Major Caribbean and Central American frog faunas originated by ancient oceanic dispersal

Matthew P. Heinicke*, William E. Duellman†, and S. Blair Hedges*‡

*Department of Biology, Pennsylvania State University, University Park, PA 16802-5301; and †Natural History Museum and Biodiversity Research Center, University of Kansas, Lawrence, KS 66045

Edited by David B. Wake, University of California, Berkeley, CA, and approved May 5, 2007 (received for review December 12, 2006)

Approximately one-half of all species of amphibians occur in the New World tropics, which includes South America, Middle America, and the West Indies. Of those, 27% (801 species) belong to a large assemblage, the eleutherodactyline frogs, which breed out of water and lay eggs that undergo direct development on land. Their wide distribution and mode of reproduction offer potential for resolving questions in evolution, ecology, and conservation. However, progress in all of these fields has been hindered by a poor understanding of their evolutionary relationships. As a result, most of the species have been placed in a single genus, Eleutherodactylus, which is the largest among vertebrates. Our DNA sequence analysis of a major fraction of eleutherodactyline diversity revealed three large radiations of species with unexpected geographic isolation: a South American Clade (393 sp.), a Caribbean Clade (171 sp.), and a Middle American Clade (111 sp.). Molecular clock analyses reject the prevailing hypothesis that these frogs arose from land connections with North and South America and their subsequent fragmentation in the Late Cretaceous (80–70 Mya). Origin by dispersal, probably over water from South America in the early Cenozoic (47–29 million years ago, Mya), is more likely.

Terrestrial breeding and direct development have allowed eleutherodactyline frogs to occupy a diversity of ecological niches and have facilitated their wide distribution (Fig. 1). Eleutherodactylines occur on almost every island in the Caribbean and display near total endemicity to single-island banks. Their elevational range also is broad, with some species occurring up to 4,400 m in the Andes of South America. Thus, they are a model group for studying Neotropical biogeography and evolution. With this in mind, we assembled samples and available sequences of 276 species of eleutherodactylines and Brachycephalus for several mitochondrial and nuclear genes. Our goal was to identify the major groups of species and their times of divergence, to better understand the historical biogeography of eleutherodactyline frogs and the region in general. Our results revealed several major and, for the most part, geographically isolated, clades of eleutherodactyline frogs and showed that the Middle American and West Indian eleutherodactylines owe their origin to Cenozoic over-water dispersal, not from land connections in the Mesozoic.

Fig. 1. Composite distribution of eleutherodactyline frogs and Brachycephalus (812 sp.). “Middle America” refers to Central America and Mexico. No evolutionary groupings are implied.
using past species–group affiliations, it was possible to assign species not included in this study to these major genetically defined clades (SI Table 2). The first major group, which we call the Caribbean Clade (*Eleutherodactylus*), consists of the West Indian members of the subgenus *Eleutherodactylus* (47 sp.), the subgenus *Pelorus* of Hispaniola (6 sp.), the West Indian subge-
nus *Euhya* (91 sp.), and the subgenus *Syrrhopus* (26 sp.) of southern North America, Middle America, and Cuba.

A second large group (111 sp.) of eleutherodactyline frogs occurs in Middle America, and already has been recognized as the subgenus or genus *Craugastor* (2, 11, 12, 15). Our analyses indicate a slightly different composition of this Middle American Clade. Previous definitions included some primarily South American species (16), which we find to form a separate clade that is, instead, most closely related to other South American eleutherodactyline (see below). The single remaining South American endemic, the distinctive *C. biforcurus*, warrants further study with DNA sequences to verify its placement in *Craugastor* (17).

The third and largest group defined in our analyses includes nearly 400 species centered in the Andes but with species also occurring elsewhere in northern South America. A few species in this group extend into Central America, including nine endemic to southern portions of that region (see SI). Also, two species occur in the southernmost islands of the Lesser Antilles. This South American Clade includes species formerly placed in the *Eleutherodactylus unistrigatus*, *conspicillatus*, and 13 other species groups (7). We use the available name *Pristimantis* Jiménez de la Espada, 1870 for this previously undefined clade.

Besides these three major clades, our analyses suggest that most of the 31 species in southeastern Brazil formerly placed in *Eleutherodactylus* form a separate, smaller clade (Figs. 2–3). Our sparse taxonomic sampling from this region makes it difficult to determine the composition of this group, but the joining of four diverse species (*E. guentheri*, *E. hoehnei*, *E. parvus*, and *E. juipoda*) in a well supported group, suggests that other species from the region believed to be closely related to them also are part of that group, which takes the available name *Ischnocnema* Reinhardt and Lürken, 1862 (see SI Text). Four southeast Brazilian species in our analysis that are not part of that clade are *E. binoatatus*, which has an unusual karyotype (18), *Holochadon bradei*, *Barycholos ternetzi*, and *Brachycephalus ephippium*. These species also branch basally among eleutherodactyline but are not closely related to other species or groups.

These major clades of species with definitive geographic patterns account for 87% of the 812 species of eleutherodactyline frogs and *Brachycephalus*. The remaining 106 species are all native to South America, mostly Andean, and are best characterized by their basal position in the phylogenetic trees (Figs. 2–3), suggesting that they represent an early stage of evolution of the group. Their relationships and those of the three major clades remain unresolved. Among these are the representatives of the (formerly *Craugastor*) *anomalus* and *bufoniformis* groups, which cluster strongly with a species in the *E. sulcatus* group. For this clade, we apply the available name *Linnophyllum* Jiménez de la Espada, 1871. The genus *Phrynopus* is polyphyletic, with species forming several independent groups, as was found elsewhere (19). Two species of *Phrynopus* cluster with species of the *E. nigrovittatus* and *E. dolops* groups. Other genera in this category of deeply branching lineages include *Oreobates* and *Phyllonastes*.

### Times of Divergence

Dates of divergence obtained by using nuclear data, mitochondrial data, or all data are similar for most nodes (Table 1 and SI Table 3). The eleutherodactyline lineage diverged from other hyloid frogs near the Mesozoic–Cenozoic boundary (57 Mya, C.I. = 78–44 Mya), as found elsewhere (20), with initial diversifications occurring among eleutherodactylines ~50 Mya (Fig. 3). The Caribbean Clade (*Eleutherodactylus*) diverged from its extant mainland relatives ~47 Mya and began diversification ~29 Mya, setting upper and lower bounds for the date that the West Indies was colonized. Assuming no extinction of the mainland source lineage, the dispersal most likely occurred early in that time interval rather than later. Similarly, the Middle American Clade (*Craugastor*) diverged 42 Mya and began diversification 31 Mya. Middle American and Cuban *Syrrhopus* split ~19 Mya. The Southeast Brazil Clade diverged from other eleutherodactylines ~50 Mya. The South American Clade (*Pristimantis*) diverged from other eleutherodactylines ~37 Mya and began an explosive diversification ~24 Mya.

### Discussion

**Major Clades of Tropical Frogs**. The discovery of three major and geographically defined groups of these tropical amphibians was unexpected. Previous studies on eleutherodactyline had been hampered by too few useful morphological characters and too

### Notes

Fig. 3. A time tree of eleutherodactyline frogs. The tree topology is derived from a ML analysis of 61 eleutherodactyline, *Brachycephalus*, and three out-group species. Support values for groups mentioned in the text are indicated at nodes (ML/ME/Bayesian posterior probability). Calibration nodes are indicated by open circle (minimum constraint), filled circles (maximum constraint), or filled square (minimum and maximum constraints). The two proposed oceanic dispersal events are on the branches leading to the Caribbean Clade (CC) and the Middle American Clade (MAC). [The South American Clade (SAC) and Southeast Brazil Clade (SBC) are indicated.] Times and credibility intervals for numbered nodes are shown in Table 1. Geologic epochs are abbreviated as follows: Paleocene (Pa), Eocene (E), Oligocene (O), Miocene (M), Pliocene (P), Pleistocene (Pl), Holocene (H).
few samples for molecular analysis. Although the Middle American Clade was known (12, 15), it had included species in South America shown here to be misclassified based on our sequence analyses. The Caribbean and South American clades, on the other hand, were unpredicted. Previous studies had assumed a close relationship between West Indian members of the subgenus *Eleutherodactylus* and the species-rich *unistrigatus* group (now in *Pristimantis*) in South America (7, 10, 11, 21). In part, this was based on shared morphological characters that may be associated with climbing habits (11). Our results show, however, that diverse morphologies and habits have evolved independently in the Caribbean and South American Clades. The geographical separation of these large clades highlights a general pattern, the greater importance of geography, revealed in many molecular phylogenetic studies (e.g., refs. 5 and 22).

**Middle America and the Caribbean.** The origin of the Middle American and West Indian terrestrial vertebrates has focused on two competing models in the context of current geologic models for the region (23–25). The vicariance model suggests that they arose in the Late Cretaceous (80–70 Mya) by fragmentation of a continuous land mass (proto-Antilles) and its biota located between North and South America (26–28). This occurred as the Caribbean tectonic plate moved eastward, carrying the West Indian fauna and isolating the Middle American fauna from its South American counterparts. This is in contrast to an origin of these faunas by dispersal, on flotsam from continental source areas. One difficulty for the vicariance model has been the great age (Cretaceous) of the groups required for this model, which is largely unsupported by the fossil record (29). Also, the fauna of the West Indies is peculiar in missing many higher-level groups, indicative of dispersal (30). Geologic evidence does not rule out the possibility of a proto-Antillean island chain or corridor, but does not favor the substantial emergence of land in the Antilles before the mid-Eocene (37–49 Mya) (24).

Molecular clock analyses have yielded mixed results, although most groups have shown Cenozoic divergences with their closest relatives on the mainland (23–32). Estimates of Cretaceous divergence between West Indian and mainland representatives of insectivores (33), xantusiid lizards (33, 34), and (in past studies) eleutherodactyline frogs (10, 35), suggested that those groups may be vicariant relics of the proto-Antilles even if most others are not. However, the relictual nature of the distribution of xantusiid lizards and West Indian insectivores raises the possibility of Cenozoic dispersal to the West Indies and subsequent extinction of those mainland source populations (34). Studies indicating Cretaceous ages for Middle American and West Indian eleutherodactylines either assumed proto-Antillean vicariance (12) or used geologic calibrations that have since been revised (10).

Based on our results, the Middle American and Caribbean clades of eleutherodactylines originated through dispersal from South America during the Cenozoic. For these clades to have originated through proto-Antillean vicariance, Mesozoic ages (e.g., 80–70 Mya) are required for divergences between these groups and their South American relatives. Instead, our data (Table 1) indicate that a single event 42–31 Mya established eleutherodactylines in Middle America, and another 47–29 Mya established eleutherodactylines in the West Indies (Fig. 4a). Early speciation in the Caribbean Clade was confined to Hispaniola and Cuba. The paleogeography of the West Indies in the

### Table 1. Times of divergence (Mya) for major nodes in Fig. 3

<table>
<thead>
<tr>
<th>Node</th>
<th>Divergence</th>
<th>Time</th>
<th>95% C.I.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Eleutherodactylines plus <em>Brachycephalus/lyrid frogs</em></td>
<td>56.79</td>
<td>(43.52, 78.13)</td>
</tr>
<tr>
<td>2</td>
<td>Southeast Brazil Clade (SBC)/other species</td>
<td>49.79</td>
<td>(37.18, 68.67)</td>
</tr>
<tr>
<td>3</td>
<td>Caribbean Clade (CC)/other eleutherodactylines</td>
<td>47.28</td>
<td>(35.09, 65.26)</td>
</tr>
<tr>
<td>4</td>
<td>Middle American Clade (MAC)/other eleutherodactylines</td>
<td>42.39</td>
<td>(30.99, 53.99)</td>
</tr>
<tr>
<td>5</td>
<td>South American Clade (SAC)/other eleutherodactylines</td>
<td>36.52</td>
<td>(26.56, 50.81)</td>
</tr>
<tr>
<td>6</td>
<td>Last common ancestor of Caribbean Clade</td>
<td>29.09</td>
<td>(20.95, 40.35)</td>
</tr>
<tr>
<td>7</td>
<td>Last common ancestor of Middle American Clade</td>
<td>30.51</td>
<td>(21.67, 43.17)</td>
</tr>
<tr>
<td>8</td>
<td>Last common ancestor of South American Clade</td>
<td>24.45</td>
<td>(17.30, 34.82)</td>
</tr>
<tr>
<td>9</td>
<td>Middle American <em>Syrrophus/Cuban Syrrophus</em></td>
<td>19.05</td>
<td>(13.06, 26.92)</td>
</tr>
</tbody>
</table>

*Times are based on the combined nuclear and mitochondrial data set (3,709 bp) and measure the divergence of the two identified lineages separated by a slash. Divergence times for all nodes (combined, nuclear, and mitochondrial data sets) are presented in SI Table 3 with a guide tree available as SI Fig. 14. The dates of the two proposed dispersal events are constrained by nodes 3 and 6 (for the Caribbean Clade) and nodes 4 and 7 (for the Middle American Clade). *Bayesian credibility interval.*
mid-Cenozoic was substantially different from that today (24). Land connections between Cuba, northern Hispaniola, and Puerto Rico probably existed in the Late Eocene (~35 Mya), facilitating dispersal among the islands. A proposed dry-land connection to South America at this time (24) lacks geologic support and remains controversial (23, 34). After subsidence in the Oligocene (23–34 Mya), land connections were broken, probably isolating the western Caribbean lineage (subgenera Euthyus plus Syrrhopus) in Cuba from the eastern Caribbean lineage (subgenera Eleutherodactylus plus Pelorius) in northern Hispaniola and Puerto Rico (Fig. 4b).

In the Early Miocene (19 Mya), an over-water dispersal occurred from western Cuba to southern North America within the subgenus Syrrhopus (Fig. 4c), as indicated by some earlier molecular studies (10, 11). It is possible that this lineage initially evolved in isolation to the north of the Middle American Clade, although the distributions of these two groups currently overlap. Dispersal from the Greater Antilles to the mainland has been found in other vertebrate groups, including turtles (36, 37) and anole lizards (38). Other Miocene dispersals of eleutherodactylines, most probably over water, occurred among islands in the West Indies (Fig. 3). The direction of some of these dispersal events would have been against the present-day water currents, which flow primarily from southeast to northwest. However, current flow within the Caribbean may have been different in the past, before the emergence of the Isthmus of Panama (39).

A striking pattern in these results is the absence of subsequent successful colonizations of eleutherodactyline frogs in Middle America and the West Indies from South America after their origin in the early Cenozoic. Of the few exceptions, two species of the South American clade now occupy the southernmost Lesser Antilles (St. Vincent and Grenada) and 18 species of the South American Clade now occur in Middle America (see SI Text). In the latter case, the presence of some or all of these species may be explained by dispersal over land after the emergence of the Isthmus of Panama (~3 Mya). Whether or not there were failed colonizations to Middle America and the West Indies as a result of competition (40) is unknown. Also, if the Middle American Clade and Caribbean Clade are later found to be closest relatives, the possibility that there was a stepwise dispersal (South America to one clade and from that clade to the other) should be considered.

South America. Most of the basal branches of eleutherodactylines, with some dating to the early Cenozoic, occur in South America (Fig. 3). This indicates that South America was the place of origin for the group, as it was for hyloid frogs in general (13, 14). However, the great diversity of species, including the South American Clade of 393 species, is associated with Andes. The Andean uplift is relatively recent, occurring mostly in the last 10–20 million years (41, 42). Rapid diversification within the South American Clade, which began 24 Mya and has continued to the present, was probably linked with this uplift. Mountain-building and associated climatic changes resulted in repeating patterns of habitat isolation, which, in turn, probably resulted in genetic isolation and speciation in these amphibians (7).

Despite the large number of South American species included in this analysis (123 spp.), we are missing a majority of species including many from southeastern Brazil. Our results indicate that the eleutherodactyline fauna of southeastern Brazil is distinct and includes several basal clades. This region is an isolated area of montane rainforest and is a region of endemism for other amphibians (43).

Methods

Taxon Sampling. Our data set encompasses ~34% of known eleutherodactyline diversity with 276 species in 12 of 18 genera, including at least one representative of every genus with more than five described species. Included species were concentrated in the largest genera, with 140 species of Eleutherodactylus, 87 of Pristimantis, 14 of Craugastor, 17 of Phrynopus, and four of the Southeast Brazil Clade. Two hylid species, Agalychnis callidryas (South America) and Litoria caerulea (Australia), were included for calibration of divergence times. Seven additional hyloid species and a more distant ranoid species (Rana catesbeiana) were included as out groups.

Data Collection. Our study included data from three mitochondrial genes: 12S ribosomal RNA (12S), 16S ribosomal RNA (16S), and intervening tRNA-Valine. In addition, fragments from two nuclear protein-coding genes were sequenced: recombination-activating gene 1 (Rag-1) and the tyrosinase gene (Tyr). Approximately 90% of the sequences used are previously uncharacterized. Data were collected as overlapping sets (SI Table 4) of 280 species (two genes), 146 species (three genes), and 65 species (five genes).

For the 280-species data set, partial 12S and 16S sequences were assembled for 277 in-group and three out-group species and used to define major clades (here recognized as genera and subgenera). This data set consists of an ~350-bp fragment of 12S concatenated with a ~800-bp fragment of 16S. For the 146-species data set, complete 12S and 16S sequences (~2.5 kb), including the intervening tRNA and fragments of the flanking tRNA sequences, were assembled for 136 species representing all major groups as defined by the partial data set, the same three out-group species, and seven additional hyloid out-group species. This data set was used to test groups found with the 280-species data set, confirm rooting within eleutherodactylines by using additional out groups, and define subgroups within the largest clades. For the 65-species data set, we also included sequences from a 493-bp region of TYr and a 639-bp region of Rag-1. This sample included representatives of most major clades and subclades, except where specimen availability or quality were limiting. Methods of sample collection, DNA extraction, amplification, and sequencing are presented in the SI Text, along with a list of primers (SI Table 5). When available, sequences for species of interest were obtained from GenBank (SI Table 4).

Phylogenetic Analyses. Reconstructions of phylogenies for all data sets were performed by using ME, ML, and Bayesian methods. For ML and Bayesian analyses, the 65-species data set was divided into three partitions: 12S and 16S, Rag-1, and Tyr. ME analyses were implemented in MEGA 3.1 (44) by using the TN + I + G model of evolution. PAUP 4b10 (45) was used to estimate the γ-parameter, and branch support was assessed with 2,000 bootstrap replicates. ML analyses were used RaxML-VI-HPC v.2.0 (46), accessed at the San Diego Supercomputing Center. For each data set, 100 alternative runs were performed under the GTRMIX model of evolution. Other parameters were maintained at default settings. Nonparametric bootstrap analysis (1,000 replicates) was used to provide branch support values for the most likely tree of 100 found in each data set. MrBayes 3.1 (47) was used to perform Bayesian analyses. Bayesian analyses used the GTR + I + G model of evolution, with all parameters unlinked in partitioned analyses. For the 65-species data set, all phylogenetic analyses were performed by using only the two nuclear genes in addition to analyses employing both the mitochondrial nuclear data, to ensure that mitochondrial and nuclear data produced results that were not significantly divergent. Additional details of analyses are available in the SI.

Divergence Timing. Times of divergence were estimated for the 65-species data set by using the T3 version of Multidivtime (48, 49). The assumed topology was from the five-gene ML analysis. The data were divided into three partitions, as in the phylogenetic analyses. In addition to estimating times by using all
available data, timing analyses were also performed by using mitochondrial and nuclear data separately. A total of five calibrations, including both upper and lower bounds within and outside the eleutherodactylines, were used based on geologic and fossil evidence. These included the earliest divergence in the Jamaican radiation of the subgenus Eleutherodactylus (10 Mya, maximum constraint), the earliest divergence in the Hispaniolan South Island radiation of Eleutherodactylus (20 Mya, maximum), the divergence of the subgenera Eleutherodactylus and Pelorius (15 Mya, minimum), and the divergence of South American and Australian hyliids (35 Mya, minimum; 70 Mya, maximum). Additional details for these and other priors are available in SI Text.

S.B.H. especially thanks Lori Fritz, Jennifer R. Grubb, Genna M. Lutz, Molly E. Means, Rebecca Montrose, Michael R. Tracy, and Cecelia Youngblood for their contributions to the initial generation of data for this project. We thank Jonathan A. Campbell and Eric N. Smith Youngblood for their contributions to the initial generation of data for this project. We thank Jonathan A. Campbell and Eric N. Smith Youngblood for their contributions to the initial generation of data for this project. We thank Jonathan A. Campbell and Eric N. Smith Youngblood for their contributions to the initial generation of data for this project. We thank Jonathan A. Campbell and Eric N. Smith Youngblood for their contributions to the initial generation of data for this project. We thank Jonathan A. Campbell and Eric N. Smith Youngblood for their contributions to the initial generation of data for this project.