Phylogeography of the Madeiran endemic lizard *Lacerta dugesii* inferred from mtDNA sequences

A. Brehm, a,∗ J. Jesus, a H. Spínola, a C. Alves, b L. Vicente, c and D.J. Harris d

a Centre of Macaronesian Studies, University of Madeira, Funchal 9000, Portugal
b IPATIMUP, University of Porto, Rua Dr. Roberto Frias, Porto 4000, Portugal
c Centro de Biologia Ambiental, Department of Zoology and Anthropology, University of Lisbon, C2 Campo Grande, Lisbon 1700, Portugal
d Centro de Estudos de Ciência Animal (CECA), ICETA-UP, Campus Agrário do Vairão, Vila do Conde 4485-661, Portugal

Received 5 March 2002; revised 27 June 2002

Abstract

Partial sequences from two mitochondrial DNA genes, cytochrome *b* and 12S rRNA, were used to assess the phylogenetic relationships of populations of *Lacerta dugesii* from the volcanic Atlantic islands of Madeira, the Desertas, Porto Santo, and the Selvagens. All four-island groups are genetically distinguishable and populations within each contain similar degrees of genetic diversity. Molecular clock estimates suggest that the islands were colonized much later after their emergence compared to other Atlantic islands, possibly due to their greater geographical isolation. Mismatch analysis of all populations is consistent with exponential growth, as expected after colonization of empty niches. The Selvagens contain genetic substructuring between the islets.

© 2002 Elsevier Science (USA). All rights reserved.

Keywords: *Lacerta dugesii*; 12S rRNA; Cytochrome *b*; Phylogeography; Madeira

1. Introduction

Geographical isolation is an important barrier to allow genetic differentiation of populations. Islands constitute perfect choices to study differentiation phenomena in lizard populations because gene flow among them may be almost non-existent if an effective sea barrier prevents the dispersal of individuals. In this case, each island may be regarded as a discrete evolutionary unit. Considerable genetic variation in reptiles has been found in inter- and intra-island studies and volcanism has been used to explain such a variation by producing rapid recurrent isolation and bottlenecking of populations (Brown et al., 2000; Gübitz et al., 2000).

The Madeira Archipelago consists of Madeira, Porto Santo Islands, the Desertas, and the Selvagens Islands (Fig. 1). Madeira is the largest island (750 km²) and lies about 700 km from the western coast of Africa (Morocco). The island is ecologically very heterogeneous due to very different weather conditions, humid subtropical in the north to hot and arid in the south and east. Porto Santo and the Desertas are almost deprived of vegetation. The Selvagens are situated about 300 km south of Madeira Island, mid-way to the Canary Islands. Geological studies have shown that all these islands have a long and complex volcanic history (Mitchel-Thomé, 1976). According to Geldmacher et al. (2000) and Galopim de Carvalho and Brandão (1991), island ages are 3.6 million years (MY) (Desertas), 4.6 MY (Madeira), 14 MY (Porto Santo), and 12 MY for the Selvagens. Volcanic activity occurred on Madeira almost exclusively from 3.9 to 4.6 MY and from 0.7 to 3.7 MY. The last eruptions are well localized and took place around 6500 YBP (Geldmacher et al., 2000). The three Desertas islands, together with Madeira, are part of the same volcanic complex and were probably connected by land bridges as recently as 18,000 years ago, before the last deglaciation periods (or the last Pleistocene glacial cycles), when the sea level rose. In fact, the sea depth between both groups is just 90 m in certain areas, otherwise the groups of islands are separated by deep channels and have never been connected above sea level.
The lizard *Lacerta dugesii* is the sole flightless vertebrate endemic to the Madeira Archipelago, although it had been introduced to some islands of the Azores and the Iberian Peninsula. In the Selvagens, the species coexists with a small gecko, *Tarentola bischoffii*. The origin of *L. dugesii* is unknown, but studies of Arnold (1989) and Harris et al. (1998) showed that its closest extant relative is *Lacerta perspicillata*, endemic to Northwest Africa, thus, suggesting that its ancestor could be from the African mainland. Other life history traits also suggest this relationship (Galán and Vicente, 2002). *L. dugesii* presents broad intra- and inter-island morphological and colour pattern polymorphism (Cook, 1979; Crisp et al., 1979). Several subspecies have been described in the past, on the basis of morphological or enzymatic polymorphisms. Mertens (1938) described a subspecies endemic to the Desertas (*Lacerta dugesii mauli*), but Bischoff et al. (1989) accepted *Lacerta dugesii dugesii* on Madeira and the Desertas, and described *Lacerta dugesii jogeri* for Porto Santo Island and *L. dugesii selvagensis* for the Selvagens. A recent study using enzymatic markers and morphological characters failed to show a clear picture of the relationships between the different island populations (Brehm et al., 2001a).

Because the age of the Madeira Archipelago, as well as the time some of the islands have been linked are well known, *L. dugesii* is a good model to study the population genetic variation among islands and compare this with the islands’ age, size, and distances separating them. The goal of this study was to present an insight into the genetic variation and relationships of *L. dugesii* populations, particularly the possible patterns that may explain the species past historic events related to the colonization process among islands. Also, to assess for intra-island variation we analyzed four different populations from sites that are geologically of different ages from Madeira Island. We use partial cytochrome *b* and 12S rRNA mtDNA gene sequences. These regions have been shown to evolve in a clocklike fashion in other reptiles, with approximately 2% divergence per million years (Carranza et al., 2000), so we can compare the levels of variation within *L. dugesii* to this calibration. The data gathered might also give insight into the subspecific status that had been proposed for some of these populations.

2. Materials and methods

Ninety-six *L. dugesii* specimens were collected from 10 localities spanning the species present-day distribution (Fig. 1). Four of the 10 sites are located in Madeira Island and coded as MP, MC, ML, and MF. Two sites are located in each of the two islands of the Selvagens Archipelago (360 km from Madeira) and coded as SG and SP. Three sites are localized in each of the three islands constituting the Desertas group (24 km from Madeira) and coded as DG, DB, and DC. Finally, PS is a sample from the island of Porto Santo (42 km north of Madeira). The number of individuals per site is indicated in Table 1. Genomic DNA was extracted from tail-tip samples of *L. dugesii* following standard phenol–chloroform protocols. For each individual, mitochondrial cytochrome *b* and 12S fragments were amplified by
PCR using the published primers and PCR conditions of Kocher et al. (1989). Sequencing of the amplified products was performed in both strands in an automated sequencer (ABI Prism). Previously published sequences of the closely related \textit{L. perspicillata} and two \textit{Podarcis hispanica} complex were included as outgroups (Harris et al., 1998; Harris and Sá-Sousa, 2002). Sequences of both genes were aligned using Clustal X (Thompson et al., 1997). Minor adjustments to the alignment of 12S rRNA sequences were made by eye with reference to the secondary structure (Hickson et al., 1996). Differences in substitution rates between gene regions can potentially result in conflicting signals if one gene is saturated. Thus, we used the partition homogeneity test (Farris et al., 1994) implemented in PAUP*4 (Swofford, 2001) to evaluate whether the two regions contained significantly different phylogenetic signals.

### Table 1

Variable sites of the cytochrome \( b \) and 12S rRNA mtDNA gene sequences in 96 individuals of \textit{Lacerta dugesii}

| Haplotype names | M | M | M | S | S | D | D | D | D | P | P | P | C | L | G | P | G | P | G | B | C | S |
| TTAAGGCTTCCGCGGTACTCGCTACGCGTTAGTACATCAAGCTCAGGCTACACTCC---TAA- | (M1) | 1 | 5 | 1 | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| TTTCTTGCCTCCGCTTCCGCTACGCGTTAGTACATCAAGCTCAGGCTACACTCC---TAA- | (M2) | 1 | 5 | 1 | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Haplotype names are shown in parentheses and digits at the top of the table indicate nucleotide positions relative to the beginning of sequences. Numbers on the right are the total number of individuals per site bearing a haplotype. Site codes are according to Fig. 1. Dots represent identical bases to the first haplotype, a dash represents a deletion. The last haplotype is from \textit{Lacerta perspicillata}. The entire sequences are deposited in GenBank.
regions ($p = 0.66$), so they were combined in all phylogenetic analyses (Table 1).

### 2.1. Data analysis

Evolutionary relationships above and below the species level are fundamentally different in nature. Relationships between genes for individuals belonging to different species are hierarchical, while those sampled from individuals within a species are not (Posada and Crandall, 2001). Methodologies for interpreting the genetic variation at these two levels are therefore also different. We expect that populations from different island groups could behave as “species” with limited gene flow, whereas, within island groups, gene flow could be extensive. We used the statistical parsimony algorithm (Templeton et al., 1992) performed in TCS (Clement et al., 2000) to estimate the maximum number of differences among haplotypes as a result of a single substitution with a 95% confidence interval and also the most probable ancestral haplotype. For each of these identified groups, inter-haplotype divergence was represented by median networks (Bandelt et al., 1995, 2000). For the phylogenetic analysis, we used maximum likelihood (ML) and Bayesian inference. For the ML analysis, we followed the approach outlined by Hulsenbeck and Crandall (1997) to test 56 alternative models of evolution, employing PAUP*4 (Swofford, 2001) and Modeltest version 3 (Posada and Crandall, 1998). L. perspicillata and two Podarcis species were designated as outgroups. A starting tree was obtained using neighbour-joining (NJ). With this tree, likelihood scores were calculated for 56 various models of evolution and then compared statistically using a $\chi^2$ test with degrees of freedom equal to the difference in free parameters between the models being tested. Once a model of evolution was chosen, it was used to estimate a tree by median networks (Bandelt et al., 1995, 2000). The Bayesian analysis was implemented using MrBayes (Huelsenbeck and Ronquist, 2001), which calculates Bayesian posterior probabilities using a Metropolis-coupled, Markov chain Monte Carlo (MC-MCMC) sampling approach. Two independent runs were conducted with random starting trees, run $0.5 \times 10^6$ generations, and sampled every 10 generations using a general-time-reversible model of evolution. In both searches, stationarity of the Markov Chain was determined as the point when sampled log-likelihood values plotted against generation time reached a stable mean equilibrium value; “burn-in” data sampled from generations preceding this point were discarded. All data collected at stationarity were used to estimate posterior nodal probabilities and a summary phylogeny. Two independent replicates were conducted and inspected for consistency to check for local optima (Huelsenbeck and Bollback, 2001). To test for demographic signatures of population expansions following colonization of each island group, we used the mismatch distribution analysis (Rogers and Harpending, 1992). To compare the observed distribution with that expected under the expansion model, we calculated the raggedness index (Rogers and Harpending, 1992) using Arlequin (Schneider et al., 2000).

### 3. Results

Amplified cytochrome $b$ and 12S rRNA partial genes yielded unambiguous sequences of 299 and 371 bp in length, respectively. Cytochrome $b$ sequences presented no indels, but the 12S sequences varied from 365 to 371 in size, depending on a variable C-stretch characteristic of this sequence. The cytochrome $b$ and 12S (including gaps as a fifth character) sequences comprised 37 and 30 variable positions, respectively. Table 1 shows the variable nucleotide positions found in both sequences. Numbers referring to these positions are based on the entire sequences that are available from the GenBank database under Accession Nos. AF543302–AF543309. Thus, nucleotide positions 6–285 refer to the partial cytochrome $b$ gene fragment sequenced. Variable positions of the 12S sequences are from nucleotides 311 to 657.

None of the networks from the four island groups (Madeira, Porto Santo, Selvagens, and Desertas) could be confidently linked using statistical parsimony, so separate networks were constructed for each group (Fig. 2). The number of haplotypes and the number of individuals carrying a particular haplotype per site are shown in Table 1. Haplotypes are labelled according to the place where animals were captured (M, Madeira Island; S, Selvagens group; D, Desertas group; and P, Porto Santo Island). Madeira Island yielded 22 different haplotypes in 39 samples from four sites. The Selvagens group had 11 haplotypes in 20 individuals, the Desertas 11 in 27 individuals, and Porto Santo Island had 5 haplotypes in 10 individuals. Table 1 also includes the cytochrome $b$ and 12S sequences of L. perspicillata (Harris et al., 1998) as a reference outgroup. In this case, only the variable positions found in L. dugesii are shown in L. perspicillata.

For the ML analysis, we concluded that the TrN (Tamura and Nei, 1993) model with a discrete approximation to a gamma distributed rate heterogeneity model ($\alpha = 0.21$) was the most appropriate model of evolution for this data set. A heuristic search incorporating this model found seven trees of $-\ln 2091$ (Fig. 3). These differed only in the position of the short branches separating within island haplotypes.
Fig. 4 shows the mismatch distribution (Rogers and Harpending, 1992) of all sequences and a comparison to Poisson expectation. The close match between the two suggests that *L. dugesii* populations have undergone a historical population increase and/or range expansion. The raggedness index (Rogers and Harpending, 1992) was low and not significantly different from expectation (Madeira: $R = 0.27$, $p = 0.35$, Selvagens: $R = 0.019$, $p = 0.91$, Porto Santo: $R = 0.26$, $p = 0.13$, Desertas: $R = 0.03$, $p = 0.974$). This indicates a smooth distribution of pairwise differences and again supports a hypothesized population increase.

Fig. 2. Medium networks showing the phylogenetic and geographic relationships between *Lacerta dugesii* mtDNA haplotypes (circles). Relationships for the different groups of islands are represented separately: Porto Santo, Madeira, Selvagens, and Desertas Islands. Circled areas are proportional to the number of individuals bearing a particular haplotype. Branch lengths are proportional to the number of mutations involved between haplotypes. Numbers on branches refer to the mutated base and are the same as in Table 1. Indels were excluded. Labels inside circles correspond to haplotypes from Table 1. Different shadows represent the proportion of haplotypes found in the different sites sampled. A star indicates the most probable root of the network.
All haplotypes found can be unambiguously assigned to one of the four major groups of islands, Madeira, Desertas, Selvagens, and Porto Santo, as none is common to two different island groups. Pairwise mismatch analysis (Fig. 4) shows single monomodal distributions for Madeira, Desertas, and Porto Santo. This is the signature expected, following the exponential growth in a panmictic population. Although we have only one population sampled from Porto Santo, we cannot make inferences of geographical isolation, for Desertas and Madeira we can infer that there is substantial gene flow between populations. For the two populations from the Selvagens, however, there is a more bimodal distribution of pairwise differences, something expected when populations are geographically subdivided (Marjoram and Donnelly, 1994). Further, when the migration rate is very low or the initial population size is high, separation of the peaks in the pairwise mismatch analysis is more extreme. Our close bimodal data are more similar to results from simulated data with a very low initial population size and a high migration rate (Marjoram and Donnelly, 1994), something unsurprising given the geographic isolation of the Selvagens coupled with the close proximity of islands within the group. According to Bravo and Coello (1978), although the Selvagens are very old, most of the emerged land dates only to the Quaternary, up to 1.5 MY. Other studies suggest a constant subaerial landmass dating to the mid-Pliocene and the probable existence of several small islands in the recent past of this island group (Geldmacher et al., 2001). Assuming a molecular clock of 1.96% divergence per million years (Carranza et al., 2000), the \L. dugesii populations on the Selvagens have been reproductively isolated nearly 2.5 MY. Although calibration of molecular clocks is complex (see Bromham, 2002), our data seem to be more consistent with the hypothesis proposed by Geldmacher et al. (2001). It will be interesting to determine the genetic diversity within the Selvagens endemic gecko \Tarentola bischoffi to test these hypotheses further.

Both the ML and Bayesian analyses support the monophyly of the Selvagens, Porto Santo, and the Desertas. The two populations from the Selvagens are nested within those from Madeira. Individuals from the Selvagens have enough fixed differences from those on Madeira that the two networks cannot be linked. However, they also contain all the fixed differences that define the Madeira populations. Therefore, in the phylogenetic analyses, populations from Madeira are paraphyletic. This is not surprising if the Selvagens were colonized from Madeira and shows the advantages of using both networks and phylogenetic analyses in combination.

Genetic diversity within islands is not correlated with island age—there is no significant difference in the number of haplotypes in each island, despite the fact that Porto Santo is approximately four times older than the Desertas. This could be due to a relatively recent radiation of \L. dugesii. The highest uncorrected divergence between populations is only approximately 5.6%. This divergence is similar between each of the four island groups, suggesting that once \L. dugesii reached the first island, it spread quickly to the others. If the ancestor of \L. dugesii colonized the islands only 2.8 mya, all the islands were already formed. This means that both Porto Santo and the Selvagens were not colonized for more than 10 MY. Studies of the herpetofauna from the Canary Islands (Carranza et al., 2000) and Cape Verde Islands (Brehm et al., 2001b; Carranza et al., 2001) suggest that these islands were colonized much more rapidly after their emergence. The late arrival of \L. dugesii could be due to a poor transmarine colonization ability of lacertids compared to geckos—geckos have made more separate colonizations within these island groups than either skinks or the lacertid lizards \Gallotia (inferred from Carranza et al., 2000; Jesus et al., 2001,
The Madeira islands are also more geographically isolated from continental source populations than either the Canary islands or the Cape Verde and this could also explain the delay in their colonization. Few other studies have assessed genetic diversity in the fauna and flora of the Madeira archipelago. Whereas *L. dugesii* shows strong differentiation among islands, the liverwort *Porella canariensis* does not, with populations from Madeira being paraphyletic with respect to both Porto Santo and the Desertas (Freitas and Brehm, 2001). More studies are needed for other taxa before any general pattern of colonization in the archipelago can be determined.

Sequences for an individual from *L. dugesii* from São Miguel Island in the Azores (Harris et al., 1998) partially overlap with those used in this analysis. Because the exact same regions are not available, it was not included in the phylogenetic analysis. However, a simple NJ analysis including this taxon (with non-overlapping sections included and coded as missing data in this sequence, analysis not shown) provides compelling evidence that the source population for this introduction was from Madeira. *L. dugesii* are found on several islands in the Azores and only further data and more variable markers will determine whether these were introduced once or several times.

Our data suggest that four genetically distinct units exist within *L. dugesii*, on Madeira, the Desertas, Porto Santo, and the Selvagens. Genetic monophyly of the Desertas populations would support the resurrection of *L. dugesii maulti* as an endemic subspecies. Gene flow is clearly restricted to be minimal or non-existent, even though the distance between the Desertas and Madeira is minimal. Colonization of the islands took place enough time ago to allow the accumulation of unique mutations in each of the island groups.

The Canary Islands have become an important model region for studying the colonization and diversification of different organisms (Juan et al., 2000), especially reptiles (e.g., Brown et al., 2000; Carranza et al., 2000; Gübitz et al., 2000). Recently, phylogenetic analyses have been used to study similar processes among the Cape Verde herpetofauna (e.g., Brehm et al., 2001b; Carranza et al., 2001; Jesus et al., 2001). Whereas, many of the patterns observed in these island archipelagos are similar, others are not. By obtaining phylogenetic data across diverse taxa in all three systems, we can understand more clearly the complex nature of colonization and subsequent radiations.

Fig. 4. Observed distribution of pairwise differences among individuals within each island group ((A) Porto Santo, (B) Madeira, (C) Selvagens, and (D) Desertas). A simulated Poisson’s distribution is indicated by a dotted line. All four groups fit a wave characteristic of an episode of population growth. The bimodal wave observed in the Selvagens (C) also indicates population fragmentation.
Acknowledgments

We thank Parque Natural da Madeira for providing some of the specimens used in this study. The authors are grateful to Professor A. Amorim from IPATIMUP, University of Porto, for his support, to Professors H.-J. Bandelt and Adalgisa Caccone for their help and comments on data analysis, two anonymous reviewers, and P. Andrade for the technical support. The work was partially financed through contract PBIC/C/BIA/2095/95 from Fundação para a Ciência e Tecnologia, Lisbon, to LV. JJ is the recipient of a Ph.D. grant PRAXIS XXI/BD/2637/94 from FCT.

References


