

# Comparing patterns of nuclear and mitochondrial divergence in a cryptic species complex: the case of Iberian and North African wall lizards (*Podarcis*, Lacertidae)

CATARINA PINHO\*, D. JAMES HARRIS and NUNO FERRAND

*CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão and Departamento de Zoologia e Antropologia, Faculdade de Ciências da Universidade do Porto, 4099-002 Porto, Portugal*

Received 31 October 2005; accepted for publication 1 July 2006

Combining different sources of information is essential for a complete understanding of the process of genetic differentiation between species. The Iberian and North African wall lizard (*Podarcis*) species complex has been the object of several studies regarding morphological and mitochondrial DNA variation but, so far, no large-scale survey of nuclear variation within this group has been accomplished. In this study, ten polymorphic allozyme loci were studied in 569 individuals collected across the Iberian Peninsula and North Africa. The obtained data were analysed using both conventional population genetic tools and recent Bayesian model-based clustering methods. Our results show that there are several well-differentiated entities corroborating the major splits observed in mtDNA analyses. These groups correspond not only to the fully recognized species *Podarcis bocagei*, *Podarcis carbonelli*, and *Podarcis vaucheri* but also to multiple forms within the polytypic *Podarcis hispanica*, all of which have a similar level of differentiation to that observed between the acknowledged species. However, relationships between forms are weakly supported both by population and individual clustering methods, suggesting a scenario of a rapid diversification that contrasts to the clear bifurcating model assumed from previous mtDNA analyses. Individual multilocus analyses report few individuals misassigned or apparently admixed, some of which are most likely explained by the persistence of high levels of ancestral polymorphism. Other admixed individuals, however, are probably the result of limited levels of gene flow between forms. © 2007 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2007, 91, 121–133.

ADDITIONAL KEYWORDS: allozyme variation – diversification – nuclear markers – population structure.

## INTRODUCTION

In the last decade, mitochondrial DNA (mtDNA) has been the major source of information used to describe genetic variability and uncover evolutionary relationships within groups of organisms. Along with a relative easiness in obtaining large data sets using conserved primers (Kocher *et al.*, 1989), theoretical approaches generalized the idea that mtDNA trees have a higher probability of recovering the correct species tree because mtDNA takes one-quarter of the time to acquire monophyly than a nuclear gene does (Moore,

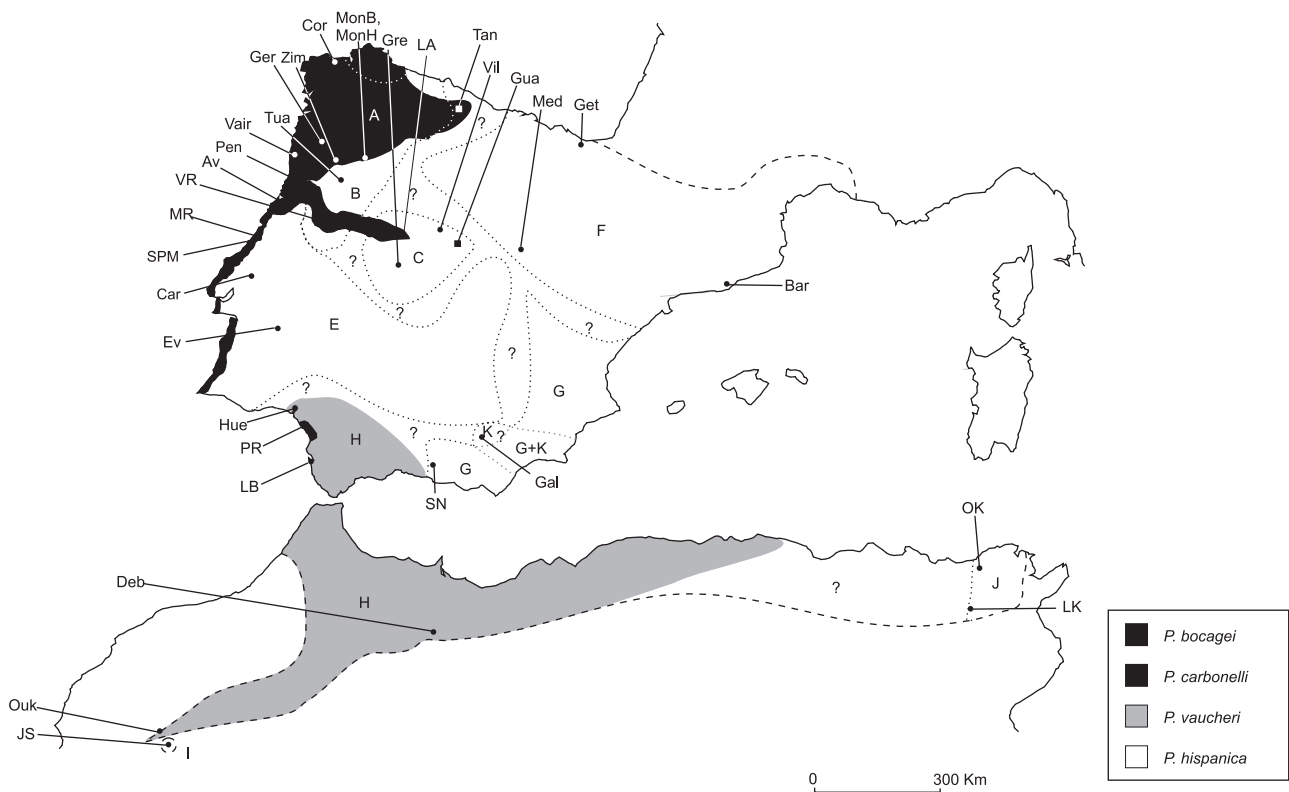
1995). However, this notion has been questioned by studies showing that the achievement of monophyly is largely dependent on stochasticity (Hudson & Turelli, 2003) and by the acknowledgement of the limitations of the use of a single locus to infer evolutionary relationships (Pamilo & Nei, 1988; Zhang & Hewitt, 2003; for a review, see Ballard & Whitlock, 2004). Unlinked loci evolve independently and may portray different evolutionary scenarios on the basis of stochastic lineage sorting or of different evolutionary forces, such as selection, acting upon them (Hey, 1997; Ballard, Chernoff & James, 2002). Fluctuating effective population sizes (Fay & Wu, 1999; Mønsen & Blouin, 2003), gender-biased gene flow (FitzSimmons *et al.*, 1997; Nyakaana & Arctander, 1999; Pierrney

\*Corresponding author. E-mail: catarina@mail.icav.up.pt

*et al.*, 2000), and introgression (DeSalle & Giddings, 1986; Shaw, 2002; Alves *et al.*, 2003; Chan & Levin, 2005) are known to affect mitochondrial and nuclear markers differently, implying that the use of mitochondrial DNA alone may result in biased estimates of the evolutionary history. These differences become increasingly important when dealing with closely-related species, in which traditional bifurcating trees usually do not accurately represent the patterns of divergence and where hybridization and introgression are a distinct possibility (Machado & Hey, 2003).

One such case is that of wall lizards (*Podarcis*) in the Iberian Peninsula and North Africa. Despite a long-standing debate about their systematics because of their extremely high morphological variability (Mertens & Müller, 1940; Klemmer, 1959), until recently, a conservative view considering only the existence of only two endemic species, *Podarcis bocagei* (Seoane, 1884) and *Podarcis hispanica* (Steindachner, 1870), prevailed (Arnold & Burton, 1978;

Barbadillo *et al.*, 1999). However, recent morphological and mitochondrial DNA studies (Geniez, 2001; Harris & Sá-Sousa, 2001, 2002; Harris *et al.*, 2002b; Sá-Sousa, Vicente & Crespo, 2002; Pinho, Ferrand, & Harris, 2006) suggested that Iberian and North African *Podarcis* are in fact a species complex and that there is broad-scale agreement between morphologically identified entities and genetic variation. Consequently, some taxonomic reevaluations were carried out, namely the elevation to the species status of *Podarcis carbonelli* Pérez-Mellado, 1981, a former subspecies of *Podarcis bocagei* (Sá-Sousa & Harris, 2002) and of the south Iberian/west Mahgrebin form *Podarcis vaucheri* (Boulenger, 1905), formerly included in *Podarcis hispanica* (Oliverio, Bologna & Mariottini, 2000; Busack, Lawson & Arjo, 2005). The remaining cryptic forms within the paraphyletic *P. hispanica* have not yet been the object of a taxonomic reassessment (Pinho *et al.*, 2006); for a tentative map of the distribution of mtDNA lineages, see Figure 1.



**Figure 1.** Map showing the predicted distribution of the mitochondrial DNA lineages and sampling sites for this study. Dashed lines represent the limits of the distribution of the Iberian and North African species complex. Dotted lines represent putative limits of the distribution areas of forms of *Podarcis hispanica* bearing distinct mitochondrial DNA lineages and were drawn connecting localities where these forms were found. It is still uncertain if and where the majority of the forms contact. A–K indicate different forms as defined by mtDNA studies (Pinho *et al.*, 2006): A, *Podarcis bocagei*; B, *Podarcis hispanica* type 1A; C, *P. hispanica* type 1B; D, *Podarcis carbonelli*; E, *P. hispanica* type 2; F, *P. hispanica* type 3; G, *P. hispanica* s.s.; H, *Podarcis vaucheri*; I, *P. hispanica* Jebel Sirwah type; J, *P. hispanica* Tunisian type; K, *P. hispanica* Galera type.

This scenario of the existence of multiple differentiated lineages within a relatively small geographical area follows a pattern that is now being acknowledged as common: southern European peninsulas probably not only provided glacial refugia for many European taxa, as has been traditionally suggested (Hewitt, 1996, 1999, 2004), but also as differentiation hotspots within themselves (the so-called 'refugia within refugia', Gómez & Lunt, 2007). This is true for many species complexes, including other *Podarcis* groups that occur in other peninsulas (Poulakakis *et al.*, 2003, 2005a, 2005b; Podnar, Mayer & Tvrtković, 2005). However, many of these studies relied only upon mitochondrial DNA differentiation and there is still no evidence that the observed groups would be corroborated by studies on nuclear variation.

The same is true for Iberian and North African *Podarcis*, for which comprehensive surveys of nuclear variation have not yet been accomplished. Preliminary analysis of allozyme markers (Pinho, Harris & Ferrand, 2003, 2004; Busack *et al.*, 2005) appear to be concordant with morphology and mitochondrial DNA data. However, such studies included only one or a few populations of each mitochondrial DNA type, did not sample some of the lineages, and did not address the possibility of hybridization and gene flow between the forms, mostly because only conventional population genetic tools were used. Issues such as introgressive hybridization can be more appropriately tackled using methods of analysis based on individual multilocus genotypes. Recent advances in this field include model-based clustering methods that do not require prior knowledge about population structure (Banks & Eichert, 2000; Pritchard, Stephens & Donnelly, 2000; Dawson & Belkhir, 2001; Anderson & Thompson, 2002; Corander, Waldmann & Sillanpää, 2003). Of these, one of the most commonly used is the Bayesian algorithm implemented in the software STRUCTURE (Pritchard *et al.*, 2000), which identifies in a given sample clusters of individuals that (as far as possible) are not in Hardy–Weinberg and linkage disequilibrium. Applied to our case study, this method has the obvious advantage of allowing a test of whether the genetic substructuring observed in previous studies (i.e. the existence of well-defined sets of populations; Pinho *et al.*, 2003, 2004), is also expressed as clearly identifiable sets of individuals. Additionally, it allows us to address the question of hybridization between forms of Iberian and North African *Podarcis*.

Taking this into consideration, in the present study, we use a set of allozyme markers to address the following questions: (1) are the groups defined on the basis of mtDNA differentiation corroborated using nuclear markers; (2) are inferred phylogenetic rela-

tionships using nuclear markers the same as those predicted by mtDNA; and (3) can hybridization be detected between the described forms?

## MATERIAL AND METHODS

A total of 569 individuals from 32 populations, including previously published data, were collected between 2000 and 2004 across Portugal, Spain, Morocco, and Tunisia (Fig. 1). MtDNA lineage identification was confirmed by using populations where at least one individual had already been sequenced (Harris & Sá-Sousa, 2001, 2002; Harris *et al.*, 2002a, b; Pinho *et al.*, 2006; C. Pinho, D. J. Harris & N. Farrand, unpubl. data). As outgroup, two populations of *Podarcis muralis*, a species that also exists in the Iberian Peninsula but has a different evolutionary origin (Harris & Arnold, 1999; Oliverio *et al.*, 2000; Harris *et al.*, 2005), were used. Sampling details (sample codes, localities, mtDNA correspondence, and sample sizes) are shown in Table 1. Samples consisted of a portion of tail tissue obtained from tail autotomy. All lizards were released after this procedure. Samples were stored frozen at  $-80\text{ }^{\circ}\text{C}$  prior to analysis. Tissue extraction, protein separation, and enzymatic detection of all loci followed the procedures given in Pinho *et al.* (2003). From the initially described battery of 11 polymorphic loci, enzymatic locus *NP* was excluded because it stopped providing consistently interpretable results, as previously explained in Pinho *et al.* (2004). Therefore, variation at ten polymorphic loci was studied by means of allozyme electrophoresis (*PEPA*, *PEPD*, *MPI*, *IDH*, *6-PGD*, *GOT*, *GPI*) and isoelectric focusing (*LDH-2*, *PGM*, *PEPB*).

Allelic frequencies were calculated directly from the observed genotypes. GENEPOP software, version 3.1b (Raymond & Rousset, 1995) probability test was used to determine whether populations were in Hardy–Weinberg and linkage equilibrium. To evaluate the partition of genetic diversity among and within groups of known mtDNA lineages, ARLEQUIN, version 2.0 (Schneider, Roessli & Excoffier, 2000) was used to calculate pairwise  $F_{st}$  values (as well as their significance) between all population pairs and to perform an analysis of molecular variance (Excoffier, Smouse & Quattro, 1992). In this analysis, groups were defined on the basis of their mtDNA ancestry, excluding the outgroup (i.e. 11 groups were considered). Pairwise  $F_{st}$  values between each of the mtDNA-defined entities were plotted to compare the magnitude of differentiation between the four recognized species (*P. bocagei*, *P. carbonelli*, *P. vaucheri*, and *P. hispanica*) with that between distinct lineages within *P. hispanica*. Genetic relationships among populations were estimated through a Neighbour-joining (NJ; Saitou & Nei, 1987) tree based on Nei's standard genetic distance (Nei,

**Table 1.** Localities and sizes of the samples examined for ten allozyme loci in this study

Species/morphotype	Sample code	Locality	Sample size
A. <i>Podarcis bocagei</i>	Vair	Vairão, Portugal*	34
	Cor	A Coruña, Spain	16
	MonB	Montesinho, Portugal*	30
	Zim	Zimão, Portugal	25
B. <i>Podarcis hispanica</i> type 1A	MonH	Montesinho, Portugal*	20
	Ger	Gerês, Portugal	14
	Tua	Tua, Portugal*	17
	Pen	Pendilhe, Portugal	21
C. <i>Podarcis hispanica</i> type 1B	Vil	Villacastin, Spain	9
	LA	La Alberca, Spain*	14
	Gre	Gredos, Spain	20
D. <i>Podarcis carbonelli</i>	Av	Aveiro, Portugal*	17
	SPM	S. Pedro de Moel, Portugal*	22
	PR	Playa del Rompeculos, Spain	12
	VR	Villasrúbias, Spain	21
E. <i>Podarcis hispanica</i> type 2	MR	Monte Real, Portugal*	20
	Car	Cartaxo, Portugal	8
	Ev	Évora, Portugal	20
F. <i>Podarcis hispanica</i> type 3	Bar	Barcelona, Spain*	14
	Med	Medinaceli, Spain	9
	Get	Getaria, Spain	20
	SN	Sierra Nevada, Spain*	18
G. <i>Podarcis hispanica sensu stricto</i>	SN	Sierra Nevada, Spain*	18
H. <i>Podarcis vaucheri</i>	Hue	Huelva, Spain	20
	LB	La Barrosa, Spain*	18
	Deb	Debdou, Morocco*	21
	Ouk	Oukaïmeden, Morocco*	26
I. <i>Podarcis hispanica</i> Jebel Sirwah type	JS	Jebel Sirwah, Morocco	8
J. <i>Podarcis hispanica</i> Tunisian type	OK	Oued Kébir, Tunisia*	16
	LK	Le Kef, Tunisia	12
	Gal	Galera, Spain	14
K. <i>Podarcis hispanica</i> Galera type	Gal	Galera, Spain	14
L. <i>Podarcis muralis</i>	Tan	Tanes, Spain	20
	Gua	Guadarrama, Spain	13

Sample codes and letters identifying mitochondrial DNA lineages correspond to those in Figure 1.

\*Previously published data (Pinho *et al.*, 2003, 2004).

1972), using PHYLIP, version 3.5 (Felsenstein, 1993). Bootstrap support was estimated using 1000 pseudoreplicates.

To evaluate genetic differentiation at the individual level, STRUCTURE, version 2.1 (Pritchard *et al.*, 2000) was employed. The initial parameter settings included the assumptions of no admixture and of independent allele frequencies between groups, based on the high genetic distances detected at the mtDNA level. The genetic structure was forced to vary from  $K = 1$  to  $K = 14$  clusters, the latter corresponding to the actual number of genetic entities (described at the mtDNA level) included in the sample plus two. STRUCTURE was run for 550 000 steps, of which the first 50 000 were discarded as burn-in. For each value of  $K$ , ten independent replicates of the Markov Chain Monte Carlo (MCMC) were conducted. The software

DISTRUCT (Rosenberg, 2002) was used to visually represent the obtained data. We used the method proposed by Evanno, Regnaut & Goudet (2005) to choose, amongst the values of  $K$ , the one that best characterized the data set. As an extension to this analysis, in the cases where STRUCTURE detected misidentified or potentially admixed individuals involving species where introgression is a strong possibility, five extra runs were performed, considering only the involved taxa, assuming admixture, and using 200 000 MCMC steps after 20 000 steps of burn-in to estimate the proportion of ancestry from each group. In addition, three runs of 100 000 steps each were performed with the software NEWHYBRIDS (Anderson & Thompson, 2002), which also implements a Bayesian-clustering algorithm based on individual genotypes, to assign the admixed individuals to a specific hybrid class (pure of

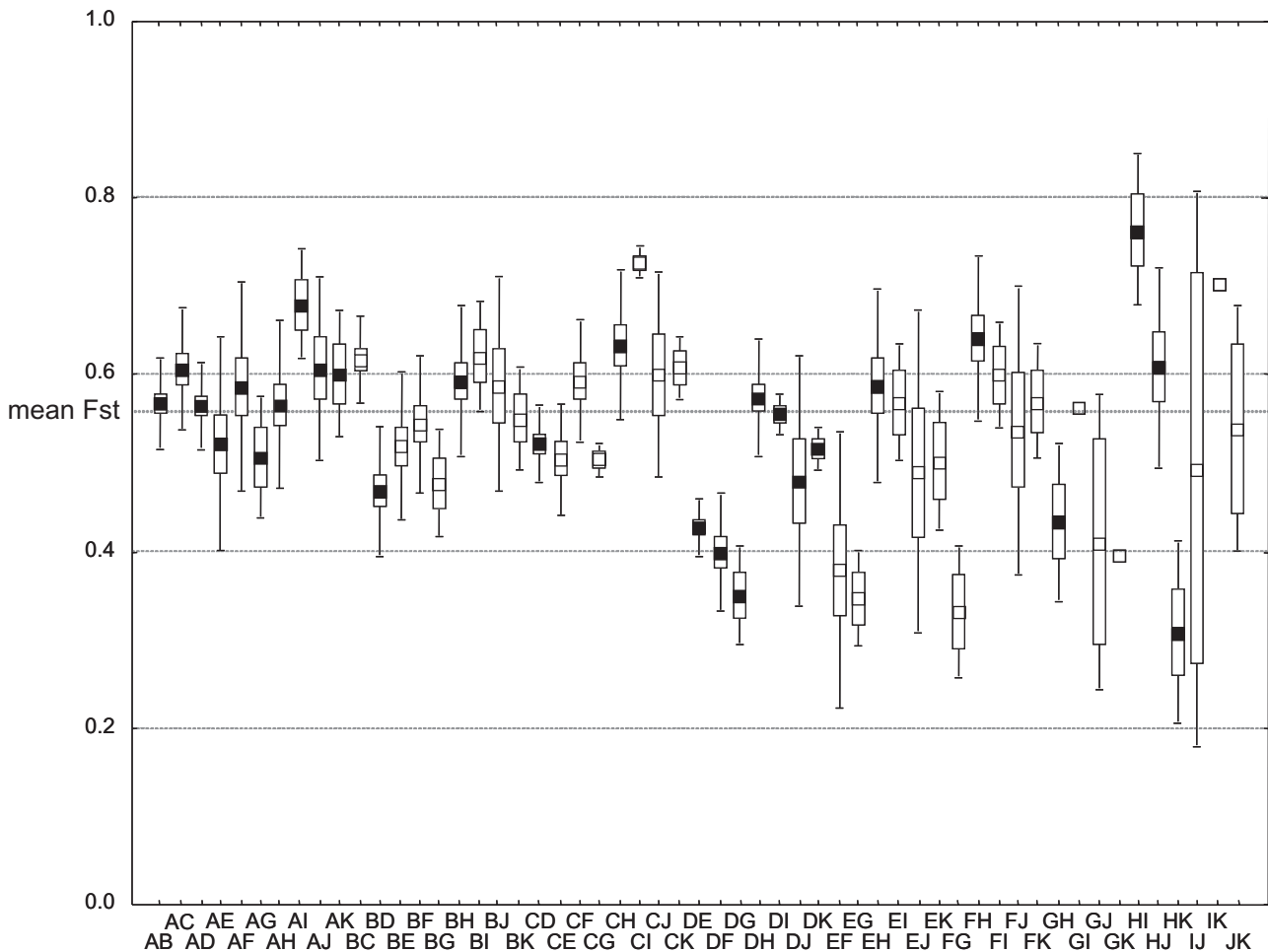
either type,  $F_1$  or  $F_2$  hybrids and backcrosses in both directions).

RESULTS

POPULATION-BASED GENETIC ANALYSIS

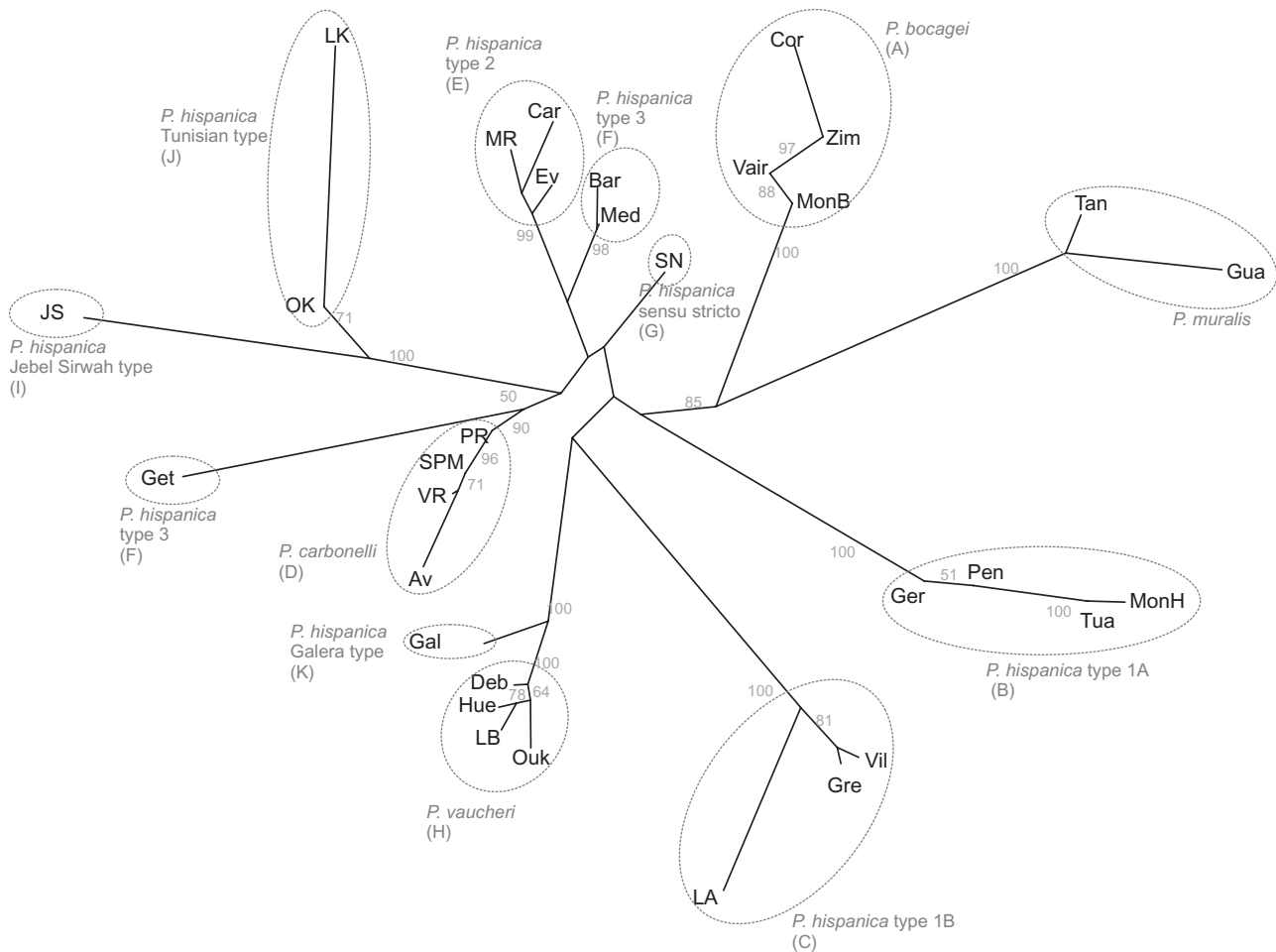
Including previously published results, a total of 89 alleles were detected among the ten studied loci. The most polymorphic locus was *PGD*, with 18 alleles detected, whereas *IDH*, with four alleles, was the least polymorphic. In addition to the cases of Hardy–Weinberg disequilibrium already reported in previous studies (locus *PEPD* in Vairão,  $N = 1$ ; Monte Real,  $N = 16$ ; and Oukaïmeden,  $N = 26$ ), significant departures from equilibrium were detected for *GOT* in La Alberca ( $N = 10$ ) and for *PEPB* in Gredos ( $N = 11$ ) ( $P < 0.05$ ). However, when considering a significance

threshold of 0.01, only one case of disequilibrium (*PEPD* in Monte Real) is maintained. No significant departure from linkage equilibrium was observed for any pair of loci. Pairwise  $F_{st}$  values between populations are presented in Table 2. These values vary from a minimum of 0.001, not significantly different from 0, between the populations of *P. hispanica* type 1A of Montesinho and Tua, to a maximum of 0.888 between *P. vaucheri* from Oukaïmeden and *P. muralis* from Guadarrama. Figure 2 represents a plot of the mean, standard error and standard deviation of  $F_{st}$  values between the 55 pairs of evolutionary groups (the outgroup *P. muralis* was excluded from this analysis). The results of the AMOVA show that the largest component of the total variance (45.79%) is due to differences among mtDNA-defined groups/species. A relatively small portion (9.05%) is found among



**Figure 2.** Box plot representation of mean  $F_{st}$  values observed between pairs of forms of *Podarcis*, including comparisons between recognized species (filled squares) and groups within *Podarcis hispanica* (white squares). Pairs of letters indicated on the x axis correspond to the forms being compared, using the same nomenclature shown in Table 1 and Figure 1. The boxes represent standard errors and whiskers indicate standard deviations. The dotted line across the graph shows the mean  $F_{st}$  value for the whole sample. *Podarcis muralis* was excluded from this analysis.





**Figure 3.** NJ tree showing the relationships between 32 populations of Iberian and North African *Podarcis* using the genetic distances of Nei (1972) based on ten allozyme loci. Letters identifying mitochondrial DNA lineages correspond to those indicated in Figure 1 and Table 1. Bootstrap values above 50% are shown.

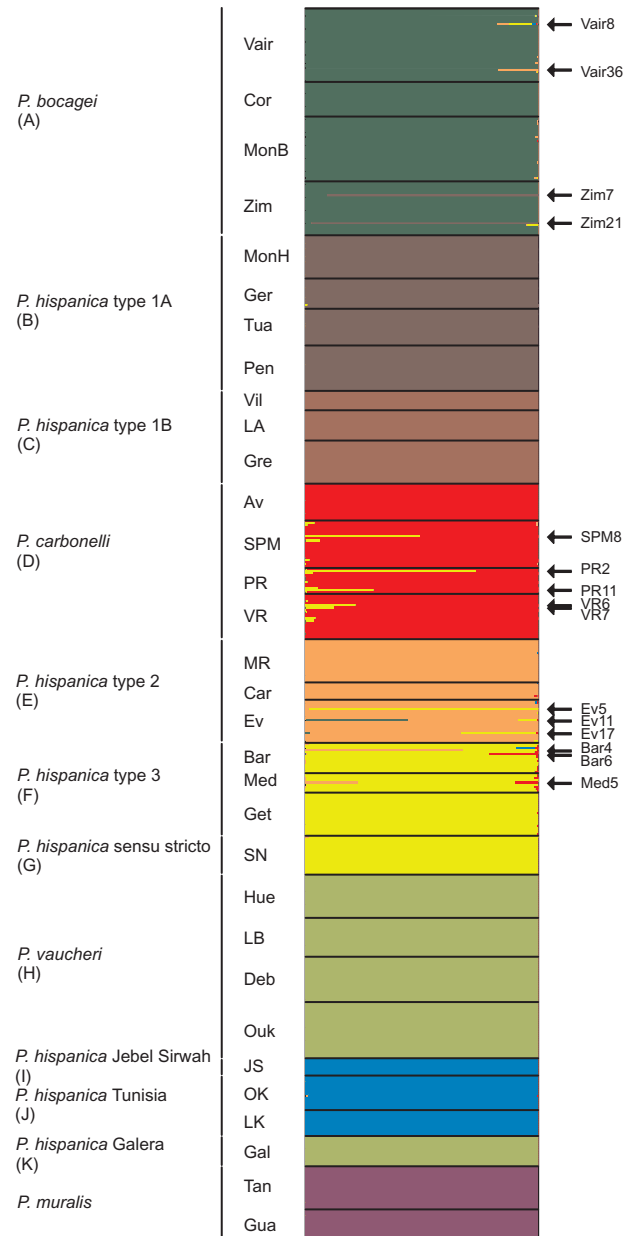
populations of the same group and, again, a high fraction of the total variance (45.15%) comes from differences within populations. Translated into fixation indexes, these values correspond to a  $\phi_{CT}$  of 0.45793, a  $\phi_{SC}$  of 0.16703, and a  $\phi_{ST}$  of 0.54854. A NJ tree showing the relationships between the studied populations, built using Nei's standard distances, is shown in Figure 3. In this tree, populations bearing the same mtDNA type also cluster together based on these nuclear markers. There is only one exception, *P. hispanica* type 3, because the population of Getaria is not grouped with the other two populations bearing its mtDNA lineage (Barcelona and Medinaceli). Bootstrap values are generally high for the clustering of populations from the same mtDNA type but support is low for relationships between forms. Exceptions are the clustering of Galera with populations of *P. vaucheri* and of Jebel Sirwah with the Tunisian populations, both with 100% bootstrap support.

#### INDIVIDUAL MULTILOCUS GENOTYPE ANALYSES

Considering the results obtained using STRUCTURE, the choice of the appropriate scenario for the data was made difficult by two particularities of the results: (1) increasingly high log probabilities of the data with increasing  $K$ , even after scenarios that are not biologically realistic were reached, and (2) inconsistencies within the same value of  $K$  (i.e. even considering the same number of assumed populations, runs differed with respect to the grouping of individuals into clusters). For example, at  $K = 3$ , individuals of *P. bocagei* were either placed in the same cluster as individuals of *P. muralis*, *P. vaucheri* + *P. hispanica* Galera type or *P. hispanica* type 1A. To try to solve the first issue, the approach described by Evanno *et al.* (2005) was used. This method searches for a mode in the distribution of  $\Delta K$ , a quantity related to the second order rate of change of the log probability of the data. However, this

procedure was inconclusive because it did not provide any obvious mode (results not shown). Although the highest value of  $\Delta K$  is found at  $K = 2$ , it is clear from the analyses that there are clusters that are biologically more meaningful with a higher partition of the data (e.g.  $K = 9$  or even  $K = 10$ ). Due to the statistical impossibility of choosing one amongst all the results, we focused on a particular scenario, obtained in four different runs at  $K = 9$ , that summarizes the most relevant aspects of the individual multilocus genotype analyses (Fig. 4). Here, all the individuals of *P. hispanica* type 1A, *P. hispanica* type 1B, and *P. muralis* are placed in a cluster of their own with probabilities higher than 90%. The same happens with all but four individuals (3.8%) of *P. bocagei*, all but five individuals (6.9%) of *P. carbonelli*, and all but three individuals (6.3%) of *P. hispanica* type 2. The remaining three clusters include individuals belonging to two distinct mtDNA lineages: all but three individuals (7.0%) of *P. hispanica* type 3 with *P. hispanica* s.s., *P. vaucheri* with *P. hispanica* Galera type and the *P. hispanica* forms from Tunisia and Jebel Sirwah. It is noteworthy that across all runs, irrespectively of the value of  $K$  and of the clusters found, individuals very rarely show probabilities smaller than 0.95 of belonging to any of the defined groups. Likewise, across all possible scenarios, individuals belonging to the same mtDNA lineage are rarely placed in separate clusters. This only happens at higher values of  $K$ , where the population of Getaria separates from its partition and there is a tendency to distinguish between *P. vaucheri* from Spain and from Morocco. Unrealistic scenarios (e.g. dividing clusters in two units but with individuals having approximately 50% probability of belonging to each) are found from  $K = 8$  onward, and become increasingly frequent at higher values of  $K$ .

Taking the chosen scenario into consideration ( $K = 9$ ), 15 individuals, highlighted in Figure 4, had less than 90% probability of belonging to their respective clade and were therefore considered to be admixed or misassigned. Considering the possibility of present, detectable gene flow, extra analyses that involved only *P. hispanica* type 1A and *P. bocagei* were performed because two misassigned individuals were detected in the population of Zimão, where these two species exist in sympatry. The genome of the two individuals (Zim7 and Zim21) with discordant ancestry in previous analyses were now attributed to both species in high proportion (0.53 *P. bocagei*/0.47 *P. hispanica* type 1A in Zim7 and 0.33 *P. bocagei*/0.67 *P. hispanica* type 1A in Zim21). The use of NEWHYBRIDS on this partial data set proved to be inconclusive. Besides these two individuals, all other specimens were correctly identified as 'pure *P. bocagei*' or 'pure *P. hispanica* type 1A' with over 90% posterior probability. However, Zim7 and Zim21 could not be



**Figure 4.** Estimated probability of ancestry of 569 individuals belonging to the Iberian and North African *Podarcis* species complex and to the outgroup *Podarcis muralis*, calculated using the software STRUCTURE, considering  $K = 9$  clusters. Each individual is represented by a horizontal line divided into nine segments of different colours, each representing a cluster. The size of the segments is proportional to the individual's estimated probability of belonging to each of the  $K = 9$  clusters. Misassigned or apparently admixed individuals are highlighted. Letters correspond to those indicated in Figure 1 and Table 1.



unambiguously assigned to any class; instead, their posterior probabilities are distributed across the four hybrid classes with a maximum of 0.38 for  $F_2$  (Zim7) and of 0.31 for  $F_1$  (Zim21). Both individuals show posterior probabilities lower than 5% of being either non-admixed *P. bocagei* or *P. hispanica* type 1A. Both individuals carry *P. bocagei* mtDNA.

## DISCUSSION

### GENETIC SUBDIVISION OF IBERIAN AND NORTH AFRICAN *PODARCIS*

The results obtained in this study largely corroborate the subdivisions already reported for mitochondrial DNA and morphology, both at the population-level and the individual-level analyses. Although reciprocal monophyly could not be evaluated in three of the studied taxa because of the inclusion of a single population (in the cases of the types from Galera and Jebel Sirwah, this was due to the fact that these two forms are not known from any other locality), monophyly was observed in all but one of the remnant forms, with near 100% bootstrap support for the grouping of populations into mtDNA-defined species. This differentiation is further supported by AMOVA, which suggests that the largest proportion of the total variance found within our data set is due to differences among these groups, whereas populations within the same group are highly homogenous. In concordance with these results, the analyses based on the individual multilocus genotypes yielded the same pattern of subdivision because six out of the 12 species/forms included are clearly identifiable and three other groups, each comprising two forms, are also observed.

On the other hand, the polyphyly of *P. hispanica* type 3 contrasted to the results from mtDNA phylogenetic analyses. Bootstrap values that support this polyphyly are very low and it could thus be an analytical artefact. Nevertheless, this discordance could also be due to the fact that the population of Getaria, which does not cluster with its conspecifics, was collected from a former island, recently connected to the mainland. Founder events associated to the colonization of islands are known to dramatically affect allele frequencies and small islands such as Getaria might be prone to rapid changes in effective population sizes. However, in the individual multilocus analyses, this population appears most often grouped with its conspecifics, only detaching from them at higher values of  $K$ .

Even acknowledging that there are some differences between the partitions defined on the basis of mitochondrial DNA and of allozyme analyses, taken together, these results constitute additional evidence

supporting the division of endemic Iberian and North African *Podarcis* into as many as 11 distinguishable genetic entities. These correspond both to acknowledged species such as *P. bocagei* or *P. carbonelli* and also to different partitions within *P. hispanica*, thus supporting previous observations on the basis of mitochondrial DNA that this taxon constitutes a cryptic species complex (Harris & Sá-Sousa, 2001, 2002; Harris *et al.*, 2002b; Pinho *et al.*, 2006). According to our data set, the genetic differentiation between the various forms of *P. hispanica*, measured by  $F_{st}$  values, falls within the same order of magnitude of those found between fully recognized species within this complex (Fig. 2), thus corroborating the idea that a taxonomical revision is needed. At a larger scale, this constitutes a validation, based on nuclear markers, of the biogeographical theory that postulates that the Iberian Peninsula functioned as a hotspot of diversification and not only as a glacial refugium for many taxa (Gómez & Lunt, 2007).

### EVOLUTIONARY RELATIONSHIPS BETWEEN DISTINCT FORMS

In a recent study based on mitochondrial DNA variation (Pinho *et al.*, 2006), the evolutionary relationships between forms of *Podarcis* and consequent biogeographical inferences are very well supported by long internal branches and bootstrap values close to 100%. In the present analyses, however, only a minority of the relationships between forms are well supported. These are the groupings of the population of Jebel Sirwah with those from Tunisia and of the population of Galera with *P. vaucheri*, both in the NJ estimates of relationships and in the analyses of individual multilocus genotypes ( $K = 9$ ), and the clustering of *P. hispanica* type 3 with *P. hispanica s.s.*, which was not observed in the NJ tree but was consistent across multiple STRUCTURE runs. In the first case, the clustering of the two forms is clearly concordant with inferences derived from mtDNA and thus corroborates the hypothesis that the forms from Tunisia and Jebel Sirwah correspond to relics of a once more widespread North African taxon that was eventually split and confined to two distant allopatric units (Harris *et al.*, 2002b). The remaining two cases are more difficult to interpret. They constitute an obvious contradiction to mtDNA estimates because each pair is formed by species that do not share close ancestors, and could result from ongoing or past gene flow between the involved forms, not detected in mtDNA analyses because few individuals were included. However, these unexpected groups could also represent evolutionary meaningful clusters that were not recovered by mitochondrial DNA analyses, or simply the studied set of markers may not be appropriate for

discriminating between these forms. Therefore, this subject needs further assessment with the study of nuclear genealogies.

Excluding the above-mentioned cases, the NJ tree provides virtually no information on the relationships between forms. This is also illustrated by the multi-locus genotype analyses, where, especially at lower values of  $K$ , many distinct but equally likely groupings of species were produced across different runs. Although we do not have genealogical data to understand the evolutionary relationships between the alleles, a potential explanation for the reported lack of information on relationships between forms could be differential lineage sorting across loci (i.e. different loci portraying distinct scenarios of the evolution of the group). This situation can easily be observed by exploring patterns of allele sharing (a frequency table is provided in Pinho *et al.*, 2003). Such cases have been well documented in closely-related *Drosophila* species by Machado *et al.* (2002) and Machado & Hey (2003), where several different genealogies were obtained using 16 independent loci. The authors note that simple bifurcating trees may thus not be suitable for describing the relationships between species that have undergone recent divergence because of the cumulative effects of differential lineage sorting of ancestral polymorphism and of introgressive hybridization that may affect distinct loci in different ways due to stochasticity or distinct selective pressures. Tracing a parallel with the present case study and taking into consideration that the calculation of a measure of genetic distance involves averaging the differentiation across loci, this situation would explain the poor resolution of the estimates of relationships between forms of *Podarcis* and the very short internal branches of the NJ tree. The pattern described is concordant with a scenario of a rapid diversification, in which all forms differentiated during a short period of time. At a first glance, this scenario would appear to be in contradiction with the clear bifurcating model assumed from the mtDNA phylogeny (Pinho *et al.*, 2006), but both observations are easily accommodated taking into account the lower effective population size of mtDNA. This hypothesis of a rapid diversification appears to be corroborated by a preliminary study of nuclear genealogies in these forms (C. Pinho, D. J. Harris & N. Ferrand, unpubl. data). Diversification events involving the formation of as many distinct entities within Iberia do not appear to be a common pattern across the widely-studied species of the Iberian herpetofauna. Nevertheless, cases such as this have been reported in other groups of organisms, such as Iberian endemic barbel fish (Callejas & Ochando, 2000; Machordom & Doadrio, 2001) or some groups of Iberian diving beetles (Ribera & Vogler, 2004).

#### HYBRIDIZATION BETWEEN FORMS OF *PODARCIS*

Considering the individual multilocus analyses, 15 individuals out of 569 were misassigned or admixed when assuming a probability threshold of 90%. There are two situations that could cause such results. An obvious one is hybridization between forms. However, when many alleles are shared between species, it is possible that apparently admixed multilocus genotypes are produced without this being a result of introgression. This is probably the case in the majority of the individuals reported. Most loci are only partially diagnostic and more than one-half of the alleles detected (46; 51.6%) are *trans*-specific. When dealing with allozyme data, there is always the inherent possibility of a lack of separation between distinct alleles that exhibit similar net charges (electromorphs; Barbadilla, King & Lewontin, 1996). This would therefore be a plausible explanation for the sharing of some alleles. However, it is also likely that the *trans*-specific nature of many alleles is a consequence of incipient speciation and that abundant ancestral polymorphisms persist across forms. This is also supported by observations on nuclear gene genealogies (C. Pinho, D. J. Harris & N. Ferrand, unpubl. data). Therefore, both of these situations most likely explain why some individuals appear to be admixed between allopatric forms that do not seem to contact and are therefore unable to exchange genes.

However, this is not the case for *P. bocagei* individuals Zim7 and Zim21. The general analyses assigned these individuals to *P. hispanica* type 1A and posterior analyses, considering only these two forms, suggested that the two individuals have approximately equal proportions of their genome originating from each of the two forms. The presence of alleles from both species on these individuals is not explainable by the hypothesis of an unsatisfactory separation of alleles or of ancestral polymorphism as in the above-mentioned cases because, in many of the most informative loci, these forms show significant differences in allele frequencies or even fixation for diagnostic alleles. Considering that these two forms are sympatric, it is therefore likely that these individuals are a product of hybridization between both species. There are few known cases of hybridization between different species of *Podarcis*. Based on genetic studies, it has been reported in Italian species *Podarcis wagleriana* and *Podarcis sicula* (Capula, 1993) and between the latter and *Podarcis tiliguerta* (Capula, 2002) and has also been shown to be possible in captivity between the Iberian *P. bocagei* and *P. carbonelli* (Galán, 2002). However, this is the first time that natural hybridization has been reported between Iberian species. In the present study, we were unable to assign these two individuals to a hybrid class, either because this would

require a higher number of markers or because these two individuals are the products of more than one backcross and do not fall in any of the a priori categories. An evaluation of how frequently hybridization occurs between these forms and of the degree of selection against hybrids requires further studies aiming to address these specific questions. Nevertheless, because *P. bocagei* and *P. hispanica* type 1A exist in sympatry across most of their distribution area and appear to maintain genetic integrity, as well as morphological identifiability, in the presence of the other species, it is likely that gene flow between the two species is limited.

Despite this preliminary evidence suggesting a low degree of gene flow between forms, our sampling scheme, biased towards the centre of the distributions, did not allow us to evaluate whether allopatric forms are exchanging genes in the areas where they meet. These questions will only be properly analysed by detecting and thoroughly studying the various suture zones within the Iberian Peninsula and North Africa.

#### CONCLUSIONS

The present data set supports the existence of multiple cryptic forms of *Podarcis* in the Iberian Peninsula and North Africa, and the partitions observed are highly concordant with previous mitochondrial DNA and morphological analyses. Moreover, the results presented evoke a scenario of a rapid diversification. Hybridization was clearly observed between two fully recognized species but appears to be a rare event. Nevertheless, these taxa have not yet become fully reproductively isolated and still share an important proportion of alleles, having probably not achieved complete monophyly in most nuclear markers, meaning that they do not fulfil the criteria imposed by some of the most important and applied species concepts such as the biological, in its strictest form (Mayr, 1963), or the genealogical (Baum & Shaw, 1995). This is particularly relevant if one takes into account that these forms probably evolved in allopatry for a long time, as suggested by sequence divergences of 8–12% found in mitochondrial genes such as cytochrome *b* or ND4. These values are several-fold higher than the boundaries traditionally accepted by phylogeneticists to define species in squamates (e.g. Hasbún *et al.*, 2005; divergences of 2–5.4% on the basis of species recognition).

Our results highlight the importance of evaluating multiple independent data sources prior to defining taxonomic units, and in particular the difficulties of determining species boundaries in this complicated species complex. *Podarcis* may therefore be a useful model for studying genetic and morphological diversification within emerging species.

#### ACKNOWLEDGEMENTS

This work was partially financed by FCT (Fundação para a Ciência e a Tecnologia) research projects POCTI/40987/BSE/2001 and POCTI/41912/BSE/2001. C. Pinho is supported by a PhD grant (SFRH/BD/4620/2001) and D. J. Harris by a postdoctoral grant (SFRH/BPD/5702/2001), both from FCT. The authors wish to thank Miguel A. Carretero, Paulo Sá-Sousa, Diana Barbosa, Vasco Batista, Elsa Froufe, Raquel Godinho, Bruno Ribeiro, Mário J. Pinho, Mohammed Znari, Jean-Christophe Mace, and Xavier Santos for their help in sample collection. Armando Gerales, Fernando Sequeira, Ricardo Pereira, and Rui Faria provided helpful comments and advice regarding the discussion of the data. Two anonymous reviewers are also acknowledged for their constructive comments on an earlier version of this article.

#### REFERENCES

- Alves PC, Ferrand N, Suchentrunk F, Harris DJ. 2003. Ancient introgression of *Lepus timidus* mtDNA into *L. granatensis* and *L. europaeus* in the Iberian Peninsula. *Molecular Phylogenetics and Evolution* **27**: 70–80.
- Anderson EC, Thompson EA. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**: 1217–1229.
- Arnold EN, Burton JA. 1978. *A field guide to the reptiles and amphibians of Britain and Europe*. London: Collins.
- Ballard JWO, Chernoff B, James AC. 2002. Divergence of mitochondrial DNA is not corroborated by nuclear DNA, morphology, or behavior in *Drosophila simulans*. *Evolution* **56**: 553.
- Ballard JWO, Whitlock MC. 2004. The incomplete natural history of the mitochondria. *Molecular Ecology* **13**: 729–744.
- Banks MA, Eichert W. 2000. WHICHRUN (version 3.2): a computer program for population assignment of individuals based on multilocus genotype data. *Journal of Heredity* **91**: 87–89.
- Barbadilla A, King LM, Lewontin RC. 1996. What does electrophoretic variation tell us about protein variation? *Molecular Biology and Evolution* **13**: 427–432.
- Barbadillo LJ, Lacomba JI, Pérez Mellado V, Sancho V, López-Jurado LF. 1999. *La guía de campo de los anfibios y reptiles de la Península Ibérica, Baleares Y Canarias*. Barcelona: Ed Planeta.
- Baum D, Shaw KL. 1995. Genealogical perspectives on the species problem. In: Hoch PC, Stephenson AC, eds. *Experimental and molecular approaches to plant biosystematics*. St Louis, MO: Missouri Botanical Garden, 289–303.
- Busack SD, Lawson R, Arjo WM. 2005. Mitochondrial DNA, allozymes, morphology and historical biogeography in the *Podarcis vaucheri* (Lacertidae) species complex. *Amphibia-Reptilia* **26**: 239–256.
- Callejas C, Ochando MD. 2000. Recent radiation of Iberian Barbel fish (Teleostei: Cyprinidae) inferred from cytochrome *b* genes. *Journal of Heredity* **91**: 283–288.

- Capula M. 1993.** Natural hybridization in *Podarcis sicula* and *P. wagleriana* (Reptilia: Lacertidae). *Biochemical Systematics and Ecology* **3**: 373–380.
- Capula M. 2002.** Genetic evidence of natural hybridization between *Podarcis sicula* and *Podarcis tiliguerta* (Reptilia: Lacertidae). *Amphibia-Reptilia* **23**: 313–321.
- Chan KMA, Levin SA. 2005.** Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution* **59**: 720–729.
- Corander J, Waldmann P, Sillanpää MJ. 2003.** Bayesian analysis of genetic differentiation between populations. *Genetics* **163**: 367–374.
- Dawson KJ, Belkhir K. 2001.** A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genome Research* **78**: 59–77.
- DeSalle R, Giddings LV. 1986.** Discordance of nuclear and mitochondrial phylogenies in Hawaiian *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* **83**: 6902–6906.
- Evanno G, Regnaut S, Goudet J. 2005.** Detecting the number of clusters of individuals using software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.
- Excoffier L, Smouse PE, Quattro JM. 1992.** Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.
- Fay JC, Wu C-I. 1999.** A human population bottleneck can account for the discordance between patterns of mitochondrial versus nuclear DNA variation. *Molecular Biology and Evolution* **16**: 1003–1005.
- Felsenstein J. 1993.** *PHYLIP (phylogeny inference package)*, Version 3.5. Seattle, WA: University of Washington.
- FitzSimmons N, Moritz C, Limpus CJ, Pope L, Prince R. 1997.** Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. *Genetics* **147**: 1843–1854.
- Galán P. 2002.** Hibridación en laboratorio de *Podarcis bocagei* y *Podarcis carbonelli*. *Boletín de la Asociación Herpetológica Española* **13**: 28–31.
- Geniez P. 2001.** Variation géographique des lézards du genre *Podarcis* (Reptilia, Sauria, Lacertidae) dans la péninsule Ibérique, l'Afrique du Nord et le sud de la France. Diplôme de l'École Pratique des Hautes Etudes, Montpellier.
- Gómez A, Lunt DH. 2007.** Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S, Ferrand N, eds. *Phylogeography in southern European refugia: evolutionary perspectives on the origin and conservation of European biodiversity*. Dordrecht: Springer, 155–188.
- Harris DJ, Arnold EN. 1999.** Relationships and evolution of wall lizards *Podarcis* (Reptilia: Lacertidae) based on mitochondrial DNA sequences. *Copeia* **3**: 749–754.
- Harris DJ, Batista V, Carretero MA, Pinho C, Sá-Sousa P. 2002a.** Mitochondrial DNA sequence data confirms the presence of *Podarcis carbonelli*, PÉREZ-MELLADO, 1981 in southern Spain. *Herpetozoa* **15**: 188–190.
- Harris DJ, Carranza S, Arnold EN, Pinho C, Ferrand N. 2002b.** Complex biogeographical distribution of genetic variation within *Podarcis* wall lizards across the Strait of Gibraltar. *Journal of Biogeography* **29**: 1257–1252.
- Harris DJ, Pinho C, Carretero MA, Corti C, Böhme W. 2005.** Determination of genetic diversity within the insular lizard *Podarcis tiliguerta* using mtDNA sequence data, with a reassessment of the phylogeny of *Podarcis*. *Amphibia-Reptilia* **26**: 401–407.
- Harris DJ, Sá-Sousa P. 2001.** Species distinction and relationships of the western Iberian *Podarcis* lizards (Reptilia, Lacertidae) based on morphology and mitochondrial DNA sequences. *Herpetological Journal* **11**: 129–136.
- Harris DJ, Sá-Sousa P. 2002.** Molecular phylogenetics of Iberian wall lizards (*Podarcis*): is *Podarcis hispanica* a species complex? *Molecular Phylogenetics and Evolution* **23**: 75–81.
- Hasbún CR, Gómez A, Köhler G, Lunt DH. 2005.** Mitochondrial DNA phylogeography of the Mesoamerican spiny-tailed lizards (*Ctenosaura quinquecarinata* complex): historical biogeography, species status and conservation. *Molecular Ecology* **14**: 3095–3107.
- Hewitt GM. 1996.** Some genetic consequences of the ice ages and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**: 247–266.
- Hewitt GM. 1999.** Postglacial recolonization of the European biota. *Biological Journal of the Linnean Society* **68**: 87–612.
- Hewitt GM. 2004.** Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* **359**: 183–195.
- Hey J. 1997.** Mitochondrial and nuclear genes present conflicting portraits of human origins. *Molecular Biology and Evolution* **14**: 166–172.
- Hudson RR, Turelli M. 2003.** Stochasticity overrules the 'three-times rule': genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution* **57**: 182–190.
- Klemmer K. 1959.** Systematische Stellung und Rassen-gliederung der spanischen Mauereidechse, *Lacerta hispanica*. *Senckenbergiana Biologica* **40**: 245–250.
- Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC. 1989.** Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America* **86**: 6196–6200.
- Machado CA, Hey J. 2003.** The causes of phylogenetic conflict in a classic *Drosophila* species group. *Proceedings of the Royal Society of London Series B, Biological Sciences* **270**: 1193–1202.
- Machado CA, Kliman RM, Markert JA, Hey J. 2002.** Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudobscura* and close relatives. *Molecular Biology and Evolution* **19**: 472–488.
- Machordom A, Doadrio I. 2001.** Evidence of a cenozoic Betic-Kabilian connection based on freshwater fish phylogeography (*Luciobarbus*, Cyprinidae). *Molecular Phylogenetics and Evolution* **18**: 252–263.
- Mayr E. 1963.** *Animal species and evolution*. Cambridge, MA: Harvard University Press.

- Mertens R, Müller L. 1940.** Die amphibien und reptilien Europas. *Abh Senck Naturfor Gesells Frankfurt* **41**: 1–62.
- Monsen KJ, Blouin MS. 2003.** Genetic structure in a montane ranid frog: restricted gene-flow and nuclear-mitochondrial discordance. *Molecular Ecology* **12**: 3275–3286.
- Moore WS. 1995.** Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49**: 718–726.
- Nei M. 1972.** Genetic distance between populations. *American Naturalist* **106**: 283–292.
- Nyakaana S, Arctander P. 1999.** Population genetic structure of the African elephant in Uganda based on variation at mitochondrial and nuclear loci: evidence for male-biased gene flow. *Molecular Ecology* **8**: 1105–1115.
- Oliverio M, Bologna M, Mariottini P. 2000.** Molecular biogeography of the Mediterranean lizards *Podarcis* Wagler 1830 and *Teira* Gray, 1838 (Reptilia, Lacertidae). *Journal of Biogeography* **27**: 1403–1420.
- Pamilo P, Nei M. 1988.** Relationships between gene trees and species trees. *Molecular Biology and Evolution* **5**: 568–583.
- Piertney SB, MacColl ADC, Bacon PJ, Racey PA, Lambin X, Dallas JF. 2000.** Matrilineal genetic structure and female-mediated gene flow in red grouse (*Lagopus lagopus scoticus*): an analysis using mitochondrial DNA. *Evolution* **54**: 279–289.
- Pinho C, Ferrand N, Harris DJ. 2004.** Genetic variation within the *Podarcis hispanica* species complex: new evidence from protein electrophoretic data. In: Pérez-Mellado V, Riera N, Perera A, eds. *The biology of lacertid lizards: evolutionary and ecological perspectives*. Menorca: Institut Menorquí d'Estudis (Recerca 8), 269–277.
- Pinho C, Ferrand N, Harris DJ. 2006.** Reexamination of the Iberian and North African *Podarcis* (Squamata: Lacertidae) phylogeny based on increased mitochondrial DNA sequencing. *Molecular Phylogenetics and Evolution* **38**: 266–273.
- Pinho C, Harris DJ, Ferrand N. 2003.** Genetic polymorphism of 11 allozyme loci in populations of wall lizards (*Podarcis* sp.) from the Iberian Peninsula and North Africa. *Biochemical Genetics* **41**: 343–359.
- Podnar M, Mayer W, Tvrtković N. 2005.** Phylogeography of the Italian wall lizard, *Podarcis sicula*, as revealed by mitochondrial DNA sequences. *Molecular Ecology* **14**: 575–588.
- Poulakakis N, Lymberakis P, Antoniou A, Chalkia D, Zouros E, Mylonas M, Valakos E. 2003.** Molecular phylogeny and biogeography of the wall-lizard *Podarcis erhardii* (Squamata: Lacertidae). *Molecular Phylogenetics and Evolution* **28**: 38–46.
- Poulakakis N, Lymberakis P, Valakos E, Pafilis P, Zouros E, Mylonas M. 2005a.** Phylogeography of Balkan wall lizard (*Podarcis taurica*) and its relatives inferred from mitochondrial DNA sequences. *Molecular Ecology* **14**: 2433–2443.
- Poulakakis N, Lymberakis P, Valakos E, Zouros E, Mylonas M. 2005b.** Phylogenetic relationships and biogeography of *Podarcis* species from the Balkan Peninsula, by bayesian and maximum-likelihood analyses of mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* **37**: 845–857.
- Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
- Raymond M, Rousset F. 1995.** GENEPOP (version 3.1): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**: 248.
- Ribera I, Vogler AP. 2004.** Speciation of Iberian diving beetles in Pleistocene refugia (Coleoptera, Dysticidae). *Molecular Ecology* **13**: 179–193.
- Rosenberg NA. 2002.** *Distruct: a program for the graphical display of structure results*. Available at <http://www.cmb.usc.edu/~noahr/distruct.html>.
- Saitou N, Nei M. 1987.** The Neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**: 406–427.
- Sá-Sousa P, Harris DJ. 2002.** *Podarcis carbonelli* is a distinct species. *Amphibia-Reptilia* **23**: 459–468.
- Sá-Sousa P, Vicente L, Crespo EG. 2002.** Morphological variability of *Podarcis hispanica* (Sauria: Lacertidae) in Portugal. *Amphibia-Reptilia* **23**: 55–70.
- Schneider S, Roessli D, Excoffier L. 2000.** *Arlequin, ver. 2.000: a software for population genetics analyses*. Geneva: Genetics and Biometry Laboratory, University of Geneva.
- Shaw K. 2002.** Conflict between nuclear and mitochondrial DNA phylogeny of recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 16122–16127.
- Zhang DX, Hewitt GM. 2003.** Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology* **12**: 563–584.