MOLECULAR ECOLOGY

Molecular Ecology (2009) doi: 10.1111/j.1365-294X.2009.04351.x

Morphology and nuclear markers reveal extensive mitochondrial introgressions in the Iberian Wall Lizard species complex

JULIEN P. RENOULT,* PHILIPPE GENIEZ,† PAUL BACQUET,‡ LAURE BENOIT§ and PIERRE-ANDRÉ CROCHET*

**CEFE-CNRS, UMR 5175, 1919 route de Mende, F-34293 Montpellier cedex 5, France, †CEFE-EPHE, UMR 5175, 1919 route de Mende, F-34293 Montpellier cedex 5, France, ‡Centre de Recherche BDIV, groupe Evolutionary Biology, Université Catholique de Louvain, Croix du sud, 4-5, Carnoy, b.108, B-1348, Louvain-la-Neuve, Belgium, §CEFE-CIRAD, UMR 5175, 1919 route de Mende, F-34293 Montpellier cedex 5, France*

Abstract

Mitochondrial markers are still often used alone to identify evolutionary units, despite widespread evidence for processes such as incomplete lineage sorting or introgressive hybridization that may blur past population history. The combination of mitochondrial DNA data with other sources of information (morphology, nuclear genes) is a powerful tool to reveal when and why mitochondrial markers are potentially misleading. In this study, we evaluate the performance of mtDNA markers to unravel the evolutionary history of Spanish lizards from the *Podarcis hispanicus* species complex. We first uncover several cases of discordance between morphological and mitochondrial data in delimitation of taxa. To assess the origin of these discordances, we analysed the same populations using several independent nuclear loci. Both morphological and nuclear markers identified the same three evolutionary units in the region, while mitochondrial data revealed four deeply divergent lineages. We suggest here that the most likely scenario to explain this discordance is ancient mitochondrial introgression originating from a fourth evolutionary unit presently absent from the study area. Notably, this resulted in a complete replacement of the original lineage in a large part of the distribution of one of the taxa investigated. We discuss the potential evolutionary scenarios leading to this complete mitochondrial replacement and suggest why the previous studies have failed to recover the correct history of this species complex.

Keywords: discordance, introgression, morphology, phylogeography, *Podarcis hispanicus*

Received 10 March 2009; revision received 2 August 2009; accepted 11 August 2009

Introduction

An accurate description of biological diversity is a prerequisite for understanding its origin. Systematic studies thus lead to direct applications in fields as diverse as medical epidemiology (Biek *et al.* 2006), theoretical ecology (Anderson *et al.* 2004) and developmental biology (Minelli 2007) and provides the basic information for many conservation programmes (Vazquez & Gittleman 1998; Bininda-Emonds *et al.* 2000; Andreasen 2005).

A necessary step in such studies is to identify and delimit species. This task is sometimes wearisome in practice, mainly because markers used to identify evolutionary units may also be subject to selection, and therefore are intrinsically inadequate for inferring population histories (Avise 2004). For example, use of morphological variation for inferring relationships among populations often leads to errors because phenotypic variation reflects both population history and local adaptation or phenotypic plasticity (Avise 1994; Parra-Olea & Wake 2001). The use of neutral molecular

Correspondence: Julien P. Renoult and Pierre-André Crochet, Fax: +33 4672138; E-mails: julien.renoult@cefe.cnrs.fr and pierre-andre.crochet@cefe.cnrs.fr

markers should thus generally provide more accurate descriptions of population history than morphological characters alone (Avise 1994), hence molecular markers are increasingly used to identify and delimit species.

Until recently, most studies in molecular systematics and phylogenetics have used mitochondrial DNA (mtDNA; see Sunnucks 2000 for a review of the advantages of mtDNA). However, because gene trees do not necessarily reflect population history (Nichols 2001), the use of mtDNA alone in evolutionary studies has been criticized (Ballard & Whitlock 2004; Edwards *et al.* 2005). Well-known causes of discordance between mitochondrial and population histories include incomplete lineage sorting (Pamilo & Nei 1988; Jennings & Edwards 2005), which is especially problematic when multiple nodes in the tree are generated over a short period of time (Moran & Kornfield 1993; Klein & Payne 1998), and introgressive hybridization, i.e. exchange of genes between two evolutionary lineages (Seehausen 2004). Introgressive hybridization of mtDNA may have long-lasting consequences, particularly as the nonrecombining mtDNA genome never becomes 'diluted' in the introgressed genetic background (Funk & Omland 2003; Morando *et al.* 2004). Moreover, introgression might be facilitated because mtDNA loci are less often linked to loci under negative selective pressure than nuclear loci (Harrison 1990) and can even undergo positive selection (i.e. Ruiz-Pesini *et al.* 2004). Intraspecific polymorphism in mtDNA lineages as a result of partial replacement of the original mtDNA by allospecific mtDNA has been documented in several organisms (e.g. Sullivan *et al.* 2004; Roca *et al.* 2005; Schelly *et al.* 2006). The most extreme outcome of complete replacement, in which the original mtDNA of a species has totally disappeared in part of its range, is much rarer (but see Bernatchez *et al.* 1995).

A classical procedure to detect misleading mtDNA relationships is to check for discordance with independent sets of markers. Agreement between results obtained with multiple types of markers provides strong support for the inferred evolutionary relationships between populations or taxa (Cummings *et al.* 1995). Conflicting results are also helpful, as they improve our understanding of the historical events that have led to the extant distribution of evolutionary lineages (Shaw 2002; Morando *et al.* 2004). This is particularly true in hotspots of biodiversity where evolutionary histories of populations are often difficult to unravel (Gillespie & Roderick 2002; Ennos *et al.* 2005).

The Iberian Peninsula was a major glacial refugium during the climatic oscillations of the late Pleistocene (Hewitt 1996; Taberlet *et al.* 1998) and it is part of one of the world's biodiversity hotspots (Myers *et al.* 2000). The Iberian reptile fauna is representative of this biodiversity: its 45 reptile species include eleven (24%) endemic species and several species include species-level genetic lineages which are still not formally recognized (e.g. Mateo *et al.* 1996; Martinez-Solano *et al.* 2006; Vaconcelos *et al.* 2006). One of the most obvious cases of such unrecognized biodiversity is the lizard genus *Podarcis* Wagler, 1830 (Reptilia, Lacertidae). Five species are currently recognized in continental Iberian Peninsula: *Podarcis muralis* (Laurenti, 1768), *P. hispanicus* (Steindachner, 1870) sensu lato (see Geniez *et al.* 2007), *P. vaucheri* (Boulenger, 1905), *Podarcis bocagei* (Seoane, 1884) and *P. carbonelli* (Pérez-Mellado, 1981). The last two species occur in most places in sympatry with *P. hispanicus*, from which they differ in morphology, coloration and ecology, as well as genetically. They are thus clearly valid species. Within the traditional *P. hispanicus*, nuclear, mitochondrial and morphological markers delimit several distinct lineages that do not form a monophyletic unit when compared with *P. carbonelli* and *P. bocagei* (Sá Sousa 2000; Harris & Sá Sousa 2001, 2002; Harris *et al.* 2002; Pinho *et al.* 2003, 2006, 2007) suggesting that *P. hispanicus* as currently understood corresponds to a species complex (Pinho *et al.* 2008).

Three distinct morphotypes of this complex occur in the southeastern part of the Iberian Peninsula. The *liolepis* morphotype ('Type 3' in Pinho *et al.* 2006; 'northeastern form' in Geniez *et al.* 2007) is found in the province of Valencia along the Mediterranean coast north of Alicante, in addition to a large region of northern Spain and southern France. The 'type 2' morphotype (Guillaume 1987; Geniez 2001; Harris & Sá Sousa 2001, 2002) occurs in the central region of the peninsula, whereas the *hispanicus* morphotype (as redefined in Geniez *et al.* 2007) is limited to the Spanish Levant in Andalusia and in the province of Murcia. Distribution ranges of these three morphotypes are contiguous, and no locality is known to accommodate more than one morphotype.

In the area occupied by these three morphotypes, four distinct mitochondrial lineages have been identified (Harris & Sá Sousa 2002; Pinho *et al.* 2006). The '*P. hispanica* Type 2' lineage only occurs in the range of the 'type 2' morphotype, but two lineages occur in the range of the *hispanicus* morphotype: '*Podarcis hispanica* sensu stricto' (called 'Valencia lineage' hereafter, Table 1) and '*P. hispanica* Galera' (called 'Hispanicus lineage' hereafter). Similarly, two mitochondrial lineages, namely the '*P. hispanica* Type 3' (called 'Liolepis lineage' hereafter) and the Valencia mitochondrial lineages, have been identified within the range of the *liolepis* morphotype.

Nuclear patterns of differentiation among these mtDNA lineages have provided conflicting results:

Table 1 Correspondence between names employed in this study to describe mitochondrial and morphological units with the names provided in the previous studies (Pinho *et al.* 2006, 2007, 2008). Corresponding symbols used in figures are indicated between parentheses. 'Reference' indicates the samples retrieved from Pinho *et al.* (2006), and used for mitochondrial lineage identification

Present study		Previous studies	
Mitochondrial lineage	Morphotype	Lineage name	Reference
—	—	*Podarcis muralis*	*Gua1*
—	—	*Podarcis bocagei*	*BTA1*
—	—	*Podarcis hispanica* type 1A/type 1B	*Mon8/Trj1*
—	—	*Podarcis carbonelli*	*PR1*
Type 2 (open triangle)	'type 2' (filled triangle)	*Podarcis hispanica* type 2	*Phv4/Tie1*
Valencia (filled diamond)	—	*Podarcis hispanica* sensu stricto	*And8,Pod12, Mot1*
—	—	*Podarcis (hispanica) vaucheri*	*Cin1/Ouk7*
—	—	*Podarcis hispanica* Tunisia/Sirwah	*LK6/Js6*
Liolepis (open square)	*liolepis* (cross)	*Podarcis hispanica* type 3 or NE type	*Bur2*
Hispanicus (open circle)	*hispanicus* (filled circle)	*Podarcis hispanica* Galera	*Gal3*

based on ten allozyme loci, Pinho *et al.* (2007) revealed a lack of nuclear genetic differentiation between individuals bearing the Valencia and the Liolepis mitochondrial lineages, but Pinho *et al.* (2008) found no gene flow between populations with Valencia haplotypes and populations with Liolepis haplotypes using nuclear introns sequence variation. They also found significant gene flow between Valencia populations and populations with Hispanicus haplotypes.

It is thus still unclear whether the limits of taxonomic units based on mitochondrial markers are discordant with those based on nuclear markers. Besides, the correspondence between evolutionary units delimited by nuclear markers and those delimited by morphological data has never been investigated in this region. Currently, all that we have is an apparent discordance between mitochondria- and morphology-based taxonomy as suggested by comparison of published maps produced by separate teams working on different specimens.

The aims of this study were thus to investigate rigorously possible discrepancies between markers in the southeastern Iberian lineages of the *P. hispanicus* complex and to elucidate their origin. We first studied the concordance between delimitations of taxa based on morphological and mitochondrial data. This revealed several cases of discordance between these two kinds of markers. Next, we assessed the origin of these discordances with nuclear DNA (nDNA): if the geographical patterns of variation in nuclear and mitochondrial DNA were concordant, then the different morphotypes would correspond to instances of local adaptation or phenotypic plasticity rather than to true evolutionary lineages. On the contrary, if the distribution of mtDNA variation did not match with morphotypes and nuclear genotypes, introgressive hybridization or incomplete lineage sorting of mtDNA must have been involved.

Material and methods

Sample origin and DNA extraction

Specimens used for morphological analyses originated from various European collections (see Table S1, Supporting information). Specimens used for molecular analyses all originated from the 'Biogéographie et Ecologie des Vertébrés' collection (Montpellier, France). Sampling locations of the analysed specimens are shown in Figure 1 (see also Table S2, Supporting information). We collected all the genetic data presented here except for the sequences of the mitochondrial control region of 16 individuals (see Table 1), which were published previously (Pinho *et al.* 2006).

Total genomic DNA was extracted from ~40 mg of muscle using the Qiagen DNeasy® Tissue kit, following the manufacturer's protocol. For several poorly preserved samples, we performed the final elution with only 100 μL buffer (half the recommended volume) to concentrate the DNA. The resulting DNA was visually quantified after migration on agarose gels and diluted for use in PCR reactions.

Morphological analyses

Twenty-four quantitative and semi-quantitative variables describing size, pigmentation and pholidosis were measured (by PG; see Table S3, Supporting information) on a total of 607 lizards. Adult individuals of both sexes were pooled because they share the main traits that separate morphotypes (results not shown). The lizards originated from a total of 70 localities represented by at

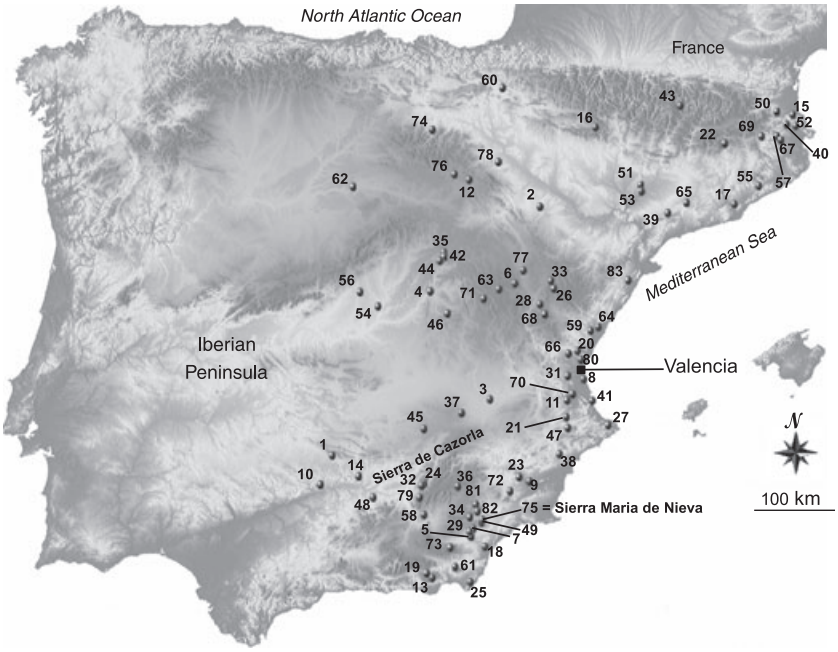

Fig. 1 The sampling sites. Map illustrating the sampling sites for morphological, mitochondrial and nuclear analyses (see also Table S2), as well as the main locations cited in this study.

least two individuals (other localities were discarded). First, a Discriminant Analysis was performed using the STATISTICA 6.1 software (StatSoft), including all of the measured specimens, to better define the limits between distribution ranges of morphotype. We defined a priori three groups of reference specimens (one per morphotype) using individuals located far from the suggested contact zones (approximately located in Geniez 2001) and that PG visually determined to be characteristic of each morphotype (localities outside the dashed square in Fig. 2). These reference specimens were used to construct two classification functions that classified individuals into the three morphological groups. Each locality was classified based on the average Mahalanobis distance from each group centroid (mean of the distances of all individuals in this locality). Each locality was attributed to the morphotype corresponding to the smallest average Mahalanobis distance. This method assumed that each locality did not accommodate more than one morphotype, in accordance with what is currently known in these lizards (see above) and with our visual examination of field or museum specimens. Finally, each lizard sample that needed to be identified (localities inside the dashed square in Fig. 2) was introduced as a supplementary item into the Discriminant Analysis.

The technique described above is suited to locate contact zones among morphotypes. However, it cannot identify more morphological groups than were defined a priori. To check whether the additional mitochondrial clade present in the study area (see results) corresponded to an overlooked fourth morphotype, we performed a Principal Component Analysis (PCA) using the same set of variables. This analysis included 298 *liolepis* and 190 *hispanicus* individuals located in an 80 km-wide and 800 km-long band running from Almeria (southern Andalusia, Spain) to Banyuls (south of France). Each locality was projected on a line running through the middle of the band to obtain a distance that was along the transect. This distance was then divided into 50 km-length sections and the first two individual principal component (PC) scores were averaged for each section. Mean PC scores were plotted against distance. We eventually compared averaged PC1 and PC2 scores between groups based on mitochondrial clade membership, and between groups based on the morphological classifications resulting from the Discriminant Analysis.

Mitochondrial DNA analyses

A 460 bp fragment of the mtDNA control region (CR) was amplified in a total of 69 specimens using the DL3F and DL4 primers (Crochet *et al.* 2004) and a protocol adjusted from Pinho *et al.* (2006). Sequencing reactions were conducted using ABI Dye Terminator kits (Applied Biosystems) following the manufacturer's recommendations. Sequencing samples were electrophoresed on an ABI Prism 310 Genetic Analyzer using DL3F as the sequencing primer. Table S4 (Supporting information) provides GenBank accession numbers of the 69 sequences.

Sequences were manually compiled and aligned using MEGA v.3.1 (Kumar *et al.* 2004). A phylogenetic tree was reconstructed with 87 sequences, including 16

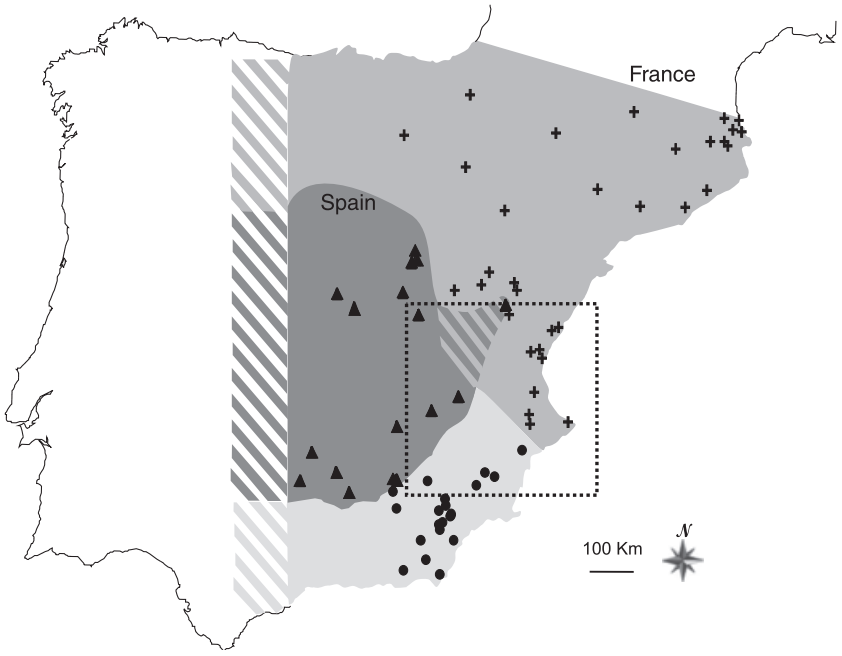

Fig. 2 Geographic distribution of the three morphotypes. Populations were assigned to a morphotype using a Discriminant Analysis. 'type 2' = filled triangles, *liolepis* = crosses and *hispanicus* = filled circles. Locations outside the dotted square were used as reference samples to compute the classification function (see text). Locations inside the dotted square were included in the analysis as supplementary items. Grey shadings are our interpretations of distributions of morphotypes derived from these results and are only used for illustrative purposes in Figs 2, 3 and 4. Grey hatching around locality 28 illustrates uncertainties about the distribution of 'type 2' in this area (see results section). Hatching on the left corresponds to the lack of precise information on distribution of morphotypes outside the study area.

sequences from Pinho *et al.* (2006), which allows us to represent the main clades known in the *Podarcis* genus and occurring in Iberia and North Africa (Table 1). Thus, we could assign each of our individuals to one of the mitochondrial lineages previously described in the *P. hispanicus* species complex. The tree was constructed using the maximum likelihood approach (ML, Felsenstein 1981). We used the heuristic search algorithm PHYML 2.4.4 (Guindon & Gascuel 2003) to estimate the ML tree. We used the ModelTest 3.06 software suite (Posada & Crandall 1998) and hierarchical likelihood ratio tests to select an appropriate model of sequence evolution. The best-fit model was the HKY, with an estimate of invariable sites (0.75) and a discrete approximation of the gamma distribution ($\alpha = 1.27$).

The nucleotide diversity (π), θ(s) computed from the number of segregating sites and haplotype diversity (h) were determined using ARLEQUIN 3.0 (Excoffier *et al.* 2005). We estimated the divergence time between Valencia haplotypes present in the *hispanicus* and in the *liolepis* morphotypes (see results) with a modified version of the coalescent-based program MDIV (Nielsen & Wakeley 2001). This program simultaneously estimates several parameters including theta ($\theta = 4\,N\mu$), migration rate ($M = 2\,Nm$) and the relative time of population divergence ($T = t/2N$) from the same data set. In this analysis, N is the gene effective population size (for mtDNA, N is equal to the effective population size of the female fraction of the population), t is the divergence time in years, and μ is the mutation rate per sequence per year. We used a modified version of MDIV (distributed by R. Nielsen) that provides the posterior distribution of θ^*T, allowing the calculation of the 90% credible interval (CI90) for this parameter. The simulations were conducted in triplicate, with runs of 5×10^6 generations and a 10% burn-in. An infinite-sites model with no migrant per generation was set to accommodate for the lack of shared haplotypes between the *hispanicus* and *liolepis* populations. The evolutionary rate of the mtDNA control region in reptiles is not well known. The CR fragment we used evolves at the same rate as cytochrome *b* in a closely related lizard genus (Crochet *et al.* 2004), but in the genus *Podarcis* it could evolve at a slower rate (C. Pinho, personal communication). On the other hand, the estimated population divergence time is almost certainly inflated because the populations we are studying violate the MDIV assumption of equal population sizes (Chapman *et al.* 2007). We thus computed an interval of T using two evolutionary rates, which probably encompass the real rate based on previously published estimates in reptiles (Crochet *et al.* 2004). This level is approximately 1% and 2% sequence divergence per million years.

Nuclear DNA analyses

Beta-fibrinogen intron 7 amplification and analysis. Pinho *et al.* (2008) amplified the *β-fibint7* intron in different *Podarcis* mitochondrial lineages. They observed an easy to detect length polymorphism between various lineages as a result of a $\sim$ 350 bp insertion. The length varies thus between $\sim$ 600 bp and $\sim$ 950 bp. For these reasons, we amplified DNA from 57 individuals using primers BF8 and Bfib (Pinho *et al.* 2008) and the PCR

conditions provided by these authors. The length polymorphism was evaluated by visualizing PCR products with UV light after electrophoresis on a 2% agarose gel.

Microsatellite amplification. A total of 104 individual samples were successfully genotyped for eight (Pb10, Pb11, Pb20, Pb37, Pb47, Pb50, Pb55 and Pb73) of the nine polymorphic microsatellite loci developed by Pinho *et al.* (2004). For multiplexing convenience, we excluded Pb66 from our analysis. Conditions of amplification are provided in Data S5 (Supporting information).

Delimitation of nuclear genetic units. The genetic structure of our samples of Iberian wall lizards was examined using the software GENELAND (Guillot *et al.* 2005). GENELAND is a spatially explicit clustering method based on a Bayesian model which assigns individuals probabilistically to clusters without a priori knowledge of population units and limits. It has been demonstrated to perform better than other spatially nonexplicit Bayesian clustering methods when genetic differentiation is weak (Fontaine *et al.* 2007) and when Wahlund effects may operate in one cluster (Coulon *et al.* 2006). In our case, samples of each lineage were collected over a wide geographic range and thus are expected to show gene-flow restrictions among them, at least because of isolation by distance, resulting in Wahlund effects within lineages. In addition, given the age of the different lineages in this complex (Pinho *et al.* 2006) and the mutation rates of microsatellites, we expected investigated taxa to be weakly differentiated because of homoplasy. Thus, the GENELAND software appears particularly appropriate for the study of the *Podarcis hispanicus* species complex. We used both microsatellites and nuclear-intron data together, with all individuals. We coded the intron as missing data for the samples in which it was not amplified. The maximum of convergence across runs was obtained from a minimum of 1.5×10^6 iterations. We performed a first series of 30 runs, allowing K to vary from 1 to 6 and using the following run parameters: 1.5×10^6 MCMC iterations sampled each 300 iterations, maximum rate of Poisson process fixed to 100, maximum number of nuclei in the Poisson-Voronoi tessellation fixed to 300, a full certainty associated with the spatial coordinates and the Dirichlet model of allelic frequencies. Moreover, we activated the newly included option that allows the algorithm to deal with null alleles (Guillot *et al.* 2008). We inferred the number of clusters from the modal value of K for these 30 runs, and then ran the MCMC 30 additional times with K fixed to this value, with the other parameters unchanged. We calculated the mean logarithm of posterior probability for each of the 30 runs. The five runs with the lowest value were discarded to avoid spurious clustering resulting from a MCMC scheme with no real statistical significance, which is always possible with Bayesian methods. The posterior probability of population membership for each pixel of the spatial domain was then computed for each of these 25 remaining runs, using a burn-in of 1.5×10^5 iterations and 125 pixels along the X axis and 95 along the Y axis. We then visually inspected the consistency of the results across the 25 runs.

Descriptive statistics within and among clusters. For microsatellite data, the mean number of alleles per locus (N_a) was calculated as well as Nei's unbiased expected heterozygosity (H_E; Nei 1978), observed heterozygozity (H_O) and Wright's F_{IS} statistic according to Weir & Cockerham (1984). The tests for departure from the Hardy–Weinberg equilibrium and for linkage disequilibrium between loci were conducted using 1000 permutations in each population. Differentiation indexes (F_{ST}) were calculated in accordance with the Weir & Cockerham (1984) procedure and the significance tested using 1000 permutations of individuals among groups. All calculations and tests were performed with GENETIX v.4.05 (Belkhir *et al.* 2001).

Results

Morphology-based clustering

All of the 51 reference localities were correctly classified using the Discriminant Analysis. Hence, we accepted the reference assignation and the classification functions as valid. One hundred and seven individuals originating from 19 localities within the contact zone (Fig. 2, inside the dashed square) were included as supplementary individuals in the classification functions to be assigned to one of the three morphotypes (*hispanicus*, *liolepis* and 'type 2').

The spatial distributions of the three morphotypes, inferred from the distribution of the classification scores averaged per locality, were nonoverlapping (Fig. 2). The *hispanicus* morphotype was restricted to the southern part of the study area, type 2 morphotype to the western part and the *liolepis* morphotype to the northern part. Individuals from Location 28, assigned on morphological grounds to 'type 2', were physically located far from the other 'type 2' populations and very close to locations attributed to *liolepis*. Because of the limited sample size for this location ($n = 2$), it is difficult to determine if the 'type 2' morphotype actually extends to this locality, or whether it should be considered as a misclassified sample.

In the *hispanicus–liolepis* contact zone, all the specimens belonging to the northernmost *hispanicus* location ($n = 6$) were individually classified as *hispanicus* (results not shown). In the two southernmost *liolepis* locations (21 and 27) located just north of the *hispanicus–liolepis* limit, 5 lizards of 21 were individually classified as *hispanicus*, although they have quite similar *hispanicus* and *liolepis* scores. Globally, in the southernmost *liolepis* samples, many individuals were separated by almost identical Mahalanobis distances from *hispanicus* and *liolepis* centroids. These results indicate that many *liolepis* specimens from the southernmost populations (close to the range limit of *hispanicus*) are morphologically close to *hispanicus*, while *hispanicus* specimens from the north of the distribution (close to *liolepis* range) do not differ in morphology from other *hispanicus* specimens. In addition, *hispanicus* populations that presented the Valencia mitochondrial haplotype (Vc) were unambiguously attributed to the *hispanicus* morphotype (see below, Fig. 3). Finally, populations 21, 27 and 47 were plotted at about equal distance from the *liolepis* centroid and the 'type 2' centroid (see Table S2 for details).

Mitochondrial sequence diversity and clustering

Our 69 specimens of *Podarcis hispanicus* from the southeastern Iberian Peninsula fell into four highly divergent mitochondrial lineages. Based on our results, we named these haplogroups the Hispanicus, Type 2, Liolepis and Valencia clades (Table 1). The overall sequence diversity was high ($h = 0.851 \pm 0.028$; $\pi = 0.018 \pm 1.1$; Table 2). Each of the 39 locations (including samples from Pinho *et al.* 2006) sampled was found to contain haplotypes of only one haplogroup. The Hispanicus, Type 2 and Liolepis haplotype groups present separate, nonoverlapping geographic distributions in the studied zone (Fig. 3a). Haplotype diversity and nucleotide diversity for these three groups varied from moderately low values in Type 2 ($h = 0.298 \pm 0.133$; $\pi = 0.9 \pm 0.7$) to very high in Liolepis ($h = 0.864 \pm 0.079$; $\pi = 3.5 \pm 1.9$). By contrast, the Valencia clade is found in several areas. This clade constitutes the only haplotype group found within a large part of southeastern Spain around the town of Valencia, while it is also found in a population (in the Sierra Maria de Nieva) surrounded by localities where we found the Hispanicus clade. In the main distribution range of the Valencia clade, only two different haplotypes were observed. One was present in most populations and individuals (23 individuals, haplotype Va in Fig. 3a, i.e. the same haplotype as Mot1 in Pinho *et al.* 2006) and the other one was only found in the two individuals from population 59 (Vb). Specimens presenting a haplotype belonging to the Valencia group sampled inside the Hispanicus range (the Sierra Maria de Nieva population) all shared a unique haplotype (Vc) distinct from both Va and Vb haplotypes by two substitutions. Haplotype and sequence diversity of the Valencia group were the lowest of all mitochondrial groups ($h = 0.458 \pm 0.094$; $\pi = 0.8 \pm 0.5$) and these values were even lower when only Valencia individuals having a *liolepis* morphotype were analysed ($h = 0.153 \pm 0.092$; $\pi = 0.2 \pm 0.2$).

The MDIV program was used to estimate the divergence time between the Va and Vc haplotypes. Maximum posterior probabilities were reached for values of $\theta = 0.18$ (CI90: 0.09–0.51), $T = 5.9$ (2.86–9.3) and $\theta^*T = 1.06$ (0.5–3), corresponding to a divergence time between the Va and Vc haplotypes of 120 000–240 000 years based on an estimated rate of sequence divergence of 2% and 1% per million years respectively, for the 431 bp sequence.

Morphological variation within and between morphotypes

The first two components of the PCA explained 44% of the total variance (Table 3). We found a significant morphological differentiation along PC1 ($F_{(1;\ 263)} = 4.39$; $P < 10^{-3}$) but not PC2 ($F_{(1;\ 263)} = 0.27$; $P = 0.79$) between *liolepis* individuals from the Valencia clade ($m_{PC1} = -0.416$ and $m_{PC2} = 0.323$) and from the Liolepis clade ($m_{PC1} = -0.815$ and $m_{PC2} = 0.356$). However, the plot with the mean PC scores (m_{PC1} and m_{PC2}) per 50 km section (Fig. 3b) illustrates that the southernmost section of the distribution range of the Valencia clade contains individuals presenting a divergent morphology. And indeed, the differentiation found on PC1 vanished ($F_{(1;\ 237)} = 1.91$; $P = 0.08$) when the individuals from the southernmost section were removed from the analysis (Valencia: $m_{PC1} = -0.61$).

Within the *hispanicus* morphotype, samples presenting the Valencia haplotype (Sierra Maria de Nieva population: $m_{PC1} = 0.75$ and $m_{PC2} = -0.33$) were not differentiated morphologically (PC1: $F_{(1;\ 61)} = -0.86$; $P = 0.38$ and PC2: $F_{(1;\ 61)} = 1.51$; $P = 0.14$) from the *hispanicus* specimens presenting the typical Hispanicus haplotypes ($m_{PC1} = 0.86$ and $m_{PC2} = -0.66$). However, they were clearly differentiated (PC1: $F_{(1;\ 43)} = 4.98$; $P < 10^{-3}$) from the *liolepis* morphotype/Valencia clade from the closest geographic section ($m_{PC1} = -0.01$).

Within the *liolepis-hispanicus* contact zone, a gap in the score distribution of PC1 was visible along the transect, and corresponded to the morphotype limits shown by the Discriminant Analysis. All size variables were strongly correlated with PC1 (Table 3). The gap between PC1 scores across the contact zone is therefore mainly due to a discontinuous transition from the small *hispanicus* to the large *liolepis* morphotype. By

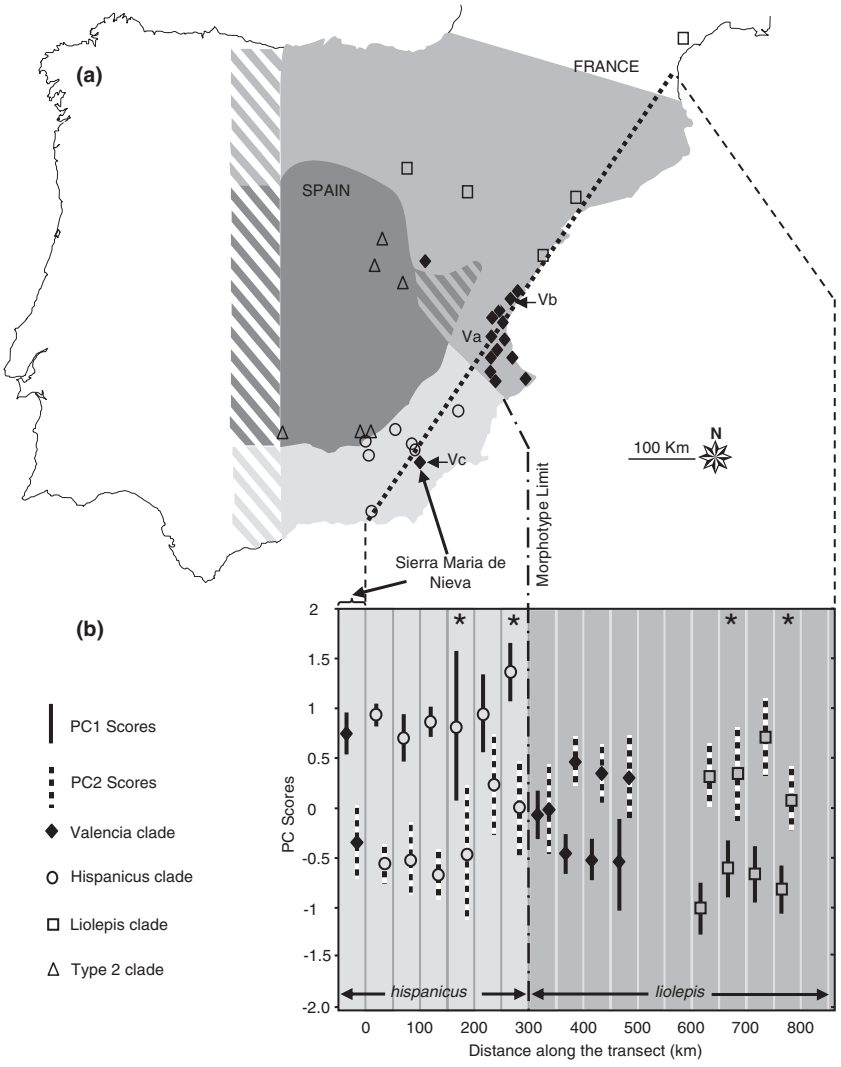

Fig. 3 Distribution of mitochondrial clades compared with the distribution of morphotypes. (a) The symbols indicate sampled locations for which a fragment of the mitochondrial Control Region was sequenced for one or several individuals. Open triangle = Type 2 mitochondrial clade, filled diamond = Valencia mitochondrial clade, open square = Liolepis mitochondrial clade, open circle = Hispanicus mitochondrial clade. No more than one mitochondrial clade was found in each location. Va, Vb and Vc indicate the distribution of the three different haplotypes occurring in the Valencia mitochondrial clade. The distribution of mitochondrial clades has to be compared with the distribution of morphotypes indicated by shading: light, medium and dark grey represent the distributions of *hispanicus*, *liolepis* and 'type 2' morphotypes respectively. Method used to construct shading is detailed in the legend of Fig. 2. The dotted line crossing Spain represents the transect along which morphological variation in *liolepis* and *hispanicus* is represented. (b) Geographical variation in PC scores extracted from a PCA performed with 24 morphological variables. Each column represents a 50-km-length section along the transect. Mean PC1 (plain line) and PC2 (dotted line) scores with 95% confidence intervals are given for each section. The symbols indicate the means and correspond to the mitochondrial clade occurring in a given section. Asterisks indicate areas where mitochondrial clade membership is inferred from haplotype distribution but not confirmed by specimen sequencing. The location 'Sierra Maria de Nieva' is represented separately, in the leftmost column. The limit between *liolepis* (medium shading) and *hispanicus* (light shading) morphotypes has been placed according to the Discriminant Analysis (Fig. 2).

contrast, PC2 scores show a gradual transition from one morphotype to another. Variables strongly correlated with PC2 concern the intensity of the ventral black pigmentation. The northernmost *liolepis* are thus less pigmented ventrally than the southernmost *hispanicus*.

Clustering based on nuclear markers

Beta-fibrinogen intron 7. A diagnostic allele presenting a 300 bp insert was observed at a homozygous state in all individuals classified morphologically as *hispanicus* by the Discriminant Analysis ($n = 18$), including the six

Table 2 Estimates of sequence diversity in southeastern Iberian mitochondrial lineages of the *Podarcis hispanicus* species complex. Hisp, Hispanicus; Lio, Liolepis; Val, Valencia; Val(*P. liolepis*), Valencia mtDNA found in *P. liolepis*. n_s, number of sampled locations including locations retrieved from Pinho *et al.* (2006); n_i, number of analysed individuals; n_h, number of observed mtDNA haplotypes; π, haplotype diversity; θ(s), computed from the number of segregating sites. Standard deviations are shown in brackets

Group	n_s	n_i	n_h	h	π (%)	θ(s) per site (%)
Hisp, Type 2, Lio,Val	39	78	21	0.851 (0.028)	1.8 (1.1)	2.0 (1.0)
Hisp	8	12	5	0.758 (0.093)	2.1 (1.3)	2.0 (1.1)
Type 2	7	19	5	0.298 (0.133)	0.9 (0.7)	1.7 (0.9)
Lio	6	13	7	0.864 (0.079)	3.5 (1.9)	3.6 (1.7)
Val	18	34	4	0.458 (0.094)	0.8 (0.5)	0.7 (0.5)
Val(*P. liolepis*)	15	26	2	0.153 (0.092)	0.2 (0.2)	0.3 (0.3)

individuals from the Sierra Maria de Nieva carrying the Valencia mitochondrial lineage. This insertion was not found anywhere else, except for the two individuals (from locations 21 and 27, Table S2) sampled within the *liolepis* range, less than 20 km north of the *hispanicus–liolepis* limit. These two individuals were heterozygous, presenting a copy of both long and short alleles.

Microsatellite variability and differentiation between morphological and mitochondrial groups. Allele numbers varied from 11 for pb11 in 'type 2' to 37 for pb55 in *liolepis* (Table 4). No linkage disequilibrium was detected between our microsatellite loci (Table S6, Supporting information). In both types of grouping (morphology vs. mitochondria-based), each group had an overall observed heterozygosity (H_O) lower than the expected heterozygosity (H_E), and presented a significant deficit in heterozygotes (Table 4). Pairwise F_{ST} estimates showed significant differentiation between morphotypes, even between the *liolepis* and the 'type 2' morphotypes that are the least differentiated. The highest F_{ST} values were observed between groups belonging to the same mitochondrial clade, but to different morphological groups (Valencia clade*/hispanicus* morphotype vs. Valencia clade*/liolepis* morphotype; $F_{ST} = 5.2\%$; $P < 10^{-2}$). On the contrary, specimens presenting the same morphotype but different mtDNA lineages were not significantly differentiated (Valencia clade*/hispanicus* morphotype vs. Hispanicus clade*/hispanicus* morphotype; $F_{ST} = 0.003$; $P < 10^{-2}$) or they were poorly differentiated (Valencia clade*/liolepis* morphotype vs. Liolepis clade*/liolepis* morphotype; $F_{ST} = 0.019$; $P < 5.10^{-2}$).

Delimitation of nuclear units using the Bayesian clustering method. Posterior distributions of the estimated number of populations (K) across the 30 runs presented a clear mode at $K = 3$ for each of the 30 runs (Fig. 4d). Twenty-four of the 25 selected runs were consistent in placing all of the *hispanicus* specimens (whatever their mitochondrial genome) into a single cluster. Indeed, individuals from the Sierra Maria de Nieva, presenting a Valencia mitochondrial haplotype, clustered with other *hispanicus* specimens, presenting a Hispanicus mitochondrial lineage. The remaining individuals were classified into two clusters following somewhat different patterns across these 24 runs. In 11 runs, individuals from the Valencia and the Liolepis clades were classified into the same cluster (Fig. 4a–c), except for

Table 3 Factor coordinates of 24 morphological variables used in the principal component analysis. * and ** indicate moderate (0.5–0.7) and strong (>0.7) correlation between the variable and the factor respectively

Variables		PC1	PC2
Scalation	Dors	−0.42	−0.23
	VenL	0.16	0.41
	Guls	0.03	−0.40
	Fpor	−0.47	−0.15
	Lame	−0.12	−0.33
	Temp	0.24	−0.46
	TeMa	−0.68*	0.33
	Mass	−0.74**	0.41
Size	ID10	−0.75**	0.34
	SVL	−0.80**	−0.10
	HeNe	−0.80**	−0.39
	PilL	−0.84**	−0.39
	PilW	−0.89**	−0.26
	HeH	−0.87**	−0.29
Pigmentation	Vert	0.15	−0.43
	Bif	0.62*	−0.33
	DoLa	0.43	−0.20
	Frag	−0.07	0.07
	SDLa	0.05	−0.49
	Pari	0.69*	−0.47
	Ponc	−0.12	−0.71**
	Rond	−0.02	−0.65*
	Tria	−0.19	−0.10
	PilP	−0.00	−0.67*
Eigenvalue		6.72	3.77
% Variance		29	16

Table 4 Genetic diversity and pairwise differentiations between clusters defined according to morphological and/or mitochondrial analyses. Mean number of overall alleles (N_a), Nei's diversity (H_E), observed heterozygosity (H_O) and fixation indices (F_{IS}) of Weir & Cockerham (1984) are given by the software GENETIX. F_{IS} values refers to Hardy–Weinberg equilibrium; F_{ST} values indicate genetic homogeneity between populations.*$P < 0.05$; **$P < 0.01$

No.	Group	Sample size	N_a	H_E	H_O	F_{IS}	Pairwise F_{ST} No. 1	2	4	6
1	*hispanicus* morphotype	33	20.63	0.88	0.75	0.17**	—	—	—	—
2	*liolepis* morphotype	48	27.50	0.94	0.70	0.26**	0.040**	—	—	—
3	'type 2' morphotype	23	16.00	0.90	0.71	0.24**	0.048**	0.022**	—	—
4	Hispanicus clade/*hispanicus* morphotype	23	17.75	0.88	0.77	0.14**	—	—	—	—
5	Valencia clade/*hispanicus* morphotype	10	10.25	0.83	0.70	0.21**	—	—	0.003	0.052**
6	Valencia clade/*liolepis* morphotype	27	20.12	0.91	0.71	0.24**	0.050**	—	—	—
7	Liolepis clade/*liolepis* morphotype	21	18.37	0.92	0.68	0.28**	0.040**	—	—	0.019*

the two individuals from location 27, which were systematically attributed to the other cluster with Type 2 clade specimens. In six runs, GENELAND assigned all remaining individuals to a single cluster and left the remaining cluster without any assigned individual. In three runs, the Valencia clade specimens clustered with all the Type 2 specimens and with individuals of the southernmost Liolepis localities, except for the westernmost Valencia specimens that clustered with the rest of the Liolepis specimens. A similar pattern was found in two runs except that southern Type 2 specimens clustered with most Valencia specimens and northern Type 2 specimens clustered with most Liolepis specimens. Two runs clustered Valencia specimens with a portion of the Type 2 and all the Liolepis specimens. Finally, a 25th run showed an aberrant clustering with no coherence with regard to either mitochondrial or morphological clustering. Most importantly, none of the 25 runs placed individuals of the Valencia clade into their own cluster. The occurrence of a 'ghost' cluster, that is to say without individuals assigned to it (because centred on the Mediterranean Sea), has already been reported by the authors of GENELAND (Guillot *et al.* 2005) and in other studies (Coulon *et al.* 2006; Fontaine *et al.* 2007) and remains a poorly understood problem.

Discussion

Mitochondria-based taxonomy revisited: cyto-nuclear discordances

In accordance with previous results (e.g. Pinho *et al.* 2006), we identified four mitochondrial lineages in south-eastern Iberia, the amounts of sequence divergence among which are typical of interspecific comparisons. As we suspected, this is incongruent with our identification of only three morphotypes in the same area, as found by Geniez (2001). The distributions of the Hispanicus, Liolepis and Type 2 mtDNA lineages closely match morphotype distributions, as they have been sampled only in populations assigned to the *hispanicus*, *liolepis* and 'type 2' morphotypes respectively. On the contrary, the Valencia clade was found in populations morphologically assigned to *liolepis* (with haplotype Va and Vb) and *hispanicus* (with haplotype Vc). It is thus the occurrence of this additional Valencia lineage in two of the three morphotypes that creates discordance between morphology and mitochondria. Moreover, GENELAND identified three genetic clusters in the study area that generally correspond to the morphological clusters (but see Discussion below). The occurrence of three nuclear clusters, concordant with morphotypes but not with the four mitochondrial lineages identified in the study area, implies that there is extensive cyto-nuclear discordance in these populations.

The first instance of discordance is the occurrence of the Valencia haplotype Vc in individuals from the Sierra Maria de Nieva population. GENELAND never associated these individuals with the other Valencia individuals but always with the specimens of the Hispanicus clade. This result is unambiguously supported by the length polymorphism of the beta-fibrinogen intron – the Sierra Maria de Nieva population shares the same alleles as the Hispanicus clade populations – as well as by the absence of significant differentiation at microsatellite loci between individuals of the Valencia and Hispanicus mtDNA clades.

The second case is the Valencia individuals presenting haplotypes Va and Vb that do not cluster into a separate group in any of the GENELAND runs. However, it should be noted that not every GENELAND run converged towards the same results (see below for a complete discussion of this issue). However, despite these evidences of occasional flaws in the nuclear clustering procedure, the fact that the most frequently obtained

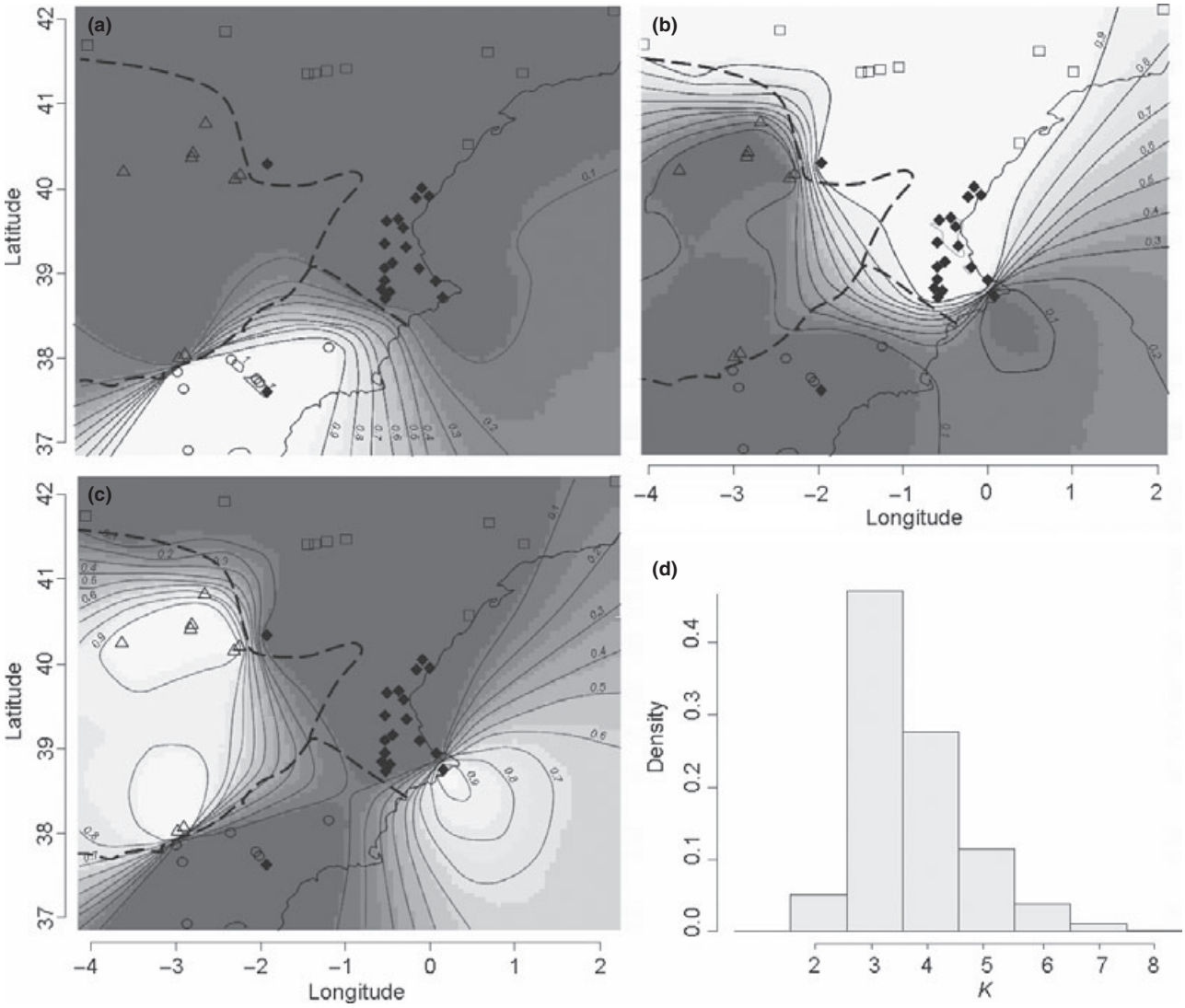

Fig. 4 Nuclear clustering performed with the GENELAND Bayesian algorithm and comparison with mitochondrial and morphological assignation. Plots (a) to (c) represent the assignment of pixels to each cluster, based on the clustering pattern showing the highest modal value among the 25 selected runs: assignment to (a) *hispanicus*; (b) *liolepis*; (c) 'type 2'. Highest membership values correspond to the palest shade and contour lines illustrate the spatial change in assignment values. Symbols for each locality correspond to their mitochondrial clade (open triangle, filled diamond, open square and open circle refer to Type 2, Valencia, Liolepis and Hispanicus mitochondrial clade respectively). Bold dashed lines represent the limits between morphological clusters according to Fig. 2. The scales indicate the geo-referenced coordinates in decimal units of degrees (WGS 84 projection system). (d). Posterior density distribution of the number of clusters along the MCMC chain after burn-in as estimated from the 30 replicate runs.

clustering pattern was congruent with morphological clustering strongly suggests that this general pattern reflects real genetic clustering.

Taken together, morphological and nuclear results thus suggest the occurrence of only three evolutionary units in the southeastern part of the Iberian Peninsula, corresponding to the previously recognized morphotypes. The three congruent morphological, mitochondrial and nuclear clusters correspond to three species: *Podarcis hispanicus*, *P. liolepis* and *P.* sp. 'type 2'. On the contrary, the occurrence of two highly divergent mtDNA lineages in two of these species (Hispanicus and Valencia mitochondrial clades in *P. hispanicus*; Liolepis and Valencia clades in *P. liolepis*) constitutes two clear cases of cyto-nuclear discordance.

The conclusions drawn here seem to contradict the results of Pinho *et al*. (2008), as they found virtually no nuclear gene flow between the Valencia mitochondrial clade and the Liolepis mitochondrial clade. However, the specimens of the Valencia clade they used originated from four distinct populations: two specimens belong to the *P. liolepis*/Valencia mitochondrial clade

populations, four specimens originate from the Sierra Nevada in southern Spain (an area isolated from *P. liolepis* populations by the distribution range of *P. hispanicus*) and belonged to an unknown taxon, one specimen originated from the Sierra Maria de Nieva (this specimen was also analysed in our study) and belonged to *P. hispanicus,* and two specimens originated from the Sierra de Cazorla (located around the *P. hispanicus*- *P.* sp. 'type 2' contact zone) and were also impossible to assign to any taxon. The estimate of nuclear gene flow between the Valencia and Liolepis lineages proposed by Pinho *et al.* was therefore based on the use of at least two taxa (*P. hispanicus* and *P. liolepis*) to represent the Valencia lineage. These gene flow estimates are thus difficult to compare with our findings of limited nuclear differentiation within *P. liolepis*.

Inconsistent nuclear clusters and variation within Podarcis liolepis

Our interpretation of the occurrence of the Valencia mitochondrial lineage in southern populations of *liolepis* as a case of cyto-nuclear discordance relies on our interpretation of the origin of the occasionally inconsistent patterns of nuclear clustering, and the weak but significant genetic and morphological differentiation observed between the northern and southern *liolepis* populations.

First, the difficulties of the software to consistently recover the same clusters may be caused by shared alleles because of incomplete lineage sorting or introgression, as evidenced by Pinho *et al.* (2008), by a high level of homoplasy in hypervariable microsatellite markers, or by an underlying genetic structure within *liolepis* and 'type 2' morphotypes (see below). A limited resolution in the GENELAND clustering procedure is indeed suggested by the systematic nuclear assignation of location 27 to *P.* sp. 'type 2'. Morphology, geography and distribution of mitochondrial lineages all indicate that these individuals are *P. liolepis* with a Valencia lineage. Nuclear intron evidences past hybridization with *P. hispanicus*, which is not unexpected given its location. However, genetic influences from *P.* sp. 'type 2' is highly unlikely as location 27 lies far from the 'type 2' – *liolepis* contact zone. Such assignation could therefore be caused by homoplasy or by incomplete lineage sorting.

Second, genetic variation within *liolepis* is likely to explain results of GENELAND runs where Va and Vb haplotypes did not cluster with all Liolepis haplotypes but where the southernmost location of Liolepis clustered with Valencia and the northernmost and westernmost Valencia location clustered with the rest of the Liolepis locations. This explanation is supported by the stronger values of heterozygote deficiency within the *liolepis* (and 'type 2') morphotypes than within the *hispanicus* morphotype. This suggests that genetic structuring among populations is more pronounced in the first two species than in the last one, in accordance with their wider distribution range. Such spatial genetic structure is expected to generate genetic differentiation between populations of *liolepis* presenting the Valencia haplotype and the Liolepis haplotype because they occupy different parts of the *liolepis* range.

In addition, the change in morphology of *liolepis* along the transect does not correspond to the change in mtDNA lineages: northern specimens carrying Valencia haplotypes are morphologically identical to typical *liolepis* and not to southern Valencia specimens. This morphological variation is thus better explained by introgression of morphological characters from *hispanicus* or by morphological convergence with this taxon as a result of adaptation to similar environments. Several arguments favour the introgression hypothesis. First, the presence of two lizards heterozygous for the beta-fibrinogen intron in the southern part of the *liolepis* range evidences past nuclear introgression between *liolepis* and *hispanicus* in their contact zone. Second, past or present gene flow may explain the occasional occurrence of typical *hispanicus* characters such as a split dark vertebral line, white belly, bluish tail in juveniles, lack of a distinct masseteric plate and shape of temporal scales in southernmost *liolepis* (see Geniez *et al.* 2007). Third, the contrasting pattern of the variation between PC1 and PC2 scores across the contact zone is also in accordance with introgression between *liolepis* and *hispanicus.* The gradual variation of ventral black punctuations revealed by PC2 is hardly explicable as an adaptation to a gradually varying environment while the size variation revealed by PC1 could be linked to climate variation, as small size is considered to be an adaptation to dry and hot habitats in lizards (Horton 1972). It should be noted that the location of the *liolepis-hispanicus* contact zone matches an abrupt transition in mean annual precipitation and temperature (Pérez Cueva 1994).

Origin of the discordances: introgressive hybridization of mitochondrial DNA

The presence of the Valencia mitochondrial lineage in *P. liolepis* and *P. hispanicus* may be due to either incomplete lineage sorting of ancestral polymorphism or introgressive hybridization. Incomplete lineage sorting has previously been suggested to explain shared polymorphism of nuclear introns in the same lizards (Pinho *et al.* 2008). However, the time needed to completely sort the ancestral polymorphism of mtDNA is on average four times smaller than for nuclear genes (Pamilo & Nei 1988), so mitochondrial gene genealogies are predicted to reach reciprocate monophyly before nuclear gene genealogies.

In our case, two points suggest that the cyto-nuclear discordances are not resulting from incomplete lineage sorting. First, the Valencia lineage is embedded within a larger cytoplasmic clade with the North African and South Iberian lineage of *P. vaucheri* (Pinho *et al.* 2006). The Liolepis and Hispanicus mitochondrial lineages are grouped into another monophyletic clade. The common ancestor to all of these clades is older than 7 Myr (Pinho *et al.* 2006). This age estimate is old enough to suggest that complete lineage sorting of the ancestral polymorphism should have been achieved under the 'no gene flow' scenario, unless the effective population size of these lizards is unrealistically large (between 2 and 7 million individuals for generation times between 1 and 3 years). Second, with the incomplete sorting hypothesis, it would be highly unlikely to find Valencia haplotypes in both *P. liolepis* and *P. hispanicus*, while finding no trace of the Type 2, *bocagei* and 'Type 1' mtDNA lineages, even though these are more closely related to Valencia than are Liolepis and Hispanicus.

On the contrary, the introgressive hybridization model is quite compatible with our results. The Iberian Peninsula was a major ice age refugium during the Pleistocene and introgressive hybridization has been predicted to occur in such locations (Avise 2000). By driving successive cycles of species range expansion followed by retraction in many taxa, the Quaternary cyclic climatic oscillations repeatedly separated, then put back into contact, many evolutionary units (Hewitt 1996; Taberlet *et al.* 1998). If we hypothesize that the haplotypes of the Valencia lineage present in *P. hispanicus* and *P. liolepis* have been isolated from each other after they 'invaded' these taxa (no current mitochondrial gene flow between *P. liolepis* and *P. hispanicus*), the age of their separation is the maximal possible age for both introgression events. The divergence time between the Valencia haplotypes found in *hispanicus* and *liolepis* was estimated to be approximately 120 000/240 000 years ago. This time indicates that their introgression indeed occurred during the last Quaternary climatic oscillations, even if there is some uncertainty over the estimates (see Methods).

The calculation of divergence time presented above relies on the hypothesis that hybridizations have occurred following contacts between the unknown Valencia taxon and both *P. hispanicus* and *P. liolepis* independently. Alternatively, hybridization could have occurred following contact with one of these taxa only. Then, hybridization between *P. hispanicus* and *P. liolepis* would have allowed the transfer of the Valencia lineage from one taxon to the other. This hypothesis of indirect transmission has been proposed in Iberian hares (Melo-Ferreira *et al.* 2007, 2009) in which the mtDNA of the arctic species *Lepus timidus* would have been captured by *L. europaeus* through hybridization with *L. granatensis*. Disentangling the two hypotheses will not be an easy task as long as we have no knowledge of the mtDNA present in the original Valencia-carrying taxon. However, if the Valencia mtDNA lineage had been passed from one taxon to the other (either from *liolepis* to *hispanicus* or the opposite), we would have expected the mtDNA polymorphism of the 'receiving' taxon to be embedded within the polymorphism of the 'donor' taxon. This is contrary to what we observe: the Valencia haplotypes that are present in *liolepis* and *hispanicus* form reciprocally monophyletic groups. We thus favour the hypothesis of two independent introgression events from the original Valencia-carrying taxon into *hispanicus* and *liolepis*. This hypothesis should nevertheless be tested by sequencing of mtDNA segments displaying more variability (such as ND4, see Pinho *et al.* 2008) to allow better examination of the relationships between the Valencia haplotypes present in *hispanicus* and *liolepis*, and by a more complete sampling of the region located south of the contact zone. Ultimately, an accurate test of the 'two introgression events' hypothesis would require locating remnant populations of the original Valencia-carrying taxon, provided they still exist. Currently, the taxon at the origin of the Valencia lineage has still not been identified.

The scenarios of mitochondrial replacement

The persistence of an alien mitochondrial lineage in a related taxon is not an uncommon phenomenon (reviewed in Mallet 2005). However, the extent of the mitochondrial introgression of the Valencia lineage within *P. liolepis* and the virtual lack of allelic polymorphism of this lineage deserve consideration. Two evolutionary scenarios can explain the current distribution of the Valencia mitochondrial lineage.

First, the distribution range of the Valencia-carrying taxon may have originally been more reduced than inferred from the current distribution of the Valencia mitochondrial lineage. After local hybridization with *P. liolepis*, this lineage would have secondarily replaced the Liolepis lineage in parts of its original range and would have spread until its current distribution in the region. In that case, the spread of the Valencia lineage over such a large area and the exclusion of the Liolepis lineage from all populations within this range are extremely unlikely by drift alone, and would rather suggest a selective sweep. Possible selection pressures on mtDNA include adaptation to climate, which appears to account for mitochondrial variation in humans (Mishmar *et al.* 2003; Ruiz-Pesini *et al.* 2004). The climate-driven selective hypothesis was originally emitted to explain the spread of the arctic *Lepus timidus* mitochondria within Iberian species of hares (Melo-Ferreira *et al.* 2005).

However, the high level of polymorphism of *timidus* mtDNA within Iberian populations of the other species and the polyphyletic origin of the introgressed haplotypes suggest multiple introgression events that do not strengthen the selective hypothesis (Melo-Ferreira *et al.* 2007). In the Valencia clade, only two haplotypes (one of which is rare and is derived from the common one by a single substitution) have been found in the southern *P. liolepis* populations, resulting in a reduced mtDNA diversity compared with the other mtDNA clades and taxa (Liolepis clade in *P. liolepis*, Hispanicus clade in *P. hispanicus*). This supports the idea of a unique introgression event from the Valencia lineage towards *P. liolepis*, which fits well with the selective hypothesis. Besides, ectotherm organisms such as reptiles are good candidates for such selection scenarios, as mitochondrial variants may exhibit marked fitness differentials (Somero 2002). Last, the areas currently occupied by the Liolepis and the Valencia clades incidentally experience contrasted climatic conditions (Pérez Cueva 1994).

In the second evolutionary scenario, the present distribution of the Valencia lineage would correspond to the former distribution range of the Valencia-carrying taxon. Direct competition with admixture in the colonization front between the invading *P. liolepis* and the invaded Valencia taxon could then have swamped most of the original nuclear genome but not the mtDNA. Recent theoretical models strongly suggest that when the range of a native species is invaded by another species, introgression of neutral genes occurs almost exclusively from the displaced species to the spreading species (Currat & Excoffier 2004; Currat *et al.* 2008). Moreover, the phenomenon would amplify as the invading species expands, leading to very high frequency or even fixation of foreign alleles in the invading species. This demographic hypothesis of competitive replacement would thus appear to fit our data well. However, simulation studies showed that complete mitochondrial replacement is reached only with a very high number of repeated hybridizations in the replacement front (Currat *et al.* 2008). The simulations are supported by empirical data on supposed competitive replacement: the foreign (captured) mitochondrial lineages are polyphyletic and the spreading lineage almost always persists in its own species, at least at low frequency (e.g. García-París *et al.* 2003; Roca *et al.* 2005; Berthier *et al.* 2006; Melo-Ferreira *et al.* 2007, 2009). None of these phenomena were observed with the Valencia mitochondrial lineage. In addition, in the case of neutral introgression, mtDNA would probably not be the only molecule to show signs of introgression. For example, in the case of the Iberian hares (Melo-Ferreira *et al.* 2009), the hypothesis of competitive replacement is favoured by signals of wide-range nuclear introgression. In the case of Iberian Wall Lizards, we would expect that populations of the invading *P. liolepis* carrying the alien mtDNA haplotypes would also exhibit signs of nuclear introgression, and thus differ significantly in nuclear genes from nonintrogressed populations (here carrying the Liolepis lineage). This was not the case in our study.

Factors responsible for reduced female-mediated gene flow compared with male-mediated gene flow, such as male-biased dispersal, would also favour the demographic scenario. In lizards, both male-biased (Ujvari *et al.* 2008) and female-biased (Olsson & Shine 2003; Ryberg *et al.* 2004) dispersal have been reported. It is therefore not possible to infer any gender-asymmetric dispersal in the genus *Podarcis* with current data. Moreover, dominance of the invading males over the invaded males could also account for differential introgression between maternally and biparentally inherited genes (Roca *et al.* 2005). However, such reproductive behaviour cannot be discussed as long as the Valencia-carrying taxon remains unidentified.

Further understanding of the processes that led to the extensive introgression of Valencia mtDNA into *P. liolepis* and *P. hispanicus* would clearly require the identification of the species at the origin of this lineage. On the one hand, the Valencia-carrying taxon could have totally disappeared. Mitochondrial introgression by a ghost lineage has been recently suspected in a fish genus (Bossu & Near 2009). On the other hand, the Valencia-carrying taxon may be extant but still undiscovered yet. To date, two mountain areas not sampled in this study are known to harbour populations with the Valencia lineage: the Sierra de Cazorla and the Sierra Nevada (Harris & Sá Sousa 2002; Pinho *et al.* 2006, 2007). One of these mountains, the Sierra de Cazorla, is already known to harbour a relict endemic lizard species, *Algyroides marchi*. During cooler glacial periods, a taxon inhabiting these mountain ranges today would probably have had a wider distribution, potentially contacting other *Podarcis* species. This putative scenario is nevertheless complicated by the result of Pinho *et al.* (2007) demonstrating the lack of nuclear genetic differentiation between *P. liolepis* and the Valencia populations of the Sierra Nevada. This finding is hardly explicable with our current knowledge of the species complex. Unravelling the origin of the Valencia taxon should be the goal of future works.

Conclusion: what confidence is there in single-marker based studies?

With the integrative approach in biology characterizing the new century (Pigliucci 2003), the drive to build taxonomic expertise is getting new impetus (Boero 2001).

The constantly growing complexity of morphology-based identification systems, coupled with the ceaseless decrease in the number of taxonomists, leads to the search for new approaches to taxon recognition. For example, the Barcode project proposes to use a single mitochondrial gene to identify and to discover new species (Hebert *et al.* 2003). In addition, some authors have explicitly assumed that for particular regions including refugia, single mtDNA sequences are good representatives of the genetic variability of such regions (e.g. Bilton *et al.* 1998). Recently, Barrowclough & Zink (2009) declared that 'if one's goal is to detect recently isolated groups of individuals or populations, i.e. discover taxa, the first step in a study ought to be the construction of an mtDNA gene tree'. One of the general messages of our study is to illustrate how misleading taxonomic conclusions would have been if based on mitochondrial markers, a message that refutes Barrowclough and Zink's statement. Indeed, the Valencia lineage has been assumed to correspond to a distinct taxon of the *P. hispanicus* complex (namely populations of *hispanicus sensu stricto*) in previous studies (Harris & Sá Sousa 2002; Harris *et al.* 2002; Pinho *et al.* 2003, 2006). None of these studies objectively analysed the morphological variation among and between lineages and therefore did not use it as a marker for systematics. The taxonomy of the *P. hispanicus* species complex is currently in revision and this revision would be inaccurate if based on mitochondrial data alone, particularly in southeastern Spain.

In addition to extensive mitochondrial introgressions shown in this study, incomplete lineage sorting is already known to affect our current view of individual nuclear gene genealogies in the *Podarcis hispanicus* species complex (Pinho *et al.* 2008). These two causes of discordance between gene trees and species trees are a major risk for single-marker taxonomy, particularly in regions where biodiversity is high (as in biodiversity hotspots or in past glacial refugia). These same areas are also those where the need to develop taxonomic work is most urgent. In other words, single-marker studies will be more prone to errors precisely in places where accurate characterization of biodiversity is most needed. In hotspots of diversity, only integrative studies will allow for a full understanding of singular events at the origin of each clade, and only an 'integrative taxonomy' (Will *et al.* 2005) will accurately recognize the specific richness in diversity of these regions.

Acknowledgements

We thank F. Kjellberg, D. McKey, D. Renoult, A. Baker and C. Pinho for constructive comments on an earlier version of this manuscript. We are also thankful for help and comments from M.-P. Dubois regarding the molecular methods. J.P. Renoult was supported by a grant from the French Ministry of Research.

References

Anderson TM, Lachance MA, Starmer WT (2004) The relationship of phylogeny to community structure: The cactus yeast community. *American Naturalist*, **164**, 709–721.

Andreasen K (2005) Implications of molecular systematic analyses on the conservation of rare and threatened taxa: Contrasting examples from Malvaceae. *Conservation Genetics*, **6**, 399–412.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman and Hall, New York.

Avise JC (2000) *Phylogeography: the History and Formation of Species*. Harvard Univ. Press, Cambridge, MA.

Avise JC (2004) What is the field of biogeography, and where is it going? *Taxon*, **53**, 893–898.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

Barrowclough GF, Zink RM (2009) Funds enough, and time: mtDNA, nuDNA and the discovery of divergence. *Molecular Ecology*, **18**, 2934–2936.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2001) *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier, Montpellier.

Bernatchez L, Glemet H, Wilson CC, Danzmann RG (1995) Introgression and fixation of arctic char (*Salvelinus alpinus*) mitochondrial genome in an allopatric population of brook trout (*Salvelinus fontinalis*). *Canadian Journal of Fisheries and Aquatic Sciences*, **52**, 179–185.

Berthier P, Excoffier L, Ruedi M (2006) Recurrent replacement of mtDNA and cryptic hybridization between two sibling bat species *Myotis myotis* and *Myotis blythii*. *Proceedings of the Royal Society B-Biological Sciences*, **273**, 3101–3109.

Biek R, Walsh PD, Leroy EM, Real LA (2006) Recent common ancestry of Ebola Zaire virus found in a bat reservoir. *PloS Pathogens*, **2**, 885–886.

Bilton DT, Mirol PM, Mascheretti S, Fredga K, Zima J, Searle JB (1998) Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards postglacial colonization. *Proceedings of the Royal Society B-Biological Sciences*, **265**, 1219–1226.

Bininda-Emonds ORP, Vazquez DP, Manne LL (2000) The calculus of biodiversity: integrating phylogeny and conservation. *Trends in Ecology & Evolution*, **15**, 92–94.

Boero F. (2001) Light after dark: the partnership for enhancing expertise in taxonomy. *Trends in Ecology & Evolution*, **16**, 266–266.

Bossu CM, Near TJ (2009) Gene trees reveal repeated instances of mitochondrial DNA introgression in orangethroat Darters (Percidae: *Etheostoma*). *Systematics Biology*, **58**, 114–129.

Chapman JW, Carlton JT, Bellinger MR, Blakeslee AMH (2007) Premature refutation of a human-mediated marine species introduction: the case history of the marine snail *Littorina littorea* in the Northwestern Atlantic. *Biological Invasions*, **9**, 995–1008.

Coulon A, Guillot G, Cosson JF *et al.* (2006) Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Molecular Ecology*, **15**, 1669–1679.

Crochet PA, Chaline O, Surget-Groba Y, Debain C, Cheylan M (2004) Speciation in mountains: phylogeography and phylogeny of the rock lizards genus *Iberolacerta* (Reptilia: Lacertidae). *Molecular Phylogenetics and Evolution*, **30**, 860–866.

Cummings MP, Otto SP, Wakeley J (1995) Sampling properties of DNA-sequence data in phylogenetic analysis. *Molecular Biology and Evolution*, **12**, 814–822.

Currat M, Excoffier L (2004) Modern humans did not admix with Neanderthals during their range expansion into Europe. *Plos Biology*, **2**, 2264–2274.

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: Massive introgression by local genes. *Evolution*, **62**, 1908–1920.

Edwards SV, Kingan SB, Calkins JD *et al.* (2005) Speciation in birds: Genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 6550–6557.

Ennos RA, French GC, Hollingsworth PM (2005) Conserving taxonomic complexity. *Trends in Ecology & Evolution*, **20**, 164–168.

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN version 3.0, an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Felsenstein J (1981) Evolutionary trees from DNA sequences – a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368–376.

Fontaine MC, Baird SJE, Piry S *et al.* (2007) Rise of oceanographic barriers in continuous populations of a cetacean: the genetic structure of harbour porpoises in Old World waters. *BMC Biology*, **5**, 30.

Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics*, **34**, 397–423.

García-París M, Alcobendas M, Buckley D, Wake DB (2003) Dispersal of viviparity across contact zones in Iberian populations of Fire Salamanders (*Salamandra*) inferred from discordance of genetic and morphological traits. *Evolution*, **57**, 129–143.

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*. Université Montpellier II & Ecole Pratique des Hautes Etudes, Montpellier.

Geniez P, Cluchier A, Sá Sousa P, Guillaume CP, Crochet PA (2007) Systematics of the *Podarcis hispanicus* complex (Sauria, Lacertidae) I: Redefinition, morphology and distribution of the nominotypical taxon. *Herpetological Journal*, **17**, 69–80.

Gillespie RG, Roderick GK (2002) Arthropods on islands: Colonization, speciation, and conservation. *Annual Review of Entomology*, **47**, 595–632.

Guillaume CP (1987) *Les petits Lacertidés du bassin méditerranéen occidental (Genera Podarcis et Archaeolacerta essentiellement). Sur quelques problèmes d'ordre systématique et biogéographique*. Université des Sciences et Techniques du Languedoc, Montpellier.

Guillot G, Mortier F, Estoup A (2005) GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.

Guillot G, Santos F, Estoup A (2008) Analysing georeferenced population genetics data with GENELAND: a new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics*, **24**, 1406–1407.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

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.

Harris DJ, Carranza S, Arnold EN, Pinho C, Ferrand N (2002) Complex biogeographical distribution of genetic variation within *Podarcis* wall lizards across the Strait of Gibraltar. *Journal of Biogeography*, **29**, 1257–1262.

Harrison RG (1990) Hybrid zones: windows on evolutionary process. *Oxford Survey in Evolutionary Biology*, **7**, 69–128.

Hebert PDN, Cywinska A, Ball SL, Dewaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B-Biological Sciences*, **270**, 313–321.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Horton DR (1972) Lizard scales and adaptation. *Systematic Zoology*, **21**, 441–443.

Jennings WB, Edwards SV (2005) Speciational history of Australian grass finches (*Poephila*) inferred from thirty gene trees. *Evolution*, **59**, 2033–2047.

Klein NK, Payne RB (1998) Evolutionary associations of brood parasitic finches (*Vidua*) and their host species: Analyses of mitochondrial DNA restriction sites. *Evolution*, **52**, 566–582.

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, **20**, 229–237.

Martinez-Solano I, Teixeira J, Buckley D, Garcia-Paris M (2006) Mitochondrial DNA phylogeography of *Lissotriton boscai* (Caudata, Salamandridae): evidence for old, multiple refugia in an Iberian endemic. *Molecular Ecology*, **15**, 3375–3388.

Mateo JA, LopezJurado LF, Guillaume CP (1996) Proteic and morphological variations in ocellated lizards (Lacertidae): A complex of species across the Strait of Gibraltar. *Comptes Rendus De L'Academie Des Sciences Serie III-Sciences De La Vie-Life Sciences*, **319**, 737–746.

Melo-Ferreira J, Boursot P, Suchentrunk F, Ferrand N, Alves PC (2005) Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Molecular Ecology*, **14**, 2459–2464.

Melo-Ferreira J, Boursot P, Randi E *et al.* (2007) The rise and fall of the mountain hare (*Lepus timidus*) during Pleistocene glaciations: expansion and retreat with hybridization in the Iberian Peninsula. *Molecular Ecology*, **16**, 605–618.

Melo-Ferreira J, Alves PC, Freitas H, Ferrand N, Boursot P (2009) The genome legacy from the extinct *Lepus timidus* to the three hare species of Iberia: contrast between mtDNA, sex chromosomes and autosomes. *Molecular Ecology*, **18**, 2643–2658.

Minelli A (2007) Invertebrate taxonomy and evolutionary developmental biology. *Zootaxa*, **1668**, 55–60.

Mishmar D, Ruiz-Pesini E, Golik P *et al.* (2003) Natural selection shaped regional mtDNA variation in humans. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 171–176.

Moran P, Kornfield I (1993) Retention of an ancestralpolymorphism in the Mbuna species flock (Teleostei, Cichlidae) of lake Malawi. *Molecular Biology and Evolution*, **10**, 1015–1029.

Morando M, Avila LJ, Baker J, Sites JW (2004) Phylogeny and phylogeography of the *Liolaemus darwinii* complex (Squamata: Liolaemidae): Evidence for introgression and incomplete lineage sorting. *Evolution*, **58**, 842–861.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Nichols R (2001) Gene trees and species trees are not the same. *Trends in Ecology & Evolution*, **16**, 358–364.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: A Markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.

Olsson M, Shine R (2003) Female-biased natal and breeding dispersal in an alpine lizard, *Niveoscincus microlepidotus*. *Biological Journal of the Linnean Society*, **79**, 277–283.

Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution*, **5**, 568–583.

Parra-Olea G, Wake DB (2001) Extreme morphological and ecological homoplasy in tropical salamanders. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 7888–7891.

Pérez Cueva A (1994) *Atlas climático de la comunidad Valenciana*. Comunidad Valenciana, Valencia.

Pigliucci M (2003) From molecules to phenotypes? The promise and limits of integrative biology. *Basic and Applied Ecology*, **4**, 297–306.

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.

Pinho C, Sequeira F, Godinho R, Harris DJ, Ferrand N (2004) Isolation and characterization of nine microsatellite loci in *Podarcis bocagei* (Squamata: Lacertidae). *Molecular Ecology Notes*, **4**, 286–288.

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 (2007) Comparing patterns of nuclear and mitochondrial divergence in a cryptic species complex: the case of Iberian and North African wall lizards (*Podarcis*, Lacertidae). *Biological Journal of the Linnean Society*, **91**, 121–133.

Pinho C, Harris DJ, Ferrand N (2008) Non-equilibrium estimates of gene flow inferred from nuclear genealogies suggest that Iberian and North African wall lizards (*Podarcis spp.*) are an assemblage of incipient species. *BMC Evolutionary Biology*, **8**, 63.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Roca AL, Georgiadis N, O'Brien SJ (2005) Cytonuclear genomic dissociation in African elephant species. *Nature Genetics*, **37**, 96–100.

Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC (2004) Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science*, **303**, 223–226.

Ryberg K, Olsson M, Wapstra E, Madsen T, Anderholm S, Ujvari S (2004) Offspring-driven local dispersal in female sand lizards (*Lacerta agilis*). *Journal of Evolutionary Biology*, **17**, 1215–1220.

Sá Sousa P (2000) A predictive distribution model for the Iberian wall lizard (*Podarcis hispanicus*) in Portugal. *Herpetological Journal*, **10**, 1–11.

Schelly R, Salzburger W, Koblmuller S, Duftner N, Sturmbauer C (2006) Phylogenetic relationships of the lamprologine cichlid genus *Lepidiolamprologus* (Teleostei : Perciformes) based on mitochondrial and nuclear sequences, suggesting introgressive hybridization. *Molecular Phylogenetics and Evolution*, **38**, 426–438.

Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology & Evolution*, **19**, 198–207.

Shaw KL (2002) Conflict between nuclear and mitochondrial DNA phylogenies of a 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.

Somero GN (2002) Thermal physiology and vertical zonation of intertidal animals: Optima, limits and cost for living. *Integrative and Comparative Biology*, **42**, 780–789.

Sullivan JP, Lavoue S, Arnegard ME, Hopkins CD (2004) AFLPs resolve phylogeny and reveal mitochondrial introgression within a species flock of African electric fish (Mormyroidea : Teleostei). *Evolution*, **58**, 825–841.

Sunnucks P (2000) Efficient genetic markers for population biology. *Trends in Ecology & Evolution*, **15**, 199–203.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Ujvari B, Dowton M, Madsen T (2008) Population genetic structure, gene flow and sex-biased dispersal in frillneck lizards (*Chlamydosaurus kingii*). *Molecular Ecology*, **17**, 3557–3564.

Vaconcelos R, Carretero MA, Harris DJ (2006) Phylogeography of the genus *Blanus* (worm lizards) in Iberia and Morocco based on mitochondrial and nuclear markers preliminary analysis. *Amphibia-Reptilia*, **27**, 339–346.

Vazquez DP, Gittleman JL (1998) Biodiversity conservation: Does phylogeny matter? *Current Biology*, **8**, R379–R381.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Will KW, Mishler BD, Wheeler QD (2005) The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology*, **54**, 844–851.

Julien P. Renoult is a PhD student in evolutionary biology, working on phylogeny and phylogeography of several animal and plant models, and on the evolution of visual signals in animal communication. Philippe Geniez is a permanent researcher in herpetology specialized in the Western Palaearctic fauna. Laure Benoit is a DNA technician and Paul Bacquet is a PhD student interested in phylogeography, chemical ecology and speciation. Pierre André Crochet is a permanent researcher working on phylogeography, phylogeny, evolutionary ecology and taxonomy of birds, reptiles and amphibians of the Western Palaearctic.

Supporting information

Additional supporting information may be found in the online version of this article:

Table S1 List of collections that provided specimens analysed in this study

Table S2 Morphological, mitochondrial and nuclear classifications for each location

Table S3 List of morphological variables used in multivariate analyses with details of the semi-quantitative transformation for qualitative variables

Table S4 List of individuals sequenced for the CR mitochondrial gene with mitochondrial lineage and corresponding GenBank accession numbers

Data S5 Protocol of microsatellite amplification.

Table S6 *P*-values of the linkage disequilibrium between the eight microsatellite loci calculated with 1000 permutations for the three morphotypes: *hispanicus*, *liolepis* and 'type 2'. NS, nonsignificant; $*P < 0.05$; $**P < 0.01$

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.