Origin and evolutionary relationships of giant Galápagos tortoises

Adalgisa Caccone1,6, James P. Gibbs3, Valerio Ketmaier1, Elizabeth Suatoni, and Jeffrey R. Powell1

1Department of Ecology and Evolutionary Biology and 2Molecular Systematics and Conservation Genetics Laboratory, Yale Institute for Biospheric Studies, Yale University, New Haven, CT 06520-8106; 3College of Environmental Science and Forestry, State University of New York, Syracuse, NY 13210; and
4Dipartimento di Biologia Animale e dell’Uomo, Università di Roma “La Sapienza”, 00185 Rome, Italy

Communicated by John C. Avise, University of Georgia, Athens, Georgia, September 13, 1999 (received for review May 28, 1999)

Perhaps the most enduring debate in reptile systematics has involved the giant Galápagos tortoises (Geochelone nigra), whose origins and systematic relationships captivated Charles Darwin and remain unresolved to this day. Here we report a phylogenetic reconstruction based on mitochondrial DNA sequences from Galápagos tortoises and Geochelone from mainland South America and Africa. The closest living relative to the Galápagos tortoise is not among the larger-bodied tortoises of South America but is the relatively small-bodied Geochelone chilensis, or Chaco tortoise. The split between G. chilensis and the Galápagos lineage probably occurred 6 to 12 million years ago, before the origin of the oldest extant Galápagos island. Our data suggest that the four named southern subspecies on the largest island, Isabela, are not distinct genetic units, whereas a genetically distinct northermost Isabela subspecies is probably the result of a separate colonization. Most unexpectedly, the lone survivor of the abingdoni subspecies from Pinta Island (“Lonesome George”) is very closely related to tortoises from San Cristóbal and Española, the islands farthest from the island of Pinta. To rule out a possible recent transplant of Lonesome George, we sequenced DNA from three tortoises collected on Pinta in 1906. They have sequences identical to Lonesome George, consistent with his being the last survivor of his subspecies. This finding may provide guidance in finding a mate for Lonesome George, who so far has failed to reproduce.

The giant tortoises of the Galápagos Islands (Geochelone nigra) are renowned both as spectacular animals and for their contribution to the development of Darwin’s theory of natural selection. They have become a symbol for efforts to conserve the unique fauna of the Galápagos Islands (whose very name refers to these creatures). Despite these tortoises’ prominence in the debate protracted over 160 years (1, 2), little is known about the origin and evolutionary history of the species. Here we present a DNA-based phylogenetic study of these animals. The results elucidate both the origin of the lineage (by identifying the closest living relative) and reveal genetic distinctiveness and relationships of the now recognized subspecies.

Giant tortoises, up to 5 feet in length, were widespread on all continents except Australia and Antarctica before and during the Pleistocene (3, 4). Now extinct from large landmasses, giant tortoises have persisted through historical times only on remote oceanic islands: the Galápagos, Seychelles, and Mascarenes. The tortoises of the Mascarenes are now extinct; the last animal died in 1804 (5). The tortoises of the Seychelles are represented by a single surviving population on the Aldabra atoll. Only in the Galápagos have distinct populations survived in multiple localities. The Galápagos tortoises remain the largest living tortoises (up to 400 kg) and belong to a pantropical genus of some 21 species (6).

The Galápagos Islands are volcanic in origin; the oldest extant island in the eastern part of the archipelago is less than 5 million years (myr) old (7); volcanic activity is ongoing, especially on the younger western islands. Because the archipelago has never been connected to the mainland, tortoises probably reached the islands by rafting from South America, 1000 km to the east. The Humboldt Current travels up the coast of Chile and Peru before diverting westward at Equatorial latitudes corresponding to the Galápagos Archipelago. Three extant species of Geochelone exist on mainland South America and are therefore the best candidates for the closest living relative of the Galápagos tortoises: Geochelone denticulata, the South American yellow-footed tortoise; Geochelone carbonaria, the South American red-footed tortoise; and Geochelone chilensis, the Chaco tortoise.

Within the archipelago, up to 15 subspecies (or “races”) of Galápagos tortoises have been recognized, although only 11 survive to the present (2, 8). Six of these are found on separate islands; five occur on the slopes of the five volcanoes on the largest island, Isabela (Fig. 1). Several of the surviving subspecies of Galápagos tortoises are seriously endangered. For example, a single male nicknamed Lonesome George represents G. nigra abingdoni from Pinta Island. The decline of the populations began in the 17th century when buccaneers and whalers collected tortoises as a source of fresh meat; the animals can survive up to six months without food or water. An estimated 200,000 animals were taken (2). More lasting, these boats also introduced exotic pests such as rats, dogs, and goats. Today, these feral animals, along with continued poaching, represent the greatest threat to the survival of the tortoises.

The designated subspecies differ in a number of morphological characters, such as carapace shape (domed vs. saddle-backed), maximum adult size, and length of the neck and limbs. These differences do not, however, permit clear discrimination between individuals of all subspecies (9). Similarly, an allozyme survey that included seven G. nigra subspecies and the three South American Geochelone failed to reveal patterns of genetic differentiation among the subspecies or to identify any of the mainland species as the closest living relative to the Galápagos tortoises (10). A robust phylogeny of the Galápagos tortoise complex and its relatives is thus unavailable currently, and it is much needed to help resolve the long-term debate over the systematics of this group, as well as to clarify subspecies distinctiveness as a basis for prioritizing conservation efforts.

Materials and Methods

The subspecies’ names, the island on which each is found, and the numbers of individuals analyzed are reported in Table 1. Samples for four subspecies, becki, microphyes, darwini, and chathamensis, were provided by Edward Louis (Henry Doorly’s Zoo, Omaha, NE); we collected the others. Two individuals were analyzed for each of the three mainland South American Geo-
chelone species. The samples of *G. carbonaria* and *G. denticulata* were provided by the Wildlife Conservation Society (New York); the Chaco tortoise samples came from the Sedgwick County Zoo (Wichita, KS). As an outgroup, we used *Geochelone pardalis*, the African leopard tortoise (whose sequences are reported in ref. 11).

DNA was extracted from blood stored in 100 mM Tris/100 mM EDTA/2% SDS buffer by using the Easy DNA extraction kit (Invitrogen). Modified primer pair 16Sar + 16Sbr (12) was used for PCR amplifications of 586 bp of the 16S rRNA gene. A 386-bp-long fragment of the cytochrome *b* (cytb) gene was amplified by using the cytb GLU: 5'-TGACATGAAAAYCAYCGTGTG (13) and cytb B2: H15149 (14) primers. The D-loop region was amplified with primers based on conserved sequences of the cytb and 12S rRNA genes, which flank the D loop in tortoises. Primer GT12STR (5'-ATCTTGCCAACCTCAGTGCC-3') is at the 5' end of the 12S ribosomal gene, and primer CYTTOR (5' GCTTAACTAAGACCACCGGCTTGT-3') is at the 3' end of the cytb gene. These primers amplify the D loop from several *Geochelone* species (unpublished observations). Internal primers specific to the D loop of *G. nigr* were used to amplify and sequence a 708-bp fragment of the D loop (corresponding to 73.7% of the region). Internal primer sequences are available from the senior author upon request. Double-stranded PCR amplifications and automated sequencing were carried out as described (11). To promote accuracy, strands were sequenced in both directions for each individual.

In addition to blood from live animals, we also obtained samples of skin from three tortoises collected on Pinta Island in 1906 and now in the California Academy of Science, San Francisco (specimen numbers CAS 8110, CAS 8111, and CAS 8113). One-half gram of skin was surface-cleaned with sterile water and subjected to 20 min of UV irradiation. The skin was pulverized in liquid nitrogen and suspended in buffer A of the Easy DNA kit. Proteinase K (100 µg/ml) was added and the sample was incubated for 24 hr at 58°C, following the Easy DNA procedure with the addition of a second chloroform extraction. The samples were washed in a Centricron 30 microconcentrator (Amicon) and suspended in 100 µl of 10 mM Tris/1 mM EDTA, pH 8.0. Only one skin sample was extracted at a time. Several rounds of PCR were performed, finally yielding four fragments of about 150 bp each, representing about 75% of the sequence obtained from blood samples. All procedures on the skin samples (until PCR products were obtained) were done in a room separate from that where all other DNA work was done.

Because of the high sequence similarity, sequences were aligned by eye. The alignment was also checked by using CLUSTAL W (15). Alignments are available from the first author. Phylogenetic analyses were carried out on each gene region and on the combined data set. *G. pardalis* was used as the outgroup. Phylogenetic inferences were made by using maximum parsimony (MP) (16), maximum likelihood (ML) (17), and neighbor joining (NJ) (18). MP trees were reconstructed by the branch-and-bound search method (19) with ACCTRAN (accelerated transformation) character–state optimization as implemented in PAUP* (20). Various weighting methods were used: all substitutions unweighted, transversions (Tv) weighted 3 times transitions (Ti), or only Tv. For cytb, MP analyses were also performed, excluding Ti from third positions of all codons. ML analyses were carried out using PAUP* with an empirically determined transition/transversion ratio (9.19) and rates were assumed to follow a gamma distribution with an empirically determined shape parameter (α = 0.149). Sequences were added randomly, with 1,000 replicates and TBR as the branch-swapping algorithm. For the NJ analysis, ML distances were calculated by PAUP* with the empirically determined gamma parameter. PAUP* was used to obtain NJ trees based on those distance matrices.

The incongruence length difference test (21) was carried out as implemented in PAUP* (in which it is called the “partition homogeneity test”). As suggested by Cunningham (22), invariant characters were always removed before applying the test. Templeton’s (23) test was used to compare competing phylogenetic hypotheses statistically, by using the conservative two-tailed Wilcoxon rank sum test (24). The significance of branch length in NJ trees was tested by using the confidence probability (CP) test as implemented in MEGA (25). The strength of support for individual nodes was tested by the bootstrap method (26) with 1,000 (MP and NJ) or 300 (ML) pseudoreplicates. Rate homogeneity among lineages was tested by Tajima’s one-degree-of-freedom method (27).

**Results**

In most cases, different individuals of the same subspecies have identical DNA sequences for cytb and 16S rRNA. The exceptions were four transitional differences among three Pinzón animals and two transitions between the two Santa Cruz samples; these were coded as polymorphisms. Therefore, in the analyses, we used a single sequence to represent each subspecies. We performed the incongruence length difference test as implemented in PAUP* and found that the data from the two regions of the mtDNA were not significantly incongruent (*P* = 0.59). Therefore, we present here the results from the combined data set.

Fig. 2 shows the 50% majority rule consensus tree for MP generated from the cytb and 16S rRNA data combined, by using

---

**Fig. 1.** Map illustrating the location of the Galápagos Islands with respect to the South American coast. On the South American continent are depicted the three mainland Geochelone, from the top: *G. denticulata*, *G. carbonaria*, and *G. chilensis*. The enlarged area includes a map of the Galápagos Archipelago with the names of islands that have extant subspecies of tortoises. The sizes of the tortoises are in proportion to one another.
therefore performed other tests to ask whether alternative trees are statistically worse than are those in Fig. 2. Table 2 presents the results of these tests. Constraining one of the other mainland South American species to be the sister taxon to the G. nigra, or using the three mainland species as a trichotomy produced significantly less parsimonious trees by Templeton’s (23) test, even with the relatively conservative two-tailed Wilcoxon rank sum test (24). For the NJ tree, the crucial branch separating the chilensis/nigra clade from the other South American species is significant at the 98% level by the confidence probability test in MEGA (25).

Estimates of genetic distances also support the sister taxa status of G. chilensis and G. nigra. Among the subspecies of G. nigra, the maximum likelihood distances range from 0 to 0.0124 with a mean of 0.0066 ± 0.004 (SD). Between subspecies of G. nigra and G. chilensis, the average distance is 0.0788 ± 0.005. Between G. nigra and G. carbonaria or G. denticulata ML distances are 0.118 ± 0.005 and 0.116 ± 0.003, respectively.

Fig. 2 also reveals some resolution of the relationships among the Galápagos subspecies. One point of interest is that the five named subspecies on Isabela do not form a monophyletic clade. The four southern Isabela subspecies are sister taxon to the subspecies from Santa Cruz, whereas the northernmost subspecies, G. n. becki, is the sister taxon to G. n. darwini on San Salvador. It is a geographically reasonable scenario for southern Isabela to be colonized from Santa Cruz and northern Isabela to be colonized from San Salvador (Fig. 1).

There is virtually no evidence for genetic differentiation among the four southern Isabela subspecies. The cyt b sequence is identical in all individuals sampled. There are only three differences in the 16S rRNA sequence among the eight samples of these four named subspecies. We have also sequenced what is generally the fastest evolving region of mtDNA, the D loop, in individuals from these four subspecies to test whether this region gives evidence of genetic differentiation (Fig. 3). Only 17 of the 708 sites varied among the 23 individuals sequenced, and there

---

**Table 1. Geochelone nigra subspecies used in this study**

<table>
<thead>
<tr>
<th>Subspecies</th>
<th>Locality</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>becki</td>
<td>Isabela–Wolf</td>
<td>2</td>
</tr>
<tr>
<td>microphyes</td>
<td>Isabela–Darwin</td>
<td>2</td>
</tr>
<tr>
<td>vandenburghi</td>
<td>Isabela–Alcedo</td>
<td>2</td>
</tr>
<tr>
<td>guntheri</td>
<td>Isabela–Sierra Negra</td>
<td>2 (5)</td>
</tr>
<tr>
<td>vicina</td>
<td>Isabela–Cerro Azul</td>
<td>2 (10)</td>
</tr>
<tr>
<td>darwini</td>
<td>San Salvador</td>
<td>2</td>
</tr>
<tr>
<td>porteri</td>
<td>Santa Cruz</td>
<td>2</td>
</tr>
<tr>
<td>abingdoni</td>
<td>Pinta</td>
<td>4*</td>
</tr>
<tr>
<td>ephippium</td>
<td>Pinzón</td>
<td>3</td>
</tr>
<tr>
<td>chatamensis</td>
<td>San Cristóbal</td>
<td>2</td>
</tr>
<tr>
<td>hoodensis</td>
<td>Española</td>
<td>6</td>
</tr>
</tbody>
</table>

* N refers to the number of individuals whose 16S rRNA and cyt b were studied; numbers in parentheses are the number sequenced for the D loop.

*One sample was from a live animal and three were from skins of dead animals.

---

**Table 2. Analyses of the robustness of the phylogenetic tree in Fig. 2**

<table>
<thead>
<tr>
<th>Node (Fig. 2)</th>
<th>MP</th>
<th>NJ</th>
<th>ML</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>88</td>
<td>94</td>
<td>81</td>
<td>98</td>
</tr>
<tr>
<td>b</td>
<td>100</td>
<td>99</td>
<td>93</td>
<td>99</td>
</tr>
<tr>
<td>c</td>
<td>99</td>
<td>100</td>
<td>92</td>
<td>97</td>
</tr>
<tr>
<td>d</td>
<td>90</td>
<td>79</td>
<td>84</td>
<td>71</td>
</tr>
<tr>
<td>e</td>
<td>84</td>
<td>92</td>
<td>82</td>
<td>89</td>
</tr>
<tr>
<td>f</td>
<td>80</td>
<td>63</td>
<td>85</td>
<td>63</td>
</tr>
</tbody>
</table>

Bootstrap support for the indicated node is tabulated as percentage of 1,000 pseudoreplicates for NJ and MP and percentage of 300 pseudoreplicates for ML. CP is the confidence probability (28) of the branch in the NJ tree leading to the indicated node.

---

Fig. 2. Phylogenetic tree based on 962 bp of the 16S ribosomal and cyt b mtDNA. The tree shown is the 50% majority rule consensus tree for MP with G. pardalis as the outgroup. Numbers above branches indicate branch lengths. All lettered nodes were obtained with all weightings for MP, NJ, and ML, and by using each gene separately. See Materials and Methods for details of the analyses; see Table 2 for statistical analyses of topology robustness.
were seven equally most parsimonious trees. The tree is only 23 steps long for the 23 sequences, with only seven nodes... vic, vicina; gun, guntheri; van, vandenburghi; and mic, microphyes.

Discussion
Origin of the Galápagos Lineage. There is good evidence that the Chaco tortoise (G. chilensis) is the closest extant relative of the Galápagos tortoises (Fig. 2, Table 3). G. chilensis’ present range extends from southwestern Bolivia, western Paraguay, and northwestern Argentina southward to about 40° S in northern Patagonia (6). It is usually a relatively small-bodied tortoise (carapace length is approximately 43 cm) compared with the other two continental species (Fig. 1). These two larger species are adapted to humid environments, living in tropical evergreen and deciduous rainforests and moist savannas (6). In contrast, the Chaco tortoise lives in a variety of dry lowland habitats, such as savanna, thorn and scrub brush, and desert, where it feeds on grasses, succulents, and cacti (6). These latter ecological conditions are very similar to those experienced by G. nigra on the Galápagos Archipelago.

Although G. chilensis is the closest living relative of the Galápagos tortoise, it is unlikely that the direct ancestor of G. nigra was a small-bodied tortoise. Several lines of reasoning (for review, see ref. 2) suggest that gigantism was a preadapted condition for successful colonization of remote oceanic islands, rather than an evolutionary trend triggered by the insular environment. Giant tortoises colonized the Seychelles at least three separate times (29). Fossil giant tortoises are known from mainland South America, and morphological analysis of these and extant species are consistent with a clade containing giant tortoise fossils and G. chilensis (30).

Further evidence that the split between the ancestral lineages that gave rise to G. chilensis and G. nigra occurred on mainland South America comes from time estimates based on a molecular clock. We applied the Tajima (27) test of the clocklike behavior of DNA sequences to pairwise comparisons between G. chilensis and Galápagos subspecies, using in turn G. carbonaria and G. denticulata as the outgroup. The tests were done on transitions and transversions together, and on transversions only. We could not reject the hypothesis of constant substitution rates for the vast majority (94%) of comparisons for both genes. Therefore, we assumed that the 16S rRNA and cyt b genes were evolving linearly with time. To calculate approximate divergence times between the lineages, we used published mtDNA rates estimated from turtles and other vertebrate ectotherms (31–33). Depending on which estimate and gene are used, the predicted time of the split between G. nigra and G. chilensis varies, but most put the date between 6 and 12 myr ago. The oldest extant islands (San Cristóbal and España) date to less than 5 myr ago (7), although sea mounts now submerged may have formed islands more than 10 myr ago (34). However, given the existence of mainland giant fossils and the argument that gigantism was required for long distance rafting, invoking colonization on now submerged islands would seem less reasonable than a split on the mainland before colonization, with the immediate ancestral lineage now extinct. The oldest split within G. nigra is estimated at no more than 2 myr ago, consistent with diversification on the existing islands.

Times of divergence and colonization of other prominent Galápagos organisms have been estimated by molecular data. The diversification of Darwin’s finches has been estimated to have occurred within the age of the extant islands (35). On the other hand, the endemic marine (Conolophorus) and land (Amblyrhynchos) iguanas are estimated to have diverged from each other between 10 and 20 myr ago (36, 37). As argued by Rasmann (37), it is likely that the split occurred on the Archipelago; therefore, it must have occurred on now-submerged islands. Similarly the diversification of the lava lizards (Tropidurus) and geckos (Phyllodactylus) was estimated to have begun around 9 myr ago, although in this case, there is some evidence indicating multiple colonizations (38, 39).

Taxonomic Status of Isabela Subspecies. From Fig. 2, it seems clear that the largest and youngest island with tortoise populations, Isabela, was colonized at least twice independently. The four
southern subspecies are sister taxa to the Santa Cruz subspecies (G. n. porteri), whereas the subspecies on the northernmost volcano (G. n. becki) is sister to the subspecies (G. n. darwinii) on San Salvador. We have found no significant genetic differentiation among the four southern Isabela subspecies (microphytes, vandenburghii, guntheri, and vicina), even for what should be the fastest evolving region of mtDNA (Fig. 3). The lack of genetic differentiation is perhaps not surprising in light of the age of the Isabela volcanoes, estimated to be less than 0.5 myr (7). For colonization by tortoises, most volcanic activity must have ceased and sufficient time must have passed for appropriate vegetation to develop. Given this relatively short time, coupled with long generation time [age of first reproduction is over 20 years (8)], significant genetic differentiation among these populations is unlikely. The genetic distinctness of the population on the northernmost volcano is accounted for by an independent colonization from another island.

The lack of genetic differentiation of these four Isabela subspecies is consistent with the morphological assessment of at least one authority. Pritchard (2) suggested that the four southern Isabela subspecies do not warrant separate subspecific status, but rather the “described differences are either attributable to environmental differences (especially of rainfall, food availability, and humidity), or do not show geographic correlation, but are artifacts of age and sex.” This, coupled with our results, would seem to warrant a reassessment of the taxonomic status of these subspecies.

The data presented here also indicate little or no genetic differentiation between or among subspecies connected to nodes c, d, and e in Fig. 2. However, faster evolving regions of the mtDNA do reveal diagnostic differences among all subspecies (unpublished data) with the exception of the four southern Isabela populations, for which none of our data indicate geographically structured differentiation. Because a major purpose of the present study was to identify the mainland sister taxon to the Galápagos lineage, we emphasize here relatively slowly evolving regions. The molecular diagnoses of subspecies, based on larger sample sizes than are available now, should be addressed in the near future.

Lonesome George. Perhaps the greatest surprise in our data was the close relationship of the single living representative of the G. n. abingdoni subspecies from Pinta to subspecies on Española and San Cristóbal. Most other relationships make biogeographic sense. The three well supported nodes in Fig. 2 (c, d, and e) all connect subspecies on islands geographically close to one another (Fig. 1). Pinta is the farthest major island from Española and San Cristóbal, being about 300 km distant. One possibility is that Lonesome George actually did originate on Española or San Cristóbal and was transported to Pinta. Morphologically, all three subspecies are considered saddle-backed, although subtle differences among them have been noted (2). Fortunately, we had available to us skin samples from three specimens collected on Pinta in 1906. The DNA sequences we obtained from these skins are identical to those of Lonesome George. Thus, it is reasonable to conclude that Lonesome George is the sole (known) living survivor of this subspecies.

Although based solely on geographic distance, it seems unlikely that the Pinta subspecies should be so closely related to those from Española and San Cristóbal, consideration of oceanic currents makes it plausible. There is a strong current running northwest from the northern coast of San Cristóbal leading directly to the area around Pinta (40). These tortoises are not strong swimmers and thus their direction of rafting in the ocean must have depended largely on currents.

Attempts to breed Lonesome George have been unsuccessful. However, he has been paired with females primarily from northern Isabela because, given its proximity, it was thought to be the most likely origin of the Pinta population (Fig. 1). Now that we see he has close genetic affinities to the Española and San Cristóbal subspecies, perhaps they would be a more appropriate source of a mate for this sole survivor.

We are grateful to Gabriele Gentile for producing Fig. 1, to Edward Louis (Henry Doorly’s Zoo, Omaha, NE) for providing DNA samples, George Amato (Wildlife Conservation Society) and the Sedgwick County Zoo (Wichita, KS) for tissue samples, and to Jens Vindum (California Academy of Sciences) for skin samples from the Pinta tortoises. Peter Pritchard made useful comments on an early version of this paper. We thank the Charles Darwin Research Station, the Parque Nacional Galápagos, and the government of Ecuador for permission to obtain samples and for assistance with field logistics; Cruz Marquez, Washington Tapia, Eduardo Vilema, Howard Snell, and Linda Cayot have been particularly helpful. This work was supported by a Consiglio Nazionale delle Ricerche Fellowship to V.K. and by a National Science Foundation grant to J.R.P.