This article was published in an Elsevier journal. The attached copy is furnished to the author for non-commercial research and education use, including for instruction at the author's institution, sharing with colleagues and providing to institution administration.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
1. Introduction

Most terrestrial vertebrates of the West Indies, especially those that cannot fly, have a high rate of endemism, including 99% of the 174 native species of amphibians and 96% of the 499 native species of reptiles (Hedges, 2001, 2006). Furthermore, they are typically found on no more than one island. However, several of the larger species of geckos are more widely distributed and represent exceptions to this rule. Here, we investigate these species using DNA sequences to gain insights into their origin, taxonomic status, and historical biogeography.

1.1. Phyllodactylus wirshingi

This species was described from specimens taken in Puerto Rico (Kerster and Smith, 1955) and additional specimens were later collected from Hispaniola, in the southern Dominican Republic and in northwestern Haiti (Schwartz, 1979) (Fig. 1a). The habitats in all three regions are xeric and have exposed rocks (Schwartz and Henderson, 1991). The only other species of Phyllodactylus known from the Greater or Lesser Antilles is P. pulcher from Barbados, a rare species. One of us (S.B.H.) has searched for P. pulcher on Barbados, unsuccessfully, although a more comprehensive effort is needed to determine its conservation status. Phyllodactylus wirshingi is considered to be a member of the P. tuberculatus group, although the relationships of the species of Phyllodactylus have not yet been fully resolved (Bauer et al., 1997; Dixon and Huey, 1970).

1.2. Tarentola americana

The gecko genus Tarentola is distributed mostly in North Africa and nearby islands in the eastern Atlantic (20 species). However, one species (T. americana) occurs in Cuba (Fig. 1b) and the Bahamas, and another (T. albertschwartzi), presumably extinct, is known from a single specimen collected in Jamaica in the 19th Century and “discovered” recently in a museum in Scotland (Sprackland and Swinney, 1998). Geckos in the genus Tarentola typically are nocturnal and inhabit dry, rocky areas.

Previous molecular analyses have investigated Tarentola phylogeny using immunological distance data from serum albumin (Joger, 1985) and DNA sequences from mitochondrial and nuclear genes (Carranza et al., 2000, 2002). The sequence analyses united the Old World species in a single
group, leaving the New World species *T. americana* as the most divergent lineage in the genus. This also agrees with its placement as the sole representative of the subgenus *Neotarentola* (Joger, 1984). No sequences have yet been obtained from the museum specimen of *T. albertschwartzi*, although morphologically it has been suggested to be most closely related to an African species, *T. deserti*, rather than *T. americana* (Carranza et al., 2000; Sprackland and Sweeney, 1998). Molecular clock analyses have suggested that *T. americana* diverged from the Old World species approximately 14–17 million years ago (Carranza et al., 2002). The New World species must have arrived by dispersal from Africa by floating on flotsam across the Atlantic on the North Equatorial Current (Hedges, 1996b), either in one event (if they are close relatives) or two separate events. The distance would have been approximately 5000 km at that time (the continents were slightly closer) and would have taken 3–6 months based on typical rates of current flow (Guppy, 1917).

One of us (S.B.H.) collected specimens of *T. americana* in Cuba, and first noted the presence of a small, cryptic species in eastern Cuba in 1987 (also noted by R. Crombie, personal communication). Subsequently S.B.H. collected additional specimens of the new species, sympatric with *T. americana*, along the southern coast of Santiago de Cuba province and Guantánamo province in eastern Cuba (Fig. 1b). Aside from a substantial difference in body size and in arrangement of tubercles, the two species are so similar morphologically that a sequence analysis was performed, using portions of the mitochondrial cytochrome *b* gene and nuclear amelogenin gene, to test the hypothesis that two species were involved. No specimens of the Bahamian taxon *T. a. warreni* were available for comparison with the Cuban material (*T. a. americana*).

### 1.3. Hemidactylus haitianus

The gecko genus *Hemidactylus* is comprised of approximately 80 species, most of which occur in Asia and Africa but with some species occurring in the New World. Geckos in this genus are especially adapted to living with humans and as a consequence are frequently dispersed around the World. Of the five species known from the West Indies, four also occur in the Old World: *H. brooki* (includes *H. ‘haitianus’*), *H. garnotii*, *H. mabouia*, and *H. turcicus*. However, without genetic data or other evidence it is sometimes difficult to distinguish between a natural colonization and human introduction. This is further confounded by a paucity of useful scale characters for taxonomy and pattern polymorphisms within populations.

In the case of *H. haitianus* Meerwarth, 1901 (treated as a subspecies of *H. brooki* by some authors), it has been argued that this West Indian taxon arose either by natural dispersal from Africa (Kluge, 1969) or by human introduction (Vanzolini, 1978). Powell and Maxey (1990) reviewed the taxonomic history of *H. b. haitianus*, and Powell (1993) and Powell et al. (1996) treated *H. b. haitianus* as a full species (*H. haitianus*) endemic to the West Indies. Recently, a DNA sequence analysis of portions of two mitochondrial genes helped to clarify some taxonomic confusion involving species of the genus *Hemidactylus* (Carranza and Arnold, 2006). Surprisingly, the two samples of
H. haitianus examined, both from western Cuba, appeared nested among populations of an African species, H. angulatus (Hallowell, 1854) (previously considered a subspecies of H. brooki). Specifically, they clustered closely (only 1.4% divergence) and significantly with a sample of that species from Bioko Island, Equatorial Guinea. For this reason, the authors concluded the H. haitianus may be conspecific with H. angulatus.

To further clarify the taxonomic status of New World geckos assigned to H. haitianus, Carranza and Arnold (2006) suggested that Hispaniolan and Puerto Rican samples of H. haitianus should also be examined. Hispaniolan material is especially pertinent because the type locality of H. haitianus is “Haïti”. To help resolve this question, we have sequenced samples of H. haitianus from Cuba, Hispaniola, and Puerto Rico using the same gene fragments used by Carranza and Arnold (2006).

2. Methods

2.1. Sample collection and DNA sequencing

Specimens were collected by hand at localities throughout the Greater Antilles. Field and laboratory research was approved by the Institutional Animal Care and Use Committee of Pennsylvania State University (#17632). Tissues were removed and frozen in liquid nitrogen, or temporarily transferred to the laboratory in 75% ethanol. The remaining specimen was preserved. Tissues were maintained in the laboratory at -80°C. The specimens sampled, localities (Fig. 1), laboratory numbers, and Genbank accession numbers are listed in the Appendix A.

DNA was extracted from tissue samples, amplified (PCR) with primers spanning defined regions of genes, and sequenced. Different combinations of four mitochondrial genes (12S rRNA, tRNA Valine, 16S rRNA, and cytochrome b) and one nuclear gene (amelogenin) were used with each project, reflecting the different questions being addressed. The Hemidactylus project involved collecting additional data for a larger project already published (Carranza and Arnold, 2006) and therefore we used the same genes (12S rRNA and cytochrome b). The Tarentola project involved the resolution of a cryptic species and therefore we selected a fast-evolving gene (cytochrome b), and added a nuclear gene (amelogenin) for corroboration. The Phyllodactylus project involved morphologically divergent populations on different islands and therefore we used genes evolving more slowly (12S rRNA, tRNA Valine, and 16S rRNA).

The primers used were (listed 5-prime to 3-prime, with gene indicated in prefix of primer name): 12L8, CAG CAGTATWAAATTA; 12G1H, CTGGYGAAGG CGGATAGAYGCC; 12G2L, AAACYWAAGGACTT GACGGTG; 12G2H, ACCATGTCRAGGATTAC GGG; 16GIL, TTAGGGACAGCYTRACGTCTCA CG; 16G1H, GCCGTTTAAAARTGGTCATCGG CA; 16L40, CGAGCCTCATGATAGCTGTTGCTCA; 16H45, GATTTRYGCTACCTTTGCACGGT TAG; 16L9, CGCCTTTTATCAAAAAACT; 16H9, CCGTCTGTA ACTCAGATCAGT; CBL14841, AAAAGCTTCCCAT CCAACATCTCAGATGATGAAA; CBH33, GGCAA ATAGGAAARTCACC; AMTA-L, ATCCAGGTAT GGCTATGAACCTA; AMTA-R, GGAAGTACAGCGG TCGATTGGGTTG. For the P. wirshingi project we used the primers 12L8/12G1H, 12G2L/12G2H to obtain approximately 1 kb of sequence from most of the 12S rRNA, the complete rRNA Valine, and the beginning of the 16S rRNA gene. We also obtained approximately 1 kb from the middle of the 16S rRNA gene using primers 16G1L/16G1H (or 16L40/16H45) and 16L9/16H9. These two fragments were combined in the analyses. A 646 bp portion of the cytochrome b gene was amplified for the T. americana and H. haitianus projects with primer pair CBL14841/CBH33. Primers used in amplification of the 12S rRNA gene of Hemidactylus included 12L8/12G1H and 12G2L/12G2H. For T. americana, a small portion of the amelogenin gene (308 bp) was amplified with the primers AMTA-L and AMTA-R. Sequence alignments have been deposited in Genbank.

2.2. Phylogenetic analyses

Following sequence alignment with CLUSTAL (Thompson et al., 1997) in MEGA 3.1 (Kumar et al., 2004), phylogenies were constructed using maximum likelihood (ML), minimum evolution (ME), and Bayesian methods of inference, using PAUP 4b10 (Swofford, 2002) and PHYML (Guindon and Gascuel, 2003), MEGA 3.1 (Kumar et al., 2004), and MrBayes 3.0 (Ronquist and Huelsenbeck, 2003), respectively. ML analysis was conducted with the best-fit model estimated using Modeltest (Posada and Crandall, 1998), which was GTR + gamma for the Phyllodactylus data set and TrN + I + gamma for the Tarentola data set; model parameters were estimated from the data. The ME analysis used the Tamura-Nei + gamma model. The Bayesian analysis was run with model parameters estimated from the data and used 10⁶ generations in four chains, sampling every 100 generations, with a burn-in of one-third of the sampled trees. The last one-third of the sampled trees was used to construct a majority-rule consensus tree. Statistical significance was evaluated with bootstrapping (ML and ME, 2000 replications) and Bayesian posterior probabilities.

For the Phyllodactylus analyses, a New World species (P. xanti) was included and Goggia lineata was added as the outgroup. For the Tarentola analyses, several Old World species were included: T. annularis (AF364322), T. boehmi (AF364320), T. boettgeri (AF184997), T. darwini (AF185047), T. deserti (AF364321), T. mauritanica (AF364327), and Pachydactylus turneri (AF184990); H. tucicus (AF184989) was added as the outgroup. The H. haitianus analysis involved simple comparison of DNA sequences and therefore did not require phylogenetic analysis.
2.3. Divergence time analyses

No fossils of West Indian Phyllodactylus were available for calibration. We could infer only rough time estimates based on a clock calibration used elsewhere for geckos and tied to a geologic event (Carranza and Arnold, 2006; Carranza et al., 2000, 2002). We adjusted that calibration so that it applied only to the 12S rRNA region in common and we used corrected sequence divergence because uncorrected divergence, frequently used in the literature, is not proportional with time. This sequence divergence was calculated between groups (Kumar et al., 2004), using the specified model. Relative rate tests (Tajima, 1993) were used to test for rate variation in the lineages to be timed.

As with Phyllodactylus, no fossils were available for direct calibration of West Indian Tarentola. Nonetheless, a calibration point was available: the divergence of T. americana from Old World species of Tarentola (13.6–17 Ma) estimated by Carranza et al. (2002) using maximum likelihood estimates of branch length. For this reason we employed the Bayesian program Multidivtime (Thorne and Kishino, 2002), using model parameter data generated by ML in PAML (Yang, 1997). Most of the Bayesian priors were based on instructions and recommendations included in the software manual. The calibration (13.6–17 Ma) and the expected time between tip and root (rttm, 24 Ma) were based on time estimates from the previous analysis of Tarentola (Carranza et al., 2002). Experimentation was made with several other priors (see discussion below).

3. Results and discussion

3.1. Phyllodactylus wirshingi

Our analyses of 1990 bp of DNA sequence data from the 12S rRNA, tRNA Valine and 16S rRNA genes yielded a robustly supported tree (Fig. 2a). Each of the three taxa are monophyletic and the two Hispaniolan taxa (P. w. hispaniolae and P. w. sommeri) are closest relatives. The clades themselves and their relationships are the same and significant, using all three methods of analysis.

These new data bear on the taxonomy of the P. wirshingi complex in the West Indies. The close genetic relationship of samples within each taxon and deep divergence among taxa provides additional support (beyond morphology) for the recognition of three distinct taxa. The close relationships of samples within P. w. wirshingi and within P. w. sommeri are not surprising because they came from essentially the same populations (the two localities of P. w. wirshingi are in close proximity). However, the two localities sampled for P. w. hispaniolae are at opposite ends of the distribution of that subspecies (~80 km apart) and thus their close relationship is of greater significance in unifying populations of that taxon.

The question then arises as to whether these three taxa should be recognized as subspecies or species. Without the evidence of coexistence (sympathy), any such decisions are by their nature subjective. However, most taxonomists rely on quantitative measures of difference, from morphology and molecules, as guides in making such decisions. The “yardstick” for what constitutes a sufficient difference between two allopatric taxa to recognize them as species usually comes from comparisons of other valid (often sympatric) species in the same group, because different groups will have different characters and different thresholds (Hedges, 2002).

In the case of P. wirshingi, all three taxa as currently known are allopatric and there is no evidence of intergradation. Their diagnostic differences include scale characters and pattern differences often used to diagnose different species (Dixon and Huey, 1970; Schwartz, 1979), although those characters are mostly not diagnostic if considered individually. In terms of molecular divergence at these two conserved ribosomal genes, the two Hispaniolan taxa are separated by 6.0% sequence divergence and the Hispaniolan clade and Puerto Rican clade are separated by 15.0% sequence divergence. Such levels of divergence are similar to or greater than differences observed between species of other geckos using sequences of these mitochondrial ribosomal genes (Hass, 1996). Considering all of this evidence, we believe the taxa are more appropriately recognized at the species level and therefore elevate the two Hispaniolan taxa to full species: P. hispaniolae (n. comb.) and P. sommeri (n. comb.).

Our time estimate for the divergence of the Antillean clade from the mainland American species (P. xanti) was 21 Ma (early Miocene), based on its sequence divergence of 39.2% and a rate of 1.85% sequence divergence per million years. Other time estimates were 8.1 Ma (late Miocene) for the split of the Hispaniolan clade and the Puerto Rican clade, and 3.2 Ma (Pliocene) for the split of the two Hispaniolan species. Without inclusion of the other West Indian species, P. pulcher from Barbados, and other mainland species of the genus, it is not possible to infer either the time of origin or source area for the Greater Antillean species. However, the comparison with P. xanti suggests that this clade arose by overwater (flotsam) dispersal during the middle to late Cenozoic and was not an ancient inhabitant of the proto-Antilles. The various models that have been proposed to explain the historical biogeography of the West Indian herpetofauna are reviewed elsewhere (Hedges, 1996a, 2001, 2006).

3.2. Tarentola americana

Our analyses of sequences from the mitochondrial cytochrome b gene define two deeply split and statistically significant clades within T. americana (Fig. 2b). One large clade comprises specimens allocated to the large species T. americana whereas the other comprises four specimens of the smaller, undescribed cryptic species of Cuban Tarentola. Within the large clade, all of the specimens from eastern Cuba form a significant group and the specimen
from central Cuba (Trinidad) forms the basal lineage. Sequences of the small fragment of the nuclear gene amelogenin were highly conserved, showing variation at only seven sites among *Tarentola*. Although the same two clades formed in the tree (not shown), corroborating the evidence from the cytochrome *b* analyses for the recognition of a cryptic species, none of the nodes were significant, as expected with such few variable sites.

The new species is completely sympatric with *T. americana* and is considerably smaller in body size. For example, the type specimen of *T. americana* is 99 mm snout-vent length (Gray, 1831; Loveridge, 1944) and the species reaches snout-vent lengths of 120 mm (Schwartz and Henderson, 1991) whereas adults of the new species are approximately 40 mm in snout-vent length. The morphological description of the new species and complete discussion of the taxonomic history of *T. americana* will be made elsewhere (L. Diaz and S.B.H., unpublished).

We have no material of the Bahamian taxon *T. americana warreni* for sequence comparisons and therefore cannot comment on its taxonomic status other than to say that its large size, up to 92 mm SVL (Schwartz, 1968), eliminates the possibility that the new cryptic Cuban species is associated with that taxon. The relatively deep split between *T. americana* from central Cuba and the clade from eastern Cuba is interesting in that distribution of the species within Cuba shows a gap of 400–500 km between those two areas, aside from localities in the southern archipelago of Jardines de al Reina (Schwartz and Henderson, 1991). This suggests that populations of *T. americana* from central and western Cuba (also large in size) may warrant recognition as a separate species, after further study. The distinctiveness of the
populations from Isla de Pinos was already noted by Schwartz (1968). Thus, West Indian Tarentola may represent a complex of species.

In estimating divergence times among Tarentola, it was discovered that two of the Bayesian priors, the mean (brownmean) and standard deviation (brownsd) of the prior for the Brownian motion constant “nu,” had unexpectedly strong affects on the estimates. There is no justification for any particular value for these priors, although most users set them so that the product of rttm and brownmean (Yoder and Yang, 2004). Nonetheless, some authors have used values as low as 0.2 for the product of rttm and brownmean (Yoder and Yang, 2004). For the Tarentola data set, setting rttm × brownmean to 1.0 resulted in time estimates that were pulled strongly towards the calibration point. The divergence of T. americana and the small, cryptic species on Cuba was 14.1 Ma (11.5–16.3 Ma, 95% credibility interval) and the divergence of the central Cuban lineage (T. americana) from the eastern Cuban lineage of T. americana was 10.6 (7.0–13.8) Ma. The bias was obvious in the time estimates for the closely related sequences, of individuals of the same species from the same population: they had small sequence divergences (1–3% of the calibration node) yet had large time estimates of 3–5 Ma (20–25% of the calibration node time). In contrast, a global clock analysis estimated the divergence times of different individuals from the same populations (among T. americana and among T. new species) as <1 Ma, which is consistent with expectations of allelic diversity among individuals within a population of a small vertebrate. Such systematic overestimation of recent divergence times, using Bayesian methods, has been noted elsewhere (Ho et al., 2005), although the bias is not yet fully understood.

Based on these findings we set rttm × brownmean to 0.2 and placed an upper constraint of 1.0 Ma on the basal nodes containing sequences of the same species from the same population. The resulting times, which we believe are more realistic, are 11.4 (7.2–15.2) Ma for the split between the large and small Cuban species and 5.5 (2.8–9.1) Ma for the split between the central and eastern Cuban large Tarentola. Corresponding dates for those two nodes based on a global clock (linearized tree) analysis (Kumar et al., 2004) were 12.3 and 5.3 Ma, respectively. These dates indicate a long occupation and diversification (~15 million years) of this genus of geckos on Cuba, following a remarkable ~5000 km journey on flotsam across the Atlantic from Africa.

3.3. Hemidactylus haitianus

We sequenced 12S rRNA (942 bp) and cytochrome b (543 bp) in five individuals of this species, resulting in a total of 1485 bp. Two samples were from eastern Cuba, two were from Hispaniola, and one was from Puerto Rico. The sequences were all identical. Carranza and Arnold (2006) sequenced shorter regions of those genes in their Hemidactylus study, including 399 bp of 12S rRNA and 303 bp of cytochrome b, but their two sequences of H. haitianus from western Cuba were identical to each other and to our sequences at all 702 common nucleotide sites. Thus, no phylogenetic analysis was needed. Such a lack of genetic variation is expected in organisms that have dispersed recently, as a result of human introduction. Members of the H. brooki complex, including this taxon, are well known for being “weedy” species that are frequently transported by humans (Carranza and Arnold, 2006).

These new data now help to resolve the taxonomic status the West Indian geckos currently called either H. brooki or H. haitianus. Because they show virtually no genetic variation throughout the Greater Antilles, and they cluster very closely with one population (Bioko Island, Equatorial Guinea) of an African species, H. angulatus, we hereby consider H. haitianus to be a junior synonym of H. angulatus. Thus, all West Indian geckos formerly referred to H. brooki or H. haitianus are assigned to H. angulatus. Given the strong, unidirectional trade in slaves that occurred between West Africa and the West Indies for several hundred years, it is likely that H. angulatus reached the West Indies by this mechanism.

Carranza and Arnold (2006) pointed out that the genetic variation observed among populations of H. angulatus suggests that a future revision of that species may reveal that it is a complex of species. If that happens, it is conceivable that resurrection of the name H. haitianus might be necessary, for the clade that includes both African and West Indian populations. The generalized type-locality of H. angulatus, “West coast of Africa” (Hallowell, 1854), although possibly restricted to Liberia based on other information in the description, does not currently offer resolution of this question.

Acknowledgments

Specimens used in this study were collected over many years and with the assistance of numerous persons in the field, especially Richard Thomas; Aaron Bauer provided material of Goggia lineata and Phyllodactylus xanti. S.B.H. is grateful to the governments of Cuba, Haiti, and the Dominican Republic for granting permission to collect and export specimens. This research was supported by grants to S.B.H. from the National Science Foundation.

Appendix A. Species, localities, and sequence accession numbers

In the following list, localities are provided for each sample, followed by the laboratory tissue collection number, and (in parentheses) sample reference number (if used in figures and if more than one individual of a taxon) and Genbank sequence accession numbers.
A.1. Phyllodactylus analyses

Goggia lineata—South Africa: Northern Cape; Richtersveld, 267107 (AY763261, AY763274).

Phyllodactylus hispaniolae—Dominican Republic: Barahona Province; 12 km E. Canoa, 102913 (1, AY763265, AY763286). Independencia Province; 5.1 km NW La Descubierta, 194517 (2, AY763272,AY763283).

P. sommeri—Haiti: L’Arribonite; 10.4 km NW Ca Soliel, 160727 (1, AY763266, AY763278), 160736 (2, AY763267, AY763279).

P. wirshingi—Puerto Rico: Playa de Tamarindo, 101728 (1, AY763262, AY763275), 101729 (2, AY763263, AY763276), 101730 (3, AY763264, AY763277), 171742 (4, AY763268, AY763280), 171743 (5, AY763269, AY763281). Puerto Rico, Bahia de la Ballena, 190729 (7, AY763271, AY763282).


A.2. Tarentola analyses

Tarentola americana—Cuba: Guantanamo Bay US Naval Station, 161873 (1, EF202100), 161874 (2, EF202101), 161875 (3, EF202102), 161876 (4, EF202103, EF202126), 161880 (5, EF202104), 161949 (6, EF202105), 171004 (7, EF202106, EF202121). Cuba: Sancti Spiritus; Trinidad, 172911 (8, EF202107, EF202128). Guantánamo; Boca de Yumuri, 190225 (9, EF202108, EF202122); Loma Redonda (5 km NW Hatibonico), 190969 (10, EF202109, EF202124); Yacabo Abajo, 190242 (11, EF202110, EF202123); 4.9 km S La Tinta 191265 (12, EF202111, EF202125); 7 km SW Baitiquiri, 191370 (13, EF202112, EF202127).

T. sp. nov—Cuba: Guantánamo; Loma Redonda (5 km NW Hatibonico), 190970 (1, EF202113, EF202119); 3.9 km N Hatibonico, 191571 (4, EF202114, EF202120). Santiago de Cuba; south side of Laguna Baconao, 190617 (2, EF202115, EF202117); 190618 (3, EF202116, EF202118).

A.3. Hemidactylus analyses


References


