</i> ^a	<i>ceii</i> ^a	<i>buengeri</i> ^a	<i>cristiani</i>	<i>leopardinus</i>
Cei, 1974; <i>L. elongatus</i> group	X	X	X								
Cei 1975; "grupo evolutivo" <i>L. elongatus</i> set + <i>L. buengeri</i> + <i>L. kriegi</i> set	X	X	X				X	X	X		
Cei, 1979; <i>L. elongatus-kriegi</i> complex	X	X	X				X	X	X		
Ortiz, 1981; <i>L. elongatus-kriegi</i> group	X	X	X				X	X	X		
Cei, 1986; <i>L. elongatus</i> group and <i>L. kriegi</i> group	X	X	X				X	X	X		
Cei, 1993; <i>L. elongatus</i> group	X	X	X	X							
Espinoza et al., 2000; <i>L. elongatus</i> group	X	X	X	X	X	X					
Schulte et al., 2000; Andean/eastern lowland clade	X	X	X	X				X	X		X
Lobo, 2001; <i>L. elongatus</i> group and <i>L. kriegi</i> group	X	X	X	X			X		X	X	

^aSpecies included in this study.

49 localities represented the majority of the species of the *L. elongatus-kriegi* complex. Populations originally recognized under the names *L. elongatus* Koslowsky, 1896, *L. kriegi* Müller and Hellmich, 1939, *L. ceii* Donoso-Barros, 1971, *L. buengeri* Werner, 1907, and *L. petrophilus* Donoso-Barros and Cei, 1971 were the focal species of this study (Wiens and Penkrot, 2002); *L. capillitas* Hulse, 1979, *L. austromendocinus* Cei, 1974, and two undescribed species (*L. sp. 1*, *L. sp. 3*) were the ingroup nonfocal species. We follow Cei's (1974) definition of the complex (Table 1) on the basis of morphological evidence. The poorly known species *L. thermarum* Videla and Cei, 1996 and *L. heliodermis* Espinoza et al., 2000 are proposed as part of this complex, but sampling from the type locality was not possible for *L. thermarum*, and at the type locality of the *L. heliodermis* we were unable to obtain samples. Other species could be included in this complex (e.g., *L. leopardinus* and related forms, *L. cristiani*, *L. curis*, Nuñez and Labra, 1985; Schulte et al., 2000; Lobo, 2001), but their relationships with this complex will be part of a later study.

Nine species were used as outgroups, with eight representing *Liolaemus* species from other groups (also nonfocal species) used to test the monophyly of the *L. elongatus-kriegi* complex, and *Phymaturus indistinctus* (considered the sister genus of *Liolaemus*; Etheridge, 1995) was used to root all trees, thus allowing the position of the outgroup nonfocal *Liolaemus* species to remain unconstrained with respect to the ingroup. The number of lizards collected per locality was maximized when possible within permit limits (which varied by province but averaged four or five individuals per species per locality). Table 2 summarizes the number of individuals sequenced per gene fragment and distributional information for all taxa used in this study.

Material was sequenced from at least one or two individuals for most of the localities (Table 2) to establish divergence profiles for the three gene markers. We in-

cluded more individuals for the cytochrome *b* (*cyt b*) region because our previous experience with other squamate groups suggests that this gene would be the most variable of the three, followed by the ND4 and 12S regions in decreasing order of divergence. Voucher specimens are deposited in the field collection of the first and second authors (LJAMM) herpetological collection (now housed in the Centro Regional de Investigaciones La Rioja [CRILAR], Argentina), Fundacion Miguel Lillo (FML), M. L. Bean Museum of Brigham Young University (BYU), Museo de La Plata (MLP.S), Museum of Vertebrate Zoology (MVZ), and San Diego State University (SDSU), and morphological observations were made on some specimens from Museo Argentino de Ciencias Naturales (MACN). Museum numbers of all voucher specimens are listed by locality in Appendix 1 (available at <http://systematicbiology.org>), and museum acronyms follow Leviton et al. (1985).

Laboratory Procedures

Total genomic DNA was extracted from liver/muscle tissues preserved in 96% ethanol following the protocol developed by Fetzner (1999). Three microliters of extraction product were electrophoresed on a 1% agarose gel to estimate the quality and amount of genomic DNA, and sample dilutions were performed where necessary prior to polymerase chain reaction (PCR) amplification. Three mitochondrial gene regions were amplified via PCR in a cocktail containing 2.0 μ l of template DNA (approximate concentration estimated on a 2% agarose gel), 8 μ l dNTPs (1.25 mM), 4 μ l 10 \times *Taq* buffer, 4 μ l each primer (10 μ M), 4 μ l MgCl (25 mM), 24 μ l distilled water, and 0.25 μ l *Taq* DNA polymerase (5 U/ μ l; Promega Corp., Madison, WI). A fragment of 782 bp from the *cyt b* gene was amplified using the light strand primers GLUDGL (5'-TGACTTGAARAACCAAYCGTTG-3'; Palumbi, 1996) and

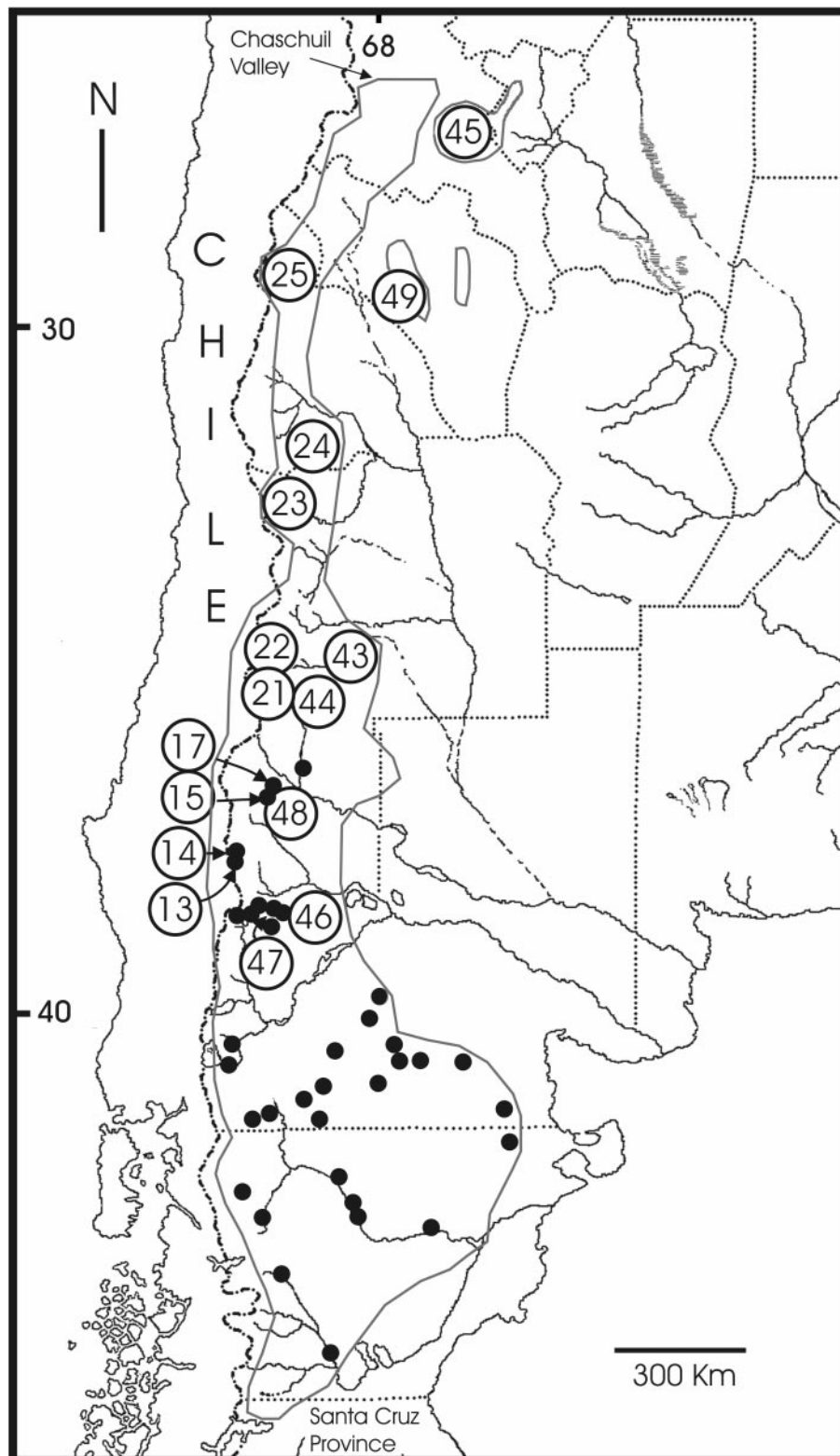


FIGURE 1. Map of western Argentina and Patagonia, showing geographic location of samples collected from the *L. elongatus-kriegi* complex for this study (see Table 2 for explanation of localities). Black dots represent populations/species described in separate maps for each group, and numbers correspond to populations/species not included in the more detailed group maps. The outline represents the approximate distribution of this complex.

TABLE 2. Number of individuals of all ingroup and outgroup taxa, by locality, used in this study; locality numbers (in parentheses) match those in Figure 1 and Appendix 1, which provide museum voucher numbers for all specimens. With the exception of *L. sp. 1* and *L. sp. 3*, which were nonfocal species, lineages identified as *L. sp.* were not obvious prior to this study and represent populations previously included in the nominate species of each group.

Province department	Locality	Genes			Coordinates	
		Cyt <i>b</i>	ND4	12S		
<i>elongatus</i> group						
<i>L. elongatus</i>						
Chubut						
	Futaleufú	17 km S Esquel (1)	5	5	4	42°57'S, 71°11'W
	Futaleufú	18 km N Tecka (2)	1	1	1	43°20'S, 70°52'W
	Río Senguer	23 km W Los Manantiales (3)	3	1	1	45°27'S, 68°43'W
	Tehuelches	22 km S Gobernador Costa (4)	4	2	1	44°12'S, 70°26'W
Rio Negro						
	Ñorquinco	Ojo de Agua (5)	6	5	3	41°32'S, 69°51'W
	Ñorquinco	23 km N El Maiten (6)	1	1	1	41°53'S, 70°59'W
	Bariloche	San Carlos de Bariloche (7)	4	2	1	41°09'S, 71°09'W
	25 de Mayo	57 km S Ingeniero Jacobacci (8)	2	1	1	41°45'S, 69°21'W
	25 de Mayo	17 km S San Antonio del Cuy (9)	8	8	7	40°17'S, 68°57'W
	25 de Mayo	64 km NE Ingeniero Jacobacci (10)	1	1	1	40°53'S, 69°17'W
Neuquén						
	Picunches	Primeros Pinos (11)	1	0	1	38°51'S, 70°33'W
	Picunches	3 km S Portal La Atravesada (12)	1	1	1	38°52'S, 70°33'W
	Aluminé	Pampa de Lonco Luan (16)	4	3	2	38°53'S, 70°58'W
	Zapala	Estacion Zapala (18)	1	1	1	38°53'S, 70°09'W
	Zapala	0.5 km N PN Laguna Blanca (19)	2	1	1	39°03'S, 70°18'W
	Los Lagos	3 km W Co. Las Ardillas (20)	1	1	1	41°45'S, 71°27'W
<i>L. sp. 6</i>						
Neuquén						
	Ñorquín	5 km E Caviahue (13)	4	5	3	37°54'S, 70°59'W
	Ñorquín	W Termas de Copahue (14)	1	1	1	37°49'S, 71°05'W
<i>L. sp. 7</i>						
Neuquén						
	Chos Malal	Los Barros (15)	9	10	8	37°06'S, 70°08'W
	Chos Malal	15 km N Los Barros (17)	4	4	4	37°02'S, 70°02'W
<i>L. sp. 5</i>						
Mendoza						
	Malargüe	Mallines Colgados (21)	8	5	7	35°39'S, 70°12'W
	Malargüe	16 km W Las Leñas (22)	4	4	2	35°06'S, 70°10'W
<i>petrophilus</i> group						
<i>L. petrophilus</i>						
Chubut						
	Paso de Indios	Valle de los Martires (26)	2	2	2	43°49'S, 67°45'W
	Paso de Indios	6 km N Cerro Condor (27)	4	2	1	43°23'S, 69°10'W
	Telsen	Quele Cura (28)	2	1	1	42°13'S, 66°21'W
	Languineo	8 km S Paso del Sapo (29)	2	1	1	42°48'S, 69°33'W
	Gastre	65 km S Paso del Sapo (30)	1	1	1	43°12'S, 69°12'W
Rio Negro						
	25 de Mayo	24 km NW Los Menucos (31)	2	3	2	40°42'S, 68°16'W
	25 de Mayo	7 km N Ingeniero Jacobacci (32)	4	2	1	41°13'S, 69°24'W
	25 de Mayo	40 km SE Maquinchao (33)	4	4	1	41°30'S, 68°33'W
	25 de Mayo	7.5 km W Los Menucos (34)	4	3	1	40°51'S, 68°10'W
	Valcheta	1 km W Chipauquil (35)	4	4	1	40°58'S, 66°39'W
	Valcheta	Arroyo Verde y Ruta Pcial. 5 (36)	4	4	1	41°45'S, 66°30'W
	9 de Julio	18 km NW Comicó (37)	4	3	1	41°00'S, 67°40'W
	Ñorquinco	40 km Ñorquinco (38)	1	1	1	41°45'S, 70°31'W
	El Cuy	El Cuy (39)	1	1	1	39°55'S, 66°24'W
<i>L. capillitas</i>						
Catamarca						
	Andalgalá	Ruta Pcial. 47, entre km 34 y 39 (45)	4	1	3	27°26'S, 68°20'W
<i>L. austromendocinus</i>						
Mendoza						
	San Rafael	9.5 km N Nihuil (43)	3	2	2	34°59'S, 68°37'W
	Malargüe	2 km N Agua Botada (44)	3	3	3	35°45'S, 69°38'W
<i>L. sp. 1</i>						
La Rioja						
	Felipe Varela	Parque Nacional Talampaya (49)	3	3	2	29°44'S, 67°45'W

(Continued on next page)

TABLE 2. Continued

Province department	Locality	Genes			Coordinates
		Cyt <i>b</i>	ND4	12S	
<i>L. sp. 2</i>					
Neuquén					
Zapala	2 km S La Amarga (46)	3	3	2	39°06'S, 69°34'W
Picunches	0.5 km W Co. Cansino (47)	2	1	1	38°52'S, 70°33'W
<i>L. sp. 4</i>					
Mendoza					
Las Heras	Vallecitos (23)	5	5	2	32°59'S, 69°20'W
Las Heras	17 km W Santa Clara (24)	5	4	1	32°09'S, 69°04'W
San Juan					
Iglesia	Quebrada Honda (25)	4	4	1	29°22'S, 29°28'W
<i>kriegi</i> group					
<i>L. kriegi</i>					
Rio Negro					
Ñorquinco	Ojo de Agua (5)	4	4	4	41°32'S, 69°51'W
Ñorquinco	23 km N El Maiten (6)	1	1	1	41°53'S, 70°59'W
25 de Mayo	40 km SE Maquinchao (33)	4	3	1	41°30'S, 68°33'W
25 de Mayo	57 km S Ingeniero Jacobacci (8)	4	3	1	41°45'S, 69°21'W
Neuquén					
Picunches	Primeros Pinos (11)	4	4	3	38°51'S, 70°33'W
Aluminé	Pampa de Lonco Luan (16)	5	5	5	38°53'S, 70°58'W
Aluminé	5 km W Portal La Atravesada (40)	2	1	1	38°52'S, 70°33'W
Los Lagos	1.5 km W Co. Las Ardillas (41)	2	1	1	41°45'S, 71°27'W
<i>L. sp. C</i>					
Neuquén					
Chos Malal	15 km N Los Barros (17)	1	1	1	37°02'S, 70°02'W
Chos Malal	Los Barros (15)	2	2	2	37°06'S, 70°08'W
<i>L. buergeri</i>					
Mendoza					
Malargüe	16 km W Las Leñas (22)	3	2	1	35°06'S, 70°10'W
Malargüe	Mallines Colgados (21)	3	2	2	35°39'S, 70°12'W
<i>L. sp. 8</i>					
Neuquén					
Ñorquín	W Termas de Copahue (14)	2	1	1	37°49'S, 71°05'W
<i>L. sp. A</i>					
Neuquén					
Ñorquín	5 km E Caviahue (13)	4	4	3	37°54'S, 70°59'W
<i>L. sp. B</i>					
Mendoza					
Malargüe	5 km N Ranquil Norte (42)	5	5	4	36°37'S, 69°50'W
<i>L. sp. 3</i>					
Neuquén					
Chos Malal	Volcan Tromen (48)	2	2	1	37°06'S, 70°08'W
Outgroups					
<i>Liolaemus bibronii</i>					
Rio Negro, 25 de Mayo	14 km W Aguada de Guerra				41°09'S, 68°30'W
<i>L. gracilis</i>					
Neuquén, Zapala	6 km NW La Amarga				39°04'S, 69°37'W
<i>L. neuquensis</i>					
Neuquén, Picunches	Primeros Pinos				38°51'S, 70°33'W
<i>L. pictus</i>					
Rio Negro, Bariloche	San Carlos de Bariloche				41°09'S, 71°09'W
<i>L. pseudoanomalous</i>					
La Rioja	3 km N Pagancillo				29°30'S, 68°07'W
<i>L. vallecurensis</i>					
San Juan, Iglesia	Llanos de la Lagunita				29°24'S, 69°25'W
<i>L. kingii</i>					
Chubut, Rio Senguer	2 km S Rio Mayo				45°42'S, 70°15'W
<i>L. lineomaculatus</i>					
Rio Negro, Bariloche	PN Nahuel Huapi, faldeo NW Piedra del Condor, Cerro Catedral				41°12'S, 71°17'W
<i>Phymaturus indistinctus</i>					
Chubut, Rio Senguer	Sierra de San Bernardo				45°27'S, 69°42'W

cyt-*b* 1 (5'-CCATCCAACATCTCAGCATGATGAAA-3'; Kocher et al., 1989) and the heavy strand primer 3 (5'-GGCAAATAGGAARTATCATTTC-3'; Palumbi, 1996). For some individuals, the 2 primer (5'-CCCTCAGAATGATATTTGTCCTCA-3'; Palumbi, 1996) was used as an internal sequencing primer. Another 789 bp from the ND4 gene was amplified with the ND4-F (5'-CACCTATGACTACCAAAAGCTCATGTAGAAGC-3') and ND4-R (5'-CATTACTTTTACTTGGATTTGCACCA-3') primers (Arévalo et al., 1994), and 824 bp of 12S gene was amplified with the 12e (5'-GTRCGCTTACC WTGTTACGACT-3') and tPhe (5'-AAAGCACRGCCTGAAGATGC-3') primers (Wiens et al., 1999). Double-stranded PCR amplified products were checked by electrophoresis on a 2% agarose gel, purified using a GeneClean III kit (BIO101, Vista, CA), and directly sequenced using the Perkin Elmer ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, Foster City, CA). Excess dye terminator was removed with CentriSep spin columns (Princeton Separations, Adelphia, NJ), and sequences were fractionated by polyacrylamide gel electrophoresis on an ABI PRISM 377 automated DNA sequencer (PE Applied Biosystems) at the DNA Sequencing Center (DNASC) at BYU.

Sequence Alignments

Most sequences were edited and aligned using the program Sequencher 3.1.1 (Gene Codes Corp., Ann Arbor, MI), and the protein coding regions of cyt *b* and ND4 were translated into amino acids for confirmation of alignment. Divergence was low for the 12S fragment, and indels were few in number (15) and in all cases only a single base in length. Alignment of this region was performed with CLUSTAL X (Thompson et al., 1997), using the default settings for gap and mismatch penalties, with subsequent manual adjustments. Missing data were coded as "?". Sequences are available on the *Systematic Biology* website.

Establishing Divergence Profiles

We began by sequencing two or four lizards (mean = 3.2) from each locality for the cyt *b* region for all focal species, one or two lizards from some localities for the 12S region (because previous experience suggested that it would be the least variable), and usually an intermediate number for ND4 (the same number of lizards was sequenced as for cyt *b* in 50% of the localities). These numbers were increased if lizards were morphologically more variable than typical for most localities and/or if sequence divergence at any gene region was detected in lizards from geographically proximate localities. The most variable region was then sequenced for almost all individuals and used to test for exclusivity of haplotypes at each locality. Once exclusivity was established, a single individual usually was sequenced per locality for the more slowly evolving (12S) region.

Phylogenetic Analyses

Single gene regions.—To apply the WP method to a large data set with multiple genes evolving at different rates, we first conducted a Bayesian analysis for the full cyt *b* data set (521 bp for all individuals without missing data, the most quickly evolving region sequenced); this analysis included 97 different haplotypes representing 199 lizards from 58 localities. Bayesian analyses were performed using MrBayes 2.0 (Huelsenbeck and Ronquist, 2001) based on the general time reversible model with invariable and variable sites with a discrete gamma distribution (GTR + I + Γ) model of evolution (Yang, 1994; Gu et al., 1995), which was selected as the best fit model of nucleotide substitution using ModelTest 3.04 (Posada and Crandall, 1998). The specific a priori parameter values were uniform and were estimated as part of the analysis. To explore the parameter space more thoroughly we ran metropolis-coupled MCMC simulations with four incrementally heated chains, using the default values. From a random starting tree, we ran 1.0×10^6 generations, and we sampled the Markov chains at intervals of 100 generations to obtain 10,000 sample points. We determined when stationarity was reached (to discard the burn-in samples) by plotting the log-likelihood (lnL) scores of sample points against generation time; when the values reached a stable equilibrium, before 50,000 generations, stationarity was assumed. The equilibrium samples (the 9,500 trees retained after burn-in) were used to generate a 50% majority rule consensus tree. The percentage of samples that recover any particular clade on this tree represents that clade's posterior probability; these are the *P* values, and $P \geq 95\%$ was considered evidence of significant support for a clade (Huelsenbeck and Ronquist, 2001).

Two additional separate analyses (including one individual per gene from each locality marked with an asterisk in Appendix 1) were conducted on the ND4 and 12S regions, to examine phylogenetic congruence between these and the cyt *b* haplotype tree (Leaché and Reeder, 2002). All analyses were performed only within a Bayesian framework to detect potential areas of strongly supported incongruence (Wiens, 1998), defined here as nodes with posterior probability (PP) values $\geq 95\%$. We again used the GTR + I + Γ model of evolution, and from a random starting tree we ran 10^6 generations in each analysis and sampled the Markov chains at intervals of 100 generations to obtain 10,000 sample points. Stationarity was reached before 50,000 generations in all cases, and after discarding these first 500 trees (burn-in), the 50% majority rule trees were obtained from the remaining 9,500 data points.

Combined gene regions.—The combined data set of 2,395 bp was used for traditional phylogenetic analyses (MP, ML, and Bayesian) on the reduced data matrix including one individual per gene (using a larger 782-bp cyt *b* fragment, with missing data for some individuals) from each locality (these individuals are marked with an asterisk in Appendix 1).

For MP analysis, all characters were equally weighted, and we conducted a heuristic search with 100 replicates of random taxon addition with tree bisection-reconnection (TBR) branch swapping and gaps coded as missing data, using PAUP* 4.0b4b (Swofford, 2001). Nodal support was assessed by nonparametric bootstrap (BS) analysis (Felsenstein, 1985), with strong support being inferred at $BS \geq 70$ (Hillis and Bull, 1993; with caveats).

For ML analysis, the combined data set was analyzed under the GTR + I + Γ model. A heuristic search with five random addition replicates using the TBR branch-swapping algorithm was performed to obtain the ML tree. Because of computational limitations imposed by ML estimation, we used PAUP* to perform five separate searches with 20 random addition replicates each in a bootstrap analysis and then combined the total 100 pseudoreplicates to obtain the bootstrap proportions. All ML analyses were performed on an IBM Sp2 supercomputer in the BYU supercomputing facility.

Using MrBayes 2.0 (Huelsenbeck and Ronquist, 2001) with the GTR + I + Γ model of evolution and specific parameter values estimated as part of the analysis, the run was performed for 2×10^6 generations with four incrementally heated chains and sampled at intervals of 100 generations to include 20,000 data points. Stationarity was reached before 40,000 generations, and after discarding these first 400 trees (burn-in) the 50% majority rule tree was obtained from the remaining 19,600 data points. To avoid a local entrapment, we ran two independent analyses and compared these for convergence to similar lnL mean values (Huelsenbeck and Bollback, 2001; Leaché and Reeder, 2002). We also compared the posterior probabilities (PP) for individual clades obtained from the separate analyses for congruence to ensure convergence of the two analyses.

Network and Nested-Clade Analyses

The complete set (no missing data) of short *cyt b* sequences (555 bp for the *elongatus* and *petrophilus* groups and 521 bp for the *kriegi* group) from the focal species was used for NCA. The program TCS 1.06 (Clement et al., 2000; available from http://bioag.byu.edu/zoology/crandall_lab/programs.htm) was used to construct haplotype networks (based on the algorithm of Templeton et al., 1992) taking into account haplotype frequencies, which are correlated with haplotype age (Castelloe and Templeton, 1994; Donnelly and Tavaré, 1995), and nesting categories were assigned following Templeton and Sing (1993) and Templeton et al. (1995). The networks were then used for NCA, which was implemented with GeoDis 2.0 (Posada et al., 2000) available from the same website. All the statistical analyses were performed using 10,000 permutations. Ambiguous connections (loops) in the networks were resolved using approaches from coalescent theory (Crandall et al., 1994), and statistically significant associations between haplotypes and geography were interpreted following the

inference key of Templeton (2001, also available from the same website).

Defining Species Boundaries

A priori criteria (Sites and Crandall, 1997) were used to delimit species in a hypothesis-testing framework by implementing the DNA tree-based method of Wiens and Penkrot (2002) and the NCA of Templeton (2001). The first method uses a dichotomous key approach to make species-level decisions (Wiens and Penkrot, 2002: fig. 1) and takes advantage of the haploid maternal inheritance of the mtDNA locus in the context of coalescent theory. The mtDNA effective population size (N_E) is smaller than that for a Mendelian locus ($0.25N_E$ on average) under many biologically plausible scenarios, which results in coalescence of mtDNA haplotypes in a species four times more rapidly (Birky et al., 1983, 1989; but for reviews of mating systems that deviate from the $0.25 N_E$ generality, see Hoelzer, 1997; Hoelzer et al., 1998). Thus, newly formed species should be distinct in their mtDNA genealogy long before they become distinct at nuclear loci (see also Palumbi et al., 2001). The first step is to determine whether the focal species is exclusive in the haplotype tree, following Baum and Shaw (1995:296): "A group of organisms is exclusive if their loci coalesce more recently within the group than between any member of the group and any organism outside the group." When haplotypes are recovered as strongly supported basal clades (i.e., the oldest splits within the focal species) and are exclusive by locality (i.e., the topology is concordant with geography), the terminal is considered distinct at the species level. Alternatively, a tree in which haplotypes from different localities interdigitate is taken as evidence of gene flow between localities and conspecificity of the populations. When haplotypes of the focal species are not exclusive with respect to other species that are distinct and exclusive, then the focal species is taken to include multiple species if its basal clades are exclusive and concordant with geography (i.e., if they show no evidence of gene flow; Wiens and Penkrot, 2002: fig. 1c). Other alternatives in the key lead to inferences of conspecificity among haplotypes of the focal species (including nonexclusive species) or possibly multiple species within the focal and at least one nonfocal species (Wiens and Penkrot, 2002: figs. 1b, 1d, or 1f, and 1e, respectively). In this study, we accepted results of the dichotomous key as evidence for provisional species status with the focal species of this study.

The WP method was used to delimit species at the deepest levels of divergence in the tree, emphasized by Wiens and Penkrot (2002:71): "by 'basal' we mean the oldest split or splits within the species," although it can be used at increasingly nested levels of divergence. In this study, we defined basal lineages as those that are separated beyond the TCS network connections (of 95% confidence) among *cyt b* haplotypes. This designation is admittedly arbitrary, but at these shallow levels of divergence, the NCA has the statistical power to test further splitting within well-sampled clades following the

method described by Templeton (2001). The method is implemented by sequentially testing two hypotheses: H_1 : organisms sampled are derived from a single evolutionary lineage; and H_2 : the distinct populations of lineages identified by rejection of H_1 are genetically, exchangeable or ecologically interchangeable. Genetic exchangeability refers to any mechanism(s) that promotes gene flow (fertilization or mate-recognition systems), and ecological interchangeability can include evidence from life-history traits, habitat requirements, etc.; both can be tested by the NCA (Templeton, 2001:787) but are beyond the objectives of this study.

Here, we calculated nested-clade distance measures to test for significant associations of nested clades with geographic locations, and the inference key was then followed to derive a plausible biological cause for each clade showing any significant pattern of geographic association shown by a clade. For postulating species boundaries, the relevant result is rejection of H_1 with a signal of allopatric fragmentation at some clade level. Data are not yet available to test H_2 in this complex, but we provide qualitative corroboration for some clades when independent data are available. In presenting results, lineages defined as species by WP criteria are identified by numbers (except *L. sp. 1* and *L. sp. 3*, previously identified as new species), and clades for which historical fragmentation was inferred by NCA or for which no unambiguous inference was possible are identified by letters (*L. sp. A*, *L. sp. B*, etc.; Table 2).

RESULTS

Phylogenetic Analysis

Sequences were deposited in GenBank under accession number AY173521–AY173929 and provided the matrices for all phylogenetic analyses. The majority rule consensus Bayesian tree obtained from all *cyt b* haplotypes (not shown) recovered almost all haplotypes as exclusive by locality, with high posterior probabilities of support ($\alpha \geq 0.95$), except for some of the shallowest nodes within the five clades analyzed with NCA. Deeper phylogenetic relationships were poorly resolved with the *cyt b* data set alone. Because the ND4 and 12S sequences showed almost no intralocality variation and all three genes are linked, this result suggested that deeper phylogenetic structure can be retrieved by including one individual per population in all further tree-reconstruction analyses. The two independent runs for each gene generated the same 50% majority rule consensus trees for all gene fragments. No nodes of strongly supported conflict ($\alpha \geq 0.95$ PP) were identified among the three partitions, so all were combined for application of the WP test.

Analysis of the combined mitochondrial genes with MP produced 4,368 most-parsimonious trees, the strict consensus of which is represented in Figure 2 (length = 3,022; consistency index = 0.449, retention index = 0.774). The MP analysis recovered a monophyletic *L. elongatus-kriegi* complex (bold branches, Figs. 2, 3) with a BS value of 98, with all a priori defined outgroup species

of *Liolaemus* external. The terminal from locality 48 (*L. sp. 3*), representing one of the nonfocal ingroup species, is the sister group to the rest of the *L. elongatus-kriegi* clade, which is supported by a BS of 92. Within this large clade, there is strong support (BS = 100) for a “*kriegi* group” + “*elongatus* group” clade, strong support (BS > 97) for each of these groups, and moderate support (BS = 75) for a “*petrophilus* group” clade. One other salient point from this analysis is that it defines three of the original focal species (*L. elongatus*, *L. kriegi*, and *L. petrophilus*) more narrowly than does the morphologically based taxonomy, with several strongly supported terminals recovered far removed from the newly defined focal species (see especially “*L. elongatus*” terminals; Figs. 2, 3, solid circles).

The topology presented in Figure 3 was estimated from a ML analysis of the combined data and is one of two very similar best trees recovered with a lnL value of -17235.01859. This topology was also representative of the majority rule consensus of 19,600 trees (mean lnL = -17119.80732) obtained with the Bayesian approach; the same result was obtained in both independent runs (not illustrated). The topologies obtained with the two approaches are almost identical and are very similar to the MP topology. In all, the most basal clade includes four nonfocal species with the topology ((*L. kingi* + *L. lineomaculatus*) (*L. pseudoanomalus* + *L. vallecurensis*)); four other nonfocal species considered part of the “*chiliensis* group” (*L. bibroni*, *L. gracilis*, *L. neuquensis*, and *L. pictus*) were recovered as basal groups of the species complex under study. All methods recovered the nonfocal ingroup terminal from locality 48 (*L. sp. 3*) as the sister taxon to the rest of the *L. elongatus-kriegi* complex (Figs. 2, 3), but ML support for monophyly of this complex is weak (BP = 64) compared with that of the Bayesian analysis (PP = 1.0). Both Bayesian and ML analyses recovered the three major clades recovered by MP analysis; the (“*kriegi* group” + “*elongatus* group”) clade is strongly supported (BP = 88, PP = 1.0; Fig. 3), and the “*petrophilus* group” is strongly supported by Bayesian posterior probability (1.0) but only weakly supported under ML assumptions (BS < 50%; Fig. 3). All tree-estimation procedures recovered the same three clades that we here recognize as the *elongatus*, *kriegi*, and *petrophilus* groups (albeit with various levels of support; Figs. 2, 3); the distribution of these groups is mapped in Figure 4.

Within the *petrophilus* group, relationships among terminals differ in the three topologies; the Bayesian and MP trees recovered the (*L. capillitas* + *L. sp. 1*) clade as the sister clade of *L. petrophilus* (PP = 0.55), whereas the ML analysis recovered the (*L. capillitas* + *L. sp. 1*) clade as most basal and the clade (*L. sp. 2* (*L. sp. 4* + *L. austromendocinus*)) as the sister clade to *L. petrophilus*. We leave these relationships unresolved because of weak support for alternative topologies (Fig. 3). Within the clade (*L. sp. 2* (*L. sp. 4* + *L. austromendocinus*)), *L. sp. 2* is the sister taxon to (*L. sp. 4* + *L. austromendocinus*) with strong support in all analyses, and within *L. petrophilus*, two clades were recovered with strong support in all analyses (clade 3.1 and 3.2; Figs. 2, 3).

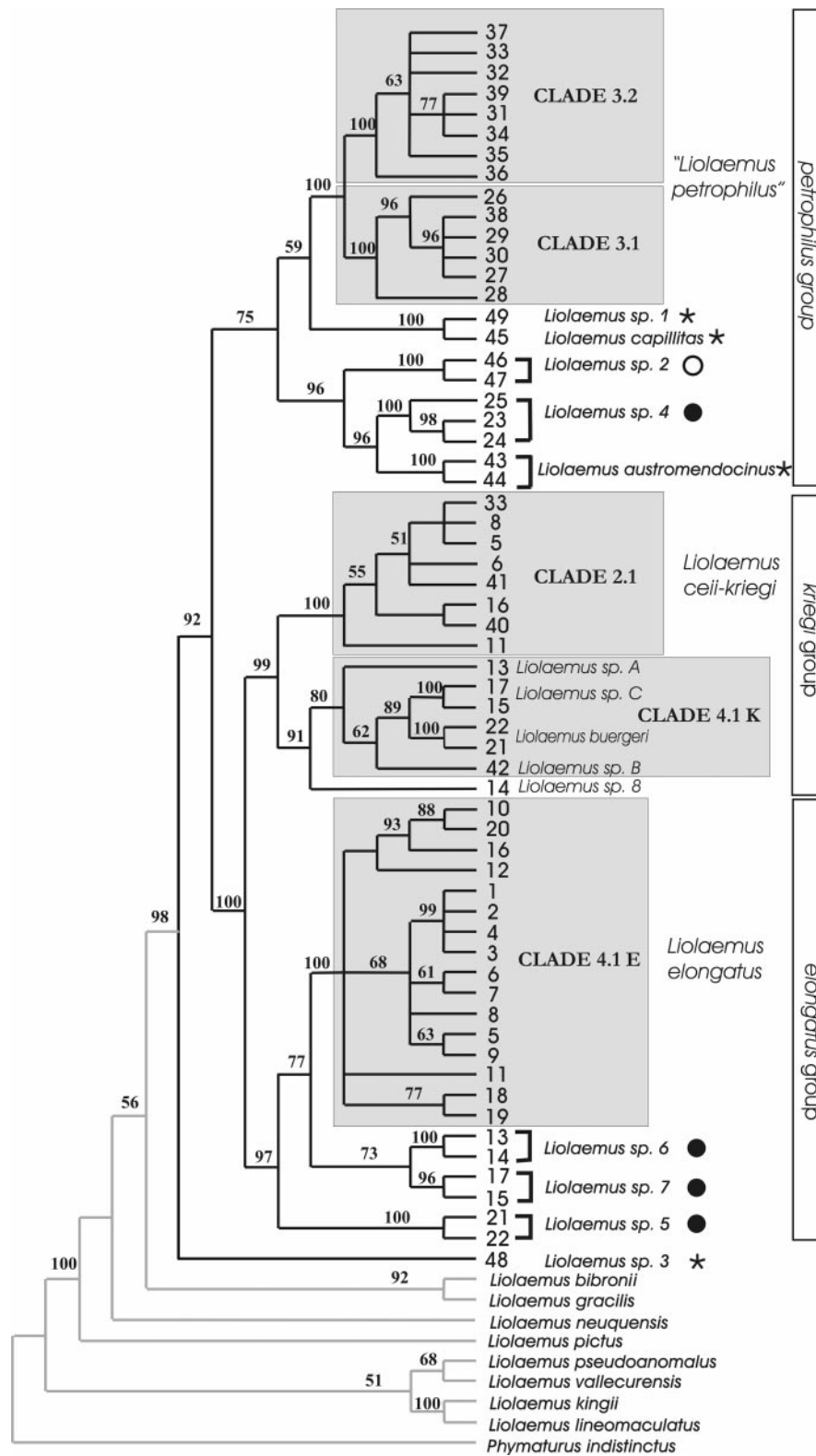


FIGURE 2. Molecular phylogeny (one of two trees obtained) for the *L. elongatus-kriegi* complex (bold branches) based on MP analysis for the combined *cyt b*, ND4, and 12S gene regions. Numbers at nodes are bootstrap values >50% (based on 100 replicates), and numbers at terminals correspond to locality numbers (Table 2). Shaded clades are those used in the NCA; terminals previously included in the species *L. elongatus* are identified by solid circles, the open circle indicates a terminal previously assigned to *L. petrophilus*, and nonfocal ingroup terminals (black lines on the tree) are identified by asterisks. Whenever two terminals have the same locality number (e.g., 14), two species were collected in sympatry at one locality (*L. sp. 6* and *L. sp. 8*). Shaded branches in the tree are outgroups.

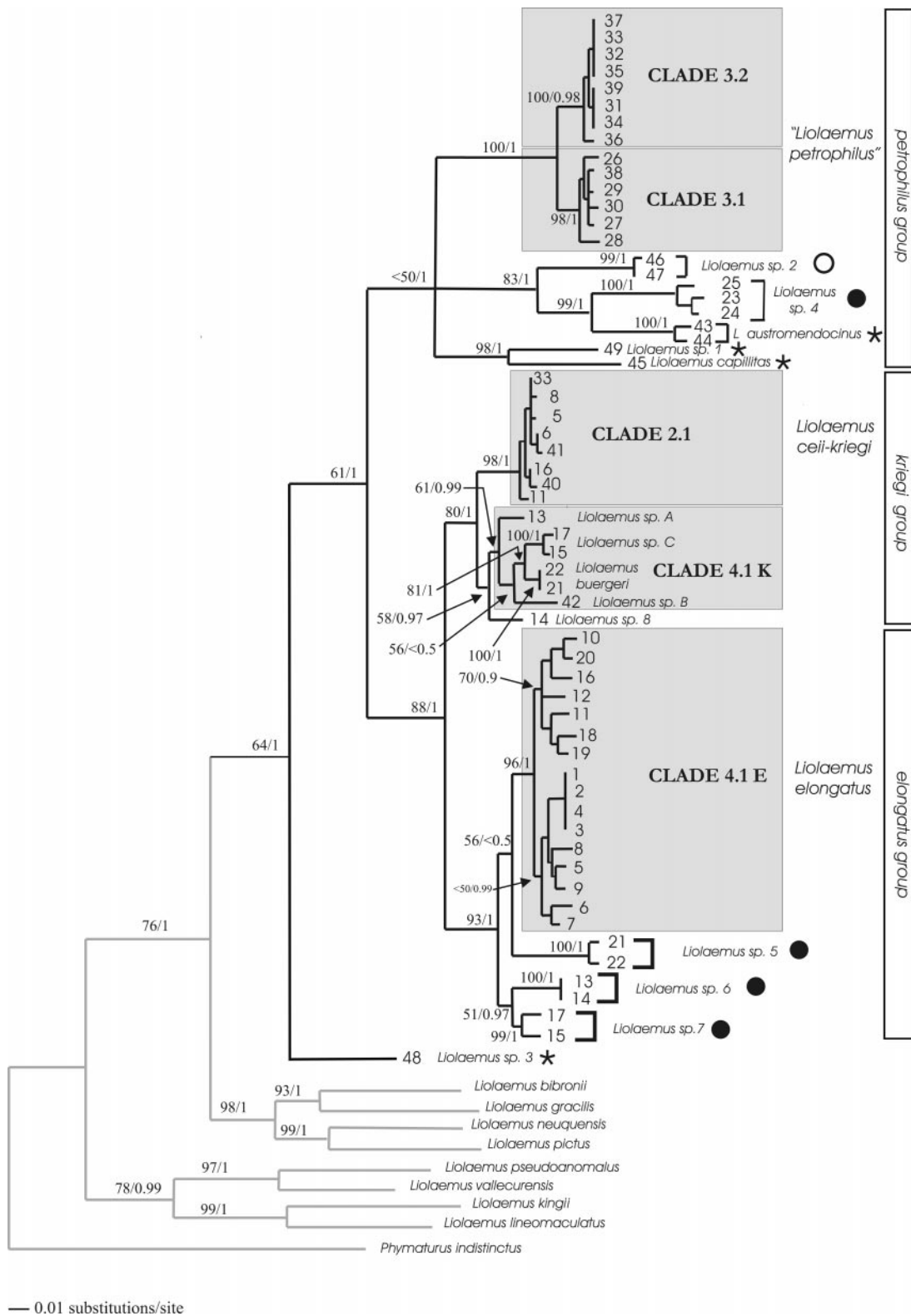


FIGURE 3. Molecular phylogeny for the *L. elongatus-kriegi* complex (bold branches) based on ML analysis (GTR + I + Γ model) for the combined *cyt b*, ND4, and 12S gene regions. Number at nodes indicate bootstrap values (>50%)/posterior probability values based on the same ML model. Numbers at terminals correspond to locality numbers (Table 2), and shaded clades, solid and open circles, asterisks, and black and shaded branches convey the same information as in Figure 2.

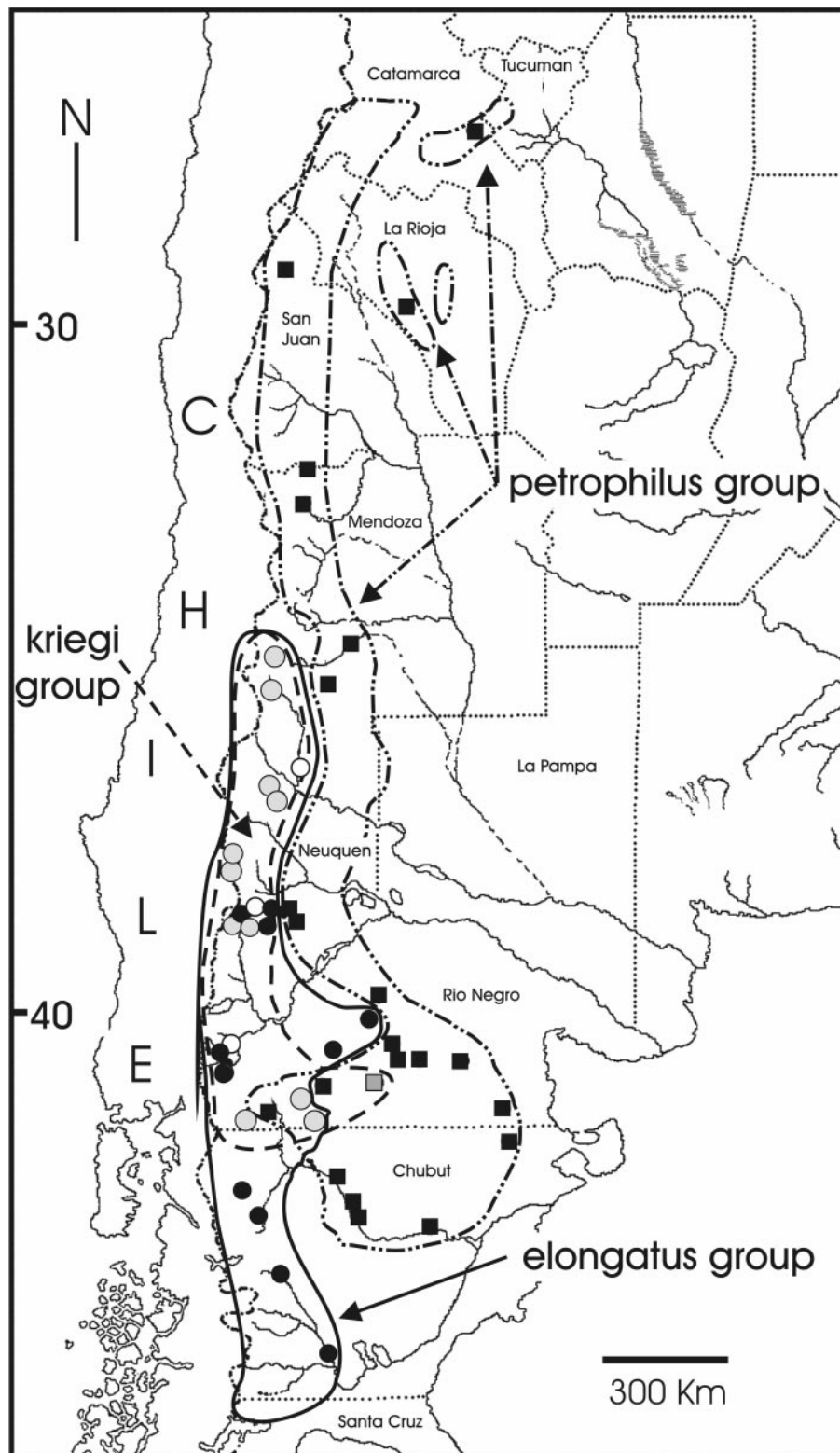


FIGURE 4. Geographic distributions of the *elongatus*, *petrophilus*, and *kriegi* groups; shaded circles represent localities where individuals from the *elongatus* and *kriegi* groups are in sympatry, and the shaded square represents sympatry between *kriegi* and *petrophilus*.

The *kriegi* group includes a strongly supported clade 2.1, and clade 4.1K was recovered with moderate to strong support in all but the ML analysis. The relationships within these two clades are almost the same in all approaches (Figs. 2, 3).

The *elongatus* group includes three clades, one of which (4.1E) includes the majority of the localities and was strongly supported in all analyses. The MP analysis recovered clade 4.1E as the sister group of the (*L. sp. 6* + *L. sp. 7*) clade with moderate support and the third clade (*L. sp. 5*) as the most basal lineage of the *elongatus* group. In the ML and Bayesian approaches, the *L. sp. 5* clade was weakly recovered as the sister group of clade 4.1E, and the (*L. sp. 6* + *L. sp. 7*) clade is basal (Fig. 3).

Nested-Clade Analysis

The tree-reconstruction methods used here cannot definitively resolve relationships within the five clades identified by light shading in Figures 2 and 3, suggesting either inadequate population sampling and/or insufficient character divergence to track splitting events. Application of the Templeton et al. (1992) algorithm to the 37 different *cyt b* haplotypes in the *elongatus* group

(Appendix 2; available at <http://systematicbiology.org>.) showed that haplotypes differing by up to 10 substitutions have at least a 0.95 probability of being parsimoniously connected; this criterion subdivided the data set into four networks (Fig. 5A). The most inclusive network represents most of the exemplars in the *elongatus* group, but haplotypes from localities 13 and 14 formed a second network, those from localities 15 and 17 formed a third, and those from localities 21 and 22 formed a fourth network. For clade 4.1E (Figs. 2, 3, 5A), the NCA revealed significant associations of haplotypes with geography in several nested clades (Fig. 5B) and indicates that the geographic association of the haplotypes at the entire clade level (clade 4.1E; Figs. 5B, 6) is the result of a contiguous range expansion along the Patagonian Steppe. However, for clade 2.1 a long distance colonization event was inferred between localities 5 and 9 (Fig. 6).

Across the distributional range of the *kriegi* group, 17 different *cyt b* haplotypes were found (Appendix 2), and parsimony analysis showed that sequences differing by nine or fewer steps have at least a 0.95 probability of being parsimoniously connected. Two networks were obtained under this criterion (Fig. 7A), one of which

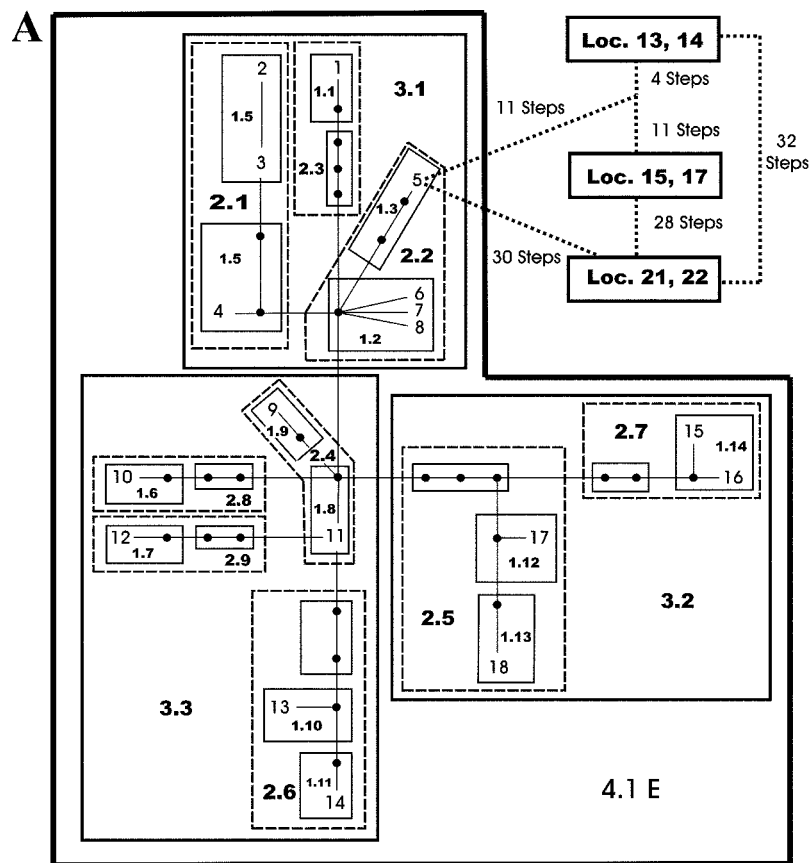


FIGURE 5. (A) Unrooted *cyt b* haplotype networks for the *elongatus* group populations and their associated nested design. The three most inclusive statistical parsimony networks (nested clades 3.1, 3.2, 3.3) are the estimated 95% plausible sets of cladograms obtained using the algorithm of Templeton et al. (1995) as implemented in TCS (Clement et al., 2000). Thin lines represent parsimonious connections between haplotypes with a probability >95%, each line segment corresponds to one mutational step, and intermediate missing haplotypes are represented by solid circles. Haplotypes from localities 13 and 14, 15 and 17, and 21 and 22 differed by >10 substitutions and were therefore beyond the 95% probability of a parsimonious connection. (Continued)

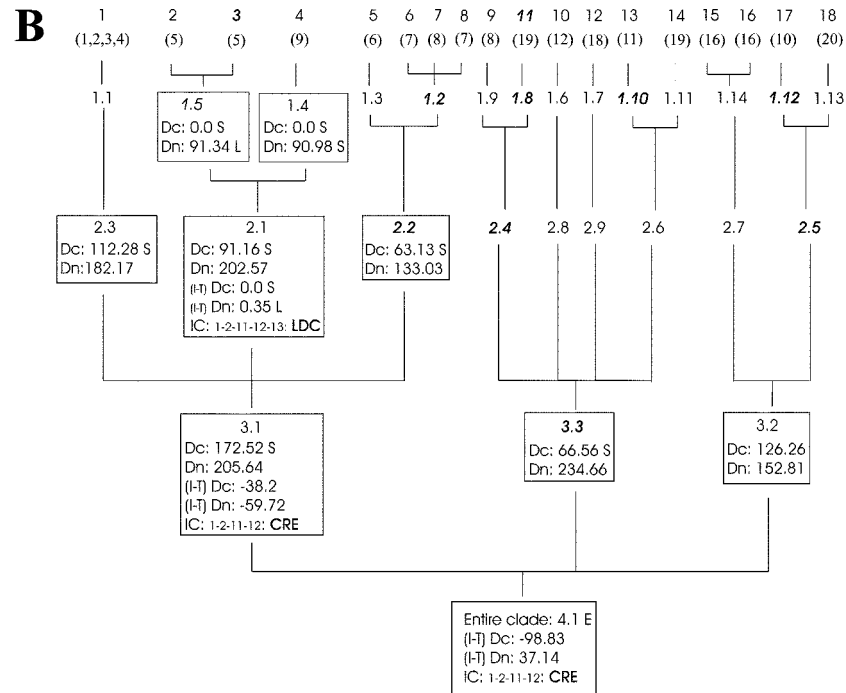


FIGURE 5. (Continued) (B) Summary of NCA results for haplotypes of the *elongatus* group. Nesting levels extend from top clades with individual haplotypes (0-step clades) to bottom more inclusive clades, counting only clades with observed haplotypes. The numbers in parentheses just below the 0-step clades (haplotypes) are the localities (Table 2) from which each haplotype is known. For clades that show a significant probability ($P < 0.05$) of association between geography and haplotype distribution, the clade distance (Dc) and the nested-clade distance (Dn) are given; distances significantly smaller or larger than expected are indicated with an S or L. Results contrasting interior versus tip clades are indicated by I-T (numbers of interior clades are italicized and bold). The inference chain (IC) from the key provided by Templeton (2001) was used to interpret the statistical results: CRE = continuous range expansion; LDC = long distance colonization.

(clade 2.1) includes all the exemplars from the *L. ceii-kriegi* clade in the phylogenies (Figs. 2, 3), and the inference for this clade is a contiguous range expansion (Figs. 7B, 8) along approximately the same distribution as that inferred for clade 4.1E (Figs. 6, 8). The second network (clade 4.1K) includes haplotypes from all exemplars of the second clade in the *kriegi* group in the phylogenies (Figs. 2, 3) except for the basal haplotype (locality 14), which is not included in either network. The inference for haplotypes from clade 3.1 (localities 15, 17, 21, and 22; Fig. 8) cannot distinguish between historical fragmentation (F) and restricted gene flow with isolation by distance (ID; Fig. 7B) between the southern Mendoza and northwestern Neuquén populations. These populations are separated by the Barrancas River valley, which may constitute a barrier (Fig. 8). For the entire clade (4.1K) a range expansion and long-distance colonization events were inferred (Fig. 7B). The distribution of this clade extends across a topographically complex landscape that includes some of the highest peaks and deepest valleys in the Andes (Figs. 4, 8) and likely has a complex evolutionary history.

For *L. petrophilus*, 15 different haplotypes were resolved (Appendix 2), and these are parsimoniously connected by ≤ 10 steps. This analysis produced two networks (Fig. 9A) that correspond to the most inclusive clades (3.1 and 3.2) recovered for this species in the

phylogenetic analyses (Figs. 2, 3). Clade 3.2 included haplotypes from localities 31–37 and 39 and reflects a pattern consistent with a contiguous range expansion (Fig. 10). For the lower level clade 2.4, there was a significant association of haplotypes and geography (Fig. 10), but resolution was insufficient to differentiate among historical fragmentation, range expansion, and restricted gene flow with isolation by distance (Fig. 9B). The inference for the second network (clade 3.1) is a contiguous range expansion, and this clade and clade 3.2 define northern and southern geographical subdivisions in “*L. petrophilus*” in the middle of the Patagonian Steppe, without apparent physical barriers separating them (Fig. 10).

Species Boundaries Inferred from the Combined Approach

The elongatus group.—This group includes several strongly supported clades nested within it (“*L. spp.* 5, 6, 7” and clade 4.1E; Figs. 2, 3), and only one of these is formally described, *Liolaemus elongatus*. Our analysis provides evidence of exclusive clades that are concordant with geography (WP inference 1a) and suggests a revised taxonomy.

The southernmost samples can be assigned to the nominal *Liolaemus elongatus* Koslowsky, 1896 and include all populations south of the Agrío and Neuquén rivers (clade 4.1E; Figs. 2, 3) and extending along the

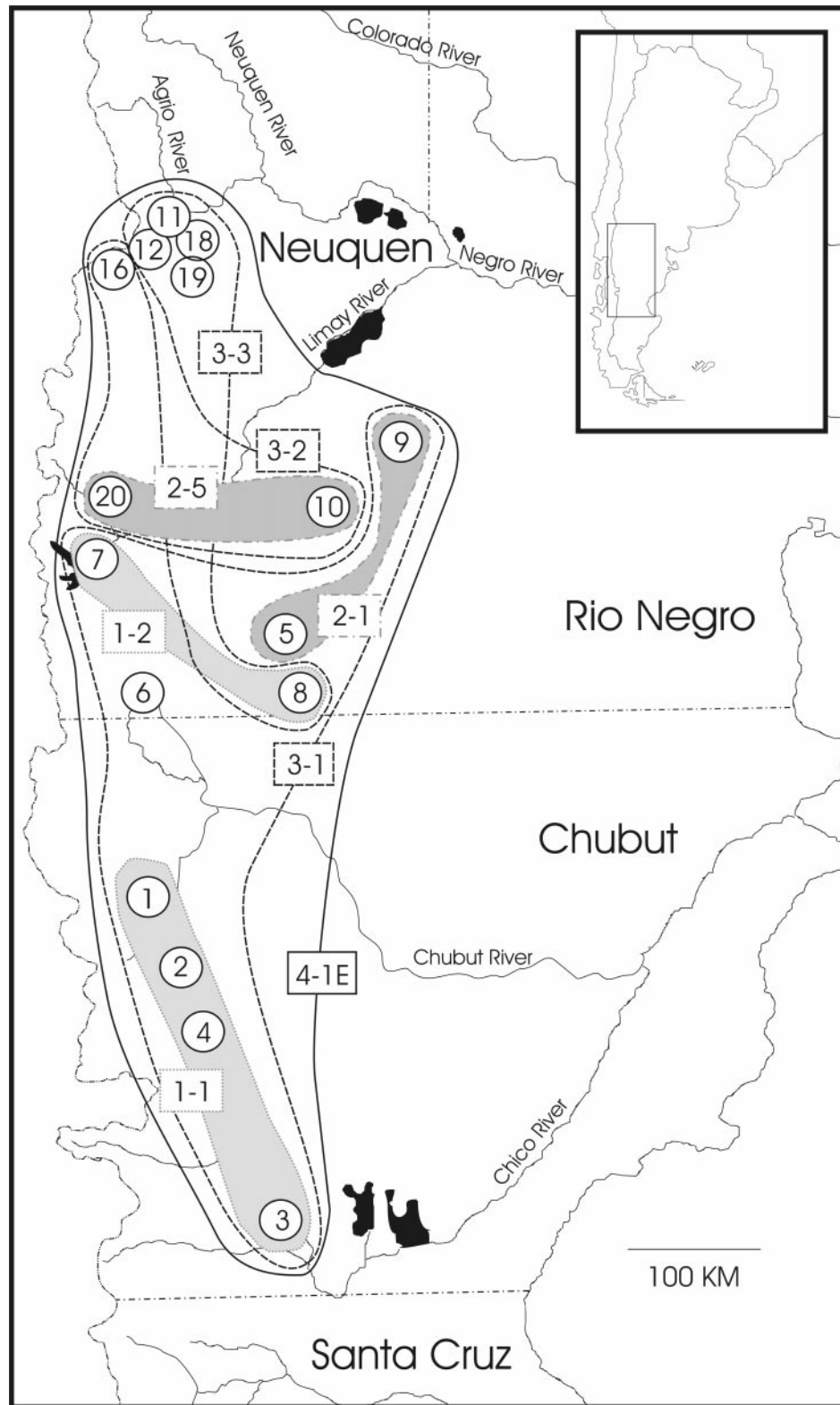


FIGURE 6. Map of the distribution of populations of the *elongatus* group, with the associated nesting design (Fig. 5A) relating the haplotypes from these localities. Numbers correspond to localities in Table 2.

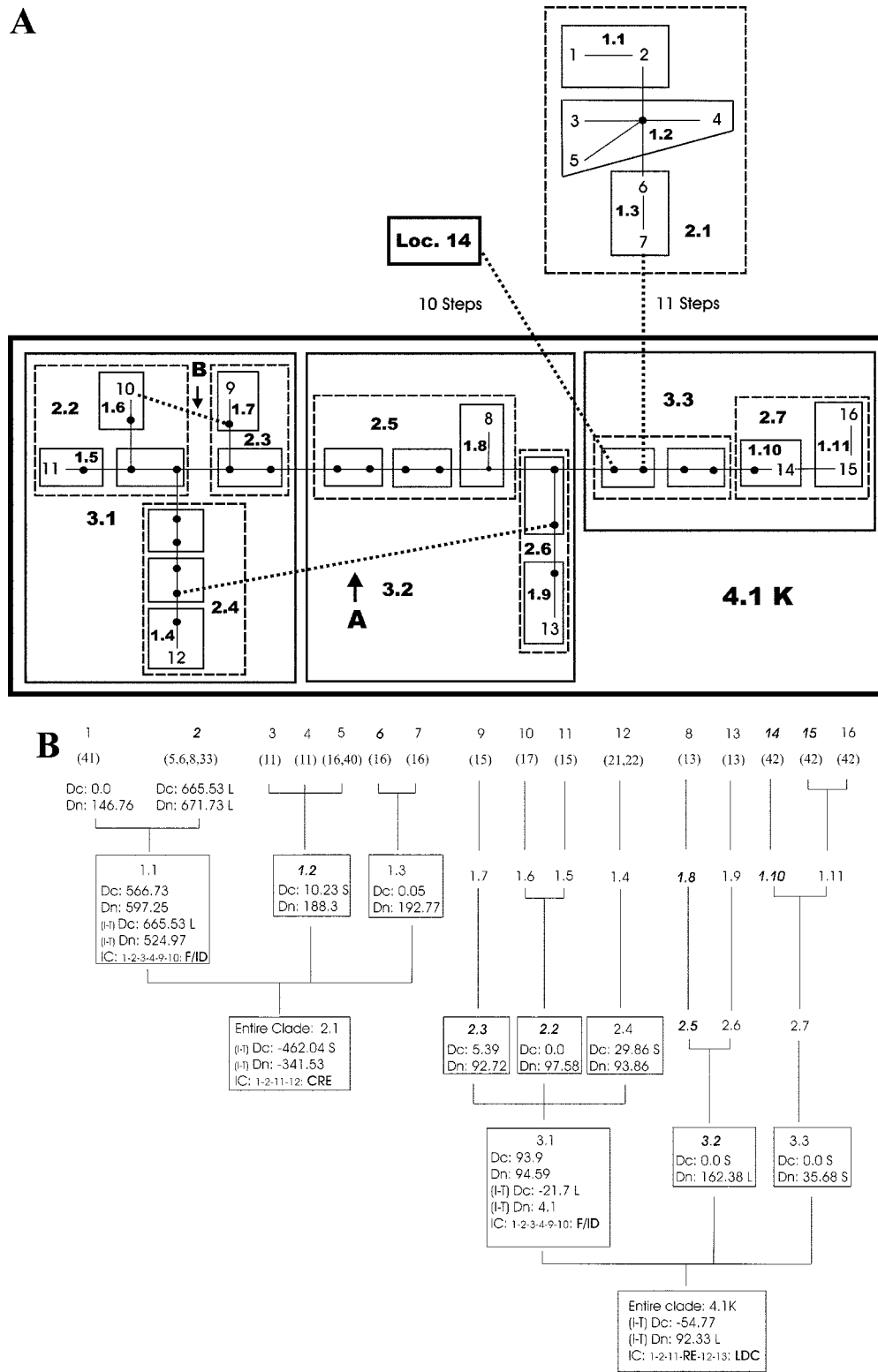


FIGURE 7. (A) Unrooted *cyt b* haplotype networks for the *kriegi* group populations and their associated nested design. The two most inclusive parsimony networks (nested clades 2.1 and 4.1K) are the estimated 95% plausible sets of cladograms, and all symbols convey the same information as in Figure 5A. Two ambiguous connections (loops) are marked A and B and were resolved according to coalescent theory (Crandall et al., 1994). Haplotypes from locality 14 and in clade 2.1 differed by more than nine substitutions and were therefore beyond the 95% probability of a parsimonious connection. (B) Summary of NCA results for haplotypes of the *kriegi* group; organization of this summary is identical to that for Figure 5B, but an additional inference is made here: F/ID = fragmentation or isolation by distance; CRE = continuous range expansion; LDC = long distance colonization.

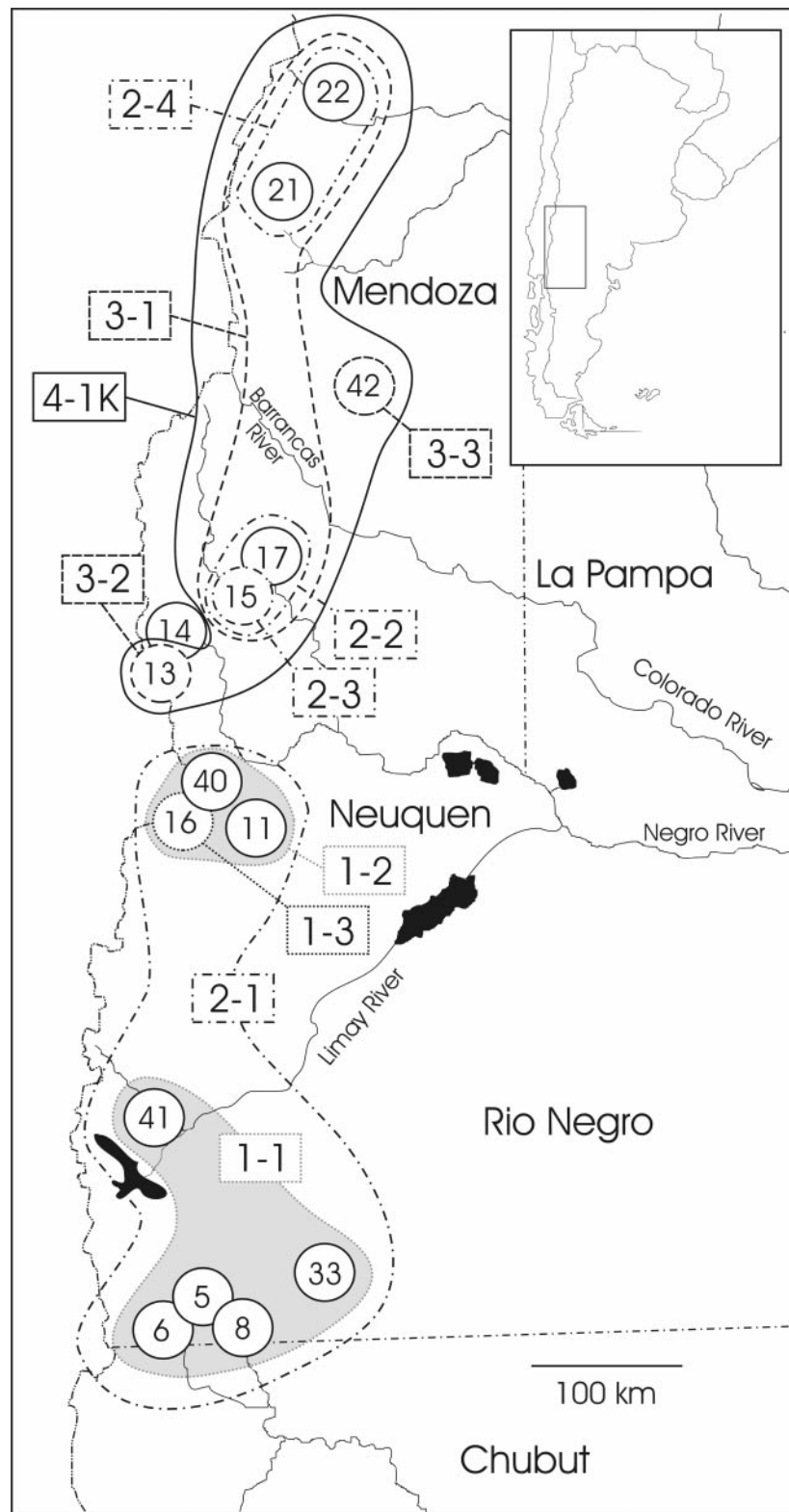


FIGURE 8. Map of the distribution of populations from the *kriegi* group, with the associated nesting design (Fig. 7A) relating the haplotypes from these localities. Numbers correspond to localities in Table 2.

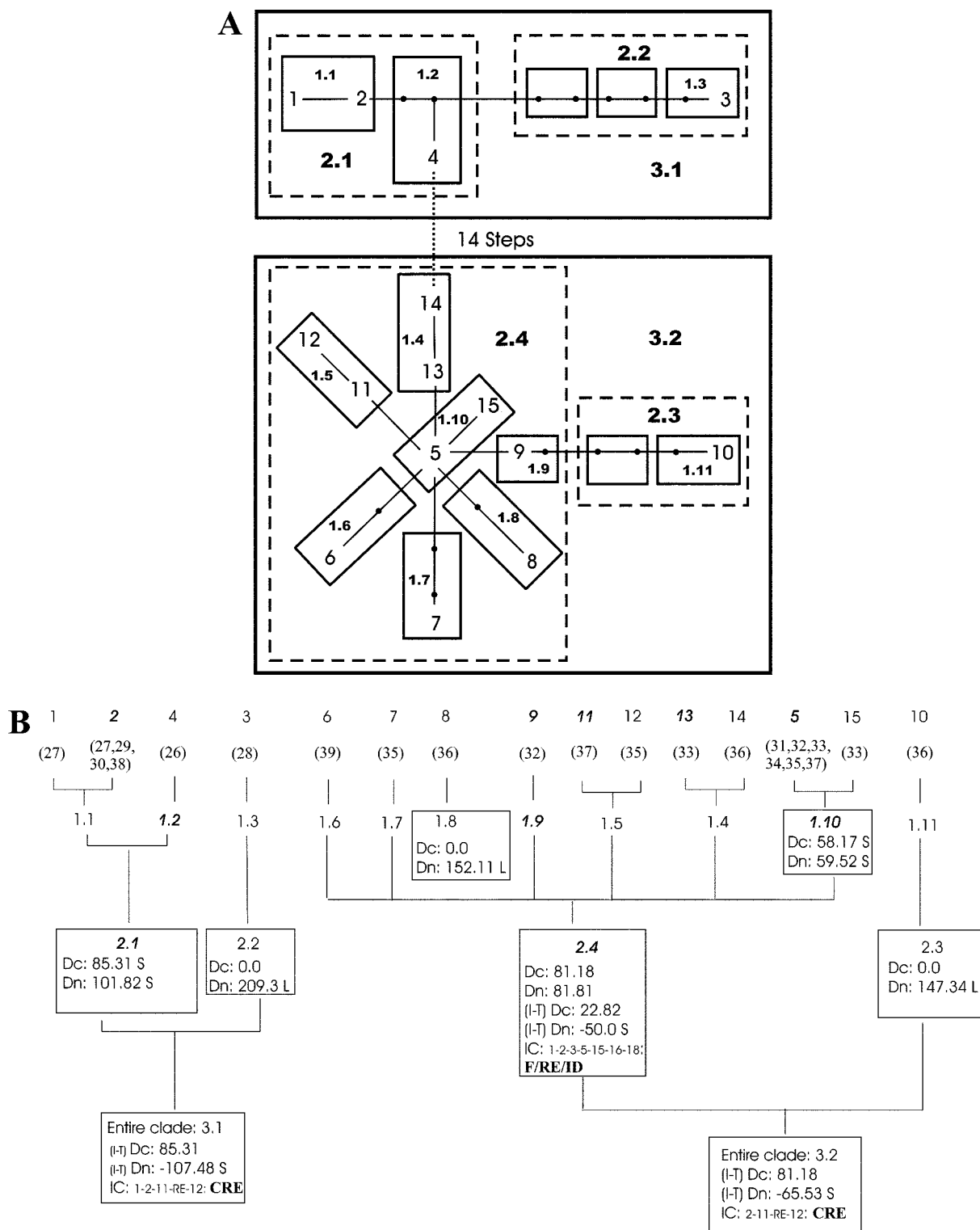


FIGURE 9. (A) Unrooted *cyt b* haplotype networks for the *petrophilus* group populations and their associated nested design; symbols convey the same information as in Figure 5A. (B) Summary of NCA results for haplotypes of the *petrophilus* group; organization is identical to that in Figure 5B, with the following inferences: F/RE/ID = fragmentation, range expansion, or isolation by distance; CRE = continuous range expansion.

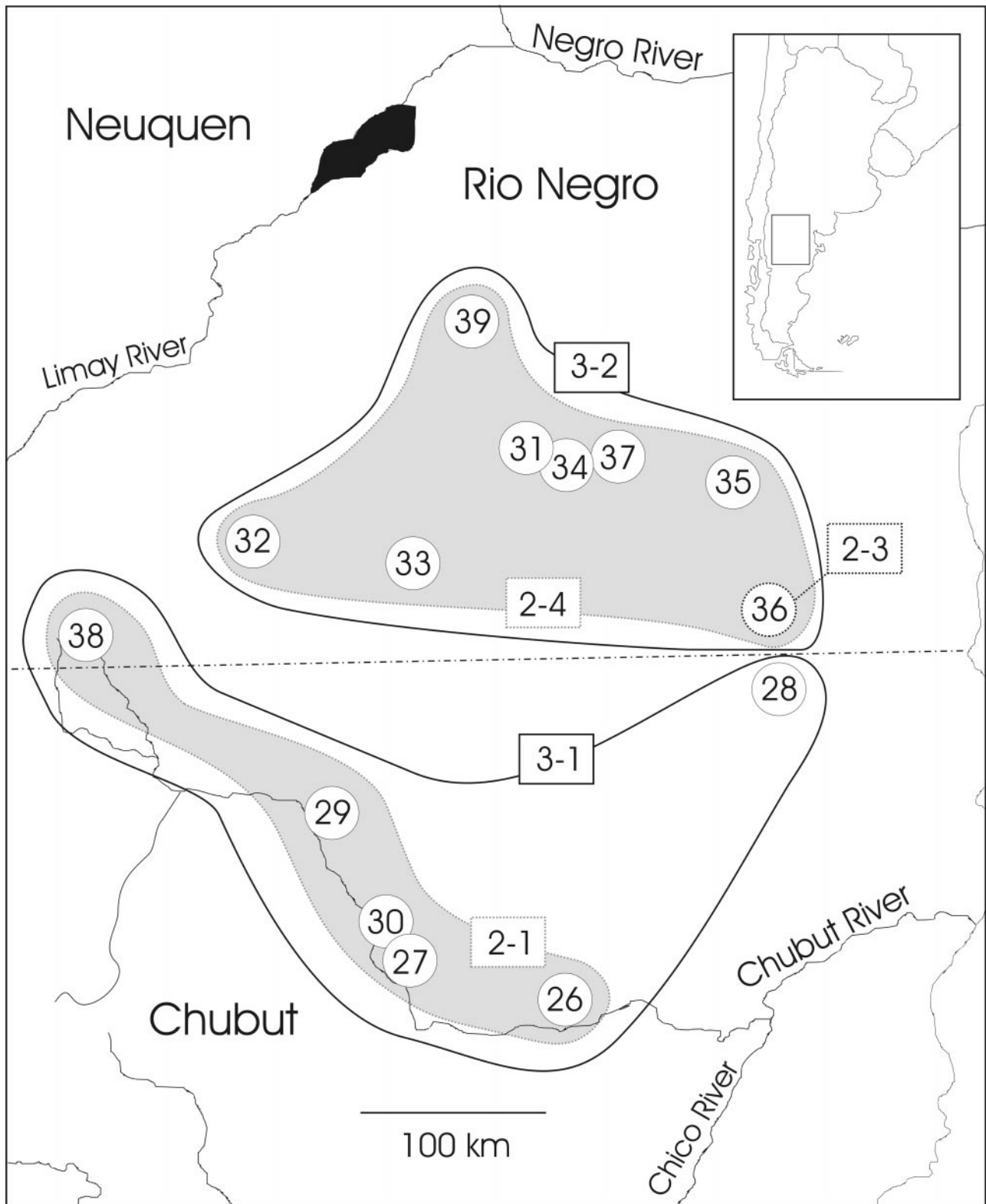


FIGURE 10. Map of the distribution of populations from the *petrophilus* group with the associated nesting design (Fig. 9A) relating the haplotypes from these localities. Numbers correspond to localities in Table 2.

Patagonian Steppe through the western part of Chubut province and probably to the northwestern edge of Santa Cruz province (Fig. 4). One of the populations included in this restricted definition of *L. elongatus* along the eastern edge of the distribution in Rio Negro province (locality 9; Fig. 6) has a very different chromatic pattern (completely black with no trace of a dorsal pattern). The NCA suggests that the genetic signature for this "black" population is consistent with its origin by long-distance colonization from western Patagonia (from locality 5 to locality 9; see clade 2-1 in Fig. 6). The NCA yields an inference of contiguous range expansion along the western Patagonian Steppe within clade 4.1E, and the type locality (territorio del Chubut cerca de la cordillera) for *L. elongatus* in the original description by Koslowsky (1896) is within the geographic range of this clade. We consider it the nominate species and the black population as a color morph of *L. elongatus*.

External to clade 4.1E, relationships of the remaining three clades within the *elongatus* group are ambiguous (*L. spp.* 5, 6, 7; Figs. 2, 3). In all phylogenies, these clades are exclusive, well supported, and concordant with geography, thus meeting WP criterion 1a for species recognition. Although no formal morphological study has been reported for these populations, some characteristics of body size and coloration appear to differentiate them (Avila et al., in prep.). *Liolaemus sp.* 7 has the larger body size and is more melanic in the head and body flanks than are any others in the *elongatus* group. *Liolaemus sp.* 6 has a slightly smaller body size than *L. elongatus* and *L. sp.* 7 and has a different color pattern. The two populations from *L. sp.* 5 (localities 21 and 22, Figs. 2, 3) are restricted to the highest valley of the Rio Tordillo–Rio Grande basin in southwestern Mendoza province (Fig. 1) and are distinguished from all other members of the *elongatus* group by smaller size and a distinct color pattern. Videla and Ceï (1996) described *Liolaemus thermarum* from a locality near our sampling sites (at a higher elevation of the same drainage system) but did not provide morphological comparisons to other species of the *elongatus* group.

Although our samples from *L. sp.* 5 may correspond to *L. thermarum*, examination of two specimens of the type series (MACN 36681, 36682) by one of us (L.J.A.) does not allow assignment of our samples to this species using morphological characters (e.g., absence of precloacal pores in *L. thermarum* whereas two to four pores are present in our samples), and sampling from the type locality for tissues was not possible in this study. Morphological and chromatic characters mentioned by Ceï (1974) and Bottari (1974) and museum specimens examined by L.J.A. (MVZ 188729, 188740–188745) suggest that *L. sp.* 5 extends into Mendoza province approximately 100 km north of our northernmost sample. Several morphological differences among *L. elongatus*, *L. sp.* 5, *L. sp.* 6, and *L. sp.* 7 were mentioned by Ceï (1974) and corroborate molecular evidence for species distinctness (Figs. 2, 3), but we consider these species provisional and in need of further corroboration with increased locality sampling and independent morphological and nuclear markers. The NCA does not lead to an inference of historical frag-

mentation and thus no additional species within clade 4.1E (Fig. 5B).

The kriegi group.—Based on morphological characters, Ceï (1986) grouped three species, *L. buergeri*, *L. ceii*, and *L. kriegi*, into a "kriegi group." Our samples cover the whole distributional range of these species, and phylogenetic analyses recovered a well-supported monophyletic group that includes two main clades, 2.1 and 4.1K (Figs. 2, 3), that are concordant with geography and show no evidence of gene flow among their basal lineages. Clade 2.1 (Figs. 2, 3) was strongly supported by all tree-reconstruction methods and includes all populations from Neuquén and Rio Negro provinces. These populations have been assigned to the species *L. kriegi* or *L. ceii* under conventional taxonomy (Table 2), but we found no evidence for this distinction at the molecular level. The NCA for clade 2.1 leads to an ambiguous inference of fragmentation/restricted gene flow with isolation by distance (in nested clade 1.1; Fig. 7) between its northernmost population (locality 41; Fig. 8) separated by the Limay River from the southern populations (localities 5, 6, 8, and 33; Fig. 8). The original description of *L. kriegi* was based on lizards from Estancia El Condor (Rio Negro province, near locality 41), whereas the type locality of *L. ceii* is Pampa de Lonco Luan (locality 16, clade 1.3; Figs. 7, 8). Both species were claimed by Ceï (1986) to occur in sympatry at both type localities and in all the distributional range of clade 2.1. We sampled both type localities and included the range of morphotypes fitting the original descriptions (which are very imprecise), and if two distinct species were present in our samples the mtDNA tree should have shown a pattern of interdigitation of two distinct sets of haplotypes among locations (WP criterion 1e). This pattern is not evident, and although additional sampling is needed, our results suggest that *L. ceii* and *L. kriegi* from these localities are conspecific, and we recommend the use of the prior name *L. kriegi*. Although the NCA inference was inconclusive, additional sampling may reveal evidence for fragmentation within clade 2.1 (Fig. 8) and the presence of additional species.

Clade 4.1K includes individuals from populations 21 and 22 that are identified as *L. buergeri* based on morphological characteristics (Figs. 2, 3), and according to the NCA inference for clade 3.1 (Fig. 7) their separation from the southern populations (localities 15 and 17; Fig. 8) is the result of either historical fragmentation or restricted gene flow with isolation by distance. Geographically, these populations (clades 2.2 and 2.4) are separated from each other by the Barrancas River (Fig. 8). All tree reconstructions recovered well-supported phylogenetic structure at most nodes within clade 4.1K and external to the *L. buergeri* terminal. We provisionally assign the label *L. sp.* C to terminals 15 and 17; this clade was strongly supported as exclusive and concordant with geography. The sample from locality 42 is provisionally recognized as *L. sp.* B because lizards from this locality are different from all others in clade 4.1K in color pattern, body shape, and some features of squamation that are strikingly similar to those of *L. austromendocinus* (*petrophilus* group);

Figs. 2, 3; Avila and Morando, unpubl. data). *Liolaemus* sp. A and *L.* sp. 8 (external to clade 4.1K; localities 13 and 14, respectively; Figs. 2, 3) are separated from each other by a straight line distance of ~20 km (Fig. 8), but locality 14 is ~300–400 m higher than locality 13. Given the deep phylogenetic structure and morphological complexity of clade 4.1K + *L.* sp. 8, our results are insufficient to draw firm conclusions about the species boundaries and the relationships among these populations.

The petrophilus group.—The *petrophilus* group includes three described species (*L. petrophilus*, *L. capillitas*, *L. austromendocinus*) plus three strongly differentiated lineages. One of these species (*L.* sp. 4; solid circle, Figs. 2, 3) is composed of populations previously assigned to *L. elongatus*. However, this clade is morphologically distinguishable (body size, coloration, and scalation) from all *L. elongatus* recognized here (clade 4.1E) and from its sister taxon *L. austromendocinus* (Figs. 2, 3; this relationship was reported by Schulte et al., 2000, under the name *L. elongatus*). This clade is distributed along Andean valleys from northern Mendoza province (Rio Mendoza valley) to south Catamarca province (southern Chaschuil Valley; Fig. 1). *Liolaemus austromendocinus* has a widespread distribution in Mendoza and Neuquén provinces between the Rio Diamante and Rio Neuquén basins (localities 43 and 44; Fig. 1) and appears to be homogeneous in morphology and coloration. However, some populations are restricted to apparently isolated geographic areas that we did not sample; research is ongoing for this species.

Liolaemus sp. 2 (open circle, Figs. 2, 3) from eastern Neuquén province (localities 46, 47; Fig. 1) was previously cited as *L. petrophilus* (Avila, 1996) based on general morphological similarity, but in life these lizards are characterized by a distinctive bright yellow-green background coloration. All known populations are restricted to a small region in central Neuquén province and thus also meet the criteria of strong nodal support for exclusivity and geographic concordance. This clade is currently being described as a distinct species (Avila et al., in prep.).

Liolaemus petrophilus includes all populations from the Central Patagonian Steppe plateaus from Rio Negro and Chubut provinces (Fig. 10) and was recovered as a single strongly supported clade by all tree-reconstruction methods (Figs. 2, 3). However, this species also includes strongly supported exclusive clades 3.1 and 3.2 (Figs. 2, 3, 9A) that are each concordant with geography (Fig. 10). The northern clade (3.2) includes populations from the Rio Negro province confined to the Somuncurá Plateau and adjacent small plateaus (localities 31–37, 39), and a southern clade (3.1) includes populations from Chubut province and the westernmost region of the Rio Negro province (localities 26–30, 38; Fig. 10). These clades both meet WP criterion 1a for species recognition, despite the apparent absence of morphological or chromatic differences between them. We therefore consider these clades as provisional species pending further study. The NCA inference suggests a continuous range expansion for each of these clades (Fig. 9), although results are ambiguous

for the largest nested clade (2.4) within clade 3.2 (Fig. 9B), so the possibility of fragmentation and additional species cannot be ruled out without additional sampling in this region.

DISCUSSION

Sampling Issues Relevant to the Combined Procedures Approach

When delimiting species is difficult because of poor historical descriptions and/or limited morphological data for the group under study, and a molecular approach is pursued, sampling design becomes important in two contexts. First, if too few individuals are sampled per locality, a false pattern of monophyly may be recovered in a gene tree when inclusion of more samples would yield different genealogies (Takahata and Slatkin, 1990; Hey, 1994; Hedin and Wood, 2002; Wiens and Penkrot, 2002). Second, insufficient geographic sampling may also be susceptible to recovery of a false signal of regional genealogical exclusivity (Neigel and Avise, 1993; Templeton et al., 1995; Templeton, 1998). Given the increased efficiency with which large molecular data sets can now be collected and analyzed, the strength of hypotheses about intraspecific phylogeographic patterns, species boundaries, and interspecific phylogenetic relationships will increasingly rest on details of sampling. Here, we examine the sampling protocol we advocate in the context of basic assumptions of coalescent theory and the limits placed on our evolutionary inferences and then suggest sampling guidelines for similar studies.

Coalescent theory describes a stochastic process that approximates a Wright–Fisher model of neutral evolution, and coalescence is often modeled as a continuous-time Markov process based on assumptions of haploidy (in its simplest form), discrete generations, random mating, no variance in reproductive success, no gene flow, and constant effective population sizes (reviewed by Nordborg, 2001). In this study, we have used linked mtDNA loci with different substitution rates to describe a sampling protocol whereby only the most rapidly evolving gene region need be sequenced for all individuals (*cyt b* in this case); we then used evidence for population exclusivity as justification for reducing the number of individuals needed for sequencing of slower regions. Thus, a key feature of our argument is based on the strength of our evidence for exclusivity of population samples.

Under the simplifying assumptions given above, the probability of sampling the deepest coalescent in a population (e.g., the most recent common ancestor [MRCA]) is given by $P_{\text{excl}} = (n - 1)/(n + 1)$, where n is the sample size per population (from Saunders et al., 1984). Our average sample size for the *cyt b* region is 3.2 lizards, and the above expression gives a P_{excl} of 0.50 with $n = 3$. Our largest sample of *cyt b* sequences (eight lizards from locality 9) gives a P_{excl} of 0.78, and increasing sample sizes to 10, 15, 20, 25, and 50 individuals per locality would raise P_{excl} to 0.82, 0.87, 0.90, 0.92, and 0.96, respectively. By conventional statistical wisdom, very large sample

TABLE 3. Some demographic assumptions influencing P_{excl} , common violations of these assumptions, and influence of violations on coalescent times (from Nordborg, 2001).

Demographic assumption	Common violation	Influence on time to MRCA
Haplotypes present in equal frequencies	Ancestral haplotype at highest frequencies	Accelerates coalescence
N_E constant	Temporal fluctuations in N_E	Accelerates coalescence
Reproductive success equal among individuals	Variance in reproductive success	Accelerates coalescence
No gene flow among populations	Gene flow present but at variable rates	Increases mean and variance in time to coalescence

sizes are required to attain $\alpha = 0.05$ level of significance, and in fact sampling 50 individuals from even a moderate number of natural populations of most organisms is prohibitive. Fortunately, the above assumptions are almost never met in nature, and many of them will reduce sample sizes needed to estimate exclusivity, although the interaction of multiple demographic factors permits only qualitative assessments here.

Table 3 summarizes the major demographic assumptions that influence P_{excl} , common violations of these assumptions, and how such violations might alter P_{excl} . The assumption of equal frequencies of haplotypes in a sample will not often be met; coalescent theory predicts that the ancestral haplotype will be present at the highest frequency, and this has been verified in many empirical data sets (Crandall and Templeton, 1994). When this is the case, a smaller sample size will be needed to capture the MRCA. The assumption of constant N_E is almost certainly never met in natural populations; small N_E accelerates coalescence, and populations that regularly fluctuate temporally are predominantly influenced by the N_E of the smallest bottleneck. Similarly, variance (σ^2) in reproductive success, in violation of the third assumption, will accelerate coalescence; the expected coalescence for a pair of lineages = N_E/σ^2 generations (see Nordborg, 2001). From a sampling perspective, both of these factors will again reduce the number of haplotypes needed to capture the MRCA per locality.

The final assumption in Table 3 is that no gene flow occurs among populations, and in contrast to the other factors that alter only branch lengths, gene flow can also alter the tree topology. Population structure is usually modeled in the coalescent context as a series of breeding groups subdivided into patches of fixed sizes under the same assumptions as described previously with an island migration model (Nordborg, 2001). This conservative migration model predicts rapid coalescence within patches under low gene flow, whereas under the reality of different patch sizes and quality, migration may be strong, that is, it may deviate significantly from an island model (e.g., nonconservative). Population structure therefore increases both the mean and the variance in time to MRCA, depending on whether it is conservative or strong. Gene flow between populations will normally increase the coalescence time for any one population because the MRCA may be influenced by a low-frequency immigrant haplotype in a sample (thus increasing time to the MRCA), and a larger sample will be needed to detect this immigration. However, in continuously distributed populations with large N_E , migration is nor-

mally expected to operate on a faster time scale than coalescence within a single population and will lead to rapid coalescence for all populations that are interconnected by gene flow. Capturing this component of tree structure will more likely depend on geographic sampling density, relative to the structure and distribution of the taxon of interest, rather than collection of a large number of individuals per locality.

From an operational perspective, the question remains: How does one sample in the field to maximize rigor of evolutionary inferences above and below the species boundary? Given the demographic aspects of coalescence, a consideration of the different role played by trees in population and phylogenetic inference can provide some guidelines. In phylogenetic inference, gene trees are used to estimate the species tree, and species are taken as lineages that are assumed to exist (Frost and Kluge, 1995) and can be treated as a model parameter. All internal branches are relied on to be long relative to within-species coalescent times (Nordborg, 2001), and there is usually little doubt about the correct interpretive model for a species tree. In most cases, the model will be one of isolation giving rise to independent lineages whose pathways of descent can be recovered as a branching tree. For population trees, none of this is true because the topology and branch lengths of a gene genealogy may depend on many model parameters (Table 3; more are discussed by Nordborg, 2001), but combining these approaches will often provide trees likely to be dominated by a small number of deeper branches, and a much larger number of very shallow branches. Our topologies (Figs. 2, 3) manifest this structure, and as a consequence of the reality of only a single genealogy, coalescent events occur much more rapidly in the shallow parts of the tree (there are many more branches to "find each other"; see Nordborg, 2001: eq. 4), and increasing population sample sizes is surprisingly ineffective at increasing resolution over the entire tree.

Empirically, the consequence of combined sampling of populations and lineages (species) suggests two things. First, because it will not be feasible to collect sufficiently large samples to demonstrate local population exclusivity at the $\alpha = 0.05$ level, we will have to settle for a lower and admittedly arbitrary P_{excl} per locality. A sample size of 5–10 individuals per locality should be adequate because these sample sizes converge quickly on P_{excl} values as a function of time since isolation (expressed in $2N_E$ generations) under the unrealistic assumption of equal N_E in ancestral and daughter lineages (see Hey, 1994: fig. 3). More is always better of course, but for many

studies these statistical limitations will be unavoidable, and the problem is strikingly parallel to the identification of a diagnostic character by estimating its “fixation” in a population with an $\alpha = 0.05$ degree of statistical confidence. The sample sizes needed become impossibly large in most cases (Wiens and Servedio, 2000).

The second point relative to field sampling and coalescence is that increased geographic density of sampling points adds more shallow branches to a gene tree; it would in our case add more terminals to the five major clades in Figures 2 and 3 and to many of the ingroup terminals outside of these clades. These short branches are expected to coalesce rapidly relative to the deeper branches, which would allow stronger inferences for geographic regions included in the tree at the cost of imprecise resolution of the exclusivity issue for each local sample. Sampling density must be considered in the context of the dispersal capabilities of the target taxon, and for widely distributed low-vagility species the ideal design may not be knowable a priori (Hedin and Wood, 2002). However, the NCA can frequently detect insufficient sampling, and every deme need not be sampled as long as gene flow is high enough to provide a signature for a regional demographic structure. A field sampling design will frequently be limited by logistical and/or permit issues, but with sufficient geographic and character sampling, the gene hierarchy design used here will be adequate to recover both deep and shallow histories (albeit with acceptance of P_{excl} at a lower limit than $\alpha = 0.05$) and will provide sufficient data for the NCA and WP criteria to yield strong inferences. For any set of evolutionary hypotheses derived from the combinational approach, both deep and shallow branches can be further tested by new characters and other methods (e.g., direct estimates of gene flow from other markers can be used to refine and extend demographic or historical inferences derived from the NCA). This approach will foster productive dialogue about how we simultaneously assess evolutionary processes at, above, and below the species boundary.

Strengths of the Combined Procedures Approach

We have emphasized sampling design limitations (which are often ignored) in a manner that may make our own data appear too preliminary to be conclusive, but this interpretation understates the value of this study. We have extended the combined approaches of CF and modified the WP protocol in both directions—above and below the species level—with a hierarchical gene sampling protocol that emphasizes use of a rapidly evolving gene region to first assess exclusivity of samples. This approach established *cyt b* as the most variable region, and *cyt b* was then used to test for genealogical exclusivity of haplotypes from all localities represented by more than a single sequence. Bayesian analysis revealed that *cyt b* haplotypes were exclusive at all localities (within our sample limits) external to the shaded clades identified in Figures 2 and 3, but deeper phylogenetic relationships were poorly resolved. In cases in which *cyt b*

haplotypes could be linked into networks with 95% probability under the assumptions of Templeton et al. (1992), the same networks were strongly supported ($PP \geq 0.95$) as exclusive by the Bayesian analysis. When haplotypes were strongly supported ($PP \geq 0.95$) as exclusive per locality, then a single haplotype was (usually) selected to represent that locality (the individuals for which we had also collected the ND4 and 12S sequences) in subsequent phylogenetic analyses. This protocol allows a reduction in the number of terminals needed for sequencing additional genes by identifying regions with lower variability and for which one or very few individuals per locality should capture deeper tree-based phylogenetic structure.

Our data set included 207 lizards (183 from 44 localities represented the focal species), for which 194 individuals were sequenced for *cyt b*, 161 for ND4, and 118 for 12S. Our analysis of the *cyt b* data matrix alone resulted in poor resolution of almost all deep nodes, and because ND4 and 12S especially were much slower, we would have gained little or no additional resolution had we sequenced every individual for all three gene regions. Our protocol saved us a minimum of 109 sequences ($194 \times 2 = 388$, minus $[161(\text{ND4}) + 118(\text{12S})] = 109$) at a cost of approximately US\$12 per gene (both strands) per lizard. This cost includes the lane charge as listed on the BYU DNASC website (\$3/lane), plus our estimates of \$6 per gene per lizard for PCR, sequencing, and cleaning (assuming that all PCR and sequencing reactions worked the first time), which equals $109 \times \$12 = \$1,308$. By extension, had we been able to sample up to five individuals per location from all ingroup taxa alone (Table 2), we would need to add a minimum of 122 more reactions to collect the *cyt b* sequences, and assuming that exclusivity was verified at each locality, we would not need any additional ND4 or 12S sequences, thus saving 244 more reactions for each of these regions (cost saved = $488 \times \$12 = \$5,856$). This protocol will thus maximize the number of localities that can be sampled and the number of gene sequences (of variable evolutionary rates) that can be included in an analysis, while minimizing the cost and labor of data collection and retaining maximum power of statistical inference across a range of divergence levels. Further, we see no reason that it could not be profitably extended to other kinds of data. For example, museum specimens may provide large series from which external characters can be collected but for which skeletal material is usually available for far fewer individuals (Wiens, 2000). Given these constraints, a similar hierarchical sampling design would be worth exploring for different classes of morphological characters that evolve at different rates.

From an analytical perspective, we have shown the feasibility of constructing population networks for multiple clades whose internal structure could not be resolved by conventional tree-building methods (our preference for the NCA is based on the larger number of assumptions required by alternatives; Emerson et al., 2001: box 3) and then combining this information with markers that are uninformative at the population level into

tree-reconstruction analyses in a manner that reduces the operational number of terminals. This approach recovered shorter trees with increased phylogenetic signal at most nodes than otherwise would have been possible, regardless of the method (Bayesian, ML, or MP) employed. On the basis of these powerful analytical tools and a deeper appreciation of the importance of sampling a larger number of individuals, populations, and characters, our protocol should provide a means to guide collection of very large data sets and explore species boundary questions with greater statistical rigor.

Species Limits in the L. elongatus-kriegi Complex

Table 4 summarizes the possible species identified by application of the combined approach to the original focal species sampled in this study. Three of the five original focal species are hypothesized to be complexes of species on the basis of WP criteria and/or an NCA inference of historical fragmentation, with or without corroboration from morphological data, and two focal species (*L. kriegi* and *L. ceii*) are hypothesized to be conspecific on the basis of all three lines of evidence. Only lineages that are formally supported by multiple lines of evidence (*L. elongatus*, *L. sp. 1*, *L. sp. 3*; Figs. 2, 3) are regarded non-provisionally as species. Some of others (*L. sp. 5*, *L. sp. 6*, *L. sp. 7*), although they meet the WP criteria for exclusivity and geographic concordance (Wiens and Penkrot, 2002: fig. 1a), are considered provisional pending completion of morphological studies. Other hypothesized species such as *L. sp. A*, *L. sp. B*, and *L. sp. 8*, are represented by single localities and are considered provisional

TABLE 4. Summary of the original focal species of this study, the species hypothesized within each of these, and the criteria supporting their recognition. Terminals that we consider provisional are indicated by a question mark. Morphological data are qualitative.

Focal <i>Liolaemus</i> species	Hypothesized species	Criteria ^a
<i>L. elongatus</i>	<i>L. elongatus</i> (restricted)	WP 1a, body size, coloration
	<i>L. sp. 4</i> (?)	WP 1a, coloration, dorsal scales?
	<i>L. sp. 5</i> (?)	WP 1a, body size, coloration, dorsal scales?
	<i>L. sp. 6</i> (?)	WP 1a, coloration
	<i>L. sp. 7</i> (?)	WP 1a, body size, coloration
<i>L. kriegi</i> and <i>L. ceii</i>	Conspecific = <i>L. kriegi</i>	WP 1f, NCA(CRE), similar morphology
<i>L. buergeri</i>	<i>L. buergeri</i> (restricted)	Morphology, NCA (F/ID) not basal
	<i>L. sp. A</i> (?)	Single locality
	<i>L. sp. B</i> (?)	Single locality
	<i>L. sp. C</i> (?)	WP criteria questionable
	<i>L. sp. 8</i> (?)	Single locality
<i>L. petrophilus</i>	<i>L. petrophilus</i> clade 3.1 (?)	WP 1a, NCA(LDC/CRE)
	<i>L. petrophilus</i> clade 3.2 (?)	WP 1a, NCA(LDC/CRE)

^aWP 1a = Wiens-Penkrot (2002) criterion of strong support for exclusivity plus geographic concordance; CRE = Continuous range expansion; F/ID = fragmentation or isolation by distance; LDC = long-distance colonization.

because WP and NCA protocols are difficult to apply in the absence of a geographic component even though these terminals are divergent at the molecular level and one is supported by morphological evidence (*L. sp. B*). The lineages defined as *L. petrophilus* clade 3.1 and *L. petrophilus* clade 3.2, which appear to be morphologically identical, are considered provisional pending some test of genetic or demographic exchangeability (Templeton, 2001).

The focal species *L. buergeri* is unique among the lineages recovered here because it is morphologically distinct and strongly supported by WP criterion 1a, but it is highly nested within clade 4.1K (Figs. 2, 3). Its non-basal position means that in a strict sense the WP method should not apply (Wiens and Penkrot were aware of the possibility of oversplitting nonbasal lineages), and this method, even though it can be applied to more nested (e.g. nonbasal) clades, "does not determine when splitting is no longer justified by the available data" (Wiens and Penkrot, 2002:73). Because we stated a priori as basal those lineages separated beyond the 95% limit given by the Templeton et al. (1992) algorithm (implemented in the NCA); we did not apply the WP method within lineages whose haplotypes are interconnected within a network (shaded clades; Figs. 2, 3). The same argument applies to *L. sp. C*. The clade to which these two terminals belong (4.1K) is characterized by an ambiguous NCA inference in which historical fragmentation cannot be distinguished from restricted gene flow with isolation by distance (Fig. 7B), which suggests inadequate sampling. If sampling of populations is implemented at geographic scales that greatly exceed individual dispersal distances (overdispersed sampling), then the NCA will be misleading because the genetic signatures of historical population fragmentation (which results from discontinuities in distributions) and isolation by distance (which results from restricted gene flow over a continuous distribution that is large relative to dispersal distances) will be identical (Hedin and Wood, 2002). This example emphasizes the importance of another aspect of sampling strategies, and the ability of the inference key to yield an ambiguous inference (clade 3.1, Fig. 7B) where sampling is inadequate is actually a strength of the method (Templeton, 2001).

The Andes Mountains and Patagonian Steppe sampled for this study include examples of populations discontinuously distributed over high-elevation landscapes of immense topological complexity and other populations that are broadly distributed across continuous habitats at lower elevations, and we strongly suspect that the ambiguities in some of our NCAs (Table 4) are artifacts of overdispersed geographic sampling for these lizards. If more complete collecting and analyses of additional data sets eventually corroborate all of the provisional species summarized here, the five original focal species, which are good examples of Good's (1994) "inertial species" within the *L. elongatus-kriegi* complex, will in reality represent at least 13 species (plus one new species recently discovered by Avila et al. [in prep.] and two new species in press [Espinoza and Lobo, 2003]). If all of these

species are eventually verified, this will triple the number of species in the original taxonomy of this complex. This complex is but 1 of approximately 16 complexes in *Liolaemus*, and a conservative estimate of doubling the species diversity in each of these complexes suggests that the true species number will be closer to 320 than the present 160+. We suspect that the same will be true for virtually all poorly studied groups in poorly collected but ecologically and topographically complex landscapes.

ACKNOWLEDGMENTS

Financial support was provided by a grant to L. J. A. (PEI 0178/98), a graduate fellowship M. M., and a postdoctoral fellowship to (L. J. A.) from Consejo Nacional de Investigaciones Científicas y Técnicas; by the Department of Integrative Biology and the M. L. Bean Museum of BYU; and by an NSF grant (DEB 98-15881) to J.W.S. We thank K. Pellegrino for laboratory training, K. Crandall for assistant with nested-clade analysis, and M. Terry and M. Porter for help with first submissions of PAUP files to the BYU supercomputer facility. J. Sullivan, A. Whitting, J. J. Wiens, and an anonymous reviewer made helpful comments on the manuscript. We thank C. H. F. Perez, N. Frutos, and D. R. Perez, for assisting in collecting lizards. M. I. Christie and L. Belver provided additional tissues samples. We also thank Y. Vilina (Geotecnica Consultores, Santiago, Chile) and Barrick Exploraciones Argentina S. A. for logistical support of the field trip to San Juan province. We thank G. Scrocchi and S. Kretschmar (FML), C. Cicero (MVZ), J. Williams (MLP), and G. Carrizo (MACN) for providing access to collections under their care.

REFERENCES

- ARÉVALO, E. S., S. K. DAVIS, AND J. W. SITES, JR. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. *Syst. Biol.* 43:387–418.
- AVILA, L. J. 1996. *Liolaemus elongatus petrophilus*: Ampliación de su distribución geográfica y primera cita para la provincia de Neuquén. *FACENA* 12:139–140.
- AVILA, L. J., AND F. LOBO. 1999. Geographic distribution: New lizard records for La Rioja and Catamarca provinces, northwestern Argentina. *Herpetol. Rev.* 30:115–117.
- AVISE, J. C., AND R. M. BALL. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. *Oxf. Surv. Evol. Biol.* 7:45–67.
- BAUM, D. A., AND K. L. SHAW. 1995. Genealogical perspectives on the species problem. Pages 289–303 in *Experimental and molecular approaches to plant biosystematics* (P. C. Hoch and A. G. Stephenson, eds.). Missouri Botanical Garden, St. Louis.
- BEHAVIDES, E., J. C. ORTIZ, AND J. W. SITES, JR. 2002. Species boundaries among the *Telmatobius* (Anura: Leptodactylidae) of the Lake Titicaca basin: Allozymes and morphological evidence. *Herpetologica* 58:31–55.
- BIRKY, C. W., JR., P. FUERST, AND T. MARUYAMA. 1989. Organelle gene diversity under migration, mutation, and drift: Equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics* 121:613–627.
- BIRKY, C. W., JR., T. MARUYAMA, AND P. FUERST. 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics* 103:513–527.
- BOTTARI, C. V. 1974. Poblaciones altitudinarias aisladas de *Liolaemus elongatus* del Volcán Payún Liso relaciones taxoserológicas y biogeográficas con el conjunto *elongatus-austromendocinus* del sureste árido de Mendoza. *Deserta* (Mendoza) 4:185–194.
- BROWER, A. V. Z. 1999. Delimitation of phylogenetic species with DNA sequences: A critique of Davis and Nixon's population aggregation analysis. *Syst. Biol.* 48:199–213.
- CASTELLOE, J., AND A. R. TEMPLETON. 1994. Root probabilities for intraspecific gene trees under neutral coalescent theory. *Mol. Phylogenet. Evol.* 3:102–113.
- CEI, J. M. 1974. Revision of the Patagonian iguanids of the *Liolaemus elongatus* complex. *J. Herpetol.* 8:219–229.
- CEI, J. M. 1975. *Herpetología patagónica*. X. El conjunto evolutivo de *Liolaemus elongatus*: Análisis serológico. *Physis* (B. Aires) C 34:203–208.
- CEI, J. M. 1979. The Patagonian herpetofauna. Pages 309–329 in *The South American herpetofauna: Its origin, evolution, and dispersal* (W. E. Duellman, ed.). Museum of Natural History, Univ. Kansas, Lawrence. Monograph 7.
- CEI, J. M. 1986. Reptiles del centro, centro-oeste y sur de la Argentina. *Herpetofauna de las zonas áridas y semiáridas*. Mus. Reg. Sci. Nat. Torino Monogr. 4:1–527.
- CEI, J. M. 1993. Reptiles del noroeste, nordeste y este de la Argentina. *Herpetofauna de las selvas subtropicales, puna y pampas*. Mus. Reg. Sci. Nat. Torino Monogr. 14:1–949.
- CLEMENT J., D. POSADA, AND K. A. CRANDALL. 2000. TCS, a computer program to estimate gene genealogies. *Mol. Ecol.* 9:1657–1659.
- CRANDALL, K. A., AND J. F. FITZPATRICK, JR. 1996. Crayfish molecular systematics: Using a combination of procedures to estimate phylogeny. *Syst. Biol.* 45:1–26.
- CRANDALL, K. A., AND A. R. TEMPLETON. 1994. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* 134:959–969.
- CRANDALL, K. A., A. R. TEMPLETON, AND C. F. SING. 1994. Intraspecific phylogenies: Problems and solutions. Pages 273–297 in *Models in phylogeny reconstruction* (R. W. Scotland, D. J. Siebert, and D. M. Williams, eds.). Clarendon Press, Oxford, U.K.
- DAVIS, J. I., AND K. C. NIXON. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. *Syst. Biol.* 41:421–435.
- DONNELLY, P., AND S. TAVARÉ. 1995. Coalecients and genealogical structure under neutrality. *Annu. Rev. Genet.* 19:401–421.
- DONOSO-BARROS, R. 1966. *Reptiles de Chile*. Ediciones de la Univ. Chile, Santiago, Chile.
- EMERSON, B. C., E. PARADIS, AND C. THEBAUD. 2001. Revealing the demographic histories of species using DNA sequences. *Trends Ecol. Evol.* 16:707–716.
- ESPINOZA, R., AND F. LOBO. 2003. Two new species of *Liolaemus* lizards from northwestern Argentina: Speciation within the northern subclade of the *elongatus* group (Iguania: Liolaemidae). *Herpetologica* 59:89–105.
- ESPINOZA, R., F. LOBO, AND F. CRUZ. 2000. *Liolaemus heliodermis*, a new lizard from northwestern Argentina with remarks on the content of the *elongatus* group (Iguania: Tropiduridae). *Herpetologica* 56:507–516.
- ETHERIDGE, R. 1995. Redescription of *Ctenoblepharis adspersa* Tschudi 1845, and the taxonomy of the Liolaeminae (Reptilia: Squamata: Tropiduridae). *Am. Mus. Novit.* 3142:1–34.
- ETHERIDGE, R., AND K. DE QUEIROZ. 1988. A phylogeny of Iguanidae. Pages 283–368 in *Phylogenetic relationships of lizard families: Essays commemorating Charles L. Camp*. (R. Estes and G. Pregill, eds.). Stanford Univ. Press, Palo Alto, California.
- ETHERIDGE, R., AND R. E. ESPINOZA. 2000. Taxonomy of the Liolaeminae (Squamata: Iguania: Tropiduridae) and a semi-annotated bibliography. *Smithson. Herpetol. Inf. Ser.* 126:1–64.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783–791.
- FETZNER, J. 1999. Extracting high-quality DNA from shed reptiles skins: A simplified method. *BioTechniques* 26:1052–1054.
- FROST, D. R., AND R. ETHERIDGE. 1989. A phylogenetic analysis and taxonomy of Iguanian lizards (Reptilia: Squamata). *Univ. Kans. Mus. Nat. Hist. Misc. Publ.* 81:1–65.
- FROST, D., AND A. G. KLUGE. 1995. A consideration of epistemology in systematic biology, with special reference to species. *Cladistics* 10:259–294.
- GOOD, D. A. 1994. Species limits in the genus *Gerrhonotus* (Squamata: Anguillidae). *Herpetol. Monogr.* 8:180–202.
- GU, X., Y. X. FU, AND W. H. LI. 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Mol. Biol. Evol.* 12:546–557.
- HEDIN, M., AND D. A. WOOD. 2002. Genealogical exclusivity in geographically proximate populations of *Hyplochilus thorelli* Marx (Araneae, Hypochilidae) on the Cumberland Plateau of North America. *Mol. Ecol.* 11:1975–1988.

- HELLMICH, W. C. 1951. On ecotypic and autotypic characters, a contribution to the knowledge of the evolution of the genus *Liolaemus* (Iguanidae). *Evolution* 5:359–369.
- HEY, J. 1994. Bridging phylogenetics and population genetics with gene tree models. Pages 435–449 in *Molecular approaches to ecology and evolution* (B. Schierwater, B. Streit, G. Wagner, and R. DeSalle, eds.). Birkhäuser-Verlag, Basel.
- HILLIS, D. M., AND J. J. BULL. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* 42:182–192.
- HOELZER, G. A. 1997. Inferring phylogenies from mtDNA variation: Mitochondrial gene-trees versus nuclear gene-trees revisited. *Evolution* 51:622–626.
- HOELZER, G. A., J. WALLMAN, AND D. J. MELNICK. 1998. The effects of social structure, geographical structure, and population size on the evolution of mitochondrial DNA: II. Molecular clocks and the lineage sorting period. *J. Mol. Evol.* 47:21–31.
- HUDSON, R. R., AND J. A. COYNE. 2002. Mathematical consequences of the genealogical species concept. *Evolution* 56:1557–1565.
- HUELSENBECK, J. P., AND J. P. BOLLBACK. 2001. Empirical and hierarchical Bayesian estimation of ancestral states. *Syst. Biol.* 50:351–366.
- HUELSENBECK, J. P., AND F. RONQUIST. 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics* 17:754–755.
- HUELSENBECK, J. P., F. RONQUIST, R. NIELSEN, AND J. P. BOLLBACK. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294:2310–2314.
- KOCHER, T. D., W. K. THOMAS, A. MEYER, S. V. EDWARDS, S. PAABO, F. X. VILLABLANCA, AND A. C. WILSON. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* 86:6196–6200.
- KOSLOWSKY, J. 1896. Sobre algunos reptiles de Patagonia y otras regiones argentinas. *Rev. Mus. La Plata* 7:447–457.
- LARGET, B., AND D. L. SIMON. 1999. Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Mol. Biol. Evol.* 16:750–759.
- LAURENT, R. F. 1983. Contribución al conocimiento de la estructura taxonómica del género *Liolaemus* Wiegmann (Iguanidae). *Bol. Asoc. Herpetol. Argent.* 1:15–18.
- LAURENT, R. F. 1984. On some iguanid genera related to or previously confused with *Liolaemus* Wiegmann. *J. Herpetol.* 18:357–373.
- LAURENT, R. F. 1985. Segunda contribución al conocimiento de la estructura taxonómica del género *Liolaemus* Wiegmann (Iguanidae). *Cuad. Herpetol.* 1:1–37.
- LEACHÉ, A. D., AND T. W. REEDER. 2002. Molecular systematics of the eastern fence lizard (*Sceloporus undulatus*): A comparison of parsimony, likelihood, and Bayesian approaches. *Syst. Biol.* 51:44–68.
- LEVITON, A. E., R. H. GIBBS, E. HEAL, AND C. E. DAWSON. 1985. Standards in herpetology and ichthyology: Part 1. Standard symbolic codes for institutional resource collections in herpetology and ichthyology. *Copeia* 1985:802–832.
- LEWIS, P. O. 2001. Phylogenetic systematics turns over a new leaf. *Trends Ecol. Evol.* 16:30–37.
- LOBO, F. 2001. A phylogenetic analysis of lizards of the *Liolaemus chilensis* group (Iguania: Tropiduridae). *Herpetol. J.* 11:137–150.
- MALLET, J. 1995. A species definition for the modern synthesis. *Trends Ecol. Evol.* 10:295–299.
- MOORE, W. S. 1995. Inferring phylogenies from mtDNA variation: Mitochondrial gene trees versus nuclear-gene trees. *Evolution* 49:718–726.
- NEIGEL, J. E., AND J. C. AVISE. 1993. Application of a non-equilibrium model to geographic variation in animal mitochondrial DNA distributions. *Genetics* 135:1209–1220.
- NORDBORG, M. 2001. Coalescent theory. Pages 1–37 in *Handbook of statistical genetics* (D. Balding, M. Bishop, and C. Cannings, eds.). Wiley, Chichester, U.K.
- NUÑEZ, H. 1992. Geographical data of Chilean lizards and snakes in the Museo Nacional de Historia Natural, Santiago, Chile. *Smithson. Herpetol. Inf. Serv.* 9:1–29.
- NUÑEZ, H., AND A. LABRA. 1985. *Liolaemus curis*, a new lizard from the Los Andes Range, Central Chile. *Copeia* 1985:556–559.
- NUÑEZ, H., AND J. C. TORRES-MURA. 1992. Adiciones a la herpetofauna de Chile. *Not. Mens. Mus. Nac. Hist. Nat. (Santiago)* 32:3–7.
- ORTIZ, J. C. 1981. Révision taxonomique et biologie des *Liolaemus* du groupe *nigromaculatus* (Squamata, Iguanidae). Doctoral Thesis, Univ. Paris.
- PALUMBI, S. R. 1996. Nucleic acids I: The polymerase chain reaction. Pages 205–247 in *Molecular systematics*, 2nd edition (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland, Massachusetts.
- PALUMBI, S. R., F. CIPRIANO, AND M. P. HARE. 2001. Predicting nuclear gene coalescence from mitochondrial data: The three-times rule. *Evolution* 55:859–868.
- POSADA, D., AND K. A. CRANDALL. 1998. ModelTest: Testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- POSADA, D., AND K. A. CRANDALL. 2001. Intraspecific gene genealogies: Trees grafting into networks. *Trends Ecol. Evol.* 16:37–45.
- POSADA, D., K. A. CRANDALL, AND A. R. TEMPLETON. 2000. GeoDis, a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Mol. Ecol.* 9:487–488.
- PUERTO, G., M. DA GRACA SALOMAO, R. D. G. THEAKSTON, R. S. THORPE, D. A. WARRELL, AND W. WUSTER. 2001. Combining mitochondrial DNA sequences and morphological data to infer species boundaries: Phylogeography of lanceheaded pitvipers in the Brazilian Atlantic forest, and the status of *Bothrops pradoi* (Squamata: Serpentes: Viperidae). *J. Evol. Biol.* 14:527–538.
- RANNALA, B., AND Z. YANG. 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *J. Mol. Evol.* 43:304–311.
- SAUNDERS, I. W., S. TAVARE, AND G. A. WATTERSON. 1984. On the genealogy of nested subsamples from a haploid population. *Adv. Appl. Prob.* 16:471–491.
- SCHULTE, J. A., II, J. R. MACEY, R. E. ESPINOZA, AND A. LARSON. 2000. Phylogenetic relationships in the iguanid lizard genus *Liolaemus*: Multiple origins of viviparous reproduction and evidence for recurring Andean vicariance and dispersal. *Biol. J. Linn. Soc.* 69:75–120.
- SCOLARO, J. A. 1993. Geographic distribution: *Liolaemus petrophilus*. *Herpetol. Rev.* 24:109.
- SITES, J. W., JR., AND K. A. CRANDALL. 1997. Testing species boundaries in biodiversity studies. *Conserv. Biol.* 11:1289–1297.
- SWOFFORD, D. L. 2001. PAUP*: Phylogenetic analysis using parsimony (*and other methods), beta version 4.0.b4b. Sinauer, Sunderland, Massachusetts.
- SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, AND D. M. HILLIS. 1996. Phylogenetic inferences. Pages 407–514 in *Molecular systematics*, 2nd edition (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland, Massachusetts.
- TAKAHATA, N., AND M. SLATKIN. 1990. Genealogy of neutral genes in two partially isolated populations. *Theor. Popul. Biol.* 38:331–350.
- TEMPLETON, A. R. 1998. Nested clade analyses of phylogeographic data, testing hypotheses about gene flow and population history. *Mol. Ecol.* 7:381–397.
- TEMPLETON, A. R. 2001. Using phylogeographic analyses of gene trees to test species status and processes. *Mol. Ecol.* 10:779–791.
- TEMPLETON, A. R., K. A. CRANDALL, AND C. F. SING. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132:619–633.
- TEMPLETON, A. R., E. ROUTMAN, AND C. A. PHILLIPS. 1995. Separating population structure from population history: A cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140:767–782.
- TEMPLETON, A. R., AND C. F. SING. 1993. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* 134:659–669.
- THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, AND D. G. HIGGINS. 1997. The clustal X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 24:4876–4882.
- VIDELA, F., AND J. M. CEI. 1996. A new peculiar *Liolaemus* species of the “chiliensis” phyletic group from the volcanic Cordilleran landscapes of southern Mendoza Province, Argentina (Iguania, Lacertilia, Reptilia). *Mus. Reg. Sci. Nat. Boll. Torino* 14:505–516.
- WHEELER, Q. D., AND R. MEIER (eds.). 2000. *Species concepts and phylogenetic theory: A debate*. Columbia Univ. Press, New York.

- WIENS, J. J. 1998. Combining data sets with different phylogenetic histories. *Syst. Biol.* 47:568–581.
- WIENS, J. J. 1999. Polymorphism in systematics and comparative biology. *Annu. Rev. Ecol. Syst.* 30:327–362.
- WIENS, J. J. (ed.). 2000. *Phylogenetic analysis of morphological data*. Smithsonian Institution Press, Washington, D. C.
- WIENS, J. J., AND T. A. PENKROT. 2002. Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Syst. Biol.* 51:69–91.
- WIENS, J. J., T. W. REEDER, AND A. NIETO MONTES DE OCA. 1999. Molecular phylogenetics and evolution of sexual dichromatism among populations of the Yarrow's spiny lizard (*Sceloporus jarrovi*). *Evolution* 53:1884–1897.
- WIENS, J. J., AND M. R. SERVEDIO. 2000. Species delimitation in systematics: Inferring diagnostic differences between species. *Proc. R. Soc. Lond. Ser. B.* 267:631–636.
- YANG, Z. 1994. Estimating the pattern of nucleotide substitution. *J. Mol. Evol.* 39:105–111.
- YANG, Z., AND B. RANNALA. 1997. Bayesian phylogenetic inference using DNA sequences: A Markov chain Monte Carlo method. *Mol. Biol. Evol.* 14:717–724.

*First submitted 6 June 2002; reviews returned 8 September 2002;
final acceptance 28 October 2002
Associate Editor: Jack Sullivan*