2006-12-07

Evolution in Neotropical Herpetofauna: Species Boundaries in High Andean Frogs and Evolutionary Genetics in the Lava Lizard Genus Microlophus (Squamata: tropiduridae): A History of Colonization and Dispersal

Edgar Benavides
Brigham Young University - Provo

EVOLUTIONARY GENETICS IN THE LAVA LIZARD GENUS *MICROLOPHUS* (SQUAMATA: TROPIDURIDAE): A HISTORY OF COLONIZATION AND DISPERSAL

by

Edgar Benavides

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

Doctor of Philosophy

Department of Integrative Biology
Brigham Young University
December 2006

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a dissertation submitted by

Edgar Benavides

This dissertation has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

_____	_____
Date	Jack W. Sites, Jr., Chair
_____	_____
Date	Leigh Johnson
_____	_____
Date	David McClellan
_____	_____
Date	Keith A. Crandall
_____	_____
Date	Michael Whiting

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the dissertation of Edgar Benavides in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

_____ _____

Date Jack W. Sites, Jr.
Chair, Graduate Committee

Accepted for the Department

_____

Keith A. Crandall
Department Chair

Accepted for the College

_____

John D. Bell
Dean, College of Biology and Agriculture

ABSTRACT

EVOLUTIONARY GENETICS IN THE LAVA LIZARD GENUS *MICROLOPHUS* (SQUAMATA: TROPIDURIDAE): A HISTORY OF COLONIZATION AND DISPERSAL

Edgar Benavides
Department of Integrative Biology
Doctor of philosophy

In this collection of papers I have summarized my investigations into the field of evolutionary genetics and more specifically into patterns of biodiversity and evolutionary processes. The lizards (and frogs) studied here share common features in that they are largely present in unique environments, which are also regions that are biologically understudied. Most of these taxa show high degrees of endemism, interesting natural history characteristics, and each group manifests distinctive adaptations of general evolutionary interest.

My work in the genus *Telmatobius* has been a progressive approach that began in my MS program, and it first focused on alpha taxonomy, morphological variation, and species boundaries. This work led to new studies initiated and completed at BYU involving further taxonomic revision (Formas et al., 2003; Chapter 1), and then revisiting and re-evaluating species boundaries established earlier (with allozyme markers) and this time with population level molecular (mitochondrial DNA) markers (Chapter 2). Our results indicate that the striking differences in size, coloration and general appearance in the various Lake Titicaca morphotypes are not genetically based. Further, there is evidence that these morphotypes have evolved very rapidly after demographic bottlenecks eroded present genetic variability. *Telmatobius* frogs of Lake Titicaca are listed by the International (IUCN) as critically endangered. We support this classification and further suggest studies to explore open questions like the possibility of adaptation along ecological resource gradients.

Lizards of the genus *Microlophus* are interesting but for different reasons, and studies of this group constitutes the bulk of my dissertation work. The genus includes both Galapagos insular species, and continental taxa distributed in a linear gradient along > 4000 km of the western coast of South America. In studying *Microlophus* I first tackled the unresolved phylogenetic relationships within the genus (Chapter 3) and then pay attention to phylogeographic aspects of the most speciose lizard radiation in the Galapagos Archipelago (Chapter 4).

Chapter 3 is a single manuscript provisionally accepted in the journal *Systematic Biology*. This paper introduces the lizard genus *Microlophus* ("lava lizards") as a study system, and includes a large nuclear data set accompanied by an equally large mitochondrial data set (7877 characters in total). This paper explicitly differentiates among sequence alignments of gene regions that vary in tempo and class of mutational events. We show that this recognition is important and we suggest ways to appropriately deal with the alignment of multi-locus non-coding DNA data sets. A secondary finding in this study is that mtDNA and nDNA topologies are discordant with each other but that both are strongly supported, and that the nuclear topology is concordant with species distribution patterns along coastal South America. We hypothesize that in this particular region of the tree, the nuclear genome recovers a topology that is closer to the species tree, and conflicts occur due to likely secondary contact of distantly related taxa, suggesting that unique taxonomic relationships in the mtDNA gene tree are the result of hybridization. This last point highlights the value of dense taxonomic and character sampling for teasing apart different aspects of evolutionary processes.

Chapter 4 is a manuscript to be submitted to the journal *Evolution*; in this study we further investigate the most speciose radiation of *Microlophus* in the Galapagos, based on an unparalleled sampling of most islands and small islets in the Archipelago. We use mtDNA sequences to both test hypothesized between-island colonization routes, as well as the expectation that within-island phylogeographic structure should be greater on older islands. Our mtDNA gene tree is strongly supported and allows rejection of previous alternatives, and we propose a novel sequence of between-island colonization events. Our results also reject the idea of phylogeographic structure been related solely to island age. Instead, we provide evidence to suggest that active volcanism as a major

player in the generation of genetic diversity in within-island environments, and this is further compounded by the seemingly stochastic nature of within-island long-distance colonization routes mediated by ocean currents. We suggest that the direction and intensity of these currents, as currently understood, are insufficient to generate *a priori* hypotheses of oceanic colonization routes and their influence on gene flow. We do show that the standard stepping-stone model of migration, where genetic interchange is only possible among neighboring localities, does not explain much of the within-island population genetic structure unraveled by this study.

From a biological conservation perspective the study of patterns of recent evolutionary history in the Galapagos provides with a window to evolutionary processes that have shaped and continue to impact the generation of biodiversity in the Galapagos Archipelago. Islands have long been viewed as natural laboratories of evolutionary change, and thus all island isolates are or could be distinctly important components of the larger, archipelago-wide processes. We provide working hypotheses for some of the demographic processes that might be generating within- and between-island biodiversity in this clade of lizards; confirmation of these explanations with independent data will have management implications for conserving the unique patterns observed in the Galapagos biota, but also the processes that generated these patterns.

ACKNOWLEDGMENTS

I am deeply indebted to so many people; some have helped me with the sampling of specimens, some through museum loans, some through academic advice, some have accompanied me while traveling, and a few have shared hours of frustration or content. All of them I consider my friends, all of them I remember well, and most of your names should be mentioned somewhere in this collection of papers. This dissertation stems from resolution, and resolution stems from simple friendships here in Provo and other parts of the US, to Chile, Peru, and Bolivia, and so many other places. There is an invisible and powerful bond made of true aspirations and honest feelings. Such is the fabric of this endeavor and I am grateful to each and every single one these friends.

I also want to acknowledge the immense love of my parents and siblings. I think of mom always helping others, always doing well for the sake of it, I think of dad and his honesty and hard work. I think of Jean Paul, Valeria and Juan Luis always dear and near to my heart, I feel humbled by their memory and I'm so happy when we are all together.

Parting ways is not easy, leaving behind people that I loved, admired or shared things with; it's both sad and somehow unreal. I want to express my admiration to Veronica because she has always been so much more than a friend and confidant, I wish you the best. I want to express my gratitude to the undergraduates that dutifully worked with me, among others; Rebecca, John, Paige and Ian helped me so much that is hard to just say thanks, but I know you all will have plentiful lives, and as far I'm concerned successful ones. I also thank my committee members for their patience and encouragement; I have appreciated your advice, even though I don't always say so.

Finally, I need to write some about my advisor Jack and his wife Joanne; like the John Mellencamp song, they are wonderful people, thanks for accepting me as I am. Thanks Jack, for showing and teaching me about commitment and dedication, not so many have the privilege of working with someone so versed in academic subjects as well as being a deeply humane person. At times it seemed that I had four instead of a single advisor; and boy you drove me crazy, but believe me, it was worth it. Every minute of it.

TABLE OF CONTENTS

Herpetologica, 59(2), 2003, 253–270

A NEW SPECIES OF *TELMATOBIUS* (ANURA: LEPTODACTYLIDAE) FROM RÍO VILAMA, NORTHERN CHILE, AND THE REDESCRIPTION OF *T. HALLI* NOBLE

J. Ramón Formas[1,4], Edgar Benavides[2,3], and César Cuevas[1]

[1]*Instituto de Zoología, Universidad Austral de Chile, Casilla 567, Valdivia, Chile*
[2]*Casilla 160-C, Departamento de Zoología, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Concepción, Chile*

Abstract: We describe a new species, *Telmatobius vilamensis,* from Río Vilama, near San Pedro de Atacama, Chile. Previously confused with *Telmatobius halli* Noble, 1938, this new species is diagnosed by distinct characteristics of the adult and larval morphology and chromosomes. The new species is the southernmost representative of the genus *Telmatobius* in Chile. *Telmatobius halli* Noble, 1938 is redescribed, and both taxa are compared with eight Chilean congeners that occur in the Andes.

Key words: Anura; Karyotype; Leptodactylidae; Osteology; San Pedro de Atacama; Tadpoles; *Telmatobius halli* redescription; *Telmatobius vilamensis* new species

Species of the genus *Telmatobius* represent one of the most diverse components of the Andean batrachofauna above 2000 m. The genus is comprised of around 50 species distributed from southern Ecuador to different latitudes on the eastern (22° 51′ S, 68° 11′ W) and western (29° 39′ S, 68° 36′ W) slopes of the Andes of northern Chile and Argentina. New species have recently been described primarily from Bolivia and Perú (de la Riva, 1994*a,b*; Lavilla and Ergueta, 1995*a,b*, 1999; Salas and Sinsch, 1996). Nevertheless, recent herpetological exploration of the southwestern limits of the range of the genus has resulted in the unexpected discovery of three new Chilean species of *Telmatobius: T. dankoi* (Formas et al., 1999), *T. philippii* (Cuevas and Formas, 2002), and *T. fronteriensis* (Benavides et al., 2002*a*). These three are in addition to five other species of *Telmatobius* previously reported from northern Chile (*T. peruvianus* Wiegmann, 1835; *T. marmoratus* [Duméril and Bibron, 1841]; *T. halli* Noble, 1938; *T. pefauri* Veloso and Trueb, 1976; and *T. zapahuirensis* Veloso et al., 1982; Fig. 1). Another new species described herein is the southernmost representative of the genus in Chile. Diagnosis of the new species is based on external morphology of adults and tadpoles, osteology, karyotype, chromosomal C-banded patterns, and Ag-NOR (nucleolar organizer region) position. Because this new species has been confused with *T. halli* (e.g., Cei, 1962; Veloso et al., 1982), a redescription, diagnosis, and illustrations of the holotype of *T. halli* are presented. Finally, we provide morphological and karyological comparisons among Chilean species as well as a key to their identification.

Materials and Methods

This description is based on 25 adult frogs, a juvenile, and 5 tadpoles collected at Río Vilama near San Pedro de Atacama, Chile (Fig. 1). Geographic coordinates were obtained using a Trimble Global Positioning System (GPS). With the help of nets, specimens were collected below dense vegetation in flowing water. Photographs, notes on the habitat, and color in life were taken in the field at the time of capture. Vouchers and additional specimens used in this study (Appendix I) are deposited in the following collections: Colección Boliviana de Fauna (CBF); Instituto de Zoología, Universidad Austral de

[3] Present Address: Department of Zoology and M. L. Bean Life Science Museum, Brigham Young University, Provo UT 84602, USA.
[4] Correspondence: e-mail, rformas@uach.cl

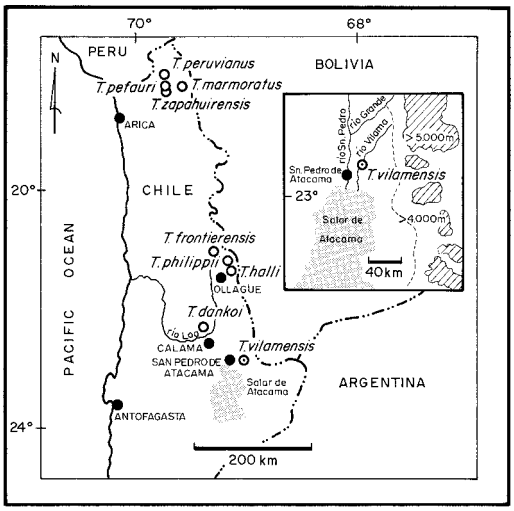


FIG. 1.—Distribution of the species of *Telmatobius* in Chile. The inset shows the type locality of *Telmatobius vilamensis*.

Chile (IZUA); Museo de Zoología, Universidad de Concepción (MZUC); Departamento de Biología Celular y Genética Universidad de Chile (DBCUCH); Departamento Biomédico, Universidad de Antofagasta (DBMUA); and Museo Nacional de Historia Natural, Chile (MNHN). The redescription of *T. halli* Noble is based on the holotype and two paratypes deposited in the American Museum of Natural History (AMNH).

The following variables were measured using dial calipers (to the nearest 0.1 mm): snout–vent length (SVL), tibia length (knee to heel), foot length (proximal edge inner metatarsal tubercle), head length (posterior corner of jaw to tip of the snout), head width (from posterior corner of jaw), interorbital distance, eye–nostril distance (from anterior corner of eye), internarial distance, eye diameter (between anterior and posterior corners of eye), hand length (proximal edge of outer metacarpal tubercle to tip of third finger), and radioulnar length (elbow to distal edge of outer metacarpal tubercle). Sexual maturity was assessed by presence of eggs in females and the degree of development of nuptial excrescences in males. Formulas for toe webbing are those of Savage and Heyer (1967), as modified by Myers and Duellman (1982). Osteological observations were made on three adult specimens (two males IZUA 3146 [SVL 44.16, 46.81 mm] and a female IZUA 3224 [SVL 47.51 mm]); cleared-and-stained skeletons were prepared using the methods of Hollister (1934) and Song and Parenti (1995). Carpal elements were identified according to Fabrezi and Alberch (1996).

Tadpoles (IZUA 3132) were staged according to Gosner (1960). The following measurements were taken: total length (distance from the tip of the snout to the tip of the tail), body length (distance from the tip of the snout to the origin of the hind limb), body depth (maximum distance between the dorsal and ventral surfaces of the body), fin depth (maximum distance between margin of dorsal and ventral fins), snout–nostril distance (distance from the tip of the snout and the anterior border of the nostril), eye diameter (maximum distance between the anterior and posterior margins of the eye), and mouth width (maximum distance between the lateral margins of the oral disc).

Chromosomal characteristics were recorded from two male specimens (IZUA 3225–26), which were injected with 0.1% colchicine. After 2 h, frogs were anaesthetized with diethyl ether. A ventral incision was made under sterile conditions, and intestines were removed. Metaphase plates were obtained by macerating intestinal epithelium fragments. These were hypotonically treated with distilled water, fixed in acetic acid-alcohol (1:3), and finally placed in 45% acetic acid. Small fragments of tissue were placed between a glass slide and cover slip and dipped in liquid nitrogen. Thereafter, the cover slip was removed with a razor blade and chromosomes were allowed to air dry. After 3 d, chromosomes were stained for 15 min in Sörensen's phosphate buffer (pH 6.8), containing 4% Giemsa solution (Formas, 1991). Chromosomes were stained to reveal C-band patterns and Ag-NOR position, following the methodologies of Sumner (1972) and Rufas et al. (1982). Centromeric positions were determined according to Levan et al. (1964).

Accounts of Species

Telmatobius ***vilamensis*** sp. nov.

Holotype.—IZUA 3080, adult male collected by C. Cuevas, J. R. Formas, and C. Jara at the Río Vilama (22° 51′ 43″ S, 68° 23′ 25″ W; 3110 m), 6.5 km NE (by road) of San Pedro de Atacama, Provincia El Loa, Región de Antofagasta, Chile, on 17 April 1997 (Fig. 1).

Paratypes.—IZUA 3081, adult male; IZUA 3224 (cleared-and-stained adult female); same data as holotype; CBF 3760–62 (two females and a juvenile); MZUC 24104–06 (three males); MZUC 25725–38 (nine females and five males); collected by E. Benavides, H. Díaz, M. Hengst, and J. C. Ortiz at the type locality on 17 December 1998.

Diagnosis.—A moderate-sized *Telmatobius* (38.36–50.81 mm) having the following combination of characters: (1) premaxillary and maxillary teeth absent, (2) vomers rudimentary or absent, (3) neopalatines reduced, (4) tongue nearly ovoid, elongate, almost adhered to the floor of mouth, posterior border free, (5) choanae large, circular, (6) tympanum, tympanic annulus, and columella absent, (7) snout depressed in lateral view, (8) postfemoral folds present, (9) dorsum olive green (in life), (10) cranium poorly ossified, (11) 26 bi-armed chromosomes, (12) toes webbed, (13) outer border of Toe V moderately fringed.

Telmatobius vilamensis differs from all other known *Telmatobius* by having a prominently pointed snout (Fig. 2A), as well as a lean, hydrodynamic body shape. Among Chilean species, *T. dankoi* most closely resembles *T. vilamensis* in overall morphology. Both taxa share the following characters: postfemoral folds (thinner and smaller in *T. vilamensis*), premaxillary and maxillary teeth absent, oral disc transangular, tympanum and tympanic annulus absent, and size moderate. *Telmatobius vilamensis* is distinguished by its smooth skin (spiny in both sexes of *T. dankoi*), tongue nearly ovoid (rounded in *T. dankoi*), snout strongly depressed (not depressed in *T. dankoi*), and cranium poorly ossified (well ossified in *T. dankoi*).

Despite obvious external differences, both *T. vilamensis* and *T. dankoi* have been confused with *T. halli* (e.g., Cei, 1962, 1986; Northland et al., 1990; Veloso et al., 1982). These misidentifications probably resulted from the incomplete nature of the original description of *T. halli* (Noble, 1938) and the lack of an illustration of the type specimen. In addition to overall external morphology, *T. halli* also is distinguished by the absence of a postfemoral fold (present in *T. dankoi* and *T. vilamensis*); small maxillary teeth (absent in *T. dankoi* and *T. vilamensis*); round, thick tongue (nearly ovoid in *T. vilamensis*); and smooth skin (spiny in *T. dankoi*).

Body size distinguishes *T. vilamensis* (SVL 38.36–50.81 mm) from the larger *T. philippii* (SVL 41.49–53.70 mm) and the smaller *T. fronteriensis* (SVL 31.62–43.20 mm). *Telmatobius vilamensis* further differs from these taxa by having a postfemoral fold (absent in *T. philippii* and *T. fronteriensis*) and lacking vomerine (present but reduced in *T. philippii* and *T. fronteriensis*), premaxillary, and maxillary teeth (present in *T. philippii* and *T. fronteriensis*). *Telmatobius vilamensis* differs from *T. marmoratus* by having a smooth dorsum and head (granular in *T. marmoratus*) and in the absence of premaxillary and maxillary teeth (present in *T. marmoratus*). *Telmatobius peruvianus, T. pefauri,* and *T. zapahuirensis* have nuptial spines on the chest, throat, and ventral surfaces of the arms; this character is absent in *T. vilamensis. Telmatobius vilamensis* differs from the northernmost species (*T. peruvianus, T. marmoratus, T. pefauri, T. zapahuirensis*; Fig. 2) by lacking premaxillary and maxillary teeth (present in the northernmost species). *Telmatobius huayra* (Lavilla and Ergueta, 1995*a*), a Bolivian species only 150 km distant from *T. vilamensis,* differs from *T. vilamensis* by having a well developed postocular fold and subcylindrical body; *T. vilamensis* additionally has a postfemoral fold (absent in *T. huayra*).

Description of holotype.—Adult male (SVL = 46.95 mm); head large, depressed, narrower than body; head length 30.5% of SVL; head broader than long (head width/

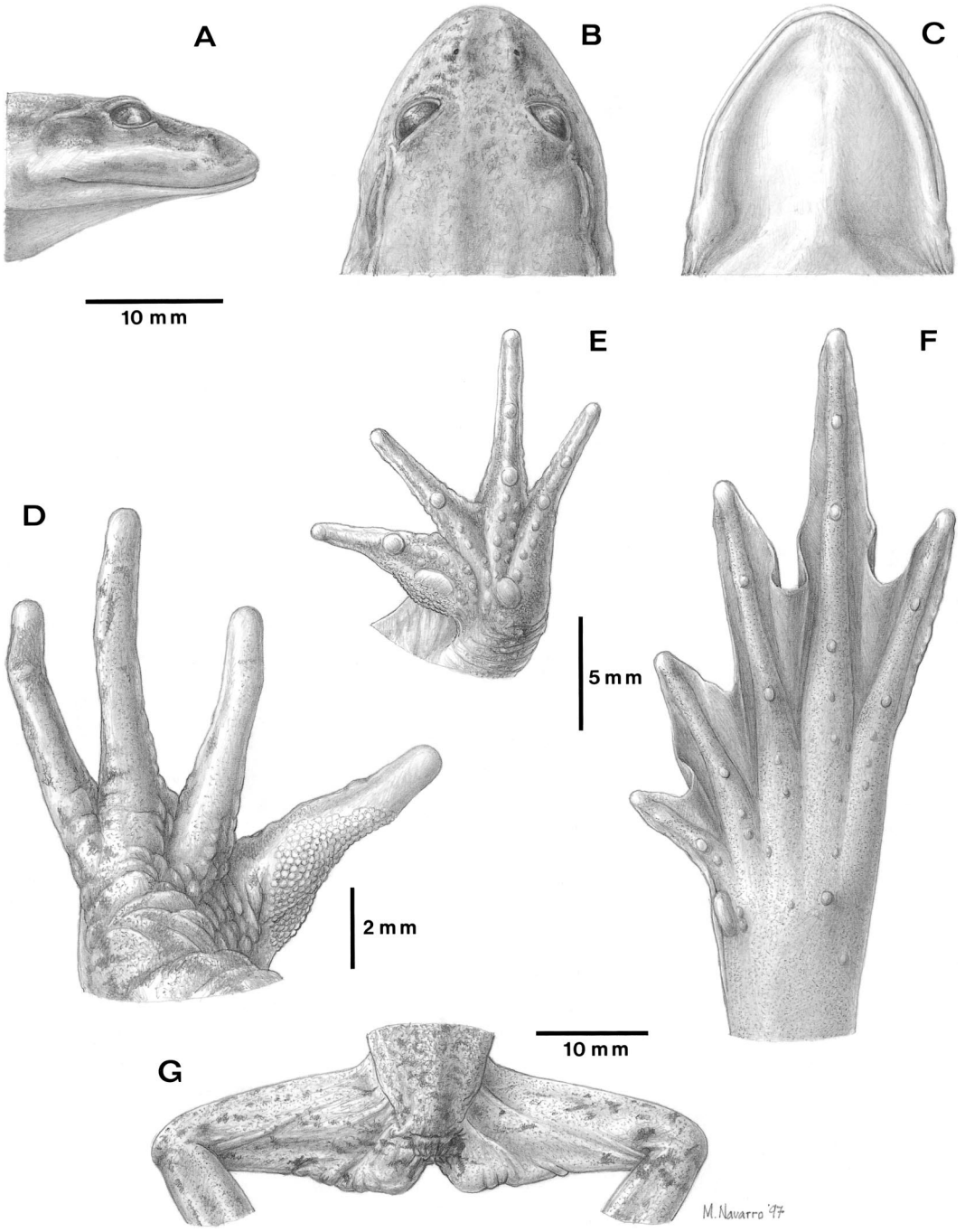


FIG. 2.—Morphological details of the male holotype (IZUA 3080) of *Telmatobius vilamensis*. Lateral (A), dorsal (B), and ventral (C) views of head; nuptial asperities (D); palmar (E) and plantar surfaces (F); and postfemoral folds (G).

SVL = 0.33; head length/head width = 0.97). Snout subovoid in dorsal view (Fig. 2B); margin of upper jaw flared; snout strongly depressed in lateral view (Fig. 2A); loreal region flat; nostrils slightly prominent, oriented anterodorsolaterally; internarial distance 24.3% head width; internarial region flat; nostrils located approximately midway between margin of eye and anterior terminus of snout; canthus rostralis indistinct, straight in dorsal view, elevated in lateral view. Eye moderately large (30.9% head length), oriented anterodorsolaterally; tympanum, tympanic annulus, and columella absent; supratympanic fold moderately developed, extending posteroventrally from posterior corner of eyelid, terminating dorsal to forelimb; lips thin; maxillary and premaxillary teeth absent; dentigerous processes of vomers absent; choanae large (50.2% internarial distance) and subcircular; tongue ovoid, elongate, with posterior border free, unnotched, attached through approximately 80% of its length anteriorly.

Forelimbs thin, with dermal antebranchial fold; dorsal surfaces of wrist and second fingers granular (Fig. 2D); ventral surfaces of arm and forearm with minute white spines; Finger I slightly longer than Finger II; relative lengths of digits on forelimb: III > IV > I > II; palmar webbing absent; tips of fingers rounded; lateral fringes present on Fingers II and III; inner palmar tubercle large, depressed and fusiform; diameter of outer metacarpal tubercle about 66.6% inner tubercle, ovoid, not depressed; subarticular tubercles simple, rounded, scarcely protruding; two subarticular tubercles on digits III and IV, one on digits I and II; numerous supernumerary tubercles present (Fig. 2E); nuptial pad present on inner surface of pollex, consisting of a weakly cornified plate with numerous, small, conical, dark spines (Fig. 2D).

Hind limbs long (approximately 158.2% SVL) and slender (Fig. 2G); toes long, thin; length of toes: IV > V − III > II > I; webbing formula: **I** (1½— 2½) **II** (1½—2) **III** (1—3½) **IV** (3½—1½) **V**; webbing diminishing distally to form wide fringes along lateral margins of toes (Fig. 2F); outer border of Toe V moderately fringed; interdigital plantar webbing with smooth margins; tips of toes rounded, slightly smaller than tips of fingers; inner metatarsal tubercle small, subelliptical and prominent; outer metatarsal tubercle small, rounded, and prominent, about 20% size of inner tubercle; subarticular tubercles rounded, small, raised; subarticular tubercle formula: I (1), II (1), III (2), IV (3), V (2); supernumerary plantar tubercles few, small; tarsal fold distinct, extending approximately 66.6% length of tarsus, continuous distally with fringe along inner margin of Toe I (Fig. 2F); postfemoral fold well developed (Fig. 2G).

Skin smooth dorsally; skin of venter granular with numerous, transparent, minute spines; throat smooth (Fig. 2C); cloacal opening directed posteriorly at dorsal level of thighs; opening round and ornamented with folds and papillae; ventral surface of postfemoral fold with small granules and minute transparent spines; ventral skin of thigh loose, folded. Measurements of holotype and paratypes are presented in Table 1.

Color in preservative.—Dorsal surfaces of body, upper arms, and legs dark gray; scattered, distinct, dark brown spots on dorsum; minute spots on dorsal surfaces of arms and legs; venter and throat cream; ventral surfaces of legs with small, irregular white spots.

Color in life.—Dorsum dark green with dark brown spots; venter and throat white.

Osteology.—Prootic ossification limited to anteromedial part of otic capsule and a pair of small dorsal ossifications associated with the lateral margins of the epiotic eminence. Dorsum of braincase with paired fontanelles (Fig. 3A), an anterior oval frontal fontanelle separated from single, small, round parietal fontanelle by taenia tecti transversalis. Rostral cartilages nearly entirely exposed with a pair of small, slender nasals located laterally and adjacent to margin of tectum nasi and anterior to planum antorbitale on each side; vomers entirely absent; neopalatines present as short, slender bones underlying middle part planum antorbitale on each side of cranium. Minute septomaxillae visible anteriorly.

TABLE 1.—Morphometric data (in mm) for the type series of *Telmatobius vilamensis* and *T. halli*, providing mean ± 1 SD and range.

	Telmatobius halli			*Telmatobius vilamensis*			
					Paratypes		
Variables	Holotype AMNH A-44753 female	Paratype AMNH A-44754 female	Paratype AMNH A-44758 juvenile	Holotype IZUA 3080 male	Male (n = 9)	Female (n = 11)	Juvenile
SVL	57.06	48.04	33.02	46.95	48.00 ± 1.72 (44.73–50.81)	44.50 ± 3.20 (38.36–48.21)	31.19
Head length	16.15	14.27	9.86	15.50	12.96 ± 0.63 (12.05–13.89)	12.27 ± 0.86 (11.05–13.36)	10.5
Head width	18.75	16.58	12.34	15.51	16.29 ± 0.66 (15.50–17.54)	15.08 ± 1.02 (12.96–16.36)	11.31
Interorbital distance	6.04	4.91	3.16	4.79	4.87 ± 0.27 (4.47–5.32)	4.68 ± 0.32 (4.27–5.26)	3.72
Internarial distance	3.65	3.03	2.43	3.24	3.40 ± 0.33 (2.79–3.81)	3.22 ± 0.22 (2.92–3.53)	2.32
Eye–nostril distance	3.65	3.31	1.91	3.20	3.40 ± 0.19 (3.16–3.73)	3.25 ± 0.16 (3.05–3.54)	2.55
Eye diameter	4.17	4.67	3.25	4.65	4.19 ± 0.34 (3.74–4.85)	3.97 ± 0.19 (3.72–4.34)	3.40
Hand length	12.20	11.98	7.78	8.57	10.34 ± 0.97 (8.18–11.41)	10.03 ± 0.78 (8.56–11.22)	7.69
Radioulnar length	12.11	10.14	6.46	10.79	10.33 ± 0.59 (9.32–11.20)	9.12 ± 1.04 (7.19–10.36)	8.58
Tibia length	24.03	20.26	24.03	19.49	21.62 ± 1.48 (19.01–23.54)	20.06 ± 1.34 (17.48–21.77)	14.11
Foot length	40.21	32.27	23.21	33.13	34.31 ± 2.10 (30.46–36.62)	32.51 ± 2.57 (28.71–36.83)	22.02

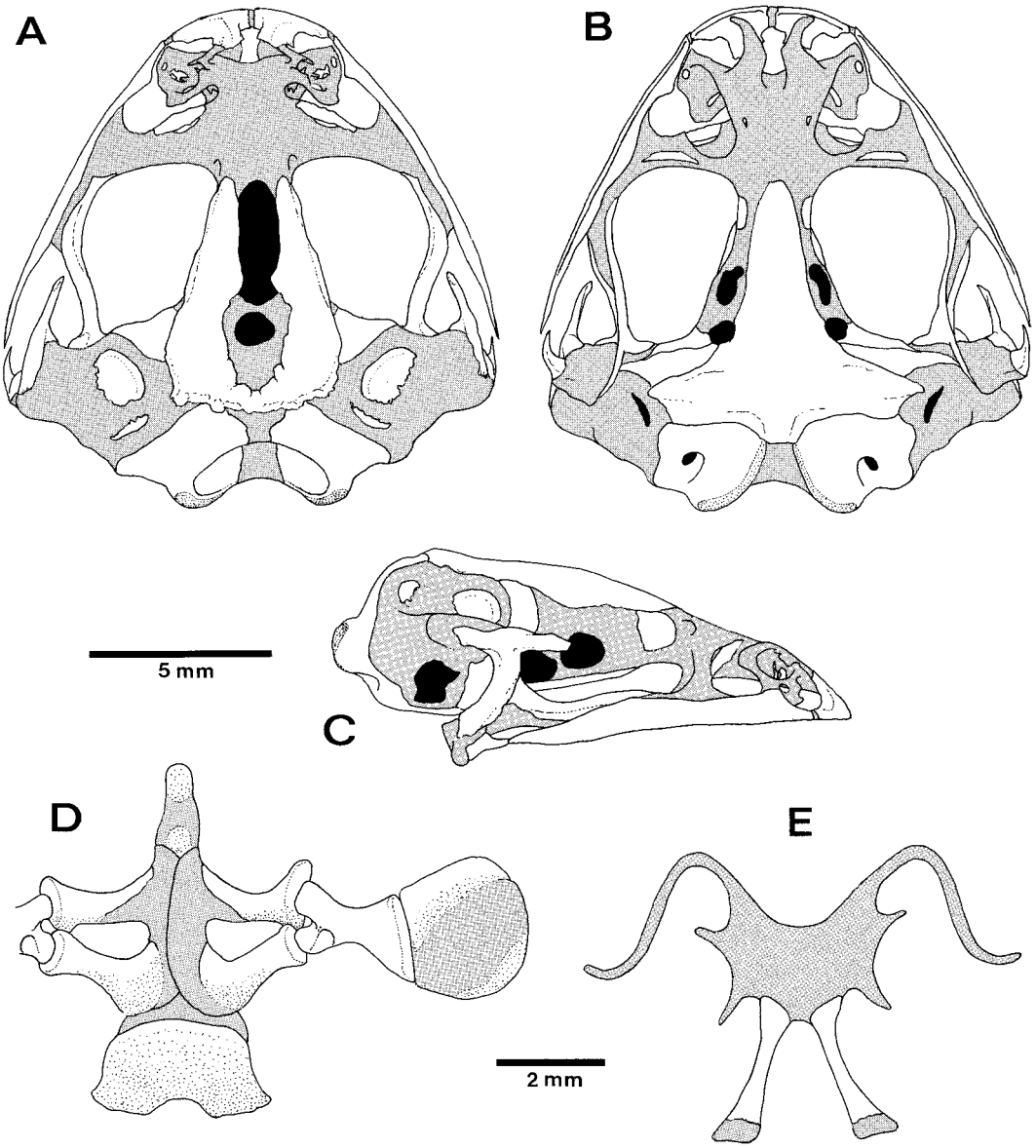


FIG. 3.—Skull of *Telmatobius vilamensis*, IZUA 3224, in dorsal (A), ventral (B) and lateral (C) views. Pectoral girdle (D) and hyoid plate (E) of *T. vilamensis*.

Frontoparietals well developed with distinct laminae perpendicularis; acuminate anteriorly and overlying dorsolateral braincase from anterior part of orbit posteriorly to angle of epiotic eminence; frontoparietals with broad medial separation except at posterior margin where the pair is fused; frontoparietals investing in medial margins of prootics and anteromedial margins of exoccipitals. Parasphenoid robust, nearly triangular; alae short, broad, and associated with prootic anteriorly and exoccipitals posteriorly; posteromedial process truncate; cultriform process underlying most of orbital region of braincase and forming ventral margin of prootic foramen (Fig. 3B). Maxillary arcade composed of robust premaxillae with massive partes dentalis

anteriorly, slender maxillae with low partes facialis, and lacking preorbital processes, quadratojugals short, not articulating with maxillae; maxillae edentate. Pterygoids slender, lacking flanges; anterior ramus terminating medial to pterygoid cartilage in anterior part of orbit; posterior ramus investing medial surface of massive palatoquadrate cartilage; medial ramus short, terminating on anteroventral margin of otic capsule and not in contact with prootic or parasphenoid bones. Squamosals T-shaped in lateral (Fig. 3C) aspect with slender otic and zygomatic rami about equal in length; otic ramus located along dorsolateral margin of cartilaginous crista parotica; ventral ramus distinctly curved (posteriorly concave), investing lateral surface of palatoquadrate cartilage and distinctly separated from quadratojugal. Jaws moderately short with jaw articulation lying anterior to level of fenestra ovalis. Operculum absent. Stapes and tympanic annulus absent.

Hyoid corpus broad, about 50% broader than length at midline; hyoglossal sinus broad, shallow; hyalia simple, uniform in shape, and lacking processes; anterolateral processes of hyoid short, slender, arcuate, directed anterolaterally; posterolateral processes short, slender, arcuate, directed posterolaterally; posteromedial processes ossified, robust, broadly separated from one another anteriorly, terminating in cartilage posteriorly (Fig. 3E).

Pectoral girdle arciferal and robust (Fig. 3D). Clavicles stocky, well ossified, apparently incorporating ossification of procoracoid cartilages laterally; transverse axis of clavicular region straight, anterior margin of bone arcuate. Coracoids short, stocky; sternal end expanded, about twice width of glenoid end; width of glenoid head about 50% greater than midshaft width of coracoid. Pectoral fenestra shallow, more than twice as broad as long, medial half of anterior margin and medial margin cartilaginous, posterior and anterolateral margins formed by coracoid and clavicle, respectively. Epicoracoid cartilages well developed; anterior termini anterior to medial end of clavicles and broadly separating these bones. Omosternum well developed with cartilaginous manubrium and terminal ossification. Sternum short and broad, mostly ossified with cartilaginous margins. Scapula short, robust, only slightly longer than coracoid; distinctly bicapitate, partes acromialis and glenoidalis approximately equal in size; pars acromialis much smaller than lateral head of clavicle. Suprascapula not broadly expanded; anterior third of blade ossified as cleithrum; posterior margin ossified.

Vertebral column composing eight, procoelus, nonimbricate, independent presacral vertebrae (Fig. 4A,B); vertebrae short and broad, overall width of neural arch about three times width of centrum, centrum width approximately equal to centrum length. Presacral I (atlas) moderately widely separated, shallow cervical cotyles; presacrals bearing neural spines; relative lengths of transverse processes and sacrum: II < sacrum < V = VI = VII = VIII < III = IV; transverse process of Presacral II distally expanded, oriented anterolaterally; transverse processes of Presacral III arcuate, oriented laterally; transverse processes of Presacral IV slightly expanded distally, with slight posterolateral orientation; orientation of transverse processes of Presacrals V–VIII approximately uniform with those of V oriented posterolaterally, those of Presacrals VI–VII oriented laterally. Sacral diapophyses round, not dilated, oriented posterolaterally; sacrum with bicondylar articulation with urostyle. Urostyle robust, bearing dorsal crest that is best developed anteriorly, flanked by lateral flanges of bone that diminish in size posteriorly.

Overall length of pelvic girdle approximately equal to length of sacrum plus presacral vertebral column. Ilial shaft robust, bearing dorsolateral crest, with low, laterally oriented dorsal prominence; interilial profile a narrow U-shape, width of U at base about half the distance between anterior ends of ilia; ilium forming anterior margin of round acetabulum; preacetabulum forming approximately a 90° angle to ilial shaft; ilia articulating with one another medially, forming posterior margin of acetabulum; ventral margin of acetabulum

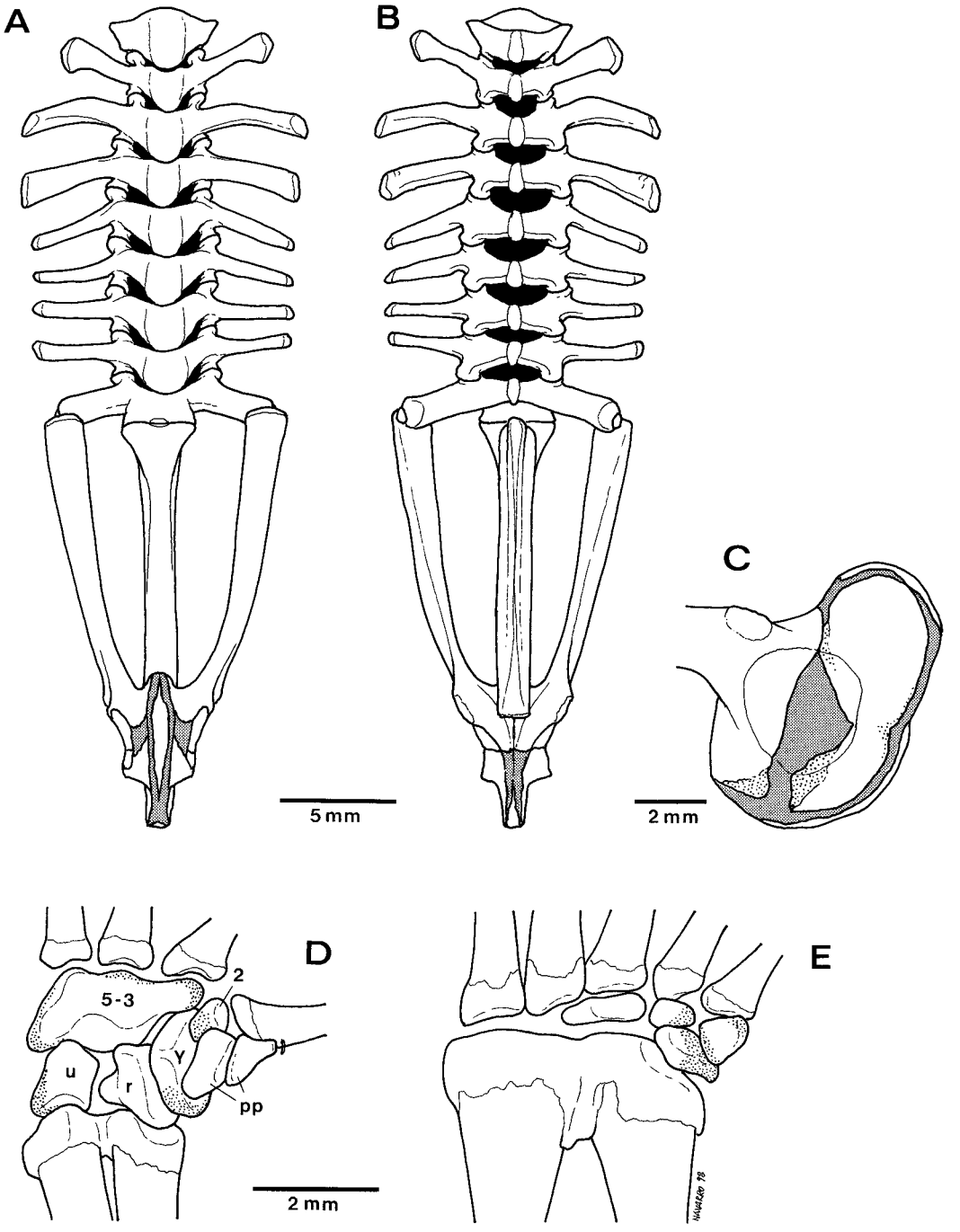


FIG. 4.—Vertebral column of *Telmatobius vilamensis* in dorsal (A) and ventral (B) views. Lateral view of pelvic girdle (C), carpus (D), and tarsus (E) of *T. vilamensis*. Abbreviations: u = ulnare; r = radiale; Y = Element-Y; pp = prepollex; 2 = Distal Carpal 2; 5–3 = Distal Carpal 5–3.

formed by slightly ossified cartilaginous pubis (Fig. 4C).

Humerus longer, more robust than radioulna; humeral crests well developed, especially crista medialis; phalangeal formula of manus: 2-2-3-3; terminal phalanges pointed; seven well ossified carpal elements present: prepollex (two elements), Distal Carpal 2 independent, Distal Carpals 3–5 fused, Element-Y, radiale, ulnare independent (Fig. 4D); femur and tibia similar in length; tibiale and fibulare fused proximally and distally, half length of femur; phalangeal formula of the foot: 2-2-3-4-3; prehallux with two elements; two additional metatarsal bones of unresolved homology (Fig. 4E).

Description of tadpoles.—Larvae (Stages 35–40) large (total length 77.71–78.15 mm; body length 31.79–34.06 mm) and robust; tail relatively short (1.5 × body length) and thick; body ovoid in lateral view (Fig. 5A); lateral profile of tip of snout gently rounded; nostrils ovoid, not protuberant; apertures situated anterodorsolaterally, with cutaneous fringe (Fig. 5B); internarial distance 150% interocular distance; nostrils closer to anterior border of eye than tip of snout; eyes circular, located in shallow depression; situated anterodorsolaterally; in life, iris with melanophores and golden tints; width of oral disc 120% interocular distance; oral disc anteroventral, transangular; rostral gap present and mental gap absent (Fig. 5C); single row of marginal papillae present along periphery of disc, marginal papillae absent in rostral region; one row of intramarginal papillae (n = 13–16) in mental area; papillae present in infra-angular and supra-angular regions 14–17, respectively; rostrodonts wider than tall well keratinized; suprarostrodonts and infrarostrodonts with serrations and dark brown pigmentation; keratodonts formula [(1) (1-1)/(1-1) (2)]; spiracular tube sinistral, short, situated laterally; aperture oval, diameter 110% eye diameter; proctodeal tube triangular, opaque (in 10% formalin), wide; vent opening dextral, not visible in ventral position; distal end ovoid (Fig. 5D); dorsal fin not extending onto body; ventral fin begins at end of proctodeal tube; tail tip rounded; maximum width of dorsal and ventral fins slightly posterior of midlength of tail; fin depth exceeding body depth; myomeres and fins with irregular, dark brown spots; body dark brown (formalin 10%); internal organs not visible; hind limbs with minute melanophores.

The tadpole of *T. vilamensis* is generalized (Orton, 1953), having characters associated with anuran larvae inhabiting lotic and benthic habitats (e.g., depressed body, anteroventral oral disc, robust caudal musculature, dorsolateral eyes, and low fins). Measurements of tadpoles are given in Table 2.

Chromosomes.—Examination of 19 metaphases from two males revealed a diploid number of 2N = 26. All chromosomes are bi-armed and the fundamental number (NF) is 52. Pairs 1, 4, 8–13 are metacentric; Pairs 2, 6, and 7 are submetacentric; and Pairs 3 and 5 are subtelocentric (Fig. 6A). Pair 6 (sm) has a secondary constriction in the smaller arm. Pairs 1–4 are large (>100 units); Pairs 5 and 6 are intermediate (between 80 and 100 units), and Pairs 7–13 are small (<80 units). A summary of the relative lengths, arm ratios, and types of chromosomes is presented in Table 3.

The C-band karyotype, based on five plates, shows constitutive heterochromatin in the pericentromeric region of all chromosomes, except in Pair 4, which shows extensive positive staining in the telomeric ends of the long arms (Fig. 6B). Thin heterochromatic bands can be discerned at some telomeres, especially on the short arms of Pairs 1, 2, and 7. A thin interstitial band was observed on the long arm of Pair 5. The Ag-NOR were revealed in the smaller arm of Pair 6 (Fig. 6C).

Distribution and ecology.—This species is only known from the type locality, the intersection of the Río Vilama and the road between San Pedro de Atacama and El Tatio (6.5 km by road from San Pedro de Atacama; Fig. 1). The type locality is a semidesert area with scarce vegetation (*Ephedra andina* and *Atriplex atacamensis*), located in the tropical marginal region (Di Castri, 1968). In San Pedro de Atacama, the annual mean temperature is

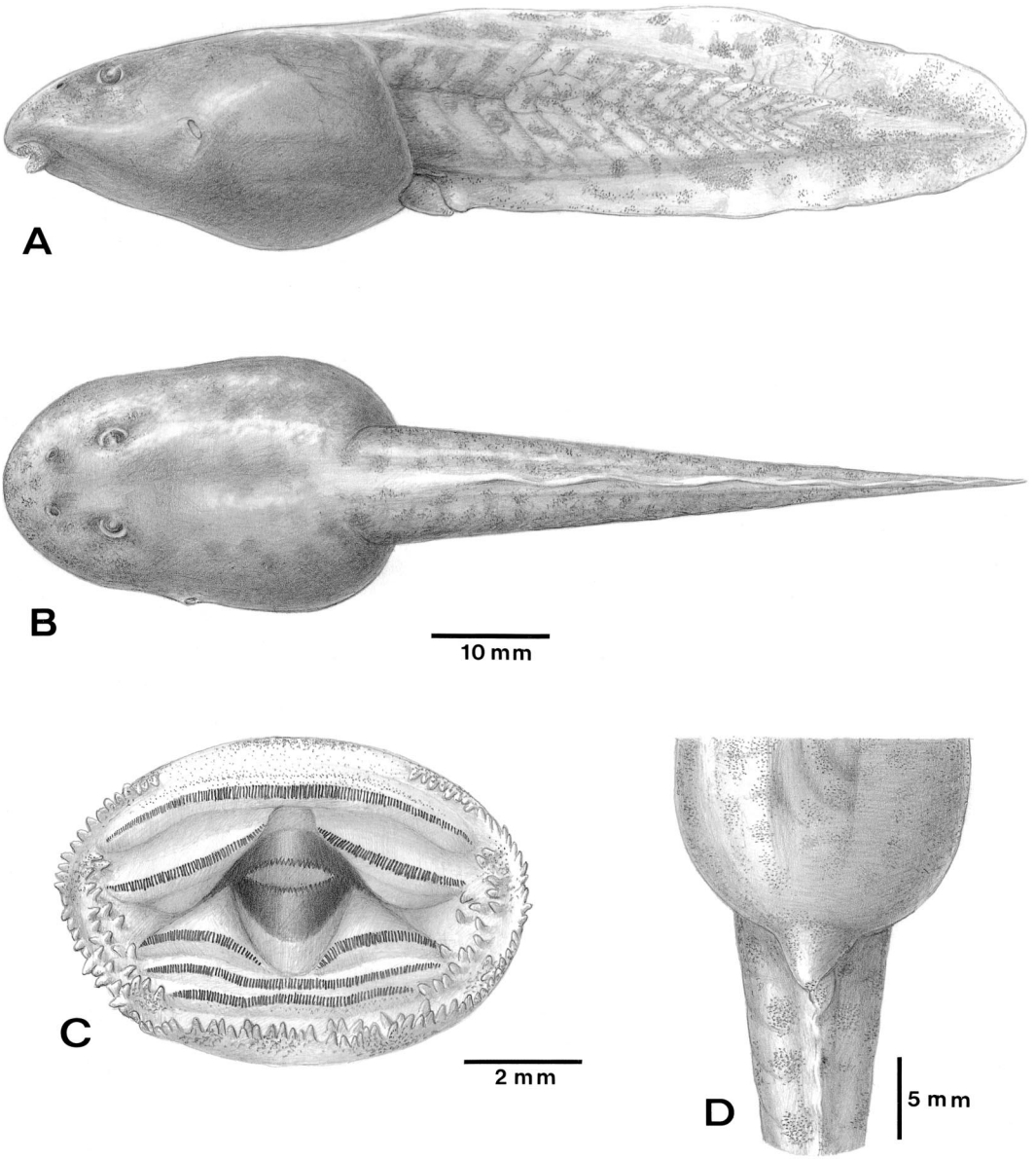


FIG. 5.—Tadpole of *Telmatobius vilamensis* (Stage 35). Lateral (A), dorsal (B), and ventral (D) views; oral disc (C).

13.3 C and rainfall is very scarce (Hajek and Di Castri, 1975). Adults and tadpoles (Stages 34–40) were collected with a net below aquatic plants on the banks of the Río Vilama (5–7 m wide, 10–12 C) during the austral autumn (11 April 1997). Male specimens had nuptial asperities, and one tadpole was near the end of metamorphosis (Stage 46, SVL = 32.0 mm). *Bufo spinulosus* and lizards of the genus *Liolaemus* were also present.

Etymology.—The name *vilamensis* is a Latin adjective and refers to the type locality, Río Vilama.

Telmatobius halli Noble

Telmatobius halli Noble, 1938:1 [Holotype: AMNH A-44753 from a warm spring

TABLE 2.—Measurements (mm) of the tadpoles of *Telmatobius vilamensis*, providing mean ± 1 SD and range.

Character	Stage 35 (n = 3)	Stage 40 (n = 1)
Total length	80.69 ± 2.98 (77.71–83.67)	78.15
Body length	32.49 ± 0.70 (31.79–33.19)	34.06
Body depth	16.76 ± 0.99 (15.77–17.75)	16.58
Fin depth	8.89 ± 0.90 (7.50–9.89)	9.73
Snout–nostril distance	6.64 ± 0.53 (6.11–7.17)	6.29
Eye diameter	2.80 ± 0.16 (2.64–2.96)	2.64
Mouth width	7.31 ± 0.26 (7.05–7.57)	7.82

near Ollagüe, 10,000 ft. altitude, Provincia El Loa, Chile].

Diagnosis.—A large *Telmatobius* (46.95–57.06 mm) having the following combination of characters: (1) premaxillary teeth absent, maxillary teeth rudimentary; (4) tongue round, thick; with posterior border free, unnotched; slightly long than wider; attached through more than 75% of its length; (5) choanae large, subrectangular; (6) tympanum and tympanic annulus absent; (7) snout truncate in dorsal view; (8) postfemoral folds absent; (9) dorsum yellow cream (in preservative); (12) toes extensively webbed; (13) outer border of Toe V widely fringed.

Redescription of holotype.—Adult female (SVL = 57.06 mm); head narrower than body; head length 28.6% of SVL; head wider than long (head width/SVL = 0.33; head length/head width = 0.86). Snout truncate in dorsal view (Fig. 7B); snout moderately short in lateral view (Fig. 7A); loreal region nearly horizontal; snout slightly truncate in lateral view; nostrils not protuberant, oriented anterodorsolaterally; internarial distance 19.6% head width; internarial region convex; nostrils closer to tip of snout than eye; canthus rostralis indistinct in dorsal profile. Eye moderately large (25.5% head length), oriented anterodorsolaterally; tympanum and tympanic annulus absent; supratympanic fold barely visible; lips thin; rudimentary maxillary teeth present, completely embedded within labial mucosa; dentigerous process of vomers absent; choanae large (60.3% internarial distance) and subrectangular; tongue thick, round, with posterior border free, unnotched, slightly longer than wide, attached through more than 75% of its length anteriorly.

Forelimbs thin, with dermal forearm fold; dermal wrist fold absent; Finger I longer than Finger II; relative lengths of digits of forelimb: III > IV > I > II; palmar webbing absent (Fig. 7C); tips of fingers round; rudimentary lateral fringes between Fingers I and II, and II and III; inner palmar tubercle oval, distinct; diameter of outer metacarpal tubercle about 66.6% inner tubercle, pear-shaped, slightly raised, distinct; one round distinct subarticular tubercle distally on Finger IV; indistinct distal subarticular tubercle on Fingers II and III; distinct subarticular tubercles proximally on each finger; one distinct supernumerary tubercle present at base of Fingers II and III. Three or more indistinct supernumerary tubercles present at base of each finger; supernumerary palmar tubercles also present along outer edge of inner metacarpal tubercle (Fig. 7C).

Hind limb, long (approximately 150.4% SVL) and robust; length of toes: IV > III > V > II > I (Fig. 7D); webbing formula **I** (1⅓—2) **II** (1⅔—3) **III** (2—3½) **IV** (3⅔—1½) **V**; webbing diminishes distally to form wide fringes along lateral margins of toes; outer border of Toe V widely fringed from base to tip; tips of toes spherical, about equal in size to finger tips; inner metatarsal tubercle elongate, raised; outer metatarsal tubercle round, small, raised, about half size of inner tubercle; subarticular tubercles round, some indistinct, distributed on toes as follows: I (1), II (1), III (6), IV (8), V (4); single, proximal tubercle distinct on Toes II and V; two proximal indistinct supernumerary tubercles on Toes IV and III; indistinct supernumerary tubercles around outer metatarsal tubercle (Fig. 7D); tarsal fold distinct, extending approximately 66.6% length of tarsus, continuous distally with fringe along inner margin of Toe I; postfemoral fold absent.

Skin smooth dorsally and ventrally, with-

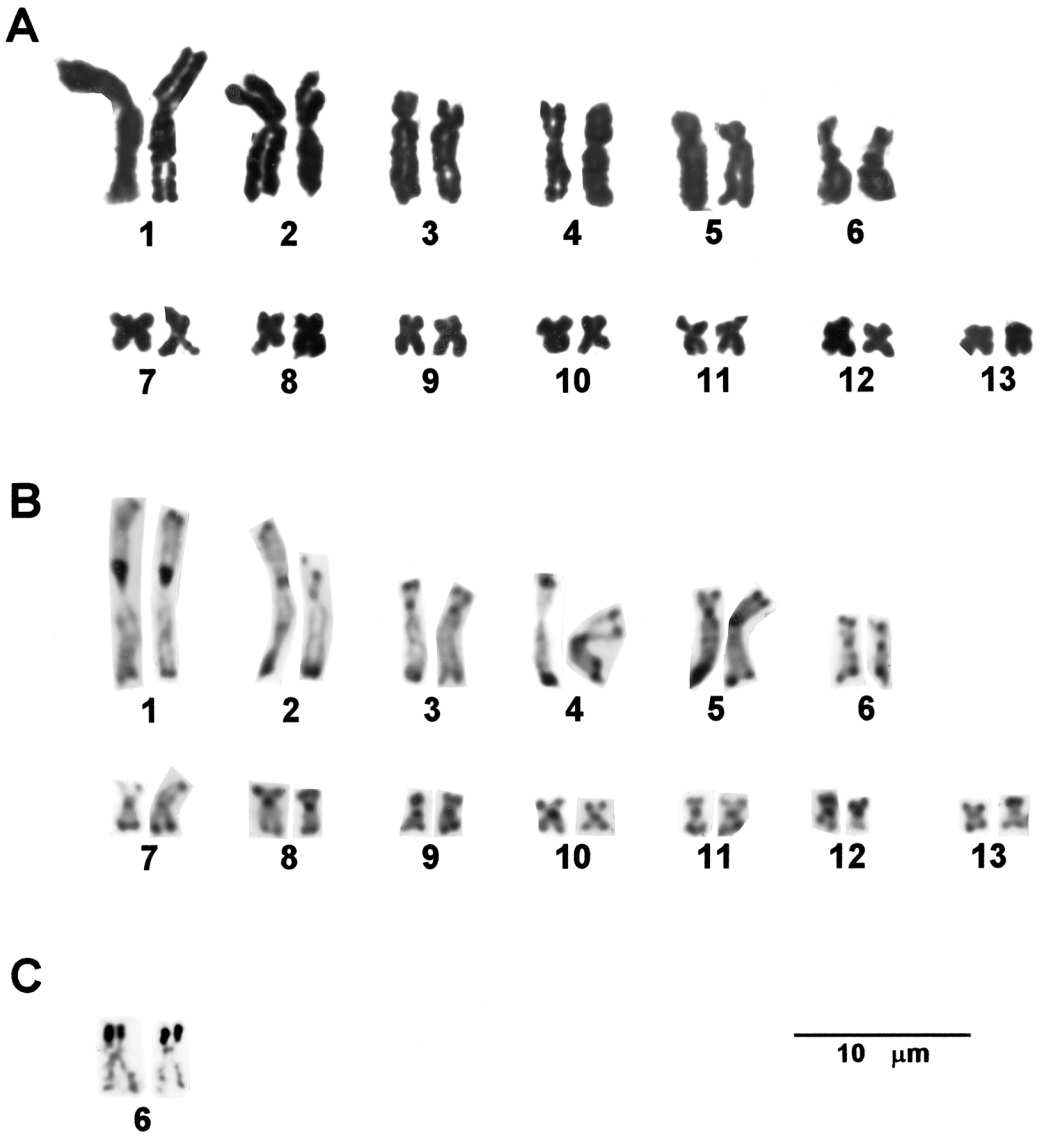


FIG. 6.—Chromosomes of *Telmatobius vilamensis*. Standard (A) and C-banded (B) karyotypes. Ag-NOR (C).

out spines or warts; cloacal opening at upper level of thigh; opening round ornamented with folds; transverse fold of skin present dorsal to the cloacal opening; ventral skin of thigh loose, folded, pustular. Measurements of the holotype and paratypes are presented in Table 1.

Color in preservative.—Dorsum of body uniform brown to tan; paler minute speckling on flanks; ventral surfaces of forearm with white speckles; venter of body uniform pale cream; tips of fingers and toes yellow; hand and foot tubercles light cream. Color in life unknown.

Distribution and ecology.—Telmatobius halli is only known from the type locality (near Ollagüe; 21° 13′ S, 68° 16′ W; 3700 m; El Loa Province, Región de Antofagasta, Chile). The type series (five adult females and six tadpoles) was collected from a warm spring in June (Noble, 1938). During two consecutive periods (March and

TABLE 3.—Relative length, arm ratio (mean and standard deviations), and type of chromosomes of *Telmatobius vilamensis*.

Pair	Relative length*	r†	Type‡
1	178.07 ± 14.52	1.22 ± 0.13	m
2	144.38 ± 10.67	2.17 ± 0.84	sm
3	124.55 ± 8.43	4.08 ± 0.52	st
4	115.70 ± 4.05	1.46 ± 0.24	m
5	97.18 ± 4.94	3.58 ± 0.61	st
6§	60.22 ± 4.68	2.12 ± 0.47	sm
7	51.85 ± 1.34	1.72 ± 0.35	sm
8	47.57 ± 1.77	1.61 ± 0.34	m
9	43.50 ± 3.19	1.64 ± 0.28	m
10	41.42 ± 2.02	1.15 ± 0.08	m
11	38.50 ± 2.59	1.32 ± 0.06	m
12	33.92 ± 2.62	1.43 ± 0.32	m
13	27.50 ± 2.47	1.38 ± 0.18	m

* Calculated according to Bogart (1970).
† Ratio of short arm divided into long arm.
‡ Chromosomic types (m = metacentric, r = 1.0–1.7; sm = submetacentric, r = 1.71–3.0; st = subtelocentric, r = 3.01–4.5) according to Levan et al. (1964).
§ Pair with secondary constriction.

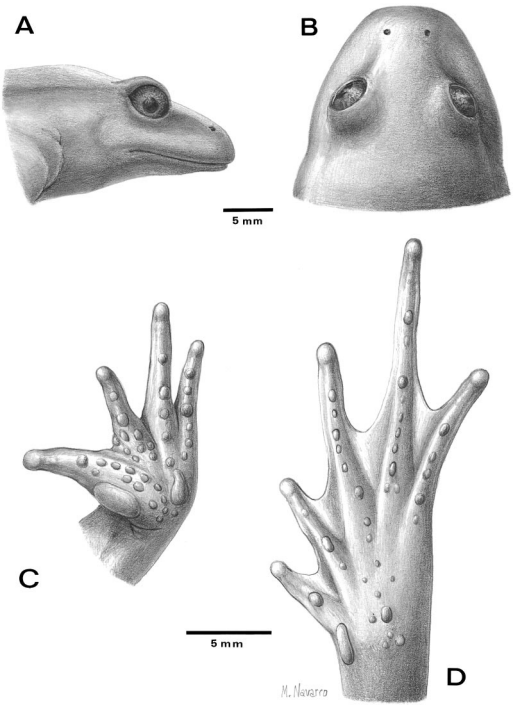


FIG. 7.—Morphological details of the female holotype (AMNH A-44753) of *Telmatobius halli*. Lateral (A) and dorsal (B) views of the head. Palmar (C) and plantar (D) surfaces.

December 1998), we searched for additional specimens at the type locality. However, no *T. halli* were collected. Ollagüe is situated in the Andean Region (Di Castri, 1968); the annual mean temperature of this region is 6.8 C and the rainfall ranges from 0.0 mm (October) to 31.8 mm (January).

DISCUSSION

Species of *Telmatobius* are characterized by a significant amount of inter- and intraspecific morphological variation (Cei, 1986; Trueb, 1979; Wiens, 1993) that precludes taxonomic arrangement based on unique character states in each species.

Discordant patterns of variation of the few informative external and osteological characters have resulted in the progressive rejection of each synapomorphy proposed for the genus. Thus, the monophyly of *Telmatobius* is suspect (see the extended discussion in Formas et al., 1999). As suggested by Wiens (1993), generic allocation is based on two character states (frontoparietals fused posteriorly, and nuptial excrescences on Finger I only; both present in *T. vilamensis*) that, although not universal, are present in most *Telmatobius* as well as in the type for the genus (*T. peruvianus* Wiegmann, 1835).

At this point, the generic assignment of *T. vilamensis* would seem questionable; nevertheless, it should be noted that zoogeographical evidence, as well as allozyme and morphometric data, support the affinity of *T. vilamensis* to other species currently referred to *Telmatobius* species of the Bolivian Altiplano and the Antofagasta Region (Benavides et al., 2002*b*).

Telmatobius vilamensis, like other *Telmatobius*, has 26 bi-armed chromosomes (Kuramoto, 1990). This 26-chromosome karyotype (NF = 52) was considered to be ancestral for telmatobiine frogs by Formas and Cuevas (2000) and Reig (1972), but Brum-Zorrilla and Sáez (1968) reported 22 chromosomes in an unidentified specimen of the *Telmatobius marmoratus* group from La Paz (Bolivia). This exceptional karyotype should be verified, because all other *Telmatobius* species have a symplesiomorphic arrangement (2N = 26; see Barbieri, 1954; Cordova et al., 1987; Cuevas and Formas, 2002; Díaz and Veloso,

1979; Formas et al., 1999; Morescalchi, 1973; Veloso and Trueb, 1976). It is also possible that the 2N = 22 karyotype is a synapomorphy for a particular clade of *Telmatobius.*

On a comparative basis, the standard karyotype of *Telmatobius* species is uniform, but the C-band karyotypes reveal some specific differences. For instance, *T. vilamensis* has an interstitial heterochromatic band in Pair 5, whereas, in *T. arequipensis,* the short arm of Pair 13 is totally heterochromatic (Cordova et al., 1987). *Telmatobius dankoi* exhibits pronounced pericentromeric bands in Pairs 1 and 9 (Formas et al., 1999), and *T. philippii* lacks pericentromeric heterochromatin in Pair 1 (Cuevas and Formas, 2002). In all cases analyzed, the secondary (nucleolar) constriction is always located in the short arm of Pair 6. The number and distribution of heterochromatic bands may offer additional phylogenetically informative characters for future studies of the group.

Finally, it is worth noting that *T. vilamensis, T. dankoi, T. halli, T. philippii,* and *T. fronteriensis* are known only from their respective type localities. In the case of *T. halli,* no additional specimens have been reported since the species was first collected in 1935 (Noble, 1938). The restricted distributions of these species can be explained by the aridity of the western slopes of the Andes, which effectively separates bodies of water as reduced independent drainages in an area of less than 250 km^2 total. Geographic isolation of small populations could have produced the surprising species diversity and unusual degree of differentiation shown by this group of *Telmatobius.* However, the evolutionary consequence of small populations is loss of heterozygosity (Hartl and Clark, 1997). Benavides et al. (2002*b*) screened levels of allozyme diversity in *T. vilamensis* and *T. fronteriensis* and found values of H_0 = 0.0058 and H_0 = 0.00, respectively. Recently, Montgomery et al. (2000) experimentally confirmed a positive correlation between allozyme diversity and population size; thus, it is plausible that these species are endangered.

Key to Adult Telmatobius of Chile

1. Snout rounded or pointed in dorsal view 2
   Snout truncated in dorsal view; tongue rounded, unnotched; attached anteriorly along 75% of its length; premaxillary teeth absent, maxillary teeth reduced; dorsum uniform brown to tan (in preservative); belly pale cream; tips of toes and fingers yellow . . *T. halli*
2. Tibio tarsal join does not reach the posterior border of the eye 3
   Tibio tarsal join reaches or exceeds the posterior border of eye 4
3. Skin on head strongly granular; nuptial spines extensive on dorsolateral surface of thumb; irregularly distributed spines on chest, inner metacarsal tubercle, lips, and ventral surface of arm; tongue attached anteriorly about 66% of its length; vomerine teeth reduced; skin on dorsum, belly, flanks and dorsal surfaces of legs strongly granular; dorsum gray (in preservative) with irregular dark spots (marmorations) . . . . . *T. marmoratus*
   Skin on head without granules; nuptial spines reduced, not reaching inner metacarpal tubercle; tongue attached anteriorly for about 80% of its length; skin on flanks strongly granular; dorsum dark gray (in preservative) with dark brown irregular spots; belly gilt copper (in life) . . . . . *T. philippii*
4. Xiphisternum notched 5
   Xiphisternum unnotched; tympanic ring absent; nuptial spines extensive on chest, throat, and dorsolateral and ventral surfaces of thumb, reaching the inner metacarpal tubercle; tongue attached 66% of its length; vomerine teeth absent; dorsum brown to gray (in preservative) . . . . . *T. peruvianus*
5. Flanks without queratinous spines 6
   Flanks with minute, brown, queratinous spines (males and females); nuptial pad extending along dorsolateral surface of thumb; minute dark spines irregularly distributed on throat, and dorsal and ventral surfaces of hand; postfemoral folds present; cranium well ossified; dorsum gray to brown (in preservative) . . . . . *T. dankoi*
6. Snout not depressed 7
   Snout strongly depressed (wedge-shaped); without vomerine, premaxillary, and maxillary teeth; nuptial

spines small, restricted on dorsal and lateral surfaces of thumb; postfemoral folds present (in preservative and life); cranium poorly ossified; dorsum gray (in preservative), olive-green (in life) . . . . . *T. vilamensis*

7. Subarticular tubercles present on all toes 8
   Subarticular tubercles present on Toes I, II, and III; nuptial spines extensive on chest and dorsolateral surfaces of thumb; tongue attached anteriorly 85% of its length; dorsum light brown (in preservative) . . . . . *T. zapahuirensis*
8. Adult size >54 mm; nuptial spines extensive on chest, throat, anterior border of arm, and dorsoventral surfaces of thumb, reaching the inner metacarpal tubercle; tongue attached 66% of its length; dorsum light brown (in preservative) with minute gray irregularly distributed spots . . . . . *T. pefauri*
   Adult size <45 mm; nuptial spines large, black, extensive on dorsolateral surface of thumb; nuptial pad continuous with inner metacarpal tubercle; throat and ventral surface of thighs and arms with leaden (in preservative) irregular spots . . . . . *T. fronteriensis*

Resumen

Describimos una nueva especie; *Telmatobius vilamensis* del Río Vilama, cerca de San Pedro de Atacama, Chile. Previamente confundida con *T. halli* Noble, 1938, esta nueva especie es descrita en base a caracteres de la morfología externa de las larvas y adultos, osteología y cromosomas. La nueva especie es el representante más sureño del género *Telmatobius* en Chile. Adicionalmente, se redescribe *T. halli* y se comparan ambos taxa con las restantes especies del género reportadas para Chile.

Acknowledgments.—We thank L. Trueb, J. Sites, and three anonymous referees for suggestions that greatly improved the manuscript; C. Cole (AMNH), J. Aparicio (CBF), and J. C. Ortiz (MZUC) for loan of specimens; and N. Bernal, C. Mehling, and R. Bain (AMNH) for pictures of the *T. halli* holotype. We also thank C. Jara for field assistance, M. Navarro for art work, and L. Brieva for drafting versions of the manuscript. This study was supported by Fondo Nacional de Ciencia y Tecnología (Proyecto S960021).

Literature Cited

Barbieri, F. D. 1954. Observaciones preliminares sobre los cromosomas de "*Telmatobius schreiteri.*" Scientia Genetica 4:223–226.

Benavides, E., J. C. Ortiz, and J. R. Formas. 2002*a*. A new species of *Telmatobius* (Anura: Leptodactylidae) from northern Chile. Herpetologica 58:210–220.

Benavides, E., J. C. Ortiz, and J. W. Sites, Jr. 2002*b*. Species boundaries among the *Telmatobius* (Anura: Leptodactylidae) of the Lake Titicaca basin: allozyme and morphological evidence. Herpetologica 58:31–55.

Bogart, J. P. 1970. Systematic problems in the amphibian family Leptodactylidae (Anura) as indicated by karyotypic analysis. Cytogenetics 9:369–383.

Brum-Zorrilla, N., and F. A. Sáez. 1968. Chromosomes of Leptodactylidae (Amphibia: Anura). Experientia 24:969.

Cei, J. M. 1962. Batracios de Chile. Ediciones de la Universidad de Chile, Santiago, Chile.

———. 1986. Speciation and adaptive radiation in Andean *Telmatobius* frogs. Pp. 374–386. *In* F. Vuilleumier and M. Monasterio (Eds.), High Altitude Tropical Biology. Oxford University Press, Oxford, U.K.

Cordova, J. H., J. Descailleaux, and W. Manya. 1987. Descripción del cariotipo de *Telmatobius arequipensis* (Anura: Leptodactylidae) y relaciones citogenéticas con otras especies del género. Revista Latinoamericana de Genética 1:44–53.

Cuevas, C. C., and J. R. Formas. 2002. *Telmatobius philippii,* una nueva especie de rana acuática de Ollagüe, norte de Chile (Leptodactylidae). Revista Chilena de Historia Natural 75:245–258.

de la Riva, I. 1994*a*. Description of a new *Telmatobius* from Bolivia (Amphibia: Anura: Leptodactylidae). Graellsia 50:161–164.

———. 1994*b*. A new aquatic frog of the genus *Telmatobius* (Anura: Leptodactylidae) from Bolivian cloud forests. Herpetologica 50:38–45.

Díaz, N. F., and A. Veloso. 1979. Sistemática y evolución de los anfibios de Chile. Archivos de Biología y Medicina Experimentales 12:59–70.

Di Castri, F. 1968. Esquisse écologique du Chili. Pp. 7–52. *In* C. D. Devoutteville and E. Rapoport (Eds.), Biologie de l'Amerique Australe. Études sur la faune du sol. Editions du Centre National de la Recherche Scientifique, Paris, France.

Fabrezi, M., and P. Alberch. 1996. The carpal elements of anurans. Herpetologica 52:188–204.

Formas, J. R. 1991. The karyotypes of the Chilean frogs *Eupsophus emiliopugini* and *E. vertebralis* (Amphibia: Anura: Leptodactylidae). Proceedings of the Biological Society of Washington 104:7–11.

Formas, J. R., and C. C. Cuevas. 2000. Comparative cytogenetic analysis of the Chilean leptodactylid frog genus *Telmatobufo,* with the description of the chromosomes of *T. venustus.* Proceedings of the Biological Society of Washington 113:890–899.

Formas, J. R., I. Northland, J. Capetillo, J. J. Nuñez, C. C. Cuevas, and L. M. Brieva. 1999. *Telmatobius dankoi,* una nueva especie de rana

acuática del norte de Chile (Leptodactylidae). Revista Chilena de Historia Natural 72:427–445.

GOSNER, K. 1960. A simplified table for staging anuran embryos and larvae with notes on identification. Herpetologica 16:183–190.

HAJEK, E. R., AND F. DI CASTRI. 1975. Bioclimatografía de Chile. Dirección de Investigación Vice-Rectoría Académica Universidad Católica de Chile, Santiago, Chile.

HARTL, D. L., AND A. G. CLARK. 1997. Principles of Population Genetics. Sinauer, Sunderland, Massachusetts, U.S.A.

HOLLISTER, G. 1934. Clearing and dyeing fish for bone study. Zoologica 12:89–101.

KURAMOTO, M. 1990. A list of chromosome numbers of anuran amphibians. Bulletin of Fukuoka University of Education 39:83–127.

LAVILLA, E. O., AND P. ERGUETA. 1995*a*. Una nueva especie de *Telmatobius* (Anura: Leptodactylidae) del Sudoeste de Bolivia. Ecología en Bolivia 24:91–101.

———. 1995*b*. Una nueva especie de *Telmatobius* (Anura, Leptodactylidae) de la ceja de montaña de La Paz (Bolivia). Alytes 13:45–51.

———. 1999. A new Bolivian species of the genus *Telmatobius* (Anura: Leptodactylidae) with a humeral spine. Amphibia-Reptilia 20:55–64.

LEVAN, A., K. FREDGA, AND A. SANDBERG. 1964. Nomenclature for centromeric position on chromosome. Hereditas 52:201–220.

MONTGOMERY, M. E., L. M. WOODWORTH, R. K. NURTHEN, D. M. GILLIGAN, D. A. BRISCOE, AND R. FRANKHAM. 2000. Relationships between population size and loss of genetic diversity: comparisons of experimental results with theoretical predictions. Conservation Genetics 1:33–43.

MORESCALCHI, A. 1973. Amphibia. Pp. 223–348. *In* A. B. Chiarelli and E. Capanna (Eds.), Cytotaxonomy and Vertebrate Evolution. Academic Press, New York, New York, U.S.A.

MYERS, C. W., AND W. E. DUELLMAN. 1982. A new species of *Hyla* from Cerro Colorado, and other tree frog records and geographical notes from western Panama. American Museum Novitates 2752:1–25.

NOBLE, G. K. 1938. A new species of the frog of the genus *Telmatobius* from Chile. American Museum Novitates 73:1–3.

NORTHLAND, I., J. CAPETILLO, P. ITURRA, AND A. VELOSO. 1990. Nuclear DNA content and karyosystematic relationships of species grouped in primitive tribes of Leptodactylidae (Amphibia-Anura). Revista Brasileira de Genetica 13:247–254.

ORTON, G. L. 1953. The systematics of vertebrate larvae. Systematic Zoology 2:63–75.

REIG, O. 1972. *Macrogenioglottus* and the South American bufonoid toads. Pp. 14–36. *In* W. F. Blair (Ed.), Evolution in the Genus *Bufo*. University of Texas Press, Austin, Texas, U.S.A.

RUFAS, J. S., P. ITURRA, W. DE SOUZA, AND P. SPONDA. 1982. Simple silver staining procedures for the location of nucleolus and nucleolar organizer under light and electron microscopy. Archivos de Biología 93:267–274.

SALAS, A., AND U. SINSCH. 1996. Two new *Telmatobius* species (Leptodactylidae, Telmatobiinae) of Ancash, Peru. Alytes 14:1–26.

SAVAGE, J. M., AND W. R. HEYER. 1967. Variation and distribution in the tree frog genus *Phyllomedusa* in Costa Rica, Central America. Beiträge zur Neotropischen Fauna 5:111–131.

SONG, J., AND L. R. PARENTI. 1995. Clearing and staining whole fish specimens for simultaneous demonstration of bone, cartilage, and nerves. Copeia 1995:114–118.

SUMNER, A. T. 1972. A simple technique for demonstrating centromeric heterochromatin. Experimental Cell Research 75:304–306.

TRUEB, L. 1979. Leptodactylid frogs of the genus *Telmatobius* in Ecuador with the description of a new species. Copeia 1979:714–733.

VELOSO, A., AND L. TRUEB. 1976. Description of a new species of Telmatobiine frog, *Telmatobius* (Amphibia: Leptodactylidae), from the Andes on northern Chile. Occasional Papers of the Museum of Natural History, University of Kansas 62:1–10.

VELOSO, A., M. SALABERRY, J. NAVARRO, P. ITURRA, J. VALENCIA, M. PENNA, AND N. DÍAZ. 1982. Contribución sistemática al conocimiento de la herpetofauna del extremo norte de Chile. Pp. 135–268. *In* A. Veloso and E. Bustos (Eds.), La Vegetación y los Vertebrados Ectotérmicos del Transecto Arica-Lago Chungará. Volumen de síntesis, Proyecto MAB 6-UNEP-UNESCO, Santiago de Chile, Chile.

WIENS, J. 1993. Systematics of the leptodactyloid frog genus *Telmatobius* in the Andes of northern Peru. Occasional Papers of the Museum of Natural History, University of Kansas 162:1–76.

Accepted: 5 June 2002
Associate Editor: Joseph Mendelson III

APPENDIX I

Specimens Examined

Telmatobius dankoi.—CHILE: Provincia El Loa: Las Cascadas, 2260 m, MNHN 3006, IZUA 2108–10, 2112 (two cleared-and-stained adults), 2113 (nine specimens), 2106 (two tadpoles), 2107 (five tadpoles), DBMUA 45, 46, 53, 57.

Telmatobius halli.—CHILE: Provincia El Loa: Ollagüe, 3050 m, AMNH A-44753–54, A-44758.

Telmatobius marmoratus.—CHILE: Provincia Parinacota: Lago Chungará, 4270 m, MZUC 24520, 24524, 24526. BOLIVIA: Provincia Murillo: Valle de La Paz, CBF 03244–56, CBF 1009–11, CBF 468–70.

Telmatobius pefauri.—CHILE: Provincia Parinacota: Quebrada de Zapahuira, 3270 m, DBCUCH 629, 632–634, 641, 652.

Telmatobius peruvianus.—CHILE: Provincia Parinacota: Putre, 3200 m, MZUC 24642–43.

Telmatobius vilamensis.—CHILE: Provincia El Loa: Rio Vilama, 3110 m, IZUA 3080–81, IZUA 3224 (cleared-and-stained adult), IZUA 3146 (two cleared-and-stained adults), IZUA 3225–26, IZUA 3132 (five tadpoles); CBF 3760–62; MZUC 24104–06, MZUC 25725–38.

Telmatobius zapahuirensis.—CHILE: Provincia

Parinacota: Quebrada de Zapahuira, 3270 m, DBCG 630, 639.

Telmatobius philippii—CHILE: Provincia El Loa: Quebrada de Amincha, 3700 m, IZUA 3093, 3087; Provincia El Loa: Quebrada del Inca, 3700 m, IZUA 3088–92, 3193–95 (three cleared-and stained specimens, chromosomes), 3094 (seven tadpoles), 3196–97 (chromosomes).

Telmatobius fronteriensis.—CHILE: Provincia El Loa: Puquios, 4150 m, MZUC 25094, 25095–103, 25261–78, 25095, and 25103 (cleared-and stained adults).

Monogr. Herpetol. (2005) 7:167-185

The *Telmatobius* species complex in Lake Titicaca: applying phylogeographic and coalescent approaches to evolutionary studies of highly polymorphic Andean frogs

EDGAR BENAVIDES

Department of Integrative Biology and M.L. Bean Museum, Brigham Young University, 401 WIDB, Provo, Utah, 84602 USA
(e-mail: eb235@email.byu.edu)

Abstract: The genus *Telmatobius* is a fascinating group of amphibians of the Andean Range, both because of its species diversity and because of its successful adaptation to high altitude aquatic environments. Additionally, most species exhibit impressive levels of phenotypic variation. In this contribution, I reexamine and extend the results obtained in a previous assessment of species boundaries of the Lake Titicaca *Telmatobius* complex by presenting genealogical information obtained from a fast evolving mitochondrial marker. I use a combination of methods that complement the phylogenetic approach and cross-validate demographic inferences drawn from a nested clade analysis (NCA). The results suggest the absence of reciprocal monophyly for the two species previously recognized as *T. culeus sensu lato* and *T. marmoratus*. Results of NCA also provided support for the recognition of two clades that are preferentially, but not exclusively, composed by "riverine" and "lacustrine" populations. Furthermore, the combination of information drawn from estimates of nucleotide diversity, neutrality tests, and haplotype mismatch distribution supported the hypothesis of significantly different demographic histories for these two clades. Lacustrine populations seem to have undergone a genetic bottleneck effect while riverine populations seem to have maintained their historical effective population sizes.
Key words: Amphibia, Bolivia, Peru, phylogeny, population genetics, *Telmatobius.*

Resumen: El complejo de especies de *Telmatobius* en el lago Titicaca: aplicando aproximaciones filogeográficas y coalescentes a estudios evolutivos de ranas andinas altamente polimórficas. – El género *Telmatobius* es un fascinante grupo de anfibios andinos. Es interesante por su diversidad específica, así como también por su exitosa adaptación a ambientes acuáticos de gran altura. Adicionalmente, este género se caracteriza por mostrar una impresionante plasticidad fenotípica. En esta contribución, se examinan y extienden los resultados obtenidos de un estudio previo acerca de los límites interespecíficos de las especies de *Telmatobius* del lago Titicaca. Se utiliza para ello información obtenida de un marcador mitocondrial de evolución rápida. Se utiliza una combinación de métodos que complementan los resultados de filogenia tradicional y de inferencias demográficas logradas aplicando un análisis de clados encajados (NCA por sus siglas en inglés). Los resultados sugieren que las poblaciones del lago Titicaca previamente asignadas a *T. culeus sensu lato* y *T. marmoratus* no constituyen grupos monofiléticos. Los resultados del NCA indican, asimismo, la existencia de dos grupos aproximadamente compuestos por poblaciones lacustres y poblaciones riparias. Adicionalmente, la información obtenida a partir de estimadores de diversidad nucleotídica, tests de neutralidad y de distribución de haplotipos pareados (MDA) apoya la conclusión de que estos dos grupos de poblaciones muestran historias demográficas distintas. Se presenta evidencia de que las poblaciones lacustres son el resultado de un efecto de cuello de botella reciente, en tanto que las poblaciones riparias han mantenido un tamaño poblacional constante.
Palabras clave: Amphibia, Bolivia, filogenia, genética poblacional, Perú, *Telmatobius.*

INTRODUCTION

The genus *Telmatobius* is one of the most interesting groups of amphibians of the Andean Range, both because of its diversity (approximately 50 species are known; BENAVIDES *et al*., 2002; DE LA RIVA, 2002) and because of its putative adaptation to high altitude aquatic environments (CEI, 1986). One notable feature of this group is its somewhat conservative morphology at a generic level, coupled with impressive levels of intra-specific variability. This phenotypic plasticity is not only related to external features such as size, coloration or external morphology (TRUEB, 1979) but also to some osteological characters (WIENS, 1993; FORMAS *et al*., 2003).

In a previous paper, BENAVIDES *et al*. (2002) used morphological and allozyme data to address the issue of species boundaries among four taxa inhabiting the Titicaca Basin; these taxa are characterized by quite remarkable differences in body sizes and by the fact that they seem to occupy well defined habitats within the lake as well as in the outer surrounding basin. Specifically, *T. culeus* and *T. albiventris* are entirely lacustrine, whereas *T. crawfordi* and *T. marmoratus* reside in small streams and ponds to the north of, and around the perimeter of Lake Titicaca, respectively. These authors found that, despite high levels of morphological variation, when body size was removed as a descriptor of morphological variability, the four putative subspecies of *T. albiventris*, and populations assigned to *T. culeus* and *T. crawfordi*, comprised a single species, whereas the putative *T. marmoratus* maintained its distinction. These results were independently confirmed by screening of 25 allozyme loci (BENAVIDES *et al*., 2002); no allelic differentiation was observed for any "species" except *T. marmoratus*, which could be diagnosed by non-fixed but exclusive (= private) alleles in three loci. BENAVIDES *et al*. (2002) concluded that only two species should be recognized in the Titicaca Basin: (1) *T. marmoratus*, found in streams and ponds around Lake Titicaca, and (2) *T. culeus* confined to Lake Titicaca and nearby interconnected lagoons.

The above-cited study provided well-supported species boundaries among the *Telmatobius* taxa of Lake Titicaca. Nonetheless, it was based on allozyme polymorphisms and morphometric characters only, thus limited genealogical information could be drawn. Furthermore, impoverished levels of genetic variability among populations of Lake Titicaca itself precluded estimates of gene flow, although it also suggested that limited allozyme variation may reflect recent bottlenecks presumably associated with dramatic late Pleistocene changes in lake levels and salinity (BENAVIDES *et al*., 2002).

In this paper, I reexamine and extend the results obtained in the previous assessment of species boundaries and relationships of the Lake Titicaca *Telmatobius* complex by presenting genealogical information obtained from a fast evolving mitochondrial marker.

MATERIALS AND METHODS

Taxon sampling

One hundred and thirteen specimens from 10 localities within the Lake Titicaca Basin, and a single outgroup taxon (*T. vilamensis* from Estero Vilama near San Pedro de Atacama, roughly 1000 km S of Lake Titicaca) were used in this analysis. All but three localities consisted of tissue vouchers previously used in BENAVIDES *et al*. (2002) (Table 1). New localities included in this paper are Río Ramis for *T. albiventris parkeri*, and Charazani and Comanche for

T. marmoratus (Fig. 1). Voucher specimens are catalogued in the Museo de Zoología de la Universidad de Concepción (MZUC), and M.L. Bean Life Science Museum, Brigham Young University (BYU). All localities and voucher numbers are described in Table 1.

Laboratory procedures

Total genomic DNA was extracted from muscle tissue preserved in 100% ethanol using a slightly modified version of the extraction procedure of FETZNER (1999) or the Quiagen DNeasy tissue kit (Quiagen Inc., Valencia, CA). I used PCR to amplify an 817 bp fragment of the mitochondrial cytochrome *b* (Cyt *b*) gene. Each 50 µl reaction volume contained 2.0 µl of template DNA (approximate concentration estimated on a 2% agarose gel), 8 µl of dNTPs (1.25 mM), 4 µl of 10x Taq buffer, 4 µl of each primer (10 µM), 4µl of MgCl2 (25 mM), 31.8 µl of distilled water and 0.20 µl of Taq DNA polymerase (5U/µ) from Promega Corp., Madison WI. Primers sequences were the light strand primer MVZ15 (5'-GAA CTA ATG GCC CAC ACW WTA CGN AA-3) (MORITZ *et al*., 1992) and the heavy strand primer Cyt *b* 3 (5'-GGC AAA TAG GAA RTA TCA TTC-3' PALUMBI, 1996).

The thermal cycler profile was one cycle at 95°C for 2 min, 40 cycles at 95°C for 1 min, 50°C for 1 min, 72°C for 1 min 30 s, and one cycle at 72°C for 7 min. PCR products were cleaned using the Millipore Montage PCR_{96} cleaning plates (Millipore Corp., Bedford, MA) and both forward and reverse strands were sequenced on an ABI 377 Automated DNA sequencer (PE Applied Biosystems, Foster City, CA). Sequences were generated using cleaned PCR product, BigDye terminator cycle sequencing mix (PE Applied Biosystems, Foster City, CA), and 5 pmol of one of the original PCR primers. The thermal cycler profile was one cycle at 96°C for 2 min, 25 cycles at 96°C for 30 s, 50°C for 30 s, and 60°C for 4 min. Sequencing products were cleaned using Centrisep Sephadex columns (Princeton Separations, Adelphia, NJ). Forward and reverse sequences for each individual were edited and aligned using Sequencher 4.0 (Gene Codes Corporation, Ann Arbor, MI). Sequence accuracy (base calling) was further assessed by aligning all sequences and crosschecking the chromatograms for all sites that differed from the consensus sequence. There were no indels in the sequence, and translation to amino acids verified the absence of stop codons.

Phylogenetic analyses

Sequence variation of 817 bp in the Cyt *b* gene region was used to infer phylogeny by Bayesian (MrBayes v3.0B4; HUELSENBECK & RONQUIST, 2001) and maximum likelihood (ML) (PAUP*; SWOFFORD, 2000) approaches. For the ML analyses, the ModelTest program ver. 3.0 (POSADA & CRANDALL, 1998) was used to determine the best substitution model. For this data set, the F81 + G (FELSENSTEIN, 1981) model was selected by the hierarchical likelihood ratio test (hLRT), while the TrN + G (TAMURA & NEI, 1993) model was chosen by the Akaike information criterion (AIC). The simplest model of nucleotide substitution (F81 + G) was implemented for ML heuristic searches consisting of 10 random additions with the TBR branch-swapping algorithm. Nodal support was assessed by nonparametric bootstrap analysis, based on a total of 100 pseudoreplicates to obtain the bootstrap proportions. Bootstrap values ≥ 70% were considered as evidence for significantly supported clades (HILLIS & BULL, 1993; with caveats). Bayesian analyses were conducted with random starting trees, run for 1.0 x 10^6 generations, and sampled every 100 generations; this search was based on a

TABLE 1. Localities, sample sizes, putative taxonomic names, and selected ecological data for populations of *Telmatobius* sampled in this study. All localities with the exception of Estero Vilama are plotted in Fig. 1. BYU: M.L. Bean Life Science Museum, Brigham Young University, Utah, USA; MNCN: Museo Nacional de Ciencias Naturales, Madrid, Spain; MZUC: Museo de Zoología de la Universidad de Concepción.

TABLA 1. Localidades, tamaños de muestra, supuestos nombres y datos ecológicos seleccionados para poblaciones de *Telmatobius* utilizadas en este estudio. Todas las localidades excepto Estero Vilama se muestran en la Fig. 1. BYU: M.L. Bean Life Science Museum, Brigham Young University, Utah, USA; MNCN: Museo Nacional de Ciencias Naturales, Madrid, Spain; MZUC: Museo de Zoología de la Universidad de Concepción.

Locality	Habitat and water depth	Putative Taxon names	Voucher Numbers
Tiquina 16º 13' 0" S 68º 48' 60" W	Lacustrine, > 10 m	*Telmatobius culeus*	BYU 46550, MZUC 25576–77, MZUC 25583, MZUC 25588, MZUC 25591, MZUC 25596, MZUC 25605–07, MZUC 25609–11
Isla Taquiri 16º 11' 60" S 68º 37' 60" W	Lacustrine, < 10 m	*Telmatobius albiventris albiventris*	BYU 46763, MZUC 25739–44, MZUC 25592–94, MZUC 25599, MZUC 25608
Playa Copani 16º 19' 0" S 69º 1' 60" W	Lacustrine, < 2 m	*Telmatobius albiventris* "copani"	BYU 46536, BYU 46539, BYU 45541, BYU 46544, MZUC 25435–36, MZUC 25441, MZUC 25449, MZUC 25454, MZUC 25456, MZUC 25459
Bahía de Puno 15º 49' 60" S 70º 1' 60" W	Lacustrine, < 10 m	*Telmatobius albiventris punensis*	BYU 46624, BYU 46630, MZUC 25421, MZUC 25423–24, MZUC 25426–27, MZUC 25429, MZUC 25482, MZUC 25484–86, MZUC 25488, MZUC 25491, MZUC 25526, MZUC 25528, MZUC 25530, MZUC 25431–32, MZUC 25533, MZUC 25537
Laguna Arapa 15º 7' 60" S 70º 7' 0" W	Lacustrine, < 10 m	*Telmatobius albiventris parkeri*	BYU 46610-12, BYU 46616, BYU 46635, MZUC 25460–61, MZUC 25463–64, MZUC 25466, MZUC 25489
Río Ramis 15º 16' 60" S 70º 1' 0" W	Stream, < 0.5 m	*Telmatobius albiventris parkeri*	MZUC 25538, MZUC 25498, MZUC 25541, MZUC 25545, MZUC 25547, MZUC 25516, MZUC 25518, MZUC 25550
Laguna Saracocha 15º 44' 52" S 70º 46' 19" W	Pond, streams, < 0.5 m	*Telmatobius crawfordi*	BYU 25551, BYU 46554, BYU 46556–57, BYU 46559, BYU 46561–64, BYU 46566, BYU 46568–69, BYU 46571, BYU 46573–74, BYU 46578-80, BYU 46582, MZUC 25542
Laguna Huayna Potosí 16º 7' 0" S 68º 1' 60" W	Stream, < 0.5 m	*Telmatobius marmoratus*	MZUC 25552–57
Charazani 15º 11' 60" S 69º 2' 60" W	Stream, < 0.5 m	*Telmatobius marmoratus*	MNCN 3459, MNCN 3685, MNCN 3847
Comanche 16º 56' 60" S 68º 25' 0" W	Stream, < 0.5 m	*Telmatobius marmoratus*	MNCN 3290
Estero Vilama 22º 51' 43" S 68º 23' 25" W	Stream, < 0.5 m	*Telmatobius vilamensis* (OG)	MZUC 25094–99

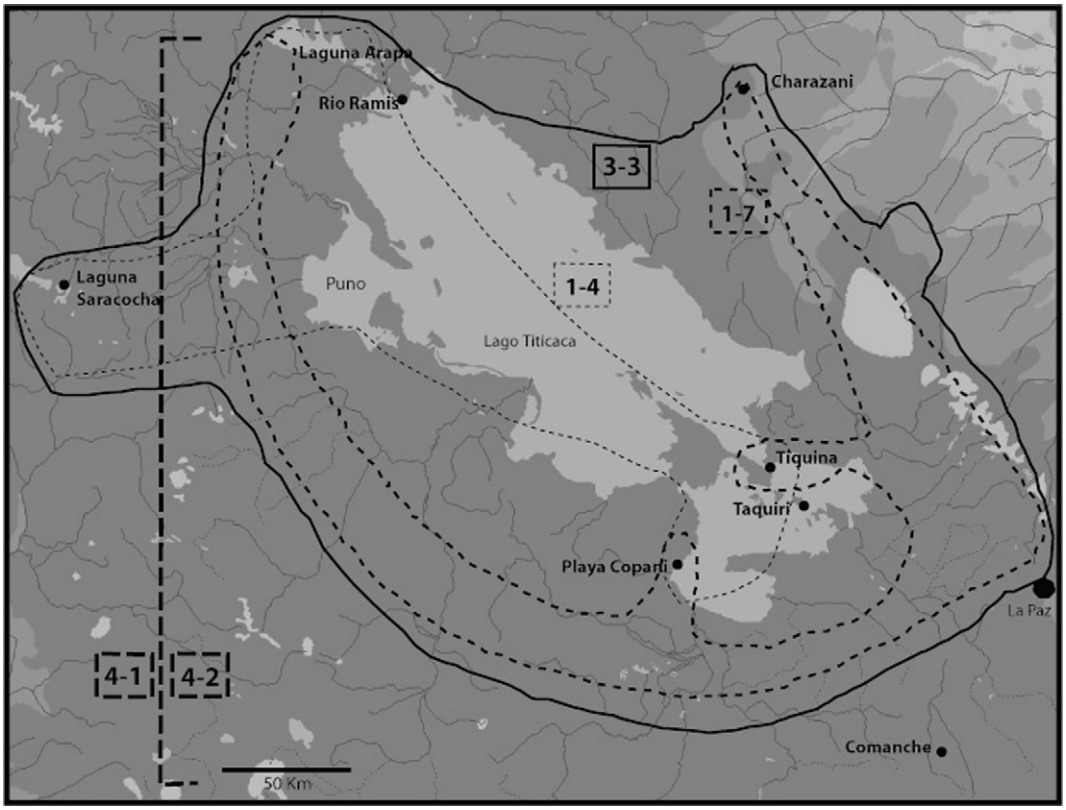


FIGURE 1. Map of the distribution of the *Telmatobius* populations considered in this study (see Table 1). Different line styles illustrate a spatial representation of the clades that show a significant relationship with geography (Table 3).

FIGURA 1. Mapa de distribución de las poblaciones de *Telmatobius* consideradas en el estudio (véase Tabla 1). Los diferentes estilos de línea muestran una representación espacial de los clados que tienen una relación significativa con la geografía (Tabla 3).

general time reversible model with variable sites with a discrete gamma distribution (GTR + Γ + G) given the discordance between models selected by ModelTest. Stationarity of the Markov chain was determined as the point when sampled log likelihood values plotted against generation time reached a stable equilibrium value; "burn-in" data sampled from generations previous to the equilibrium plateau were discarded. The equilibrium samples were used to generate a 50% majority rule consensus tree as a best estimate of the phylogeny that included branch lengths. The percentage of samples that recover any particular clade represents the posterior probability (p values) for that clade, and a value of $p \geq 95\%$ was considered as evidence for a significant support for a clade (HUELSENBECK & RONQUIST, 2001).

Phylogeographic analyses: haplotype network estimation and geographical associations

In order to estimate gene genealogies the original 817 bp matrix was trimmed down to 470 bp, thus eliminating missing data while

preserving the largest number of individuals and the most haplotype variation. I used TCS (CLEMENT *et al*., 2000) to implement the statistical parsimony procedure developed by TEMPLETON *et al*. (1992). The TCS program produced a network linking different haplotypes only if they have a 95% probability of being justified by an algorithm that enforces a parsimony criterion. Interconnections among plausibly non-nested networks were estimated from a pairwise matrix of absolute number of differences generated in PAUP* (SWOFFORD, 2000). Once the haplotype network was available, I used the rules given in TEMPLETON *et al*. (1987) and TEMPLETON & SINGH (1993) to define a clade hierarchy, or nested design, that was subsequently used to test the null hypothesis of haplotype random geographic distribution for all clades within each nesting level by using Geodis ver. 2.0 (POSADA *et al*., 2000). For those clades in which the null hypothesis is rejected, population processes underlying the non-random haplotype/geography association were inferred by using the inference key of TEMPLETON *et al*. (1995; an updated version is available at http://bioag.byu.edu/zoology/crandall_lab/geodis.htm). To cross-validate the nested clade analyses (NCA) inferences (ALTHOFF & PELLMYR, 2002; PFENNINGER & POSADA, 2002; TEMPLETON, 2002; MASTA *et al*., 2003), I estimated two measures of nucleotide variability: gene and nucleotide diversity indices (TAJIMA, 1983; NEI, 1987) for the original populations and the nested categories provided by the NCA. I also calculated Tajima's D (TAJIMA, 1989) and Fu's Fs (FU, 1997) tests to independently assess phylogeographic history for those clades for which evolutionary inferences were drawn after NCA. A third independent assessment of historical changes of population sizes was performed for these same clades by implementing a mismatch distribution analysis (MDA) (ROGERS & HARPENDING, 1992; HARPENDING & ROGERS, 2000). Genetic diversity estimates, neutrality tests, and mismatch distribution analyses were implemented in the Arlequin ver. 2.0 package (SCHNEIDER *et al*., 2000).

Finally, current levels of gene flow and contemporary population structure for lacustrine populations of Lake Titicaca (clade 1-4 in Figs. 1, 3) were studied by means of an analysis of molecular variance (AMOVA) (EXCOFFIER *et al*., 1992). Standard variance components were generated at three levels of subdivision established a priori; among major geographical regions (northern embayment [Laguna Arapa, Puno, Río Ramis, and Laguna Saracocha]) versus southern embayment [Playa Copani, Tiquina, and Taquiri]); among populations (i.e. localities) within regions (northern or southern embayment), and among individuals within populations within regions. The significance of the variant components was obtained through 1023 nonparametric permutations in Arlequin (SCHNEIDER *et al*., 2000).

RESULTS

Haplotype variation

Nineteen haplotypes were found in the total of 113 frogs screened for the short fragment (470 bp) (Table 2); estimates of nucleotide diversity ranged from 0.00 (Bahía de Puno, Isla Taquiri, Río Vilama) to 0.015 (Río Ramis), and the mean number of pairwise distances ranged from 0.00 (Río Vilama) to 8.22 (Playa Copani). Two haplotypes were found in more than three localities (VI [7 localities] and XII [5 localities]), five haplotypes were found in two localities (I, II, IV, VIII, X), and 12 haplotypes were restricted to a single locality (singletons) (Table 2). Haplotype VI is the

TABLE 2. Distribution of cytochrome *b* haplotypes (I–XIX) across 11 populations of *Telmatobius* based on the 470 bp used in the nested clade analysis, and genetic parameter estimates of the original populations. N: nucleotide diversity (± standard deviation), Mean: mean number of pairwise differences (π), n: sample size.

TABLA 2. Distribución de haplotipos del citocromo *b* (I-XIX) en 11 poblaciones de *Telmatobius* basada en los 470 pares de bases usados en el análisis de clados encajados, y estimas de parámetros genéticos de las poblaciones originales. N: diversidad de nucleótidos (± desviación típica), Mean: número medio de diferencias por pares (π), n: tamaño muestral.

	Freq.	Laguna Saracocha (n = 20)	Playa Copani (n = 11)	Río Ramis (n = 8)	Laguna Arapa (n = 11)	Bahía de Puno (n = 22)	Isla Taquiri (n = 12)	Tiquina (n = 13)	Laguna Huayna Potosi (n = 6)	Charazani (n = 3)	Comanche (n = 1)	Río Vilama (n = 6)
I	3	2	1									
II	6	5		1								
III	1			1								
IV	3		1					2				
V	1				1							
VI	53	1	3	1	7	22	10	9				
VII	1							1				
VIII	3		1				2					
IX	2				2							
X	13	12	1									
XI	1		1									
XII	13		1	4	1			1	6			
XIII	1			1								
XIV	2									2		
XV	1									1		
XVI	1		1									
XVII	1		1									
XVIII	1										1	
XIX	6											6
N		0.011 ± 0.006	0.014 ± 0.008	0.015 ± 0.009	0.004 ± 0.002	0.000 ± 0.000	0.000 ± 0.000	0.003 ± 0.002	–	0.001 ± 0.001	–	0.000 ± 0.000
Mean		8.100	8.218	8.107	2.654	0.090	0.303	2.256	–	1.333	–	0.000

most geographically widespread and its frequency is particularly high in lacustrine localities (Laguna Arapa, Bahía de Puno, Isla Taquiri, and Tiquina). Haplotypes X and XII were the second most frequent haplotypes (n = 13), but haplotype X was almost entirely restricted to Laguna Saracocha, whereas haplotype XII was more frequent in populations from Laguna Huayna Potosi (n = 6), and Río Ramis (n = 4) (Table 2).

Phylogenetic analyses

An exploratory phylogenetic analysis was first conducted using the Bayesian approach and the original data set (817 bp and 113 terminals) that recovered three large polytomies (not shown). Because unresolved clades considerably increase computation time for a ML approach, I pruned and reduced the number of terminals of the original data set by calculating the absolute number of pairwise differences and deleting haplotypes differing by less than 2 bp. This procedure reduced the matrix to 59 exemplar haplotypes and two outgroup terminals. Both ML and Bayesian analyses were applied to this data set. In the Bayesian analysis the Markov chain appeared to reach stationarity by 8.0×10^4 generation; these "burn-in" generations were discarded, and the remaining 9.2×10^6 generations were used for the estimation of a mean likelihood score (ln = –1904.6183933) and posterior nodal probabilities. For the ML analysis, I obtained an island of 10 most probable trees (ln = –1739.8595) with almost identical topologies. Likewise, ML and Bayesian inferences showed the same general topology with no major differences (ML tree) (Fig. 2). The most obvious feature of this tree is the absence of phylogenetic structure at shallow levels of divergence and the absence of reciprocal monophyly for the two species delimited by BENAVIDES *et al*. (2002) (i.e. *T. culeus sensu lato* which includes populations from Puno, Copani, Tiquina, Arapa, Taquiri, and Laguna Saracocha, and *T. marmoratus* which includes populations from Laguna Huayna Potosi, Comanche and Charazani). Moreover, not even single population samples show reciprocal monophyly, as evident with haplotypes from Laguna Saracocha for which there is good support for three well differentiated clades (2-5, 2-6, and 3-1 in Fig. 2). Likewise, it appears that two main haplotype groups coexist in Playa Copani (clades 2-5 and 3-4 in Fig. 2), since haplotypes of this locality are alternatively placed in and out the primarily "lacustrine" clade 2-5. Overall, the deepest split occurs between a group of haplotypes belonging to Laguna Saracocha and Río Ramis populations (clade 3-1 in Fig. 2), and haplotypes from all other localities, notably including a second haplotype group from Laguna Saracocha and Río Ramis. The second deepest split provides only partial evidence to discriminate between lacustrine (clade 2-5) and riverine haplotypes (clades 2-5, 2-7, and 3-4). Interdigitation of lacustrine haplotypes into riverine haplotypes and vice versa, renders both major clades geographically inconsistent.

Phylogeographic analysis

Application of statistical parsimony to the complete matrix of haplotypes (470 nucleotides) showed that those connected by ≤ 9 substitutions have at least a 95% probability of being parsimoniously connected (i.e. $p < 0.05$ for multiple substitutions at any variable nucleotide positions). The nested clade design was largely congruent with the phylogenetic inference, and revealed that clade 3-1 (haplotypes I, II and III), and haplotype group XIX (corresponding entirely to the outgroup *T. vilamensis*) were separated from all other

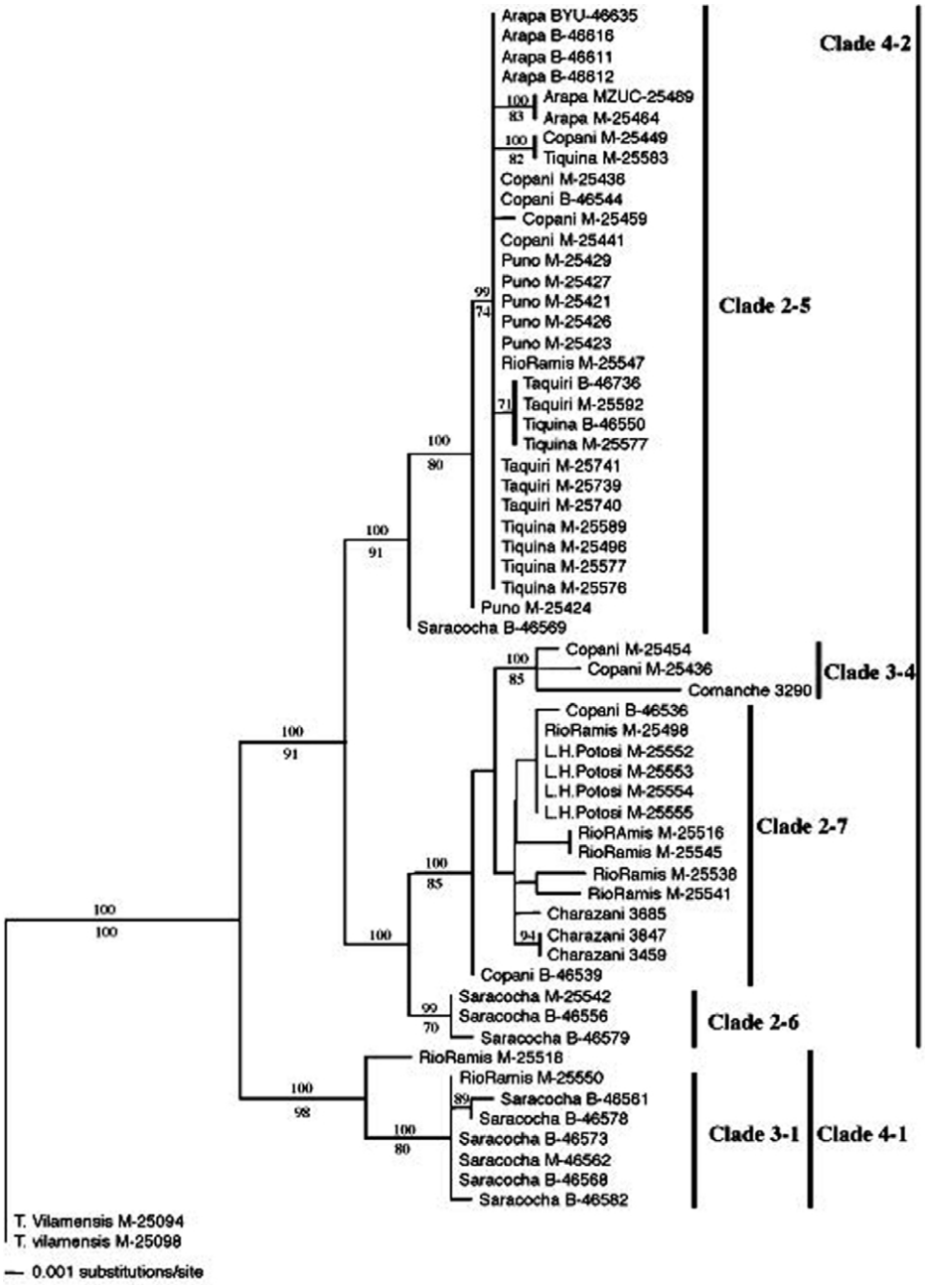


FIGURE 2. One of ten most probable ML trees (ln = –1739.8595) recovered with a F81 + G model of sequence evolution. This tree was identical to a 50% majority rule consensus tree from the Bayesian analysis (GTR + Γ + G). Nodal support is given as likelihood bootstrap values (below) and posterior probability values (above) branches.

FIGURA 2. Se representa uno de los diez árboles ML más probables (ln = –1739.8595) obtenidos siguiendo el modelo de evolución F81 + G. La topología de este árbol resultó ser idéntica a la obtenida en el árbol consenso (50% majority rule) en el análisis Bayesiano (GTR + Γ + G). Los valores de bootstrap (abajo) y los valores de probabilidad posterior (arriba) aparecen indicados en los nodos de las ramas.

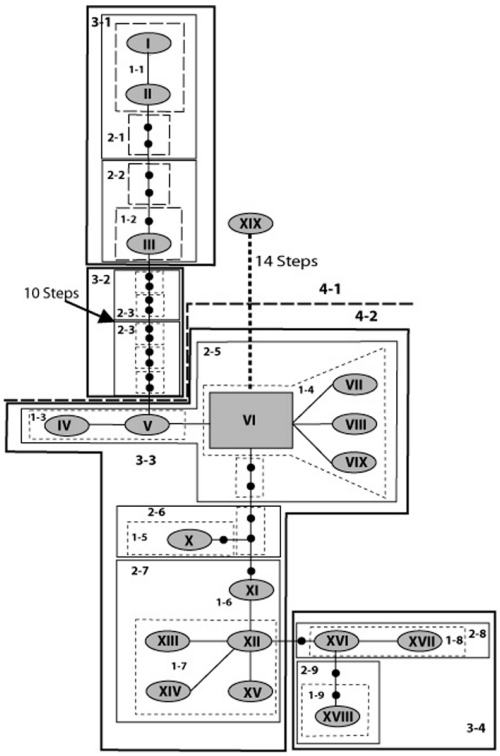


FIGURE 3. Hierarchically nested haplotype network of Titicaca Basin *Telmatobius* constructed by estimating 95% plausibility sets (≤ 9 substitutions) with statistical parsimony, and interconnecting sets from an pairwise absolute difference matrix. Sampled haplotypes are indicated by Roman numerals (Table 2) and inferred haplotypes by solid dots. The square denotes the most frequent haplotype (haplotype VI).

FIGURA 3. Red de los haplotipos encajados jerárquicamente de *Telmatobius* en la cuenca del lago Titicaca, construida mediante el cálculo de grupos con 95% de probabilidad (≤ 9 sustituciones) con parsimonia estadística, y relacionando los grupos obtenidos a partir de una matriz de diferencias absolutas por pares. Los haplotipos representados están indicados con números romanos (Tabla 2), y los inferidos con puntos sólidos. El cuadrado denota el haplotipo más frecuente (haplotipo VI).

clades by 10 and 14 mutational steps respectively (Fig. 3). All other clades are part of a single network. The nested design showed haplotype VI being nested in a central, interior position due to its presence in all localities within Lake Titicaca, as well as in some outer lake (riverine) localities (Río Ramis, Laguna Saracocha, and Laguna Arapa) (Table 2). The nested network also confirmed that the Laguna Saracocha and Playa Copani localities seem to have a mixed origin. Both populations are characterized by one haplotype group nested in clade 2-5, whereas a second haplotype group is nested either in clade 3-1 (10 mutational steps separated from the main network in the case of Laguna Saracocha), or within riverine clades 3-4 and 2-7 in the case of Playa Copani. No clade incongruences (loops) were detected. GeoDis revealed significant deviations from the null hypothesis of no geographic structure in five nested clades; two one step clades (1-4 and 1-7) and three higher-level clades (3-3, 4-2, and total cladogram) (Table 3 and Fig. 1). Restricted gene flow with isolation-by-distance was inferred for clade 1-7 (that includes mostly riverine individuals with the exception of two samples from Copani and Río Ramis), and clade 1-4 (that includes 51 individuals from lacustrine localities [Puno, Taquiri, Tiquina, Copani, and Laguna Arapa], and only two from riverine localities [Río Ramis and Laguna Saracocha]) (Table 2 and Fig. 1). Conversely, range expansion was inferred for clade 3-3, and continuous range expansion was inferred for clade 4-2. No clear inference could be drawn from the total cladogram because geographical sampling did not include a sufficient number of locations between Laguna Saracocha and Lake Titicaca to distinguish between historical fragmentation and restricted gene flow with isolation-by-distance.

Molecular diversity, neutrality tests and contemporaneous population structure

Table 4 summarizes two estimates of nucleotide diversity in four nested clades for which the NCA allowed inferences of one or

TABLE 3. Nested contingency analysis of geographical associations (* significant at the 0.05 level, 10 000 permutations) for cytochrome *b* data from *Telmatobius* populations, and evolutionary inferences. CRE: continuous range expansion, IBD: isolation-by-distance, LDC: long distance colonization, RE: range expansion.

TABLA 3. Análisis de contingencia encajado de asociaciones geográficas (* significativo a nivel 0.05, 10 000 permutaciones) para datos del citocromo *b* de poblaciones de *Telmatobius,* e inferencias evolutivas. CRE: expansión continua del rango de distribución, IBD: aislamiento por distancia, LDC: colonización a larga distancia, RE: expansión del rango de distribución.

Clade	X^2	p	Inference Chain	Inferred Pattern
1-4	25.05	0.2270	1-2-3-4-NO	Restricted gene flow with IBD
1-7	19.35	0.3480	1-2-3-4-NO	Restricted gene flow with IBD
3-3	145.09	0.000*	1-2:NO-11:RE-12-13-14:NO	Range expansion but sampling inadequate to discriminate between CRE or LDC
4-2	43.61	0.002*	1-2-11:YES-RE-12:NO 10:NO	CRE
Total cladogram	24.86	0.004*	1-2-3-4-9-	Sampling inadequate to discriminate between historical fragmentation and IBD

more processes. Gene and nucleotide diversity estimates are moderate and show increasingly higher values with more inclusive clades. In all cases haplotype diversity values are higher relative to nucleotide diversity (H [0.573-4.665], % [0.001-0.009]). Tajima's D was marginally significant for clade 1-4, and non-significant for clades 1-7, 3-3, and 4-2. Conversely, all clades showed highly significant Fs negative values. The results of the mismatch analysis on clades 1-4 (lacustrine haplotypes), and 1-7 (riverine haplotypes) provides additional information. While there is good fit of the mismatch distribution of pairwise distances to a model of sudden population expansion in the case of the lacustrine clade 1-4 (Fig. 4A), a flatter mismatch distribution for clade 1-7 is more consistent with a historically stable population size (Fig. 4B). Mismatch distribution of clades 3-3, and 4-2 showed bimodal distributions (not shown), that are consistent with lineages with demographically stable populations. The AMOVA assessment of contemporaneous population structure of lacustrine haplotypes/localities included in clade 1-4 shows that the great majority of mitochondrial genetic variation occurs within single populations, rather than between populations of the same lake region (populations within the southern or northern embayment), or between the two major Lake Titicaca regions (southern and northern embayment) (Table 5).

DISCUSSION

General phylogenetic patterns

The most surprising result of the phylogenetic inference is the failure in recovering a single clade for populations putatively assigned to *T. marmoratus* or *T. culeus sensu lato*. Instead, the strongly supported basal branching shows the deepest split between clade 3-1 (Río Ramis and Laguna Saracocha haplogroup), and two additional clade groups previously described as typically lacustrine (clade 2-5) and primarily riverine clades (2-6, 2-7, and 3-4). Also notable is the contrast between riverine populations' deeper branches and the "star-like" phylogeny of the lacustrine populations.

TABLE 4. Two estimates of nucleotide diversity (mean ± SE) and neutrality tests (Tajima's D and Fu's Fs) for clades with significant haplotype-geographic associations identified with NCA (Table 3). n: sample size, NS: non significant, *: $p < 0.05$, **: $p < 0.001$.

TABLA 4. Dos estimaciones de diversidad de nucleótidos (media ± SE) y pruebas de neutralidad (D de Tajima y Fs de Fu) para clados con asociaciones significativas entre haplotipo y geografía, identificadas con NCA (Tabla 3). n: tamaño muestral, NS: no significativo, *: $p < 0.05$, **: $p < 0.001$.

	n	Gene Diversity	Nucleotide diversity (% p)	Tajima's D	Fu's Fs
Clade 1-7	17	0.573 ± 0.487	0.0014 ± 0.0010	–1.576 (NS)	–13.02**
Clade 1-4	59	0.362 ± 0.355	0.0005 ± 0.0000	–1.771*	–34.02 x 1038**
Clade 3-3	94	3.986 ± 2.011	0.0073 ± 0.0410	0.384(NS)	–11.78**
Clade 4-2	103	4.665 ± 2.305	0.0099 ± 0.0054	–0.658(NS)	–11.40**

The rather old divergence of clade 3-1, and the absence of reciprocal monophyly between *T. marmoratus* and *T. culeus* populations, differs from the topology obtained through a step matrix frequency-parsimony approach of 25 allozyme loci by BENAVIDES *et al*. (2002) (see Fig. 5). In this tree, *T. culeus sensu lato* populations were monophyletic whereas *T. marmoratus* and *T. vilamensis* sorted as the first and second outgroup, respectively.

Discordances between hypotheses of relationships among taxa based on mitochondrial DNA and nuclear genes may be present by multiple reasons, ranging from natural selection to the population biology of the organism involved (FUNK & OMLAND, 2003; GARCÍA-PARÍS *et al*., 2003). The most frequently cited explanation for such incongruence is based on the time elapsed since the divergence of two or more

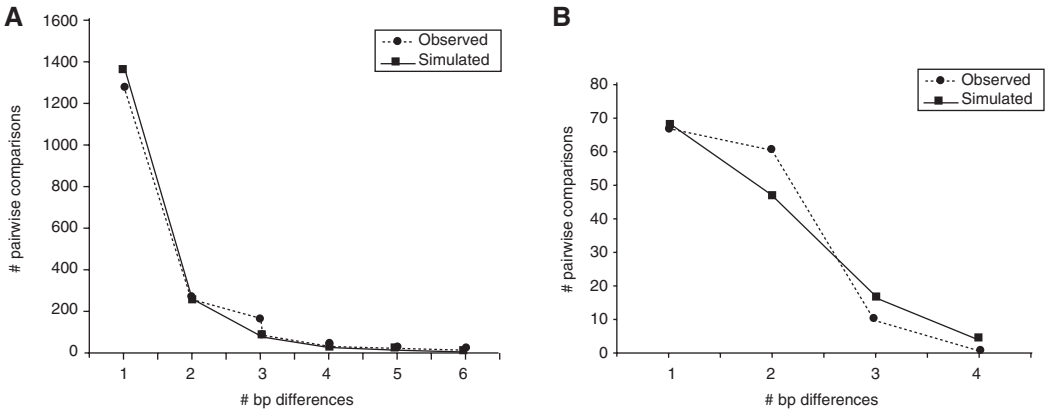


FIGURE 4. Mismatch distributions for pairwise combinations of all individuals nested in clade 1-4 (A) and clade 1-7 (B). The observed unimodal distribution in clade 1-4 conforms to a model of population growth or sudden expansion, whereas the ragged distribution of the mismatch distribution in clade 1-7 conforms to a model of stable population sizes.

FIGURA 4. Distribuciones dispares de las combinaciones por pares de los individuos de los clados 1-4 (A) y 1-7 (B). La distribución unimodal observada en el clado 1-4 se ajusta un modelo de crecimiento poblacional o a una expansión repentina, mientras que la distribución truncada en el clado 1-7 se ajusta a un modelo poblacional de tamaños estables.

TABLE 5. Results of an analysis of molecular variance as implemented in Arlequin. F-statistics are analogous to Wright's *F*-statistics and identify the correlation of alleles at three *a priori* established hierarchical levels: among geographical regions (northern [Laguna Arapa, Puno, Río Ramis, and Laguna Saracocha] *versus* southern [Playa Copani, Tiquina, and Taquiri] groups of localities nested in clade 1-4), among populations (i.e. localities) within regions, and among individuals within populations. ***: $p < 0.0001$.

TABLA 5. Resultados de un análisis de la varianza molecular tal como se implementa con Arlequín. Los estadísticos F son análogos a los Wright *F* e identifican la correlación de alelos a tres niveles jerárquicos establecidos *a priori*: entre regiones geográficas (norte [Laguna Arapa, Puno, Río Ramis y Laguna Saracocha] *versus* sur [Playa Copani, Tiquina y Taquiri]), entre poblaciones (localidades) dentro de regiones y entre individuos dentro de poblaciones. ***: $p < 0.0001$.

Source of variation	Degrees of freedom	Variation	Φ-statistics
Among regions	1	-28.37%	$\Phi_{CT} = -0.28$
Among populations within regions	5	59.63%	$\Phi_{SC} = 0.464$***
Among individuals within populations	74	68.74%	$\Phi_{ST} = 0.312$***

populations. In recent divergences it is difficult to assure the convergence of mitochondrial gene trees to the organismal tree through the achievement of reciprocal monophyly. In such cases, conflict between nuclear and mtDNA gene trees of closely related species is usually explained as a consequence of ancestral polymorphism retention, incomplete lineage sorting, or gene introgression (CRANDALL & TEMPLETON, 1994; MADDISON, 1997; SULLIVAN *et al*., 2002; FUNK & OMLAND, 2003). In the present case, monophyly of the Lake Titicaca *Telmatobius* based on allozyme polymorphisms is very weakly supported, as is the placement of *T. marmoratus* as a separate species (bootstrap values < 70%; see Fig. 5 in BENAVIDES *et al*., 2002). Conversely, the phylogenetic signal of the mtDNA haplotype tree appears to be strong and is supported by high bootstrap (> 70%) and posterior probability (> 90%) values. The mtDNA tree reconstructed here does provide with meaningful evidence of recent population history (see SLATKIN & HUDSON, 1991; HARPENDING *et al*., 1998). On one hand, the fact that some haplotypes in this data set did not sort phylogenetically by locality implies that; (1) the process of lineage sorting is not complete, (2) is prevented by gene flow or (3) indicate the possibility that introgression or secondary contact is currently taking place (e.g. Laguna Saracocha in clades 3-1 and 2-6 in Fig. 2). On the other hand, star-like topologies like clade 2-5 (Fig. 2) can be interpreted as the result of a population bottleneck followed by rapid population growth (GRANT & BOWEN, 1998; HARPENDING *et al*., 1998; AVISE, 2000).

Phylogeographic patterns and demographic inferences

The maximum parsimony network effectively recovers the majority of terminals from localities within or near Lake Titicaca in two 2 step clades (2-5, 2-6) (Figs. 2, 3); and validates: (1) the mixed genealogical origin of the shore-lake population from Playa Copani (partitioned in clades 2-5 ["lacustrine"] and 2-7 ["riverine"]), and (2) the possibility that Laguna Saracocha is actually a complex of two currently introgressing populations.

In addition, NCA, provides with statistical support for non-random geographical associations for clades 1-4 (nested within clade 2-5) and 1-7 (nested within clade 2-7), and in both cases, the inference key suggests that restricted gene flow with isolation by distance is the most likely explanation for the genetic structure of these clades (Table 3).

Closer examination of the frequency and geographic distribution of the haplotypes included in clades 1-4 and 1-7 shows that haplotype VI in clade 1-4 is the most frequent as well as most widespread haplotype. In a genealogical context, and according to coalescence theory, the most frequent alleles are also the oldest haplotypes (CRANDALL *et al*., 1994; DONNELLY & TAVARE, 1996) thus the rank of the alleles by age is equal to the rank of alleles by frequency (DONNELLY & TAVARE, 1996). In terms of the NCA, the oldest allele will then be placed at interior nodes, whereas rare or newer haplotypes will be expected to occur at the tips of any particular network. Likewise, older haplotypes will also show a greater number of mutational connections with other haplotypes, and they will be the most widespread geographically (WATTERSON, 1985; EXCOFFIER & LANGANEY, 1989; PFENINGER & POSADA, 2002).

Interestingly, in spite of the apparent longer genetic history of lacustrine populations of clade 1-4 (and by extension clade 2-5), none of these populations show any deep split in the phylogenetic reconstruction shown in Fig. 2, nor show higher levels of gene and nucleotide diversity (Table 4) that are distinctive of older lineages. In fact, lacustrine populations and clade 1-4 in particular is almost one order of magnitude less variable than riverine populations embedded in clade 1-7 (Table 4). Several reasons could lead to depauperate levels of genetic variability in the lacustrine populations, including: non-equal mutation rates between lineages, differences in selective pressures among populations, historically small effective population sizes or recent population size reductions (FAUVELOT *et al*., 2003). The assumption that the star-like topology of clade 1-4 suggests a population bottleneck could be confirmed only if local small effective population sizes throughout the lake are not the outcome of strongly reduced levels of gene flow that otherwise balance out the homogenizing effects of genetic drift (TEMPLETON *et al*., 2001). The result of the AMOVA of clade 1-4 explicitly tests this alternative. There is no significant population structure in "lacustrine" populations; moreover, negative statistics show that haplotypes drawn at random from one population have a higher probability of being identical to haplotypes of another population rather than the original population (EXCOFFIER *et al*., 1992). The AMOVA results are therefore consistent with little population structure and this means that lacustrine frogs are virtually all interconnected by gene flow.

Two other lines of evidence seem to support the occurrence of a population bottleneck in lacustrine populations. The first line of evidence is historical. Lake Titicaca has a record of dramatic oscillations in water and salinity levels throughout the last 20 000 years as shown by palinological evidence (WIRRMANN *et al*., 1991). Such changes have promoted several extinction and recolonization events confirmed by differences in historical and contemporaneous macrophyte communities. Massive macrophyte turnovers undoubtedly impacted animal communities. Remarkably, the most extreme lowstand of Lake Titicaca (100 m below the present water level) occurred very recently; in the mid-Holocene between 3800 and 6000 years ago (BETANCOURT *et al*., 2000). Such a fluctuation effectively reduced 42% of the lake surface and 30% of the water volume (emptying the southern embayment and Puno Bay; WIRRMANN *et al*., 1991), as well as drastically increased salinity levels. The lake only returned to a freshwater ecosystem about 3600 years ago (YBERT, 1991). The second line of evidence is the survey of demographic history by means of the analysis of mismatch

distributions and haplotype and nucleotide diversity. Both approaches convey independent corroboration from tests that are based on entirely different assumptions. First, mtDNA unimodal mismatch distribution of clade 1-4 (Fig. 4A) shows the typical genetic signature of a model of population growth or sudden expansion, in which mutation-genetic drift equilibrium has not been reached (NEE *et al*., 1995; EMERSON *et al*., 2001). Second, significantly negative values for Tajima's D, and Fu's Fs are also highly significant for clade 1-4 (Table 4). Third, the high haplotype diversity relative to nucleotide diversity is also indicative of a population bottleneck followed by rapid population growth (GRANT & BOWEN, 1998; AVISE, 2000).

Contrasting population histories emerge when clade 1-7 and 1-4 are compared. Mismatch distribution for clade 1-7, shows a pattern consistent with a historically stable population size (Fig. 4B). Likewise, a significantly higher "riverine" diversity might result from a significantly longer genetic history, further supported by the results of the phylogenetic analyses. The likelihood that riverine, or more generally, basin *Telmatobius* bypassed the effects of desiccation and subsequent increment of the lake salinity by retracting to higher altitude, mountain-fed streams or lakes is actually reasonable. It is noteworthy that *Isoetes,* a common element of the lake aquatic macroflora 10 000 years ago, disappeared after the drawdown and dramatic peak in salinity. This plant is no longer present in the lake, although is commonly found in almost all other lakes in the surrounding mountains (DEJOUX, 1994). Some other plants like the widespread aquatic cane *Schoenoplectus totora,* appeared in Lake Titicaca only around 3000 years ago. Thus the molecular evidence is consistent with the original interpretation that lacustrine *Telmatobius* populations have recently expanded from a severe bottleneck that ended perhaps just 3600 to 3800 years ago.

In summary, lacustrine haplotypes are more widespread and probably older than riverine haplotypes despite the fact that they appear to be recently founded populations. Lacustrine haplotypes represent a small subset of the overall genetic diversity (HEWITT, 2000; WARES & CUNNINGHAM, 2001). In fact, they are one order of magnitude less genetically variable than riverine populations (Tables 2, 4). At the same time, lacustrine haplotypes nest interior in the NCA network indicating that lacustrine populations are part of a much older clade. This last conclusion only makes sense if a reduction in effective population size resulted in reduced allelic diversity and heterozygosity among previously widespread lacustrine haplotypes (STORZ *et al*., 2002). Apparently, the drought-provoked salinity turnovers of Lake Titicaca had extraordinarily different effects on the otherwise genetically and geographically contiguous riverine and lacustrine populations. At the present time, all Lake Titicaca Basin *Telmatobius* populations continue to be in the process of population growth and geographical range expansion as indicated by the demographic inferences made for higher category clades like 3-3 and 4-2 (Table 4). Some of these recent expansions may explain the apparent mixed origin of some populations as the natural consequence of secondary contact between previously isolated populations.

Acknowledgments

The author thanks Drs. D. McClellan, J.W. Sites, Jr., and Dr. J.C. Ortiz for their academic support, J. Aparicio, and E. Pérez from Colección Boliviana de Fauna (CBF), and Dr. L. Lauzanne (Orstom-Bolivia) for logistical support in the field. M. Morando (BYU) for

NCA insights, and to I. Stehmeier, J. Wells, and J. Marshall for lab assistance. Sequences of *T. marmoratus* specimens from Charazani and Comanche were kindly provided by Ignacio De la Riva, from the Museo Nacional de Ciencias Naturales, Madrid, Spain (MNCN). Financial support was provided by NSF awards DEB 98-15881 (J. Sites) and DEB 03-09111 (J. Sites and E. Benavides), and the Department of Integrative Biology and M.L. Bean Museum at BYU.

REFERENCES

ALTHOFF, D.M. & PELLMYR, O. (2002): Examining genetic structure in a bogus Yucca moth: a sequential approach to phylogeography. *Evolution*, 56: 1632-1643.

AVISE, J.C. (2000): *Phylogeography, the History and Formation of Species*. Harvard University Press, Cambridge, MA.

BENAVIDES, E., ORTIZ, J.C. & SITES, J.W. (2002): Species boundaries among the *Telmatobius* (Anura: Leptodactylidae) of the Lake Titicaca Basin: allozyme and morphological evidence. *Herpetologica*, 58: 31-35.

BETANCOURT, J.L., LATORRE, C., RECH, J.A., QUADE, J. & RYLANDER, K.A. (2000): A 22,000-year record of moonsonal precipitation from northern Chile Atacama Desert. *Science*, 289: 1542-1546.

CEI, J.M. (1986): Speciation and adaptive radiation in Andean *Telmatobius* frogs. Pp. 374-383, *in*: Vuillieumier, F. & Monasterio, M. (eds.), *High Altitude Tropical Biology*. Oxford University Press, Oxford, U.K.

CLEMENT, J., POSADA, D. & CRANDALL, K.A. (2000): TCS, a computer program to estimate gene genealogies. *Molecular Ecology*, 9: 1657-1659.

CRANDALL, K.A.C. & TEMPLETON, A.R. (1994): Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics*, 134: 959-969.

CRANDALL, K.A., TEMPLETON, A.R. & SING, C.F. (1994): Intraspecific phylogenies: problems and solutions. Pp. 273-297, *in*: Scotland, R.W., Siebert, D.J. & Williams, D.M. (eds.), *Models in Phylogeny Reconstruction*. Clarendon Press, Oxford, England.

DE LA RIVA, I. (2002): Rediscovery and taxonomic status of *Telmatobius marmoratus gigas* Vellard, 1969 "1968" (Anura: Leptodactylidae). *Herpetologica*, 58: 220-228.

DEJOUX, C. (1994): Lake Titicaca. Pp. 35-42, *in*: Martens, K., Goddeeris, B. & Coulter, G. (eds.) *Speciation in Ancient Lakes*. Archiv für Hydrobiologie-Advances in Limnology Vol. 44, E. Schweizerbart'sche Verlagsbuchhandlung, Science Publishers, Stuttgart.

DONNELLY, P. & TAVARE, S. (1996): The ages of alleles and a coalescent. *Advances in Applied Probability*, 18: 1-19.

EMERSON B.C., PARADIS, E. & THEBAUD, C. (2001): Revealing the demographic histories of species using DNA sequences. *Trends in Ecology and Evolution*, 16: 707-716.

EXCOFFIER, L. & LANGANEY, A. (1989): Origin and differentiation of human mitochondrial DNA. *American Journal of Human Genetics*, 44: 73-85.

EXCOFFIER, L., SMOUSE, P.E. & QUATTRO, J.M. (1992): Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131: 479-491.

FAUVELOT, C., BERNARDI, G. & PLANES, S. (2003): Reductions in the mitochondrial

DNA diversity of coral reef fish provide evidence of population bottlenecks resulting from Holocene sea-level change. *Evolution*, 57: 1571-1583.

FELSENSTEIN, J. (1981): Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, 17: 368-376.

FETZNER, J. (1999): Extracting high-quality DNA from shed reptile skins: a simplified method. *BioTechniques*, 26: 1052-1054.

FORMAS, J.R., BENAVIDES, E. & CUEVAS, C.C. (2003): A new species of *Telmatobius* (Anura: Leptodactylidae) from Río Vilama, northern Chile, and the redescription of *T. halli* Noble. *Herpetologica*, 59: 253-270.

FU, Y.X. (1997): Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147: 915-925.

FUNK, D.J. & OMLAND, K.E. (2003): Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology and Systematics*, 34: 397-423.

GARCÍA-PARÍS, M., ALCOBENDAS, M., BUCKLEY, D. & WAKE, D.B. (2003): Dispersal of viviparity across contact zones in Iberian populations of fire salamanders *(Salamandra)* inferred from discordance of genetic and morphological traits. *Evolution*, 57: 129-143.

GRANT, S.W. & BOWEN, B.W. (1998): Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, 89: 415-426.

HARPENDING, H. & ROGERS, A. (2000): Genetic perspectives on human origins and differentiation. *Annual Review of Genomics and Human Genetics*, 1: 361-385.

HARPENDING, H.C., BATZER, B.A., GURVEY, M., JORDE, L.B., ROGERS, A.R. & SHERRY, S.T. (1998): Genetic traces of ancient demography. *Proceedings of the National Academy of Sciences USA*, 95: 1961-1967.

HEWITT, G.M. (2000): The genetic legacy of the quaternary ice ages. *Nature*, 405: 907-913.

HILLIS, D.M. & BULL, J.J. (1993): An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, 42: 182-192.

HUELSENBECK., J.P. & RONQUIST, F. (2001): MrBayes: Bayesian inference of phylogeny. *Bioinformatics*, 17: 754-755.

MADDISON, W. (1997): Gene trees in species trees. *Systematic Biology*, 46: 523-536.

MASTA, S.E., LAURENT, N.M. & ROUTMAN, E.J. (2003): Population genetic structure of the toad *Bufo woodhousii*: an empirical assessment of the effects of haplotype extinction on nested clade analysis. *Molecular Ecology*, 12: 1541-1554.

MORITZ, C., SCHNEIDER, C.J. & WAKE, D.B. (1992): Evolutionary relationships within the *Ensatina escholtzi* complex confirm the ring species interpretation. *Systematic Biology*, 41: 273-292.

NEE, S., HOLMES, E.C., RAMBAUT, A. & HARVEY, P.H. (1995): Inferring population history from molecular phylogenies. *Philosophical Transactions of the Royal Society of London B*, 349: 25-31.

NEI, M. (1987): *Molecular Evolutionary Genetics*. Columbia University Press, New York.

PALUMBI, S.R. (1996): Nucleic acids I: the polymerase chain reaction. Pp. 205-247, *in*: Hillis, D.M., Moritz, C. & Mable, B.K. (eds.), *Molecular Systematics,* 2nd ed. Sinauer Associates, Sunderland, MA.

PFENNINGER, M. & POSADA, D. (2002):

Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution*, 56: 1776-1788.

POSADA, D. & CRANDALL, K.A. (1998): Modeltest: testing the model of DNA substitution. *Bioinformatics*, 14: 817-818.

POSADA, D., CRANDALL, K.A. & TEMPLETON, A.R. (2000): GeoDis, a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, 9: 487-488.

ROGERS, A.R. & HARPENDING, H. (1992): Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, 9: 552-569.

SCHNEIDER, S.D., ROESSLI, D. & EXCOFFIER, L. (2000): *Arlequin vers. 2.0: A Software for Population Genetic Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

SLATKIN, M. & HUDSON, R.R. (1991): Pairwise comparisons of mitochondrial DNA secuences in stable and exponentially growing populations. *Genetics*, 129: 555-562.

STORZ, J.F., RAMAKRISNAN, U. & ALBERTS, S.C. (2002): Genetic effective size of a wild primate population: influence of current and historical demography. *Evolution*, 56: 817-829.

SULLIVAN, J.P., LAVAQUE, S. & HOPKINS, C.D. (2002): Discovery and phylogenetic analysis of a riverine species flock of African electric fishes (Mormyridae: Teleostei). *Evolution*, 56: 597-616.

SWOFFORD, D.L. (2000): PAUP*: *Phylogenetic Analysis Using Parsimony (*and other Methods)*. Sinauer Associates, Sunderland, MA.

TAJIMA, F. (1983): Evolutionary relationships of DNA sequences in finite populations. *Genetics*, 105: 437-460.

TAJIMA, F. (1989): Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123: 585-595.

TAMURA, K. & NEI, M. (1993): Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10: 512-526.

TEMPLETON, A.R. (2002): Out of Africa again and again. *Nature*, 416: 45-51.

TEMPLETON, A.R. & SING, C.F. (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.

TEMPLETON, A.R., BOEWRWINKLE, E. & SING, C.F. (1987): A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*, 117: 343-351.

TEMPLETON, A.R., CRANDALL, K.A. & SING, C.F. (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., ROUTMAN, E. & PHILLIPS, C.A. (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., ROBERSTON, R.J., BRISSON,

J. & STRASBURG, J. (2001): Disrupting evolutionary processes: the effect of habitat fragmentation on collared lizards in the Missouri Ozarks. *Proceedings of the National Academy of Sciences USA*, 10: 5426-5432.

TRUEB, L. (1979): Leptodactylidae frogs of the genus *Telmatobius* in Ecuador with the description of a new species. *Copeia*, 1979: 714-733.

WARES, J.P. & CUNNINGHAM, C.W. (2001): Phylogeography and historical ecology of the north Atlantic intertidal. *Evolution*, 55: 2455-2469.

WATTERSON, G.A. (1985): The genetic divergence of two populations. *Theoretical Population Biology*, 27: 298-317.

WIENS, J.J. (1993): Systematics of the leptodactylidae frog genus *Telmatobius* in the Andes of Northern Peru. *Occasional Papers of the Museum of Natural History, University of Kansas*, 162: 1-76.

WIRRMANN, D., YBERT, J.P. & MOURGUIART, P. (1991): Una evaluación paleohidrológica de 20 000 años. Pp. 61-67, *in*: Dejoux, C. & Iltis, A. (eds.), *El Lago Titicaca. Síntesis del Conocimiento Limnológico Actual*. ORSTOM-Hisbol, La Paz, Bolivia.

YBERT, J.P. (1991): Los paisajes lacustres antiguos según el análisis palinológico. Pp. 69-79, *in*: Dejoux, C. & Iltis, A. (eds.), *El Lago Titicaca. Síntesis del Conocimiento Limnológico Actual*. ORSTOM-Hisbol, La Paz, Bolivia.

MOLECULAR PHYLOGENETICS OF THE LIZARD GENUS *MICROLOPHUS* (SQUAMATA: TROPIDURIDAE): ALIGNING AND RETRIEVING INDEL SIGNAL FROM NUCLEAR INTRONS

EDGAR BENAVIDES[1], REBECCA BAUM[2], DAVID MCCLELLAN[1], AND JACK W. SITES, JR.[1]

[1]*Department of Integrative Biology, Brigham Young University, Provo, Utah, 84602, USA. Email: eb235@email.byu.edu*

[2]*Department of Chemistry, Brigham Young University, Provo, Utah, 84602, USA.*

Corresponding author: Edgar Benavides, Department of Integrative Biology, Brigham Young University, Provo, UT 84602; ph: 801/422-2203; fax: 801/422/0090; email: eb235@email.byu.edu

Abstract.—Most recent advances of alignment techniques have focused on improving score-based heuristic algorithms, and most studies align multi-locus, multi-genome data sets in ways that do not explicitly differentiate among gene regions that vary in tempo and class of mutational events (i.e., point vs. small indel vs. large length mutations). We present a multi-gene data set (14 gene regions) to test phylogenetic relationships in the South American "lava lizards" (genus *Microlophus*), and describe a strategy for aligning ribosomal, protein-coding and non-coding sequences that accounts for these differences. We show that seven nuclear introns show considerable variation in size and frequencies of multi-base length mutations, which can be derived as insertions and deletions. Heuristic alignment programs ignore this distinction and provide ambiguous solutions regardless of the range of gap penalties used. We employed a manual alignment strategy to identify large indels across the entire data matrix, independent of any input tree, and a tree-based progressive-alignment (PRANK) algorithm that uses outgroups to distinguish insertions and deletions. We show that while manual alignment is successful in finding the most obvious phylogenetically informative indels, many regions can only be subjectively aligned, and there are obvious limits to the size of a data matrix for which this approach can be taken. PRANK alignment identified more phylogenetically informative indels, and improved overall alignments by reducing the number of homoplasious characters in the aligned nucleotide blocks. We show that, while the coded indel data partition does not influence tree topology, it does contribute significantly to nodal support and branch length estimates. We present a well-corroborated phylogenetic hypothesis for *Microlophus* relationships based on the full data set that supports previous hypotheses for two colonizations of the Galapagos Archipelago. Finally, we highlight strongly supported conflict between mitochondrial and nuclear genomes in a portion of the tree, suggesting that the combined data set for this part of the tree provides a topology and levels of support that are misleading. We hypothesize that in this particular region of the tree, the nuclear genome supports a topology that is closer to the species tree, and conflicts occur due to secondary contact of distantly related taxa, suggesting that unique taxonomic relationships in the mtDNA gene tree are the result of hybridization. This last point highlights the value of dense

taxonomic and genetic sampling for teasing apart different aspects of evolutionary processes

Key words: nuclear introns, alignment, length mutations, *Microlophus*, secondary contact, mitochondrial-nuclear conflict, phylogenetics.

Sequence homology statements are key to formulating sound phylogenetic hypotheses, and multiple alignment programs based on either progressive (Thompson et al., 1994 [CLUSTAL]; Edgar, 2004 [MUSCLE]) or consistency-based scoring alignments (Notredame et al., 2000 [T-Coffe]) are widely used by systematists. Improvements such as iterative optimizations to correct errors introduced early in the alignment process (Wallace et. al., 2005) have greatly increased the accuracy and sensitivity of these methods (Thompson et al., 1999; Hickson et al., 2000; Pollard et al., 2004; Wallace et al., 2005; Lunter et al., 2005), but many studies apply a single heuristic alignment algorithm to data sets with drastically different mutation mechanisms and patterns (ribosomal vs. protein-coding sequences), implying that all can be aligned under the same set of assumptions.

Many of the intricacies of quality alignments refer to the appropriate placement of insertion-deletion (indel) events and therefore indel positions with unclear positional homology are usually excluded from subsequent analyses (Lutzoni et al., 2000; Gatesy et al., 1993). Some authors have advocated the use of ambiguously aligned regions and proposed methods to incorporate them into phylogenetic analyses (Wheeler et al., 1995; Lutzoni et al., 2000; Castresana, 2000; Lee, 2001; Geiger, 2002). Other approaches attempt to explore alignment space more thoroughly and they do so by parameter perturbation of some heuristic algorithm in order to identify areas of conflict (Leache and Reeder, 2002; Wiens et al., 2005), or take an alternative conceptual route by using an algorithm developed for "direct optimization" through simultaneous exploration of alignment and tree-search space (Wheeler, 1996; also see Aagesen et al., 2005).

In general, concern for obtaining optimal tree topologies means that issues of optimal sequence alignments may be overlooked in phylogenetic studies, yet different alignment methods and/or parameter combinations can produce different topologies (Morrison and Ellis, 1997; Hickson et al., 2000), which in turn can be the basis of conflicts between trees obtained from different alignments (Cognato and Vogler, 2001). The most pervasive and often ignored effect of suboptimal alignments is likely to be in the incorrect estimation of other parameters of interest, including branch lengths and/or nodal support, which then compromise the use of such phylogenies in comparative studies and/or divergence time estimates.

In our view, the indiscriminate application of a single heuristic alignment method to protein-coding, non-coding, and ribosomal sequences, ignores both algorithm appropriateness (Jeanmougin et al., 1998), and differences in substitution rates and mutation dynamics. For instance, ribosomal gene substitution patterns (indel and point mutations in both stem and loop regions) are tightly constrained (or not) by secondary structure, whereas protein gene substitutions are constrained by reading frame and codon conservation (Li, 1997). In contrast to both of these gene classes, larger length mutations and random point mutations characterize the evolution of intron regions (Belshaw and Bensasson, 2006).

The increased use of novel nuclear gene regions in phylogenetic studies, alone or combined with more traditional markers such as ribosomal and mitochondrial gene regions, will likely raise new theoretical and empirical issues (e.g., parameter and tree optimization [Pagel and Meade, 2004], data partitioning [Castoe et al., 2004; Brandley et al., 2005], model choice [Sullivan and Joyce, 2005], and topological incongruence [Gatesy and Baker, 2005; Phillips et al., 2004)]). Likewise, some investigators have empirically compared performance of different alignment methods (see recent examples in Pollard et al., 2004; Wallace et al., 2005; Creer et al., 2006; Whiting et al., 2006; Kjer et al., 2006), but to our knowledge, a clear distinction of issues affecting alignment procedures across genes that vary in tempo and class of mutational event (point vs. small indel vs. large length mutations) has not been presented. Our goal in this study is to empirically differentiate the alignment of protein coding and ribosomal gene regions from nuclear introns characterized by length mutations of varying sizes and frequencies, and compare phylogenetic hypotheses generated from alignments that make these distinctions with those derived from one widely-used method that does not accommodate the differences in mutation dynamics. We profit from recent theoretical insights for global alignments of complex large-indel substitution patterns (Loytinoja and Goldman, 2005), and use these to recover an improved phylogenetic hypothesis for the South American lizard genus *Microlophu*s.

The South American "Lava Lizards" (Genus Microlophus)

The "lava lizards" (genus *Microlophus*; Tropiduridae) display an unusual geographic distribution among terrestrial vertebrates; the 22 recognized species include nine taxa endemic to the Galapagos Archipelago, and 13 species mostly confined to a linear strip of 5,000 km along rain-shadowed western coastal deserts of South America (Figs. 1 and 2). Monophyly of the genus is well supported (Frost, 1992; Harvey & Gutberlet, 2000; Frost et al., 2001), but the currently recognized *Occipitalis* (Dixon & Wright, 1975; Frost, 1992) and *Peruvianus* groups (Van Denburgh & Slevin, 1913; Frost, 1992; Heise, 1998) are less well supported. The 12 species of the *Occipitalis* group include the aforementioned Galapagos endemics, which represent two independent colonization events from the continent (Lopez et al., 1992; Wright, 1983; Heise, 1998; Kizirian et al., 2004). The remaining three species of this group and the 10 species of the *Peruvianus* group are confined to the continent. This study is the first to include all recognized species and multiple localities of widespread species, and thus improves on earlier studies by encompassing two critical variables of phylogenetic experimental design - the number of taxa and amount of data (Rokas et al., 2005). Our objectives are to: (1) test alternative alignment strategies based on an algorithm that offers realistic considerations to the specific problem of alignment of large length variable mutations typical of nuclear introns; and (2) obtain a well-supported phylogeny for the genus *Microlophus*, which will provide the framework for ongoing studies of inter-island colonization patterns in the Galapagos Archipelago, and patterns of speciation within the *Peruvianus* group on the western edge of South America.

MATERIALS AND METHODS

Taxon Sampling

This study includes 73 terminal taxa, of which 70 are ingroup samples from one or more localities for each of the 22 species of *Microlophus*; Appendix 1 summarizes locality information for each terminal, and provides details on museum vouchers. All localities are plotted in Figures 1 and 2.

Gene Sampling and Laboratory Methods

Genomic DNA was extracted from muscle tissue preserved in 100% ethanol using either a slightly modified version of the procedure of Fetzner (1999), or the Quiagen extraction kit (Quiagen Inc., Valencia, CA). Amplifications were performed (under varying profiles) in 20µl reaction volumes using TaKaRa hotstart *Taq* DNA polymerase and 10X reaction buffer (100 mM Tris-HCl [pH 8.3], 500 mM KCl, 15 mM $MGCl_2$). Fourteen gene regions representing five mitochondrial (cyt-*b*, ND4, 12S, 16S, and $tRNA^{his}$ plus $tRNA^{ser}$) and nine nuclear gene regions (recombination activating gene 1 [Rag-1], nuclear oocyte maturation factor gene [Cmos], α-enolase [Enol], glyceraldehyde-phosphate dehydrogenase [Gapdh], creatin kinase [CK], ribosomal protein 40 [RP40], B-crystallin [Cryba], Alpha-tropomysin [Atrp], and an anonymous [Anon] region), were sampled for all terminals. Only the Rag-1 and Cmos genes are coding genes while the remaining seven are non-coding introns. The primers used for each gene region are summarized in Appendix 2, and specific PCR conditions are given in Appendix 3. The Anon locus was found while optimizing temperature PCR conditions for the Atrp gene; a BLAST search did not match it to any known gene, and it displayed no homoplasy or conflict when used in phylogenetic analyses. It was therefore presumed to contain phylogenetic signal and thus included in further analyses (see Shaw [2002] and Dolman and Moritz [2006] for recent examples).

Purified double-stranded products were used directly in 1/4 volume dideoxy-termination sequencing reactions using BigDye Terminator v3.1 (Applied Biosystems). Both strands were sequenced for all PCR products, and sequences were edited and proofread with Sequencher v. 4.1 (Gene Codes Corp.). Several taxon-specific internal sequencing primers were designed to allow full amplification of length variable introns (Appendix 2). All sequenced gene regions were used as queries in (BLAST) searches of GenBank to confirm homology. All sequences have been deposited in GenBank with accession numbers (Appendix 4 - TO BE COMPLETED) and in TreeBase.

Alignments

Different strategies were used for the alignment of protein-coding regions, ribosomal genes, and for non-coding intron regions. First, protein-coding regions (Cyt-b, ND4 [mtDNA], Cmos, and Rag-1 [nuclear]) were aligned using MUSCLE (Edgar, 2004)

and open reading frames checked with SeAl (Rambaut, 1996). Second, ribosomal (12S, 16S and tRNAs) gene regions were also aligned with MUSCLE and the 'refine' command was used twice to further improve the existing alignment. However, final sequence positional homology was derived by visual adjustment to secondary structure models developed by Van de Peer et al. (1994; [for the 12S gene]); Gutell and Fox (1988; [for the 16S gene]), and Kumazawa and Nishida (1993; [for the tRNAs]). In these cases blocks of sequence were sometimes shifted by a few bases pairs to conserve base identity with blocks inferred to be stem motifs (from the models), thus constraining small indels to appear in loop regions.

Third, the alignment of nuclear indel regions aimed to incorporate information from both point and length mutations that characterizes regions presumed to experience limited or no selective pressures (Thorne et al., 1992; but see Belshaw and Bensasson, 2006; Roy and Gilbert, 2006). All seven introns displayed length variation among ingroup terminals when indels were not included (see Cryba example in Fig. 3), and here we took two approaches to align length-variable introns.

Manual Alignment of Introns.

Manual alignments were implemented under two assumptions. First, all indels represent single mutational events (Graham et al., 2000; Kelchner, 2000), and second, alignments are most parsimonious when indels are placed to conserve blocks of sequence integrity (defined as indel-free regions with maximum base pair identity) in the indel-flanking regions (Britten et al., 2003; Ogden and Rosenberg, 2006). The "single mutation" assumption represents a fundamental difference from any algorithm-based optimization of positional homology (Kelchner, 2000), and we justify this assumption based on recent descriptions of mutational hotspots, secondary structure configurations, and motifs of variable complexity (Graham et al., 2000; Kelchner, 2000). These studies suggest that some introns evolve under structural constraints in a non-random, non-independent fashion in which large length differences between sequences are best explained as single mutational events (Kelchner, 2000; Lohne and Borsch, 2005; Andolfatto, 2005). We manually aligned the seven nuclear introns individually, on the basis of these two assumptions, using the "3 – step" procedure outlined in Fig. 3.

First, we grouped conspecific ingroup terminals side-by-side in the data matrix, which easily permitted identification of blocks of sequence integrity in some regions. In other parts of the sequence for which such blocks were not obvious, we shifted sequences to create indels having identical positional extension – indels of the same length whose insertion into a sequence created additional blocks of sequence integrity on either side of the indel. These are the 'entire' indels described by Graham et al. (2000), and their placement usually reduced/eliminated the number of base substitutions (Fig. 3; step 1).

After completing step 1 for conspecific sequences, we then compared heterospecific blocks side-by-side (independently for ingroup and outgroup terminals), to identify either the same indels, or those for which positional extension differed slightly between different taxa ('overlapping' indels [Graham et al., 2000]; see step 3 below). Heterospecific terminals with identical alignments were grouped together to facilitate subsequent rounds of comparison, in which the more complex overlapping indels could be readily identified. Here sequences were manually shifted to maximize integrity of flanking blocks (Fig. 3; step 2), and these more complex arrangements were usually located after the first step, and always inferred between terminals above the species level. The iterative collapsing of identical alignments identified in steps 1 and 2 is analogous to the population aggregation analysis described by Davis and Nixon (1992), which is used to identify diagnostic character differences between species. Our approach allows easy visualization of blocks of sequence integrity within and across multiple taxa, which at a coarse scale are evident as taxonomically consistent blocks of nucleotide colored columns in the unaligned Cryba intron depicted in the upper panel of Fig. 3.

The third step in manual alignment required a number of alternative options to deal with more complex rearrangements identified in steps 1 and 2. At this step, indel placement was further improved (i.e., the overall alignment made more stringent with respect to nucleotide identity in conserved blocks) by comparing alignments made in steps 1 and 2 to specific intron micro-structural changes first described by Golenberg et al. (1993) and Gu and Li (1995), and later employed in manual alignments by Kelchner (2000), Kelchner and Clark (1997), Borsch et al (2003), and Lohne and Borsch (2005). The use of an *a priori* set of rules further reduces subjectivity of the alignment (Kelchner, 2000; Lohne and Borsh, 2005), and enhances its repeatability (Sanchis et al., 2001). Step

3 emphasizes the identification of simple sequence repeats motifs (SSR), and/or possible inversions, as they are evidence of either slipped-strand mispairing (SSM; Levinson and Gutman, 1987), and/or simple hairpin structures (Kelchner, 2000; Lohne and Borsh, 2005). The following specific "sub-routines" of step 3 assisted in finding the most parsimonious accommodation of complex indels in relationship to full-length sequence matrices.

First, a common observation is the repeat of a neighboring sequence. In some instances, multiple repeats are observed that generate length-variable mononucleotide strings (Fig. 3; step 3.1 [see also positions 265 to 288 for the Ck gene, in online PAUP file; Appendix 6A]). Because mononucleotide length variable strings could be caused by PCR inaccuracies (Kelchner, 2000), we considered them as regions of uncertain homology (Kelchner, 2000; Borsch et al., 2003) and discarded them from the final matrix.

A second observation was the presence of indels that could result from the deletion/insertion of multinucleotide repeats of a neighboring sequence, and here we directed our attention to the location of repeat motifs and their flanking sequences. In the example illustrated in Fig. 3 (step 3.2), two alternative indel placements hide a multinucleotide insertion as a repeat unit (in alternatives 1 and 2), and we positioned the indel to emphasize flanking motif integrity (alternative 3).

Third, indels shared among several terminals and with the same positional extension, were aligned in the same column(s) so long as they shared the same flanking sequences. In cases where sequence composition adjacent to the inferred indel indicates multiple solutions (*i. e.,* alternative indel placements in Fig. 3; step 3.3), all indels were placed in a single arbitrary column so that a single mutational event is inferred (Lohne and Borsch, 2001).

Fourth, in cases of overlapping indels, sequence columns were rearranged to minimize the number of inferred length mutations. In the example illustrated in Fig. 3 (step 3.4), alternatives 1 and 2 require the origin of four and three independent indel rearrangements, respectively, while alternative 3 requires only two.

Single indel events in the context of microstructural changes were assumed to be less frequent than single nucleotide substitutions (Lohne and Borsch, 2001), but we

inserted an additional indel if such placement reduced or eliminated two or more base substitutions among closely adjacent nucleotides (Fig. 3; step 3.5).

In the final step, all aligned introns were concatenated to each other and to the two protein-coding regions to create the final nuclear sequence data matrix (Fig. 3; step 3.6). Although time consuming, the manual alignment based on the premise of primary homology assessment (De Pinna, 1991) provided a reasonable comparative framework for input-tree based computer-generated alignments (Sanchis et al., 2001; Creer et al., 2006). Our protocol emphasizes flanking motif conservation, and thereby reinforces the basic assumptions of length mutations as single evolutionary events (Borsch et al., 2003), and maximization of base pair identity in flanking sequences (Britten et al., 2003; Ogden and Rosenberg, 2006) as a proxy of primary homology (De Pinna, 1991)

Progressive Alignment of Introns

The alternative to the manual alignment of non-coding regions was found in a modification of current progressive alignment algorithms (Thompson et al., 1994; Edgar, 2004; Notredame et al., 2000; Katoh et al., 2000; Keightley and Johnson, 2004; Do et al., 2005) that provides for a better handling of multinucleotide insertion and deletion events. This component of alignment is fundamental in the context of intron sequences (Keightley and Johnson, 2004), and we used the Probabilistic Alignment Kit (PRANK; Loytynoja and Goldman, 2005), which improves in earlier algorithms by explicitly distinguishing insertions from deletions in the alignment process. In most other multiple alignment implementations indels are penalized relative to nucleotide changes, and such arbitrarily chosen penalties might produce highly fragmented sequences with multiple indels and few nucleotide differences, or very few indels coupled with a large number of nucleotide differences. Additionally, insertions and deletions are thought to be drastically different mutational events, and as such they should be handled differently in alignments (Loytynoja and Goldman, 2005). Most multiple alignment programs penalize deletions only where they occur, while single insertion events are penalized multiple times in every alignment iteration between their original occurrence and the root of the guide tree (Loytynoja and Goldman, 2005). The PRANK implementation uses outgroups to explicitly distinguish between deletions and insertions, and it avoids repeated

penalization of insertions through the use of 'storage matrices' and a hidden Markov match/gap scoring algorithm that has theoretical advantages for the correct alignment of intron sequences.

Because this method considers indels as phylogenetic information, the algorithm may be sensitive to the order in which sequences are added (i.e., the initial guide tree topology). We did not consider individual gene guide trees because of the potential bias length mutations might introduce in the topology of the input tree based on and indel-rich intron. We chose to generate a starting tree by concatenating intron fragments that were shown to be unambiguously aligned from multiple algorithms, and implemented a five-step protocol to obtain an "unbiased" input tree (one not influenced by indel placement) for the PRANK alignment protocol (Fig. 4).

First, the matrix used to obtain the guide tree included fragments of the seven nuclear introns that were first purged of positions of uncertain homology, as identified through parameter perturbation. In this step, we used the program SOAP (Loytynoja and Milinkovitch, 2001) to align each intron under different parameters, and then to purge the reference alignment of all unstable blocks (i.e., those sensitive to parameter perturbation). Three alignment programs were used in order to generate these 'conditional' alignments; parameter perturbation was conducted on CLUSTAL W alignments produced with gap-opening penalties ranging from 1 to 20, and gap-extension penalties ranging from 0.1 to 0.5 (4 x 2 combinations). Two additional alignments, obtained with default parameters, were included in the alignment pool (for a total of 10 pooled alignment matrices) including: (1) a tree-based partitioning algorithm coupled with multiple iterations (MUSCLE; Edgar, 2004); and (2) a Bayesian probabilistic sequence alignment using hidden Markov models (HMM, ProAlign; Loytynoja and Milinkovitch, 2003). All blocks or positions supported by less than 95% in the set of ten 'conditional' alignments were excluded from each intron to produce a "purged" final alignment (Fig. 4; step 1).

Second, the seven purged intron sequences were concatenated to each other and to the Cmos and Rag-1 sequences, resulting in a final matrix of 2568 bp (i.e., roughly 54% of the original nuclear region data matrix of 4691 bp.; Fig. 4; step 2), for construction of the input tree.

Third, a Bayesian analysis (10 million generations) of this concatenated data set yielded a single tree, based on a GTR + I + G substitution model that accommodates biases in base frequencies and skewed transition/transversion ratios (common in non-coding sequences, Kelchner, 2000; Fig. 4; step 3). This tree was then used as the input tree for a PRANK alignment of each of the seven introns (Fig. 4; step 4). Unless indicated, all PRANK alignments were made with default parameters and the HKY substitution model. Once alignments for all indels were completed by this procedure, the introns were concatenated with each other, and with the two protein genes to produce the final PRANK-aligned nuclear matrix (Fig. 4; step 5).

To establish *a priori* "baseline" against which to compare both manual and PRANK aligned sequences, we constructed trees from introns aligned by CLUSTAL, with default parameters and gaps coded as missing. These were compared to the "species tree" derived from the total data set (see below). Our goal is not to provide an intensively "bench-marked" baseline on optimal CLUSTAL parameters (see Odgen and Whiting, 2003), but rather to qualitatively evaluate tree topology, nodal support, and indexes of homoplasy in the manual and PRANK aligned sequences, relatively to a widely used alternative.

Phylogenetic Methods

All introns were concatenated with both the nuclear protein coding (Cmos, Rag-1) and the mitochondrial DNA (cyt *b*, 16s, 12s, and tRNAs) data set for phylogenetic analyses. Inferred indels were codified as an additional partition, and their effect on the final topology was recorded through partitioned and overall measures of clade support (ML bootstrap and Bayesian posterior probability.

For phylogeny reconstruction the following partitions were analyzed: (1) 13 gene regions individually (tRNAs [135 bp] excluded); (2) the combined mitochondrial-only (five partitions, including the tRNAs) and nuclear-only (nine partitions) matrices; and (3) a combined matrix of all 14 gene partitions, plus coded indel regions (a fifteenth partition). We use more parameter-rich models here because under-parameterization may result in biased (increased) clade support (Huelsenbeck and Rannala, 2004; Lemmon and

Moriarty, 2004). Maximum parsimony (MP) analyses were run under equal character weighted heuristic searches with 1000 replicates of random addition, and tree bisection and reconnection (TBR) branch swapping, as implemented in PAUP* 4.0b5 (Swofford, 2002). For individual genes the limit of trees saved was set to 10,000, and 100 for concatenated genes. Measures of branch support were obtained by 500 nonparametric bootstrap pseudoreplicates based on five runs of 100 replicates (each with 10 random addition replicates). Replicated Bayesian analyses coupled with Markov Chain Monte Carlo (MCMC) simulations were run for 10 million generations using four incrementally heated Markov chains sampled every 1000 generations, and started from independent random trees. We used DT-ModSel (Minin et al., 2003) for the partitions defined *a priori* in this study because it typically selects simpler (less parameter rich) models, which estimate branch lengths with less error and topologies with the same accuracy as the more complex models selected by hLRTs, AIC, and BIC approaches (as inferred from simulations; see Abdo et al., 2005). Convergence of phylogenetic parameters and stationarity of likelihood scores was assessed using the program Tracer v1.2 (Rambaut and Drummond, 2003).

All analyses were run using the parallel version of Mr. Bayes (v3.04b; Ronquist and Huelsenbeck, 2003) on the Beowulf cluster housed in the Department of Integrative Biology at Brigham Young University. Nodal support was considered strong with bootstrap values > 70% (Hillis and Bull, 1993; but see their caveats) and posterior probability values > 95% (Alfaro et al., 2003; but see Lewis et al., 2005 for caveats), and gene trees recovering strongly supported conflicting nodes are interpreted as real conflict (Wiens, 1998). Conflict between gene trees might be due to differences in coalescent histories, recombination, non-orthology, or pseudogene amplification, and the absence of conflict by these criteria is taken here to imply that these potentially confounding factors have minimal influence on the genealogies of the regions used in this study.

Coding Length Mutations for Phylogenetic Analyses

Single length mutational events inferred from manual and PRANK alignments were coded as individual binary characters following Simmons and Ochoterena (2000). More complex methods (Simmons et al., 2001; Muller, 2006) use multistate coding to

accommodate sequentially evolving gaps, but in our judgment these require assumptions that do not seem justified at present (Muller, 2006), and at the same time they can only be implemented through step matrices currently unavailable for model-based phylogenetic inference. Indel information was incorporated in both MP and Bayesian tree searches, with the Bayesian analyses of indel partitions based on the maximum likelihood model (Markov k) developed for morphological characters (Lewis, 2001). In all cases we assumed that the indel partitions had unequal rates among characters and so we incorporated a gamma distribution (Mk + G). We assessed the effect of including the indel partition, inferred by both manual and algorithm approaches, by using Bremer partitioned support values (PBS) over the preferred Bayesian topologies. We did this using the TreeRot program (Sorenson, 1999) implementing rather intensive heuristic search parameters (addseq = random nreps=1000 swap= tbr hold=10).

We are fully aware of the limitations of the decay index in general (DeBry, 2001), and do not interpret any of these values as support for a given node. Given that individual mitochondrial and nuclear gene tree topologies were congruent within genomes (see Results) but showed significant between-genome conflict (Wiens 1998) at a few nodes, we use PBS here to identify character incongruence localized to specific nodes, because this conflict has significant evolutionary ramifications.

RESULTS

Patterns of Sequence Variability

Fourteen genomic regions were collected for all terminals, including two small $tRNA^{his}$ and $tRNA^{ser}$ genes (135 bp; considered as a single partition hereafter) and after coding for length mutations, a fifteenth partition was added. Both mitochondrial and nuclear genes show moderate to high levels of variation, with ND4 and the tRNAs being the most variable among mitochondrial regions (as estimated by maximum pairwise uncorrected distance values; Table 1), and Cryba the most variable nuclear gene. BLAST searches did not reveal any match to known genes for the Anon locus, and patterns of variation showed that the nuclear intron regions evolve roughly at the same rate as ribosomal genes and two to four times more rapidly than the two nuclear exons (Cmos and Rag-1; Table 1).

Protein and Ribosomal Coding Alignments

Alignment of mitochondrial protein and nuclear protein coding genes (Cyt-*b* and ND4; Cmos and Rag-1, respectively) was facilitated by conservation of the codon reading frames. Alignment of the ribosomal 12s, 16s, and $tRNA^{his}$ + $tRNA^{ser}$ regions identified conserved blocks that upon comparison with secondary structure models, allowed the identification of stem and loop partitions. Regions of questionable homology – commonly found in loops – were excluded; nine of 10 excluded regions in the 12s gene corresponded to loop regions (80 bp in total), and four of the five excluded regions in the 16s gene were located in regions identified as loops (32 bp in total). Only four tRNA base positions were deemed ambiguously aligned and removed from phylogenetic analyses.

Manual Indel Alignments

We inferred a total of 115 length mutation events (across all taxa) in the seven intron regions (summarized in Table 2), including Enol (6), Gapdh (11), CK (11), RP40 (14), Cryba (51), Atrp (10), and Anon (12). Appendix 5 lists the number of indel events and the type of length mutation recorded for each gene. The frequency of the indel occurrence varies broadly, with Enol and Cryba being the least and most variable introns, respectively. The total number of base pairs inferred as length substitutions across ingroup and outgroup taxa may exceed the total number of nucleotides in the complete gene. For example, the length of the Cryba locus varied from 185 (*M. stolzmanni*) to 839 bp (*T. plica*), and the manual alignment revealed 51 separate indel events (Table 2) from which the majority were length deletions (45), followed by insertions (5), and a single simple sequence repeat (SSR) insertion (indel 15; positions 429-431; Appendix 5 and Appendix 6A. Large length modifications are also apparent in the RP40 intron; here the original intron sequence length (357 bp) has been increased by a insertion that characterizes terminals belonging to *M. occipitalis* (continent) and Eastern Galapagos species *M. habeli* and *M. bivittatus* (Marchena and San Cristobal Islands). This insertion (447 bp), sequenced through the use of additional primers over the same PCR templates, has since recorded novel point and length mutations that further characterize the two Eastern Galapagos species (indels 71 and 72 [Appendix 5], and positions 1572 and 1732 [Appendix 6A]).

On a single gene basis, we were unable to discern nucleotide positional homology in two gene regions (CK and Cryba). In the first case, CK presents a SSR trinucleotide repeat (TCC) between positions 76 to 95 (these and the positions identified below can be viewed in the PAUP file [available online] in Appendix 6A). The motif appears to be evolving randomly in the *Peruvianus* group and has reached maximum length (six repeats) only in southern populations of *M. thoracicus* (indel 3, Appendix 5). The CK alignment shows two continuous momonucleotide strings (cytosine and guanines) between positions 267 to 288) in most taxa except *Tropidurus* and *M. thoracicus* terminals that have independently lost this repeat. Homononucleotides can be considered as stepwise indels (Lohne and Borsh, 2005), but we ignored this region (it was coded as missing data) because we could not reject the possibility that the variable length string repeat may have resulted from inaccurate *Taq* enzymatic activity during PCR (Kelchner, 2000). In the second case, the Cryba intron shows an insertion of a non-repeat 8 bp segment in the *Occipitalis* group (positions 1001 to 1007) that has no apparent positional homology with remaining ingroup or outgroup taxa, and it was also considered as missing data.

Additionally, we found evidence for both independent and non-independent origin of overlapping deletions in related clades in the Cryba intron. Non-independent overlapping indels were inferred in several clades: (*M. occipitalis* + (*M. bivittatus* + *M. habeli*)); positions 582 to 602, and positions 1026 to 1070 [these and all positions identified below can be viewed in Appendix 6A]); (*M. theresiae* LY + (*M. theresiae* Ai + *M. theresiae* K514S)); positions 631 to 637), and (Western Galapagos + (Eastern Galapagos + *M. occipitalis));* positions 2088 to 2126). On the other hand, indels that show positional overlap in unrelated taxa (i.e., putative independent origin) were less frequent (e.g., *M. duncanensis* and *M. occipitalis*; Cryba positions 424 to 431; or *M. stolzmanni* and Eastern Radiation positions 1909 to 1947).

In many cases extensive variation between the ingroup and outgroup forced indel placements that produced some blocks of putatively homologous sites that were dis-continuous with blocks that showed the same degree of nucleotide identity. Two general cases were recorded; first, tentatively aligned segments may result due to high divergence between ingroup and outgroup taxa (see for example positions 762 to 773 and 991 to

997). In the second case, a particular segment of outgroup sequence that is included within a larger indel in the ingroup (e.g., positions 790 to 811; 825 to 829; 837 to 859; 871 to 880; or 974 to 983) can only be shifted within the boundaries of the larger indel, limiting the number of possible homology statements and hence causing the overmatch (Cline et al., 2002) of outgroup blocks with the only available ingroup positions within the larger indel.

PRANK Alignments

To capitalize on the PRANK's ability to distinguish among insertions and deletions, the alignment of all intron sequences was optimized under the same guiding tree obtained from the 'filtered' unambiguous alignment of the nine nuclear regions (2568 bp in total; see the Material and Methods section). The input tree (not shown) successfully recovered nodes at deep and intermediate levels of divergence, but more nested nodes were not fully resolved. Because PRANK ignores unresolved nodes we used a neighbor-joining tree based on uncorrected distances of this same data set. The NJ tree is fully resolved and completely congruent with the topology resolved in the 'filtered' tree. PRANK allows the use of simpler models of substitution (Jukes-Cantor [JC] and Hasegawa-Kishino-Yano [HKY]) to accommodate models previously selected by DT-ModSel for the manually aligned partitions. Overall, results of the progressive alignment generally contain more indels than the manual solution (up to 131 coded gaps; Table 2 and Appendix 6B [PAUP file online available]); more specifically the number of indel characters after the pairwise alignment remained the same for two genes (Enol, and RP40), decreased for one (Anon), increased slightly in three (Gapdh, Atrp, and CK), and was substantially higher in the most variable nuclear intron (Cryba; Table 2). The quality of the alignment was effectively influenced by the use of sequence substitution parameters corresponding to the JC or HKY models, and we illustrate this difference with the alignment of the Cryba intron (see below).

On a general basis, PRANK found one third of the indels that we inferred manually (38 out of 115; Appendix 5 but also produced alternative solutions that accommodated hard-to-align regions in more parsimonious solutions than we recovered (e.g., pos 185 to 196 in online Appendixes 6A and 6B); these alignments increased the number of identical base pairs (2482 to 2571) and decreased the number of parsimony

informative sites (Table 2). On an individual gene basis, PRANK increased the number of identical base pairs in the alignment from 248 to 254, and the number of indels from 11 to 15, while decreasing slightly the number of parsimony informative sites from 84 to 76 (Table 2). Specific differences were the rearrangement of the tri-nucleotide simple sequence repeat (SSR) of positions 76 to 95 (these and all positions identified below can be viewed in Appendix 6B), and single base pair insertions restricted to the outgroup taxa (positions 191 and 275). Other arrangements included insertions in one or a few related terminals for the alignment of the guanine mononucleotide repeat (positions 269, and 278-279).

The Cryba Intron as an Example

The alignment of the Cryba intron involved the most numerous changes regarding manual alignment (Table 2) and, unlike previous intron alignments, it also underscored the need for a thoughtful selection of a substitution model and the correction parameters implemented in PRANK. Our best alignment (i.e., highest likelihood score and the one that maximizes the number of invariant base pairs) was obtained when using the HKY model (kappa=1.830 and empirical base frequencies 0.229/0.276/0.230/0.263 for A/C/G/T), and forcing insertions to be always skipped.

Compared to the manual solution, our best alignment significantly reduced the number of parsimony informative sites (from 204 to 178; Table 2), increased the number of invariable positions (from 583 to 657; Table 2; implying fewer point mutations after indel optimization), and retained higher likelihood scores (Table 3). PRANK succeeded in opening a number of gaps and insertions already present in the manual approach (positions 429 to 431 vs. positions 436 to 438; Appendix 7A and 7B, respectively), but refined the manual alignment by recovering more parsimonious solutions for the location of indels. For example, PRANK recovered blocks with primary homology that were previously overlooked (positions 632 to 635 and 642 to 656; Appendix 6B), and maintained, single base indels as unmatched, instead of either not opening these at all or creating small-length gaps with no phylogenetic consistency (compare positions 804-828 [Appendix 7A] and 784-810 [Appendix 7B], and similar examples in positions 687 to 706; 1072 to 1073; 1142, and 1150 to 1151; Appendix 6B). This alignment also

overturned inferred non-independent overlapping indels, for solutions that increased the number of identical nucleotides in neighboring indel positions (e.g., positions 1086 to 1134 [1026 and 1070 in Appendix 7A]). Our best alignment kept also unmatched the three manually inferred insertions of indels 15 (positions 437 to 439 in Appendix 6B), 53, and 58 (Table 2), but most importantly, it distributed the decaying sequence fragments of *M. stolzmanni* (185 bp vs. 839 bp) in a fully homologous solution, without modifying the PRANK default values for the gap extension probability (gaprate) and the gap opening rate (gaptext) options.

To help illustrating the results of PRANK progressive alignments, we re-aligned the Cryba intron under CLUSTAL (default parameters), MUSCLE (default), and PRANK (default parameters); and compared these alignments to see what influence these might have on tree topology, nodal support, and indexes of homoplasy. Figure 5 illustrates four gene trees (plus the manual alignment of Cryba), and shows that the PRANK (default) topology (Tree D; TL = 330 steps; -ln L = -3543.619; CI = 0.8077; RI = 0.9668) is the only one to recover the *Occipitalis* and *Peruvianus* groups as monophyletic. Gene trees are longer and show lower likelihood values under the CLUSTAL alignment (tree A; TL = 563; -ln L = -3730.23; CI = 0.8202; RI = 0.9591), and these values improve successively through the MUSCLE (tree B; TL = 393; -ln L = -3672.755; CI = 8219; RI = 0.9624), and the manual alignments (tree C; TL = 359 steps; -ln L = -3568.648; CI = 0.8106; RI = 0.9631). Additionally, the alignment of Cryba can still be improved by considering both, unequal base frequencies and a transition/transversion ratio (HKY model of substitution) and the use of an unbiased guide tree. This alignment produces a topology that is 126.2 likelihood units better (Table 3), 18 steps shorter (TL = 312) and has lower homoplasy indexes (CI = 0.8020; RI = 0.9668) than any of the previous options and hence is the alignment we choose for further use.

Phylogenetic Analyses - Mitochondrial DNA

A single alignment hypothesis based on either codon or secondary structure was used for all mitochondrial partitions. Table 1 summarizes substitution models selected by DT-ModSel as well as the length of used in phylogenetic analyses. Individual Bayesian gene trees were estimated for the four largest mtDNA gene regions (12s, 16s, Cyt-*b*, and

ND4), and Figure 6 shows the Bayesian topology estimated for all mitochondrial regions combined. There is an increase of > 200 log likelihood units when using full structural partitions (12) instead of the five nominal gene partitions regardless of the number of parameters estimated (Table 1). Nine major clades relevant to the major evolutionary and/or phylogenetic issues in the genus are recovered with strong support in the combined mtDNA gene tree, and most of these were also recovered in some or all of the individual gene trees, albeit with varying degrees of support (Fig. 6). These are nodes 3, 4, 15, 28, 30, 31, 32, 50, and 58 in the combined tree (Fig. 6); no conflicting nodes are recovered with strong support in any of the individual gene trees (individual gene trees are shown in Appendix 8A).

Node 3 includes all members of the *Peruvianus* group except *M. thoracicus* (node 58), and node 4 (Fig. 6) includes coastal and inland populations of Chilean taxa, from Caleta Meca, Peru (M. quadrivittatusCM) to Tres Playitas, Chile (M. atacamensisTPL; Fig. 1). Within this node two subclades are particularly well supported (nodes 5 and 8), but some relationships within each of these are poorly resolved, and in node 8 heterospecific terminals are better supported (nodes 11 and 13) than conspecifics. Node 15 includes several species of the *Peruvianus* group, but does not recover strongly supported haploclades for either *M. peruvianus* or *M. tigris*, and terminals of *M. quadrivittatus* are inter-digitated with other taxa (*M. tigris* and *M. heterolepis*, nodes 22 and 25, respectively). The combined mtDNA data set do not resolve the polytomy between the three subclades of node 15, and while strongly supporting node 28, *M. theresiae*, its placement in the context of the genus phylogeny is less certain (node 14 is weakly supported). Node 58 (*M. thoracicus*) is strongly supported across all gene regions, but its placement within the genus tree is uncertain (node 2 is weakly supported). Within *M. thoracicus*, three well resolved haploclades of northern (node 60), central (node 63), and southern Peru terminals (node 64) are strongly supported.

In contrast to the results for the *Peruvianus* group, the monophyly of the *Occipitalis* group is recovered with strong support, and in several of the individual gene trees. Within the *Occipitalis* group, three other clades are notable. First, the "Eastern Galapagos" clade hypothesized in earlier studies is strongly supported in our combined tree (node 50), and the mainland species *M. occipitalis* is recovered as the sister group to

the clade of Galapagos endemics (*M. habeli* and *M. bivittatus*) from Marchena and San Cristobal islands, respectively. External to this clade, there is only weak support for the positions of the mainland species *M. koepckeorum* and *M. stolzmanni* (nodes 48 and 49). Second, all remaining species hypothesized to comprise the *Occipitalis* group are recovered in a strongly supported clade that includes the seven species of the "Western Galapagos" radiation proposed in earlier studies (node 31; Wright, 1983; Lopez et al., 1992; Heise, 1998; Kizirian et al., 2004), and *M. delanonis* is strongly supported as the sister species to the remaining taxa in this clade (node 32). Although support is high for monophyly of conspecific terminals in all cases, the polytomy of node 32 prevents a full understanding of the relationships between species of the Western Galapagos radiation on the basis of mitochondrial haplotypes only.

Phylogenetic Analyses - Nuclear DNA

Table 1 summarizes the substitution models selected by DT-ModSel, and the size of each nuclear gene partition used in the phylogenetic analyses. Individual Bayesian gene trees were estimated for 13 partitions of the nine nDNA gene regions aligned under both criteria and topologies were largely congruent among them. Support for major clades by gene region as well as concatenated data sets is summarized in Appendix 9. Figure 7 shows the Bayesian topology of the combined nuclear data set, and recovers several strongly supported nodes of particular interest. These include the *Peruvianus* group (node 2) and several groups within it, and the *Occipitalis* group (node 35) and several groups within it. Among the single gene trees, all loci but RP40 recover the *Occipitalis* and *Peruvianus* groups (and RP40 does recover the first of these two), and several of the gene trees support these nodes at or close to PP = 1.0 (Appendix 9).

Within the *Peruvianus* group, seven genes recover the *M. thoracicus* clade (node 27, six genes recover *M. theresiae* as the sister taxon of the remaining *Peruvianus* group species (node 3), and three genes recover all other *Peruvianus* group taxa as the sister group of the *M. theresiae* clade (node 4). At a more nested level, northern coastal terminals of *M. peruvianus* are recovered with strong support as sister taxa to paraphyletic northernmost populations of *M. tigris* (LP and SCH; node 18), and this clade is in turn paraphyletic to the remaining terminals of *M. tigris* (CT, DRQ and PU; node 16). The *M. tigris* terminals from southern Peru (CT, DRQ, and PU) are paraphyletic to a

clade now comprised of terminals of the Chilean clade that also includes populations of *M. tigris* (Yura; basal terminal at node 6) and *M. heterolepis* (Rospigliozzi; basal terminal at node 7). These two interior populations of the *M. peruvianus* complex *sensu lato* appear as basal to the Chilean clade and are sequentially related to all interior populations of the Chilean clade (*M. yanezi* [Pocon], and *M. theresioides* [Pica, and Calama]) from which coastal populations of *M. atacamensis* (ISM, COL, TAL, TPL, and DL; node 14 [most]) and *M. quadrivittatus* (IQC, CZ, CM, IL, and CB; nodes 11, 13, 15) seem to have been derived.

Within the *Occipitalis* group there is strong support for *M. stolzmanni* as the sister clade of the two Galapagos radiations (node 36), and *M. occipitalis* as the sister group of the Eastern Galapagos radiation (*M. bivittatus* and *M. habeli*; node 56). Resolution of a monophyletic Western Galapagos radiation is strongly supported (node 37), and as in the combined mtDNA data set, *M. delanonis* is the sister clade to all others in this radiation (nodes 37 and 38). Within this radiation, some heterospecific clades are evident due to non-monophyly *M. albemarlensis* from Isabela (nodes 38) and *M. jacobi* from Santiago (nodes 44 and 46).

Combined Data and a Phylogenetic Hypothesis for Microlophus

Partitioned Bayesian analysis of the combined mtDNA and nuclear DNA (with coded indels) matrix produced a single tree with an average lnL of -42033.714 in the case of introns aligned manually and an average of lnL -42116.161 in the case introns that had been aligned by PRANK (Table 3). Figure 8 displays the topology recovered from this analysis, and Table 4 summarizes different measures of clade support for all nodes in the phylogram, under both manual and PRANK inferred alignments. There are no topology differences between both alignment strategies.

The combined data topology recovers the majority of the clades previously recovered with either the mtDNA or nuclear data sets (see exceptions below). At the deepest levels of divergence, these include the *Peruvianus* and *Occipitalis* groups (nodes 2 and 40, respectively), each of which is strongly supported by both BS and PP values, as well as multiple data partitions (Table 4). Similar levels of support are evident for node three that unites all *Peruvianus* group species except *M. thoracicus*, nodes 5 (Chilean *Microlophus*), 15 (the "*peruvianus*" clade), 30 (*M. theresiae*), and the *M. thoracicus*

clade (node 32). Within the *Occipitalis* group, there is similar support for the Western and Eastern Galapagos clades (nodes 42 and 62, respectively), and in both of the most inclusive clades (nodes 2 and 40), nuclear partitions generally support the deepest splits with various mtDNA partitions frequently in conflict, and this pattern is reversed in more recent divergences (Table 4).

Incorporating nDNA loci had a significant effect in the placement of *M. thoracicus* within the *Peruvianus* group. The mtDNA data weakly support this clade as the sister group to all other terminals of the *Peruvianus* group plus the *Occipitalis* groups (node 2 in Fig. 6), while the nuclear loci independently strongly supported this clade as being sister group to the all *Peruvianus* group species with (node 2 in Fig. 7), a result confirmed by the combined data set (Fig. 8) with evidence of minor conflict due to two mtDNA loci (Table 4). The mtDNA data weakly support the placement of *M. theresiae* as sister to some but not all members of the *Peruvianus* group (node 14 in Fig. 6), while the nDNA (node 3 in Fig. 7) strongly corroborated the combined data set in placement of *M. theresiae* as sister to all remaining *Peruvianus* group terminals but *M. thoracicus* (node 3 in Fig. 8).

We note some topological discordance between mitochondrial and nuclear genomes stemming from: (1) suboptimal node resolution due to markers resolving different temporal scales; and (2) conflict due to significant incongruence between the mitochondrial and nuclear data sets (following the criterion of Wiens, 1998). An example of suboptimal node resolution is the ambiguous placement of *M. stolzmanni* and *M. koepckeorum* within the *Occipitalis* group; this appears to be due to a limited number of characters and weak conflict between data sets. Support for placement of *M. stolzmanni* and *M. koepckeorum* is poor in the mtDNA topology, but stronger for a (*M. koepckeorum* (*M. occipitalis* (Eastern Galapagos))) clade in the nDNA topology. In contrast, the combined data weakly recover the same clade (node 61; Fig. 8), but strongly support (across many partitions; Table 4) *M. stolzmanni* as the sister taxon to the remaining *Occipitalis* group terminals (nodes 40 and 41; Fig. 8).

The most obvious example of strongly supported mtDNA vs. nDNA conflict is found at several clades nested within the *Peruvianus* group; compare the symmetrical mtDNA topology (Fig. 6) with the strongly pectinate nDNA topology (Fig. 7), versus the

combined data that differs slightly from both (Fig. 8). As an example, node 15 is strongly supported in the combined topology, while mtDNA sequences alone recover this clade with weak support (node 14). Likewise, terminals in the strongly supported Chilean *Microlophus* clade in the combined tree are grouped in the mtDNA topology into an arrangement that is discordant with the geographic distribution of some species (especially *M. atacamensis* and *M. quadrivittatus*; node 4 in Fig. 6; see Discussion).

DISCUSSION

Indel Alignment

As larger and more varied molecular data sets become routine in phylogenetic analyses, some studies have begun to focus attention on alignment strategies (Creer et al., 2006; Whiting et al., 2006). The sudden wealth of intron-rich data sets should not be a surprise given that non-coding sequences comprise the vast majority of metazoan genomes and vertebrates in particular (Roy and Gilbert, 2006), but to our knowledge alignment strategies have not explicitly differentiated among genes that vary in tempo and class of mutational events (point vs. small indel vs. large length mutations). For example, in contrast to protein and ribosomal gene regions, there are no sources of independent evidence, i.e., secondary structure or codon conservation, for the correct alignment of non-coding DNA (Pollard et al., 2004). Further, heuristic alignments are obscured by the occurrence of length mutations that can potentially bias any initial guide (input) tree upon which the entire alignment is built or, in the case of manual alignments, make very difficult the search for strings of homologous blocks. We suggest that the use of a filtered tree based exclusively on unambiguously aligned sequences, provides reasonably 'independent' evidence that, despite counter-arguments (Giribet and Wheeler, 1999), yields working hypotheses for explaining some of the dynamics of length mutations that support phylogenetic relationships (Loytynoja and Goldman, 2005).

Aside from the direct optimization approach (Wheeler, 1996) and/or manual alignments, there is no algorithm solution that circumvents the use of a guide tree. In practice, when aligning non-coding regions common heuristic methods always find unambiguously aligned regions first, and then try to optimize gap placement. At the same time most programs relying on such an approach would regard insertions

introduced early in the matching of two obviously similar sequences as a nuisance because, unlike deletions, they are not penalized only at the place where they occur but rather have to be handled in all subsequent iterations. As a result, insertions end up being penalized repeatedly as the algorithm progresses down to the root of the guide tree (Loytinoja and Goldman, 2005). Such common heuristic designs (i.e., implemented in the CLUSTAL, MUSCLE, and T-COFFE programs, among others) do find approximate solutions to the alignment of both point and length mutations, but with the caveat that alternative gap placements are ignored as long as heuristic programs ignore the difference between deletion and insertion events. In reality, there is no way to distinguish insertions from deletions in the absence of phylogenetic information, and for this particular problem the only possible solution is to explicitly recognize outgroup terminals. Further, since the guide tree determines the order in which sequences are progressively added into the alignment, the use of guide trees that are themselves derived from sequences that can have a wide range of length differences (Cryba and RP40 genes in Table 2) do have unpredicted consequences for the inference of phylogenetically informative insertion-deletion (INDELS) events. Thus the alignment of regions that has both length and point mutations often start from wrong assumptions.

In our opinion, two sources of conflict should be addressed explicitly in order to adequately align non-coding genome regions. First, inferring indels must be given increased attention because progressive methods that rely on the construction of an initial guide tree assume that the best information on gap placement will be found among the most similar sequences, when in reality there may be better information in the alignment of all sequences considered in the group of interest. No computer approach is fully capable of correcting a particular alignment based on a global solution, but the first step in this direction should be a program that handles insertions and deletions differently, and this is the advantage of the PRANK HMM implementation. We attempted to make this distinction in our manual alignments, and while partially successful, PRANK alignments inferred more indels, which were more parsimonious solutions in that they required fewer mutational events elsewhere in the same sequences (Table 2). The method of Loytynoja and Goldman (2005) distinguishes insertions and deletions as separate events, and thus provides with a mechanism that allows for the correction of early mistakes in the

placement of gaps (indels) in a progressive alignment context. The innovation of this development is that it does more than simply maximize a similarity score; it rather seeks a solution that is phylogenetically consistent across the entire data matrix.

In effect, most of the indels inferred by PRANK and our manual assessments are phylogenetically consistent (i.e., they recover the same synapomorphies for the same nodes), but perhaps more important is the fact that PRANK aligned regions that were clearly over-aligned either by CLUSTAL, MUSCLE, or manual heuristics, at least for the Cryba intron (Fig. 5). Our manual effort shows that trying to resolve variable nucleotide regions flanked by conserved blocks is a daunting task that requires sliding and partitioning of these blocks in a sometimes highly subjective manner. The CLUSTAL and MUSCLE algorithms also fail because they simply ignore the difference between an indel was caused by an insertion or deletion, and in which sequences. For example, the CLUSTAL alignment of the Cryba intron is characterized by outgroup overmatched sequences that are more or less in the same positions as in our manual alignment. Applying different gap costs do not solve the issue; gap costs set too high result in few or no gaps in the matrix (and hence the overaligned sequences), or if too low the alignment will be extensively fragmented into many small gaps with little or no phylogenetic consistency among sequences (i.e., identical placement of the same indels).

Our manual alignment was partially successful in minimizing the number of variable sites by accommodating more invariant sites. For example, the manual, CLUSTAL, and MUSCLE Cryba alignments find 583, 418, and 550 invariant nucleotide positions respectively, whereas the PRANK alignment finds 657 identical sites (Table 2). We think that the relative success of manual alignment is in fact, result of the moderate divergence within our study group, but it also reflects the utility of an *a priori* set of rules for the manual alignment of introns. One logical extension of these study is that manual alignments, although impractical for large data sets, retain some accuracy, and that PRANK alignments constructed from an unbiased guide tree likely contain fewer homoplasious point mutations while recovering indels that are phylogenetically informative.

Finally, the ability of PRANK to handle sequences characterized by base compositional biases is very important. Implementing different substitution models (JC

and HKY in the version we used), resulted in slightly different alignment outputs. For the Cryba intron the HKY model alignment produced an improved likelihood score by considering the two parameters (unequal base frequencies and ti/tv ratios) that are usually associated with systematic biases and strong non-historical signal in non-coding nuclear DNA (Phillips et al., 2004). Applying modeling techniques that account for such biases in alignment issues presumably allows for a more realistic reconstruction of evolutionary history (Loytynoja and Goldman, 2005), and is an important step forward towards linking alignment relative penalties for substitution and indel events to actual parameter estimates of sequence evolution.

Phylogenetic Information Content of Intron Length Substitutions

Several important observations emerged from our inclusion of the length mutations inferred from the seven nuclear introns included in this study. First, inclusion of both manual and PRANK aligned gap-coded partitions had no effect on the final topology, as indicated by identical topologies obtained for the nuclear data set with and without gap partitions (Fig. 8). Second, inclusion of these partitions resulted in lowered likelihood scores, and the associated need for more intense searches since an entire new partition was added. On the positive side we also note that in the absence of topological differences, the likelihood scores of the nuclear tree inferred using PRANK-aligned sequences is higher than the manually aligned tree (Table 3). Gap-coded partitions collectively contributed positive support to 21 nodes in the combined topology, while contributing no support for 44 nodes, and showing conflict in only 6 nodes (Table 4). Comparing manual and PRANK aligned partitions, these ratios (positive/zero/negative support) are 21/44/6 and 21/40/10, respectively, and these contributions were independent of the differences in the number of parsimony informative sites in both data sets. The total number of parsimony informative characters gained after coding the indel characters from manual and PRANK alignments is 82 and 94, respectively, suggesting that the overall contribution of the gap partitions in the context of the remaining nuclear and the mitochondrial data is surprisingly important (Table 4). With eventual modeling of these kinds of mutations (see above section), coded indel partitions should also

improve estimates of branch lengths as non-coding sequences become more common in phylogenetic studies.

Phylogenetic and Evolutionary Implications

Within the genus *Microlophus*, both the *Occipitalis* (Dixon and Wright, 1975; Frost, 1992) and *Peruvianus* groups (Van Denburgh and Slevin, 1913; Frost, 1992; Heise, 1998) are only moderately supported on the basis of morphological data, and previous molecular studies have not included all recognized species, thereby precluding rigorous independent tests of the monophyly of each clade. In this study, nuclear and the combined data sets always recover both the *Occipitalis* and *Peruvianus* groups with strong bootstrap and PP support, and this support was spread across many data partitions (Table 4). The mtDNA partition corroborates the nDNA data in strong support of the *Occipitalis* group, but it fails to recover the *Peruvianus* group (Fig. 6), a result not likely due to retention of ancestral polymorphisms at this deep node. One alternative is that terminals that are topologically unstable between the mtDNA vs. nDNA data sets may show base compositional bias and therefore be attracted to branches sharing similar base frequencies, rather than ancestry (Wiens and Hollingsworth, 2000). Because the mtDNA data were also the most inconsistent with recognized species boundaries (a result also consistent with base compositional bias), we re-evaluated phylogenetic relationships using a Log-determinant method designed to accommodate this bias, and thus avoid artifactual branch attraction (Jordan and Hewitt. 2004). This analysis of the combined mtDNA data set (Log-det tree not shown) recovered a well-supported monophyletic *Peruvianus* group that places *M. thoracicus* as the sister species to all others in the group, suggesting that compositional bias partially accounts for failure of the mixed-model mtDNA data to corroborate the nuclear genes in support of the *Peruvianus* group.

Elsewhere, the resolution of evolutionary relationships within *Peruvianus* and *Occipitalis* groups depends on the combined effect of all markers. Resolution of relationships among the 12 species recognized within the *Occipitalis* group generally corroborates earlier studies supporting origin of the Galapagos endemics on the basis of two independent colonization events from the continent (Lopez et al., 1992; Wright,

1983; Heise, 1998; Kizirian et al., 2004). Specifically, there is strong support for a small "Eastern Galapagos Radiation" consisting of the sister species *M. bivittatus* and *M. habeli* endemic to the islands of San Cristobal and Marchena, respectively (Fig. 2), and sharing a sister clade relationship to the mainland species *M. occipitalis* (Fig.8). Our results also confirm monophyly of the larger "Western Galapagos Radiation" that includes the remaining seven Galapagos endemics that occupy most of the remaining islands in the archipelago (Fig. 2), and the absence of a close relationship of this radiation to any individual mainland species (Fig. 8). A close examination of the mtDNA data set shows that there is conflict between markers for all nodes that resolve interspecific (inter-island) relationships of the western radiation (nodes 47, 48, 53, and 53; Table 4). Because the mtDNA gene regions are all linked, the source of this conflict must reside in homoplasy or other confounding factors that are not fully accommodated by our methods. The contribution of individual nuclear markers for the same nodes is negligible (Table 4), probably because this radiation is geologically relatively young, but the combined nDNA do resolve the mtDNA polytomy at node 32 (Fig. 6; see node 38 in Fig. 7). We accept the combined data topology (node 42; Fig. 8) as the best available working hypothesis for colonization of the Galapagos Archipelago, and are conducting studies to estimate the timing of both colonization events, and to test alternative inter-island colonization hypotheses for the Western radiation (Benavides et al., in prep.).

Similarly, the combined data tree resolves almost all nodes within the *Peruvianus* group with strong support (Fig. 8), but this topology masks the aforementioned well-supported mtDNA (symmetrical topology at node 3; Fig. 6) vs nDNA conflict (pectinate topology at node 5 in Fig. 7). The nuclear tree is qualitatively "better" in the sense that its topology is more concordant with both currently recognized species and their distributions (neither of which holds for the mtDNA tree), but the match is not exact, and several processes likely are or have contributed to the discordance. In Fig. 9 we have juxtaposed the relevant clades from the mtDNA and nDNA trees, and mapped the distributions of the relevant terminals on the western edge of South America.

Examination of these topologies suggests that one cause for this conflict could be mtDNA introgression between *M. quadrivitattus* from Caleta Ballenita and *M. tigris* from Caleta Meca (terminals quaCB and tigCT in tree B [nuclear] in Fig. 9). These species are

sympatric and barely segregated by habitat (specimens were collected for this study < 0.5 km apart), and secondary contact with mtDNA introgression could explain why *M. quadrivitattus* is recovered with a geographically overlapping terminal of *M. tigris*, yet remains a clearly distinct clade that more closely matches current species boundaries on the basis of the nDNA data. Three nuclear markers (Cryba, CK, Gapdh) support a phylogenetic pattern of differentiation of the Chilean *Microlophus* (nodes 12 to 15 in Fig. 9), via derivation from an ancestor shared within a paraphyletic *M. tigris* - *M. heterolepis* – *M. yanezi* – *M. theresioides* group (nodes 6 to 11 in Fig. 9). The nDNA topology further implies that *M. atacamensis* diverged via dispersal through the Rio Loa valley to the coast, and then continued south, and as *M. quadrivittatus* dispersed north until overlapping and hybridizing with *M. tigris* at a point of secondary contact in southern Peru (Fig. 9).

The combined data (Fig. 8) support the mtDNA topology, but some nodes (9 and 10 in Fig. 8) show relatively high PP values associated with low bootstrap values (Table 4). Although PP values are generally thought to be less biased estimators of confidence (Alfaro et al., 2003; Erixon et al., 2003), relatively high PP values for short internodes (nodes 5 to 14 in Fig. 8) have been shown to be clear overestimates, especially if they show conflict with other support values (Alfaro et al., 2003; Brandley et al., 2005). It is noteworthy that the nuclear signal for this alternative arrangement is not only provided by point substitutions but also provided by length mutations (nodes 8 to 14; Table 4). We suggest that the combined data topology of the Chilean *Microlophus* clade (Fig. 8) is both misleading and its support is overestimated, and that, based on the match between the nDNA topology and current taxonomy, that this alternative (Fig. 7) offers the best working hypothesis for this region of the tree. Paraphyly of some taxa in the nDNA tree (B in Fig. 9) suggests that recent speciation, some mixing or incomplete sorting of nuclear haplotypes, poorly defined species, or some combination of these processes has contributed to the geographic complexity implied by this topology, but detailed population studies are in progress to clarify the dynamics of this interaction (Benavides et al.; in progress)

CONCLUSION

There are challenges in any alignments of large, multi-gene data sets, but we suggest that a clear distinction be made among the classes of gene regions used, alignments should be based on a strict consideration of mutational patterns, and the use of length mutations based on methods tailored to their unique features. On a more operational level, selection of an optimal guiding tree should not be seen as a circular practice because there is no alternative to objectively differentiate insertions and deletions, and this distinction enhances the quality of the alignment by reducing homoplasy and increasing phylogenetic signal. In summary, greater effort should be devoted to the alignment of indels (also see Metzler, 2003; Miklos et al., 2004; Holmes, 2005; Redelings and Suchard, 2005; Fleisner et al., 2005; Ogden and Rosenberg, 2006 for new developments) with an eye to solutions that are phylogenetically consistent and incorporating their mutational dynamics into branch length estimates.

ACKNOWLEDGMENTS

We thank Heidi & Howard Snell, and the staff of the Charles Darwin Research Station and the Galapagos National Park service for logistical support in Galapagos; L. Coloma, M.T. Rodrigues, H. Snell, and J. C. Ortiz, A. Catenazzi, S. Kelez, and J. Cordoba for tissue loans; F. Torrez, P. Victoriano, R. Grams, and M. Vidal for field work in Chile and Peru. In Provo, the authors thank P. Alsbury, I. Stehmeier, and J. Wells. Lizards were collected in the field under permits issued by INRENA-Peru (TUPA 0113-2002-AG); Parque Nacional Galápagos Ecuador (PT 7.5 FR 28 to H. Snell) and Servicio Agrícola y Ganadero in Chile (Permit 3112 to J. C. Ortiz and M. Vidal). EB was supported by various units of Brigham Young University, including the Department of Integrative Biology and the M.L. Bean Life Science Museum, BYU graduate mentoring and graduate research fellowships, a Larsen Scholarship, and a B.F. Harrison Scholarship. Additional research support was provided by a graduate research award from the Society of Systematic Biologists, and NSF awards DEB 0309111 (doctoral dissertation improvement award to J.W. Sites, Jr. and E. Benavides), and EF 0334966 ("Assembling the Tree of Life: The Deep Scaly Project: Resolving Higher Level Squamate Phylogeny Using Genomic and Morphological Approaches") to T. Reeder, M. Kearney, J. Wiens, and J.W. Sites, Jr.

.

REFERENCES

Aagesen, L., G. Peterson, and O. Seberg. 2005. Sequence length variation, indel costs, and congruence in sensitivity analysis. Cladistics 21:15-30

Abdo, Z., V. Minin, P. Joyce, and J. Sullivan. 2005. Accounting for uncertainty in the tree topology has little effect on the decision theoretic approach to model selection in phylogeny estimation. Mol. Biol. Evol. 22:691–703.

Alfaro, M. E., S. Zoller, and F. Lutzoni. 2003. Bayes or Bootstrap?. A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstraping in assessing phylogenetic confidence. Mol. Biol. Evol. 20:255-266.

Arévalo, E., 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 (Sauria, Phrynosomatidae) in central Mexico. Syst. Biol. 43:387-418.

Andolfatto, P. 2005. Adaptive evolution of non-coding DNA in *Drosophila*. Nature 437: 1149-1152.

Belshaw, R., and D. Bensasson. 2006. The rise and fall of introns. Heredity 96:208-213.

Borsch T., K. W. Hilu, D. Quandt, V. Wilde, C. Neinhuis, and W. Bartlott. 2003. Noncoding plastid *trnT-trnF* sequences reveal a well resolved phylogeny of basal angiosperms. J. Evol. Biol. 16:558-576.

Brandley, M. C., A. Schmitz, and T. W. Reeder. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Syst. Biol. 54:373-390.

Britten, R. J., L. Rowen, J. Williams, and R. A. Cameron. 2003. Majority of divergence between closely related DNA samples is due to indels. Proc. Natl. Acad. Sci. USA 100:4661-4665.

Castresana. J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17:540-552.

Castoe, T.A., T. M. Doan, and C. L. Parkinson. 2004. Data partitions and complex models in Bayesian analysis: the phylogeny of gymnophthalmid lizards. Syst. Biol. 53:448-459.

Cline, M., R. Hugheyook, and K. Karplus. 2002. Predicting reliable regions in protein

sequence alignments. Bioinformatics. 18:306-314.

Cognato, A. I., and A. P. Vogler. 2001. Exploring Data Interaction and Nucleotide Alignment in a Multiple Gene Analysis of *Ips* (Coleoptera: Scolytinae). Syst. Biol. 50:758-780.

Creer, S., C. E. Pook, A. Malhotra, and R. S. Thorpe. 2006. Optimal intron analyses in the *Trimeresurus* radiation of Asian pitvipers. Syst. Biol. 55:57-72.

Davis, J. I., and K. C. Nixon. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. Syst. Biol. 41:421-435.

DeBry, R.W. 2001. Improving the interpretation of the decay index for DNA sequence data. Syst. Biol. 50:742-752.

De Pinna, M. C. C. 1991. Concepts and tests of homology in the cladistic paradigm. Cladistics 7:367-394.

Dixon, J. R., and J. W. Wright. 1975. A review of the lizards of the iguanid genus *Tropidurus* in Peru. Contributions in Science of the Natural History Museum of Los Angeles County 271:1-39.

Do, C.B., M.S.P., Mahabhashyam, M. Brudno, and S. Batzoglou. 2005. PROBCONS: Probabilistic Consistency-based Multiple Sequence Alignment. *Genome Res.* 15: 330-340.

Dolman, G., and B. Phillips. 2004. Single copy nuclear DNA markers characterized for comparative phylogeography in Australian wet tropics rainforest skinks. Mol. Ecol. Notes 4:185-187.

Dolman, G., and C. Moritz. 2006. A multi-locus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). Evolution 60:573-582.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Erixon, P., B. Svennblad, T. Britton, and B. Oxelman. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst. Biol. 52:665–673.

Fetzner, J. 1999. Extracting high-quality DNA from shed reptiles skins: a simplified method. BioTechniques 26:1052-1054.

Fleissner, R., D. Metzler, and A. von Haeseler. 2005. Simultaneous statistical multiple

alignment and phylogeny reconstruction. Syst. Biol. 54:548-561.

Friesen V. L., B. C. Congdon, H. E. Walsh, and T. P. Birt. 1997. Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. Mol. Ecol. 6: 1047-1058.

Friesen, V. L., B. C. Congdon, M. G. Kidd, and T. P. Bird. 1999. Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrates. Mol. Ecol. 8:2147-2149.

Frost, D. 1992. Phylogenetic analysis and taxonomy of iguanian lizards (Reptilia:Squamata). Am. Mus. Novitates 3033:1-68.

Frost, D. R., M. T. Rodrigues, T. Grant, and T. A. Titus. 2001. Phylogenetics of the lizard genus *Tropidurus* (Squamata: Tropiduridae: Tropidurinae): direct optimization, descriptive efficiency, and sensitivity analysis of congruence between molecular data and morphology. Mol. Phylogenet. Evol. 21:352-371.

Gatesy, J., R. DeSalle, and W. Wheeler. 1993. Alignment-ambiguous nucleotide sites and the exclusion of systematic data. Mol. Phylogenet. Evol. 2:152-157.

Gatesy J., and R. H. Baker. 2005. Hidden likelihood support in genomic data: can forty-five wrongs make a right? Syst. Biol. 54: 483-492.

Geiger, D. L. 2002. Stretch coding and block coding: two new strategies to represent questionably aligned DNA sequences. J. Mol. Evol. 54:191–199.

Giribet, G., and W. C. Wheeler.1999. On gaps. Mol. Phylogenet. Evol. 13:132-143.

Golenberg, E. M., M. T. Clegg, M. L Durbin, J. Doebley, and D. P. Ma. 1993. Evolution of a non-coding region of the chloroplast genome. Mol. Phylogenet. Evol. 2:52-64.

Graham, S. W., and R. G. Olmstead. 2000. Evolutionary significance of an unusual chloroplast DNA inversion found in two basal angiosperm lineages. Curr. Genet. 37:183-188.

Graham, S. W., P. A. Reeves, A. C. E. Burns, and R. G. Olmstead. 2000. Microstructural changes in noncoding chloroplasts DNA: interpretation evolution and utility of indels and inversions in basal angiosperm phylogenetic inference. Int. J. Plant. Sci. 161:S83-S96.

Gu, X., and W-H. Li. 1995. The size distribution of insertions and deletions in human and

rodent pseudogenes suggests the logarithmic gap penalty for sequence alignment. J. Mol. Evol. 40:464-473.

Gutell, R. R., and G. E. Fox. 1988. A compilation of large subunit RNA sequences presented in structural format. Nucleic Acids Res. 16:r175-r269.

Harvey, M. B., and R. L. Gutberlet Jr. 2000. A phylogenetic analysis of tropidurine lizards (Squamata: Tropiduridae), including new characters of squamation and epidermal microstructure. Zool. J. Linn. Soc. 128:189-233.

Heise, P. J. 1998. Phylogeny and biogeography of Galapagos lava lizards (*Microlophus*) inferred from nucleotide sequence variation in mitochondrial DNA. A dissertation presented for the Doctor of Philosophy Degree. The University of Tennessee, Konxville, Tennessee. 206 p.

Hickson, R. E., C. Simon and S. W. Perrey. 2000. The Performance of Several Multiple-Sequence Alignment Programs in Relation to Secondary-Structure Features for an rRNA Sequence. Mol. Biol. Evol. 17:530–539.

Hillis, D. M., and J. J. Bull. 1993. An empirical testing of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42:182-192.

Holmes, I. 2005. Using evolutionary expectation to estimate indel rates. Bioinformatics 21: 2294-2300.

Huelsenbeck, J. P., and B. Rannala. 2004. Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. Syst. Biol. 53:904-913.

Janke, A., D. Erpenbeck, M. Nilsson, and U. Arnason. 2001. The mitochondrial genomes of the iguana (*Iguana iguana*) and the caiman (*Caiman crocodylus*): implications for amniote phylogeny. Proc. Biol. Sci. 268:623-631.

Jeanmougin, F., J. D. Thompson, M. Gouy, D. G. Higgins, and T. J. Gibson. 1998. Multiple sequence alignment with CLUSTAL X. Trends Biochem. Sci. 23:403–405.

Jordan, B. H., and G. M. Hewitt. 2004. The origin and radiation of Macaronesian beetles breeding in *Euphorbia*: the relative importance of multiple data partitions and population sampling. Syst. Biol. 53:711-734.

Katoh, K., K. Misawa, K. Kuma, and T. Miyata. 2002. MAFFT: a novel method for

rapid multiple sequence alignment based on fast Fourier transform. Nucleic. Acids. Res. 30:3059-3066.

Kelchner, S. A. 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. Ann. Missouri Bot. Gard. 87:482-498.

Kelchner, S. A., and L. G. Clark. 1997. Molecular evolution and phylogenetic utility of the chloroplast *rpl*16 intron in *Chusquea* and the Bambusoideae (Poaceae). Mol. Phylogenet. Evol. 8:385-397.

Keightley, P. D., and T. Johnson. 2004. MCALIGN: stochastic alignment of non-coding DNA sequences based on an evolutionary model of sequence evolution. Genome Research 14:442-450.

Kjer, K. M., Gillespie, J. J., and K. A. Ober. 2006. Opinions on multiple sequence alignment, and an empirical comparison of repeatability and accuracy between POY and structural alignment. Syst. Biol. (in press).

Kizirian, D., A. Trager, M. A. Donnelly, and J. W. Wright. 2004. Evolution of Galapagos Island Lizards (Iguania: Tropiduridea: *Microlophus*). Mol. Phylogenet. Evol. 32: 761-769.

Kumazawa, Y., and M. Nishida. 1993. Sequence evolution of mitochondrial tRNA genes and deep branch animal phylogenetics. J. Mol. Evol. 37:380-398.

Leache, 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.

Lee, M. S. Y. 2001. Unalignable sequences and molecular evolution. Trends Ecol. Evol. 16:681–685.

Lemmon, A. R., and E. C. Moriarty. 2004. The importance of proper model assumption in Bayesian phylogenetics. Syst. Biol. 53:265-277.

Levinson, G., and G. A. Gutman. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Bio. Evol. 4:203-221.

Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50:913-925.

Lewis, P.O., M.T. Holder, and K.E. Holsinger. 2005. Polytomies and Bayesian and phylogenetic inference. Syst. Biol. 54:241-253.

Li, W-H. 1997. Molecular Evolution. Sinauer, Sunderland, MA. USA

Loytynoja, A., and M. C. Milinkovitch. 2001. SOAP. Cleaning multiple alignments from unstable blocks. Bioinformatics 17: 277-286.

Loytynoja, A., and M. C. Milinkovitch. 2003. ProAlign, a probabilistic multiple alignment. Bioinformatics 19: 1505-1513.

Loytynoja, A., and N. Goldman. 2005. An algorithm for progressive multiple alignment of sequences with insertions. Proc. Natl. Acad. Sci. USA 102:10557-10562.

Lohne, C., and T. Borsch. 2005. Molecular evolution and phylogenetic utility of the petD group II intron: A case study in basal angiosperms. Mol. Biol. Evol. 22:317-322.

Lopez, T. J., E. D. Hauselman, L. R. Maxson, and J. W. Wright. 1992. Preliminary analysis of phylogenetic relationships among Galapagos Island Lizards of the genus *Tropidurus*. Amphibia-Reptilia 13: 327-339.

Lunter, G., A. J. Drummond, I. Miklos., and J. Hein. 2005. Statistical alignment: Recent progress, new applications, and challenges. Pages 382-411 *in* Statistical Methods in Molecular Evolution, 1st edition (R. Nielsen, ed.). Springer, New York, New York.

Lutzoni F, P. Wagner, V. Reeb, and S. Zoller. 2000. Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. Syst Biol. 49:628–651.

Metzler. D. 2003. Statistical alignment based on fragment insertion and deletion models. Bioinformatics 19: 490-499.

Miklos, I., G. A. Lunter, and I. Holmes. 2004. A "long-indel' model for evolutionary sequence alignment. Mol. Biol. Evol. 21: 529–540.

Minin, V., Z. Abdo, P. Joyce, and J. Sullivan. 2003. Performance-based selection of likelihood models for phylogeny estimation. Syst. Biol. 52:674-683.

Metzler. D. 2003. Statistical alignment based on fragment insertion and deletion models. Bioinformatics 19: 490-499.

Morrison, D. A., and J. T. Ellis. 1997. Effects of Nucleotide Sequence Alignment on Phylogeny Estimation: A Case Study of 18s rDNAs of Apicomplexa. Mol. Biol. Evol. 14:428-441.

Müller K. 2006. Incorporating information from length-mutational events into

phylogenetic analysis. Mol. Phylogenet. Evol. 38:667-676.

Notredame, C., D. G. Higgins, and J. Heringa. 2000. T-coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302: 205–217.

Odgen, T. H., and M. Whiting. 2003. The problem with "the Paleoptera problem:" sense and sensitivity. Cladistics 19:432-442.

Ogden, T. H., and M. S. Rosenberg. 2006. Multiple sequence alignment accuracy and phylogenetic inference. Syst. Biol. 55:314-328.

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 Associates, Sunderland, MA.

Pagel, M., and A. Meade. 2004. A phylogenetic mixture model for detecting pattern heterogeneity in gene sequence or character state data. Syst. Biol. 53:571–81.

Phillips M. J., F. Delsuc. and D. Penny. 2004. Genomescale phylogeny and the detection of systematic biases. Mol. Biol. Evol. 21:1455–58.

Pollard, D. A., C. M. Bergman, J. Stoye, S. E. Celniker, and M. B. Eisen. 2004. Benchmarking tools for the alignment of functional noncoding DNA. BMC Bioinformatics 5: 6.

Rambaut, A. 1996. Se-Al: Sequence Alignment Editor. Available at http://evolve.zoo.ox.ac.uk/.

Rambaut, A., and A. J. Drummond. 2003. Tracer v1.2, http:// evolve.zoo.ox.ac.uk/.

Redelings, B. D., and M. A. Suchard. 2005. Joint Bayesian Estimation of Alignment and Phylogeny. Syst. Biol. 54:401–418.

Rokas, A., D. Kruger, and S. B. Carrol. 2005. Animal evolution and the molecular signature of radiations compressed in time. Science 310:1933-1938.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Roy, S.W., and W. Gilbert. 2006. The evolution of spliceosomal introns: patterns, puzzles, and progress. Nat. Rev. Genet. 7:211-221.

Saint, K. M., C. C. Austin, S. C. Donnellann, and M. N. Hutchinson. 1998. C-mos, a nuclear marker useful for squamate phylogenetic analysis. Mol. Phylogenet. Evol. 3:240-247.

Sanchis, A, J. M. Michelena, A. Latorre, D. L. J. Quicke, U. Gardenfors, and R. Belshaw. 2001. The phylogenetic Analysis of variable-length sequence data: elongation factor–1a introns in European populations of the parasitoid wasp genus *Pauesia* (Hymenoptera: Braconidae: Aphidiinae). Mol. Biol. Evol. 18:1117-1131.

Shaw, K. L. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proc. Natl. Acad. Sci. USA 99: 16122-12127.

Simmons, M. P., and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49:369-381.

Simmons, M. P., H. Ochoterena, and T. G. Carr. 2001. Incorporation, relative homoplasy, and effect of gap characters in sequence-based phylogenetic analyses. Syst. Biol. 50:454-462.

Sorenson, M. D. 1999. TreeRot, version 2. Boston University, Boston, MA.

Sullivan, J., and P. Joyce. 2005. Model selection in phylogenetics. Ann. Review of Ecol. and Syst. 36:445-466.

Swofford, D.L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods), Beta Version 4.0.b5b. Sinauer, Sunderland, MA.

Thompson, J.D., D.G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673–4680.

Thomson, J.D., Plewniak, F., and Poch, O. 1999. A comprehensive comparison of multiple sequence alignment programs, Nucleic Acids Res. 27: 2682–2690.

Thorne, J. L., H. Kishino, and J. Felsenstein. 1992. Inching towards reality: an improved likelihood model of sequence evolution. J. Mol. Evol. 34:3-16.

Van Denburgh, J., and J. R. Slevin. 1913. Expedition of the California Academy of Sciences to the Galapagos Islands, 1905-1906. IX. The Galapagoan lizards of the genus *Tropidurus* with notes on iguanas of the genera *Conolophus* and *Amblyrhynchus*. Proc. California Acad. Sci., ser. 4, 2:132-202.

Van de Peer, Y., I. Van den Broek, P. de Rijk, and R. de Watcher. 1994. Database on the structure of small ribosomal subunit RNA. Nucleic Acids Res. 22: 3488-3494.

Wallace, I. A., O. O'Sullivan, and D. G. Higgins. 2005. Evaluation of iterative alignment algorithms for multiple alignment. Bioinformatics 21:1408-1414.

Wheeler, W., J. Gatesy, and R. DeSalle. 1995. Elision: A method for accommodating multiple molecular sequence alignments with alignment-ambiguous sites. Mol. Phylogenet. Evol. 4:1–9.

Wheeler, W. C. 1996. Optimization alignment: The end of multiple sequences alignment in phylogenetics? Cladistics 12:1–9.

Whiting A. S., A. M. Bauer, and J. W. Sites, Jr.. 2003. Phylogenetic relationships and limb loss in Sub-Saharan South African scincine lizards (Squamata:Scincidae). Mol. Phylogenet. Evol. 3:582-592.

Whiting, A.S., J.W. Sites, Jr., K.C.M. Pellegrino, and M.T. Rodrigues. 2006. Comparing alignment methods for inferring the history of the New World lizards genus *Mabuya* (Squamata: Scincidae). Mol. Phylogenet. Evol. 38:719-730.

Wiens, J. J. 1998. Combining data sets with different phylogenetic histories. Syst. Biol. 47: 568–581.

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 jarrovii*). Evolution 53:1884-1897.

Wiens, J. J., and B. D. Hollingsworth. 2000. War of the iguanas: conflicting molecular and morphological phylogenies and long-branch attraction. Syst. Biol. 49:143-159.

Wiens, J. J., T. W. Reeder, J. W. Fetzner, C. L. Parkinson, and W. E. Duellman. 2005. Hylid frog phylogeny and sampling strategies for speciose clades. Syst. Biol. 54:719-748.

Wright, J. W. 1983. The evolution and biogeography of the lizards of the Galapagos Archipelago: evolutionary genetics of *Phyllodactyllus* and *Tropidurus* populations. In: Bowman, R. I. Et al. (Eds.), Patterns of Evolution in Galapagos Organisms. AAAS Symposium Volume, San Francisco, USA, pp. 123-155.

LIST OF FIGURES

original indel positional differences in left panel (overlapping indels), and subsequent placement of indels to improve sequence base-pair identify of flanking regions. 3.1. Identification of length-variable mononucleotide strings of uncertain positional homology (in the third and fourth sequences); these were eliminated from the matrix. 3.2. Alternative indel placements and motif recognition (alignments 1 and 2); only the third alternative reflects a simple sequence repeat (SSR) event that is consistent with hypothesized mechanisms. 3.3. Sequence composition adjacent to a indel that could not resolve the position of a putative indel; in alternative 1 the indel is placed in three different blocks and sequence similarity is preserved, in alternative 2 the indel is placed arbitrarily but it ensures inference of a single mutational event across all ingroup taxa (under the assumption that it represents a single mutational event). 3.4. Length mutations may overlap with one another to create a series of overlapping step indels, but here only the third alternative reduces the number of possible mutations that create overlapping indels. 3.5. In this example two indels (alternative 2) instead of one (alternative 1) are inserted to eliminate the two inferred base substitutions in the flanking sequence required by inference of a single indel. 3.6. Manually aligned introns are concatenated to each other and the two coding regions (Cmos and Rag-1), to produce the nuclear gene matrix.

Figure 4. The software-assisted alignment to obtain input tree for PRANK analyses. 1. Selection of unambiguously aligned blocks (in gray) through parameter perturbation of multiple alignments of each intron using SOAP and a 95% cutoff level (see text for details). 2. Concatenation of unambiguously aligned intron blocks plus nuclear coding genes (Cmos, Rag-1; 2568 bp). 3. Building of a guide tree (Bayesian) based on concatenated sequences from step 2. 4. Use of PRANK with the input tree obtained in step 3, and used to align each of the seven introns. 5. Concatenation of PRANK aligned introns plus coding regions (Cmos and Rag-1) to produce final nuclear sequence matrix.

Figure 5. Resulting Bayesian and MP support values (Posterior probabilities on top; bootstrap values below) from phylogenetic analyses of four alignment hypotheses of the Cryba intron; A. CLUSTAL with default parameters (TL = 563; CI = 0.8202; RI = 0.9591); B. MUSCLE with default parameters (TL = 393; CI = 0.8219; RI = 0.9624); C. Manual alignment (TL = 359; CI=0.8106; RI = 0.9631); and D. PRANK progressive alignment with JC model of substitution (TL = 330; CI = 0.8030; RI = 0.9631). White, gray and black bars represent terminals from interior Chile, *M. theresiae*, and *M. thoracicus* respectively and help differentiate tree topologies.

Figure 6. 50% majority rule consensus of Bayesian MCMC trees from five mtDNA gene regions analyzed as 12 partitions (see Table 1 for the best-fit model to each partition); nodes are numbered below branches (enlarged node numbers are discussed in the text), and posterior probability values are given above branches.

Figure 7. 50% majority rule consensus of Bayesian MCMC trees from nine nDNA gene regions analyzed as 13 partitions (see Table 1 for the best-fit model to each partition); nodes and support values are numbered as in Fig. 6.

Figure 8. 50% majority rule consensus phylogram of 9001 Bayesian MCMC trees (10 million generations) from 15 data partitions for the combined data sets (all gene regions plus the coded indels); nodes and support values are numbered as in Figs. 6 and 7. Topologies resulting from manual (115 indels; 7682 characters total) and PRANK (131 indels; 7902 characters) alignments are identical; Table 4 summarizes several measures of support for all nodes under both alignments, and the "*" identify nodes with PP $\geq$ 0.99.

Figure 9. Conflicting mtDNA (tree A, re-drawn from node 3 in Fig. 6) and nuclear DNA (tree B, from node 4 in Fig. 7) topologies (with node numbers and PP values as in the original Figures. The map enlarges the southern Peru – northern Chile region,

and shows the approximate distributions of currently recognized species, and possible routes of dispersal/speciation in *M. atacamensis* (ata) from *M. theresioides* (the) implied (with caveats) by the nDNA topology. Note the strongly supported sister-group relationship between terminals (quaCB + tigCT) at node 25 in the mtDNA tree (A), versus strong support for discordant placement of these terminals in the nDNA topology (nodes 15 and 17 in B).

TABLE 1. Models of evolution inferred from DT-ModSel for *a priori* partitions discussed in the text. Also shown is the number of characters included for each partition used in the final Bayesian mixed model analyses. All introns were aligned manually and with the progressive approach implemented in the probabilistic alignment kit software (PRANK; indicated by "*"). The two strategies sometimes differ in the length of the final alignment as well as in the model of substitution selected by DT-ModSel. Estimates of maximum uncorrected sequence divergence were calculated across all taxa for single genes only.

Data partitions	Number of characters	DT-ModSel substitution models	% Pairwise sequence divergence
mtDNA			
ND4 all	662	TrN +Γ +G	27.08
ND4 1st position	221	HKY + G	
ND4 2nd position	221	HKY + Γ + G	
ND4 3rd position	220	GTR + G	
Cyt-b all	829	TrN + Γ + G	24.12
Cyt-b 1st position	277	K80 + G	
Cyt-b 2nd position	276	HKY + Γ	
Cyt-b 3rd position	276	TrN + Γ + G	
12s all	774	GTR + Γ + G	16.10
12s stems	408	TVMef + Γ + G	
12s loops	365	TVMef + Γ + G	
16s all	479	GTR + Γ + G	14.81
16s stems	213	HKY + Γ + G	
16s loops	267	SYM + Γ + G	
tRNA (Ser + His)	135	TrN + G	45.56
tRNA stems	73	K80 + G	
tRNA loops	62	K81uf + G	
nDNA			
Anon	415	K80+G	15.39
Anon*	415	K80+G	15.39
Atrp	258	TrN+G	16.82
Atrp*	260	TIM + G	15.75
Ck	358	K80 + G	14.62
Ck*	363	TrNef + G	14.77
Cryba	841	K81 + G	30.35

Cryba*	897	K81+G	23.08
Enol	284	K80	12.00
Enol*	284	K81+G	12.00
Gapdh	341	HKY + G	13.69
Gapdh*	342	HKY + G	13.56
RP40	806	K80 + G	14.91
RP40*	806	K80 + G	14.91
Rag-1 all	801	K80 + Γ	4.23
Rag-1 1st position	281	K80	
Rag-1 2nd position	280	K80	
Rag-3 3rd position	280	K80 + G	
Cmos all	547	K80	6.49
Cmos 1st position	183	K80	
Cmos 2nd position	182	K80	
Cmos 3rd position	182	K80	
Gaps-Manually inferred	117	Mk + G	56.17
Gaps-Prank inferred	131	Mk + G	56.82

TABLE 2. Summary of indel variability in the seven length-variable nuclear introns inferred by manual (M) and PRANK (PR) alignments, and sequence parameters obtained from each. Ti/Tv columns refer to the transition/transversion ratios obtained under both alignment strategies, as well as the numbers of parsimony informative characters and the numbers of invariant sites across all terminals.

Gene region	# indels and parsimony informative indels (in parentheses)		# characters after alignment		Sequence size range (bp)	Ti/Tv (κ)		# parsimony informative sites		# invariant sites	
	M	PR	M	PR		M	PR	M	PR	M	PR
Enol	6 (5)	6 (5)	284	284	147 -- 284	1.815	1.792	36	36	224	224
Gapdh	11 (4)	16 (6)	341	342	289 -- 341	2.367	2.292	56	55	256	258
Atrp	10 (4)	13 (5)	258	260	237 -- 248	2.344	2.384	42	41	197	203
Anon	12 (10)	11 (11)	415	415	360 -- 414	1.896	1.896	97	97	309	310
RP40	14 (7)	14 (7)	806	806	327 -- 765	1.572	1.572	87	87	665	665
Cryba	51 (39)	56 (38)	841	897	185 -- 839	1.830	2.020	204	178	583	657
CK	11 (9)	15 (13)	358	363	310 -- 355	3.626	3.277	84	76	248	254
Total	115(78)	131 (85)	3303	3367	1855 3246			606	570	2482	2571

TABLE 3. Summary statistics of Bayesian MCMC runs of separate and combined mitochondrial (including the tRNA partition [135 bp]) and nuclear gene regions considered in this study. The number of substitution model parameters includes total tree lengths, and * identifies substitution model parameters obtained from introns aligned with PRANK. Complete data sets including introns aligned with PRANK and manually aligned are available online.

Gene regions	# of partitions	# of estimated parameters	- ln L ± SD	- ln L median
mtDNA				
total	5	62	-24519.783 ± .232	-24519.383
total	12	110	-24257.128 ± 0.27	-24256.777
Cyt-*b*	3	26	-8980.807 ± 0.235	-88990.538
ND4	3	26	-7040.242 ± 0.235	-7039.915
16s	2	21	-2780.426 ± 0.410	-2884.465
12s	2	25	-4185.293 ± 0.691	-4184.836
nDNA				
Anon	1	8	-1548.278 ± 0.269	-1547.746
Anon*	1	8	-1639.186 ± 0.294	-1638.744
Atrp	1	13	-1005.092 ± 0.218	-1004.725
Atrp*	1	13	-988.44 ± 0.176	-988.06
CK	1	8	-1285.173 ± 0.183	-1284.834
CK*	1	8	-1404.889 ± 0.182	-1404.889
RP40	1	8	-2361.236 ± 0.207	-2360.951
Cryba	1	13	-3543.619 ± 0.69	-3543.219
Cryba*	1	8	-3417.017 ± 0.204	-3470.725
Enol	1	7	-924.098 ± 0.299	-923.841
Enol*	1	7	-932.782 ± 0.255	-932.782
Gapdh	1	8	-1309.269 ± 0.194	-1308.827
Gapdh*	1	8	-1294.639 ± 0.169	-1294.639
RAG	3	18	-1955.559 ± 0.317	-1954.98
Cmos	3	17	-1406.969 ± 0.35	-1406.816
nDNA total				
(4691 bp)	13	79	-15462.596 ±0.299	-15462.194
(4755 bp)*	14	79	-15394.903 ±0.339	-15394.484
nDNA + indels				

(4806 bp)	13	80	-16065.554 ±0.333	-16065.071
(4886 bp)*	14	80	-16113.739 ± 0.318	-16113.468
mtDNA + nDNA				
(7680 bp)	25	187	-41420.862 ±0.335	41420.391
(7746 bp)*	25	187	-41367.878 ±03.85	-41367.447
combined + indels				
(7797 bp)	26	187	-42033.714 ±0.367	-42033.170
(7877 bp)*	26	188	-42116.161 ±0.364	-42115.712

TABLE 4. Summary of bootstrap (BS; run only for PRANK alignments and shown only if > 50), posterior probabilities (PP), and partitioned Bremer support values for clades recovered through Bayesian MCMC searches for all partitions combined; 15 nominal partitions were considered for the calculation of Bremer supports, including: five mtDNA regions (2991 bp), nine nuclear regions (4691 [manual] and 4755 [Prank] bp), and coded indels ("gaps"; n = 115 [manual] and 131 [PRANK]). The first and second lines of values for each node indicate manual and PRANK alignments, respectively; bolded nodes are discussed in the text and highlighted in Fig. 7.

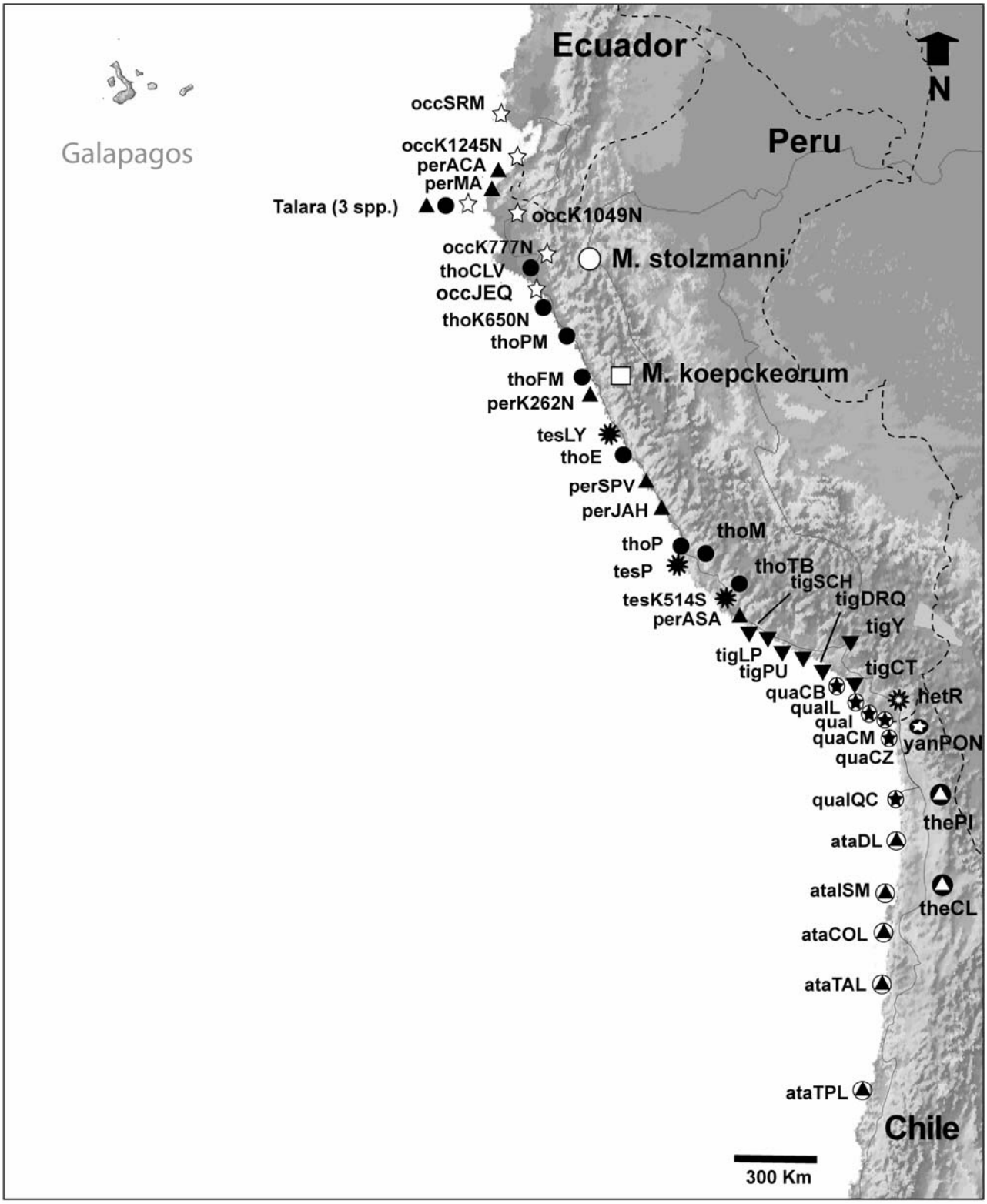
#	Node Description	Partitioned Bremer support																
		BS	PP	12s	16s	ND4	tRNA	Cyt-*b*	Ck	Cryba	RP40	Enol	Gapd	Anon	Atrp	Cmos	Rag-1	Gaps
1	Genus *Microlophus*		1.0	17	12	9	6	17	1	0	1	1	-1	-1	-1	0	6	7
		100	1.0	17	14	8	7	19	1	-1	2	1	-2	-3	-3	1	8	14
2	***Peruvianus* group**		**1.0**	**0**	**1**	**1**	**-1**	**-1**	**5**	**1**	**1**	**1**	**2**	**4**	**1**	**4**	**3**	**-1**
		100	**1.0**	**0**	**1**	**1**	**-2**	**-2**	**6**	**4**	**1**	**2**	**1**	**6**	**2**	**5**	**4**	**2**
3	**Node 4 + node 30**		**1.0**	**5**	**0**	**-2**	**-1**	**5**	**1**	**12**	**2**	**4**	**3**	**2**	**1**	**1**	**1**	**1**
		100	**1.0**	**3**	**-1**	**-4**	**-1**	**7**	**2**	**12**	**2**	**4**	**4**	**3**	**3**	**1**	**1**	**1**
4	**Node 55 + node 15**		**1.0**	**-3**	**-1**	**3**	**-3**	**-1**	**4**	**2**	**1**	**0**	**0**	**3**	**1**	**2**	**2**	**1**
		87	**1.0**	**-5**	**-1**	**3**	**-3**	**0**	**6**	**3**	**1**	**0**	**1**	**4**	**1**	**3**	**2**	**2**
5	**Node 9 + 6 (Chilean *Microlophus*)**		**1.0**	**25**	**10**	**29**	**4**	**29**	**1**	**3**	**1**	**0**	**0**	**0**	**0**	**0**	**1**	**1**
		100	**1.0**	**32**	**14**	**34**	**6**	**38**	**1**	**2**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	**1**
6	M. atacamensisISM + node 7		1.0	3	0	2	1	-2	0	0	0	0	0	0	0	0	0	0
		98	1.0	3	0	3	1	-3	1	1	1	0	1	1	0	0	0	-1
7	M. atacamensisTAL + node 8		.81	0	-1	1	0	0	0	0	1	0	0	-1	0	0	0	0

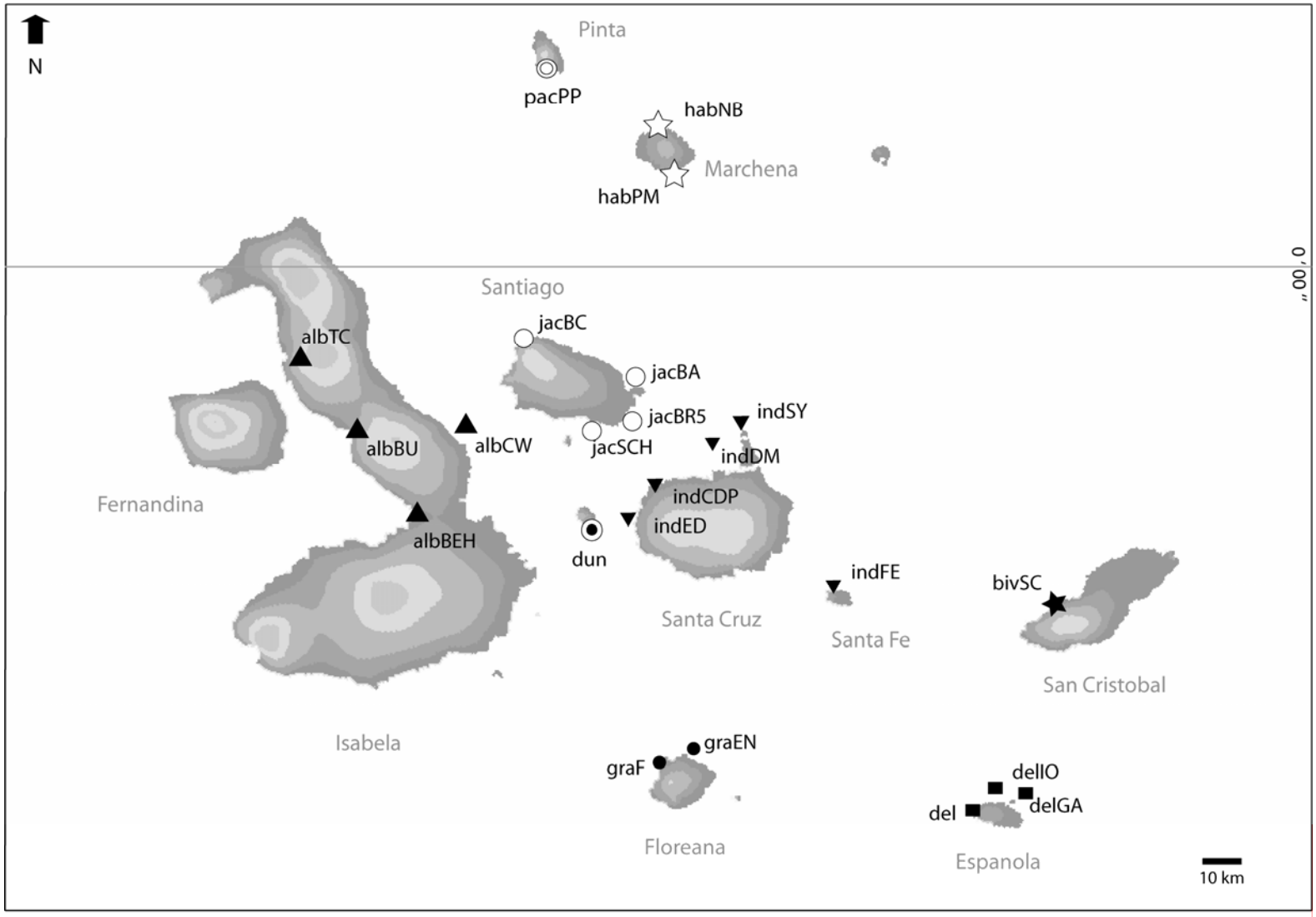
			.80	-2	0	0	0	2	0	0	0	0	0	0	0	0	0	0
8	M. atacamensisTPL + COL		.79	3	0	0	0	-2	0	0	0	0	0	0	0	0	0	0
			.77	2	0	0	0	-2	0	0	0	0	0	0	0	0	0	0
9	Node 10 + node 14		1.0	-1	1	3	0	-2	0	0	0	0	0	0	0	0	0	1
		93	.98	-2	1	4	0	-2	0	1	0	0	0	0	0	0	0	1
10	M. quadrivittatusIQC +node 11		1.0	1	0	1	0	-2	0	1	0	0	0	0	0	0	0	1
		65	.98	-2	0	0	0	1	0	1	0	0	0	0	0	0	0	1
11	M. atacamensisDL + node 12		1.0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
			.97	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
12	M. yaneziPON + node 13		1.0	0	0	0	0	-1	0	1	0	0	0	0	0	0	0	1
			1.0	-2	0	0	0	1	0	1	0	0	0	0	0	0	0	1
13	M. theresioidesCL + M. thePi		1.0	1	0	2	0	-1	0	0	0	0	-1	0	0	0	0	0
			1.0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
14	M. quadrivittatusCM + CZ		1.0	2	0	2	0	2	0	0	0	0	0	0	0	0	0	0
		100	1.0	1	0	3	0	5	0	0	0	0	0	0	0	0	0	0
15	**“peruvianus”**		**1.0**	**16**	**7**	**16**	**-0**	**12**	**-2**	**-3**	**-1**	**0**	**-2**	**-1**	**0**	**-1**	**0**	**-1**
		100	**1.0**	**21**	**10**	**21**	**0**	**17**	**-3**	**-5**	**-1**	**-1**	**-2**	**-1**	**0**	**-1**	**0**	**-1**
16	Node 17 + node 23		1.0	-1	0	-3	0	1	0	3	0	0	1	0	2	1	0	0
		74	1.0	-5	0	-5	0	5	0	3	0	0	2	0	3	1	0	0
17	Node 18 + node 21		1.0	1	-2	2	0	0	0	1	0	0	0	0	0	0	0	0
		59	1.0	0	-2	1	0	2	0	1	0	0	0	0	0	0	0	0
18	M.peruviuanusK262N + node 19		1.0	2	0	9	0	15	0	0	-1	0	0	1	0	0	0	1
		100	1.0	2	2	12	1	25	0	0	-2	0	0	1	0	0	0	1
19	M.peruvianusTAL + node 20		1.0	0	0	1	0	0	0	0	1	0	2	0	0	0	0	0
		87	1.0	0	0	0	0	1	0	0	1	0	3	0	0	0	0	0
20	M.peruvianusMA + ACA		1.0	2	1	2	0	3	0	0	0	0	0	0	0	0	0	0
		100	1.0	1	1	1	0	4	0	0	0	0	0	0	0	0	0	0
21	M. peruvianusASA + node 22		1.0	-2	2	3	1	6	0	2	1	0	1	0	2	1	0	0
		100	1.0	-5	3	3	1	11	0	3	2	0	2	0	3	1	0	0
22	M. peruvianusJAH + SPV		1.0	1	-2	1	0	1	0	0	0	0	0	2	0	0	1	0
		97	1.0	0	-2	0	0	3	0	1	0	0	0	3	0	0	1	0
23	M. tigrisLP + SCH		1.0	4	1	9	0	6	0	0	-1	0	0	0	0	0	-1	0
		100	1.0	3	1	9	0	10	0	0	-1	0	0	0	0	-1	-1	0
24	Node 25 + node 26		1.0	2	1	1	0	1	0	-1	0	0	0	0	0	0	0	0
		96	1.0	2	3	0	0	4	0	-1	0	0	0	0	0	0	0	0
25	M. hetR + M. quadrivittatusIL		1.0	1	0	6	0	5	-1	1	0	0	0	0	0	0	0	0
		100	1.0	1	1	9	0	10	-1	1	0	0	0	0	0	0	0	0

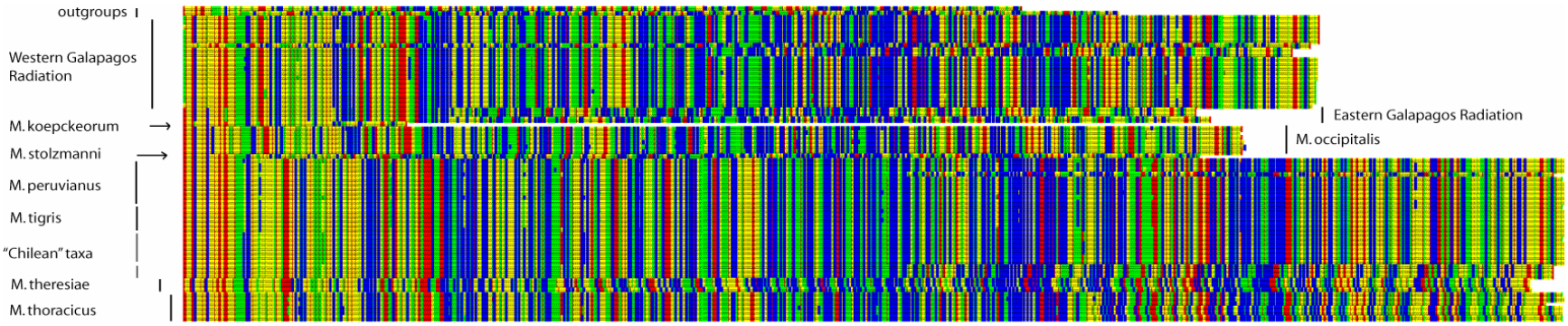
26	M. tigrisY + node 27		1.0	3	0	7	0	6	0	-3	0	0	1	0	0	0	0	0
		100	1.0	2	0	7	0	8	0	-4	0	0	1	0	0	0	0	0
27	M. quadrivittatusCB + node 28		.99	1	0	1	0	-1	0	0	0	0	1	0	0	0	0	0
		62	1.0	0	0	1	0	1	0	0	0	-1	1	0	0	0	0	0
28	M. tigrisCT + node 29		1.0	0	0	0	-1	-3	0	6	0	0	-1	0	0	0	0	0
		70	1.0	-1	0	-1	-1	-2	0	9	0	0	-1	0	0	0	0	0
29	M. tigrisDRQ + PU		1.0	1	0	1	0	3	0	0	0	-1	0	0	0	0	1	0
		93	1.0	1	0	1	0	3	0	0	0	-1	0	0	0	0	1	0
30	**M. theresiaeK514S + node 31**		**1.0**	**8**	**4**	**19**	**1**	**23**	**2**	**5**	**3**	**2**	**0**	**2**	**0**	**3**	**2**	**4**
		100	**1.0**	**11**	**6**	**25**	**1**	**32**	**3**	**5**	**4**	**3**	**1**	**4**	**0**	**4**	**2**	**7**
31	M. theresiaeLY + Ai		1.0	1	-1	0	0	-1	0	0	1	0	0	0	0	0	0	0
			1.0	0	0	-1	0	-1	1	-1	1	0	1	1	0	0	0	-1
32	**M. thoracicus**		**1.0**	**16**	**5**	**13**	**1**	**11**	**4**	**5**	**11**	**1**	**5**	**5**	**5**	**4**	**4**	**2**
		100	**1.0**	**21**	**8**	**17**	**1**	**17**	**6**	**10**	**15**	**2**	**8**	**8**	**7**	**6**	**6**	**4**
33	Node 34 + node 37		1.0	7	-2	6	1	5	0	0	1	-1	0	0	0	0	0	0
		100	1.0	8	-2	8	2	9	0	-1	2	-1	0	0	0	0	0	0
34	M. thoracicusPM + node 35		1.0	-3	3	15	2	17	1	2	0	0	1	0	1	-2	0	0
		100	1.0	-4	4	16	2	21	1	2	0	0	1	0	1	-2	0	0
35	M. thoracicusK650N + +node 36		1.0	7	1	1	0	0	0	0	0	0	0	0	0	0	0	-1
		98	1.0	8	1	0	0	1	0	0	0	0	0	0	0	0	1	0
36	*M. thoracicus* TAL + CLV		1.0	3	-1	3	0	-2	0	1	1	0	0	0	0	0	0	0
		96	1.0	2	-1	2	0	-1	0	1	1	0	0	0	0	0	0	0
37	M. thoracicusFM + E		1.0	2	3	10	1	6	0	0	0	0	-2	0	0	0	-1	-1
		100	1.0	0	4	9	1	11	0	1	0	0	-2	0	0	0	0	-3
38	M. thoracicusTB + node 39		1.0	11	3	8	1	9	1	3	1	0	0	0	0	0	1	2
		100	1.0	12	5	8	2	11	2	4	1	0	1	1	0	0	1	2
39	M. thoracicusM + P		1.0	6	2	15	2	11	0	0	0	0	0	1	0	1	1	0
		100	1.0	5	3	14	2	14	0	0	0	0	-1	1	0	1	1	0
40	**Occipitalis group**		**1.0**	**3**	**-1**	**1**	**1**	**-1**	**3**	**0**	**3**	**2**	**2**	**1**	**-1**	**1**	**2**	**1**
		100	**1.0**	**6**	**0**	**1**	**1**	**-5**	**6**	**0**	**5**	**5**	**6**	**5**	**-1**	**3**	**5**	**3**
41	Node 42 + node 61		1.0	0	-1	-1	1	3	0	-1	2	0	0	-2	-1	0	2	-1
		51	1.0	-3	-1	-3	2	9	0	-2	3	0	1	-3	-1	1	3	-2
42	**Western Galapagos radiation**		**1.0**	**7**	**-3**	**7**	**2**	**7**	**2**	**3**	**3**	**1**	**1**	**2**	**0**	**3**	**2**	**1**
		100	**1.0**	**13**	**-6**	**12**	**2**	**13**	**7**	**6**	**3**	**3**	**3**	**5**	**1**	**6**	**3**	**3**
43	Node 44 + node 50		1.0	6	4	8	0	6	0	2	0	0	0	0	0	0	-1	0
		100	1.0	8	7	12	1	12	0	3	1	0	1	0	0	1	-1	0
44	Node 45 + node 49 (*M. grayii*)		.98	1	-1	2	0	-2	0	0	1	0	0	0	0	1	0	0

		99	.99	1	0	2	0	-3	0	0	1	0	0	0	0	1	-1	0
45	*M. pacificus* + node 46		1.0	3	-1	1	0	5	0	0	1	0	0	0	0	0	0	0
		70	1.0	5	0	1	-1	13	0	-1	1	0	0	0	0	1	0	0
46	M. albemarlensisCW + node 47		1.0	2	-2	2	0	-1	0	0	0	0	0	0	0	0	-1	0
		97	.99	4	-3	2	0	1	-1	0	0	0	0	0	0	0	-1	0
47	M. albemarlensisI + node 48		1.0	1	2	0	0	3	0	1	0	0	-1	0	0	0	0	0
		69	1.0	1	3	-1	0	5	0	1	0	0	-2	0	0	0	0	0
48	M. albemarlensisBEH + F		.97	1	0	0	0	2	-1	0	0	0	0	-1	0	0	0	0
			.53	0	0	-1	0	3	-1	1	0	0	0	-1	0	0	0	0
49	M. grayiiEN + F		1.0	8	1	14	0	5	0	2	2	0	2	3	1	0	-1	1
		100	1.0	9	1	19	0	10	0	1	3	0	3	4	2	1	-1	1
50	Node 51 + node 55		1.0	3	-3	0	1	4	0	0	0	0	0	0	0	0	1	0
		81	1.0	4	-5	1	1	6	0	1	-1	0	0	1	0	0	0	0
51	*M. duncanensis* + node 52		1.0	3	-1	4	0	0	0	-1	-1	0	-1	1	0	-1	0	0
		72	.99	2	-1	4	0	2	0	-1	-1	0	0	1	0	0	0	0
52	M. jacobiBC + node 53		1.0	6	1	19	0	15	0	0	-1	0	1	1	0	0	0	0
		100	1.0	5	1	18	0	15	0	0	0	0	2	1	0	1	0	0
53	M. jacobiBA + node 54		1.0	2	0	0	0	2	1	0	0	0	0	0	0	-1	0	0
		81	1.0	1	0	-1	0	4	1	0	0	0	0	0	0	-1	0	0
54	M. jacobiSCH + BR5		.83	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0
			.84	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
55	M. indefatigabilisFE + node 56		1.0	2	2	2	0	-1	0	0	0	0	0	0	0	0	0	0
		100	1.0	3	3	3	0	1	0	0	1	0	0	0	0	1	-1	0
56	M. ndefatigabilisCDP + node 57		1.0	4	1	8	0	10	0	1	0	0	0	2	0	0	0	0
		100	1.0	4	2	10	0	16	0	1	1	0	0	2	0	1	0	0
57	M. indefatigabilisED + node 58		91	1	0	0	0	0	0	-1	0	0	0	1	0	0	0	0
			.92	0	1	0	0	0	0	-1	-1	0	0	2	0	0	0	0
58	M. indefatigabilisDM + SY		1.0	1	1	2	0	-2	0	0	0	0	0	2	0	0	-1	0
		99	1.0	1	2	2	0	-2	0	0	1	0	0	2	0	1	-1	0
59	***M. delanonis* + node 60**		**1.0**	**10**	**8**	**5**	**4**	**15**	**2**	**3**	**0**	**1**	**1**	**0**	**1**	**0**	**1**	**2**
		100	**1.0**	**12**	**10**	**7**	**6**	**23**	**2**	**4**	**-1**	**2**	**1**	**0**	**1**	**0**	**1**	**2**
60	M. delanonisES + IO		.57	1	0	1	0	-1	0	0	0	0	0	-1	0	0	0	0
			1.0	0	0	-1	-1	1	0	-1	1	0	0	0	0	0	0	0
61	***M. koepckeorum* + node 62**		**.87**	**-2**	**0**	**3**	**0**	**3**	**0**	**-1**	**1**	**0**	**0**	**-1**	**-1**	**0**	**1**	**-1**
		100	**.72**	**-4**	**-1**	**3**	**0**	**8**	**0**	**-1**	**0**	**0**	**0**	**-1**	**0**	**0**	**1**	**-2**
62	**Node 63 + node 65**		**1.0**	**-1**	**2**	**3**	**2**	**4**	**2**	**0**	**5**	**0**	**0**	**-1**	**1**	**1**	**3**	**1**
		100	**1.0**	**-2**	**5**	**7**	**3**	**7**	**4**	**0**	**10**	**1**	**0**	**0**	**3**	**2**	**5**	**5**

63	*M. bivittatus* + node 64		1.0	-8	1	18	1	1	3	4	0	0	-1	6	3	0	0	2
		100	1.0	-11	2	23	2	3	3	5	0	0	-1	8	4	1	0	1
64	M. habeliNB + M.habeliPM		1.0	22	0	9	1	102	0	1	9	0	1	0	0	0	0	3
		100	1.0	24	1	10	1	113	0	1	10	0	1	0	0	1	-1	3
65	Node 66 + node 68		1.0	12	6	14	2	-2	3	3	2	0	4	0	2	1	2	1
		100	1.0	17	10	19	3	-1	2	4	3	0	6	1	4	1	4	3
66	Node 67 + M. occ_K1049N + TAL		1.0	17	3	24	0	24	0	-1	0	0	0	1	0	0	-1	0
		100	1.0	18	5	26	0	27	0	-2	0	0	0	1	0	0	-1	0
67	M. occipitalisJEQ + K777N		1.0	1	0	6	0	-2	0	0	0	0	0	-1	0	0	1	0
		87	1.0	1	1	8	0	-2	0	1	0	0	0	-2	0	0	3	0
68	M. occipitalisSRM + K1245N		1.0	10	5	21	1	25	-2	-1	0	0	2	0	-1	0	1	0
		100	1.0	9	4	22	2	37	-1	-3	2	0	4	0	0	1	3	-3


Ecuador
Peru
Chile
N
Galapagos
occSRM
occK1245N
perACA
perMA
Talara (3 spp.)
occK1049N
occK777N
M. stolzmanni
thoCLV
occJEQ
thoK650N
thoPM
thoFM
M. koepckeorum
perK262N
tesLY
thoE
perSPV
perJAH
thoP
thoM
tesP
thoTB
tigSCH
tesK514S
tigDRQ
perASA
tigY
tigLP
tigPU
tigCT
quaCB
hetR
qualL
qual
quaCM
yanPON
quaCZ
qualQC
thePI
ataDL
ataISM
theCL
ataCOL
ataTAL
ataTPL
300 Km


N
Pinta
pacPP
habNB
habPM
Marchena
0°00'
Santiago
jacBC
albTC
jacBA
jacBR5
indSY
albBU
albCW
jacSCH
indDM
Fernandina
indCDP
indED
albBEH
dun
indFE
bivSC
Santa Cruz
Santa Fe
San Cristobal
Isabela
graEN
graF
dellO
delGA
del
Floreana
Espanola
10 km



Manual alignment;

Motif recognition and gap placement

1.

```
lquique  CTCCCACTTTTGCCTTTTGCCTTCAGC        CTCCCACTTTTGCCTTTTGCCTTCAGC
Taltal   CTCCCACTTTTGCCTTTTGCCTTCAGC        CTCCCACTTTTGCCTTTTGCCTTCAGC
Calama   CTCCCCCTTTTGCCTTCAGC               CTCCC-------CCTTTTGCCTTCAGC
Pica     CTCCCCCTTTTGCCTTCAGC               CTCCC-------CCTTTTGCCTTCAGC
               *          **
```

2.

```
pac      TGTTCTGTGGCTTTGGCTAGCGGCTGGGCTCT  TGTTCTGTGGCTTTGGCTAGCGGCTGGGCTCT
albCW    TGTTCTGTGGCTTTGGCTAGCGGCTGGGCTCT  TGTTCTGTGGCTTTGGCTAGCGGCTGGGCTCT
habNB    CGTTCTGCTCG----------------------  CGTTCT----------------------GCTCG
habPM    CGTTCTGCTCG----------------------  CGTTCT----------------------GCTCG
bivSC    CGTTCTGCGGCTGGGCTCG--------------  CGTTCT--------------GCGGCTGGGCTCG
occJEQ   CGTTCTGCGGCTGGGCTCT--------------  CGTTCT--------------GCGGCTGGGCTCT
               ***  **  ****
```

3.1.

alternative (1)

```
GCTGGGGGGGGGGTCTT
ACTGGGGG-----TCTT
ACTGGGGGGGGG-TCTT
ACTGGGGGGGG--TCTT
```

3.2.

alternative (1)	alternative (2)	alternative (3)
GGTTAT-----GAATTAADA	GGTTA-----TGAATTAACA	GGTTATGA-----ATTAACA
GGTTAT-----AAATTAACA	GGTTAT-----AAATTAACA	GGTTATAA-----ATTAACA
GGTTATAA**tataa**ATTAACA	GGTTATAA**tataa**ATTAACA	GGTTATAA**tataa**ATTAACA
GGTTATAA**tataa**ATTAACA	GGTTATAA**tataa**ATTAACA	GGTTATAA**tataa**ATTAACA
*	*	

3.3.

alternative (1)	alternative (2)
CAGATTGATTGATTATTATACTGATTATGC	CAGATTGATTGATTTTATACTGATTATGC
CAGATT----------**gatt**ATGC	CAGATT**gatt**----------ATGC
CAGATT**gatt**----------ATGC	CAGATT**gatt**----------ATGC
CAGATT----**gatt**----------ATGC	CAGATT**gatt**----------ATGC

3.4.

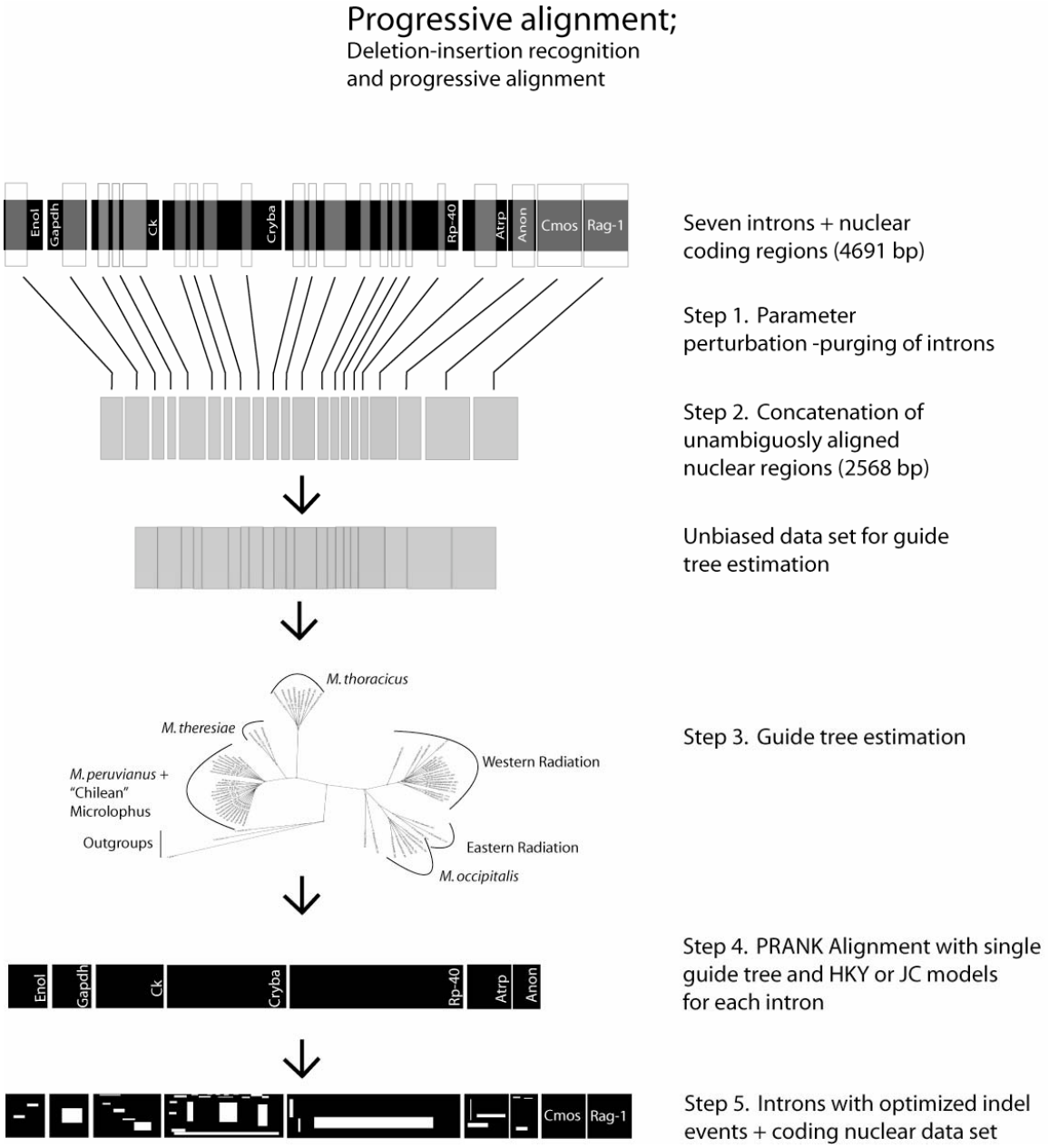
alternative (1)	alternative (2)	alternative (3)
TCCTCCTCCTCCTCCTCACC	TCCTCCTCCTCCTCCTCACC	TCCTCCTCCTCCTCCTCACC
TCCTCCTCCTCC---TCACC	TCCTCCTCC---TCCTCACC	TCCTCCTCCTCC---TCACC
TCCTCCTCCTCCTCCTCACC	TCCTCCTCCTCCTCCTCACC	TCCTCCTCCTCCTCCTCACC
TCCTCCTCC---TCCTCACC	TCCTCCTCCTCC---TCACC	TCCTCCTCCTCC---TCACC
TCCTCCTCCTCC---TCACC	TCCTCCTCCTCC---TCACC	TCCTCCTCCTCC---TCACC
TCCTCC---TCC---TCACC	TCCTCCTCC------TCACC	TCCTCCTCC------TCACC

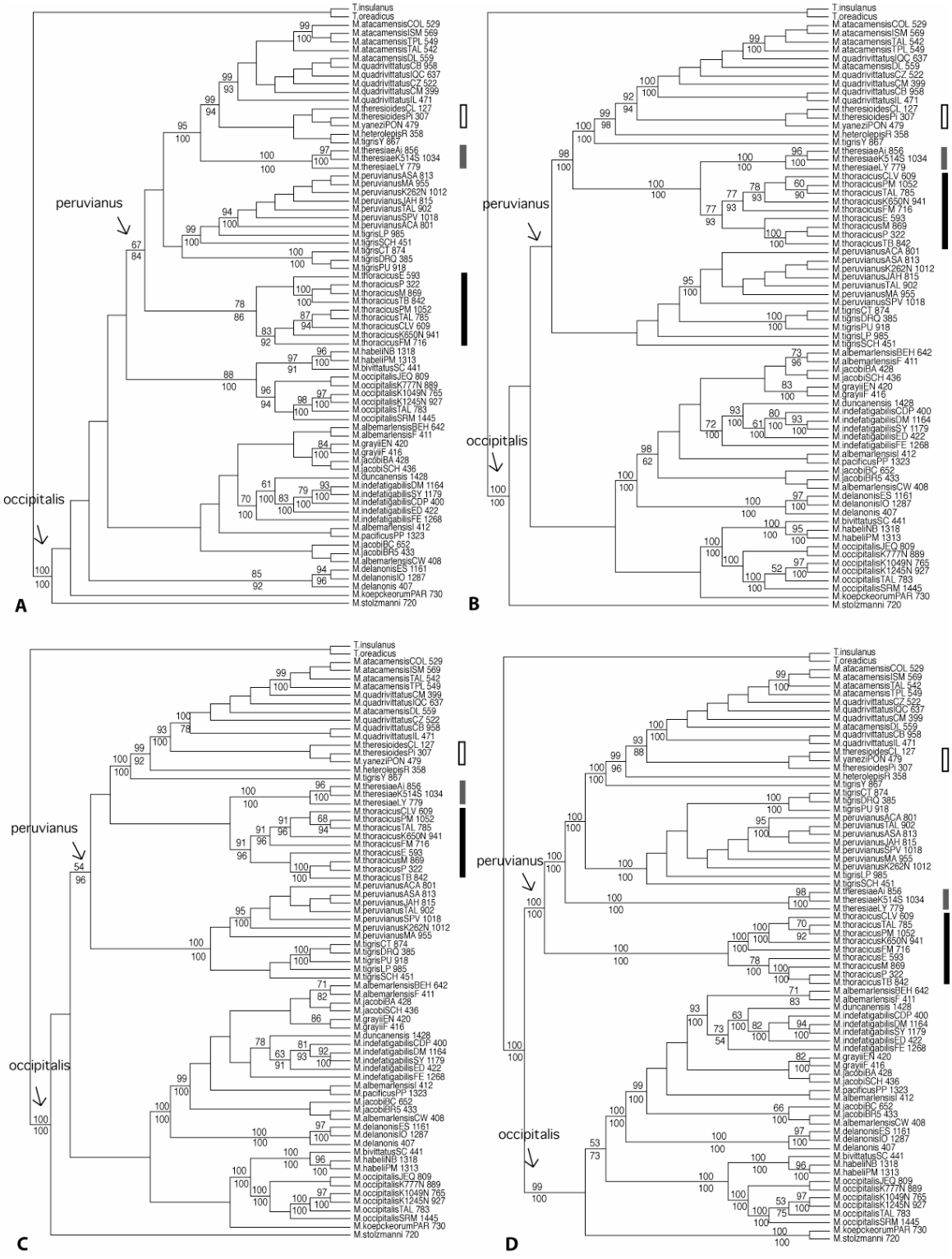
3.5.

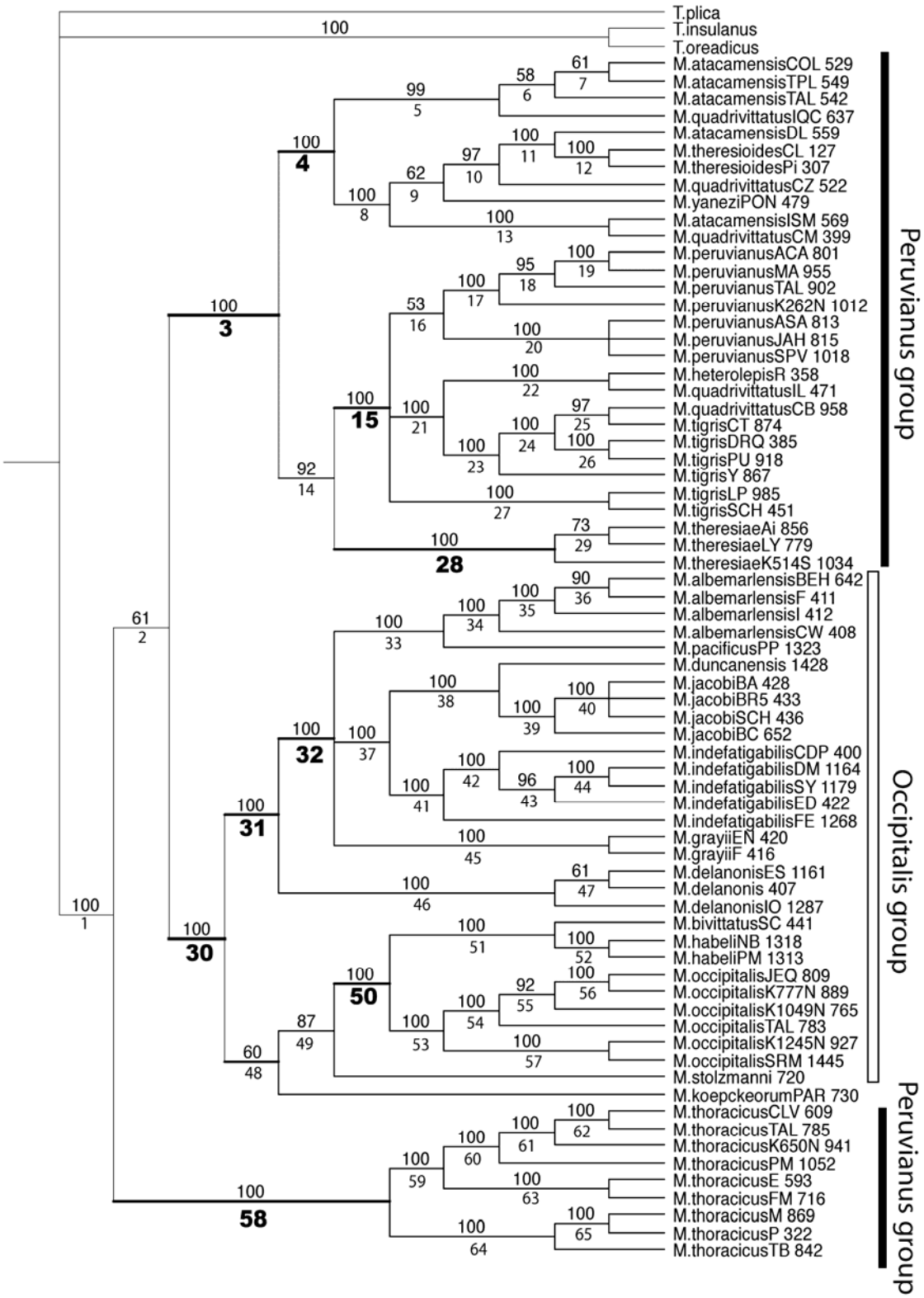
alternative (1)	alternative (2)
ATGACGAGATAGTAGTGC	ATGACGAGATAGTAGTGC
ATGA------GATAGTGC	ATGA---GATAGT---GC
ATGACGAGATAGTAGTGC	ATGACGAGATAGTAGTGC
ATGA------GATAGTGC	ATGA---GATAGT---GC
ATGACGAGATAGTAGTGC	ATGACGAGATAGTAGTGC
**	

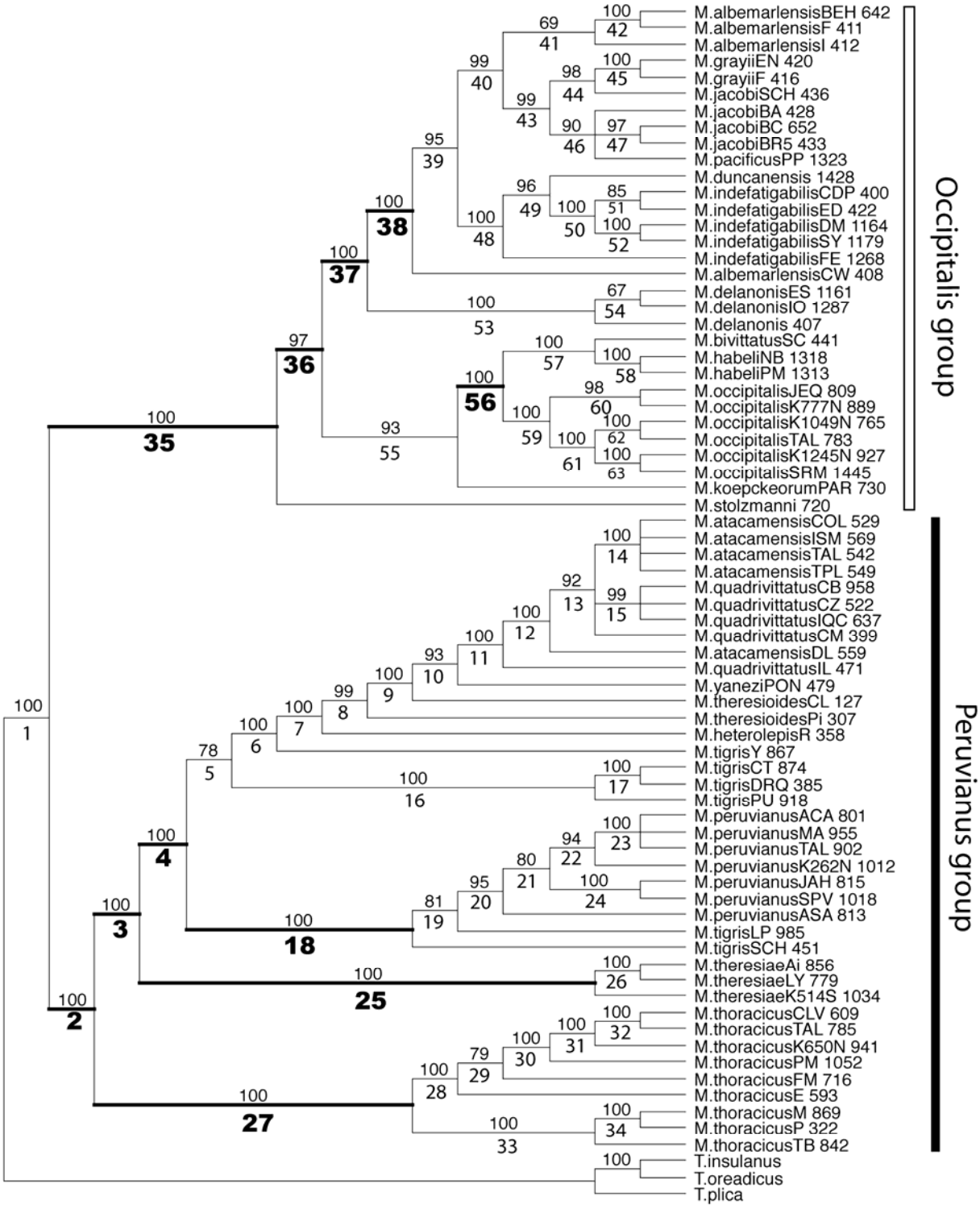
3.6.

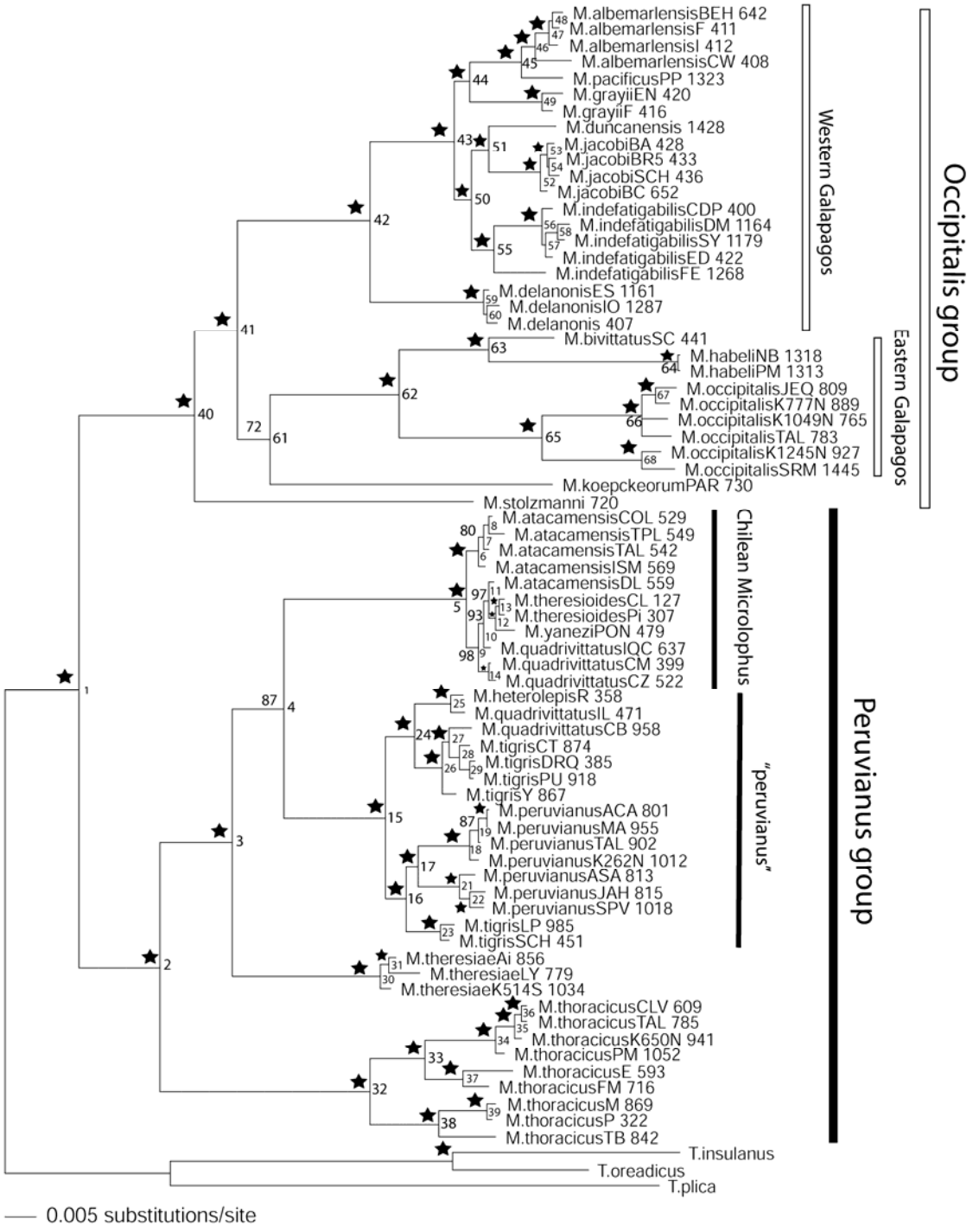
Introns with optimized indel events + coding regions nuclear data set

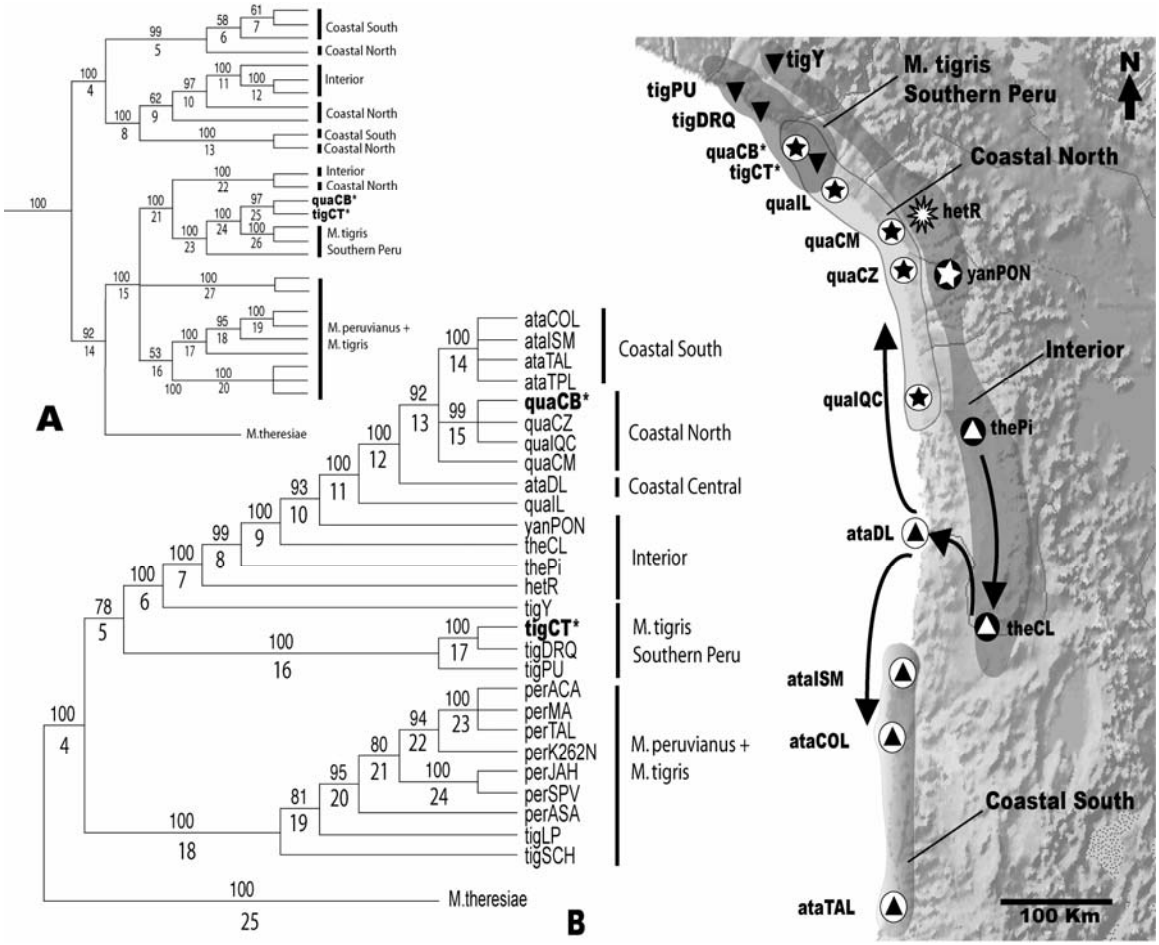

Progressive alignment;
Deletion-insertion recognition
and progressive alignment
Enol
Gapdh
Ck
Cryba
Rp-40
Atrp
Anon
Cmos
Rag-1
Seven introns + nuclear coding regions (4691 bp)
Step 1. Parameter perturbation -purging of introns
Step 2. Concatenation of unambiguosly aligned nuclear regions (2568 bp)
Unbiased data set for guide tree estimation
M. thoracicus
M. theresiae
M. peruvianus + "Chilean" Microlophus
Outgroups
Western Radiation
Eastern Radiation
M. occipitalis
Step 3. Guide tree estimation
Step 4. PRANK Alignment with single guide tree and HKY or JC models for each intron
Step 5. Introns with optimized indel events + coding nuclear data set


A
B
C
D
peruvianus
occipitalis
T.insulanus
T.oreadicus
M.atacamensisCOL 529
M.atacamensisISM 569
M.atacamensisTPL 549
M.atacamensisTAL 542
M.atacamensisDL 559
M.quadrivittatusCB 958
M.quadrivittatusIQC 637
M.quadrivittatusCZ 522
M.quadrivittatusCM 399
M.quadrivittatusIL 471
M.theresioidesCL 127
M.theresioidesPI 307
M.yaneziPON 479
M.heterolepisR 358
M.tigrisY 867
M.theresiaeAi 856
M.theresiaeK514S 1034
M.theresiaeLY 779
M.peruvianusASA 813
M.peruvianusMA 955
M.peruvianusK262N 1012
M.peruvianusJAH 815
M.peruvianusTAL 902
M.peruvianusSPV 1018
M.peruvianusACA 801
M.tigrisLP 985
M.tigrisSCH 451
M.tigrisCT 874
M.tigrisDRQ 385
M.tigrisPU 918
M.thoracicusE 593
M.thoracicusP 322
M.thoracicusM 869
M.thoracicusTB 842
M.thoracicusPM 1052
M.thoracicusTAL 785
M.thoracicusCLV 609
M.thoracicusK650N 941
M.thoracicusFM 716
M.habeliINB 1318
M.habeliPM 1313
M.bivittatusSC 441
M.occipitalisJEQ 809
M.occipitalisK777N 889
M.occipitalisK1049N 765
M.occipitalisK1245N 927
M.occipitalisTAL 783
M.occipitalisSRM 1445
M.albemarlensisBEH 642
M.albemarlensisF 411
M.grayiiEN 420
M.grayiiF 416
M.jacobiBA 428
M.jacobiSCH 436
M.duncanensis 1428
M.indefatigabilisCDP 400
M.indefatigabilisDM 1164
M.indefatigabilisSY 1179
M.indefatigabilisED 422
M.indefatigabilisFE 1268
M.albemarlensisI 412
M.pacificusPP 1323
M.jacobiBC 652
M.jacobiBR5 433
M.albemarlensisCW 408
M.delanonisES 1161
M.delanonisIO 1287
M.delanonis 407
M.koepckeorumPAR 730
M.stolzmanni 720


T.plica
T.insulanus
T.oreadicus
M.atacamensisCOL 529
M.atacamensisTPL 549
M.atacamensisTAL 542
M.quadrivittatusIQC 637
M.atacamensisDL 559
M.theresioidesCL 127
M.theresioidesPi 307
M.quadrivittatusCZ 522
M.yaneziPON 479
M.atacamensisISM 569
M.quadrivittatusCM 399
M.peruvianusACA 801
M.peruvianusMA 955
M.peruvianusTAL 902
M.peruvianusK262N 1012
M.peruvianusASA 813
M.peruvianusJAH 815
M.peruvianusSPV 1018
M.heterolepisR 358
M.quadrivittatusIL 471
M.quadrivittatusCB 958
M.tigrisCT 874
M.tigrisDRQ 385
M.tigrisPU 918
M.tigrisY 867
M.tigrisLP 985
M.tigrisSCH 451
M.theresiaeAi 856
M.theresiaeLY 779
M.theresiaeK514S 1034
M.albemarlensisBEH 642
M.albemarlensisF 411
M.albemarlensisI 412
M.albemarlensisCW 408
M.pacificusPP 1323
M.duncanensis 1428
M.jacobiBA 428
M.jacobiBR5 433
M.jacobiSCH 436
M.jacobiBC 652
M.indefatigabilisCDP 400
M.indefatigabilisDM 1164
M.indefatigabilisSY 1179
M.indefatigabilisED 422
M.indefatigabilisFE 1268
M.grayiiEN 420
M.grayiiF 416
M.delanonisES 1161
M.delanonis 407
M.delanonisIO 1287
M.bivittatusSC 441
M.habeliNB 1318
M.habeliPM 1313
M.occipitalisJEQ 809
M.occipitalisK777N 889
M.occipitalisK1049N 765
M.occipitalisTAL 783
M.occipitalisK1245N 927
M.occipitalisSRM 1445
M.stolzmanni 720
M.koepckeorumPAR 730
M.thoracicusCLV 609
M.thoracicusTAL 785
M.thoracicusK650N 941
M.thoracicusPM 1052
M.thoracicusE 593
M.thoracicusFM 716
M.thoracicusM 869
M.thoracicusP 322
M.thoracicusTB 842
Peruvianus group
Occipitalis group
Peruvianus group


M.albemarlensisBEH 642
M.albemarlensisF 411
M.albemarlensisI 412
M.grayiiEN 420
M.grayiiF 416
M.jacobiSCH 436
M.jacobiBA 428
M.jacobiBC 652
M.jacobiBR5 433
M.pacificusPP 1323
M.duncanensis 1428
M.indefatigabilisCDP 400
M.indefatigabilisED 422
M.indefatigabilisDM 1164
M.indefatigabilisSY 1179
M.indefatigabilisFE 1268
M.albemarlensisCW 408
M.delanonisES 1161
M.delanonisIO 1287
M.delanonis 407
M.bivittatusSC 441
M.habeliNB 1318
M.habeliPM 1313
M.occipitalisJEQ 809
M.occipitalisK777N 889
M.occipitalisK1049N 765
M.occipitalisTAL 783
M.occipitalisK1245N 927
M.occipitalisSRM 1445
M.koepckeorumPAR 730
M.stolzmanni 720
M.atacamensisCOL 529
M.atacamensisISM 569
M.atacamensisTAL 542
M.atacamensisTPL 549
M.quadrivittatusCB 958
M.quadrivittatusCZ 522
M.quadrivittatusIQC 637
M.quadrivittatusCM 399
M.atacamensisDL 559
M.quadrivittatusIL 471
M.yaneziPON 479
M.theresioidesCL 127
M.theresioidesPi 307
M.heterolepisR 358
M.tigrisY 867
M.tigrisCT 874
M.tigrisDRQ 385
M.tigrisPU 918
M.peruvianusACA 801
M.peruvianusMA 955
M.peruvianusTAL 902
M.peruvianusK262N 1012
M.peruvianusJAH 815
M.peruvianusSPV 1018
M.peruvianusASA 813
M.tigrisLP 985
M.tigrisSCH 451
M.theresiaeAi 856
M.theresiaeLY 779
M.theresiaeK514S 1034
M.thoracicusCLV 609
M.thoracicusTAL 785
M.thoracicusK650N 941
M.thoracicusPM 1052
M.thoracicusFM 716
M.thoracicusE 593
M.thoracicusM 869
M.thoracicusP 322
M.thoracicusTB 842
T.insulanus
T.oreadicus
T.plica
Occipitalis group
Peruvianus group


M.albemarlensisBEH 642
M.albemarlensisF 411
M.albemarlensisI 412
M.albemarlensisCW 408
M.pacificusPP 1323
M.grayiiEN 420
M.grayiiF 416
M.duncanensis 1428
M.jacobiBA 428
M.jacobiBR5 433
M.jacobiSCH 436
M.jacobiBC 652
M.indefatigabilisCDP 400
M.indefatigabilisDM 1164
M.indefatigabilisSY 1179
M.indefatigabilisED 422
M.indefatigabilisFE 1268
M.delanonisES 1161
M.delanonisIO 1287
M.delanonis 407
M.bivittatusSC 441
M.habeliNB 1318
M.habeliPM 1313
M.occipitalisJEQ 809
M.occipitalisK777N 889
M.occipitalisK1049N 765
M.occipitalisTAL 783
M.occipitalisK1245N 927
M.occipitalisSRM 1445
M.koepckeorumPAR 730
M.stolzmanni 720
M.atacamensisCOL 529
M.atacamensisTPL 549
M.atacamensisTAL 542
M.atacamensisISM 569
M.atacamensisDL 559
M.theresioidesCL 127
M.theresioidesPi 307
M.yaneziPON 479
M.quadrivittatusIQC 637
M.quadrivittatusCM 399
M.quadrivittatusCZ 522
M.heterolepisR 358
M.quadrivittatusIL 471
M.quadrivittatusCB 958
M.tigrisCT 874
M.tigrisDRQ 385
M.tigrisPU 918
M.tigrisY 867
M.peruvianusACA 801
M.peruvianusMA 955
M.peruvianusTAL 902
M.peruvianusK262N 1012
M.peruvianusASA 813
M.peruvianusJAH 815
M.peruvianusSPV 1018
M.tigrisLP 985
M.tigrisSCH 451
M.theresiaeAi 856
M.theresiaeLY 779
M.theresiaeK514S 1034
M.thoracicusCLV 609
M.thoracicusTAL 785
M.thoracicusK650N 941
M.thoracicusPM 1052
M.thoracicusE 593
M.thoracicusFM 716
M.thoracicusM 869
M.thoracicusP 322
M.thoracicusTB 842
T.insulanus
T.oreadicus
T.plica
0.005 substitutions/site
Western Galapagos
Eastern Galapagos
Occipitalis group
Chilean Microlophus
"peruvianus"
Peruvianus group


A
Coastal South
Coastal North
Interior
Coastal North
Coastal South
Coastal North
Interior
Coastal North
quaCB*
tigCT*
M. tigris
Southern Peru
M. peruvianus +
M. tigris
M.theresiae
B
ataCOL
ataISM
ataTAL
ataTPL
quaCB*
quaCZ
qualQC
quaCM
ataDL
qualL
yanPON
theCL
thePi
hetR
tigY
tigCT*
tigDRQ
tigPU
perACA
perMA
perTAL
perK262N
perJAH
perSPV
perASA
tigLP
tigSCH
M.theresiae
Coastal South
Coastal North
Coastal Central
Interior
M. tigris
Southern Peru
M. peruvianus +
M. tigris
tigY
tigPU
tigDRQ
M. tigris
Southern Peru
N
quaCB*
tigCT*
Coastal North
qualL
hetR
quaCM
quaCZ
yanPON
Interior
qualQC
thePi
ataDL
theCL
ataISM
ataCOL
Coastal South
ataTAL
100 Km

Appendix 1. Species and populations of all ingroup and outgroup taxa used in this study. Country acronyms in locality column are: Chile (CH), Ecuador (EC), and Peru (PE); while abbreviations identify localities plotted in Figs. 1 and 2 Voucher specimens (if collected) are cataloged at Museo de Historia Natural Universidad Nacional Mayor de San Marcos, Lima, Peru (MNSM), Museo de Zoología de la Universidad de Concepción, Concepción, Chile (MZUC), and Museo de Zoología, Pontificia Universidad Católica, Quito, Ecuador (QCAZ). Outgroup *Tropidurus* tissue samples were provided by M. T. Rodrigues from Museu de Zoologia Universidade de São Paulo (MZUSP). Tissue samples from the Galápagos are catalogued in the Museum of Southwestern Biology (MSB), University of New Mexico, Albuquerque.

Species	Abbr.	Locality	Geographical coordinates	Museum number
INGROUP				
***Peruvianus* group**				
M. atacamensis	ataISM	Isla Santa Maria, Antofagasta. CH	23° 25 60" S 70° 37' 0" W	M569, M640
	ataCOL	Coloso, Playa Rocosa, 18 km S Antofagasta. CH	23°45'24" S 70°27'53" W	M529
	ataTAL	35 km N de Taltal, Camino a Paposo, Roquerios Costa. CH	25° 3' 0" S 70° 28' 60 W	M153, M542
	ataTPL	Tres Playitas, 15 km N Huasco, III Region. CH	28° 28' 0" S 71° 10' 60" W	M547 – M549
	ataDL	Desembocadura Rio Loa, interior quebrada. CH	21° 25' S 70° 04' W	M141
M. quadrivittatus	quaCM	Caleta Meca, ~11 km al S de Ite, Depto. Tacna. PE	17° 58' 32" S 70° 51 04" W	M399
	quaCZ	Playa Corazones, 15 km S Arica. CH	18° 28' 60 S 70° 19' 60" W	M522 – M523
	quaIQC	Iquique, costa. CH	20° 13' 0" S 70° 10' 0" W	M480, M576
	quaIL	Ilo (Km 1186 Panam S) Roquerios al sur de, Moquegua. PE	17° 38' S 71° 20' W	M471 – M472
	quaCB	Caleta la Ballenita (intertidal). PE	17° 00' 56.5" S 72° 02' 28.9" W	M905, M958 – M959
M. theresioides	thePi	Pica, Iquique, 4666ft. CH	20° 28 80.9" S 69° 19.161" W	M306 – M307
	theCL	Calama. CH	22° 28' 0" S 68° 55' 60" W	M127

M. yanezi	yanPON	Pocon, Valle de Lluta, Arica. CH	18° 26' 60" S 70° 40' 00" W	M474 – M479
M. peruvianus	perACA	Acapulco (Km 1227 Panam N). PE	03° 42' 0" S 80° 43' 60" W	M801, M1063
	perMA	Mancora (Km 1160 Panam N; Depto. Piura). PE	04° 06' 15" S 81° 02' 45" W	M955 – M958
	perTAL	Puerto Talara (118 km N de Piura). Piura. PE	04° 34' 33' S 81° 16' 9" W	M900 – M902
	perK262N	Km 262 Panam. N. (~20 km S de Huarmey). PE	10° 04' 00" S 78° 10 00" W	M723, M1012
	perSPV	Km 69.4 Panam. S. Salinas de Puerto Viejo. PE	12° 33' 20.6" S 76° 42' 41.9" W	M1018
	perJAH	Jahuay (km 179 Panam. S; Desemb. Rio Topaca). PE	13° 19' 19.7" S 76° 14 38.4" W	M815, M1039
	perASA	Agua Salada 26 km al N de Chala (Km 587 Panam S). PE	15° 47' 6.9" S 74° 24' 55.2" W	M811 – M813
M. tigris	perSCH	10 km S de Chala, Quebrada Interior, Arequipa. PE	15° 52' S 74° 16' W	M449 – M451
	tigY	Yura, Arequipa. PE	16° 13' 4" S 71° 42' 10" W	M198, M867
	tigLP	Los pescadores (Km 749 Panam. S). PE	16° 23' 16.1" S 73° 15' 45.6" W	M985 – M987
	tigPU	Pucchun, Valle de Camana (Km 820 Panam S). PE	16° 36' 42.4" S 72° 47' 35.4" W	M918 – M919
	tigDRQ	Quebrada Seca/Desemb. Rio Quilca, Km 11 Panam., Arequipa. PE	16° 43' S 72° 25' W	M385 – M340
	tigCT	Catarindo (cerca de Mollendo). PE	17° 00' 59.9" S 72° 02' 00.8" W	M874
M. heterolepis	hetR	Fundo Rospigliosi, S de Tacna. PE	18° 0' 20" S 70° 14' 54" W	M358 – M359
M. thoracicus	thoTAL	Puerto Talara (118 km N Piura). Piura. PE	04° 14' 53" S 81° 13' 46" W	M784 – M785
	thoCLV	Cerro la Vieja Km 52 entre Jayanca y Motupe, Lambayeque. PE	06° 9' 0" S 79° 43' 60" W	M609
	thoK650N	Km 650 Panam N. (~7 km S de San Pedro de Lloc; Lambayeque). PE	07° 27' 42.0" S 79° 26' 53.5" W	M941
	thoPM	Km 526 al E de Puerto Morin (Tillandsia-dunas). PE	08° 22' 30.9" S 78° 51' 25.0"W	M1052 – M1059
	thoFM	Fundo Marianela (Km 331 Panam. N). PE	10° 18' 17.2" S 78 2' 57.7" W	M716
	thoE	Peaje Paraiso (~50 km N de Chancay), Depto. de Lima. PE	11° 19' 0" S 77° 30' 0" W	M591 – M593
	thoP	Pozo santo, Depto. De Ica. PE	13° 54' 0" S 76° 4' 60" W	M 10, M322
	thoM	San Jose de Los Molinos. PE	13° 56' 25.0" S 75° 40' 14.9" W	M630, M869
	thoTB	Tierras Blancas (Km 9 Nazca a Puquios). PE	14° 49' 60" S 74° 52' 60" W	M842 – M846
M. theresiae	tesLY	Punta Lachay (Km 105 Panam N.)(Lima). PE	11° 17 45" S 77° 34 35" W	M779 – M780
	tesAi	Asia, Km 96.5 Panam S. al lado de la carretera, Depto Lima. PE	12° 45' 13.5" S 76° 36' 26.0" W	M856 – M857
	tesK514S	Km 514 Panam S. al sur de Nazca. PE	15° 23' 58.2" S 74° 53' 13.9" W	M1034 – M1035
***Occipitalis* group**				
M. occipitalis	occSRM	Guayas, Canton Santa Helena, Santa Rosa de Murillo.	02° 13 0" S 80° 58' 0" W	QCAZ 6057

		EC		
	occK1245N	Caleta Grau (Km 1245 Panam N), Tumbes. PE	03° 40' 0" S 80° 39' 0" W	M922 – M927
	occTAL	Puerto Talara (118 km N de Piura). Piura. PE	04° 34' 33" S 81° 16' 9" W	M784
	occK1049N	Salida de Piura (Km 11 Norte; Piura Km 1038 Panam. N). PE	05° 12' 0" S 80° 37' 60" W	M765
	occK777N	Salida de Chiclayo (Km 777 Panam N), Lambayeque. PE	06° 44' 12.5" S 78° 52' 51.3" W	M889 – M890
	occJEQ	Jequetepeque (N de La Libertad), PE	07° 20' 32.3" S 79° 34' 46.2" W	M809-M613
M. koepckeorum	kopPAR	Pariakoto (Km 56 Casma-Huaraz, Depto Ancash). PE	09° 33' 43.5" S 77° 53' 21.7" W	M729 – M731
M. stolzmanni	stz	Salida de Jaen. PE	05° 42' 18" S 78° 48' 23" W	M705 – M706
M. indefatigabilis	indCDP	Santa Cruz, Cerro Dragon Path. EC	00° 31' 44.32" S 90° 29' 11.97" W	MSB 63535 – 63536
	indDM	Santa Cruz, Islet Daphne Major. EC	00° 25' 20.38" S 90° 22' 21.68" W	MSB 64256
	indSY	Santa Cruz, Isla Seymour Norte Western bench. EC	00° 23' 33.14" S 90° 16' 40.94" W	MSB 68904
	indFE	Isla Santa Fe: Isla Santa Fé, NE corner of island at tourist site, on beach. EC	00° 48' 11.98" S 90° 02' 28.17" W	MSB 69824
	indED	Santa Cruz, Islote Eden. EC	00° 33' 42.87" S 90° 32' 13.34" W	MSB 63627 – 63628
M. albemarlensis	albBU	Isabela, Volcan Alcedo, Bahia Urvina. EC	00° 23' 37.39" S 91° 13' 49.90" W	MSB 63579 – 63579
	albTC	Isabela,Volcan Darwin, W slope, E of Tagus Cove. EC	0° 14' 26.23" S 91° 21' 27.50" W	MSB 63595 – 63596
	albCW	Islote Cowley. EC	00° 23' 5.20" S 90° 57' 46.90" W	MSB 63552 – 64229
	albBEH	Bahia Elizabeth, Isabela. EC	00° 35' 47.68" S 91° 04' 15.99" W	MSB 63287
M. pacificus	pacPP	Isla Pinta, Playa posada, costa Sur W/plaque. EC	00° 36' 42.48" S 90° 47' 11.83" W	M1323 – 1324
M. duncanensis	dun	Isla Pinzon, Islote Onan, Playa Escondida frente a placa. EC	00° 35' 56.11" S 90° 39' 16.48" W	M1428 – 1429
M. jacobi	jacBA	Santiago, Isla Bartolome. EC	00° 17' 7.87" S 90° 33' 17.35" W	MSB 63648 – 63649
	jacSCH	Santiago, Sombrero Chino. EC	00° 21' 58.57"S 90° 35' 9.20" W	MSB 63672
	jacBR5	Santiago, Bainbridge Rock #5. EC	00° 21' 50.70" S 90° 33' 59.79" W	MSB 63660
	jacBC	Santiago, Buccaneer Cove. EC	00° 9' 52.30" S 90° 49'17.07" W	MSB 63895 – 63886
M. delanonis	del	Española. EC	1° 21' 23.36" S 89° 39' 8.92" W	MSB 63550 – 63549
	del	Espanhola: Isla Gardner, coast above W cove (anchorage) S of cliff. EC	1° 20' 40.09" S 89° 35' 50.52" W	MSB 70195 – 70206
	delIO	Espanhola, Isote Oeste. EC	1° 15' 40.75" S 89° 39' 44.17" W	MSB 69783 – 69785
M. grayii	graF	Floreana. EC	1° 15' 40.75" S 90° 22' 6.92" W	MSB 63610 – 63611
	graEN	Floreana, Islote Enderby. EC	1° 13' 56.64" S 90° 37' 43.68" W	MSB 63622 – 63623

M. bivittatus	bivSC	San Cristobal. EC	00° 53' 59.1" S 89° 33' 56.95" W	MSB 63708 - 63709
M. habeli	habPM	Isla Marchena, Playa de los Muertos. EC	00° 18' 1.83" S 90° 30' 28.61" W	M1318 - 1319
	habNB	Isla Marchena, North Beach. EC	00° 23' 17.95" S 90° 28' 25.67" W	M1313 - 1314
OUTGROUP				
T. insulanus		Vila Rica, MT, Brasil		MZUSP 978131
T. oreadicus		Porto Nacional, TO, Brazil		MZUSP LG1053
T. plica		Aripuana, MT, Brazil		MZUSP 968014

Appendix 2. Primers used to amplify and sequence mtDNA and nuclear gene regions from *Microlophus* and *Tropidurus*; abbreviations for gene regions are defined in the text. Positions of the mitochondrial primers refer to the *Iguana iguana* mitochondrial genome (Janke et al., 2001).

Primer	Primer sequence 5'-3'	Source
mtDNA		
ND4 F (10807-10838)	CACCTATGACTACCAAAAGCTCATGTAGAAGC	Arévalo et al., 1994
ND4 R (11733-11759)	CATTACTTTTACTTGGATTTGCACCA	Arévalo et al., 1994
16s F.1 (1917-1942)	TGT TTA CCA AAA ACA TMR CCT YTA GC	Whiting et al., 2003
16s R.0 (2495-2516)	TAG ATA GAA ACC GAC CTG GAT T	Whiting et al., 2003
Cyt-b WWF	AAAYCAYCGTTGTWATTCAACTAC	A. Whiting (unpubl.)
Cyt-b C3	GGCAAATAGGAARTATCATTC	Palumbi 1996
12s tPhe	AAAGCACRGCACTGAAGATGC	Wiens et al., 1999
12s 12e	GTRCGCTTACCWTGTTACGACT	Wiens et al., 1999
nDNA		
Enol-H	CCAGGCACCCCAGTCTACCTGGTCAAA	Friesen et al., 1997
Enol-L	TGGACTTCAAATCCCCCGATGATCCCAGC	Friesen et al., 1997
GapdH	CATCAAGTCCACAACACGGTTGCTGTA	Friesen et al., 1997
GapdL	ACCTTTAATGCGGGTGCTGGCATTGC	Friesen et al., 1997
RP40-F	GGGCCTGATGTGGTGGATGCTGGC	Friesen et al., 1999
*RP40.1F	ATTTCCCGTGAACATCCATGG	This study
[1]**RP40intFW1	GCTCCGTACTGCTATGAGCGG	This study
RP40-R	GCT TTCTCAGCAGCAGCCTGCTC	Friessen et al., 1999
*RP40.2R	AGTATCACTTGAAATTACAC	This study
[1]**RP40intrev2	GGCAGCTCYGCGACTTACCTTTCC	This study
Anon (ATRPL.1F)	TTGGACCAGTGACRTTAGCCAG	This study
Anon (ATRPL.1R)	TGAAGAGGCTGACCAAGACCCTA	This study
α-ATRP-F	GAG TTGGATCGCGCTCAGGAGCG	Friessen et al., 1999
ATRP-2F	GCCACTGCKCTRCAGAAGCTGG	This study
α-ATRP-R	CGGTCAGCCTCYTCAGCAATGTGCTT	Friesen et al., 1999
ATRP-2R	CCTCTTCAGCAATGTGCTTGGCC	This study
Cryba-Fr1	CGCCTGATGTCTTTCCGCC	Dolman and Phillips, 2004
Cryba-Fr2	CCAATGAAGTTCTCTTTCTCAA	Dolman and Phillips, 2004
[2]**Cryba-Fr2.1r	TGCTTAAATCCGGCCCACTGAAGA	This study
Rag-1F	AGYCAR TAYCAAYAARATGTA	D. Vietes UC Berkeley
Rag-1R	GCRTTNCCDATRTCRCARTG	D. Vietes UC Berkeley
Cmos-73	GCGGTAAAGCAGGTGAAGAAA	Saint et al., 1998
Cmos-78	AGRGTGATRGCAAAVGARTARATG	Saint et al., 1998
Ck-6	GAC CAY MCC MGA GTSA TCT CCA TG	Palumbi, 1996
Ck-7	SAGG YGY TCR TTC CAC ATG AA	Palumbi, 1996
[3]**Ckrevlseq	TGY GCC TGG TCC CCA CAG GCC	This study

*Optimized PCR primers for *Microlophus* and *Tropidurus*.

[1]** Additional sequencing primers for the complete amplification of the RP40 gene in *M. occipitalis*, *M. bivittatus*, and *M. habeli* (*Occipitalis* group in part).

[2]**Additional sequencing primer for the complete sequence of CRYBA in *M. thoracicus, M. theresiae*, *M. peruvianus, M. heterolepis, M. tigris, M. yanezi, M. theresioides, M. quadrivittatus*, and *M. atacamensis* (*Peruvianus* group).

[3]**Additional sequencing primers for the complete amplification of the CK gene in *M. peruvianus, M. heterolepis, M. tigris, M. yanezi, M. theresioides, M. quadrivittatus, and M. atacamensis* (*Peruvianus* group, in part).

Appendix 3. PCR conditions for the amplification of four mitochondrial and nine nuclear gene regions used in this study.

Gene region	PCR cycle
mtDNA	
ND-4 (+ tRNAs)	95º C for 1 min, 50º for 1 min, 72º C 1 min (35 cycles)
Cyt-*b*	95º C for 1 min, 48º for 1 min, 72º C 1 min (35 cycles)
16s	94º C for 1 min, 50º for 1 min, 72º C 1 min (35 cycles)
12s	94º C for 1 min, 50º for 1 min, 72º C 1 min (35 cycles)
nDNA	
Gapdh	95º C for 1 min, 52º for 1 min, 72º C 1 min (40 cycles)
Cmos	94º C for 1 min, 53º for 1 min, 72º C 1 min (40 cycles)
Enol	94º C for 1 min, 53º for 1 min, 72º C 1 min (40 cycles)
Cryba	95º C for 1 min, 60º for 1 min, 72º C 1 min (40 cycles)
Ck	94º C for 1 min, 58º for 1 min, 72º C 1 min (40 cycles)
RP40	95º C for 1 min, 50º for 1 min, 72º C 1 min (40 cycles)
Atrp	94º C for 1 min, 60º for 1 min, 72º C 1 min (40 cycles)
Anon	95º C for 1 min, 60º for 1 min, 72º C 1 min (40 cycles)
RAG	94º C for 1 min, 55º for 1 min, 72º C 1 min (40 cycles)

Appendix 4 – Summary of GenBank accession nos. for each gene and taxon (online).

Appendix 5. -List of length mutations inferred by manual alignment of seven nuclear introns. Abbreviations are: ins = insertion; del = deletion; SSR = simple sequence repeat; del^{ni} = non-independent origin of overlapping indel; del^{i} = independent origin of overlapping indel; del* = secondary deletion. Eastern Galapagos = *M. habeli* + *M. bivittatus*; Western Galapagos = *M. albemarlensis, M. jacobi, M. indefatigabilis*, *M. duncanensis, M. pacificus*, *M. grayii, M. delanonis.* The complete data matrix is available online (Appendix 6A). Bolded rows indicate indels inferred by both, manual and progressive (PRANK) approaches.

Intron	Indel no.	Kind of length mutation	Position in combined nuclear matrix	Length	Indel sequence or motif	Clades or taxa sharing the indel
CK (11)	1	SSR	89-94	6	TCC	Chilean taxa, *M. peruvianus M. thoracicus*
	2	Ins	89-108	14	TCACCTCCCCCTCC	All *Microlophus*
	3	**SSR**	**92-94**	**3**	**TCC**	***M. thoracicus* M, P, TB**
	4	Del	187-187	1	C	*T. insulanus, T. oreadicus*
	5	Del	193-194	2	CG	*T. plica*
	6	**Del**	**251-295**	**45**	CGTGTGTGTTTGCAGGGGCA GGGGCCTTCCTCCCCTGGGCTT	***M. thoracicus***
	7	**Ins**	**253-254**	**2**	**AT**	
	8	**Del**	**253-256**	**4**	**ATGT**	***M. koepckeorum***
	9	Del	267-267	1	-	*M. occipitalis*
	10	Del	273-273	1	-	*M. stolzmanni* + *M. theresiae*
	11	Del	273-276	4	-	Eastern and Western Galapagos
Cryba (51)	12	Del	416-433	18	TTAAGGAATGGGAGATGA	*T. oreadicus*
	13	**Del**	**418-418**	**1**	**-**	***M. theresiae***
	14	Del	424-431	8	-	*M. duncanensis*
	15	SSR	429-431	3	GAT	*M. occipitalis*
	16	**Del**	**460-473**	**14**	**GGCCGACGGGGCAG**	***Occipitalis* group**
	17	Del	482-482	1	-	*M. theresiae*

18	**Ins**	**512-522**	**11**	**AGAGAGCGTGC**	***Microlophus***
19	Del	513-694	182	-	*M. stolzmanni*
20	**Del**	**519-527**	**9**	**GTGCTTTCT**	***M. koepckeorum***
21	**Del**	**576-586**	**12**	**-**	***T. insulanus***
22	Del	576-587	11	-	*T. oreadicus*
23	**del** [ni]	**582-594**	**13**	**GTGGCTTTGGCTA**	***M. bivittatus* + *M. occipitalis***
24	del [ni]	582-602	21	GTGGCTTTGGCTAGCGGCTGG	*M. habeli* NB + PM
25	Del	594-684	91	-	*Tropidurus*
26	Del	615-616	2	CA	*M. koepckeorum*
27	del [ni]	631-637	7	ATCTCCA	*M. theresiae* LY (Northern range)
28	del [ni]	631-636	6	ATCTCC	*M. theresiae* Ai + K514S (Southern range)
29	Del	638-694	57	-	*Occipitalis* except *M. stolzmanni*
30	Del	652-652	1	-	*M. thoracicus* E
31	Del	688-688	1	-	*Tropidurus*
32	Del	694-694	1	-	*Tropidurus*
33	Del	701-705	5	CTTAT	*Tropidurus*
34	Del	701-1157	456	-	*M. stolzmanni*
35	Del	715-717	3	-	*Tropidurus*
36	Del	722-758	37	-	*Tropidurus*
37	Del	740-760	20	CCTGGTTATTCCTCTGCCGT	*M. koepckeorum*
38	Del	775-775	1	-	*Tropidurus*
39	Del	778-789	12	-	*Tropidurus*
40	Del	778-792	15	TTTCAAATGTCCCGG	*M. grayii* EN + F
41	**Del**	**811-817**	**7**	**ACTTTTG**	***M. theresioides* (CL + PI) + *M. yanezi* + *M. heterolepis***
42	Del	812-816	5	-	*Tropidurus*
43	Del	818-818	1	-	*Tropidurus*
44	Del	824-824	1	-	*Tropidurus*
45	Del	833-836	4	TTAC	*Tropidurus*
46	Del	860-870	11	TCGCTTGTACT	*Tropidurus*
47	Del	881-886	6	TTTTTG	*Tropidurus*
48	**Del**	**895-901**	**7**	**TAAATTA**	***M. occipitalis* + Eastern Galapagos**
49	Del	906-908	3	-	*Tropidurus*
50	**Del**	**919-970**	**52**	**-**	***Occipitalis* group – *M. stolzmanni***
51	Del	923-923	1	-	*M. thoracicus* M, P, TB (southern range) + E (Central range)
52	Del	963-973	11	-	*Tropidurus*

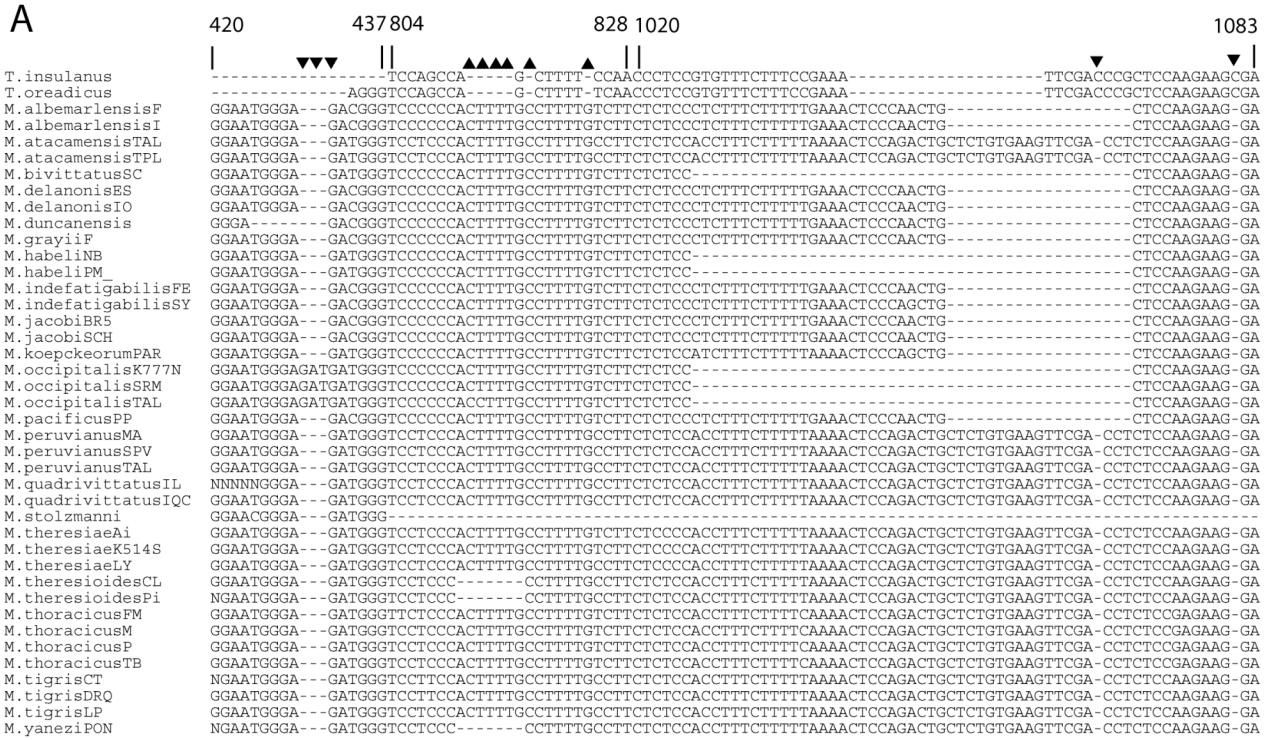
	53	Ins	1001-1004	4	-	*Occipitalis* group minus *M. stolzmanni*
	54	Del	1012-1013	2	-	*Microlophus*
	55	del[ni]	1026-1070	45	-	*M. occipitalis* + Eastern Galapagos
	56	Del	1042-1061	20	-	*Tropidurus*
	57	del[ni]	1052-1070	19	-	Western Galapagos
	58	Ins	1067-1067	1	-	*Tropidurus*
	59	Ins	1081-1081	1	-	*Tropidurus*
	60	Ins	1088-1089	2	-	*Tropidurus*
	61	**Del**	**1118-1131**	**14**	**CCGGATTTAAGCAG**	**Eastern Galapagos**
	62	**Del**	**1142-1149**	**8**	**-**	***Tropidurus***
RP40 (14)	63	Del	1208-1208	1	-	*M. theresiae*
	64	Del	1265-1265	1	-	*T. plica*
	65	**Del**	**1282-1289**	**8**	**-**	***T. plica***
	66	**Del**	**1288-1288**	**1**	**-**	***T. insulanus***
	67	Del	1301-1301	1	-	*T. plica*
	68	**Del**	**1320-1345**	**26**	**-**	***T. insulanus***
	69	**Del**	**1358-1362**	**5**	**-**	***T. insulanus***
	70	**Ins**	**1372-1818**	**447**	**-**	***M. occipitalis* + Eastern Galapagos**
	71	del*	1512-1550	39	-	Eastern Galapagos. *deletion within insertion
	72	**del***	**1732-1792**	**61**	**-**	**Eastern Galapagos. *deletion within insertion**
	73	Del	1836-1837	2	-	*M. habeli* (NB + PM)
	74	**Ins**	**1866-1867**	**2**	**-**	***T. plica* + *T. insulanus***
	75	**del[ni]**	**1909-1947**	**39**	**-**	***M. occipitalis* + Eastern Galapagos**
	76	**SSR**	**1930-1932**	**3**	**AGA**	***M. stolzmanni***
Enol (6)	77	Del	2056-2192	137	-	*T. insulanus*+ *T. oreadicus*
	78	**Del**	**2066-2066**	**1**	**-**	***M. theresiae***
	79	**del[ni]**	**2088-2125**	**38**	**-**	***M. occipitalis* + Eastern Galapagos**
	80	Del	2106-2106	1	-	*M. theresiae*
	81	**del[i]**	**2118-2125**	**8**	**TTCCCAGT**	**Western Galapagos**
	82	**Ins**	**2189-2192**	**4**	**AAAA**	***T. plica***
Gapd (11)	83	Del	2362-2374	13	-	*T. plica*
	84	**Del**	**2412-2419**	**8**	**GAATAGTT**	**Chilean clade + *M. peruvianus* (Southern range)**
	85	**del[i]**	**2490-2500**		**TGAGGAGATAG**	***M. indefatigabilis* FE (Isla Santa Fe)**
	86	del[i]	2493-2498	8	CGAGA	*M. thoracicus* except Southern range (M, P, TB + FM)
	87	Del	2506-2546	41	-	*Occipitalis* group

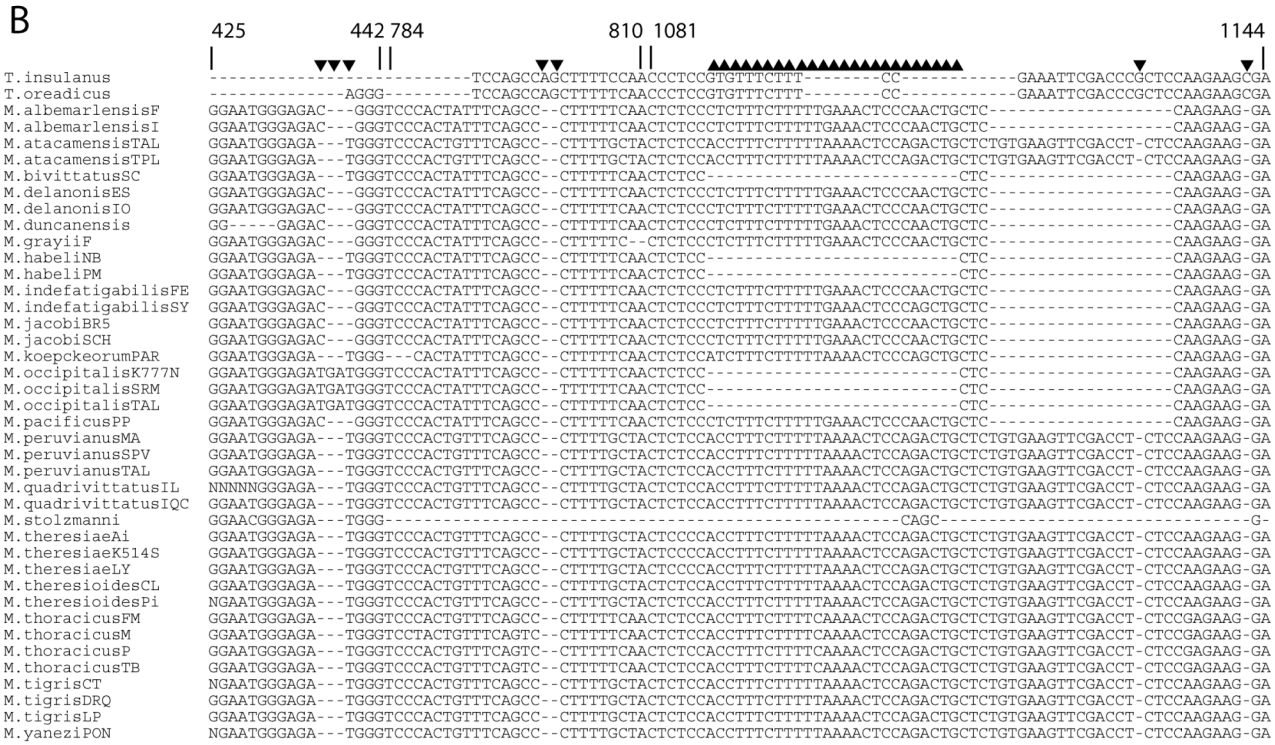
	88	Del	2508-2517	10	-	*T. insulanus*
	89	Del	2527-2530	4	-	*T. plica*
	90	Del	2546-2546	1	-	*T. insulanus*
	91	Del	2555-2555	1	-	*T. insulanus*
	92	**Del**	**2559-2565**	**7**	**-**	**Chilean clade + *M. peruvianus* (Southern range)**
	93	**Del**	**2569-2575**	**7**	**-**	***T. insulanus***
ANON (12)	94	Ins	2686-2686	1	-	*Peruvianus* group except *M. thoracicus*
	95	**Del**	**2699-2702**	**4**	**TGGT**	***M. delanonis***
	96	**Del**	**2721-2730**		**TGTCACATTGT**	***M. delanonis***
	97	Del	2789-2793	5	CTGAA	*Occipitalis* group
	98	**Ins**	**2801-2801**	**1**	**-**	***Microlophus***
	99	Del	2808-2810	3	CAT	*M. koepckeorum*
	100	**Ins**	**2845-2845**	**1**	**-**	***Microlophus***
	101	Ins	2868-2869	2	GC	*M. peruvianus* K262N, TAL. MA, and ACA (Northern clade)
	102	**Del**	**2932-2964**	**33**	**TTTCATTCC-GGGGTCATGGATGCCCCCCGTCA**	***M. stolzmanni***
	103	Del	2933-2965	34	TTTCATTCC-GGGGTCATGGATGCCCCCCGTCAT	*Occipitalis* group except *M. stolzmanni*
	104	**Ins**	**2941-2941**	**1**	**A**	***M. thoracicus***
	105	**Ins**	**2988-2988**	**1**	**-**	***Microlophus***
ATRP (10)	106	Ins	3148-3153	5	TTGGGT	*T. insulanus*
	107	Del	3148-3154	6	TTGGGTG	*Microlophus*
	108	**Ins**	**3167-3167**	**1**	**-**	***Microlophus* except *M. stolzmanni and M. koepckeorum***
	109	Ins	3172-3172	1	-	*Microlophus*
	110	Del	3196-3209	12	GCTGCCCAAGGCGT	*M. occipitalis* SRM (Ecuador)
	111	Del	3207-3218	10	CGCGCTCCCCTC	*T. plica*
	112	Del	3217-3217	1	-	*Peruvianus* group except *M. thoracicus*
	113	Ins	3238-3240	3	GGC	*M. occipitalis* SRM (Ecuador)
	114	Ins	3245-3245	1	-	*Occipitalis* group
	115	Del	3245-3250	6	ACCCCG	*M. grayii* EN (Enderby islet)

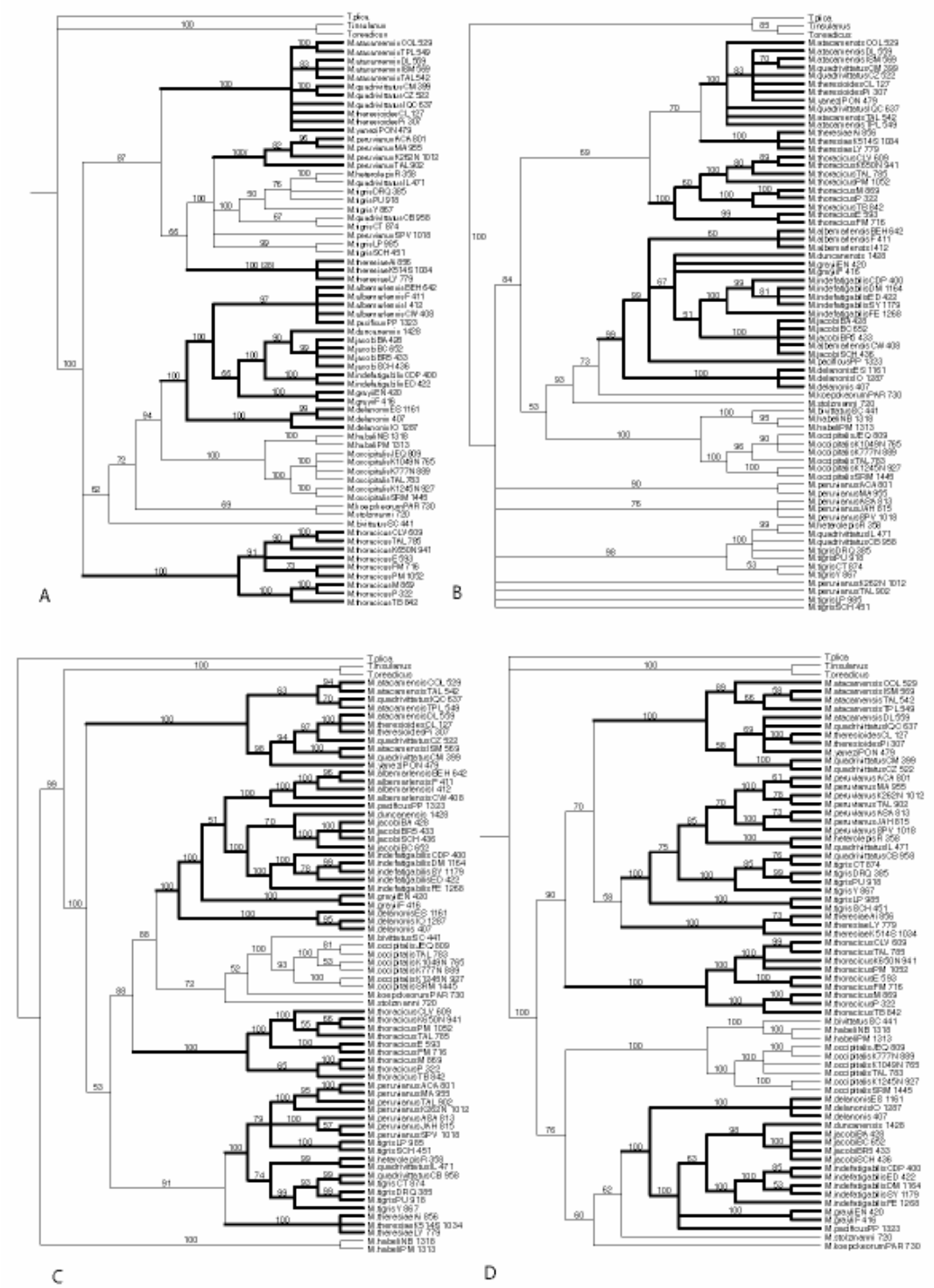
Appendix 6A. Data matrix of nuclear genes aligned by hand (4691 bp). [PAUP FILE_online]

Appendix 6B. Data matrix of nuclear genes aligned using the progressive method implemented in PRANK (4755 bp) (Loytynoja and Goldman, 2005) [PAUP FILE_online]

Appendix 7. Selected Cryba intron regions that show disparate alignment solutions. A. Manual alignment based on the procedure outlined in Fig. 3. B. Progressive alignment using the Progressive alignment kit (PRANK) (Loytynoja and Goldman, 2005).







Appendix 8A. 50% majority rule consensus Bayesian MCMC trees for the 4 largest mitochondrial partitions. Numbers above branches are posterior probability values. Each tree was run for 10 x 10^6 generations with 3 (cyt-*b*; ND4) or 2 (12s; 16s) partitions. Models used are shown in Table 1. Gene trees are: **A**. 12s; **B** 16s; **C** Cyt-*b*; **D** ND4. Bolded branches represent nodes also present in the combined data set in Fig. 6.

Appendix 8B.-Summary of posterior probabilities values (PP) for clades recovered through Bayesian MCMC searches of four of the five single mtDNA gene regions, and the combined mitochondrial data set (2991 bp) that includes $tRNA^{his}$ and $tRNA^{ser}$ genes as a single partition (135 bp). Bolded rows correspond to clades of importance identified in the combined data analyses that were alternatively found in the Bayesian MCMC searches of some or all individual genes.

Node	Description of node	12s	16s	ND4	Cyt-*b*	Combined data
1	*Microlophus* (node 58 + node 2)					1.0
2	Node 3 + node 30					61
3	**Node 4 + node 14**	**.87**		**.70**		**1.0**
4	**Node 8 + node 5**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
5	M. quadrivittatusIQC +node 6				.63	.99
6	M. atacamensisTAL + node 7					.58
7	M. atacamensisTPL + M. atacamensisCOL	1.0			.94	.61
8	Node 9 + node 13					1.0
9	M. yaneziPON + node 10					.62
10	M. quadrivittatusCZ + node 11				.97	.97
11	M. atacamensisDL +node 12				1.0	1.0
12	M. theresioidesCL + M. theresioidesPi			1.0		1.0
13	M. atacamensisISM + M. quadrivittatusCM				1.0	1.0
14	Node 15 + Node 14	92		.58	1.0	92
15	Nodes 16 + 21 + 27	**1.0**	**.76**	**1.0**		**1.0**
16	Node 20 + node 17 [*M. peruvianus*]			.70		53
17	M. peruvianusK262N + node 18	1.0			1.0	1.0
18	M. peruvianusTAL + node 19	.82			.95	.95
19	M. peruvianusACA + M. peruvianusMA	.96		.61	1.0	1.0
20	M. peruvianusASA, JAH , SV			1.0	1.0	1.0
21	Node 22 + node 23	1.0			.74	1.0
22	M. heterolepisR + M. quadrivittatusIL	1.0	.99	1.0	.99	1.0
23	M. tigrisY + node 24	76		1.0	..99	1.0
24	Node 25 + node 26			.85	.93	1.0
25	M. quadrivittatusCB + M. tigrisCT	.67		.76	.99	.97
26	M. tigrisDRQ + M. tigrisU			.99	.99	1.0
27	M. tigrisLP + M. tigrisSCH	.99		1.0	1.0	1.0
28	***M. theresiae***	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
29	M. theresiaeAi + M. theresiaeLY			.73		.73
30	***Occipitalis* group**	**.94**	**.83**		**.88**	**1.0**
31	**Western Galapagos radiation**	**1.0**	**.99**	**1.0**		**1.0**
32	Nodes 33 + 37 + 45		**.99**		**1.0**	**1.0**
33	*M. pacificus* + node 34 [*M. albemarlensis*]	.97			1.0	1.0
34	M. albemarlensisCW + node 35				1.0	1.0
35	M. albemarlensisI +node 36				1.0	1.0
36	M. albemarlensisBEH + M. albemarlensisF				.96	.90
37	Node 38 + node 41			.63	1.0	1.0

38	*M. duncanensis* + node 39 [*M. jacobi*]	.90	1.0	.98	70	1.0
39	M. jacobiBC + node 40				1.0	1.0
40	M. jacobiBA + BC + BR5	.99	1.0	1.0	1.0	1.0
41	M. indefatigabilisFE + node 42		.99	1.0	1.0	1.0
42	M. indefatigabilisCD + node 43		1.0		1.0	.96
43	M. indefatigabilisED + node 44				1.0	1.0
44	M. indefatigabilisDM + M. indefatigabilisSY			.53	1.0	1.0
45	M. grayiiEN + M. grayiiF	1.0		1.0	1.0	1.0
46	M. delanonisIO + node47	1.0	1.0	1.0	1.0	.61
47	M. delanonisES + M. delanonis407	.99		1.0		.60
48	*M. koepckeorum* + node 49			.60		.87
49	*M. stolzmanni* + node 50			.87		1.0
50	**Eastern Galapagos radiation**		**1.0**	**1.0**		**1.0**
51	M. bivittatusSC + node 52		1.0	1.0		1.0
52	M. habeliNB + M. habeliPM	1.0	.95	1.0	1.0	1.0
53	Node 57 + node 54 [*M. occipitalis*]	1.0	1.0	1.0		1.0
54	M. occipitalisTAL + node 55		.96	1.0		.92
55	M. occipitalisK1049N + node 56		.90		1.0	1.0
56	M. occipitalisJEQ + M. occipitalisK777N			1.0		1.0
57	M. occipitalisK1245N + M. occipitalisSRM	1.0	1.0	1.0	1.0	1.0
58	***M. thoracicus***	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
59	Node 63 + node 60	.91		1.0	1.0	1.0
60	M. thoracicusM + node 61		1.0	1.0		1.0
61	M. thoracicusK650N + node 62	.90	.80	1.0	.55	1.0
62	M. thoracicusCLV + M. thoracicusTAL	1.0		.99		1.0
63	M. thoracicusE + M. thoracicusFM (center)		.99	1.0	1.0	1.0
64	M. thoracicusTB + node 65	1.0	1.0	1.0	65	1.0
65	M. thoracicusP + M. thoracicusM	1.0	1.0	1.0	1.0	1.0

Appendix 9. Summary of posterior probability support values for clades recovered through Bayesian MCMC searches for 9 nuclear gene regions aligned manually. Also shown are results for the Bayesian MCMC of manually aligned combined data (4691 characters), and the same matrix including coded length polymorphisms inferred from manual indel alignments (4808 characters [115 binomial characters]). *Heuristic alignments inferred by PRANK (4755 characters) plus coded length polymorphism (4886 characters [131 binomial characters]). Bolded nodes are discussed in the text and highlighted in Figure 7.

Node	Node description	Anon	Atrp	CK	Cmos	Cryba	Enol	Gapdh	Rag-1	RP40	4808	4691	4755*	4886*
1	Node 2 + node 35	1.0		.96				1.0	1.0		1.0	1.0	1.0	1.0
2	**Node 27 + node 3**	**1.0**		**1.0**	**1.0**			**1.0**	**1.0**		**1.0**	**1.0**	**1.0**	**1.0**
3	**Node 25 + node 4**	**.96**	**1.0**			**1.0**				**.93**	**1.0**	**1.0**	**1.0**	**1.0**
4	**Node 18 + node 5**		**.73**		**1.0**	**.99**					**1.0**	**1.0**	**1.0**	**1.0**
5	Node 16 + node 6										91	.77	.66	.78
6	M. tigrisY + node 7					1.0					1.0	1.0	1.0	1.0
7	*M. heterolepis* + node 8					1.0					1.0	1.0	1.0	1.0
8	M. theresioidesPi + node 9					.93					.92	.99	.99	.99
9	M. theresioidesCL + node 10										.92	1.0	1.0	1.0
10	M. yaneziPON + node 11										.86	.93	.93	.93
11	M. quadrivittatusIL + node 12					1.0					1.0	1.0	1.0	1.0
12	M. atacamensisDL + node 13										1.0	1.0	1.0	1.0
13	M. quadrivittatusCM + node 15 + node 14										.87	.93	.93	.92
14*	M. atacamensisCOL+ ISM+TAL+TPL					.99					1.0	1.0	1.0	1.0
15*	M. quadrivittatusCZ + IQC +CB										.93	.99	.99	.99
16	M. tigrisPU + node 17					1.0					1.0	1.0	1.0	1.0
17	M. tigrisCT + M. tigrisDRQ										1.0	1.0	1.0	1.0
18	**M. tigrisSCH + node 19**										**1.0**	**1.0**	**1.0**	**1.0**
19	M. tigrisLP + node 20										.92	.79	.63	.81
20	M. peruvianusASA + node 21										.98	.96	.84	.95

21	Node 20 + node 24										.81	.82	.54	.80
22	M. peruvianusK262N + node 23										1.0	.95	.62	.94
23*	M. peruvianusTAL+MA+ACA	.99						.93			1.0	1.0	1.0	1.0
24	M. peruvianusJAH + M. peruvianusSPV	1.0							.96		1.0	1.0	1.0	1.0
25	**M. theresioidesK514S + node 26**	**1.0**	**70**	**1.0**	**1.0**	**1.0**	**.99**			**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
26	M. theresioidesAi + M. theresioidesLY	97		1.0				.98		.97	1.0	1.0	1.0	1.0
27	**Node 33 + node 28**	**1.0**	**.96**	**1.0**	**1.0**		**.93**	**1.0**	**1.0**		**1.0**	**1.0**	**1.0**	**1.0**
28	M. thoracicusE + node 29									1.0	1.0	1.0	1.0	1.0
29	M. thoracicusFM + node 30							1.0			.89	.77	1.0	.79
30	M. thoracicusPM + node 31	.90									1.0	1.0	1.0	1.0
31	M. thoracicusK650N + node 32			.99							1.0	1.0	1.0	1.0
32	M. thoracicusCLV + M. thoracicusTAL									1.0	1.0	1.0	1.0	1.0
33	M. thoracicusTB + node 34	.84		.99		1.0					1.0	1.0	1.0	1.0
34	M. thoracicusM + M. thoracicusP	.50			.98						1.0	1.0	1.0	1.0
35	***M. stolzmanni* + node 36**	**1.0**		**1.0**				**1.0**	**1.0**		**1.0**	**1.0**	**1.0**	**1.0**
36	**Node 37 + node 55**										**.98**	**.97**	**.95**	**.97**
37	**Node 38 + node 53**					**1.0**					**1.0**	**1.0**	**1.0**	**1.0**
38	**M. albemarlensisCW + node 39**							**.97**			**1.0**	**1.0**	**1.0**	**1.0**
39	Node 40 + node 48										.96	.95	.94	.95
40	Node 43 + node 41										1.0	.99	1.0	.99
41	M. albemarlensisI + node 42										.66	70	.75	.69
42	M. albemarlensisF + M. albemarlensisBEH										1.0	1.0	1.0	1.0
43	Node 44 + node 46										.99	.99	.99	.99
44	M. jacobiSCH + node 45										.98	.99	.99	.98
45	M. grayiiEN + M. grayiiF	1.0								1.0	1.0	1.0	1.0	1.0
46*	*M. jacobi* + *M. pacificus*										.93	.90	.88	.90
47	M. jacobiBC + jacobiBR5										.96	.97	.96	.97
48	M. indefatigabilisFE + node 49										1.0	1.0	1.0	1.0
49	*M. duncanensis* + node 50										.96	.96	.97	.96
50	Node 51 + node 52										1.0	1.0	1.0	1.0
51	M. indefatigabilisED + M.indefatigabilisCDP										.86	.84	.87	.85
52	M. indefatigabilisDM + M.indefatigabilisSY										1.0	1.0	1.0	1.0
53	M. delanonis407 + node 54					1.0	.96				1.0	1.0	1.0	1.0
54	M. delanonisES + M. delanonisIO			1.0		.97					.65	.64	.65	.67
55	*M. koepckeorum* + node 56								87		.90	.93	.95	-
56	**Node 57 + node 59**		**1.0**	**1.0**	**.99**		**.88**		**.99**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
57	M. bivittatusSC + node 58	1.0	1.0	1.0	.97	1.0		.99		1.0	1.0	1.0	1.0	1.0

58	M. habeliNB + M. habeliPM					.96	.99		1.0	1.0	1.0	1.0	1.0
59	Node 60 + node 61		1.0	1.0	.95		1.0	1.0	1.0	1.0	1.0	1.0	1.0
60	M. occipitalisJEQ + M. occipitalisK777N							1.0		.97	.98	.97	.98
61	Node 62 + node 63									1.0	1.0	1.0	1.0
62	M. occipitalisK1049N + M. occipitalisTAL	.94								1.0	1.0	1.0	1.0
63	M. occipitalisSRM + M. occipitalisK1245N							1.0		1.0	1.0	1.0	1.0

SMALL-SCALE PHYLOGEOGRAPHY OF GALAPAGOS LAVA LIZARDS (TROPIDURIDAE: *MICROLOPHUS*): PATTERNS OF COLONIZATION AND DEMOGRAPHIC HISTORY OF ISLANDS WITHIN ISLANDS

EDGAR BENAVIDES[1], REBECCA BAUM[2], HOWARD L. SNELL[3,4], HEIDI K. SNELL[3,4], AND JACK W. SITES, JR.[1]

[1]*Department of Integrative Biology, Brigham Young University, Provo, Utah, 84602, USA. Email: eb235@email.byu.edu*

[2]*Department of Chemistry, Brigham Young University, Provo, Utah, 84602, USA.*

[3]*Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque, New Mexico 87131, USA.*

[4]*Charles Darwin Research Station. Puerto Ayora, Isla Santa Cruz, Galápagos, Ecuador.*

Corresponding author: Edgar Benavides, Department of Integrative Biology, Brigham Young University, Provo, UT 84602; ph: 801/422-2203; fax: 801/422/0090; email: eb235@email.byu.edu

ABSTRACT— The "lava lizards" (*Microlophus*) are the most extensively distributed group of terrestrial vertebrates in the Galapagos Archipelago, and in this study we tested alternative phylogenetic hypotheses for within-archipelago colonization histories of the "Western Radiation" (a clade that includes seven of the nine recognized species endemic to Galapagos). Our sampling design was extremely dense (614 animals representing 78 localities), and allowed us to also assess predicted within-island phylogeographic patterns of the same clade. We used > 1000 bp of the mtDNA cytochrome-b gene to test the general east-to-west patterns (oldest to youngest islands) predicted for both the species colonization history, and the population structure within islands. Results of the phylogenetic study are consistent with this general pattern, but reject details of previous hypotheses suggested by Heise (1998) and Kizirian et al. (2004). We present a modified colonization hypothesis based on a well-supported cyt-b gene tree. Estimates of within-island phylogeographic structure based on the same locus was not consistent with the expectation that phylogeographic structure should be better developed on the oldest islands. Younger islands revealed more complex structures, but these are also the largest and geologically most complex, and we inferred complex phylogeographic histories mediated by volcanic eruptions, sea-level fluctuations, and shifts in ocean currents. Overall, two features are the most outstanding of Galapagos phylogeography. First, the recurrent and repeated evidence of within-island long distance dispersal in three major islands. In all of them, long distance colonizations occurred repeatedly and from random sources and hardly ever from neighboring localities contradicting the stepping-stone island model of migrant exchange. Second, our results are interesting in that they reject the idea of phylogeographic structure been related solely to island age. We provide evidence to support active volcanism as a major player in the generation of genetic diversity in island environments.

Key words: Galapagos, mtDNA, phylogeography, nested clades, colonization routes, volcanism, lava lizards.

Island environments are ideal for disentangling the spatial and temporal elements of genetic differentiation, adaptation, and speciation based on some inherent conditions; (i) islands are discrete geographical entities with defined oceanic boundaries, (ii) gene flow between individual islands is reduced and directed by oceanic barriers and currents (Calsbeek and Smith, 2003; Cowie and Holland, 2006), (iii) often their small geographic size has made population surveys easier (Emerson, 2002; Blondel et al., 1999), (iv) despite their size they can show a diversity of habitats (Emerson and Oromi, 2005), and (v) island environments do have the potential to be geologically and ecologically very dynamic (Geist, 1996; Grant and Grant, 2002; Thorpe et al., 2005). Most oceanic islands exhibit increased levels of endemism, and patterns specific to a particular island group are frequently the outcome of replicated evolutionary processes with sometimes predictable outcomes (i.e., the well-documented adaptive radiations of Darwin's Finches [Grant, 1999]; the Hawaiian silversword alliance [Baldwin and Sanderson, 1998], and West Indian *Anolis* lizards [Losos et al., 1998]). More recently, some studies have also shown accelerated rates of molecular evolution in island populations due to small effective sizes relative to mainland sister taxa (Chiba, 1999; Fry, 1999; Johnson and Seger, 2000; Barrier et al., 2001; Vandergast et al., 2004; Gantenberg and Keightley, 2004).

The Galapagos Archipelago, located about 960 km due west from the coast of Ecuador, is an interesting example of oceanic island isolation and the effect of ocean currents mediating island colonization and subsequent evolutionary processes. The Galapagos Islands are not homogeneous in age, size, and plant community composition (White et al., 1993; Geist, 1996; Snell et al., 1996), and they harbour a number of vertebrates (e.g., land and marine iguanas, giant tortoises, finches, mockingbirds) and marine species that exhibit different distributions, levels of local adaptation, and obvious morphological differences within and between islands. Interestingly, very few of these groups are present on most of the 60 named islands, islets, and rocks that constitute the Archipelago, and lava lizards of the genus *Microlophus* are one of these exceptions. Lava lizards are probably the most common and widespread terrestrial vertebrates throughout the Archipelago (Jackson, 1993; Snell et al., 1996).

The genus *Microlophus* (Dumeril and Bibron, 1837) consists of a moderately sized group of Tropidurine lizards restricted to the western edge of South America. The monophyly of the group is supported by morphological synapomorphies such as the presence of apical disks on the hemipenes (Frost, 1992; Harvey and Gutberlet, 2000; Frost et al., 2001), and multiple nuclear gene regions (Benavides et al., in revision). The genus includes 22 species distributed along the Pacific coastal deserts of southern Ecuador, Peru and Chile, and it has also colonized the Galapagos Archipelago. Two main clades have been recognized as the *Peruvianus* (mainland) and the *Occipitalis* (insular and mainland) groups by Van Denburgh and Slevin (1913), Dixon and Wright (1975), and Frost (1992), and these are strongly supported by an extensive molecular data set (Benavides et al., in revision).

A total of 12 species are recognized within the *Occipitalis* group (Van Denburgh and Slevin, 1913; Frost, 1992; Heise, 1998), and the extant Galapagos *Microlophus* consists of nine species that have radiated as a consequence of two independent colonizations from the mainland. Insular *Microlophus* are not a monophyletic group as evidenced by several studies based on allozymes (Wright, 1983), immunological distances (Lopez et al., 1992), mtDNA sequences (Heise, 1998; Kizirian et al., 2004), and we have recently confirmed the "two colonization hypothesis" with multiple nuclear sequences (Benavides et al., in revision). These two colonizations include a small "Eastern Radiation" consisting of only two species endemic to the islands of San Cristobal (*M. bivittatus*) and Marchena (*M. habeli*), and a more extensive "Western Radiation" that comprises the remaining seven species that are widely distributed across almost all of the remaining islands and islets (Fig. 1). Both radiations appear to have been established on older islands in the eastern part of the archipelago (San Cristobal [Eastern Radiation] and Española [Western Radiation]), and westward sequential colonization of younger islands followed by subsequent divergence is hypothesized to be responsible for the current species diversity. No two species occupy the same island, but despite the apparent allopatric nature of the radiation, sexual selection, as suggested by pronounced patterns of sexual dimorphism (Werner, 1978; Watkins, 1996, 1997, 1998), may have accelerated and contributed to the speciation process (Snell et al., 1988; Miles et al., 2001; Panhuis, 2001; M. Jordan et al., 2005).

Beyond the issue of initial colonization of the Galapagos Archipelago, however, the sequence of dispersal events among islands is unresolved for the Western Radiation, and multiple studies are discordant with each other with regard to the number and direction of inter-island colonization events (Wright, 1983; Lopez et al., 1992; Heise, 1998; Kizirian et al., 2004; Benavides et al., in revision). A number of recent studies have used molecular data to infer the evolutionary history of Galapagos endemics (Arbogast et al., 2006 [mockingbirds]; Rassmann, 1997 [marine iguanas]; Beheregaray et al., 2004; Caccone et al., 2002 [tortoises]; Petren et al. 1999, 2005; Sato et al., 1999, 2001; Burns et al., 2002; Tonnis et al., 2005 [finches]; Bollmer et al. 2006 [hawks]; Friesen et al. 2006 [petrels]), and we complement these studies with a phylogenetic/phylogeographic investigation of the *Microlophus* Western Radiation. We focus on the Western Radiation to better understand the evolutionary events chronologically associated with the sequence of island emergence, and the history of evolutionary divergence both between and within islands in the Galapagos Archipelago.

We address two issues in this study. First, we present a mtDNA-based phylogenetic reconstruction of the Western Radiation to derive a strongly-supported gene tree, and we test this topology against previously published alternatives for the sequence of between-island colonization/speciation events. The null hypothesis for the direction of multiple colonization events is based on a biogeographic scenario of island migration in an approximately east-to-west (older vs. younger islands) direction (Cox, 1983; White et al., 1993), and two earlier studies have provided phylogenetic resolution sufficient to offer unambiguous sequences of species derivation. Second, we evaluate within-island population structure predicted by this same geological history; lizard populations should show decreasing phylogeographic structure in the same geographic direction (east-to-west) as that predicted for colonization/speciation history (Beheregaray et al., 2004). However this expectation is less clear because the simple east-to-west history of island formation does not describe the stochastic patterns of dispersal, migration, population expansion, local extinction, re-colonization, and/or secondary contact within island isolates shaped by strong volcanism (Beheregaray et al., 2003*a*; Nielsen, 2004; Ciofi et al., 2006; Brown et al., 2006; Bloor and Brown, 2005), and the often unpredictable direction of ocean currents (Beheregaray et al., 2003*b*; Rusello et al., 2005). Here we

describe intra-island phylogeographic structure and demographic history based on dense population sampling and the use of a combination of coalescent and phylogeographic methods suited to the retrieval of signal from a range of divergence events.

MATERIALS AND METHODS

Taxon and Geographic Sampling Design

Multiple populations of six of the seven species (one locality only for *M. pacificus*) of the Western Radiation were sampled, primarily through loans from the tissue collection of Galapagos *Microlophus* currently archived at the University of New Mexico (Museum of Southwestern Biology; MSB). The MSB collection was augmented by field work in 2004 designed to fill distributional gaps for species endemic to large islands or groups of islands, including populations under the names of: *M. albemarlensis* (Isabela and Fernandina Islands, and Cowley, Cuatro Hermanos, South Marielas, and Tortugas islets), *M. indefatigabilis* (Santa Cruz, Santa Fe, Seymour, and Baltra islands/islets), *M. jacobi* (Santiago, Rabida Islands, and offshore islets), *M. delanonis* (Española Island, and Champion, Gardner and Enderby islets), *M. grayii* (Floreana Island, and Xarifa, Osborn and 'Islote Oeste' islets), *M. duncanensis* (Pinzon Island), and *M. pacificus* (Pinta Island). Samples were taken non-destructively (tail tips or toe clips stored on silica or in ethanol), and lizards were then immediately released at their capture points. We obtained tissues from 614 lizards representing 78 localities (field collections plus UNM samples), represented by 1 to 20 animals per locality. These localities are plotted in Fig. 1, and details (tissue voucher numbers, name of location, coordinates, etc.) for all samples used in this study are summarized in Appendix 1. All tissue voucher samples are deposited in the Museum of Southwestern Biology (MSB).

Both empirical and theoretical studies show that under-sampling of either the number of individuals per population, or density of geographic samples within a species, may yield a false signal of genealogical exclusivity (Neigel and Avise, 1993; Hey, 1994; Templeton et al. 1995; Hedin and Wood 2002). Under simplifying assumptions of coalescent theory, the probability of sampling the deepest coalescent in a population (the

most recent common ancestor; MRCA) is given by: P = (n–1)/(n+1), where n is the sample size per population (Nordborg, 2001). A sample of five lizards/locality gives P = 0.67, and increasing n to 10, 20, and 50 individuals gives P = 0.82, 0.90, and 0.96, respectively. Thus prohibitively large sample sizes are needed to attain $\alpha = 0.05$ level of certainty of exclusivity, so we follow the sampling protocols described by Morando et al. (2003) and Carstens et al. (2004a), and attempted to sample at least five lizards per locality.

Gene Sampling and Laboratory Procedures

Both nuclear and mitochondrial genes have previously been evaluated for their informativeness at different levels of divergence within *Microlophus*, and available nuclear genes are phylogenetically uninformative at lower hierarchical levels (i.e., within species; Benavides et al., in revision). We used the cytochrome-*b* gene region (cyt-b) because the mtDNA locus is expected to sort rapidly to monophyly (Moore, 1995), and many of its limitations are well characterized (Funk and Omland, 2003; Rubinoff and Holland, 2005). The cyt-b region is especially informative at lower hierarchical levels, and our primers amplify almost the entire gene region in two sequencing reactions.

Total genomic DNA was extracted from tissues preserved in ethanol or dried on a silica gel strip, following the protocol developed by Fetzner (1999). Four microliters of extraction product were electrophoresed on 1% agarose gel to estimate the quality and amount of genomic DNA. Most of the cyt - b gene region was amplified via PCR in a cocktail containing 3.0 *ul* of template DNA (approximate concentration estimated on a 2% agarose gel), 8 *ul* dNTPs (1.25 mM), 4 *ul* 10x *Taq* buffer, 4 *ul* each primer (10 *ul*), 4 *ul* MgCl (25 mM), 22 *ul* distilled water, and 0.25 *ul Taq* DNA polymerase (5 U/ *ul*; Promega Corp., Madison, WI). A fragment of 1050 bp from the cyt *b* gene was amplified using the light strand primers WWF (5’ AAAYCAYCGTTGTWATTCAACTAC 3’) and WWR (5’ TGGCCRATGATRATAAATGGGT 3’; both from Broadley et al., 2006). Sequencing primers used included Cytb-F1 (5’ TGAGGACARATATCHTTYTGRGG 3’; Whiting et al., 2003), Cytb-C2 (5’ CCCTCAGAATGATATTTGTCCTCA 3’) and Cytb-C3 (5’ GGCAAATAGGAARTATCATTC 3’; both from Palumbi, 1996). For some

localities, additional species or island-specific primers were needed and are summarized in Table 1).

Double-stranded PCR amplified products were checked by electrophoresis on a 1% agarose gel (the size of the target region estimated using a molecular weight marker), purified using a GeneClean III kit (BIO101, Inc, Vista, CA), and directly sequenced in both directions on a Perkin Elmer ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction (PE Applied Biosystems, Foster City, CA). Excess of Dye Terminator was removed with CentriSep spin columns (Princeton Separations, Inc., Adelphia, NJ), and sequences were fractionated by polyacrylamide gel electrophoresis on a ABI PRISM 377 automated DNA sequencer (PE Applied Biosystems) at the DNA Sequencing Center at Brigham Young University.

Alignment and Phylogenetic Analyses

Forward and reverse sequences for each individual were edited and aligned using Sequencher 4.2 (Gene Codes Corporation, Ann Arbor, MI). There were no indels in the sequence and translation to amino acids verified the absence of stop codons. The program COLLAPSE v.1.1 (http://bioag.byu.edu/zoology/crandall_lab/programs.htm) was used to eliminate redundant haplotypes to reduce tree landscape space in phylogenetic analyses. The program DT-ModSel (Minin et al., 2003) was used to select substitution models for phylogenetic inference of a reduced data set. The DT-ModSel algorithm is expected to select simpler models that nevertheless provide better branch-length estimates than alternative strategies (Sullivan and Joyce, 2005). Bayesian analyses (Huelsenbeck et al., 2001) were implemented with random starting trees, run for 10 million generations, and sampled every 1000 generations. Burn-in and optimized parameter values were obtained through TRACER (Rambaut and Drummond, 2003), and equilibrium samples (9000 trees) were used to generate a 50% majority-rule consensus tree. The program PHYML (Guindon and Gascuel, 2003) was used to generate a maximum likelihood (ML) tree based on a single data partition and bootstrap values (500 replicates). The percentage of samples that recover any particular clade represents the posterior probability (*P* values) for that clade, and normally a value of $P \geq 95\%$ is taken

as for significant support for a clade (Huelsenbeck and Ronquist, 2001), and bootstrap values ≥ 70% are usually considered as evidence for significantly supported clades (Hillis and Bull, 1993; with caveats). Because Bayesian methods may resolve bifurcations with strong support when relationships are really unresolved (a polytomy is not considered as a possible outcome, see Lewis et al., 2005), we considered nodes as strongly supported only if both the PP and BS values exceeded the above values. All trees were rooted with outgroup taxa *M. bivittatus* and *M. koepckeorum*, which were constrained to monophyly (see justification in Benavides et al., in revision).

Testing Alternative Colonization Hypotheses

Figure 2 shows the alternative routes of colonization hypothesized in earlier studies that sampled all relevant island and mainland species. Wright (1983: his Fig. 11, p. 150) proposed a colonization history that attempts to reconcile geological evidence with an unresolved topology inferred from an allozyme-based genetic distance tree. In Wright's hypothesis, Española Island was colonized first, and the endemic *M. delanonis* populations provided founders for lizards on Santa Cruz (*M. indefatigabilis*), which was then the source for the colonization of the conspecific Santa Fe population, and the Santiago (*M. jacobi*), and Pinzon (*M. duncanensis*) species; the sequence in which these islands were colonized cannot be established from Wright's data. Santiago (*M. jacobi*) in turn provided for the colonization of Pinta (*M. pacificus*) and Isabela (*M. albemarlensis*) from which populations on Fernandina (*M. albemarlensis*) and Floreana (*M. grayii*) are descended. More recently, Heise (1998) and Kizirian et al. (2004) provided fully resolved colonization hypotheses for this same radiation on the basis of mtDNA sequence data (Fig. 2). Both studies also show that Española Island was colonized first (*M. delanonis*), but there are differences between these two topologies, as well as placement of some taxa in comparison to Wright's colonization scenario (i.e., *M. grayii* is the second species derived in the colonization sequence of Heise [1998: his Fig. 3.4, p. 65], but it is the last in Wright's [1983] hypothesis). Lopez et al. (1992) presented another colonization scenario for this same group based on albumen immunological distance data, but they did not sample all relevant taxa. Both the Lopez et al. (1992) and Wright (1983)

topologies are not fully resolved and thus their hypotheses are not considered further. Our best-supported topology will be used in pairwise comparisons with the two fully-resolved *a priori* hypotheses (Goldman et al., 2000) using SH99 tests (Shimodaira and Hasegawa, 1999) on likelihood trees constrained to the topologies of Heise (1998), and Kizirian et al. (2004).

Nested Clade Phylogeographic Analyses

To estimate relationships among recently diverged populations, we used nested clade phylogeographic analyses (NCPA; Templeton, 2004) to test associations of haplotypes with geography, and then to infer historical patterns of colonization and dispersal when a statistically significant association is demonstrated. Uncritical use of the NCPA with a single gene tree may lead to over-interpretation (Knowles, 2004; Knowles and Maddison, 2002), but convincing arguments have been made that this analysis is ideal for generating genealogical hypotheses that can be further tested, if inferences are cautiously made (Morando et al., 2003). Networks of related haplotypes were first constructed using the statistical parsimony algorithm (Posada et al., 2006) as implemented in the TCS program v.1.13 (Clement et al., 2000; available at: http//:inbio.byu.edu/Faculty/kac/crandall_lab/Computer.html). Nesting categories were assigned following Templeton and Singh (1993) and Templeton et al. (1987, 1995). Ambiguous connections (loops, which represent homoplasy) in some networks were resolved using predictions from coalescent theory, as validated with empirical data sets (Crandall and Templeton, 1993; Crandall, 1996). These predictions can be summarized as three criteria that can be applied to make decisions about which equally parsimonious connection to break; they include: (1) *frequency* criterion – haplotypes are more likely to be connected to haplotypes with higher frequencies than to singletons; (2) *topological* criterion – haplotypes are more likely to be connected to interior haplotypes than to tip haplotypes; and (3) *geographical* criterion – haplotypes are more likely to be connected to haplotypes from the same population or region than to haplotypes from distant populations. Non-nested networks were interconnected by calculating the absolute

number of nucleotide differences drawn from a pairwise matrix generated in PAUP* (Swofford, 2000).

The nested designs were then used for the NCPA, which was implemented with Geodis 2.5 (Posada et al., 2000; Posada et al., 2006). All statistical analyses were performed using 10,000 Monte Carlo replications, and for clades in which the null hypothesis is rejected, population processes underlying the non-random haplotype/geography association were derived from the inference key of Templeton (2004; an updated version is also available at the above website).

Cross-validating NCPA demographic inferences

In order to assess population equilibrium independent of NCPA inferences (Templeton, 2004; Benavides, 2005), we implemented Tajima's *D* (Tajima, 1989) and Fu's *Fs* (Fu, 1997) neutrality tests, both of which assume that populations are in mutation-drift and migration-drift equilibrium. Significant values for either may indicate that sequences are not evolving in a neutral manner (they are not in mutation-drift equilibrium) or that they were previously subdivided and/or have experienced past demographic growth (they are not in migration-drift equilibrium). The *Fs* test appears to be more sensitive to detection of population expansion (Fu, 1997), and we evaluated the significance of this test by comparing the statistic against a distribution generated from 1,000 random samples under the hypotheses of selective neutrality and population equilibrium. A second independent assessment, this time of historical changes in the effective population sizes, was performed by employing mismatch distribution analyses (MDA; Rogers and Harpending, 1992; Harpending and Rogers, 2000). Both of these analyses were implemented in Arlequin 3.0 (Excoffier et al. 2005).

Coalescent Estimates of Divergence Times and Migration Rates

The pairwise coalescent method of isolation-with-migration of Hey and Nielsen (2004; IM) was used to simultaneously estimate effective population sizes, amount and direction of maternal gene flow, and time since divergence for selected pairwise

combinations of clades showing a significant relationship between haplotype distribution and geography within islands. The IM application provides a non-equilibrium approach that uses a Bayesian framework to estimate the above parameters by calculating their posterior probabilities given the genealogy (Nielsen and Wakeley 2001), under a finite sites model of sequence evolution. IM uses Markov-Chain Monte Carlo simulations to estimate likelihoods of parameters Θ ($N_{ef}\mu$), M ($N_{ef}m$), and T (t/ N_{ef}), where N_{ef} is the female effective population size, *m* is the migration rate, and μ is the mutation rate per locus per generation. Estimated mutation rates for iguanian lizard mtDNA (Macey et al., 1998, 1999), and in cyt-*b* for other squamate reptiles (Malhotra and Thorpe, 2000; Thorpe et al., 2005) give pairwise divergence estimates in the range of 1.3% - 2.0%, so we calculated N_{ef} from μ estimates for a range of values of $\mu = 3.94 \times 10^{-6} - 6.07 \times 10^{-6}$ per lineage per locus, per generation, assuming a generation time of one year (see Morando et al., 2006). These parameters were used to obtain estimates of population and gene divergence times, after initial population divergence. In this analysis we chose to run several pairwise combinations (see below for further information); (a) southwestern-northwestern clades and (b) western and Eastern clades in Isabela; (c) Santa Fe – southern Santa Cruz clades in Santa Cruz, and (d) Rabida – Santiago clade in Santiago. All of them should give independent estimates of divergence times that cannot be obtained through geology for reasons explained elsewhere. However, the IM method requires running repeated analysis of the same data to assure appropriate convergence of Markov-chains, and when a particular replicate is chosen it might take anywhere from 30 to 60 days to see convergence on every single parameter. Electrical blackouts and the nature of this study has prevented us from completing all results of this section, however, analyses are underway and they will provide with a number of points to further improve our inferences about demographic histories of selected haploclades.

Molecular Diversity Analyses

For the main clades identified in the phylogenetic and NCPA analyses, we summarized patterns of genetic variability by calculating the mean number of pairwise differences (using PAUP; Swofford, 2002). We also estimated haplotype (gene) diversity

(*h*; Nei 1987, p. 180)) and nucleotide diversity (π, the mean number of pairwise sequence differences; Nei 1987, p. 257), as well as the number of polymorphic sites for these same clades. Finally, we calculated patterns of genetic structure across the seven major islands, and within some of these, using a hierarchical analysis of molecular variance (AMOVA) implemented in Arlequin 3.0 (Excoffier et al., 2005).

RESULTS

Phylogenetic Analyses

We used a total of 614 ingroup terminals representing 78 localities across the range of the Western Radiation, and these reduce to 187 unique haplotypes. Figure 3 shows the consensus topology of 9001 trees based on a partitioned Bayesian approach recovered after 10 million generations. The models of substitution for the three codon positions were TrN + I + G (1st codon position), HKY + I (2nd codon position), and TrN + I + G (3rd codon position). This topology was completely concordant with the single-partition approach obtained through the use of PHYML with equilibrium base frequencies of α_A=0.358, α_C = 0.314, α_G = 0.125, α_T = 0.203; transition-transversion ratio = 15.52; α -distribution shape parameter = 0.600; proportion of invariant sites = 0.452, and unequal substitution rates (TrN + I + G; *ln*L = -6559.41714).

In our phylogenetic hypothesis (Fig. 3), *M. delanonis* (Española) appears as a sister group to a well-supported clade (PP/BS = 100/90) of *M. grayii* (Floreana) plus all remaining species. The sister group to *M. grayii* is comprised of two large clades with the following topologies: (1) one clade with strong support (100/100; Fig. 3) includes *M. pacificus* (Pinta) and *M. albemarlensis* (Isabela and Fernandina)]; and (2) the second clade (99/80) includes the remaining three western radiation species; *M. indefatigabilis* (Santa Cruz + Santa Fe), *M. jacobi* (Santiago) and *M. duncanensis* (Pinzon). Divergence among all island/species clades is deep and most are strongly corroborated (but see the *M. duncanensis* and *M. jacobi* clade for an exception). Extensive phylogenetic structure is also seen among subclades of Isabela, Santa Cruz, and Santiago Islands, and this is described in more detail below.

We compared our best topology to previously estimated topologies, and Table 2 summarizes the results of Shimodaira-Hasegawa tests. Because meaningful statistical tests require that the trees being compared are equally resolved (Shaw, 2002), we only evaluated the alternative topologies of Heise (1998), and Kizirian et al. (2004) to the fit of our topology reduced to a species tree (Fig. 2; our tree is the *a posteriori* hypothesis; Goldman et al., 2002). Constraints for phylogenetic tests were built using MacClade v4.03, and none of the topologies tested represented a likely alternative solution; both were significantly worse that our best estimate (Table 2).

Parsimony Networks, Phylogeographic Structure, and Population History

Ten independent networks were recovered based on the statistical parsimony algorithm of Templeton et al. (1992), and haplotypes separated by up to 14 mutational steps had greater than 95% probability of being connected parsimoniously (i.e., no superimposed mutations). Three separate networks were recovered for populations of *M. albemarlensis*; two for populations of *M. indefatigabilis* and two for populations of *M. grayii* (Appendix 2). All other independent networks were related to the major islands from which species were originally hypothesized (Van Denburgh and Slevin, 1913). These include (from east to west): Española (*M. delanonis*), Pinzon (*M. duncanensis*), and Pinta *(M. pacificus*). All TCS networks are given in Appendix 2, and those with sufficient structure to implement NCPA analyses are considered in detail below.

Santiago Network. Twenty-nine haplotypes defined by 16 segregating sites were identified in 13 samples ($n = 95$; Appendix 1) of *M. jacobi* from Isla Santiago and its seven adjacent islets (Rabida, Sombrero Chino, Bartolome, and Brainbridge Rocks 3, 4, 5 and 6), and these were represented by one to eight haplotypes per locality. A single parsimony network connecting all haplotypes contained six loops (Appendix 2), and we recovered two 4-step and three 3-step haploclades in the nesting design (Fig. 4). When mapped onto the island's geography, several of these show partially overlapping distributions. One widespread western haploclade (3-1) includes northern Santiago samples (jacPES, jacAL, jacBC, jacBO, jacSU, and jacBA [nested clades 2-1, 2-2, and 2-

3; gray and yellow in Fig. 4]), and wraps around the western edge of this island to also include haplotypes from southwestern Santiago (jacSSCH and jacSCH), and eastern Santiago (jacBR3, BR4, BR5 and BR6; inset in Fig. 4). A second haploclade (3-3; nested clades 2-6, 2-7 and 2-8 [orange in Fig. 4]) includes all haplotypes from Isla Rabida (albRAB in the inset map), but unpredictably it also includes haplotypes from the northernmost sampling site on Santiago (jacBC). A third haploclade (3-2; nested clades 2-4 and 2-5 [dark blue]) consists of genetically well-differentiated haplotypes of Isla Bartolome (jacBA), and northern Santiago's Bahia Sullivan (jacSU) and La Bomba (jacBO) sites, both of which overlap with clade 3-1 haplotypes.

Results of the contingency tests showed significant associations between haplotypes and geographic distributions at various levels in eight of 15 nested clades (Table 3). The NCPA inferences from these associations suggest that there is ongoing but restricted gene flow among Brainbridge Rocks localities that comprise the eastern haploclade (yellow clade 2-1). These localities are in turn connected through restricted gene flow or long distance dispersal, with northern populations from haploclade 3-1 (gray), but there is no contact with the neighboring populations of Sombrero Chino (jacSSCH and jacSCH). The Sombrero Chino sample (jacSCH), which is only ~2.0 km straight-line distance from the nearest population (jacBR6) of the Brainbridge Rocks haploclade, is part of nested clade 2-2 that is connected at a higher nesting level through long distance colonization or past fragmentation to the same northern ancestral haplotype as those from Brainbridge Rocks (haplotype 5 in Fig. 4).

At the most inclusive nesting levels, restricted gene flow and long distance dispersal seems to have structured populations on mainland Santiago (clade 2-2), between Santiago and the networks of the satellite island haplotypes (Table 3). Most unexpectedly, the presence of Rabida-like haplotypes in the Buccaneer Cove population (jacBC and jacRAB; clade 3-3) is inferred to reflect long distance dispersal (total nesting cladogram in Table 3) from Rabida to northern Santiago.

Santa Cruz Network. Fifty-three haplotypes defined by 85 segregating sites were identified in 22 samples of *M. indefatigabilis* (n = 209; Appendix 1) from Isla Santa Cruz and 12 nearby islets. The number of haplotypes per sampling site ranged from one to 12

(Cerro Colorado; indCC in Fig. 5), and two separate parsimony networks were recovered (Appendix 2). The most extensive network included haplotypes from Santa Cruz and the 12 associated islets, while the Isla Santa Fe and Islote Santa Fe samples were recovered as an independent network. The Santa Cruz network had five loops and is separated from the Santa Fe network by 49 mutational steps (Appendix 2).

Populations from northern mainland Santa Cruz have a comparatively higher number of haplotypes (localities indCC [12], indCD [10], indITA [4] and indBL [3] in Fig. 5) than southern localities (indRF [2], and indLFE [1]). Conversely, eight islets are characterized by a single haplotype (indBWN, indGFS, indGFN, indED, indPZS, indVE, indiFE, and indSY), three others segregate for two haplotypes (indGFE, indBWS, and indDM), while three islets segregate for more than two haplotypes (indGFW [5], indGFE [4] and indPZN [3]).

The Santa Cruz clade is composed of six 3-step subclades (3-1 to 3-6; Fig. 5) that when mapped onto the island's geography shows the following distributions. Subclades 3-1 (pink) and 3-2 (orange) encompass haplotypes that are restricted to one eastern (indCC) and three Santa Cruz mainland localities (indCD, indRF, and indLFE) localities. Related to the latter, clade 3-3 (green, inset map) includes geographically proximate haplotypes from northwestern Santa Cruz (indCDP and indVE) and two neighboring islets (indGFW and indGFS). Clade 3-4 (yellow) is geographically more extended and encompasses the northern Santa Cruz shore from Punta Bowditch (indBWS) to Cerro Colorado (indCC), including the Guy Fawkes islets West and South (indGFS and indGFW), and 'Santa Cruz' (indSC), but not localities in between (e.g., indVE [< 1.0 km air distance from indSC]). Clade 3-5 (light blue) includes haplotypes only found on three northern Santa Cruz islets; Baltra (indBL), Seymour (indSY), and Daphne Major (indDM; inset in Fig. 5), but it also includes the westernmost haplotype of Eden islet (indED; full map in Fig. 5) while "bypassing" clades 3-3 and 3-4. Finally, clade 3-6 (dark blue) is restricted to four localities in the northeastern corner of Santa Cruz (indITA, indBL, indPZN, and indPZS).

Several subclades within these six show significant haplotype/geography associations; and contingency tests allowed demographic inferences at various levels in 19 of 27 nested clades (Table 3). The population history of *M. indefatigabilis* is

characterized by well-defined clades that overlap geographically, and islet isolates that do not conform to relationships predicted by geography. All 1-step clades connect haplotypes from adjacent islets for which there is evidence for three demographic inferences. Restricted gene flow is inferred between islets Bowditch North and South (clade 1-16); past fragmentation is inferred between Guy Fawkes West and South, Seymour and Baltra, and Plaza Sur and Norte (clades 1-12, 1-21, 1-26 respectively); and allopatric fragmentation is inferred for the historical relationship among haplotypes of Guy Fawkes North and South islets (clade 1-19; Fig. 5). Long-distance colonization is also inferred between some of these localities (Table 3).

At the next nesting level, clade 2-7 further clarifies the relationship of clade 1-12 (indGFW and indGFS) with Santa Cruz haplotypes (indVE, indCDP) through restricted gene flow and isolation-by-distance. Likewise, clade 2-11 clarifies the relationship of clade 1-19 (indGFN and indGFE) with mainland Santa Cruz haplotypes, although in this case, by past fragmentation and long distance colonization from Cerro Colorado (indCC) eastern Santa Cruz (see arrow in clade 2-11 in Fig. 5). Similarly, clade 2-15 connects populations of Plaza Sur and Norte (clade 1-26) to haplotypes from Canal Itabaca (indITA) across Baltra Islet through restricted gene flow and isolation-by-distance. Clade 2-13 modifies the inference of past fragmentation for clade 1-21 (Baltra and Seymour) for one of long distance dispersal that includes Daphne Major (indDM; Fig. 5), the most distant islet from Santa Cruz.

With the exception of range expansion inferred for clade 3-2, all five 3-step clades infer restricted gene flow and isolation-by-distance. The inference of restricted gene flow of clade 3-3 (Fig. 5) supports the same inference made in lower clades (clade 2-7), but includes more haplotypes from the same localities (indVE, indCDP, and indGFW, and indGFS). Clades 3-1 and 3-2 summarize haplotype relationships of the southern part of Santa Cruz, and link them by restricted gene flow (indCC, indRF, indCD), and eventually range expansion from Cerro Colorado (indCC) to La Fe (indFE). Finally, restricted gene flow with isolation-by-distance explains the relationship of the easternmost Santa Cruz haplotype (indED) with clade 2-13 (northern islets Daphne Major, Baltra and Seymour; Fig. 5). Higher category clades yield either inferences of past fragmentation and long distance colonization (4-1) and range expansion (4-2), or coincide in the common

inference of restricted gene flow with isolation-by-distance (5-step and total cladograms; Table 3).

Isabela and Fernandina Network. Seventy-nine haplotypes defined by 100 segregating sites were identified in 27 populations (n = 210; Appendix 1) from both Isabela and Fernandina islands. Three separate parsimony networks were recovered (Western, Eastern, and Cuatro Hermanos), which contained 5, 9, and 0 loops, respectively; (Appendix 2). The "Western Clade" includes 47 haplotypes from 20 localities from all of Fernandina and northern and western Isabela (Fig. 6). Three 3-step subclades are nested within the western clade, one of which (3-1; orange) extends from a northeastern Isabela locality (albCMR; Fig. 6) across the northern coast of this island (but anomalously skipping one site [albPB]), and southwest to include all of Fernandina. Most haplotypes in this clade are connected by one step to a high frequency ancestral haplotype (#7; Fig. 6). A second subclade (3-2; yellow) and the third (3-3; light green) are both restricted to western Isabela, and in clade 3-3 most haplotypes are also connected by one step to a high frequency ancestral haplotype (# 34; Fig. 6).

Isabela's "Eastern Clade" is recovered from the second TCS network, and includes 31 haplotypes from eight localities spanning a considerable distribution and also characterized by limited overlap among subclades. This clade encompasses southern (albCSP, albTOR, and albPMO), eastern (albBA, albPUA, albCW, and albCMR, but not alb4H), and a single northern Isabela locality (albPB; Fig. 7). The third "network" includes a single specimen from Cuatro Hermanos islet off of Isabela's southeastern shore (alb4H; Fig. 7), which is 22 mutational steps distant from the Eastern Clade sample at Caleta San Pedro (albCSP).

The eastern and western Isabela clades overlap geographically but not at nearby localities where geographic contact might be expected (the southwestern sites albCWE and albCSP, or between northern localities albAL and albPB; Figs. 6, 7). Rather, haploclade sympatry occurs in two areas that cannot be explained by geography alone (i.e., albCMR and albPMO), while the Piedra Blanca sample (albPB) is comprised exclusively of haplotypes that are part of the eastern Isabela clade, despite being geographically "embedded" within the range of the western clade (Fig. 7).

Significant haplotype/geography associations were found in seven of 16 subclades present in the western Isabela clade (Table 3). Demographic inferences for subclade 1-1 include continuous range expansion, with the possibility of long-distance colonization from northern Isabela to some localities of Fernandina (Table 3). At the 2-step nesting level, restricted gene flow (clade 2-2) adds all remaining haplotypes of Fernandina and the three additional, but geographically non-continuous, north/north-eastern Isabela subclade (localities albCMR, albRC, and albAL; Fig. 6) into the same nested clade. Overall, restricted gene flow with isolation-by-distance (clade 3-1) is the underlying demographic hypothesis inferred to explain the relationships of all northwestern Isabela haplotypes. The smaller southwestern clade 1-22 is inferred to be structured by restricted gene flow with isolation-by-distance among four localities (Table 3). A range expansion is inferred at the 2-step next level (clade 2-6), and the 3-step clades (3-1, 3-2) are inferred to be structured by restricted gene flow and either long distance colonization (3-1), or restricted gene flow with isolation-by-distance for the entire western Isabela clade (3-2; Table 3).

The eastern Isabela clade (Fig. 7) is simpler in the number of inferences made; only five of eight subclades showed significant associations (Table 3), and four of these five are hypothesized to be best explained by long distance colonization events. There is evidence for restricted gene flow among some localities in clade 2-3 (albTOR and albBA in Fig. 7), and at more inclusive levels, clade 3-1 extends the distribution of clade 2-3 to include two haplotypes from Punta Moreno (albPMO), and clade 3-2 connects haplotypes from Cowley (albCW) to those at the distant locality of Caleta San Pedro (Fig. 7). All 3-step clades lead to inferences of long distance colonization or past fragmentation, and the total cladogram suggests restricted gene flow and some long distance colonization.

Cross-validation of NCPA Inferences

Significant Fu's tests suggest that on Fernandina (*Fs*=-28.42), Española (*Fs* = -27.01), Pinzon (*Fs* = -14.40), and Santiago (*Fs* = -25.03), *Microlophus* populations are either in a demographic expansion, or the mtDNA locus is not evolving in an approximately neutral manner (Table 4). In contrast, pooled samples from Isabela, Santa

Cruz, and Floreana do not show a clear support for the sudden expansion model (Table 4) and this is emphasized by the distinctly bimodal shape of the mismatch distributions (MMD) of *M. albemarlensis* (in Isabela), and *M. indefatigabilis* (Santa Cruz), the slightly bimodal MMB of *M. jacobi* (Santiago), and the multimodal MMD of *M. grayii* (Floreana; Fig. 8). In contrast, Fu's *Fs* and Tajima's D are rather consistent with a signature of population stability or neutral evolution for populations of Isabela, Santa Cruz and Floreana, but on closer inspection, the MMD of Santiago and Santa Cruz show two patterns. One of these fits a sudden expansion model, but this is based only on pairwise differences from the main islands alone, whereas the second mode is based on pairwise differences across haplotypes from the main islands as well as to the divergent haplotypes from offshore islets. In the case of Isabela populations neither of the two MMD modes fit the model of a recent expansion, but pairwise haplotypes distribution of Fernandina alone are consistent with this interpretation (Fig. 8).

Another result is the multimodal mismatch in populations of Floreana (Fig. 8). The pattern shows hidden population structure and is further substantiated by the haplotype network of *M. grayii* (Appendix 2) that shows at least three of the four Floreana Islets as moderately (Champion Islet) to highly differentiated (Caldwell, Enderby and Gardner-by-Floreana); these are separated by anywhere from three to >12 inferred mutational steps respectively (Appendix 2). In contrast, MMD of pooled populations of Española and associated islets (Osborn, Xarifa, Oeste and Gardner-by-Española) fit the model of sudden expansion while showing a reduced number of current and inferred haplotypes (Appendix 2).

Partitioning of Genetic Diversity

The largest number of haplotypes corresponds to populations of *M. albemarlensis* in Isabela and Fernandina (81 combined; Table 4); and it is followed by Santa Cruz (56 haplotypes) and Santiago (26 haplotypes). Gene diversity is high on all islands, and the mean number of pairwise differences is > 16 on Isabela and Santa Cruz (this is also reflected in the number of polymorphic sites; Table 4). Such a pattern reflects deep phylogeographic structure on both of these islands, even though the Isabela-Fernandina

complex is the youngest in the archipelago (Geist, 1996). The smaller sized Pinta and Española islands show the lowest number of haplotypes (3 and 5 respectively, but the Pinta sample represents 100% of the lizards surveyed; Table 4), while Santiago shows an intermediate value (26 haplotypes; Table 4). Overall, genetic variability seem to be correlated with island size rather than island age and the most extreme case is the number of haplotypes found on the youngest island Fernandina (n = 9) versus the number of haplotypes found in the oldest island Española (n = 5).

The effects of genetic drift can alternatively explain a lower number of haplotypes in small islands/islets/rocks. In small populations there is a higher likelihood of allele fixation that could be the result of either selection or random genetic drift. In both cases the net demographic effect is lower levels of gene and nucleotide diversity. Such effect is hardly noticeable in populations from the main islands but it is in an individual population basis in almost all islets and rocks of the Archipelago. In fact, regardless of geographic distance to their main island, most islets/rocks of the archipelago show comparatively lower gene diversity values.

Patterns of between vs. within-island population differentiation and genetic variability are summarized in a hierarchical analysis of molecular variance (AMOVA) for all islands but Pinzon and Pinta (Table 5). There is significant among-island variation for all combinations, but the relative proportion of within-island variation ranges from 0.0 on the oldest island of Española, through modest values on Santiago (14.60%) and Isabela (25.68%), to the largest proportion of total variability on Fernandina (>85%; Table 5). The inverse pattern is also evident; the highest among-island variation occurs for the oldest islands (Española and Floreana), while the lowest belongs to Fernandina. The clearest pattern is that among-island genetic variation increases in parallel with estimated island ages (Table 5). Finally, and with the exception of populations from Española (*M. delanonis*), none of the islands show a demographic signature that significantly differs from a simulated model of sudden population expansion (Table 4).

DISCUSSION

Phylogeny and Colonization History of the Western Radiation of Microlophus

The earliest proposed colonization history of the western radiation of Galapagos *Microlophus* is that of Wright (1983), but the hypothesis is based on an unresolved topology (a UPGMA dendrogram of allozyme genetic distances), and as a result, his actual colonization scenario requires additional assumptions about ocean currents and the order of speciation events (see our Fig. 2). The Heise (1998) topology represents the first true phylogenetic hypothesis for this group, but it remains unpublished and is based on a sampling design limited both in the number of localities (24), individuals (36), and number of gene regions and base-pairs sampled (469 from 16s; 307 from cyt-b). The recent study of Kizirian et al. (2004), improves on the previous sampling design by increasing the number of mitochondrial gene regions (ND1, ND2, COI, and tRNAs) and base-pairs included (1956), but it is based on a smaller number of localities (16 total vs 24 in Heise [1998]), and remains very limited with respect to the number of individuals sampled within each locality (n = 1 - 2). Given the complexity of island systems and the potential for strong evolutionary forces to operate over relatively short distances on single islands (Caccone et al., 1999; Ogden and Thorpe, 2002; Beheregaray et al., 2003*a*; 2003*b*; Jordan et al. 2005a), we attempted here to extend the above-mentioned studies by increasing both the number of localities sampled on the larger islands, and the number of individuals sampled per site on all islands.

Our tree-based methods recover a fully resolved topology of all but very recent nodes for the Western Radiation (Fig. 3), and when this topology is constrained to the fully-resolved alternatives of Heise (1998) and Kizirian et al. (2004) and evaluated with pairwise tests, both were significantly worse that our best estimate (Table 2). Based on the phylogenetic hypothesis, we present an alternative colonization hypothesis for the Western Radiation of *Microlophus* species (Fig. 9). In agreement with all previous studies, Española was the first colonized island (represented by the extant *M. delanonis*), followed by a successful colonization to Floreana (*M. grayii*). Given that prevailing ocean currents run in a northwesterly direction going from the older to younger islands for much of the year, (Pak and Zaneveld 1973; Wyrtki, 1976), this colonization pattern is not surprising.

Floreana then served as a source for two subsequent radiations. One of these founded the Isabela/Fernandina populations (*M. albemarlensis*), and this group included the colonists to Pinta Island (*M. pacificus*). The second group derived from *M. grayii* ancestors on Floreana Island included the common ancestor of extant species on Pinzon (*M. duncanensis*) and Santiago (*M. jacobi*), and the extant populations of *M. indefatigabilis* on Santa Cruz and Santa Fe Islands. There is no obvious ancestral land mass that could serve as a single "jump off" point for the first split needed to explain the strongly supported sister group relationship between the (*M. duncanensis* + *M. jacobi*) + (*M. indefatigabilis* [Santa Cruz] + *M. indefatigabilis* [Santa Fe]) clades (Fig. 9), but alternative sequences of colonization are possible for either Pinzon or Santa Cruz first, with subsequent sequential dispersal to the other islands. This unresolved component of our hypothesis can be further tested by coalescent modeling under relatively simple alternative colonization histories (see Carstens et al. 2004*b*, for a recent example).

The within-archipelago colonization histories for other radiations of Galapagos organisms have been studied from molecular perspectives, including finches (Sato et al., 2001; Burns et al., 2002), hawks (Bollmer et al., 2006), geckos (Wright, 1983) mockingbirds (Arbogast et al., 2006), marine iguanas (Rassmann et al., 1997), giant tortoises (Caccone et al., 1999), plants (Schilling et al., 2004; Nielsen, 2004), beetles (Sequeira, et al., 2000), and warblers (Collins, 2003). All of these radiations are characterized by single colonizations from the mainland, but they differ in the approximate timing of the first colonization event. Estimates range widely from a low of 0.3 Myr for the Galapagos hawk (Bollmer et al., 2006) to a high of 15-20 Myr for marine iguanas (Rassmann, 1997), with a spread of estimates between these extremes (e.g., 2.3 Myr for finches [Sato et al., 2001; Burns et al., 2002]; 2.5 Myr for warblers [Collins, 2003]; 2.0 Myr for tortoises [Caccone et al 1999]; 9 Myr for *Microlophus* [Wright, 1983]; and 11 Myr for beetles [Sequeira et al., 2000]). In some cases local extinctions prevent strong inferences of the first island to be colonized from mainland (Bollmer et al., 2006), but most others show a pattern of colonization parallel to what we report in this paper for the Western Radiation of *Microlophus*; older lineages usually inhabit eastern islands while younger lineages occupy western islands (Beheregaray et al., 2004; Bollmer et al., 2006; Rassmann et al., 1997; Sequeira et al., 2000). This phylogenetic pattern is also

reflected in greater between-island genetic divergence among older lineages; in tortoises for example, among-island differences explain 97% of mitochondrial molecular variance for older islands and only 60 % for younger islands (Beheregaray et al., 2004).

On a final note, our study rejects the hypothesis of Kizirian et al. (2004) that *M. albemarlensis* is paraphyletic with respect to *M. pacificus* (see their Fig. 4; p. 767). Those authors did note the high mtDNA sequence divergence among localities on Isabela and the likelihood of ". . . independent evolution of Lava Lizards on Isabela", but the topology upon which they based their colonization scenario is likely biased by limited sampling on this large island. Our results show strong support (i.e., high posterior probability and bootstrap values) for each of the seven recognized species, and for some of the subclades within some of these (e.g., the Santa Cruz and Santa Fe samples of *M. indefatigabilis*; Fig. 3). Some of these subclades are geographically localized and all but *M. albemarlensis* are represented by separate TCS networks, which in other groups have been shown to be concordant with morphologically recognized species (Cardoso and Vogler, 2005). Proper quantification and possible adjustment of species boundaries in this complex will require independent evidence and is beyond the scope of this study, but it will be important to resolve these issues from both perspectives of understanding evolutionary processes on oceanic islands (Emerson, 2002), and to avoid overlooking taxa or distinct lineages of conservation significance (Russello et al., 2005).

Climatic and Geological Influences on the Evolution of Microlophus *Within Islands*

The second objective of this study was to compare and evaluate within-island population structure predicted by the east-to-west colonization history, in which the null hypothesis is that lizard populations should show decreasing phylogeographic structure in the same direction (east-to-west) as that predicted for colonization/speciation history (Beheregaray et al., 2004). However this expectation is overly-simplistic because islands in the Galapagos Archipelago are not strictly linear in age and appearance (unlike Hawaii), and evolutionary histories of endemic groups may be strongly influenced by stochastic processes shared by many or all oceanic island archipelagos (Censky et al., 1998; Spiller et al., 1998; Juan. et al., 2000; Carson et al. 2004; Calsbeek and Smith,

2003, de Queiroz, 2005; Jordan et al., 2005b; Gübitz et al., 2005; Cowie and Holland, 2006). For example, genetic variability in local populations may be dramatically increased by multiple colonizations (Kolbe et al., 2004), and on islands this may contribute to within-island speciation (Jordal et al., 2006). More relevant for the Western Radiation of *Microlophus* in Galapagos is that the simple east-to-west history of island formation does not describe the confounding influences of within-island dispersal, population expansion, local extinction, re-colonization, and/or secondary contact on islands shaped by volcanism (Beheregaray et al., 2003*a*; Vandergast et al., 2004; Nielsen, 2004; Ciofi et al., 2006; Brown et al., 2006; Bloor and Brown, 2005), and the often unpredictable direction of ocean currents (Beheregaray et al., 2003*b*; Calsbeek and Smith, 2003; Glor et al., 2005; Rusello et al., 2005).

In Galapagos the Humboldt Current flows in a predominantly northwesterly direction (Pak and Zaneveld, 1973; Wyrtki et al., 1976; Caccone et al., 2002), and as a rule would be expected to transport debris, organisms, etc., in this same direction. Further, this prevailing current would likely prevent "back colonization" in the opposite direction (from younger back to older islands), but the Humboldt Current is temporally 'reversed' during El Niño Southern Oscillation (ENSO) events (Guilderson et al., 1998), and thus may periodically provide for back colonization in an easterly or southeasterly direction. Further, ENSO events significantly alter timing of seasonal rainfall and increase total precipitation, and Kizirian et al. (2004) hypothesized that excessive rains in ENSO years caused flooding on high-elevation islands. This flooding would presumably increase the frequency with which mats of vegetation, and presumably "stowaway" lizards, would be washed into the ocean and available for transport among islands. Kizirian et al. (2004; p. 768) proposed this mechanism specifically "for maintaining gene flow in *M. albemarlensis* complex and explain the weak diversification within that clade." We return to this point below.

In addition to currents, ocean-level fluctuations have been proposed by Hickman and Lipps (1984) and Geist (1996) for the Galapagos Archipelago, and the importance of these events in promoting allopatric fragmentation and/or allowing gene flow through recent land bridges has been only recently explored in oceanic islands (see S. Jordan et al., 2005b, for a case in Hawaii). In Galapagos, the most recent glacial maximum is

estimated to have lowered sea level about 130 m below the present sea level; this drop would extend the areas of all islands, and for many islands the majority of satellite islets would be connected to the main body land. However, there is no evidence that this drop in sea level formed temporary land bridges between the main islands of the Archipelago (Geist, 1996).

The geological history of the Galapagos is relatively well understood, and the archipelago sits on a crustal plate (Nazca Plate) moving from the east-central Pacific towards the western edge of South America (Nordlie, 1973; Cox, 1983; White et al., 1993; Geist, 1996; Werner et al., 1999; Naumann and Geist, 2000). Some islands are aligned by age along the west-to-east direction of plate movement, from Española (3.3 My) to Floreana (1.5 My) to Isabela and Fernandina (< 700,000 and 300,000 yr, respectively; these two are over the "hotspot" and still have active volcanoes). Others, however, are grouped into clusters of similar age that orient more at a perpendicular angle to the direction of plate movement, such as Española and Marchena (3.3 My each), and a Floreana (1.5 My) – Pinzon (1.5 My) – Santiago (1.2 My) – Pinta (1.0 My) cluster. Thus, predicting simple east-to-west colonization routes is not straight-forward in the context of island history, shifting ocean currents, and fluctuation sea levels.

One other issue is relevant here, and that is the history of volcanic activity within the complex centrally located younger islands, including Santa Cruz, Santiago, Isabela, and Fernandina. Portions of all of these islands except Santa Cruz (the oldest) are still covered by extensive lava fields of various ages, and their extent is inversely correlated with the vegetated area on each island (Aadversen, 1988; see Table 5). Isabela is the second-youngest island, and the largest and most geologically complex; it is comprised of five volcanoes (Wolf, Darwin, Alcedo, Sierra Negra and Cerro Azul; Fig. 1) of various sizes and ages. The northern three cones (first three above) have probably always been separated from the southern two (S. Negra and Azul), and from each other, by the sea or barren lava fields (Willerslev, 2002; Geist, 1996). No dates for the extent or duration of these separations have been published; these volcanoes orient perpendicular to the direction of plate motion, and multiple eruptions challenge conventional dating methods (Naumann and Geist, 2000). There is some support for separation of a northern block (comprised of Wolf, Darwin, and Alcedo volcanoes) from a southern block (Sierra Negra

and Cerro Azul; Fig. 1) of Isabela (Eliasson, 1974; Beheregaray et al., 2004; Nielsen, 2004), and Nordlie (1973) proposed a sequence of origin of these volcanoes based on the morphology of each. In Nordlie's model, Volcan Sierra Negra is hypothesized to be the oldest part of Isabela, and Cerro Azul and Volcan Wolf the youngest. The Galapagos chain continues to the east but transforms into submerged seamounts (Christie et al., 1992) and sub-aerial substrate dates to ~ 14 My ago (Werner et al., 1999).

Population Structure and Evolutionary Processes Within Islands

Our results show general concordance with similarly dense sampling studies of other Galapagos radiations (Beheregaray et al., 2004; Nielsen, 2004), which is manifest as the same trend in the distribution of molecular variance. For the mtDNA locus, molecular variance is preferentially apportioned among populations in the oldest islands and within populations in the younger islands (Table 5). Further, this pattern fully coincides with putative island ages and has no relationship to island area. Conversely, estimates of haplotype and gene diversity, as well as the mean number of pairwise differences, have no relationship with phylogenetic patterns, age, or island area, but they show positive association with the degree of complexity of island topography (general patterns of mtDNA complexity are captured in the genealogical networks summarized in Appendix 2).

Ignoring the smallest islands represented by few lizards (Pinzon, Pinta), the simplest haplotype networks are recovered for the islands of Española and Floreana. Both islands show relatively similar gene and haplotype indexes (Table 4) but they significantly differ in the mean number of pairwise differences, and demographic histories inferred from mismatch distributions. Results of this test shows that pooled localities of Española – the oldest island sampled – are consistent with a model of recent population expansion in *M. delanonis*, while the same test suggests demographic equilibrium in *M. grayii* on Floreana (Fig. 8). This result was unexpected because the two islands are in the same relative geographic position with regard to prevailing ocean currents, both have a similar number of offshore islets (Table 5) at the same approximate distances of the main island. However, TCS networks for these two islands suggest that

there is gene flow between mainland and offshore populations of *M. delanonis*, whereas deep phylogeographic structure (up to 31 missing haplotypes between Gardner-by-Floreana and Champion Islets Appendix 2) within *M. grayii* suggests that ocean currents may be sufficiently constant and unidirectional to Floreana from its offshore islets. If mainland – islet gene flow in *M. delanonis* is due in part to frequent population crashes and re-colonization to these smaller islets, the demographic signature might resemble a recently expanded population.

Two independent networks define phylogeographic structure in Santa Cruz, and this divergence is putatively old as 49 missing haplotypes separate the two networks. Within Santa Cruz there is complex set of demographic inferences that involve multiple nested levels. Most 1-step level demographic inferences involve the majority of the 14 associated islets, and NCPA inferences suggest that past fragmentation (for Plaza Norte and Plaza Sur and the islets of Baltra and Seymour [clades 1-21 and 1-26 in Fig. 5]) or allopatric fragmentation (Guy Fawkes North and East [clade 1-19; Fig. 5]); these events likely derived from Pleistocene sea level fluctuations (Jordan et al., 2005a; Jordan et al., 2005b; Geist, 1996). Two of the four 2-step clades give inferences of colonization by long distance rafting events (Daphne Major; clade 2-13; Fig. 5). For the remaining islets the predominant inference is also long distance colonization, but with the distinction that the source populations are almost never from adjacent islets, but rather from unexpected parts of the island following seemingly random colonization routes. For instance, both Guy Fakes North and East and Plaza Sur and Norte were founded by migrants from Itabaca (indITA in Fig. 5), while populations from the northern coast of Santa Cruz have founded neighboring Guy Fawkes South and West. In contrast, Eden Islet was colonized by haplotypes related to clade 2-13 (Daphne Major, Seymour and Baltra) and not from neighboring localities. Localities on the southern edge of Santa Cruz return a NCPA inference of restricted gene flow and range expansion among distant localities (clades 3-1 and 3-2 Fig. 5), and the number of inferred haplotypes between southern Santa Cruz and Santa Fe (21 steps; Appendix 2) suggest a relatively early colonization of Santa Fe; and these patterns collectively suggest a predominant role for ocean currents both very early in the evolution of these populations (colonization of Santa Fe), and more recently in the 1-step and 2-step clades.

In Santiago a single northern haplotype (# 3 in Fig. 4) fits the prediction that ancestral haplotypes should be located in less volcanically active regions, while southern Santiago localities show the demographic inference of restricted gene flow (clade 3-3; Fig. 4) in populations characterized by lower diversity on the only remaining lava fields in the island (Willerslev et al., 2002). Two examples illustrate the effect of long distance colonization in Santiago (Brainbridge Rocks and Rabida Islet Fig. 4). In the first, all four Brainbridge Rocks are colonized by haplotypes from northern Santiago and there is no evidence for gene exchange with neighboring Sombrero Chino localities (Fig. 4). The second example shows haplotypes of Rabida islet and Buccaneer Cove in the same nested clade (clade 3-3 in Fig. 4). Both examples represent startling examples of long distance colonization from non-neighbor sources, and in contrast to Santa Cruz, NCPA analyses return inferences of long-distance colonization (or dispersal) at all nesting levels for Santiago (Table 3). Ocean-mediated transport of lizards has been a dominant force throughout the history of *M. jacobi* on this island.

Comparing Santiago and Santa Cruz is also instructive because their phylogeographic structure correlates with the putative age (Santiago, 1.2 My; Santa Cruz, 2.2 My). While Santiago still preserves nested ancestral haplotypes, on Santa Cruz most high frequency haplotypes are rather islet tip clades (Appendix 2). The high frequency of a single haplotype in the numerous Santa Cruz islets could be alternatively explained by very recent colonization or by the higher likelihood of allele fixation of genetic drift. This effect is hardly noticeable in the pooled samples across the eight main islands, but it becomes apparent on an individual population basis. For example, all Santa Cruz (haplotypes 15, 36, 35, 17, 18, 47, 48, 27 and 28; Appendix 2), and Santiago Islets (haplotypes 2, 3, and 4 Appendix 2) show this pattern, and this is true regardless of the nesting level of the islet haplotypes (e.g., the deeply nested Eden Islet haplotype at site indED in Fig. 5).

The large Isabela - Fernandina complex best reflects the many influences of volcanic events on within-island phylogeographic structure and demographic history. TCS analyses recovered two separate networks that we have labeled “Eastern” and “Western” clades (Figs. 6 and 7), and the two reveal a complex history of *M. albemarlensis* on these landmasses. The eastern clade shows a pattern consistent with an

initial colonization event followed by dispersal to the north or south (depending on the initial landfall), and NCPA inferences show that long-distance colonization (or long-distance dispersal) have been dominant events throughout the history of this clade (i.e., at 2-step, 3-step, and complete nesting levels; Table 3). This history is consistent with an "island hopping" dispersal mode that may have resulted in colonization of separate volcanoes before subsequent eruptions coalesced to unit formerly separate landmasses. Similarly, the odd distribution of a single clade 3-1 locality on the western flank of Sierra Negra (light blue in Fig. 6) might be expected via oceanic drift "through the gap" between Isabela northern and southern blocks, prior to coalescence. Both the "stepping stone north" and "though the gap" hypotheses are plausible if the founding lizard population initially made landfall on the eastern or southeastern flank of Sierra Negra. This appears plausible to us given that: (1) Sierra Negra is the oldest and the most voluminous volcano, and (2) this is the expected direction of rafting in the prevailing Humboldt Current (directly NW from Floreana; see Fig. 9). For similar reasons this landfall has been postulated as the initial colonization point on Isabela by the giant tortoises (Beheregaray at al., 2004). NCPA inferences also suggest a history of restricted gene flow with isolation-by-distance throughout the history of this clade (Table 3), but distinguishing among alternatives awaits further study with other genetic markers.

Following the above argument, we hypothesize that the western clade of Isabela also originated via a "through the gap" colonization shortly after lizards made first landfall on the eastern/southeastern flanks of Sierra Negra. At the highest nesting level this clade is characterized by a signature of restricted gene flow with isolation-by-distance (total cladogram, Table 3), but clade 3-2 (yellow in Fig. 7) also has a history of long-distance dispersal. Nesting clade 1-1 is inferred to have a history of long-distance colonization, which likely accounts for its distribution along the western edge of Fernandina, where as nesting clade 1-22 is inferred to have a history of restricted gene flow with isolation by distance, and it is distributed across all of this island. Unlike the eastern clade, the nesting designs of the western clade reveal a north and southern ancestral haplotypes (# 7 and # 34; Fig. 7) connected by single steps to many others in the networks at 1-step and 2-step levels. This structure implies a very different demographic history for the western clade, and we hypothesize that most dispersal along the western

edge of Isabela was land-based, and followed shoreline gradients north and south of the original landfall point for the group.

In summary, patterns of within island phylogeographic structure in Isabela seem to mirror geology and ongoing volcanic activity that creates new land at both northern and southern tips of Isabela as well as land bridges between volcanoes (Nordlie, 1973). Volcanic activity also isolated the western and eastern shores of Isabela preventing further gene exchange between the Eastern and Western clades. These two clades do overlap geographically (in two places; albPMO, albCMR; Fig. 7) but not in the spatial edges of both clades where contact might be expected (i.e., between albCWE and albCSP; Fig. 7). In these two localities, the genetic effect of between-clade sympatry is readily observed because the mean number of pairwise differences is abnormally high in both, albPMO and albCMR (13.16 and 13.26, respectively). On the other hand, long distance colonization also took migrants from the eastern clade to areas well within the western clade (albPB north of Wolf Volcano in Figs. 1 and 7) but remarkably no secondary contact is resolved for this locality and this may be associated with the young age of Wolf (Nordlie, 1973). A detailed study of interactions at these sites would be instructive with regard to speciation potential between eastern and western clades.

The surprising pattern of haploclade diversity and genealogical relationships within Isabela parallels the patterns and rates of speciation recently reported for the Hawaiian cricket genus *Laupala* (Mendelson and Shaw, 2005). In this study, the fastest diversification rates occurred on the youngest island (Hawaii), which is the largest and also the most geologically active of the archipelago. Unlike Hawaiian crickets where sexual selection is mentioned as a likely cause of this sudden diversification, the genetic differentiation of Isabela clades is probably driven by volcanic history in conjunction with oceanic currents.

Overall, two features are the most outstanding of Galapagos phylogeography. First, the recurrent and repeated evidence of long distance dispersal in three major islands; Isabela, Santa Cruz, and Santiago. In all of them, long distance colonizations occurred repeatedly and from random sources and hardly ever from neighboring localities empirically contradicting the stepping-stone island model of migrant exchange. Second, our results are interesting in that they reject the idea of phylogeographic structure been

related solely to island age. We provide evidence to support active volcanism as a major player in the generation of genetic diversity in island environments

As a final point, we want to stress the use of lava lizards as a model vertebrate to study evolutionary processes in the Galapagos. Lava lizards are by far the most widely distributed vertebrate group in the Archipelago, and unlike some other vertebrates there is no recorded extinctions (Caccone et al, 2002), they are entirely allopatric and unlike finches there is no evidence of hybridization (Grant and Grant, 1996), they are abundant, and according to preliminary data (Wright, 1983, Benavides et al., unpublished) they have been in the Galapagos for most of the putative age of the islands. Sampling at finer geographic scales and with high-resolution nuclear markers will be necessary to further test evolutionary processes in lava lizards.

ACKNOWLEDGMENTS

We thank Susana Cardenas and the staff of the Charles Darwin Research Station and the Galapagos National Park Service for logistical support in the field, and the crew of the Queen Mabel. We thank Luis Coloma, Alessandro Catenazzi, Shaleyla Kelez, and Jose Cordoba for tissue loans of mainland species, BYU students John Wells, Paige Alsbury and Ian Stehmeier for laboratory assistance, and Ray Grams for field assistance in Peru. E. Benavides thanks BYU for support, including the Department of Integrative Biology and the M.L. Bean Life Science Museum, BYU graduate mentoring and graduate research fellowships, a Larsen Scholarship, and a B.F. Harrison Scholarship. Additional support was provided by a graduate research award to E. Benavides from the Society of Systematic Biologists, and NSF awards DEB 0309111 (doctoral dissertation improvement award to J.W. Sites, Jr. and E. Benavides), and EF 0334966 ("Assembling the Tree of Life: The Deep Scaly Project: Resolving Higher Level Squamate Phylogeny Using Genomic and Morphological Approaches") to T. Reeder, M. Kearney, J. Wiens, and J.W. Sites, Jr.

LITERATURE CITED

Adsersen, H. 1988. Null hypotheses and species composition in the Galapagos Islands In: During, H. J., M. J. A. Werger, and J. H. Willems, J. H. (eds), Diversity and pattern in plant communities. SPB Academic Publ., pp. 37–46

Arbogast, B.S., S V. Drovetski, R. L. Curry, P. T. Boag, G. Seutin, P. R. Grant, R. Grant and D. J, Anderson. 2006. The origin and diversification of Galapagos mockingbirds. Evolution 60:370-382.

Baldwin, B. G., and M. J. Sanderson. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proc. Nat. Acad. Sci. USA. 95:9402-9406.

Barrier, M., R. H. Robichaux, and M. D. Purugganan. 2001. Accelerated regulatory gene evolution in an adaptive radiation. Proc. Nat. Acad. Sci. USA 98:10208-10213.

Beheregaray, L. B., C. Ciofi, D. Geist, J. P. Gibbs, A. Caccone, and J. R. Powell. 2003a. Genes record a prehistoric volcanic eruption in Galapagos. Science 302:75.

Beheregaray, L. B., C. Ciofi, A. Caccone, J. P. Gibbs, and J. R. Powell. 2003b. Genetic divergence, phylogeography and conservation units of giant tortoises from Santa Cruz and Pinzon, Galápagos islands. Conserv. Genet. 4:31–46.

Beheregaray, L. B., C. Ciofi, D. Geist, J. P. Gibbs, A. Caccone, and J. R. Powell. 2004. Giant tortoises are not so slow: rapid diversification and biogeographic consensus in the Galápagos. Proc. Nat. Acad. Sci. USA 101:6514-6519.

Benavides, E. 2005. The *Telmatobius* species complex in Lake Titicaca: applying phylogeographic and coalescent approaches to evolutionary studies of highly polymorphic Andean frogs. Pp. 167-185 *In* Lavilla, E. O., and I. de la Riva, eds. Studies on Andean frogs of the genera *Telmatobius* and *Batrachophrynus* (Anura: Leptodactylidae). Monografías de Herpetología. Valencia, Spain.

Blondel, J., et al. 1999. Selection-based biodiversity at a small spatial scale in a low-dispersing insular bird. Science 285:1399-1402.

Bloor, P., and R. P. Brown. 2005. Morphological variation in *Gallotia atlantica* from the volcanic island of Lanzarote: subspecies designations and recent lava flows. Biol. J. Linn. Soc. 85:395-406.

Bollmer, J.L., R. T. Kimball, N. K. Whiteman, J. H. Sarasola, and P.G. Parker. 2006. Phylogeography of the Galapagos hawk (*Buteo galapagoensis*): a recent arrival to the Galapagos Islands. Mol. Phylogen. Evol. 1:237-247.

Broadley, D.G., A.S. Whiting, and A.M Bauer. 2006. A revision of the East African species of *Melanoseps* Boulenger (Sauria: Scincidae: Feyliniinae). African Journal of Herpetology (in press).

Brown, R. P., P. A. Hoskisson, J-H. Welton, and M. Baez. 2006. Geological history and within-island diversity: a debris avalanche and the Tenerife lizard *Gallotia galloti*. Mol. Ecol. (ONLINE EARLY)

Burns, K. J., S. J. Hackett, and N. J. Klein. 2002. Phylogenetic relationships and morphological diversity in Darwin's finches and their relatives. Evolution 56:1240–1252.

Calsbeek, R., and T. B. Smith. 2003. Ocean currents mediate evolution in island lizards. Nature 426:552-555.

Caccone, A., J. P. Gibbs, V. Keitmaier, E. Suatoni, and J. R. Powell. 1999. Origin and evolutionary relationships of giant Galapagos tortoises. Proc. Natl. Acad. Sci. USA. 96: 13223–13228

Caccone, A., G. Gentile, J. P. Gibbs, T. H. Fritts, H. Snell, J. Betts., and J. Powell. 2002. Phylogeography and history of the giant Galapagos tortoises. Evolution 56:21052-2066.

Cardoso and A. P. Voger. 2005. DNA taxonomy, phylogeny and Pleistocene diversification of the *Cicindela hybrida* species group (Coleoptera: Cicindelidae). Mol. Ecol. 14: 3531-3546

Carranza, S., E. N. Arnold, J. A. Mateo, and L. F. Lopez-Jurado. 2000. Long-distance colonization and radiation in geckonid lizards, *Tarentola* (Reptilia: Geckonidae), revealed by mitochondrial DNA sequences. Proc. R. Soc. London Series B 267:637-649.

Carranza, S., E. N. Arnold, J. A. Mateo, and L. F. Lopez-Jurado. 2001. Parallel gigantism and complex colonization patterns in the Cape Verde scincid lizards *Mabuya* and *Macroscincus* (Reptilia:Scincidae) revealed by mitochondrial DNA sequences. Proc. Royal Soc. London Series. B 268:1595-1603.

Carpenter, C. C. 1966. Comparative behavior of the Galapagos lava lizards (*Tropidurus*). Pp. 269-273.*In* Bowman, I. R , ed. The Galapagos. Proceedings of the symposia of the Galapagos international scientific project. University of California Press. CA., U.S.A.

Carson, H. L., J. P. Lockwood., and E. M. Craddock. 1990. Extinction and recolonization of local populations on a growing shield volcano. Proc. Natl. Acad. Sci. USA. 87:7055-7057.

Carstens, B. C., A. L. Stevenson, J. D. Degenhardt, and J. Sullivan. 2004a. Testing nested phylogenetic and phylogeographic hypotheses in the *Plethodon vandykei* species group. Syst. Biol.53:781-792.

Carstens, B. C, J. Sullivan, L. M. Davalos, P. A. Larsen, and S. C. Pedersen. 2004b. Exploring population genetic structure in three species of Lesser Antillean bats. Mol. Ecol. 13:2557-2566.

Carstens, B. C., J. D. Degenhardt, A. L. Stevenson, and J. Sullivan. 2005. Accounting for coalescent stochasticity in testing phylogeographic hypotheses: modeling Pleistocene population structure in the Idaho giant salamander *Dicamptodon aterrimus*. Mol. Ecol. 14:255–265.

Censky, E. J., J. K. Hodge, and J. Dudley. 1998. Over-water dispersal of lizards due to hurricanes. Nature 395:556.

Chiba, S. 1999. Accelerated evolution of land snails *Mandarina* in the oceanic Bonin Islands: evidence from mitochondrial DNA sequences. Evolution 53:460-471.

Ciofi, C., G. A. Wilson., L. B. Beheregeray, J. P. Gibbs, C. Marquez, W. Tapia, H. L. Snell, A. C., and J. R. Powell. 2006. Phylogeographic history and gene flow among giant Galápagos tortoises on southern Isabela Island. Genetics 172:1727-1744.

Clement, M., D. Posada and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9: 1657–1659.

Crandall, K. A. C. 1996. Multiple interspecies transmissions of human and simian T-cell leukemia/lymphoma virus type I sequences. Mol Biol Evol. 13:115-131

Crandall, K.A., and A. R., Templeton. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134:959-969.

Collins, E.I., 2003. Genetic Variation of Yellow Warblers, *Dendroica petechia,* in Galápagos, Ecuador. MS thesis, Wake Forest University, Winston-Salem, North Carolina

Cowie, R.H., and B.S. Holland. 2006. Dispersal is fundamental to evolution on oceanic islands. J. Biogeog. 33:193-198

Cox, A. 1983. Ages of the Galapagos Islands. Pp. 11-24 *in* Bowman, R. I., M. Berson, and A. E. Leviton, eds., Patterns of Evolution in Galápagos Organisms. Pacific Division of the American Association for the advancement of Science, San Francisco, CA.

Davison, A., and S. Chiba. 2006. The recent history and population structure of five *Mandarina* snail species from subtropical Ogasura (Bonin Islands, Japan). Mol. Ecol. 15:2905-2919.

De Queiroz, A. 2005. The resurrection of oceanic dispersal in historical biogeography. Trends Ecol. Evol. 20:69-73

Dixon, J. R., and J. W. Wright. 1975. A review of the lizards of the iguanid genus *Tropidurus* in Peru. Contributions in Science of the Natural History Museum of Los Angeles County 271:1-39.

Edwards, S. V., and P. Beerli. 2000. Gene divergence, population divergence, and the variance in coalescent time in phylogeographic studies. Evolution 54: 707-716.

Eliasson, U. 1974. Studies in Galapagos Plants XIV. The Genus *Scalesia* Arn. Opera Bot 36: 1–117.

Emerson, B. C. 2002. Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. Mol. Ecol. 11:951-966.

Emerson, B. C., and P. Oromi. 2005. Diversification of the forest beetle genus *Tarphius* on the Canary Islands, and the evolutionary origin of island endemics. Evolution 59:586-598.

Emerson, B. C., S. Forgie, S. Goodacre, and P. Oromi. 2006. Testing phylogeographic predictions on an active volcanic island: *Brachyderes rugatus* (Coleptera: Curculionidae) on La Palma (Canary Islands). Mol. Ecol. 15:449-458.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

Friesen, V.L., J.A. Gonzalez, and F. Cruz-Delgado. 2006. Population genetic structure and conservation of the Galápagos petrel (*Pterodroma phaeopygia*). Conserv. Genet. 7:105-115.

Frost, D. 1992. Phylogenetic analysis and taxonomy of iguanian lizards (Reptilia: Squamata). Am. Mus. Novitates 3033:1-68.

Frost, D. R., M. T. Rodrigues, T. Grant, and T. A. Titus. 2001. Phylogenetics of the lizard genus *Tropidurus* (Squamata:Tropiduridae:Tropidurinae): direct optimization, descriptive efficiency, and sensitivity analysis of congruence between molecular data and morphology. Mol. Phylogenet. Evol. 21:352-371.

Fry, A. J. 1999. Mildly deleterious mutations in avian mitochondrial DNA: evidence from neutrality tests. Evolution 53:1617-1620.

Funk, D. J., and K. E. Olmstead. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Sys. 34:397-403.

Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925

Gantenberg, B., and P. D. Keightley. 2004. Rates of molecular evolution in nuclear genes of east Mediterranean scorpions. Evolution 58:2486-2497.

Geist, D. 1996. On the emergence and submergence of the Galápagos Islands. Noticias de Galápagos 56:5-8.

Glor R. E., M.E. Gifford, A. Larson, J. B. Losos, L. R. Schettino, A. R. C. Lara, and T. R. Jackman. 2004. Partial island submergence and speciation in an adaptive radiation: a multilocus analysis of the Cuban green anoles. Proc. Royal Soc. London SERIES B-Biological Sciences 271: 2257-2265.

Glor R.E., J. B. Losos, and A. Larson. 2005. Out of Cuba: overwater dispersal and speciation among lizards in the *Anolis carolinensis* subgroup. Mol. Ecol. 14: 2419-2432.

Goldman, N., J. P. Anderson., and A. G. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49:652-670.

Grant, P. R. 1999. Ecology and evolution of Darwin's finches. 2nd edition. Princeton University Press , Princeton, N. J., U.S.A.

Grant, P. R., and B. R. Grant. 1996. High survival of Darwin's finch hybrids: effects of beak morphology and diet. Ecology 77:500-509.

Grant, P. R., and B. R. Grant. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. Science 296:707-711.

Gubitz, T., R. S. Thorpe., and A. Malhotra. 2000. Phylogeography and natural selection in the Tenerife gecko *Tarentola delalandii*: testing historical and adaptive hypotheses. Mol. Ecol. 9:1213-1221.

Gubitz, T., R. S. Thorpe., and A. Malhotra. 2005. The dynamics of genetic and morphological variation on volcanic islands. Proc. Royal Soc. London, Series B. 272:751-757.

Guindon, S. and O. Gascuel. 2003. A simple fast and accurate algorithm to estimate large phylogenies by maximum likelihood approach. Mol. Biol. 52:696-704.

Guilderson, T. P., and D. P. Schrag. 1998. Tropical Pacific Associated with Changes in El Niño Abrupt Shift in Subsurface Temperatures in the. Science 281:241-243

Harpending, H., and A. Rogers. 2000. Genetic perspectives on human origins and differentiation. Ann. Rev. Gen. Hum. Genet. 1:361-385

Harvey, M. B., and R. L. Gutberlet Jr. 2000. A phylogenetic analysis of tropidurine lizards (Squamata: Tropiduridae), including new characters of squamation and epidermal microstructure. Zool. J. Linn. Soc. 128:189-233.

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.

Heise, P. J. 1998. Phylogeny and biogeography of Galapagos lava lizards (*Microlophus*) inferred from nucleotide sequence variation in mitochondrial DNA. Ph.D. disssertation, University of Tennessee, Knoxville, TN. 206 p.

Hey, J. 1994. Bridging phylogenetics and population genetics with gene tree models. Pp. 435-449 *in.* B. Schierwater, B. Streit, G. Wagner and R. DeSalle, eds. Molecular Approaches to Ecology and Evolution. Birkhäuser-Verlag, Basel.

Hickman, C. S, and J. H. Lipps. 1984. Geologic youth of Galapagos Islands confirmed by marine stratigraphy and paleontology. Science 227:1578-1580.

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.

Huelsenbeck., J.P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. Bionformatics 17:754-755.

Jackson, M. H. 1993. Galapagos: A Natural History. Univ. of Calgary Press, Calgary, Canada.

Johnson, K. P., and J. Seger. 2001. Elevated rates of nonsynonymous substitution in island birds. Mol. Biol. Evol. 18:874-881.

Jordal, B.H., B.C. Emerson, and G.M. Hewitt. 2006. Apparent 'sympatric' speciation in ecologically similar herbivorous beetles facilitated by multiple colonizations of an island. Mol. Ecol. 15:2935-2947.

Jordan, M. A., R. L. Hammond, H. L. Snell, H. M. Snell., and W. C. Jordan. 2002. Isolation and characterization of microsatellite loci from Galapagos lava lizards (*Microlophus* spp.). Mol. Ecol. Notes 2:349-350

Jordan, M. A., H. L. Snell, H. M. Snell., and W. C. Jordan. 2005a. Phenotypic divergence despite high levels of gene flow in Galapagos lava lizards (*Microlophus albemarlensis*). Mol. Ecol. 14:859-867.

Jordan, S., C. Simon, D. Foote., and R. A. Englund. 2005b. Phylogeographic patterns of Hawaiian *Megalagrion* damselflies (Odonata: Coenagrionidae) correlate with Pleistocene island boundaries. Mol. Ecol. 14:3457-3470

Joy, D. A., and J. E. Conn. 2001. Molecular and morphological analysis of an insular radiation in pacific black flies (*Simulium*). Syst. Biol. 50:18-38.

Juan, C., B. C. Emerson, P. Oromí, and G. M. Hewitt. 2000. Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. Trends. Ecol. Evol. 15:104-109.

Kizirian, D, A., and M. A. Donnelly. 2004. The criterion of reciprocal monophyly and classification of nested diversity at the species level. Mol. Phylogenet. Evol. 32:1072-1076.

Kizirian, D., A. Trager, M. A. Donnelly, and J. W. Wright. 2004. Evolution of Galapagos island lizards (Iguania: Tropiduridae: *Microlophus*). Mol. Phylogenet. Evol. 32: 761-769.

Knowles, L, L. and W. P. Maddison. 2002. Statistical phylogeography. Mol. Ecol. 11:2623-2635.

Kolbe, J.J., R.E. Glor, L. Rodriguez-Schettino, A.C. Lara, A. Larson, and J.B. Losos. 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature 431:177-180

Kricher, J. 2002. Galapagos. Smithsonian Natural History Series. Smithsonian. Washington. 221 pp.

Lewis, P. O., M. T. Holder, and K. E. Holsinger. 2005. Polytomies and Bayesian phylogenetic inference. Syst. Biol. 54:241-253.

Lopez, T.J., E.D. Hauselman, L.R. Maxson, and J.W. Wright. 1992. Preliminary analysis of phylogenetic relationships among Galapagos Island lizards of the genus *Tropidurus*. Amphibia-Reptilia 13: 327–339.

Losos, J.B., T.R. Jackman, A. Larson, K. de Queiroz, and L. Rodríguez-Schettino. 1998. Historical Contingency and determinism in replicated adaptive radiations of island lizards. Science. 279:2115-2118.

Macey, J.R., J. A. Schulte II, N. B. Ananjeva A. Larson, N. Rastegar-Pouyani, S. M. Shammakov, T. J. Pappenfuss. 1998. Phylogenetic relationships among agamid lizards of the *Laudakia caucasia* species group: testing hypotheses of biogeographic fragmentation and an area cladogram for the Iranian Plateau. Mol. Phylogenet. Evol. 10: 118–131.

Macey, J.R., Y., Wang, N. B. Ananjeva, A., Larson, and T.J., Pappenfuss. 1999. Vicariant patterns of fragmentation among gekkonid lizards of the genus *Teratoscincus* produced by the Indian collision: a molecular phylogenetic perspective and an area cladogram for Central Asia. Mol. Phylogenet. Evol. 12: 320–332.

Malhotra, A., and R. S. Thorpe. 2000. The dynamics of natural selection and vicariance in the Dominican anole: patterns of within-island molecular and morphological divergence. Evolution 54:245-258.

Mendelson, T. C., and K. L. Shaw. 2005. Rapid speciation in an arthropod. The Likely force behind an explosion of new Hawaiian cricket species is revealed. Nature 433: 375-376.

Miles, D. B., H. L. Snell, and H. M. Snell. 2001. Intrapopulation variation in endurance of Galapagos lava lizards (*Microlophus albemarlensis*): evidence between natural and sexual selection. Evol. Ecol. Res. 3:795-804.

Minin, V., Z. Abdo, P. Joyce, and J. Sullivan. 2003. Performance-based selection of likelihood models for phylogenetic estimation. Syst. Biol. 52:674-683

Moore, S. W. 1995. Inferring phylogenies from mtDNA variation: Mitochondrial gene trees versus nuclear-gene trees. Evolution 49:718-726

Morando, M., Avila, L.J., C. Turner and J. W. Sites, Jr. 2006. Molecular evidence for a species complex in the patagonian lizard *Liolaemus bibronii* and phylogeography of the closely related *Liolaemus gracilis* (Squamata: Liolaemidae). Mol. Phylogenet. Evol. 28:xxx-xxx.

Moritz, C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. Syst. Biol. 51:238-254.

Moya , O., H G. Contreras-Diaz, P. Oromi, and C. Juan. 2004. Genetic structure, phylogeography and demography of two ground-beetle species endemic to the Tenerife laurel forest (Canary Islands). Mol. Ecol. 13:3153-3167.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia Univ. Press, New York.

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.

Nielsen, R., and J. Wakeley. 2001. Distinguishing migration from isolation. A Markov chain Monte Carlo approach. Genetics 158:885-96.

Nielsen, L. R. 2004. Molecular differentiation within and among island populations of the endemic plant *Escalesia affinis* (Asteraceae) from the Galapagos Islands. Heredity 93:434-442.

Nichols, H. 2004. One of a kind. Nature 429:498-500.

Nordborg, M. 2001. Coalescent theory. Pages 1-37 *in* D. Balding, M. Bishop, and C. Cannings, eds. Handbook of Statistical Genetics. Wiley, Chichester, UK.

Rambaut, A., and A. J. Drummond. 2003. Tracer v1.2, http:// evolve.zoo.ox.ac.uk/.

Rambaut, A., and N. C. Grassly. 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comp. Appl. Biosci. 13:235-238.

Rassmann, K. 1997. Evolutionary age of the Galapagos iguanas predates the age of the present Galapagos Islands. Mol. Phyl. Evol. 7:158-172.

Raxworthy, C. J., Forstner, M. R. J., and R. A. Nussbaum. 2002. Chameleon radiation by oceanic dispersal. Nature 415:784-787.

Rees, D. J., Emerson, R.C., Oromi, P., and G. M. Hewitt. 2001. Reconciling gene trees with organism history: the mtDNA of three *Nesotes* species (Coleoptera: Tenebronidae) on the western Canary islands. J. Evol. Biol. 14:129-147.

Rogers, A. R., and H. Harpending. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9:552-569.

Rubinoff, D., and B.S. Holland. 2005. Between the two extremes: Mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst. Biol. 54: 92-961.

Russello, M. A., S. Glaberman, J. P. Gibbs, C. Marquez, J. R. Powell, and A. Caccone. 2005. A cryptic taxon of Galapagos tortoise in conservation peril. Biol. Lett. 1:287-290.

Ogden, R., and R.S. Thorpe. 2002. Molecular evidence for ecological speciation in tropical habitats. Proc. Nat. Acad. Sci. USA 99:13612-13615

Pak, H., and J. R. V. Zaneveld. 1973. The Cromwell Current on the east side of the Galapagos Islands. J. Geophys. Res. 20: 7845–7859

Palumbi, S. R. 1996. Nucleic acids I: The polymerase chain reaction. Pp. 205–247 *in* D. M. Hillis, C. Moritz, and B. K. Mable, eds. Molecular Systematics. Sinauer, Sunderland, Massachusetts.

Panhuis, T.M., R. Butlin, M. Zuk, T. Tregenza. 2001. Sexual selection and speciation. Trends Ecol. Evol. 16: 364-337.

Pestano, J., and R. P. Brown. 1999. Geographical structuring of mitochondrial DNA in *Chalcides sexlineatus* within the island of Gran Canaria. Proc. R. Soc. London Series B 266:805-812.

Petren, K., B.R. Grant, P.R. Grant. 1999. A phylogeny of Darwin's finches based on microsatellite DNA length variation. Proc. Roy. Soc. Lond. B 266: 321-329.

Petren, K., B.R. Grant & P.R. Grant, & L. F. Keller. 2005. Comparative landscape genetics and the adaptive radiation of Darwin's finches: the role of peripheral isolation. Molecular Ecology 14: 2943-2957.

Pfenninger, M., and D. Posada. 2002. Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. Evolution 56:1776–1788.

Posada, D. A., K. A. C. 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.

Posada, D. A., K. A. C. Crandall, and A. R. Templeton. 2006. Nested clade analysis statistics. Mol. Ecol. Notes 6: 590-593.

Sato, A., C. Ohigin, F. Figueroa, P. R. Grant, B. R. Grant, H. Tichy, and J. Klein. 1999. Phylogeny of Darwin's finches as revealed by mtDNA sequences. Proc. Natl. Acad. Sci. 96:5101-5106.

Sato, A., H. Tichy, C. Ohigin, P. R. Grant, B. R. Grant, and J. Klein. 2001. On the origin of Darwin's finches. Mol. Biol. Evol. 18:299-311.

Schilling, E. E., J. L. Panero, and U. H. Eliasson. 1994. Evidenc e from chloroplast DNA restriction site analysis on the relationships of Scalesia (Asteraceae: Heliantheae). Am J Bot 81: 248–254

Shimodaira, H., and M. Hasegawa.1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16:1114-1116.

Schneider, S. D., Roessli D., and L. Excoffier. 2000. Arlequin ver. 2.0: a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Shaw, K. L. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proc. Natl. Acad. Sci. 9:16122-16127.

Snell, H. M., P.A. Stone, and H.L. Snell. 1996. A summary of geographical characteristics of the Galapagos Islands. J. Biogeog. 23:619-624.

Snell, H. M., R. D. Jennings, H. M. Snell, and S. Harcourt. 1988. Intrapopulation variation in predator-avoidance performance of Galapagos lava lizards: the interaction of sexual and natural selection. Evolutionary Ecology 2:353-369.

Spiller, D. A., J. B. Losos, and T. W. Schoener. 1998. Impact of catastrophic hurricane on island populations. Science 281:695-697.

Sullivan, J., and P. Joyce. 2005. Model selection in phylogenetics. Ann. Review of Ecol. and Syst. 36:445-466.

Swofford, D. L. 2000. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, MA.

Tajima, F. 1989. The effect of change in population size on DNA polymorphism. Genetics 123:597–601

Templeton, A. R., E., Boewrwinkle, and C. F. Sing. 1987. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. Genetics 117:343-351.

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.

Templeton, A. R., E. Routman, and 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 N. J. Georgiadis. 1996. A Landscape Approach to Conservation Genetics: Conserving Evolutionary Processes in the African Bovidae. Pp. 419-

429 *in* J. C. Avise, and J. L. Hamrick eds. Conservation Genetics: Case Histories from Nature. Chapman & Hall, NY.

Templeton, A. R., Robertson, R. J., Brisson, J., and J. Strasburg. 2001. Disrupting evolutionary processes: the effect of habitat fragmentation on collared lizards in the Missouri Ozarks. Proc. Natl. Acad. Sci. 98:5426-5432.

Templeton, A. R. 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors. Mol. Ecol. 13: 789-809.

Thorpe, R. S., L. Leadbeather, and C. E. Pook. 2005. Molecular clock and geological dates: cytochrome *b* of *Anolis exremus* substantially contradicts dating of Barbados emergence. Mol. Ecol. 14:2087-2096.

Tonnis B, Grant PR, Grant BR, Petren K .2005. Habitat selection and ecological speciation in Galápagos warbler finches (*Certhidea olivacea* and *C. fusca*). Proc. Roy. Soc. of London. Ser. B, Bio. Sci. 272:819–826.

Van Denburgh, J., and J. R. Slevin. 1913. Expedition of the California Academy of Sciences to the Galapagos Islands, 1905-1906. IX. The Galapagoan lizards of the genus *Tropidurus* with notes on iguanas of the genera *Conolophus* and *Amblyrhinchus*. Proc. California Acad. Sci., ser. 4, 2:132-202.

Vandergast A. G., R. G. Gillespie, and G. K . Roderick. 2004. Influence of volcanic activity on the population genetic structure of Hawaiian *Tetragnatha* spiders: fragmentation, rapid population growth and the potential for accelerated evolution. Mol. Ecol. 13:1729-1743.

Watkins, G. G. 1996. Proximate causes of sexual size dimorphism in the iguanian lizard *Microlophus occipitalis*. Ecology 77:1473-1482.

Watkins, G. G. 1997. Inter-sexual signaling and the functions of female coloration in the tropidurid lizard *Microlophus occipitalis*. Anim. Behav. 53:843-852.

Watkins, G. G. 1998. Function of a secondary sexual ornament: the crest in the South American iguanian lizard *Microlophus occipitalis* (Peters, Tropiduridae). Herpetologica 54:161-169.

Werner, D. I. 1978. On the biology of *Tropidurus delanonis*, Baur (Iguanidae). Z. Tierpsychol. 47:337-395.

White, W. M., A. R. McBirney, and R.A. Duncan. 1993. Petrology and geochemistry of the Galapagos Islands: portrait of a pathological mantle. J. Geophys. Res. 98:533-563.

Whiting, A. S., A. M. Bauer, and J. W. Sites, Jr. 2003. Phylogenetic relationships and limb loss in sub-Saharan African scincine lizards (Squamata: Scincidae). Mol. Phylogenet. Evol. 29:582-598.

Wiens, J. J., and T. Penkrot. 2002. Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). Syst. Biol. 51:69-91.

Wright, J. W. 1983. The evolution and biogeography of the lizards of the Galapagos Archipelago: evolutionary genetics of *Phyllodactylus* and *Tropidurus* populations. Pp. 123-155. *In*. Bowman, R. I., M. Berson, and A. E. Leviton eds. Patterns of evolution in Galápagos Organisms. Pacific Division of the American Association for the advancement of Science, San Francisco

Wright, J. W. 1984. The origin and evolution of lizards of the Galapagos Islands. Terra 21-27.

Wyrtki, K., E. Stroup, W. Patzert, R. Williams, and W. Quinn, W. 1976. Predicting and observing El Niño. Science 191 (4225), 343–346

TABLE 1. Internal PCR primers used for amplification/sequencing of selected samples in this study; the 'jai' primers were designed for *M. albemarlensis, M. jacobi,* and *M. indefatigabilis*; 'gd' primers for *M. delanonis* and *M. grayii*; and 'dunpac" primers for *M. duncanensis* and *M. pacificus*.

Primer name		Sequence
Ctb_Wfjai	5'	AAATCATCGTTGTWATTCAACTAC
Ctb_F1jai1	5'	TGRGGRCARATRTCYTTCTGRGG
Ctb_F1jai2	5'	CTRCTWGTAATAGCRACAGCCTTC
Ctb_WRjai2	5'	GAA GGC TGT YGC TAT TAC WAG YAG
Ctb_WRjai1	5'	CTR CTW GTA ATA GCR ACA GCC TTC
Ctb_C3jai2	5'	TGT ARG AGA ART ATG GGT GAA ATG G
Ctb_C3jai1	5'	AAR TAY CAY TCT GGT TTR ATG TG
Ctb_WFgd1	5'	TTATTCAACTACAAAAACMAATGAC
Ctb_F1gd1	5'	TAATAGCAACAGCCTTCATAGG
Ctb_F1gd2	5'	AATGAATCTGRGGCGGGTTCTC
Ctb_C3gd1	5'	ACT CTG GCT TAA TGT GKG GTG GTG
Ctb_WRgd1	5'	GCCAACCAGTAGAAGACCCWTTTAT
Ctb_C3dunpac	5'	TGG GGT GTA GTT TTC TGG GTC GC
Ctb_F1dunpac	5'	TGAGGACAAATATCYTTCTGAGG
Ctb_WRdunpac	5'	TGATGATAAATGGGTCTTCTACTGG

TABLE 2. Results of pairwise topological constraint tests of our best-supported phylogenetic hypothesis (reduced to a species-only topology in Fig. 2) against those proposed by Heise (1998: his Fig 3.4, p. 65), and Kizirian et al. (2004: their Fig 3, p. 766). The one-tailed Shimodaira-Hasegawa test results are based on a model calculated with DT-ModSel (TrN + I + G), and a RELL distribution and 10,000 bootstrap replicates, as implemented in PAUP*. Alternative topological constraints are given below for each topology inside of *M. delanonis*, which is basal in all trees; species abbreviations are: alb – *albemarlensis*; dun – *duncanensis*; gray – *grayii*; indef – *indefatigabilis*; jac – *jacobi*; and pac - *pacificus*.

Tree (source)	Shimodaira-Hasegawa		Proposed phylogenetic relationship
	-lnL	*P*	
1. (This study)	6587.11525		(gray (((dun + jac) indef) + (alb + pac)))
2. (Heise, 1998)	11064.85419	0.000	(gray ((pac + alb) + ((indef + jac) dun)))
3. (Kizirian et al., 2004)	22701.90465	0.000	(((gray (alb+ pac)) + ((jac + dun) indef)))

TABLE 3. Clades identified by resolved networks/nesting designs (Figs. 4 – 7) for which significant clade distances (Dc) or nested clade distances (Dn) were obtained. Four main islands showed significant population structure; Isabela and Fernandina (*M. albemarlensis*), Santa Cruz (*M. indefatigabilis*), and Santiago (*M. jacobi*).

Haploclade groups	χ^2 statistic	*P*	Chain of inference	Demographic inference
Clades nested within Isla Santiago (*M. jacobi*)				
Clade 1-2	11.1522	0.0070	1-2-3-5-N-6-N-8-Y	Restricted gene flow with some long distance dispersal
Clade 1-3	43.0490	0.0230	1-N-2-Y-3-5-15-N	Past fragmentation or long distance colonization
Clade 2-1	35.0000	0.0000	1-2-11-RE-12-CRE	Continuous range expansion
Clade 2-2	70.8996	0.0000	1-2-3-5-15-N-21	Past fragmentation or long distance colonization
Clade 2-4	7.0000	0.0420	1-2-3-4-N	Restricted gene flow with isolation-by-distance
Clade 3-1	87.7061	0.0050	1-2-3-5-6-7-Y	Restricted gene flow/dispersal with some long distance dispersal
Clade 4-1	70.8210	0.0000	1-2-3-5-6-7-8-Y	Restricted gene flow/dispersal with some long distance dispersal
Total cladogram	72.8210	0.0000	1-2-3-5-6-7-Y	Restricted gene flow/dispersal with some long distance dispersal
Clades nested within Isla Santa Cruz (*M. indefatigabilis*)				
Clade 1-12	22.0000	0.0000	1-2-Y-3-Y-5-Y-15-N	Past fragmentation and/or long distance colonization
Clade 1-16	28.0000	0.0000	1-2-Y-3-Y-5-6-N-7-N-8-Y	Restricted gene flow/dispersal, with some long distance dispersal
Clade 1-19	22.0000	0.0000	1-Y-19-N-	Allopatric fragmentation
Clade 1-21	11.0000	0.0000	1-2-Y-3-Y-5-Y-15-N	Past fragmentation or long distance colonization
Clade 1-26	13.5354	0.0040	1-2-Y-3-Y-5-N	Past fragmentation
Clade 2-7	9.9167	0.0340	1-N-2-Y-3-4-N	Restricted gene flow with some isolation-by-distance
Clade 2-11	25.0000	0.0000	1-2-3-Y-5-15-N	Past fragmentation or long distance colonization
Clade 2-13	18.0000	0.0000	1-2-Y-3-Y-5-Y-15-N	Past fragmentation and/or long distance colonization
Clade 2-15	17.1542	0.0110	1-2-Y-3-N-4-N	Restricted gene flow with some isolation-by-distance
Clade 3-1	21.4268	0.0000	1-N-2-3-5-N-6-N-7-8-Y	Restricted gene flow/dispersal or past gene flow followed

				by extinction of intermediate populations
Clade 3-2	14.0000	0.0020	1-2-N-11-Y-RE	Range expansion
Clade 3-3	22.587	0.0000	1-2-3-5-N-6-N	Restricted gene flow/dispersal
Clade 3-4	117.0000	0.0000	1-2-Y-3-N-4-N	Restricted gene flow with some isolation-by-distance
Clade 3-5	30.0000	0.0000	1-2-Y-3-N-4-N	Restricted gene flow with isolation-by-distance
Clade 4-1	30.8518	0.0000	1-2-Y-3-5-Y-15-N	Past fragmentation and/or long distance colonization
Clade 4-2	84.0000	0.0000	1-2-N-11-Y	Range expansion
Clade 5-2	49.7432	0.0000	1-2-Y-3-N-4-N	Restricted gene flow with isolation-by-distance
Clade 5-1	111.6312	0.0000	1-2-Y-3-N-4-N	Restricted gene flow with isolation-by-distance
Total cladogram	1.0000	0.0000	1-2-Y-3-N-4-N	Restricted gene flow with isolation-by-distance
Clades nested within Isla Isabela and Fernandina (*M. albemarlensis*)				
Western Clade				
Clade 1-1	120.5031	0.0000	1-2-N-11-Y-RE-12-Y-13--Y	Continuous range expansion and long distance colonization possibly coupled with subsequent fragmentation
Clade 1-22	80.0781	0.0000	1-2-3-N-4-N	Restricted gene flow with isolation-by-distance
Clade 2-2	176.8515	0.0000	1-2-3-Y-5-N-6-N-7-Y	Restricted gene flow/dispersal
Clade 2-6	47.7840	0.0157	1-2-N-11-Y	Range expansion
Clade 3-1	72.4483	0.0102	1-2-3-N-4-N	Restricted gene flow with isolation-by-distance
Clade 3-2	13.0000	0.0037	1-2-3-5-N-6-N-7-Y	Restricted gene flow/dispersal but with some long distance dispersal
Total cladogram	124.4339	0.0000	1-2-4-N	Restricted gene flow with isolation-by-distance
Eastern Clade				
Clade 2-3	17.0000	0.0000	1-2-Y-3-N-4-N	Restricted gene flow with isolation-by-distance
Clade 2-7	8.0000	0.0194	1-2-3-Y-5-15-N	Past fragmentation or long distance colonization
Clade 3-1	21.000	0.0000	1-2-3-Y-5-Y-15-N	Past fragmentation or long distance colonization
Clade 3-2	18.000	0.0002	1-2-3-5-Y-13-N	Past fragmentation or long distance colonization
Clade 3-3	32.3042	0.0000	1-2-3-Y-5-Y-15-N	Past fragmentation or long distance colonization
Total cladogram	65.000	0.0000	1-2-3-5-N-6-7-Y	Restricted gene flow/dispersal but with some long distance dispersal

TABLE 4. Genetic diversity measures of Galapagos lava lizards from eight major islands (and associated sampled islets) inhabited by the seven species in the Western Radiation. The table shows the number of individuals and number of haplotypes per population, with nucleotide diversity estimates and Tajima's *D* and Fu's *Fs* neutrality tests.

Island	Number of populations	No. of haplotypes/ No. of specimens	No. of polymorphic Sites	Gene diversity	Nucleotide diversity	Mean # of pairwise differences	Tajima's *D*	Fu's *Fs*
Fernandina	8	9/50	16	0.9784 +/- 0.0017	0.001371 +/- 0.000935	1.398504 +/- 0.861884	-1.23986 (P=0.10351)	-28.42 (P=0.0000)
Isabela	19	72/157	100	0.9928 +/- 0.0001	0.017275 +/- 0.008511	16.480731 +/- 7.343860	0.86257 (P=0.84000)	-23.21 (P=0.03000)
Pinzon	2	7/14	7	0.9684 +/- 0.0224	0.001588 +/- 0.001113	1.510526 +/- 0.947963	-0.77080 (P=0.24500)	-14.40 (P=0.00000)
Santiago	13	26/93	38	0.9878 +/- 0.0004	0.004164 +/- 0.002305	4.001546 +/- 2.003218	-0.68394 P=0.29500	-25.03 (P=0.00000)
Pinta	1	3/3	4	1.0000 +/- 0.2722	0.002669 +/- 0.002396	2.666667 +/- 1.918994	0.0000 0.82600	-0.34093 (P=0.19700)
Santa Cruz	23	56/214	85	0.9942 +/- 0.0001	0.018495 +/- 0.009104	16.737761 +/- 7.452396	1.72561 (P=0.94400)	0.00000 (P=0.00000)
Floreana	7	9/47	56	0.9784 +/- 0.0019	0.012620 +/- 0.006347	11.913281 +/- 5.414329	0.69947 (P=0.79300)	-5.34977 (P=0.16500)
Española	5	5/36	9	0.9763 +/- 0.0026	0.002437 +/- 0.001480	2.368595 +/- 1.299547	1.00270 (P=0.86600)	-27.01 (P=0.00000)

TABLE 5. Analyses of molecular variance (AMOVA); the islands of Pinzon (*M. duncanensis*) and Pinta (*M. pacificus*) are not included because of limited sampling. Also shown are estimates of island age (from Geist, 1996; in parentheses are Geist "best estimates" pers. com., 2006), island area, the percentage each island covered by vegetation (Pinta = 62.0 % and Pinzon = 95.2 %), and the mismatch distribution sum of squared deviations and raggedness indexes. NS: non significant, ***: $p < 0.0001$

Island (# offshore islets sampled)	Source of variation	df	% variation	*P*-value	Island age My.	Island area (km^2)	% of island vegetated	*P*(SSD)	*P*(*H*rag)
Fernandina (1)	among	7	14.41	<0.0001	300,000	647.6	30.5	0.95	0.90
(M. albemarlensis)	within	200	85.59		(0.06)			NS	NS
	total	207							
Isabela (4)	among	18	74.32	<0.0001	< 700,000	4710.7	66.5	0.34	0.75
(M. albemarlensis)	within	1644	25.68		(400,000)			NS	NS
	total	1662							
Santiago (7)	among	12	85.40	<0.0001	~ 1.2	577.5	68.6	0.04	0.16
(M. jacobi)	within	630	14.60		(0.8)			NS	NS
	total	642							
Floreana (4)	among	6	98.98	<0.0001	~ 1.5	172.5	--	0.001	0.001
(M. grayii)	within	177	1.02		(1.5)			***	***
	total	183							
Santa Cruz – Santa Fe (14)	among	22	95.68	<0.0001	~ 2.2	984.1-24.8	100.0-	0.14	0.26
(M. indefatigabilis)	within	2466	4.32		(2.2)		100.0	NS	NS
	total	2488							
Española (4)	among	4	100.00	<0.0001	~ 3.3	61.1	98.2	0.45	0.53
(M. delanonis)	within	116	0.00		(2.7)			NS	NS
	total	120							

LIST OF FIGURES

inferences for selected clades with significant values: LDC = long-distance colonization, RGF = restricted gene flow, RE = range expansion (all inferences are given in Table 3).

Figure 5. Nesting design for the TCS networks of cyt-*b* haplotypes observed in 22 samples of *M. indefatigabilis* from Santa Cruz and associated islets; symbolism is identical to that in Fig. 4, and the dotted line provisionally links haplotypes from Santa Fe and Santa Fe Islet to the Santa Cruz network. The map depicts the distributions of haploclades, and inferences for selected clades: LDC = long distance colonization, PF = past fragmentation, RE = recent expansion, RGF = restricted gene flow (all inferences are given in Table 3).

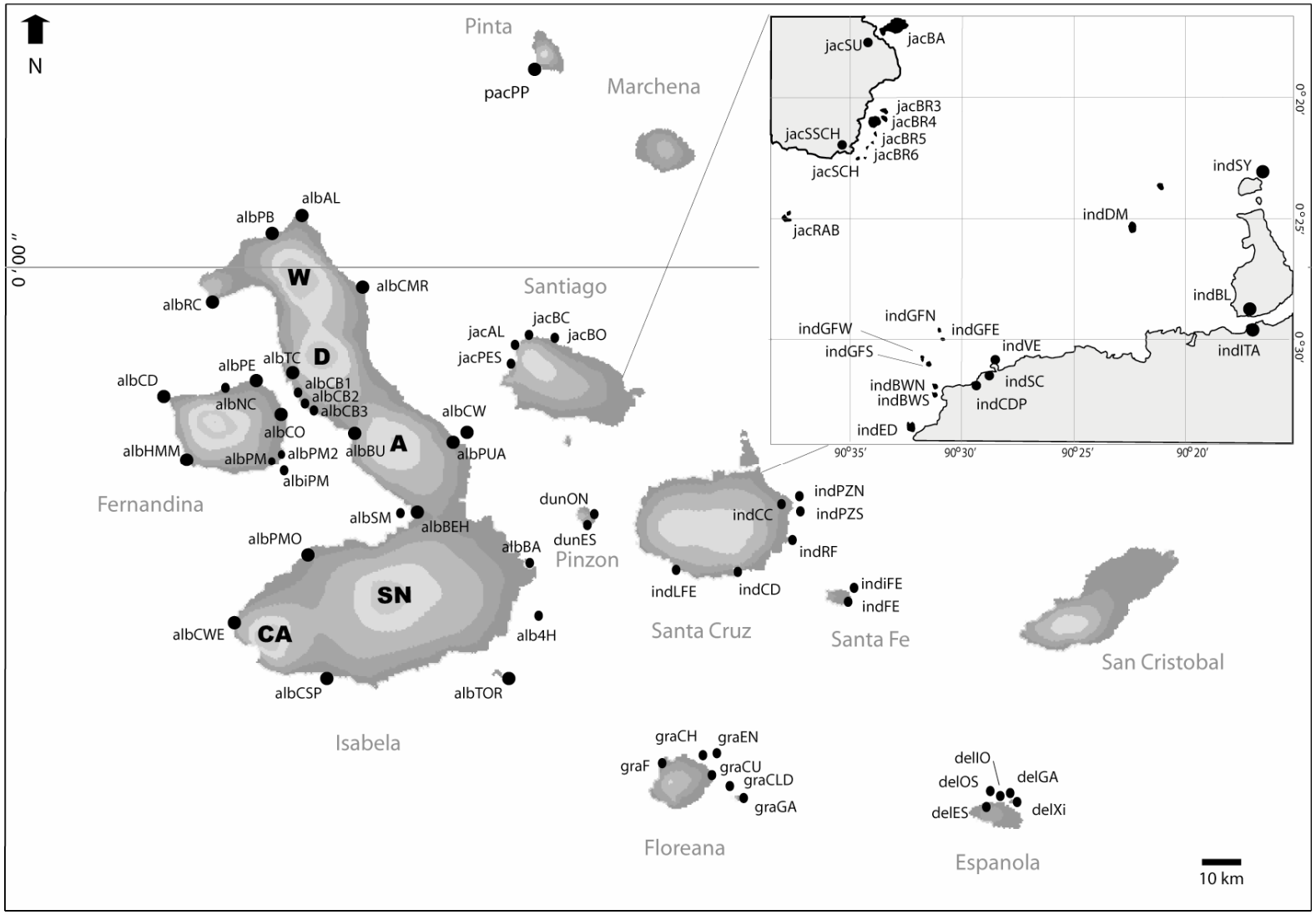
Figure 6. Nesting design for the TCS network of cyt-b haplotypes observed in the western clade of *M. albemarlensis* from Isabela and, Fernandina Islands, and associated islets; symbolism is identical to that in Fig. 4. The map depicts the distribution of haploclades and inferences made for selected clades: CRE = continuous range expansion, RE = recent expansion, RGF = restricted gene flow, RGF-IBD = restricted gene flow + isolation-by-distance (all inferences are given in Table 3).

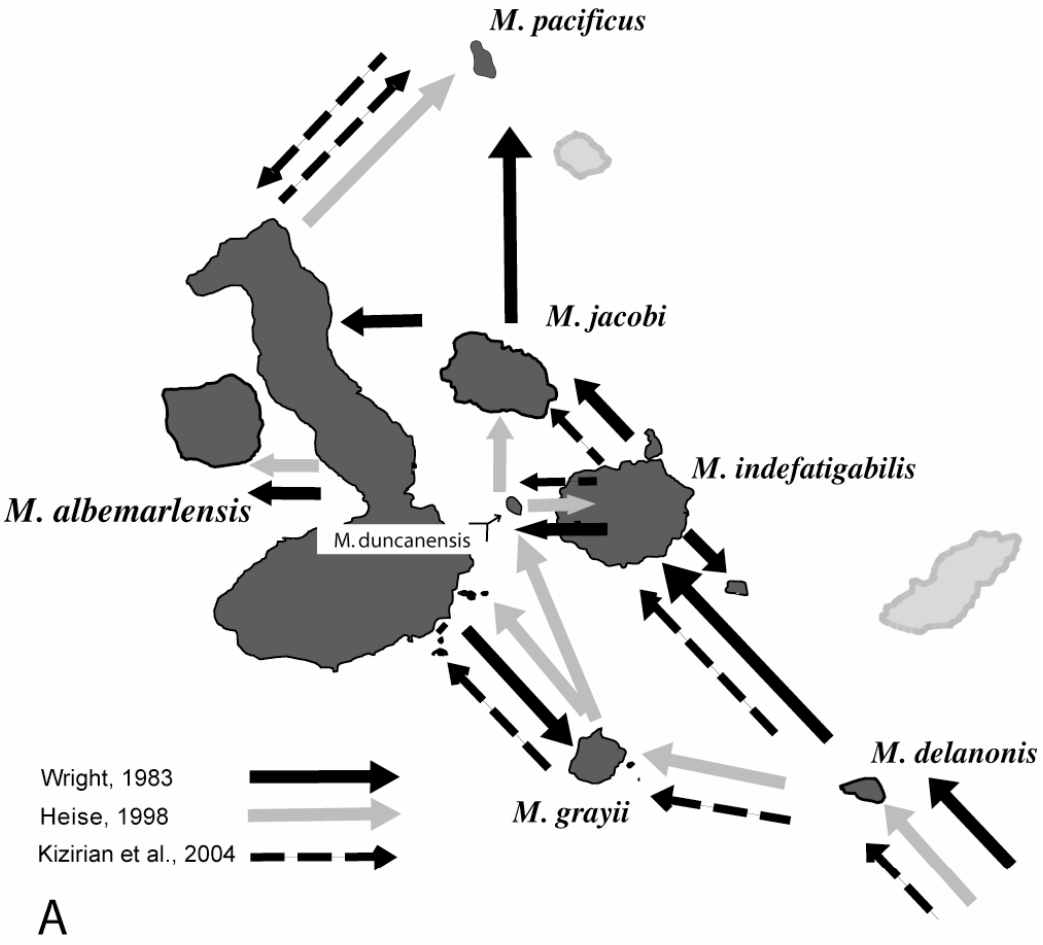
Figure 7. Nesting design for the TCS networks of cyt-b haplotypes observed in the eastern clade of *M. albemarlensis* from Isabela and associated islets; symbolism is identical to that in Fig. 4. The map depicts the distribution of haploclades and inferences made for selected clades: LDC = long distance colonization, RGF-IBD = restricted gene flow + isolation-by-distance (all inferences are given in Table 3).

Figure 8. Mismatch distributions for five species representing six major islands in the Western Radiation; this analysis is only possible for islands from which multiple populations were sampled. In each graph, diamonds represent the observed relative frequencies of nucleotide differences between pairs of individuals, and the red curve display the distribution fitted to the data under a model of population expansion.

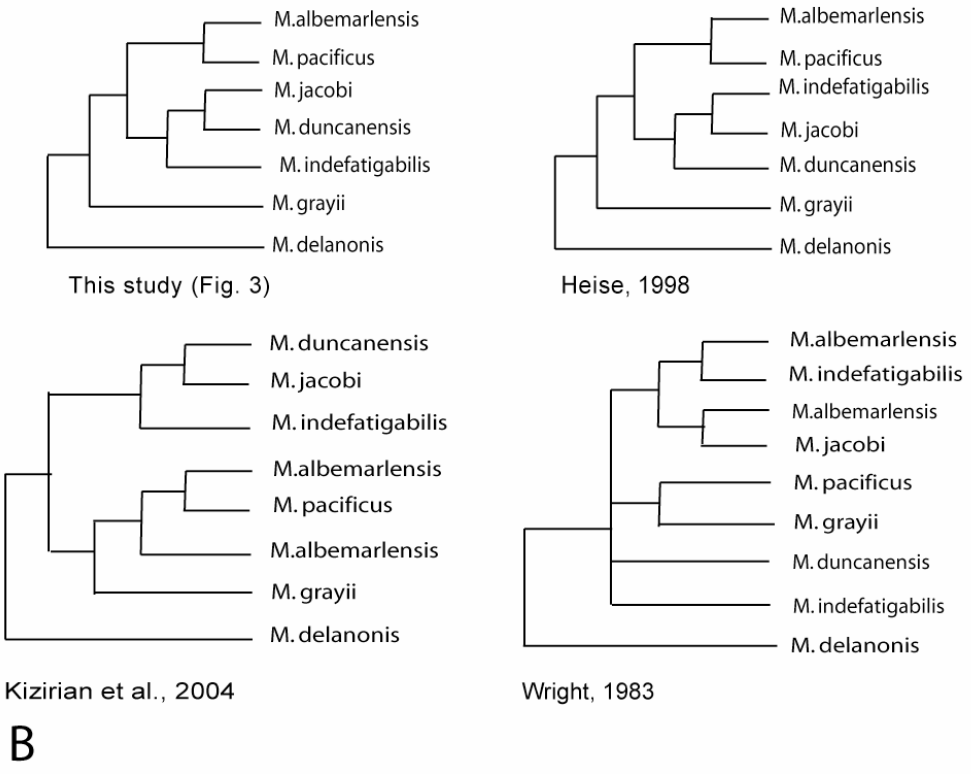
Figure 9. Hypothesized sequence of colonization events of the Western Radiation *Microlophus* implied from the phylogenetic hypothesis (simplified from Fig. 3); solid

arrows show unambiguous colonization routes inferred from the tree (numbered 1 and 2), while dashed lines indicate possible alternative routes (several for branches 3, 3a, and 3b). The solid circles on the tree identify nodes with strong support (as defined in the text).

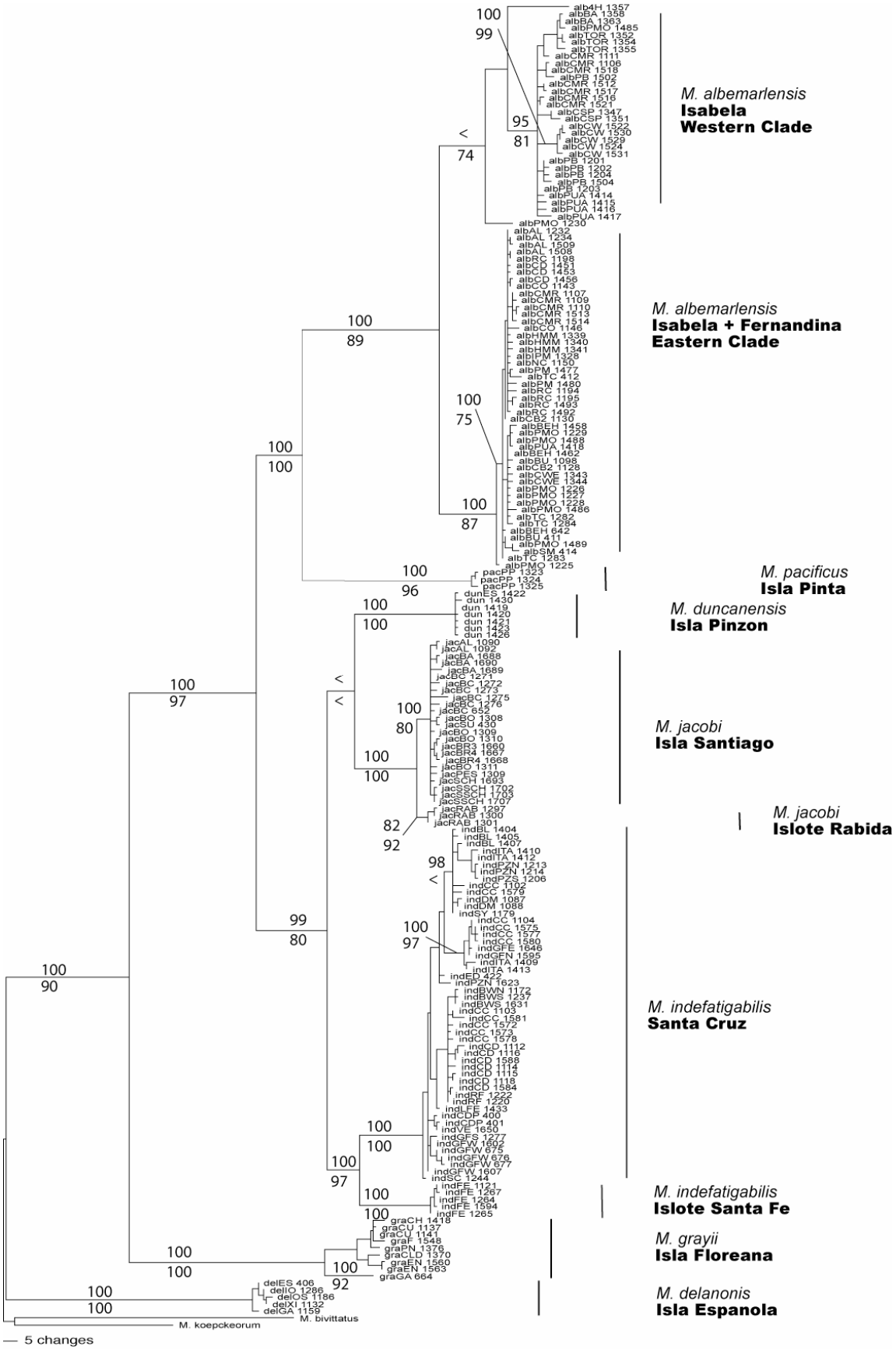

N
0 ° 00 ′ ″
Pinta
pacPP
Marchena
albAL
albPB
W
albCMR
albRC
Santiago
jacBC
jacAL
jacBO
jacPES
albTC
D
albPE
albCD
albCB1
albCB2
albNC
albCB3
albCW
albCO
albBU
A
albPUA
albHMM
albPM
albPM2
albiPM
Fernandina
dunON
albSM
albBEH
dunES
albPMO
albBA
Pinzon
SN
CA
albCWE
alb4H
albCSP
albTOR
Isabela
indPZN
indCC
indPZS
indRF
indLFE
indCD
Santa Cruz
indiFE
indFE
Santa Fe
San Cristobal
graCH
graEN
graF
graCU
graCLD
graGA
Floreana
delIO
delOS
delGA
delES
delXi
Espanola
10 km
jacSU
jacBA
jacBR3
jacBR4
jacSSCH
jacBR5
jacBR6
jacSCH
jacRAB
indSY
indDM
indBL
indGFW
indGFN
indGFE
indVE
indGFS
indITA
indBWN
indSC
indBWS
indCDP
indED
90°35′
90°30′
90°25′
90°20′
0°20′
0°25′
0°30′

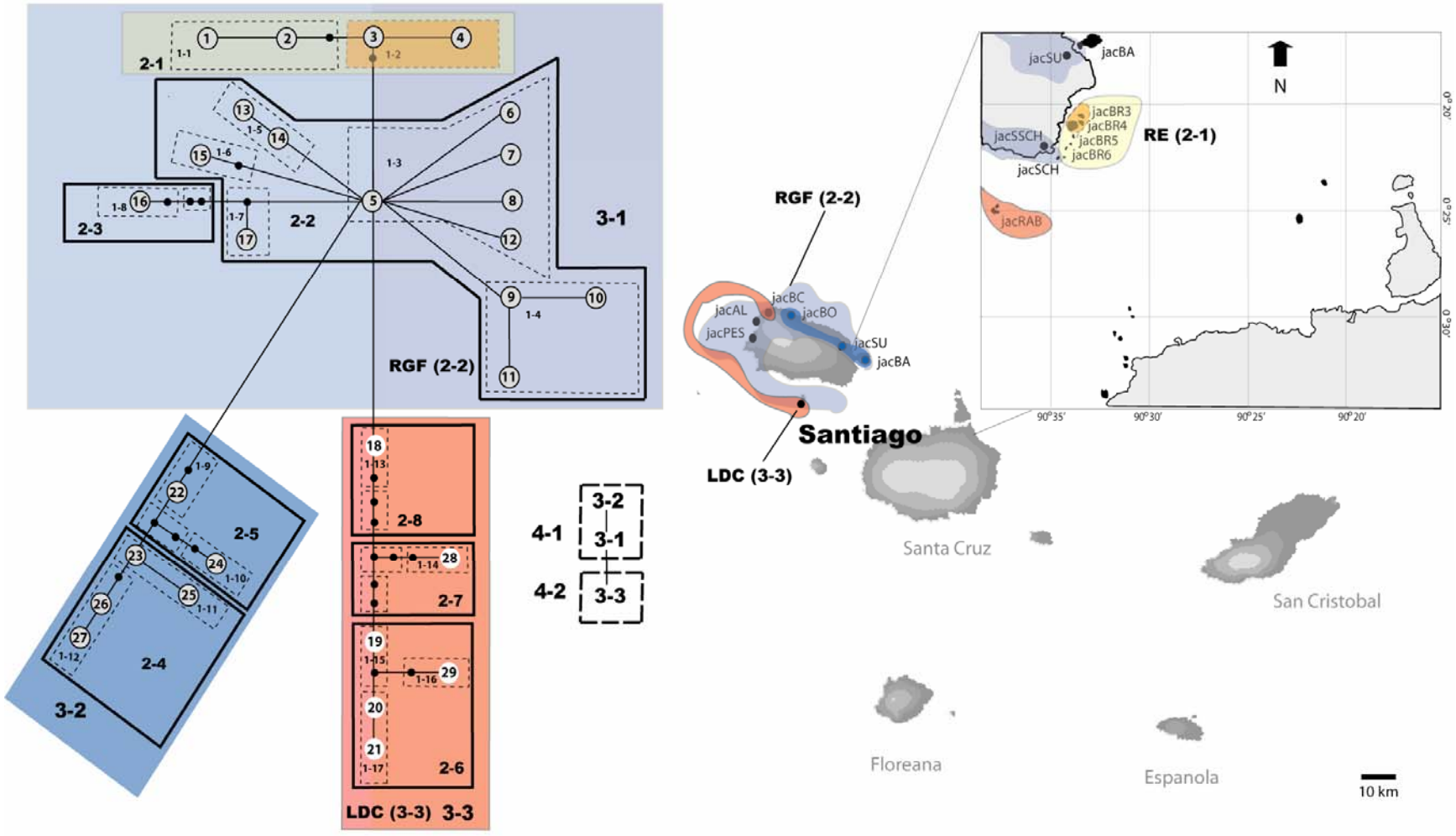

M. pacificus
M. jacobi
M. indefatigabilis
M. albemarlensis
M. duncanensis
M. delanonis
M. grayii
Wright, 1983
Heise, 1998
Kizirian et al., 2004

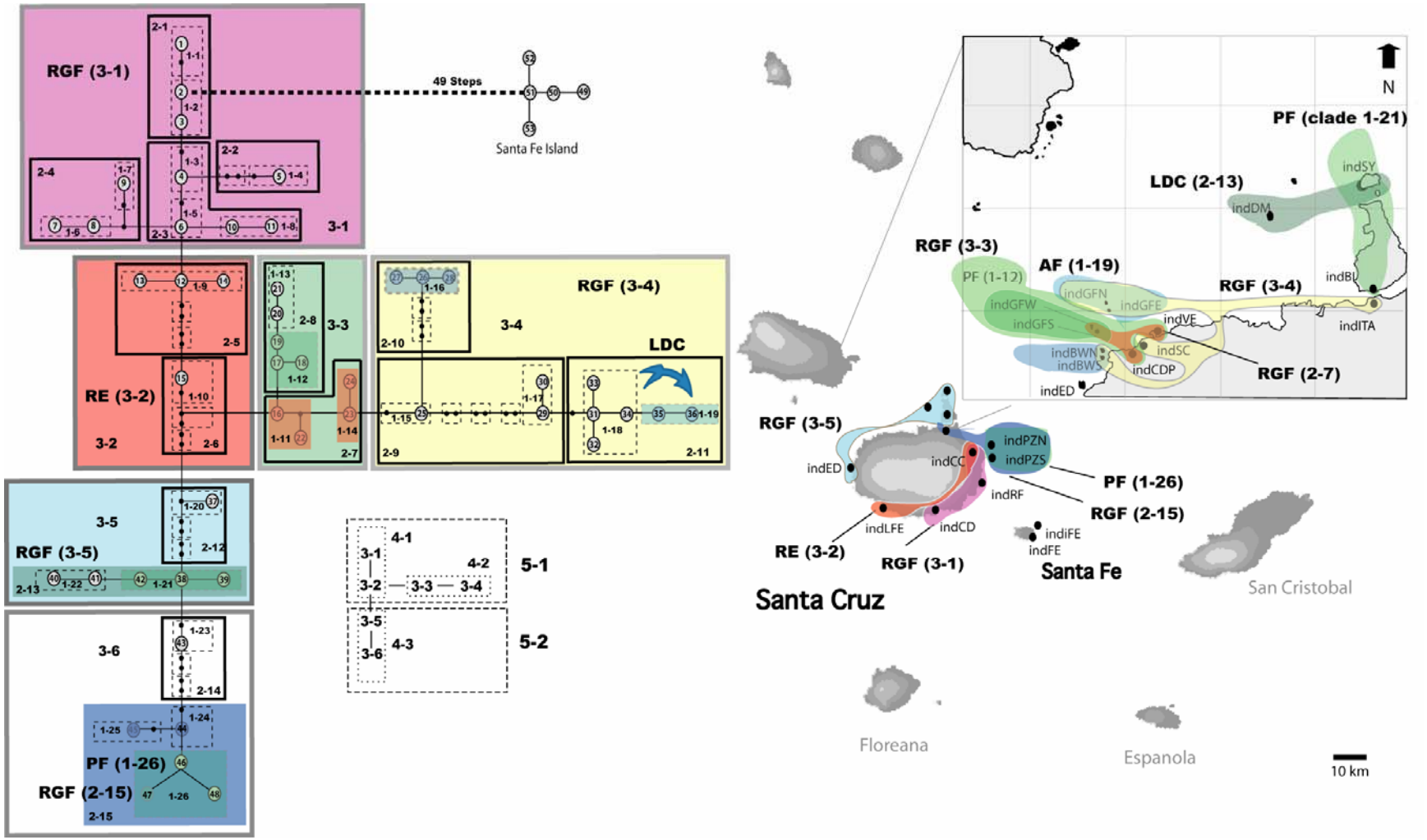
A

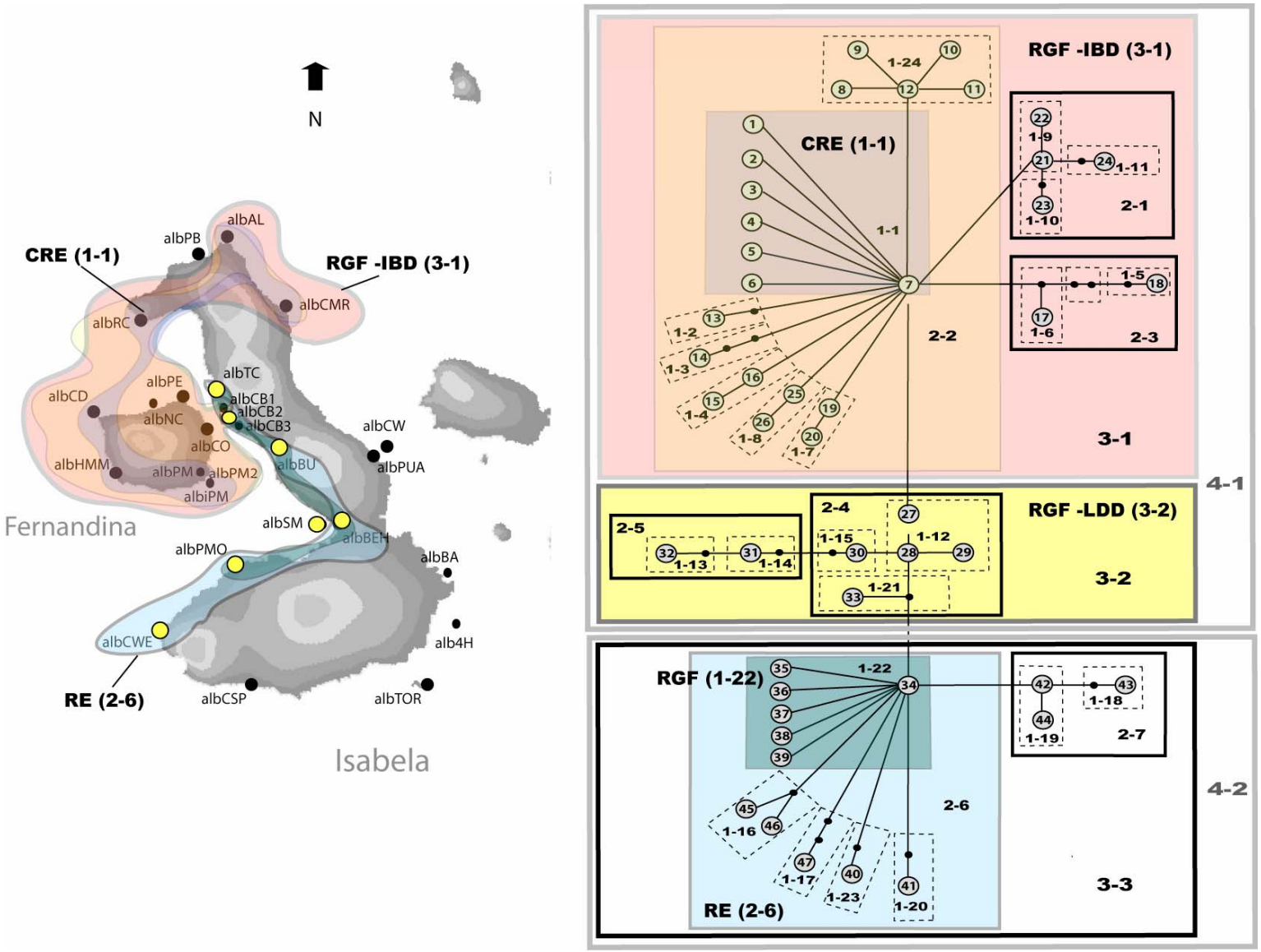

M.albemarlensis
M. pacificus
M. jacobi
M. duncanensis
M. indefatigabilis
M. grayii
M. delanonis
This study (Fig. 3)
M.albemarlensis
M. pacificus
M. indefatigabilis
M. jacobi
M. duncanensis
M. grayii
M. delanonis
Heise, 1998
M. duncanensis
M. jacobi
M. indefatigabilis
M.albemarlensis
M. pacificus
M.albemarlensis
M. grayii
M. delanonis
Kizirian et al., 2004
M.albemarlensis
M. indefatigabilis
M.albemarlensis
M. jacobi
M. pacificus
M. grayii
M. duncanensis
M. indefatigabilis
M. delanonis
Wright, 1983

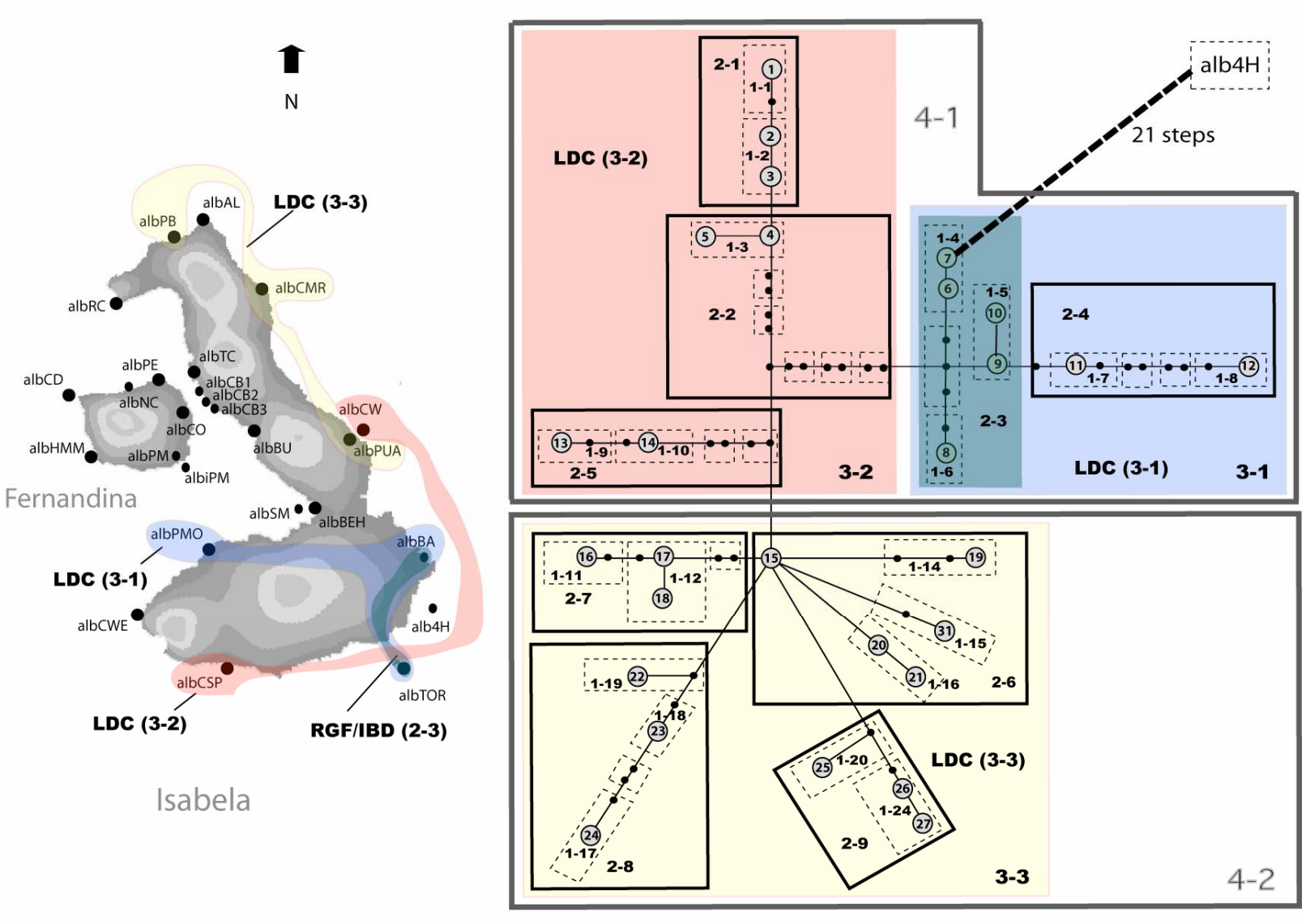
B

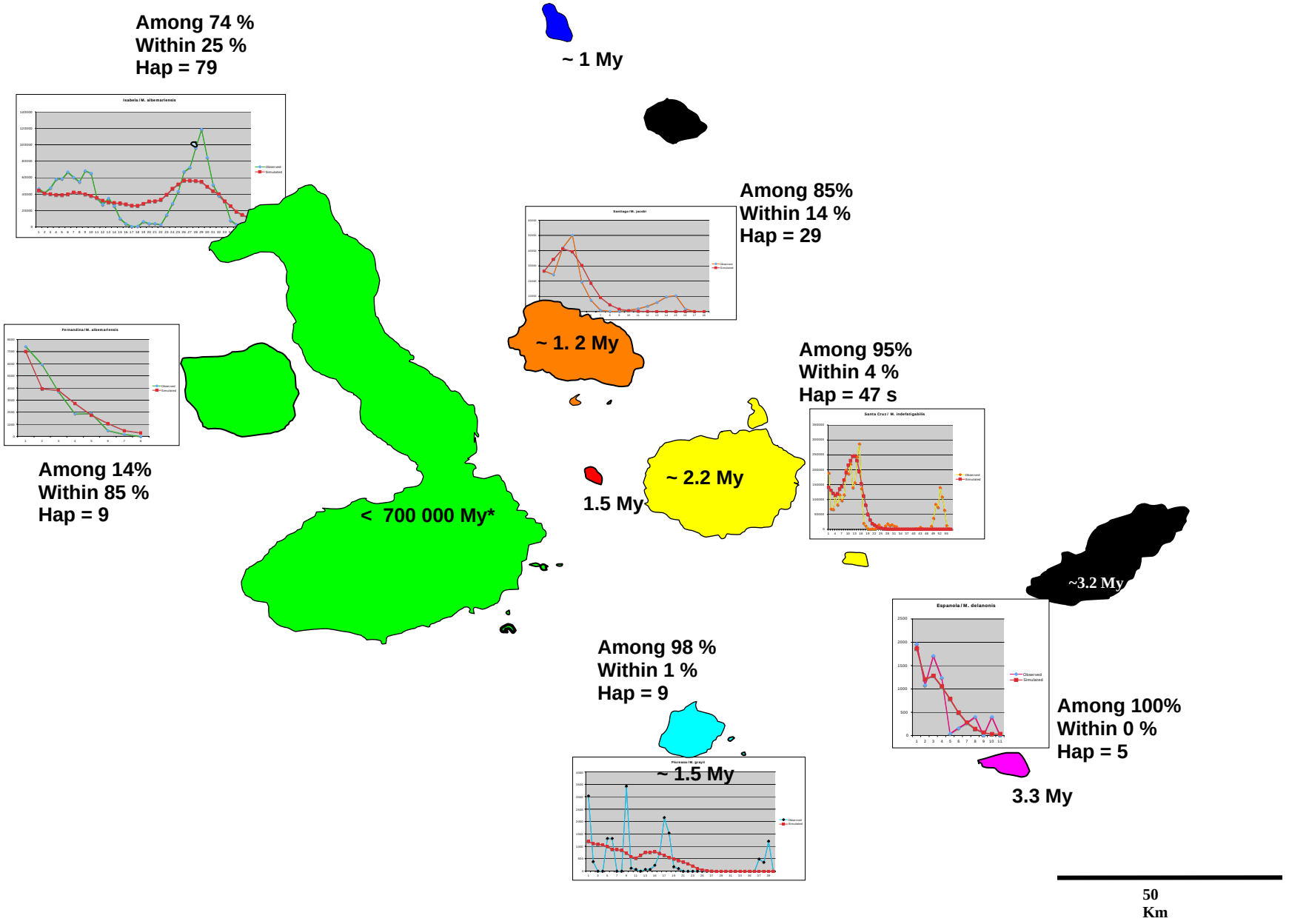

M. albemarlensis
Isabela
Western Clade
M. albemarlensis
Isabela + Fernandina
Eastern Clade
M. pacificus
Isla Pinta
M. duncanensis
Isla Pinzon
M. jacobi
Isla Santiago
M. jacobi
Islote Rabida
M. indefatigabilis
Santa Cruz
M. indefatigabilis
Islote Santa Fe
M. grayii
Isla Floreana
M. delanonis
Isla Espanola
M. bivittatus
M. koepckeorum
5 changes

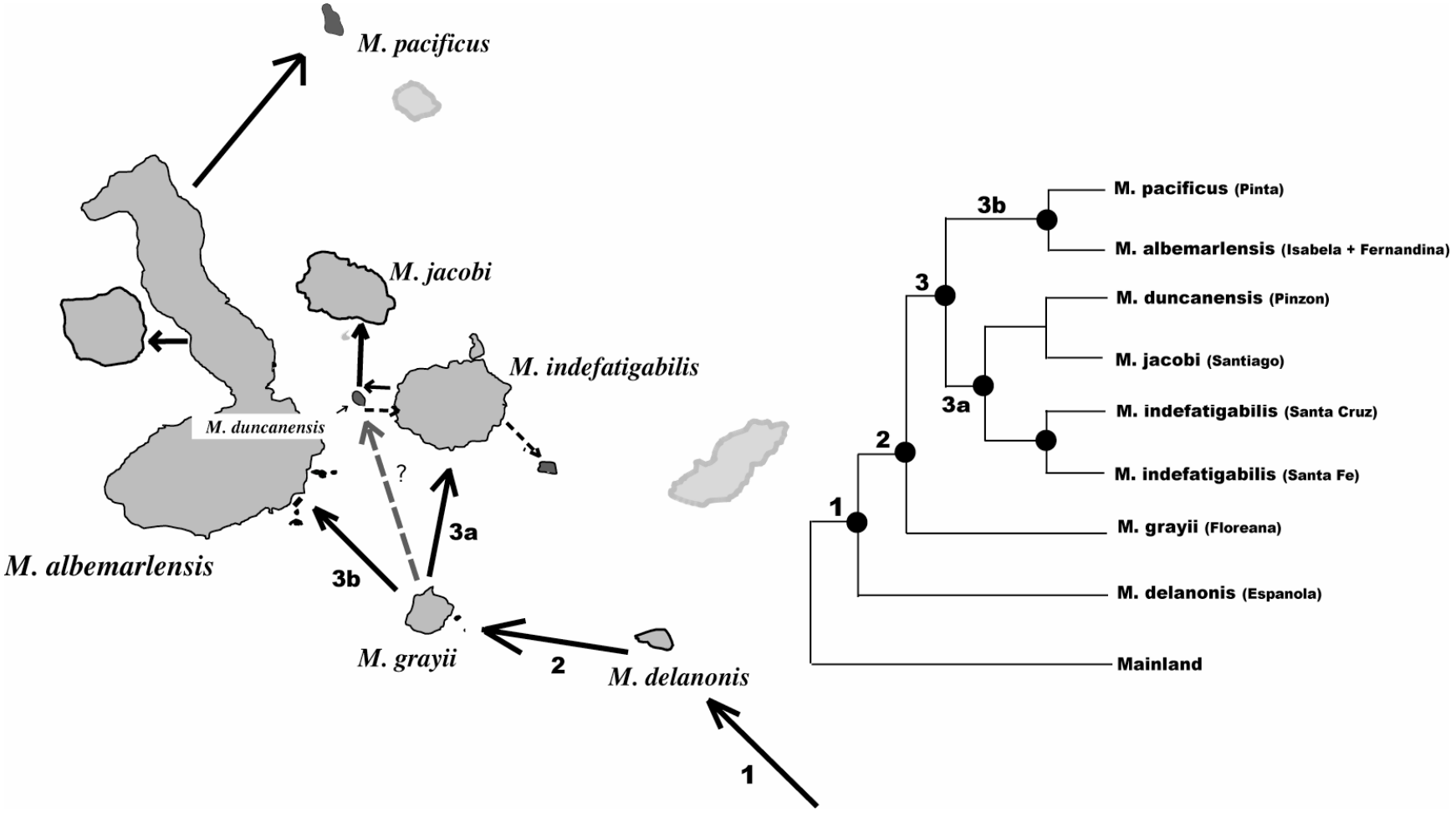

2-1
1-1
1-2
1-3
1-4
1-5
1-6
1-7
1-8
2-2
2-3
3-1
RGF (2-2)
1-9
1-10
1-11
1-12
2-5
2-4
3-2
1-13
1-14
1-15
1-16
1-17
2-8
2-7
2-6
LDC (3-3) 3-3
4-1
4-2
3-2
3-1
3-3
RGF (2-2)
jacBC
jacAL
jacBO
jacPES
jacSU
jacBA
LDC (3-3)
Santiago
Santa Cruz
San Cristobal
Floreana
Espanola
jacSU
jacBA
jacBR3
jacBR4
jacBR5
jacBR6
jacSSCH
jacSCH
jacRAB
RE (2-1)
N
90°35'
90°30'
90°25'
90°20'
0°20'
0°25'
0°30'
10 km


RGF (3-1)
49 Steps
Santa Fe Island
RE (3-2)
RGF (3-4)
LDC
RGF (3-5)
PF (1-26)
RGF (2-15)
4-1
4-2
5-1
5-2
N
PF (clade 1-21)
LDC (2-13)
indSY
indDM
RGF (3-3)
AF (1-19)
PF (1-12)
indGFN
indGFE
indGFW
indGFS
indVE
indBL
indITA
RGF (3-4)
indBWN
indBWS
indSC
indCDP
RGF (2-7)
indED
RGF (3-5)
indPZN
indPZS
indCC
PF (1-26)
indRF
RGF (2-15)
indLFE
indCD
indiFE
indFE
RE (3-2)
RGF (3-1)
Santa Fe
Santa Cruz
San Cristobal
Floreana
Espanola
10 km


N
CRE (1-1)
RGF -IBD (3-1)
RE (2-6)
albAL
albPB
albCMR
albRC
albTC
albPE
albCD
albCB1
albCB2
albNC
albCB3
albCW
albCO
albBU
albPUA
albHMM
albPM
albPM2
albiPM
albSM
albBEH
albPMO
albBA
albCWE
alb4H
albCSP
albTOR
Fernandina
Isabela
RGF -IBD (3-1)
1-24
CRE (1-1)
1-1
2-1
1-9
1-11
1-10
2-2
1-5
1-6
2-3
1-2
1-3
1-4
1-8
1-7
3-1
4-1
RGF -LDD (3-2)
2-4
2-5
1-15
1-12
1-13
1-14
1-21
3-2
RGF (1-22)
1-22
2-7
1-18
1-19
2-6
1-16
1-17
1-23
1-20
RE (2-6)
3-3
4-2


N
albAL
albPB
LDC (3-3)
albCMR
albRC
albTC
albPE
albCD
albCB1
albCB2
albNC
albCB3
albCO
albCW
albHMM
albBU
albPUA
albPM
albiPM
Fernandina
albSM
albBEH
albPMO
albBA
LDC (3-1)
albCWE
alb4H
albCSP
albTOR
LDC (3-2)
RGF/IBD (2-3)
Isabela
alb4H
21 steps
4-1
LDC (3-2)
2-1
1-1
1-2
1-3
2-2
2-5
1-9
1-10
3-2
1-4
1-5
2-4
1-7
1-8
2-3
1-6
LDC (3-1)
3-1
1-11
2-7
1-12
1-14
1-15
2-6
1-16
1-19
1-18
1-17
2-8
LDC (3-3)
1-20
1-24
2-9
3-3
4-2


Among 74 %
Within 25 %
Hap = 79
~ 1 My
Among 85%
Within 14 %
Hap = 29
~ 1. 2 My
Among 95%
Within 4 %
Hap = 47 s
~ 2.2 My
1.5 My
Among 14%
Within 85 %
Hap = 9
< 700 000 My*
~3.2 My
Among 98 %
Within 1 %
Hap = 9
~ 1.5 My
Among 100%
Within 0 %
Hap = 5
3.3 My
50
Km


M. pacificus
M. jacobi
M. indefatigabilis
M. duncanensis
?
M. albemarlensis
3a
3b
M. grayii
2
M. delanonis
1
3b
3
3a
2
1
M. pacificus (Pinta)
M. albemarlensis (Isabela + Fernandina)
M. duncanensis (Pinzon)
M. jacobi (Santiago)
M. indefatigabilis (Santa Cruz)
M. indefatigabilis (Santa Fe)
M. grayii (Floreana)
M. delanonis (Espanola)
Mainland

APPENDIX 1. Species and populations of all localities considered in this study (Western Galapagos radiation). Abbreviations identify localities plotted in Figure 1. Museum numbers correspond to the Museum of Southwestern Biology (MSB), University of New Mexico, and represent individual samples from each locality. Non-cataloged specimens are flagged by an M. N = number of individuals per population.

Species	Abbr.	Locality	N	Museum Numbers (MSB)	Geographical coordinates
INGROUP					
M. albemarlensis	1. alb_4H	Isla Isabela, Isla Cuatro Hermanos	1	M-1357	S 0° 52' 7.536" W 90° 46' 46.7041"
	2. alb_AL	Isla Isabela, Volcan Wolf, Punta Albemarle. Old radar base	10	MSB678156, MSB678161, MSB678151, MSB678155, MSB678153, MSB678136, MSB678159, MSB678458, MSB678146, MSB678150	S0° 8' 20.67880" W 91° 22' 40.188"
	3. alb_BA	Isla Isabela, Cerro Ballena	9	M-1358, M-1362, M-1363, M-1364, M-1365, M-1366, M-1367, M-1368, M-1369	S 0° 48' 52.272" W 90° 49' 42.384"
	4. alb_BEH	Isla Isabela, Bahia Elizabeth	12	MSB63297, MSB63281, MSB63291, MSB63305, MSB63285, MSB63304, MSB63293, MSB63288, MSB63287, MSB63295, MSB63547, MSB63548	S 0° 10' 22.5840" W 90° 50' 48.8401"
	5. alb_BU	Isla Isabela, Volcán Alcedo, Bahia Urvina, area along beach	7	MSB63399, MSB63405, MSB63409, MSB63415, MSB63402, MSB63578, MSB63579	S 0° 23' 37.392" W 91° 13' 49.908"
	6. alb_CB1	Isla Isabela, Islote Crater Beagle # 1	9	MSB63735, MSB63745, MSB63743, MSB63746, MSB63747, MSB63738, MSB63748, MSB63755, MSB63737	S 0° 17' 0.672" W 91° 21' 11.664"
	7. alb_CB2	Isla Isabela, Islote Crater Beagle # 2	8	MSB63767, MSB63760, MSB63763, MSB63764, MSB63765, MSB63759,	S 0° 16' 48.1439" W 91° 21' 1.26"

			MSB63766, MSB63758	
8. alb_CB3	Isla Isabela, Islote Crater Beagle # 3	10	MSB63528, MSB63525, MSB63529, MSB63530, MSB63526, MSB63761, MSB63776, MSB63771, MSB63773, MSB63772	S 0° 16' 59.7719" W 91° 20' 58.236”
9. alb_CD	Isla Fernandina, Cabo Douglas	10	MSB63782, MSB63783, MSB63780, MSB63784, MSB63779, MSB63787, MSB63786, MSB63789, MSB63793, MSB63794	S 0° 18' 21.852" W 91° 39' 3.0241"
10. alb_CMR	Isla Isabela; Volcán Wolf, Cabo Marshall, on shore W of point	17	MSB63853, MSB63844, MSB63852, MSB63847, MSB63843, MSB63851, MSB63848, MSB63829, MSB63835, MSB63830, MSB63864, MSB63860, MSB63861, MSB63857, MSB63866, MSB63862, MSB63846,	S 0° 23' 57.912" W 90° 59' 23.2439”
11. alb_CO	Isla Fernandina: middle of east coast at Campamento Copiano	7	MSB65132, MSB65138, MSB65128, MSB65120, MSB65127, MSB65117, MSB65137	S 0° 20' 32.496" W 91° 23' 10.968”
12. alb_CSP	Isla Isabela, Caleta San Pedro, SW of Cerro Azul	5	M-1347, M-1348, M-1349, M-1350, M-1351	S 1° 1' 54.696" W 91° 13' 21.072"
13. alb_CW	Isla Isabela, Islote Cowley	14	MSB64218, MSB64238, MSB64217, MSB64240, MSB64230, MSB64220, MSB64215, MSB64223, MSB64214, MSB64234, MSB63553, MSB63552, MSB64227, MSB64229	S 0° 23' 5.208" W 90° 57' 46.908"
14. alb_CWE	Isla Isabela, Caleta Webb, near Radar, West of Cerro Azul	4	M-1343, M-1344, M-1345, M-1346	S 0° 50' 52.9801" W 91° 29' 18.3841"
15. alb_HMM	Isla Fernandina, Cabo Hammond	5	M-1338, M-1339, M-1340, M-1341, M-1342	S 0° 28' 13.152" W 91° 36' 39.6719"

16. alb_IPM	Isla Fernandina, Islote Punta Mangle	5	M-1328, M-1329, M-1330, M-1331, M-1332	S 0° 27’ 13.9320" W 91° 23' 10.7520"
17. alb_NC	Isla Fernandina, North coast, midway between Cabo Douglas & Punta Espinosa	5	MSB65246, MSB65245, MSB65247, MSB65251, MSB65252,	S 0° 17' 4.488" W 91° 31' 15.6721"
18. alb_PB	Isla Isabela: Volcán Wolf, N coast at Piedra Blanca, at rock beach	10	MSB678262, MSB678263, MSB678257, MSB678267, MSB678266, MSB678270, MSB678261, MSB678260, MSB678255, MSB678272	S 0° 1' 9.696" W 91° 12' 56.7"
19. alb_PE	Isla Fernandina, Punta Espinosa, around mangroves behind visitor site dock	5	MSB65158, MSB65171, MSB65157, MSB65148, MSB65170	S 0° 17' 4.488" W 91° 31' 15.6721"
20. alb_PM	Isla Fernandina, Punta Mangle	11	MSB678186, MSB678189, MSB678187, MSB678185, MSB678181, MSB678174, MSB678184, MSB678193, MSB678192, MSB678172, MSB678191	S 0° 26' 19.644" W 91° 23' 28.284"
21. alb_PM2	Isla Fernandina, Punta Mangle (2)	5	M-1333, M-1334, M-1335, M-1336, M-1337	S 0° 26' 25.3680" W 91° 23' 23.9640"
22. alb_PMO	Isla Isabela, Punto Moreno	14	MSB678204, MSB678218, MSB678205, MSB678208, MSB678217, MSB678210, MSB678214, MSB678211, MSB678207, MSB678213, MSB678203, MSB678219, MSB678215, MSB678216	S 0° 26' 19.644" W 91° 23' 28.284"
23. alb_PUA	Isla Isabela, Punta Alfaro, Southeast of Volcan Alcedo	5	M-1414, M-1415, M-1416, M-1417, M-1418	No satellite signal
24. alb_RC	Isla Isabela: Cerro Ecuador, Punta Vicente Roca	18	MSB678231, MSB678224, MSB678233, MSB678228, MSB678223, MSB678237, MSB678232, MSB678226, MSB678240, MSB678229, MSB678230, MSB678234,	S 0° 3' 7.92" W 91° 33' 37.692"

				MSB678239, MSB67822, MSB678225, MSB678221, MSB678227, MSB69557,	
	25. alb_TC	Isla Isabela: Volcan Darwin, W slope, E of Tagus Cove	6	MSB69557, MSB69578, MSB69579, MSB69576, MSB63595, MSB63596	S 0° 14' 26.232" W 91° 21' 27.504"
	26. alb_TOR	Isla Isabela, Isla Tortugas	8	M-1352, M-1353, M-1354, M-1355, M-1356, M-1359, M-1360, M-1361	S 1° 0' 30.816" W 90° 52' 40.584"
	27. alb_SM	Isla Isabela, Islote South Marielas	6	MSB63604, MSB63605, MSB66023, MSB66080, MSB66113, MSB66114	S 0° 35' 55.4641" W 91° 5' 27.672"
M. delanonis	1. del_ES	Isla Española	5	MSB63549, MSB63550, MSB64809, MSB64827, MSB64849	S 1° 21' 23.364" W 89° 39' 8.9281"
	2. del_OS	Isla Española, Osborn Islet	12	MSB67071, MSB67078, MSB67087, MSB67085, MSB67082, MSB67095, MSB67093, MSB67088, MSB67077, MSB67076, MSB67096, MSB67094	S 1° 21' 23.364" W 89° 39' 8.9281"
	3. del_GA	Isla Española: Isla Gardner, coast above W cove (anchorage) S of cliff	5	MSB70152, MSB70193, MSB70195, MSB70206, MSB70209	S 0° 8' 8.3364" W 89° 39' 8.9281"
	4. del_IO	Española, Isote Oeste	10	MSB69773, MSB69785, MSB69783, MSB69784, MSB69772, MSB69597, MSB69600, MSB69601, MSB69599, MSB69594, MSB67088	S 1° 20' 47.256" W 89° 39' 44.1721"
	5. delXI	Isla Española, Xarifa Islet	9	MSB69729, MSB69701, MSB69730, MSB69700, MSB69699, MSB69726, MSB69705, MSB69731, MSB69724	S 1° 21' 26.28" W 89° 38' 39.444"
M. duncanensis	1. dun_es	Isla Pinzon, Playa Escondida	8	M-1419, M-1420, M-1421, M-1422, M-1423, M-1424,	S 1° 21' 7.308" W 89° 38' 56.508"

				M-1425, M-1426	
	2. dun_On	Isla Pinzon, Islote Onan, a 50 m de playa escondida frente a placa	6	M-1427, M-1428, M-1429, M-1430, M-1431, M-1432	S 1° 21' 7.308" W 89° 38' 56.508"
M. grayii	1. gra_CH	Isla Floreana, Islote Champion	8	MSB64202, MSB64198, MSB64190, MSB64192 MSB63615, MSB63616, MSB64183, MSB64194	S 1° 14' 6.036" W 90° 23' 5.28"
	2. gra_CLD	Isla Floreana: Islote Caldwell	5	M-1370, M-1371, M-1372, M-1373, M-1374	No satellite signal
	3. gra_CU	Isla Floreana: Las Cuevas, area around beach into canyon	8	MSB65111, MSB65112, MSB65113, MSB65110, MSB65114, MSB65108, MSB65115, MSB65109	S 1° 15' 40.752" W 90° 22' 6.920"
	4. gra_EN	Isla Floreana: Las Cuevas, area around beach into canyon	9	MSB64736, MSB64720, MSB64737, MSB64739, MSB64744, MSB63622, MSB63623, MSB64731, MSB64734	S 1° 13' 56.64" W 90° 21' 46.368"
	5. gra_PN	Isla Floreana: Playa Negra. 1km E of Puerto Velasco Ibarra	12	M-1375, M-1376, M-1377, M-1378, M-1379, M-1380, M-1381, M-1382, M-1383, M-1384, M-1385, M-1387	S 1° 16' 53.6519" W 90° 29' 34.1881"
	6. gra_F	Isla Floreana	10	MSB65331, MSB65324, MSB65329, MSB65339, MSB65323, MSB65316, MSB65318, MSB65332, MSB63610, MSB63611	S 1° 13' 53.364" W 90° 26' 37.536"
	7. gra_GA	Isla Floreana, Isla Gardner	2	MSB65685, MSB65689	S 1° 20' 0.888" W 90° 17' 56.796"
M. Indefatigabilis	1. ind_BL	Isla Baltra, at dock- Canal Itabaca	5	M-1404, M-1405, M-1406, M-1407, M-1408	S 0° 29' 15.1441" (5 of 5) W 90° 16' 47.5319"
	2. ind_BWN	Isla Santa Cruz, Islet Punta Bowditch Norte	14	MSB67325, MSB67327, MSB67287, MSB67345, MSB67332, MSB67282, MSB67276, MSB67509, MSB67503, MSB67513, MSB67510, MSB67518,	S 0° 32' 0.7799" W 90° 31' 3.1441" (14 of 14)

			MSB67516, MSB67512	
3. ind_BWS	Isla Santa Cruz, Islet Punta Bowditch Sur	14	MSB68927, MSB68936, MSB68939, MSB68941, MSB68937, MSB68950, MSB68938, MSB67245, MSB67253, MSB67255, MSB67257, MSB67250, MSB67241, MSB67242	S 0° 32' 9.5999" W 90° 31' 2.4601" (14 of 14)
4. ind_CC	Isla Santa Cruz; Cerro Colorado, 100 N of hill on coast before cliff	15	MSB70285, MSB70298, MSB70287, MSB70266, MSB70309, MSB64013, MSB64023, MSB64040, MSB64031, MSB64026, MSB64019, MSB64021, MSB64020, MSB64070, MSB64077	S 0° 34' 40.224" W 90° 10' 23.988" (11 out of 15)
5. ind_CD	Isla Santa Cruz: Puerto Ayora, Charles Darwin Research Station	17	MSB64173, MSB64109, MSB64175, MSB64150, MSB64149, MSB64096, MSB64168, MSB64155, MSB64114, MSB64139, MSB64105, MSB64144, MSB64157, MSB64100, MSB64177 MSB64145, MSB64165	S 0° 44' 33.0719" W 90° 18' 13.212" (12 of 17)
6. ind_CDP	Isla Santa Cruz, Cerro Dragon Path	2	MSB63535, MSB63536	S 0° 31' 21.0721" (2 of 2) W 90° 28' 54.228"
7. ind_DM	Isla Santa Cruz, Islet Daphne Major	7	MSB64295, MSB64383, MSB64427, MSB64256, MSB64423, MSB64271, MSB64257	S 0° 25' 20.388" W 90° 22' 21.684" (7 of 7)
8. ind_ED	Isla Santa Cruz, Islet Eden	12	MSB69626, MSB69638, MSB69637, MSB69611, MSB69635, MSB69625, MSB69645, MSB69632 MSB63627, MSB63628, MSB64564, MSB64573	S 0° 33' 42.876" W 90° 32' 13.3439" (12 of 12)

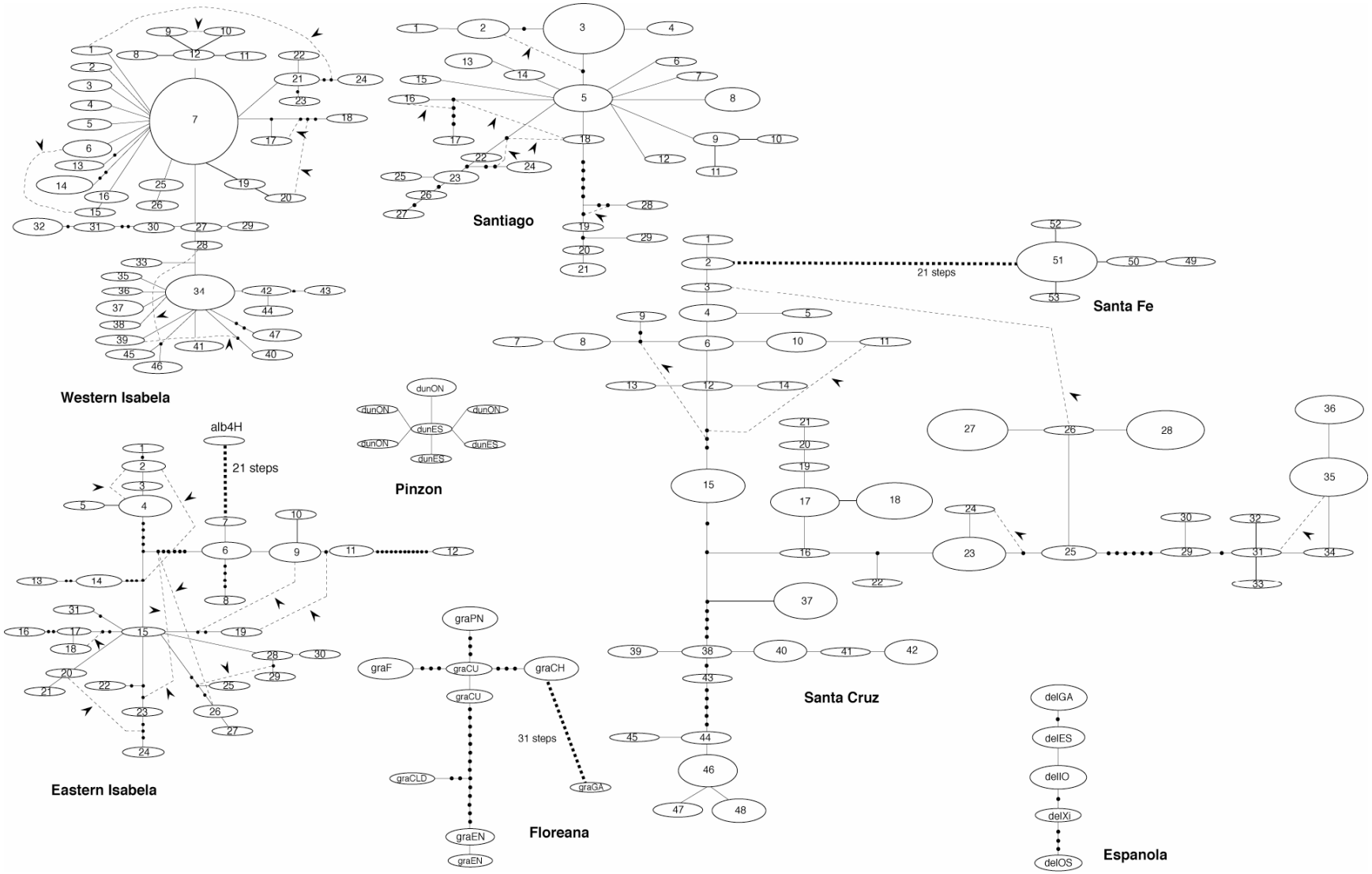
9. ind_FE	Isla Santa Fe: Isla Santa Fé, NE corner of island at tourist site, on beach	10	MSB69939, MSB69926, MSB69946, MSB69940, MSB69928, MSB69947, MSB69964, MSB69979, MSB69972, MSB69960	S 0° 48' 11.988" (10 of 10 W 90° 2' 28.176"
10. ind_GFE	Isla Santa Cruz, Southeast Guy Fawkes (Easternmost)	10	MSB65858, MSB65863, MSB658B69, MSB65861 MSB63645, MSB63646, MSB65841, MSB65842, MSB65844, MSB65845	S 0° 29' 57.1201" (9 of 10) W 90° 30' 49.14"
11. ind_GFN	Isla Sant Cruz, Northeast Guy Fawkes (Northernmost)	13	MSB69678, MSB696786, MSB696780, MSB69692, MSB69679, MSB696782, MSB69690, MSB63538, MSB63539, MSB65827, MSB65829, MSB65830, MSB65831	S 0° 29' 46.7161" (13 of 13) W 90° 30' 55.728"
12. ind_GFS	Isla Santa Cruz, Islet Guy Fawkes South	12	MSB69146, MSB69156, MSB69149, MSB69151, MSB69159, MSB69990, MSB69993, MSB69989, MSB69995, MSB69987, MSB69988, MSB69991	S 0° 30' 59.4" W 90° 31' 34.1401" (12 of 12)
13. ind_GFW	Isla Santa Cruz, Northeast Guy Fawkes (Westernmost)	13	MSB70007, MSB70002, MSB70006, MSB70004, MSB70005, MSB69998, MSB69997, MSB63642, MSB63643, MSB65925, MSB65926, MSB65927, MSB65928	S 0° 30' 52.092" W 90° 31' 43.6801" (13 of 13)
14. ind_iFE	Islote Santa FE, Isla Santa FE	10	MSB69809, MSB6790, MSB69826, MSB69798, MSB69824, MSB69805, MSB68511, MSB68509, MSB68522, MSB69804	S 0° 48' 11.988" (10 of 10) W 90° 2' 28.176"
15. ind_ita	Isla Santa Cruz, al frente de Baltra-Canal Itabaca	5	M-1409, M-1410, M-1411, M-1412, M-1413	S 0° 29' 15.1441" (4 of 5) W 90° 16' 47.5319"

	16. ind_Lfe	Isla Santa Cruz, Islote La Fe, East of Cerro Gallina	11	M-1433, M-1434, M-1435, M-1436, M-1437, M-1438, M-1439, M-1440, M-1441, M-1441, M-1442, M-1443	S 0° 45' 42.516" W 90° 24' 57.024"
	17. ind_PZN	Isla Santa Cruz, Islet Plaza Norte	12	MSB70350, MSB70347, MSB70315, MSB70342, MSB70355, MSB70357, MSB70346, MSB67625, MSB67662, MSB67637, MSB67663, MSB67647	S 0° 34' 48.864" (12 of 12) W 90° 9' 31.716"
	18. ind_PZS	Isla Santa Cruz, Isla Plaza Sur	10	MSB70932, MSB70910, MSB70891, MSB70931, MSB70961, MSB70912, MSB70965, MSB67666, MSB67693, MSB67679	S 0° 35' 0.1321" (9 of 10) W 90° 9' 55.188"
	19. ind_RF	Isla Santa Cruz: Punta Roca Fuerte, SE corner of island	10	MSB70406, MSB70435, MSB70376, MSB70421, MSB70419, MSB67827, MSB68346, MSB68342, MSB68334, MSB68351	S 0° 34' 40.224" (8 of 10) W 90° 10' 23.988"
	20. ind_SC	Isla Santa Cruz, on coast, 1 km of Islote Venecia	5	MSB68404, MSB68406, MSB68398, MSB68409, MSB68403	S 0° 31' 21.0721" (3 of 5) W 90° 28' 54.228"
	21. ind_SY	Isla Santa Cruz, Isla Seymour Norte Western bench	7	MSB68904, MSB68898, MSB68892, MSB68910, MSB68864, MSB68870, MSB68891	S 0° 23' 33.144" W 90° 16' 40.9439"
	22. ind_VE	Isla Santa Cruz, Venecia	14	MSB69514, MSB69505, MSB69520, MSB69506, MSB69522, MSB69499, MSB69510, MSB69501, MSB69515, MSB69525, MSB65985, MSB65986, MSB65988, MSB65989	S 0° 31' 21.0721" (11 of 14) W 90° 28' 54.228"
M. jacobi	1. jac_AL**	Isla Santiago, Isla Albany	7	MSB63184, MSB63205, MSB63210, MSB63195,	S 0° 10' 22.584" W 90° 50' 47.3641"

			MSB63206, MSB63190, MSB63201	
2. jac_BA	Isla Santiago, Isla Bartolome	9	MSB63476, MSB63511, MSB63512, MSB643678, MSB63471, MSB63648, MSB63649, MSB63472, MSB63475	S 0° 6' 43.704" W 90° 25' 12.9" (8 of 9)
3. jac_BC	Isla Santiago: Caleta Bucanero, N of peninsula, ESE Islote Mao	9	MSB69029, MSB69008, MSB69002, MSB69012, MSB69014, MSB69048, MSB69025, MSB63886, MSB63895	S 0° 9' 52.308" W 90° 49' 17.076"
4. jac_BO	Isla Santiago: la Bomba	5	M-1308, M-1309, M-1310, M-1311, M-1312	S 0° 6' 43.704" W 90° 25' 12.9"
5. jac_BR3	Isla Santiago, Bainbridge Rock #3	8	MSB63451, MSB63444, MSB63462, MSB63447, MSB63456, MSB63446, MSB63449, MSB63667	S 0° 20' 59.856" W 90° 33' 48.672" (7 of 8)
6. jac_BR4	Isla Santiago, Bainbridge Rock #4	8	MSB63312, MSB63324, MSB63313, MSB63315, MSB63316, MSB63319, MSB63311, MSB63666	S 0° 21' 30.996" W 90° 33' 51.948"
7. jac_BR5	Isla Santiago, Bainbridge Rock #5	8	MSB63345, MSB63330, MSB63333, MSB63344, MSB63336, MSB63343, MSB63337, MSB63660	S 0° 21' 50.76" W 90° 33' 59.796"
8. jac_BR6	Isla Santiago, Bainbridge Rock #6	11	MSB63387, MSB63389, MSB63366, MSB63357, MSB63355, MSB63382, MSB63385, MSB63657, MSB63351, MSB63373, MSB63384	S 0° 22' 2.892" W 90° 34' 13.188" (10 of 11)
9. jac_PES	Isla Santiago: Playa Espumilla	5	MSB71300, MSB71301, MSB71302, MSB71303, MSB71304	S 0° 7' 10.056" W 90° 29' 51.00" (4 of 5)
10. jac_RAB	Isla Rabida, sitio turistico	6	MSB71294, MSB71295, MSB71296, MSB71297,	S 0° 23' 59.1" W 90° 42' 24.192"

				MSB71298, MSB71299	
	11. jac_SCH	Isla Sombrero Chino	9	MSB69231, MSB69241, MSB69233, MSB69258, MSB69248, MSB69256, MSB69254, MSB63672, MSB63673	S 0° 22' 9.696" W 90° 35' 10.1401"
	12. jac_SU	Isla Santiago, Bahia Sullivan	2	MSB63652, MSB63653	S 0° 17' 14.748" W 90° 33' 57.708"
	13. jac_SSCH	Isla Santiago/on older lava across from Sombrero Chino	12	MSB63676, MSB63677 MSB88689, MSB68995 MSB68992, MSB8994 MSB68983, MSB68977 MSB68985, MSB68982 MSB68980, MSB68979	S 0° 22' 9.696" W 90° 35' 10.1401" (11 of 12)
M. pacificus	1. pac_PP	Isla Pinta, en playa posada, costa Sur w/plaque	5	M-1324, M-1325, M-1326, M-1327	S 0° 36' 42.4799" W 90° 47' 11.8319"
OUTGROUPS					
M. bivittattus	biv	Isla San Cristobal	1	MSB 63708	S 0° 51' 35.0999" W 89° 33' 56.952"
M. koepckeorum	koep	Pariakoto, PERU,	1	M-729	S 09° 33' 43.5" W 77° 53' 21.7"

Appendix 2. TCS networks for Cyt *b* haplotypes sampled from 78 localities representing the western Galapagos radiation. Networks were constructed using the statistical parsimony algorithm of Templeton et al. (1992), under a 95% limit of 14 steps. In all networks the size of each oval is proportional to the frequency of each haplotype, and solid dots represent unsampled haplotypes. Dashed lines and arrows pointing at them indicate discarded loops (see text for details).


Western Isabela
Santiago
21 steps
Santa Fe
Pinzon
dunON
dunES
alb4H
21 steps
Eastern Isabela
graPN
graF
graCU
graCH
graCLD
graEN
31 steps
graGA
Floreana
Santa Cruz
delGA
delES
delIO
delXi
delOS
Espanola