Short communication

A phylogeographically distinct and deep divergence in the widespread Neotropical turnip-tailed gecko, *Thecadactylus rapicauda*

D.J.C. Kronauer\textsuperscript{a,*}, P.J. Bergmann\textsuperscript{b,1}, J.M. Mercer\textsuperscript{a}, A.P. Russell\textsuperscript{b}

\textsuperscript{a} Department of Biology, Duke University, Science Drive, Durham, NC 27708, USA
\textsuperscript{b} Department of Biological Sciences, University of Calgary, 2500 University Dr. N.W., Calgary, Alberta T2N 1N4, Canada

Received 7 May 2004; revised 12 October 2004

1. Introduction

Although the Neotropics are recognized as a region of high biodiversity, mainland Neotropical gekkotans are a notable exception to this rule. There are comparatively few species, representing a number of disparate genera (Abdala and Moro, 1996) with independent evolutionary origins, distributions, and histories (Vanzolini, 1968). This is surprising given the high worldwide gekkonid diversity of over 600 species (Kluge, 2001). The turnip-tailed gecko, *Thecadactylus rapicauda* (Houttuyn, 1782), demonstrates this low diversity as the only nocturnal, arboreal lizard over much of its range (Vitt and Zani, 1997).

Considering its ostensible monotypy, *T. rapicauda* has one of the most extensive ranges of any lizard (Mattison, 1989), occurring from Yucatán to Bolivia, throughout the Amazon Basin, east to the Guyanan region, and on the Lesser Antilles (Fig. 1; Russell and Bauer, 2002). Based on morphology, Burt and Myers (1942) postulated that the taxon might be split into multiple entities in the future and Mijares-Urrutia and Arends (2000) suggested that the karyotypes of specimens from different populations differ.

Whether *T. rapicauda* represents a single cohesive species or multiple distinct species is of considerable interest. Although a phylogenetic reconstruction may not in itself provide sufficient information for deciding species status (Harrison, 1991), nucleotide substitution data showing evidence of substantial genetic divergence between populations can suggest hypotheses of species-level differentiation, and may indicate when and where barriers to gene flow existed in the past. Due to its extensive range, analysis of this taxon may also shed additional light on more general biogeographical patterns in the Neotropics. We employ Bayesian inference (BI), maximum likelihood (ML), and maximum parsimony (MP) analyses of sequences from the mitochondrial cytochrome b (cyt b) gene to address these issues.

2. Materials and methods

2.1. Sampling and molecular protocols

Thirty specimens of *Thecadactylus* from 18 localities distributed across the range, with a sample size of 1–4 for each locality (see Fig. 1 and Table 1), were examined. One specimen of *Blaesodactylus boivini* was sequenced as the outgroup. Tissues were stored in 95% ethanol and DNA was extracted with either phenol-chloroform extraction (Werman et al., 1996) or the Qiagen DNA extraction kit. A 584 base pair (bp) fragment of the mitochondrial *cyt b* gene was amplified via standard PCR using a customized primer pair \[3p: 5'-GGAAT(AG)GCCAC(AC)GCATTCTGT-3'; \ TRA3r: 5'-GGGTCTTCTAC(CT)GG(CT)TG(AG)CC-3'\] at an optimal annealing temperature of 53°C. This fragment corresponds to nucleotide positions 14,543–15,126 of the mitochondrial genome of the scincid lizard *Eumeces*.
egregius (Kumazawa and Nishida, 1999). Sequencing reactions were done using a BigDye Terminator Cycle Sequencing kit (Applied Biosystems) and run on a PRISM ABI 373A sequencer.

2.2. Data analysis

Alignment was checked visually as well as with translation to amino acid sequence using the software ProSeq 2.9 (Filatov, 2002). Pairwise $p$-distances were calculated using PAUP* 4.0b10 (Swofford, 1998) and converted to percent sequence divergences. Subsequent to phylogenetic analyses, within-clade distances were compared to between-clade distances for identified clades. ML distances were calculated for saturation analyses using parameter estimates from Modeltest v. 3.06 (Posada and Crandall, 1998) for the general time reversible model with gamma distributed sites and a proportion of invariant sites (GTR + G + I).

BI phylogenetic analysis was executed in MrBayes v. 3.0b4 (Huelsenbeck and Ronquist, 2001) under the GTR + G + I model as selected using hLRT in ModelTest v. 3.06 (Posada and Crandall, 1998). The Bayesian analysis was done with default uninformative priors using four Metropolis-Coupled Monte Carlo Markov Chains, a temperature of 0.2, and four replicate runs starting from a random initial tree. Trees and parameters were sampled every 100 generations for 1,000,000 generations. Tree and parameter summaries were made from samples after an initial 100,000 generation burn-in.

MP analysis with 1000 bootstrap replicates was implemented in PAUP* 4.0b10 (Swofford, 1998) using the tree bisection-reconnection heuristic search algorithm, with random taxon addition sequence, 10,000 replicates, and the MULTREES option deactivated. Character states were optimized using accelerated transformation (ACCTRAN). Empirically determined weightings from saturation analysis were applied in MP analysis, differentially weighting codon positions and applying the ti:tv ratio following Tan and Wake (1995).

Tests of phylogeographic hypotheses were done by ML parametric bootstrapping (SOWH test; Goldman et al., 2000; Huelsenbeck et al., 1996). To speed computation of the null distributions, the original dataset of 33 OTUs (ingroup 30; outgroup 3) was reduced to a total of 15 including one outgroup (marked by asterisks in

Fig. 1. The distribution of *T. rapicula*, with localities of origin of tissue samples plotted. Samples belonging to the southwestern clade are underlined.
Table 1 and Fig. 2). The taxa that were eliminated were very closely related (as evidenced by genetic distance) and geographically close with respect to the larger phylogeographic hypotheses being tested, while a representative sampling of localities was retained. Constraint trees for phylogeographic hypotheses were constructed in MacClade 4.05 (Maddison and Maddison, 2000). Null distributions for the log likelihood ratio test statistic were estimated from 100 replicate datasets simulated using Seq-Gen 1.2.7 (Rambaut and Grassly, 1997) with the SG Runner graphical interface 1.5.2 (Wilcox, 2004). GTR + G + I was used as the model to estimate the best ML tree consistent with each constraint tree. Two iterations of estimating and 2xing parameters followed by heuristic searches were found to be sufficient for accurate estimation of both the tree and the model parameters both with the original dataset and the datasets generated by simulation. Most ancillary file manipulations were done with standard UNIX file utilities.

Table 1

<table>
<thead>
<tr>
<th>ID No.</th>
<th>Locality</th>
<th>GenBank Accession No.</th>
<th>Collection No./collector</th>
</tr>
</thead>
<tbody>
<tr>
<td>7080/81</td>
<td>Rattan, St. Croix</td>
<td>AY604507; AY604508</td>
<td>USNM 561446; 561447</td>
</tr>
<tr>
<td>7091/93</td>
<td>Maria Major Island, St. Lucia</td>
<td>AY604505; AY604506</td>
<td>PJB/APR; USNM 561448</td>
</tr>
<tr>
<td>7097/100/101</td>
<td>Levera, Grenada</td>
<td>AY604510; AY604512; AY604511</td>
<td>PJB/APR</td>
</tr>
<tr>
<td>7105/08/09/11</td>
<td>CEIBA Biological Center, Demerara, Guyana</td>
<td>AY604496; AY604494; AY604497; AY604495</td>
<td>USNM 561449; 561450; 561451; 561452</td>
</tr>
<tr>
<td>Trin1/5</td>
<td>Trinidad</td>
<td>AY604498; AY604499</td>
<td>H. Kaiser</td>
</tr>
<tr>
<td>Sel1</td>
<td>Speyside, Tobago</td>
<td>AY604500</td>
<td>S. Rose</td>
</tr>
<tr>
<td>Sel4</td>
<td>Blanchisseuse, Trinidad</td>
<td>AY604501</td>
<td>S. Rose</td>
</tr>
<tr>
<td>Tra1</td>
<td>Felipe Carrillo Puerto, Yucatán, Mexico</td>
<td>AY604492</td>
<td>USNM 561442</td>
</tr>
<tr>
<td>Tra2/3</td>
<td>Tikal, Guatemala</td>
<td>AY604489; AY604490</td>
<td>USNM 561443; 561444</td>
</tr>
<tr>
<td>Tra4</td>
<td>San Jacinto, Loreto, Perú</td>
<td>AY604486</td>
<td>UKNHM KU222144</td>
</tr>
<tr>
<td>Tra5</td>
<td>Near Puerto Maldonado, Madre de Dios, Perú</td>
<td>AY604488</td>
<td>UKNHM KU214928</td>
</tr>
<tr>
<td>Tra7</td>
<td>Drake’s Bay, Osa Peninsula, Costa Rica</td>
<td>AY604503</td>
<td>USNM 561445</td>
</tr>
<tr>
<td>Tra8</td>
<td>Morales, Sierra de Caral, Izabal, Guatemala</td>
<td>AY604491</td>
<td>UTACV R-39897</td>
</tr>
<tr>
<td>Tra9</td>
<td>Gringo Perdido, Petén, Guatemala</td>
<td>AY604493</td>
<td>UTACV R-50343</td>
</tr>
<tr>
<td>Tra11</td>
<td>Long Bay Water Tank, Anguilla</td>
<td>AY604504</td>
<td>R. Powell 6764</td>
</tr>
<tr>
<td>Tra12</td>
<td>Fazenda Nova Esperanca, Roraima, Brazil</td>
<td>AY604502</td>
<td>LSUMZ H-12314</td>
</tr>
<tr>
<td>Tra13</td>
<td>Reserva Faunistica Cuyabeno, Sucumbios, Ecuador</td>
<td>AY604483</td>
<td>LSUMZ H-12565</td>
</tr>
<tr>
<td>Tra14</td>
<td>Near Porto Walker, Acre, Brazil</td>
<td>AY604487</td>
<td>LSUMZ H-13562</td>
</tr>
<tr>
<td>Tra15</td>
<td>Near Santarem, Pará, Brazil</td>
<td>AY604509</td>
<td>LSUMZ H-14379</td>
</tr>
<tr>
<td>Tra16</td>
<td>Amazonas, Brazil</td>
<td>AY604485</td>
<td>LSUMZ H-16450</td>
</tr>
<tr>
<td>Tra17</td>
<td>Parque Estadual Guajara-Mirim, Rondônia, Brazil</td>
<td>AY604484</td>
<td>LSUMZ H-17797</td>
</tr>
<tr>
<td>B. boivini</td>
<td>Madagascar</td>
<td>AY604513</td>
<td>PJB</td>
</tr>
</tbody>
</table>

Note: For all specimens, their identifying number in this study (ID No.), locality of origin, as well as GenBank accession number are given. Accession numbers marked with an asterisk are sequences used in SOWH hypothesis tests. For specimens deposited in museums or tissue collections, the collection number is given, and for those not deposited, the collector name is given (Smithsonian National Museum of Natural History (USNM), Museum of Natural Science, Louisiana State University (LSUMZ), Collection of Vertebrates, University of Texas at Arlington (UTACV), University of Kansas Natural History Museum (UKNHM)).

Blasodactylus boivini was defined as the outgroup in BI (as analyses in MrBayes are limited to a single outgroup) and ML analyses, and two sequences from GenBank that overlapped our alignment by the first 400 bp (AY217802: Gehyra mutilata and AY217801: Hemi- dactylus frenatus) were used as additional outgroups in MP analysis. Due to sequencing problems with B. boivini, its sequence lacked 65 bp of the final alignment (45 and 20 bp at the 5’ and 3’ end, respectively), and this was partly compensated for by the inclusion of the additional outgroups.

3. Results

3.1. Phylogenetic analyses

A pronounced guanine deficiency and the absence of premature stop codons, frameshifts, or indels are supportive of the idea that mtDNA has been sequenced rather than nuclear pseudogenes.

All four replicates of the BI analysis resulted in the same summary tree, which was topologically identical to the ML bootstrap tree. Bootstrap support was highly correlated with, but somewhat less than corresponding posterior probabilities, which is expected for this gener-
ally more conservative assessment of support. Nevertheless, high posterior probabilities with at least moderate ML bootstrap support can be a good indicator of phylogenetic accuracy (Alfaro et al., 2003). For MP analysis, the weighting scheme for codon positions was determined empirically (Tan and Wake, 1995) as 1:3.5:9.5 (3rd:1st:2nd). A weighting scheme calculated for just the ingroup samples would differ minimally from this at 1:3.6:10.5. Transitions are weighted at one-third of transversions.

All three analyses recover two well-supported major clades, one consisting of Amazonian Ecuador, Peruvian Loreto and Madre de Dios, and Brazilian Acre, Amazonas, and Rondônia (posterior probability: 1, ML and MP bootstrap: 100%) and the other consisting of all remaining OTUs (posterior probability: 0.99, ML bootstrap: 68%, MP bootstrap: 95%) (Fig. 2). Sequence divergences between the two clades (23.0–26.9%, which is comparable to divergences between Thecadactylus OTUs and the outgroup (24.12–30.3%)) are almost always higher than those within a clade (≤13% within the southwestern clade and ≤23.7% within the clade representing the remainder of T. rapicauda). Generally, relationships between OTUs obtained from Bayesian or ML analyses on the one hand, and MP analysis on the other, are congruent where nodal support is high, but they disagree in several cases of poor nodal support on the exact topology, especially within the non-southwestern clade.

3.2. Phylogeographic hypotheses

For testing phylogeographic hypotheses for the sampling represented in our study we divided the non-southwestern clade into four geographical regions: Central America, northern Brazil plus Guyana, Trinidad, and the Lesser Antilles. Trinidad is recognized separately from the Lesser Antilles because it is geologically distinct in its origin and in its continental-shelf connection to the mainland, but we also evaluate its association both with the mainland and with the other Caribbean islands. We tested the following hypotheses of monophyly: (a) Caribbean islands, (b) Lesser Antilles, (c) Central America, (d) northern Brazil plus Guyana, (e) northern Brazil plus Guyana plus Trinidad, and (f) the non-southwest clade excluding Antilles (i.e. “e” plus “c” above).

For a nominal value of $P = 0.05$ with six tests, the most restrictive critical value with sequential Bonferroni correction (Rice, 1989) is approximately 0.01. Values of the test statistics allowed rejection of monophyly with $P << 0.01$ for each of the following four sets of samples: (a) Caribbean islands, (c) Central America, and (d, e) northern Brazil plus Guyana, with or without Trinidad. Thus, consistent with the wide separation of Guatemalan and Costa Rican samples on the phylogenetic trees, Central American samples do not form the cohesive group that geographic proximity might otherwise suggest, and Trinidad is not closely linked with the Lesser Antillean islands.

Monophyly could not be rejected, however, for samples from the Lesser Antilles ($P = 0.30$), or the non-southwest clade excluding Antilles ($P = 1.0$), which are mutually complementary groups. These results suggest that targeting portions of the species distribution within the mainland region stretching from Yucatán across northern South America to Pará for additional sampling may clarify the phylogeographic substructure of this area in the future. Trinidad associates more closely with samples from the mainland than with other Caribbean islands, and although the islands of the Lesser Antilles do not group consistently together on phylogenetic trees, the hypothesis remains viable that the Antillean samples are monophyletic.

4. Discussion

A number of phylogeographic patterns emerge from the phylogeny of T. rapicauda, of which most obvious is
its major division into a southwestern Amazonian clade and a clade containing all remaining OTUs (Figs. 1 and 2). An examination of sequence divergences within and between these two clades further supports their distinctness.

A similar southwestern area has been identified in phylogeographic studies of other taxa, including lizards (Avila-Pires, 1995), primates (da Silva and Oren, 1996), and anurans (Ron, 2000). However, patterns of geographic diversification in Amazonia can differ even between similar taxa (e.g., Anolis lizards; Glor et al., 2001).

While some groups within the two major clades are well supported, others are not unequivocally resolved, and many subclades are not geographically coherent. Striking examples of geographic incoherence are the well supported clade of samples from Yucatán, Guatemala, Guyana, Trinidad, and Tobago, or the samples from Amazonian Ecuador and Rondônia that are sister OTUs, but are geographically quite distant, with more phylogenetically distinctly related OTUs intervening spatially. The grouping of Trinidad and Tobago samples with samples from the mainland, rather than other Caribbean islands, is consistent with their separation from Venezuela by only a narrow, shallow channel that was probably exposed during the last glacial maximum. Proximity of Trinidad and Tobago to the Orinoco River delta may also have enhanced opportunities for dispersal from the mainland.

Our results clearly show that neither mainland South American samples as a whole, nor Middle American, nor Caribbean (including Trinidad and Tobago) OTUs form monophyletic groups (Fig. 2; see also the results of SOWH tests described in Section 3.2). While documentation of gene flow would require a larger sampling of individuals, a general discordance between haplotype clades and the well-delimited geographical area in which they occur is potential evidence for gene flow (Wiens and Penkrot, 2002) and may indicate repeated faunal exchange between these geographical areas. By contrast, the southwestern clade comprises a well-supported ancient and genetically distinct entity that is concordant with geography, which could suggest the absence of recurrent genetic exchange with other clades.

Estimated evolutionary rates of cyt b in squamate reptiles typically range from 0.5 to 1.4% per million years (Gübitz et al., 2000; Malhotra and Thorpe, 2000; Zamudio and Greene, 1997). Given these rates, even a conservative estimate would suggest that the two major Thecadactylus lineages have been separated at least since the early Miocene, more than 16 million years ago, predating postulated Pleistocene refugia and the Pliocene (Moritz et al., 2000). A similar late Tertiary origin has been estimated for lowland rodent and marsupial species, which are typically much older than their Andean counterparts (reviewed in Moritz et al., 2000).

Several mechanisms that may have promoted vertebrate differentiation in Amazonia have been proposed (Hauffer, 1997), some of which might apply to Thecadactylus. The geographical boundary between the two major clades may follow parts of the course of the Amazon River (the sample from Pará, which was collected just south of the Amazon and close to the estuary, is an exception). But there are many species of vertebrates that do not show an obvious genetic divergence across riverine barriers (Patton et al., 2000). Moreover, it is not evident why a river system would be an effective isolation barrier for such a widely dispersed species. Instead, the deep split may have alternative explanations. Sedimentary, fossil, and palynomorph evidence from the South American Solimões/Pebas formation points to an early Miocene to late middle Miocene interval in which the Andean foreland basins were dominated by marine incursions and extensive fluvi–lucustrine environments (Lundberg et al., 1998). The incursions were perhaps the result of high eustatic sea levels in the Burdigalian, Langhian, and Serravallian (Haq et al., 1987), as well as some tectonic subsidence. Particularly from the late middle Miocene to the early late Miocene, 11.8–10.0 Mya, the Andean foreland basins were again dominated by marine transgressions, with the Pebsian seaway from the Caribbean in the north reaching as far south as Acre (at approximately 10° S), and the Paraná seaway reaching from the south Atlantic (to approximately 17° S). There is no evidence for an epicontinental seaway connecting the Caribbean to the South Atlantic, but the evidence is such that much of the Andean foreland basin was filled at times with extensive marine, brackish, and freshwater environments. By 11 Mya, with the onset of a new bout of tectonism ultimately giving rise to the Colombian Eastern Cordillera and Mérida mountains, the Caribbean marine connection was cut off and the paleo-Orinoco and paleo-Amazonas established their modern drainage systems into the Atlantic; the Pebsian sea, if it still existed, was drained (Lundberg et al., 1998). Our early to middle Miocene divergence therefore pre-dates the Amazonas system and points to a possible role for extensive water barriers in the foreland basin systems as a source of vicariance between the southwestern and non-southwestern clade.

Notwithstanding the difficulties in using genetic divergence as an indicator of species (Harrison, 1991), the long persistence of a well supported and geographically concordant southwestern clade in apparent isolation is suggestive of its distinctness at the specific level (Lessios et al., 1999; Morando et al., 2003; Wiens and Penkrot, 2002). The degrees of sequence divergence between the two clades of T. rapicauda accord completely with reported between-species divergences for cyt b in other reptiles. For example, divergences between the viperid snake Lachesis muta and species of the related genus Agkistrodon (13.8–17.0%; Zamudio and Greene, 1997),
between species of the gekkonid genera Phelsuma (8.9–28.9%; Radtkey, 1996), Pachydactylus (15–43%; Lamb and Bauer, 2002), and Tarentola (mean 23.6±4.2% SE; Nogales et al., 1998), as well as between species of the lacertid genus Gallotia (mean 15.7±2.4% SE; González et al., 1996). Accordingly, the evidence presented herein refutes the long unquestioned taxonomic unity of T. rapicauda and is suggestive of the southwestern clade representing a cryptic species. The case of Amazonian Thecadactylus may be comparable to that of Anolis nitens, another widespread and genetically, as well as morphologically, highly differentiated Amazonian lizard species of apparently similar age (Glor et al., 2001). Examination of the morphological variation of Thecadactylus is required prior to making any formal taxonomic decisions.

Acknowledgments

We are grateful to the Museum of Natural Science, Louisiana State University, for tissues, and to the following for collecting and contributing samples to this study: J.P. Caldwell, J.A. Campbell (UTAVC), W.E. Duellman (UKNHM), B. Hedges, H. Kaiser, G. Köhler, A. Miranda, R. Powell, S. Rose, and L.J. Vitt (supported by NSF). We are also indebted to those people who facilitating obtaining permits. Research permit #020515BR035 and export permit #020716SP035 were issued by the Environmental Protection Agency, Guyana. An unnumbered collection and export permit for St. Lucia was issued by B. James, the Chief Forestry Officer, Ministry of Agriculture, Forestry and Fisheries. Collection and export permits were also issued for specimens collected on St. Croix by W. Coles and M. Evans (USFW). Dr. G. Bourne (Guyana) assisted greatly in obtaining necessary permits and facilitating work in Guyana, and A. Robart (University of Calgary) gave advice about various molecular protocols. Sampling procedures were approved in protocol BI 2002-003 by the University of Calgary Animal Care Committee. V.L. Roth deserves our appreciation for her helpful comments on various drafts of this manuscript. Partial financial support for this work was provided by NSF DEB 9726855 to V.L. Roth and J.M.M. and by NSERC and the University of Calgary to P.J.B. and A.P.R. The comments of three anonymous reviewers have helped to significantly improve the manuscript.

References


