UNIVERSIDADE DO PORTO
FACULDADE DE CIÊNCIAS

Evolution of wall lizards (*Podarcis* spp.) in the Iberian Peninsula and North Africa

Catarina Lopes Pinho

Dissertação apresentada à Faculdade de Ciências da Universidade do Porto para obtenção do grau de Doutor em Biologia

Este trabalho foi apoiado pela Fundação para a Ciência e a Tecnologia (FCT) através da atribuição de uma bolsa de doutoramento (SFRH/BD/4620/2001)

À minha família

Nota prévia

Na elaboração desta dissertação, e nos termos do nº 2 do artigo 8º do Decreto-Lei nº 388/70, foi efectuado o aproveitamento total dos resultados de trabalhos já publicados, os quais integram alguns capítulos da presente dissertação. Em todos estes trabalhos, a candidata participou na obtenção, análise e discussão dos resultados, bem como na elaboração da sua forma publicada.

Agradecimentos

A realização desta tese de doutoramento teria sido impossível sem a ajuda e o apoio de diversas pessoas e instituições, às quais quero manifestar o meu reconhecimento.

Gostaria de agradecer em primeiro lugar ao meu orientador, Professor Nuno Ferrand de Almeida, que, ainda como meu professor na faculdade, estimulou o meu interesse pelo estudo da Evolução e me acolheu no centro de investigação que dirige. O seu entusiasmo contagiante, os seus ensinamentos, as discussões científicas enriquecedoras e a liberdade que sempre me garantiu contribuíram de forma decisiva para levar este trabalho a bom termo e para o meu desenvolvimento pessoal como investigadora.

Queria também expressar o meu agradecimento ao meu co-orientador, Professor D. James Harris, pelo interesse, companheirismo, encorajamento e ajuda constantes, bem como pela confiança que depositou em mim durante toda a nossa colaboração. Agradeço-lhe sobretudo pela partilha dos seus vastos conhecimentos e pontos de vista sempre originais, que foram uma grande inspiração e uma fonte constante de novas perspectivas para este trabalho.

Queria também agradecer de forma muito especial ao Miguel Angel Carretero pelo enorme jeito para apanhar lagartixas, pela boa disposição durante o trabalho de campo e pela colaboração preciosa em várias etapas deste trabalho. Os seus sólidos conhecimentos e abordagem multidisciplinar da biologia das lagartixas permitiram contextualizar de uma forma mais abrangente os resultados desta tese.

Gostaria de expressar um agradecimento especial a todos quantos colaboraram nas campanhas de amostragem ou cederam amostras: Miguel A. Carretero, Ana Luz, Anna Perera, Antigoni Kaliontzopoulou, Armando Loureiro, Bruno Ribeiro, Daniel Cunha, Diana Barbosa, Elsa Froufe, Gonçalo Themudo, Jean-Christophe Mace, Mário J. Pinho, Mohammed Znari, Mohsen Kalboussi, Neftalí Sillero, Paulo Sá-Sousa, Pedro Galán, Pierre-André Crochet, Philippe Geniez, Raquel Godinho, Raquel Ribeiro, Ricardo Pereira, Said Larbes, Salvador Carranza, Sara Rocha, Vasco Batista e Xavier Santos.

À Antigoni Kaliontzopoulou agradeço ainda a imprescindível contribuição para um dos artigos que compõem esta tese, quer na produção e análise de dados quer na redacção do mesmo.

À Raquel Godinho e à Elsa Froufe agradeço os ensinamentos e o apoio em algumas fases do trabalho laboratorial.

Ao multifacetado biólogo/designer/músico Vasco Batista, agradeço o design da capa desta tese.

Ao fotógrafo Jorge Coutinho agradeço a permissão para utilizar uma imagem de sua autoria na capa desta tese.

Ao Guillaume Queney agradeço a ajuda na tradução do resumo para francês.

Não posso deixar de manifestar um agradecimento muito especial aos meus colegas Sara Rocha, Ricardo Pereira, Miguel Carneiro, Hélder Freitas e José Melo-Ferreira pela partilha de bibliografia, opiniões e conhecimentos em diversas fases da escrita desta dissertação. Um agradecimento ainda mais especial é devido ao Rui Faria, à Xana Sá Pinto, ao Armando Geraldes e ao Fernando Sequeira, com quem discuti muitos dos meus resultados e que me ensinaram grande parte do que sei sobre genética e evolução, tendo contribuído decisivamente para a concretização desta tese.

Aos restantes colegas do CIBIO agradeço a boa-disposição, o companheirismo, a amizade e o espírito de entreajuda, que contribuiram para tornar o trabalho diário numa missão sempre agradável.

Gostaria também de agradecer ao Professor Craig Moritz, do Musem of Vertebrate Zoology da Universidade de Berkeley, na Califórnia (EUA), pelas valiosas discussões e por me ter recebido no laboratório que dirige. Apesar de curta, esta visita foi uma experiência enriquecedora tanto em termos profissionais como pessoais. Agradeço também a toda a equipa do MVZ por me ter feito sentir em casa durante a minha estadia.

À Sara Ferreira e à Sandra Rodrigues, gestoras de ciência e tecnologia do CIBIO, agradeço a boa-disposição e o indispensável apoio nos assuntos burocráticos.

De uma forma muito pessoal, agradeço sinceramente aos meus amigos, que me ajudaram a ultrapassar os momentos difíceis e sem os quais a conclusão desta tese teria sido bem mais complicada!

Ao Daniel Cunha agradeço a amizade, o carinho e a paciência sem fim com que aturou as minhas súbitas mudanças de humor, a falta de horários e os monólogos entusiasmados sobre lagartixas.

Aos meus pais, irmã e avós agradeço os mimos, os exemplos de perserverança, o encorajamento para aceitar todos os desafios e a compreensão e apoio incondicionais, mesmo perante os longos períodos de ausência e de saudades que este trabalho acarretou. Obrigada por tudo.

Em último lugar, gostaria de agradecer à Fundação para a Ciência e a Tecnologia o apoio financeiro através da concessão de uma bolsa de Doutoramento (SFRH/BD/4620/2001) e dos projectos de investigação POCTI/BSE/34547/2000, POCTI/BSE/40987/2001, POCTI/BSE/41912/2001, e POCI/BIA-BDE/60911/2004.

Contents

Summary

Iberian and North African *Podarcis* wall lizards constitute a cryptic species complex for which different morphological studies have produced largely conflicting results regarding the number, taxonomic rank and distribution of evolutionary entities. Recently, the assessment of mitochondrial DNA (mtDNA) diversity demonstrated the existence of highly differentiated and geographically consistent clusters, some of which corresponded neatly to partitions previously suggested by morphological analyses. As a result of this concordance, systematic reevaluations were carried out, and at present five different species are recognised: *P. bocagei*, *P. carbonelli*, *P. vaucheri*, *P. hispanica* and *P. atrata*, although the latter has been shown to be closely related to one of the various phylogroups that constitute the polytypic (and, from a mitochondrial DNA perspective, paraphyletic) *P. hispanica*.

Given this framework, the first goal of this thesis was to characterize in more detail the dynamics of genetic differentiation among forms of this species complex. Previous descriptions of genetic variability within the complex relied on a single marker – mitochondrial DNA –, and biogeographical inferences were heavily dependent upon poorly supported estimates of relationships. On one hand, we aimed at obtaining a more robust mtDNA-based phylogenetic tree in order to validate previous hypotheses regarding the biogeographical processes governing variability within the clade. On the other hand, we sought to complement these data with information from the nuclear genome, in order to independently assess the distinctiveness of the various clades previously described. Additionally, by examining patterns of nuclear-mitochondrial concordance and analysing in detail the dynamics of a contact zone between two species, we were expecting to evaluate levels of reproductive isolation among forms of the species complex.

The reevaluation of mitochondrial phylogenetic relationships was accomplished by selecting a few individuals from each lineage and extending the amount of included sequence data. In addition to the two mtDNA gene regions already published, we sequenced three other fragments, yielding a total of 2425bp. A robust phylogenetic tree was recovered, with a significant proportion of nodes receiving bootstrap support close to 100%. This rigorous assessment of mtDNA relationships between lineages confirmed some of the results obtained in previous studies, such as the paraphyly of *P. hispanica*. However, other aspects of the phylogeny differ, offering alternative scenarios for the timing and polarity of the colonization of North Africa by wall lizards. In particular, we suggest that the opening of the Strait of Gibraltar might have caused vicariant isolation of Iberian from North African forms, and that a subsequent transmarine colonization could have occurred from North Africa to Iberia. An additional important contribution of this study was the detection of a previously undescribed lineage inhabiting southern Spain, which further exemplifies the evolutionary complexity within this group.

Two major lines of research were pursued regarding the description of nuclear gene variation: allozyme electrophoresis and sequencing of nuclear introns. The study of allozyme variation in 30 populations representing all known mtDNA lineages documented a high degree of genetic divergence between most of them, with groups within *P. hispanica* generally presenting similar levels of differentiation to those observed between fully recognised species. Although this constitutes evidence corroborating the validity of mtDNA lineages as evolutionary distinct units, the application of recent model-based clustering approaches revealed some important discordant patterns with respect to the mtDNA, which may be indicative of the occurrence of gene flow between forms. In this context, two hybrid individuals between the broadly sympatric *P. bocagei* and *P. hispanica* type 1 were detected.

Conversely, the study of two nuclear gene genealogies (*β-fibint7* and *6-Pgdint7*), demonstrated that, in general, species and mtDNA lineages are strikingly non-monophyletic, which was unexpected given the high levels of differentiation detected by the use of other markers. An important issue that required clarification was, then, if this pattern emerged as the result of incomplete lineage sorting of ancestral polymorphism (motivated by the four-times higher effective population size of nuclear genes when compared to the mtDNA) or if it resulted from high levels of gene flow. In order to disentangle between these two non- mutually exclusive hypotheses, we estimated levels of gene flow between all pairs of lineages using recent methods of analyses based on coalescent theory that allow the study of the dynamics of admixture and divergence in the same framework. In contrast to the information obtained using classic, *F*-statistic-based estimators of gene flow, this approach revealed that only a few pairs of lineages have exchanged genes since their divergence, and even fewer show important levels of gene flow, suggesting that in general the various lineages are overall differentiated. Therefore, although gene flow has certainly played an important role in the evolution of the species complex (and as a result some mtDNA lineages may have even lost their nuclear identity), the polyphyletic pattern observed in nuclear genealogies is especially influenced by the persistence of ancestral polymorphism.

These studies addressed the subject of gene flow among forms superficially, mostly based on the patterns of mitochondrial/nuclear discordance. In order to characterize in more detail the dynamics of gene flow, a battery of nuclear and mitochondrial markers was studied along a transect crossing the contact zone between *P. bocagei* and *P. carbonelli*. This information was complemented with analyses of morphology and fertility. Applying model-based individual multilocus genotype clustering approaches, we document abundance of individuals showing signs of admixture in the locality where the two species meet. However, strong Hardy-Weinberg and linkage disequilibria were observed. This clear bimodality suggests the existence of strong barriers to gene flow, the nature of which still remain obscure. For example, our data suggests neither obvious reductions in hybrid fertility nor the verification of Haldane's rule. However, divergent natural selection acting on morphology is suggested by the fact that genetically admixed individuals are clearly assigned to one species or the other based on morphological characters. Bimodality within a hybrid zone is usually suggestive of a nearly complete process of speciation; while these results cannot be promptly generalized to other contact zones, they are in accordance with previous suggestions that, although *Podarcis* lineages are not fully reproductive isolated, levels of gene flow among them do not seem to challenge the ongoing process of differentiation.

A second major goal of this thesis was to describe the phylogeographic structure within selected forms of the species complex using multiple molecular markers. In particular, we attempted to evaluate the response of the species to the Pleistocene climatic oscillations, comparing the patterns detected with phylogeographic scenarios described for other species occupying the same regions, and interpreting contrasting signatures left by the Ice Ages on distinct species of *Podarcis*.

The study of mtDNA variation in *P. bocagei*, *P. carbonelli* and *P. vaucheri,* three species with parapatric distributions that replace each other along a latitudinal gradient, allowed us to test two predictions relative to the differential influence of Quaternary climatic oscillations on distinct latitudes: i) northerly distributed species are expected to bear lower levels of genetic subdivision and diversity than species distributed in the south because they were able to survive in fewer, smaller and less stable patches of favourable habitat during glaciations; ii) species distributed in the south should retain signs of long-term effective population size stability, while northern species, because they were more

confined during glacial stages, should have experienced a rapid demographic growth concomitant with a post-glacial colonization of newly available habitats. Our results show that *P. bocagei* presents remarkably low levels of diversity and subdivision, a shallow coalescent history (the estimated coalescence time was ~100,000-~70,000 years) and a strong signature of demographic growth. *P. vaucheri*, on the other hand, presents large levels of genetic diversity even at small geographic scales, strong geographic subdivision, no evidence of demographic growth and an ancient coalescence time, probably dating back to the initial stages of the Pleistocene. The intermediately distributed *P. carbonelli* presents average values of all studied variables. Taken together, these results entirely fit to our main predictions and demonstrate that different effects of the Ice Ages can be detected at small geographic scales.

The analyses of nuclear markers (allozymes and a set of microsatellite loci specifically developed during this study) in populations of *P. bocagei* and *P. carbonelli* confirmed, on one hand, the low levels of population subdivision observed in both species from the study of mitochondrial DNA. Moreover, levels of differentiation are higher among populations of *P. carbonelli* than among populations of *P. bocagei*, which is also in accordance with higher levels of persistence and a consequent longer time for differentiation in the former inferred from mtDNA. In *P. bocagei*, we report a progressive loss of genetic diversity in a northwards fashion, consistent with a recent geographic expansion from a reduced source as suspected from mtDNA-based evidence of a rapid demographic growth. Moreover, based on a spatial interpolation of genetic data and on a Bayesian model-based multilocus genotype clustering algorithm, we were able to identify probable expansion routes and to pinpoint with some degree of certainty the area that has probably functioned as a glacial refugium for the species during the last glaciations. Applying the same methodology in *P. carbonelli* we document some degree of association between genetic clusters and geography, but with only partial correspondence to inferences based on mtDNA; to explain these discordant patterns, we hypothesise that recent fragmentations in the species' distribution might have led to a loss of the genetic signatures that are typically found in co-distributed species, and that the patterns that are presently observed are more a by-product of recent genetic drift than of Pleistocene climatic oscillations.

Globally, these results constitute a major improvement regarding previous knowledge on evolutionary relationships, stages of differentiation, levels of gene flow and patterns of intraspecific subdivision in Iberian and North African *Podarcis*. At a more general level, the large-scale study of nuclear gene variation on these lizards, documenting polyphyly and gene flow among mtDNA lineages, illustrates that these correspond to very incipient species with permeable boundaries, highlighting the need for the assessment of multiple genetic markers prior to any taxonomic reevaluations. Moreover, the results have a double importance for the understanding of the Iberian and North African Quaternary biogeography: on one hand, by helping to confirm the transversality of patterns that appear in response to climatic fluctuations, such as glacial fragmentation and post-glacial expansion; on the other hand, by documenting how species-specific such patterns may be even among closely-related species.

Resumo

As lagartixas do género *Podarcis* da Península Ibérica e do Norte de África constituem um complexo de espécies crípticas para o qual diferentes estudos de caracterização morfológica produziram resultados controversos no que diz respeito à delimitação de entidades evolutivas. Recentemente, a descrição da variabilidade genética ao nível do DNA-mitocondrial (mtDNA) demonstrou a existência de vários grupos altamente diferenciados e com uma forte associação com a geografia, alguns dos quais correspondendo claramente a entidades sugeridas pelos estudos morfológicos. Como resultado desta concordância, foi levada a cabo uma revisão taxonómica, reconhecendo-se presentemente cinco espécies: *P. bocagei*, *P. carbonelli*, *P. vaucheri*, *P. hispanica* e *P. atrata*, se bem que esta última seja extremamente similar do ponto de vista genético a um dos vários clados detectados dentro de *P. hispanica*, espécie reconhecidamente politípica e, do ponto de vista do DNA-mitocondrial, parafilética.

Com base neste cenário, estabeleceu-se como primeiro objectivo desta tese a caracterização detalhada dos padrões de diferenciação genética entre formas do complexo *Podarcis*, uma vez que as descrições prévias da variabilidade genética neste género se baseavam totalmente num único marcador molecular – o mtDNA – e que as inferências biogeográficas eram altamente dependentes de estimativas de relações evolutivas relativamente mal suportadas estatisticamente. Por um lado, tentou-se obter uma árvore filogenética mais robusta para o DNA-mitocondrial, de forma a testar as hipóteses biogeográficas colocadas anteriormente. Por outro lado, procurou-se complementar estes dados com informação do genoma nuclear, para averiguar de modo independente o grau de diferenciação das linhagens anteriormente descritas. Adicionalmente, através do exame dos padrões de concordância entre DNA-mitocondrial e genes nucleares e do estudo detalhado da dinâmica de uma zona de contacto entre duas espécies, tentou-se avaliar o grau de isolamento reprodutivo entre formas do complexo *Podarcis*.

A reavaliação das relações filogenéticas foi concretizada através da selecção de alguns indivíduos de cada linhagem e da sequenciação de uma porção maior de DNA-mitocondrial. Para além dos dois genes estudados em trabalhos anteriores, foram sequenciados fragmentos de outras três regiões, obtendo-se um total de 2425 pares de bases. Da análise destes dados resultou uma árvore filogenética robusta, na qual a maioria dos clados obtém elevados valores de *bootstrap*. Esta avaliação rigorosa da filogenia mitocondrial confirma algumas das hipóteses levantadas anteriormente, como por exemplo a parafilia de *P. hispanica*. No entanto, alguns aspectos da filogenia recuperada diferem em relação a estimativas anteriores, propondo-se cenários alternativos para a colonização do Norte de África. Em particular, sugere-se que a abertura do Estreito de Gibraltar possa ter causado o isolamento vicariante das formas Ibéricas e Norte-Africanas, e que uma colonização posterior tenha ocorrido do Norte de África para a Península Ibérica. Uma contribuição importante deste estudo foi também a descrição de uma linhagem endémica do Sudeste da Península Ibérica, não detectada em trabalhos anteriores, que constitui um exemplo adicional da complexidade evolutiva deste grupo de espécies.

No que diz respeito à descrição da variabilidade genética ao nível do genoma nuclear, foram seguidas duas linhas de investigação: a descrição da variação electroforética de aloenzimas e a sequenciação de intrões nucleares. O estudo da variabilidade genética ao nível das aloenzimas em 30 populações representativas de todas as linhagens mitocondriais revelou que, em geral, os grupos definidos com base no mtDNA se encontram bem diferenciados e que as várias linhagens observadas em *P. hispanica* apresentam entre si graus de diferenciação semelhantes aos observados entre espécies reconhecidas. Apesar de este facto constituir uma corroboração da validade das linhagens

mitocondriais como entidades evolutivas claramente distintas, a aplicação de métodos de análise recentes que agrupam indivíduos com base no seu genótipo *multilocus* demonstrou alguns padrões discordantes que poderão resultar da ocorrência de fluxo génico entre formas. A este nível, foram detectados dois indivíduos que muito provavelmente serão híbridos entre *P. bocagei* e *P. hispanica* tipo 1, formas largamente simpátridas.

Em oposição ao elevado grau de diferenciação observado através da análise do DNA-mitocondrial e das proteínas, o estudo de duas genealogias nucleares (*β-fibint7* e *6-Pgdint7*) revelou uma inesperada ausência de monofilia das diversas espécies e linhagens mitocondriais. Tornou-se necessário, assim, esclarecer se este padrão polifilético resultaria da persistência de polimorfismo ancestral, incompletamente distribuído pelas diversas linhagens, ou de um elevado grau de fluxo génico. No sentido de avaliar a plausibilidade destes dois cenários (que não são mutuamente exclusivos), recorreu-se a métodos de análise baseados na teoria da coalescência que permitem estudar migração e diferenciação no mesmo contexto. De acordo com esta abordagem, detectaram-se níveis de miscigenação importantes entre muito poucos pares de linhagens, o que sugere que a maior parte das formas se encontram isoladas reprodutivamente. Assim sendo, apesar de o fluxo génico poder ter desempenhado um papel importante na evolução deste complexo de espécies (sendo que algumas linhagens mitocondriais poderão inclusivamente ter perdido a sua identidade nuclear), o padrão polifilético observado nas genealogias nucleares resulta sobretudo da persistência de níveis elevados de polimorfismo ancestral.

Estes estudos abordaram a caracterização dos níveis de fluxo génico entre linhagens de uma forma superficial, através do estudo da discordância entre padrões evolutivos revelados por marcadores mitocondriais e nucleares. Para caracterizar com mais detalhe os processos de miscigenação entre linhagens de *Podarcis*, estudou-se uma bateria de marcadores nucleares e mitocondriais ao longo de um transecto atravessando a zona de contacto entre *P. bocagei* e *P. carbonelli.* Esta informação foi complementada com análises de morfologia e fertilidade. Aplicando métodos de agrupamento individual, demonstra-se uma abundância de indivíduos apresentando sinais de miscigenação na localidade onde as duas espécies coexistem. No entanto, observam-se também desvios significativos aos equilíbrios de Hardy-Weinberg e de ligação. Esta bimodalidade sugere a existência de barreiras impeditivas do fluxo génico entre as duas espécies, apesar da natureza das mesmas permanecer obscura. Por exemplo, não foi possível demonstrar reduções na fertilidade dos híbridos nem a verificação da regra de Haldane. No entanto, a ausência clara de indivíduos de morfologia intermédia, mesmo em indivíduos apresentando sinais de miscigenação, parece implicar a existência de forças selectivas actuando sobre a morfologia. Este tipo de zona híbrida bimodal é normalmente sinal de um processo de especiação praticamente concluído. Estes resultados, apesar de não poderem ser generalizados a outras zonas de contacto, estão de acordo com a sugestão de que o grau de fluxo génico que ocorre entre linhagens de *Podarcis* não parece ser suficiente para impedir o processo de diferenciação.

Um segundo objectivo desta tese prendeu-se com a descrição da estrutura filogeográfica intraespecífica em três espécies de lagartixa através do uso de múltiplos marcadores moleculares. Em particular, pretendeu-se avaliar a resposta das espécies escolhidas às oscilações climáticas que caracterizaram o período Quaternário, comparando os padrões observados com os cenários filogeográficos descritos para outras espécies distribuídas nas mesmas regiões e interpretando as histórias evolutivas contrastantes observadas em diferentes espécies de *Podarcis.*

O estudo da variação genética ao nível do mtDNA em *P. bocagei*, *P. carbonelli* e *P. vaucheri*, três espécies com distribuições parapátricas que se substituem ao longo de um gradiente latitudinal, permitiu-nos testar duas hipóteses relativas à influência diferencial

das oscilações climáticas do Quaternário de acordo com a latitude: i) espera-se que espécies distribuídas a Norte apresentem níveis diminuídos de diversidade e de subestruturação genética quando comparadas com espécies distribuídas a Sul, devido ao facto de durante as glaciações as condições climáticas permitirem apenas a subsistência de um pequeno número de manchas de habitat favorável a Norte, provavelmente de menores dimensões e estabilidade; ii) espécies distribuídas a Sul deverão evidenciar sinais de estabilidade demográfica a longo-prazo, contrastando com espécies distribuídas a Norte, que, por terem permanecido confinadas em refúgios reduzidos durante as fases glaciais, deverão ter sofrido um crescimento demográfico súbito concomitante com a colonização pós-glacial de regiões anteriormente inóspitas. Os resultados obtidos demonstram que *P. bocagei*, a espécie com uma distribuição mais nortenha, apresenta níveis relativamente baixos de diversidade e subdivisão, uma história coalescente recente (o tempo de coalescência estimado é de 70000 a 100000 anos) e uma clara evidência de crescimento demográfico. Pelo contrário, *P. vaucheri,* a espécie com distribuição mais a Sul, apresenta uma grande diversidade genética (mesmo a uma escala geográfica reduzida), forte subestruturação geográfica, evidência de estabilidade demográfica e um tempo de coalescência antigo, coincidente com as fases iniciais do Pleistoceno. *P. carbonelli*, espécie com uma distribuição intermédia entre as duas anteriores, apresenta valores médios para todas as variáveis estudadas. Em conjunto, estes resultados ajustam-se totalmente às previsões iniciais e demonstram que efeitos contrastantes das glaciações de acordo com a latitude podem ser detectados a escalas geográficas relativamente pequenas.

A análise de marcadores nucleares (aloenzimas e um conjunto de microssatélites especificamente desenvolvido durante este estudo) em populações de *P. bocagei* e *P. carbonelli* confirmou os baixos níveis de subestruturação populacional inferidos para ambas as espécies através da análise do mtDNA. Para além disso, os níveis de diferenciação são mais elevados em *P. carbonelli* do que em *P. bocagei*, o que também está de acordo com o mtDNA ao sugerir um maior grau de persistência e maior tempo de divergência na primeira do que na segunda espécie. Em *P. bocagei* detectou-se uma perda progressiva da variabilidade genética no sentido Sul-Norte, consistente com uma expansão geográfica recente a partir de um refúgio, que está de acordo com o cenário de expansão demográfica inferida a partir do mtDNA. Adicionalmente, através da interpolação espacial dos dados genéticos e da aplicação de um método Bayesiano de agrupamento individual, foi possível identificar vias prováveis para a expansão geográfica e apontar com um considerável grau de detalhe a área que terá funcionado como refúgio para a espécie durante as últimas glaciações. Aplicando a mesma metodologia, em *P. carbonelli* verifica-se um grau moderado de associação entre a variabilidade genética e a geografia, mas apenas uma correspondência parcial em relação aos grupos inferidos com base no mtDNA; como forma de acomodar estes padrões discordantes, apresenta-se a hipótese de que fragmentações recentes na distribuição da espécie poderão ter levado à perda das assinaturas genéticas tipicamente encontradas em espécies co-distribuídas, e que os padrões presentemente observados se fiquem a dever mais a deriva génica recente do que às oscilações climáticas do Pleistoceno.

No seu conjunto, estes resultados constituem um importante avanço em relação ao conhecimento prévio das relações evolutivas, estados de diferenciação, níveis de fluxo génico e padrões de subdivisão intraespecífica nas lagartixas do género *Podarcis* da Península Ibérica e do Norte de África. De um ponto de vista mais geral, o estudo em larga escala da variação em genes nucleares nestes organismos, ilustrando polifilia e fluxo génico entre linhagens altamente diferenciadas do ponto de vista do DNA-mitocondrial, sugere que estas correspondem a espécies muito incipientes, com limites ainda

permeáveis e documenta a necessidade do estudo de múltiplos marcadores genéticos antes de reavaliações taxonómicas. Adicionalmente, estes resultados revestem-se de uma dupla importância para o conhecimento da biogeografia Ibérica e Magrebina: por um lado, ao confirmar a transversalidade de padrões que surgem como resposta às oscilações climáticas (tais como fragmentação glacial e expansão pós-glacial); por outro lado, ao documentar o quão específicos estes padrões podem ser, mesmo entre espécies muito aparentadas.

Résumé

Les lézards du genre *Podarcis* de la Péninsule Ibérique et de l'Afrique du Nord forment un complexe d'espèces cryptiques pour lequel différentes études de caractérisation morphologique ont donné des résultats contreversés en ce qui concerne l'identification des limites des entités évolutives. Récemment, la description de la variabilité génétique au niveau de l'ADN mitochondrial (mtDNA) a montré l'existence de divers groupes très différenciés et clairement associés à la géographie, quelques-uns correspondant à des entités suggérées par les études morphologiques. En raison de cette concordance, une révision taxonomique a été réalisée et cinq espèces sont reconnues à présent: *P. bocagei*, *P. carbonelli*, *P. vaucheri*, *P. hispanica* et *P. atrata*, même si cette dernière soit extrêmement similaire du point de vue génétique à un des différents clades détectés au sein de *P. hispanica*, une espèce typiquement politypique et, du point de vue mitochondrial, paraphylétique.

Basé sur ce scénario, nous avons établi comme premier objectif de cette thèse la caractérisation détaillée des patrons de différenciation génétique entre les formes du complexe *Podarcis*. D'un côté, nous avons essayé d'obtenir un arbre phylogénétique plus robuste pour le mtDNA de façon à tester les hypothèses biogéographiques décrites antérieurement. D'un autre côté, nous avons aussi essayé de compléter ces donnés avec l'information du génome nucléaire pour analyser de façon indépendante le degré de différenciation des lignées connues. En plus, nous avons évalué le degré d'isolement reproducteur entre les formes du complexe *Podarcis* à travers l'analyse des patrons de concordance entre mtDNA et gènes nucléaires, et de l'étude détaillée de la dynamique d'une zone de contact entre deux espèces.

Le ré-évaluation des relations phylogénétiques a été faite à travers la sélection de quelques individus de chaque lignée et du séquençage d'une plus grande région du mtDNA. En plus des deux gènes étudiés dans des travaux précédents, des fragments de trois autres régions ont aussi été séquencés et un total de 2425 pb (paires de bases) a été obtenu. L'analyse de ces donnés a permis de reconstruire un arbre phylogénétique robuste, où la grande majorité des clades montre des valeurs élevées de *bootstrap*. Cette rigoureuse évaluation de la phylogénie mitochondriale confirme quelques hypothèses antérieurement décrites, comme par exemple la paraphylie de *P. hispanica*. Cependant, quelques aspects de cette phylogénie diffèrent des résultats déjà connus, ce qui nous a mené à la proposition de scénarios alternatifs pour expliquer la colonisation de l'Afrique du Nord. En particulier, nous suggérons que l'ouverture du Détroit de Gibraltar a pu être la cause de l'isolement vicariant des formes ibériques et nord-africaines, et qu'une colonisation ultérieure s'est déroulée de l'Afrique du Nord vers la Péninsule Ibérique. Une importante contribution de cette étude a été aussi la description d'une lignée endémique du sud-est de la Péninsule, non détectée dans les études antérieures, ce qui constitue un exemple additionnel de la complexité évolutive de ce groupe d'espèces.

En ce qui concerne la description de la variabilité génétique au niveau du génome nucléaire, nous avons poursuivi deux voies de recherche: la description de la variation électrophorétique des allozymes et le séquençage des introns nucléaires. L'étude de la variabilité génétique au niveau des allozymes dans 30 populations représentatives de toutes les lignées mitochondriales a montré qu'en général, les groupés basés sur l'ADN mitochondrial se trouvent bien différenciés et que les différentes lignées observées en *P. hispanica* présentent entre elles un degré de différenciation similaire à celle observée entre les espèces correctes. Bien que ce fait corrobore la validité des lignées mitochondriales comme entités évolutives clairement distinctes, l'application de méthodes d'analyse récentes qui regroupent les individus basés sur leur génotype multilocus a montré

quelques patrons discordants qui peuvent résulter de l'occurrence de flux géniques entre les différentes formes. A ce niveau, deux individus très vraisemblablement hybrides entre *P. bocagei* et *P. hispanica* type 1, formes largement sympatriques, ont été détectés.

En opposition au haut degré de différenciation observé au travers de l'analyse du mtDNA et des protéines, l'étude de deux généalogies nucléaires (*β-fibint7* et *6-Pgdint7*) a révélé une surprenante absence de monophylie des différentes espèces et lignées mitochondriales. Il s'est alors avéré nécessaire de clarifier que ce patron polyphylétique avait résulté de la persistance d'un polymorphisme ancestral, incomplètement distribué dans les différentes lignées, ou alors d'un fort taux de flux géniques. Pour évaluer la vraisemblance de ces deux scénarios (qui ne sont pas mutuellement exclusifs), nous avons utilisé des méthodes d'analyse basés sur la théorie de la coalescence, qui permettent d'étudier la migration et la différenciation dans le même contexte. En accord avec cette méthodologie, nous avons détecté très peu de paires de lignées montrant des niveaux de mélange importants, ce qui suggère que la plupart des formes se trouvent reproductivement isolées. Dans ces conditions, et malgré le rôle significatif du flux génique dans l'évolution de ce complexe d'espèces (qui pourrait même être à l'origine de la perte d'identité nucléaire de certes lignées mitochondriales), le patron polyphylétique observé dans les généalogies nucléaires est surtout le résultat de la persistance de hauts niveaux de polymorphisme ancestral.

Ces études ont analysé la caractérisation des niveaux de flux géniques entre lignées d'une façon relativement simple à travers l'étude de la discordance entre patrons évolutifs révélés par des marqueurs mitochondriaux et nucléaires. Pour caractériser avec plus de détail les processus de mélange entre lignées de *Podarcis*, nous avons étudié une batterie de marqueurs nucléaires et mitochondriaux tout au long d'un transect traversant la zone de contact entre *P. bocagei* et *P. carbonelli.* Cette information a été complétée par des analyses morphologiques et de fertilité. En utilisant des méthodes de groupement individuel, nous avons montré une abondance d'individus qui présentent des signes de mélange dans l'endroit où les deux espèces coexistent. Cependant, nous observons aussi des déviations significatives aux équilibres de Hardy-Weinberg et de liaison. Cette bimodalité suggère l'existence de barrières qui empêchent le flux génique entre les deux espèces bien que la nature de ces barrières demeurent inconnues. Par exemple, il n'a pas été possible de démontrer une réduction de la fertilité des hybrides ni la vérification de la règle de Haldane. Pourtant, l'absence d'individus de morphologie intermédiaire, même d'individus qui montraient des signes de mélange, semble impliquer l'existence de forces sélectives qui jouent sur la morphologie. Ce type de zone hybride bimodal est normalement le signe d'un processus de spéciation pratiquement terminé. Même si ces résultats ne peuvent pas être généralisés à d'autres zones de contact, ils sont en accord avec des suggestions préalables indiquant que, malgré un isolement reproductif incomplet entre les lignées de *Podarcis*, le degré de flux génique entre elles est si faible qu'il n'empêche pas le processus de différenciation.

Un deuxième objectif de cette thèse était la description de la structuration phylogéographique intraspécifique en trois espèces de *Podarcis* à travers l'utilisation de marqueurs moléculaires multiples. En particulier, nous avons essayé d'évaluer la réponse de ces espèces aux oscillations climatiques qui ont caractérisé le Quaternaire en comparant les patrons observés avec les scénarios phylogéographiques décrits pour d'autres espèces réparties dans les mêmes régions et en interprétant les histoires évolutives contradictoires observées dans des différentes espèces de *Podarcis*.

L'étude de la variation génétique au niveau du mtDNA pour *P. bocagei*, *P. carbonelli* et *P. vaucheri*, trois espèces avec des distributions parapatriques qui se substituent le long d'un gradient latitudinal, nous a permis de tester deux hypothèses liées à l'influence

différentielle des oscillations climatiques du Quaternaire en rapport avec la latitude: i) on s'attend que les espèces réparties plus au nord présentent des faibles niveaux de diversité et de sous-structuration génétique quand elles sont comparées aux espèces présentes plus au sud, en raison des conditions climatiques pendant les glaciations ayant permis la persistance d'un petit nombre de régions d'habitat favorable dans le nord, probablement de petites dimensions et faible stabilité; ii) par contre, les espèces présentes plus au sud devront montrer des signes de stabilité démographique à long terme, en opposition avec les espèces réparties plus au nord qui, pour avoir persistées dans des petits refuges durant les périodes glaciaires, devront avoir subie une croissance démographique très rapide, étant le résultat de la colonisation post-glaciale de régions antérieurement inhospitalières. Les résultats obtenus montrent que *P. bocagei*, l'espèce avec une distribution plus septentrionale, présente des niveaux de diversité et de sous-structuration relativement faibles, une histoire coalescente récente (le temps de coalescence estimé est de 70000 a 100000 ans) et une évidence claire de croissance démographique. Par contre, *P. vaucheri,* l'espèce avec une distribution plus méridionale, présente une grande diversité génétique (même à une échelle réduite), une forte sous-structuration géographique, indiquant une stabilité démographique et un temps de coalescence ancien, coincidant avec le début du Pléistocène. *P. carbonelli*, une espèce avec une distribution intermédiaire entre les deux précédentes, présente des valeurs moyennes pour toutes les variables étudiées. Ensemble, ces résultats s'ajustent très bien aux prévisions initiales et montrent que des effets contrastants des glaciations en rapport avec la latitude peuvent être détectés à des échelles géographiques relativement petites.

L'analyse de marqueurs nucléaires (allozymes et un ensemble de microsatellites spécifiquement développés pour cette étude) dans des populations de *P. bocagei* et de *P. carbonelli* a confirmé les faibles niveaux de sous-structuration populationnel inférés pour les deux espèces à travers l'analyse du mtDNA. De plus, les niveaux de différenciation sont plus élevés pour *P. carbonelli* par rapport à *P. bocagei*, ce qui est aussi en accord avec le mtDNA en raison d'une persistance plus grande et d'un temps de divergence plus profond dans la première espèce. Pour *P. bocagei* nous avons détecté une perte progressive de variabilité génétique dans le sens sud-nord, cohérente avec une expansion géographique récente à partir d'un refuge, ce qui est en accord avec un scénario d'expansion démographique inféré à partir du mtDNA. Par ailleurs, l'utilisation de l'interpolation spatiale des données génétiques et l'application d'une méthode Bayesienne de groupement individuel ont permis l'identification des voies probables d'expansion géographique et suggèrent avec un détail considérable la région qui a pu avoir fonctionné comme refuge pour l'espèce pendant les dernières glaciations. L'application d'une méthodologie identique pour *P. carbonelli* a montré un degré modéré d'association entre variabilité génétique et géographie, mais une correspondance partielle en rapport avec les groupes inférés au travers du mtDNA; de façon à évaluer ces patrons discordants, nous émettons l'hypothèse que des fragmentations récentes dans la distribution de l'espèce pourront avoir conduit à la perte des signatures génétiques typiquement trouvées dans les espèces co-distribués, et que les patrons observés à présent résultent plus de phénomènes de dérive génétique récente que d'oscillations climatiques du Pléistocène.

Dans l'ensemble, ces résultats constituent une avancée significative dans l'état des connaissances préalables des relations évolutives, des états de différenciation, des niveaux de flux génique et des patrons de sous-structuration intraspécifique chez les lézards du genre *Podarcis* dans la Péninsule Ibérique et l'Afrique du Nord. D'un point de vue plus général, l'étude à large échelle de la variation de gènes nucléaires dans ces organismes, illustrant la polyphylie et des flux géniques entre lignées fortement différenciés du point de vue du mtDNA, suggère qu'elles puissent correspondre à des espèces naissantes, avec des

limites encore très perméables, et met en évidence aussi la nécessité de l'étude de marqueurs génétiques multiples avant toute ré-évaluation taxonomique. En plus, ces résultats ont une importance double pour la connaissance de la biogéographie Ibérique et Maghrébine: d'une part, parce qu'ils confirment la transversalité de patrons qui apparaissent en réponse aux oscillations climatiques (comme la fragmentation glaciaire et l'expansion post-glaciaire); d'autre part, parce qu'ils documentent la spécificité de ces patrons, même s'ils sont étudiés entre espèces phylogénétiquement très proches.

Chapter 1

General introduction

General introduction

Understanding the mechanisms that are responsible for the isolation of populations, genetic differentiation and ultimately speciation is one of the main aims of evolutionary biology. In recent years, genetic tools have been widely used to clarify evolutionary relationships between species, relate genetic variation to specific geographical or geological events, describe population structure, make inferences on past and present demographic dynamics and a wide number of other applications. One of the most important contributions has been that of uncovering hidden genetic variation among morphologically cryptic organisms, as exemplified by the species complex that is the object of this study, *Podarcis* wall lizards. The initial motivation for the work presented in this thesis was indeed related to this basic desire of understanding the diversity within this group of organisms. However, it soon became clear that *Podarcis* could be a useful model for the study of evolutionary processes, from the dynamics of the responses to geological or climatic events to the acquisition of reproductive isolation between closely-related taxa. In other words, by describing the patterns that are presently observed, we also hope that this work will be a valuable contribution towards understanding the processes that generated those patterns.

1.1. The Western Mediterranean: a complex biogeographic set point

It is now well known that major geological and climatic events can be responsible for vicariance, isolation and consequently differentiation of taxa (Avise 2000). The Western Mediterranean region, situated at the convergence of two major continental areas (Europe and Africa), has had a particularly eventful and diverse geological and climatic history, which is reflected in a present-day high number of endemisms and complex intraspecific phylogeographic structures. For their implications in the research presented in this thesis, we will focus this review on events that occurred in the past 15 million years and address three types of phenomena: those related to the evolution of the Strait of Gibraltar area, those that are related to the uplift of mountains and formation of river basins and Pleistocene climatic oscillations.

1.1.1. Geological evolution of the area surrounding the Strait of Gibraltar

During Alpine orogenesis, in the Oligocene and Early Miocene, a wide zone in the interface between Africa and Europe underwent extension, governed by subduction rollback. This resulted from a slower rate of convergence between the Eurasian and African plates starting around 30 million years ago (Mya) (Rosembaum *et al.* 2002). As a result of subduction rollback, extension in the Early Miocene led to the breakup and drifting of continental fragments formerly attached to southern France and Iberia, which are now scattered throughout the Western Mediterranean (the Betic region, in the Iberian Peninsula, the Rif and the Kabylies in North Africa, and also the Balearic Islands, Sardinia, Corsica and part of the Italian Peninsula (Calabria); Lornegan and White 1997, Rosembaum *et al.* 2002). The land mass now forming the Betic and Rif regions eventually collided with South Iberia and North Africa ca. 15 Mya (Weijermars 1991), promoting a temporary connection between both continents, which provided a land corridor for biotic dispersal (Figure 1.1A). This has been suggested to have happened in *Pleurodeles* ribbed salamanders from Iberia to North Africa (Veith *et al.* 2004), although this is not consensual (for example, based on a different molecular clock calibration, Carranza and Arnold (2004) claim that this dispersal did not occur until the Messinian salinity crisis), and in *Buthus* scorpions in the reverse direction (Gantenbein and Lagiardèr 2003). In both genera, a vicariant event 14-15 Mya probably resulting from the disjunction of the Betic region from Iberia caused allopatric divergence from then on and is responsible for present-day high genetic distances between Iberian and North African populations. Additional evidence for Iberian-North African dispersal and subsequent isolation around this time is given by a palaeogeographic study on beetles (Palmer and Cambefort 2000).

There is also evidence for later events of allopatric isolation in that area, related to the opening of the Betic marine corridor (~10-8 Mya; Figure 1.1B) or to the fragmentation of the Betic region (~8-6 Mya), suggesting that this area functioned as a focal point of speciation and endemism. This pattern has been invoked to explain differentiation among species of fire salamanders (*Salamandra;* Steinfarz *et al.* 2000, but see Escoriza *et al.* (2006) for a contrasting scenario), painted frogs (*Discoglossus* spp.; García-Paris and Jockush 1999, Fromhage *et al.* 2004, although this phenomenon was inferred to cause divergence between different sets of species; see also Zangari *et al.* 2006 for yet another different interpretation) and midwife toads (*Alytes*), although different molecular clock calibrations suggest distinct timings and patterns of allopatric differentiation in this genus (Fromhage *et al.* 2004; Martínez-Solano *et al.* 2004). In *Alytes*, this scenario explains why species of this genus inhabiting

southern Iberia are more closely related to North African species than to other Iberian counterparts, since parts of this landmass became later connected to both Iberia and North Africa.

Other possible connections between southern Iberia and land masses in the Gibraltar region are thought to have occurred during the Late Tortonian, ca. 7.8-7.6 Mya, when a salinity crisis led to the deposition of sediments in the eastern Betics region (Krijgsman *et al.* 2000). A similar event occurred in the Rif region 6.7 – 6 Mya (Krijgsman and Langereis 2000). Examples of a putative speciation event related to the formation of both of these land bridges are given by Paulo (2001), in lizards of the *Lacerta lepida* species complex.

One of the most dramatic events in the geological history of the Mediterranean was the Messinian salinity crisis (Hsü 1977, Krijgsman *et al.* 1999, Duggen *et al.* 2003, Rouchy and Caruso 2006). Around 5.96 Mya (Krijgsman *et al.* 1999), the connection between the Mediterranean Sea and the Atlantic, previously ensured by the Betic and Rif marine corridors, was closed (Figure 1.1C). The causes for this remain obscure, although several scenarios have been hypothesized, including tectonic uplifts and a decrease of sea level caused by a world-wide glacial stage (Duggen *et al.* 2003, Rouchy and Caruso 2006). Without the input from Atlantic waters, a succession of evaporitic events caused a rapid desiccation of the Mediterranean basin. Towards the end of the Messinian, a geodynamical event, probably related to faulting along the Gibraltar arch, abruptly reopened the connections with the Atlantic Ocean through the Gibraltar gate ca. 5.33 Mya, roughly at its present location, definitively separating the Betic and Rif areas, which became part of Iberia and North Africa, respectively. The replenishment of the Mediterranean is thought to have occurred extremely rapidly at a geological scale (about 100 years). During the salinity crisis, a land bridge existed between Iberia and North Africa which allowed the dispersal of terrestrial organisms; when the Mediterranean refilled, the Strait of Gibraltar became a barrier for biotic dispersal, separating inhabiting taxa into allopatric units.

Despite periodic episodes of lower sea levels, the Strait of Gibraltar has remained practically constant since its formation. In this context, several studies have attempted to quantify the role of this barrier in shaping genetic divergence of closely related taxa inhabiting its margins. The classic study of Busack (1986) compared genetically amphibian and reptile populations distributed on both sides of the Strait, showing a remarkable variability on the observed genetic distances across it and suggesting that the effectiveness of this barrier was taxon-dependent. Many species divergences were inferred to have an origin coinciding with the opening of the Strait of Gibraltar: *Acanthodactylus* lizards (Harris *et al.* 2004), *Pleurodeles* salamanders (Carranza and Arnold 2004; but see Veith *et al.* 2004), *Blanus* worm lizards (Vasconcelos *et al.* 2006),

Salamandra salamanders (Escoriza *et al.* 2006; but see Steinfarz *et al.* 2000), *Discoglossus* frogs (Zangari *et al.* 2006; but see Fromhage *et al.* 2004), *Lacerta* lizards (Paulo 2001) and *Podarcis vaucheri* (Busack *et al.* 2005; but see Harris *et al.* 2002b). Unexpectedly, despite its narrowness (~14Km), the Strait of Gibraltar also seems to constitute an effective barrier against gene flow between deep evolutionary lineages in volant organisms like bats (*Myotis*, Castella *et al.* 2000; *Plecotus*, Juste *et al.* 2004)) or birds (*Ficedula* flycatchers, Saetre *et al.* 2001; *Parus* blue tits, Salzburger *et al.* 2002) although the actual role played by the opening of the Strait in the origin of these species groups remains unclear.

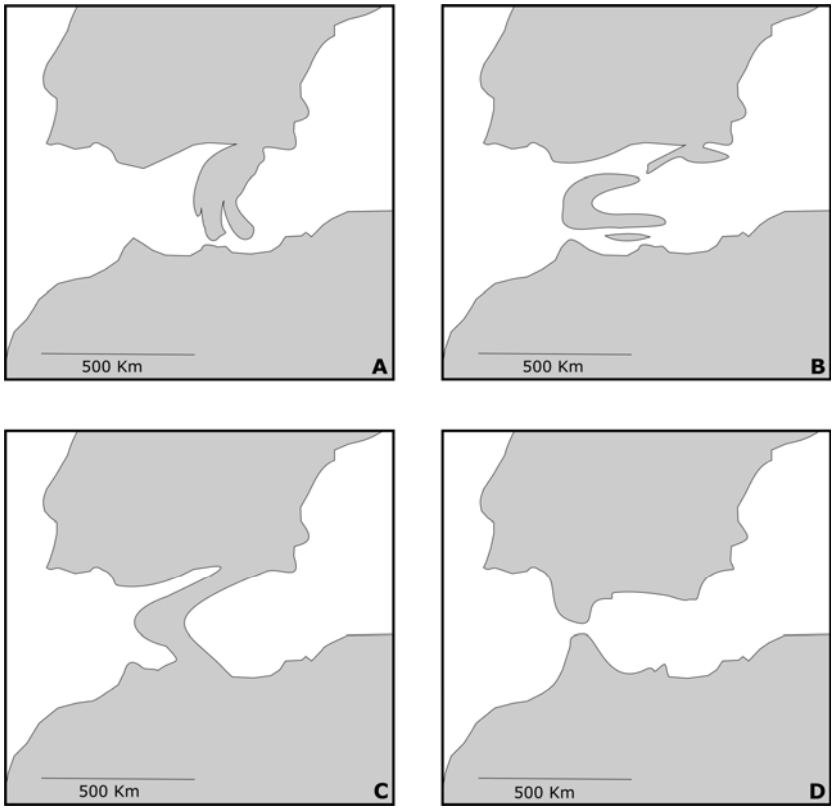

Figure 1.1. Reconstruction of the Western Mediterranean, from 15 Mya to the present. **A. 15Mya**: the Betic-Rif land masses establish a transient land connection between Iberia and North Africa. **B. 10–8Mya**: fragmentation along this massif leads to the formation of the Betic Strait. Subsequent fragmentation of these land masses promoted the appearance of several islets. **C. 5.96-5.33Mya**: during the Messinian salinity crisis, Iberia and North Africa were again connected. This connection abruptely ended when the Strait of Gibraltar opened at its definitive location. **D. present configuration.** Adapted from Weijermars (1991), Rosembaum *et al.* (2002) and Duggen *et al.* (2003).

However, a significant number of terrestrial species have been shown to have crossed the Strait after its formation, presenting closely-related forms on both

sides of the Strait. This pattern has been recognised in salamanders (*Pleurodeles waltl*, Veith *et al.* 2004, Carranza and Arnold 2004), lizards (*Podarcis* spp., Harris *et al.* 2002b; *Psammodromus algirus*, Carranza *et al.* 2006a), snakes (*Macroprotodon brevis,* Carranza *et al.* 2004b; *Malpolon monspessulanus,* Carranza *et al.* 2006b; *Hemorrhois hippocrepis*, Carranza *et al.* 2006b), worm lizards (*Blanus cinereus*, Vasconcelos *et al.* 2006), tortoises (*Testudo graeca*, Álvarez *et al.* 2002), terrapins (*Emys orbicularis*, Lenk *et al.* 1999; *Mauremys leprosa,* Fritz *et al.* 2006), chamaleons (*Chamaleo chamaeleon,* Paulo *et al.* 2002a), larks (*Galerida cristata*, Guillaumet *et al.* 2006), woodmice (*Apodemus sylvaticus*, Michaux *et al.* 2003) and shrews (*Crocidura russula,* Cosson *et al.* 2005). Although some of these could represent human-mediated introductions (particularly probable are the cases of chamaleons and tortoises (Paulo *et al.* 2002a, Álvarez *et al.* 2000)), in other cases such a hypothesis is difficult to corroborate, either because the species involved are usually secretive and not related to human settlements (particularly the cases of snakes and worm lizards) or because the species display considerable variation and form monophyletic lineages on both sides of the Strait, suggesting a relatively more ancient colonization (in *Podarcis* lizards or *Apodemus* woodmice, for example). Carranza *et al.* (2006b) suggested that some of these recent colonization events might have occurred during Pleistocene glaciations, when sea levels at the Strait of Gibraltar dropped by approximately 130m, probably exposing temporary islands that could have provided a colonization route for some species.

1.1.2. Mountain uplift and formation of major river basins

Geographical barriers other than the Strait of Gibraltar may have played an equally important role in causing allopatric fragmentation or preventing gene flow among Western Mediterranean taxa. Examples of such barriers are mountain chains or river basins.

Most of the orogenic events shaping the Western Mediterranean landscape occurred during the Oligocene-Miocene boundary. However, during mid-late Miocene, the compression of the African and Eurasian plates further left to the uplift of more recent mountain ranges. An example of this is the formation of the Moroccan Atlas. This process has been invoked to explain differentiation between major lineages within the *Agama impalearis* species complex (Brown *et al.* 2002). Although the origin of its phylogroups is probably not related to this uplift, the role of the Atlas as a barrier to gene flow was also evidenced in a recent study of genetic variation in terrapins *Mauremys leprosa* (Fritz *et al.* 2006). Another important uplift occurring during this period was the shaping of the Neo-Pyrenees (~10Mya), an event that is thought to have caused the

vicariance of, for example, *Alytes obstetricans almogavarii* from other *A. obstetricans* (Fromhage *et al.* 2004, Martínez-Solano *et al.* 2004), of major lineages of *Discoglossus* (Fromhage *et al.* 2004; again, see Zangari *et al.* 2006 for a discordant perspective) and also of the genera *Triturus* and *Calotriton* (Carranza and Amat 2005).

Most of the modern river basins in Iberia and North Africa were formed more recently, during the Pliocene. As expected, the separation of different watersheds is attested in concordant genetic differentiation detected between several ichthyofaunal species (e.g. Carmona *et al.* 2000, Machordom and Doadrio 2001, Robalo *et al.* 2006). However, rivers often constitute important barriers for the dispersal of terrestrial organisms as well, in particular those with limited dispersal abilities. Although other rivers have been described as barriers to gene flow between taxa (Alexandrino *et al.* 2000) for their well-documented importance two rivers stand out: the Guadalquivir, in Iberia, and the Moulouya, in North Africa. Whether or not the formation of this basin was the trigger for differentiation, several species are believed to be, or to have been in some point of their history, restricted to the area located south of the Guadalquivir. These are the cases of the painted frog *Discoglossus jeanneae* (Martínez-Solano 2004, Fromhage *et al.* 2004), the midwife toad *Alytes dickhillenii* (Arntzen and García-Paris 1995, Martínez-Solano *et al.* 2004), the salamander subspecies *S. salamandra longirostris* (García-Paris *et al.* 1998), the rabbit subspecies *Oryctolagus cuniculus algirus* (Branco *et al.* 2000, 2002), the Scots pine subspecies *Pinus sylverstris nevadensis* (Sinclair *et al.* 1999) and probably also distinct unnamed phylogroups within the lizard *Psammodromus algirus* (Carranza *et al.* 2006a), the salamander *Pleurodeles waltl* (Carranza and Arnold 2004), the worm lizard *Blanus cinereus* (Vasconcelos *et al.* 2006), the ocellated lizard *Lacerta lepida* (Paulo 2001) and white oaks *Quercus* (Petit *et al.* 2002). In some cases there is a remarkable coincidence of estimated divergence times of around 2-3 million years, coinciding with the Pliocene formation of the river basin (in the valley resulting from the closing of the former Betic Strait). With respect to the Moulouya, it has been suggested as a primary barrier acting against gene flow between *Discoglossus pictus* and *D. scovazzi* (Zangari *et al.* 2006), *Lacerta pater* and *Lacerta tangitana* (Paulo 2001), and between distinct lineages within *Testudo graeca* (Álvarez *et al.* 2000) and *Salamandra algira* (Escoriza *et al.* 2006).

1.1.3. Pleistocene glaciations

Several lines of evidence suggest that the global climate suffered profound alterations since the Plio-Pleistocene boundary. Recent studies regarding levels

of oxygen and carbon isotopes, magnetic and CO_2 measures, animal, vegetable and mineral remains in cores from the seabed, land and ice, are concordant in providing strong support for the hypothesis that the earth's climate has been dominated by major glacial periods interrupted by relatively short interglacial stages (reviewed in Hewitt 1996). Although the boundaries of this glacial period are controversial (see for example Zagwijn 1992a), the first large-scale glaciation can be dated at 2.4 - 2.3 Mya (Webb and Bartlein 1992, Zagwijn 1992a).

The effects of such climatic changes are believed to have had a major impact on species survival and distribution. In Europe, during the Ice Ages, many species such as large carnivores were condemned to extinction because small patches of available habitat no longer sustained viable populations (e.g. O'Regan *et al.* 2002). Pollen and fossil data provide strong support for a constant pattern of migration of fauna and flora species tracking suitable habitat change (Zagwijn 1992a,b): during cold stages, temperate taxa tended to migrate south and expanded northwards during warm interglacials. Palaeontological, palynological and palaeolimnological studies on southern European Peninsulas, including the Iberian Peninsula, suggest that some areas within these regions might have been relatively ecologically stable throughout the Quaternary, allowing the persistence of temperate species (e.g. Tzedakis *et al.* 2002, Roucoux *et al.* 2005). The importance of southern European peninsulas as survival refugia for temperate biota is further documented by their high level of endemism and by phylogeographic studies on inhabiting organisms, that also identify such regions as sources for post-glacial recolonization of Northern areas (reviews in Taberlet *et al.* 1998 and Hewitt 1999; more recent examples can be found for example in Griswold and Baker 2002; Michaux *et al.* 2003, 2005; Schmitt *et al.* 2003, 2005; Brito 2005).

However, the same features that made these peninsulas favourable regions for the survival of populations throughout the Pleistocene (high physiographic complexity and multiplicity of climatic influences) also make it unlikely that this survival occurred at single homogenous and continuous refugium (Gómez and Lunt 2007). In accordance with this expectation, recent work has disclosed complex intraspecific phylogeographic patterns within each of these regions, and the Iberian Peninsula constitutes no exception. As in above-mentioned European-wide studies (Hewitt 1996), the genetic consequences of the climatic oscillations in Iberia can be summarized by a set of common patterns: i) complex phylogeographic structures related to isolation in distinct glacial refugia, with varying degrees of genetic divergence related to the period during which populations remained isolated; ii) signatures of post-glacial demographic growth and range expansions; iii) establishment of secondary contact zones with

various degrees of admixture as a result of range expansions from glacial refugia.

Genetic differentiation resulting from isolation in distinct glacial refugia is the pattern that is often easier to infer from a general description of mitochondrial DNA variation, which currently constitutes the main tool for addressing most of the phylogeographic questions. Therefore, complex phylogeographic structures that can be related to Pleistocene climatic oscillations have been described in a wide variety of species, from small plants to trees, in invertebrates and nearly every group of vertebrates. Among the species that have been described to show such complex structures are the ragwort *Senecio gallicus* (Comes and Abbot 1998), the brown trout *Salmo trutta* (García-Marin *et al.* 1999; Machordom *et al.* 2000), the European rabbit *Oryctolagus cuniculus* (Branco *et al.* 2000, 2002), the golden-striped salamander *Chioglossa lusitanica* (Alexandrino *et al.* 2000, 2002), Schreiber's green lizard *Lacerta schreiberi* (Paulo *et al.* 2001, 2002), the viviparous lizard *Lacerta vivipara* (Surget-Groba *et al.* 2001), the white oak *Quercus* (Olalde *et al.* 2002), the field vole *Microtus agrestis* (Jaarola and Searle 2002), the painted frog *Discoglossus galganoi* (Martínez-Solano 2004), the butterfly *Erebia triaria* (Vila *et al.* 2005), the natterjack toad *Bufo calamita* (Rowe *et al.* 2006), Bosca's newt *Lissotriton boscai* (Martínex-Solano *et al.* 2006) and several species of diving beetles (Dysticidae; Ribera and Vogler 2004). Although the inferred location of putative refugia shows in some cases remarkable similarities among species (reviewed in Gómez and Lunt 2007), these studies also point to obviously different trends, suggesting that the response to the glacial ages was significantly distinct across species. In some of these and other species, signatures of a rapid post glacial expansion from one or more Iberian refugia have also been inferred directly from genetic data (Branco *et al.* 2002, Alexandrino *et al.* 2000, 2002, Paulo *et al.* 2001, Surget-Groba *et al.* 2001, Olalde *et al.* 2002, Martínez-Solano 2004, Martínez-Solano *et al.* 2006).

The detection of the third pattern, the formation of suture zones resulting from secondary contact of forms that were restricted during glacial stages and that subsequently expanded their distribution after the climate amelioration is dependent upon detailed sampling schemes because of the narrowness of many such zones. Therefore only a few well documented cases have been reported (Branco *et al.* 2000, 2002, Alexandrino *et al.* 2000, 2002, Surget-Groba *et al.* 2001, Godinho *et al.* 2006) and even fewer have been studied in detail (Sequeira *et al.* 2005, Pereira 2005, Geraldes *et al.* 2006, Godinho *et al.* 2006, Sequeira 2006). Nevertheless, these few cases clearly illustrate different dynamics, from high levels of neutral admixture to hints of incipient speciation (see also Ribera and Vogler 2004), suggesting that confinement in glacial refugia led to varying degrees of reproductive isolation.

In North Africa, there is relatively little information on how habitats and organisms reacted to climatic changes. Pollen and lake data suggest that during colder phases there was an increase of shrub and grass vegetation characteristic of drier climates and a decrease of forest-type vegetation (Jolly *et al.* 1998). There are only a few phylogeographic studies documenting vicariant isolation caused by such climatic changes (in land snails *Helix* (Guiller *et al.* 2001), *Agama* lizards (Brown *et al.* 2002), *Pleurodeles* salamanders (Carranza and Arnold 2004), *Crocidura* shrews (Cosson *et al.* 2005), and *Galerida* larks (Guillaumet *et al.* 2006)).

1.2. The lizards of genus *Podarcis*

1.2.1. Distribution, biology and taxonomy

Podarcis (Wagler 1830) is a genus belonging to the family Lacertidae, which in turn belongs to the order Squamata. It represents one of the most abundant, conspicuous and widely distributed reptile genera in Europe and North Africa (see figure 1.2), and has a mostly circum-mediterranean distribution. All species belonging to the genus are diurnal and small body sized lizards that share a high degree of morphological and ecological similarity.

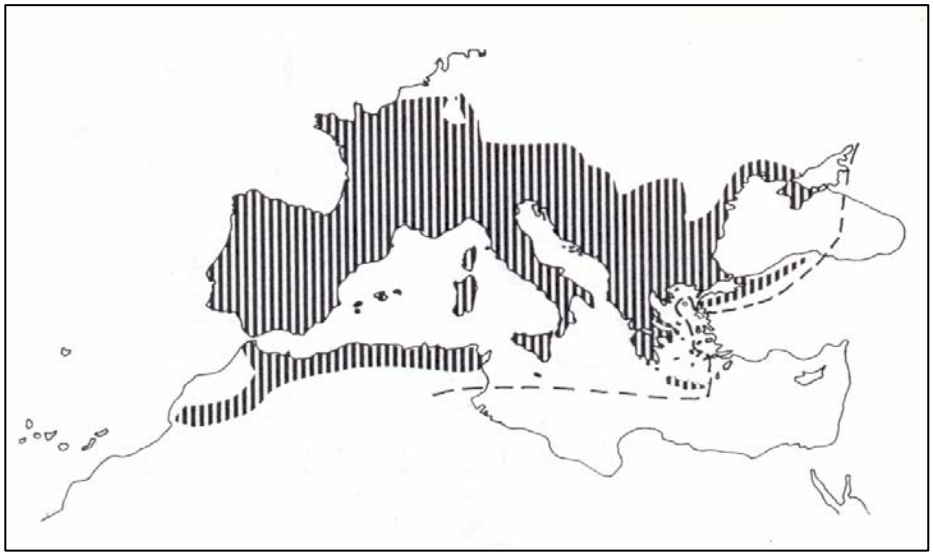

Figure 1.2. Distribution of genus *Podarcis*. Reproduced from Arnold (1973).

It took until the early 1970's, with Arnold's (1973) thorough morphological revision of the Palaearctic Lacertidae, for the generic status of *Podarcis* to be widely accepted by herpetologists. This study clearly showed that members of this genus share a consistent set of derived (although not unique) features. Prior to this date, specimens of the genus were included in *Lacerta,* although their distinctiveness as a subgenus had been proposed by previous authors (Boulenger 1916, Mertens and Wermuth 1960, Böhme 1971).

Nineteen species of *Podarcis* are presently recognised: *P. atrata*, *P. bocagei*, *P. carbonelli, P. erhardii*, *P. filfolensis*, *P. gaigeae, P. hispanica, P. lilfordi, P. melisellensis, P. milensis, P. muralis, P. peloponnesiaca, P. pityusensis, P. raffonei, P. sicula, P. taurica, P. tiliguerta P. vaucheri* and *P. wagleriana*. Some authors (e.g. Böhme 1986, Gasc *et al.* 1997) also include in the genus the Menorca wall lizard *Lacerta perspicillata* and the Madeiran wall lizard *L. dugesii* under a separate subgenus (*Teira*), which are placed as sister taxa to *Podarcis* in phylogenetic analyses based on morphology and molecular data (Harris *et al.* 1998, Fu *et al.* 2000).

Podarcis are a good example of unstable and controversial taxonomy. Not only relationships between species have been difficult to determine, but species delimitation on the basis of morphology itself has been a matter of debate. The main reasons for this are an overall uniform morphology reflected in the lack of variable characters for use in phylogenetic reconstruction, extensive intraspecific variation obscuring species differentiation, lability of structural features (leading to character homoplasy) and the lack of consistent outgroup information to understand the polarity of morphological changes (Arnold 1990). Because of this, although species relationships have been the object of speculation, formal phylogenetic reconstructions from morphology have rarely been attempted.

After several partial attempts to document species relationships based on allozyme variation (e.g. Mayer and Tiedemann 1980, Capula 1994a,b) a recent plethora of mitochondrial DNA sequencing studies focusing on the whole genus (Oliverio *et al.* 1998, Harris and Arnold 1999, Oliverio *et al.* 2000, Carranza *et al.* 2004a, Harris *et al.* 2005) or on particular geographical groups (Castilla *et al.* 1998, Harris and Sá-Sousa, 2001, 2002; Podnar *et al.* 2004, 2005; Poulakakis *et al*. 2003, 2005a,b, Terrasa *et al.* 2004) have clarified some aspects of the intrageneric phylogeny of *Podarcis*. In some aspects, these studies produced largely conflicting results, suggesting that the systematics of the genus would benefit from a large-scale phylogenetic study. Nevertheless, a few clades are consistently recovered: a Balearic clade, consisting of *P. lilfordii* and *P. pityusensis*; a Balkan clade, including *P. erhardii*, *P. taurica, P. milensis, P. gaigeae* and *P. melisellensis* (but the position of *P. peloponnesiaca* within or outside this clade remains controversial, see Carranza *et al.* 2004a, Harris *et al.* 2005); an Iberian/Maghrebin clade comprising *P. bocagei, P.carbonelli*, *P.*

hispanica, P. atrata and *P. vaucheri* (although analyses by Poulakakis *et al.* (2003) place *P. filfolensis* within this clade). Phylogenetic relationships among these clades and among other species of the genus remain, however, poorly known. Oliverio *et al.* (2000) suggested that radiation of wall lizards into the various clades occurred between 10–16 Mya, dates which are corroborated by other studies (Poulakakis *et al.* 2005a).

Although there has been relatively little effort in clarifying this issue, hybridization between distinct species of the genus has been shown to occur. Relying on morphological evidence, Arnold (1973) and Bischoff (1973) described the existence of apparent hybrids between *P. erhardii* and *P. muralis* and between the latter and *P. melisellensis*, as well as between *P. sicula* and *P. wagleriana* in areas of sympatry. Many other species pairs were shown to be able to produce viable hybrids in captivity. More recently, using protein electrophoretic techniques, Capula (1993, 2000) has confirmed natural hybridization between *P. sicula* and *P. wagleriana* and between the former and *P. tiliguerta*. It should be noted that according to all the phylogenetic studies involving these species (Harris and Arnold 1999, Oliverio *et al.* 2000, Carranza *et al.* 2004a, Harris *et al.* 2005), they do not seem to be closely related, implying that *Podarcis* species remain permeable to gene flow with congenerics after a long time of divergence.

1.2.2. The Iberian and North African clade of *Podarcis*

1.2.3.1. A controversial species complex

Currently, six distinct species are recognised in the Iberian Peninsula and North Africa: *P. atrata, P. bocagei, P. carbonelli, P. hispanica, P. vaucheri* and the less closely related *P. muralis*. Apart from *P. muralis*, which can be easily diagnosed on the basis of morphology, the delimitation of Iberian and North African species was a long, debated and controversial process, since successive studies disagreed on the number, distribution and proposed taxonomic rank of evolutionary entities (Mertens and Muller 1940, Klemmer 1959, Salvador 1974). The first taxonomic reevaluation that was actually widely accepted was Arnold and Burton's (1978) formal separation of *P. bocagei* from *P. hispanica* (although in 1989 Crespo and Oliveira still used the term "*P. bocagei-P. hispanica* complex" to refer to Portuguese forms of *Podarcis*, which is illustrative of the difficulties in distinguishing between both species). Studies attempting to provide accurate morphological, ecological and corological characterization of the two species soon followed (Pérez-Mellado 1981a, Galán 1986, Pérez-Mellado and Galindo 1986, Guillaume 1987).

However, instead of helping in establishing guidelines for the identification of the two species, the pattern that arose was that both could be divided into numerous entities. In *P. bocagei,* this is illustrated by the description of novel subspecies, such as *P. b. carbonelli*, primarily considered an endemism of the Spanish Central System (Pérez-Mellado 1981b) and, later, *P. b. berlengensis* from the Berlengas Archipelago (Vicente 1985). According to several publications (Pérez-Mellado 1981a, Galán 1986, Pérez-Mellado and Galindo 1986, Barbadillo *et al.* 1999), the distribution of *P. b. bocagei*, mostly located in the Northwest of the Iberian Peninsula would reach as far south as Lisboa along a narrow strecht of coastline (Figure 1.3A). The division of *P. bocagei* into these subspecies and their proposed distribution were consensually accepted, although controversy over the distribution of *P. b. carbonelli* was raised when Magraner (1986) detected this form in the Doñana National Park, in southern Spain. Because this information was in clear disagreement with the distribution maps accepted at the time, Magraner's report was generally discredited.

The description of subspecies within *P. hispanica* was not as consensual. Pérez-Mellado and Galindo (1986) demonstrated that this species was highly morphologically variable, both at the intra and interpopulation level, but the authors did not find geographic congruence in the distribution of morphological variability. Consequently, they recommended a conservative split of the species into *P. hispanica hispanica,* in the Iberian Peninsula and France, and *P. hispanica vaucheri,* in North Africa (Figure 1.3A). In contradiction with this view, Guillaume (1987) suggested the division of *P. hispanica* into eight different morphotypes, designated both as "forms" and subspecies (when there were previously available names), with apparently non overlapping geographic distributions (Figure 1.3B). Because Guillaume's work was not widely divulged, the conservative view of Pérez-Mellado and Galindo was adopted by subsequent publications (e.g. Barbadillo *et al.* 1999).

This system prevailed for over a decade, until Sá-Sousa (1995, 2001) performed a systematic revision of portuguese forms of *Podarcis.* These studies confirmed the distinctiveness of the two continental subspecies within *P. bocagei*, but the suggested distribution maps (Sá-Sousa 1998, 1999, 2000a) strongly disagreed with previous reports. For example, *P. b. carbonelli* which was supposedly restricted to the Spanish Central System, was found in several localities in the Portuguese side of this mountain system and also along the western coast from the locality of Espinho to the Algarve, giving new credit to the early reports of the presence of *P. b. carbonelli* in southern Spain (Magraner 1986). By contrast, *P. b. bocagei* was found to have a more restricted distribution, located mostly north of the Douro river and only penetrating territories to the south in a small coastal area (Figure 1.3C).

With respect to *P. hispanica*, Sá-Sousa (1995, 2001) and Sá-Sousa *et al.* (2002) described two morphotypes with possibly parapatric distributions, which were given the provisional designations "type 1" and "type 2". Type 1 was found mainly in northern Portugal, existing in sympatry with *P. b. bocagei*, north of Douro river, and with *P. b. carbonelli,* south of this river. Type 2 was found in southern Portugal, reaching however higher latitudes in a coastal fringe with Mediterranean climate. This division of portuguese *P. hispanica* into two morphotypes disagreed both with the works of Pérez-Mellado and Galindo (1986) and Guillaume (1987).

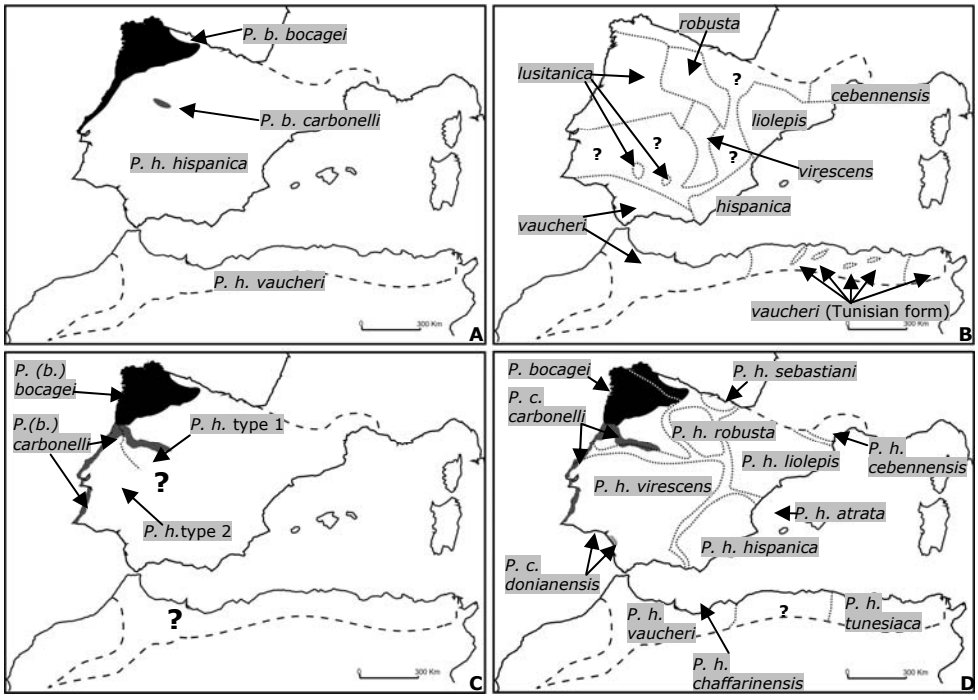

Figure 1.3. Four different proposals for the partition of *Podarcis* from the Iberian Peninsula and North Africa into different morphotypes. **A.** Pérez-Mellado and Galindo (1986). **B.** Guillaume (1987) (only *P. hispanica*). **C.** Sá-Sousa (1998, 1999, 2000a,b). **D.** Geniez (2001). Dashed lines indicate the distribution limits of the species complex. Dotted lines indicate the limits of different subspecies/forms of *P. hispanica*. The distribution of *P. (b.) bocagei* is presented in black and that of *P. (b.) carbonelli* in grey. Question marks indicate areas not surveyed by the authors.

More recently, Geniez (2001) analysed *P. hispanica* populations from the complete distribution area. Despite some concordant features with the proposals previously put forward by Guillaume (1987) and Sá-Sousa (2000b, 2001), this work shows remarkable differences in the division of *P. hispanica* into morphotypes and their distribution (Figure 1.3D), definitively showing that morphological analyses alone could not fully resolve the complexity underlying

variation within this species. Additionally, Geniez (2001) described a novel subspecies within *P. carbonelli*, which he names *P. c. donianensis*, endemic from southern Iberian Peninsula (validating the observations of Magraner in 1986).

1.2.3.2. Evolutionary genetics of Iberian and North African wall lizards and recent taxonomic reevaluations

The first applications of genetic tools to Iberian *Podarcis'* taxonomy were partial because they were intended to solve particular evolutionary problems rather than the taxonomy of the whole group. For example, Guillaume (1977) and Béa *et al.* (1986) characterized Iberian *P. muralis* and *P. hispanica* based on allozyme markers and then used this information to assign the Monte Urgull (Basque Country, Spain) population of *Podarcis*, previously described as *P. muralis sebastiani* Klemmer 1964, to *Podarcis hispanica.* Also using electrophoretic techniques, Busack (1986) and Capula (1997) studied differentiation of *Podarcis* lizards from both sides of the Strait of Gibraltar reaching surprisingly different results: the first author found a low level of genetic divergence consistent with intraspecific variation, whereas Capula (1997) reported high genetic distances, suggesting the occurrence of two different species. In 1993, Almeida reported low genetic differentiation of *P. bocagei berlengensis* from mainland forms of the species using protein electrophoretic data. Later, Sá-Sousa *et al.* (2000) used the same data set, along with information from morphology and ecology, to first propose the elevation of *P. b. carbonelli* to the species level. Based on mitochondrial DNA sequencing, Castilla *et al.* (1998) suggested the elevation to the species status of the insular lizard *Podarcis hispanica atrata* from Columbretes Islands, based on high divergence from mainland forms. *P. atrata* is now a fully recognised species which has been given special conservational attention (Castilla 2002). Also on the basis of high mitochondrial DNA divergence from Iberian lizards, consistent with previous findings by Capula (1997), Oliverio *et al.* (2000) suggested that North African forms of *Podarcis* deserved the species status (as *Podarcis vaucheri*). However, this study involved a single individual from each side of the Strait, which was obviously insufficient for this recommendation to be followed.

Despite their contribution to the knowledge on genetic variation within Iberian and North African *Podarcis,* these studies raised more questions than the ones they answered, mainly because of their partial approach to the problem. A more widespread sampling scheme was accomplished in the works of Harris and Sá-Sousa (2001, 2002). These studies sampled individuals from various representative localities in Iberia and North Africa and confirmed through

mitochondrial DNA sequencing the genetic distinctiveness of the four Portuguese forms (*P. b. bocagei*, *P. b. carbonelli*, *P. hispanica* type 1 and *P. hispanica* type 2) described by Sá-Sousa (1995, 2001). Furthermore, the authors reported the existence of two other divergent cryptic lineages within the Iberian Peninsula, one in the southeast and another in the northeast. Interestingly, the latter lineage was shown to be extremely closely related and probably conspecific to *P. atrata*, suggesting that colonization of the Columbretes occurred recently. This contradicts previous suggestions of the endemic nature of *P. atrata* (Castilla *et al.* 1998), which had been based on a comparison of insular lizards to the specimens belonging to the southeastern lineage within the Iberian Peninsula.

A major outcome of these studies was the suggestion of *P. carbonelli* as a different species, not only because of the high level of differentiation observed both morphologically and genetically, but also because of the paraphyly of the former «*bocagei*» clade. *P. carbonelli* was later formally described as a distinct species (Sá-Sousa and Harris 2002), and is nowadays fully recognised as such (e.g. Ferrand de Almeida *et al.* 2001, Pleguezuelos *et al.* 2002, Loureiro *et al.* in press). The subspecies *berlengensis* Vicente 1985 is now assigned to *P. carbonelli*.

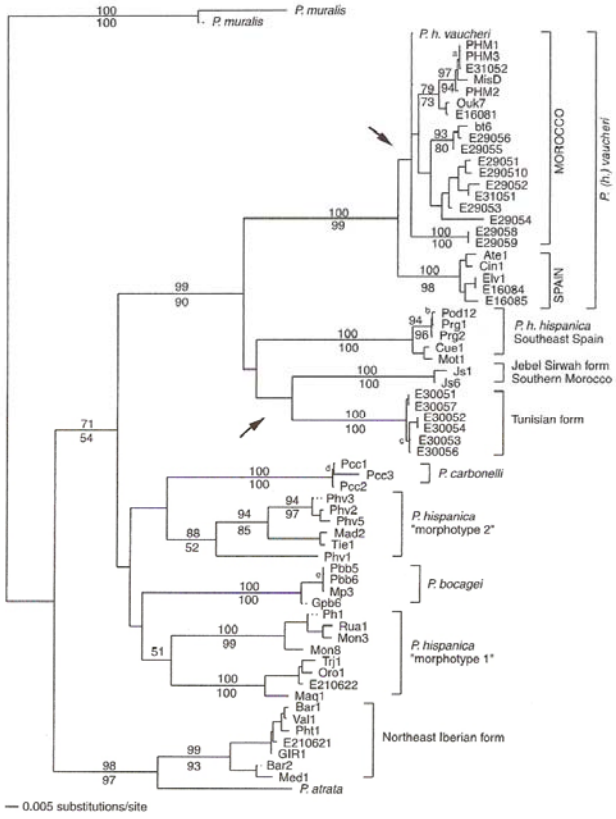

Figure 1.4. Minimum-evolution tree depicting relationships among species and forms of the Iberian and North African species complex based on mitochondrial DNA sequence data (cytochrome *b* and 12S rRNA genes) (reproduced from Harris *et al.* 2002b). Arrows indicate putative independent colonizations from the Iberian Peninsula to North Africa.

A more thorough study of the variation within southern Iberian Peninsula and North Africa (Harris *et al.* 2002b; Figure 1.4) revealed complex patterns of variation on both sides of the Strait of Gibraltar. This study described three distinct lineages inhabiting North Africa: one in Tunisia, one in Jebel Sirwah (an isolated mountain in southern Morocco) and another (corresponding to the *vaucheri* form of Boulenger (1905)) being present on both sides of the Strait. The southeastern form of *P. hispanica* appears as sister taxa to the group formed by *P. hispanica* from Tunisia and Jebel Sirwah. This scenario led the authors to put forward the hypothesis of two independent colonizations of North Africa, both occurring after the opening of the Strait of Gibraltar. More recently, Busack *et al.* (2005) followed the recommendation of Oliverio *et al.* (2000) and implicitly elevated *P. (h.) vaucheri* to the species level, describing variation within this species based on morphology, allozyme and mitochondrial DNA variation. These authors claim that *P. vaucheri* is itself a species complex, presenting two highly differentiated lineages that were probably separated by the opening of the Strait of Gibraltar. This view strongly contradicts the scenario suggested by Harris *et al.* (2002b) and reflects different calibrations of the molecular clock.

Besides helping to elucidate the number of taxonomic entities existing within the clade of Iberian and North African *Podarcis*, studies relying on mitochondrial DNA sequencing allowed the confirmation of previous reports (Sá-Sousa 1998, 1999, 2000a) and the collection of more precise data regarding the distribution of forms. The confirmation of the presence of *P. carbonelli* in Southern Spain (Harris *et al.* 2002a) is such an example. Figure 1.5 summarizes the knowledge available at the beginning of this work regarding the distribution of species and forms of Iberian and North African *Podarcis*.

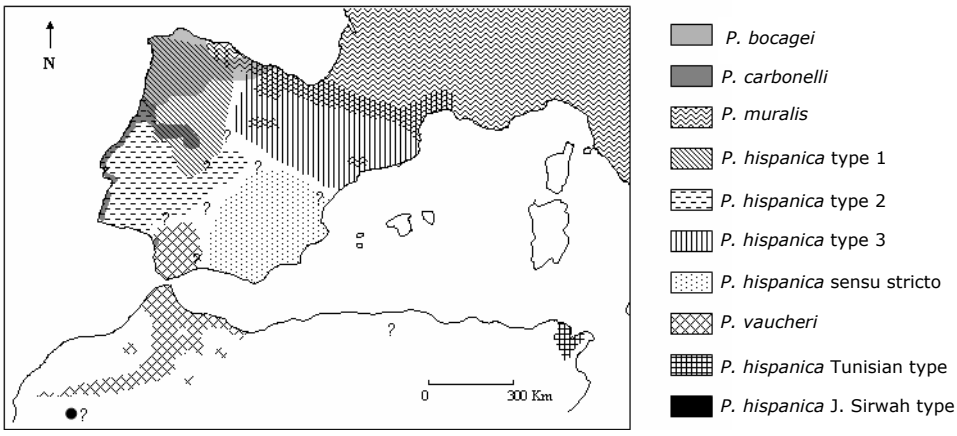

Figure 1.5. Distribution of *Podarcis* forms in the Iberian Peninsula and North Africa based on mtDNA sequencing and morphology data known at the beginning of this work (adapted from Pinho (2002)).

In terms of conservation, most species of Iberian and North African wall lizards are abundant and widely distributed and are therefore not considered to be threatened. There is one exception, *P. carbonelli*, which, due to its reduced and fragmented distribution area and latent risk of habitat destruction, is classified as "vulnerable" by the Portuguese vertebrate Red Data Book (Cabral *et al.* 2005).

1.3. Recent advances in evolutionary genetics

Evolutionary and population genetics began as almost strictly theoretical disciplines. Works by the modern synthesis founding fathers, such as Fisher (1930), Wright (1931) and Haldane (1932), used simple representations of ideal populations from which mathematical equations describing the forces that drive evolution, such as mutation, genetic drift, migration and natural selection, could be derived. By creating this conceptual framework and establishing the theoretical grounds for a better understanding of evolutionary processes, these studies opened the door for modern population biology; however, their work was entirely theory driven. It was not until the mid 1960's that the possibility of directly obtaining genetic data for organisms allowed evolutionary and population genetics to become fully-developed experimental sciences.

This genetic revolution was first driven by the development of allozyme electrophoresis techniques (Harris 1966, Lewontin and Hubby 1966). The generalization of these approaches allowed multilocus genotypic data to be readily gathered for a large number of organisms; these data could be analysed in the framework of traditional population genetics theory and were also soon noticed to have systematic value, allowing a clarification of population and species relationships, thus contributing to the emergence of the field of molecular phylogenetics. A second revolution took place in the mid 1970's, with the demonstration of the possibility of generating restriction site maps for animal mtDNA (Brown and Vinograd 1974), coupled with the discovery of a fast pace of molecular evolution for this molecule (Brown *et al.* 1979). Descriptions of mitochondrial DNA variability in nature (e.g. Brown and Wright 1975, Avise *et al.* 1979), and concomitant analytical innovations (e.g. Watterson 1975) soon became abundant. The possibility of describing genetic variation in nature allowed testing and application of "old-school" models of population genetics, and, by showing their limitations, opened avenues for new, less simplistic and more enlightening methods of analyses. In particular, practical advances such as the advent of the polymerase chain reaction (PCR) and the possibility of obtaining the sequence of genomic portions (and more recently, of complete genomes) caused a complete shift with traditional thinking in population

genetics, from the analyses of variation in the form of anonymous alleles to the possibility of directly observing gene genealogies and the historical information contained therein. Nowadays, evolutionary biology is evolving at an extraordinarily rapid pace. New markers, new methods, new concepts and whole new disciplines are emerging. At the same time, evolutionary genetics presently shows a welcome tendency for a new synthesis – integrating different sources of information, combining analytical approaches and bridging disciplines and fields of research that have traditionally been kept completely separated – to more accurately describe the genetic processes that drive the evolution of organisms.

1.3.1. Major disciplines of evolutionary biology and their scope

Within the field of evolutionary biology, genetic tools are now widely used to approach a multitude of questions. These range from macroevolutionary processes such as determining species relationships to phenomena at the population or even the individual level (microevolution). Traditionally, these fields of investigation have been kept apart and relied on entirely different analytical approaches and assumptions.

An active field of research, *molecular phylogenetics*, is devoted to the determination of species relationships, and its ultimate goal is to produce trees that accurately describe the relationships between organisms. In general, however, phylogenetic studies describe the phylogeny of a gene or a set of genes in the hope that it is representative of the evolutionary relationships of the organisms under study. Under many circumstances (see next section), this may not be the case. There are highly elaborated analytical methods to produce gene trees, varying on the optimality criteria used and on the nature of the genetic data. Recent analytical methods are model-based and mathematically complex (Felsenstein 1981, Huelsenbeck *et al.* 2001). Methods aiming at determining the underlying species tree from a set of discordant gene genealogies are, on the other hand, much less developed (Maddison 1997, Page and Charleston 1997, Maddison and Knowles 2006).

On the other hand, *population genetics* is the study of genetic variation at a microevolutionary level and the characterization of the primary forces (mutation, migration, genetic drift, selection) that produce genetic changes in populations over time. Unlike phylogenetics, which is entirely tree-based, population genetics relies mainly on obtaining interpretable summary statistics from the data and depends on mathematical models that depict population dynamics and that are unavoidable simplifications of reality. Natural populations are strikingly dynamic over space and time; when compared to the complexity of natural processes, most models used in population genetics are therefore far too

simplistic. In response to the growing notion that assumptions in models are frequently violated in natural populations, in recent years several more realistic models have been developed and adopted. Another important trend in population genetics is the growing reliance on genealogical data. Early-developed summary statistics were usually based on allele frequency data; the possibility of obtaining large sequence data sets and the inherent historical information contained in the patterns of DNA sequence variation (but also in other markers for which evolution can be modelled and therefore superimposed on a genealogy) have opened up new areas of mathematical model development (Hey and Machado 2003), which make use of coalescent theory (Kingman 1982). According to this approach, the inferential analysis is retrospective: we seek to understand aspects of the samples's evolutionary past through analysis of the present-day sample (Hein *et al.* 2005). The coalescent approach can be simply used to generate simulated data which, when compared to the actual data, can provide valuable insights (Nordborg 2001). Other methods of analyses are more sophisticated and may incorporate complex parameter estimations under various models. In these approaches, the gene tree is usually treated as a nuisance parameter which is estimated alongside other parameters of interest.

The phylogenetic and population genetics approaches were traditionally kept apart, but are not irreconcilable. *Phylogeography* is a remarkable example of an interface between these two perspectives. In its essence, phylogeography is devoted to the study of processes governing the geographical distribution of genealogical lineages (Avise *et al*. 1987, Avise 2000). Nearly every species whose genetic variation has been characterized appears to have some degree of phylogeographic structure, which arises as a result of the interaction between demographic and genealogical processes and landscape level dynamics; therefore, by applying phylogenetic methods to intraspecific data, phylogeography forms the bridge between the study of macro and microevolutionary forces. One of the main concerns stemming from the phylogeographic approach is that it relies entirely on the estimation of a gene tree and that it often lacks the consistent model-based statistical framework that characterizes population genetics (Knowles 2004). Moreover, although phylogeography is concerned with the spatial distribution of genetic variation, only a handful of methods make actual use of geographically explicit frameworks (Templeton 1998, Dupanloup *et al.* 2002, Currat *et al.* 2004, Guillot *et al.* 2005, Miller 2005). Both concerns are being dealt with through the effort development of new tools for data analyses, such as the use of complex population genetic models incorporating subdivision under a coalescent framework (e.g. Nielsen and Wakeley 2001); the spatially explicit coalescent is also a promising avenue for future investigations (Baird, personal communication; Weiss and Ferrand 2007).

1.3.2. The study of different genomic compartments: contrasting perspectives on evolutionary processes

Different regions of the genome have particular modes of evolution and inheritance and therefore different properties for evolutionary inference. Although the first descriptions of genetic variation were carried out using allozymes (i.e. nuclear genes), mitochondrial DNA soon became the primary source of inference in evolutionary biology.

There are several properties of mitochondrial DNA that make it an attractive source of information in an evolutionary context. One is its relatively high rate of molecular evolution when compared to nuclear genes. Several hypotheses have been advanced to explain the rapid evolution of animal mtDNA, such as the relaxation of functional constraints, a high mutation rate derived from the inefficiency of repair mechanisms, high exposure to mutagenic free radicals or fast replicative turnover within cell lineages and the absence of associated histone complexes (Avise 2000). Other important particularities of mtDNA reside on its haploid nature and peculiar mode of inheritance. In vertebrates, mitochondrial DNA has almost exclusive maternal transmission (there are well-documented cases of paternal leakage that nevertheless seem to represent exceptions (e.g. Gyllensten *et al.* 1991, Kvist *et al.* 2003)). Because of this, recombination within the mitochondrial genome, if existent (see e.g. Tsaousis *et al.* 2005), is rather limited, implying that mtDNA evolution is not reticulated and lineages can be viewed as hierarchical branches in a genealogy, which constitutes a major advantage for phylogenetic inference. Additionally, the absence of recombination and uniparental inheritance reduce the effective population size of mtDNA to (on average) a quarter of that of a nuclear gene. Based on this fact, Moore (1995) provided theoretical evidence for preferring mtDNA over nuclear gene trees to infer species relationships due to the fact that it has a higher probability of tracking splitting events. Additional reasons for the popularity of mtDNA in evolutionary studies are technical ones: on one hand, there are several mitochondria in each cell, which allowed the obtention of high quantities of DNA for analyses, even before the advent of PCR; on the other hand, because the organization of the mitochondrial genome is relatively conserved across taxa, it is easy to obtain large sequence data sets using conserved primers (e.g. Kocher *et al.* 1989), even for non-model organisms.

However, there are several drawbacks associated to the single use of mitochondrial DNA (reviewed in Zhang and Hewitt 2003 and Ballard and Whitlock 2004). Because mtDNA does not recombine, it behaves as a single locus. Thus, studies that rely entirely on mitochondrial DNA variation will suffer from the insufficiency of observing only one among several possible independent replicate witnesses (i.e. genealogies) of the evolutionary process.

The rationale behind trying to obtain data from as many independent loci as one can is that the genealogy of a single gene may not be consistent with the evolutionary history of the organisms being studied (the classical gene-tree vs. species-tree problem, Pamilo and Nei 1988); this happens because different genes might be under the influence of different selective pressures or respond differently to introgression (Machado and Hey 2003), or simply because the sorting of lineages in a genealogy may occur in a random fashion, without necessarily corresponding to population splitting events. It has been demonstrated that the use of more than one gene for the estimation of population parameters produces significantly more reliable estimates (Edwards and Beerli 2000).

The disadvantages of analysing one gene are therefore valid not only for studies completely based on mtDNA but also for any study considering a single nuclear gene genealogy. However, because of the specific particularities of mtDNA, there are yet other well-documented forces that could provide non-reliable estimates of relationships, differentiation, genetic diversity, demography and historical processes because mtDNA only tells one side of the story (that is, the history of females rather than that of the whole population). These forces are, for example, selection, to which the mtDNA molecule seems particularly prone (Ballard *et al.* 2002, Bazin *et al.* 2006), fluctuating effective population sizes (Fay and Wu 1999, Monsen and Blouin, 2003), gender-biased gene flow (e.g. FitzSimmons *et al*. 1997, Nyakaana and Arctander 1999, Piertney *et al*. 2000) and introgression (e.g. DeSalle and Giddings 1986, Shaw 2002, Alves *et al*. 2003, Chan and Levin 2005).

By contrast, nuclear autosomal genes reveal the history of both males and females and do not suffer from many of the above mentioned biases. In particular, it is impossible from the perspective of mitochondrial variation to address issues like introgressive hybridization, which need the assessment of diploid markers. Another important particularity of the nuclear genome is that it recombines. Because of this, even genes that are located on the same chromosome may have completely independent evolutionary histories. Therefore, the nuclear genome is a virtually unlimited source of independent genetic markers. A wide array of nuclear markers, with varying evolutionary rates, can be used to address different evolutionary questions.

The study of other genomic compartments, namely the sexual chromosomes, such as the X or Y chromosomes in mammals or *Drosophila*, or the Z and W chromosomes in birds and butterflies, offer other contrasting properties in terms of inheritance modes, recombination and effective population size that could provide valuable insights for the comprehension of evolutionary processes, in particular those that are sex-specific. Moreover, it is now being acknowledged that these chromosomes harbour genes of great importance for sexual

reproduction and, in the particular case of the X chromosome, for speciation (Coyne and Orr 2004). However, while this important role is clear for mammalian and *Drosophila* species and studies of their molecular evolution are beginning to accumulate, for many non-mammalian organisms the amount of knowledge regarding the existence of sexual chromosomes, their organization and sexual determination itself is insufficient and usually prevents studies of these genetic markers.

1.3.3. Speciation

Although there is an active debate over species definitions and concepts (e.g. Hey 2001, 2006), studies on speciation are generally centred on the acquisition of reproductive isolation. This follows the Biological Species Concept (Dobzhansky 1935, Mayr 1942), according to which species are "groups of actually or potentially interbreeding natural populations that are reproductively isolated from other such groups" (Mayr 1942).

1.3.3.1. Biogeography of speciation

One of the most active controversies about the process of speciation is whether it requires complete geographic isolation (allopatric speciation) or if it is able to occur in populations that have overlapping distributions and that therefore are able to exchange genes throughout the process (in a restricted fashion, if their distribution only partially overlaps – parapatric speciation – or freely, if their distributions overlap to a great extent – sympatric speciation). Although there seems to be a trend towards considering allopatric speciation as the rule and other modes of speciation as exceptions, this intuition is highly biased by the easiness of demonstrating speciation in allopatry when compared to that in sympatry or parapatry (Coyne and Orr 2004). Sympatric speciation became increasingly accepted only when models were developed that substantiated its plausibility (e.g. Kondrashov 1983); the acceptance of theories invoking allopatric speciation was never dependent upon such demonstrations. One of the main reasons for a general disbelief in non-allopatric modes of speciation is that they require a major role for selective forces in divergence, outweighting the merging influence of gene flow, whereas in allopatric circumstances virtually any forces causing divergence can eventually yield speciation. Nevertheless, there is growing evidence for the plausibility of non-allopatric speciation in nature (Via 2001).

1.3.3.2. The genetic basis of speciation

In a way, speciation can be regarded as a by-product of normal population genetic processes such as selection or drift; a distinctive feature of speciation genetics is, however, epistasis (Coyne and Orr 1998). The incompatible interaction between two or more loci has been invoked as a major mechanism leading to reproductive isolation (by causing hybrid inviability or sterility or by promoting pre-zygotic incompatibilities) long before this was actually tested on real data (Dobzhansky 1937, Muller 1940, 1942). Major questions of interest are, therefore, which and how many loci are involved in such incompatibilities and where, in a genomic context, they are located. Identifying genes that are directly responsible for causing speciation has therefore become one of the most ambitious endeavours of modern evolutionary genetics. By definition, a speciation gene is one that can be shown to cause some degree of ecological, sexual or post-mating isolation between young species (Wu and Ting 2004). Although several studies have identified genomic locations that putatively comprise such genes (e.g. Harr 2006), only a handful have been clearly identified and characterized so far, all but one in *Drosophila* (Wittbrodt *et al.* 1989, Ting *et al.* 1998, Presgraves *et al.* 2003, Barbash *et al.* 2003; see also Noor and Feder 2006 for other candidate speciation genes). Common features across most of these genes are that they are generally located on the X-chromosome and seem to have evolved through positive selection (see also Greenberg *et al.* 2006).

Besides epistatic interactions, other genetic forces have been described to cause speciation. In plants, polyploidy causing "instantaneous" speciation is a well known phenomenon, whether it involves hybridization between distinct species (allopolyploydy) or not (autopolyploydy). Chromosomal rearrangements have also been suggested to promote speciation. Structural changes in chromosomes might directly cause reproductive isolation (chromosomal speciation), but evidence for this is controversial. Alternatively, rearrangements could help to tie up sets of genes that are responsible for reproductive isolation (Noor *et al.* 2001, Rieseberg 2001, Navarro and Barton 2003) and therefore slow the progression of gene flow. In particular taxa, a well-documented pattern causing speciation is cytoplasmic incompatibility. In *Drosophila* and other insects, it has been shown that males infected with cytoplasmically inherited endosymbionts (such as *Wolbachia*) do not produce viable offspring when crossed with uninfected females, which may cause reproductive isolation between taxa in nature.

1.3.3.3. The study of speciation

The stage that emergent species have achieved in the process of differentiation is directly related to the degree of reproductive isolation between them. It is therefore critical to assess levels of gene flow and investigate putative barriers impeding introgression. Traditionally, studies on speciation have followed two major approaches: laboratory controlled interspecific crosses and the study of hybrid zones. Both approaches are complementary because they provide different perspectives on the nature of isolating mechanisms: while laboratory crosses allow a direct quantification of hybrid viability and fertility, are conducted under controlled conditions and are therefore repeatable, the study of hybrid zones allows the assessment of multiple (instead of a few) generations of recombination and of the fitness of hybrids under natural conditions. An extensive theory regarding the behaviour of hybrid zones has been developed (e.g. Barton and Hewitt 1985) which deals primarily with the quantification of the relative effects of dispersal and selection against hybrids and allows the assessment of the strength of barriers acting against gene flow.

1.4. Objectives and thematic organization of the thesis

This thesis is organized in five chapters. The first is a general introduction containing basic information on the biogeographic complexity of the Western Mediterranean, the organisms used as models (*Podarcis* wall lizards) and their controversial taxonomy, and also current trends and methods in evolutionary biology, providing a conceptual framework for the work presented in this thesis. The next two chapters describe the two major research directions pursued: chapter two concerns the study of genetic divergence and speciation in wall lizards and addresses the definition of evolutionary units, concordance amongst markers, description of species relationships and assessment of the level of reproductive isolation; chapter three is devoted to the description of patterns of intraspecific variation and microevolutionary patterns in selected taxa within the *Podarcis* species complex. Chapter four consists of a general discussion that summarizes and contextualizes the major findings drawing from the work presented in the former chapters. Finally, in chapter five we refer to the major conclusions that can be drawn from this work and also to the work that has been left undone or that we would like to follow in the near future.

This thesis consists of eight articles, included in chapters two and three. The first article is entitled "**Reexamination of the Iberian and North African *Podarcis* (Squamata: Lacertidae) phylogeny based on increased**

mitochondrial DNA sequencing" and has been published in *Molecular Phylogenetics and Evolution*. Our main objective with this paper was to accurately describe the mitochondrial DNA phylogeny for the Iberian and North African species complex. In previous works employing mtDNA sequencing, major lineages within this complex were readily revealed (Harris and Sá-Sousa 2001, 2002; Harris *et al.* 2002); however, evolutionary relationships between those lineages remained a mystery. This was a serious drawback when attempting to infer biogeographic patterns responsible for differentiation within this group. Therefore we extended our sequencing effort to a total of 2425bp of mitochondrial DNA, including portions of five different regions. A well-resolved and robust phylogenetic tree, evidencing between-species relationships with high resolution, was the major outcome of this study. We also described a previously unidentified and highly divergent lineage that additionally demonstrates the evolutionary complexity of this species group. These results were important because they establish a well defined scenario for the mitochondrial DNA evolution in this group that can be compared with scenarios suggested by other markers. Moreover, by applying a molecular clock to the estimates of species divergence, this study raised new hypotheses on the role played by the opening of the Strait of Gibraltar and other geological events in shaping the biogeographic history of Iberian and Maghrebian *Podarcis*.

Acknowledging that a single locus may not truly represent the evolutionary history of a group of taxa, we turned to allozymes to obtain an independent assessment of population subdivision based on different genomic compartments. In our second paper ("**Genetic polymorphism of 11 allozyme loci in populations of wall lizards (*Podarcis* sp.) from the Iberian Peninsula and North Africa**"), which has been published in *Biochemical Genetics*, we present the results from the development and characterization of genetic variation at protein loci in these lizards. Using conventional starch gel electrophoresis and isoelectric focusing, we accomplished a preliminary evaluation of differentiation among forms of the species complex, showing that species and other lineages defined on the basis of mitochondrial DNA are also distinguishable using nuclear markers, reinforcing the idea that a taxonomic revision is clearly needed.

We extended this study by analysing the same genetic loci in individuals from 32 populations more representative of the total distribution of the species complex. The outcome of these analyses is presented in article III ("**Comparing patterns of nuclear and mitochondrial divergence in a cryptic species complex: the case of Iberian and North African wall lizards (*Podarcis*, Lacertidae)**"), which is currently in press in the *Biological Journal of the Linnean Society.* In this article we apply both conventional population genetic tools, such as F_{ST} estimation, and recently developed model-based clustering algorithms that operate on individual multilocus genotypes. By doing this we

intended to investigate whether mitochondrial DNA lineages could be identified not only as clusters of populations, as suggested by the previous article, but also if they represented clearly observable clusters of individuals. In doing so, we were also aiming to detect putative cases of hybridization or introgression. The results obtained show high levels of concordance with mtDNA with respect to the delimitation of evolutionary units, both when using a population tree to investigate major partitions or when using individual clustering methods. We demonstrate, for example, that pairs of forms within the *P. hispanica* complex and pairs of fully recognized species show similar F_{ST} values. However, not all mtDNA lineages correspond to clearly distinct clades; we discuss several hypotheses relative to the nature of these discrepancies, emphasizing the possibility of introgression. Another striking difference obtained from the analysis of allozyme loci is that, with a few exceptions, relationships between forms are poorly supported. This contrasts with the clearly bifurcating tree inferred from the analyses of mitochondrial DNA variation and probably reflects the different effective population sizes that characterize mitochondrial and nuclear markers. This study also allowed us to identify two putative hybrids between *P. bocagei* and *P. hispanica* type 1A, which constitutes the first clear evidence of the occurrence of gene flow between forms of the species complex. This article therefore illustrates that, despite nuclear and mitochondrial markers describe roughly concordant scenarios, the evolutionary dynamics of Iberian and North African *Podarcis* might not be as simple as depicted by the mitochondrial DNA phylogeny alone and that gene flow between forms is an evolutionary force that should not be disregarded.

Although analyses based on allozyme variation were useful in assessing genetic variation at nuclear markers and understanding the general agreement from two independent sources of information, they lack the genealogical framework that allows a direct comparison to the scenario portrayed by the mtDNA phylogeny. In this context, we turned to the study of nuclear gene genealogies to test some of the hypotheses that were raised by the previous paper. This is described in the fourth article, "**Abundant ancestral polymorphism at the nuclear loci *β-fibint7* and *6-Pgdint7* characterizes different *Podarcis* species from the Iberian Peninsula and North Africa**", which is still in preparation. We studied genetic variation at two unlinked intronic loci. The most striking result of this article is that in general, species are not monophyletic with respect to the nuclear introns' genealogies. Surprisingly, not even the oldest splitting events inferred for the history of the genus by the mtDNA analyses were recovered. We then asked the question: is this pattern completely explained by incomplete lineage sorting of ancestral polymorphism or did gene flow between species contribute to the interspecific sharing of closely related alleles? Our analyses revealed that incomplete lineage sorting is indeed

the main cause for the lack of species monophyly, although some pairs of species have clearly engaged in gene exchange at some point during their divergence process.

With this set of articles we described the major evolutionary units within Iberian and North African *Podarcis* and how genetic variation was partitioned among them, moreover demonstrating that species and forms of *Podarcis* are still not completely reproductively isolated. However, these studies were carried out using the whole group of species as the unit of study; by doing so, the dynamics of gene flow between them were only superficially addressed. In order to study in more detail if and how species boundaries are maintained, we studied a contact zone between two species. The results of this research are presented in the form of article V, "**Genetic admixture between the Iberian endemic lizards *Podarcis bocagei* and *P. carbonelli*: evidence for limited natural hybridization and a bimodal hybrid zone**", which is currently submitted. In this article we studied a battery of allozyme, microsatellite and intron loci that perform well in discriminating between the two species, and also mitochondrial DNA. By using multilocus clustering methods, we evaluated the degree of admixture of individuals from the contact zone. Although it is clear that the two species hybridise, the majority of sampled individuals show either no signs of admixture or close to "pure" multilocus genotypes, yielding large levels of Hardy-Weinberg and linkage disequilibria. This bimodality clearly suggests strong barriers to gene flow. We also characterized morphologically individuals from the contact zone and compared their assignment based on both methods (i.e. morphology and genetics). These analyses reveal a striking absence of morphological intermediates, even among individuals that were shown to be admixed with respect to their genetic origin, suggesting that probably there is a selective mechanism, intrinsic or extrinsic, acting against intermediate phenotypes. Although the conclusions of this study cannot be readily generalized to other contact zones, it provides hints on the processes that are helping in keeping these species apart.

The articles included in the third chapter deal with the assessment of intraspecific variation using multiple molecular markers. In particular, we were hoping to understand the effects that Pleistocene glaciations had in shaping the patterns of geographical subdivision and demography. With this approach, we intended not only to compare observed patterns with phylogeographic scenarios described for other species occupying the same regions, but also to investigate how different were the genetic signatures left by the Ice Ages on distinct *Podarcis* species. We began by assessing intraspecific genetic variation in mitochondrial genes in three species of wall lizards (*Podarcis bocagei, P. carbonelli* and *P. vaucheri*) that replace each other along a latitudinal gradient. The outcome of this study is presented as article VI ("**Contrasting patterns of**

population subdivision and historical demography in three Western Mediterranean lizard species inferred from mitochondrial DNA variation") and is now in press in *Molecular Ecology*. In this study we test hypotheses relative to the differential influence of the Ice Ages on different latitudes. We conducted several analyses regarding patterns of nucleotide diversity, geographic subdivision, coalescence times and demographic trends in all three species. *P. bocagei* shows low levels of diversity, no geographic substructure and a relatively shallow coalescent history resulting from a recent bottleneck, suggesting that only one of the putatively multiple refugia to which this species was confined was able to persist until the present. Accordingly, we detected a clear signature of population growth concomitant with a post-glacial colonization of recently available habitat. On the other extreme of our transect, Moroccan *P. vaucheri* shows complex geographic substructure resulting from isolation in distinct glacial refugia. Some of these population groups persist since the beginning of the Ice Ages (around 1.8 million years ago), but subsequent glacial cycles also left their imprint, suggesting much lower extinction levels across multiple glacial cycles than in *P. bocagei*. Populations of this species show no sign of a demographic expansion. *P. carbonelli* presents intermediate levels of diversity and subdivision; the coalescent history of this species is not as shallow as in *P. bocagei* but denotes that persistence levels were lower than in *P. vaucheri*. Only one of its recent phylogroups evidences signs of population growth. Taking these results together, this study demonstrates that different effects of the Pleistocene glaciations can be felt even within a small study area such as ours.

Once again, we also tried to corroborate the inferred patterns by the study of nuclear markers. A first step was to develop loci that are more appropriate for studying recent divergence - microsatellites. Although microsatellite loci had been previously developed for a species of the same genus (Boudjemadi *et al.* 1999), there were still no microsatellites developed specifically for *Podarcis* from the Iberian-Maghrebian clade, which constitute the models for this work. In article VII, "**Isolation and characterization of nine microsatellite loci in *Podarcis bocagei* (Squamata: Lacertidae)**", published in *Molecular Ecology Notes*, we report the development of nine dinucleotide microsatellite loci for *P. bocagei* through an enrichment protocol. The preliminary study of these loci in a single population of the species revealed remarkably high levels of polymorphism and suggested that these markers would be useful in uncovering population structure within this and probably other species.

Finally, in the last article ("**Combining mtDNA, allozyme polymorphisms and microsatellites to assess the evolutionary histories of two endemic lacertids (*Podarcis bocagei* and *P. carbonelli*) from Western Iberian Peninsula**"), which is still in preparation, we screen genetic variation at

allozyme and microsatellite loci in two species for which phylogeographic scenarios had already been inferred from the analyses of mitochondrial DNA. In *P. bocagei*, we describe a clear decrease in genetic variation from southern to northern populations, concordant with the previous description of a post-glacial expansion. By applying model-based clustering approaches enforcing successive levels of subdivision we were able to identify two putative northward expansion routes and pinpoint a probable refugial area. In *P. carbonelli*, we found higher, but still low, levels of subdivision and no obvious trend in the patterns of variability across geography; the inferred clusters only roughly correspond to the subdivision suggested by the mitochondrial DNA analyses. In this species, our data strongly suggest that genetic drift related to post-glacial changes in the species distribution has probably masked the signatures that are traditionally detected as response to the Pleistocene climatic oscillations.

1.5. References

Alexandrino J, Arntzen JW, Ferrand N (2002) Nested clade analysis and the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Heredity*, **88**, 66–74.

Alexandrino J, Froufe E, Arntzen JW, Ferrand N (2000) Genetic subdivision, glacial refugia and postglacial recolonization in the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Molecular Ecology*, **9**, 771–781.

Almeida AP (1993) *Diferenciação Genética de* Podarcis bocagei *do Arquipélago da Berlenga*, Relatório de Estágio para a obtenção da Licenciatura em Bioquímica, Faculdade de Ciências de Lisboa.

Álvarez Y, Mateo JA, Andreu AC, Díaz-Paniagua C, Diez A, Bautista JM (2000) Mitochondrial DNA haplotyping of *Testudo graeca* on both continental sides of the Straits of Gibraltar. *Journal of Heredity,* **91,** 39–49.

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.

Arnold EN (1973) Relationships of the Palaearctic lizards assigned to the genera *Lacerta, Algyroides* and *Psammodromus* (Reptilia, Lacertidae). *Bulletin of the British Museum of Natural History (Zoology)*, **25,** 291-366

Arnold EN (1990) Why do morphological phylogenies vary in quality? An investigation based on the comparative history of lizard clades. *Proceedings of the Royal Society of London B: Biological Sciences*, **240,** 135–172.

Arnold EN, Burton JA (1978) *A field guide to the reptiles and amphibians of Britain and Europe*. Collins, London.

Arntzen JW, García-París M (1995) Morphological and allozyme studies of midwife toads (Genus *Alytes*), including the description of two new taxa from Spain. *Contributions to Zoology*, **65,** 5–34.

Avise JC, Giblin-Davidson C, Laerm J, Patton JC, Lansman RA (1979) Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. *Proceedings of the National Academy of Sciences USA,* **76,** 6694–6698.

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18,** 489–522.

Avise JC (2000) *Phylogeography*. *The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

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,** 527–545.

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

Barbadillo LJ, Lacomba JI, Pérez Mellado V, Sancho V, López-Jurado LF (1999) *La Guia de Campo de los Anfíbios y reptiles de la Península Ibérica, Baleares y Canárias*. Ed. Planeta, Barcelona

Barbash DA, Slino DF, Tarone AM, Roote J (2003) A rapidly evolving MYB-related protein causes species isolation in *Drosophila*. *Proceedings of the National Academy of Sciences USA,* **100,** 5302–5307.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16,** 113–148.

Bazin E, Glémin S, Galtier N (2006) Population size does not influence mitochondrial genetic diversity in animals. *Science,* **312,** 570–572.

Bea A, Gosa A, Guillaume CP, Geniez P (1986) *Podarcis hispanica sebastiani* (Klemmer, 1964), *nomen novum pro Podarcis hispanica hispanica* (Steindachner, 1870) del Monte Urgull e Isla de Santa Clara (San Sebastian, España), *Revista Española de Herpetología*, **1,** 188–205.

Bischoff W (1973) Lacertenbastarde II. *Zoologischer Garten NF, Berlin,* **43,** 278-291.

Böhme W (1971) Über das Stachelepithel am Hemipenis lacertider Eidechsen und seine systematische Bedeutung. *Zeitschrift fur Zoologische Systematik und Evolutionsforschung*, **9,** 187–223.

Böhme W (ed.) (1986) *Handbuch der reptilien und amphibian Europas. Band 2/II. Echsen (Sauria) III (Lacertidae III: Podarcis)*. Aula, Wiesbaden.

Boudjemadi K, Martin O, Simon JC, Estoup A (1999) Development and cross-species comparison of microsatellite markers in two lizard species, *Lacerta vivipara* and *Podarcis muralis*. *Molecular Ecology,* **8,** 518-520.

Boulenger GA (1916) On the lizards allied to *Lacerta agilis* and *Lacerta parva*. *Transactions of the Zoological Society of London*, **21,** 1–90.

Branco M, Ferrand N, Monnerot M (2000) Phylogeography of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome *b* gene. *Heredity*, **85**, 307–317.

Branco M, Monnerot M, Ferrand N, Templeton AR (2002) Postglacial dispersal of the European rabbit (*Oryctolagus cuniculus*) on the Iberian Peninsula reconstructed from nested clade and mismatch analyses of mitochondrial DNA genetic variation. *Evolution*, **56**, 792–803.

Brito PH (2005) The influence of Pleistocene glacial refugia on tawny owl genetic diversity and phylogeography in western Europe. *Molecular Ecology,* **14,** 3077–3094.

Brown RP, Suárez NM, Pestano J (2002) The Atlas mountains as a biogeographical divide in Northwestern Africa: evidence from mtDNA evolution in the Agamid lizard *Agama impalearis*. *Molecular Phylogenetics and Evolution*, **24,** 324–332.

Brown WM, Vinograd J (1974) Restriction endonuclease cleavage maps of animal mitochondrial DNAs. *Proceedings of the National Academy of Sciences, USA*, **71,** 4617–4621.

Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences, USA*, **76,** 1967–1971.

Brown WM, Wright JW (1975) Mitochondrial DNA and the origin of parthenogenesis in whiptail lizards (*Cnemidophorus*). *Herpetological Review,* **6,** 70–71.

Busack SD (1986) Biogeographic analysis of the herpetofauna separated by the formation of the Strait of Gibraltar. *National Geographic Research,* **2,** 17–36.

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.

Cabral MJ (coord.), Almeida J, Almada PR, Dellinger T, Ferrand de Almeida N, Oliveira ME, Palmeirim JM, Queiroz AI, Rogado L, Santos-Reis M (eds.) (2005) *Livro Vermelho dos Vertebrados de Portugal*. Instituto de Conservação da Natureza, Lisboa.

Capula M (1993) Natural hybridization in *Podarcis sicula* and *P. wagleriana* (Reptilia: Lacertidae). *Biochemical Systematics and Ecology,* **3,** 373–380.

Capula M (1994a) Genetic variation and differentiation in the lizards, *Podarcis wagleriana* (Reptilia: Lacertidae). *Biological Journal of the Linnean Society* **52,** 177–196.

Capula M (1994b) Evolutionary relationships of *Podarcis* lizards from Sicily and the Maltese Islands. *Journal of Zoological Systematics and Evolutionary Research,* **32,** 180–192.

Capula M (1997) Genetic divergence and phylogenetic inferences in the genus *Podarcis.* (Reptilia: Lacertidae) In: Roček, Z. & Hart, S. (eds.) *Herpetology 97. Abstrs. III World Congress of Herpetology*, Prague, Czech Republic, 35.

Capula M (2002) Genetic evidence of natural hybridization between *Podarcis sicula* and *Podarcis tiliguerta* (Reptilia: Lacertidae). *Amphibia-Reptilia,* **23,** 313–321.

Carmona JA, Domínguez J, Doadrio I (2000) Congruence between allozyme and cytochrome *b* gene sequence data in assessing genetic differentiation within the Iberian endemic *Chondrostoma lemmingii* (Pisces: Cyprinidae). *Heredity*, **84,** 721–732.

Carranza S, Amat F (2005) Taxonomy, biogeography and evolution of *Euproctus* (Amphibia: Salamandridae), with the resurrection of the genus *Calotriton* and the description of a new endemic species from the Iberian Peninsula. *Zoological Journal of the Linnean Society*, **145,** 555–582.

Carranza S, Arnold EN (2004) History of West Mediterranean newts, *Pleurodeles* (Amphibia: Salamandridae) inferred from old and recent DNA sequences. *Systematics and Biodiversity* **1,** 327–337.

Carranza S, Arnold EN, Amat F (2004a) DNA phylogehy of *Lacerta* (*Iberolacerta)* and other lacertine lizards (Reptilia: Lacertidae): did competition cause long-term mountain restriction? *Systematics and Biodiversity,* **2,** 57–77.

Carranza S, Arnold EN, Pleguezuelos JM (2006a) Phylogeny, biogeography and evolution of two Mediterranean snakes, *Malpolon monspessulanus* and *Hemorrhois hippocrepis* (Squamata, Colubridae), using mtDNA sequences. *Molecular Phylogenetics and Evolution,* **40,** 532 – 546.

Carranza S, Arnold EN, Wade E, Fahd S (2004b) Phylogeography of the false smooth snakes, *Macroprotodon* (Serpentes, Colubridae): mitochondrial DNA sequences show European populations arrived recently from Northwest Africa. *Molecular Phylogenetics and Evolution,* **33,** 523–532.

Carranza S, Harris DJ, Arnold EN, Batista V, Gonzalez de la Vega JP (2006b) Phylogeography of the lacertid lizard *Psammodromus algirus*, in Iberia and across the Strait of Gibraltar. *Journal of Biogeography,* **33,** 1279–1288.

Castella V, Ruedi M, Excoffier L, *et al*. (2000) Is the Gibraltar Strait a barrier to gene flow for the bat *Myotis myotis* (Chiroptera: Vespertilionidae)? *Molecular Ecology* **9**, 1761–1772.

Castilla AM, Fernandez-Pedrosa V, Harris DJ, Gonzalez A, Latorre A, Moya A (1998) Mitochondrial DNA divergence suggests *Podarcis hispanica atrata* from the Columbretes Islands merits specific distinction. *Copeia* **4,** 1037–1040

Castilla AM (2002) *Podarcis atrata.* In: *Atlas y libro rojo de los anfíbios y reptiles de España* (Pleguezuelos JM, Márquez R, Lizana M, eds.). Dirección General de Conservación de la Naturaleza – Asociación Herpetologica Española, 2nd Edition, Madrid, 238–239.

Chan KMA, Levin SA (2005) Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution* **59,** 720–729.

Comes HP, Abbott RJ (1998) The relative importance of historical events and gene flow on the population structure of a Mediterranean ragwort, *Senecio gallicus* (Asteraceae). *Evolution*, **52**, 355–367.

Cosson JF, Hutterer R, Libois R, *et al.* (2005) Phylogeographical footprints of the Strait of Gibraltar and Quaternary climatic fluctuations in the Western Mediterranean: a case study with the greater white-toothed shrew, *Crocidura russula* (Mammalia: Soricidae). *Molecular Ecology,* **14,** 1151–1162.

Coyne JA, Orr HA (1998) The evolutionary genetics of speciation. *Philosophical Transactions of the Royal Society of London, B: Biological Sciences,* **353,** 287–305.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, MA, USA.

Crespo EG, Oliveira ME (1989) *Atlas da distribuição dos Anfíbios e Répteis de Portugal Continental*. Ed. Serviço Nacional de Parques, Reservas e Conservação da Natureza, Lisboa.

Currat M, Ray N, Excoffier L (2004) SPLATCHE: A program to simulate genetic diversity taking into account environmental heterogeneity. *Molecular Ecology Notes,* **4,** 139–142.

DeSalle R, Giddings LV (1986) Discordance of nuclear and mitochondrial phylogenies in Hawaiian *Drosophila*. *Proceedings of the National Academy of Sciences, USA,* **83,** 6902–6906.

Dobzhansky T (1935) A critique of the species concept in biology. *Philosophy of Science*, **2,** 344–355.

Dobzhansky T (1937) *Genetics and the Origin of Species.* Columbia Univ. Press, New York, USA.

Duggen S, Hoernie K, van den Bogaard P, Rupke L, Morgan JP (2003) Deep roots of the Messinian salinity crisis. *Nature*, **422**, 602–606.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**, 2571–2581.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution,* **54,** 1839–1854.

Escoriza D, Comas MM, Donaire D, Carranza S (2006) Rediscovery of *Salamandra algira* Bedriaga, 1833 from the Beni Snassen massif (Morocco) and phylogenetic relationships of North African *Salamandra*. *Amphibia-Reptilia*, **27,** 448–455.

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

Ferrand de Almeida N, Ferrand de Almeida P, Gonçalves H, Sequeira F, Teixeira J, Ferrand de Almeida F (2001) *Guia FAPAS - Anfíbios e répteis de Portugal*. Ed. FAPAS. Porto

Fisher RA (1930) *The genetical theory of natural selection.* Oxford University Press, Oxford, UK.

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.

Fritz U, Barata M, Busack SD, Fritzsch G, Castilho R (2006) Impact of mountain chains, sea straits and peripheral populations on genetic and taxonomic structure of a freshwater turtle, *Mauremys leprosa* (Reptilia, Testudines, Geoemydidae). *Zoologica Scripta* **35,** 97–108.

Fromhage L, Vences M, Veith M (2004) Testing alternative vicariance scenarios in Western Mediterranean discoglossid frogs. *Molecular Phylogenetics and Evolution*, **31,** 308–322.

Fu J (2000) Toward the phylogeny of the family Lacertidae – why 4708 base pairs of mtDNA sequences cannot draw the picture. *Biological Journal of the Linnean Society,* **71,** 203–217.

Galán P (1986) Morfología y distribución del género *Podarcis* Wagler, 1830 (Sauria, Lacertidae) en el noroeste de la Península Ibérica. *Revista Española de Herpetología,* **1,** 87-142.

Galán P (2002) Hibridación en laboratorio de *Podarcis bocagei* y *Podarcis carbonelli*. *Boletin de la Asociacion Herpetologica Española,* **13,** 28–31.

Gantenbein B, Lagiardèr CR (2003) The phylogeographic importance of the Strait of Gibraltar as a gene flow barrier in terrestrial arthropods: a case study with the scorpion *Buthus occitanus* as model organism. *Molecular Phylogenetics and Evolution*, **28**, 119–130.

García-Marin J-L, Utter FM, Pla C (1999) Postglacial colonization of brown trout in Europe based on distribution of allozyme variants. *Heredity*, **82,** 46–56.

Garcia-Paris M, Alcobendas M, Alberch P (1998) Influence of the Guadalquivir river basin on mitochondrial DNA evolution of *Salamandra salamandra* (Caudata: Salamandridae) from southern Spain, *Copeia,* **1998,** 173–176.

García-París M, Jockush EL (1999) A mitochondrial DNA perspectiva on the evolution of Iberian *Discoglossus* (Amphibia: Anura). *Journal of Zoology*, **248,** 209–218.

Gasc JP, Cabela A, Crnobrnja-Isailovic J. *et al.* (eds.) (1997). *Atlas of amphibians and reptiles in Europe.* Muséum National d'Histoire Naturelle, Paris, France.

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. Mémoire présenté pour l'obtention du diplôme de l'École Pratique des Hautes Etudes, Montpellier.

Geraldes A, Ferrand N, Nachman MW (2006) Contrasting patterns of introgression at X-linked loci across the hybrid zone between subspecies of the European rabbit (*Oryctolagus cuniculus*). *Genetics,* **173,** 919–933.

Godinho R, Mendonça B, Crespo EG, Ferrand N (2006) Genealogy of the nuclear b-fibrinogen locus in a highly structured lizard species: comparison with mtDNA and evidence for intragenic recombination in the hybrid zone. *Heredity,* **16,** 454–463.

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origin and Conservation of European Biodiversity* (Weiss S, Ferrand N, eds.). Springer, Dordrecht, The Netherlands, 155–188.

Greenberg AJ, Moran JR, Fang S, Wu C-I (2006) Adaptive loss of an old duplicated gene during incipient speciation. *Molecular Biology and Evolution,* **23,** 401–410.

Griswold CK, Baker AJ (2002) Time to the most recent common ancestor and divergence times of populations of common chaffinches (*Fringilla coelebs*) in Europe and North Africa: insights into Pleistocene refugia and current levels of migration. *Evolution,* **56,** 143–153.

Guillaume CP (1977) Premiers résultats d'une étude comparative entre *Lacerta muralis* et *Lacerta hispanica*. *Bulletin de la Société Zoologique de France*. **102,** nº 4.

Guillaume CP (1987) *Les petits lacértides du bassin méditerranéen occidental (Genera* Podarcis *et* Archaeolacerta *essentiellement)*. Thèse de Doctorat. Montpellier, Univ. Sciences et Techniques du Languedoc.

Guillaumet A, Pons J-M, Godelle B, Crochet P-A (2006) History of the Crested Lark in the Mediterranean region as revealed by mtDNA sequences and morphology. *Molecular Phylogenetics and Evolution*, **39,** 645–656.

Guiller A, Coutellec-Vreto MA, Madec L, Jeuniff J (2001) Evolutionary history of land snail *Helix aspersa* on the Western Mediterranean: preliminary results inferred from mitochondrial DNA sequences. *Molecular Ecology*, **10,** 81–87.

Guillot G, Estoup A, Mortier F, Cosson JF (2005) A spatial statistical model for landscape genetics. *Genetics*, **170,** 1261–1280.

Gyllensten U, Wharton D, Josefson A, Wilson A (1991) Paternal inheritance of mitochondrial DNA in mice. *Nature*, **352,** 255–257.

Haldane JBS (1932) The time of action of genes, and its bearing on some evolutionary problems. *The American Naturalist,* **66,** 5–24.

Harr B (2006) Genomic islands of differentiaion between house mouse subspecies. *Genome Research*, **16,** 730–737.

Harris DJ, Arnold EN (1999) Relationships and evolution of wall lizards *Podarcis* (Reptilia: Lacertidae) based on mitochondrial DNA sequences. *Copeia,* **1999**, 719–754.

Harris DJ, Arnold EN, Thomas RH (1998) Relationships of lacertid lizards (Reptilia: Lacertidae) estimated from mitochondrial DNA sequences and morphology. *Proceedings of the Royal Society of London B: Biological Sciences*, **265,** 1939–1948.

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,** 1257–1262.

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, 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, Batista V, Carretero MA (2004) Assessment of genetic diversity within *Acanthodactylus erythrurus* (Reptilia: Lacertidae) in Morocco and the Iberian Peninsula using mitochondrial DNA sequence data. *Amphibia-Reptilia,* **25,** 227–232.

Harris H (1966) Enzyme polymorphisms in Man. *Proceedings of the Royal Society of London, B: Biological Sciences,* **164,** 298–310.

Hein J, Schierup MH, Wiuf C (2005) *Gene genealogies, variation and evolution. A primer in coalescent theory.* Oxford University Press, Oxford, UK.

Hey J (2001) The mind of the species problem. *Trends in Ecology and Evolution,* **7,** 326–329.

Hey J (2006) On the failure of modern species concepts. *Trends in Ecology and Evolution,* **21,** 447–450.

Hey J, Machado CA (2003) The study of structured populations – new hope for a difficult and divided science. *Nature Reviews Genetics,* **4,** 535–543.

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.

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, **294,** 2310–2314.

Hsü KJ, Montadert L, Bernoulli D, *et al.* (1977) History of the Mediterranean salinity crisis. *Nature*, **267,** 399–403.

Jaarola M, Searle JB (2002) Phylogeography of field voles (*Microtus agrestis*) in Eurasia inferred from mitochondrial DNA sequences. *Molecular Ecology*, **11,** 2613–2621.

Jolly D, Harrison SP, Damnati B, BonneWlle R (1998) Simulated climate and biomes of Africa during the late quaternary: comparison with pollen and lake status data. *Quaternary Science Reviews*, **17,** 629–657.

Juste J, Ibáñez C, Muñoz J *et al.* (2004) Mitochondrial phylogeography of the long-eared bats (*Plecotus*) in the Mediterranean Palaearctic and Atlantic Islands. *Molecular Ecology*, **31,** 1114–1126.

Kingman JFC (1982) The coalescent. *Stochastic Processes and Their Applications,* **13,** 235–248.

Klemmer K (1959) Systematishe Stellung und Rassengliederung der spanishen Mauereidechse, *Lacerta hispanica*. *Senckenbergiana Biologica*, **40,** 245–250.

Knowles LL (2004) The burgeoning field of statistical phylogeography. *Journal of Evolutionary Biology*, **17,** 1–10.

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 USA,* **86,** 6196–6200.

Kondrashov AS (1983) Multilocus model of sympatric speciation. II. Two characters. *Theoretical Population Biology,* **24,** 136–144.

Kvist L, Martens J, Nazarenko AA, Orell M (2003) Paternal leakage of mitochondrial DNA in the Great Tit (*Parus major*). *Molecular Biology and Evolution,* **20,** 243–247.

Krijgsman W, Garcés M, Agustí J, Raffi I, Taberner C, Zachariasse WJ (2000) The "Tortonian salinity crisis" of the Eastern Betics (Spain). *Earth and Planetary Science Letters,* **181,** 497–511.

Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature,* **400,** 652–655.

Krijgsman W, Langereis CG (2000) Magnetostratigraphy of the Zobzit and Koudiat Zarga sections (Taza-Guercif basin, Morocco): implications for the evolution of the Rifian Corridor. *Marine and Petroleum Geology*, **17,** 359–371.

Lenk P, Fritz U, Joger U, Winks M (1999) Mitochondrial phylogeography of the European pond turtle, *Emys orbicularis* (Linnaeus 1758). *Molecular Ecology*, **8,** 1911–1922.

Lewontin RC, Hubby JL (1966) A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudobscura*. *Genetics*, **54,** 595–609.

Lornegan L, White N (1997) Origin of the Betic-Rif mountain belt. *Tectonics*, **16,** 504–522.

Loureiro A, Carretero MA, Ferrand N, Paulo OS, eds. (in press) *Atlas de Anfíbios e Répteis de Portugal*. Instituto de Conservação da Natureza, Lisboa, Portugal.

Machado CA, Hey J (2003) The causes of phylogenetic conflict in a classic *Drosophila* species group. *Proceedings of the Royal Society of London B,* **270,** 1193–1202.

Machordom A, Suárez J, Almodóvar A, Bautista JM (2000) Mitochondrial haplotype variation and phylogeography of Iberian brown trout populations. *Molecular Ecology*, **9,** 1325–1338.

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.

Maddison WP (1997) Gene trees in species trees. *Systematic Biology,* **46,** 523–536.

Maddison WP, Knowles LL (2006) Inferring phylogenies despite incomplete lineage sorting. *Systematic Biology*, **55,** 21–30.

Magraner J (1986) Nouvelle donnée sur la répartition de *Podarcis bocagei carbonelli* (V. Pérez Mellado, 1981), dans la Peninsule Ibérique et observations sur son écologie à Doñana (Andalousie, Espagne). *Bulletin de la Societé Herpétologique de France,* **38**, 6–12.

Martínez-Solano I (2004) Phylogeography of Iberian *Discoglossus* (Anura: Discoglossidae). *Journal of Zoological Systematics and Evolutionary Research*, **42,** 298–305.

Martínez-Solano I, Gonçalves H, Arntzen JW, García-París M (2004) Phylogenetic relationships and biogeography of midwife toads (Discoglossidae: *Alytes*). *Journal of Biogeography,* **31,** 603–618.

Martínez-Solano I, Teixeira J, Buckley D, Garcia-París M (2006) Mitochondrial DNA phylogeography of *Lissotriton boscai* (Caudata, Salamandridae): evidence for old, multiple refugia in an Iberian endemic. *Molecular Ecology,* **13,** 3375–3388.

Mayer W, Tiedemann F (1980) Elektrophoretische Untersuchungen an europäischen Arten der Gattungen *Lacerta* and *Podarcis*. I. Die *Podarcis*-Formen der griechischen Inseln Milos und Skiros. *Zeitschrift fur Zoologische Systematik und Evolutionsforschung*, **18,** 147–152.

Mayr E (1942) *Systematics and the Origin of Species*. Columbia University Press, New York, USA.

Mertens R, Müller L (1940) Die Amphibien und Reptilien Europas. *Abh. Senckenbergische Naturforschende Gesellschaft*, **41,** 1–62.

Mertens R, Wermuth H (1960) *Die Amphibien und Reptilien Europas.* Kramer, Frankfurt-am-Main, Germany.

Michaux JR, Libois R, Filippucci M-G (2005) So close and so different: comparative phylogeography of two small mammal species, the Yellow-necked fieldmouse (*Apodemus flavicollis*) and the Woodmouse (*Apodemus sylvaticus*) in the Western Palearctic region. *Heredity*, **94**, 52–63.

Michaux JR, Magnanou E, Paradis E, Nieberding C, Libois R (2003) Mitochondrial phylogeography of the Woodmouse (*Apodemus sylvaticus*) in the Western Palearctic region. *Molecular Ecology,* **12,** 685–697.

Miller MP (2005) Alleles in space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. *Journal of Heredity*, **96,** 722–724.

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.

Muller HJ, (1940) Bearing of the *Drosophila* work on systematics, pp.185-268 in *The new systematics*. Clarendon press, Oxford, U.K.

Muller HJ (1942) Isolating mechanisms, evolution and temperature. *Biological Symposium,* **6,** 71-125.

Navarro A, Barton NH (2003) Chromosomal speciation and molecular divergence – accelerated evolution in rearranged chromosomes. *Science*, **300,** 321–324.

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

Nordborg M (2001) Coalescent theory. In: *Handbook of Statistical Genetics* (Balding DJ, Bishop M, Cannings C, eds.). Wiley, Chichester, UK, 179–212.

Noor MAF, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal invertions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences, USA,* **98,** 12084–12088.

Noor MAF, Feder JL (2006) Speciation genetics: evolving approaches. *Nature Reviews in Genetics*, **7,** 851–861.

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.

Olalde M, Herrán A, Espinel S, Goicoechea PG (2002) White oaks phylogeography in the Iberian Peninsula. *Forest Ecology and Management*, **156,** 89–102.

Oliverio M, Bologna MA, Mariottini P (2000) Molecular biogeography of the Mediterranean lizards *Podarcis* Wagler, 1830 and *Teira* Gray, 1838 (Reptilia, Lacertidae). *Journal of Biogeography*, **27,** 1403–1420.

Oliverio M, Bologna MA, Monciotti A, Annesi F, Mariottini P (1998) Molecular phylogenetics of the Italian *Podarcis* lizards (Reptilia, Lacertidae). *Italian Journal of Zoology,* **65,** 315–324.

O'Regan HJ, Turner A, Wilkinson DM (2002) European quaternary refugia: a factor in large carnivore extinction? *Journal of Quaternary Science,* **17,** 789–795.

Page RDM, Charleston MA (1997) From gene to organismal phylogeny: reconciled trees and the gene tree/species tree problem. *Molecular Phylogenetics and Evolution*, **7,** 231–240.

Palmer M, Cambefort Y (2000) Evidence for reticulate palaeogeography: beetle diversity linked to connection-disjunction cycles of the Gibraltar strair. *Journal of Biogeography,* **27,** 403–416.

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.

Paulo OS (2001) *Phylogeography of reptiles of the Iberian Peninsula*. PhD Thesis. University of London.

Paulo OS, Dias C, Bruford MW, Jordan WC, Nichols RA (2001) The persistence of Pliocene populations through the Pleistocene climatic cycles: evidence from the phylogeography of an Iberian lizard. *Proceedings of the Royal Society of London*. *Series B, Biological Sciences*, **268,** 1625–1630.

Paulo OS, Jordan WC, Bruford MW, Nichols RA (2002a) Using nested clade analysis to assess the history of colonization and the persistence of populations of an Iberian lizard. *Molecular Ecology*, **11,** 809–819.

Paulo OS, Pinto I, Bruford MW, Jordan WC, Nichols RA (2002b) The double origin of Iberian peninsular chameleons. *Biological Journal of the Linnean Society*, **75,** 1–7.

Pereira R (2005) *Combinação de marcadores moleculares e sistemas de informação geográfica no estudo da zone híbrida de* Lacerta schreiberi. MSc thesis, Faculdade de Ciências, Universidade do Porto.

Pérez Mellado V (1981a) La lagartija de Bocage, *Podarcis bocagei* (Seoane, 1884): Primeros datos sobre su distribución, colorido y ecología. *Amphibia-Reptilia,* **1,** 253–268.

Pérez Mellado V (1981b) Nuevos datos sobre la sistemática y distribución de *Podarcis bocagei* (Seoane, 1884) (Sauria, Lacertidae) en la Península Ibérica. *Amphibia-Reptilia,* **2,** 259–265.

Pérez-Mellado V, Galindo MP (1986) Sistemática de *Podarcis* (Sauria, Lacertidae) ibéricas y norteafricanas mediante técnicas multidimensionales. *Série Manuales Universitários,* Univ. Salamanca, Spain.

Petit RJ, Brewer S, Bordacs S, *et al*. (2002) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management* **156**, 49-74.

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 (2002) *Differentiation between and within forms of* Podarcis *lizards in the Iberian Peninsula and North Africa: evidence from protein electrophoretic data.* Relatório de Estágio, Faculdade de Ciências, Universidade do Porto.

Pleguezuelos JM, Márquez R, Lizana M (eds). (2002) *Atlas y libro rojo de los anfíbios y reptiles de España*. Dirección General de Conservación de la Naturaleza – Asociación Herpetologica Española, 2nd Edition, Madrid,

Podnar M, Mayer W, Tvrtković N (2004) Mitochondrial phylogeography of the Dalmation wall lizard, *Podarcis melisellensis* (Lacertidae). *Organisms, Diversity and Evolution,* **4,** 307–317.

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 Península, by bayesian and maximum-likelihood analyses of mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution,* **37,** 845–857.

Presgraves DC, Bagalavan L, Abmayr SM, Orr HA (2003) Adaptive evolution drives divergence in a hybrid inviability gene between two species of *Drosophila. Nature,* **423,** 715–719.

Ribera I, Vogler AP (2004) Speciation of Iberian diving beetles in Pleistocene refugia (Coleoptera, Dysticidae). *Molecular Ecology,* **13,** 179–193.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution*, **16,** 351–358.

Robalo JI, Santos CS, Almada VC, Doadrio I (2006) Paleobiogeography of two Iberian endemic cyprinid fishes (*Chondrostoma arcasii-Chondrostoma macrolepidotus*) inferred from mitochondrial DNA sequence data. *Journal of Heredity,* **97,** 143–149.

Rosembaum G, Lister GS, Duboz C (2002) Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. *Journal of the Virtual Explorer*, **8,** 107–130.

Rouchy JM, Caruso A (2006) The Messinian salinity crisis in the Mediterranean basin: a reassessment of the data and an integrated scenario. *Sedimentary Geology,* **188-189,** 35–67.

Roucoux KH, de Abreu L, Shackleton NJ, Tzedaki PC (2005) The response of NW Iberian vegetation to North Atlantic climate oscillations during the last 65 kyr. *Quaternary Science Reviews*, **24**, 1637–1653.

Rowe G, Harris DJ, Beebee TJC (2006) Lusitania revisited: a phylogeographic analysis of the natterjack toad *Bufo calamita* across its entire biogeographical range. *Molecular Phylogenetics and Evolution*, **39,** 335–346.

Sá-Sousa P (1995) Distinção das lagartixas do género *Podarcis* em Portugal. *Folha Herpetológica (SPH)* **5,** 4-6.

Sá-Sousa P (1998) Distribución de la lagartija *Podarcis b. bocagei* en Portugal. *Boletin de la Asociación Herpetológica Española,* **9,** 2-4.

Sá-Sousa P (1999) New data on the distribution of *Podarcis bocagei carbonelli* Pérez-Mellado, 1981 in Portugal. *Herpetozoa,* **12,** 87-90.

Sá-Sousa P (2000a) Distribución de la lagartija *Podarcis bocagei carbonelli* Pérez-Mellado 1981 en Portugal. *Boletin de la Asociación Herpetológica Española* **11,** 12–16.

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

Sá-Sousa P (2001) *A Controversa Sistemática das Lagartixas do género* Podarcis *Wagler, 1830 (Sauria: Lacertidae) em Portugal*. Dissertação de Doutoramento. Fac. Ciências, Univ. Lisboa.

Sá-Sousa P, Almeida AP, Rosa HD, Vicente L, Crespo EG (2000) Genetic and morphological relationships of the Berlenga wall lizard (*Podarcis bocagei berlengensis*: Lacertidae). *Journal of Zoological Systematics and Evolutionary Research*, **38,** 95-102

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,** 55–69.

Saetre G-P, Borge T, Moum T (2001) A new bird species ? The taconomic status of 'the Atlas flycatcher' assessed from DNA sequence analysis. *Ibis,* **143,** 494–497.

Salvador A (1974) *Guía de los Anfíbios y Reptiles Españoles*. ICONA, Madrid.

Salzburger W, Martens J, Sturmbauer C (2002) Paraphyly of the blue tit (*Parus caeruleus*) suggested from cytochrome *b* sequences. *Molecular Phylogenetics and Evolution*, **24,** 19–25.

Schmitt T, Giessl A, Seitz A (2003) Did *Polyommatus icarus* (Lepidoptera: Lycaenidae) have distinct glacial refugia in southern Europe? Evidence from population genetics. *Biological Journal of the Linnean Society,* **80,** 529–528.

Schmitt T, Röber S, Seitz A (2005) Is the last glaciation the only relevant event for the present genetic population structure of the meadow brown butterfly *Maniola jurtina* (Lepidoptera: Nymphalidae)? *Biological Jounrla of the Linnean Society* **85,** 419–431.

Sequeira F, Alexandrino J, Rocha S, Arntzen JW, Ferrand N (2005). Genetic exchange across a hybrid zone within the Iberian endemic golden-striped salamander, *Chioglossa lusitanica*. *Molecular Ecology* **14,** 245–254.

Sequeira F (2006) História evolutiva da salamandra-lusitânica, *Chioglossa lusitanica*. Filogenia, reconstrução dos padrões filogeográficos e análise dos processos de miscigenação na zona híbrida através do uso de múltiplos marcadores moleculares. Dissertação de Doutoramento, Fac. Ciências, Univ. Porto.

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, USA,* **99,** 16122–16127.

Sinclair WT, Morman JD, Ennos RA (1999) The postglacial history of Scots pine (*Pinus sylvestris* L.) in western Europe: evidence from mitochondrial DNA variation. *Molecular Ecology*, **8,** 83–88.

Steinfarz S, Veith M, Tautz D (2000) Mitochondrial sequence analysis of *Salamandra* taxa suggests old splits of major lineages and postglacial recolonizations of Central Europe from distinct source populations of *Salamandra salamandra. Molecular Ecology*, **9,** 397–410.

Surget-Groba Y, Heulin B, Guillaume C-P, Thorpe RS, Kupryianova L, Vogrin N *et al.* (2001) Intraspecific phylogeography of *Lacerta vivipara* and the evolution of viviparity. *Molecular Phylogenetics and Evolution* **18**, 449–459.

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

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypothese about gene flow and population history. *Molecular Ecology,* **7,** 381–397.

Terrasa B, Picornell A, Castro JA, Ramon MM (2004) Genetic variation within endemic *Podarcis* lizards from the Balearic Islands inferred from partial cytochrome *b* sequences. *Amphibia-Reptilia,* **25,** 407–414.

Ting C-T, Tsaur S-C, Wu M-L, Wu C-I (1998) A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science*, **282**, 1501–1504.

Tsaousis AD, Martin DP, Ladoukakis ED, Posada D, Zouros E (2005) Widespread recombination in published animal mtDNA sequences. *Molecular Biology and Evolution*, **22,** 925–933.

Tzedakis PC. Lawson IT, Frogley MR, Hewitt GM, Preece RC (2002) Buffered tree population changes in a Quaternary refugium: evolutionary implications. *Science*, **297,** 2044–2047.

Vasconcelos 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.

Veith M, Mayer C, Samroui B, Donaire Barroso D, Bogaerts S (2004) From Europe to Africa and vice-versa: evidence for multiple intercontinental dispersal in ribbed salamanders (Genus *Pleurodeles*). *Journal of Biogeography*, **31,** 159–171

Via S (2001) Sympatric speciation in animals: the ugly duckling grows up. *Trends in Ecology and Evolution*, **16,** 381–390.

Vicente LA (1985) Description d'une nouvelle sous-espèce de *Podarcis bocagei* (Seoane, 1884) (Sauria, Lacertidae) de l'île de Berlenga: *Podarcis bocagei berlengensis. Bulletin du Muséum National d'Histoire Naturelle de Paris, 4e sér.* **7,** 267–274.

Vila M, Vidal-Romaní JR, Björklund M (2005) The importance of time scale and multiple refugia: incipient speciation and admixture of lineages in the butterfly *Erebia triaria* (Nymphalidae). *Molecular Phylogenetics and Evolution,* **36,** 249–260.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology,* **7**, 256–276.

Webb T, Bartlein PJ (1992) Global changes during the last 3 million years: climatic controls and biotic responses. *Annual Review of Ecology and Systematics*, **23,** 141–173.

Weijermars R (1991) Geology and tectonics of the Betic Zone, SE Spain. *Earth-Science Reviews,* **31,** 153–236.

Weiss S, Ferrand N (2006) Current perspectives in phylogeography and the significance of South European refugia in the creation and maintenance of European biodiversity. In: *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origin and Conservation of European Biodiversity* (Weiss S, Ferrand N, eds.). Springer, Dordrecht, The Netherlands, 341–357.

Wittbrodt J, Adam D, Malitschek B *et al.* (1989) Novel putative receptor tyrosine kinase encoded by the melanoma-inducing *Tu* locus in *Xiphophorus*. *Nature*, **341,** 415–421.

Wright S (1931) Evolution in Mendelian populations. *Genetics,* **28,** 114–138.

Wu C-I, Ting C-T (2004) Genes and Speciation. *Nature Reviews in Genetics*, **5,** 114–122.

Zagwijn WH (1992a) The beginning of the Ice Age in Europe and its major subdivisions. *Quaternary Science Reviews,* **11**, 583–591.

Zagwijn WH (1992b) Migration of vegetation during the Quaternary in Europe. *Courier Forschunginstitut Senckenberg*, **153**, 9–20.

Zangari F, Cimmaruta R, Nascetti G (2006) Genetic relationships of the western Mediterranean painted frogs based on allozymes and mitochondrial markers: evolutionary and taxonomic inferences (Amphibia, Anura, Discoglossidae). *Biological Journal of the Linnean Society,* **87,** 515–536.

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology,* **12,** 563–584.

Chapter 2

Differentiation, evolutionary relationships and gene flow among forms of the Iberian and North African *Podarcis* species complex

Article I. Reexamination of the Iberian and North African *Podarcis* (Squamata: Lacertidae) phylogeny based on increased mitochondrial DNA sequencing.

CATARINA PINHO, NUNO FERRAND & D. JAMES HARRIS, 2006
Molecular Phylogenetics and Evolution, **38,** 266–273.

Article II. Genetic polymorphism of 11 allozyme loci in populations of wall lizards (*Podarcis* sp.) from the Iberian Peninsula and North Africa.

CATARINA PINHO, D. JAMES HARRIS & NUNO FERRAND, 2003
Biochemical Genetics, **41,** 343–359.

Article III. 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 & NUNO FERRAND
Biological Journal of the Linnean Society, in press.

Article IV. Abundant ancestral polymorphism at the nuclear loci *β-fibint7* and *6-Pgdint7* characterizes different *Podarcis* species from the Iberian Peninsula and North Africa.

CATARINA PINHO, D. JAMES HARRIS & NUNO FERRAND
In preparation

Article V. Genetic admixture between the Iberian endemic lizards *Podarcis bocagei* and *P. carbonelli*: evidence for limited natural hybridization and a bimodal hybrid zone.

CATARINA PINHO, ANTIGONI KALIONTZOPOULOU, MIGUEL A. CARRETERO, D. JAMES HARRIS & NUNO FERRAND
Submitted

Reexamination of the Iberian and North African *Podarcis* (Squamata: Lacertidae) phylogeny based on increased mitochondrial DNA sequencing

CATARINA PINHO, NUNO FERRAND, D. JAMES HARRIS

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

Keywords: *Podarcis*; phylogeny; cytochrome *b*; 12S rRNA; 16S rRNA; control region; ND4; biogeography;

Introduction

In recent years, molecular studies have revealed complex patterns of differentiation within many European herpetological taxa. Wall lizards (*Podarcis* spp.) were one of the most studied groups, perhaps due to the complexity of defining taxonomical entities within this genus by morphological analysis only, due to an extreme intraspecific variability coupled with low variation between species. Regarding Iberian forms, this difficulty is illustrated by the variety of different proposals for the classification of these lizards (Mertens and Müller 1940, Klemmer 1959, Arnold and Burton 1978, Geniez 2001, Sá-Sousa 2001). Despite the controversy, the most common view (Arnold and Burton 1978) accepted three species within the Iberian Peninsula: *P. muralis* Laurenti 1760, *P. hispanica* (Steindachner 1870) and *P. bocagei* (Seoane 1884). According to some authors (Barbadillo *et al.* 1999), *P. hispanica* was further subdivided into two subspecies: the nominal form, in the Iberian Peninsula, and *P. h. vaucheri*, in North Africa. Oliverio *et al.* (2000) also proposed the elevation of the African populations to the species rank as *P. vaucheri* on the basis of high genetic distances between single individuals from each side of the Strait of Gibraltar, obtained using the 12S rRNA gene.

Genetic studies addressing the phylogeny of the whole genus *Podarcis* (Harris and Arnold 1999, Oliverio *et al.* 2000, Harris *et al.* 2005) have shown that all Iberian and North African lizards form a clade, with the exception of *P. muralis*, which exists nowadays in northern Iberia and is widely distributed across Europe, most likely after an expansion from the Italian Peninsula (Harris and Arnold, 1999; Oliverio *et al.* 2000). Recently, Iberian forms have been characterized using morphology (Geniez 2001, Sá-Sousa *et al.* 2002), sequencing of partial cytochrome *b* and 12S rRNA mitochondrial genes (Harris and Sá-Sousa 2001, 2002, Harris *et al.* 2002b) and allozyme data (Pinho *et al.* 2003, 2004). All these studies clearly showed that the actual variation within Iberian lizards is much more complex than that concealed behind species' names, with multiple forms being present both in the Iberian Peninsula and North Africa. The first outcome of these studies was the elevation of *P. carbonelli* Pérez-Mellado 1981 to species status after morphological and molecular data corroborated its distinctiveness from *P. bocagei*, in which this form had been included (Sá-Sousa and Harris 2002). The second outcome was the recognition of *Podarcis hispanica* as a species complex, since the forms included in this designation do not form a clade relative to *P. bocagei* and *P. carbonelli* (Harris and Sá-Sousa 2002, Harris *et al.* 2002b, Pinho *et al.* 2003, 2004).

However, although the existence of multiple forms within *P. hispanica* from the Iberian Peninsula and North Africa was well documented both by mitochondrial DNA (mtDNA) and allozyme studies, neither data set convincingly

supported relationships between forms, moreover suggesting slightly different evolutionary scenarios. This fact was interpreted as either being caused by the independent evolution of each kind of marker, leading to different phylogenetic relationships, or by the forms having split during a short period of time, leading to an unresolved phylogeny and short internal branches (Pinho *et al.* 2003). To address this question and to obtain a more reliable estimate of the mitochondrial phylogeny, we extended the sampling of the mitochondrial genes studied. Besides 12S rRNA and cytochrome *b*, previously studied in other works, we obtained sequences for three other mitochondrial DNA regions: 16S rRNA, NADH dehydrogenase subunit 4 (ND4) and adjacent t-RNAs, and the control region in a set of 32 individuals representing all known lineages. By sequencing different portions of mitochondrial DNA, we also reduce the probability of putative nuclear copies influencing the final estimates of the phylogeny, because the studied genes are scattered along the mtDNA molecule.

Materials and Methods

Sampling

Thirty-two specimens, including all the known morphotypes of *P. hispanica, P. bocagei, P. carbonelli* and the outgroup *P. muralis*, were analysed (Figure 1 and Table 1). Morphological identification or prior sequencing of at least one mitochondrial DNA gene allowed the assignment of individuals to morphotypes/lineages as described in Harris and Sá-Sousa (2002). One of the individuals included in the study (Gal3) was found not to belong to any genetic lineage described to date and is herein referred to as "unidentified type". Whenever possible, more than one individual for each form was analysed.

DNA extraction, amplification and sequencing

Total genomic DNA was extracted from small portions of alcohol-preserved tail muscle following standard methods (Sambrook *et al.* 1989). We amplified a portion of five mitochondrial regions: cytochrome *b*, 12S rRNA, 16S rRNA, ND4-$tRNA^{LEU}$ and the control region. The chosen fragments were amplified using published primers (Table 2), except the ND4-$tRNA^{LEU}$ fragment, for which an additional specific primer was designed in the $tRNA^{HIS}$ region (HisR primer, sequence 5'-CTAGAGTCACAGTCTAGTGTTTT-3'). Amplifications of 12S, 16S, cytochrome *b* and the control region were carried out in 25µL volumes, containing 2.5µL 10X reaction buffer (Ecogen), 3.0mM $MgCl_2$, 0.4mM each dNTP, 0.4µM each primer, 1 unit of *Ecotaq* DNA polymerase (Ecogen) and approximately 100 ng of genomic DNA. Polymerase chain reaction (PCR) conditions consisted of a pre-denaturing step of 3 min at 94ºC, and 30-35 cycles

of denaturing (94ºC, 30 s), annealing (52ºC for the control region, 50ºC for the rest of the fragments, 30s) and extension (72ºC, 30s). A final extension was conducted at 72ºC for 3 min. Amplifications of the ND4-$tRNA^{LEU}$ and ND4-$tRNA^{HIS}$ fragments were conducted in 25µL volumes, containing 2.5µL 10X reaction buffer (Ecogen), 3.2mM $MgCl_2$, 0.4mM each dNTP, 0.2µM each primer, 1 U of *Ecotaq* DNA polymerase (Ecogen) and approximately 50 ng genomic DNA. Amplification conditions consisted of a pre-denaturing step of 3 min at 94ºC followed by 35 cycles of a denaturing step of 30s at 94ºC, annealing at 54ºC for 30s and extension at 72ºC for 40s. The final extension was accomplished at 72ºC for 4 min.

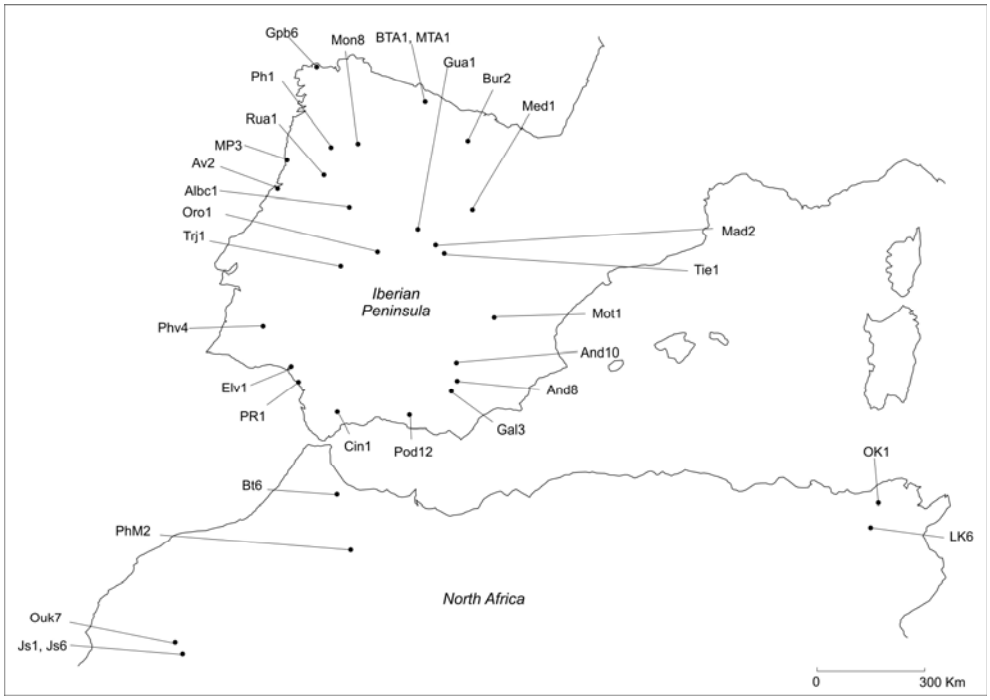

Figure 1. Map showing the geographical origin of samples analysed in this study.

The PCR products were enzymatically purified and sequenced using the ABI Prism BigDye Terminator Cycle sequencing protocol in an ABI Prism 310 automated sequencer (Applied Biosystems) with the same primers used for amplification.

Table 1. Specimen type, locality data and GenBank accession numbers for samples included in this study

Specimen type	Sample code	Locality	GenBank accession numbers				
			cytochrome *b*	12S rRNA	16S rRNA	ND4 and tRNAs	control region
P. bocagei	BTA1	Tanes, Asturias, Spain	DQ081139	DQ081064	DQ081077	DQ081153	DQ081109
	MP3	Madalena, Portugal	AF469424*	AF469423*	DQ081076	DQ081152	DQ081108
	Gpb6	Malpica, Galicia, Spain	AF469426*	AF469425*	DQ081075	DQ081151	DQ081107
P. carbonelli	Av2	Aveiro, Portugal	DQ081140	DQ081065	DQ081078	DQ081154	DQ081110
	Albc1	La Alberca, Castilla y León, Spain	DQ081142	DQ081066	DQ081080	DQ081156	DQ081112
	PR1	Playa del Rompeculos, Andalucía, Spain	DQ081141	AY214449*	DQ081079	DQ081155	DQ081111
P. hispanica morphotype 1	Mon8	Montesinho, Portugal	AF469447*	AF469446*	DQ081086	DQ081162	DQ081118
	Rua1	Vila da Rua, Portugal	AF469445*	AF469444*	DQ081085	DQ081161	DQ081117
	Ph1	Vila Real, Portugal	AF372084*	AF469443*	DQ081087	DQ081163	DQ081119
	Oro1	Oropesa, Castilla-La Mancha, Spain	AF469453*	AF469452*	DQ081089	DQ081165	DQ081121
	Trj1	Trujillo, Extremadura, Spain	AF469451*	AF469450*	DQ081088	DQ081164	DQ081120
P. hispanica morphotype 2	Phv4	Beja, Portugal	AF469455*	AF469454*	DQ081083	DQ081159	DQ081115
	Mad2	Madrid, Spain	AF469460*	AF469459*	DQ081082	DQ081158	DQ081114
	And10	Benatae, Andalucia, Spain	DQ081143	DQ081067	DQ081084	DQ081160	DQ081116
	Tie1	Tielmes, Madrid, Spain	AY134680*	AY134715*	DQ081081	DQ081157	DQ081113
P. hispanica morphotype 3	Bur2	Burgos, Castilla y León, Spain	DQ081144	DQ081068	DQ081090	DQ081166	DQ081122
(NE Iberia)	Med1	Medinaceli, Castilla y León, Spain	AF469436*	AF469435*	DQ081091	DQ081167	DQ081123
P. hispanica sensu stricto	And8	Puebla de D. Fadrique, Andalucía, Spain	DQ081145	DQ081069	DQ081093	DQ081169	DQ081125
	Pod12	Granada, Andalucía, Spain	AF469428*	AF469427*	DQ081094	DQ081170	DQ081126
	Mot1	Motilla del Palancar, Castilla-La Mancha, Spain	AY134677*	AY134712*	DQ081092	DQ081168	DQ081124
P. (hispanica) vaucheri	PhM2	Mid Atlas Mts, Morocco	AF372083*	AF469416*	DQ081097	DQ081173	DQ081129
	Bt6	Bab Taza, Morocco	AY134681*	AY134716*	DQ081098	DQ081174	DQ081130
	Ouk7	Oukaïmeden, Morocco	AY134678*	AY134713*	DQ081096	DQ081172	DQ081128
	Cin1	Guadalcacin, Andalucía, Spain	AY134679*	AY134714*	DQ081099	DQ081175	DQ081131
	Elv1	Huelva, Andalucía, Spain	AY134674*	AY134709*	DQ081100	DQ081176	DQ081132
P. hispanica Jebel Sirwah	Js1	Jebel Sirwah, Morocco	AY134672*	AY134707*	DQ081104	DQ081180	DQ081136
	Js6	Jebel Sirwah, Morocco	AY134682*	AY134717*	DQ081103	DQ081179	DQ081135
P. hispanica Tunisia	OK1	Oued Kébir, Tunisia	DQ081148	DQ081072	DQ081102	DQ081178	DQ081134
	LK6	Le Kef, Tunisia	DQ081147	DQ081071	DQ081101	DQ081177	DQ081133
*P. hispanica***	Gal3	Galera, Andalucia, Spain	DQ081146	DQ081070	DQ081095	DQ081171	DQ081127
P. muralis	MTA1	Tanes, Asturias, Spain	DQ081150	DQ081074	DQ081106	DQ081182	DQ081138
	Gua1	Guadarrama, Madrid, Spain	DQ081149	DQ081073	DQ081105	DQ081181	DQ081137

*indicates gene region previously published (Harris *et al.* 2002a,b; Harris and Sá-Sousa 2002)
**unidentified type

Table 2. Primers used in this study

Gene fragment	Primers used	Reference
Cytochrome *b*	cytochrome *b* 1	Kocher *et al.* (1989)
	cytochrome *b* 2	Kocher *et al.* (1989)
12s rRNA	12s*a*	Kocher *et al.* (1989)
	12s*b*	Kocher *et al.* (1989)
16s rRNA	16sL1	Hedges and Bezy (1993)
	16sH1	Hedges and Bezy (1993)
ND4-tRNALEU	ND4	Arévalo *et al.* (1994)
	Leu	Arévalo *et al.* (1994)
	HisR	this study
Control region	DL3F	Crochet *et al.* (2004)
	DL4R	Crochet *et al.* (2004)

Sequences were aligned manually using Bioedit v. 5.0.9 (Hall 1999). The ND4-tRNA^{LEU} fragment was subdivided into the partial ND4 sequence, tRNA^{HIS} and tRNA^{SER}. The tRNA^{LEU} portion was excluded from the analysis because sequences were incomplete. The cytochrome *b*, ND4 and tRNA^{SER} sequences contained no indels. Alignment of the 12S, 16S, control region and tRNA^{HIS} required insertions or deletions in five, eight, two, and two places, respectively.

Phylogenetic analysis

Sequences were imported into PAUP* 4.0b10 (Swofford 2002). For the phylogenetic analysis of the combined data, we used maximum likelihood (ML), maximum parsimony (MP) and Bayesian inference. When estimating phylogenetic relationships among sequences, one assumes a model of evolution. We used the approach outlined by Huelsenbeck and Crandall (1997) to test 56 alternative models of evolution, employing PAUP* 4.0b10 and Modeltest 3.06 ((Posada and Crandall (1998), described in detail in Posada and Crandall (2001)). Once a model of evolution was chosen, it was used to estimate a tree using ML (Felsenstein 1981) with random sequence addition (10 replicate heuristic searches). The MP analysis was also performed with random sequence addition (10 replicate heuristic search), and support for nodes was estimated using the nonparametric bootstrap technique (Felsenstein 1985) with 100 replicates. The Bayesian analysis was implemented using MrBayes (Huelsenbeck and Ronquist 2001), which calculates Bayesian posterior probabilities using a Metropolis-coupled, Markov chain Monte Carlo (MC-MCMC) sampling approach. Bayesian analyses were conducted with random starting trees, run 0.5 X 10^6 generations, and sampled every 100 generations using a general-time-reversible model of evolution with a gamma model of among-site rate variation. In both searches, stationarity of the Markov Chain was determined as the point when sampled ln-likelihood values plotted against generation time reached a stable mean equilibrium value; "burn-in" data sampled from generations preceding this point were discarded. All data collected at stationarity were used to estimate

posterior nodal probabilities and a summary phylogeny. Two independent replicates were conducted and inspected for consistency to check for local optima (Huelsenbeck and Bollback 2001).

We used the cytochrome *b* data set to infer divergence times. Molecular clock assumptions were tested using a likelihood-ratio test.

Results and discussion

Including the two outgroups, 32 individuals from 12 differentiated species/entities were analysed and sequenced for the 5 mitochondrial regions. Several measures were taken to evaluate the possibility of the presence of nuclear copies among the data. Although the base composition (less than 5% of guanines in the third position, see Harris 2002) and the absence of stop codons in protein-coding genes conformed to the expectations of a mitochondrial origin, visual inspection of the sequences revealed that the ND4 sequence from sample Gal3 seemed to have an excess of amino acid substitutions when compared to the others. This was confirmed by plotting synonymous versus non-synonymous distances for the complete data set, which showed that for all the comparisons involving this particular sample there was an excess of non-synonymous substitutions (data available from the authors upon request). Also, in the phylogenetic analyses of the ND4 gene alone, the length of the branch leading to this sample was much higher than all the others. Taken together, these observations suggested that this sequence might correspond to a nuclear pseudogene. As an attempt to confirm this hypothesis, we designed an internal primer in the $tRNA^{HIS}$ region (HisR, see Materials and Methods), which was used to amplify a fragment of about 750bp when used together with primer ND4 (Arévalo *et al.* 1994). The sequence of Gal3 obtained using these primers, showed several heterozygous-like positions which were interpreted as resulting from an overlap of the mitochondrial and the nuclear sequence. From this, we were able to reconstruct the putative mitochondrial sequence. When plotting synonymous versus non-synonymous distances in this gene using this corrected sequence we did not find the bias that had previously been found. The putative nuclear copy sequence has been deposited in GenBank, Accession No. DQ081063.

Aligned sequences of the combined mitochondrial DNA genes were 2425bp long. From these, 306bp correspond to cytochrome *b*, 381bp to 12S rRNA, 509bp to 16S rRNA, 419bp to the control region, 675bp to ND4, 69bp to $tRNA^{HIS}$ and 66bp to $tRNA^{SER}$. The length of the sequences used in this study was approximately the same for all the samples, with the exception of the above mentioned Gal3, for which sequences lack around 30bp of the ND4 gene and the

complete tRNA genes due to the procedures described above. Of these 2425bp, 597 characters were variable and 530 parsimony informative. The most appropriate model of evolution for the data was the GTR model with an estimate of invariable sites (0.67) and a discrete approximation of the gamma distribution (4 rate categories, alpha=1.67). Using this model we recovered a single ML tree of -ln 10,181 (Figure 2). Maximum parsimony analysis recovered 4 equally parsimonious trees, the strict consensus of which differed from the ML tree only in being less well resolved. Since Bayesian analysis recovered the same tree as that derived from ML, only the ML tree is depicted in Figure 2, but MP bootstraps and Bayesian posterior probabilities have been overlaid onto this estimate of phylogeny.

The clade formed by the Northeastern type of *P. hispanica* and the newly found *P. hispanica* form from Galera, in the Baza Depression of SE Spain, is the sister taxon to all other Iberian and North African *Podarcis*. The remaining taxa are split into two very well supported clades: one comprising all the Western Iberian forms and one including all Southern Iberian and North African morphotypes. Within the first clade, there is strong support for two groups: one that encompasses *P. carbonelli* and morphotype 2 of *P. hispanica* (sensu Sá-Sousa *et al.* 2002) and another that comprises *P. bocagei* and morphotype 1 of *P. hispanica*. This morphotype is further subdivided into two geographically separated and well-supported clades, one corresponding to the Northwestern part of its distribution, whereas the other seems to be confined to the Spanish central system and surrounding areas. The second clade is composed of both Iberian and North African forms. As sister taxon to all the other forms in this clade is *P. hispanica* sensu stricto, from SE Iberia. The other forms can be subdivided into two groups: one including *P. (hispanica) vaucheri*, which is further subdivided into Iberian and North African forms, and another including sequences of *P. hispanica* from Tunisia and the Jebel Sirwah form from Southern Morocco.

Our estimates of phylogeny support most of the topology of previous works with less included characters (Harris *et al.* 2002b). However, bootstrap values have significantly increased, approaching 100% in branches where previous analyses had retrieved less than 50%. Therefore, besides confirming the monophyly of previously described groups, almost all the relationships between those forms are now well supported. There is, therefore, strong support for the monophyly of *P. hispanica* morphotype 1, which is formed by two very divergent lineages (around 9% cytochrome *b* divergence). The close relationship between this morphotype and *P. bocagei*, as well as that between *P. hispanica* type 2 and *P. carbonelli,* is also confirmed through the analysis of this data set. This is relevant in understanding the evolutionary patterns within the Iberian *Podarcis* because the current knowledge on the distribution of the lineages points to the

fact that the only forms with distributions that overlap extensively are precisely these sister taxa (Sá-Sousa 2000, Geniez 2001).

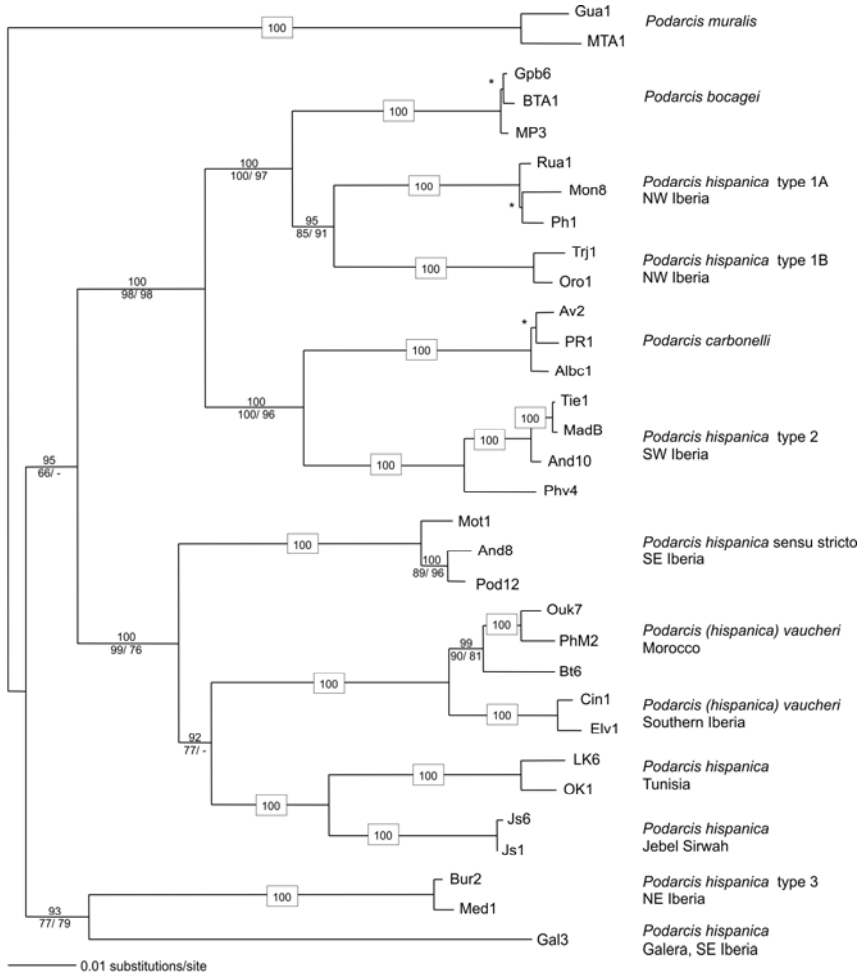

Figure 2. Estimate of relationships derived from a ML search using the GTR model with an estimate of invariable sites (0.67) and a discrete approximation of the gamma distribution (4 rate categories, α=1.67). Bayesian analysis produced the same estimate of relationships. The tree was rooted using *P. muralis*. Bayesian posterior probabilities are given above nodes, ML and MP bootstraps, respectively, below nodes. When all three were identical one value is given in a box. (-) indicates bootstrap values lower than 50%. (*) indicates branches where different methods yielded different topologies.

In this study we describe a previously undetected lineage of *Podarcis* from the Iberian Peninsula. This lineage, represented in our work by sample Gal3, has been found only in the flat, arid depression between Sierra de Cazorla and Sierra de Baza, in the Betic region, south-eastern Spain. Previous works including more

samples from Southeastern Spain (Harris *et al.* 2002b) only detected *P. hispanica* sensu stricto. Therefore, taken together, our data seem to indicate that the distribution of this lineage may be very restricted. The same happens with the lineage from Jebel Sirwah, in Southern Morocco, which is known only from one locality. In our estimates of phylogeny, the newly described lineage clusters with the Northeastern type of *P. hispanica* with 79% bootstrap support in the ML analysis. The clustering of these two forms may imply that in the past the ancestor of this group had a widespread distribution throughout eastern Iberia, and that the Betic form might be a relic of that ancient distribution.

Harris et al. (2002b) proposed a biogeographic scenario for the colonization of North Africa by *Podarcis* in which these lizards crossed the Strait of Gibraltar twice after its formation, both times from the Iberian Peninsula to North Africa. This analysis thus discarded the hypothesis of vicariance caused by the opening of the Strait, occurred ca. 5.3 million years ago (Mya), after the Messinian salinity crisis, which began ca. 5.6 Mya and allowed land connection between Iberia and North Africa (Hsü *et al.* 1977). However, according to the phylogeny published by these authors, based on cytochrome *b* and 12S rRNA gene sequences, *P. hispanica* sensu stricto was not the sister taxon to all North African forms, but it clustered within this group. These relationships were weakly supported (less than 50% bootstrap support). Moreover, the authors used a calibration for the molecular clock based on the two genes combined derived from geckos (Carranza *et al.* 2000). In this work we chose to base our analysis on cytochrome *b* distances only, since more calibrations have been accomplished for this gene than for any of the others analysed. A compilation of these (Paulo 2001) suggested a standard rate of substitution of around 2.0% per million years for cytochrome *b* gene evolution in lacertids. The likelihood-ratio test performed on the cytochrome *b* data set showed no significant difference between the log-likelihood of phylogenetic trees with (-ln 1776) and without (-ln 1760) the molecular clock enforced. Although the ML estimate of phylogeny based on the full data set differs in topology from that obtained using only the cytochrome *b* data, we forced the latter to fit the topology of the former. Since this tree was not significantly less likely using a Shimodaira and Hasegawa (1999) test than the best estimate of phylogeny derived from cytochrome *b* (p=0.112), we calculated divergence times from cytochrome *b* distances only. Our estimates of phylogeny place *P. hispanica* sensu stricto as the sister taxon to all North African forms. The mean uncorrected pairwise distance between this form and all other lineages in this group, computed using Mega v2.1 (Kumar *et al.* 2001) is 0.1101 ± 0.0129 (minimum 0.088, maximum 0.13). Using the above-mentioned rate of substitution, this yields a divergence time around 5.50 ± 0.65 Mya, roughly coincident with the opening of the Strait of Gibraltar. Our estimates of relationships would then indicate that *Podarcis (h.) vaucheri*

crossed back across the Strait from North Africa to the Iberian Peninsula around 2.81 Mya. This scenario thus represents an alternative biogeographical hypothesis to that suggested by Harris *et al.* (2002b).

The divergence time between the western Iberian clade (comprising *P. bocagei, P. hispanica* type 1, *P. carbonelli* and *P. hispanica* type 2) and the Southern Iberian/North African clade, estimated from uncorrected pairwise distances of cytochrome *b*, is about 7.02 ± 0.65 Mya. Around 7.8 - 7.6 Mya, a salinity crisis (the Tortonian salinity crisis; Krijgsman *et al.* 2000) caused the formation of land bridges between the mainland and small islets in what is now the Betic region. These areas may have been colonized by lizards and could have acted as speciation hotspots, as has been suggested in other lacertids (Paulo 2001). Later when the two continents became connected (around 5.6 Mya), these were the lizards that colonized North Africa. This scenario of a greater similarity between Southern Spanish and North African forms than that between southern Iberia and other regions in the Peninsula has also been reported in other herpetofaunal species that exist on both sides of the Strait of Gibraltar (e.g. *Pleurodeles,* (Batista *et al.* 2004; Veith *et al.* 2004)).

Conclusions

Although previous works suggested that Iberian and North-African *Podarcis* were a species complex, the phylogenetic relationships between these forms were weakly supported. By increasing the extent of sequencing included in the analysis, we confirm the existence of cryptic forms and have now obtained a robust phylogeny for these lizards. We also suggest that, unlike previous inferences had reported, the opening of the Strait of Gibraltar and other geological events may well have been important factors in the diversification process.

Acknowledgements

This work was partially financed by FCT (Fundação para a Ciência e a Tecnologia) research projects POCTI/34547/BSE/2000 and POCTI/41912/BSE/2001. C. Pinho is supported by a PhD grant (SFRH/BD/4620/2001) and D.J. Harris by a post-doctoral grant (SFRH/BPD/5702/2001), both from FCT. The authors wish to thank M.A. Carretero, P. Sá-Sousa, D. Barbosa, V. Batista and E. Froufe for help in sampling campaigns. We would also like to thank Xana Sá Pinto for her help in the phylogenetic analysis.

References

Arévalo E, Davis SK, Sites Jr JW, (1994) Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. *Systematic Biology,* **43,** 387–418.

Arnold EN, Burton JA. (1978) *A field guide to the reptiles and amphibians of Britain and Europe*. Collins, London.

Barbadillo LJ, Lacomba JI, Pérez Mellado V, Sancho V, López-Jurado LF (1999) *La Guia de Campo de los Anfíbios y reptiles de la Península Ibérica, Baleares y Canárias*. Ed. Planeta, Barcelona

Batista V, Harris DJ, Carretero MA, (2004) Genetic variation in *Pleurodeles waltl* Michahelles, 1830 across the Strait of Gibraltar derived from mitochondrial DNA sequences. *Herpetozoa,* **16,** 166–168.

Carranza S, Arnold EN, Mateo JA, Lopez-Jurado LF (2000) Long-distance colonization and radiation in gekkonid lizards, *Tarentola* (Reptilia: Gekkonidae), revealed by mitochondrial DNA sequences. *Proceedings of the Royal Society of London - Series B, Biological Sciences,* **267,** 637–649.

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

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

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution,* **39,** 783–791.

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. Mémoire présenté pour l'obtention du diplôme de l'École Pratique des Hautes Etudes, Montpellier.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series,* **41,** 95–98.

Harris DJ (2002) Reassessment of comparative genetic distance in reptiles from the mitochondrial cytochrome *b* gene. *Herpetological Journal*, **12,** 85–86.

Harris DJ, Arnold EN (1999) Relationships and evolution of wall lizards *Podarcis* (Reptilia: Lacertidae) based on mitochondrial DNA sequences. *Copeia* **1999**, 719–754.

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,** 1257–1262.

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, 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.

Hedges SB. Bezy RL (1993) Phylogeny of xantusiid lizards: concern for data and analysis. *Molecular Phylogenetics and Evolution,* **2,** 76–87.

Hsü KJ, Montadert L, Bernoulli D, Cita MB, Erickson A, Garrison RE, Kidd RB, Mèliéres F, Müller C, Wright R (1977) History of the Mediterranean salinity crisis. *Nature*, **267,** 399–403.

Huelsenbeck JP, Bollback JP (2001) Empirical and hierarchical Bayesian estimation of ancestral states. *Systematic Biology*, **50,** 351–366.

Huelsenbeck JP, Crandall KA, (1997) Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review of Ecology and Systematics,* **28,** 437–466.

Huelsenbeck JP, Ronquist F (2001) Mr. Bayes: Bayesian inference of the phylogeny. *Bioinformatics*, **17,** 754–755.

Klemmer K (1959) Systematishe Stellung und Rassengliederung der spanishen 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 USA,* **86,** 6196–6200.

Krijgsman W, Garcés M, Agustí J, Raffi I, Taberner C, Zachariasse WJ (2000) The "Tortonian salinity crisis" of the Eastern Betics (Spain). *Earth and Planetary Science Letters,* **181,** 497–511.

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: Molecular Evolutionary Genetics Analysis software, Arizona State University, Tempe, Arizona, USA.

Mertens R, Müller L (1940) Die Amphibien und Reptilien Europas. *Abh. Senckenbergische Naturforschende Gesellschaft*, **41,** 1–62.

Oliverio M, Bologna MA, Mariottini P (2000) Molecular biogeography of the Mediterranean lizards *Podarcis* Wagler, 1830 and *Teira* Gray, 1838 (Reptilia, Lacertidae). *Journal of Biogeography*, **27,** 1403–1420.

Paulo OS (2001) *Phylogeography of reptiles of the Iberian Peninsula*. PhD Thesis. University of London.

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, Ferrand N, Harris DJ (2004) Genetic variation within the *Podarcis hispanica* species complex: new evidence from protein electrophoretic data. In: *The Biology of Lacertid Lizards: Evolutionary and Ecological Perspectives* (Pérez-Mellado V, Riera N, Perera A, eds.). Institut Menorquí d'Estudis. *Recerca*, **8,** 269–277.

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

Posada D, Crandall KA (2001) Selecting models of nucleotide substitution: and application to human immunodeficiency vírus 1 (HIV-1). *Molecular Biology and Evolution,* **18,** 897–916.

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

Sá-Sousa P (2001) *A Controversa Sistemática das Lagartixas do género* Podarcis *WAGLER, 1830 (Sauria: Lacertidae) em Portugal*. Dissertação de Doutoramento. Fac. Ciências, Univ. Lisboa.

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,** 55–69.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Press, New York.

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*, **13,** 964–969.

Swofford DL (2002) *PAUP* ver 4.0.b10. Phylogenetic analysis using parsimony and other methods.* Sinauer Associates, Sunderland, Sunderland, MA.

Veith M, Mayer C, Samroui B, Donaire Barroso D, Bogaerts S (2004) From Europe to Africa and vice-versa: evidence for multiple intercontinental dispersal in ribbed salamanders (Genus *Pleurodeles*). *Journal of Biogeography*, **31,** 159–171

Genetic polymorphism of 11 allozyme loci in populations of wall lizards (*Podarcis* sp.) from the Iberian Peninsula and North Africa

CATARINA PINHO[1,2], D. JAMES HARRIS[1], NUNO FERRAND[1,2]

[1]*Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO/UP), Campus Agrário de Vairão, 4485-661 Vairão, Portugal*
[2] *Departamento de Zoologia-Antropologia, Faculdade de Ciências da Universidade do Porto, 4099-002 Porto, Portugal*

Abstract
The taxonomy of Iberian and North African wall lizards (*Podarcis* sp.) has been controversial. Recently, morphological and mtDNA sequence data have provided new information on differentiation within these lacertids. In order to compare these results to those provided by nuclear markers, we investigated variation at 11 polymorphic protein loci using conventional electrophoresis and isoelectric focusing, in 11 populations belonging to seven different mtDNA lineages. A total of 62 alleles were found. Populations belonging to the same mtDNA type presented high genetic similarity, whereas strong differentiation was observed between groups. These results are consistent with those previously obtained from morphological and mtDNA analysis and support the idea that Iberian and North African *Podarcis* are composed by several well-differentiated entities, some of which are already recognised as species, whereas others (belonging to the *P. hispanica* complex) clearly need taxonomic revision.

Keywords: Lacertidae; *Podarcis hispanica; Podarcis bocagei; Podarcis carbonelli; Podarcis vaucheri;* electrophoresis; isoelectric focusing; genetic polymorphism.

Introduction

Prior to recent studies based on mitochondrial DNA sequence data, three species of wall lizards were recognized from the Iberian Peninsula: *Podarcis muralis* (Laurenti, 1768), *P. hispanica* (Steindachner, 1870) and *P. bocagei* (Seoane, 1884), of which only *P. hispanica* is present in North Africa. With the exception of *P. muralis*, the taxonomy of Iberian *Podarcis* has been complex and subject to several revisions, because most forms display considerable intraspecific morphological variation (Guillaume 1987, Sá-Sousa 2001). On the basis of morphological and mitochondrial DNA sequence data, Harris and Sá-Sousa (2001) suggested that the two continental subspecies of *P. bocagei* were not closely related, leading to the specific recognition of *P. carbonelli* Pérez-Mellado 1981 (Sá-Sousa and Harris 2002), and that *P. hispanica* appeared to be a species complex (Harris and Sá-Sousa 2002). With respect to North African forms, Oliverio *et al*. (2000) recommended the elevation of these to the specific level under the designation of *P. vaucheri* (Boulenger 1905) on the basis of high mitochondrial DNA divergence, whereas previous electrophoretic studies (Busack 1986) described very low genetic distances between southern Iberian and North African lizards. Recently, Harris *et al.* (2002) suggested the presence of multiple forms in North Africa, one of which is also present in the Iberian Peninsula, and pointed out that differential sampling might have accounted for different results among the previous studies.

Although so far there has been agreement between two independent lines of evidence (morphology and mitochondrial DNA), the study of differentiation within and between forms of *Podarcis* in the Iberian Peninsula and North Africa would not be complete without the analysis of nuclear markers. Some of the lineages identified by mtDNA have not been distinguished using morphological characters, and there are many cases where mtDNA divergence is not corroborated by nuclear DNA or morphology (e.g. Ballard *et al*. 2002). Thus there is a need for assessment of nuclear markers prior to any taxonomic revision. The electrophoretic study of proteins has been previously applied to the taxonomical revision or the clarification of uncertain evolutionary relationships within some lacertid species complexes (e.g. Capula 1994a,b, Mayer and Arribas 1996, MacCulloch *et al.* 2000).

The present paper describes biochemical variation at 11 polymorphic protein loci in eleven populations belonging to different forms of *Podarcis*.

Materials and Methods

A total of 239 specimens were collected from 11 populations belonging to different mitochondrial DNA lineages of *Podarcis* (Fig. 1). Sampling details are shown in Table 1. Samples consisted of a portion of tail muscle obtained from tail autotomy. All lizards were released after this procedure. Samples were stored frozen at –80ºC.

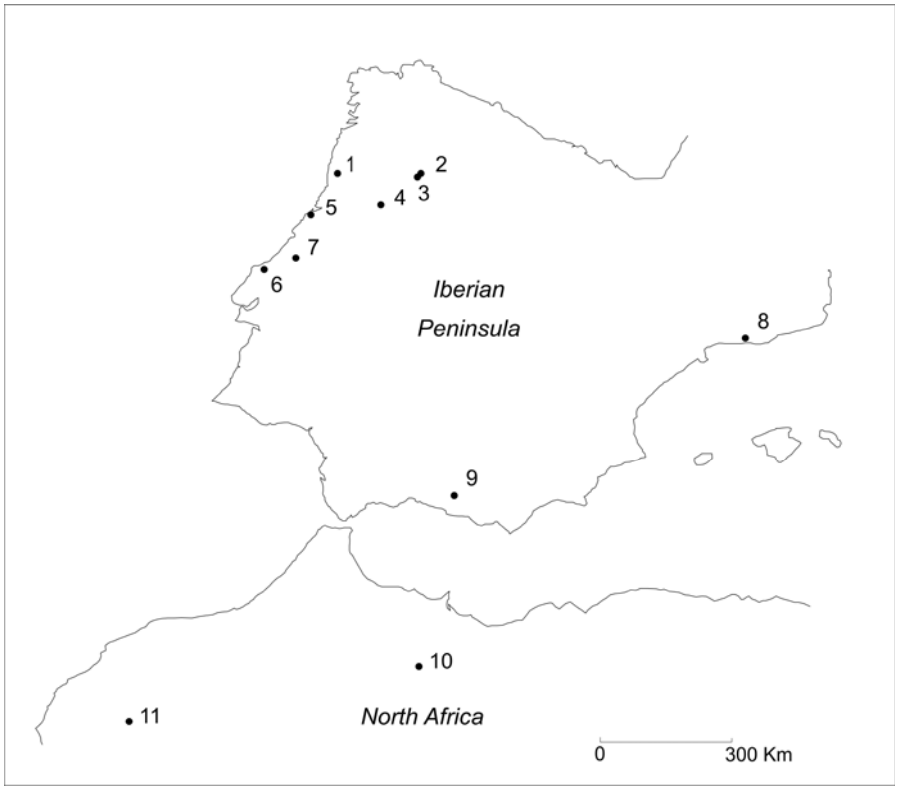

Figure 1. Location of sampling sites. 1 – Vairão; 2, 3 – Montesinho; 4 – Tua; 5 – Aveiro; 6 – S. Pedro de Moel; 7 – Monte Real; 8 – Barcelona; 9 – Sierra Nevada; 10 – Debdou; 11 – Oukaïmeden.

Tissue extracts were obtained by grinding a portion of the tail in about equal volume of homogenizing buffer (Tris-EDTA-HCl, pH 7.0). Homogenates were then submitted to a brief ultrasound treatment in order to disrupt membranes, followed by centrifugation at 14,000 rpm for 15 min at –4ºC. The supernatant was then applied to gels. When performing analysis of peptidases, samples were

submitted to a reducing treatment with dithiothreitol 120mM for 1 h at 37ºC prior to their application, after Ferrand (1995).

We studied 11 enzyme systems representing 15 structural loci, by means of conventional 15% starch gel electrophoresis and isoelectric focusing (see tables 2 and 3 for details).

Starch gel electrophoreses were carried out using three buffer systems: SGE1, Tris-citrate buffer, pH 7.6 (Amorim and Siebert 1982); SGE2, Tris-NaH_2PO_4, pH 7.6 (Branco *et al.* 1999); and SGE3, citrate-NaOH-His/HCl, pH 6.0 (Ferrand and Amorim 1990). Samples were applied in 1:1 Sephadex (PHARMACIA) medium, except in the case of peptidases, in which better results were accomplished by applying the samples in 8.5 × 6.5 mm pieces of chromatography paper (Whatman number 3). Electrophoreses were performed at 6-8 V/cm for 18-20hrs at 8ºC.

Protein separation by isoelectric focusing was carried out in polyacrylamide nondenaturing gels (T=5%, C=3%, 20% saccharose; 6% (v/v) carrier ampholytes; 0.1% (v/v) TEMED; 0.7% of a 0.44M ammonium persulfate solution; 230 × 100 × 0.3mm). Three systems were used: IEF1 (mixture of the ampholytes 5-8 Pharmalyte, 6-8 Ampholine and 7-9 Ampholine, 2:1:1), IEF2 (6-8 Ampholine and 6.7-7.7 Pharmalyte, 1.4:1) and IEF3 (5-6 Pharmalyte). Aspartic acid 0,04M and NaOH 1M were used as anode and cathode solutions, respectively. Gels were prefocused at constant power setting limits at 1500V, 25mA, and 1W (30min), 2W (15min) and 3W (15min). After prefocusing, 10μl of each diluted sample was applied 1.5 cm from the cathode in a silicone strip (Serva). Focusing was performed at constant power setting limits at 1500V, 25mA and 4W (1hr), 2000V, 25mA and 5W (1hr) and 2500V, 25mA and 6W (1hr) for systems IEF1 and IEF2, whereas for system IEF3 the conditions used were 1500V, 25mA and 4W (2.5hrs) and 2000V, 25mA and 5W (0.5hrs).

Table 2. Enzyme systems analysed by means of starch gel electrophoresis and electrophoretic conditions employed

Enzyme/ locus	E.C.	Electrophoretic system	V/Cm	Time (h)	Variation
Malate dehydrogenase (*MDH1*)	1.1.1.37	SGE3	7	18	No
Malate dehydrogenase (*MDH2*)	1.1.1.37	SGE3	7	18	No
Isocitrate dehydrogenase (*IDH1*)	1.1.1.42	SGE1	7	20	Yes
6-Phosphogluconate dehydrogenase (*PGD*)	1.1.1.44	SGE1	8	20	Yes
Purine nucleoside phosphorylase (*NP*)	2.4.2.1	SGE3	7	18	Yes
Glutamate-oxaloacetate transaminase (*GOT1*)	2.6.1.1	SGE1	8	20	Yes
Glutamate-oxaloacetate transaminase (*GOT2*)	2.6.1.1	SGE1	8	20	No
Peptidase A (*PEPA*)	3.4.11	SGE2	6	20	Yes
Peptidase D (*PEPD*)	3.4.11	SGE2	6	20	Yes
Mannose-phosphate isomerase (*MPI*)	5.3.1.8	SGE1	7	20	Yes
Glucose-phosphate isomerase (*GPI*)	5.3.1.9	SGE3	7	18	Yes

Table 3. Separation conditions of proteins examined by isoelectric focusing

Enzyme/ locus	E.C.	Electrophoretic system	Variation
Lactate dehydrogenase (*LDH1*)	1.1.1.27	IEF1	No
Lactate dehydrogenase (*LDH2*)	1.1.1.27	IEF1	Yes
Phosphoglucomutase (*PGM1*)	2.7.5.1	IEF2	Yes
Peptidase B (*PEPB*)	3.4.11	IEF3	Yes

Enzymatic detection of all loci was carried out according to the methods described by Harris and Hopkinson (1976), using leucylalanine, leucylglycylglycine and phenylalanylproline as substrates for *PEPA, PEPB* and *PEPD*, respectively. The agar-overlay method was used to detect all zymograms except *GOT*.

Allelic frequencies were calculated directly from the observed genotypes. We used the GENEPOP software (v. 3.1b; Raymond and Rousset 1995) probability test to determine whether populations were in Hardy-Weinberg equilibrium. To evaluate genetic relationships among the populations, we calculated Nei's standard genetic distance (Nei 1972) and obtained a tree based on the neighbor-joining method (Saitou and Nei 1987), using the PHYLIP package (v. 3.5; Felsenstein 1993). Bootstrap support was estimated using 1000 pseudoreplicates.

Results

From the 15 presumptive loci analysed, 11 were found polymorphic, with the number of allelic variants ranging from 2 (*IDH1*) to 10 (*PGD*). A total of 62 alleles were detected among the polymorphic loci. The zymograms of the studied proteins or their schematic representations are shown in Figs. 2 and 3. Although no progeny testing was performed to confirm the mode of inheritance of genetic variants, zymograms conformed with simple patterns of codominant inheritance and observed phenotypic distributions were generally in agreement with Hardy-Weinberg expectations. However, *PEPD* (in Vairão, p=0.0262; Monte Real, p=0.0062; and Oukaïmeden, p=0.0416) showed significant deviations from the expected distributions under the assumption of equilibrium.

PGM1 and *MPI* showed a pattern of variation that was concordant with a monomeric subunit structure, while *IDH1, PGD, GOT1, PEPA, PEPB, PEPD* and *GPI* were found to have a dimeric structure. Heterozygous individuals for *NP* presented 4 bands, thus revealing a trimeric structure, whereas *LDH2* was tetrameric. These subunit structures were consistent with those reported for other vertebrate species (e.g. Alexandrino *et al*. 1997, Harris and Hopkinson 1976).

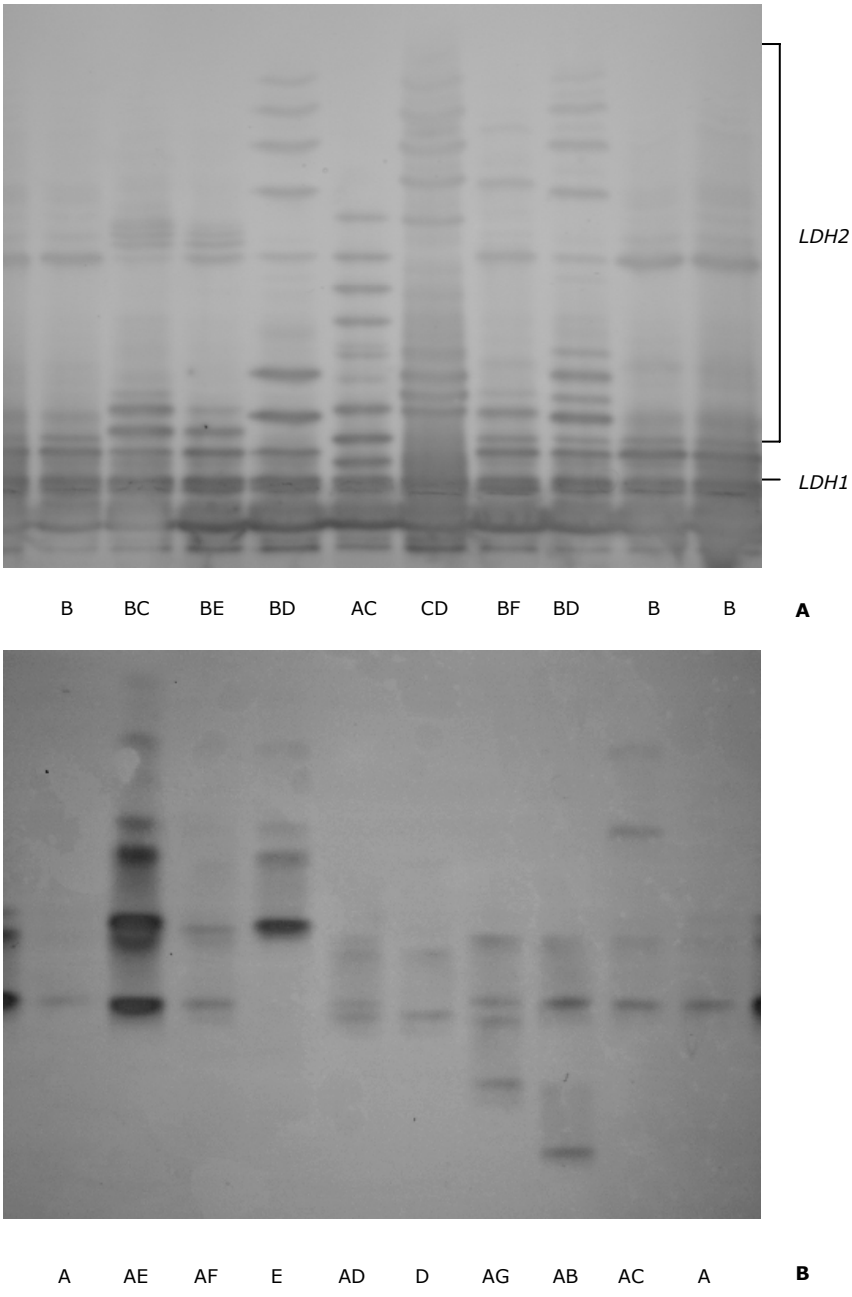

Figure 2. Electrophoretic patterns of *LDH2* (A) and *PGM1* (B) after isoelectric focusing. Alleles *LDH2**E and *LDH2**F correspond to low activity and null variants, respectively.

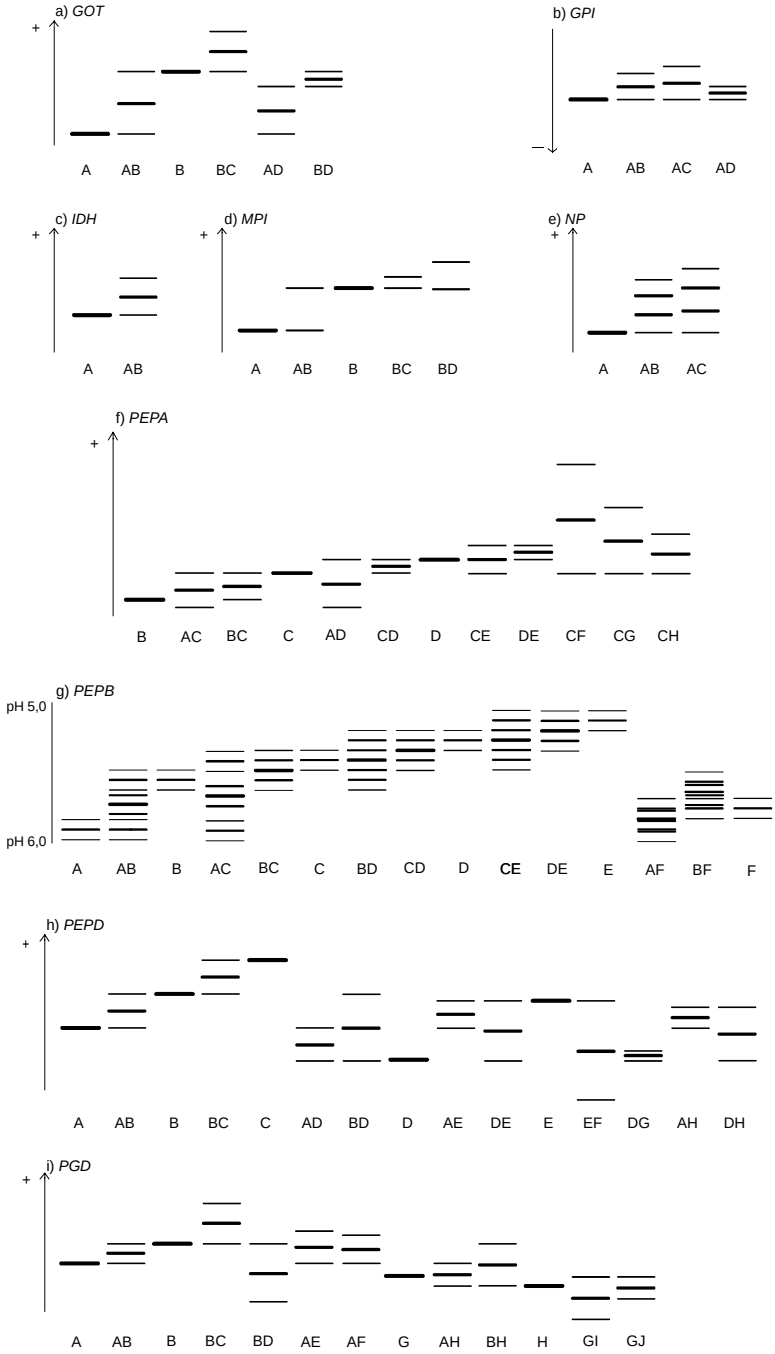

Figure 3. Schematic representation of the observed electrophoretic patterns of *GOT* (a), *GPI* (b), *IDH* (c), *MPI* (d), *NP* (e), *PEPA* (f), *PEPB* (g), *PEPD* (h), and *PGD* (i).

Allele frequencies are presented in table 4. The observed genetic distances ranged from 0.009, between Montesinho (*P. hispanica*) and Tua, to 0.398, between Montesinho (*P. hispanica*) and Oukaïmeden (see table 5). The estimate of relationships derived from a neighbor-joining analysis based on this data is shown in Fig. 4.

Discussion

The obtained results show high levels of concordance with previous analyses based on morphological and mtDNA sequence data.

In this study four species/morphotypes are represented by more than one population, and these conspecific populations cluster together with 100% bootstrap support. Although values of genetic distance obtained in different studies may not be fully comparable, they provide an idea of the level of differentiation that can be expected between closely related species. Considering the values reported in the literature (e.g. Mayer and Tiedemann 1982, Busack 1987, Capula 1994a,b, MacCulloch *et al.* 2000), the observed distances between all seven analysed morphotypes are in general equivalent to those found between recognised lacertid species. Our results thus support the separation of *P. bocagei* and *P. carbonelli* (Sá-Sousa and Harris 2002) and of the two Portuguese morphtypes of *P. hispanica* (type 1 and type 2; Sá-Sousa *et al*. 2002). Moreover, the paraphyly of *P. hispanica* (Harris and Sá-Sousa 2001, 2002) is also suggested from the analysis of nuclear markers. With respect to North African wall lizards, we found higher genetic distances across the Strait of Gibraltar than Busack (1986; Nei's D of 0.05-0.08, while in our study we obtain 0.132 between Debdou and Sierra Nevada and 0.173 between Oukaïmeden and Sierra Nevada). This discrepant result is probably due to differential sampling: while we sampled *P. hispanica* sensu stricto in southeastern Iberia, Busack (1986) probably sampled a *P. (hispanica) vaucheri* population, which is known to occur in southwestern Spain (Harris *et al.* 2002).

Both evidence from mtDNA sequence (Harris and Sá-Sousa 2001, 2002) and electrophoretic data strongly support monophyly of morphologically identified forms. However, considerable differences between trees derived from mtDNA and electrophoretic data were found, namely in what concerns relationships between forms. These differences can be due to intrinsic characteristics of both types of markers or to the low number of populations analysed in the present study. However, neither marker provides strong bootstrap support for relationships between forms. Thus, historical factors like the occurrence of a rapid radiation at this point of these lizards' evolution may also account for these short internal branches.

Table 4. Allele frequencies at 11 polymorphic loci in *Podarcis* populations from the Iberian Peninsula and North Africa.

		Population										
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
GOT1	n	32	20	19	17	17	22	20	14	18	21	26
	A	0.95 (±0.03)	0.93 (±0.04)	0.74 (±0.07)	0.59 (±0.08)					0.47 (±0.08)	1.00	1.00
	B	0.03 (±0.02)	0.07 (±0.04)	0.21 (±0.07)	0.41 (±0.08)	1.00	1.00	1.00	0.96 (±0.04)	0.53 (±0.08)		
	C								0.04 (±0.04)			
	D	0.02 (±0.02)		0.05 (±0.04)								
GPI	n	33	24	20	17	17	22	20	13	18	21	26
	A	0.98 (±0.02)	1.00	1.00	1.00	1.00	1.00	0.98 (±0.02)	1.00	1.00	0.95 (±0.03)	0.98 (±0.02)
	B	0.02 (±0.02)										
	C										0.05 (±0.03)	0.02 (±0.02)
	D							0.02 (±0.02)				
IDH1	n	33	28	19	14	14	22	20	14	18	21	18
	A	0.95 (±0.03)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	B	0.05 (±0.03)										
LDH2	n	34	28	20	17	17	22	20	13	18	21	26
	A		0.07 (±0.03)	0.35 (±0.08)	0.15 (±0.06)			0.13 (±0.05)			0.02 (±0.02)	
	B	0.31 (±0.06)	0.16 (±0.05)	0.13 (±0.05)	0.15 (±0.06)	0.38 (±0.08)	0.59 (±0.07)	0.87 (±0.05)	0.46 (±0.10)	0.94 (±0.04)	0.88 (±0.05)	1.00
	C	0.10 (±0.04)	0.34 (±0.06)	0.42 (±0.08)	0.58 (±0.08)	0.21 (±0.07)						
	D	0.59 (±0.06)	0.43 (±0.07)	0.10 (±0.05)	0.09 (±0.05)	0.41 (±0.08)	0.39 (±0.07)		0.54 (±0.10)	0.06 (±0.04)	0.10 (±0.05)	
	E				0.03 (±0.03)							
	F						0.02 (±0.02)					
MPI	n	26	30	20	16	15	22	19	14	18	21	26
	A	0.08 (±0.04)		0.80 (±0.06)	0.81 (±0.1)				0.11 (±0.06)		0.17 (±0.06)	0.02 (±0.02)
	B	0.90 (±0.04)	1.00	0.20 (±0.06)	0.19 (±0.1)	1.00	1.00	0.95 (±0.04)	0.89 (±0.06)	1.00	0.83 (±0.06)	0.98 (±0.02)
	C	0.02 (±0.02)										
	D							0.05 (±0.04)				
NP	n	20	14	14	17	16	22	20	14	18	20	25
	A	0.85 (±0.06)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98 (±0.02)
	B	0.15 (±0.06)										
	C											0.02 (±0.02)
PEPA	n	34	22	20	17	17	22	20	12	18	21	26
	A					0.09 (±0.05)	0.07 (±0.04)					
	B	0.12 (±0.04)	0.09 (±0.04)		0.03 (±0.03)							
	C	0.88 (±0.04)	0.78 (±0.06)	1.00	0.97 (±0.03)	0.29 (±0.08)	0.80 (±0.06)	0.95 (±0.03)	0.92 (±0.06)	1.00	1.00	1.00
	D		0.11 (±0.05)			0.56 (±0.09)	0.11 (±0.05)					
	E					0.06 (±0.04)	0.02 (±0.02)					
	F								0.08 (±0.06)			
	G							0.05 (±0.03)				
	H		0.02 (±0.02)									

Table 4. (cont.)

		Population										
		1	2	3	4	5	6	7	8	9	10	11
PEPB	n	21	22	15	17	15	22	19	12	17	20	19
	A				0.03 (±0.03)	0.03 (±0.03)		0.42 (±0.08)				
	B	0.02 (±0.02)	0.05 (±0.03)	0.87 (±0.06)	0.76 (±0.07)		0.09 (±0.04)	0.34 (±0.08)				
	C	0.98 (±0.02)	0.95 (±0.03)	0.13 (±0.06)	0.21 (±0.07)	0.87 (±0.06)	0.84 (±0.06)	0.03 (±0.03)	0.04 (±0.04)	0.32 (±0.08)	0.60 (±0.08)	1.00
	D					0.10 (±0.05)	0.07 (±0.04)		0.83 (±0.08)	0.59 (±0.08)	0.40 (±0.08)	
	E								0.13 (±0.07)	0.09 (±0.05)		
	F							0.21 (±0.07)				
PEPD	n	30	21	20	16	17	22	20	14	17	20	25
	A	0.38 (±0.06)	0.60 (±0.08)	1.00	0.94 (±0.04)	0.47 (±0.09)	0.25 (±0.07)	0.57 (±0.08)	0.96 (±0.04)	0.47 (±0.09)	0.67 (±0.07)	0.46 (±0.07)
	B	0.52 (±0.06)	0.40 (±0.08)		0.06 (±0.04)			0.30 (±0.07)				
	C	0.10 (±0.04)										
	D					0.47 (±0.09)	0.55 (±0.08)	0.13 (±0.05)		0.21 (±0.07)	0.33 (±0.07)	0.54 (±0.07)
	E						0.07 (±0.04)			0.29 (±0.08)		
	F								0.04 (±0.04)	0.03 (±0.03)		
	G					0.06 (±0.04)						
	H						0.13 (±0.05)					
PGD	n	34	29	17	17	17	22	20	13	17	21	26
	A			1.00	0.94 (±0.04)	0.85 (±0.06)	0.88 (±0.05)		0.15 (±0.07)	0.15 (±0.06)		
	B	0.97 (±0.02)	1.00			0.15 (±0.06)	0.07 (±0.04)	0.97 (±0.03)	0.77 (±0.08)	0.17 (±0.07)		
	C	0.03 (±0.02)										
	D							0.03 (±0.03)				
	E						0.05 (±0.03)		0.08 (±0.05)			
	F				0.06 (±0.04)							
	G										0.91 (±0.04)	1.00
	H									0.68 (±0.08)		
	I										0.02 (±0.02)	
	J										0.07 (±0.04)	
PGM1	n	32	26	17	17	17	22	20	14	18	21	26
	A	0.90 (±0.03)	1.00	1.00	1.00	0.88 (±0.06)	0.98 (±0.02)	0.87 (±0.05)	1.00	0.97 (±0.03)	0.98 (±0.02)	1.00
	B	0.08 (±0.03)										
	C	0.02 (±0.02)										
	D					0.12 (±0.06)						
	E							0.13 (±0.05)				
	F									0.03 (±0.03)	0.02 (±0.02)	
	G						0.02 (±0.02)					

NOTE: For population names see Table 1.

Table 5. Values of genetic distance (Nei 1972) among populations of *Podarcis* from Iberia and North Africa.

Population[a]	1	2	3	4	5	6	7	8	9	10	11
1	-										
2	0.015	-									
3	0.376	0.334	-								
4	0.366	0.321	0.009	-							
5	0.309	0.265	0.365	0.310	-						
6	0.279	0.270	0.327	0.281	0.038	-					
7	0.252	0.234	0.391	0.349	0.253	0.201	-				
8	0.260	0.234	0.347	0.310	0.212	0.192	0.110	-			
9	0.230	0.223	0.339	0.317	0.223	0.156	0.159	0.131	-		
10	0.196	0.195	0.308	0.311	0.337	0.271	0.307	0.275	0.132	-	
11	0.194	0.199	0.398	0.394	0.319	0.247	0.337	0.386	0.173	0.024	-

[a] Populations: 1 – Vairão; 2 – Montesinho (*P.bocagei*); 3 – Montesinho (*P. hispanica*); 4 – Tua; 5 – Aveiro; 6 – S. Pedro de Moel; 7 – Monte Real; 8 – Barcelona; 9 - Sierra Nevada; 10 – Debdou; 11 – Oukaïmeden. See figure 1 for locations.

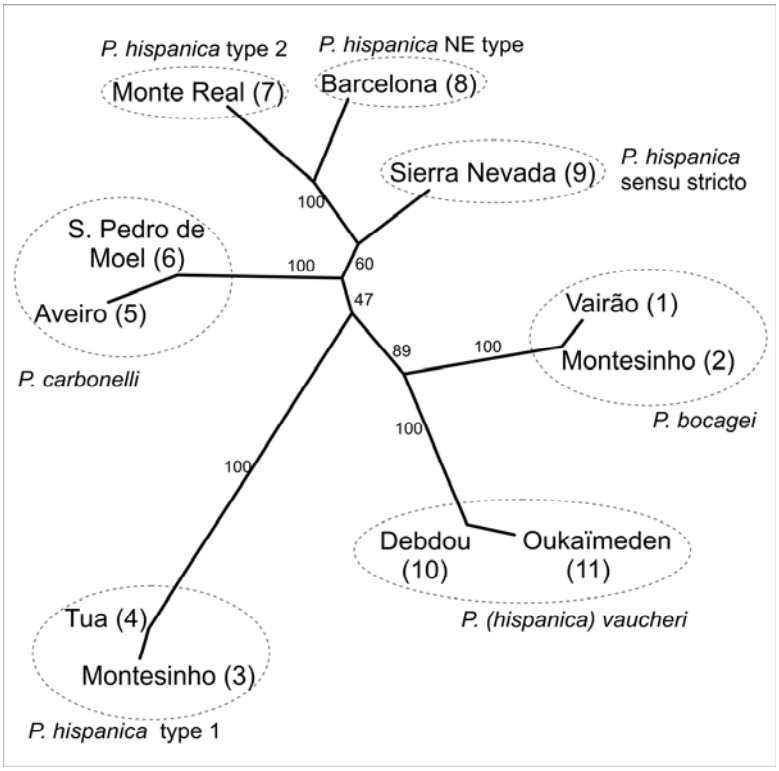

Figure 4. NJ tree based on Nei (1972) genetic distances for 11 populations of *Podarcis*. Percent bootstrap values are shown

Some authors suggest that allozymes are not useful markers for population genetics, mainly because of low polymorphism levels, which decrease the ability to detect population structure and differentiation (e.g. Piel and Nutt 2000). Barbadilla *et al.* (1996) showed that the resolution power of electrophoresis decreases with the number of polymorphic amino acid positions and that therefore underestimation of genetic variability is common, especially in loci with higher levels of polymorphism. In this study, we tried to minimize such underestimations by applying highly resolving techniques, such as isoelectric focusing, which can detect cryptic variations in protein net charge, and also by performing long electrophoreses (18-20h) in proteins in which isoelectric focusing did not provide satisfactory results. Compared to other investigations involving Iberian *Podarcis* (Busack 1986, Sá-Sousa *et al.* 2000), we detected higher levels of variation, not only because we included populations from several different morphotypes but also because of the use of more powerful techniques. Indeed, we found a higher average number of alleles per polymorphic loci (5.64) than many reports using microsatellites. Therefore, this study shows that the analysis of allozymes can still be an effective tool to evaluate genetic differentiation and evolutionary history as long as proper screening methods are applied.

In conclusion, this investigation provides a new set of evidence concerning differentiation within Iberian and North African *Podarcis.* This preliminary study shows that the several entities within *Podarcis hispanica*, recognizable from morphological and mitochondrial DNA evidences, can also be strongly differentiated using nuclear markers. The congruence between these three different lines of evidence clearly suggests that the taxonomy of Iberian and North African Wall lizards needs revision.

Acknowledgements

This work was partially financed by FCT (Fundação para a Ciência e a Tecnologia) research projects POCTI/34547/BSE/2000 and POCTI/41912/BSE/2001. C. Pinho is supported by a Ph. D grant (SFRH/BD/4620/2001) and D.J. Harris by a post-doctoral grant (SFRH/BPD/5702/2001), both from FCT. The authors wish to thank D. Barbosa, M.A. Carretero, E. Froufe, S.B. Ribeiro and M. Znari for help in sampling campaigns.

References

Alexandrino J, Ferrand N, Arntzen JW (1997) Genetic variation in some populations of the Golden-striped Salamander, *Chioglossa lusitanica* (Amphibia: Urodela), in Portugal. *Biochemical Genetics* **35,** 371-381.

Amorim A, Siebert G (1982) Glutamate pyruvate transaminase, esterase D, glyoxalase I, and phosphoglucomutase 1 polymorphisms in Porto district (Portugal). *Human Heredity* **32,** 293–300.

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,** 527–545.

Barbadilla A, King LM, Lewontin RC (1996) What does electrophoretic variation tell us about protein variation? *Molecular Biology and Evolution,* **13,** 427–432.

Branco M, Machado JC, Ferrand N (1999) Extensive genetic polymorphism of peptidases A, B, C, and D, in wild rabbit (*Oryctolagus cuniculus*) populations from the Iberian Peninsula. *Biochemical Genetics,* **37,** 237–249.

Busack SD (1986) Biogeographic analysis of the herpetofauna separated by the formation of the Strait of Gibraltar. *National Geographic Research,* **2,** 17–36.

Busack SD (1987) Morphological and biochemical differentiation in Spanish and Moroccan populations of the lizard, *Lacerta lepida*. *Journal of Herpetology,* **21,** 277–284.

Capula M (1994a) Genetic variation and differentiation in the lizards, *Podarcis wagleriana* (Reptilia: Lacertidae). *Biological Journal of the Linnean Society* **52,** 177–196.

Capula M (1994b) Evolutionary relationships of *Podarcis* lizards from Sicily and the Maltese Islands. *Journal of Zoological Systematics and Evolutionary Research* **32,** 180–192.

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) Version 3.5. Distributed by the author. University of Washington, Seattle.

Ferrand N (1995). *Variação genética de proteínas em populações de coelho (*Oryctolagus cuniculus*). Análise da diferenciação subespecífica, subestruturação, expansão geográfica e domesticação.* Doctoral thesis, Universidade do Porto.

Ferrand N, Amorim A (1990). Genetic polymorphism of δ-aminolaevulinic acid dehydratase (E.C. 4.2.1.24, ALAD) in the domestic rabbit. *Animal Genetics* **21,** 217–219.

Guillaume CP (1987) *Les petits lacértides du bassin méditerranéen occidental (Genera Podarcis et Archaeolacerta essentiellement)*. Doctoral thesis, Univ. Sciences et Techniques du Languedoc.

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

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 H, Hopkinson DA (1976) *Handbook of enzyme electrophoresis in human genetics*. North-Holland, Amsterdam.

MacCulloch RD, Fu J, Darevsky IS, Murphy R (2000) Genetic evidence for species status of some Caucasian rock lizards in the *Darevskia saxicola* group. *Amphibia-Reptilia,* **21,** 169–176.

Mayer W, Tiedemann F (1982) Chemotaxonomical investigations in the collective genus *Lacerta* (Lacertidae, Sauria) by means of protein electrophoresis. *Amphibia-Reptilia* **2,** 349–355.

Mayer W, Arribas OJ (1996) Allozyme differentiation and relationship among the Iberian-Pyrenean Mountain Lizards (Squamata: Sauria: Lacertidae). *Herpetozoa,* **9,** 57–61.

Nei M (1972) Genetic distance between populations. *American Naturalist,* **106,** 283–292.

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.

Piel WH, Nutt KJ (2000) One species or several? Discordant patterns of geographic variation between allozymes and mtDNA sequences among spiders in the genus *Metepeira* (Araneae: Areneidae). *Molecular Phylogenetics and Evolution,* **15,** 414–418.

Raymond M, Rousset F (1995) GENEPOP (version 3.1): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86,** 248–249.

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 (2001) *A Controversa Sistemática das Lagartixas do género* Podarcis *Wagler, 1830 (Sauria: Lacertidae) em Portugal*. Doctoral thesis, Faculdade de Ciências, Univ. Lisboa.

Sá-Sousa P, Almeida AP, Rosa HD, Vicente L, Crespo EG (2000) Genetic and morphological relationships of the Berlenga wall lizard (*Podarcis bocagei berlengensis*: Lacertidae). *Journal of Zoological Systematics and Evolutionary Research* **38,** 95–102.

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.

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, 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

Abstract

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, 10 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 that corroborate the major splits observed in mtDNA analyses. These groups correspond not only to the fully recognized species *P. bocagei, P. carbonelli* and *P. vaucheri* but also to multiple forms within the polytypic *P. hispanica*, all of which have a similar level of differentiation to that observed between the acknowledged species. Relationships between forms are, however, 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.

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 (e.g. 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 a 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 and Turelli 2003) and by the acknowledgement of the limitations of the use of a single locus to infer evolutionary relationships (e.g. Pamilo and Nei 1988, Zhang and Hewitt 2003; for a review, see Ballard and 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 (e.g. Hey 1997, Ballard, Chernoff and James 2002). Fluctuating effective population sizes (Fay and Wu 1999; Monsen and Blouin 2003), gender-biased gene flow (e.g. FitzSimmons *et al.* 1997, Nyakaana and Arctander 1999, Piertney *et al.* 2000) and introgression (e.g. DeSalle and Giddings 1986, Shaw 2002; Alves *et al.* 2003; Chan and Levin 2005) are known to affect mitochondrial and nuclear markers differently, which implies 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 (e.g. Machado and 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 (e.g. Mertens and Müller 1940, Klemmer 1959), until recently a conservative view considering only the existence of only two endemic species, *P. bocagei* (Seoane 1884) and *P. hispanica* (Steindachner 1870), prevailed (Arnold and Burton 1978; Barbadillo *et al.* 1999). However, recent morphological and mitochondrial DNA studies (Geniez 2001; Harris and Sá-Sousa 2001, 2002, Harris *et al.* 2002b, Pinho, Ferrand, and Harris 2006; Sá-Sousa, Vicente and Crespo 2002) 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 *P. carbonelli* Pérez-Mellado, 1981, a former subspecies of *P. bocagei* (Sá-Sousa and Harris 2002) and of the south

Iberian/west Mahgrebin form *P. vaucheri* (Boulenger, 1905), formerly included in *P. hispanica* (Oliverio, Bologna and Mariottini 2000; Busack, Lawson and 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; see figure 1 for a tentative map of the distribution of mtDNA lineages).

This scenario of the existence of multiple differentiated lineages within a relatively small geographic area follows a pattern that is now being acknowledged as common: southern European peninsulas probably did not only provide 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 and Lunt 2007, and references therein). This is true for many species complexes, including other *Podarcis* groups that occur in other peninsulas (Poulakakis *et al.* 2003, 2005a, 2005b; Podnar, Mayer and 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 and Ferrand 2003; Pinho, Ferrand and Harris 2004; Busack *et al.* 2005) seem 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 like 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 (e.g. Banks and Eichert 2000; Pritchard, Stephens and Donnelly 2000; Dawson and Belkhir 2001; Anderson and Thompson 2002; Corander, Waldmann and Sillanpaa 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 (that is, 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 this work 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 relationships using nuclear markers the same as those predicted by mtDNA; and (3) can hybridization be detected between the described forms?

Materials 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 (Figure 1). MtDNA lineage identification was confirmed by using populations where at least one individual had already been sequenced (Harris and Sá-Sousa, 2001, 2002; Harris *et al.* 2002a, Harris *et al.* 2002b, Pinho *et al.* 2006, Pinho *et al.* unpublished data). As outgroup, two populations of *P. muralis,* a species that also exists in the Iberian Peninsula but has a different evolutionary origin (Harris and 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 -80ºC 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 since it stopped providing consistently interpretable results, as previously explained in Pinho *et al.* (2004). Therefore, variation at 10 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 and 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 and 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 and 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 in order to compare the magnitude of differentiation between the four recognised 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

Neighbor-Joining (NJ, Saitou and Nei 1987) tree based on Nei's standard genetic distance (Nei 1972), using PHYLIP, version 3.5 (Felsenstein 1993). Bootstrap support was estimated using 1000 pseudoreplicates.

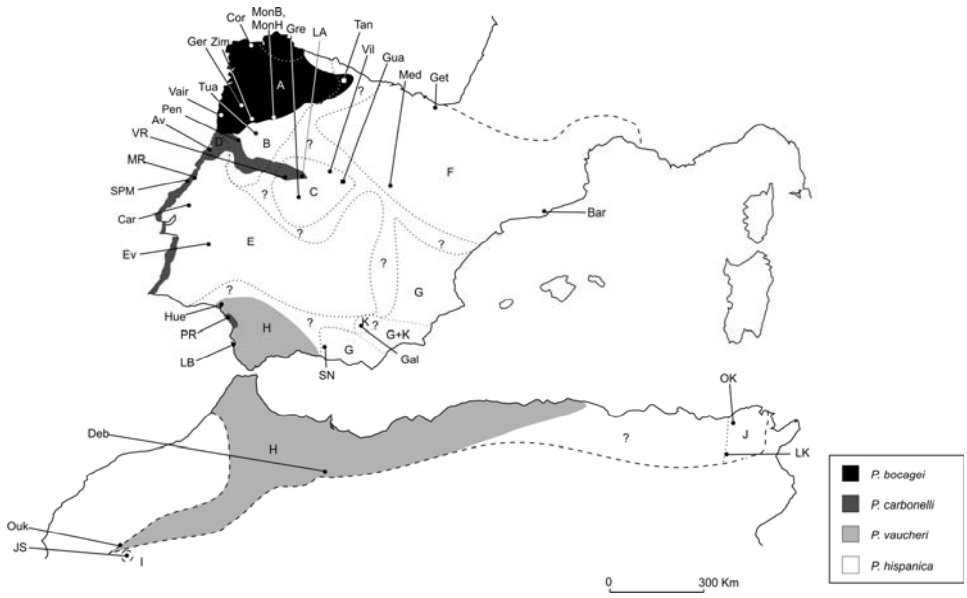

Figure 1. Map showing the predicted distribution of the mitochondrial DNA lineages and sampling sites for this study. Dashed lines represent the limits of the distribution of the Iberian and North African species complex. Dotted lines represent putative limits of the distribution areas of forms of *P. 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. Letters from A to K indicate different forms as defined by mtDNA studies (Pinho *et al.* 2006): A – *Podarcis bocagei*, B – *P. hispanica* type 1A, C – *P. hispanica* type 1B, D – *P. carbonelli*, E – *P. hispanica* type 2, F – *P. hispanica* type 3, G – *P. hispanica* sensu stricto, H – *P. vaucheri*, I – *P. hispanica* Jebel Sirwah type, J – *P. hispanica* Tunisian type, K – *P. hispanica* Galera type. Samples from *P. muralis* are represented by squares.

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, 10 independent replicates of the MCMC were conducted. The software DISTRUCT (Rosenberg 2002) was used to visually represent the obtained data. We used

the method proposed by Evanno, Regnaut and Goudet (2005) to choose amongst the values of K the one that best characterised 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 and Thompson 2002), that also implements a Bayesian clustering algorithm based on individual genotypes, to assign the admixed individuals to a specific hybrid class (pure of either type, F1 or F2 hybrids and backcrosses in both directions).

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

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

Results

Population-based genetic analysis

Including previously published results, a total of 89 alleles were detected among the 10 studied loci. The most polymorphic locus was *PGD*, with 18 alleles detected, whereas *IDH*, with 4 alleles, was the least polymorphic. In addition to the cases of Hardy-Weinberg disequilibrium already reported in previous studies (locus *PEPD* in Vairão (1), Monte Real (16) and Oukaïmeden (26)), significant departures from equilibrium were detected for *GOT* in La Alberca (10) and for *PEPB* in Gredos (11) ($p<0.05$). However, when considering a significancy 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 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 ϕ_{CT} of 0.45793, a ϕ_{SC} of 0.16703 and a ϕ_{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, *Podarcis hispanica* type 3, since 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.

Table 2. Pairwise F_{ST} values between populations of *Podarcis* from the Iberian Peninsula and North Africa based on ten polymorphic protein loci.

		A				B				C			D				E			F			G	H				I	J		K	L	
		Vair	Cor	MonB	Zim	MonH	Ger	Tua	Pen	Vil	LA	Gre	Av	SPM	PR	VR	MR	Car	Ev	Bar	Med	Get	SN	Hue	LB	Deb	Ouk	JS	OK	LK	Gal	Tan	Gua
A	Vair	0.000																															
	Cor	0.132	0.000																														
	MonB	0.055	0.262	0.000																													
	Zim	0.035	0.122	0.142	0.000																												
B	MonH	0.567	0.672	0.590	0.590	0.000																											
	Ger	0.512	0.633	0.514	0.545	0.122	0.000																										
	Tua	0.539	0.639	0.564	0.563	0.001*	0.061	0.000																									
	Pen	0.497	0.599	0.523	0.525	0.073	0.051	0.039	0.000																								
C	Vil	0.584	0.722	0.627	0.664	0.688	0.623	0.627	0.550	0.000																							
	LA	0.487	0.606	0.530	0.576	0.628	0.588	0.590	0.548	0.366	0.000																						
	Gre	0.553	0.689	0.604	0.644	0.692	0.646	0.647	0.582	0.052	0.274	0.000																					
D	Av	0.518	0.612	0.509	0.553	0.534	0.390	0.472	0.380	0.472	0.539	0.492	0.000																				
	SPM	0.513	0.619	0.529	0.557	0.552	0.414	0.495	0.382	0.470	0.568	0.506	0.103	0.000																			
	PR	0.522	0.654	0.550	0.575	0.553	0.415	0.492	0.372	0.492	0.579	0.542	0.173	0.056	0.000																		
	VR	0.537	0.647	0.558	0.583	0.596	0.474	0.542	0.440	0.499	0.595	0.522	0.115	0.014*	0.056	0.000																	
E	MR	0.459	0.636	0.503	0.547	0.603	0.521	0.555	0.440	0.490	0.570	0.508	0.451	0.438	0.425	0.478	0.000																
	Car	0.535	0.740	0.589	0.634	0.664	0.568	0.594	0.472	0.543	0.594	0.536	0.459	0.418	0.427	0.453	0.192	0.000															
	Ev	0.325	0.518	0.351	0.422	0.524	0.428	0.480	0.377	0.432	0.441	0.427	0.382	0.389	0.376	0.427	0.081	0.168	0.000														
F	Bar	0.477	0.643	0.509	0.527	0.581	0.480	0.531	0.424	0.574	0.608	0.598	0.403	0.415	0.341	0.446	0.298	0.477	0.257	0.000													
	Med	0.421	0.609	0.470	0.496	0.597	0.490	0.541	0.440	0.480	0.559	0.520	0.351	0.351	0.273	0.370	0.211	0.333	0.188	0.033*	0.000												
	Get	0.670	0.781	0.696	0.727	0.673	0.598	0.635	0.535	0.662	0.670	0.667	0.471	0.456	0.423	0.489	0.552	0.601	0.494	0.503	0.482	0.000											
G	SN	0.437	0.594	0.472	0.523	0.544	0.452	0.507	0.409	0.487	0.499	0.523	0.396	0.349	0.277	0.383	0.361	0.391	0.288	0.329	0.260	0.407	0.000										
H	Hue	0.460	0.605	0.449	0.537	0.576	0.503	0.554	0.478	0.587	0.496	0.595	0.496	0.495	0.494	0.536	0.531	0.567	0.412	0.558	0.518	0.622	0.377	0.000									
	LB	0.514	0.703	0.509	0.609	0.656	0.594	0.634	0.551	0.717	0.589	0.705	0.601	0.595	0.620	0.643	0.629	0.727	0.500	0.663	0.644	0.724	0.468	0.098	0.000								
	Deb	0.445	0.601	0.476	0.528	0.566	0.520	0.550	0.486	0.615	0.525	0.622	0.541	0.524	0.493	0.552	0.560	0.601	0.456	0.551	0.517	0.635	0.349	0.143	0.192	0.000							
	Ouk	0.579	0.769	0.602	0.682	0.745	0.705	0.723	0.650	0.767	0.662	0.732	0.660	0.625	0.665	0.658	0.701	0.774	0.597	0.757	0.728	0.780	0.541	0.227	0.352	0.169	0.000						
I	JS	0.621	0.765	0.668	0.664	0.695	0.611	0.633	0.546	0.738	0.709	0.737	0.540	0.534	0.566	0.579	0.580	0.627	0.499	0.592	0.546	0.663	0.564	0.681	0.809	0.706	0.859	0.000					
J	OK	0.472	0.593	0.510	0.533	0.565	0.462	0.509	0.415	0.469	0.505	0.520	0.374	0.322	0.330	0.380	0.367	0.322	0.307	0.412	0.336	0.430	0.293	0.453	0.558	0.489	0.618	0.273	0.000				
	LK	0.623	0.777	0.653	0.699	0.742	0.683	0.703	0.628	0.727	0.659	0.720	0.608	0.578	0.623	0.626	0.656	0.724	0.566	0.704	0.660	0.682	0.527	0.596	0.732	0.637	0.785	0.716	0.283	0.000			
K	Gal	0.541	0.694	0.552	0.614	0.612	0.529	0.576	0.483	0.619	0.570	0.636	0.514	0.504	0.498	0.549	0.520	0.571	0.418	0.553	0.516	0.640	0.397	0.200	0.346	0.258	0.430	0.704	0.442	0.637	0.000		
L	Tan	0.648	0.794	0.664	0.684	0.778	0.731	0.751	0.686	0.847	0.784	0.822	0.667	0.662	0.718	0.700	0.757	0.862	0.680	0.749	0.749	0.819	0.686	0.703	0.816	0.732	0.868	0.849	0.645	0.841	0.756	0.000	
	Gua	0.669	0.823	0.701	0.712	0.745	0.698	0.715	0.645	0.851	0.786	0.827	0.670	0.682	0.687	0.709	0.719	0.845	0.637	0.643	0.668	0.768	0.631	0.706	0.830	0.710	0.888	0.839	0.648	0.866	0.742	0.702	0.000

Sample codes and mitochondrial DNA lineage identification letters correspond to those in table 1.
*nonsignificant (P>0.05)

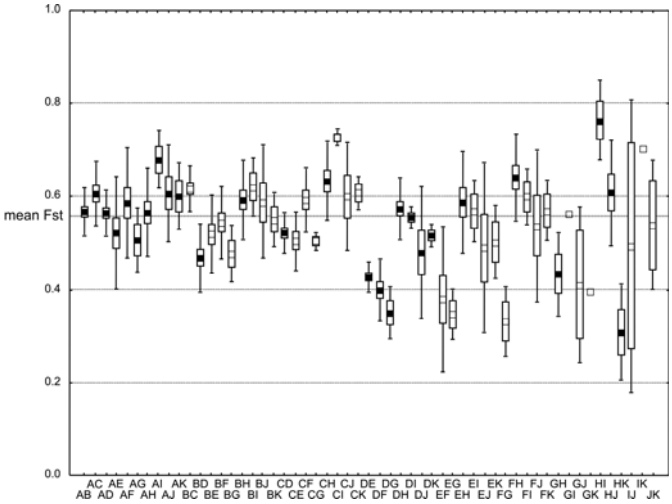

Figure 2. Box plot representation of mean F_{ST} values observed between pairs of forms of *Podarcis*, including comparisons between recognized species (filled squares) and groups within *P. hispanica* (white squares). Pairs of letters indicated on the x axis correspond to the forms being compared, using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

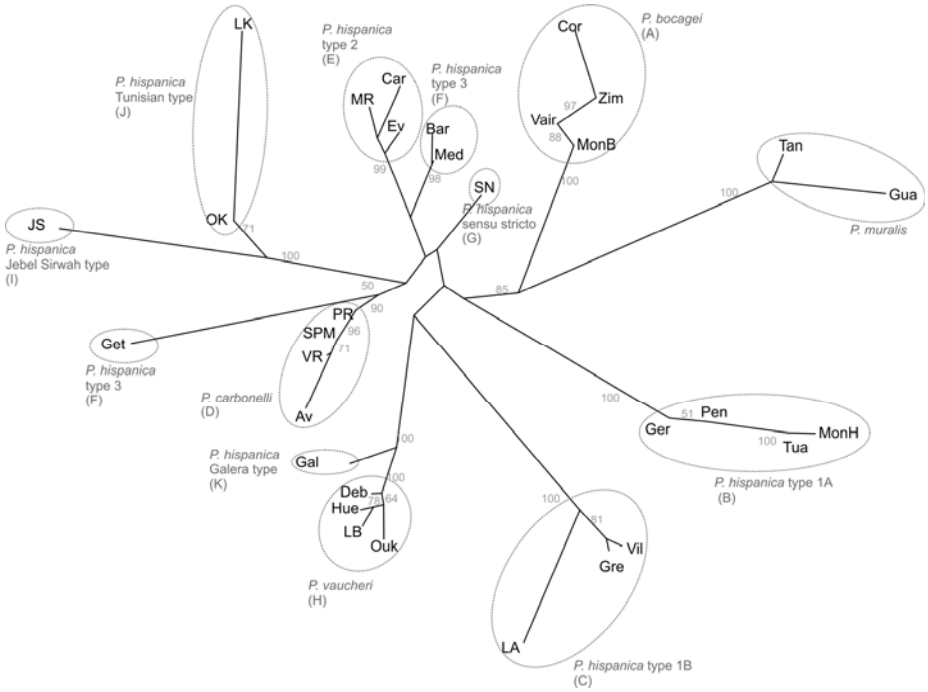

Figure 3. NJ tree showing the relationships between 32 populations of Iberian and North African *Podarcis* using the genetic distances of Nei (1972) based on 10 allozyme loci. Letters identifying mitochondrial DNA lineages correspond to those indicated in Figure 1 and Table 1. Bootstrap values above 50% are shown.

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 ΔK, a quantity related to the second order rate of change of the log probability of the data. This procedure, however, was inconclusive since it did not provide any obvious mode (results not shown). Although the highest value of ΔK 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). In this figure, 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 4 individuals (3.8%) of *P. bocagei*, all but 5 individuals (6.9%) of *P. carbonelli*, and all but 3 individuals (6.3%) of *P. hispanica* type 2. The remaining three clusters include individuals belonging to two distinct mtDNA lineages: all but 3 individuals (7.0%) of *P. hispanica* type 3 with *P. hispanica* sensu stricto, *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 around 50% probability of belonging to each) are found from K=8 on, and become increasingly frequent at higher values of K.

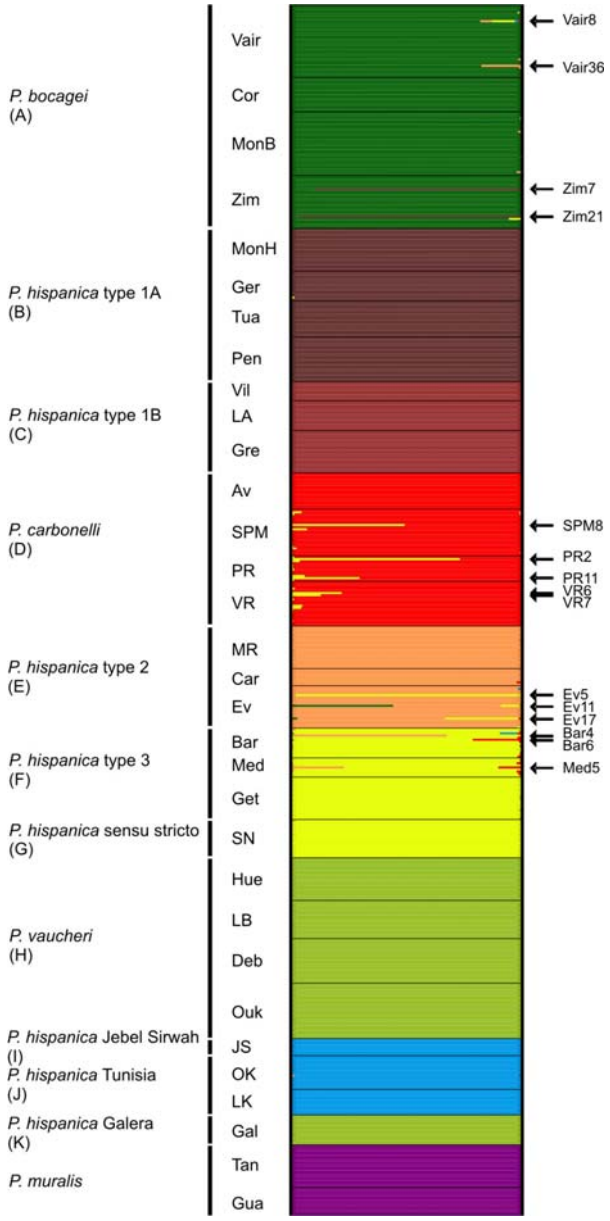

Figure 4. Estimated probability of ancestry of 569 individuals belonging to the Iberian and North African *Podarcis* species complex and to the outgroup *P. muralis*, calculated using the software STRUCTURE, considering K=9 clusters. Each individual is represented by a horizontal line divided into 9 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 in Figure 1 and Table 1.

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 unambiguously assigned to any class; instead, their posterior probabilities are distributed across the four hybrid classes with a maximum of 0.38 for F2 (Zim7) and of 0.31 for F1 (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), in all but one of the remnant forms monophyly was observed, with near 100% bootstrap support for the grouping of populations into mtDNA-defined species. This differentiation is further supported by the analyses of molecular variance (AMOVA), which suggest 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 since six out of the twelve species/forms included are clearly identifiable and three other groups, each comprising two forms, are also observed.

On the other hand, the polyphyly of *Podarcis 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. In the individual multilocus analyses, however, 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 and 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 (figure 3), thus corroborating the idea that a taxonomical revision is needed. At a larger scale, this constitutes a validation, based on nuclear markers, of the biogeographic theory that postulates that the Iberian Peninsula functioned as a hotspot of diversification and not only as a glacial refugium for many taxa (Gómez and 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* sensu stricto, 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 since 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 multilocus 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 (see Pinho *et al.* 2003 for a frequency table). Such cases have been well documented in closely related *Drosophila* species by Machado *et al.* (2002) and Machado and Hey (2003), which obtained several different genealogies using 16 independent loci. The authors call attention for the fact that simple bifurcating trees may thus not be suitable to describe 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 seems so far to be corroborated by a preliminary study of nuclear genealogies in these forms (Pinho, Harris and Ferrand, unpublished data). Diversification events involving the formation of as many distinct entities within Iberia do not seem 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 and Ochando 2000, Machordom and Doadrio 2001) or some groups of Iberian diving beetles (Ribera and 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 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; see for example Barbadilla, King and 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 (Pinho, Harris and Ferrand, unpublished). 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.

This is not the case, however, of *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 as explainable by the hypothesis of an unsatisfactory separation of alleles or of ancestral polymorphism as in the above-mentioned cases, since 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 *P. wagleriana* and *P. sicula* (Capula 1993) and between the latter and *P. 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 this study we were not able 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, directed to these

specific questions. Nevertheless, since *P. bocagei* and *P. hispanica* type 1A exist in sympatry across most of their distribution area and seem to maintain genetic integrity, as well as morphological identifiability, in the presence of the other species, it seems 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 if 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 seems to be a rare event. Nevertheless, the facts that 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, means 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 and 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 times higher than the boundaries traditionally accepted by phylogeneticists to define species in squamates (see for example 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 post-doctoral 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 Geraldes, 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 a 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,** 527–545.

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 Guia de Campo de los Anfíbios y reptiles de la Península Ibérica, Baleares y Canárias*. 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, Sillanpaa 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, USA,* **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. Distributed by the author. University of Washington, Seattle

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*. *Boletin de la Asociacion Herpetologica 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. Mémoire présenté pour l'obtention du 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: *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origin and Conservation of European Biodiversity* (Weiss S, Ferrand N, eds.). Springer, Dordrecht, The Netherlands, 155–188.

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.

Harris DJ, Arnold EN (1999) Relationships and evolution of wall lizards *Podarcis* (Reptilia: Lacertidae) based on mitochondrial DNA sequences. *Copeia,* **3,** 719–754.

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,** 1257-1262.

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.

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 B Biology 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) Systematishe Stellung und Rassengliederung der spanishen 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 (USA),* **86,** 6196–6200.

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.

Machado CA, Hey J (2003) The causes of phylogenetic conflict in a classic *Drosophila* species group. *Proceedings of the Royal Society of London B,* **270,** 1193–1202.

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 Univ. Press.

Mertens R, Müller L (1940) Die Amphibien und Reptilien Europas. *Abh. Senckenbergische Naturforschende Gesellschaft*, **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.* 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 Península, 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,** 248–249.

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. http://www.cmb.usc.edu/~noahr/distruct.html.

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.

Saitou N, Nei M (1987) The Neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution,* **4,** 406–427.

Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2.000: A software for population genetics analyses. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

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, USA,* **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.

Abundant ancestral polymorphism at the nuclear loci *β-fibint7* and *6-Pgdint7* characterizes different *Podarcis* species from the Iberian Peninsula and North Africa

CATARINA PINHO, D. JAMES HARRIS, NUNO FERRAND

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

Abstract

In the study of recently diverged species, analyses of mitochondrial DNA variation have shown to be useful in identifying major lineages and establishing evolutionary hypotheses. However, it has been widely demonstrated that such single locus perspectives may be erroneous or provide an over-simplified realization of the processes underlying species divergence. Iberian and North African wall lizards (*Podarcis*) constitute a cryptic species complex for which previous assessments of mitochondrial DNA (mtDNA) and allozyme variation have been published. Albeit concordant in describing the existence of several highly differentiated evolutionary units, some of which are currently accepted as different species, previous studies on mitochondrial and nuclear gene variation showed some important differences that suggested that gene flow may have occurred among these forms. To further understand the level of differentiation, evolutionary relationships and overall evolutionary dynamics among *Podarcis* species, we studied sequence variation at two nuclear introns, *β-fibint7* and *6-Pgdint7*, from individuals representing all known mtDNA lineages. Both nuclear gene genealogies reveal a striking absence of monophyly of mtDNA-defined species and a general lack of branching patterns presenting signs of concordance with those suggested by mtDNA. Because a mere lack of resolution due to the absence of variation cannot be invoked to explain discordant evolutionary relationships portrayed by different markers, only two non-mutually exclusive scenarios could account for the species polyphyly: persistence of ancestral polymorphism and high levels of gene flow among forms. To discriminate between the effects of these factors, we estimated migration rates among mtDNA-defined lineages using both classic estimators of gene flow, which do not take into account the possibility of ancestral polymorphism sharing, and a coalescent-based multilocus approach that simulates a model of divergence with gene flow. Both classes of estimates suggested very different levels of gene flow: overall non-zero introgression in the case of classic

estimators and virtually zero gene flow between most (but not all) species pairs in the case of coalescent estimates. This suggests that although gene flow among forms may have occurred, the main cause for species polyphyly is incomplete lineage sorting of ancestral polymorphism, thus suggesting that most forms have not exchanged genes since their divergence. This is therefore in accordance with previous reports of strong isolation based on mtDNA and allozyme data. Nevertheless, a few cases of confirmed gene flow between species were detected, of which only two, previously reported in other studies, indicate important amounts of admixture between species. Taken together, these results constitute further evidence validating most forms of Iberian and North African *Podarcis* as differentiated, although incipient, species.

Keywords: *Podarcis*, ancestral polymorphism, gene flow, isolation with migration, Iberian Peninsula, North Africa, Lacertidae, *β-fibint7, 6-Pgdint7.*

Introduction

The study of emerging species poses several challenges for evolutionary biologists. From a phylogenetic perspective, for example, attempts to reconstruct relationships among such taxa are often hampered by a poor resolution in relationship estimates, by lack of monophyly inferred from individual gene genealogies due to incomplete lineage sorting or, when multiple loci are analysed, by discordant scenarios portrayed by distinct genealogies. When a species splits into two, the diverging forms will share much of their genetic variation for a long period of time and the pre-existing variation may be sorted in a stochastic fashion or reflecting differential selective pressures, not necessarily tracking species-splitting events. Moreover, closely related species are likely to retain some permeability to the exchange of genes, which may affect some loci more than others (e.g. Wu 2001, Machado *et al.* 2002). Taking all of the above together, many authors have highlighted the importance of analysing multiple genes when making inferences on evolutionary relationships because the genealogy of a single locus may not reflect the true history of the species being analysed (Pamilo and Nei 1988, Nichols 2001, Zhang and Hewitt 2003), especially when dealing with closely-related species.

Despite this now commonly acknowledged notion that a single gene tree may not be appropriate to describe evolutionary relationships among taxa, many taxonomic reevaluations and phylogeographic studies still rely solely on mitochondrial DNA (mtDNA). MtDNA has several properties that make it an attractive source of information for phylogenetic and phylogeographic inference. For example, rates of molecular evolution for mitochondrial genes are generally

higher than for their nuclear counterparts. Another important feature is its haploid nature combined with uniparental inheritance, which reduces the effective population size of mtDNA to (on average) a quarter of that of a nuclear gene. Based on this principle, Moore (1995) provided theoretical evidence for preferring mtDNA over nuclear gene trees to infer species relationships due to the fact that it has a higher probability of tracking splitting events. Nevertheless, there are situations that may reduce the effective population size of nuclear genes relative to that of mtDNA (e.g. Slatkin and Hudson 1991, Hoelzer 1997) and may offset its apparent advantage for phylogeny and phylogeographic estimation. Additionally, the relationship between the time needed to achieve monophyly in mtDNA and in a nuclear gene is highly unpredictable because of the variance of the coalescence process (Hudson and Turelli 2003). Decoupled mtDNA and species evolution can moreover be caused by several factors besides stochasticity: for example, several studies have documented natural selection driving the evolution of the mtDNA molecule (Nachman *et al.* 1994, Ballard and Whitlock 2004, Bazin 2006). Additionally, fluctuating effective population sizes (Fay and Wu 1999, Monsen and Blouin 2003), gender-biased gene flow (e.g. FitzSimmons *et al*. 1997, Nyakaana and Arctander 1999, Piertney *et al*. 2000) and introgression (e.g. DeSalle and Giddings 1986, Shaw 2002, Alves *et al*. 2003, Chan and Levin 2005) may also contribute to biased estimates of relationships and differentiation if only mtDNA variation is taken into account. Moreover, single locus perspectives do not provide a precise estimation of parameters such as effective population sizes or divergence times (Edwards and Beerli 2000). For these reasons, although mtDNA is a practical starting point to establish evolutionary hypotheses, it is important to combine that information with that from loci from the nuclear genome. In response to this need, a whole set of new methods of analyses based on coalescence theory have been recently developed that allow the estimation of parameters of interest and hypothesis testing from multilocus data (see e.g. Knowles and Maddison 2002, Hey and Machado 2003, Pearse and Crandall 2004).

Podarcis wall lizards in the Iberian Peninsula and North Africa constitute a cryptic species complex that has been studied using both mtDNA (Harris and Sá-Sousa 2001, 2002, Harris *et al.* 2002, Pinho *et al.* 2006) and allozyme data (Sá-Sousa *et al.* 2000, Pinho *et al.* 2003, 2004, in press a). In concordance with emerging morphological evidence (Sá-Sousa 2001, Geniez 2001, Sá-Sousa *et al.* 2002) both types of markers have documented the existence of several highly differentiated population groups, some of which correspond to the currently accepted species (*P. bocagei, P. carbonelli* and *P. vaucheri*), whereas others constitute different forms within the polytypic and, from a mitochondrial perspective, paraphyletic *P. hispanica*. Reported genetic distances at the mitochondrial level are higher than traditional species-delimitation thresholds

proposed for squamates (see e.g. Harris 2002). However, not all the mitochondrial DNA lineages seem to belong to clearly distinct morphological entities. Moreover, although both nuclear and mitochondrial markers provide roughly concordant species delimitation, a detailed comparison of both analyses reports two significant differences (Pinho *et al.* in press a). First, not all of the forms detected using mtDNA are readily distinguishable using allozymes, which could result from massive introgression of the nuclear genome or from the inability of allozyme data to detect all the differentiated entities. Secondly, evolutionary relationships between species and forms within *P. hispanica* as depicted by mtDNA are very well supported, suggesting a "step-by-step" speciation scenario, whereas the analyses of allozymes produces only a few well-supported multi-species clusters, implying that the diversification process was more or less simultaneous. These discordant results were interpreted as reflecting the shorter time required by mtDNA to achieve monophyly. Despite their utility in this case of corroborating the major partitions inferred from mtDNA, allozyme analyses have several drawbacks, of which the most important for comparing the patterns of mitochondrial and nuclear levels divergence is the lack of a genealogical framework to understand the relationships between alleles and the evolutionary processes underlying the distribution of genetic variation.

In this work we aim to analyse the evolutionary history of the Iberian and North African *Podarcis* species complex from a novel perspective: that of nuclear genealogies. We studied nucleotide variation at two nuclear introns in individuals representing all known morphotypes and mitochondrial lineages. Our main goals were to: a) investigate whether species and forms of *Podarcis* are monophyletic with respect to nuclear genealogies; b) if not, to assess the relative roles of incomplete lineage sorting and gene flow in shaping the observed patterns. With this approach we expect to obtain a better picture of the dynamics that shaped the evolution of this species complex.

Materials and Methods

Sampling and DNA extraction

Currently, 11 evolutionary units have been identified using mitochondrial DNA within the Iberian Peninsula and North Africa clade of *Podarcis*. The species used here as an outgroup, *P. muralis*, is phylogenetically distinct from the other forms (Harris and Arnold 1999; Oliverio *et al.* 2000) and has a wide distribution range throughout Europe, including Northern Iberian Peninsula. Our sampling scheme was designed to incorporate a minimum of five individuals per species, morphotype or mtDNA lineage (table 1). The majority of individuals were preliminarily identified and assigned to an evolutionary group using

Table 1. Samples analysed in this study and the corresponding nuclear gene genotypes. For the geographical origin of samples, see figure 1. Alleles correspond to those in figure 2.

Species/Moprhotype	Sample Code	Locality	Country	Genotype	
				6-Pgdint7	*β-fibint7*
P. bocagei	Mad11*	Madalena	Portugal	PR1-PR2	BR1-BR2
	Vair12*	Vairão	Portugal	PR2-PR2	BR1-BR3
	Cor6*	Coruña	Spain	PR1-PR1	BR1-BR1
	M6*	Montesinho	Portugal	PR1-PR1	BR1-BR4
	BTA5*	Tanes	Spain	PR1-PR2	BR1-BR1
P. carbonelli	MC3*	Monte Clérigo	Portugal	PR3-PR3	BR5-BR5
	PR3*	Playa del	Spain	PR5-PR5	BR5-BR8
	A11*	Aveiro	Portugal	PR3-PR7	BR6-BR9
	SPM2*	S. Pedro de Moel	Portugal	PR6-PR6	BR7-BR7
	LA7*	La Alberca	Spain	PR3-PR4	BR10-BR11
P. vaucheri	LB2*	La Barrosa	Spain	PR37-PR40	-
	LB4*	La Barrosa	Spain	PR38-PR39	BR45-BR47
	LB5*	La Barrosa	Spain	-	BR46-BR47
	LB7*	La Barrosa	Spain	PR37-PR38	-
	BT6*	Bab Taza	Morocco	PR41-PR42	BR25-BR25
	Mis3*	Mischliffen	Morocco	PR41-PR42	BR48-BR49
	Ouk7*	Oukaimeden	Morocco	PR41-PR43	BR49-BR49
P. hispanica type 1A	Mon1	Montesinho	Portugal	PR9-PR9	BR12-BR25
	Mon2	Montesinho	Portugal	PR8-PR10	BR13-BR24
	Anc2	Los Ancares	Spain	PR8-PR8	BR15-BR18
	FT12	Tua	Portugal	-	BR21-BR23
	PG2	Ria de Arosa Is.	Spain	PR1-PR1	BR18-BR18
	Pen2	Pendilhe	Portugal	PR11-PR12	BR17-BR26
	Pen8	Pendilhe	Portugal	PR1-PR8	BR14-BR25
	Rua1*	Vila de Rua	Portugal	-	BR18-BR22
P. hispanica type 1B	Trj1*	Trujillo	Spain	PR20-PR22	BR27-BR30
	Oro1*	Oropesa	Spain	PR21-PR21	BR29-BR31
	Vil3	Villacastin	Spain	-	BR20-BR20
	Vil8	Villacastin	Spain	PR17-PR20	BR19-BR20
	GuaI1	Guadarrama	Spain	PR18-PR19	BR20-BR20
	HLA1	La Alberca	Spain	PR13-PR14	BR16-BR28
P. hispanica type 2	CV1	Castelo de Vide	Portugal	PR23-PR24	BR32-BR33
	SM1	S. Mamede	Portugal	PR24-PR25	BR34-BR34
	Ev4	Évora	Portugal	PR27-PR29	BR35-BR35
	Mad1	Madrid	Spain	PR26-PR28	-
	Mad2	Madrid	Spain	PR15-PR15	BR36-BR36
	And9	Saucedilla	Spain	-	BR38-BR39
	And10*	Benatae	Spain	PR26-PR27	-
	CR1	Castaño de Robledo	Spain	PR16-PR30	BR37-BR37
P. hispanica sensu stricto	Cue2	Cuenca	Spain	-	BR43-BR44
	Mot1*	Motilla del Palancar	Spain	PR31-PR33	-
	Pod12*	Granada	Spain	PR3-PR34	BR41-BR41
	And8*	Puebla de D.	Spain	PR32-PR34	BR40-BR43
	SN2	Sierra Nevada	Spain	PR33-PR34	BR25-BR42
	SN10	Sierra Nevada	Spain	PR35-PR35	BR25-BR25
	SN11	Sierra Nevada	Spain	PR36-PR36	BR25-BR42
	BEV7337	Sta. Maria de Nieva	Spain	PR56-PR58	BR62-BR66
P. hispanica type 3	Barc5	Barcelona	Spain	PR33-PR50	BR56-BR57
	Burg2*	Burgos	Spain	PR33-PR33	BR25-BR25
	Get1	Getaria	Spain	PR52-PR52	BR70-BR70
	Med1*	Medinaceli	Spain	PR3-PR51	BR58-BR59
	Pht1	Tarragona	Spain	PR53-PR54	BR60-BR60
P. hispanica Galera	Gal1x	Galera	Spain	PR55-PR57	BR64-BR65
	Gal1	Galera	Spain	-	BR63-BR63
	Gal3*	Galera	Spain	PR55-PR55	-
	Gal3x	Galera	Spain	PR55-PR56	BR67-BR67
	Gal5x	Galera	Spain	-	BR65-BR67
	BEV7353	Puebla de D.	Spain	PR55-PR56	BR68-BR69
	Gal2	Galera	Spain	PR55-PR55	-
	Gal7x	Galera	Spain	PR55-PR55	BR61-BR67
P. hispanica Jebel Sirwah	JS1*	Jebel Sirwah	Morocco	PR44-PR46	BR50-BR51
	JS2	Jebel Sirwah	Morocco	PR45-PR46	BR51-BR51
	JS3	Jebel Sirwah	Morocco	PR45-PR46	BR51-BR51
	JS6*	Jebel Sirwah	Morocco	PR44-PR46	BR51-BR51
	PH184	5Km SE of J. Sirwah	Morocco	PR45-PR46	BR52-BR53
	PH186	5Km SE of J. Sirwah	Morocco	PR44-PR44	BR51-BR51
P. hispanica Tunisia	OK1*	Oued Kébir	Tunisia	PR44-PR49	BR54-BR54
	OK8	Oued Kébir	Tunisia	PR44-PR49	BR55-BR55
	OK11	Oued Kébir	Tunisia	PR47-PR48	BR54-BR54
	LK5	Le Kef	Tunisia	PR49-PR49	BR55-BR55
	LK6*	Le Kef	Tunisia	PR44-PR49	BR55-BR55
P. muralis	MTA1*	Tanes	Spain	-	BR71-BR71
	MTA2	Tanes	Spain	-	BR71-BR72
	MTA3	Tanes	Spain	PR59-PR59	BR71-BR71
	MTA4	Tanes	Spain	PR60-PR60	BR71-BR71
	Gua1*	Guadarrama	Spain	PR61-PR61	-
	Gua2	Guadarrama	Spain	PR61-PR61	BR71-BR71
	Gua13	Guadarrama	Spain	PR61-PR61	BR71-BR71

*indicates ND4 sequence already published (Pinho *et al.* 2006, in press b)

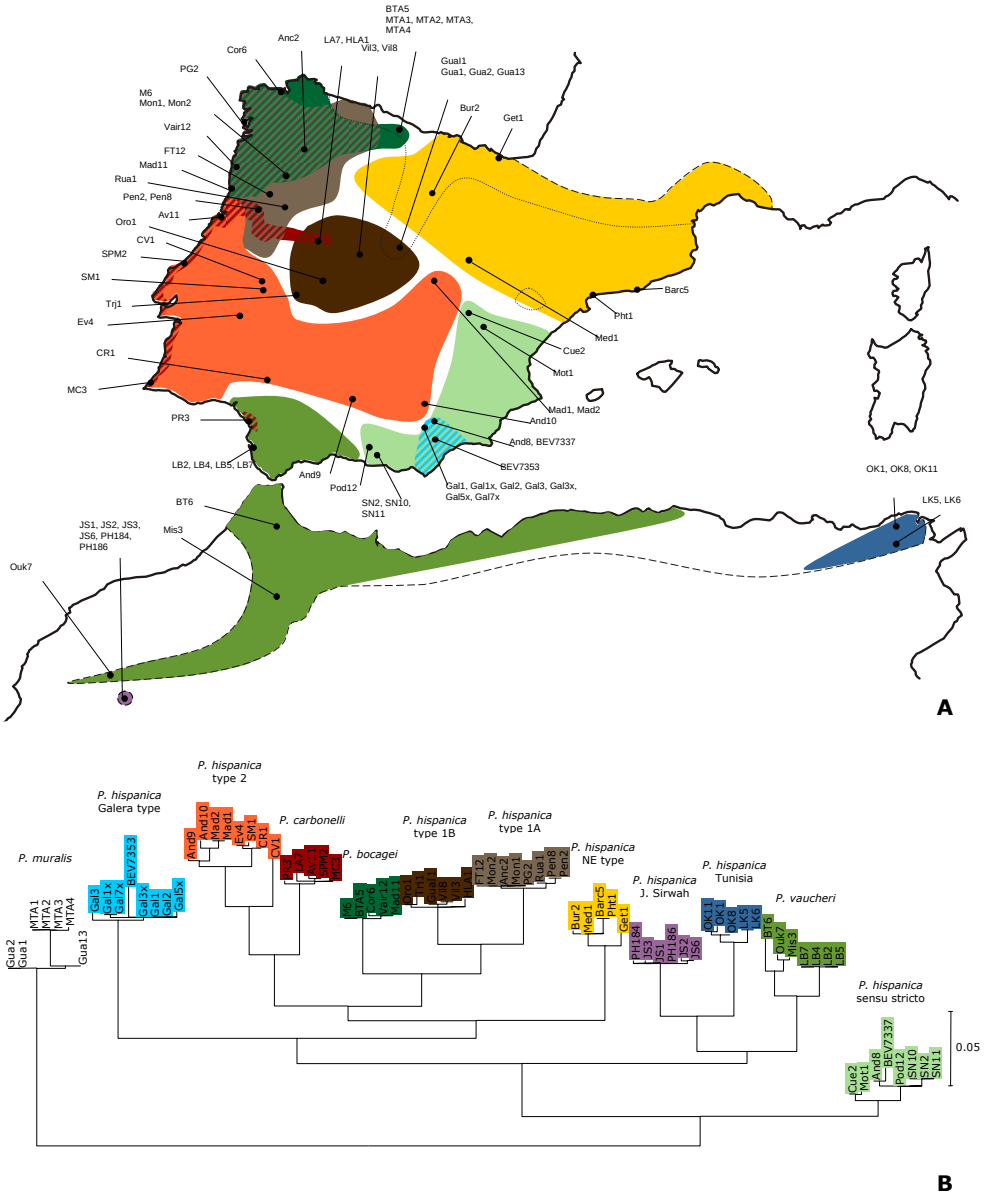

Figure 1. Geographical origin and mitochondrial DNA assignment of samples used in this study. **A**. Map of the Iberian Peninsula and the Maghreb showing the putative distribution of mtDNA lineages based on a compilation of available data and the samples analysed in this study. The dashed line represents the distribution of the Iberian-Mahgrebian clade of *Podarcis*; the dotted line represents the limits of the Iberian distribution of *P. muralis*. **B**. Maximum-likelihood tree of the mitochondrial gene ND4 for the samples analysed.

morphological criteria. Because some mtDNA lineages have not been studied in detail morphologically and others that have remain indistinguishable, our final assignment was based on mitochondrial DNA sequencing (see below). We will, from here on, refer to the defined groups as species for a matter of simplicity. It should be noted, however, that the specific status of some of these groups still remains to be evaluated. The sampling scheme was as representative of the species distributions as possible. In total, 78 individuals were analysed. The geographical locations of the samples, as well as a preliminary distribution map of mitochondrial lineages are shown in figure 1. Samples consisted of a small tail clip, obtained taking advantage of the lizard's natural tail autotomy capacity. All lizards were released after sample collection. Samples were stored in ethanol or frozen at -80ºC prior to DNA extraction, which was accomplished following standard protocols (Sambrook *et al.* 1989). All mitochondrial DNA sequences from *P. bocagei*, *P. carbonelli* and *P. vaucheri* individuals, as well as several individuals from other species (identified in table 1) were generated for previous studies on mitochondrial variation (Pinho *et al.* 2006, in press b).

DNA amplification, sequencing and haplotype determination

We generated sequences from two nuclear genes: *β*-fibrinogen intron 7 (*β-fibint7)* and 6-phosphogluconic acid dehydrogenase intron 7 (*6-Pgdint7*). To establish a correct assignment of all individuals to a mitochondrial DNA lineage, we further sequenced a mitochondrial fragment comprising the 3' end of the NADH dehydrogenase subunit 4 gene and adjacent tRNAs (from here on referred to as ND4). A list of the primers that were developed for amplification of the three studied genes is given in table 2.

Table 2. Primers developed for this study

Gene region	Primer name	Primer sequence
ND4	GalND4F	5'-TGC TAA AAC TAG GTG GCT ATG GCT TAA TCC GCA TC-3'
	GalND4R	5'-TCT CGA GTG TGG GTG GGA GGA AGG AGT CGA AT-3'
6-Pgdint7	PgdP7F	5'-GAC ATG CAG CTG ATC TGT GAG GCC-3'
	PgdP8R	5'-GAG TCC AGC TCA GTC TTA TTC CAC-3'
	PGD500	5'-CAT TTG CTC TTA AGA AAA TAG GAA G-3'
β-fibint7	BF8	5'-CAC CAC CGT CTT CTT TGG AAC ACT G-3'
	BfibR	5'-CAG GGA GAG CTA CTT TTG ATT AGA C-3'

Amplification and sequencing of the ND4 gene was accomplished using primers ND4 and Leu (Arévalo *et al.* 1994) and followed the conditions described by Pinho *et al.* (2006). In this work the authors reported the existence of a nuclear pseudogene of ND4 in the *P. hispanica* morphotype from the population of Galera that was amplified instead of the mitochondrial fragment (see Pinho *et al.* 2006 for details). To avoid the amplification of this nuclear copy we

developed primers targeting only the mitochondrial sequence observed in this morphotype (primers GalND4F and GalND4R, described in appendix 1). Sequences obtained through the use of these primers were apparently single-copy and similar to the mitochondrial sequence inferred by Pinho *et al.* (2006).

At a preliminary stage, amplification of the *β-fibint7* gene was accomplished using primers FIB-B17U and FIB-B17L (Prychitko and Moore 1997) under the conditions described in Godinho *et al.* 2005 (lowering the annealing temperature to 50ºC). However, because amplification and sequencing success was low with these primers, we tried several other combinations of primers, both already published and newly designed for this study. The definitive amplification and sequencing of this gene was accomplished using primer BFXF described in Sequeira *et al.* (2006), as well as a new primer (BF8) designed considering the alignment of sequences from *Podarcis* with those from several *Lacerta* species obtained from Godinho *et al.* (2005). Amplification was carried out in 20μL volumes, containing 2μL 10X reaction buffer (Ecogen), 2mM $MgCl_2$, 0.4mM each dNTP, 0.2μM each primer, 1 unit of *Ecotaq* DNA polymerase (Ecogen) and approximately 100 ng of genomic DNA. Amplification conditions consisted of a pre-denaturing step of 3 min at 92ºC followed by 40 cycles of a denaturing step of 30s at 92ºC, annealing at 53ºC for 30s and extension at 72ºC for 90s. The final extension was accomplished at 72ºC for 5 min. This protocol failed to amplify some of the samples, including all *P. hispanica* with the Galera mtDNA type. We therefore designed an internal primer, BfibR, which was used together with primer BF8 under the same conditions described above.

The locus *6-Pgdint7* was chosen for this work based on very high polymorphism levels observed for the studied lizards in the allozyme encoded by the *6-Pgd* gene (Pinho *et al.* 2003, 2004, in press a). We began by aligning mRNA sequences for several organisms (*Homo sapiens,* accession number BC000368, *Mus musculus*, accession number BC011329, *Danio rerio,* accession number AY391449 and *Lasaea australis,* accession number AF345495). Highly conserved regions situated on human exons 7 and 8 were targeted to design multiple degenerate primers for exon-primed, intron-crossing (EPIC) PCR (Palumbi 1995). Several combinations of these were tried on a limited number of *Podarcis* samples under gradients of annealing temperatures and $MgCl_2$ concentrations. The beginning, supposedly exonic portions of the amplified fragments revealed a high degree of homology with the mRNA sequences used to construct the initial alignment, suggesting that the sequenced genomic portion was indeed intron 7 of the *6-Pgd* gene. New primers, named PgdP7F and PgdP8R, specific for *Podarcis,* were then designed in the exons and allowed successful amplification of this gene region. Optimal amplification was achieved using the same conditions described above for *β-fibint7*, with the exception that

some samples required annealing temperatures of 60ºC. An internal reverse primer, PGD500, was also designed for sequencing purposes.

The PCR products were enzymatically purified and sequenced using the ABI Prism BigDye Terminator Cycle sequencing protocol in an ABI Prism 310 automated sequencer (Applied Biosystems) with the same primers used for amplification.

We used two approaches to resolve the haplotype phase of nuclear DNA sequences. First, for sequences that were heterozygous for insertions or deletions, we used the method recently described by Flot *et al.* (2006). Second, we used the Bayesian algorithm implemented in the PHASE software (Stephens *et al.* 2001), using known phases of haplotypes determined by the previous method. Several individuals were sequenced for which the haplotype phase could not be determined; these individuals were not included in subsequent analyses and for simplicity are not represented in table 1 and figure 1. Sequences were aligned manually using BIOEDIT v. 7.0.5.2 (Hall 1999). For nuclear genes, each allele was coded with the name of the individual carrying it following by the letters A or B.

Analytical methods

Mitochondrial DNA phylogeny. We used an alignment of 536bp of the ND4 gene to carry out phylogenetic analyses under a maximum likelihood approach using GARLI version 0.95 (Zwickl 2006; available at http://www.bio.utexas.edu/faculty/antisense/garli/Garli.html). This software allows simultaneous estimates of tree topology, branch lengths and model parameters under the general-time-reversible model of sequence evolution.

Nuclear gene genealogies. To illustrate relationships between haplotypes, a median-joining network (Bandelt *et al.* 1999) was built using NETWORK (available at http://www.fluxus-technology.com/sharenet.htm). We preferred this method of representing genealogies rather than phylogenetic trees because the levels of polymorphism observed at both nuclear genes were much lower than those observed in mitochondrial DNA. For these situations, haplotype networks constitute a better way of representing relationships (Posada and Crandall 2001). Because multiple-base insertions or deletions are likely to have resulted from a single evolutionary step, we pruned the data in order to leave only the first base of the indel. We chose this approach rather than completely removing indels because this would significantly reduce the number of polymorphic sites used to build the network and disregard much of the information contained in the data sets. Nevertheless, some segregating positions located within areas of indels were removed by this pruning method.

In order to test alternative hypotheses (see Results) relative to the topology of the nuclear gene genealogies, we also searched for ML trees that better represented the phylogeny of the observed haplotypes using GARLI. Under the model of sequence evolution derived from these analyses, we performed new ML searches enforcing a particular topology using PAUP* 4.0b10 (Swofford 2000) and compared the resulting trees' likelihoods with those obtained from unconstrained analyses using Shimodaira-Hasegawa tests (Shimodaira and Hasegawa 1999) under the RELL approximation with 1000 bootstrap replicates. This step was conducted using the full data sets (i.e. without removing indels from the alignment).

DNA sequence polymorphism, population subdivision and recombination. All the following analyses were performed using reduced data sets from which all indels were eliminated. We calculated summary diversity statistics (number of haplotypes, number of segregating sites, nucleotide diversity, π, and the population mutation parameter θ) for the three data sets and separately for each partition. We also computed Tajima's D (Tajima 1989) to test for non-neutral evolution of the analysed data-sets. We also computed values of Hudson *et al.*'s (1992) F_{ST} between all species pairs. All these analyses were conducted in DNASP (Rozas *et al.* 2003). To evaluate the possibility of recombination in the nuclear genes, we computed Hudson and Kaplan's (1985) Rm statistic (representing the minimum number of recombination events) using DNASP. Because this statistic is likely to be highly affected by homoplasy, we also used the software PHIPACK (available at http://www.mcb.mcgill.ca/~trevor) to test for the presence of recombination using the pairwise homoplasy index (Φ_w statistic) of Bruen *et al.* (2006). This statistic is a powerful detector of the presence of recombination and has been shown by simulation studies to be less sensitive to the effects of mutation rate correlation than other available statistics, which are prone to falsely infer recombination when levels of recurrent mutation are high (Bruen *et al.* 2006). PHIPACK provides *P*-values for the acceptance of the null hypothesis of no recombination. We used the two available options to estimate such *P*-values: an analytical approach and a permutation test (1000 permutations). In order to assess the presence or absence of recombination with robustness, we used three different window sizes: 50bp, 100bp (the default value) and 150bp.

Estimation of migration rates between species. We used two different approaches to assess levels of gene flow between forms. First, *Nm* values were inferred from F_{ST} according to Wright's (1951) island model of population structure using the expression $F_{ST} = 1/(1-2Nm)$ for mitochondrial DNA and $F_{ST} = 1/(1-4Nm)$ for nuclear autosomal loci. *Nm* is a combination of parameters that

indicates the relative strengths of gene flow and genetic drift; when $Nm<1$, genetic drift will result in differentiation, whereas if $Nm>1$ the homogenizing influence of gene flow will lead to a lack of differentiation (Wright 1931). Traditional, F_{ST}–based methods of estimating *Nm* assume equilibrium between drift and migration. However, species that have recently diverged often share much of their genetic variation not only due to their permeability to gene flow but also because of incomplete lineage sorting of ancestral polymorphism. In these cases, obtaining realistic estimates of migration between populations is challenging. Recently, Hey and Nielsen (2004) developed a model which takes into account population divergence and gene flow in the same framework, thus being appropriate to disentangle between the relative effects of isolation and migration in shaping the patterns of variation among diverging species. This model is implemented in the IM computer program, which works under a two-population model and uses a Markov Chain Monte Carlo approach to estimate six parameters scaled by the neutral mutation rate: effective population sizes for both extant populations and their ancestor, time since divergence and per-gene migration rates on both directions. This model was developed for a simple two-population scenario and therefore does not directly apply to the complexity inherent to divergence within Iberian and North African *Podarcis*, in which eleven mitochondrial DNA lineages have been described. Nevertheless, because no applications of this model have been developed to deal with more than two populations, we performed pairwise analyses to assess historical migration rates between all species pairs. We included the outgroup, *P. muralis*, in these analyses, because preliminary inspection of sequence data for one of the studied gene regions suggested that introgression with the ingroup could have occurred (see Results). Several analyses involving more than two species have been successfully performed using IM or similar methods (Griswold and Baker 2002, Won and Hey 2005, Won *et al.* 2005, Kronforst *et al.* 2006, Dolman and Moritz 2006). It should be kept in mind, however, that with such a two population model, influence from other populations might have an unpredictable effect in our estimates. In order to minimize such effect, we performed a second set of runs for comparisons involving *P. hispanica* sensu stricto and *P. hispanica* type 3, excluding alleles that were obviously introgressed (see Results).

Since IM cannot accommodate gaps in DNA sequence data, we pruned all gaps and missing data from the data sets. Another assumption made by IM is no recombination; because we had no evidence for recombination in our data sets using Bruen *et al.* 's (2006) method (see Results), we used the complete data sets in the analyses. Because the data did not fit the infinite sites model of sequence evolution, we used the HKY (Hasegawa *et al.* 1985) model in all analyses. After several experimental runs to assess appropriate parameter settings, for each comparison, IM was run for 10–20 million steps along the

Markov Chain after 1 million steps of burn-in (using the ramped heating scheme, option –bh), with 5 Metropolis-coupled chains with linear heating. A second run with the same options and a different random seed was used to confirm convergence of parameter estimates.

Results

Mitochondrial DNA assignment

The mitochondrial data set included 78 sequences and was trimmed to a common fragment of 534bp. This data set contained 187 polymorphic positions, of which 179 were parsimony informative. None of these polymorphisms correspond to insertions or deletions, which conforms to the coding nature of the region analysed. An ML tree depicting the relationships between the observed haplotypes is shown in figure 1. Because it includes less informative characters, the depicted tree does not reflect the same well-supported relationships that were inferred in a previous phylogenetic study including 2425bp of mitochondrial DNA sequence data (Pinho *et al.* 2006), although exactly the same major phylogroups were recovered. This tree was used to assign individuals to a mitochondrial group. All the individuals that had a prior morphological assignment clustered within the expected mitochondrial lineage, with the exception of the individual BEV7337, which morphologically resembles a "Galera type" individual (P.A. Crochet, pers. comm.) but that nevertheless clustered within the sympatric *P. hispanica* sensu stricto.

Nuclear gene variability and recombination

Alignment of the *6-Pgdint7* sequences for 136 chromosomes was trimmed to a common fragment of 503bp. This 503bp alignment required 10 insertions or deletions, from 1bp to 23bp long and included a small poly-G repetitive region with size variability. A total of 65 different alleles were detected. Alignment of *β-fibint7* was pruned to a common fragment of total length 951bp for 140 chromosomes; this alignment required several insertions and deletions, including a very large (~350bp) insertion in all samples bearing the "Galera" mtDNA lineage plus sample BEV7337 which was probably responsible for the failure to amplify the complete intron in these samples using primers BF8 and BFXF. Eighty-one alleles were detected. The "Galera" insertion showed a poly-A motif within it, with size variability, but each sample's number of repeats could not be resolved. Sites with alignment gaps were eliminated from the data sets in order to conduct further analyses. Polymorphism levels calculated from these reduced data sets are presented in table 3.

Table 3. Summary statistics and tests of recombination for the three gene regions analysed in this study. Alternative values of Tajima's D and F_{ST} are for estimates including or excluding *P. muralis*. N, number of haplotypes sequenced; S, number of segregating sites; π, nucleotide diversity; θ, population mutation parameter, calculated according to Watterson (1975); Rm, minimum number of recombination events (Hudson and Kaplan 1985); Φ_w statistic p-value calculated according to Bruen *et al.* (2006).

Gene	Taxa	Length	N	Polymorphism				Tajima's D	Recombination		F_{ST}
		(bp)		*H*	*S*	*π*	θ		Rm	Φ_w *P*-value	
ND4	All	534	78	54	187	0.10738	0.07107	0.67941 / 0.65100	-	-	0.89885 / 0.89707
	P. bocagei	534	5	3	2	0.00187	0.00180	0.24314			
	P. carbonelli	534	5	5	7	0.00562	0.00629	-0.74682			
	P. vaucheri	534	7	4	32	0.02836	0.02446	0.45857			
	P. hispanica type 1A	534	8	3	5	0.00461	0.00361	1.26023			
	P. hispanica type 1B	534	6	5	9	0.00687	0.00738	-0.41545			
	P. hispanica type 2	534	8	7	45	0.03023	0.03250	-0.48372			
	P. hispanica sensu stricto	534	8	6	17	0.01318	0.01228	0.37704			
	P. hispanica type 3	534	5	4	14	0.01273	0.01258	-0.40617			
	P. hispanica Galera	534	8	5	18	0.00876	0.01300	-1.68765*			
	P. hispanica Jebel Sirwah	534	6	4	3	0.00225	0.00246	-0.44736			
	P. hispanica Tunísia	534	5	4	9	0.00974	0.00809	1.44761			
	P. muralis	534	7	4	14	0.01427	0.01070	1.86716			
6-Pgdint7	All	414	136	60	91	0.01721	0.04007	-1.89302** / -2.02372*	7	0.276	0.61498 / 0.46435
	P. bocagei	414	10	2	1	0.00129	0.00085	1.30268			
	P. carbonelli	414	10	5	7	0.00644	0.00598	0.46601			
	P. vaucheri	414	12	7	11	0.00659	0.00880	-0.87808			
	P. hispanica type 1A	414	12	6	11	0.00853	0.00880	-0.04145			
	P. hispanica type 1B	414	10	8	16	0.01390	0.01366	-0.03458			
	P. hispanica type 2	414	14	10	24	0.01680	0.01823	-0.62704			
	P. hispanica sensu stricto	414	14	9	15	0.00805	0.00993	-0.77168			
	P. hispanica type 3	414	10	7	12	0.00859	0.01025	-0.49424			
	P. hispanica Galera	414	12	3	3	0.00131	0.00205	-1.17901			
	P. hispanica Jebel Sirwah	414	12	3	6	0.00677	0.00480	1.63651			
	P. hispanica Tunísia	414	10	6	4	0.00462	0.00512	-0.40924			
	P. muralis	414	10	3	2	0.00215	0.00171	0.83017			
β-fibint7	All	507	140	67	100	0.01269	0.03576	-2.15863** / -2.29815**	8	0.194	0.51125 / 0.40491
	P. bocagei	507	10	4	5	0.00197	0.00349	-1.74110*			
	P. carbonelli	507	10	7	18	0.01267	0.01255	-0.20511			
	P. vaucheri	507	10	6	20	0.01179	0.01394	-0.72606			
	P. hispanica type 1A	507	16	11	25	0.00830	0.01486	-1.72813			
	P. hispanica type 1B	507	12	8	13	0.00529	0.08490	-1.59698			
	P. hispanica type 2	507	12[+]	8	18	0.01133	0.01176	-0.15948			
	P. hispanica sensu stricto	507	14	9	16	0.00759	0.01013	-0.61276			
	P. hispanica type 3	507	10	7	20	0.01315	0.01394	-0.26800			
	P. hispanica Galera	507	12	6	9	0.00493	0.00588	-1.30703			
	P. hispanica Jebel Sirwah	507	12	2	1	0.00033	0.00065	-1.14053			
	P. hispanica Tunísia	507	10	1	0	0.00000	0.00000	-			
	P. muralis	507	12	2	1	0.00033	0.00065	-1.14053			

* p <0.05; ** p<0.01; [+]2 alleles were discarded because of a very large deletion.

The nuclear loci analysed in this study, despite being introns, show polymorphism levels considerably lower than those observed in the mitochondrial DNA fragment analysed (e.g. $\pi_{6\text{-}Pgdint7}$=0.01721; $\pi_{\beta\text{-}fibint7}$=0.01269; π_{ND4}=0.10738). Interestingly, both nuclear genes show significantly negative values of Tajima's D, indicating a skew towards rare alleles. Both genes failed to pass the four-gamete test of Hudson and Kaplan (1985) and were inferred to have suffered several recombination events (at least 7 in the case of *6-Pgdint7* and at least 8 in *β-fibint7*. However, when applying the more sensitive test of Bruen *et al.* (2006) we were not able to reject the null hypothesis of no recombination (*P*-values shown in table 3), suggesting that recurrent mutations are frequent in both genes.

Nuclear gene genealogies

The haplotype networks inferred for both nuclear genes are shown in figure 2. These networks were built after removing all but the first base of each indel (see Methods), which resulted in 100 variable characters and 61 alleles in *6-Pgdint7* and 113 characters and 72 alleles in *β-fibint7*.

The most evident result is a complete lack of monophyly of the mitochondrial-defined groups. Many alleles are even shared by distinct species. Moreover, although the outgroup *P. muralis* represents a relatively distinct lineage according to both genes, the group formed by all Iberian and North African *Podarcis* is not monophyletic with respect to this species according to the *β-fibint7* genealogy, since one individual carrying *P. hispanica* type 3 mtDNA exhibits an allele that falls within the clade formed by *P. muralis.* The only species group that is consistently monophyletic is the clade formed by *P. hispanica* from Tunisia and Jebel Sirwah.

Because this lack of monophyly could result from an insufficiency of data to recover correct relationships, we tested alternative hypotheses based on mitochondrial DNA phylogenetic relationships using a Shimodaira-Hasegawa test (Shimodaira and Hasegawa 1999). Using datasets without removing indels (but excluding the poly-A region in "Galera" samples), we enforced two distinct topological constraints for each gene: i) monophyly of all species; ii) monophyly of the three major groups recovered for Iberian and North African *Podarcis* by Pinho *et al.* (2006) (group 1: *P. bocagei, P. carbonelli*, *P. hispanica* type 1A and 1B, *P. hispanica* type 2; group 2: *P. hispanica* sensu stricto, *P. vaucheri*, *P. hispanica* Tunisia and *P. hispanica* Jebel Sirwah; group 3: *P. hispanica* type 3 and *P. hispanica* Galera). In both of these analyses *P. muralis* was included also enforcing its monophyly. Because a few cases of allele proximity could at a first glance be interpreted as suggestions of recent gene flow, we also conducted the same analyses excluding these alleles (BEV7337A and BEV7337B in *6-Pgdint7*; BEV7337A, BEV7337B, Get1A and Get1B in *β-fibint7*). All these tests suggest

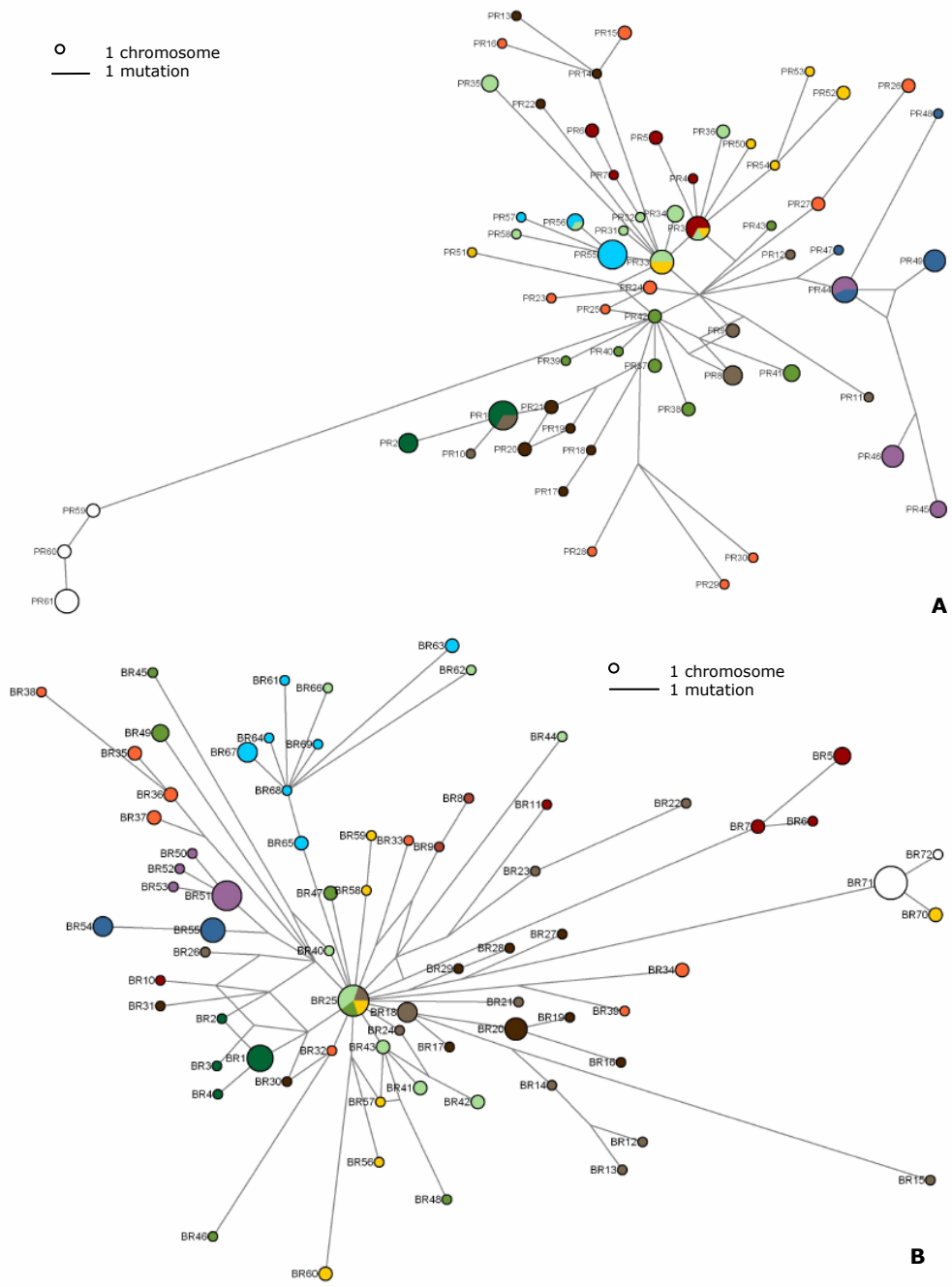

Figure 2. Gene genealogies for two nuclear introns in Iberian and North African *Podarcis.* Allele names correspond to those in table 1. Colours represent mitochondrial DNA lineages and correspond to those used in figure 1. Alleles detected in *P. muralis* are represented in white. **A.** *6-Pgdint7.* **B**. *β-fibint7.*

that enforced topologies are significantly less likely than the ones observed (p<0.01), with the exception of the division into three groups in *β-fibint7* excluding the putative introgressed alleles.

Genetic differentiation and classic gene flow estimates between groups

We evaluated overall genetic differentiation among mtDNA-defined species by computing F_{ST} for the two data sets. These results are shown in table 3. Values obtained for the mitochondrial DNA are also shown for comparison (noting, however, that "species" are, by definition, monophyletic with respect to the mtDNA, leading to necessarily high levels of differentiation). Despite the lack of monophyly, differentiation between the 12 species accounts for around 61% of genetic variation in *6-Pgdint7* and 51% in *β-fibint7*; these values fall to around 46% and 40%, respectively, when the outgroup *P. muralis* is excluded from the analyses.

Examination of pairwise *Nm* values in more detail (table 4) reveals distinct trends across species pairs; in general *Nm* values based on nuclear genes are different from 0, ranging from 0.01 (between *P. muralis* and both *P. bocagei* and *P. hispanica* Tunisian type) to 1.76 (between *P. hispanica* types 1A and 1B). Besides this last case, other species pairs exhibit *Nm* values higher than 1, suggesting large levels of gene flow: *P. vaucheri* vs. *P. hispanica* types 1A and 1B, *P. hispanica* type 1B vs. *P. hispanica* type 2 and *P. hispanica* sensu stricto vs. *P. hispanica* type 3.

Application of the isolation with migration model to Iberian and North African Podarcis

Independent runs using IM converged on the same marginal posterior probability distributions. We were able to obtain reliable estimates for θ1, θ2 and migration rates across most of the tested pairs; however, some 90% highest posterior density (HPD) intervals could not be reliably assessed either for θ or migration. The posterior probability distributions for the ancestral population sizes were generally flat or had a maximum at the lowest bin; our data does not therefore seem to incorporate enough information for these estimates to be reliable. Time since divergence scaled by the mutation rate (t) was also not recovered with confidence for a few species pairs because the distributions were flat (albeit non-zero), particularly in the comparisons involving species that are reciprocally monophyletic for both genes. Most of the curves, however, were well resolved or presented a clear peak although their tails did not reach 0; in these cases we obtained a value for the HPD but not a reliable 90% HPD interval.

Table 4. Pairwise F_{ST} (below the diagonal) and corresponding *Nm* values (calculated according to Wright (1951)) between mtDNA-defined species of Iberian and North African *Podarcis*. *Nm v*alues higher than 1 are shown in bold. Species names are abbreviated as follows: *Pb* – *P. bocagei* ; *Pc* – *P. carbonelli*; *Pv* – *P. vaucheri*; *Ph*1A – *P. hispanica* type 1A; *Ph*1B – *P. hispanica* type 1B; *Ph*2 – *P. hispanica* type 2; *Ph*ss – *P. hispanica* sensu stricto; *Ph*3 – *P. hispanica* type 3; *Ph*Gal – *P. hispanica* Galera type; *Ph*JS – *P. hispanica* Jebel Sirwah type; *Ph*Tun – *P. hispanica* Tunisian type; *Pm* – *P. muralis.*

		Pb	***Pc***	***Pv***	***Ph*1A**	***Ph*1B**	***Ph*2**	***Phss***	***Ph*3**	***Ph*Gal**	***Ph*JS**	***Ph*Tun**	***Pm***
Pb	mtDNA	-	0.02	0.09	0.02	0.03	0.11	0.04	0.04	0.03	0.01	0.03	0.04
	nDNA	-	0.15	0.28	0.51	0.43	0.32	0.22	0.24	0.07	0.06	0.04	0.01
Pc	mtDNA	0.9649	-	0.10	0.02	0.03	0.14	0.06	0.05	0.03	0.02	0.03	0.04
	nDNA	0.6240	-	0.36	0.38	0.39	0.48	0.73	0.87	0.22	0.18	0.15	0.05
Pv	mtDNA	0.8428	0.8277	-	0.08	0.09	0.18	0.14	0.13	0.09	0.11	0.12	0.09
	nDNA	0.4728	0.4084	-	**1.69**	**1.19**	0.89	0.66	0.69	0.23	0.26	0.25	0.06
***Ph*1A**	mtDNA	0.9572	0.9565	0.8649	-	0.05	0.09	0.04	0.04	0.03	0.02	0.02	0.04
	nDNA	0.3274	0.3952	0.1288	-	**1.76**	0.79	0.79	0.73	0.25	0.25	0.25	0.07
***Ph*1B**	mtDNA	0.9348	0.9432	0.8435	0.9072	-	0.10	0.05	0.05	0.03	0.02	0.03	0.04
	nDNA	0.3674	0.3880	0.1742	0.1244	-	**1.02**	0.88	0.89	0.25	0.24	0.23	0.06
***Ph*2**	mtDNA	0.8175	0.7790	0.7380	0.8522	0.8265	-	0.13	0.12	0.08	0.09	0.10	0.10
	nDNA	0.4391	0.3427	0.2187	0.2413	0.1965	-	0.69	0.72	0.32	0.33	0.33	0.08
Phss	mtDNA	0.9186	0.8919	0.7858	0.9253	0.9028	0.7986	-	0.07	0.05	0.05	0.06	0.07
	nDNA	0.5286	0.2550	0.2752	0.2404	0.2206	0.2651	-	**1.64**	0.51	0.20	0.20	0.05
***Ph*3**	mtDNA	0.9235	0.9089	0.7966	0.9282	0.9079	0.8006	0.8760	-	0.06	0.03	0.06	0.05
	nDNA	0.5119	0.2236	0.2648	0.2557	0.2200	0.2568	0.1322	-	0.34	0.22	0.22	0.08
***Ph*Gal**	mtDNA	0.9389	0.9373	0.8454	0.9426	0.9360	0.8578	0.9090	0.8982	-	0.03	0.04	0.04
	nDNA	0.7910	0.5367	0.5255	0.5041	0.5038	0.4352	0.3299	0.4240	-	0.08	0.07	0.02
***Ph*JS**	mtDNA	0.9787	0.9685	0.8143	0.9703	0.9614	0.8496	0.9113	0.9349	0.9523	-	0.04	0.04
	nDNA	0.8123	0.5837	0.4924	0.4974	0.5132	0.4324	0.5534	0.5272	0.7614	-	0.12	0.02
***Ph*Tun**	mtDNA	0.9471	0.9453	0.8044	0.9529	0.9368	0.8283	0.8867	0.8963	0.9322	0.9276	-	0.05
	nDNA	0.8492	0.6325	0.4977	0.5034	0.5229	0.4287	0.5603	0.5328	0.7892	0.6794	-	0.01
Pm	mtDNA	0.9321	0.9237	0.8414	0.9310	0.9189	0.8387	0.8740	0.9025	0.9229	0.9338	0.9169	-
	nDNA	0.9498	0.8372	0.8094	0.7928	0.8164	0.7576	0.8265	0.7494	0.9198	0.9381	0.9557	-

In order to convert IM estimates into biologically meaningful values (effective population sizes, divergence times in years or migration rates) one needs an estimate of mutation rate. For the two genes included in the analyses, we have no information regarding the mutation rate and no straightforward way of calibrating a molecular clock. We therefore focused on estimates that do not require the assumption of a particular evolutionary rate: the population migration rate (*2Nm*) and estimates of relative divergence times (t = *tμ*) between species pairs. These results are given in tables 5 and 6, respectively. In these tables, results involving *P. hispanica* sensu stricto and *P. hispanica* type 3 were estimated excluding alleles presumably introgressed from *P. hispanica* Galera type and *P. muralis*, respectively (obviously, the comparisons with the species from which the alleles were introgressed were performed on the complete data sets). Confirmation that these alleles were introgressed from these particular species was obtained when runs excluding them yielded

Table 5. Maximum-likelihood estimates and 90% highest posterior distributions of population migration rates (*2Nm*) between species of Iberian and North African *Podarcis*, calculated using the IM software. Values higher than 0.05 are highlighted. Species abbreviations are the same as in table 4

		from										
		Pb	***Pc***	***Pv***	***Ph*1A**	***Ph*1B**	***Ph*2**	***Phss***	***Ph*3**	***Ph*Gal**	***Ph*JS**	***Ph*Tun**	***Pm***
to	***Pb***	-	0.001 0.000-1.280	0.001 0.000-0.796	0.001 0.000-2.884^{+}	0.000 0.000-1.787^{+}	0.001 0.000-0.222	0.015 0.005-4.637^{+}	0.001 0.000-1.110	0.016 0.000-0.646	0.001 0.000-0.757	0.001 0.000-0.782	0.001 0.000-0.348
	Pc	0.035 0.002-2.165	-	0.005 0.002-1.622	0.086 0.002-1.875	**0.145** 0.002-1.934	0.006 0.002-1.525	0.004 0.002-4.968	0.005 0.002-4.603	0.005 0.003-1.883	0.006 0.003-1.852	0.006 0.003-1.983	0.006 0.003-0.617
	Pv	0.016 0.006-12.278	0.011 0.005-3.325	-	0.015 0.002-23.282^{+}	0.008 0.001-18.577^{+}	0.011 0.004-3.654	**1.246** 0.003-9.682	**0.228** 0.001-12.485^{+}	0.010 0.004-8.070	0.009 0.004-2.861	0.009 0.004-3.761	0.010 0.005-1.014
	***Ph*1A**	**6.376** 0.009-50.859*$^{+}$	0.014 0.005-2.504	0.015 0.007-9.151	-	0.017 0.007-27.424*	0.018 0.010-1.874	0.014 0.006-11.321	0.017 0.009-3.993	0.013 0.007-2.034	0.017 0.008-16.265^{+}	0.018 0.008-34.647*	0.011 0.006-0.839
	***Ph*1B**	0.012 0.006-26.036*	0.010 0.006-2.069	0.036 0.004-58.638*$^{+}$	0.009 0.004-6.992	-	0.012 0.006-3.797	**0.336** 0.004-18.837	0.012 0.005-7.917	0.009 0.005-2.336	0.012 0.005-6.189^{+}	0.012 0.005-5.007	0.009 0.005-0.880
	***Ph*2**	0.018 0.010-11.758	0.020 0.011-1.830	0.018 0.010-3.050	0.017 0.010-1.788	0.017 0.008-5.149	-	0.016 0.009-1.942	0.017 0.009-2.209	0.018 0.010-2.335	0.018 0.009-2.196	0.017 0.009-2.553	0.015 0.009-0.891
	Phss	0.001 0.000-0.417	**0.109** 0.002-5.794	0.005 0.002-1.973	0.003 0.002-1.481	0.005 0.002-5.716	0.005 0.002-0.814	-	0.003 0.001-13.098^{+}	**1.156** 0.129-6.072	0.005 0.002-1.424	0.005 0.002-1.719	0.005 0.003-0.501
	***Ph*3**	0.012 0.005-8.500^{+}	0.014 0.005-19.796	0.012 0.004-47.275*$^{+}$	0.008 0.003-3.467	0.011 0.004-9.743	0.010 0.005-1.527	**5.484** 0.004-75.707*$^{+}$	-	0.013 0.005-10.353*	0.009 0.004-6.929*	0.010 0.004-9.566	**0.134** 0.004-2.073
	***Ph*Gal**	0.002 0.001-0.665	0.002 0.001-0.722	0.002 0.001-0.880	0.002 0.001-0.542	0.002 0.001-0.653	0.002 0.001-0.553	0.002 0.001-0.856	0.002 0.001-0.640	-	0.002 0.001-0.708	0.002 0.001-0.740	0.002 0.001-0.391
	***Ph*JS**	0.001 0.000-0.737	0.001 0.000-1.435^{+}	0.001 0.000-1.219^{+}	0.000 0.000-2.432	0.001 0.000-1.823^{+}	0.001 0.000-0.750	0.001 0.000-0.899	0.001 0.000-1.171^{+}	0.003 0.000-0.743	-	**0.107** 0.000-1.548^{+}	0.001 0.000-0.277
	***Ph*Tun**	0.003 0.000-0.721	0.001 0.000-1.116	0.001 0.000-0.914	0.001 0.000-0.645	0.001 0.000-0.809	0.001 0.000-0.746	0.001 0.000-0.900	0.001 0.000-0.983	0.001 0.000-0.752	0.002 0.000-1.142	-	0.001 0.000-0.309
	Pm	0.003 0.000-0.267	0.001 0.000-0.350	0.018 0.000-0.500	0.017 0.000-0.461	0.010 0.000-0.457	0.001 0.000-0.432	0.016 0.000-0.354	0.000 0.000-0.658^{+}	0.005 0.000-0.264	0.002 0.000-0.219	0.001 0.000-0.228	-

*, $^{+}$: the actual interval could not be reliably estimated because the likelihood surfaces for θ (*) or m ($^{+}$) were relatively flat.

Table 6. Maximum-likelihood estimates of time since divergence (t = *t*μ) between species of Iberian and North African *Podarcis*, obtained using IM. Only estimates corresponding to a clear peak in the likelihood surface are shown. Species abbreviations are the same as in table 4.

	Pb	***Pc***	***Pv***	***Ph*1A**	***Ph*1B**	***Ph*2**	***Phss***	***Ph*3**	***Ph*Gal**	***Ph*JS**	***Ph*Tun**	***Pm***
Pb	-											
Pc	?	-										
Pv	2.163 1.063-24.988*	2.488 1.063-10.338	-									
***Ph*1A**	?	3.163 1.338-5.862	2.288 1.588-3.188	-								
***Ph*1B**	?	2.913 0.838-20.963*	2.313 0.388-7.688	2.413 0.338-6.238	-							
***Ph*2**	3.338 0.788-15.313*	3.913 2.388-5.388	3.363 2.213-4.663	3.438 2.513-4.438	2.788 1.288-4.238	-						
Phss	1.888 1.063-24.963*	1.913 0.688-22.888*	2.188 1.088-23.288*	2.513 1.013-22.813*	2.288 0.688-20.413*	2.963 1.088-9.463	-					
***Ph*3**	1.9875 0.738-23.163*	1.838 0.463-10.838*	2.112 0.658-10.513	2.663 1.538-3.963	2.313 1.488-3.363	3.088 1.838-4.238	1.688 0.663-24.988*	-				
***Ph*Gal**	?	2.663 1.713-24.913*	2.238 1.088-23.938*	2.838 0.936-14.338*	2.313 0.813-23.763*	3.313 0.988-10.238	1.938 0.963-24.413*	1.913 0.713-23.513*	-			
***Ph*JS**	?	?	2.738 1.238-24.538*	2.938 1.088-24.988*	2.488 0.838-24.038*	2.913 0.713-10.913	2.063 1.338-24.938*	2.288 0.838-22.463*	?	-		
***Ph*Tun**	?	?	2.363 1.063-23.538*	2.738 1.163-24.888*	2.413 0.613-23.537*	3.113 0.663-19.063*	2.038 1.163-24.588*	2.138 0.813-22.988*	?	?	-	
Pm	?	?	?	?	?	?	?	?	?	?	?	-

*the credibility interval could not be reliably estimated because the likelihood surface was relatively flat.

migration rates of 0, in contrast to the moderate levels estimated when they were present. In other cases of introgression, alleles that were obviously introgressed could not be a priori pinpointed with confidence.

Population migration rates obtained using this method are strikingly different from those calculated using classic estimators of gene flow. In fact, *2Nm* values between species pairs are in general very low, close to zero. Most of these values are at the lower limit of resolution because the obtained migration parameter estimates (*m1* and *m2*) correspond to the first bin of the surveyed parameter space. However, clearly non-zero estimates of gene flow were obtained between ten species pairs. High population migration rates were documented from *P. bocagei* into *P. hispanica* type 1A and from *P. hispanica* sensu stricto into *P. hispanica* type 3 (when excluding introgressed alleles; when these were included, gene flow between these two forms was inferred to be close to 0, results not shown). In this last case, it should be pointed out that the estimates are not completely reliable because migration curves from both runs showed two peaks: one, corresponding to the maximum likelihood estimate, reflecting high levels of migration, and a much smaller peak close to zero. That is, the program could not confidently choose between the hypotheses of high or zero gene flow, although higher levels of support were found for high migration rates. A relatively high migration rate was also detected from *P. hispanica* "Galera" type to *P. hispanica* sensu stricto and from the latter to *P. vaucheri*; lower but still non-zero rates were also observed from *P. carbonelli* into *P. hispanica* sensu stricto, from *P. hispanica* type 1B into *P. carbonelli*, from *P. hispanica* sensu stricto to *P. hispanica* type 1B, from *P. hispanica* type 3 into *P. vaucheri*, from *P. hispanica* Tunisian type to Jebel Sirwah type, and from *P. muralis* to *P. hispanica* type 3. Curiously, all these cases involve instances of asymmetrical gene flow; moreover, five out of these ten situations involve *P. hispanica* sensu stricto.

Reliable estimates of relative divergence times range from 1.69 between *P. hispanica* sensu stricto and *P. hispanica* type 3, to 3.91 between *P. carbonelli* and *P. hispanica* type 2.

Discussion

Lack of population structure in nuclear genealogies

The most striking result emerging from this study is a complete lack of monophyly at nuclear genes in Iberian and North African lizards. Both of the studied nuclear genealogies fail to detect obvious population structure that relates to that observed in mitochondrial DNA. In fact, there is a single common pattern emerging from the study of mtDNA, allozymes and both the nuclear

genealogies, which is a close relationshp between *P. hispanica* from Tunisia and Jebel Sirwah.

This situation of polyphyly in nuclear genealogies relative to mitochondrial data is not uncommon (see for example Buckley *et al.* 2006, Bull *et al.* 2006, Dolman and Moritz 2006, amongst many others, for similar results). At a first glance, this discordance could suggest that mitochondrial DNA-defined species of Iberian and North African *Podarcis* do not correspond to true evolutionary entities and that this differentiation is the result of stochastic or deterministic effects acting only on the mitochondrial genome. In fact, it has been demonstrated that quite deep phylogeographic breaks in single gene genealogies may appear in the absence of historical barriers to gene flow simply because of stochastic lineage sorting (Neigel and Avise 1993; Irwin 2002). It has also been shown that mitochondrial DNA divergence may not be necessarily corroborated by nuclear divergence because of selection acting on mitochondrial DNA (Ballard *et al.* 2002). However, there is a remarkable degree of concordance between the units defined based on mitochondrial DNA (Harris and Sá-Sousa 2001, 2002; Harris *et al.* 2002; Pinho *et al.* 2006) and those observed by morphological analyses (Sá-Sousa 2001, Geniez 2001). Moreover, multilocus nuclear data (Pinho *et al.* 2003, 2004, in press a) also suggest the same pattern. Since both mitochondrial and nuclear data suggest high levels of differentiation, it is therefore unlikely that subdivision within Iberian and North African *Podarcis* is nonexistent. Why, then, do nuclear genealogies apparently not support the partitions observed in mtDNA and, in particular, in allozyme data?

A preliminary explanation could be a low level of diversity and consequently of resolution in estimates of relationships. However, monophyly of *Podarcis* mtDNA lineages has been observed even in mitochondrial genes evolving at a very slow rate (such as the control region data set analysed by Pinho *et al.* 2006, for which ~14.5% of the positions are variable (compared to ~22% in *6-Pgdint7* and ~19.7% in *β-fibint7*); also, if a lack of resolution was responsible for the polyphyly of mtDNA groups, than we would expect that no significant differences would be observed in tree likelihoods with and without enforced species' monophyly. Tests involving topological constraints, however, demonstrated that this is not the case.

Distinguishing between incomplete lineage sorting of ancestral polymorphism and gene flow

Given that lack of informative variation is not a satisfactory explanation for the discrepant results observed between mtDNA and nuclear genealogies, we are left with two possible scenarios, which are not mutually exclusive: a) incomplete

lineage sorting of ancestral polymorphism and b) gene flow, particularly male-biased.

In order to investigate whether, despite their polyphyly, species of wall lizards are differentiated with respect to nuclear genealogies, it becomes critical to discriminate between the influences of these two scenarios in shaping the patterns of allele sharing. Our data strongly suggest that, although gene flow may be important between some species pairs, the polyphyletic pattern mostly derives from the incomplete lineage sorting of ancestral polymorphism. The persistence of ancestral polymorphism in nuclear loci is a likely scenario because nuclear genes take on average four-times as much time to reach monophyly than mitochondrial DNA does (Moore 1995). Therefore, when differentiation is recent, there is a distinct possibility that mitochondrial genes may be monophyletic (in this case, with respect to morphological traits) whereas their nuclear counterparts are not. Three major lines of evidence support our conclusion of a greater effect of incomplete lineage sorting:

Comparison with allozyme variation. If gene flow was the main cause for the non-monophyly of species when considering nuclear gene variation, then allozyme data would also fail to observe any differentiation. This is certainly not the case in Iberian and North African *Podarcis*. A recent study documented not only that species are in general highly differentiated and monophyletic when considering a population tree, but also that most species can be individualized as discrete clusters by taking into account individual multilocus genotypes (Pinho *et al.* in press a). There are nevertheless exceptions with this regard, since some species pairs could not be distinguished in such clustering approaches; even if these exceptions resulted solely from introgression, the correspondent pattern in nuclear genealogies would be the sharing of alleles between these particular indistinguishable or hybridizing forms and not others, which is not verified.

Relationships between allele age and transpecificity. A second line of evidence suggesting a major role for incomplete lineage sorting relies on the simple observation of patterns of allele relationships, particularly on the *β-fibint7* data set. If gene flow was the main cause for the observed species polyphyly, we would expect that both "ancestral" (alleles that have a central position in the haplotype network) and "derived" alleles were transpecific or closely related to alleles found in other species. However, we do not detect this pattern; instead, we find that putatively older alleles are more widespread among species than alleles placed as tips, which are likely to have arisen after the species were separated. In order to visualize this trend, we plotted for the *β-fibint7* data set the number of mutations separating a given haplotype from haplotype BR25, placed in the centre of the network, against the number of mutations separating

that particular haplotype from the nearest haplotype found in another species (fig. 3). The ancestral nature of haplotype BR25 was confirmed using an alternative method of haplotype network construction (TCS 1.21, Clement *et al.* 2000), which pinpoints the haplotype that is most probably the ancestral within a network. Although figure 3 is a very rough representation, it clearly illustrates the relationship between allele "ancestrality" and "sharedness". Exceptions to this "rule" therefore constitute the most obvious cases of gene flow, as documented by the position of the haplotypes BR62 and BR66 (putatively introgressed from *P. hispanica* Galera type into *P. hispanica* sensu stricto). The results are not as clear for *6-Pgdint7*, both because a central haplotype cannot be pinpointed with confidence and because this gene shows more segregation between species groups. Nevertheless, there is still a positive correlation between both measures (results not shown).

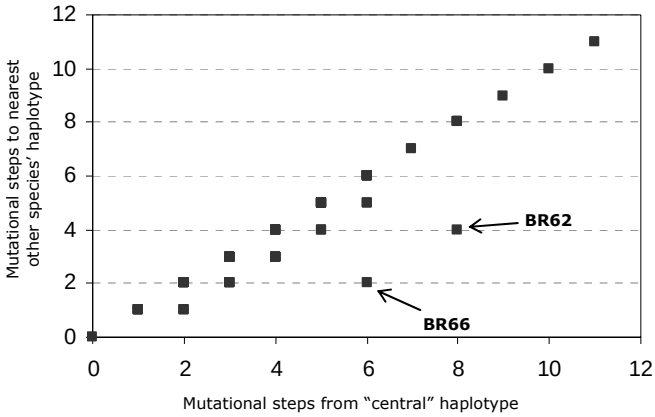

Figure 3. Relationship between allele "ancestrality", measured as the distance to haplotype BR25, and trans-species allele proximity, measured as the distance to the closest allele found in a different species, for locus *β-fibint7*. Alleles belonging to *P. muralis* (including the supposedly introgressed allele found in *P. hispanica* type 3) were excluded. Other putatively introgressed alleles are highlighted.

Estimates of gene flow between species. Although classic estimates of migration rates suggest the existence of non-zero gene flow between almost all species pairs, inferences based on coalescent simulations in a divergence with gene flow framework suggest the opposite scenario: all but ten species pairs show virtually zero levels of historical gene flow. More over, only two of these cases point to important levels of gene flow since divergence, with most estimates falling close to zero. Differences between both classes of estimators reflect the fact that classic estimators are immune to the effects of stochastic lineage sorting, whereas coalescence-based approaches incorporate information on the species divergence time and therefore take into account the possibility of allele sharing by simple incomplete lineage sorting. This result and its interpretation are similar to a study involving comparisons between equilibrium and non-equilibrium estimates of migration rates between recently expanded chaffinch populations (Griswold and Baker 2002).

Taken together, these results illustrate that, even though they are non-monophyletic, Iberian and North African wall lizard species are indeed overall differentiated with respect to nuclear genealogies.

Evidence for historical gene flow among Podarcis *species*

Although our results argue in favour of a major role for incomplete lineage sorting in shaping the observed polyphyletic patterns in nuclear gene genealogies, gene flow between some species pairs has unequivocally occurred.

Gene flow estimates based on an isolation with migration model revealed very high levels of admixture from *P. bocagei* to *P. hispanica* type 1A and most likely from *P. hispanica* sensu stricto to *P. hispanica* type 3. Albeit lower, still important levels were detected from *P. hispanica* "Galera" type to *P. hispanica* sensu stricto and from *P. hispanica* sensu stricto into *P. vaucheri*. Excluding the latter, these cases do not constitute a surprise since suggestions of hybridization and introgression between these species pairs have been described before. Concerning *P. bocagei* and *P. hispanica* type 1A, these were based on individual multilocus genotype clustering (Pinho *et al.* in press a) and the observation of interspecific matings in nature (R. Ribeiro and M. A. Carretero, pers. comm.). Gene flow between *P. hispanica* sensu stricto and *P. hispanica* type 3 had also been suggested to explain why these two forms were indistinguishable based on allozyme multilocus genotype clustering (Pinho *et al.* in press a) and based on morphology-mitochondrial DNA discordance (Renoult 2006). The same discordant trend (with some additional evidence from nuclear data) also previously suggested hybridization between *P. hispanica* "Galera" type and *P. hispanica* sensu stricto (Renoult 2006). Notably, the trend for unidirectional admixture recovered in this study is not sustained by these previous reports, which suggest gene flow in the opposite or even both directions.

Remarkably, we detected the occurrence of gene flow between forms that are sympatric at least in part of their distribution (*P. bocagei*/*P. hispanica* type 1A; *P. hispanica* "Galera"/*P. hispanica* sensu stricto; *P. hispanica* type 3/*P. muralis*; *P. carbonelli*/*P. hispanica* type 1B), confirmedly or putatively parapatric (*P. hispanica* sensu stricto/*P. hispanica* type 3; *P. vaucheri*/*P. hispanica* sensu stricto) and completely allopatric, sometimes with other species distributed in between (*P. carbonelli*/*P. hispanica* sensu stricto; *P. hispanica* type 3/*P. vaucheri*; *P. hispanica* sensu stricto/*P. hispanica* type 1B; *P. hispanica* Jebel Sirwah/*P. hispanica* Tunisia). On one hand, this suggests that present contact is not a prerequisite for gene exchange; given that species have most likely suffered dramatic and repeated pulses of range expansions and contractions due to geological or climatic events, these allopatric pairs of species may have met at some point in their past and then hybridized (it should also be noted that

such situations could artificially arise as an effect of applying a two-population model to our multi-population system; a third species hybridizing with both *P. carbonelli* and *P. hispanica* sensu stricto, for example, could in theory promote such results). On the other hand, this also suggests that gene flow going on across contact zones, if present, is likely spatially-restricted, because only two cases corresponding to such a probable situation were detected.

Besides spatial considerations, other results regarding introgression between species are noteworthy. First, in the case of *P. bocagei* and *P. hispanica* type 1A, the inferred levels of admixture were surprisingly high, above the values considered as thresholds for the maintenance of genetic differentiation (Wright 1931). However, the two species are sympatric and nevertheless maintain their morphological identifiability. It has also been shown that males discriminate conspecific from heterospecific females based on chemical stimuli, which suggests at least some degree of assortative mating (Barbosa *et al.* 2006). Furthermore, the two species correspond to clearly distinct genetic clusters as shown by almost perfect assignment using individual allozyme multilocus genotype clustering approaches (Pinho *et al.* in press a). Taking this into consideration, it seems that the present estimates of gene flow between these two species are slightly inflated; this unexpected result could result from our limited sampling and therefore needs further clarification.

Second, the present analyses suggest that *P. muralis* may have engaged in gene exchange with *P. hispanica* type 3. This result was unexpected given that *P. muralis* belongs to a very distinct clade within the genus *Podarcis*, with estimated divergence times from the Iberian and North African clade of at least 10 million years (Oliverio *et al.* 2000; hence its use for outgroup purposes in this study). This suggests that species of *Podarcis* may take a long time to acquire complete reproductive isolation.

Third, *P. hispanica* sensu stricto seems to be more prone than other species to introgression with congenerics. It has indeed been suggested, mainly based on discordance between morphological and mitochondrial DNA data (but also with some, less solid, insights from microsatellite and intron data (Renoult 2006)) that the divergent mitochondrial clade that we refer to as *P. hispanica* sensu stricto is a "ghost" lineage; in other words, this divergent mtDNA clade would be an extant signature of a nowadays extinct species whose mtDNA was captured by other species inhabiting its former area of distribution (namely *P. hispanica* Galera type and *P. hispanica* type 3), while its nuclear background was diluted through introgression. Our results partially agree with this scenario; on one hand, we find that gene flow has indeed occurred between the mtDNA-defined *P. hispanica* sensu stricto and both *P. hispanica* Galera type and type 3. However, inferred gene exchange with these forms is not similar: with respect to *P. hispanica* type 3, we find evidence for extensive gene flow, which is largely

corroborated by the analyses of allozyme data, in which these two forms were found to be non-diagnosable (Pinho *et al.* in press a). This relative lack of differentiation between these two forms at the nuclear level could indeed agree with the scenario proposed by Renoult (2006). Similar scenarios of mtDNA persistence through interspecific capture and nuclear dilution have in fact been shown to occur in several other cases (e.g. Alves *et al.* 2003, Martínez-Solano *et al.* 2004). However, the directionality inferred for gene flow among these forms does not suggest that *P. hispanica* sensu stricto lost its nuclear background to *P. hispanica* type 3, as suggested by Renoult (2006), but rather the opposite. On the other hand, the same scenario of mtDNA capture and nuclear introgression does not seem to hold regarding comparisons between *P. hispanica* sensu stricto and *P. hispanica* Galera type. Although significant levels of gene flow were inferred between these two forms, they do not seem to have been large enough for the two forms to stand out as undifferentiated. The same situation is suggested by analyses of allozyme variation, which suggest high levels of differentiation and a complete lack of gene flow between nearby interspecific populations of these two taxa. Because data with respect to the nuclear differentiation of mitochondrial DNA lineages from southeastern Iberian Peninsula thus remains controversial, this subject clearly needs further assessment.

Implications for the origin and evolution of Iberian and North African Podarcis

A primary observation that can be drawn from this work is that single nuclear genealogies do not reflect a branching pattern related to putative isolation in allopatry that led to the formation of the observed species, mainly because of shared ancestral polymorphism. The only clearly observable pattern is the isolation of this group of species from other *Podarcis*, represented here by *P. muralis*, suggested by the highly differentiated alleles detected in this species. The nuclear genes surveyed are therefore probably tracking the evolution of the genus with a significant delay. The observed lack of population structure furthermore implies that this clade of *Podarcis* was able to maintain a relatively high overall effective population size throughout its presumably eventful history that allowed the fixation (and consequent monophyly) of single mtDNA lineages but not the corresponding pattern in nuclear genes.

Pinho *et al.* (in press a) suggested that contrasting evolutionary scenarios inferred for the evolution of the genus on the basis of mitochondrial vs. nuclear markers (i.e. a step-by-step speciation scenario vs. a rapid simultaneous diversification) could be accommodated taking into account the different effective population sizes that characterize both classes of markers. Indeed, speciation may have occurred at a rate fast enough so that nuclear gene genealogies are uninformative regarding species relationships but mitochondrial

genealogies are not. In fact, although multi-species clades in Pinho *et al.*'s (2006) robust mitochondrial phylogenetic assessment of relationships are supported by high bootstrap values, internal branches are short, which is consistent with a fast speciation rate. Accordingly, our estimates of divergence time between mtDNA-defined lineages on the basis of nuclear markers do not correlate to mitochondrial DNA branching patterns; for example, the largest time since divergence was inferred for the mitochondrial sister pair *P. carbonelli*/*P. hispanica* type 2.

An intriguing characteristic of both nuclear gene genealogies is the detection of a skew towards rare alleles, as suggested by largely negative and significant Tajima's D. Although such a multilocus pattern can usually be interpreted as a signature of demographic growth, it has been shown that the same patterns may also arise as a consequence of fine-scale population structure, which is translated as a bias into sampling schemes such as ours, in which few individuals from multiple populations are analysed (Ptak and Przeworski 2002).

Concluding remarks and future prospects

With the analyses of nuclear genealogies in *Podarcis* wall lizards from the Iberian Peninsula and North Africa, we have provided compelling evidence for general historical isolation among distinct lineages despite their non reciprocal monophyly. These results are therefore in accordance with emergent morphological data and previous reports of mitochondrial and nuclear gene variation suggesting that these forms may correspond to different, almost completely reproductively isolated species. We have also provided evidence for the existence of limited to important levels of gene flow among a few species pairs, also in strong accordance with previously published data. However, because we only sampled 10-16 chromosomes per species, the patterns described are clearly the ones that have had the most obviously detectable effects in each species overall genetic diversity. Therefore, the inference that historical admixture between species has been generally close to zero does not necessarily mean that the species are completely reproductively isolated. Preliminary work on the contact zone between *P. bocagei* and *P. carbonelli*, for example, suggests spatially-restricted gene flow and a bimodal hybrid zone concordant with the existence of strong barriers to gene flow (Pinho *et al.* submitted). A necessary future step in the evaluation of the degree of gene flow among species will be therefore the detection and thorough characterization of other contact zones. *Podarcis* wall lizards in the Iberian Peninsula and North Africa present a wide variety of distribution patterns, including sympatric, parapatric and allopatric forms and will therefore constitute an excellent model for testing hypotheses related to the origin and maintenance of reproductive isolation in closely related species.

Acknowledgements

This work was partially financed by FCT research projects POCTI/40987/BSE/2001 and POCI/BIA-BDE/60911/2004. CP was supported by a PhD grant (SFRH/BD/4620/2001) and DJH was funded by a post-doctoral grant (SFRH/BPD/5702/2001), all from FCT. The authors would like to thank all the people that participated in sampling campaigns or provided samples for this work (Miguel A. Carretero, Antigoni Kaliontzopoulou, Bruno Ribeiro, Diana Barbosa, Elsa Froufe, Jean-Christophe Mace, Mohammed Znari, Paulo Sá-Sousa, Pedro Galán, Pierre-André Crochet, Philippe Geniez, Raquel Godinho, Vasco Batista and Xavier Santos). We would also like to thank Toni Gabaldón for help in confirming the identity of the *6-Pgdint7* gene region and Xana Sá-Pinto, Sara Rocha, Fernando Sequeira, Armando Geraldes, Miguel A. Carretero, Pierre-André Crochet and Craig Moritz for valuable suggestions on data analyses and discussion.

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.

Arévalo E, Davis SK, Sites Jr JW (1994) Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. *Systematic Biol*ogy, **43,** 387–418.

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,** 527–545.

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

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution,* **16,** 37–48.

Barbosa D, Font E, Desfilis E, Carretero MA (2006) Chemically mediated species recognition in closely related *Podarcis* wall lizards. *Journal of Chemical Ecology,* **32,** 1587–1598.

Buckley TR, Cordeiro M, Marshall DC, Simon C (2006) Differentiation between hypotheses of lineage sorting and introgession in New Zealand alpine cicadas (*Maoricicada* Dugdale). *Systematic Biology,* **55**, 411 – 425.

Bull V, Beltrán M, Jiggins CD, McMillan WO, Bermingham E, Mallet J (2006) Polyphyly and gene flow between non-sibling *Heliconius* species. *BMC Biology,* **4,** 11.

Bruen TC, Philippe H, Bryant D (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics,* **172,** 2665–2681.

Chan KMA, Levin SA (2005) Leaky prezygotic isolation and porous genomes: rapid introgression of

maternally inherited DNA. *Evolution,* **59,** 720–729.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology,* **9,** 1657–1659.

DeSalle R, Giddings LV (1986) Discordance of nuclear and mitochondrial phylogenies in Hawaiian *Drosophila*. *Proceedings of the National Academy of Sciences, USA,* **83,** 6902–6906.

Dolman G, Moritz C (2006). A multilocus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). *Evolution,* **60**, 573–582.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution,* **54,** 1839–1854.

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.

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.

Flot J-F, Tillier A, Samadi S, Tillier S (2006) Phase determination from direct sequencing of length-variable DNA regions. *Molecular Ecology Notes,* **6,** 627–630.

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. Mémoire présenté pour l'obtention du diplôme de l'École Pratique des Hautes Etudes, Montpellier.

Geraldes A, Ferrand N, Nachman MW (2006) Contrasting patterns of introgression at X-linked loci across the hybrid zone between subspecies of the European rabbit (*Oryctolagus cuniculus*). *Genetics,* **173,** 919–933.

Godinho R, Crespo EG, Ferrand N, Harris DJ (2005) Phylogeny and evolution of the green lizards, *Lacerta* spp. (Squamata: Lacertidae) based on mitochondrial and nuclear DNA sequences. *Amphibia-Reptilia*, **26,** 271-286.

Griswold CK, Baker AJ (2002) Time to the most recent common ancestor and divergence times of populations of common chaffinches (*Fringilla coelebs*) in Europe and North Africa: insights into Pleistocene refugia and current levels of migration. *Evolution,* **56,** 143–153.

Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution,* **56,** 2383–2394.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series,* **41,** 95–98.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution,* **22,** 160–174.

Harris DJ (2002) Reassessment of comparative genetic distance in reptiles from the mitochondrial cytochrome *b* gene. *Herpetological Journal*, **12,** 85–86.

Harris DJ, Arnold EN (1999) Relationships and evolution of wall lizards *Podarcis* (Reptilia: Lacertidae) based on mitochondrial DNA sequences. *Copeia,* **1999,** 719–754.

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.

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.

Hey J, Machado CA (2003) The study of structured populations – new hope for a difficult and divided science. *Nature Reviews Genetics,* **4,** 535–543.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics,* **167,** 747–760.

Hoelzer GA (1997) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees revisited. *Evolution,* **51,** 622–626.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics,* **111,** 147–164.

Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics*, **132,** 583-589.

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.

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology,* **11,** 2623–2625.

Kronforst MR, Young LG, Blume LM, Gilbert LE (2006) Multilocus analyses of admixture and introgression among hybridizing *Heliconius* butterflies. *Evolution,* **60,** 1254–1268.

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.

Martínez-Solano I, Gonçalves HA, Arntzen JS, García-París M (2004) Phylogenetic relationships and biogeography of midwife toads (Discoglossidae: *Alytes*). *Journal of Biogeography,* **31,** 603–618.

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.

Nachman MW, Boyer SN, Aquadro CF (1994) Non-neutral evolution of the NADH dehydrogenase subunit 3 gene in mice. *Proceedings of the National Academy of Sciences USA,* **91,** 6364-6368.

Neigel JE, Avise JC (1993) Application of a random-walk model to geographical distribution of animal mitochondrial DNA variation. *Genetics,* **135,** 1209–1220.

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

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.

Palumbi SR (1995) Nucleic acids II: the polymerase chain reaction. In: *Molecular systematics* (Hillis D, Moritz C, eds.), Sinauer, Sunderland, MA, 205–247.

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

Pearse DE, Crandall KA (2004) Beyond F_{ST}: analysis of population genetic data for conservation. *Conservation Genetics,* **5,** 585–602.

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: *The Biology of Lacertid Lizards: Evolutionary and Ecological Perspectives* (Pérez-Mellado V, Riera N, Perera A, eds.). 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.

Pinho C, Harris DJ, Ferrand N (in press a) 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.* In press.

Pinho C, Harris DJ, Ferrand N (in press b) Contrasting patterns of population structure and historical demography in three western Mediterranean lizard species inferred from mtDNA variation. *Molecular Ecology*. In press.

Pinho C, Kaliontzopoulou A, Carretero MA, Harris DJ, Ferrand N (submitted) Genetic admixture between the Iberian endemic lizards *Podarcis bocagei* and *P. carbonelli*: evidence for limited natural hybridization and a bimodal hybrid zone.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networs. *Trends in Ecology and Evolution,* **16,** 37–45.

Prychitko TM, Moore WS (1997) The utility of DNA sequences of an intron from the beta-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). *Molecular Phylogenetics and Evolution*, **8,** 193–204.

Ptak SE, Przeworski M (2002) Evidence for population growth in humans is confounded by fine-scale population structure. *Trends in Genetics*, **18,** 559–563.

Renoult J (2006) *Histoire évolutive du Lézard ibérique* Podarcis hispanicus *(Squamata: Lacertidae): discordances entre morphologie, ADN nucléaire et ADN mitochondrial*. Master 2 Recherche. Université Montpellier II Sciences et Techniques du Languedoc.

Rozas J, Sánchez-Del Barrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics,* **19,** 2496–2497.

Sá-Sousa P (2001) A Controversa Sistemática das Lagartixas do género Podarcis WAGLER, 1830 (Sauria: Lacertidae) em Portugal. Dissertação de Doutoramento. Fac. Ciências, Univ. Lisboa.

Sá-Sousa P, Vicente L, Crespo EG (2002) Morphological variability of *Podarcis hispanica* (Sauria: Lacertidae) in Portugal. *Amphibia-Reptilia,* **23,** 55–70.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Press, New York.

Sequeira F, Ferrand N, Harris DJ (2006). Assessing the phylogenetic signal of the nuclear ß-Fibrinogen intron 7 in salamandrids (Amphibia: Salamandridae). *Amphibia-Reptilia,* **27,** 409–418.

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, USA,* **99,** 16122–16127.

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution,* **13,** 964–969.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics,* **129,** 555–562.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics,* **68,** 978–989.

Swofford DL (2002) *PAUP* ver 4.0.b10. Phylogenetic analysis using parsimony and other methods*. Sinauer Associates, Sunderland, Sunderland, MA.

Tajima F (1989) Statistical methods for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics,* **123,** 585–595.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology,* **7,** 256–276.

Won Y-J, Sivasundar A, Wang Y, Hey J (2005) On the origin of Lake Malawi cichlid species: a population genetic analysis of divergence. *Proceedings of the National Academy of Sciences, USA.* **102,** 6581–6585.

Won Y-J, Hey J (2005) Divergence population genetics of chimpanzees. *Molecular Biology and Evolution,* **22,** 297–307.

Wright S (1931) Evolution in Mendelian populations. *Genetics,* **28,** 114–138.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics,* **15,** 323–354.

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology,* **12,** 563–584.

Zwickl DJ (2006) *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion*. Ph.D. dissertation, The University of Texas at Austin.

Genetic admixture between the Iberian endemic lizards *Podarcis bocagei* and *P. carbonelli*: evidence for limited natural hybridization and a bimodal hybrid zone

CATARINA PINHO[1,2], ANTIGONI KALIONTZOPOULOU[1,3], MIGUEL A. CARRETERO[1], D. JAMES HARRIS[1,2] AND NUNO FERRAND[1,2]

[1]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal
[2]Departamento de Zoologia e Antropologia, Faculdade de Ciências da Universidade do Porto, 4099-002 Porto, Portugal.
[3] Departament de Biologia Animal (Vertebrats), Facultat de Biologia, Universitat de Barcelona. Avgda. Diagonal, 645, 08028 Barcelona, Spain

Abstract

When recently diverged taxa come into contact, the extent of introgression between them is directly related to the stage of differentiation that they have achieved. The study of contact zones is therefore essential to understand if differentiated taxa are reproductively isolated and, ultimately, to assess if they are likely to remain distinct. Recent work on differentiation within Iberian and North African wall lizards (*Podarcis*) has shown that multiple evolutionary units exist, diagnosable both by genetic markers and morphology, some of which have been recognised as full species. These studies also suggest that gene flow between distinct forms may occur. Therefore, we were interested in evaluating how species boundaries are maintained in the areas where they meet. In this work, we present data relative to the contact zone between *P. bocagei* and *P. carbonelli*. We studied a transect including the only known locality where these two lizards are known to occur in syntopy and analysed a battery of 15 unlinked nuclear genetic markers (11 allozymes, three microsatellites and one nuclear intron), mitochondrial DNA and morphology. We also conducted a preliminary analysis of fertility. Using model-based clustering approaches we show that the two species hybridize in the population where they contact, but evidences of introgression are low for nearby populations. Although a significant number of individuals show evidence of admixture, this hybrid zone is clearly bimodal, suggesting strong barriers to gene flow, which putative nature are discussed. Interestingly, morphological analyses revealed a complete lack of intermediate forms, even among individuals that are admixed genetically. Taken together,

these results constitute further evidence validating the species status of *P. bocagei* and *P. carbonelli*.

Keywords: *Podarcis bocagei, Podarcis carbonelli*, hybrid zone, allozymes, microsatellites, mtDNA, morphology, fertility, reproductive isolation, bimodality.

Introduction

Speciation is usually thought of as the development of reproductive isolation between diverging taxa. Early views on this subject (Dobzhansky 1937, Muller 1940, 1942) and recent empirical work (e.g. Presgraves *et al.* 2003) suggest that the achievement of reproductive isolation is essentially led by the acquisition of epistatic incompatibilities in genes responsible for ecological, physiological or behavioural differentiation. If only a few loci have developed such incompatibilities between diverging populations, then gene flow between these populations is extensive, except for these few loci. The more genes that are involved in functional divergence between populations, the more difficult it becomes for them to exchange genes, since different loci become co-adapted within these populations (e.g. Orr 1995), and the less-likely it is that populations will fuse in the future. At some point in this process (Wu (2001) calls it "the point of no return"), the populations will have diverged sufficiently so that they will not merge even if the initial barrier that lead to their separation is removed. The process of divergence continues thereon even in the presence of some degree of gene flow, until full reproductive isolation is achieved. Contact zones provide us with the unique opportunity of assessing if species are reproductively isolated and thus of understanding the stage that they have reached in the process of differentiation. When differentiated genetic entities come into contact, different situations may occur: in one extreme of the divergence process, they do not interbreed or, if they do, are unable to produce viable or fertile offspring and are thus considered to be fully reproductively isolated; in the other extreme, they may interbreed freely and eventually merge into a single unit. In between these two situations, they may often interbreed in a narrow hybrid zone, the width of which is dependent upon a balance between dispersal and selection (a tension hybrid zone, e.g. Barton 1983, Barton and Hewitt 1985), and maintain their genetic integrity over most of their distribution. By analysing the dynamics of contact zones between recently diverged taxa, we may therefore predict how the future of the two populations will be in terms of fusion or divergence.

Contact zones between forms of wall lizards in Iberian Peninsula and North Africa constitute one such opportunity. These animals' taxonomy has been a long standing matter of debate and is in the process of re-examination, mainly

drawing from recent studies on morphology and genetic variation. Analyses of mitochondrial DNA sequences (Harris and Sá-Sousa 2001, 2002; Harris *et al.* 2002; Pinho *et al*. 2006) suggest the existence of as many as 11 differentiated forms, that were suggested to deserve species status based on the high genetic distances observed between them. Four species are presently recognised: *P. bocagei, P. carbonelli, P. vaucheri* and *P. hispanica.* Although these studies are ongoing, *P. hispanica* is currently a paraphyletic assemblage of divergent monophyletic mtDNA lineages distinguished as "types" by Harris and Sá-Sousa (2001) and subsequent works. The distribution of species and forms within the whole complex does not seem to overlap extensively, except for the pairs *P. bocagei/P. hispanica* type 1 and *P. carbonelli*/*P. hispanica* type 2, which are essentially sympatric and even syntopic. Although distribution maps are still incomplete and most contact zones have not yet been described, the distribution of lineages is most likely parapatric. In general, these different partitions observed in mitochondrial DNA are also identifiable on the basis of multiple nuclear markers (Sá-Sousa *et al.* 2000, Pinho *et al.* 2003, 2004, in press a), but nuclear intron genealogies reveal that both fully recognized species and groups within *P. hispanica* are not reciprocally monophyletic, due to incomplete lineage sorting of ancestral polymorphism coupled with the existence of limited gene flow between forms (Pinho *et al.* submitted). It has therefore been suggested that forms of *Podarcis* are not completely reproductively isolated but the extent and dynamics of gene exchange between forms remains to be completely understood. As a result, the characterization of contact zones between the various forms is of utmost importance in assessing stages of differentiation and evaluating if and how species boundaries are maintained (Pinho *et al.* in press a).

One of the first contact zones to be identified was that between *P. bocagei* and *P. carbonelli* (Carretero *et al.* 2002). Both species are mostly ground-dwelling, exhibit similar ecological requirements, and reproduce at the same time of the year (Carretero *et al.* 2006, Carretero *et al.* unpublished data). Formerly regarded as conspecific because of their morphological and ecological similarities, these taxa are now considered to be separate species (Sá-Sousa and Harris 2002) and indeed are not sister taxa with respect to mtDNA (Harris and Sá-Sousa 2001, 2002, Pinho *et al.* 2006). Their distributions are parapatric, overlapping in a narrow zone, with only a few kilometres of width. Based on inferred biogeographic and phylogeographic patterns (Sá-Sousa 2001a, Pinho *et al.* in press b), it seems likely that *P. bocagei* and *P. carbonelli* have come into contact after a post-glacial expansion of their distributions, which implies a relatively recent origin for this contact. *P. bocagei* and *P. carbonelli* have been shown to hybridize in captivity (Galán 2002), but there is no further information on the sterility or fitness of their hybrids. Thus far, there are no field

observations suggesting that the two species could interbreed in natural conditions. Instead, an apparent lack of morphological intermediate individuals between both species has been reported (Carretero *et al.* 2002, Kaliontzopoulou 2004, Kaliontzopoulou *et al.* in press). It has also been demonstrated that males from both species are able to recognize conspecific or heterospecific females based on chemical cues (Barbosa *et al.* 2005), suggesting assortative mating.

In this work, we studied a battery of mitochondrial and nuclear markers in individuals coming from populations situated along a North-South transect that crosses the contact zone. We also investigated morphological and fertility aspects of individuals collected in the area. Our main goals were a) to determine if these species are able to exchange genes; in the case that they are, b) to assess if hybrids are viable and fertile; c) to investigate if there are barriers to gene flow or if hybrids are formed freely; d) to investigate the hybrids' morphological characteristics; e) to assess which factors might influence the dynamics of this contact zone; f) to preliminarily describe the geographical extent of the introgression and, ultimately, g) to clarify whether or not *P. bocagei* and *P. carbonelli* are good species.

Materials and Methods

Genetic analyses

Sampling. Samples were obtained along a North-South transect of the Portuguese coast, centered in the locality of Espinho, which is the only locality where the two species have been reported in strict syntopy (Carretero *et al.* 2002). For the genetic characterization we analysed a total of 146 individuals from five localities (Figure 1): Vairão (N=30), Madalena (N=22), Espinho (N=59), Esmoriz (N=18) and Aveiro (N=17). Individuals from Vairão and Madalena were morphologically identified as *Podarcis bocagei*, whereas individuals from Esmoriz and Aveiro were assigned to *P. carbonelli*. The sample from Espinho included only adult individuals that were examined by the authors immediately after capture and were assigned to one of the two species in question considering various empirical diagnostic characters that principally include size, colour pattern, and head dimensions and shape (Sá-Sousa 2000, 2001b). This classification was used throughout the study to *a priori* classify individuals into species. The sample included 26 individuals assigned to *P. bocagei* and 28 to *P. carbonelli*. Both sexes (18 females and 36 males) were represented. Five samples of unknown sex and specific origin, that were collected in the area by non-experts and consisted only of tail tissue, were also included. Samples consisted of tail muscle and were stored frozen at -80ºC and in 96% ethanol.

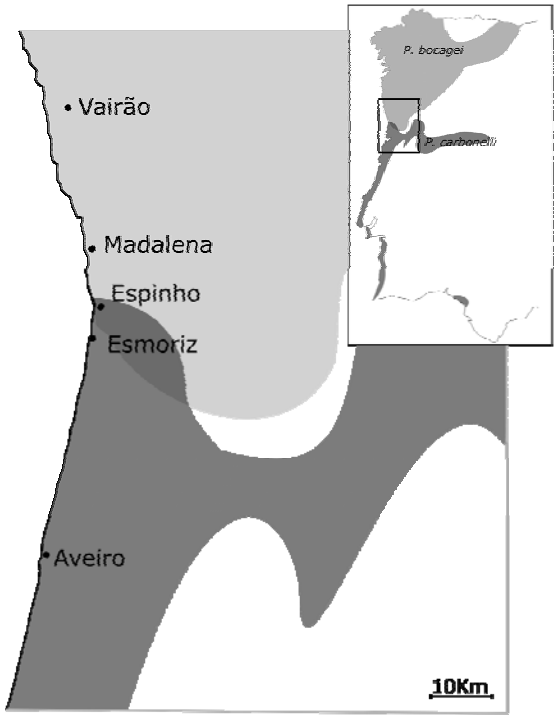

Figure 1. Distribution of *P. bocagei* and *P. carbonelli* and localities sampled for genetic analyses.

Data collection. Allozyme data from the populations of Vairão and Aveiro have been obtained previously and are included in publications (Pinho *et al.* 2003, 2004a, in press a). Tissue extraction, protein separation and enzymatic detection of all loci in the remaining populations followed the procedures described in these papers. Variation at 10 polymorphic loci was studied by means of allozyme electrophoresis (*PEPA, PEPD, MPI, IDH, 6-PGD, GOT, GPI*) and isoelectric focusing (*LDH-2, PGM, PEPB*). This set of markers was previously shown to allow the distinction between *P. bocagei* and *P. carbonelli* individuals in a Bayesian assignment framework (Pinho *et al.* in press a).

For microsatellite analyses, DNA was extracted following standard procedures (Sambrook *et al*. 1989). We used three from the battery of nine microsatellites developed for *P. bocagei* by Pinho *et al*. (2004b) (*Pb11, Pb50* and *Pb66),* which were amplified according to the described conditions with the exception of the annealing temperature, that was lowered to 53ºC in all cases in order to amplify difficult samples. The electrophoretic separation of the amplified fragments was carried out in 6% denaturing polyacrylamide gels and silver stained as described in Pinho *et al.* (2004b).

We also analysed genetic variation at a nuclear intron using Single Strand Conformation Polymorphism (SSCP). Pinho *et al.* (submitted) described a novel nuclear intron (intron 7 of the 6-phosphogluconate dehydrogenase (*6-Pgdint7*)

locus) that was found to be diagnostic between *P. bocagei* and *P. carbonelli*. We used specific primers *PgdA* (5'-GGAATTCCCCATCTCTGACTTAG-3') and *PgdB* (5'-GCATGCAAAGACAGGTTCTGGTG-3') to target a 160bp fragment containing 3 diagnostic SNPs, based on Pinho *et al*. (submitted) data. Polymerase chain reaction (PCR) was carried out in 10 µL volumes, containing 1.5 mM $MgCl_2$, 0.2mM each dNTP, 0.2µM each primer, 0.5U of *Ecotaq* DNA polymerase (Ecogen) and approximately 25 ng of genomic DNA. Amplification conditions were as follows: an initial denaturation step for 3 min at 94ºC, 40 cycles consisting on 30 sec at 94ºC, 20 sec at 53ºC and 20 sec at 72ºC, and a final extension of 72ºC for 2 min. Before loading PCR products into SSCP gels, they were diluted 1:5 in denaturing loading buffer (95% deionized formamide, 10mM NaOH, 0.01% bromophenol blue), denaturated for 5 min at 94ºC and kept on ice until loading. Preliminary SSCP tests were performed with samples known to carry different sequences in order to optimize separation conditions. For the small fragment of 160bp, all *P. bocagei* sequenced by Pinho *et al.* (submitted) exhibit one allele, whereas *P. carbonelli* possesses four distinct alleles, two of which (differing in 1SNP and corresponding to alleles PR3 and PR4 in Pinho *et al.* (submitted)) were impossible to separate under all tested conditions. Optimal resolution for the separation of *P. bocagei* vs *P. carbonelli* alleles was obtained using 12% polyacrilamide gels (29:1 acrylamide:methylbisacrylamide) with 1x TBE on a vertical electrophoresis system, run at a constant voltage of 170V at 15ºC for 16 hours. Previously genotyped samples were always added to gels in order to ensure correct genotype scoring. Previously undetected alleles were PCR-amplified for the whole intron and sequenced following Pinho *et al.* (submitted).

As an addition to the study of nuclear markers, we also developed a protocol for discriminating the mitochondrial DNA lineage carried by each individual, consisting of a Restriction Fragment Length Polymorphism (RFLP) analysis. Based on previously published sequences of the 12S rRNA mitochondrial DNA gene (Harris and Sá-Sousa 2002, Harris *et al.* 2002, Pinho *et al.* 2006), we used BioEdit v. 7.0.5.2 (Hall 1999) to produce restriction maps and to select restriction enzymes that allow species distinction. Although a single enzyme would have sufficed, we selected two, *Mse* I and *Taq* I, to reduce the possibility of misclassification due to homoplasy. Each of these enzymes produces a single restriction profile for each species. PCR amplification followed the conditions described in Pinho *et al.* (2006). PCR products were digested with each of the enzymes separately and run on 3% high resolution agarose gels side by side with previously typed samples.

Analytical methods. Allelic frequencies were calculated directly from the observed genotypes. We used ARLEQUIN v 2.000 (Schneider *et al.* 2000) to

perform Analyses of Molecular Variance (Excoffier *et al.* 1992) to evaluate and compare the usefulness of the markers in discriminating between both species by computing the partition of the total genetic diversity that is observed between species, between populations within species and within populations. For this purpose we used data from the four populations outside the contact zone. The GENEPOP software (v. 3.1b; Raymond and Rousset 1995) probability test was used to determine whether populations were in Hardy-Weinberg and linkage equilibrium.

To analyse individual assignment based on multilocus genotypes, we followed three main directions. First, we used GENETIX v. 4.05.2 (Belkhir *et al*. 1996-2004) to perform a Factorial Correspondence Analysis. Second, we used STRUCTURE (Pritchard *et al.* 2000) to estimate the proportion of each individual's genome originating in each parental species. This software implements a Bayesian model-based clustering algorithm that creates population structure in our data set and attempts to find clusters of individuals that minimize Hardy-Weinberg and linkage disequilibrium. The parameter settings included the assumption of admixture and of independent allele frequencies. We forced the number of populations (K) to 2 in order to evaluate if individuals from the two species could be discriminated. In exploratory runs we did not provide the software with prior population information regarding the origin of the individuals, but in final runs individuals from the two populations located in the extremes of the transect (Vairão and Aveiro), assumed to represent pure *P. bocagei* and *P. carbonelli,* were used as a proxy for determining the degree of admixture of all the other individuals. STRUCTURE was run for 550,000 steps, of which the first 50,000 were discarded as burn-in and we conducted 5 independent replicates of the MCMC. We also performed analyses using the correlated allele frequencies model to test for robustness of our conclusions to the violation of prior assumptions. Thirdly, we used NEWHYBRIDS (Anderson and Thompson 2002), which calculates the posterior probability, under a model-based Bayesian approach, of multilocus individual genotypes belonging to one of six different categories (pure of each parental type, F1 or F2 hybrids and F1-parental backcrosses on both directions). This software was run three times, each with 250,000 steps along the Markov Chain. Individuals with more than 20% of missing data were discarded in these analyses.

Morphological analyses

In order to examine the possible effect of hybridization on the morphology of the studied species, we typified morphologically individuals from the different populations studied. Morphological data were not available for all the populations examined genetically. From these, we included the populations of Vairão, representing "pure" *P. bocagei* to be used as the reference for this species, as

well as individuals of both species from Espinho. No morphological data were available for Aveiro or Esmoriz, the allopatric populations of *P. carbonelli* analysed genetically, but we used individuals from Torreira, a locality lying about 15km to the north of Aveiro. Our principal goal was to represent "pure" *P. carbonelli* morphology, and therefore using the same populations as reference for genetic and morphological analyses was not indispensable. Some of the individuals studied (those from Espinho) were previously used for a morphological study (Kaliontzopoulou *et al*. in press) and details on the populations' morphological properties, as well as the detailed protocols used for morphological characterization, can be found therein. In short, we used linear body measurements, as well as geometric morphometric methods, to record size and shape properties of the populations in question. Then, a discriminant and canonical variate analysis was applied to establish differences between the two species in a multivariate space. In continuation, discriminant functions produced were used to classify the individuals from Espinho into species, calculate posterior probabilities and examine the position of hybrids in the morphological space defined. For these analyses, individuals were considered to be hybrids when they showed posterior probabilities of assignment to a species (using STRUCTURE) lower than the lowest value observed in reference ("pure") populations (94.2%). For the individuals from the contact zone analysed both genetically and morphologically, we examined frequencies of coincidence in classification based on genetic and morphological methods. We correlated the attributed probabilities of specific identity to examine if individuals with a high probability of being hybrids were also erroneously classified based on morphological data. Due to a marked sexual dimorphism being present in both species (Kaliontzopoulou *et al.* in press) we examined both sexes separately.

Analysis of fertility

Adult specimens from Vairão, Espinho and Torreira analysed morphologically were sacrificed humanely and dissected for visual inspection of the reproductive organs (Carretero *et al.* unpublished). For those collected during the reproductive season (February-July), gonads were measured and weighed according to standard protocols (Carretero 2006). Furthermore, the density of spermatozoa in testis was estimated following Carretero *et al.* (2006) and vitellogenic follicles and oviductal eggs were also counted. Comparisons of gonad size and number of spermatozoa, follicles and egg between "pure" specimens and putative hybrids were performed through ANCOVA. Since such variables were dependent on body size we used SVL (snout-vent length) as a covariate. Because the slopes between SVL and those variables were not homogeneous between species (Carretero *et al.* 2006), we used a model for separate slopes.

All variables were log-transformed prior to the analyses to ensure normality and homoscedasticity.

Results

Genetic polymorphism outside the contact zone and distinction between species

Allozymes. A total of 37 alleles were detected amongst the four populations outside the contact zone. AMOVA analyses showed that using this data set differentiation between the two species accounts for 45.28% of the total genetic variation (table 1). The most informative loci in discriminating between the two species were *GOT* and *PGD*. Although data from previous studies (Pinho *et al.* 2003, 2004a, in press a) suggest that allele *GOT*2* is fixed in northern *P. carbonelli* populations, the population of Esmoriz shows a frequency of 22% of allele *GOT*1*, the most frequent in *P. bocagei* and fixed in Madalena. In *PGD*, the most frequent allele in *P. carbonelli* is *PGD*1*, which was absent from all *P. bocagei* populations studied to date but that is present at a very low frequency (2%) in Madalena. Other than these cases, both populations (Esmoriz and Madalena) are typical of each species for the allozyme genotypes observed.

Microsatellites. Cross-amplification of the microsatellites developed for *P. bocagei* was successful in *P. carbonelli*. All three microsatellites were highly polymorphic and showed mostly overlapping allele sizes between both species. Sixteen alleles were found in *Pb11*, 24 in *Pb50* and 19 in *Pb66*. Microsatellite loci appear not to be as good discriminators between *P. bocagei* and *P. carbonelli* as the other markers, since only 12.53% of the variation is explained by between species differentiation, but this is most likely due to the huge amount of within-population variability (Hedrick 1999), probably coupled with some degree of homoplasy (as e.g. in Queney *et al.* 2001). Nevertheless, the combination of the three markers performs well in discriminating between the two species (using STRUCTURE, results not shown). In particular, *Pb66* showed high frequencies of diagnostic alleles, which is reflected in a higher than average 24.35% of variation found between species. Alleles with size 130 and 140 are highly frequent in *P. carbonelli*, and absent from *P. bocagei*. Other alleles that appear in *P. carbonelli* are a small subset of those detected in *P. bocagei*.

6-Pgdint7 SSCP genotyping. As expected from earlier sequence data, this marker was fully diagnostic between *P. bocagei* and *P. carbonelli*. Only one allele was found in *P. bocagei* populations outside the contact zone, whereas two of the previously detected *P. carbonelli* alleles, corresponding to alleles PR3/PR4

and PR7 defined by Pinho *et al.* (submitted) were found in *P. carbonelli* populations. Allele PR3/PR4 was the most abundant, found at an average frequency of 92.9%. Because of its high level of diagnosis, over 93% of the variation at this locus is explained by differences between the two species.

Mitochondrial DNA RFLP analyses. Only the *P. bocagei* mtDNA type was found in populations situated north of the contact zone, whereas the *P. carbonelli* haplotype was fixed in the southern populations.

Table 1. Utility of the nuclear markers analysed in the distinction between *P. bocagei* and *P. carbonelli*, as shown by Hierarchical Analyses of Molecular Variance (AMOVA) using populations outside the contact zone

	Percentage of variation		
	Among species	**Within species, among populations**	**Within populations**
Allozymes	45.28	3.70	51.03
GOT	87.49	2.01	10.50
GPI	1.50	-2.31	100.81
IDH	4.91	1.75	93.33
LDH-2	-4.77	10.62	94.15
MPI	10.23	-1.07	90.84
PEPA	47.37	5.35	47.28
PEPB	-0.49	11.21	89.27
PEPD	30.45	1.50	68.06
PGD	66.37	6.34	27.29
PGM	0.39	3.23	96.38
Microsatellites	12.53	3.57	83.90
Pb11	9.05	3.02	87.92
Pb50	3.21	5.82	90.97
Pb66	24.35	1.92	73.72
6-Pgdint7	93.43	0.22	6.36
Total	38.10	3.31	58.59

Genetic composition of the contact zone

In the locality where the two morphotypes contact, Espinho, both mtDNA types were detected. In the allozyme loci, six alleles not detected in the other surveyed populations were observed, one of which in locus *LDH-1*, in which polymorphism had not been previously observed (Pinho *et al.* 2003). Genetic variants not found in the remaining populations of the transect were also detected in the microsatellites (six alleles) and in the nuclear intron (one new allele; after sequencing, it was shown to be allele PR8 from Pinho *et al.* (submitted), which had previously not been found in *P. bocagei* or *P. carbonelli* but was detected in *P. hispanica* type 1A). In the nuclear markers, in loci exhibiting contrasting allele frequencies between both species, intermediate frequencies were observed (results not shown). Individuals simultaneously

carrying alleles that were diagnostic for both species were detected, suggesting hybridization.

Hardy-Weinberg and linkage equilibrium

Most loci did not show significant ($p<0.05$) deviations to the Hardy-Weinberg expectations outside the contact zone. There were a few exceptions: *PEPD* in Vairão (already reported in previous studies; Pinho *et al.* 2003, 2004a, in press a) and microsatellite *Pb50* in the same locality. However, in the population of Espinho, five significant cases of disequilibria were detected (*GOT, PEPD, Pb50*, *Pb66* and *6-Pgdint7*). Linkage disequilibrium (LD) follows the same pattern: outside the contact zone there are no significant cases of association between different loci. This is true even for the allozyme *PGD,* which is encoded by the *6-PGD* gene, and the intron 6-*Pgdint7*. This probably happens because of recombination outside the intron, leading to a lack of association between the allozyme and the intron genotype. Because of this, we were able to assume both loci as independent in subsequent analyses. In the contact zone LD was detected between several pairs of loci ($p<0.05$): *PEPA* and *PEPD*, *PEPA* and *GOT*, *PEPD* and *GOT*, *IDH* and *GOT*, *PEPA* and *PGD*, *PEPD* and *PGD*, *GOT* and *PGD*, *PEPA* and *LDH1*, *GOT* and *LDH1*, *GOT* and *LDH2*, *PEPA* and *Pb50*, *PEPD* and *Pb50*, *GOT* and *Pb50*, *PGD* and *Pb50*, *GOT* and *Pb66*, *PEPA* and *6-Pgdint7*, *PEPD* and *6-Pgdint7, IDH* and *6-Pgdint7, GOT* and *6-Pgdint7, PGD* and *6-Pgdint7,* and *Pb50* and *6-Pgdint7.*

Individual multilocus genotype analyses

Factorial Correspondence Analyses results show that the populations from outside the contact zone are clearly separable (Figure 2). However, some individuals from Espinho are placed in between these two, suggesting an admixed origin. This was confirmed in the analyses using STRUCTURE and NEWHYBRIDS (Figure 3). Initial runs not using any prior information clearly identified two sets of individuals corresponding to the two species, suggesting that the used set of markers performs well in their separation. Using prior population information by marking individuals from Aveiro and Vairão as pure of each species, a high proportion of the genome of the majority of individuals from Madalena and Esmoriz is assigned to *P. bocagei* and *P. carbonelli*, respectively, suggesting that in general these represent "pure" individuals of each species. However, using STRUCTURE, one individual from Esmoriz shows a low but non-neglectable portion of its genome assigned to *P. bocagei*. The same individual fails to be classified as "pure" with high posterior probability using NEWHYBRIDS. This is probably due to the presence of alleles *GOT*1* and *Pb66*132* in this individual, which were not detected in nearby populations of *P. carbonelli* and seem to be characteristic of *P. bocagei*.

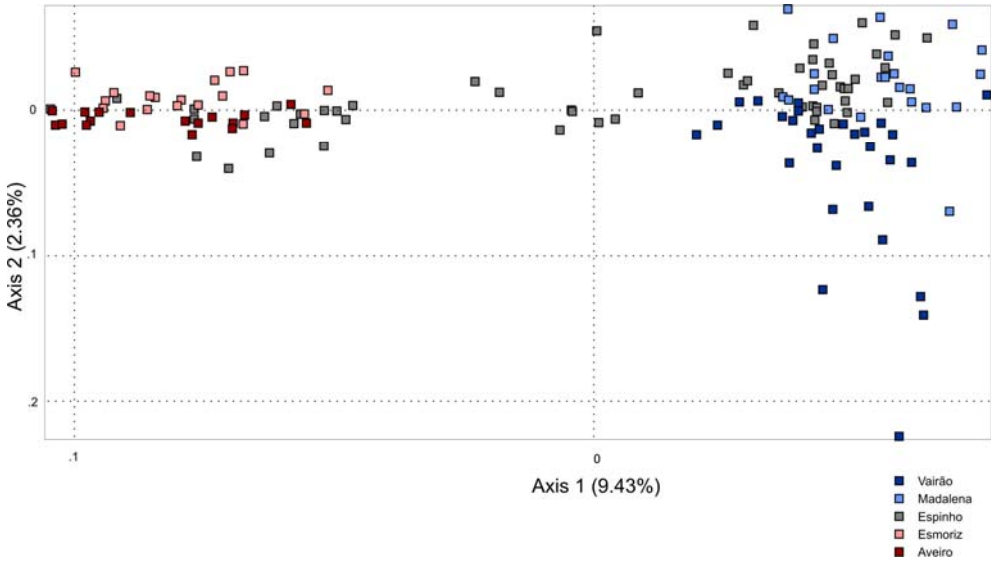

Figure 2. Results of the factorial correspondence analyses performed using GENETIX. Each square represents an individual.

Other individuals from Esmoriz also present these alleles, but not simultaneously. Unlike the general trend outside the contact zone, in the population of Espinho many individuals show signs of admixture. Few individuals, however, have similar proportions of their genome originating in both species; the most common admixed individuals, although being obviously "not pure", show a high proportion of the genome assigned to one of the species (Figure 4). The hybrid individuals identified using STRUCTURE are rarely straightforwardly assigned to a category by NEWHYBRIDS; usually the posterior probability of assignment is scattered throughout the four hybrid classes, either because they correspond to a multiple generation backcross or because the markers do not provide sufficient resolution. Five individuals show over 75% posterior probability of representing a backcross of a F1 with a pure *P. bocagei* and two with *P. carbonelli*. Clear first generation hybrids were not detected in the sample, although three individuals show over 50% posterior probability of being one. Information regarding the mtDNA type carried by each individual is shown in figure 3.

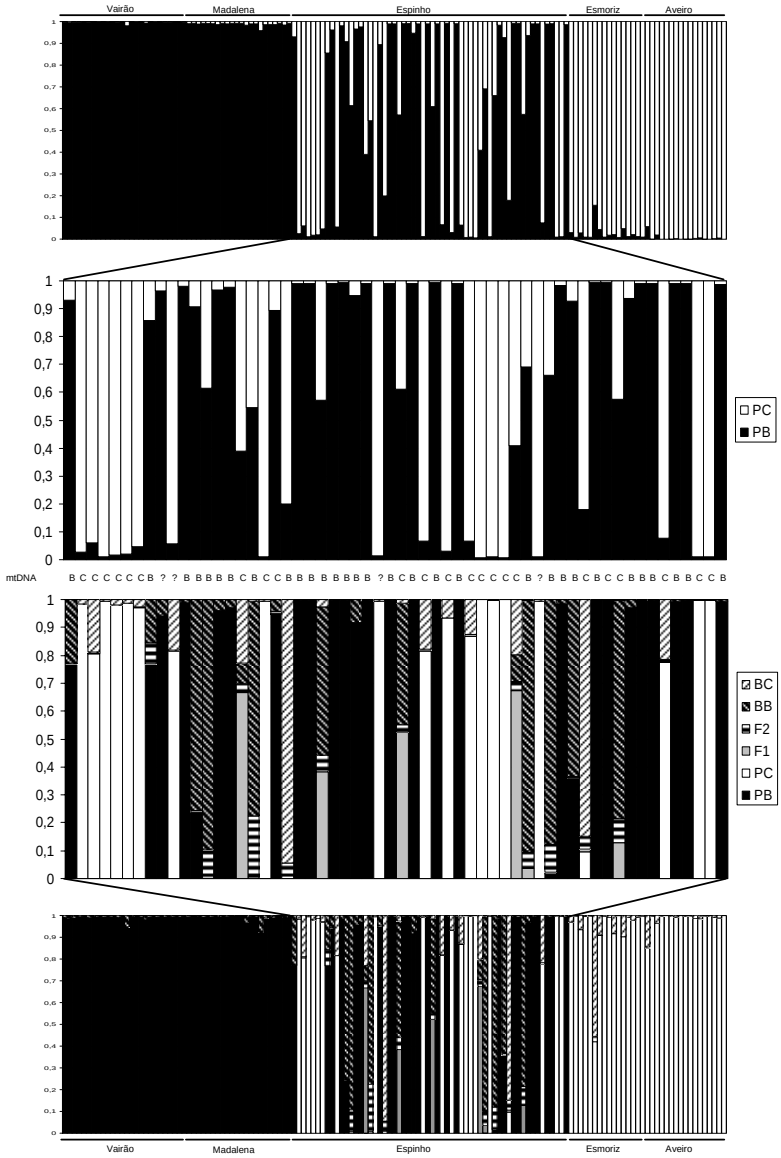

Figure 3. Model-based multilocus genotype analyses and mitochondrial DNA results. Upper graphs correspond to the proportion of the genome of each individual originating in each of the two species, estimated using STRUCTURE under a model with admixture. Each individual is represented by a vertical bar divided in two segments, which length is proportional to the estimated proportion of the genome originating in *P. bocagei* (PB) or *P. carbonelli* (PC). On the bottom, results obtained using NEWHYBRIDS are shown. Again, each individual corresponds to a bar divided in six portions, each proportional to the estimated posterior probabilities of being either a pure *P. bocagei* (PB), pure *P. carbonelli* (PC), F1, F2, and backcross of a F1 hybrid with a pure *P. bocagei* (BB) or with a pure *P. carbonelli* (BC). The graphs shown in the centre of the figure highlight the results obtained in the contact zone. Letters B or C presented between these graphs correspond to the mitotype (*P. bocagei* and *P. carbonelli,* respectively) detected in the individuals. ? means that no information is available for the mtDNA type presented by the individual.

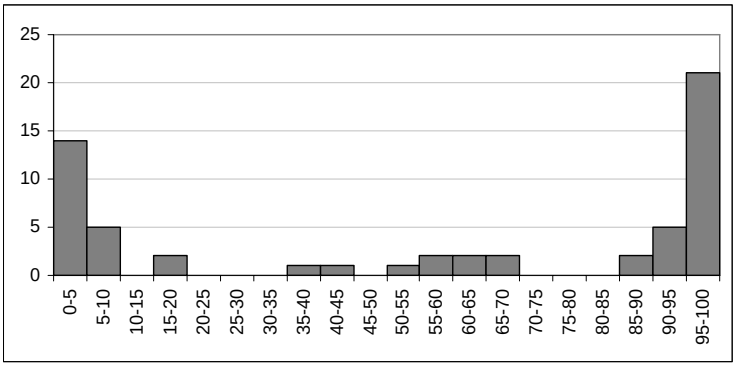

Figure 4. Genetic bimodality within the hybrid zone. The X axis of the histogram represents the proportion of the genome attributed to *P. bocagei* estimated using STRUCTURE and is divided in classes of 5%. The Y axis corresponds to the absolute number of individuals falling in each class.

Comparison with the morphology

Discriminant analyses based on linear and shape variables resulted in highly correct, but not perfect, classifications of individuals, with a notably higher discriminatory capacity of geometric morphometrics methods over those based on linear body measurements (Table 2). Using linear measurement methods, frequencies of correct classification were higher for individuals from allopatric populations than for those from the contact zone, but there is no evidence of hybrids being more frequently classified erroneously than genetically pure individuals.

Table 2. Classification matrix obtained by discriminant analysis using morphological data (linear and geometric morphometrics). PB: *Podarcis bocagei*, PC: *Podarcis carbonelli*. Numbers in parentheses indicate classification of individuals from the hybrid zone only.

		Males			Females		
		% correct	PB	PC	% correct	PB	PC
Linear measurements	**PB**	89.74 (83.33)	35 (15)	4 (3)	85.29 (78.95)	29 (15)	5 (4)
	PC	87.50 (85.00)	5 (3)	35 (17)	90.00 (90.00)	4 (2)	36 (8)
	Total	88.61 (84.21)	40 (18)	39 (20)	87.84 (84.62)	33 (17)	41 (22)
GM Dorsal	**PB**	100.00 (100.00)	41 (20)	0 (0)	100.00 (100.00)	36 (20)	0 (0)
	PC	97.50 (100.00)	1 (0)	39 (20)	100.00 (100.00)	0 (0)	40 (20)
	Total	98.77 (100.00)	42 (20)	39 (20)	100.00 (100.00)	36 (20)	40 (20)

In fact, both potential hybrids and pure individuals (figure 5A) were in some cases erroneously classified using linear measurements. Therefore, erroneous classifications should be attributed to the high intraspecific morphological

variability, usual in members of the genus *Podarcis* (Arnold 1973), and the high level of morphological similarity between the two species examined (Sá-Sousa 2001b, Sá-Sousa and Harris 2002, Kaliontzopoulou 2004, Kaliontzopoulou *et al.* 2005), rather than to the presence of morphological intermediates. It is interesting to note that using geometric morphometric data for dorsal head shape an almost perfect classification of individuals was obtained, with the exception of one male *P. carbonelli*, which was erroneously classified as *P. bocagei*. Nevertheless, this individual came from Torreira, i.e. the allopatric population for *P. carbonelli*. The results obtained by examining dorsal head shape with geometric morphometric methods indicate that hybrids are always classified as either *P. carbonelli* or *P. bocagei*, with a complete absence of morphological intermediate individuals. As a result, correlation values between the individuals' posterior probabilities of species identity evaluated by genetic and morphometric methods were not significant for any of the species or methods used (Table 4).

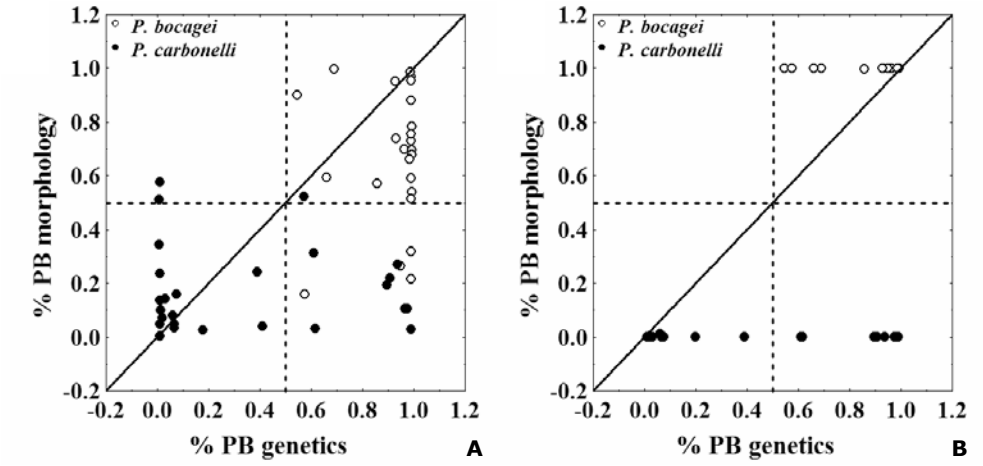

Figure 5. Correspondence between genetic and morphological assignments. In both graphs, the X axis represents the proportion of the genome of individuals that was attributed to *P. bocagei* according to STRUCTURE. The Y axis represents the probabilities of assignment to *P. bocagei* by discriminant analysis using linear biometric variables (**A**) and using dorsal head shape variables provided be geometric morphometrics (**B**). The continuous line represents a slope of one for the relation between the two probabilities and the dotted lines indicate assignment probabilities of 0.5. Blank and filled circles indicate the species to which individuals were assigned by visual inspection.

Fertility

Data were available for 49 lizards from the contact zone (6 male and 9 female "pure" *P. carbonelli*; 13 male and 3 female "pure" *P. bocagei* and 14 male and 4 female putative hybrids) as well as for 86 *P. bocagei* (51 males and 35 females) from Vairão and 71 *P. carbonelli* (37 males and 34 females) from Torreira. None

of them displayed any apparent abnormality in the sexual organs. In the contact zone, male hybrids had similar testis relative to their body size to "pure" *P. carbonelli* but bigger than *P. bocagei* (volume ANCOVA, $F_{3,27}$= 21.17, $P < 10^{-6}$, Scheffé's post-hoc test $p < 0.02$; weight ANCOVA, $F_{3,27}$= 21.14, $P < 10^{-6}$, Scheffé's post-hoc test $p < 0.03$). The number of spermatozoa also displayed variation but it could not be attributed to any specific pairwise comparison (ANCOVA, $F_{3,27}$= 9.69, $P = 0.0002$, Scheffé's post-hoc test $p > 0.32$). However, both *P. carbonelli* and *P. bocagei* carried more spermatozoa in the contact zone than in the allopatric populations (Carretero *et al.* 2006). Regarding females, no differences between hybrids and "pure" specimens (allopatric or sympatric) were recorded (ovary weight, number of follicles and eggs, ANCOVAs $p > 0.15$).

Table 3. Absolute frequencies of possible hybrids suggested by genetic analyses erroneously or correctly classified by morphometric methods (HB_EC and HB_CC respectively), as well as non-hybrid individuals erroneously classified (NOT HB_EC), coming from allopatric populations and the hybrid zone. N: total sample size, Ng: number of individuals also genetically analyzed, Hybrids: total number of potential hybrids.

FEMALES	N	Ng	Erroneously classified	Potential hybrids	HB_EC	HB_CC	NOT HB_EC	
							Allopatry	Sympatry
Linear measurements								
P. bocagei	34	6	5	3	1	2	1	3
P. carbonelli	39	10	3	3	0	3	2	1
Total	73	16	8	6	1	5	3	4
Geometric morphometrics								
P. bocagei	36	5	0	3	0	3	0	0
P. carbonelli	40	8	0	3	0	3	0	0
Total	76	13	0	6	0	6	0	0
MALES	N	Ng	Erroneously classified	Potential hybrids	HB_EC	HB_CC	NOT HB_EC	
							Allopatry	Sympatry
Linear measurements								
P. bocagei	39	18	4	4	0	4	1	3
P. carbonelli	40	16	5	10	1	9	2	1
Total	79	34	9	14	1	13	3	4
Geometric morphometrics								
P. bocagei	41	15	0	4	0	4	0	0
P. carbonelli	39	13	1	8	0	8	1	0
Total	80	28	1	12	0	12	1	0

Table 4. Correlation matrices between individual probabilities of species identity from genetic and morphological data. Both Spearman Rank Order Correlation Coefficients and Kendall Correlation Probabilities and their respective *P*-values are presented.

Morphology	**Genetics**			
	Spearman R	***P*-value**	**Kendall Tau**	***P*-value**
P. bocagei				
Linear measurements	-0.003	0.990	-0.007	0.959
GM dorsal	0.229	0.331	0.151	0.351
P. carbonelli				
Linear measurements	-0.052	0.798	-0.043	0.753
GM dorsal	-0.101	0.664	-0.077	0.626

Discussion

Our preliminary results are unambiguous in showing that *Podarcis bocagei* and *Podarcis carbonelli* hybridize in the area where they contact, since several individuals from the contact zone show a multilocus genotype presenting signs of admixture between both species. By a simple observation of multilocus genotypes, we are also able to answer our second question: hybrids are not only viable but also show at least some degree of fertility, since successive generations of hybrid individuals were found in the contact zone.

Genetic and morphological bimodality within the hybrid zone
Although a large number of individuals seem to have a hybrid origin (Figure 3), more individuals than expected under a scenario of free admixture were found to be genetically pure. In addition, the majority of hybrid individuals are not intermediate in composition; instead, a paucity of F1s (and a complete lack of F2s) was described, and most individuals bearing signs of admixture present a large fraction of their genome assigned to one of the species (Figure 4). Both of these observations result in strong Hardy-Weinberg and linkage disequilibria. This contact zone therefore entirely fits the description of a bimodal hybrid zone (Jiggins and Mallet 2000). Although the excess of pure individuals could be due to extensive dispersal from neighbouring populations into the contact zone (Barton and Hewitt 1985), it seems more likely that it implies a reproductive barrier impeding frequent introgression between the two species.

Possible pre-zygotic isolating mechanisms. It has been suggested that pre-zygotic isolating mechanisms generally evolve faster than post-zygotic barriers (see example from *Drosophila* in Coyne and Orr 1997), although this remains controversial (see for a thorough review Coyne and Orr 2004). Many bimodal hybrid zones show strong prezygotic isolation as a result of assortative mating (Coyne and Orr 1997, Rieseberg *et al.* 1998, Mallet *et al.* 1998, Howard *et al.* 1998, Price and Bouvier 2002). In the contact zone, the two species occupy similar habitats in close contact and breed during the same period of the year (Carretero *et al.* 2002, 2006); in other words, there are no obvious ecological or temporal barriers preventing gene flow. It is therefore likely that the predicted segregation between *P. bocagei* and *P. carbonelli* results from ethological barriers. The dynamics of mate recognition between these two species have been studied from a chemosensory point of view (Barbosa *et al.* 2005). It has been shown that both *P. bocagei* and *P. carbonelli* males tend to show a higher rate of tongue-flicks when presented with chemical stimuli from conspecific females, which is usually a signature of a larger interest in pursuing those stimuli. It seems therefore likely that at least female chemical cues and male

chemosensory responses have coevolved in allopatry and interactions between these traits are now acting as barriers to gene flow.

These and other behavioural interactions may tend to impede initial heterospecific matings and be sufficient to maintain bimodality within the hybrid zone. However, the first hybrids are often the most difficult to produce and once this first barrier is overcome it is likely that backcrosses and F2s may hybridize more freely since behavioural incompatibilities tend to be attenuated (Coyne and Orr 2004, Mallet 2005). It is therefore reasonable to assume that the bimodal nature of this hybrid zone is also maintained by post-zygotic isolating mechanisms.

Putative post-zygotic isolating mechanisms. Possible selective forces operating against hybrids have not been studied in these species through controlled cross-species matings. In this context, we were interested in making a preliminary evaluation of patterns indicating decreased hybrid survival or reproductive success. First, we evaluated whether Haldane's (1922) rule was verified on this hybrid zone. This empirical rule states that, when there is absence, rarity or sterility of first generation hybrids of one sex in a population of hybridizing taxa, that sex is usually the heterogametic. As in most lacertids, females of *Podarcis* are the heterogametic sex (Olmo and Gelo Signorino 2003). If Haldane's rule was verified, we would expect to see 1) no females as F1 (in the case of female inviability) and/or 2) backcrosses to a parental species presenting mtDNA from that species (in the case of complete F1 female sterility, resulting from the mating of a male hybrid with a pure female and not the opposite). In fact, the three individuals with higher probability of being F1s are all males. However, one out of the five individuals with over 75% of being a backcross with *P. bocagei* has *P. carbonelli* mtDNA, whereas one out of two backcrosses with *P. carbonelli* have the other species' mitotype. We may therefore rule out the possibilities of complete female inviability or sterility, but the possibility that the survival and reproductive success of F1 females might be lower than those of F1 males remains.

Secondly, we compared fertility indexes in hybrids to those in pure individuals. Although sample size is a limiting factor, our study failed to detect a decrease in the fertility of the hybrids. Hybrid males have apparently normal gonads and their numbers of spermatozoa fall within the usual values of the species they were morphologically assigned to (Carretero *et al.* 2006). Although less represented in the sample, hybrid females also do not appear to show reduced ovary weights, number of follicles or eggs. Therefore, if selection is acting against hybrid fertility, it is probably doing so at later stages of the life cycle, such as embryo development.

We further explored the nature of post-zygotic isolating mechanisms by looking at hybrid morphology. Although intermediate individuals were detected by the genetic analyses, those individuals' morphology is clearly assigned to one species or the other, at least by one of the techniques applied (figure 5II), indicating that intermediate phenotypes are scarce in adults and resulting in a completely bimodal morphological pattern. We may think of two hypotheses to explain the observed lack of morphological intermediates: either 1) the inheritance of morphological features typical of both species is done in a completely linked fashion (i.e. there is Mendelian inheritance of head shape characteristics as a whole) and one of the morphologies is acting as dominant, which is rather unlikely given the usually complex genetic regulation of morphological characteristics; or 2) there is divergent selection acting against juvenile intermediate morphotypes; only those hybrids that resemble a pure individual are able to reach adulthood. Nevertheless, there is data suggesting that there might be different trends within different pairs of hybridizing *Podarcis* species with regard to morphology: in a hybrid zone between the Italian *P. sicula* and *P. tiliguerta*, hybrid individuals clearly exhibited intermediate morphotypes (Capula 2002), which does not seem to be the case in our data. An also interesting trend observed in our study is the existence of genetically pure *P. bocagei* individuals with *P. carbonelli* morphology but not the opposite. Given the present knowledge, the causes for this asymmetry remain obscure.

All these observations on mechanisms that might be acting to prevent gene flow between species are preliminary. Low sample sizes in comparisons between genetic and morphological and fertility data clearly restrict the validity of present assessments. Moreover, to correctly study and identify post-zygotic mechanisms impeding gene flow it is crucial to produce hybrid offspring in captivity and analyse their morphological and reproductive characteristics. Nevertheless, taking these data together, although the mechanisms that are preventing gene flow between both species are still not readily identified, it is clear that they exist. According to Jiggins and Mallet (2000), "(…) bimodality within a local population indicates that speciation of the parental forms is nearly complete." Therefore, although it is clear that *P. bocagei* and *P. carbonelli* exchange genes, they probably represent "good" species and it is unlikely that they will ever fuse in the future. Further examination of other Iberian *Podarcis* lineages, both to detect contact zones and to investigate levels of gene flow across them, is necessary to determine how many of the described mtDNA lineages may also deserve recognition as full species.

Geographic extension of the introgression

Our sampling scheme does not allow strongly supported observations regarding the geographic extent of the introgression between both species. Although nearby populations are nearly pure, suggesting that the hybrid zone is narrow, there is evidence of immigrant alleles outside the area where the two morphotypes overlap. This evidence consists on the presence of allele *GOT*1* and *Pb66*132*, characteristic of *P. bocagei*, in individuals from Esmoriz, around 5km to the south of the contact zone, and the presence of allele *PGD*1*, otherwise absent from *P. bocagei*, in the population of Madalena. Alternative hypotheses for the presence of these alleles would be the persistence of ancestral polymorphism or homoplasy. However, given that they were strictly found in the populations near the limits of the other species' distribution, the hypothesis of gene flow seems rather likely. Why, then, was the introgression of these particular alleles not accompanied by that of others from the species' gene pool? Cline theory predicts that when there is strong selection acting on hybrid genotypes, most clines will be concordant and different loci act as if they were tightly linked (Barton 1983, Baird 1995). Otherwise, stochasticity will allow distinct loci to produce different clines. In the populations surrounding the contact zone, most of the alleles found are typical from the respective species, and therefore clines are apparently concordant (but denser sampling between the center of the contact zone and our neighbouring populations would be needed before assessing this with certainty). Foreign alleles that are found outside the contact zone have been allowed to break apart from their genomic background and are introgressing more readily than others, suggesting that they may be advantageous alleles under positive selection or tightly linked to loci that are (Piálek and Barton 1997). However, this could also be an effect of stochasticity. A more dense and widespread sampling design would be desirable to investigate cline concordance, determine the width of this hybrid zone and identify other loci under the putative effects of positive selection.

Concluding remarks

With this study we have demonstrated that *P. bocagei* and *P. carbonelli*, two wall lizard species that have been estimated to have diverged around 5 million years ago, are still able to exchange genes. The bimodal nature of this hybrid zone clearly indicates the existence of strong barriers to gene flow. Although at this stage we are still not able of correctly identifying the mechanisms that are involved in keeping these species apart, we have described interesting patterns suggesting divergent selection acting on intermediate morphologies.

Although this is the first analysis of a contact zone in the genus, these results do not constitute the first evidence of natural hybridization between species of the genus *Podarcis*. In fact, species of this genus seem particularly prone to

hybridize with congenerics. Capula (1993, 2002) reported natural hybridization between Italian species *P. sicula* and both *P. wagleriana* and *P. tiliguerta*; Pinho *et al*. (in press a) found individuals evidencing admixture from *P. bocagei* and *P. hispanica* type 1; Pinho *et al.* (submitted) reported clear evidences of introgression between several other pairs of Iberian lineages, including the distantly related *P. muralis* and *P. hispanica* type 3. Arnold (1973) and Bischoff (1973) also reported the existence of hybrids between other species pairs, but these were not confirmed using genetic tools. Considering only fully recognized species (and thus assuming all forms of the polytypic *P. hispanica* as one) and hybridization situations confirmed through genetics, this means that at least seven species of *Podarcis* are able to hybridize with congenerics. Presently, 18 species of *Podarcis* are recognized, which yields a percentage of 38.9% that are capable of hybridization. This is clearly an underestimation since most contact zones and areas of sympatry have not been surveyed using multiple genetic markers and many other species are isolated island endemics. Mallet (2005) estimated that around 10% of animal species may be involved in hybridization with closely related taxa. This suggests that wall lizards are way above the average, showing values similar to those observed in well studied groups such as *Heliconius* butterflies (Mallet *et al*. 1998), birds of paradise, American warblers and tits (see Mallet 2005 and references therein). By presenting a mosaic of distributions and varying patterns of geographical contact (from small areas such as the one studied herein to almost complete sympatry), *Podarcis* wall lizards thus provide good models to study the role played by natural selection in promoting speciation despite ongoing gene flow.

Acknowledgements

This work was partially financed by FCT research projects POCTI/40987/BSE/2001 and POCTI/BSE/45664/2002. CP was supported by a PhD grant (SFRH/BD/4620/2001), AK was supported by the project POCI/BIA-BDE/55865/2004, MAC and DJH were funded by post-doctoral grants (SFRH/BPD/3596/2000 and SFRH/BPD/5702/2001, respectively), all from FCT. Collecting permits were provided by Instituto de Conservação da Natureza (ICN, Portugal). The authors would like to thank all the people that participated in sampling campaigns (Raquel Ribeiro, Bruno Ribeiro, Diana Barbosa, Ricardo Pereira, Sara Rocha, Gonçalo Themudo, Raquel Godinho and Elsa Froufe). We would also like to thank Armando Loureiro for providing accurate distribution maps, Ettore Olmo for information on sexual chromosomes in *Podarcis*, and Rui Faria and Hélder Freitas for helpful advice regarding data analyses and discussion.

References

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics,* **160,** 1217–1229

Arnold EN (1973) Relationships of the Palaearctic lizards assigned to the genera *Lacerta, Algyroides* and *Psammodromus* (Reptilia, Lacertidae). *Bulletin of the British Museum of Natural History (Zoology)*, **25,** 291-366

Baird SJE (1995) A simulation study of multilocus clines. *Evolution,* **49,** 1038-1045.

Barbosa D, Desfils E, Carretero MA, Font E (2005) Chemical stimuli mediate species recognition in *Podarcis* wall lizards. *Amphibia-Reptilia*, **26,** 257–263.

Barton NH (1983). Multilocus clines. *Evolution*, **37,** 454–471.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16,** 113–148.

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F (1996-2004) GENETIX, logiciel sous WindowsTM pour la génétique des populations.

Bischoff W (1973) Lacertenbastarde II. *Zoologischer Garten NF, Berlin,* **43,** 278-291.

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.

Carretero MA (2006) Reproductive cycles in Mediterranean lacertids: Plasticity and constraints. In: *Mainland and insular lizards. A Mediterranean perspective* (Corti C, Lo Cascio P, Biaggini M, eds.). Firenze University Press. Florence. (in press)

Carretero MA, Sá-Sousa P, Barbosa D, Harris DJ, Pinho C (2002) Sintopía estricta entre *Podarcis bocagei* y *P. carbonelli*. Bolletin de la Asociasión Herpetológica Española, **13,** 20-24.

Carretero MA, Ribeiro R, Barbosa D, Sá-Sousa P, Harris DJ (2006) Spermatogenesis in two Iberian *Podarcis* lizards : relationships with male traits. *Animal Biology*, **56,** 1–12.

Coyne JA, Orr HA (1997) Patterns of speciation in *Drosophila* revisited. *Evolution,* **51,** 295–303.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, MA.

Dobzhansky T (1937) *Genetics and the Origin of Species.* Columbia Univ. Press, New York.

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.

Galán P (2002) Hibridación en laboratorio de *Podarcis bocagei* y *Podarcis carbonelli*. *Boletin de la Asociacion Herpetologica Española,* **13,** 28–31.

Haldane JBS (1922) Sex-ratio and unisexual sterility in hybrid animals. *Journal of Genetics,* **12,** 101-109.

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,** 1257-1252.

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.

Hedrick PW (1999). Highly variable loci and their interpretation in evolution and conservation. *Evolution,* **53,** 313–318.

Howard DJ, Gregory PG, Chu J, Cain ML (1998) Conspecific sperm precedence is an effective barrier to hybridization between closely related species. *Evolution,* **52,** 511–516

Jiggins C, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology and Evolution,* **15,** 250–255.

Kaliontzopoulou A (2004) *Efecto de la simpatría en la morfología de dos especies del género* Podarcis *en Portugal.* DEA Thesis, University of Barcelona.

Kaliontzopoulou A, Carretero MA, Llorente G (2005) Differences in the pholidotic patterns of *Podarcis bocagei* and *P. carbonelli* and its implications for species determination. *Revista Española de Herpetología,* **19,** 71-86.

Kaliontzopoulou A, Carretero MA, Llorente G (in press) Multivariate and geometric morphometrics in the analysis of sexual dimorphism variation in *Podarcis* Lizards. *Journal of Morphology*

Mallet J, McMillan WO, Jiggins CD (1998) Estimating the mating behavior of a pair of hybridizing *Heliconius* species in the wild. *Evolution,* **52,** 503-510.

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

Muller HJ (1940) Bearing of the *Drosophila* work on systematics. In: *The new systematics.* Clarendon press, Oxford, U.K., 185-268.

Muller HJ (1942) Isolating mechanisms, evolution and temperature. *Biological Symposium,* **6,** 71-125.

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

Olmo E, Gelo Signorino G (2003) Chromorep database, http://193.206.118.100/professori/chromorep.pdf

Orr HA (1995) The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics,* **139,** 1805–1813.

Piálek J, Barton NH (1997) The spread of an advantageous allele across a barrier: the effects of random drift and selection against heterozygotes. *Genetics*, **145,** 493–504.

Pinho C, Ferrand N, Harris DJ (2004a) Genetic variation within the *Podarcis hispanica* species complex: new evidence from protein electrophoretic data. In: *The Biology of Lacertid Lizards: Evolutionary and Ecological Perspectives* (Pérez-Mellado V, Riera N, Perera A, eds.). 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* spp.) from the Iberian Peninsula and North Africa. *Biochemical Genetics* 41, 343–359.

Pinho C, Harris DJ, Ferrand N. (in press a) 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,* in press.

Pinho C, Harris DJ, Ferrand N. (in press b) Contrasting patterns of population structure and historical demography in three western Mediterranean lizard species inferred from mtDNA variation. *Molecular Ecology,* in press.

Pinho C, Harris DJ, Ferrand N. (submitted) Abundant ancestral polymorphism at the nuclear loci βfibint7 and 6-PGDint7 characterizes different *Podarcis* species from the Iberian Peninsula and North Africa.

Pinho C, Sequeira F, Godinho R, Harris DJ, Ferrand N (2004b) Isolation and characterization of 9 microsatellite loci in *Podarcis bocagei* (Squamata: Lacertidae). *Molecular Ecology Notes*, **4,** 286-288.

Presgraves DC, Bagalavan L, Abmayr SM, Orr HA (2003) Adaptive evolution drives divergence in a hybrid inviability gene between two species of *Drosophila. Nature,* **423,** 715–719.

Price TD, Bouvier MM (2002) The evolution of F1 postzygotic incompatibilities in birds. *Evolution,* **56,** 2083–2089

Pritchard JK, Stephens M, Donnelly P. (2000). Inference of population structure using multilocus genotype data. *Genetics,* **155,** 945–959.

Queney G, Ferrand N, Weiss S, Mougel F and Monnerot M (2001) Stationary distributions of microsatellite loci between divergent population groups of the European rabbit (*Oryctolagus cuniculus*). *Molecular Biology and Evolution*, **18,** 2169-2178.

Raymond M, Rousset F (1995) GENEPOP (version 3.1): population genetics software for exact tests and ecumenicism. *Journal of Heredity,* **86,** 248–249.

Rieseberg LH, Baird S, Desrochers A (1998) Patterns of mating in wild sunflower hybrid zones. *Evolution,* **52,** 713–726.

Sá-Sousa P (2001a) Comparative chorology between *Podarcis bocagei* and *P. carbonelli* (Sauria: Lacertidae) in Portugal. *Revista Española de Herpetología,* **15,** 85–97.

Sá-Sousa, P (2001b) A controversa sistemática das lagartixas do género *Podarcis* Wagler, 1830 (Sauria, Lacertidae) em Portugal. PhD Thesis, University of Lisbon.

Sá-Sousa P, Almeida AP, Rosa HD, Vicente L, Crespo EG (2000) Genetic and morphological relationships of the Berlenga wall lizard (*Podarcis bocagei berlengensis*: Lacertidae). *Journal of Zoological Systematics and Evolutionary Research*, **38,** 95-102

Sá-Sousa P, Harris DJ (2002). *Podarcis carbonelli* is a distinct species. *Amphibia-Reptilia,* **23,** 459–468.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor Press, New York.

Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2.000: A software for population genetics analyses. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Wu C-I (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology,* **14,** 851–861.

Chapter 3

Genetic variation within species of wall lizards: processes of population subdivision, historical demography and post-glacial dynamics in *P. bocagei, P. carbonelli* and *P. vaucheri.*

Article VI. Contrasting patterns of population subdivision and historical demography in three Western Mediterranean lizard species inferred from mitochondrial DNA variation.

CATARINA PINHO, D. JAMES HARRIS & NUNO FERRAND, 2007
Molecular Ecology, in press

Article VII. Isolation and characterization of 9 microsatellite loci in *Podarcis bocagei* (Squamata: Lacertidae)

CATARINA PINHO, FERNANDO SEQUEIRA, RAQUEL GODINHO, D. JAMES HARRIS & NUNO FERRAND, 2004
Molecular Ecology Notes, **4,** 286–288.

Article VIII. Combining mtDNA, allozyme polymorphisms and microsatellites to assess the evolutionary histories of two endemic lacertids (*Podarcis bocagei* and *P. carbonelli*) of Western Iberian Peninsula

CATARINA PINHO, D. JAMES HARRIS AND NUNO FERRAND
In preparation

Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation

CATARINA PINHO, D. JAMES HARRIS AND NUNO FERRAND

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

Abstract

Pleistocene climatic oscillations were a major force shaping genetic variability in many taxa. We analyse the relative effects of the Ice Ages across a latitudinal gradient in the Western Mediterranean region, testing two main predictions: 1) species with historical distributions in northern latitudes should have experienced greater loss of suitable habitat, resulting in higher extinction of historical lineages than species distributed in southern latitudes, where the effects of the Ice Ages were not as drastic. This would be reflected in the observation of lower diversity and number of differentiated lineages in northern areas. 2) a signature of demographic expansion following the climate amelioration should be obvious in northern species, whereas in the south evidence of long-term effective population size stability should be observed. We used as models three species of wall lizards (*Podarcis bocagei, P. carbonelli* and *P. vaucheri*) that replace each other along the study area. We investigated the patterns of mitochondrial DNA diversity and subdivision and obtained demographic parameter estimates for each species. Our results suggest that *P. bocagei*, the northernmost species, bears low genetic diversity, a shallow coalescent history and marks of a demographic expansion. In contrast, *P. vaucheri*, the species with a southernmost distribution, shows deeper coalescence events, complex geographic substructure and no evidence for population growth. The species with an intermediate distribution, *P. carbonelli*, shows average levels of diversity, substructure and population growth. Taken together, these results conform to our main predictions and are explained by a differential influence of the Ice Ages on distinct latitudes.

Keywords: *Podarcis*; glaciations; Iberian Peninsula; North Africa; phylogeography; population structure; demography; latitudinal gradient

Introduction

Pleistocene climatic oscillations are generally believed to have played a major role in shaping genetic diversity across a wide number of taxa (Hewitt 1996, 1999). In recent years, following the rise of phylogeography as a formal discipline (Avise *et al.* 1987, Avise 2000), a large number of surveys of genetic variation have disclosed patterns of genetic subdivision related to isolation in glacial refugia during the cold stages and demographic and geographic expansions following the retreat of the ice sheets (Taberlet *et al.* 1998, Milá *et al.* 2000, Lessa *et al.* 2003). Although this paradigm is one of the most consensual in modern biogeography, some aspects are still under debate, namely whether glacial times had an as severe influence on the historical distributions of non-temperate species as has been demonstrated for Europe and North America (e.g. Willis and Whittaker 2000). In this context, Lessa *et al*. (2003) investigated the effects of the Pleistocene glacial ages across the American continent by exploring the genetic signatures of post glacial demographic expansion in mammal populations from both North America and Amazonia. These signs of expansion were found to be present in North America, suggesting post-glacial colonization of these areas from a reduced source, but absent from tropical species, suggesting long-term effective population size stability. Although fundamental to understand the relative effects of the Ice Ages across latitudinal and ecological gradients, studies such as this have scarcely been repeated across the globe.

In Europe, several phylogeographic studies highlighted the importance of southern Peninsulas as glacial refugia for many taxa, functioning as survival pockets from which northern areas were later colonized (Hewitt 1996, 1999, 2000; Taberlet *et al*. 1998). Recent studies have also disclosed complex phylogeographic patterns within these southern regions, consistent with the isolation of populations during Pleistocene glaciations, leading to the acknowledgement of these regions as hotspots of diversification instead of mere historical survival areas (e.g. Alexandrino *et al.* 2000, Gómez and Lunt 2007). Furthermore, the biogeographic history of Western Europe is undoubtedly linked to that of North Africa. Although these two regions have been separated by the Strait of Gibraltar for over five million years (Duggen *et al.* 2003), resulting on vicariant differentiation of several Iberian and North African taxa (e.g. Castella *et al.* 2000, Harris *et al*. 2004a, Pinho *et al.* 2006, Vasconcelos *et al.* 2006), recent studies have documented cases of natural trans-Gibraltar colonization of either Iberia or the Maghreb, suggesting some degree of permeability of this barrier (Harris *et al.* 2002, 2004b, Carranza *et al.* 2004, Cosson *et al.* 2005, Carranza *et al.* 2006, Guillaumet *et al.* 2006). All these features contribute to

establish the biogeography of the Western Mediterranean region as actually very complex and difficult to interpret in the light of simple patterns (De Jong 1998).

Nevertheless, we expect some common patterns to emerge from studies on different organisms. In this context, we were interested in testing two main predictions: i) in northern latitudes, where the effects of glaciations were more severe, fewer and smaller patches of suitable habitat were left for the survival of populations across multiple glaciation cycles, thus leading to a higher degree of extinction of historical lineages; this should be reflected in overall lower diversity, and number of differentiated lineages in northern than in southern areas (as in for example, Michaux *et al.* (2005), Deffontaine *et al.* (2005)); ii) the effects of climatic changes on the effective population sizes were more dramatic in northern than in southern regions, meaning that northern populations should bear the signature of a rapid demographic expansion following the climate amelioration (as in Lessa *et al.* 2003), whereas southern populations should evidence marks of more stable long-term effective population size.

To test these predictions, we focused on wall lizards *Podarcis* spp. (Wagler, 1830) as model organisms, taking advantage of a recent well-established mitochondrial DNA phylogeny for the Iberian/Maghrebian part of the genus (Pinho *et al.* 2006), which major subdivisions are corroborated by a study of nuclear markers (Pinho *et al.*, in press). We selected three species that replace each other along a latitudinal gradient that spreads from Northwestern Iberia to West-Central Morocco (Figure 1) (*P. bocagei* (Seoane 1884), *P. carbonelli* Pérez-Mellado 1981, and *P. vaucheri* (Boulenger 1905)) and investigated mitochondrial DNA variation and patterns of population subdivision and historical demography in each of them. These three species display similar morphological and ecological characteristics: they are small (average snout-vent length of about 6-7 cm), tend to avoid arid habitats, prey mainly on invertebrates and show reduced dispersal capabilities. Although such a comparative phylogeographic approach could be performed on any non-related organisms, working with closely related species makes the inferences simpler and more straightforward, since general ecological requirements, dispersal abilities and important quantities for the estimation of demographic parameters (such as mutation rates or generation times) can be treated as similar without strongly biasing the analyses.

Materials and Methods

Sampling and mitochondrial DNA sequencing

We sampled a total of 247 individuals from 57 localities across the Iberian Peninsula and Morocco. Three species were collected: *P. bocagei* (N=82), an endemism of the Northwestern corner of the Iberian Peninsula, *P. carbonelli* (N=84), endemic to Central and Southern Iberia and *P. vaucheri* (N=81), distributed both in the Maghreb and in southern Spain. According to the biogeographic scenario suggested by Pinho *et al.* (2006), it is likely that *P. bocagei* and *P. carbonelli* speciated in the Iberian Peninsula, whereas *P. vaucheri* originated in North Africa, having colonized southern Spain ca. 2.8 million years ago (MYA). Although the distributions of *P. bocagei* and *P. carbonelli* are fairly well known (Pleguezuelos *et al.* 2002, A. Loureiro, pers. comm.), the distribution of *P. vaucheri* is still far from being completely described, both in the Iberian Peninsula and in North Africa (e.g it is unknown whether Algerian forms of *Podarcis* are representative of this or of other species). Sampling details are given in table 1 and figure 1. Sequences from 10 individuals were analysed in a previous study (Pinho *et al.* 2006, GenBank accession numbers DQ081151-6, DQ081172, DQ081174-6). Samples consisted of a portion of tail muscle, using the lizards' natural autotomy capacity. Samples were stored in 96% ethanol. Total genomic DNA was extracted following standard methods (Sambrook *et al.* 1989). We used primers ND4 and Leu (Arévalo *et al.* 1994) to amplify through Polymerase Chain Reaction (PCR) a portion of ~950bp of the mitochondrial DNA, comprising the 3' end of the NADH dehydrogenase subunit 4 gene and adjacent tRNAs. The amplification and sequencing procedures followed those described by Pinho *et al.* (2006). Sequences were aligned manually using BIOEDIT v. 7.0.5.2 (Hall 1999).

Data analyses

Detection of population subdivision. In order to visualize the relationships between haplotypes, NETWORK v. 4.1.1.2 was used to generate Median-Joining (MJ) networks (Bandelt *et al.* 1999). To investigate the partition of genetic variability, we performed a two-level hierarchical analysis of molecular variance (AMOVA; Excoffier *et al*. 1992) using ARLEQUIN v.2.000 (Schneider *et al.* 2000). Based on spatial analyses of molecular variance (SAMOVA; Dupanloup *et al.* 2002) each species was divided according to geographic and/or genetic criteria into different partitions and subsequent AMOVA was performed considering a three-level structure (among groups, among populations within groups and within population). DNASP (Rozas *et al.* 2003) was used to calculate summary

diversity statistics (haplotype (*Hd*) and nucleotide (π) diversity) for each partition of the data as well as Nei's D_A (Nei 1987) between the partitions.

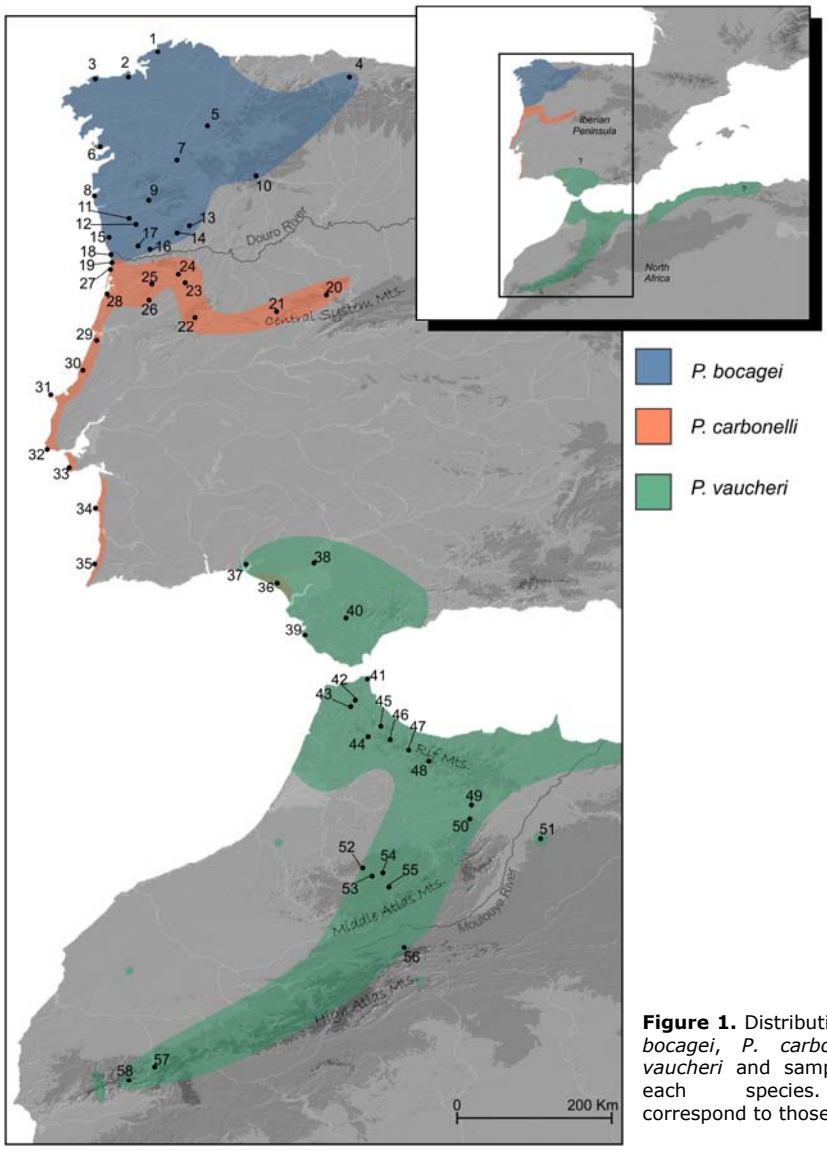

Figure 1. Distribution of *Podarcis bocagei*, *P. carbonelli* and *P. vaucheri* and sampling sites for each species. Numbers correspond to those in Table 1.

Table 1. Samples analysed in this study. Sample numbers correspond to those in figure 1.

Species	Population ID	Locality	Country	Latitude/ Longitude	Sample size
P. bocagei	1	Valdoviño	Spain	43º 36′ N 08º 07′ W	4
	2	A Coruña	Spain	43º 22′ N 08º 23′ W	7
	3	Malpica	Spain	43º 19′ N 08º 48′ W	1*
	4	Tanes	Spain	43º 12′ N 05º 24′ W	3*
	5	Sarria	Spain	42º 47′ N 07º 24′ W	6
	6	Sanxenxo	Spain	42º 24′ N 08º 49′ W	7
	7	Taboadela	Spain	42º 14′ N 07º 49′ W	2
	8	Moledo	Portugal	41º 51′ N 08º 50′ W	7
	9	Gerês	Portugal	41º 43′ N 08º 10′ W	2
	10	Montesinho	Portugal	41º 55′ N 06º 46′ W	7
	11	Braga	Portugal	41º 32′ N 08º 25′ W	6
	12	Caldas das Taipas	Portugal	41º 28′ N 08º 20′ W	1
	13	Zimão	Portugal	41º 27′ N 07º 40′ W	6
	14	Alvão	Portugal	41º 21′ N 07º 52′ W	1
	15	Vairão	Portugal	41º 19′ N 08º 40′ W	8
	16	Marco de Canavezes	Portugal	41º 10′ N 08º 09′ W	1
	17	Penafiel	Portugal	41º 12′ N 08º 16′ W	1
	18	Madalena	Portugal	41º 06′ N 08º 39′ W	8*
	19	Espinho	Portugal	41º 00′ N 08º 37′ W	4
					Total: 82
P. carbonelli	20	La Alberca	Spain	40º 28′ N 06º 05′ W	7*
	21	Villasrúbias	Spain	40º 19′ N 06º 37′ W	8
	22	Serra da Estrela	Portugal	40º 23′ N 07º 31′ W	8
	23	Sátão	Portugal	40º 43′ N 07º 43′ W	1
	24	Pendilhe	Portugal	40º 53′ N 07º 49′ W	1
	25	S. Pedro do Sul	Portugal	40º 45′ N 08º 04′ W	1
	26	Tondela	Portugal	40º 31′ N 08º 04′ W	1
	27	Esmoriz	Portugal	40º 57′ N 08º 38′ W	8
	28	Aveiro	Portugal	40º 37′ N 08º 45′ W	8*
	29	Carriço	Portugal	39º 58′ N 08º 48′ W	2
	30	S. Pedro de Moel	Portugal	39º 45′ N 09º 01′ W	10
	31	Berlenga Island	Portugal	39º 23′ N 09º 30′ W	2
	32	Cabo Raso	Portugal	38º 42′ N 09º 28′ W	8
	33	Meco	Portugal	38º 28′ N 09º 10′ W	1
	34	Sines	Portugal	37º 57′ N 08º 52′ W	1
	35	Monte Clérigo	Portugal	37º 19′ N 08º 48′ W	9
	36	Playa del Rompeculos	Spain	37º 06′ N 06º 45′ W	8*
					Total: 84
P. vaucheri	37	Huelva	Spain	37º 15′ N 06º 57′ W	5*
	38	Mairena del Aljarafe	Spain	37º 19′ N 06º 04′ W	1
	39	La Barrosa	Spain	36º 22′ N 06º 13′ W	5
	40	Guadalcacín	Spain	36º 38′ N 05º 39′ W	1*
	41	Jebel Musa	Morocco	35º 52′ N 05º 24′ W	1
	42	8Km SW of Zinat	Morocco	35º 23′ N 05º 28′ W	1
	43	15Km SW of Zinat	Morocco	35º 19′ N 05º 29′ W	2
	44	El Had	Morocco	35º 00′ N 05º 23′ W	1
	45	Chefchaouene	Morocco	35º 10′ N 05º 16′ W	1
	46	Bab Taza	Morocco	35º 03′ N 05º 12′ W	7*
	47	Bab Berred	Morocco	34º 58′ N 04º 55′ W	2
	48	Ketama	Morocco	34º 55′ N 04º 34′ W	5
	49	Taza	Morocco	34º 13′ N 04º 01′ W	2
	50	Jebel Tazzeka	Morocco	34º 06′ N 04º 10′ W	7
	51	Debdou	Morocco	33º 59′ N 03º 03′ W	6
	52	Balcon d'Ito	Morocco	33º 31′ N 05º 17′ W	1
	53	Azrou	Morocco	33º 25′ N 05º 13′ W	3
	54	Mischliffen	Morocco	33º 31′ N 05º 05′ W	8
	55	Foum Kheneg	Morocco	33º 12′ N 05º 03′ W	5
	56	Midelt	Morocco	32º 41′ N 04º 44′ W	5
	57	Oukaïmeden	Morocco	31º 12′ N 07º 51′ W	7*
	58	Tizi n' Test	Morocco	30º 49′ N 08º 20′ W	5
					Total: 81

*includes one sequence previously published (Pinho *et al.* 2006)

Detection of population growth. Coalescence theory is a powerful framework that provides models that relate the shape of the genealogy to the demographic history of populations. We used the maximum likelihood (ML) approach implemented in the software FLUCTUATE (Kuhner *et al.* 1998) to simultaneously estimate the model parameters (growth parameter *g* and theta $\theta=2N_f\mu$, where N_f is the female effective population size and μ is the mutation rate expressed in mutations per site per generation) that better suited each data set. We used an initial growth factor of 100 and Watterson (1975) estimate of theta was given as the starting value for computations. A random tree was chosen as the starting point. We ran 50 short chains, with a sampling increment of 5 and each short chain consisted of 10000 steps. Fifteen long chains were performed, each 50000 steps long and sampled every 5 steps. To ensure stability of the parameter estimates, we checked for convergence by performing five runs with different random seeds. We followed the approach described by Lessa *et al.* (2003), that considers significant the values of *g* with associated standard deviations that are lower than one-third of the value of *g*. Also according to Lessa *et al.* (2003), very significant values are those in which $g > 6\text{xSD}(g)$. Because the ML computations of the growth parameter tend to be affected by an upward bias (Kuhner *et al.* 1998) we also used more conservative tests of population growth. Fu's (1997) F_S and Ramos-Onsins and Rozas' (2002) R_2 statistics are some of the most powerful tests available to detect demographic expansions (Ramos-Onsins and Rozas, 2002). F_S represents the probability of observing a similar or higher number of haplotypes in a random neutral population given the observed value of theta; in populations that have undergone recent expansion an excess of rare alleles is expected, leading to large negative values of this statistic. R_2 illustrates the difference between the number of singleton mutations and the average number of pairwise differences. A recent population expansion will lead to low values of this statistic. We used DNASP to calculate both statistics and to assess their significance using 10000 coalescent simulations. Since a high level of population structure may confound the inferences derived from all of these tests or provide unclear results because they are designed to deal with panmictic populations, we partitioned the data set into different subsets and performed independent analyses for each partition where polymorphism was observed.

Coalescence times. As an addition to the analyses about historical demography and to estimate the coalescence times and ages of relevant mutations, we used GENETREE v9.0 (R.C. Griffiths, available in http://www.stats.ox.ac.uk/~griff/software.html). This program uses a standard coalescent framework to produce ML estimates of mutation, migration and growth rates, as well as the time to the most recent common ancestor (TMRCA) and ages of mutations in a given gene tree. This program assumes an infinite-

sites model of evolution for sequence data. Therefore, we had to make several adjustments in order to accommodate the data sets to this model of evolution. These consisted in removing homoplasic positions, which led to the merging of some haplotypes, or unfolding them into independent mutations based on the evolutionary path suggested by the haplotype network (as in Wilder *et al.* 2004). We analysed each species separately and only considered Moroccan *P. vaucheri*, leaving out the less well sampled Spanish populations of this species. Ancestral states were defined based on the consensus sequence of each species and, when required, on comparisons to the nearest outgroup (based on Pinho *et al.* 2006). GENETREE works under two different demographic models: one of constant effective population size, which requires the estimation of a single parameter, θ and one that incorporates exponential growth, in which β, the intrinsic population growth parameter, is additionally estimated. GENETREE does not estimate the ML values of both parameters simultaneously. We therefore estimated θ over a wide range of values of β and progressively narrowed the interval for each parameter in order to reach the joint ML estimate of both. Each analysis was carried out for 10^6 steps and different random seeds were used in each run. After achieving an estimation of the ML values of θ and β, we tested the null hypothesis of constant population size by computing the likelihood of the gene trees constructed assuming no growth and comparing it to that of the gene trees assuming exponential growth through likelihood ratio tests. The TMRCA were estimated after choosing the appropriate demographic model for each data set. Generation time, which is needed to calculate the TMRCA, was assumed to be on average 2.5 years based on published data on sexual maturity and longevity of these lizards (Caetano *et al.* 1986, Carretero *et al.* 2006).

Mutation rate. Since there are no calibrations available for the mitochondrial DNA fragment studied in *Podarcis*, we estimated a mutation rate using the ND4 and tRNA data sets from Pinho *et al.* (2006), trimmed to the same size as the fragment analysed here. First, a model of sequence evolution was chosen using PAUP 4.0b10 (Swofford 2002) and MODELTEST 3.7 (Posada and Crandall 1998). Based on this model, we computed average ML distances between Iberian and North African *Podarcis* and the outgroup *P. muralis.* These two groups belong to distinct lineages among the several differentiated clades that constitute the genus, which are believed to have radiated during the Miocene, 10 to 16 million years ago (Oliverio *et al.* 2000, Poulakakis *et al.* 2005). Although these dates constitute a large interval, they are the only historical dates available for the divergence between Iberian/Maghrebian *Podarcis* and *P. muralis*, and were therefore used to compute the rate of sequence evolution.

These analyses can only be performed under the assumption of a molecular clock. In order to test this, we used relative rate tests (Takezaki *et al.* 1995)

implemented in the software PHYLTEST v2.0 (Kumar 1996). To test for rate constancy between the Iberian/Maghrebian *Podarcis* and *Podarcis muralis* we had to use an external outgroup. In the absence of available Lacertidae sequences for the same stretch of the ND4-tRNA(Leu) fragment we used a sequence of the Teiidae *Cnemidophorus vanzoi* from GenBank (accession number DQ168991). In this test we used the data set of Pinho *et al.* (2006). In order to be able to use the same rate of sequence evolution in the TMRCA estimates for the *P. bocagei, P. carbonelli* and *P. vaucheri* data sets, we also tested for rate constancy between the three species using pairwise analyses with *P. muralis* sequences from Pinho *et al.* (2006) as outgroups. We used the Kimura-2-parameter correction throughout the PHYLTEST analyses.

Results

DNA sequence variation and population subdivision

In this study, we generated sequences of the mitochondrial fragment ND4-tRNA(Leu) for 247 individuals from the three different species. After alignment, sequences were trimmed to a fragment of 797bp. No indels were detected. A total of 81 haplotypes were detected among the three studied species: 19 in *P. bocagei*, 21 in *P. carbonelli* and 41 in *P. vaucheri*. These haplotypes have been coded with the letter B, C, or V, representative of the species where they were found, followed by a number. The sequences of each haplotype have been deposited in GenBank, accession numbers EF081075 to EF081155. MJ networks for each species are shown in figure 2. The correspondence between the studied populations and the haplotypes detected, as well as their frequency are shown in appendixes 1, 2 and 3.

In *P. bocagei*, a frequent and widespread haplotype was found (B1). All other 18 haplotypes detected differ from this by one or two mutations. A second widespread haplotype, B2, differing in one mutation from B1, was found to be nearly fixed in individuals sampled from the eastern area of the distribution, although also being present in a southwestern population (Espinho). Many populations have private, sometimes fixed haplotypes. This translates into a relatively high Φ_{ST} value (0.663), although overall genetic diversity is low (*Hd*=0.789, π=0.00168). Even acknowledging this high Φ_{ST} value, considering that there are no major geographic or genetic discontinuities in the distribution of *P. bocagei*, for practical purposes all subsequent analyses on population growth regarded this species as a single evolutionary unit and were performed on the complete data set.

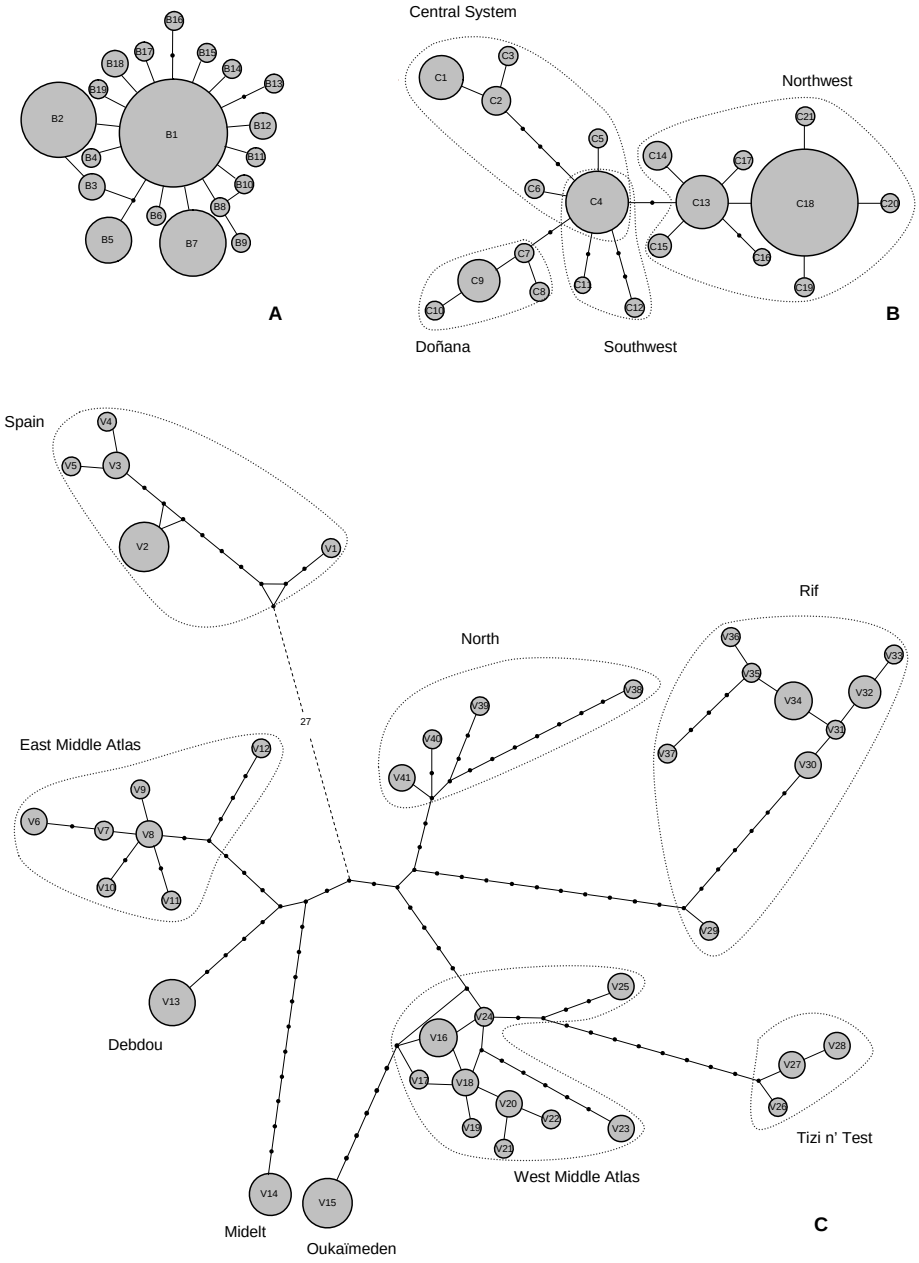

Figure 2. Median-Joining networks obtained for the three species under study: **A.** *Podarcis bocagei*. **B.** *Podarcis carbonelli.* **C.** *Podarcis vaucheri*. The area of each circle is proportional to the frequency of the haplotype it represents. Haplotype codes correspond to those in appendixes 1, 2 and 3. Numbers on dashed branches refer to number of mutations between the connected nodes. Dotted lines delimit the geographic partitions of the data sets used for demographic analyses.

Compared to this species, *P. carbonelli* revealed a higher level of genetic variability (*Hd*=0.823, π=0.00456) and differentiation (Φ_{ST}=0.819). There is a clear geographic pattern in the distribution of haplotypes, although some haplotypes are shared between distant regions. Based on this geographical distribution of haplotypes, the *P. carbonelli* data set was subdivided into four geographical subsets: "Northwest" (populations 22-32), "Central System" (populations 20 and 21), "Doñana" (population 36) and "Southwest" (populations 33-35). Genetic variation among these defined groups accounts for 71.8% of the total variance. Average Nei's D_A between groups is 0.0032.

Contrasting to the lower levels of genetic diversity detected in the previous species, *P. vaucheri* exhibits remarkably high haplotype and nucleotide diversity (*Hd* = 0.971, π =0.03024), the latter an order of magnitude higher than those found in the strictly Iberian species. In accordance with the findings of previous studies, two major groups were detected, corresponding to Spanish and Moroccan samples. These two monophyletic groups show a high degree of differentiation (D_A=0.0388) and 61.7% of the total variance is due to differentiation among them. Still, a considerable amount of variation is found within each clade, particularly in Morocco, where several highly divergent and geographically non overlapping haplotype groups were found: one comprises only individuals sampled in the north of the country (populations 41–44); another includes samples from the Rif Mountain system (populations 45–48); samples from the Eastern Middle Atlas Mountains (49 and 50), although clearly differentiated, appear to be evolutionarily closer to those from Debdou (51), where a single haplotype was detected, than to other localities in the Middle Atlas (52–55); instead, these seem to form a clade with samples from the High Atlas Mountains (57 and 58), although sequences from these two populations do not group together. In the population of Midelt (56) a single highly differentiated haplotype was found (V14). Based on these results, we established 8 subsets of the Moroccan data, that we named "North", "Rif", "East Middle Atlas", "Debdou", "West Middle Atlas", "Oukaïmeden", "Tizi n'Test" and "Midelt". Considering these partitions, Φ_{CT} is 0.836 and average D_A between groups is 0.0219.

For each level of subdivision (into four populations in *P. carbonelli* and into eight in Moroccan *P. vaucheri*) there was support from SAMOVA for the proposed partitions although higher levels of subdivision would increase Φ_{CT}. Since further dividing of the data sets would yield very small sample sizes for some subsets, we adopted the conservative division presented above for the benefit of statistical power in subsequent analyses.

Table 2. Summary statistics, measures of population differentiation and population growth for three species of wall lizards and respective data subsets, based on mitochondrial DNA variation.

Species	Partition	N	Polymorphism and population differentiation								Population growth		
							2-level AMOVA	3-level AMOVA					
			H	*Hd*	*π*	θ	Φ_{ST}	Φ_{CT}	Φ_{ST}	Φ_{SC}	*Fs*	R_2	*g* ± SD
P. bocagei	-	82	19	0.789	0.00168	0.04468	0.663**	-	-	-	-15.092**	0.0347*	5345.75 ± 336.94‡‡
P. carbonelli	-	84	21	0.823	0.00456	0.01343	0.819**	0.718**	0.865**	0.522**	-6.179	0.0593	516.07 ± 178.13
	Northwest	50	9	0.569	0.00111	0.03294	0.502**	-	-	-	-4.902**	0.0547*	9509.44 ± 1047.30‡‡
	Southwest	11	3	0.345	0.00115	0.04116	1.000*	-	-	-	0.401	0.1975	10000.00 ± 2777.23‡
	Central System	15	6	0.762	0.00327	0.00337	0.528*	-	-	-	-0.114	0.1469	148.52 ± 244.63
	Doñana	8	4	0.643	0.00117	0.06429	-	-	-	-	-1.387	0.1610	9009.43 ± 2109.01‡
P. vaucheri	-	81	41	0.971	0.03024	0.05570	0.908**	0.617**	0.949**	0.868**	-2.119	0.1140	32.15 ± 16.74
	Spain	12	5	0.667	0.00384	0.00436	0.747**	-	-	-	0.820	0.1917	-40.43 ± 115.38
	Morocco	69	36	0.970	0.02312	0.04949	0.869**	0.836**	0.878**	0.258**	-3.290	0.1107	58.14 ± 24.59
	North	5	4	0.900	0.00812	0.02407	-0.868	-	-	-	0.883	0.2000	301.94 ± 121.63
	Rif	15	9	0.905	0.00438	0.01108	0.223	-	-	-	-2.276	0.1213	215.23 ± 127.06
	West Middle Atlas	17	10	0.926	0.00532	0.00951	0.047	-	-	-	-2.176	0.1209	159.26 ± 120.69
	Oukaïmeden	7	1	0	0	0	-	-	-	-	-	-	-
	Tizi n' Test	5	3	0.800	0.00176	0.00532	-	-	-	-	0.061	0.2848	1659.56 ± 1129.44
	East Middle Atlas	9	7	0.944	0.00471	0.04013	-0.053	-	-	-	-1.960	0.1413	960.85 ± 228.61‡
	Debdou	6	1	0	0	0	-	-	-	-	-	-	-
	Midelt	5	1	0	0	0	-	-	-	-	-	-	-

N, number of individuals; *H*, number of haplotypes detected; *Hd*, haplotype diversity; *π*, nucleotide diversity; θ, population mutation parameter, estimated jointly with *g*, growth rate parameter using FLUCTUATE (Kuhner *et al*. 1998); Φ_{CT}, Φ_{ST}, Φ_{SC}, fixation indexes; *Fs*, Fu's (1997) Fs; R_2, Ramos-Onsins and Rozas' (2002) R_2; 2-level AMOVA were performed considering all the populations of the data set as a single group; 3-level AMOVA were performed dividing the data set into subsets as explained in the text. AMOVA analyses were carried out using uncorrected pairwise distances; the use of more parameter-rich models of substitution did not yield significantly different values (results not shown).**, significant at p<0.01; *, significant at p<0.05; ‡‡, *g* over 6 standard deviations (SD) above 0; ‡, *g* between 3 and 6 SD above 0.

Population growth

Table 2 summarizes most of the information described above and also the results of the tests used to detect population expansion. ML growth parameter estimates obtained using FLUCTUATE are positive in all cases except in Spanish *P. vaucheri*. However, although large positive values were detected across populations of all three species, both large and significant (sensu Lessa *et al.* 2003) growth parameter estimates are found only in a few cases: *P. bocagei,* the northwestern, southwestern and Doñana geographical partitions of *P. carbonelli* and the East Middle Atlas clade of *P. vaucheri*. The computations of the more conservative Fu's (1997) F_S and Ramos-Onsins and Rozas' (2002) R_2 statistics are concordant with these analyses; significant values of both of these statistics were only detected in *P. bocagei* and the northwestern region of the distribution of *P. carbonelli.*

Calculation of coalescence times

According to MODELTEST, the most appropriate model of evolution for the Pinho *et al.* (2006) ND4-tRNA(Leu) data set was the TrN model with an estimate of invariable sites (0.6018) and a discrete approximation of the gamma distribution (4 rate categories, shape parameter α=2.091).

All relative rates tests performed showed that the null hypothesis of rate constancy between all tested lineages was not rejected at the 5% level. We were therefore able to estimate a rate of sequence evolution based on the corrected divergence to *P. muralis* and to assume the same evolutionary rate in the estimates of the TMRCA for all three data sets. Our analyses suggest that for the ND4-tRNA(Leu) mitochondrial fragment differences accumulate at a rate that ranges from 1.74 to 2.78% per million years, which yields a value of μ ranging from 2.175×10^{-8} to 3.475×10^{-8} mutations per site per lineage per generation.

Estimates of θ_{ml}, β_{ml} and TMRCA obtained using GENETREE are shown in table 3. Likelihood ratio tests show that only the *P. bocagei* data set fits an exponential growth model. Gene trees for the three data sets are shown in figure 3, and were drawn under the assumption of exponential growth for *P. bocagei* and constant population size for both *P. carbonelli* and Moroccan *P. vaucheri.*

Discussion

In this work we tested two biogeographic expectations related to the geographic distribution of the species under study and their response to climatic oscillations during the Pleistocene. Our data strongly suggests that species with historical

distributions on different latitudes were affected in distinct ways by climatic changes, since both of our initial predictions were verified.

Table 3. Population parameters estimated using GENETREE, assuming constant population size (upper row) and exponential growth (lower row) for each of the three groups analysed. The coalescence times were estimated using two extreme rates of evolution for the studied mitochondrial DNA fragment: a slow rate of 2.175×10^{-8} and a fast rate of 3.475×10^{-8} (see text for details).

Species/group	θ_{ml}	β_{ml}	Likelihood Score ± SE	LRS	TMRCA ± SD in 10^3 years	
					Slow rate	Fast rate
P. bocagei	0.00715	–	$2.20 \times 10^{-9} \pm 4.15 \times 10^{-10}$	16.3**	280.3±84.2	175.4±52.3
	0.04266	121.3	$7.66 \times 10^{-6} \pm 5.19 \times 10^{-6}$		113.4±16.1	71.0±10.1
P. carbonelli	0.00866	–	$4.65 \times 10^{-23} \pm 1.25 \times 10^{-24}$	2.7	500.0±132.5	312.9±82.3
	0.01531	6.7	$1.78 \times 10^{-22} \pm 1.37 \times 10^{-23}$		343.2±47.9	214.8±30.0
P. vaucheri	0.03915	–	$4.56 \times 10^{-72} \pm 3.54 \times 10^{-72}$	3.7	1615.1±82.7	1010.9±51.8
(Morocco)	0.04755	1.9	$2.89 \times 10^{-71} \pm 2.10 \times 10^{-71}$		1481.1±138.8	927.0±86.9

LRS, likelihood ratio score; ** significant ($p<0.005$).

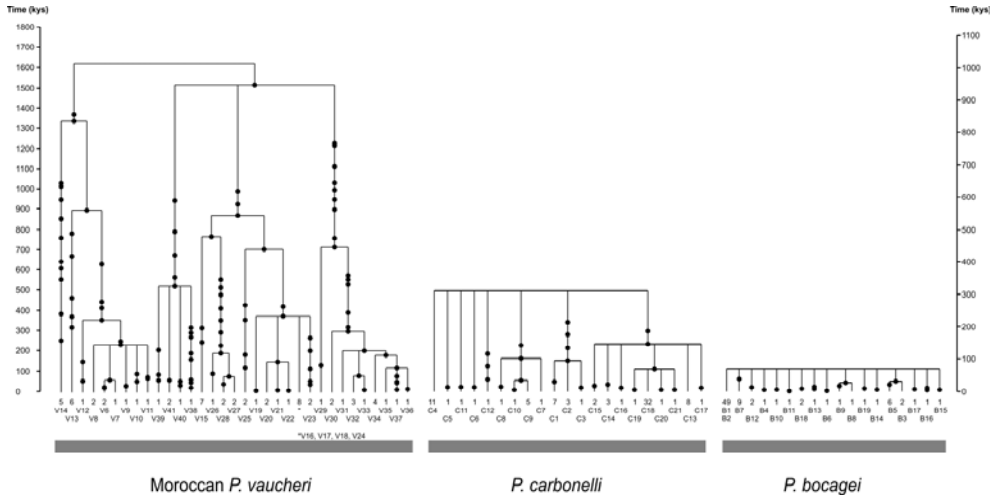

Figure 3. Gene trees for *P. bocagei, P. carbonelli* and the Moroccan populations of *P. vaucheri*, obtained from parameters estimated using GENETREE, assuming constant population size in *P. carbonelli* and *P. vaucheri* and an exponential growth model in *P. bocagei*, following the results of the likelihood ratio tests presented in table 3. Mutations along the genealogy are represented by black circles. Parameters used to estimate these gene trees, their respective TMRCA and ages of mutations are represented in table 3. The time scale is the same in the three genealogies. The scale represented on the left is based on a slow rate of evolution ($\mu=2.175 \times 10^{-8}$) and on the right on a fast rate of evolution ($\mu=3.475 \times 10^{-8}$).

Effects of the Ice Ages on the species historical distributions

The Western Mediterranean region is known to have been severely affected by climatic oscillations. Although patches of temperate deciduous forest appear to

have been maintained throughout cold phases (Roucoux *et al.* 2005), Iberian landscape became characterized by the extensive presence of tundra, steppe and, in northernmost regions, ice sheets (Costa Tenorio *et al.* 1998). In North Africa, pollen data suggest an increase of semi-desertic and decrease of forest-type vegetation during cold stages (Jolly *et al.* 1998). These changes in vegetation were tracked by co-distributed fauna, and, although in many cases we cannot model paleodistributions (as in Teixeira 1999, or Hugall *et al.* 2002) to understand those changes, we can make inferences by looking at the genetic signature left by the Ice Ages.

Podarcis bocagei, the species with the northernmost distribution among the three studied species, shows a very low level of variability and of population subdivision. Our analyses based on the coalescent show that the most recent common ancestor of all *P. bocagei* mtDNA haplotypes occurred 71 - 113 ky ago, which is quite recent assuming that this species diverged from its closest living relative, *P. hispanica* type 1, some 4-6 million years ago (based on Pinho *et al.* (2006) data). This recent common ancestry suggests that the genetic signature of earlier events on the history of this species has been erased by an extreme bottleneck that reset the variation to an almost zero level, concordant with a severe shrinkage of its distribution (and consequently of its effective population size) during repeated episodes of cooling climatic conditions. Moreover, variability was kept at a low level until the last glacial maximum since almost all the variation in this species is extremely recent and arose after the last climate warming. Using the slowest rate of evolution, thirteen out of the nineteen mutations observed in the *P. bocagei* data set occurred (according to GENETREE calculations) less than 20000 years ago (Figure 3). However, since the molecular clock was calibrated taking into account an ancient splitting event (10-16 million years ago), these and other time estimates, which refer to very recent events, should be taken with caution because the extrapolation of rates across different timescales could provide non-reliable inferences (e.g. Ho and Larson 2006). Nevertheless, taking these data together, it seems reasonable to infer that *P. bocagei* recently survived near-extinction.

On the other hand, the southernmost species, *Podarcis vaucheri*, contains a very large amount of inter and intra population variability, which make an extreme contrast to the depletion of genetic diversity observed in *P. bocagei*. The fact that this species shows a profound genetic discontinuity resulting from the colonization of southern Spain, about 2.8 million years ago (Pinho *et al.,* 2006), does not solely account for this discrepancy. Disregarding the effect that this colonization has had on the species' variability by focusing only on Moroccan populations, we find that nucleotide diversity within this group is over one order of magnitude higher than in *P. bocagei* (0.0231 and 0.0018, respectively). The coalescence time for Moroccan *P. vaucheri* was estimated to be between 1.010 –

1.615 MYA. Although the relationship between sequence coalescence and population divergence times is complex and dependent upon a large number of factors (Rosenberg and Feldman 2002), these values may indicate that the mechanism that triggered differentiation in this species was habitat fragmentation related to the beginning of the Ice Ages. However, not all of the presently observed structure originated simultaneously (Figure 3). After an early split, subsequent fragmentation probably arose after the successive cycles of cooling and warming climatic conditions that characterize the Pleistocene (Zagwijn 1992). Unlike the scenario observed in the other species, many of these refugial populations were able to survive until the present thanks to the maintenance of suitable climatic conditions, which conforms to our first prediction of a lower degree of extinction in southern territories. Substructure within this species seems to be associated with different mountain ranges and distinct areas within them. Presently, *P. vaucheri* seems to be highly dependent upon humid areas since its Moroccan distribution appears to be correlated to rainfall (see Bons and Geniez, 1996). We may therefore speculate that during cold and dry stages, moist valleys within these mountain ranges may have acted as glacial refugia for this species.

Podarcis carbonelli represents somewhat of an intermediate stage between the two above-mentioned cases. There is clear but shallow geographical substructure which does not appear to have arisen during the first stages of the Pleistocene as in *P. vaucheri;* coalescence time for this species (313 – 500 kya) falls within the Cromerian complex stage (Early Middle Pleistocene), which comprised several cold and warm phases (Zagwijn 1992); all variability prior to that time has been erased. After that, this species was allowed to disperse within southern Iberia, probably during warm interglacials.

Evidence for population growth and recolonization after the Ice Ages

The star-shaped genealogy observed in *P. bocagei* is consistent with the hypothesis of a rapid population expansion from a reduced refugium. This was suggested by several different tests of population expansion, which consistently yielded significant instances of demographic growth for this species. The associated range expansion appears to have occurred northwards, as expected (Hewitt 1996, 1999, 2000; Taberlet *et al.* 1998).

In *P. vaucheri*, however, there is little support for population growth (or decline) in any of the tests performed, both when considering the species as a whole and after dividing it into smaller geographical units. We may conclude that a relative demographic stability has been a feature of this species' evolution. Smoother climatic changes allowed the colonization of new habitats a long time ago and permitted the maintenance of high effective population sizes throughout the Pleistocene, even during repeated episodes of fragmentation, allowing for

the achievement of large levels of within-population variability and rendering impossible the detection of any sign of population expansion. Long range recolonizations are not likely to have occurred since all the groups detected are apparently strictly allopatric.

Lying between these two extremes, in *P. carbonelli* there is not consistent evidence for population growth except for its NW clade. As observed in *P. bocagei,* a range expansion of this subgroup may have occurred in a northwards direction, with a progressive loss of mtDNA variability that led to an almost fixation of haplotype C18. Other phylogroups within this species may not have had the chance to expand geographically.

Concordance of scenarios in the context of the Western Mediterranean Quaternary biogeography

Despite our finding of a low level of genetic substructure both in *P. bocagei* and *P. carbonelli*, these two situations are not the rule regarding most of the Iberian taxa analysed. Instances of deep population fragmentation related to isolation in glacial refugia have been described for a wide range of Iberian species, including for example the European rabbit, *Oryctolagus cuniculus* (Branco *et al.* 2000, 2002), the golden-striped salamander, *Chioglossa lusitanica* (Alexandrino *et al.* 2000, 2002) and the Schreiber's green lizard, *Lacerta schreiberi* (Paulo *et al.* 2001, 2002). Many more such cases have been reported (see for a comprehensive review Gómez and Lunt 2007). Phylogeographic data are scarce for other Iberian *Podarcis* species, but the few mitochondrial DNA sequences available (e.g. Harris *et al.* 2002, Pinho *et al.* 2006) suggest that this may also be the case of the closely related *Podarcis hispanica* types 1 and 2. However, despite this general trend for the observation of differentiation within Iberia across a wide range of different taxa, usually no more than two or three refugial populations are inferred, suggesting some degree of lineage extinction.

With respect to the phylogeography of *Podarcis bocagei*, there are several reports in the literature for similar evolutionary histories of confinement and posterior expansion within north-western Iberian herpetofauna, if not for a whole species at least for a specific phylogroup within it. Two of the better studied case-studies that are good examples of such historical processes are the salamander *Chioglossa lusitanica* (Alexandrino *et al.* 2000, 2002), the lizard *Lacerta schreiberi* (Paulo *et al.* 2001, 2002). In what concerns the most recent stages of the Pleistocene, a similar pattern of a northwards expansion tracking suitable climatic conditions can be observed in both species. Another interesting feature is that the Douro River, in Northern Iberia, might have functioned as a barrier to the dispersal of herpetofauna since its crossing has been associated to a major loss of genetic variability. The present distribution of *P. bocagei,* situated mostly north of this river, may indicate that the species was unable,

during the last cold stages, to move south tracking the changing distribution of suitable habitat (Sá-Sousa 2001).

Adding to these case-studies, another concordant scenario is herein described for the NW group of *P. carbonelli*, which seems to have recently expanded its distribution from a southern location and that has not been able to cross the above-mentioned river. At a more general level, however, it seems difficult to establish a parallel between the phylogeography of *P. carbonelli* and that of other species because of the uniqueness of this species' distribution in the context of Iberian herpetofauna. There appears to be a rough temporal concordance with the substructure observed within the coastal clade in *Lacerta schreiberi* (0.4 – 0.6 MYA), that has probably resulted from fragmentation also occurred during a "Cromerian extended cold phase" (Paulo *et al.* 2001, 2002). These processes may have ultimately led to the formation of southern geographical isolates in both species. Another common feature is that the area near the mountains of the Spanish Central System probably provided refugium for both species, although at very different time-scales, a pattern that is also found in the Iberian endemic newt *Lissotriton boscai* (Martínez-Solano *et al.* 2006).

Isolation in glacial refugia has been invoked to explain differentiation in a few Maghrebian species, not only in squamates (*Agama* lizards, Brown *et al.* 2002) but also in invertebrates (*Helix* land snails, Guiller *et al.* 2001), amphibians (*Pleurodeles* salamanders, Carranza and Arnold 2004) mammals (*Crocidura* shrews, Cosson *et al.* 2005) and birds (*Galerida* larks, Guillaumet *et al.* 2006). Curiously, in most of these cases (snails, shrews and larks) there is a consistent geographic pattern of vicariance between Eastern (Algeria and Tunisia) and Western (Morocco) populations, which is also found in *Podarcis* (Harris *et al.* 2002, Pinho *et al.,* 2006) and in other species (*Timon* spp., Paulo 2001, *Pleurodeles* spp.*,* Veith *et al.* 2004) at higher taxonomic levels. To our knowledge, examples of Pleistocene differentiation comparable to that in *P. vaucheri* in a small regional scale as Morocco are absent from the literature, but this might be a consequence of the fact that most of the studies performed in this area are essentially phylogenetic and attempt to solve higher-level taxonomy (e.g. Carranza *et al.* 2004, Harris *et al.* 2004a,b). It is interesting to notice that the dry basin of the Moulouya river, which has been proposed as a major biogeographic divide in North Africa (e.g. Álvarez *et al.* 2000) and separates the population of Debdou from all the others, does not seem to stand out in importance from other putative barriers in the diversification of *P. vaucheri*.

Although phylogeographic scenarios that conform to those inferred for each of the three lizard species analysed have been therefore described, few studies have actually documented differences related to the species' latitudinal

distributions. A recent study involving the newt *Lissotriton boscai* (Martínez-Solano *et al.* 2006) showed that persistence levels across the Pleistocene glaciations, as shown by older phylogeographic splits, seem to be relatively higher within the cryptic lineage inhabiting southern Iberian Peninsula than for its northern counterparts. Another study, involving the freshwater turtle *Mauremys leprosa* (Fritz *et al.* 2006), describes instances of rapid demographic expansion in its northern clade, whereas a southern clade restricted to North Africa shows signs of demographic stability.

Exploring alternative hypotheses

It is well understood that the genetic imprint of selective and demographic processes can be difficult to distinguish. For example, the pattern observed in *P. bocagei* would also be consistent with a very recent selective sweep acting on its mitochondrial genome. What therefore makes us prefer a demographic explanation rather than a selective one? Not only the previously referred examples of similar findings in co-distributed species reinforce the idea of a demographic and climatic-related expansion, but we also find evidence for that expansion using nuclear markers (allozymes and microsatellites, C. Pinho*,* unpublished results). Similarly, an alternative hypothesis to explain the comparatively enormous level of mtDNA variation in *P. vaucheri* would be an acceleration of the substitution rate in this species; however, we were not able to reject the hypothesis of rate constancy between this species and the others, which leaves us with a strictly demographic scenario to explain the differences obtained.

Concluding remarks

It is well known that the effects of the Quaternary glaciations were not uniform across the globe. Latitudinal differences on the way that species responded to these climatic changes have been well described across large areas (Lessa *et al.* 2003), but it was uncertain if these effects could be detected on a smaller scale such as the region presently studied (which has a north-south extension of less than 1500Km). Obviously, the ability to detect such differences is also correlated to the dispersal potential of the organisms under study, which make *Podarcis* a suitable model to study evolutionary processes at a smaller scale due to their reduced dispersal abilities. To our knowledge, comparative analyses such as this study, including a demographic characterization of both Iberian and Maghrebian populations, are lacking for the Western Mediterranean. A reason for this might be the lack of suitable model-taxa on which to perform comparisons. Many species are distributed only on one side of the Strait or were shown to have colonized one of the continents too recently to distinguish between the effects of a recent founder event from a response to Quaternary climatic oscillations

(Harris *et al.* 2004b, Carranza *et al.* 2004, Cosson *et al.* 2005, Guillaumet *et al.* 2006, Carranza *et al.* 2006). However, examples of more ancient colonizations or vicariant events have also been described and, although so far some of these systems have only been analysed from a strictly phylogenetic perspective, they are potential candidates for replicating our gradient analyses and testing our main conclusions (e.g *Buthus* scorpions, Gantenbein and Lagiardèr, 2003; *Alytes* toads, Martínez-Solano *et al.* 2004; *Salamandra* salamanders, Escoriza *et al.,* 2006; *Acanthodactylus* lizards, Harris *et al.* 2004a; *Blanus* worm lizards, Vasconcelos *et al.* 2006). In any case, our data, taken together with other studies that address these issues (Fritz *et al.* 2006, Martínez-Solano *et al.* 2006) provide substantial evidence for a direct relationship between latitude and the patterns of subdivision and historical demography.

Acknowledgements

This work was partially financed by FCT research projects, POCTI/40987/BSE/2001 and POCTI/41912/BSE/2001. C Pinho was supported by a PhD grant (SFRH/BD/4620/2001) and DJ Harris by a post-doctoral grant (SFRH/BPD/5702/2001), both from FCT. The authors wish to thank Salvador Carranza for sharing samples and Miguel A. Carretero, Ana Luz, Anna Perera, Antigoni Kaliontzopoulou, Armando Loureiro, Bruno Ribeiro, Daniel Cunha, Diana Barbosa, Elsa Froufe, Gonçalo Themudo, Jean-Christophe Mace, Mário J. Pinho, Mohammed Znari, Neftalí Sillero, Paulo Sá-Sousa, Raquel Godinho, Raquel Ribeiro, Ricardo Pereira, Sara Rocha and Vasco Batista for sample collection. Armando Geraldes, Fernando Sequeira, Miguel Carneiro, Sara Rocha, Rui Faria and Xana Sá Pinto provided helpful comments and advice regarding the analyses and discussion of the data. We also acknowledge anonymous reviewers and the Subject Editor Angus Davison for their constructive comments on earlier versions of this manuscript.

References

Alexandrino J, Arntzen JW, Ferrand N (2002) Nested clade analysis and the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Heredity*, **88**, 66–74.

Alexandrino J, Froufe E, Arntzen JW, Ferrand N (2000) Genetic subdivision, glacial refugia and postglacial recolonization in the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Molecular Ecology*, **9**, 771–781.

Álvarez Y, Mateo JA, Andreu AC, Díaz-Paniagua C, Diez A, Bautista JM (2000) Mitochondrial DNA haplotyping of *Testudo graeca* on both continental sides of the Straits of Gibraltar. *Journal of Heredity,* **91,** 39–49.

Arévalo E, Davis SK, Sites Jr JW (1994) Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. *Systematic Biol*ogy, **43**, 387–418.

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18,** 489–522.

Avise JC (2000) *Phylogeography*. *The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution,* **16**, 37–48.

Bons J, Geniez P (1996) *Amphibiens et Reptiles du Maroc (Sahara Occidental compris). Atlas biogéographique*. Asociación Herpetológica Española, Barcelona.

Branco M, Ferrand N, Monnerot M (2000) Phylogeography of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome *b* gene. *Heredity*, **85**, 307–317.

Branco M, Monnerot M, Ferrand N, Templeton AR (2002) Postglacial dispersal of the European rabbit (*Oryctolagus cuniculus*) on the Iberian Peninsula reconstructed from nested clade and mismatch analyses of mitochondrial DNA genetic variation. *Evolution*, **56**, 792–803.

Brown RP, Suárez NM, Pestano J (2002) The Atlas mountains as a biogeographical divide in Northwestern Africa: evidence from mtDNA evolution in the Agamid lizard *Agama impalearis*. *Molecular Phylogenetics and Evolution*, **24**, 324–332.

Caetano MH, Castanet J, Crespo EG (1986) Estimation à l'aide de la squelettochronologie de l'âge de *Podarcis hispanica,* (Steindachner, 1870), Sauria, Lacertidae, provenant d'une population portugaise. *Revue Suisse de Zoologie,* **93**, 117–127.

Carranza S, Arnold EN (2004). History of West Mediterranean newts, *Pleurodeles* (Amphibia: Salamandridae) inferred from old and recent DNA sequences. *Systematics and Biodiversity* **1,** 327–337.

Carranza S, Arnold EN, Pleguezuelos JM (2006) Phylogeny, biogeography and evolution of two Mediterranean snakes, *Malpolon monspessulanus* and *Hemorrhois hippocrepis* (Squamata, Colubridae), using mtDNA sequences. *Molecular Phylogenetics and Evolution,* **40**, 532 - 546

Carranza S, Arnold EN, Wade E, Fahd S (2004) Phylogeography of the false smooth snakes, *Macroprotodon* (Serpentes, Colubridae): mitochondrial DNA sequences show European populations arrived recently from Northwest Africa. *Molecular Phylogenetics and Evolution,* **33**, 523–532.

Carretero MA, Ribeiro R, Barbosa D, Sá-Sousa P, Harris DJ (2006) Spermatogenesis in two Iberian *Podarcis* lizards: relationships with male traits. *Animal Biology*, **56**, 1–12.

Castella V, Ruedi M, Excoffier L, *et al*. (2000) Is the Gibraltar Strait a barrier to gene flow for the bat *Myotis myotis* (Chiroptera: Vespertilionidae)? *Molecular Ecology* **9**, 1761–1772.

Cosson JF, Hutterer R, Libois R, *et al.* (2005) Phylogeographical footprints of the Strait of Gibraltar and Quaternary climatic fluctuations in the Western Mediterranean: a case study with the greater white-toothed shrew, *Crocidura russula* (Mammalia: Soricidae). *Molecular Ecology,* **14**, 1151–1162.

Costa Tenorio M, Morla Juaristi C, Sainz Ollero H (eds.) (1998) *Los bosques ibéricos. Una interpretación geobotánica.* Editorial Planeta, Barcelona

Deffontaine V, Libois R, Kotlík P *et al.* (2005) Beyond the Mediterranean peninsulas : evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*). *Molecular Ecology*, **14**, 1727–1739.

Duggen S, Hoernie K, van den Bogaard P, Rupke L, Morgan JP (2003) Deep roots of the Messinian salinity crisis. *Nature*, **422**, 602–606.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**, 2571–2581.

Escoriza D, Comas MM, Donaire D, Carranza S (2006) Rediscovery of *Salamandra algira* Bedriaga, 1833 from the Beni Snassen massif (Morocco) and phylogenetic relationships of North African *Salamandra*. *Amphibia-Reptilia*, **27,** 448–455.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics,* **131**, 479–491.

Fritz U, Barata M, Busack SD, Fritzsch G, Castilho R (2006) Impact of mountain chains, sea straits and peripheral populations on genetic and taxonomic structure of a freshwater turtle, *Mauremys leprosa* (Reptilia, Testudines, Geoemydidae). *Zoologica Scripta* **35,** 97–108.

Fu Y-X (1997) Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics,* **133**, 693–709.

Gantenbein B, Lagiardèr CR (2003) The phylogeographic importance of the Strait of Gibraltar as a gene flow barrier in terrestrial arthropods: a case study with the scorpion *Buthus occitanus* as model organism. *Molecular Phylogenetics and Evolution*, **28**, 119–130.

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography in southern European refugia: evolutionary perspectives on the origin and conservation of European biodiversity* (Weiss S, Ferrand N, eds). Pp. 155-188. Springer, Dordrecht, The Netherlands.

Guillaumet A, Pons J-M, Godelle B, Crochet P-A (2006) History of the Crested Lark in the Mediterranean region as revealed by mtDNA sequences and morphology. *Molecular Phylogenetics and Evolution*, **39,** 645–656.

Guiller A, Coutellec-Vreto MA, Madec L, Jeuniff J (2001) Evolutionary history of land snail *Helix aspersa* on the Western Mediterranean: preliminary results inferred from mitochondrial DNA sequences. *Molecular Ecology*, **10**, 81–87.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series,* **41**, 95–98.

Harris DJ, Batista V, Carretero MA (2004a) Assessment of genetic diversity within *Acanthodactylus erythrurus* (Reptilia: Lacertidae) in Morocco and the Iberian Peninsula using mitochondrial DNA sequence data. *Amphibia-Reptilia,* **25,** 227–232.

Harris DJ, Batista V, Lymberakis P, Carretero MA (2004b). Complex estimates of evolutionary relationships in *Tarentola mauritanica* (Reptilia: Gekkonidae). *Molecular Phylogenetics and Evolution,* **30**, 855–859.

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.

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 (2000) The genetic legacy of the quaternary Ice Ages. *Nature*, **405**, 907–913.

Ho SYW, Larson G (2006) Molecular clocks: when times are a-changin'. *Trends in Genetics*, **22**, 79–83.

Hugall A, Moritz C, Moussalli A, Stanisic J (2002) Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875). *Proceedings of the National Academy of Sciences, USA.* **99,** 6112–6117.

Jolly D, Harrison SP, Damnati B, BonneWlle R (1998) Simulated climate and biomes of Africa during the late quaternary: comparison with pollen and lake status data. *Quaternary Science Reviews*, **17**, 629–657.

Jong H de (1998) In search of biogeographic patterns in the western Mediterranean terrestrial fauna. *Biological Journal of the Linnean Society*, **65**, 99–164.

Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, **149**, 429–434.

Kumar S (1996) PHYLTEST: a program for testing phylogenetic hypotheses. Version 2.0. Institute of Molecular Evolutionary Genetics and Department of Biology, Pennsylvania State University, University Park.

Lessa E, Cook JA, Patton JL (2003) Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. *Proceedings of the National Academy of Sciences, USA*, **100**, 10331–10334.

Martínez-Solano I, Gonçalves H, Arntzen JW, García-París M (2004) Phylogenetic relationships and biogeography of midwife toads (Discoglossidae: *Alytes*). *Journal of Biogeography,* **31**, 603–618.

Martínez-Solano I, Teixeira J, Buckley D, Garcia-París M (2006) Mitochondrial DNA phylogeography of *Lissotriton boscai* (Caudata, Salamandridae): evidence for old, multiple refugia in na Iberian endemic. *Molecular Ecology* **13,** 3375–3388.

Michaux JR, Libois R, Filippucci M-G (2005) So close and so different: comparative phylogeography of two small mammal species, the Yellow-necked fieldmouse (*Apodemus flavicollis*) and the Woodmouse (*Apodemus sylvaticus*) in the Western Palearctic region. *Heredity*, **94**, 52–63.

Milá B, Girman DJ, Kimura M, Smith TB (2000) Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American songbird. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **267**, 1033–1040.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Oliverio M, Bologna MA, Mariottini P (2000) Molecular biogeography of the Mediterranean lizards *Podarcis* Wagler, 1830 and *Teira* Gray, 1838 (Reptilia, Lacertidae). *Journal of Biogeography,* **27**, 1403–1420.

Paulo OS (2001) *Phylogeography of reptiles of the Iberian Peninsula*. PhD Thesis. University of London.

Paulo OS, Dias C, Bruford MW, Jordan WC, Nichols RA (2001) The persistence of Pliocene populations through the Pleistocene climatic cycles: evidence from the phylogeography of an Iberian lizard. *Proceedings of the Royal Society of London*. *Series B, Biological Sciences*, **268**, 1625–1630.

Paulo OS, Jordan WC, Bruford MW, Nichols RA (2002) Using nested clade analysis to assess the history of colonization and the persistence of populations of an Iberian lizard. *Molecular Ecology*, **11**, 809–819.

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 (in press) 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*, in press.

Pleguezuelos JM, Márquez R, Lizana M (eds). (2002) *Atlas y libro rojo de los anfíbios y reptiles de España*. Dirección General de Conservación de la Naturaleza – Asociación Herpetologica Española, 2nd Edition, Madrid,

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

Poulakakis N, Lymberakis P, Valakos E, Pafilis P, Zouros E, Mylonas M (2005) Phylogeography of Balkan wall lizard (*Podarcis taurica)* and its relatives inferred from mitochondrial DNA sequences. *Molecular Ecology,* **14**, 2433–2443.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

Roucoux KH, de Abreu L, Shackleton NJ, Tzedaki PC (2005) The response of NW Iberian vegetation to North Atlantic climate oscillations during the last 65 kyr. *Quaternary Science Reviews*, **24**, 1637–1653.

Rosenberg NA, Feldman MW (2002) The relationship between coalescence times and population divergence times. In: *Modern Developments in Theoretical Population Genetics: the Legacy of Gustave Malécot* (Slatkin M, Veuille M, eds.), pp. 130–164. Oxford University Press, Oxford.

Rozas J, Sánchez-Del Barrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Sá-Sousa P (2001) Comparative chorology between *Podarcis bocagei* and *P. carbonelli* (Sauria: Lacertidae) in Portugal. *Revista Española de Herpetología* **15**, 85–97.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Press, New York.

Schneider S, Roessli D, Excoffier L (2000). *Arlequin ver. 2.000: A software for population genetics data analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Swofford DL (2002) *PAUP* ver 4.0.b10. Phylogenetic analysis using parsimony and other methods*. Sinauer Associates, Sunderland, Sunderland, MA.

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

Takezaki N, Rzhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized trees. *Molecular Biology and Evolution,* **12**, 823–833.

Teixeira J (1999) *Biogeografia da salamandra-lusitânica,* Chioglossa lusitanica*. Utilização de Sistemas de Informação Geográfica na modelação da distribuição*. MSc Thesis. University of Porto.

Vasconcelos 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.

Veith M, Mayer C, Samroui B, Donaire Barroso D, Bogaerts S (2004) From Europe to Africa and vice-versa: evidence for multiple intercontinental dispersal in ribbed salamanders (Genus *Pleurodeles*). *Journal of Biogeography*, **31**,159–171

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology,* **7**, 256–276.

Wilder JA, Mobasher Z, Hammer MF (2004) Genetic Evidence for Unequal Effective Population Sizes of Human Females and Males. *Molecular Biology and Evolution,* **21**, 2047–2057.

Willis K, Whittaker RJ (2000) Paleoecology: the refugial debate. *Science* **287**, 1406–1407.

Zagwijn WH (1992) The beginning of the Ice Age in Europe and its major subdivisions. *Quaternary Science Reviews,* **11**, 583–591.

Appendix 1. Geographical distribution of the ND4-tRNA(Leu) haplotypes found in *P. bocagei*. See table 1 and figures 1 and 2 for population and haplotype number correspondence

Haplotype	GenBank	Populations																		total
	Accession No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
B1	EF081116	2	-	-	-	-	6	-	-	-	-	4	-	6	1	6	-	-	6	2	33
B2	EF081117	-	-	-	3	6	-	-	-	-	5	-	-	-	-	-	-	-	-	2	16
B3	EF081118	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	2
B4	EF081119	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	1
B5	EF081120	-	-	-	-	-	-	-	6	-	-	-	-	-	-	-	-	-	-	-	6
B6	EF081121	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	1
B7	EF081122	2	7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9
B8	EF081123	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	1
B9	EF081124	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	1
B10	EF081125	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	1
B11	EF081126	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	1
B12	EF081127	-	-	1	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	2
B13	EF081128	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	1
B14	EF081129	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1
B15	EF081130	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1
B16	EF081131	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1
B17	EF081132	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	1
B18	EF081133	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	2
B19	EF081134	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	1

Appendix 2. Geographical distribution of the ND4-tRNA(Leu) haplotypes found in *P. carbonelli*. See table 1 and figures 1 and 2 for population and haplotype number correspondence.

Haplotype	GenBank	Population																	total
	Accession No.	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
C1	EF081135	7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7
C2	EF081136	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3
C3	EF081137	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
C4	EF081138	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	9	-	11
C5	EF081139	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
C6	EF081140	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
C7	EF081141	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
C8	EF081142	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
C9	EF081143	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5	5
C10	EF081144	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
C11	EF081145	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	1
C12	EF081146	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	1
C13	EF081147	-	-	-	-	-	-	-	-	-	1	1	-	6	-	-	-	-	8
C14	EF081148	-	-	-	-	-	-	-	-	-	-	2	-	1	-	-	-	-	3
C15	EF081149	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	2
C16	EF081150	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	1
C17	EF081151	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1
C18	EF081152	-	-	8	-	1	-	1	8	8	1	5	-	-	-	-	-	-	32
C19	EF081153	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	1
C20	EF081154	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1
C21	EF081155	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	1

Appendix 3. Geographical distribution of the ND4-tRNA(Leu) haplotypes found in *P. vaucheri*. See table 1 and figures 1 and 2 for population and haplotype number correspondence.

Haplotype	GenBank	Population																						total
	Accession No.	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	
V1	EF081075	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
V2	EF081076	1	1	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7
V3	EF081077	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2
V4	EF081078	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
V5	EF081079	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
V6	EF081080	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	2
V7	EF081081	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	1
V8	EF081082	-	-	-	-	-	-	-	-	-	-	-	-	1	1	-	-	-	-	-	-	-	-	2
V9	EF081083	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	1
V10	EF081084	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	1
V11	EF081085	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	1
V12	EF081086	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	1
V13	EF081087	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6	-	-	-	-	-	-	-	6
V14	EF081088	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	5
V15	EF081089	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7	-	7
V16	EF081090	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2	-	-	-	4
V17	EF081091	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1
V18	EF081092	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	1	-	-	-	2
V19	EF081093	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1
V20	EF081094	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1	-	-	-	-	2
V21	EF081095	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1
V22	EF081096	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1
V23	EF081097	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	2
V24	EF081098	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	1
V25	EF081099	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	2
V26	EF081100	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1
V27	EF081101	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2
V28	EF081102	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2
V29	EF081103	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	1
V30	EF081104	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	2
V31	EF081105	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	1
V32	EF081106	-	-	-	-	-	-	-	-	1	2	-	-	-	-	-	-	-	-	-	-	-	-	3
V33	EF081107	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	1
V34	EF081108	-	-	-	-	-	-	-	-	-	4	-	-	-	-	-	-	-	-	-	-	-	-	4
V35	EF081109	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	1
V36	EF081110	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1
V37	EF081111	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1
V38	EF081112	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
V39	EF081113	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
V40	EF081114	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
V41	EF081115	-	-	-	-	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2

Isolation and characterization of nine microsatellite loci in *Podarcis bocagei* (Squamata: Lacertidae)

CATARINA PINHO[1,2], FERNANDO SEQUEIRA[1,2], RAQUEL GODINHO[1,3], D. JAMES HARRIS[1] AND NUNO FERRAND[1,2]

[1]*CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal*
[2]*Departamento de Zoologia e Antropologia – Faculdade de Ciências da Universidade do Porto, Praça Gomes Teixeira, 4099-002, Porto, Portugal*
[3]*Centro de Biologia Ambiental, Faculdade de Ciências, Universidade de Lisboa, Bloco C2, Piso 3, 1749-116 Lisboa, Portugal*

Abstract

Nine dinucleotide microsatellite loci were developed through an enrichment protocol for Bocage's wall lizard, *Podarcis bocagei* Seoane 1884, a lacertid endemic to the Iberian Peninsula. Nineteen primer pairs were designed and tested. From these, nine loci yielded satisfactory results and were screened on 15–19 individuals. These loci revealed a high level of polymorphism (8–15 alleles) and heterozygosity (0.611–0.947) and will certainly be useful in the study of population structure and evolutionary history of this species.

Keywords: Lacertidae; microsatellites; *Podarcis*; polymorphism; wall lizards.

Podarcis is a lacertid genus comprising at least 15 species, which is widely distributed across Europe and North Africa, including some Mediterranean and Atlantic islands. In the Iberian Peninsula, these lizards have been extensively studied using morphological (Harris & Sá-Sousa 2001), mitochondrial DNA (Harris & Sá-Sousa 2002) and protein electrophoretic (Pinho *et al.* 2003) data. Although these studies were useful in differentiating evolutionary entities, pointing out the need for a taxonomical revision, the applied markers gave little or no information on differentiation within each of the described forms. Neutral molecular markers with higher variability levels are therefore required to study in more detail the evolutionary history of each species, especially in determining population structure, gene flow, glacial refugia and potentially important conservation areas.

To this end we have developed a set of microsatellite markers in *Podarcis bocagei,* an endemic species from northwest Iberia, with the expectation that in the future these markers could be applied in other species of this genus.

Genomic DNA, extracted from tail tissue of a single individual, was digested with *Mbo*I. Fragments with the desired size (300-750 bp) were isolated and submitted to an enrichment protocol for dinucleotide microsatellite loci (GT and AC motifs) following Armour *et al.* (1994). A partial genomic library containing sequences enriched for these particular repeats was constructed by ligation of the selected fragments to pUC19 vector followed by transformation in *E. coli* competent cells. Colonies were transferred to Hybond-N+ nylon membranes (Amersham) by colony-lift and hybridized with labelled probes containing the GT/CA and AG/TC tandem arrays. Labelling of probes and detection of positive clones were performed using ECL direct nucleic acid labelling and detection system (Amersham-Pharmacia Biotech) based on enhanced chemiluminescence. Positive clones were polymerase-chain reaction (PCR)-amplified using universal M13 primers and re-hybridised in order to exclude false positives. A total of 29 positive clones were sequenced using the ABI Prism BigDye Terminator Cycle sequencing protocol in an ABI Prism 310 automated sequencer (Applied Biosystems). Primers were designed by eye in the flanking regions of 19 microsatellite loci.

DNA was extracted from individuals of *P. bocagei* belonging to the same population, following standard procedures (Sambrook *et al.* 1989). PCR was carried out in 10 µL volumes, containing 1 µL of reaction buffer [166 mM $(NH_4)_2SO_4$, 670 mM Tris-HCl, 0.1% Tween-20, pH=8.8; Ecogen], 1–1.5 mM $MgCl_2$, 5 µM each dNTP, 0.2 µM each primer, 0.5 U of *Ecotaq* DNA polymerase (Ecogen) and approximately 25 ng of genomic DNA. For the microsatellite *Pb47*, better results were accomplished adding 5% dimethyl sulfoxide in PCR reactions. PCR products were visualized in 2% agarose gels stained with ethidium bromide. Thirteen loci were successfully amplified within the expected size range and

were run on 6% denaturing polyacrylamide gels and visualized by silver staining. Four of these did not show interpretable patterns and were excluded from further analyses. The results obtained for the remaining nine loci are shown in Table 1.

Two microsatellite loci are compound (*Pb10* and *Pb73*) and the other seven are simple dinucleotides. In six out of the nine microsatellites typed (*Pb10, Pb20, Pb37, Pb47, Pb55, Pb66*), both even and odd-sized alleles were detected, suggesting the occurrence of insertions or deletions in the flanking regions.

Observed and expected heterozygosities were calculated using the GENETIX software (version 4.04, Belkhir *et al.* 1996-2002). Tests for Hardy-Weinberg and linkage disequilibria were performed using GENEPOP (version 3.3, Raymond & Rousset 1995). The observed allelic richness ranged from 8 (*Pb55*) to 15 alleles (*Pb10* and *Pb47*), with a mean of 11.4, and observed heterozygosities ranged from 0.611 (*Pb11*) to 0.947 (*Pb47* and *Pb66*), averaging 0.827 across loci. Of the nine markers, one (*Pb73)* showed a deviation to Hardy-Weinberg's expectations ($p<0.05$). It is common that the presence of null alleles is invoked to explain such deviations in microsatellite markers. However, in this marker, heterozygote deficiency, tested using GENEPOP option 1.1, was not significant ($P < 0.05$). An alternative hypothesis can be an insufficient sampling, especially if one considers the high number of alleles at this locus. Linkage disequilibrium between pairs of loci was not detected ($P < 0.05$).

The high levels of polymorphism observed at these markers contrast with the very low levels of variation described in mitochondrial DNA and morphology (Harris & Sá-Sousa 2001, 2002). Therefore, this set of loci will certainly be useful in uncovering the yet to be studied evolutionary history of *Podarcis bocagei*.

Acknowledgements

This work was partially financed by Fundação para a Ciência e a Tecnologia (FCT), through research projects POCTI/BSE/34547/2000 and POCTI/BSE/40987/2001, and SFRH/BD/4620/2001, SFRH/BD/3365/2000, PRAXIS XXI/BD/16087/98 and SFRH/BPD/5702/2001 grants to CP, FS, RG and DJH, respectively. We are further grateful to Steven Weiss and Barbara Wallner for their help in laboratory work and guidance in the cloning protocol.

Table 1. Characterization of nine microsatellite loci isolated from *Podarcis bocagei*.

Locus	Accession numbers	Repeat motif	Primer sequences (5′ – 3′)	T_a (°C)	Allele size (bp)	*N*	No. of alleles	H_o	H_E	H-W (*P*-value)
Pb10	AY545220	$(GT)_nGC(GT)_nGC(GT)_n(AG)_n$	AGT GGA ATC GGC TGC AAT AC ACC AGT CCC AGG AAT TTA GG	56	178 – 204	18	15	0.889	0.864	0.123
Pb11	AY545221	$(TG)_n$	TTT CTG GGA GGA GAA GAC AC CTG GAA GAA CAC AGC AGG AG	56	152 – 180	18	9	0.611	0.711	0.217
Pb20	AY545222	$(AC)_n$	ACG CAA AGT CTC TCC ACA CC CTT TGG CAG CTT CTT GCT TC	57	124 – 155	18	10	0.889	0.843	0.914
Pb37	AY545223	$(CA)_n$	GAG AGT ATA CCA ACC GTG CTA ATG CTG GAA CTA TCC	54	129 – 158	18	11	0.778	0.858	0.153
Pb47	AY545224	$(GT)_n$	CTT GGT GGT TAA CAA TGT TGG C GTG AGC TAA TAC AAC TCT CCA C	56	203 – 238	19	15	0.947	0.895	0.917
Pb50	AY545225	$(CA)_n$	GGA TGT TTC AGC ATG CTT GG AGA CCT CAC TGG GCC ATT AC	54	113 – 135	18	12	0.833	0.877	0.063
Pb55	AY545226	$(TG)_n$	CCC ATC CTA ACC CTT ACC TTT G GCA GCT CCA TCA CTG GCC CTG	55	228 – 242	15	8	0.667	0.822	0.127
Pb66	AY545227	$(TG)_n$	GGA CAG CTA GTC CCA TGG CTT AC GGA TTG CTG TCA CCA GTC TCC CC	55	138 – 171	19	12	0.947	0.889	0.968
Pb73	AY545228	$(CA)_nCT(CA)_n$	GCC CAT GTC ACT TCA GGT AGA AGC GAA AAC TAG GAG TTA GGG AGA AGG	58	146 – 178	18	11	0.889	0.892	0.017

T_a, annealing temperature; *n*, number of individuals analysed; H_o, observed heterozygosity; H_E, expected heterozygosity; H-W (*P*-value), Hardy-Weinberg probability

References

Armour JAL, Neumann R, Gobert S, Jefferys AJ (1994) Isolation of human simple repeat loci by hybridization selection. *Human Molecular Genetics,* **3**, 599–605.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2002) GENETIX 4.04, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

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.

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.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): a population genetics software for exact tests and ecumenicism. *Journal of Heredity,* **86**, 248–249.

Sambrook J, Fritsch EF, Maniatis T, eds (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Press, New York.

Combining mtDNA, allozyme polymorphisms and microsatellites to assess the evolutionary histories of two endemic lacertids (*Podarcis bocagei* and *P. carbonelli*) of Western Iberian Peninsula

CATARINA PINHO, D. JAMES HARRIS AND NUNO FERRAND

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

Abstract

Podarcis bocagei and *P. carbonelli* are two Iberian-endemic wall lizard species for which a detailed phylogeographic study based on mitochondrial DNA (mtDNA) variation has been accomplished, showing that the two species responded differently to climatic oscillations during the Quaternary. In this study, our main goal was to assess whether nuclear markers reveal similar evolutionary patterns to those inferred from the study of mtDNA variation. We studied a battery of allozyme and microsatellite loci in populations from both species. For each species we evaluated overall levels of differentiation, the patterns of variation in genetic variability, genetic relationships among populations, changes in allele frequencies across the species' distribution area and applied model-based individual multilocus genotype clustering approaches to detect hidden population structure. Our results for *Podarcis bocagei* are highly concordant with the phylogeographic scenario inferred from mtDNA variation: not only we find very low levels of population differentiation, consistent with survival in a single glacial refugium, but we detect signatures of a rapid demographic and geographic expansion, as suggested from mtDNA analyses. The analyses of nuclear markers furthermore helped in the identification of the probable refugial area, as well of expansion routes. Also with similarity to observations based on mtDNA variation, a low level of population differentiation was observed in *P. carbonelli.* In concordance with mitochondrial DNA, differentiation levels were slightly higher in *P. carbonelli* than those observed in *P. bocagei*. However, the geographic basis for differentiation in *P. carbonelli* is highly inconsistent between mtDNA and nuclear markers, suggesting a complex, albeit recent, history of isolation. Furthermore, a recent reduction on the species distribution has probably erased the signatures of glacial isolation and post-glacial expansion that are normally found in other Iberian species, suggesting

that the currently observed pattern of genetic differentiation in this species was shaped more by recent genetic drift than by the Pleistocene climatic oscillations.

Keywords: *Podarcis bocagei, Podarcis carbonelli*, allozymes, microsatellites, population structure, glaciations, glacial refugia.

Introduction

The Pleistocene climatic oscillations left considerable signatures on the genetic structure of temperate organisms (Hewitt 1996, 1999, 2000). This period was characterized by an alternation of cold (glacial) and warm (interglacial) stages, which dramatically shifted, over short periods of time, habitat conditions across the globe. Although these changes increased extinction rates, many species were able to survive by moving along with the changing habitat. The genetic consequences of such range contractions and expansions have been well characterized in many organisms. Among the patterns that most commonly emerge are signatures of long-term isolation and persistence in distinct refugia, leading to high levels of population subdivision and, in a more debatable perspective, even speciation (e.g. Avise *et al.* 1998). In the case of European-wide distributed temperate taxa, for example, a common pattern is the presence of highly divergent lineages, signatures of glacial isolation, in southern European peninsulas, and evidence for post-glacial recolonization of northern Europe (classical examples of such scenarios are, for example, the grasshopper *Chorthippus paralellus* (Cooper *et al.* 1995), the hedgehog *Erinaceus europaeus* (Santucci *et al.* 1998) and the brown bear *Ursus arctos* (Taberlet *et al.* 1994); reviewed in Taberlet *et al.* (1998) and Hewitt (1999, 2000)). As a result of post-glacial expansion from these source populations, northern areas often show depleted levels of genetic diversity, a paradigm often known as "northern purity vs. southern richness" (Hewitt 1996, 2000, 2001). On the other hand, cases have also been documented in which northern areas evidence high levels of genetic diversity because of the mixing of waves expanding from different sources (Comps *et al.* 2001, Petit *et al.* 2003). In a number of cases, these differentiated lineages establish hybrid suture zones, the location and orientation of which often reveal remarkably similar patterns amongst different organisms (Taberlet *et al.* 1998, Hewitt *et al.* 2000).

Adding to their important role in functioning as repositories of European-wide genetic variability, studies focusing on southern European Peninsulas have shown that the processes that are seen at a larger scale (isolation in different refugia, post-glacial expansion and secondary contact) are mirrored at smaller scales within Peninsulas. This "refugia-within-refugia" model has been

particularly well-documented in the Iberian Peninsula. Various species have been shown to present highly complex population structure consistent with isolation in distinct glacial refugia (Comes and Abbott 1998, Alexandrino *et al.* 2000, Branco *et al.* 2000, Paulo *et al.* 2001, amongst many others; reviewed in Gómez and Lunt 2007). Accordingly, strong signatures of post-glacial demographic expansion and establishment of secondary contact zones have also been detected (Alexandrino *et al.* 2000, Surget-Groba *et al.* 2001, Branco *et al.* 2002, García-Paris *et al.* 2003, Sequeira *et al.* 2005, Godinho *et al.* 2006a,b). Although most studies indicate a strong level of fragmentation, consistent with long-term survival in relatively stable refugia, the depth of divergence between lineages varies widely among species; furthermore, other endemics reveal a low degree of substructure and a shallow coalescent history, indicating lower persistence levels across the Quaternary glaciations (Pinho *et al.* in press a). Rather than failing to corroborate the paradigm of multiple refugia within Iberia, these cases illustrate the variety of species-specific responses to the same climatic phenomenon.

The objects of this study, Iberian endemic wall lizards *Podarcis bocagei* and *P. carbonelli* are good examples of species exhibiting shallow population subdivision, at least from a mitochondrial DNA perspective (Pinho *et al.* in press a). This lack of phylogeographic structure is extreme in the case of *P. bocagei*, with an inferred coalescence time of around 70,000 – 100,000 years, and has been correlated to its northern distribution within the Iberian Peninsula. In fact, species with natural distribution ranges at more northern latitudes are expected to have experienced more severe habitat changes than species distributed in the south, affecting persistence levels throughout the Ice Ages. Accordingly, the signature of population demographic recovery is also expected to be stronger in these northern populations than in southern populations that were allowed to maintain a relatively high long-term effective population size throughout glacial cycles. This signature is also evident in *Podarcis bocagei*. On the other hand, *P. carbonelli*, a species distributed in Central and Southern Iberia, shows higher levels of genetic variation and subdivision, although reduced when compared to those observed within other Iberian lizard species (e.g. Paulo *et al.* 2001). Some differentiation was observed between haplotypes found in the Iberian Central System, the Northwestern region of the distribution and a geographical isolate in Southern Spain. The southern Portuguese coast was less well sampled but haplotypes detected in this region did not fall into any of the previous groups, although it shared a haplotype, placed in a central position within the network, with a population from the Central System. Only the northwestern clade within this species evidenced strong signatures of post-glacial demographic expansion.

In this study we investigate in more detail the evolutionary history of these two species using two different sets of nuclear markers: allozymes and

microsatellites. By doing so, we test whether the scenarios inferred from mitochondrial DNA variation are validated from an independent source of information. Specifically, we were interested in i) evaluating whether the apparent low degree of genetic substructure (when compared to that observed in several other species, including *Podarcis vaucheri*, (Pinho *et al.* in press a)) is reflected by a lack of subdivision in nuclear loci; ii) testing hypotheses related to the partition of genetic variability in both species, namely whether mtDNA-defined clusters in *P. carbonelli* correspond to distinct entities from a nuclear gene perspective; iii) evaluate the likelihood of the scenarios of post-glacial expansion inferred for *P. bocagei* and some geographic regions in *P. carbonelli*, which could be alternatively attributed to a recent selective sweep affecting their mitochondrial genome instead of to a simple demographic effect; iv) compare the patterns of genetic diversity observed in both species, testing whether the differences in their distribution (*P. bocagei* is distributed northerly to *P. carbonelli* and has a continuous distribution range, whereas *P. carbonelli* shows a much more fragmented range and at least one geographical isolate) are reflected in different evolutionary scenarios.

Materials and Methods

Sampling. We sampled individuals from a total of 8 populations (N=176) for *P. bocagei* and 9 (N=157) for *P. carbonelli*. Samples consisted of a portion of tail tissue obtained from the lizards' natural autotomy. All individuals were released after sample collection. Samples were stored frozen at -80ºC (for allozyme analyses) and in 96% ethanol (for DNA extraction). Sampling details (location, sample sizes and geographical coordinates) are given in table 1 and figure 1. Individuals from all the sampled populations were included in a previous article addressing the mitochondrial DNA phylogeography of these two species (Pinho *et al.* in press). Some of the data included in this work (allozyme data for the *P. bocagei* populations of Madalena, Vairão, Zimão and A Coruña and for the *P. carbonelli* populations of Villasrúbias, Esmoriz, Aveiro, S. Pedro de Moel and Playa de Rompeculos; partial microsatellite data sets for Zimão, Vairão, Madalena, Esmoriz and Aveiro) were also used in previous publications (Pinho *et al.* 2003, 2004a,b; Pinho *et al.* in press b; Pinho *et al.* submitted).

Table 1. Sampling details for this study. Sample codes refer to those in figure 1.

Species	Pop ID	Locality	Country	Latitude/longitude	Sample size	
					Allozymes	STRs
P. bocagei						
	Mad	Madalena	Portugal	41º 06′ N 08º 39′ W	22*	25*
	Vair	Vairão	Portugal	41º 19′ N 08º 40′ W	33*	27*
	Bra	Braga	Portugal	41º 32′ N 08º 25′ W	11	21
	Zim	Zimão	Portugal	41º 27′ N 07º 40′ W	18*	18*
	Mon	Montesinho	Portugal	41º 55′ N 06º 46′ W	30*	29
	San	Sanxenxo	Spain	42º 24′ N 08º 49′ W	10	10
	Cor	A Coruña	Spain	43º 22′ N 08º 23′ W	16*	16
	Sar	Sarria	Spain	42º 47′ N 07º 24′ W	20	20
					160	166
P. carbonelli						
	LA	La Alberca	Spain	40º 28′ N 06º 05′ W	16	16
	VR	Villasrúbias	Spain	40º 19′ N 06º 37′ W	21*	22*
	SE	Serra da Estrela	Portugal	40º 23′ N 07º 31′ W	15	15
	Esm	Esmoriz	Portugal	40º 57′ N 08º 38′ W	18*	18
	Av	Aveiro	Portugal	40º 37′ N 08º 45′ W	17*	17*
	SPM	S. Pedro de Moel	Portugal	39º 45′ N 09º 01′ W	22*	22
	CR	Cabo Raso	Portugal	38º 42′ N 09º 28′ W	19	19
	MC	Monte Clérigo	Portugal	37º 19′ N 08º 48′ W	16	16
	PR	Playa del Rompeculos	Spain	37º 06′ N 06º 45′ W	12*	12
					156	157

*includes previously published data (Pinho et al. 2003, Pinho et al. 2004b, Pinho et al. in press b)

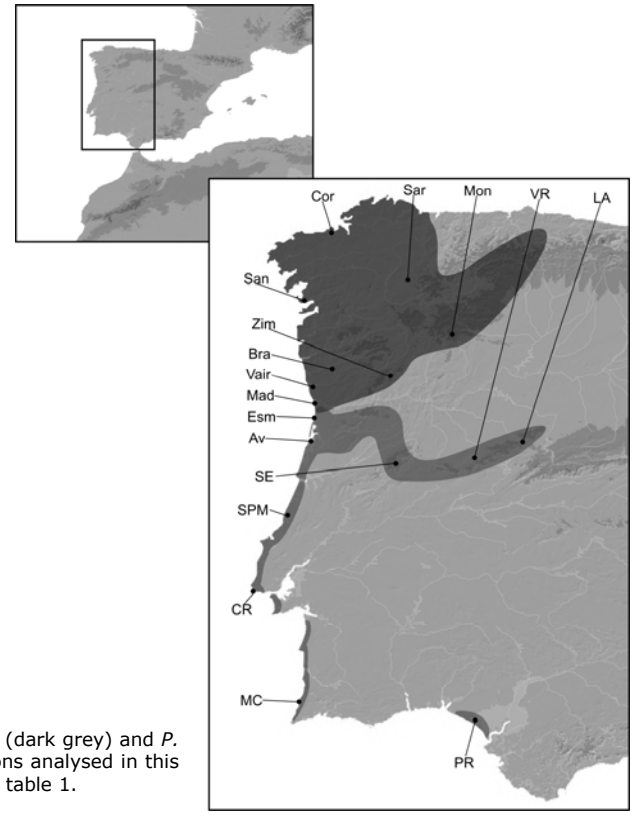

Figure 1. Distribution of *P. bocagei* (dark grey) and *P. carbonelli* (light grey) and populations analysed in this study. Sampling details are given in table 1.

Data collection. For allozyme analyses, tissue extraction, protein separation and enzymatic detection of all loci followed the procedures given in Pinho *et al.* (2003). Variation at a total of 10 allozyme loci was screened by means of conventional starch gel electrophoresis (*GOT*, *GPI*, *IDH*, *MPI*, *PEPA*, *PEPD*, *6-PGD*) and isoelectric focusing (*LDH-2*, *PEPB*, *PGM*). For microsatellite analyses, DNA was extracted following standard procedures (Sambrook *et al*. 1989). For *P. bocagei* individuals, we studied six dinucleotide microsatellite loci from the battery of nine developed for this species by Pinho *et al*. (2004b) (*Pb11, Pb37, Pb47, Pb50, Pb66* and *Pb73)*. Only a subset of these loci (*Pb11*, *Pb47*, *Pb50*, *Pb66*) were analysed in *P. carbonelli*. These loci were amplified according to the described conditions with the exception of the annealing temperature, that was lowered to 53ºC in all cases in order to amplify difficult samples. The electrophoretic separation of the amplified fragments was carried out in 6% denaturing polyacrylamide gels and silver stained as described in Pinho *et al.* (2004b).

Analytical methods. Allele frequencies and heterozygosity for each population were calculated using GENETIX v. 4.05.2 (Belkhir *et al*. 1996-2004). In order to correctly compare genetic diversity across populations of each species, we computed allelic richness (sensu El Mousadik and Petit 1996, Petit *et al.* 1998) by applying a rarefaction methodology using the software CONTRIB 1.01 (Petit *et al.* 1998). This method diminishes (but does not completely eliminate) the effect of differential sample sizes on the estimates of diversity. We used 18 and 22 chromosomes (for *P. bocagei* and *P. carbonelli*, respectively) as rarefaction sizes in these analyses, corresponding to the smallest analysed population size across all markers. Variability measures were averaged across allozyme and microsatellite loci independently. The GENEPOP software (v. 3.1b; Raymond and Rousset 1995a) probability test was used to determine whether populations were in Hardy-Weinberg and linkage equilibrium and also to perform an exact test for evaluating the significance of allele frequency differences across all populations and also between each population pair (Raymond and Rousset 1995b). We also used this software to calculate overall F_{ST} values (according to Weir and Cockerham 1984) for each marker in each species. The software ARLEQUIN v. 2.000 (Schneider *et al.* 2000) was used to compute F_{ST} statistics between all population pairs (as well as their significance). Genetic relationships among populations of each species were estimated using Neighbor-Joining (NJ, Saitou and Nei 1987) trees based on Cavalli-Sforza and Edwards (1967) chord distance obtained from the combined data sets, using PHYLIP 3.6 (Felsenstein 1993). Bootstrap support was evaluated through 1000 pseudoreplicates. To test hypotheses related to the geographic partition of genetic variation we used ARLEQUIN to perform hierarchical Analyses of Molecular Variance (AMOVA,

Excoffier *et al.* 1992) using the combined allozyme and microsatellite data sets. In these analyses we partitioned each data set into several combinations of subsets, searching for structures that produced the highest fraction of variation among groups (that is, that maximized Φ_{CT}). We tested hypotheses related to the influence of geographical barriers, mitochondrial DNA-defined subdivision and relevant geographic clusters in the total genetic diversity. Because we were aiming at detecting genetic substructure concordant with geographic partitions, we focused mainly on clusters involving geographically close populations. A second approach attempting to correlate genetic variation with geography consisted of using GENETIX to perform a Factorial Correspondence Analysis on populations of each species using information from all markers simultaneously; the data retrieved was used to construct synthetic maps using a kriging method. This was performed using Surfer 8.0 (Golden Software).

These analyses used populations as operational units. To further explore the patterns of intraspecific variability we used a model-based individual multilocus genotype clustering method implemented in the software STRUCTURE 2.1 (Pritchard *et al.* 2000). By doing so we were able to search for hidden population structure and evaluate how this population structure is distributed across geographic space. For these analyses, in order to minimize missing data, we used only individuals analysed for both classes of markers, a total of 150 *P. bocagei* and 151 *P. carbonelli* individuals. The parameter settings included the assumption of admixture and a correlated allele frequencies model, which assumes that allele frequencies are similar across populations and that variation between them arose mainly by genetic drift rather than mutation. Although this model seemed appropriate for our data given both the shallow substructure inferred for both species from mtDNA variation and the low differentiation among populations detected in this study (see Results), because of recent suggestions that the choice of the model might strongly influences the outcome of the clustering algorithm (Serre and Pääbo 2004, Rosenberg *et al.* 2005) we repeated the runs without assuming correlated allele frequencies. We tested scenarios assuming from one to 10 or 11 clusters (in *P. bocagei* and *P. carbonelli*, respectively, representing the actual number of populations sampled plus two). STRUCTURE was run for 500,000 steps after 200,000 iterations, discarded as burn-in. For each value of K, five independent replicates of the MCMC were conducted. We used the software DISTRUCT (Rosenberg *et al.* 2002) to obtain graphical representations of the inferred genetic structure. To choose amongst the various possible values of K the one that more accurately characterized each data set, we used the method suggested by Evanno *et al.* (2005), which searches for a mode in the distribution of ΔK, a quantity related to the second order rate of change of the log probability of the data.

Results

Genetic diversity.
All ten allozyme loci studied were polymorphic in *P. bocagei*, whereas in *P. carbonelli* locus *GPI* was found to be monomorphic and was therefore not included in analyses of differentiation within this species. A total of 31 alleles were detected in the 10 polymorphic loci in *P. bocagei*, with variability levels ranging from two (in *GPI, IDH, PEPB*) to four (in *LDH-2*, *PEPA*, *PEPD*, *PGD*) alleles per locus. Thirty-six alleles were detected in the studied populations of *P. carbonelli* amongst the 9 polymorphic loci, ranging from two allelic variants in *IDH* and *MPI* to six alleles observed in *PEPD*. Allelic frequencies for each population are given in tables 2 and 3 for *P. bocagei* and *P. carbonelli*, respectively. Microsatellite loci showed remarkably high levels of variation, with a minimum of 19 alleles observed per locus in *P. bocagei* (*Pb11*, *Pb50*) and 15 in *P. carbonelli* (*Pb47*, *Pb66*), whereas the loci with the highest number of alleles detected were *Pb37* in *P. bocagei* (36 alleles) and *Pb50* in *P. carbonelli* (31 alleles). Microsatellite allele frequencies obtained for all populations from both species are given in figures 2 and 3 (for *P. bocagei* and *P. carbonelli*, respectively). With the exception of *Pb11*, all loci analysed showed both even and odd allele sizes. This suggests that insertions or deletions in the flanking regions may have occurred. Taking these results together with the multimodal distribution of overall allele frequencies in most markers, this additionally implies that a strict stepwise mutation model of evolution (Ohta and Kimura 1973) cannot be properly applied to our microsatellite data sets. We thus abstained from performing analyses that assume this model to compute divergence or differentiation between populations.

Several cases of strong deviations from Hardy-Weinberg expectations ($p<0.01$) were detected, all in microsatellite loci in *P. bocagei*: *Pb37* in Madalena, Montesinho and Sarria; *Pb50* in Vairão and Montesinho; and *Pb66* in A Coruña. Because these deviations could be signatures of the presence of null alleles, we also performed an exact test for heterozygote deficit, which yielded significant results among the above reported cases (*Pb37* in Madalena and Montesinho and *Pb50* in Vairão and Montesinho). Other cases of less strong but still significant ($p<0.05$) deviations were also detected in allozyme (*PEPD* in Vairão, *PGD* in La Alberca) and microsatellite loci (*Pb11* in Monte Clérigo, *Pb37* in Braga, *Pb47* in Serra da Estrela, *Pb50* in Braga, *Pb73* in Zimão and Sarria). All loci were in linkage equilibrium across populations.

Table 2. Allele frequencies for polymorphic allozyme loci in *P. bocagei* populations.

		Population							
		Mad	**Vair**	**Bra**	**Zim**	**Mon**	**San**	**Cor**	**Sar**
GOT1	*N*	22	31	11	15	20	10	16	20
	A	1.00	0.95	1.00	0.97	0.93	1.00	1.00	0.83
	B		0.03		0.03	0.07			0.17
	D		0.02						
GPI	*N*	22	33	11	18	28	10	16	20
	A	0.98	0.98	0.91	1.00	1.00	1.00	1.00	1.00
	B	0.02	0.02	0.09					
IDH	*N*	22	33	11	17	28	10	16	20
	A	0.86	0.95	0.91	1.00	1.00	1.00	1.00	1.00
	B	0.14	0.05	0.09					
LDH2	*N*	22	33	10	18	28	10	15	17
	A					0.07			
	B	0.11	0.30	0.15	0.05	0.16			0.35
	C	0.07	0.09	0.20	0.03	0.34	0.05	0.03	0.03
	D	0.82	0.61	0.65	0.92	0.43	0.95	0.97	0.62
MPI	*N*	22	25	9	18	30	10	16	20
	A	0.16	0.08	0.17	0.22				
	B	0.84	0.90	0.83	0.78	1.00	0.75	0.50	0.85
	C		0.02				0.25	0.50	0.15
PEPA	*N*	22	33	11	15	22	10	16	20
	B		0.12			0.09			
	C	1.00	0.88	1.00	1.00	0.78	1.00	1.00	1.00
	D					0.11			
	H					0.02			
PEPB	*N*	20	20	9	16	22	9	15	20
	B	0.23	0.02	0.06	0.09	0.05			
	C	0.77	0.98	0.94	0.91	0.95	1.00	1.00	1.00
PEPD	*N*	22	29	10	16	21	10	16	20
	A	0.20	0.38	0.55	0.44	0.60	0.45	0.41	0.65
	B	0.64	0.52	0.45	0.56	0.40	0.55	0.59	0.35
	C		0.10						
	H	0.16							
PGD	*N*	22	33	11	18	29	10	15	20
	A	0.02							
	B	0.64	0.97	0.73	1.00	1.00	1.00	1.00	1.00
	C	0.34	0.03	0.23					
	K			0.04					
PGM	*N*	20	31	11	18	26	10	16	20
	A	1.00	0.93	0.91	1.00	1.00	1.00	1.00	1.00
	B		0.05	0.09					
	C		0.02						

Table 3. Allele frequencies for polymorphic allozyme loci in *P. carbonelli* populations.

		Population								
		LA	**VR**	**SE**	**Esm**	**Av**	**SPM**	**CR**	**MC**	**PR**
GOT1	*N*	16	21	15	16	17	22	18	15	12
	A			0.10	0.22					0.04
	B	1.00	1.00	0.90	0.78	1.00	1.00	0.86	1.00	0.96
	E							0.14		
IDH	*N*	16	21	15	15	14	22	19	16	12
	A	1.00	1.00	1.00	1.00	1.00	1.00	0.45	1.00	1.00
	B							0.55		
LDH2	*N*	16	18	15	18	17	22	19	15	12
	B	0.81	0.53	0.33	0.22	0.38	0.59	0.21	0.90	0.58
	C					0.21				
	D	0.19	0.47	0.67	0.78	0.41	0.39	0.79	0.10	0.42
	F						0.02			
MPI	*N*	16	21	15	16	15	22	19	16	12
	B	1.00	1.00	1.00	1.00	1.00	1.00	0.79	1.00	1.00
	C							0.21		
PEPA	*N*	14	19	14	15	17	22	15	16	12
	A			0.11	0.07	0.09	0.07			
	C	0.79	0.74	0.54	0.23	0.29	0.80	0.77	0.97	0.96
	D	0.14	0.13	0.28	0.30	0.56	0.11	0.23	0.03	0.04
	E			0.07	0.40	0.06	0.02			
	I	0.07	0.13							
PEPB	*N*	16	20	15	15	15	22	19	16	12
	A					0.03				
	B	0.03		0.40			0.09	0.34		
	C	0.59	0.80	0.53	1.00	0.87	0.84	0.53	0.81	0.54
	D	0.13	0.20	0.07		0.10	0.07	0.13	0.19	0.46
	E	0.25								
PEPD	*N*	16	18	13	17	17	22	14	16	12
	A	0.44	0.16	0.58	0.44	0.47	0.25	0.07	0.62	0.38
	D	0.56	0.78	0.31	0.38	0.47	0.55	0.89	0.38	0.62
	E						0.07			
	G				0.15	0.06		0.04		
	H		0.03	0.11	0.03		0.13			
	I		0.03							
PGD	*N*	16	21	15	16	17	22	19	15	12
	A	0.53	0.88	0.83	0.88	0.85	0.88	0.90	1.00	0.75
	B	0.47	0.12	0.14	0.09	0.15	0.07	0.05		
	E				0.03		0.05			0.25
	F			0.03						
	H							0.05		
PGM	*N*	16	21	15	17	17	22	19	16	12
	A	1.00	1.00	1.00	0.94	0.88	0.98	1.00	0.97	1.00
	C								0.03	
	D					0.12				
	G				0.06		0.02			

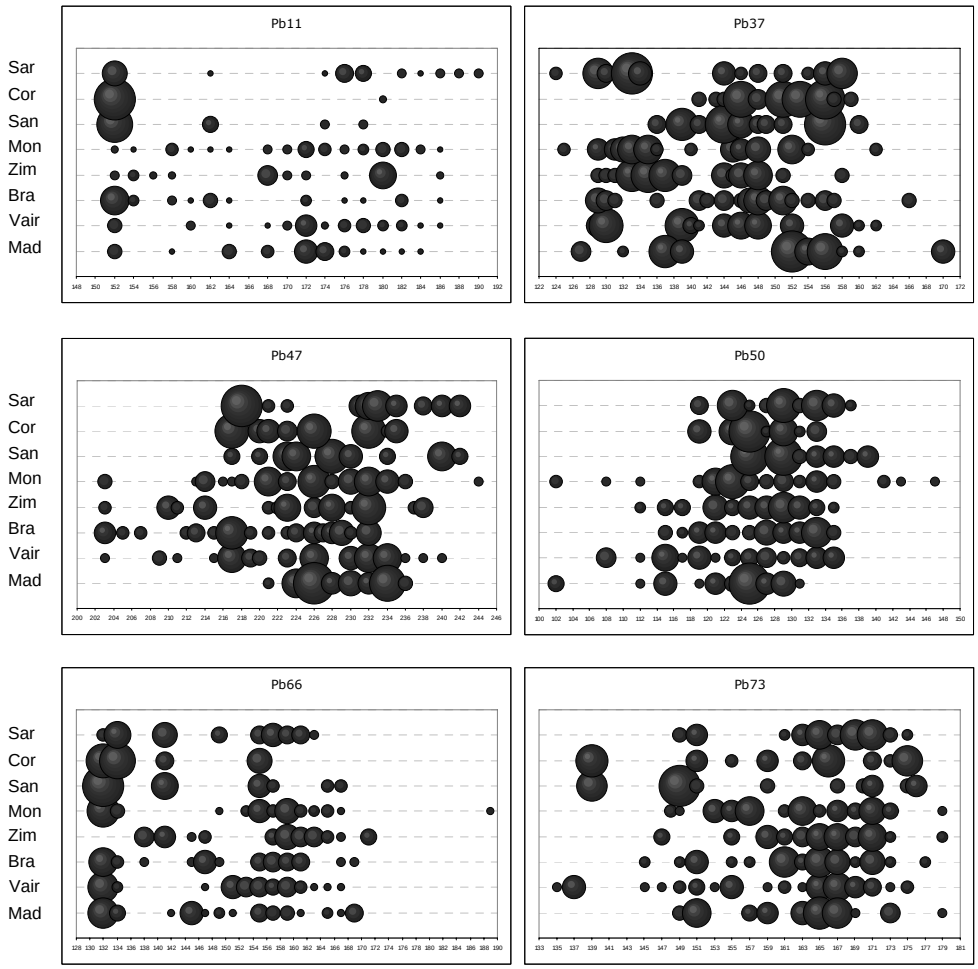

Figure 2. Microsatellite allele frequencies in populations of *Podarcis bocagei*.

Expected heterozygosity and allelic richness levels are given in figures 4 and 5. In *P. bocagei*, a trend suggesting a decrease of genetic diversity from southern to northern populations is observed in both classes of markers. In *P. carbonelli*, however, there appears to be no such observable trend; interestingly, comparative variability levels differ between both classes of markers: for example, the populations of Aveiro and Esmoriz exhibit comparatively high allozyme but relatively low microsatellite diversity, while the opposite is observed in Playa del Rompeculos. There is one exception to these discordant patterns, the population of Monte Clérigo, which consistently exhibits low diversity levels.

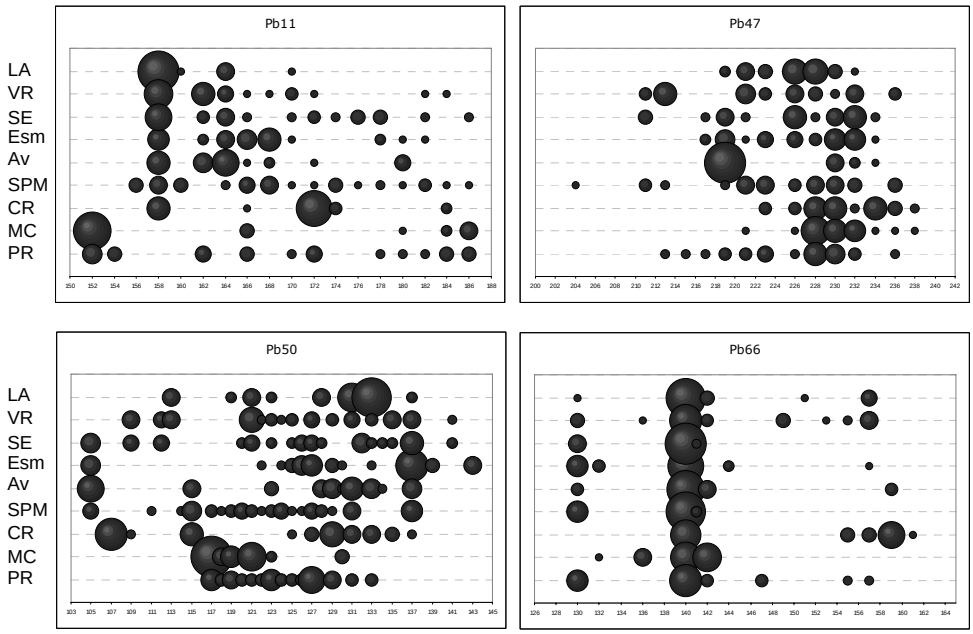

Figure 3. Microsatellite allele frequencies in populations of *Podarcis carbonelli*.

Genetic differentiation among populations

An exact test (Raymond and Rousset 1995b) showed that there are significant differences in allele frequencies, both at an overall level and between each pair of populations for each species. Overall F_{ST} estimates for all markers reveal low levels of differentiation in both species, with an average of 0.0772 for *P. bocagei* and 0.1364 for *P. carbonelli* (Table 4). For comparative purposes, we also computed for *P. bocagei* F_{ST} for the same panel of 13 loci that were polymorphic in *P. carbonelli*, which yielded a value of 0.0868*.* Tables 5 and 6 represent pairwise F_{ST} values between populations of *P. bocagei* and *P. carbonelli*, respectively.

In *P. bocagei*, these range from a minimum of virtually zero, between Vairão and Braga, to a maximum of 0.187, between A Coruña and Montesinho. In *P. carbonelli*, the minimum value was obtained between Villasrúbias and S. Pedro de Moel (0.007), whereas the maximum value (0.246) was observed between the populations of Esmoriz and Monte Clérigo.

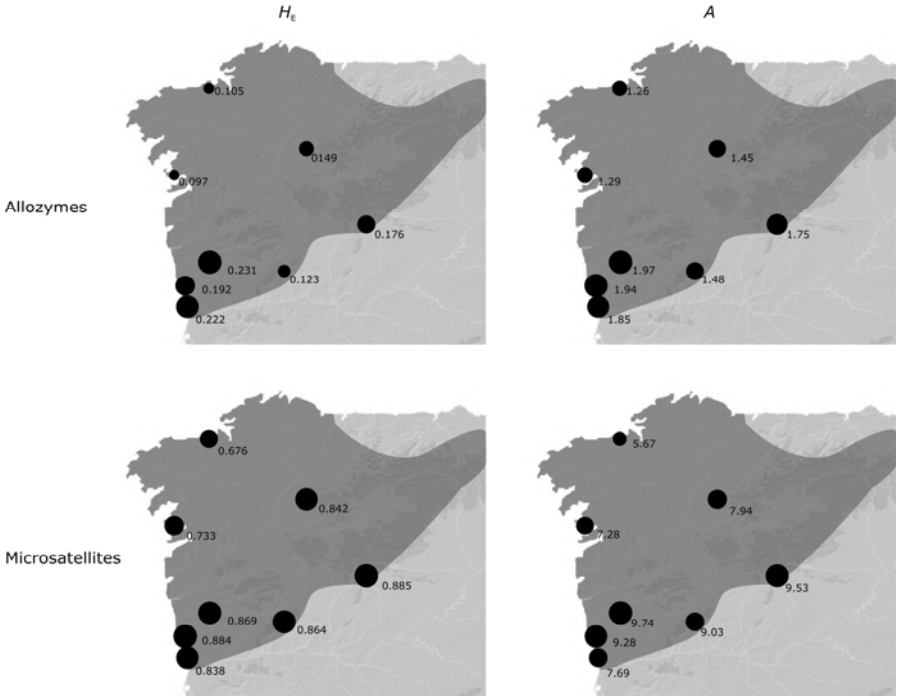

Figure 4. Mean diversity measures (expected heterozygosity (H_E) and allelic richness (A)) in populations of *P. bocagei*.

Table 4. Overall genetic differentiation (F_{st}) in *P. bocagei* and *P. carbonelli*, using allozymes, microsatellites and mitochondrial DNA

		P. bocagei	***P. carbonelli***
Allozymes	*Got1*	0.0549	0.1120
	Gpi	0.0240	-
	Idh	0.0594	0.5313
	Ldh2	0.1322	0.2000
	Mpi	0.1653	0.1823
	PepA	0.0896	0.2001
	PepB	0.0813	0.1359
	PepD	0.0440	0.1034
	Pgd	0.2451	0.1174
	Pgm	0.0223	0.0368
	All	0.1068	0.1706
Microsatellites	*Pb11*	0.1649	0.1737
	Pb37	0.0533	-
	Pb47	0.0517	0.0983
	Pb50	0.0401	0.0793
	Pb66	0.0529	0.0771
	Pb73	0.0464	-
	All	0.0678	0.1089
Nuclear markers		0.0772	0.1364
Mitochondrial DNA			
nucleotide		0.7036	0.8348
haplotype		0.6039	0.6334

Mitochondral DNA data were obtained from article VI. Only data from the same populations analysed in this study were used for comparison. Overall F_{st} were computed using nucleotide distances among haplotypes ("nucleotide") and only considering haplotypes, independently of the nucleotide variation ("haplotype").

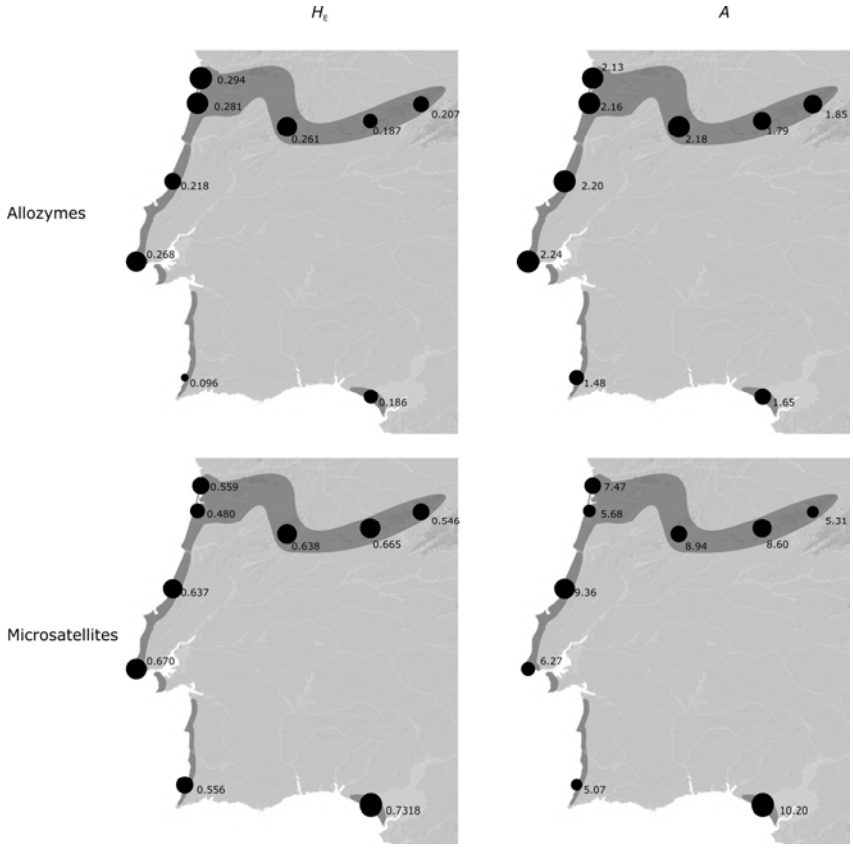

Figure 5. Mean diversity measures (expected heterozygosity (H_E) and allelic richness (A)) in populations of *P. carbonelli*.

Table 5. Pairwise F_{ST} values for populations of *P. bocagei*.

	Mad	Vair	Bra	Zim	Mon	San	Cor	Sar
Mad	-	0.042	0.018	0.083	0.080	0.112	0.154	0.122
Vair	**	-	-0.005	0.036	0.030	0.091	0.139	0.041
Bra	**	ns	-	0.047	0.053	0.065	0.102	0.023
Zim	**	**	**	-	0.065	0.109	0.162	0.073
Mon	**	**	**	**	-	0.138	0.187	0.051
San	**	**	**	**	**	-	0.053	0.093
Cor	**	**	**	**	**	**	-	0.129
Sar	**	**	**	**	**	**	**	-

ns - non significant; ** - significant ($p<0.05$)

Table 6. Pairwise F_{ST} values for populations of *P. carbonelli*

	LA	VR	SE	Esm	Av	SPM	CR	MC	PR
LA	-	0.093	0.123	0.187	0.163	0.105	0.236	0.217	0.134
VR	**	-	0.069	0.092	0.108	0.007	0.147	0.169	0.035
SE	**	**	-	0.033	0.077	0.039	0.144	0.202	0.088
Esm	**	**	**	-	0.084	0.078	0.153	0.246	0.131
Av	**	**	**	**	-	0.107	0.187	0.244	0.148
SPM	**	ns	**	**	**	-	0.139	0.132	0.034
CR	**	**	**	**	**	**	-	0.289	0.149
MC	**	**	**	**	**	**	**	-	0.097
PR	**	**	**	**	**	**	**	**	-

ns - non significant; ** - significant ($p<0.05$)

NJ trees depicting the relationships among the observed populations are shown in figure 6A (*P. bocagei*) and 6B (*P. carbonelli*). A common observation between both trees is the lack of long internal branches supporting strong differentiation between population groups. In *P. bocagei*, the Galician populations appear as a relatively well-supported clade, within which the grouping of A Coruña and Sanxenxo is also well-supported. On the opposite end of the tree, the populations from the southeastern area of the distribution (Zimão and Montesinho) also form a clade, although this is not as well supported. In *P. carbonelli*, a striking feature is the long branch leading to the population of Cabo Raso. A few groups are also recovered: a cluster comprising the two southernmost populations, Monte Clérigo and Playa del Rompeculos, stands out with high bootstrap support; the two populations from the Iberian Central System (La Alberca and Villasrúbias) also appear as a group. The clades formed by Aveiro and Espinho and by these two populations and Serra da Estrela are less well supported but still present bootstrap values above 50%.

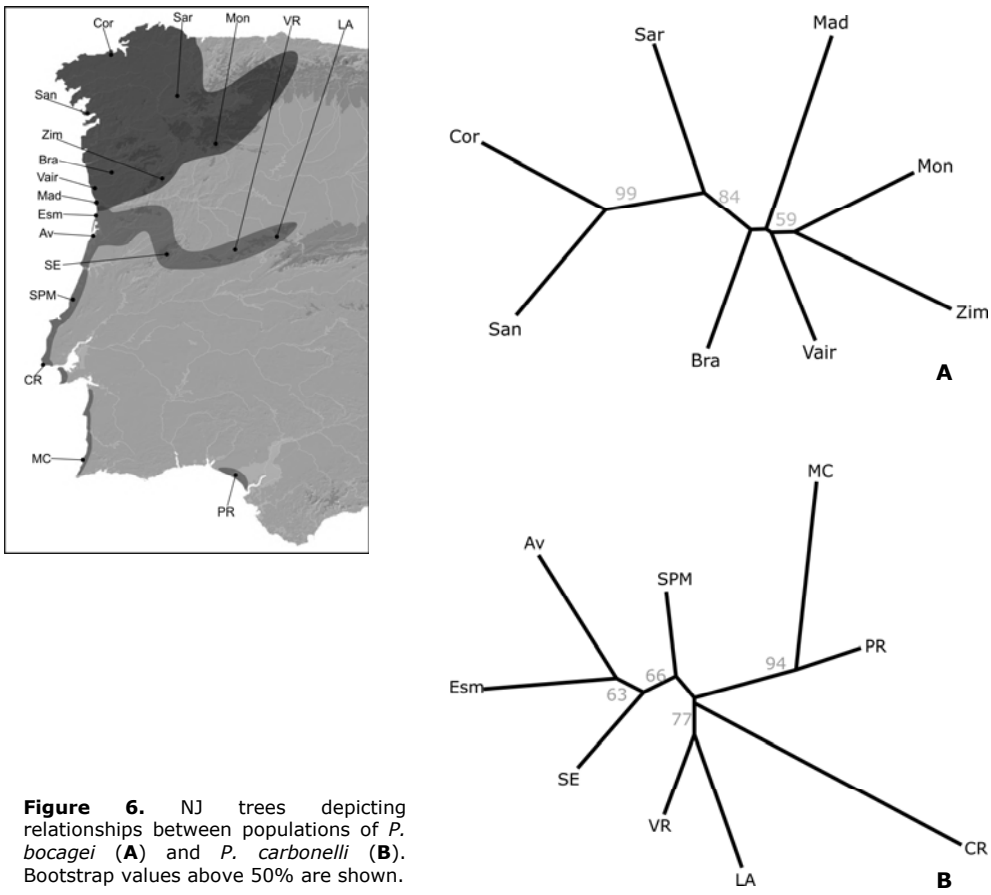

Figure 6. NJ trees depicting relationships between populations of *P. bocagei* (**A**) and *P. carbonelli* (**B**). Bootstrap values above 50% are shown.

We also conducted analyses of molecular variance (AMOVA) to test hypotheses regarding the partition of genetic variability in both species. We did not intend to perform an exhaustive search of all possible combinations of populations within species. Instead, we were hoping to find signatures of some degree of correlation of genetic variability with geographical groups and therefore focused on groups involving geographically close populations. Only the partitions producing maximum values of Φ_{CT} plus important hypotheses to be tested (the influence of the Douro river in *P. bocagei* and mtDNA-based partition in *P. carbonelli*) are shown in table 7. In both species only shallow genetic variation was found among groups. In *Podarcis bocagei* the highest Φ_{CT} values were obtained when Galician populations (especially A Coruña, but also Sanxenxo and Sarria albeit at a lower levels) were placed in separate groups from the remaining populations, suggesting that they are somewhat derived in the context of *P. bocagei* genetic variability, which is in accordance with the relatively longer branches leading to these populations in the NJ tree. Grouping the populations of Vairão and Braga and considering each of the remaining populations as belonging to a distinct group also produced a significant and relatively high Φ_{CT}. To test the influence of the river Douro, a well-known biogeographic barrier in the context of Northwestern Iberian herpetofauna, we also established the population of Madalena as a group versus all the other populations, which yielded a very low and non-significant Φ_{CT} of 0.023. In *P. carbonelli*, none of the tested groups exhibits higher levels of variation among them than individual populations among themselves when no group is considered, suggesting that no geographic group satisfactorily represents genetic variability within this species. The highest value of Φ_{CT} is found when all populations are considered separately with the exception of Esmoriz and Serra da Estrela. Within this species, four geographic groups were detected at the mtDNA level, although some haplotypes were shared between them (Pinho *et al.* in press a). We also tested the significance of such partition of genetic variability on the data sets. Only 3.64% of the total genetic variation could be ascribed to differences between these groups. Also based on mitochondrial DNA variation, we tested partitions involving the grouping of the populations of Villasrúbias and Monte Clérigo because, despite their distant geographic situation, these populations share one haplotype (haplotype C4 in Pinho *et al.* in press a). None of these hypotheses produced significant Φ_{CT} values (results not shown).

Synthetic maps produced using the results obtained for axes 1 and 2 from the Factorial Correspondence Analyses (FCA) for both species are shown in figures 7 and 8. In *P. bocagei*, axes 1 and 2 account for 21.74% and 19.16%, respectively, of the total variation in allele frequencies. Visual inspection of the maps reveals, for axis 1, a North-South oriented gradient in allele frequencies in this species. Axis 2 shows instead an East-West gradient. In *P. carbonelli*, the

percentages of explained variance are 25.22% and 20.38%, for axis 1 and 2, respectively. The first axis clearly illustrates the differentiation (also observable from the NJ tree) of the population of Cabo Raso; the second shows a NE-SW gradient in allele frequencies.

Table 7. Testing hypotheses concerning the geographic distribution of genetic variation in *P. bocagei* and *P. carbonelli*.

Defined population structure	K	Percentage of variation		
P. bocagei				
No structure		AP		WP
	1	7.53		92.47
Groups		AG	AP/WG	WP
Cor + Others	2	8.58	5.36	86.06
Mad + (Vair, Bra) + Zim + Mon + San + Cor + Sar	7	8.56*	-0.60	92.04
(San, Cor) + Others	2	8.37*	4.64	86.99
San + Cor + Others	3	8.06*	4.65	87.29
Mad + (San, Cor) + Others	3	6.97*	3.61	89.42
(San, Cor) + Sar + Others	3	6.48	4.12	89.40
San + Cor + Sar + Others	4	6.48*	4.05	89.47
Mad + San + Cor + Sar + Others	5	6.38*	2.94	90.68
Mad + (Vair, Bra, Zim) + Mon + San + Cor + Sar	6	6.30	1.98	91.72
Mad + (San, Cor) + Sar + Others	4	6.12*	3.20	90.68
Mad + Others	2	2.27	6.76	90.97
P. carbonelli				
No structure		AP		WP
	1	13.08		86.92
Groups		AG	AP/WG	WP
LA + VR + (SE, Esm) + Av + SPM + CR + MC + PR	8	10.20	3.28	86.52
LA + (VR, SE, Esm, Av, SPM) + CR + MC + PR	5	9.15*	6.49	84.36
LA + (VR, SE, Esm, Av, SPM) + CR + (MC, PR)	4	8.96*	6.77	84.27
MC + Others	2	8.58	10.42	81.01
CR + (MC, PR) + Others	3	8,05*	8.37	83.58
(LA, VR) + (SE, Esm) + Av + SPM + CR + MC + PR	7	7.95*	5.52	86.53
LA + (VR, SE) + Esm + Av + SPM + CR + MC + PR	8	7.84	5.45	86.71
CR + MC + PR + Others	4	7.77*	8.42	83.81
LA + VR + (SE, Esm) + Av + SPM + CR + (MC, PR)	7	7.69*	5.68	86.63
CR + outras	2	7.46	10.55	81.99
(LA + VR) + (SE, Esm, Av, SPM, CR) + MC + PR	4	3.64	10.50	85.86

K, number of groups considered; AP, among populations; WP, within populations; AG, among groups; AP/WG, among populations within groups; *, significant ($p<0.05$). See table 1 for population name abbreviations.

Individual multilocus genotype clustering

Results obtained using the software STRUCTURE are presented in figures 9 and 10 for *P. bocagei* and *P. carbonelli*, respectively, using a model with admixture and correlated allele frequencies. Results from runs employing an independent allele frequencies model produced highly similar results for both data sets, albeit with lower degrees of "clusteredness" (sensu Rosenberg *et al.* 2005; results not shown). We present the partitions that for each value of K (from 2 to 5, since higher levels of subdivision produces complex and uninterpretable scenarios)

showed the highest values of Ln Probability. Because STRUCTURE does not integrate over the possible values of K, we employed the method suggested by Evanno *et al.* (2005) to choose the number of partitions that better represented the variation present in the data sets; these results are also presented in figures 9 and 10 and the chosen values of K represented with respect to geography.

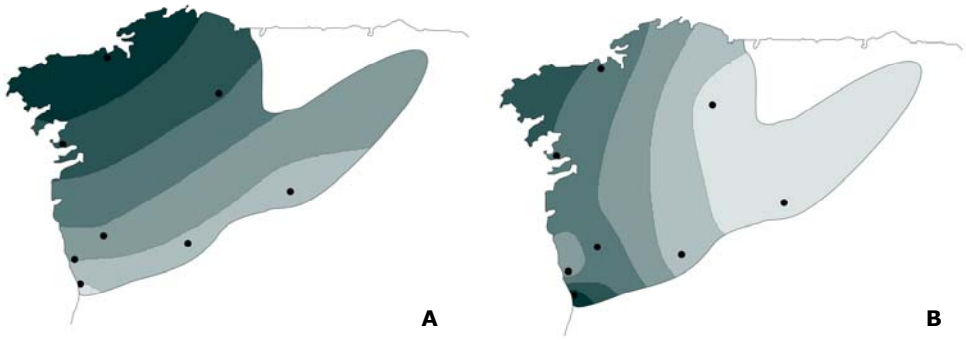

Figure 7. Synthetic maps showing the variation in allele frequencies across geography based on the interpolation of the values calculated for each population of *P. bocagei* for the first (**A**) and second (**B**) axis of a Factorial Correspondence Analysis.

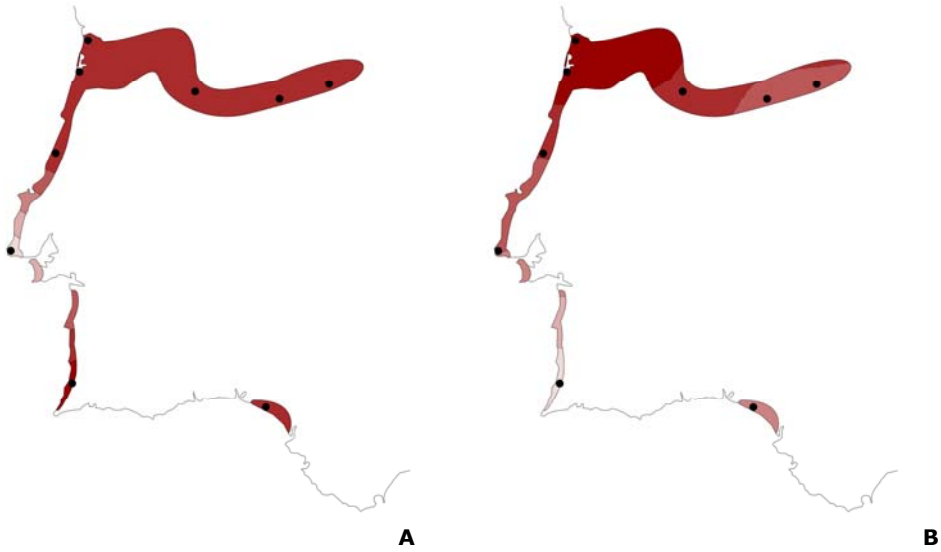

Figure 8. Synthetic maps showing the variation in allele frequencies across geography based on the interpolation of the values calculated for each population of *P. carbonelli* for the first (**A**) and second (**B**) axis of a Factorial Correspondence Analysis.

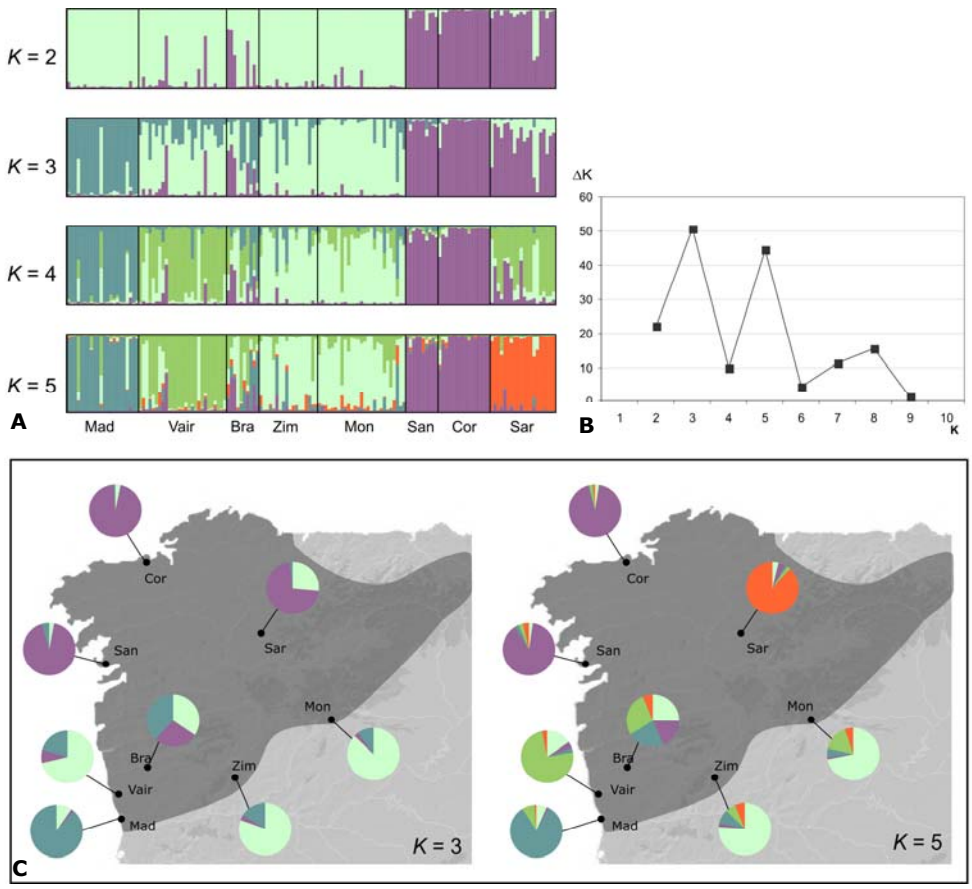

Figure 9. Individual multilocus genotype clustering analysis for *P. bocagei*. **A.** Inferred population structure from K=2 to K=5. These plots were obtained from the runs producing the highest values of Ln Probability for each value of K, assuming correlated allele frequencies. In these plots, each individual is represented by a column divided into K segments, the size of each corresponding to the individual's estimated membership fraction in each of the K clusters. See table 1 for population name abbreviations. **B.** Variation of the value of ΔK with the number of clusters, following Evanno *et al.* (2005). **C.** Pie charts representing the mean proportion of membership for K=3 and K=5 (chosen by the previous method) for each population.

In *P. bocagei*, the first differentiation to appear (with K=2) is that between Galician populations and southern ones; nevertheless, the discrimination is not perfect, with apparently admixed individuals in the populations of Vairão, Braga and Sarria. Increasing the number of subpopulations (K=3), A Coruña and Sanxenxo appear as almost "pure" for one of the inferred clusters. A second cluster appears at higher proportions in individuals from Zimão and Montesinho; individuals from Sarria are highly admixed with respect to these first two

clusters, with a predominance of the Galician component. A third cluster is more frequent in the population of Madalena, with nearly all individuals "pure" for this population. Individuals from Vairão and, in particular, Braga present a high level of admixture between all three clusters. With K=4 a clade formed mainly by individuals from Vairão and Sarria but also present in other populations emerges. With K=5, the clusters roughly correspond to Madalena, Vairão, the two populations from the southeast (Zimão and Montesinho), the two western Galician populations (Sanxenxo and Coruña) and Sarria. Braga remains as admixed, with high proportions of the first four clusters. Evanno *et al.*'s test (figure 8B) produced inconclusive results, with both K=3 and K=5 receiving high support.

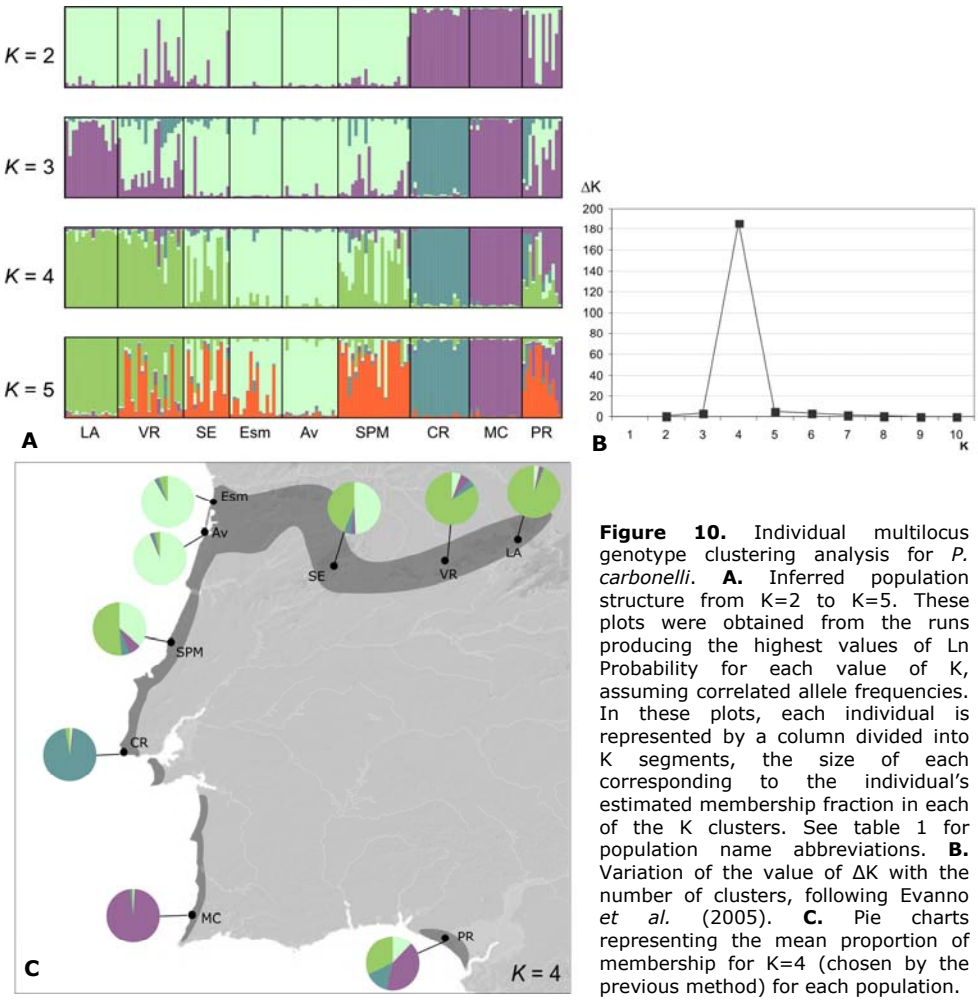

Figure 10. Individual multilocus genotype clustering analysis for *P. carbonelli*. **A.** Inferred population structure from K=2 to K=5. These plots were obtained from the runs producing the highest values of Ln Probability for each value of K, assuming correlated allele frequencies. In these plots, each individual is represented by a column divided into K segments, the size of each corresponding to the individual's estimated membership fraction in each of the K clusters. See table 1 for population name abbreviations. **B.** Variation of the value of ΔK with the number of clusters, following Evanno *et al.* (2005). **C.** Pie charts representing the mean proportion of membership for K=4 (chosen by the previous method) for each population.

Introducing a two-population structure into the *P. carbonelli* data set, Cabo Raso and Monte Clérigo appear as a group versus all other populations. The population of Playa del Rompeculos remains, however, admixed with respect to these two clusters, as well as some individuals from Villasrúbias. With K=3, Cabo Raso becomes clearly detached from all other populations; a second cluster is formed by individuals mainly from the populations of Monte Clérigo and La Alberca, but is also represented in Playa del Rompeculos, Villasrúbias and, to a lesser extent, S. Pedro de Moel. A third cluster is widely distributed among all populations (except Cabo Raso and Monte Clérigo), with high levels of homogeneity particularly in the populations of Serra da Estrela, Aveiro and Esmoriz. With K=4, individuals from Cabo Raso fall once again in a cluster of their own; individuals from Monte Clérigo likewise, although this cluster is also represented in some individuals from Playa del Rompeculos. Individuals from La Alberca and Villasrúbias, on one hand, and from Aveiro and Esmoriz, on the other hand, are placed with a low degree of admixture in a third and a fourth clusters, respectively. The populations of Serra da Estrela and S. Pedro de Moel appear admixed between these two clusters. Increasing the number of populations to five, La Alberca, Aveiro, Cabo Raso and Monte Clérigo each constitute a highly homogenous cluster; a fifth group appears at high frequency in the genome of individuals from S. Pedro de Moel and also with admixture from other clusters in the populations of Villasrúbias, Serra da Estrela, Esmoriz and Playa del Rompeculos. For this species, results from Evanno *et al*. 's test show a clear mode for K=4.

Discussion

Evolutionary history of Podarcis bocagei

In general, low levels of subdivision were found in nuclear markers across populations of *P. bocagei*. This is in accordance with evidence based on mitochondrial DNA differentiation. In fact, there is no sign of a strong geographic substructure that would be concordant with isolation and divergence in multiple glacial refugia, as has been described using nuclear markers in several other species from the Northwest of the Iberian Peninsula (e.g. Alexandrino *et al.* 2000, Godinho 2003, Godinho *et al.* 2003, 2006a, 2006b). Interestingly, levels of divergence based on nuclear multilocus data are much lower than those observed using mitochondrial DNA when the same populations and only allele (rather than nucleotide distance) data are considered ($\Phi_{ST(nuc)}$=0.07 whereas $\Phi_{ST(mit)}$=0.60). This discrepancy probably reflects the different effective population sizes (and probably different overall variability levels) that characterize each class of marker.

The phylogeographic study of Pinho *et al.* (in press a), based solely on mtDNA, described a pattern consistent with an extremely rapid population demographic growth from a single refugial source, but was inconclusive with respect to the location of the glacial refugium for *P. bocagei*. Nevertheless, from the explosive magnitude of the demographic expansion inferred from mtDNA variation, we expect that this refugial region would be relatively small in comparison to the species' present-day distribution area. We may therefore assume that the majority of the species' current distribution range has been colonized after the climate warming, that is, within the past 10,000 years. It is therefore surprising that based on a panel of 16 loci (including only a few highly variable markers) geographic-consistent genetic differentiation can be recovered within this short evolutionary time-frame. This genetic differentiation, albeit shallow, can be observed from multiple lines of evidence: the NJ tree, AMOVA results, synthetic maps and, in particular, the multilocus genotype individual clustering approach. All of these methods point, for example, to a moderate degree of differentiation of the Galician, northernmost populations, and to a close relationship between the two Southeastern populations (Zimão and Montesinho). In concordance with the findings based on mtDNA variation, we believe that this pattern of differentiation, which is particularly obvious in STRUCTURE analyses with K=3, is a by-product of a recent post-glacial expansion. The differentiated population groups, rather than representing true "clusters" in the sense that they result from independent evolutionary pathways, correspond to the extremes of clines deriving from the area that functioned as a glacial refugium. Corroborating this evidence, the populations of Braga and, to a lesser extent, Vairão, present evidence of "admixture". This particularity therefore identifies the southwestern region of the distribution as the source for geographic expansion and helps pinpointing the probable location of the area to which this species was confined during the last glaciations. The dynamics of post-glacial expansion are furthermore documented by a marked reduction in variability levels in a northwards fashion, consistent with a scenario of "northern purity" (Hewitt 2001).

A much denser sampling scheme would be necessary to identify with precision the routes by which the expansion was carried out. Nevertheless, an interesting result is obtained by comparing mitochondrial DNA haplotype distribution with the nuclear NJ tree. A single mitochondrial DNA haplotype is shared between the populations located on the northeastern area of the species' distribution, including Sarria (Pinho *et al.* in press a). This haplotype is also present with high frequencies in Montesinho but at very low frequencies elsewhere. This suggests that colonization of the northeastern areas of *P. bocagei*'s distribution was carried out through the east. However, in the NJ tree the population of Sarria clusters with Sanxenxo and A Coruña rather than with

Montesinho, as observed in the mtDNA analyses. These two incompatible observations are reconciled by taking into consideration the results of the multilocus genotype clustering algorithm. Individuals from Sarria, although ascribed in their majority to the Galician group, present some signatures of admixture with the southeastern clade, probably resulting from the merging of two waves of expansion. Both of these waves are represented in the synthetic maps built from the interpolation of the values assumed by each population along axes one and two of the FCA. This situation therefore mirrors, at a much shallower scale, the process of northern admixture of different expansion waves described by Comps *et al.* (2001). Interestingly, the river Douro, which has been described as a major biogeographic break in other species with similar distributions (*Chioglossa lusitanica,* Alexandrino *et al.* 2000; *Lacerta schreiberi*, Godinho 2003; probably *Discoglossus galganoi*, Martínez-Solano 2004), does not appear to have significantly contributed to genetic differentiation within *P. bocagei*. Nevertheless, the population of Madalena stands out as mildly differentiated by emerging as a clade of its own in STRUCTURE analyses when three or more populations are enforced. It has been suggested (Sá-Sousa 2001) that the colonization of the area south of this river occurred in the past two-hundred years via human-built bridges. Because the population of Madalena is characterised by relatively high variability levels, moderate differentiation relative to other populations and does not appear in a particularly derived position in the NJ tree, our data are not consistent with the hypothesis of an extremely recent colonization and therefore do not support this suggestion.

Evolutionary history of Podarcis carbonelli

Although a scenario with some degree of isolation was inferred for *P. carbonelli* based on mitochondrial DNA variation (Pinho *et al.* in press a), the observed level of differentiation is not consistent with early isolation within the first stages of the Pleistocene, as inferred for many other species of the Iberian herpetofauna (Alexandrino *et al.* 2000, 2002; Paulo *et al.* 2001, 2002). In fact, instead of deeply divergent lineages, we find only a few mutations between each haplotype group; moreover, the haplotype inferred to be ancestral to all other haplotypes still exists and is notably present in two distant populations. Applying a molecular clock, coalescence time for *P. carbonelli* mtDNA was estimated around 500,000 – 300,000 years (Pinho *et al.* in press a).

According to allozymes and microsatellites, differentiation within this species is quite low when compared to mtDNA variation ($\Phi_{ST(nuc)}$=0.13 whereas $\Phi_{ST(mit)}$=0.63), which probably derives from the same factors already noted for *P. bocagei*. Still, this value is higher than differentiation within *P. bocagei*, which is in accordance with the older mtDNA coalescence time inferred for this species, implying a longer persistence of its populations.

However, we find that genetic variation in nuclear loci is not highly correlated to that in mtDNA. In fact, genetic differentiation between the four groups described at the mtDNA level merely accounts for 3.64% of the total genetic variation observed in this species. Moreover, none of the possible geographic clusters that were tested yield higher levels of variability among groups than the ones observed among populations when these are considered individually. The topology of the NJ tree is mostly star-like, with short internal branches inconsistent with long term isolation in distinct refugia. There is, nevertheless, a tendency for the clustering of geographically close populations. In accordance with mtDNA data, the two populations from the Spanish Central System cluster together; some populations from the northwestern clade also appear as closely-related (Esmoriz, Aveiro, Serra da Estrela and S. Pedro de Moel), albeit with lower bootstrap support. However, the population of Cabo Raso, which is assigned to the northwestern clade with the mtDNA, is unrelated to other populations in this clade in the NJ tree obtained from nuclear markers. Interestingly, the two southernmost populations, Monte Clérigo and Playa del Rompeculos, cluster together with high bootstrap support, which was not suggested by mtDNA analyses.

These relationships among populations, however, are not clearly depicted on the individual clustering analyses; despite high similarities in the assignment of individuals belonging to populations that cluster together in the NJ tree, these correspondences are not perfect, with some populations appearing as highly admixed between distinct clusters. Although these results could be the outcome of complex patterns of fragmentation and admixture, an alternative explanation is that these analyses have instead identified populations or groups of closely related populations (such as Cabo Raso, Monte Clérigo, Aveiro + Esmoriz and La Alberca + Villasrúbias) that have experienced higher levels of genetic drift. This interpretation thus suggests that populations that appear as admixed are probably those that were able to maintain higher effective population sizes throughout climatic changes. One such populations is the geographical isolate from southern Spain, which is unexpected because isolated populations usually display low effective population sizes. However, given the high variability levels detected in microsatellite loci in this population, it is likely that it sustained high effective population sizes despite the progressive isolation. In contrast, the population of Monte Clérigo shows depleted levels of genetic variability and was most likely strongly affected by genetic drift. This could be the result of its peripheral location regarding the species distribution, which has probably caused repeated local extinctions and recolonizations.

The lack of concordance between patterns of genetic differentiation inferred from mtDNA and nuclear genes may be ascribed to several possible causes: i) the two types of markers are portraying different time-frames of differentiation;

ii) rather than allopatric isolation, differentiation within this species results from a gradual change along geography (i.e. isolation by distance) coupled with some degree of genetic drift, affecting mitochondrial and nuclear loci differently because of their intrinsic characteristics or simply because of stochasticity, creating apparent phylogeographic breaks in mtDNA without this being caused by real isolation (Neigel and Avise 1993, Irwin 2002); iii) the evolutionary history of *P. carbonelli* corresponds to a complex sequence of events of fragmentation and admixture that left different signatures on different markers. Even though these hypotheses are not mutually exclusive, the present analyses do not allow a straightforward evaluation of the roles that allopatric fragmentation, genetic drift and gene flow have played in this species' evolution. Additionally, unlike in *P. bocagei*, it is difficult to pinpoint areas that might have functioned as glacial refugia. The currently fragmented distribution of *P. carbonelli,* the uniqueness of its distribution in the context of Iberian herpetofauna and the difficulties to model such distribution (Sá-Sousa 2001a) suggest that this species has probably had a much larger distribution area in the past; it is possible that the genetic signatures of isolation, demographic recovery and admixture usually associated to the alternation of glacial and interglacial stages have been erased as a consequence of recent changes in the species distribution.

The mitochondrial DNA-based phylogeographic study carried out by Pinho *et al.* (in press a) detected a strong signature of population demographic expansion within the northwestern clade of *P. carbonelli*. This was particularly evident not only from statistical tests designed to detect demographic changes but also from a progressive loss of genetic variability in a northwards fashion. Given the unequivocal data for this expansion, our preliminary expectation was that nuclear markers would reveal a similar pattern. Instead, we find no consistent decrease in levels of genetic variability. In particular, the population of Serra da Estrela, which because of its geographic position was expected to bear the lowest levels of genetic variation within its clade, shows high variability levels in both allozyme and microsatellite loci. Which processes may be underlying such discrepant results? A possible explanation resides on differences in current effective population sizes between putative source populations and recently colonized territories. In fact, *P. carbonelli* seems to be nowadays more abundant in the northern area of its distribution than in the south, where it becomes rare and more locally distributed (Sá-Sousa 2001b, pers. obs.). It is possible that, at the same time that *P. carbonelli* expanded to the North, southern habitats became progressively less suitable and thus sustained lower effective population sizes and consequently lower diversity, whereas the demographic expansion in the north compensated for the loss of diversity via genetic drift. This hypothesis, however, does not explain why high mtDNA variability levels were maintained in,

for example, S. Pedro de Moel. A second scenario could be that of a slower expansion. Petit *et al.* (1999), for example, found no evidence of loss of variability in the post-glacial colonization of noctule bats, probably because the expansion was carried out in a slow fashion tracking forest expansion. In the case of *P. carbonelli*, the expansion would have to have been carried out in a slow enough fashion for drift to affect mtDNA but not nuclear markers because of different effective population sizes. A third scenario would imply extensive genetic admixture (from, for example, the Central System phylogroup) in northern populations but not in southern ones; however, we find no corroborating evidence for this using mtDNA. These hypotheses are purely demographic. However, scenarios that invoke natural selection cannot be ruled out. For example, a recent selective sweep in the mitochondrial genome of *P. carbonelli* could be responsible for the positive signature of demographic expansion, since both phenomena leave similar imprints on genetic variation. Similarly, balancing selection acting on nuclear loci could be preventing the loss of genetic variation expected under a scenario of rapid geographic expansion. Nevertheless, although balancing selection has been shown to affect allozyme variation in some species (e.g. Karl and Avise 1992), microsatellite loci are usually regarded as selectively neutral and less prone to be affected by such phenomena.

Concluding remarks

This study highlights the need for the assessment of multiple loci prior to making inferences on the evolutionary history of organisms. Although general evolutionary trends were concordant between mtDNA and nuclear genes (e.g. low level of subdivision inferred for both species albeit higher levels of differentiation in *P. carbonelli;* a history of recent demographic and geographic expansion in *P. bocagei*), the study of mtDNA alone fails to detect significant particularities in the evolutionary history of both species. For example, only the analyses of nuclear markers detected signatures of multiple expansion routes in *P. bocagei*. In *P. carbonelli*, although the pattern of genetic variation in nuclear loci is not straightforwardly interpreted in the light of Pleistocene climatic alterations, mtDNA portrays an oversimplified history of isolation which does not reflect the necessarily complex dynamics of this species evolution. In this particular case, the inclusion of more samples from the southern range of the species distribution could eventually provide important cues for a better understanding of this species evolutionary history and, eventually, help in the adoption of appropriate conservation measures.

Acknowledgements

This work was partially financed by FCT research projects POCTI/40987/BSE/2001 and POCI/BIA-BDE/60911/2004. CP was supported by a PhD grant (SFRH/BD/4620/2001) and DJH was funded by a post-doctoral grant (SFRH/BPD/5702/2001), all from FCT. The authors would like to thank all the people that participated in sampling campaigns (Miguel A. Carretero, Antigoni Kaliontzopoulou, Bruno Ribeiro, Daniel Cunha, Diana Barbosa, Elsa Froufe, Gonçalo Themudo, Mário J. Pinho, Raquel Godinho, Raquel Ribeiro, Ricardo Pereira, Sara Rocha and Vasco Batista). We would also like to thank Neftali Sillero's invaluable help on producing synthetic maps and Fernando Sequeira's helpful advice on the analyses of data.

References

Alexandrino J, Arntzen JW, Ferrand N (2002) Nested clade analysis and the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Heredity*, **88**, 66–74.

Alexandrino J, Froufe E, Arntzen JW, Ferrand N (2000) Genetic subdivision, glacial refugia and postglacial recolonization in the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Molecular Ecology*, **9**, 771–781.

Avise JC, Walker D, Johns GC (1998). Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society of London B: Biological Sciences*, **265**, 1707–1712.

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F. (1996-2004) GENETIX, logiciel sous WindowsTM pour la génétique des populations. Laboratoire Génome et Populations, CNRS, Université de Montpellier II, Montpellier, France

Branco M, Ferrand N, Monnerot M (2000) Phylogeography of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula inferred from RFLP analysis of the cytochrome *b* gene. *Heredity*, **85**, 307–317.

Branco M, Monnerot M, Ferrand N, Templeton AR (2002) Postglacial dispersal of the European rabbit (*Oryctolagus cuniculus*) on the Iberian Peninsula reconstructed from nested clade and mismatch analyses of mitochondrial DNA genetic variation. *Evolution*, **56**, 792–803.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution* **32:** 550-570

Comes HP, Abbott RJ (1998) The relative importance of historical events and gene flow on the population structure of a Mediterranean ragwort, *Senecio gallicus* (Asteraceae). *Evolution*, **52**, 355–367.

Comps B, Gömöry T, Letouzey J, Thiébaut B, Petit RJ (2001) Diverging trends between heterozygosity and allelic richness in the post-glacial colonization of the European beech. *Genetics* **157,** 389–397.

Cooper SJB, Ibrahim KM, Hewitt GM (1995) Postglacial expansion and genome subdivision in the European grasshopper C*horthippus parallelus*. *Molecular Ecology*, **4**, 49–60.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* **92**, 832-839

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.

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) Version 3.6. Distributed by the author. University of Washington, Seattle

Garcia-Paris 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.

Godinho R (2003) Evolutionary history of the Schreiber's green lizard, *Lacerta schreiberi*: Phylogenetic analysis, reconstruction of processes of population expansion, contraction and miscigenation, and integration in the biogeographic context of Northwestern Iberian Peninsula. PhD thesis: University of Lisbon.

Godinho R, Domingues V, Crespo EG, Ferrand N (2006a). Extensive intraspecific polymorphism detected by SSCP at the nuclear C-mos gene in the endemic Iberian lizard *Lacerta schreiberi*. *Molecular Ecology* **15**, 731–738.

Godinho R, Mendonça B, Crespo EG, Ferrand N (2006b) Genealogy of the nuclear b-fibrinogen locus in a highly structured lizard species: comparison with mtDNA and evidence for intragenic recombination in the hybrid zone. *Heredity* **16**, 454–463.

Godinho R, Paulo OS, Ferrand N, Luis C, Rosa HD, Crespo EG (2003). Major patterns of population differentiation in the Iberian Schreiber's Green Lizard (*Lacerta schreiberi*) inferred from protein polymorphism. *Herpetological Journal* **13**, 35–42.

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography in southern European refugia: evolutionary perspectives on the origin and conservation of European biodiversity* (Weiss S, Ferrand N, eds). Pp. 155-188. Springer.

Hedrick PW (1999). Highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**, 313–318.

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 (2000) The genetic legacy of the quaternary Ice Ages. *Nature*, **405**, 907–913.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology,* **10**, 537–549.

Irwin DE (2002) Phylogeographic breacks without geographic barriers to gene flow. *Evolution* **56:** 2383–2394.

Karl SA, Avise JC (1992) Balancing selection at allozyme loci in oysters: Implications from nuclear RFLPs. *Science* **256**, 100-102

Martínez-Solano I (2004) Phylogeography of Iberian *Discoglossus* (Anura: Discoglossidae). *Journal of Zoological Systematics and Evolutionary Research*, **42**, 298–305.

Neigel JE, Avise JC (1993) Application of a random-walk model to geographical distribution of animal mitochondrial DNA variation. *Genetics* **135:** 1209–1220.

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population *Genetic Research* **22,** 201–204.

Paulo OS, Dias C, Bruford MW, Jordan WC, Nichols RA (2001) The persistence of Pliocene populations through the Pleistocene climatic cycles: evidence from the phylogeography of an Iberian lizard. *Proceedings of the Royal Society of London*. *Series B, Biological Sciences*, **268**, 1625–1630.

Petit E, Excoffier L, Mayer F (1999) No evidence of bottleneck in the post-glacial colonization of Europe by the noctule bat (*Nyctalus noctula*). *Evolution*, **53,** 1247–1258.

Petit RJ, Aguinagalde I, Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**, 1563–1565.

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology* **12**, 844-855

Pinho C, Ferrand N, Harris DJ (2004a) Genetic variation within the *Podarcis hispanica* species complex: new evidence from protein electrophoretic data. In: *The Biology of Lacertid Lizards: Evolutionary and Ecological Perspectives* (Pérez-Mellado V, Riera N, Perera A, eds.). Institut Menorquí d'Estudis. *Recerca*, **8,** 269–277.

Pinho C, Harris DJ, Ferrand N (2003) Genetic polymorphism of 11 allozyme loci in populations of wall lizards (*Podarcis* spp.) from the Iberian Peninsula and North Africa. *Biochemical Genetics* **41**, 343–359.

Pinho C, Harris DJ, Ferrand N (in press a) Contrasting patterns of population structure and historical demography in three western Mediterranean lizard species inferred from mtDNA variation. *Molecular Ecology* (in press)

Pinho C, Harris DJ, Ferrand N (in press b) 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,* in press.

Pinho C, Sequeira F, Godinho R, Harris DJ, Ferrand N (2004b) Isolation and characterization of 9 microsatellite loci in *Podarcis bocagei* (Squamata: Lacertidae). *Molecular Ecology Notes*, **4**, 286-288.

Pritchard JK, Stephens M, Donnelly P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Raymond M, Rousset F (1995a) GENEPOP (version 3.1): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248–249.

Raymond M, Rousset F (1995b). An exact test for population differentiation. *Evolution* **49**: 1280-1283

Rosenberg NA. (2002) *Distruct*: a program for the graphical display of *structure* results. http://www.cmb.usc.edu/~noahr/distruct.html

Rosenberg NA, Mahajan S, Ramachandran S, Chengfeng Z, Pritchard JK, Feldman MW (2005). Clines, clusters and the effect of study design on the inference of human population structure. *PLoS Genetics* **1**, e70

Sá-Sousa P (2001a) Comparative chorology between *Podarcis bocagei* and *P. carbonelli* (Sauria: Lacertidae) in Portugal. *Revista Española de Herpetología* **15**, 85–97.

Sá-Sousa P (2001b) A Controversa Sistemática das Lagartixas do género Podarcis WAGLER, 1830 (Sauria: Lacertidae) em Portugal. Dissertação de Doutoramento. Fac. Ciências, Univ. Lisboa.

Saitou N, Nei M (1987) The Neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4,** 406–427.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor Press, New York.

Santucci F, Emerson B, Hewitt G (1998) Mitochondrial DNA phylogeography of European hedgehogs. *Molecular Ecology*, **7**, 1163–1172.

Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2.000: A software for population genetics analyses. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Sequeira F, Alexandrino J, Rocha S, Arntzen JW, Ferrand N (2005). Genetic exchange across a hybrid zone within the Iberian endemic golden-striped salamander, *Chioglossa lusitanica*. *Molecular Ecology* **14:** 245–254.

Serre D, Päabo S (2004) Evidence for gradients of human genetic diversity within and among continents. Genome Research **14**, 1679–1685.

Surget-Groba Y, Heulin B, Guillaume C-P, Thorpe RS, Kupryianova L, Vogrin N *et al.* (2001) Intraspecific phylogeography of *Lacerta vivipara* and the evolution of viviparity. *Molecular Phylogenetics and Evolution* **18**, 449–459.

Taberlet P, Bouvet J (1994) Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear (*Ursus arctos*) in Europe. *Proceedings of the Royal Society of London B*, **255**, 195–200.

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

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

Chapter 4

General discussion

General discussion

4.1. Differentiation and gene flow in Iberian and North African wall lizards

4.1.1. Mitochondrial phylogeny and biogeography of Iberian and North African *Podarcis*

4.1.1.1. A new mitochondrial DNA scenario

The study of mitochondrial DNA variation has proven to be an invaluable tool in defining evolutionary units among species complexes in which traditional approaches to systematics have failed to do so (e.g. Bickford *et al.* 2007). Cryptic species, that is, groups of species that show limited differentiation with respect to morphological characters, often remain undetected unless studies on their genetic variation are carried out. In the particular case of Iberian and North African *Podarcis*, evolutionary units have remained cryptic not simply because of a high degree of homogeneity in morphology across the group's range. Decades of studies on morphological variation in these lizards had demonstrated an extreme degree of plasticity in morphology and colouration and it was therefore suspected that Iberian and North African *Podarcis* constituted a species complex long before the first analyses of genetic variability were carried out (Arnold and Burton, 1978). Nevertheless, the lack of concordance among different classification proposals prevented the achievement of a stable taxonomy for these lizards.

The studies of mitochondrial DNA variation carried out by Harris and Sá-Sousa (2001, 2002) and by Harris and co-workers (2002b) were the first large-scale attempts to solve the systematic puzzle which Iberian and North African *Podarcis* had constituted for decades. These studies clearly highlighted an unexpected degree of genetic variability within both of the formerly recognised species (*P. bocagei* and *P. hispanica*), suggesting that these two species could indeed represent cryptic species complexes. Nevertheless, although these first studies of mitochondrial DNA variation were crucial in the identification of major lineages and in helping pinpoint their distribution, they provided only mild support for assemblages comprising more than one form. In fact, the only group that receives consistent support in the studies by Harris and Sá-Sousa (2002) and Harris *et al.* (2002b) is the cluster including *P. hispanica* sensu stricto, *P. vaucheri* and other North-African forms. Relationships within this group,

although poorly supported, were suggestive of a complex history of colonization across the Strait of Gibraltar. According to these authors, two North African species clades had their closest relatives in southern Iberia. Applying a molecular clock, the authors inferred divergence times between sister groups inhabiting both sides of the Strait to be smaller than the time since the reopening of the Strait (ca. 5.3 MYA), thus suggesting at least two colonizations from the Iberian Peninsula into North Africa after the Strait was formed. Because other groups of species were not supported, biogeographic scenarios for the origin of the multiple forms could not be further explored. The doubt then remained: were poorly resolved relationships a consequence of insufficient sampling of mitochondrial DNA sequence or did they reflect a scenario of a rapid radiation in the origin of the observable variation? For example, Fu (2000) used large portions of the mitochondrial DNA molecule to establish a phylogeny of the Lacertidae; despite the data included many informative characters, only a few nodes were resolved with confidence, which was interpreted as resulting from almost simultaneous multiple speciation events.

To investigate whether this was the case in Iberian and North African *Podarcis,* and because obtaining a well-supported phylogeny is crucial prior to the establishment of any biogeographic scenario, we selected a few individuals from each of the previously defined clades plus a newly identified lineage and extended the amount of DNA sequence data incorporated. We combined partial data sets from five different mitochondrial DNA portions: cytochrome *b*, 12S ribosomal RNA (used in previous studies), 16S ribosomal RNA, NADH dehydrogenase subunit 4 (ND4) and adjacent tRNAs and the control region, yielding a total of 2425bp.

Using this combined mitochondrial DNA data set, our estimates of phylogeny were concordant independent of the method used to resolve it; moreover, most nodes above species level obtained bootstrap support close to 100% (figure 4.1). These estimates of relationships within Iberian and North African *Podarcis* confirm, on one hand, the monophyly of forms inferred from previous descriptions of mitochondrial DNA variability; on the other hand, they provide a robust assessment of evolutionary relationships among forms. In this context, three major groups can be observed. The first group incorporates forms mainly found in western Iberian Peninsula, including a group formed by *P. bocagei* and *P. hispanica* type 1 (which in turn incorporates two highly divergent and possibly allopatric lineages), and another clade grouping *P. carbonelli* and *P. hispanica* type 2. A second clade encompasses all North African forms and *P. hispanica* from South Eastern Iberian Peninsula, referred to in this work as *P. hispanica* sensu stricto. Within this clade, a sister-taxon relationship is confirmed for the forms from Jebel Sirwah and Tunisia. Interestingly, unlike the estimates of relationships reported by Harris *et al.* (2002b), *P. vaucheri* groups with these

forms with high bootstrap support. *P. hispanica* sensu stricto therefore appears as sister to all North African forms. A third major clade which appears as sister to all other Iberian and North African *Podarcis* incorporates two forms: *P. hispanica* from the North East of the Iberian Peninsula and a previously unidentified lineage found only in South Eastern Spain.

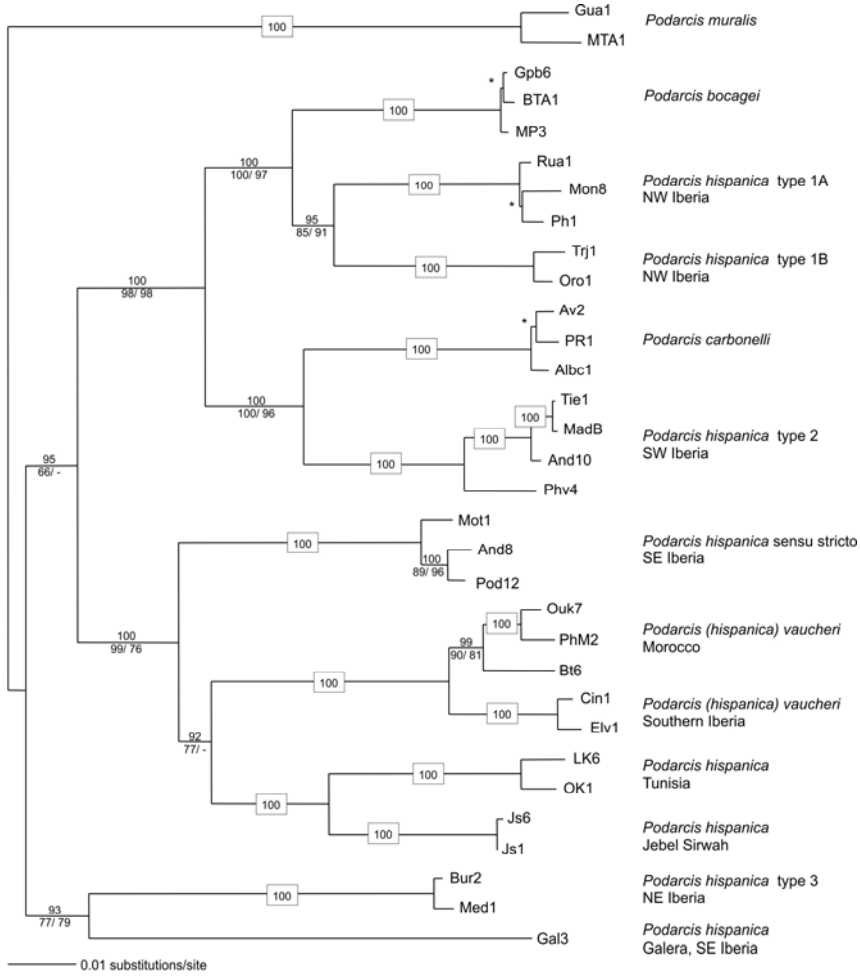

Figure 4.1. Estimate of relationships among samples of Iberian and North African *Podarcis* based on 2425bp of mitochondrial DNA, derived from a ML search using the GTR model with an estimate of invariable sites (0.67) and a discrete approximation of the gamma distribution (4 rate categories, α=1.67). Bayesian analysis produced the same estimate of relationships. The tree was rooted using *P. muralis*. Bayesian posterior probabilities are given above nodes, ML and MP bootstraps, respectively, below nodes. When all three were identical one value is given in a box. (-) Indicates bootstrap values lower than 50%. (¤) Indicates branches where different methods yielded different topologies.

A first conclusion to be retrieved from these results is the demonstration that mitochondrial DNA relationships within Iberian and North African *Podarcis* are recoverable and that previous unresolved phylogenetic estimates probably resulted from a lack of informative characters included in the studied data sets.

On the other hand, these results clearly suggest important biogeographic patterns:

a) Phylogeography of genus Podarcis *across the Strait of Gibraltar*

Because of their poor dispersal abilities, reptiles and amphibians have been widely used as models to test the relative effects of vicariance and dispersal in shaping genetic variability across geographic barriers (e.g. Oosterbroek and Arntzen 1992). The Strait of Gibraltar is a case in point. It has long been acknowledged that Iberian and Maghrebian faunas share affinities as the result of a close geographical proximity and several land connections in the past. The definitive opening of the Strait of Gibraltar at its present location occurred around 5.33 MYA, immediately after the Messinian salinity crisis, during which a dramatic decrease in sea levels allowed the contact between the two continents. Because there is no geological evidence for posterior connections, an institutionalized paradigm was that the opening of the Strait was the major cause for differentiation within organisms distributed on both sides. The belief in this generalization still induces the use of the opening of the Strait of Gibraltar as a molecular clock calibration point in relatively recent studies (e.g. Gantebein 2004, Zangari *et al.* 2006). However, a growing body of evidence suggests that patterns of genetic variation across the Strait are largely organism-specific instead of concordant, as would be expected if the opening of the Strait had had a major impact. Several species of reptiles and amphibians exhibit complex partitions of genetic variability both within and across the two regions. While a few species were indeed inferred to have started to diverge as the Strait opened (e.g. *Acanthodactylus erythrurus* (Harris *et al.* 2004a), and probably *Blanus cinereus* (Vasconcelos *et al.* 2006)), most of the studied organisms show patterns that do not conform to a simple scenario of vicariance related to the opening of the Strait. For example, several pre-Messinian divergence scenarios have been suggested (García-Paris and Jockush 1999, Veith *et al.* 2004, Martínez-Solano *et al.* 2004; Fromhage *et al.* 2004, Paulo 2001). In contrast, divergence times much shallower than the opening of the Strait have also been reported, suggesting that transmarine colonisations or human-mediated introductions might not be as infrequent as previously suspected (e.g. Álvarez *et al.* 2002, Paulo *et al.* 2002a, Carranza and Arnold 2004, Carranza *et al.* 2004b, Carranza *et al.* 2006a,b).

A similar but more complex scenario has been reported for *Podarcis* by Harris *et al.* (2002b), involving not one but two transmarine colonisations after the

Strait was formed. Our re-examination of the phylogenetic relationships within Iberian and North African *Podarcis* is concordant with a complex pattern of genetic variation across the Strait of Gibraltar. However, the establishment of a robust evolutionary scenario, different from that suggested by Harris *et al.* (2002b), allowed us to put forward alternative hypotheses to explain genetic variation (figure 4.2):

Hypothesis I: Two colonisations from Iberia into North Africa. The differentiation between *P. hispanica* sensu stricto and the ancestor of other forms in its clade occurred in southern Iberian Peninsula (or former islands near the Betic region). The ancestor of the forms from Jebel Sirwah and Tunisia colonised North Africa, promoting their separation from *P. vaucheri*. *P. vaucheri* invaded North Africa more recently.

Hypothesis II: First colonisation from Iberia into North Africa; second in the opposite direction. The ancestor of this clade colonised North Africa, promoting the separation between *P. hispanica* sensu stricto and the remaining forms. *P. vaucheri* and the forms from Jebel Sirwah and Tunisia speciated within North Africa. *P. vaucheri* secondarily invaded the Iberian Peninsula.

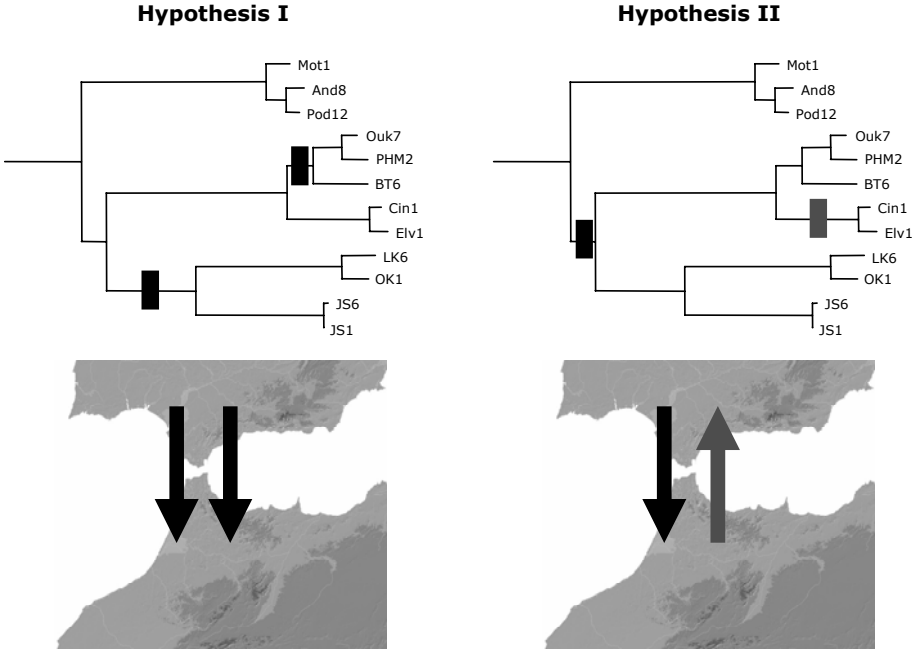

Figure 4.2. Two hypotheses for the colonization of North Africa compatible with the mitochondrial DNA-based estimates of relationships among *Podarcis*. See text for details.

Obviously, the word "colonisation" could be replaced by "vicariant event" in any of these formulations.

Without the estimation of divergence dates between clades, we would have no proper way of choosing amongst these two hypotheses taking into account the present knowledge. However, the use of an evolutionary rate of 2% sequence divergence per million years for the cytochrome *b* gene (which has been suggested for several reptiles) on uncorrected pairwise distances suggests a divergence time of around 5.5 MYA between *P. hispanica* sensu stricto and the remaining taxa within this clade. Because this differentiation is consistent with vicariance related to the opening of the Strait of Gibraltar, we favour the second hypothesis relative to the first. That is, our results suggest that, unlike previously thought, the opening of the Strait of Gibraltar may well be related to the partition of genetic variability in *Podarcis.*

This scenario should however be taken with caution because it is highly dependant upon the molecular clock calibration, which for several reasons might not be the most appropriate (see for a recent review Pulquério and Nichols 2006). An example of the importance of the molecular clock calibration is given by the discordant biogeographic scenarios inferred by Veith *et al.* (2004) and Carranza and Arnold (2004) to explain genetic variation on ribbed newts (*Pleurodeles*) although both studies inferred the same phylogenetic relationships.

b) Evolutionary relationships among western Iberian lizards

Prior to the evaluation of their evolutionary relationships using molecular tools, *Podarcis carbonelli* and *P. bocagei* were considered as two subspecies within *P. bocagei*. The reasons for this clustering were based on several similarities, both morphological (e.g. both species have thicker heads than *P. hispanica*) and ecological (both are ground-dwelling species, unlike *P. hispanica* forms, which are generally rock-dwelling). It was therefore expected that these two lizards would come out as sister taxa by genetic analyses. However, the first studies on genetic variation within *Podarcis* suggested that this was not the case (Harris and Sá-Sousa 2001, 2002), which was confirmed by our robust assessment of evolutionary relationships: *P. bocagei* is a sister group to *P. hispanica* type 1 whereas *P. carbonelli* is sister to *P. hispanica* type 2. Why these pairs of sister taxa are precisely among the few that are able to coexist in the same environment is therefore an important question that needs further evaluation. Competitive exclusion between distinct *Podarcis* may be an important factor conditioning the species' distributions (Pérez-Mellado and Galindo, 1986); probably, *P. carbonelli* and *P. bocagei*, by being able to explore different ecological niches, developed the potential for coexistence without competitive exclusion (with *P. hispanica* type 1A, in the case of *P. bocagei* and with *P.*

hispanica type 1A, type 1B and in particular type 2 in the case of *P. carbonelli*). However, our phylogenetic assessment strongly suggests that this adaptation arose in the *P. carbonelli* and *P. bocagei* lineages independently. An alternative scenario that should not be disregarded is that of a falsely inferred sister-relationship between either *P. bocagei* or *P. carbonelli* with *P. hispanica,* resulting instead from an ancient introgression of the mitochondrial DNA long ago in the species past. This pattern was described, for example, in Hawaiian crickets (Shaw 2002); in these species, mitochondrial DNA relationships suggest a within-island (i.e. sympatric) sister relationship, whereas both morphology and nuclear DNA suggest multiple colonisations and among-island sister-group relationships. However, this hypothesis remains largely speculative in the light of present knowledge.

Pairwise distances between these forms (around 10 – 12% cytochrome *b* divergence) suggest almost simultaneous divergence between species pairs, the time-frame of which may be in accordance with phylogeographic breaks inferred for the co-distributed newt *Lissotriton boscai*. Within this species two cryptic phylogroups may have started to diverge around 5.8 Mya (Martínez-Solano *et al.* 2006). This differentiation could reflect a common pattern of isolation due to changes in habitat availability at the time; a possibility is that the Messinian salinity crisis increased aridity all around the Mediterranean, creating islands of humid habitat where both *Podarcis* and *Lissotriton* species became isolated.

c) Separation between the southeastern Iberian-Mahgrebian clade and other groups

Applying the same molecular clock calibration as before, we obtained an estimate of around 7.0 Mya for the split between the clade comprising *P. bocagei, P. hispanica* type 1, *P. carbonelli* and *P. hispanica* type 2 and the clade including *P. hispanica* sensu stricto, *P. vaucheri*, *P. hispanica* "Jebel Sirwah" and *P. hispanica* "Tunisia". The rough temporal and geographic concordances could suggest that depositional events occurred around this time (7.8 – 7.6 Mya) in southern Iberia, described as the "Tortonian salinity crisis" (Krijgsman *et al.* 2000), could be responsible for the isolation between these two clades, as has been suggested in lizards of genus *Timon* (Paulo 2001). However, we should again stress that these inferences are highly dependent on a molecular clock calibration; given the highly eventful history of the region surrounding the Strait of Gibraltar (see section 1.1), other scenarios could also seem plausible.

d) A new lineage in Southeastern Iberia could be the nominotypical taxon of Podarcis hispanica

This study provided evidence for the existence of a mitochondrial DNA lineage (*P. hispanica* Galera type) endemic to Southeastern Spain that was not detected

in previous studies. This newly described lineage is sister to *P. hispanica* from the Northeast of the Iberian Peninsula, although divergence between them is ancient. Because previous studies had detected only *P. hispanica* sensu stricto in the area (Harris and Sá-Sousa 2002, Harris *et al.* 2002b), it was suspected that this lineage would have a very relictic distribution. However, recent studies (Renoult 2006) have demonstrated that in fact the distribution area of this mtDNA lineage is larger than previously thought. This raises some concerns relative to the misuse of the name *P. hispanica* sensu stricto in previous studies. The type locality for *P. hispanica* (described by Steindachner 1870) is situated near Murcia, southeastern Spain; because previous studies had only identified one lineage in southern Iberia, this clade within *P. hispanica* was automatically referred to as "*P. hispanica* sensu stricto" (Harris and Sá-Sousa 2002). Recent studies suggest, however, that both lineages are present near the Murcia region. Geniez and collaborators (in press) and Renoult (2006) consider the "Galera" form to be the nominotypical taxon. Because this subject is still not completely clarified and for a matter of correspondence with the nomenclature used in previous articles, we have used the name "*P. hispanica* sensu stricto" to designate the form more related to North African lizards throughout this work, keeping however in mind that this nomenclature could change in future studies.

4.1.1.2. Assessing the phylogenetic position of Podarcis *from Algeria*

A key gap in sampling systems in all former phylogenetic studies in *Podarcis* was the lack of sequences from Algeria, mainly for political reasons. Previously described genetic variation within North African *Podarcis* is complex, with two differentiated entities detected in Morocco and a third one in Tunisia. Therefore, in theory Algeria could contain representative of all three forms. Moreover, many species of reptiles (e.g. *Acanthodactylus* (Harris *et al.* 2004a); *Tarentola* (Harris *et al.* 2004b) show a very high degree of geographic substructure within North Africa, so the hypothesis of completely new lineages being found in this territory was not farfetched. We obtained samples of Algerian *Podarcis* from several localities in June 2006; 14 of these samples were sequenced for the mitochondrial gene ND4 for a preliminary assignment to a previously described mtDNA lineage. The results from these analyses are shown in figure 4.3 A (see also figure 4.4 and the appendix for the geographic origin of studied samples). Samples from Western and Central Algeria cluster within *P. vaucheri*, with low differentiation from its Moroccan relatives. However, samples from two localities in Eastern Algeria, although clustering with samples from Tunisia, show a considerable amount of differentiation, not only relative to Tunisian samples (Dxy of 0.063 and 0.065 for the southern and northern locality, respectively) but

also among themselves (Dxy=0.075). These values are higher than levels of differentiation found in this gene between, for example, *P. hispanica* types 1A and 1B.

Figure 4.3. Phylogenetic position of *Podarcis* from Algeria. The trees represent relationships among Iberian and North African *Podarcis* based on partial sequences of the ND4 gene (left) and a combination of ND4 and 12S rRNA sequences for a subset of the samples (right). In the ND4 tree, individuals from Algeria are represented by an asterisk. Bootstrap values above 50% are shown. Colours correspond to those in figure 4.4.

To confirm these lizards' phylogenetic position, we added sequence data for the 12S rRNA gene for one representative of each of these two new lineages. Maximum likelihood estimates of relationships obtained with this data set are shown in figure 4.3. B. Within the Jebel Sirwah-Algerian-Tunisian clade, estimates of relationships remained unchanged. These results clearly suggest that further sampling in Algeria is needed to assess the distribution of the detected lineages; our preliminary data suggest that that *P. vaucheri* might be sympatric with one of the divergent forms at least in Eastern Algeria, but this needs further investigation.

4.1.1.3. Using mitochondrial DNA sequence data to build a preliminary distribution map of Podarcis *lineages*

In order to assess with accuracy the distribution of mitochondrial DNA lineages, we compiled all published information including mitochondrial sequence data for Iberian and North African *Podarcis* (Castilla *et al.* 1998a,b, Oliverio *et al.* 2000, Harris and Sá-Sousa, 2001, 2002, Harris *et al.* 2002a, Harris *et al.* 2002b; Carranza *et al.* 2004a; Busack *et al.* 2005; Renoult 2006, Arntzen and Sá-Sousa, in press, articles I, IV and VI of this thesis) and added large amounts of unpublished data produced during the course of this PhD work (see Appendix for a complete list of included samples). The map resulting from this compilation is presented in figure 4.4. Although many areas have only been sparsely sampled (e.g. Algeria, some regions in Andalucia, Castilla-la Mancha and northern Castilla y León), the amount of data assembled is already highly representative of the lineages' distribution. Excluding the pair *P. hispanica* "Galera type" and *P. hispanica* sensu stricto, which are clearly sympatric, a remarkable trend is an almost complete geographic segregation of *P. hispanica* forms. Because most samples usually consist of one or a few individuals per locality and large territories remain unsampled, we might however be overlooking areas of distribution overlap. Nevertheless, a few contact zones are beginning to delineate; given the present data, it seems that these contact zones are most likely narrow. A further interesting feature is that areas of transition between forms do not seem to be marked by any geographical barrier or discontinuity. Whether these transitions result from a differential adaptation of lineages to different environmental backgrounds or from competitive exclusion remains to be investigated. As more data become available, a welcome step in this respect will therefore be the modelling of the species/lineages distributions.

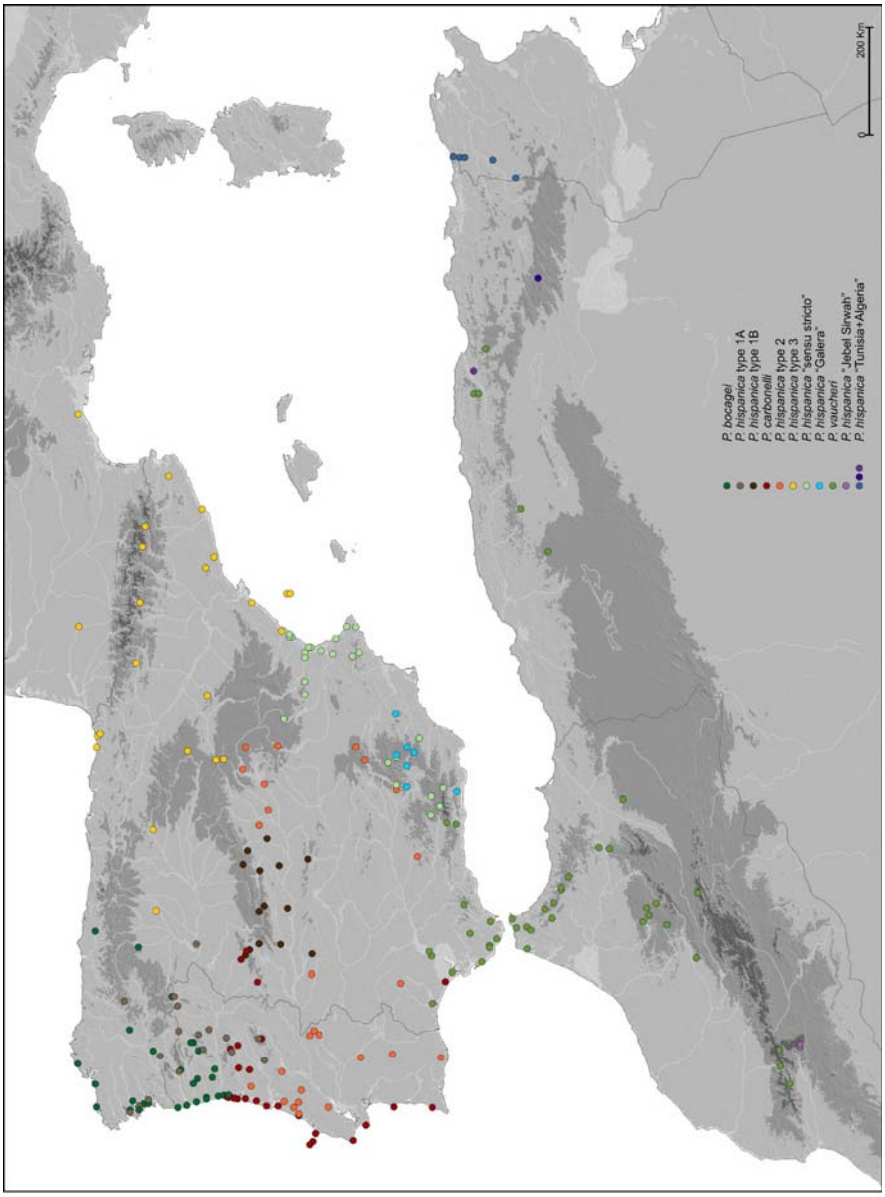

Figure 4.4. Distribution of *Podarcis* mtDNA lineages in the Iberian Península and North Africa based on an assemblage of published and unpublished data.

4.1.2. Nuclear gene variation in Iberian and North African *Podarcis*

The analysis of genetic variability at the mitochondrial level provided evidence for the existence of several cryptic and well-differentiated entities within Iberian and North African *Podarcis*. However, in the analyses of evolutionary processes, the study of genetic variation at more than one locus is of utmost importance. On one hand, it has long been realized that single locus-based inferences may be erroneous because of stochastic lineage sorting (Neigel and Avise 1993, Irwin 2002) or selection (Ballard and Whitlock 2004) and may also be less precise in the estimation of divergence times or effective population sizes (Edwards and Beerli 2000). As data from different loci accumulate, it thus becomes easier to disentangle patterns that are locus-specific from those that reflect general evolutionary trends, while reducing variance in parameter estimation. On the other hand, mitochondrial DNA has several properties (such as its haploid nature and exclusively maternal inheritance) that may lead to highly biased perspectives on evolutionary processes if genetic variation at nuclear loci is not taken into account (Nyakaana and Arctander 1999, Evans *et al.* 2003).

In this work, two different approaches were used to study genetic variation at the nuclear level across *Podarcis* species: allozyme electrophoresis and nuclear intron sequencing. Both approaches are complementary: on one hand, the study of allozymes enabled us to screen a large number of loci for many individuals; on the other hand, the study of nuclear gene sequences provides a genealogical perspective on nuclear gene evolution which can be directly compared to mitochondrial DNA variation.

A first general goal of the analyses of nuclear markers in wall lizards was therefore to evaluate whether the existence of several evolutionarily distinct entities suggested by mitochondrial DNA variation could be confirmed. A second goal was to assess whether these forms were reproductively isolated and if they corresponded to distinct species under a biological species concept (Mayr 1942). It has been shown that species of *Podarcis* may take a long time of divergence before becoming fully reproductive isolated (Capula 1993, 2000). Given this scenario, the possibility of gene flow among forms of the Iberian and North African *Podarcis* species complex should not be neglected. Because of its haploid nature, the study of mtDNA alone does not provide information on hybridization or introgression unless there are compelling morphological, ethological or ecological differences among the putatively hybridizing taxa that can be used to assess patterns of concordance with mtDNA variation. Since this is not the case for most of the forms within the *Podarcis* species complex, the study of nuclear gene variation becomes the only way of assessing levels of admixture between forms.

Traditionally, the investigation of the level of gene flow between emerging species is challenging because of the possibility of incomplete sorting of ancestral polymorphism; two species may share alleles not because they are hybridizing but simply because they have diverged recently and still share ancestral variation. To cope with these difficulties, in recent years refined analytical methods have been developed to investigate admixture between recently diverged populations. One class of these methods takes into account multilocus data (such as allozymes, microsatellites or AFLP) to infer discrete genetic clusters and estimate admixture between them (e.g. Pritchard *et al.*'s (2000) STRUCTURE software). Another class of approaches, such as the coalescent simulation method of Hey and Nielsen (2004) make use of multilocus genealogical data to estimate current and ancestral effective population sizes, divergence times and migration rates between pre-defined populations. We applied both of these approaches to measure the extent of contemporary and historical gene flow among *Podarcis* species.

4.1.2.1. Variation at allozyme loci

Although their use as a source of information has become less popular with the easiness of developing and screening other genetic markers (such as microsatellites and, more recently, multilocus nuclear genealogies), allozyme loci have been used for decades for assessing genetic differentiation among cryptic organisms, including several lacertid species complexes (Capula 1994a,b, MacCulloch *et al.* 2000, Mayer and Arribas 1996). A particularly welcome advantage of this type of analyses is that large amounts of data can be obtained even for non-model organisms, whereas the development of primers for amplification and sequencing of nuclear genes often require genomic resources that are only now beginning to accumulate for reptiles.

In this work, we studied genetic variation at 10 polymorphic allozyme loci, screened by means of conventional starch gel electrophoresis (*GOT-1*, *GPI*, *IDH*, *MPI*, *PEPA, PEPD, PGD*) and isoelectric focusing (*LDH-2*, *PEPB*, *PGM*) and analysed a total of 569 individuals from 30 populations representative of all major mtDNA lineages known at the time (recently detected lineages from Algeria were not known at the time this work was carried out). Two populations of *Podarcis muralis* were also added as outgroup.

Our results suggest that a significant proportion of genetic variation lies within mtDNA-defined lineages (Φ_{CT}=0.458). Most of them are highly homogenous, clearly well-differentiated and supported by high bootstrap values when evolutionary relationships are depicted in a Neighbor-Joining tree (NJ, Saitou and Nei 1987) (figure 4.5A). The only exception is *P. hispanica* type 3

(from northeast Iberia), which appears as paraphyletic, with one population (Getaria) clustering closer to *P. carbonelli* than to other populations carrying its mtDNA type. Although this paraphyly is supported by very low bootstrap values, the differentiation of Getaria probably results from the fact that this population was a former island, which only recently became connected to the mainland. Island populations often bear the signs of strong founder effects coupled with low effective population sizes, which may cause drift to produce dramatic changes in allele frequencies over short periods of time. It should be noted that only one population of *P. hispanica* Galera type, *P. hispanica* Jebel Sirwah type and *P. hispanica* sensu stricto were sampled, which prevented the assessment of monophyly for these three species. Nevertheless, the concordance between the partitions inferred from allozyme and mitochondrial DNA strongly suggests the validity of the multiple lineages within the *Podarcis* species complex as distinct evolutionary entities. To compare the level of differentiation observed among fully-recognised species within the species complex (*P. bocagei, P. carbonelli*, *P. vaucheri* and *P. hispanica*) to that observed among different "types" within *P. hispanica*, we plotted average F_{ST} values for all pairs of mtDNA-defined lineages (Figure 4.5B). The observation of this figure clearly suggests that levels of differentiation among forms of *P. hispanica* falls within the same values observed between recognised species, providing unequivocal evidence that a taxonomic revision of the *P. hispanica* complex is needed.

We further explored the subject of the distinctiveness among forms of *Podarcis* by conducting a model-based clustering analysis of individual multilocus genotypes using the software STRUCTURE. With these analyses we attempted to investigate whether the differentiation observed by the grouping of populations into well-defined clusters was corroborated at the individual level. Notably, we were able to detect nine homogenous clusters, with high levels of "clusteredness" (sensu Rosenberg *et al.* 2005) (figure 4.5C). Six of these groups neatly correspond to mtDNA-defined lineages, whereas the remaining three include more than one form: the pairs *P. hispanica* type 3/*P. hispanica* sensu stricto, *P. hispanica* "Jebel Sirwah"/*P. hispanica* "Tunisia" and *P. hispanica* "Galera"/*P. vaucheri* are indistinguishable. The former two cases furthermore correspond to the only multiple-lineage clusters that receive high bootstrap support in the NJ tree. Except the pair "Jebel Sirwah"/"Tunisia", which are evidently closely related (as attested by the mtDNA-nuclear concordance), none of these clusters have correspondence with the estimates of relationships inferred from mtDNA.

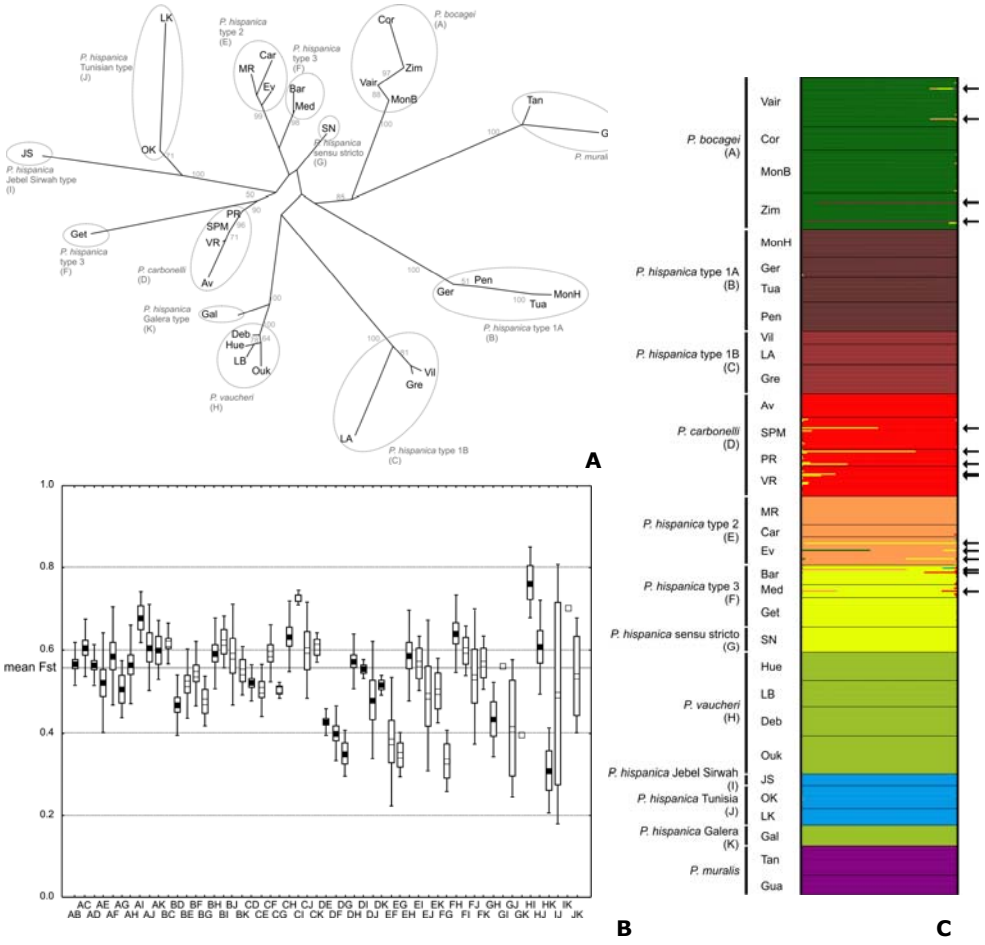

Figure 4.5. Inferences from allozyme variation in Iberian and North African *Podarcis*. **A.** NJ tree showing the relationships between 32 populations of Iberian and North African *Podarcis* using Nei (1972) genetic distances based on 10 allozyme loci. Populations corresponding to the same mitochondrial DNA lineages are encircled. Bootstrap values above 50% are shown. **B.** Box plot representation of mean Fst values observed between pairs of forms of *Podarcis*, including comparisons between recognized species (filled squares) and groups within *P. hispanica* (white squares). Pairs of letters indicated on the x axis correspond to the mitochondrial lineages being compared, using the same nomenclature shown in C. The boxes represent the standard errors and the whiskers the standard deviations. The dotted line across the graph shows the mean F_{ST} value for the whole sample. *P. muralis* was excluded from this analysis. **C.** Estimated probability of ancestry of 569 individuals belonging to the Iberian and North African *Podarcis* species complex and to the outgroup *P. muralis*, calculated using the software STRUCTURE, considering K=9 clusters. Each individual is represented by a horizontal line divided into 9 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.

Several explanations can be suggested to account for these two cases of nuclear/mitochondrial discrepancy. First, these clusters could result from

massive introgression in the nuclear genome while the mtDNA remains distinct; even assuming non-male-biased gene flow, this hypothesis is not farfetched as the mtDNA, because of its smaller effective population size, is more prone to the effects of drift and may have locally fixed the original divergent haplotypes. A similar situation has been recently described in the midwife toad *Alytes obstetricans* (Martínez-Solano *et al.* 2004). In this species, the subspecies *almogavarii* presents a well-differentiated mitochondrial DNA lineage but little differentiation from conspecific populations at the nuclear level resulting from progressive dilution through introgression. Second, mtDNA could be portraying an erroneous picture of evolutionary relationships because of chance or selection, and these two pairs of species could in fact be closely related. A third explanation could be related to balancing selection acting on allozyme loci, as suggested by Karl and Avise (1992), and Pogson *et al.* (1995) to explain low levels of differentiation observed in allozymes when compared to other markers. A fourth hypothesis is that the used set of markers is insufficient to provide a clear separation between the referred species pairs; since only ten allozyme loci were surveyed, this set of markers could be at its limit of detecting population structure.

Apart from these cases where lineages appear to have a close relationship, clusters in the NJ tree involving more than one form are not supported by high bootstrap values. In fact, although branches leading to species or forms are usually long, internal branches in the tree are extremely short. This situation is mirrored in the individual clustering approach, in which the clusters produced when lower numbers of populations were enforced were highly inconsistent, albeit each form remained homogenous. This implies that from an allozyme perspective, relationships among forms cannot be confidently resolved. It is likely that this results from the fact that *Podarcis* species are young and diverged almost simultaneously, which precludes the representation of their evolutionary relationships by a clearly bifurcating tree. Instead of a step-by-step speciation scenario, as portrayed by mtDNA estimates of relationships, allozymes thus suggest a rapid radiation. This discrepancy probably reflects differences in the markers' effective population size.

Besides above-reported cases of potential introgression, the multilocus clustering analyses revealed only a few individuals with potentially admixed or misassigned genotype. Most of these cases likely result from the persistence of shared ancestral polymorphism; because many alleles are still shared between species, apparently admixed multilocus genotypes can be produced in the absence of hybridization. However, we found two individuals that appear as intermediate between *P. bocagei* and *P. hispanica* type 1A, which are not easily explainable taking into account the hypothesis of shared ancestral polymorphisms. In fact, these two species do not share as many alleles at

informative loci as other species pairs; they are, moreover, sympatric. Taking these aspects into account, these two individuals most likely represent hybrids between both species. Albeit natural hybridization between distinct *Podarcis* species has been suggested to occur between other species within the genus (Capula 1993, 2002), this constitutes one of the first reports of natural hybridization among species of the Iberian and North African *Podarcis* species complex. Recently, interspecific matings between these two species have been observed in the field (Ribeiro and Carretero, pers. comm.).

In conclusion, the study of allozyme variation in Iberian and North African wall lizards provided strong support for the distinctiveness of mtDNA-defined groups at the nuclear level. Nevertheless, discrepancies in species delimitation and patterns of species divergence may suggest that evolutionary dynamics within *Podarcis* may be more complex than previously appreciated.

4.1.2.2. Nuclear gene genealogies

To further explore the patterns of concordance (or lack thereof) of nuclear and mitochondrial genetic variation in Iberian and North African *Podarcis,* we turned to the study of nuclear gene genealogies. Apart from a few recent examples (e.g. Creer *et al.* 2003, 2006, Godinho *et al.* 2006a, b, Dolman and Moritz 2006), the study of nuclear DNA sequences in reptiles is still in development due to the lack of genomic resources for these organisms. Prior attempts to study genetic variation at the nuclear gene C-*mos* in *Podarcis* (D.J. Harris, unpublished results) were hampered by very low levels of sequence polymorphism and therefore of utility in resolving relationships among species. For this study, we therefore chose two gene regions for which previous reports of genetic variation suggested a fast evolutionary rate: intron 7 of the *β*-fibrinogen gene (*βfibint-7*) and intron 7 of the 6-phosphogluconate dehydrogenase gene (*6-Pgdint7)*. The first gene region was chosen for this study because it has been successfully applied to both phylogeographic or phylogenetic assessments among birds, in particular (e.g. Prychitko and Moore 1997, 2003, Bonaccorso and Peterson 2007), but also in mammals (Krajewsky *et al.* 2004, Yu and Zhang 2006), amphibians (Sequeira *et al.* 2006) and reptiles (Creer *et al.* 2003, 2006, Godinho *et al.* 2005, 2006) and shows remarkable levels of nucleotide diversity among the studied organisms. On the other hand, the study of the second gene region was developed especially for this survey, since reports of allozyme variation at the *6-PGD* locus had demonstrated outstanding levels of genetic variability within *Podarcis* (articles II, III, Pinho *et al.* 2004).

We studied sequence data variation for these two genes in a total of 78 individuals for which a portion of the mitochondrial DNA (ND4) was also sequenced in order to allow the assignment of all individuals to a mitochondrial DNA lineage. Sequence data for these two gene regions revealed moderate levels of genetic variability, lower than some mitochondrial DNA genes (such as the fast-evolving cytochrome *b* and ND4) but higher, for example, than the fragment of the mitochondrial control region studied in article I.

Surprisingly, our estimates of relationships between alleles for both genealogies reveal a striking pattern of polyphyly of mitochondrial DNA-defined lineages (figure 4.6). The only concordant feature observed between both genealogies and mitochondrial branching patterns is the monophyly of the clade formed by *P. hispanica* from Tunisia and from Jebel Sirwah; no single species forms a consistent clade, and not even the three main divergent clades inferred from mtDNA variation are recovered. This is surprising for organisms that were inferred to have started to diverge at least 4 – 5 million years ago, even for some of the most recently diverged species pairs.

Taking into account a simple comparison of mtDNA-nuclear DNA, we would be led to think that simply mtDNA-defined species do not correspond to distinct evolutionary entities. In fact, with a few exceptions, these lineages are indistinguishable with regard to nuclear gene genealogies. It has indeed been demonstrated using coalescent simulations that rather deep phylogeographic breaks may arise within a species distribution even in the absence of barriers to gene flow, simply because of stochastic lineage sorting; this argument has been invoked to explain why species that are structured according to their mtDNA variation do not have correspondingly structured patterns in phenotype or other traits (Irwin 2002). Likewise, it has been shown that divergent mtDNA lineages in *Drosophila simulans* do no find correspondence in nuclear variation, morphology or behaviour, but instead could result from selective mechanisms acting on the mitochondrial genome alone (Ballard *et al.* 2002).

These hypotheses could eventually be satisfactory to explain the patterns observed in *Podarcis* had we not other sources of information corroborating the distinctiveness of mitochondrial DNA lineages. In fact, albeit large-scale morphological studies are still lacking for the species complex, the available data shows remarkably concordant patterns with mtDNA-defined entities (Harris and Sá-Sousa 2001, 2002, Geniez 2001, Sá-Sousa *et al.* 2002; but see Renoult 2006). Moreover, other nuclear markers (allozymes) also support the general distinctiveness of such forms (articles II and III). Therefore, we have to seek other explanations that can elucidate why nuclear gene genealogies are not portraying the history of differentiation that was inferred from other markers.

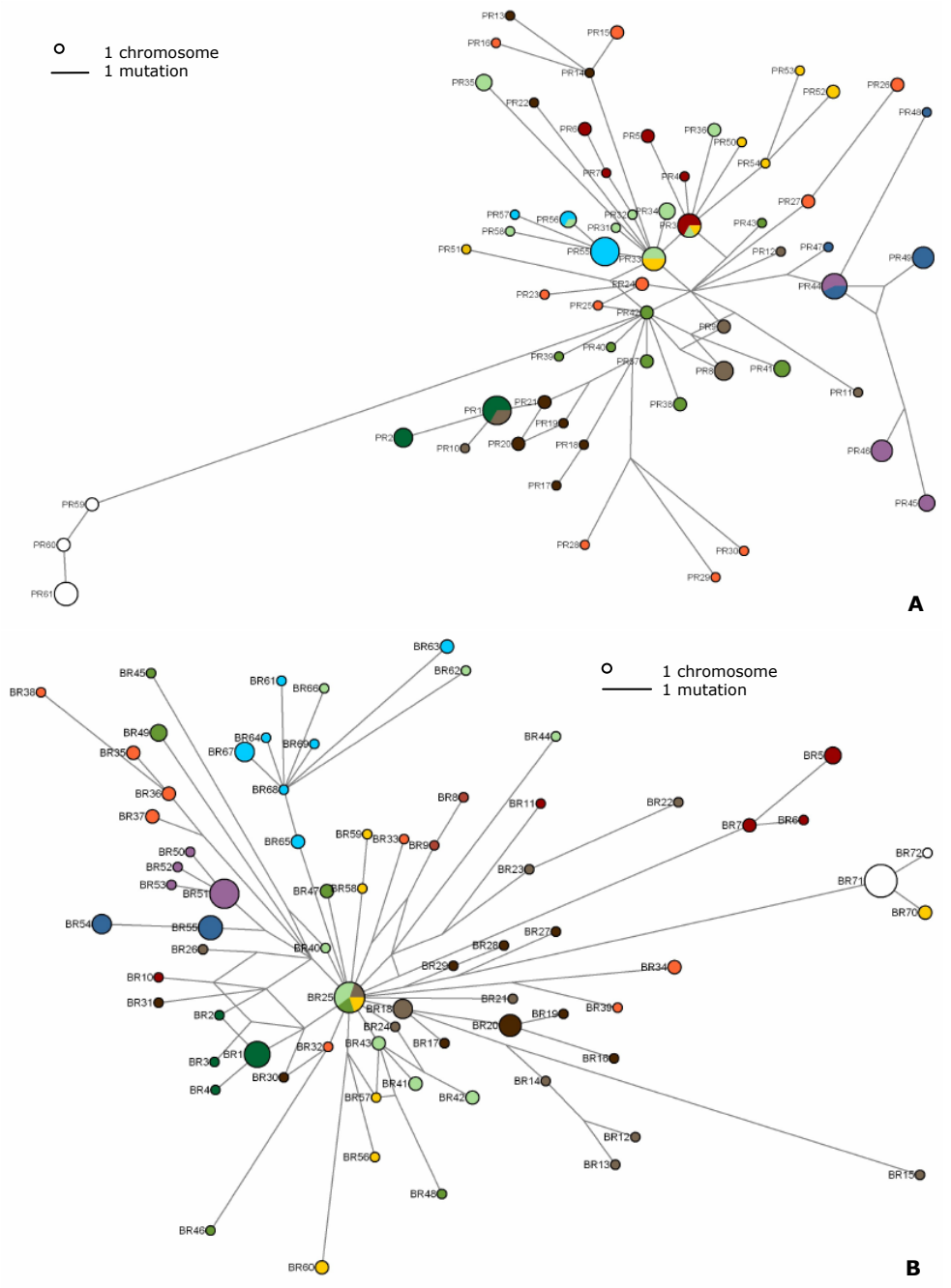

Figure 4.6. Gene genealogies for two nuclear introns in Iberian and North African *Podarcis.* Colours represent mitochondrial DNA lineages and correspond to those used in figure 4.4. Alleles detected in *P. muralis* are represented in white. **A.** *6-Pgdint7.* **B**. *β-fibint7.*

A possible preliminary hypothesis to account for such discrepancy could be related to the lower levels of variability observed in both nuclear genes when compared to their mitochondrial counterparts, yielding low resolution for relationship estimates. However, the portion of the control region analysed in article I, although presenting a lower percentage of variable positions, still renders all lineages monophyletic (results not shown). Moreover, if lack of informative characters was the main reason for the observed discrepant results, we would expect that the enforcement of species monophyly in maximum-likelihood searches would not produce significantly different topologies than our best unconstrained estimates. However, when compared through Shimodaira-Hasegawa tests (Shimodaira and Hasegawa 1999), trees constrained for species monophyly are significantly less likely.

Based on these observations, the lack of informative genetic variation cannot be accounted for as a major explanation for the observed discordant patterns. We are then left with two non-mutually exclusive hypotheses: a) the observed patterns emerge as a consequence of incomplete lineage sorting of ancestral polymorphism; b) there are significant levels of gene flow going on between lineages (particularly male-biased, since differentiation at mitochondrial loci is high) that explain the almost random distribution of genetic variation across forms. Again, both hypotheses seem preliminarily plausible. Mitochondrial DNA takes, in theory, a quarter of the time a nuclear gene does to reach reciprocal monophyly (Moore 1995; in this case, we are defining mtDNA monophyly considering morphological units); therefore, when differentiation is recent, a possibility is that mtDNA might have already reached monophyly whereas nuclear genes have not. Gene flow across lineages is also a theoretical possibility; it has been shown that *Podarcis* species may take a long divergence time to acquire full reproductive isolation, as documented by natural hybridization reports between the distantly related *P. sicula/P. tiliguerta* and *P. sicula/P. wagleriana* (Capula 1993, 2002).

Distinguishing between the effects of these two forces is challenging when dealing with closely-related species, and it has been proven to be a difficult task across several organisms (Hey *et al.* 2004, Muir and Schlötterer 2005, 2006, Buckley *et al.* 2006, Bull *et al.* 2006, Lexer *et al.* 2006). In the case of our study organisms, this task is of major importance for understanding of whether forms of *Podarcis* represent true species in the framework of the biological species concept, as suggested by other lines of evidence. Our data strongly suggest that the polyphyletic pattern derives mostly from the incomplete lineage sorting of ancestral polymorphism, although gene flow may occasionally occur among forms. Three distinct lines of evidence support this observation:

a) comparison with allozyme data. As demonstrated in article III, most forms of *Podarcis* are recognisable taking into account not only population relationships but also individual multilocus genotypes, which suggest a general pattern of isolation among forms.

b) relationship between allele age and trans-species relatedness. If gene flow was the main factor causing species polyphyly, we would expect that both old and young alleles would be similarly shared among forms. This is, however, not the case, at least in the *βfibint-7* gene. Alleles that are placed near the network centre are also alleles that are more prone to be shared or closely-related to alleles found in different species, whereas alleles that are placed as tips are likely to have arisen after the species were formed and are generally not shared and distantly related to other species' alleles.

c) *estimates of migration rates among forms*. Obtaining realistic estimates of gene flow among recently diverged species is challenging because of the potential for variation to be shared due to incomplete lineage sorting. Classic F_{ST}-based admixture estimates assume equilibrium and therefore overlook this possibility; on the other hand, recently-developed methods based on coalescent theory consider divergence and admixture in the same framework and are therefore suitable to estimate rates of gene flow even in the presence of shared ancestral polymorphism (Hey and Nielsen 2004). We applied both classes of methods to our data sets. Both of them suggest generally low levels of migration among forms; however, classic estimates point to non-zero admixture whereas coalescent estimates suggest virtually zero gene flow among most species pairs (table 4.1). This strongly suggests that despite their polyphyly species of Iberian and North African *Podarcis* are indeed isolated.

Although several lines of evidence suggest an important role for ancestral polymorphism sharing in the polyphyly of nuclear gene genealogies in wall lizards, clearly non-zero levels of historical gene flow were detected among some species pairs. Two of these cases stand out, as they suggest levels of admixture that would seem to prevent species differentiation (that is, they represent *Nm* values above 1 (Wright 1931)). These two cases correspond to previously described instances of gene flow, inferred from the study of other markers: between *P. bocagei* and *P. hispanica* type 1A (suggested in article III) and between *P. hispanica* sensu stricto and *P. hispanica* type 3 (suggested in the same article due to the impossibility of discriminating between these forms and by Renoult (2006)). A third case of important levels of gene flow was detected from *P. hispanica* Galera type to *P. hispanica* sensu stricto and from the latter to *P. vaucheri*. Low (but non-zero) levels of gene flow were detected between

Table 4.1. Maximum-likelihood estimates and 90% highest posterior distributions of population migration rates (*2Nm*) between species of Iberian and North African *Podarcis*, calculated using an Isolation with Migration (IM) model (Hey and Nielsen 2004). Values higher than 0.05 are highlighted. Species names are abbreviated as follows: *Pb* – *P. bocagei* ; *Pc* – *P. carbonelli*; *Pv* – *P. vaucheri*; *Ph*1A – *P. hispanica* type 1A; *Ph*1B – *P. hispanica* type 1B; *Ph*2 – *P. hispanica* type 2; *Ph*ss – *P. hispanica* sensu stricto; *Ph*3 – *P. hispanica* type 3; *Ph*Gal – *P. hispanica* Galera type; *Ph*JS – *P. hispanica* Jebel Sirwah type; *Ph*Tun – *P. hispanica* Tunisian type; *Pm* – *P. muralis*.

		from											
		Pb	***Pc***	***Pv***	***Ph*1A**	***Ph*1B**	***Ph*2**	***Phss***	***Ph*3**	***Ph*Gal**	***Ph*JS**	***Ph*Tun**	***Pm***
to	***Pb***	-	0.001 0.000-1.280	0.001 0.000-0.796	0.001 0.000-2.884^{+}	0.000 0.000-1.787^{+}	0.001 0.000-0.222	0.015 0.005-4.637^{+}	0.001 0.000-1.110	0.016 0.000-0.646	0.001 0.000-0.757	0.001 0.000-0.782	0.001 0.000-0.348
	Pc	0.035 0.002-2.165	-	0.005 0.002-1.622	0.086 0.002-1.875	**0.145** 0.002-1.934	0.006 0.002-1.525	0.004 0.002-4.968	0.005 0.002-4.603	0.005 0.003-1.883	0.006 0.003-1.852	0.006 0.003-1.983	0.006 0.003-0.617
	Pv	0.016 0.006-12.278	0.011 0.005-3.325	-	0.015 0.002-23.282^{+}	0.008 0.001-18.577^{+}	0.011 0.004-3.654	**1.246** 0.003-9.682	**0.228** 0.001-12.485^{+}	0.010 0.004-8.070	0.009 0.004-2.861	0.009 0.004-3.761	0.010 0.005-1.014
	***Ph*1A**	**6.376** 0.009-50.859*$^{+}$	0.014 0.005-2.504	0.015 0.007-9.151	-	0.017 0.007-27.424*	0.018 0.010-1.874	0.014 0.006-11.321	0.017 0.009-3.993	0.013 0.007-2.034	0.017 0.008-16.265^{+}	0.018 0.008-34.647*	0.011 0.006-0.839
	***Ph*1B**	0.012 0.006-26.036*	0.010 0.006-2.069	0.036 0.004-58.638*$^{+}$	0.009 0.004-6.992	-	0.012 0.006-3.797	**0.336** 0.004-18.837	0.012 0.005-7.917	0.009 0.005-2.336	0.012 0.005-6.189^{+}	0.012 0.005-5.007	0.009 0.005-0.880
	***Ph*2**	0.018 0.010-11.758	0.020 0.011-1.830	0.018 0.010-3.050	0.017 0.010-1.788	0.017 0.008-5.149	-	0.016 0.009-1.942	0.017 0.009-2.209	0.018 0.010-2.335	0.018 0.009-2.196	0.017 0.009-2.553	0.015 0.009-0.891
	Phss	0.001 0.000-0.417	**0.109** 0.002-5.794	0.005 0.002-1.973	0.003 0.002-1.481	0.005 0.002-5.716	0.005 0.002-0.814	-	0.003 0.001-13.098^{+}	**1.156** 0.129-6.072	0.005 0.002-1.424	0.005 0.002-1.719	0.005 0.003-0.501
	***Ph*3**	0.012 0.005-8.500^{+}	0.014 0.005-19.796	0.012 0.004-47.275*$^{+}$	0.008 0.003-3.467	0.011 0.004-9.743	0.010 0.005-1.527	**5.484** 0.004-75.707*$^{+}$	-	0.013 0.005-10.353*	0.009 0.004-6.929*	0.010 0.004-9.566	**0.134** 0.004-2.073
	***Ph*Gal**	0.002 0.001-0.665	0.002 0.001-0.722	0.002 0.001-0.880	0.002 0.001-0.542	0.002 0.001-0.653	0.002 0.001-0.553	0.002 0.001-0.856	0.002 0.001-0.640	-	0.002 0.001-0.708	0.002 0.001-0.740	0.002 0.001-0.391
	***Ph*JS**	0.001 0.000-0.737	0.001 0.000-1.435^{+}	0.001 0.000-1.219^{+}	0.000 0.000-2.432	0.001 0.000-1.823^{+}	0.001 0.000-0.750	0.001 0.000-0.899	0.001 0.000-1.171^{+}	0.003 0.000-0.743	-	**0.107** 0.000-1.548^{+}	0.001 0.000-0.277
	***Ph*Tun**	0.003 0.000-0.721	0.001 0.000-1.116	0.001 0.000-0.914	0.001 0.000-0.645	0.001 0.000-0.809	0.001 0.000-0.746	0.001 0.000-0.900	0.001 0.000-0.983	0.001 0.000-0.752	0.002 0.000-1.142	-	0.001 0.000-0.309
	Pm	0.003 0.000-0.267	0.001 0.000-0.350	0.018 0.000-0.500	0.017 0.000-0.461	0.010 0.000-0.457	0.001 0.000-0.432	0.016 0.000-0.354	0.000 0.000-0.658^{+}	0.005 0.000-0.264	0.002 0.000-0.219	0.001 0.000-0.228	-

*, $^{+}$: the actual interval could not be reliably estimated because the likelihood surfaces for θ (*) or m ($^{+}$) were relatively flat.

several other species pairs, either sympatric, parapatric or allopatric, suggesting that present contact is not a prerequisite for gene exchange. These last cases of observation of gene flow between species that are separated by the distribution of other species, suggest that probable changes in distributions promoted contacts in the past, allowing for gene exchange to occur. Noteworthy is the finding of gene flow between *P. hispanica* type 3 and *P. muralis*, which was surprising given that *P. muralis* has been diverging from other *Podarcis*, including the Iberian and North African clade, for over 10 million years, according to Oliverio *et al.* (2000).

Although nuclear genealogies are therefore concordant with other markers in defining species of *Podarcis* as generally reproductively isolated, with exceptions to this rule that are also in agreement with previously documented cases of introgression, this study constitutes a very preliminary effort regarding the description of nuclear genealogies in these lizards, both because of the low number of individuals analysed for each species, and because it included only two nuclear genealogies. The inclusion of more samples and genes will certainly improve our understanding of the evolution of the genus, as well as increase our confidence in the estimates of historical gene flow among forms.

4.1.2.3. Implications of the study of nuclear markers for the understanding of the origin and evolution of Podarcis *lineages*

Looking with detail into results obtained through the analyses of nuclear markers (both allozymes and nuclear gene genealogies) in Iberian and North African *Podarcis*, we are able to produce a few generalizations that describe patterns of nuclear variation in these lizards:

a) Podarcis *mtDNA lineages are generally well-differentiated with respect to the nuclear genome*

One of the main goals of analysing nuclear markers in Iberian and North African *Podarcis* was to assess, by examining patterns of nuclear-mitochondrial concordance, whether the highly divergent lineages observed in mtDNA phylogenetic studies could be candidates for species status recognition. In this respect, although there are some important differences in species delimitation that may suggest that not all the mtDNA lineages are equally differentiated and deservers of the specific status (e.g. *P. hispanica* from Tunisia and *P. hispanica* from Jebel Sirwah and, in particular, *P. hispanica* sensu stricto and *P. hispanica* type 3, seem rather undifferentiated) nuclear markers are fairly concordant in describing most forms within the studied organisms as divergent, clearly identifiable from multilocus perspectives and generally reproductively isolated.

Because more loci were analysed, allozymes provided more information than nuclear genealogies with respect to the first two observations (high divergence levels and species identifiability). Nevertheless, both types of nuclear markers strongly suggest that reproductive isolation amongst mtDNA-defined species is a general trend within Iberian and North African *Podarcis*.

b) Gene flow exists among lineages of wall lizards

The study of nuclear genes also helped us to preliminarily detect and characterize gene flow among species, which constituted one of the most important questions that were left unaddressed by studies of mitochondrial DNA variation, partly because of this marker's haploid nature and maternal inheritance, partly because of the lack of a solid morphological framework that would allow a direct comparison between both sources of information. In fact, gene flow among distinct forms within the Iberian Peninsula and North Africa was not unexpected, given that hybridization had been previously documented among other species of the genus that supposedly had started to diverge long before diversification within this clade (Capula 1993, 2002); moreover, it was also shown that some Iberian species (namely *P. bocagei* and *P. carbonelli*) were able to be crossed in captivity (Galán 2002). By the analyses of nuclear markers, we have demonstrated that some Iberian and North African wall lizard species are indeed exchanging or have exchanged genes in their past. Notably, some of these cases are attested by different studies using different markers and sampling schemes (for example, gene flow between *P. bocagei* and *P. hispanica* type 1A, suggested both by articles III and IV; between *P. hispanica* sensu stricto and *P. hispanica* from Galera, suggested both by article IV and by Renoult (2006); between *P. hispanica* sensu stricto and *P. hispanica* type 3, suggested by articles III and IV and by Renoult (2006)), implying that important levels of gene flow might be occurring among these forms (see discussion below).

In the context of gene flow among forms, the case of *P. bocagei* and *P. hispanica* type 1A deserves our attention since different lines of evidence seem somewhat contradictory. On one hand, these two species are broadly sympatric, yet are well-distinguishable morphologically (Sá-Sousa 2001a). Clustering of individual allozyme multilocus genotypes for 105 *P. bocagei* and 72 *P. hispanica* type 1A from across their distribution areas report the existence of only two individuals presenting signs of introgression, which suggests that in general these two species also maintain their genetic integrity even in sympatric conditions (article III). In a behavioural study of these two species in strict syntopy (i.e. on the same wall), only one inter-specific mating was observed (R. Ribeiro and M.A. Carretero, personal communication). Moreover, it has been shown that males from these two species are able to recognise conspecific females, which suggests at least some degree of assortative mating (Barbosa *et*

al. 2006). Taken together, these observations would suggest limited amounts of gene flow among these forms. However, estimates of migration rates based on nuclear gene genealogies suggest much higher levels of introgression, probably beyond the level that would allow the maintenance of species differentiation (article IV). This discordance could be a bias of the low sample size employed in the study of nuclear genealogies; nevertheless, more detailed studies regarding the levels of gene flow among these species are clearly needed.

c) Some lineages within Podarcis *remain evolutionary puzzles*

Also in the context of gene exchange with other forms, we should discuss with more detail the situation of *P. hispanica* sensu stricto. A recent survey of genetic and morphological variation in *Podarcis* from southeastern Iberian Peninsula (Renoult 2006) found no evidence of the existence of this form based on morphology; all specimens carrying its mtDNA lineage could be morphologically attributed either to *P. hispanica* type 3 or to *P. hispanica* "Galera" type. This observation, coupled with some information from nuclear genes, led this author to suggest this mtDNA lineage was the only extant signature of an extinct species which nuclear background and morphology became diluted through massive introgression with the above-referred forms. Our results are again contradictory with this respect. On one hand, according to the allozymes we did not detect evidences of introgression of this form with the Galera lineage in a population that, based on the distribution map presented by Renoult (2006), should present the Galera typical morphology. However, this same population was indistinguishable from *P. hispanica* type 3 (article III). Taking into account nuclear gene genealogies, *P. hispanica* sensu stricto was inferred to be involved in limited to moderate gene flow with several other species, including *P. hispanica* "Galera" type, but these situations of gene exchange were inferred to have occurred at low levels that would not compromise this form's distinctiveness. On the other hand, high levels of gene flow between *P. hispanica* sensu stricto and *P. hispanica* type 3 were again suspected (unfortunately, reliable unimodal curves could not be obtained), in similarity to the findings based on allozyme differentiation (or, more precisely, the lack thereof). This could mean that these two mtDNA lineages do not currently correspond to distinct entities from a nuclear level, resulting from massive introgression of the nuclear genome and a persistence of divergent mtDNA lineages, as shown to have occurred between the Iberian and Brown hares with the Mountain hare in northern Iberia (Alves *et al.* 2003, Melo-Ferreira *et al.* 2005, Freitas 2006) and between *Alytes obstetricans almogavarii* and other subspecies (Martínez-Solano *et al.* 2004). Taking into account the directionality of admixture inferred between these two forms, our results would seem to indicate that migration occurred from *P. hispanica* sensu stricto into *P. hispanica* type 3, which

disagrees with the scenario suggested by Renoult (2006). Adding extra complexity to this already puzzling scenario, the form of *P. hispanica* from Galera seems to be indistinguishable from *P. vaucheri*, which also may be present in the region (see figure 4.4), using allozyme data (article III); however, no evidence for this was observed studying nuclear gene genealogies, where gene flow between these two species was found to be zero (article IV).

Given the highly contradictory nature of different studies analysing genetic variation within the southeastern corner of the Iberian Peninsula, more sampling (both in terms of genes and individuals) and a novel assessment of morphological distinctiveness using more refined techniques (e.g. Kaliontzopoulou *et al.* 2006) are clearly in need to accommodate the contrasting scenarios inferred from different studies and markers, prior to any taxonomic reevaluations of the involved mtDNA lineages.

d) Assessing relationships between forms and biogeographic scenarios using nuclear markers is difficult

A striking difference between patterns of genetic differentiation inferred from mitochondrial DNA and nuclear markers lies in the estimates of species relationships: whereas the study of mitochondrial DNA suggests a step-by-step speciation scenario, nuclear markers (either allozymes or nuclear genealogies) fail to provide any clues on clusters above species level. There is a single exception with this respect, which is the clustering of *P. hispanica* from Jebel Sirwah and Tunisia, constant across mtDNA, allozymes and both nuclear gene genealogies. This lack of signal for recovering phylogenetic relationships can be interpreted as resulting from a rapid and almost simultaneous diversification of all forms, causing, on one hand, the non-monophyly of any mtDNA-defined multispecies cluster, either young or old, in individual gene genealogies, and, on the other hand, discordance among distinct nuclear gene trees. This contrasting scenario between mitochondrial and nuclear patterns of divergence is most likely due to differences in both markers' effective population size. In fact, hints that speciation occurred at a fast rate can be found also in mitochondrial DNA: for example, although multispecies clusters are well supported by high bootstrap values, internal branches are quite short; in fact, applying the controversial "three-times rule" of Palumbi and collaborators (2001), we would actually not expect nuclear monophyly of *any* of the mtDNA-defined multispecies cluster (results not shown). Interestingly, this lack of utility of individual nuclear gene genealogies for inferring species relationships could also imply that overall effective population size in this group of lizards was low enough for stochastic lineage extinction cause the mtDNA to become fixed several times in the genus' evolution, but also high enough to avoid the fixation of a single nuclear allele throughout several successive episodes of isolation.

This observation has an obvious practical disadvantage: we have no way of testing whether species relationships and biogeographic scenarios inferred for mitochondrial DNA are valid from a nuclear perspective; moreover, it is unlikely that we will ever have, unless a very large number of genes is analysed in the future and that patterns of species relatedness start to emerge from the overall polyphyly.

4.1.3. Genetic characterization of the contact zone between *P. bocagei* and *P. carbonelli*

Speciation is commonly thought of as the process underlying the acquisition of reproductive isolation among diverging taxa. The degree of reproductive isolation can be correlated to the number of genes that are involved in epistatic incompatibilities (Dobzhansky 1937, Muller 1940, 1492); if only a few genes are responsible for functional, physiological or behavioural differentiation, than gene flow between diverging forms is extensive and the forms remain undifferentiated. As more genes become involved in such interactions, the more difficult it is for emerging species to engage in gene flow and more portions of the genome will become definitively isolated (Wu 2001); eventually, the species will not merge even if the barrier that led to their separation is removed. The assessment of the degree of reproductive isolation between closely related taxa thus provides important clues regarding the stage that they have achieved in the process of speciation. By analysing the dynamics of gene exchange across contact zones between recently diverged taxa, we may therefore attempt to predict how the future of the two populations will be in terms of fusion or divergence.

In the particular case of Iberian and North African wall lizards, evidence for general reproductive isolation between forms lies mainly on overall descriptions of genetic variability which uncover remarkably concordant patterns among different types of markers (articles I – IV). However, there is also evidence that in the genus *Podarcis* the potential for gene exchange with congenerics remains even after over ten million years of divergence, as shown by the observation of hybridization or introgression between distantly related taxa such as, for example, *P. sicula* and *P. wagleriana* (Capula 1993), *P. sicula* and *P. tiliguerta* (Capula 2002) and, more recently, *P. muralis* and *P. hispanica* (article IV). In addition, general surveys of genetic variation at the nuclear and mitochondrial level suggested the occurrence of gene flow among forms of the Iberian and North African species complex (article III, IV; Renoult 2006). However, the possibility of gene flow occurrence across contact zones cannot be correctly evaluated by simply describing general patterns of genetic variation. Moreover,

the study of contact zones may provide important insights regarding the mechanisms that are driving speciation and that are maintaining species boundaries, particularly when genetic data are put into context with other sources of evidence such as behaviour, physiology or ecology (Jiggins and Mallet 2000).

We therefore sought to complement our work regarding general patterns of differentiation and genetic exchange among wall lizard species by directly studying levels of gene flow across contact zones. As shown by figure 4.4, presently very few of these suture zones among *Podarcis* forms have been detected; one of the first such areas to be described was the contact zone between *P. bocagei* and *P. carbonelli* (Carretero *et al.* 2002). Individuals from both species were found to co-exist syntopically in a very narrow zone with only a few kilometres of width, near the locality of Espinho, in northern Portugal. We selected this area as a starting point for the description of levels of gene exchange across *Podarcis* contact zones. For this purpose, we sampled a transect along the Portuguese coast in the area where these two species replace each other. The precise locality where the two species are known to exist in syntopy was also sampled. We analysed a battery of 15 nuclear markers, including 10 allozyme loci (*GOT, GPI, IDH, LDH-2, MPI, PEPA, PEPB, PEPD, 6-PGD, PGM)* three microsatellites (*Pb11, Pb50* and *Pb66*) and one nuclear intron (a portion of the *6-Pgdint7* gene analysed through Single Strand Conformation Polymorphism), and also mitochondrial DNA (based on Restriction Fragment Length Polymorphism of the 12S rRNA gene).

4.1.3.1. Genetic evidence for admixture

Although only one of the nuclear loci analysed is fully diagnostic between the two species (*6-Pgdint7*), many others show species-specific alleles that help in putative hybrid identification. A preliminary observation of multilocus genotypes clearly showed a high prevalence of individuals showing signs of admixture between the two species in the locality where they are sympatric. In order to further explore patterns of introgression, we used, amongst other approaches, the model-based individual multilocus genotype clustering method implemented in software STRUCTURE (Pritchard *et al.* 2000). Both species are clearly distinguishable; individuals from populations in allopatry are assigned with high probability levels to their putative species of origin. In the contact zone, however, many individuals present signs of admixture, which confirms that the two species hybridize (figure 4.7B).

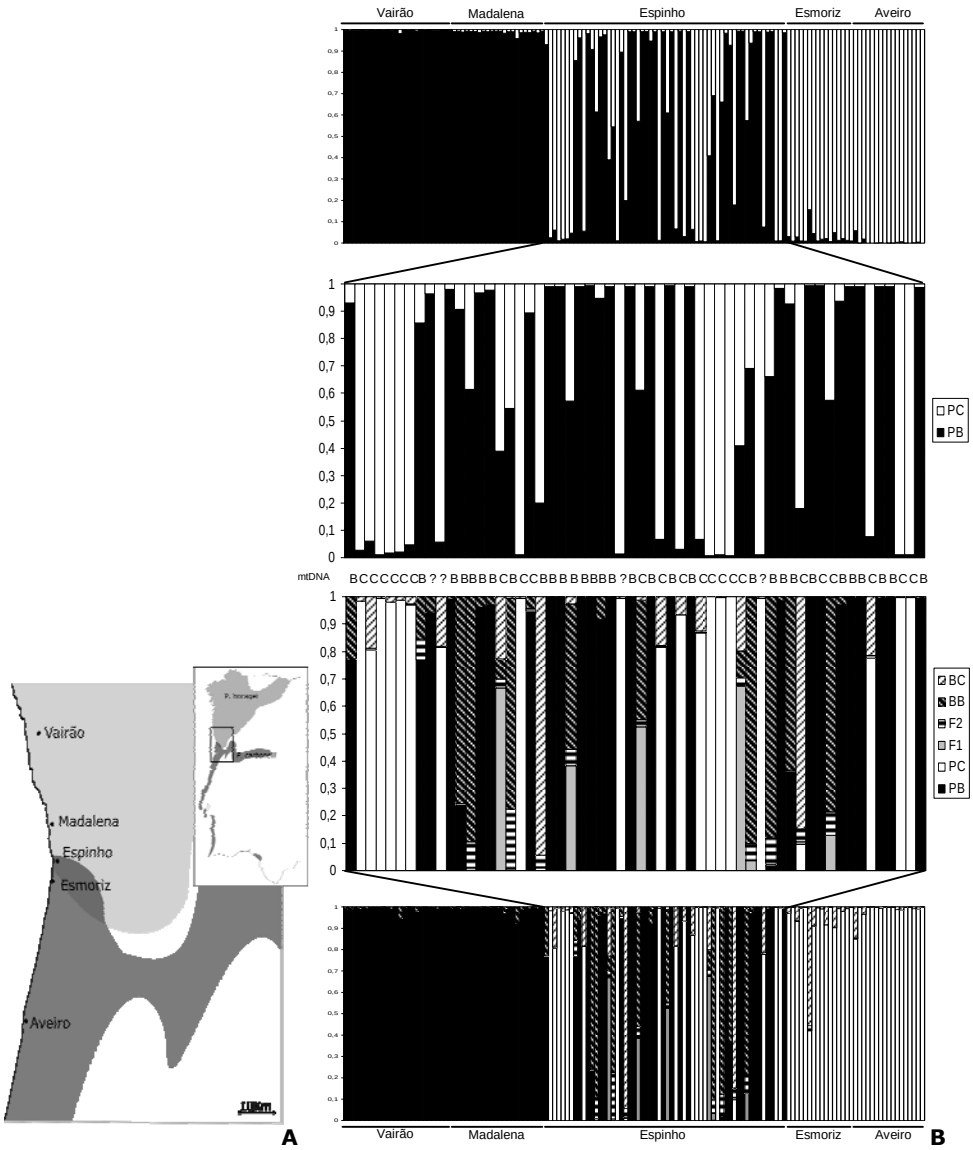

Figure 4.7. Genetic characterization of the contact zone between *P. bocagei* and *P. carbonelli.* **A**. Sampled localities. **B**. Model-based multilocus genotype analyses and mitochondrial DNA results. Upper graphs correspond to the proportion of the genome of each individual originating in each of the two species, estimated using STRUCTURE. Each individual is represented by a vertical bar divided in two segments, which length is proportional to the estimated proportion of the genome originating in *P. bocagei* (PB) or *P. carbonelli* (PC). On the bottom, results obtained using NEWHYBRIDS are shown. Again, each individual corresponds to a bar divided in six portions, each proportional to the estimated posterior probabilities of being either a pure *P. bocagei* (PB), pure *P. carbonelli* (PC), F1, F2, and backcross of a F1 hybrid with a pure *P. bocagei* (BB) or with a pure *P. carbonelli* (BC). The graphs shown in the centre of the figure highlight the results obtained in the contact zone. Letters B or C presented between these graphs correspond to the mitotype (*P. bocagei* and *P. carbonelli,* respectively) detected in the individuals. ? means that no information is available.

To investigate whether these individuals could be clearly assigned to a hybrid category (e.g. first or second generation hybrids or backcrosses to one of the parental species), we applied another Bayesian multilocus clustering approach (software NEWHYBRIDS, Anderson and Thompson 2002). These analyses showed that most individuals presenting signs of admixture probably result from multiple generations of admixture and cannot be assigned to a specific category; a few individuals, however, could be identified as backcrosses to one of the species. Only three individuals presented a probability higher than 50% of being an F1 hybrid. Taken together, these results strongly suggest that hybrids between both species are not only viable but also fertile.

4.1.3.2. Bimodality within the hybrid zone and its implications

At a first glance, the high frequency of individuals presenting signs of admixture could suggest that *P. bocagei* and *P. carbonelli* are simply not reproductively isolated. However, detailed analyses of STRUCTURE results suggest that more individuals than expected under a scenario of free admixture were found to be genetically "pure". These two species clearly do not form a hybrid swarm, as shown to occur in many hybrid zones. Moreover, even in individuals that were found to be admixed, a large portion of the genome is usually assigned to one of the species, with very few individuals presenting intermediate genotypes between the two species (Figure 4.8A). Both these features result in strong Hardy-Weinberg and linkage disequilibria within the contact zone. Taking these results together, it appears that the contact between *P. bocagei* and *P. carbonelli* entirely fits the description of a bimodal hybridzone (Jiggins and Mallet 2000). Although this situation could arise as the result of extensive migration of pure individuals into the hybrid zone from neighbouring populations (Barton and Hewitt 1985), it is therefore likely that there are strong reproductive barriers impeding introgression between the two species.

The two species occupy the same habitat, live in close contact and reproduce during the same period of the year. Therefore, there are no obvious ecological and temporal barriers preventing gene flow among them. There is, however, evidence from behavioural studies suggesting that males from both species react differently to chemical stimuli from conspecific and heterospecific females. This implies that there are species-specific recognition cues that could be acting as barriers to interspecific gene flow, suggesting that there is, at least to a certain extent, assortative mating between the two species. Could these and probably other behavioural interactions be responsible for the observed bimodal pattern? From our present knowledge on this subject, it remains unclear; in most bimodal hybrid zones in which this link has been studied, prezygotic

isolation mechanisms are certainly important (Coyne and Orr 1997, Rieseberg *et al.* 1998, Mallet *et al.* 1998, Howard *et al.* 1998, Price and Bouvier 2002). However, the first hybrids are often the most difficult to produce; once the first prezygotic barriers are overcome, behavioural differences tend to be attenuated and backcrossing is much more straightforward (Mallet 2005).

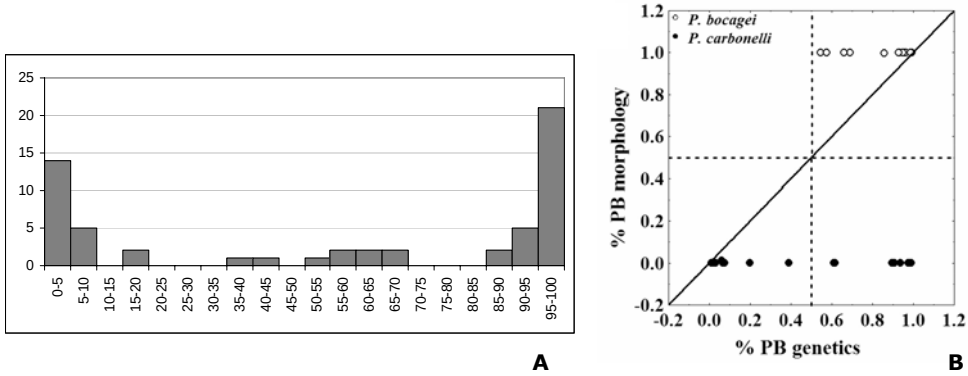

Figure 4.8. Genetic and morphological bimodality in the contact zone. **A.** Division of the individuals analysed in the contact zone according to their estimated proportion of *P. bocagei* ancestry obtained using STRUCTURE. **B.** Correspondence between morphological and genetic assignments of individuals. The X axis represents the proportion of the genome of individuals that was attributed to *P. bocagei* according to STRUCTURE. The Y axis represents the probabilities of assignment to *P. bocagei* by discriminant analysis using dorsal head shape variables provided be geometric morphometrics. The continuous line represents a slope of one for the relation between the two probabilities and the dotted lines indicate assignment probabilities of 0.5. Blank and filled circles indicate the species to which individuals were assigned by visual inspection.

Although speculative, this could suggest that the bimodal nature of this hybrid zone is also enforced by post-zygotic isolating mechanisms. To explore this possibility, we began by evaluating whether one of the best known generalizations related to post-zygotic isolation, Haldane's (1922) rule, was verified in the *P. bocagei – P. carbonelli* contact zone. This empirical rule states that when there is absence, rarity or sterility of first generation hybrids of one sex in a population of hybridizing taxa, that sex is usually the heterogametic. In *Podarcis*, as in most lacertids, females are the heterogametic sex; in the case Haldane's rule was verified, we would therefore expect to find no female F1 hybrids, for example, in the case of female hybrid inviability; in the case of female F1 hybrid sterility, we would expect that individuals inferred to be backcrosses with one of the species (i.e. resulting from the mating of a first generation hybrid with a non-admixed individual) would not carry the other species' mtDNA type, since it would imply that the mother was an F1 hybrid. By observing the correspondence between the hybrid category assigned to an individual and its mitochondrial DNA lineage, we may rule out the hypotheses of complete female inviability or sterility in this contact zone. However, the low

number of individuals that could be straightforwardly assigned to a hybrid class prevents us from testing with statistical robustness whether deviations to the expected proportions are suggestive of a partial loss of viability of reproductive success in female F1 hybrids. In order to achieve this, many more individuals from this contact zone would have to be surveyed.

A second step in the evaluation of possible post-zygotic barriers was the assessment of a possible decrease in hybrid fertility. Although fertility data (testis size and density of spermatozoa, in the case of males, and ovary size, number of follicles and oviductal eggs, in the case of females) were available for only a few of the individuals analysed (Carretero *et al.*, unpublished results), preventing strongly supported statistical inferences, there were no significant differences between "pure" and admixed individuals. This suggests that either natural selection is not acting against hybrid fertility or that it is doing so at later stages of the life cycle that were not studied (such as embryo development).

These two approaches failed to detect any form of decrease in hybrid viability or fertility that could indicate the effect of post-zygotic isolating mechanisms in the prevention of admixture between the species. However, clear evidence for the existence of such barriers was found when comparing morphological to genetic data. Using geometric morphometric methods, pure individuals from the two species are readily distinguishable, both in allopatry and in sympatry. A surprising result was, however, the finding that individuals that present an admixed genotype do not appear as morphologically intermediate (Figure 4.8B). That is, hybrid individuals are clearly morphologically assigned to one species or the other. This suggests that natural selection may be acting against embryo or juvenile intermediate phenotypes and that only those individuals that mostly resemble a "pure" specimen of one of the species are able to reach adulthood.

These considerations regarding the nature of both pre-zygotic and post-zygotic isolating mechanisms remain fairly speculative and need further assessment both by analysing more individuals and by performing laboratory controlled crossings. However, independently of the specific processes that are keeping these two species apart, it is clear that they exist; otherwise bimodality would not be observed. Bimodality within a hybrid zone implies that speciation is nearly complete (Jiggins and Mallet 2000), which means that *P. bocagei* and *P. carbonelli* are at the final stages of becoming "good" species.

4.1.3.3. The hybrid zone and geography

Because our sampling scheme was limited, we were not able to apply a formal analysis of clinal variation in allele frequencies across the transect. However, the fact that we did not detect evidence of major levels of introgression outside the

zone where the two morphotypes contact (even in populations as close as 5Km) suggests that the hybrid zone is rather narrow and that clines in allele frequencies are concordant. However, alleles that are private from the other species were indeed detected both to the North and the South of the contact zone, indicating that some alleles might have escaped the central barrier to gene flow, either because they are selectively favourable (Piálek and Barton 1997) or because of stochastic effects. Although this study might provide some initial insights, a complete assessment of the geographic behaviour of the introgession between the two species will require a much denser sampling scheme.

4.1.4. Are Iberian and North African forms of *Podarcis* true species? Implications for the study of multiple molecular markers in the analysis of species complexes.

Although a taxonomic reevaluation of Iberian and North African wall lizards was not regarded as a priority of this thesis, it is inevitable to analyse the potential of the different lineages for representing distinct species. This poses a significant problem right from the start, since one of the most debated questions in evolutionary biology, often with philosophical contours, concerns precisely the definition of species. For example, controversy stems from whether a species is as much of a real and natural entity as it is a human construct (Hey 2001): some opinions are sceptical regarding the biological validity of species in general, considering that this is yet another artificial division of biological diversity such are genera and families (e.g. Hendry *et al.* 2000). Most evolutionary biologists, however, agree that, although delineating species is a difficult task, "the living world is comprised of more or less distinct entities which we call species" (Mayr 1957).

Because completely diagnostic criteria for identifying species limits do not exist, a multitude of species concepts have been suggested, with variable scopes and applicabilities, but all with debatable particularities (see e.g. Howard and Berlocher 1998, Hey 2001, Mallet 2006, for reviews of the various concepts). In recent years, a tendency to rely solely on phylogenetic criteria for the definition of species has arisen, mostly drawing on the easiness of obtaining species-level phylogenies using mitochondrial DNA data. The application of such a "Phylogenetic Species Concept" (in which a species is defined as "(…) an irreducible (basal) cluster of organisms, diagnosably distinct from other such clusters, and within which there is a parental pattern of ancestry and descent"; Cracraft 1989) based on mitochondrial DNA data has led to attempts, largely unfruitful, to correlate genetic divergence at particular genes to specific status as means of defining a practical boundary for the definition of species (Johns

and Avise 1998, Avise and Walker 1999). By looking only at mitochondrial DNA variation, *Podarcis* morphotypes clearly seem to deserve species status, since they all form highly divergent and identifiable mtDNA clades, with average pairwise divergences that clearly fall within values traditionally observed between fully-recognised species (figure 4.9). However, if one tries to apply this criterion to nuclear gene genealogies, it is clear that *Podarcis* wall lizards will not be regarded as different species because they do not form monophyletic units (and hence fail to be recognised species according to the genealogical species concept (Baum and Shaw 1995)). Again, it is now well understood that differentiated species may share genetic variation for a long time; according to a strict genealogical species concept, for example, humans and chimpanzees would be conspecific. In the search for a more applicable criterion for diagnosing species based on genetic variation, Mallet (1995) suggested a "genotypic species cluster definition", according to which species are "(…) distinguishable groups of individuals that have few or no intermediates when in contact." According to this genotypic view, we would lean towards considering most of the forms as distinct species based on observations of multilocus genotype distinctiveness reported in articles III and V.

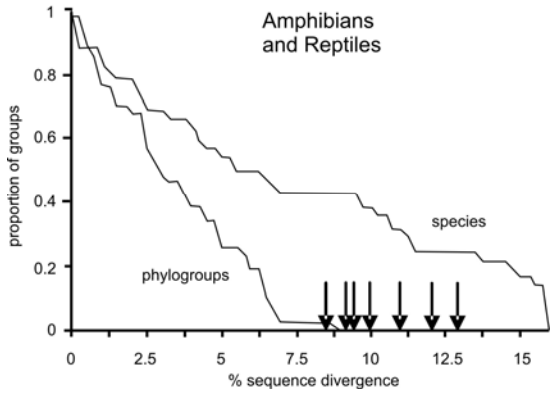

Figure 4.9. Average sequence divergence at mitochondrial gene cytochrome *b* among *Podarcis* sister taxa (represented by arrows), compared to the proportions of reptile and amphibian sister species pairs and phylogroup pairs exceeding different levels of sequence divergence based on a compilation of published data (adapted from Hendry *et al.* 2000). This comparison suggests that the levels of divergence found among *Podarcis* sister lineages are amongst the highest found between herpetofauna sister species.

These definitions of species tend to focus on distinct evolutionary processes in the species history; other species concepts rely on present mechanisms keeping the species apart. Under the latter class we find the most consensual and popular of species concepts: the Biological Species Concept (BSC), first formulated by Mayr (1942). According to this view, "species are groups of interbreeding natural populations that are reproductively isolated from other such groups". It is easily understood that a strict application of this criteria immediately excludes the validity of distinct *Podarcis* forms as species, since we have shown that several forms are able to hybridize. Given their relative

permeability to gene flow and our own experience from the analyses of a contact zone, it seems moreover highly likely that the ability of most forms to exchange genes in areas of overlap may be a general feature of these lizards; in fact, a probably underestimated ~40% of *Podarcis* fully-recognised species hybridize with at least one congener, which makes *Podarcis* one of most hybridization-prone vertebrate genera to be described so far (see for comparison Mallet 2005 and references therein). The acknowledgement that, as happens in *Podarcis*, fully-recognised species often have fuzzy borders (see e.g. Mallet 2005) has made the BSC evolve to accommodate such cases of limited gene flow. In the case of Iberian and North African *Podarcis*, it seems that, with some exceptions, phenomena of introgression do not have a major impact on the species overall genetic differentiation and that the various hybridizing forms are not sentenced to merge in the future. This is suggested, on one hand, by a high degree of concordance between partitions defined on morphology, mitochondrial DNA and nuclear genes; on the other hand, the maintenance of species borders despite ongoing gene flow is clearly documented by the study of the contact zone between *P. bocagei* and *P. carbonelli*. It therefore seems that at least some forms within the Iberian and North African *Podarcis* species complex could deserve specific status in the light of the modern, more flexible version of the BSC.

Just from this brief review of the applicability of different species concepts to the objects of our study, it is clear that this is a rather unfruitful exercise; while at the light of some concepts *Podarcis* forms may have the potential to be described as distinct species, they constitute violations to others. A consensual view could therefore be that the various lineages within *Podarcis* are emerging species, which are on the way of becoming fully distinct, illustrating the progressive – rather than sudden – nature of the process of speciation, in similarity to well known case-studies such as heliconiine butterflies (Mallet *et al.* 2007).

On the other hand, our data also seem to indicate that this may not be true for all mitochondrial DNA lineages. It is clear from our work that at least two species pairs might be less differentiated than the rest. One such case is *P. hispanica* Tunisian type and *P. hispanica* Jebel Sirwah. All markers suggest that these forms are closely-related. Unfortunately, because these two forms probably do not overlap geographically, it will be difficult in the future to properly evaluate the degree of distinctiveness that these species have achieved. Another case is that between *P. hispanica* sensu stricto and *P. hispanica* type 3. Unlike the former case, these two taxa are not sister at the mtDNA level; it is therefore unlikely that the lack of differentiation between these forms is due to a recent divergence. It seems plausible that these two mtDNA lineages have been exchanging genes throughout their evolution, leading

to a homogenization of their nuclear background. However, because genetic data is still controversial with respect to the nature of the introgression between these and yet other forms, we recommend that more studies directed at the study of the dynamics of gene flow between these two forms are performed before any taxonomic reevaluations are carried out.

Although our data point to a generally high degree of reproductive isolation, it should also be noted that there could be a wide variation in the extent of gene flow across different contact zones. This possibility has not been correctly evaluated in this study because most of the descriptions of genetic variability herein performed were not focused on directly quantifying levels of gene flow among forms. A future characterization of the evolutionary dynamics of other contact zones will certainly help in the clarification of whether species limits are successfully maintained in areas of overlap.

4.2. Phylogeographic structure within species of wall lizards: processes of population subdivision, historical demography and post-glacial expansion *in P. bocagei, P. carbonelli* and *P. vaucheri*.

A second major goal of this study was to describe phylogeographic patterns within selected species of wall lizards. Recent studies on both Iberian and North African organisms have disclosed complex patterns of genetic variation which, in a large number of cases, can be correlated to the effects of the Pleistocene climatic oscillations. Since ca. 2.3 – 1.8 million years ago, the Earth's climate has been dominated by major glacial periods interrupted by relatively short interglacial stages (reviewed in Hewitt 1996), which produced dramatic changes in species distributions over short periods of time. Although the effectiveness of such changes remain controversial for non-temperate taxa (Willis and Whitakker 2000), in well-studied regions such as Europe and North America these consequences have been well documented, both by studies based on the pollen and fossil records, which directly assess changes in species distributions, and by descriptions of genetic variability, which provide a more indirect way of detecting such changes. Although responses to climatic oscillations are largely species-specific, a set of well-established phylogeographic generalizations for temperate taxa can be recovered from different studies (Hewitt 1996, 1999, 2000, 2004; Taberlet *et al.* 1998): 1) during cold phases, species distributions became restricted to small areas where climatic conditions were less severe and favourable habitat was maintained, generally in southern regions; 2) after climate amelioration, species expanded from these southern refugia and

recolonized areas that were previously unsuitable. Both these historical patterns left their signatures on organismal genetic variation. On one hand, habitat fragmentation and isolation in distinct glacial refugia gave rise to profound spatially-structured genetic discontinuities; on the other hand, rapid recolonizations from glacial refugia left signatures of post-glacial demographic expansion coupled with a progressive loss of genetic diversity from refugia to newly colonized territories (usually from south to north), and, in a number of cases, to the formation of secondary contact zones between lineages isolated in distinct refugia.

In Europe, these patterns were first described at a continental scale, with southern peninsulas functioning as refugial areas (reviewed in Hewitt 1996, Taberlet *et al.* 1998). However, in recent years several studies have documented that at smaller scales, within Peninsulas, the same patterns can be found. This is particularly true for the well-studied Iberian Peninsula (reviewed in Gómez and Lunt 2007). At present, a wide array of species has been studied in detail, at least from a mitochondrial DNA perspective, which provides a well-established comparative framework for future phylogeographic analyses. It was therefore a preliminary goal of this study to assess if the phylogeographic structure of Iberian and North African *Podarcis* would fit the previously described scenarios for co-distributed taxa.

4.2.1. Contrasting effects of the Ice Ages across a latitudinal gradient

Highlighting the features that are common across species is important because it can help identifying the influence of common geographic or temporal phenomena. However, analysing the factors that cause species responses to be different is also essential. These differences can be related, for example, to different habitat requirements, dispersal abilities and life history traits (e.g. Campbell *et al.* 2006), or to different latitudinal distributions, influencing the magnitude of the effects of the Ice Ages. In fact, it is highly unexpected to find the exact same responses across different geographic regions. For example, Lessa *et al.* (2003) compared patterns of historical demography in mammalian faunas from North America and Amazonia and showed that the signatures of demographic expansion resulting from post-glacial recolonizations from reduced source populations were only present in the first region because Amazonia suffered only mild alterations during the Quaternary glaciations, allowing inhabiting organisms to maintain long-term stable effective population sizes.

In this work, we tried to investigate whether such different latitude-dependent effects could be detected at a smaller scale (i.e. the Iberian Peninsula and North Africa). Even within a relatively small region, we expect that northern

areas experienced more dramatic changes in habitat availability, with detectable effects on the genetic diversity of inhabiting species. More specifically, we aimed to test two main predictions: 1) in northern latitudes, climatic conditions during cold stages were more severe, resulting in fewer and smaller patches of favourable habitat for the survival of species than in southern regions, where the effects were milder; 2) species with distribution in areas located in northern areas were probably more confined during cold stages and expanded more rapidly in response to the climate amelioration, whereas species distributed in the south maintained more stable effective population sizes. The expected signatures left by the Ice Ages on species genetic diversity are therefore different according to the latitude: in species distributed in northern areas we expect overall lower genetic variability and less complex phylogeographic subdivision due to higher levels of lineage extinction, whereas southern species should bear higher levels of diversity; on the other hand, strong signatures of demographic expansion should be expected in species distributed to the north, and not in southern-distributed species. To test these hypotheses, we focused this work on three species with parapatric distributions which replace each other along a latitudinal gradient: *Podarcis bocagei*, distributed in Northwestern Iberian Peninsula, *P. carbonelli*, found in central and southwestern Iberia, and *P. vaucheri*, which is distributed both in southern Iberian Peninsula and in North Africa. Because these three species are closely-related and very similar from an ecological point of view, we expect that the differences we observe are indeed related to their different distribution and not to other factors susceptible of altering the species' response to climatic fluctuations.

We studied genetic variation at a fragment of mitochondrial DNA gene NADH dehydrogenase subunit 4 (ND4) in a representative sample from the distribution of the three species, and found remarkably different patterns of genetic diversity. *P. bocagei* presented a widespread and abundant haplotype with 18 rarer haplotypes only differing from this by one or two mutations (figure 4.10A), yielding a very low nucleotide diversity (π=0.00168; table 4.2). Moreover, no major geographic substructure could be detected within this species' range. Based on coalescent simulations, we were able to pinpoint an interval for this species' mtDNA coalescence between 71 and 113 thousand years (kyr); most of the mutations observed in this species were inferred to have occurred in the past 20 kyr, suggesting that all the variability found arose very recently, most likely after the climate amelioration (figure 4.11). This contrasts with the high degree of genetic variability found in *P. vaucheri*. Not only did we find a high level of differentiation between Iberian and North African specimens of these species, as had already been reported in previous studies (e.g. Harris *et al.* 2002b, Busack *et al.* 2005, article I), but also very high amounts of genetic and geographic subdivision were detected when focusing only in Morocco. For

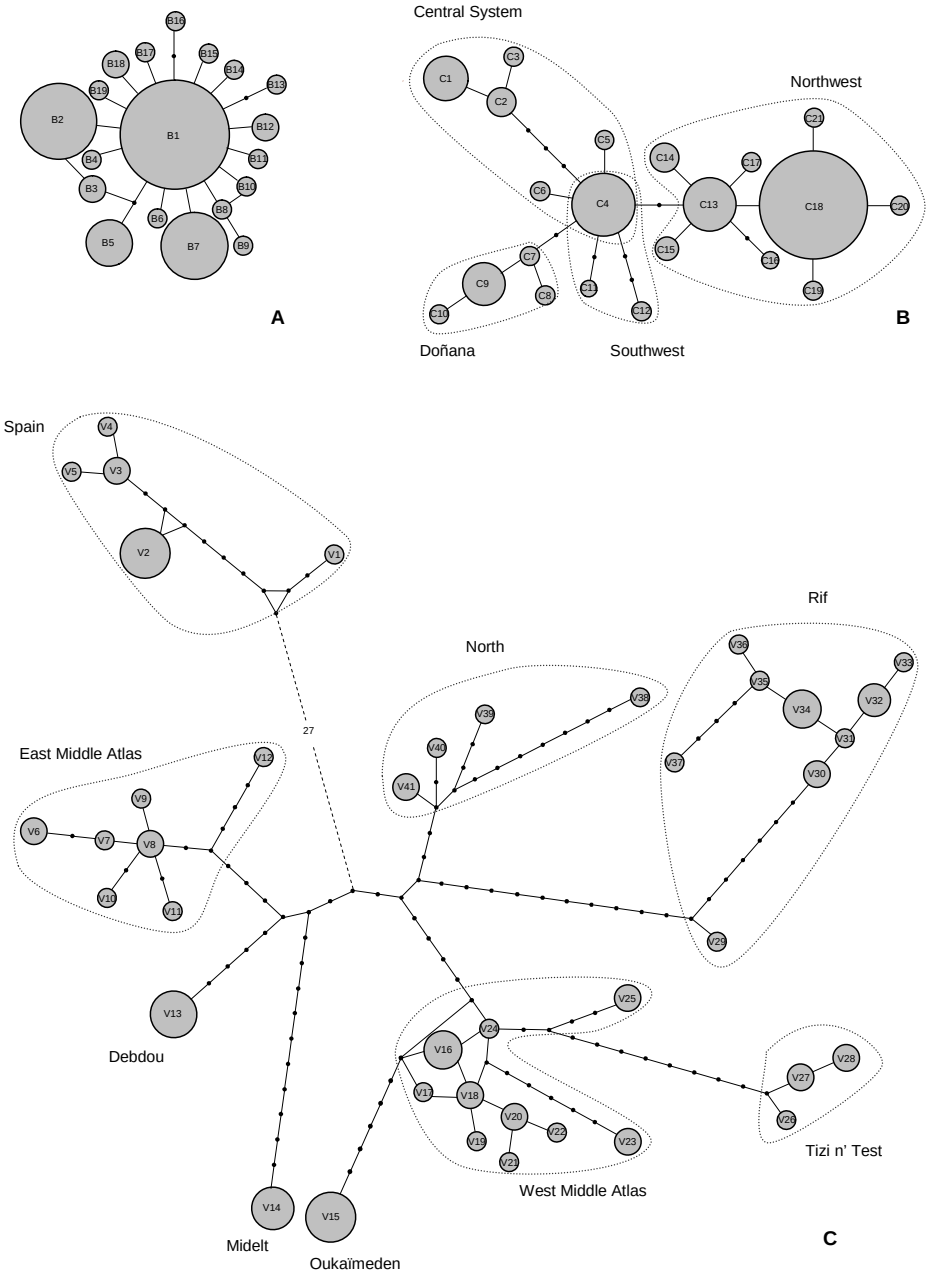

Figure 4.10. Median-Joining networks representing relationships between mitochondrial DNA haplotypes detected in three studied species. **A.** *Podarcis bocagei*. **B.** *Podarcis carbonelli*. **C.** *Podarcis vaucheri*. The area of each circle is proportional to the frequency of the haplotype it represents. Numbers on dashed branches refer to number of mutations between the connected nodes. Dotted lines group haplotypes found in the same geographic region.

Table 4.2. Summary statistics, measures of population differentiation and population growth for three species of wall lizards and respective data subsets, based on mitochondrial DNA variation.

Species	Partition	N	Polymorphism and population differentiation								Population growth		
							2-level AMOVA	3-level AMOVA					
			H	*Hd*	*n*	θ	Φ_{ST}	Φ_{CT}	Φ_{ST}	Φ_{SC}	*Fs*	R_2	*g* ± SD
P. bocagei	-	82	19	0.789	0.00168	0.04468	0.663**	-	-	-	-15.092**	0.0347*	5345.75 ± 336.94‡‡
P. carbonelli	-	84	21	0.823	0.00456	0.01343	0.819**	0.718**	0.865**	0.522**	-6.179	0.0593	516.07 ± 178.13
	Northwest	50	9	0.569	0.00111	0.03294	0.502**	-	-	-	-4.902**	0.0547*	9509.44 ± 1047.30‡‡
	Southwest	11	3	0.345	0.00115	0.04116	1.000*	-	-	-	0.401	0.1975	10000.00 ± 2777.23‡
	Central System	15	6	0.762	0.00327	0.00337	0.528*	-	-	-	-0.114	0.1469	148.52 ± 244.63
	Doñana	8	4	0.643	0.00117	0.06429	-	-	-	-	-1.387	0.1610	9009.43 ± 2109.01‡
P. vaucheri	-	81	41	0.971	0.03024	0.05570	0.908**	0.617**	0.949**	0.868**	-2.119	0.1140	32.15 ± 16.74
	Spain	12	5	0.667	0.00384	0.00436	0.747**	-	-	-	0.820	0.1917	-40.43 ± 115.38
	Morocco	69	36	0.970	0.02312	0.04949	0.869**	0.836**	0.878**	0.258**	-3.290	0.1107	58.14 ± 24.59
	North	5	4	0.900	0.00812	0.02407	-0.868	-	-	-	0.883	0.2000	301.94 ± 121.63
	Rif	15	9	0.905	0.00438	0.01108	0.223	-	-	-	-2.276	0.1213	215.23 ± 127.06
	West Middle Atlas	17	10	0.926	0.00532	0.00951	0.047	-	-	-	-2.176	0.1209	159.26 ± 120.69
	Oukaïmeden	7	1	0	0	0	-	-	-	-	-	-	-
	Tizi n' Test	5	3	0.800	0.00176	0.00532	-	-	-	-	0.061	0.2848	1659.56 ± 1129.44
	East Middle Atlas	9	7	0.944	0.00471	0.04013	-0.053	-	-	-	-1.960	0.1413	960.85 ± 228.61‡
	Debdou	6	1	0	0	0	-	-	-	-	-	-	-
	Midelt	5	1	0	0	0	-	-	-	-	-	-	-

N, number of individuals; *H*, number of haplotypes detected; *Hd*, haplotype diversity; *n*, nucleotide diversity; θ, population mutation parameter, estimated jointly with *g*, growth rate parameter using FLUCTUATE (Kuhner *et al*. 1998); Φ_{CT}, Φ_{ST}, Φ_{SC}, fixation indexes; *Fs*, Fu's (1997) Fs; R_2, Ramos-Onsins and Rozas' (2002) R_2; 2-level AMOVA were performed considering all the populations of the data set as a single group; 3-level AMOVA were performed dividing the data set into subsets. AMOVA analyses were carried out using uncorrected pairwise distances.**, significant at p<0.01; *, significant at p<0.05; ‡‡, *g* over 6 standard deviations (SD) above 0; ‡, *g* between 3 and 6 SD above 0.

example, nucleotide diversity within this region was one order of magnitude higher than that observed in *P. bocagei* (π=0.02312). Moreover, several highly divergent clades, coinciding with different mountain ranges or distinct areas within these, were inferred to exist. Levels of genetic variability within some of these groups were again higher than those inferred for the whole distribution of *P. bocagei*. The time to the most recent common ancestor inferred for the Moroccan phylogroup within this species was consequently found to be much older than that estimated for *P. bocagei* (1.010 – 1.615 million years ago), suggesting that the observable differentiation within this species started in the early stages of the Pleistocene. Regarding *P. carbonelli*, levels of genetic variation were again low, although still higher than those observed in *P. bocagei* (π=0.00456). We also documented a high degree of association between geography and genetic differentiation, although one haplotype was shared between distant localities. TMRCA inferred for this species was intermediate between the previous two species (313 – 500 kyr).

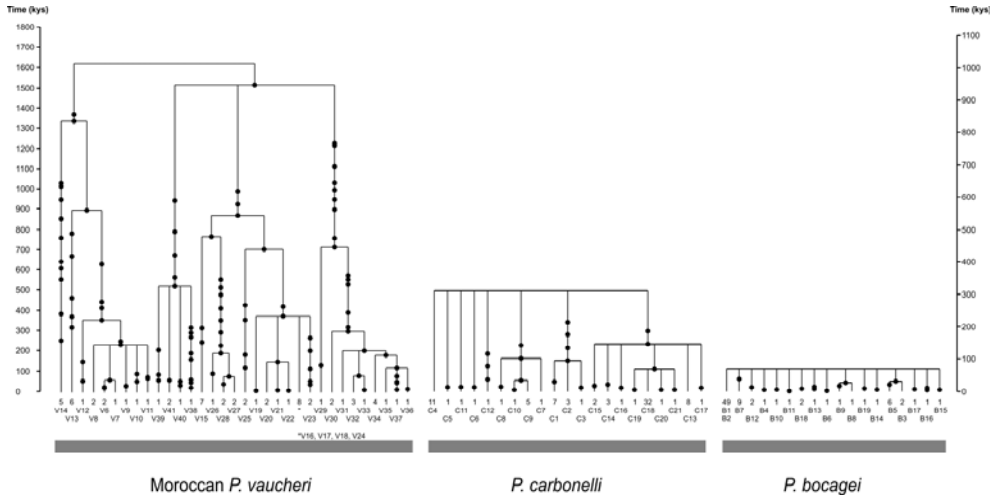

Figure 4.11. Gene trees for *P. bocagei, P. carbonelli* and the Moroccan populations of *P. vaucheri*, obtained assuming constant population size in *P. carbonelli* and *P. vaucheri* and an exponential growth model in *P. bocagei*. Mutations along the genealogy are represented by black circles. The time scale is the same in the three genealogies. The scale represented on the left is based on a slow rate of evolution (μ=2.175x10^{-8}) and on the right on a fast rate of evolution (μ=3.475x10^{-8}).

These results provide unquestionable evidence regarding the validity of our first prediction. *P. bocagei,* more northerly distributed, suffered the effects of more severe habitat changes than the other studied species; as a result, it nowadays presents depleted levels of diversity resulting from the extinction of historical lineages, and no signs of geographic discontinuity in the distribution of genetic variability, which are concordant with the survival of a single refugial

population that remained confined until the recent climate warming. Conversely, the distribution of *P. vaucheri*, our southernmost case-study, was probably repeatedly fragmented during successive glacial stages. However, unlike in *P. bocagei*, many of these geographically separated refugial populations were able to persist until the present, which is a probable consequence of the fact that more areas of suitable habitat remained available during climatic changes. Because the different clades within this species generally exhibit relatively high levels of variability, refugial populations probably maintained high effective population sizes even throughout repeated glacial periods. *P. carbonelli* presents intermediate values of subdivision and overall variability, which are concordant with its intermediate distribution in the context of the three species analysed: on one hand, this species certainly suffered from a dramatic loss of historical lineages because all of its variability was reset to zero by a strong bottleneck occurred around to ~500 Kya; on the other hand, it persisted through the last glacial stages in multiple areas of its distribution, since a certain degree of association between geography and genetic diversity was observed.

In order to test our second prediction, we investigated whether these three species exhibited the signatures of a rapid demographic expansion concomitant with a population recovery after the Ice Ages. We performed several tests designed to describe demographic scenarios. We applied these not only on full data sets of each species but also on geographic partitions within *P. carbonelli* and *P. vaucheri* in order to minimize the effects of applying methods that assume panmixia in species that are obviously partitioned. Our results suggest that among the three species as a whole, only *P. bocagei* shows the genetic imprint of a rapid demographic expansion, as shown by largely negative Fu's (1997) *Fs* values and large positive values of the growth parameter *g* as estimated using the software FLUCTUATE (table 4.2). Among the geographic partitions of *P. carbonelli* and *P. vaucheri*, only the northwestern clade within *P. carbonelli* presents these same signatures, suggesting that most phylogroups within these species were able to maintain stable effective population sizes throughout the Pleistocene.

Because both of our main predictions were strongly verified, we may therefore conclude that different effects of the Quaternary glaciations across different latitudes can be detected even at a rather small scale. Obviously, the ability to detect such differential effects depends on the species dispersal capacity. Terrestrial invertebrates, reptiles and amphibians thus constitute good models to perform such comparisons. However, few studies aiming at detecting these differential effects have been performed, particularly regarding the Western Mediterranean. For example, a study involving the freshwater turtle *Mauremys leprosa* showed the same pattern inferred in *Podarcis* of a rapid population expansion in northern populations whereas a southern clade

restricted to North Africa presents signs of effective population size stability (Fritz *et al.* 2006). Several other species groups could therefore be suitable candidates for replicating this gradient analysis and investigating whether similar trends could be reported for other species. The concordance, or lack thereof, of scenarios for distinct species could provide valuable clues for the interpretation of biogeographic patterns across the Western Mediterranean region.

4.2.2. Concordance of mtDNA phylogeographic scenarios in the context of the western Mediterranean Quaternary biogeography

A concordant feature across several phylogeographic studies performed on the Iberian Peninsula is the observation of strong genetic discontinuities related to isolation in distinct glacial refugia during the Pleistocene. This feature has been observed in nearly every species studied, from small plants to trees, from invertebrates to every group of vertebrates (reviewed in Gómez and Lunt 2007). Moreover, a set of species with strong biogeographic affinities with both *P. bocagei* and, to a lesser extent, with *P. carbonelli* (*Lacerta schreiberi, Rana iberica, Discoglossus galganoi, Chioglossa lusitanica* and *Lissotriton boscai*) also shows signs of surviving in distinct refugia throughout the Ice Ages (Paulo *et al.* 2001, 2002b; Alexandrino *et al.* 2000, 2002; Martínez-Solano 2004; Martínez-Solano *et al.* 2006; J. Teixeira, unpublished results). The lack of similar variation observed in *Podarcis bocagei*, although expected from its northerly distribution, is therefore rather unique. It has been suggested that the river Douro has probably played an important role in restricting this species from moving southwards (Sá-Sousa 2001b). This river has also been demonstrated to coincide with a marked discontinuity in genetic variability at least in some of the above-mentioned species, such as *C. lusitanica* (Alexandrino *et al.* 2000, 2002; Sequeira 2006), *L. schreiberi* (Godinho 2003) and probably also *D. galganoi* (Martínez-Solano 2004). We may therefore speculate that the Douro river might have prevented *P. bocagei* from moving southwards tracking the regression of suitable habitat, which caused the species to be confined to a single and very restricted area north of this river. However, the biogeography of *Podarcis* lizards might not be simply interpretable in the light of climatic change as in other species, which are not replaced by a closely-related form along the latitudinal gradient (an exception could be, for example, cryptic species within *L. boscai*); in the case of *P. bocagei*, for example, competition with the ecologically similar *P. carbonelli* might itself reduce the species potential for dispersal into southern territories tracking the changing distribution of suitable habitat. These explanations remain, at present, speculative; more effort regarding modelling the distributions of these species to understand to which extent they are

influenced by habitat conditions (as opposed to the presence of other species) would be desirable to properly address these issues.

The detection of a strong signature of population expansion in *P. bocagei* is, however, completely concordant with previously described scenarios for co-distributed species, since similar signatures have been found both in *L. schreiberi* (Paulo *et al.* 2001, 2002b) and *C. lusitanica* (Alexandrino *et al.* 2000, 2002), particularly for phylogroups which show geographic distributions similar to that of *P. bocagei.* This therefore seems to be a general trend for species that became severely confined during the glaciations.

The evolutionary scenario inferred for *Podarcis carbonelli* is much more difficult to analyse in a comparative phylogeography framework because this species has a unique distribution in the context of Iberian herpetofauna. Mitochondrial DNA variation within this species differs from descriptions of strong fragmentation and persistence in separate refugia since the first stages of the Pleistocene as has been described in other species (namely *L. schreiberi*, possibly one of the species with more similarity in the distribution); differentiation is, as discussed above, rather shallow at a temporal level. There is a remarkable degree of association between different haplotypes and distinct geographic regions: a set of haplotypes is private of the Iberian Central System; another clade is found only in the geographic isolate in southern Spain; and a third group of haplotypes was detected only in the Northwestern region of the distribution. However, the haplotype which is placed as ancestral in the network is found in distant areas (both in the southern Portuguese coast and in the Central System), probably reflecting incomplete lineage sorting. Although a close parallel cannot be traced with the evolutionary history of co-distributed taxa, some concordant features were found between the scenario inferred for *P. carbonelli* and other species, which confirm the importance of some areas as glacial refugia for distinct species. For example, the mountains of the Iberian Central system seem to have played an important role in the diversification of both *P. carbonelli* and also *L. schreiberi* (Paulo *et al.* 2001, 2002b), *L. boscai* (Martínez-Solano *et al.* 2006) and *R. iberica* (J. Teixeira, unpublished), although at different time scales for different organisms. Our results also suggest a rough temporal concordance (around 500 kya) regarding differentiation within *P. carbonelli* and within the coastal clade of *L. schreiberi* (Paulo *et al.* 2001, 2002b); the same processes may also have ultimately led to the formation of southern geographical isolates in both species.

The Quaternary biogeography of North Africa is certainly not as well understood as that of the Iberian Peninsula. On one hand, there are few reports describing how habitats changed in this region across the Ice Ages (but see Jolly *et al.* 1998). On the other hand, most studies addressing genetic variation within North African species aim at solving higher level taxonomy, which seems to be

rather complex in this region (Carranza *et al.* 2004b, Harris *et al.* 2004a, 2004b). Only a handful of studies address intraspecific phylogeography in North African organisms. Nevertheless, these studies clearly describe the same signatures of isolation in distinct glacial refugia as portrayed for Iberian taxa (Brown *et al.* 2002, Guiller *et al.* 2001, Carranza and Arnold 2004, Cosson *et al.* 2005, Guillaumet *et al.* 2006). Interestingly, none of the above-referred species seems to present such a complex distribution of genetic variability as the one we observe in *P. vaucheri*. We identified several major regions that may have provided refugium for this species across the Ice Ages, but even within these refugial areas genetic variation seems to be rather fragmented. There is, furthermore, a strong association of genetic variability to different mountain chains or distinct areas within them, although our sampling scheme still does not allow supported observations regarding the distribution of the major phylogroups. Although the correlation of *P. vaucheri*'s distribution with environmental variables has not been studied in detail, a simple observation of this species distribution maps suggests that it might be closely dependent on areas of high humidity (see e.g. Bons and Geniez 1996). We may therefore speculate that during cold stages, which seem to be associated to the expansion of arid habitats (Jolly *et al.* 1998), the species was able to survive in warm and moist valleys within mountain chains. It is noteworthy to analyse the role of the river Moulouya in the diversification of this species, since this river has been suggested as a major biogeographic barrier for North African taxa (Álvarez *et al.* 2000, Paulo 2001, Zangari *et al.* 2006). In article VI, we analysed a single population of *P. vaucheri* located east to this river (Debdou). The level of differentiation found between this and the other Mahgrebian populations of this species is moderate, suggesting that this river does not stand out in importance from other putative barriers to the dispersal of *P. vaucheri*. This trend is also confirmed by the analyses of Algerian populations of this species (results not shown).

4.2.3. Intraspecific variation in *Podarcis bocagei* and *P. carbonelli*: insights from nuclear markers

As discussed several times throughout this thesis, it is dangerous to base all inferences regarding the evolutionary history of organisms on a single gene. According to mtDNA variation, when compared to other Iberian species, *P. bocagei* and *P. carbonelli* present relatively shallow differentiation because they only recently occupied their present area of distribution. There are however differences in the degree of this lack of differentiation: *P. carbonelli* has probably persisted in several areas of its distribution range across the last stages of the

glaciations because each geographical region has its own set of mildly differentiated mtDNA haplotypes; in the case of *P. bocagei*, however, the shallowness of differentiation is extreme, and we expect that any process of divergence between geographic regions arose after the species recolonized the northern areas of the Iberian Peninsula (that is, around 13000 years ago). In Iberian endemics where strong discontinuities are found at the mitochondrial level, accordingly high levels of divergence have been reported also at the nuclear level (e.g. in *Lacerta schreiberi* (Godinho 2003, Godinho *et al.* 2006a,b), *Chioglossa lusitanica* (Alexandrino *et al.* 2000) and *L. boscai* (J. Teixeira, unpublished). Therefore, a first goal of applying nuclear markers to the characterization of intraspecific variation within these two species was, on one hand, the confirmation of the overall uniformity in genetic variation, particularly in *P. bocagei*, and, on the other hand, the investigation of whether the geographic substructure observed at the mitochondrial level within *P. carbonelli* was confirmed by the study of nuclear polymorphisms. A second goal was to confirm demographic dynamics inferred for both species. Our study of mitochondrial DNA variation detected signatures of exponential growth in *P. bocagei* and in the northwestern clade of *P. carbonelli*. Although such signatures are also found in codistributed taxa and thus most likely refer to strictly demographic phenomena, single locus approaches do not clearly distinguish between the effects of demography and selection, as selective sweeps may leave exactly the same imprint; however, selective sweeps are locus specific, while demographic phenomena tend to affect the whole genome. Thus, the study of nuclear markers provides a way of testing the hypotheses of demographic (and concomitant geographic) expansions for both species.

Two different classes of markers were studied: allozymes and microsatellites. Allozyme loci surveyed for both species were the same 10 polymorphic markers that were used in the study of diversification within the whole clade of Iberian and North African *Podarcis*. These markers are characterized by a relatively low level of diversity at the intraspecific level and are thus unsuitable as the single means to study phenomena of very recent divergence. We therefore developed markers with higher mutation rates, susceptible of detecting differentiation even at shallower temporal time-frames – microsatellites (article VII). We studied genetic variation at six microsatellite loci in *P. bocagei* and four in *P. carbonelli*. As expected, these markers demonstrated much higher variability levels than those observed in allozyme loci. Both classes of markers are concordant in describing low levels of differentiation for both species (table 4.3). Nevertheless, differentiation levels, as measured by F_{ST} values, are higher for *P. carbonelli* than for *P. bocagei*, which may indicate that diversification within the former is more ancient than within the latter, as suggested also by mtDNA analyses. Interestingly, in both species, F_{ST} values are much lower when considering

nuclear markers than those suggested by mtDNA, which probably results from the higher effective population size that characterizes the nuclear genome. Another interesting comparison is that among both classes of nuclear markers surveyed. In general, differentiation depicted by microsatellite markers is apparently lower than that observed for allozymes, which is expected given the higher levels of diversity (Hedrick 1999) and homoplasy (Estoup *et al.* 2002, O' Reilly *et al.* 2004) that are usually associated to microsatellite variation.

Table 4.3. Overall genetic differentiation (F_{st}) in *P. bocagei* and *P. carbonelli*, using allozymes, microsatellites and mitochondrial DNA

		P. bocagei	***P. carbonelli***
Allozymes	*Got1*	0.0549	0.1120
	Gpi	0.0240	-
	Idh	0.0594	0.5313
	Ldh2	0.1322	0.2000
	Mpi	0.1653	0.1823
	PepA	0.0896	0.2001
	PepB	0.0813	0.1359
	PepD	0.0440	0.1034
	Pgd	0.2451	0.1174
	Pgm	0.0223	0.0368
	All	0.1068	0.1706
Microsatellites	*Pb11*	0.1649	0.1737
	Pb37	0.0533	-
	Pb47	0.0517	0.0983
	Pb50	0.0401	0.0793
	Pb66	0.0529	0.0771
	Pb73	0.0464	-
	All	0.0678	0.1089
Nuclear markers		0.0772	0.1364
Mitochondrial DNA			
nucleotide		0.7036	0.8348
haplotype		0.6039	0.6334

Mitochondral DNA data were obtained from article VI. Only data from the same populations analysed in this study were used for comparison. Overall F_{st} were computed using nucleotide distances among haplotypes ("nucleotide") and only considering haplotypes, independently of the nucleotide variation ("haplotype").

4.2.3.1. Evolutionary history of P. bocagei

While mitochondrial DNA variation strongly suggested a scenario of demographic growth for *P. bocagei*, the precise location of glacial refugia and putative recolonization routes were not known with certainty. An indication that the demographic expansion was simultaneous with a northwards geographic expansion, as expected if the species had been trapped in a refugium located in the south, is only suggested by slightly higher levels of variability in the south (but this could also result from a sampling bias) and the fixation in the northeastern region of the distribution of one haplotype present mostly in the southeast (B2 in the haplotype network presented in figure 4.10A), which also

provides hints that the colonization of this area was conducted independently from that of western Galicia.

Allozyme and microsatellite variation provides support for the hypothesis of demographic growth by clearly demonstrating a rapid geographic expansion. This inference is supported by different signatures left by this process on genetic variation:

a) Progressive loss of variability in a northwards fashion

When populations expand rapidly, dispersal in the leading edge causes successive founder events that imply a progressive loss in the number of alleles and heterozygosity (Ibrahim *et al.* 1996, Hewitt 1999). This pattern is rather evident in *Podarcis bocagei* from the analysis of the patterns of diversity in both allozyme and microsatellite loci (figure 4.12). Populations located in the north of the distribution, particularly those located in western Galicia, consistently show depleted levels of genetic diversity as opposed to those found in southern populations.

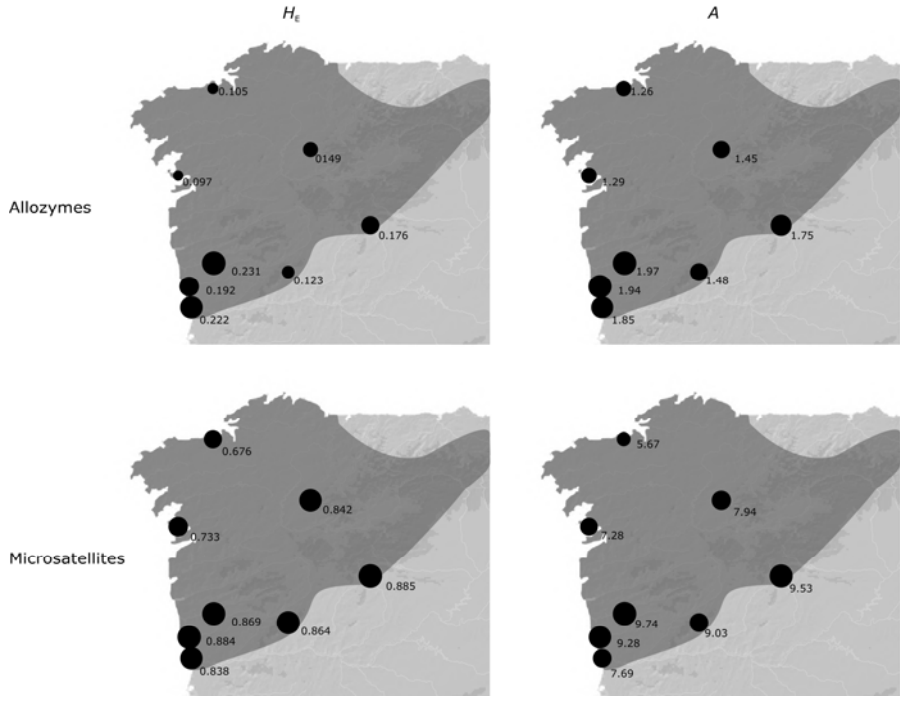

Figure 4.12. Mean diversity measures (expected heterozygosity (H_E) and allelic richness (A)) in populations of *P. bocagei*.

b) Differentiation increases towards the extremes of the range expansion

Differentiation in *P. bocagei* is, as stated above, rather shallow (F_{ST}=0.0772); however, northernmost populations are more differentiated from the rest, suggesting that they have been more affected by genetic drift as expected if they corresponded to recently colonized areas. Moreover, when clustering populations in a NJ tree, the populations of the southwestern part of the distribution are placed in a more central position, whereas populations from the north of the distribution, on one hand, and from the east, on the other hand, are placed in the extremes of the tree (figure 4.13A). A similar result is obtained when applying a model-based individual multilocus genotype clustering method, as implemented in the software STRUCTURE (Pritchard *et al.* 2000; figure 4.13B and C). This method was applied with the intent of detecting hidden population structure, since it works by finding clusters of individuals that minimize Hardy-Weinberg and linkage disequilibria. With K=3, for example, individuals from peripheral populations appear as mostly pure for one of the clusters (Madalena (Mad), corresponding to one cluster, Zimão (Zim) and Montesinho (Mon), corresponding to a second cluster, and the Galician populations corresponding to a third cluster), while the southwestern populations of Vairão (Vair) and Braga (Bra) emerge as rather admixed. Instead of corresponding to three clusters in the sense that they result from some degree of isolation, these three groups most likely correspond to the extremes of clines that progressively emerged from a range expansion with epicentre in the southwestern populations. With K=5, more geographically consistent groups emerge, corresponding to more recent events of differentiation, but the population of Braga still remains admixed for all the five clusters in approximately equal proportions, as if it corresponded to the original gene pool from which the range expansion was carried out.

Taken together, these results not only confirm the validity of the inference of demographic growth associated to a range expansion based on mtDNA variation, but also help in pinpointing with a notable degree of precision the populations of the southwestern area of the distribution as a putative glacial refugium that acted as the source for geographic expansion. Interestingly, we were able to detect two possible colonization routes: one through the west and another through the east, the latter inferred mostly by complementing the inferences derived from nuclear markers with information based on mtDNA variation, but also on the findings of admixture between the northern and eastern clusters (with K=3) in the population of Sarria (Sar). This situation therefore mirrors at a very narrow temporal and geographic scale the observations of admixture between distinct population groups and consequent increase in genetic variability resulting from the merging of independent waves of colonization, detected at a European scale for several species (e.g. Comps *et al.* 2001).

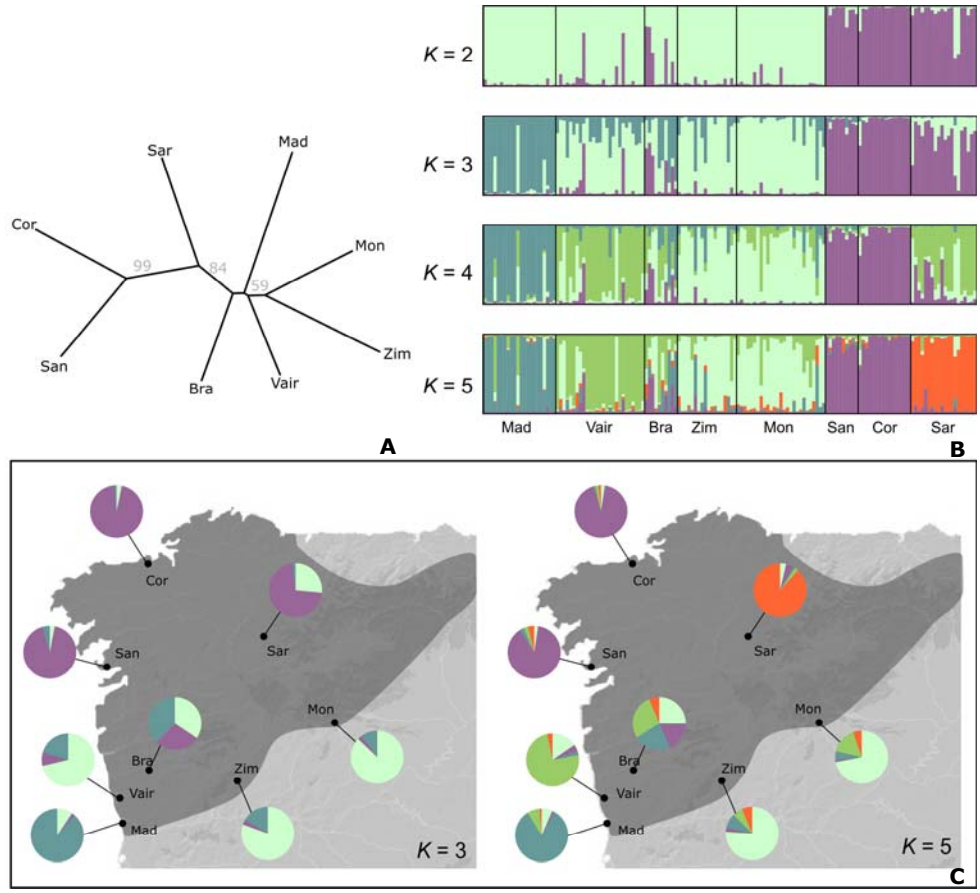

Figure 4.13. Differentiation within *P. bocagei* using nuclear markers. **A.** NJ tree depicting relationships between populations. Bootstrap values above 50% are shown. **B.** Inferred population structure (from from K=2 to K=5) using the model-based individual clustering approach implemented in software STRUCTURE. In these plots, each individual is represented by a column divided into K segments, the size of each corresponding to the individual's estimated membership fraction in each of the K clusters. **C.** Pie charts representing the mean proportion of membership for K=3 and K=5 (chosen as the best estimates of subdivision within this species applying the method described by Evanno *et al.* 2005) for each population.

The distribution of *P. bocagei* is situated mostly north of the river Douro, and it has been speculated that the colonization of territories to the south did not occur until very recently, through bridges built over this river by humans (Sá-Sousa 2001b). If true, this very recent colonization would imply very low levels of genetic diversity in populations located south of this river. Although we only sampled one population located in this area, our results do not conform to this suggestion. Estimated levels of diversity and differentiation for the population of Madalena are moderate in the context of the distribution of the species, implying a more ancient colonization of this region.

Globally, inferences about the evolutionary history of *P. bocagei* derived from mitochondrial DNA and nuclear markers are concordant, but each type of marker offers particular levels of information that are complementary, reinforcing the idea that the study of multiple markers for establishing phylogeographical scenarios is always beneficial.

4.2.3.2. Evolutionary history of P. carbonelli

A particular feature of *P. carbonelli*'s distribution is its uniqueness. No other species of the Iberian herpetofauna presents a similar distribution. Moreover, attempts to model this species distribution, at least for its Portuguese range, failed to detect habitat characteristics that could be significantly correlated to the species distribution (Sá-Sousa 2001b). Therefore, if for *P. bocagei* the similarity of its distribution with co-distributed taxa could be used for generating a priori hypotheses about its evolutionary history, for *P. carbonelli* few such hypotheses could be put forward. Descriptions of mitochondrial DNA variation suggested that this species persisted in diverse areas within Iberia for the past 300-500 Kyr. Moreover, because the distribution of this species seems to be somewhat fragmented, we were expecting a certain degree of concordance between mitochondrial DNA and nuclear markers, suggesting a scenario of isolation in distinct refugia.

Surprisingly, genetic variation at nuclear markers does not reveal the same patterns of isolation that were inferred for mtDNA. On one hand, genetic differentiation between the four population groups inferred to exist from mitochondrial variation, accounts only for 3.64% of the total genetic variation at allozymes and microsatellites. No other geographical clustering of populations produces values of differentiation as high as those obtained when considering each population separately, suggesting that there is not a clear geographic trend in the distribution of genetic variability in nuclear markers in *P. carbonelli*.

However, the neighbor-joining tree for populations of this species (figure 4.14A) displays a tendency to group geographically closer populations, with some notable differences from mitochondrial estimates of relationships: the populations of the central system appear as one clade; the same is observed for the populations of the northwestern area of the distribution, but, unlike observations based on mtDNA, the population of Cabo Raso (CR) is not grouped within it. Another difference is the grouping of the two southernmost populations, Monte Clérigo (MC) and Playa del Rompeculos (PR), of which one (PR) corresponds to a geographical isolate presenting a set of differentiated mtDNA haplotypes. Inferences of population structure based on a Bayesian clustering approach do not clearly reveal these similarities (figure 4.14B and C).

Instead, they identify populations that are the most differentiated from the rest, possibly those that have been more affected by genetic drift due to recent fragmentation of the species range. It is noteworthy that the geographical isolate in southern Spain does not emerge as a highly distinct population nor does it bear reduced levels of variability as expected if it suffered from a strong reduction in effective population size.

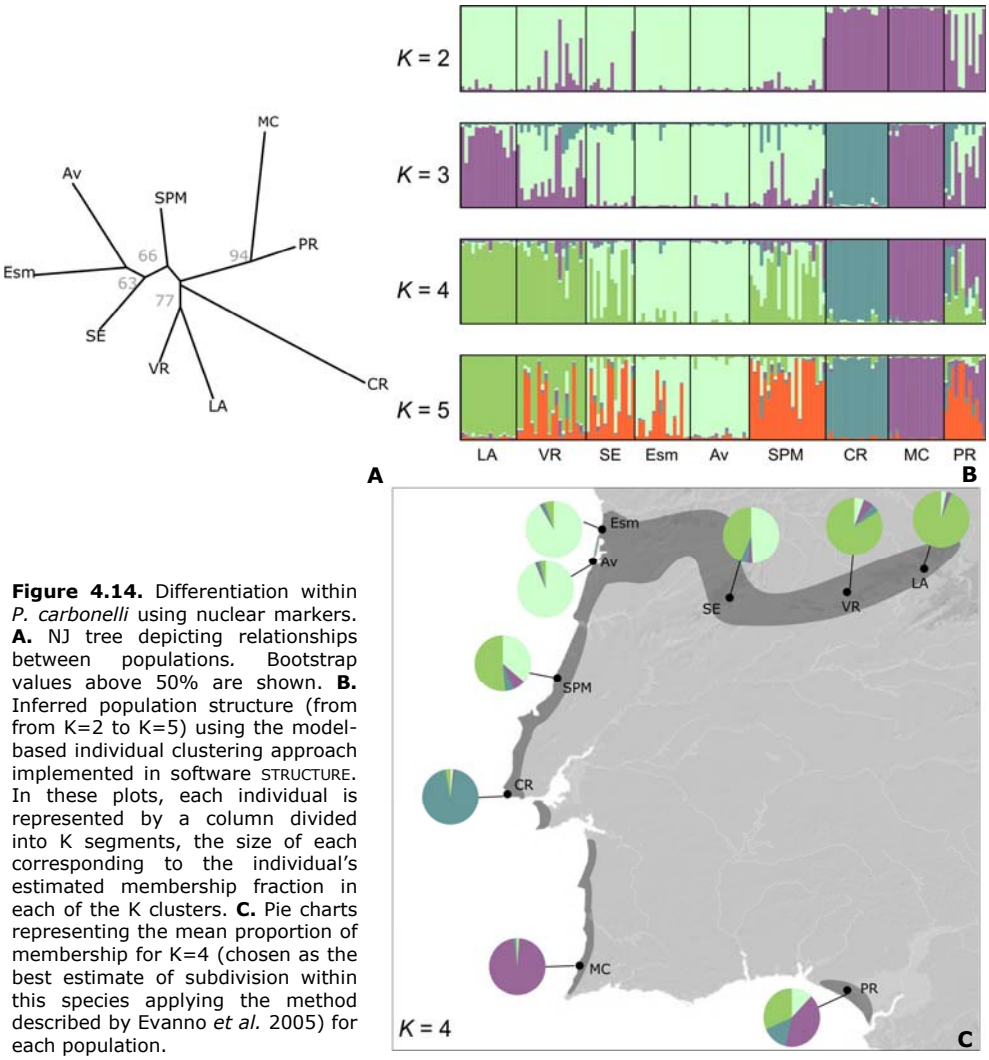

Figure 4.14. Differentiation within *P. carbonelli* using nuclear markers. **A.** NJ tree depicting relationships between populations. Bootstrap values above 50% are shown. **B.** Inferred population structure (from from K=2 to K=5) using the model-based individual clustering approach implemented in software STRUCTURE. In these plots, each individual is represented by a column divided into K segments, the size of each corresponding to the individual's estimated membership fraction in each of the K clusters. **C.** Pie charts representing the mean proportion of membership for K=4 (chosen as the best estimate of subdivision within this species applying the method described by Evanno *et al.* 2005) for each population.

The lack of concordance between nuclear and mtDNA variation anticipates difficulties in the understanding of the evolutionary history of this species. Reconciling information from these two sets of markers is not an easy task,

since differences among the evolutionary patterns depicted by different genomic compartments can be due to a wide number of causes (e.g. stochasticity, demography, selection, and admixture). Both types of markers are concordant in the inference of a shallow history of differentiation, which, by itself, increases the possibility that stochasticity might preside to the different partition of genetic variability observed in mtDNA and nuclear markers (Irwin 2002, Kuo and Avise 2005). Furthermore, we cannot rule out the possibility that a very recent (i.e. post-glacial) fragmentation in the species range might have completely masked the complex signatures of glacial isolation, expansion and admixture that are often reported for other species (Godinho 2003, Sequeira 2006). Such a hypothesis is not farfetched, since several lines of evidence suggest that this species has had a much larger distribution in the past, particularly the uniqueness of its distribution, the presence of geographical isolates and the difficulty in modelling its distribution. In this context, it is impossible to pinpoint for *P. carbonelli* areas that might have acted as glacial refugia: for example, while the isolated population in southern Spain bears a set of differentiated mtDNA haplotypes that could indicate a history of isolation, the close relationship inferred between this population and another southern locality argues in favour of a fixation of divergent haplotypes due to a recent stochastic lineage extinction.

Another discordant feature between the evolutionary trends depicted by mtDNA and nuclear markers concerns the demographic history of the northwestern clade within this species. Its mitochondrial DNA genealogy is clearly star-shaped, and estimates of demographic growth suggest that this clade suffered a dramatic population expansion. This is enforced by a progressive loss of mitochondrial DNA variability from south to north, leading to an almost fixation of haplotype C18 (figure 4.10B) in northernmost populations. However, there is not a concordant reduction in allozyme and microsatellite variability across the same populations (figure 4.15), which was therefore highly unexpected. Several factors could help explain this pattern. For example, the expansion could have occurred in a slow enough fashion for this trend not to be detected in nuclear markers but to be present in mtDNA; likewise, it is also possible that, again, present population dynamics might be masking the effects of a sudden demographic growth, since the area that has been inferred to have been recently colonized is precisely the same that now bears the most impressive species abundance (Sá-Sousa 2001a, personal observation), while the populations that supposedly have functioned as a source for post-glacial expansion are now locations where the species is scarce. Nevertheless, this does not explain why a clear-cut pattern of variability decrease was observed in mitochondrial DNA. Another possibility is, therefore, a selective sweep acting on the mitochondrial genome.

Taken together, these results suggest that the evolutionary history of *P. carbonelli*, although not readily understood due to recent contractions of its distribution, has been probably complex; without insights from nuclear markers, we would be led to think that mitochondrial variation within this species was consistent with a simple scenario of isolation in several glacial refugia, which would therefore constitute an oversimplified description of this species' evolution.

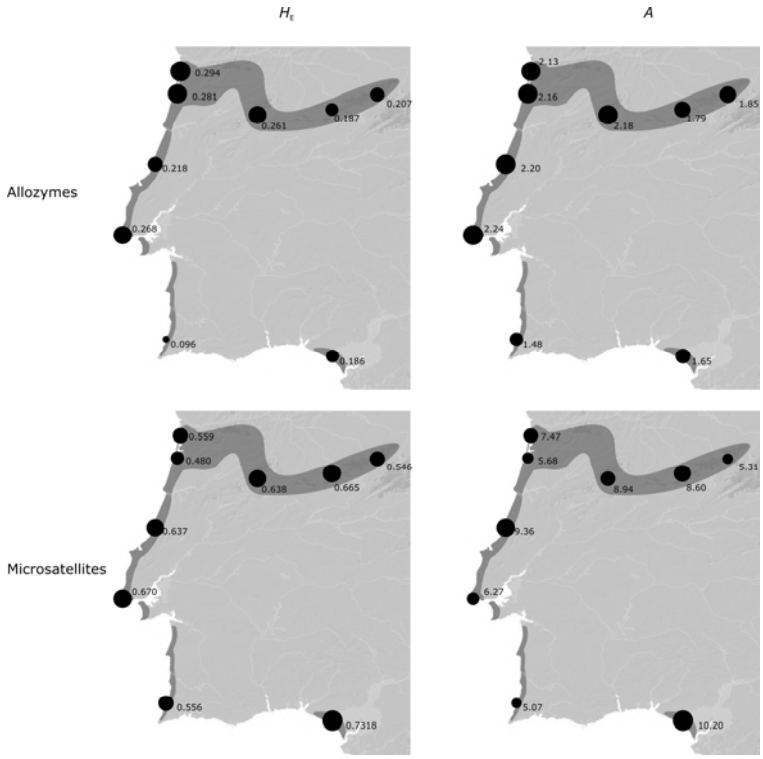

Figure 4.15. Mean diversity measures (expected heterozygosity (H_E) and allelic richness (A)) in populations of *P. carbonelli*.

4.3. References

Alexandrino J, Arntzen JW, Ferrand N (2002) Nested clade analysis and the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Heredity*, **88**, 66–74.

Alexandrino J, Froufe E, Arntzen JW, Ferrand N (2000) Genetic subdivision, glacial refugia and postglacial recolonization in the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Molecular Ecology*, **9**, 771–781.

Álvarez Y, Mateo JA, Andreu AC, Díaz-Paniagua C, Diez A, Bautista JM (2000) Mitochondrial DNA haplotyping of *Testudo graeca* on both continental sides of the Straits of Gibraltar. *Journal of Heredity,* **91,** 39–49.

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.

Arntzen JW, Sá-Sousa P (in press) Morphological and genetical differentiation of lizards (*Podarcis bocagei* and *P. hispanica*) in the Ria de Arosa archipelago (Galicia, Spain) resulting from vicariance and occasional dispersal

Avise JC Walker D (1999) Species realities and numbers in sexual vertebrates: Perspectives from an asexually transmitted genome. *Proceedings of the National Academy of Sciences USA*, **96**, 992–995.

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,** 527–545.

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

Barbosa D, Desfils E, Carretero MA, Font E (2005) Chemical stimuli mediate species recognition in *Podarcis* wall lizards. *Amphibia-Reptilia*, **26,** 257–263.

Barbosa D, Font E, Desfilis E, Carretero MA (2006) Chemically mediated species recognition in closely related *Podarcis* wall lizards. *Journal of Chemical Ecology,* **32,** 1587–1598.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16,** 113–148.

Baum D, Shaw KL (1995) Genealogical perspectives on the species problem. In: *Experimental and molecular approaches to plant biosystematics* (Hoch PC, Stephenson AC, eds.) Missouri Botanical Garden, St. Louis, MO, USA. 289–303.

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, **22,** 148–155.

Bonaccorso E, Peterson AT (2007) A multilocus phylogeny of New World jay genera. *Molecular Phylogenetics and Evolution*, **42,** 467–476.

Bons J, Geniez P (1996) *Amphibiens et Reptiles du Maroc (Sahara Occidental compris). Atlas biogéographique*. Asociación Herpetológica Española, Barcelona.

Brown RP, Suárez NM, Pestano J (2002) The Atlas mountains as a biogeographical divide in Northwestern Africa: evidence from mtDNA evolution in the Agamid lizard *Agama impalearis*. *Molecular Phylogenetics and Evolution*, **24,** 324–332.

Buckley TR, Cordeiro M, Marshall DC, Simon C (2006) Differentiation between hypotheses of lineage sorting and introgession in New Zealand alpine cicadas (*Maoricicada* Dugdale). *Systematic Biology* **55,** 411 – 425.

Bull V, Beltrán M, Jiggins CD, McMillan WO, Bermingham E, Mallet J (2006) Polyphyly and gene flow between non-sibling *Heliconius* species. *BMC Biology* **4,** 11.

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.

Campbell P, Schneider CJ, Adnan AM, Zubaid A, Kunz TH (2006) Comparative population structure of *Cynopterus* fruit bats in peninsular Malaysia and southern Thailand. *Molecular Ecology,* **15,** 29–47.

Capula M (1993) Natural hybridization in *Podarcis sicula* and *P. wagleriana* (Reptilia: Lacertidae). Biochemical *Systematics and Ecology,* **3,** 373–380.

Capula M (1994a) Genetic variation and differentiation in the lizards, *Podarcis wagleriana* (Reptilia: Lacertidae). *Biological Journal of the Linnean Society* **52,** 177–196.

Capula M (1994b) Evolutionary relationships of *Podarcis* lizards from Sicily and the Maltese Islands. *Journal of Zoological Systematics and Evolutionary Research,* **32,** 180–192.

Capula M (2002) Genetic evidence of natural hybridization between *Podarcis sicula* and *Podarcis tiliguerta* (Reptilia: Lacertidae). *Amphibia-Reptilia* **23,** 313–321.

Carranza S, Arnold EN (2004) History of West Mediterranean newts, *Pleurodeles* (Amphibia: Salamandridae) inferred from old and recent DNA sequences. *Systematics and Biodiversity* **1,** 327–337.

Carranza S, Arnold EN, Pleguezuelos JM (2006a) Phylogeny, biogeography and evolution of two Mediterranean snakes, *Malpolon monspessulanus* and *Hemorrhois hippocrepis* (Squamata, Colubridae), using mtDNA sequences. *Molecular Phylogenetics and Evolution,* **40,** 532 – 546

Carranza S, Arnold EN, Amat F (2004a) DNA phylogehy of *Lacerta* (*Iberolacerta)* and other lacertine lizards (Reptilia: Lacertidae): did competition cause long-term mountain restriction? *Systematics and Biodiversity,* **2,** 57–77.

Carranza S, Arnold EN, Wade E, Fahd S (2004b) Phylogeography of the false smooth snakes, *Macroprotodon* (Serpentes, Colubridae): mitochondrial DNA sequences show European populations arrived recently from Northwest Africa. *Molecular Phylogenetics and Evolution,* **33,** 523–532.

Carranza S, Harris DJ, Arnold EN, Batista V, Gonzalez de la Vega JP (2006b) Phylogeography of the lacertid lizard *Psammodromus algirus*, in Iberia and across the Strait of Gibraltar. *Journal of Biogeography,* **33,** 1279–1288.

Carretero MA, Sá-Sousa P, Barbosa D, Harris DJ, Pinho C (2002) Sintopía estricta entre *Podarcis bocagei* y *P. carbonelli*. *Bolletin de la Asociasión Herpetológica Española* **13,** 20-24.

Castilla AM, Fernandez-Pedrosa V, Backeljau T, Gonzalez A, Latorre A, Moya A (1998a) Mitochondrial DNA variability and conservation of an endangered lizard from the Columbretes islands. *Molecular Ecology,* **7,** 1047–1051.

Castilla AM, Fernandez-Pedrosa V, Harris DJ, Gonzalez A, Latorre A, Moya A (1998b) Mitochondrial DNA divergence suggests *Podarcis hispanica atrata* from the Columbretes Islands merits specific distinction. *Copeia* **4,** 1037–1040.

Comps B, Gömöry T, Letouzey J, Thiébaut B, Petit RJ (2001) Diverging trends between heterozygosity and allelic richness in the post-glacial colonization of the European beech. *Genetics* **157,** 389–397.

Cosson JF, Hutterer R, Libois R, *et al.* (2005) Phylogeographical footprints of the Strait of Gibraltar and Quaternary climatic fluctuations in the Western Mediterranean: a case study with the greater white-toothed shrew, *Crocidura russula* (Mammalia: Soricidae). *Molecular Ecology,* **14,** 1151–1162.

Coyne JA, Orr HA (1997) Patterns of speciation in *Drosophila* revisited. *Evolution,* **51,** 295–303.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, MA, USA.

Cracraft J (1989) Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In: *Speciation and its Consequences* (Otte D, Endler JA, eds.). Sinauer Associates, Sunderland, MA, USA, 28–59.

Creer S, Malhotra A, Thorpe RS (2003) Assessing the phylogenetic utility of four mitochondrial genes and a nuclear intron in the Asian Pit Viper genus, *Trimeresurus*: separate, simultaneous, and conditional data combination analyses. *Molecular Biology and Evolution*, **20,** 1240-1251.

Creer S, Pook CE, Malhotra A, Thorpe RS (2006) Optimal Intron Analyses in the Trimeresurus Radiation of Asian Pitvipers. *Systematic Biol*ogy **55,** 57-72.

Dobzhansky T (1937) *Genetics and the Origin of Species.* Columbia Univ. Press, New York, USA.

Dolman G, Moritz C (2006). A multilocus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). *Evolution,* **60,** 573–582.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution,* **54,** 1839–1854.

Estoup A, Jarne P, Cornuet J-M (2002) Homaplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. *Molecular Ecology,* **11,** 1591–1604.

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.

Evans BJ, Supriatna J, Andayani N, Melnick DJ (2003) Diversification of Sulawesi macaque monkeys: decoupled evolution of mitochondrial and autosomal DNA. *Evolution*, **57,** 1931–1946.

Freitas H (2006) *Natural hybridization between the Iberian hare (*Lepus granatensis*) and the brown hare (*L. europaeus*) in northern Iberian Peninsula*. MSc thesis, Faculdade de Ciências, Universidade do Porto.

Fritz U, Barata M, Busack SD, Fritzsch G, Castilho R (2006) Impact of mountain chains, sea straits and peripheral populations on genetic and taxonomic structure of a freshwater turtle, *Mauremys leprosa* (Reptilia, Testudines, Geoemydidae). *Zoologica Scripta* **35,** 97–108.

Fromhage L, Vences M, Veith M (2004) Testing alternative vicariance scenarios in Western Mediterranean discoglossid frogs. *Molecular Phylogenetics and Evolution*, **31,** 308–322.

Fu J (2000) Toward the phylogeny of the family Lacertidae – why 4708 base pairs of mtDNA sequences cannot draw the picture. *Biological Journal of the Linnean Society,* **71,** 203–217.

Fu Y-X (1997) Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics,* **133,** 693–709.

Galán P (2002) Hibridación en laboratorio de *Podarcis bocagei* y *Podarcis carbonelli*. *Boletin de la Asociacion Herpetologica Española,* **13,** 28–31.

Gantenbein B (2004) The genetic population structure of *Buthus occitanus* (Scorpiones: Buthidae) across the Strait of Gibraltar: calibrating a molecular clock using nuclear allozyme variation. *Biological Journal of the Linnean Society,* **81,** 519–534.

García-París M, Jockush EL (1999) A mitochondrial DNA perspectiva on the evolution of Iberian *Discoglossus* (Amphibia: Anura). *Journal of Zoology*, **248,** 209–218.

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. Mémoire présenté pour l'obtention du diplôme de l'École Pratique des Hautes Etudes, Montpellier.

Geniez P, Cluchier A, Sá-Sousa P, Guillaume CP, Crochet P-A (in press) Systematics of the *Podarcis hispanicus*-complex (Sauria, Lacertidae): I – redefinition, morphology and distribution of the nominotypical taxon. *Herpetological Journal* (in press)

Godinho R (2003) Evolutionary history of the Schreiber's green lizard, *Lacerta schreiberi*: Phylogenetic analysis, reconstruction of processes of population expansion, contraction and miscigenation, and integration in the biogeographic context of Northwestern Iberian Peninsula. PhD thesis: University of Lisbon.

Godinho R, Crespo EG, Ferrand N, Harris DJ (2005) Phylogeny and evolution of the green lizards, *Lacerta* spp. (Squamata: Lacertidae) based on mitochondrial and nuclear DNA sequences. *Amphibia-Reptilia*, **26,** 271-286.

Godinho R, Domingues V, Crespo EG, Ferrand N (2006a). Extensive intraspecific polymorphism detected by SSCP at the nuclear C-mos gene in the endemic Iberian lizard *Lacerta schreiberi*. *Molecular Ecology,* **15,** 731–738.

Godinho R, Mendonça B, Crespo EG, Ferrand N (2006b) Genealogy of the nuclear b-fibrinogen locus in a highly structured lizard species: comparison with mtDNA and evidence for intragenic recombination in the hybrid zone. *Heredity,* **16,** 454–463.

Godinho R, Paulo OS, Ferrand N, Luis C, Rosa HD, Crespo EG (2003). Major patterns of population differentiation in the Iberian Schreiber's Green Lizard (*Lacerta schreiberi*) inferred from protein polymorphism. *Herpetological Journal,* **13,** 35–42.

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origin and Conservation of European Biodiversity* (Weiss S, Ferrand N, eds.). Springer, Dordrecht, The Netherlands, 155–188.

Guillaumet A, Pons J-M, Godelle B, Crochet P-A (2006) History of the Crested Lark in the Mediterranean region as revealed by mtDNA sequences and morphology. *Molecular Phylogenetics and Evolution*, **39,** 645–656.

Guiller A, Coutellec-Vreto MA, Madec L, Jeuniff J (2001) Evolutionary history of land snail *Helix aspersa* on the Western Mediterranean: preliminary results inferred from mitochondrial DNA sequences. *Molecular Ecology*, **10,** 81–87.

Haldane JBS (1922) Sex-ratio and unisexual sterility in hybrid animals. *Journal of Genetics* **12,** 101-109.

Harris DJ, Batista V, Carretero MA (2004a) Assessment of genetic diversity within *Acanthodactylus erythrurus* (Reptilia: Lacertidae) in Morocco and the Iberian Peninsula using mitochondrial DNA sequence data. *Amphibia-Reptilia,* **25,** 227–232.

Harris DJ, Batista V, Lymberakis P, Carretero MA (2004b). Complex estimates of evolutionary relationships in *Tarentola mauritanica* (Reptilia: Gekkonidae). *Molecular Phylogenetics and Evolution,* **30,** 855–859.

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,** 1257–1262.

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.

Hedrick PW (1999) Highly variable loci and their interpretation in evolution and conservation. *Evolution* **53,** 313–318.

Hendry AP, Vamosi SM, Latham SJ, Heilbuth JC, Day T (2000) Questioning species realities. *Conservation Genetics* **1,** 67–76

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 (2000) The genetic legacy of the quaternary Ice Ages. *Nature*, **405,** 907–913.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B Biology Sciences,* **359,** 183–195.

Hey J (2001) The mind of the species problem. *Trends in Ecology and Evolution,* **7,** 326–329.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics,* **167,** 747–760.

Hey J, Won Y-J, Sivasundar A, Nielsen R, Markert JA (2004) Using nuclear haplotypes with microsatellites to study gene flow between recently separated cichlid species. *Molecular Ecology.* **13,** 909–919.

Howard DJ, Berlocher SH (1998) *Endless Forms: Species and Speciation*. Oxford University Press, Oxford, UK.

Howard DJ, Gregory PG, Chu J, Cain ML (1998) Conspecific sperm precedence is an effective barrier to hybridization between closely related species. *Evolution,* **52,** 511–516

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity,* **77,** 282–291.

Irwin DE (2002) Phylogeographic breacks without geographic barriers to gene flow. *Evolution* **56,** 2383–2394.

Jiggins C, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology and Evolution* **15,** 250–255

Johns GC, Avise JC (1998) A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b g*ene. *Molecular Biology and Evolution*, **15,** 1481–1490.

Jolly D, Harrison SP, Damnati B, BonneWlle R (1998) Simulated climate and biomes of Africa during the late quaternary: comparison with pollen and lake status data. *Quaternary Science Reviews*, **17,** 629–657.

Kaliontzopoulou A, Carretero MA, Llorente G (in press) Multivariate and geometric morphometrics in the analysis of sexual dimorphism variation in *Podarcis* Lizards. *Journal of Morphology*

Karl SA, Avise JC (1992) Balancing selection at allozyme loci in oysters: Implications from nuclear RFLPs. *Science* **256,** 100-102

Krajewsky C, Moyer GR, Sipiorski JT, Fain MG (2004) Molecular systematics of the enigmatic 'phascolosoricine' marsupials of New Guinea. *Australian Journal of Zoology,* **52,** 389–415.

Krijgsman W, Garcés M, Agustí J, Raffi I, Taberner C, Zachariasse WJ (2000) The "Tortonian salinity crisis" of the Eastern Betics (Spain). *Earth and Planetary Science Letters* **181,** 497–511.

Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, **149,** 429–434.

Kuo C-H, Avise JS (2005) Phylogeographic breaks in low-dispersal species: the emergence of concordance across gene trees. *Genetica*, **124,** 179–186.

Lessa E, Cook JA, Patton JL (2003) Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. *Proceedings of the National Academy of Sciences, USA*, **100,** 10331–10334.

Lexer C, Kremer A, Petit RJ (2006) Shared alleles in sympatric oaks: recurrent gene flow is a more parsimonious explanation than ancestral polymorphism. *Molecular Ecology*, **15,** 2007–2012.

MacCulloch RD, Fu J, Darevsky IS, Murphy R (2000) Genetic evidence for species status of some Caucasian rock lizards in the *Darevskia saxicola* group. *Amphibia-Reptilia,* **21,** 169–176.

Mallet J (1995) A species definition for the Modern Synthesis. *Trends on Ecology and Evolution,* **10,** 294–299.

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

Mallet J (2006) Species concepts. In: *Evolutionary Genetics: Concepts and Case Studies* (Fox CW, Wolf JB, eds.). Oxford University Press, Oxford, UK, 367–373.

Mallet J, Beltrán M, Neukirchen W, Linares M (2007) Natural hybridization in helinoniine butterflies: the species boundary as a continuum. *BMC Evolutionary Biology*, **7,** 28.

Mallet J, McMillan WO, Jiggins CD (1998). Estimating the mating behavior of a pair of hybridizing *Heliconius* species in the wild. *Evolution,* **52,** 503-510.

Martínez-Solano I (2004) Phylogeography of Iberian *Discoglossus* (Anura: Discoglossidae). *Journal of Zoological Systematics and Evolutionary Research*, **42,** 298–305.

Martínez-Solano I, Gonçalves H, Arntzen JW, García-París M (2004) Phylogenetic relationships and biogeography of midwife toads (Discoglossidae: *Alytes*). *Journal of Biogeography,* **31,** 603–618.

Martínez-Solano I, Teixeira J, Buckley D, Garcia-París M (2006) Mitochondrial DNA phylogeography of *Lissotriton boscai* (Caudata, Salamandridae): evidence for old, multiple refugia in na Iberian endemic. *Molecular Ecology,* **13,** 3375–3388.

Mayer W, Arribas OJ (1996). Allozyme differentiation and relationship among the Iberian-Pyrenean Mountain Lizards (Squamata: Sauria: Lacertidae). *Herpetozoa* **9,** 57–61.

Mayr E (1942) *Systematics and the Origin of Species*. Columbia University Press, New York, USA.

Mayr E (1957) Species concepts and definitions. In: *The Species Problem* (Mayr E, ed.). American Association for the Advancement of Science, Washington, USA. 1–22.

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.

Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution,* **49,** 718–726.

Muir G, Schlotterer C (2005) Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Molecular Ecology*, **14,** 549–561.

Muir G, Schlotterer C (2006) Moving beyond single-locus studies to characterize hybridization between oaks (*Quercus* spp.). *Molecular Ecology*, **15,** 2301–2304.

Muller HJ, (1940) Bearing of the *Drosophila* work on systematics, pp.185-268 in *The new systematics*. Clarendon press, Oxford, U.K.

Muller HJ (1942) Isolating mechanisms, evolution and temperature. *Biological Symposium,* **6,** 71-125.

Nei M (1972) Genetic distance between populations. *American Naturalist,* **106,** 283-292.

Neigel JE, Avise JC (1993) Application of a random-walk model to geographical distribution of animal mitochondrial DNA variation. *Genetics,* **135,** 1209–1220.

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 MA, Mariottini P (2000) Molecular biogeography of the Mediterranean lizards *Podarcis* Wagler, 1830 and *Teira* Gray, 1838 (Reptilia, Lacertidae). *Journal of Biogeography,* **27,** 1403–1420.

Oosterbroek P, Arntzen JW (1992) Area-cladograms of circum-Mediterranean taxa in relation to Mediterranean palaeogeography. *Journal of Biogeography*, **19,** 3–20.

O'Reilly PT, Canino MF, Bailey KM, Bentzen P (2004) Inverse relationship between F_{ST} and microsatellite polymorphism in the marine fish, walleye Pollock (*Theragra chalcogramma*): implications foro resolving weak population structure. *Molecular Ecology*, **13,** 1799–1814.

Paulo OS (2001) *Phylogeography of reptiles of the Iberian Peninsula*. PhD Thesis. University of London.

Paulo OS, Dias C, Bruford MW, Jordan WC, Nichols RA (2001) The persistence of Pliocene populations through the Pleistocene climatic cycles: evidence from the phylogeography of an Iberian lizard. *Proceedings of the Royal Society of London*. *Series B, Biological Sciences*, **268,** 1625–1630.

Paulo OS, Jordan WC, Bruford MW, Nichols RA (2002a) Using nested clade analysis to assess the history of colonization and the persistence of populations of an Iberian lizard. *Molecular Ecology*, **11,** 809–819.

Paulo OS, Pinto I, Bruford MW, Jordan WC, Nichols RA (2002b) The double origin of Iberian peninsular chameleons. *Biological Journal of the Linnean Society*, **75,** 1–7.

Pérez-Mellado V, Galindo MP (1986) Sistemática de *Podarcis* (Sauria, Lacertidae) ibéricas y norteafricanas mediante técnicas multidimensionales. *Série Manuales Universitários,* Univ. Salamanca, Spain.

Piálek J, Barton NH (1997) The spread of an advantageous allele across a barrier: the effects of random drift and selection against heterozygotes. *Genetics*, **145,** 493–504.

Pinho C, Ferrand N, Harris DJ (2004) Genetic variation within the *Podarcis hispanica* species complex: new evidence from protein electrophoretic data. In: *The Biology of Lacertid Lizards: Evolutionary and Ecological Perspectives* (Pérez-Mellado V, Riera N, Perera A, eds.). Institut Menorquí d'Estudis. *Recerca*, **8,** 269–277.

Pogson GH, Mesa KA, Boutilier RG (1995) Genetic population structure and gene flow in the Atlantic cod, *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. *Genetics,* **139**, 375–385.

Price TD, Bouvier MM (2002) The evolution of F1 postzygotic incompatibilities in birds. *Evolution,* **56,** 2083–2089

Pritchard JK, Stephens M, Donnelly P. (2000). Inference of population structure using multilocus genotype data. *Genetics,* **155,** 945–959.

Prychitko TM, Moore WS (1997) The utility of DNA sequences of an intron from the beta-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). *Molecular Phylogenetics and Evolution*, **8,** 193–204.

Prychitko TM, Moore WS (2003) Alignment and phylogenetic analysis of β-Fibrinogen intron 7 sequences among avian orders reveal conserved regions within the intron. *Molecular Biology and Evolution*, **20**, 762-771.

Pulquério MJ, Nichols RA (in press) Dates from the molecular clock: how wrong can we be? *Trends in Ecology and Evolution* (in press).

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19,** 2092–2100.

Renoult J (2006) *Histoire évolutive du Lézard ibérique* Podarcis hispanicus *(Squamata: Lacertidae): discordances entre morphologie, ADN nucléaire et ADN mitochondrial*. Master 2 Recherche. Université Montpellier II Sciences et Techniques du Languedoc.

Rieseberg LH, Baird S, Desrochers A (1998) Patterns of mating in wild sunflower hybrid zones. *Evolution* **52,** 713–726.

Rosenberg NA, Mahajan S, Ramachandran S, Chengfeng Z, Pritchard JK, Feldman MW (2005). Clines, clusters and the effect of study design on the inference of human population structure. *PLoS Genetics* **1,** e70

Sá-Sousa P (2001a) A Controversa Sistemática das Lagartixas do género Podarcis WAGLER, 1830 (Sauria: Lacertidae) em Portugal. Dissertação de Doutoramento. Fac. Ciências, Univ. Lisboa.

Sá-Sousa P (2001b) Comparative chorology between *Podarcis bocagei* and *P. carbonelli* (Sauria: Lacertidae) in Portugal. *Revista Española de Herpetología* **15,** 85–97.

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.

Saitou N, Nei M (1987) The Neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution,* **4,** 406–427.

Sequeira F (2006) História evolutiva da salamandra-lusitânica, *Chioglossa lusitanica*. Filogenia, reconstrução dos padrões filogeográficos e análise dos processos de miscigenação na zona híbrida através do uso de múltiplos marcadores moleculares. Dissertação de Doutoramento, Fac. Ciências, Univ. Porto.

Sequeira F, Ferrand N, Harris DJ (2006). Assessing the phylogenetic signal of the nuclear ß-Fibrinogen intron 7 in salamandrids (Amphibia: Salamandridae). *Amphibia-Reptilia,* **27,** 409–418.

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, USA,* **99,** 16122–16127.

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution,* **13,** 964–969.

Steindachner F (1870) Herpetologische Notizen II. Über einige neue oder seltene Reptilien des Wiener Museums. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien. Mathematisch-Naturwissenschaftliche classe,* **62,** 336–350.

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

Vasconcelos 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.

Veith M, Mayer C, Samroui B, Donaire Barroso D, Bogaerts S (2004) From Europe to Africa and vice-versa: evidence for multiple intercontinental dispersal in ribbed salamanders (Genus *Pleurodeles*). *Journal of Biogeography*, **31,** 159–171

Willis K, Whittaker RJ (2000) Paleoecology: the refugial debate. *Science,* **287,** 1406–1407.

Wright S (1931) Evolution in Mendelian populations. *Genetics,* **28,** 114–138.

Wu C-I (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology,* **14,** 851–861.

Yu L, Zhang YP (2006) Phylogeny of the caniform carnivore: evidence from multiple genes. *Genetica,* **127,** 65–79.

Zangari F, Cimmaruta R, Nascetti G (2006) Genetic relationships of the western Mediterranean painted frogs based on allozymes and mitochondrial markers: evolutionary and taxonomic inferences (Amphibia, Anura, Discoglossidae). *Biological Journal of the Linnean Society,* **87,** 515–536.

Chapter 5

Concluding remarks and future perspectives

Concluding remarks and future perspectives

The work presented in this thesis constitutes a major advance regarding the previous knowledge on the evolutionary history of Iberian and North African *Podarcis,* and is also an important contribution towards the understanding of the evolutionary dynamics of the Western Mediterranean biota in general. First, we have established a solid scenario for the mitochondrial DNA evolution of these lizards, confirming the distinctiveness of previously described lineages, offering alternative biogeographic scenarios for the origin of biological diversity within the clade and describing a novel lineage that had remained undetected. Second, we have corroborated the overall distinctiveness of mitochondrial clades from a nuclear marker perspective, demonstrating that distinct phylogroups within *P. hispanica* are, as suggested by the mitochondrial DNA, as differentiated as fully recognised species. Third, by detecting cases of clear mitochondrial-nuclear discordance and by directly assessing levels of gene flow across a contact zone, we have documented for the first time that forms within the species complex are not fully reproductively isolated; although most cases of gene flow are limited and do not seem to compromise the overall differentiation of the species/lineages, we have documented important levels of miscegenation among some species pairs, which may have led to a lack of nuclear distinction of distantly-related mitochondrial DNA clades. Fourth, we have shown that, as observed in other Western Mediterranean organisms, the genetic variability within species of *Podarcis* has been strongly affected by Pleistocene climatic oscillations, and that the patterns that are typically found in co-distributed taxa (signatures of glacial isolation and post-glacial expansion) can also be detected in these species, additionally demonstrating the transversality of such patterns. Finally, we have also shown that patterns such as those above-mentioned have different magnitudes on distinct species, caused by an uneven, largely species-specific response to the climatic oscillations.

By sequencing a larger portion of mitochondrial DNA than previous studies, we were able to obtain a highly supported estimate of relationships between the various forms, which allowed the establishment of alternative biogeographic scenarios to the ones previously proposed to explain the partition of genetic variability across the Strait of Gibraltar. Although we now suggest that the opening of the Strait might have caused vicariant isolation between Iberian and North African forms, we also document with confidence an episode of transmarine colonization after the Strait was formed, which adds up to a growing body of evidence suggesting that this geographic discontinuity is not as effective a barrier to gene flow between both continents as previously thought.

Because our suggestions regarding the timing and polarity of the colonization of North Africa are highly dependent upon a molecular clock calibration, the results obtained in this respect should be taken with caution, moreover because we were not able to obtain enough information from nuclear markers to help enlighten this subject. Therefore, future work with respect to the testing of these hypotheses should aim both to calibrate a molecular clock specifically for these species and to investigate faster-evolving nuclear gene genealogies in order to obtain an independent assessment of population subdivision.

By studying mtDNA variation, we also detected previously undescribed and highly differentiated lineages (namely in southern Spain and in Algeria). Taken together with the apparent geographic confinement of some phylogroups within *P. hispanica*, this could indicate that despite the recent effort in obtaining samples from as many unsampled regions as possible, overall diversity within the clade may still be underestimated. Moreover, the restricted distribution ranges of some phylogroups suggest that they may need special conservation attention. Therefore, a necessary future step is the ecological characterization and the evaluation of potential threats, not only for fully recognised species but also for lineages within *P. hispanica*.

The description of genetic variation at nuclear markers, the first large-scale study on the subject ever conducted in *Podarcis*, suggested that, although most mtDNA lineages are not identifiable from the analyses of single genealogies due to the persistence of high levels of ancestral polymorphism, they are clearly differentiated from a multilocus perspective, as shown by our analyses of allozyme variation. Moreover, despite their polyphyly, nuclear gene genealogies suggest few cases of nuclear introgression between mtDNA lineages. Both allozyme and nuclear intron data thus suggest that historical gene flow among forms must have been kept at low levels throughout the process of divergence and that mtDNA lineages are overall differentiated with respect to the nuclear genome. Recognised species do not stand out as more differentiated than phylogroups within *P. hispanica*, which is indicative that the present classification is inappropriate. Nevertheless, it is clear, from the persistence of ancestral polymorphism and the confirmation of the permeability of forms to gene exchange with congeners, that forms of *Podarcis* arose quite recently and have not yet completed the process of speciation. Conventional classification systems do not deal well with such intermediate stages of differentiation. On one hand, it is clear that several well-differentiated entities exist; on the other hand, they have poorly-defined boundaries. Our data suggest that a wide spectrum of introgression patterns, unrelated to mitochondrial DNA distances, may be detected among lineages, from virtually zero levels of gene flow to large amounts of admixture compromising the forms' distinctiveness. This implies that attempts to establish a correlation between mitochondrial DNA divergence and

the specific status in this and other species complexes are likely to fail. *Podarcis* therefore constitute a clear example of the limitations of the single use of mtDNA to define evolutionary units and conduct taxonomic reevaluations. At a more general level, our results add up to recent evidence stemming from the study of other incipient species groups, clearly illustrating that the acquisition of reproductive isolation (i.e. speciation) is a gradual process.

Regarding the description of nuclear gene variation, the work presented in this thesis clearly represents the beginning. More accurate estimates of divergence times and rates of gene flow, and probably some insights into evolutionary relationships among forms will be obtained with the inclusion of both more individuals and more gene genealogies in the analyses. This task, which would seem impossible to perform a few years ago due to the lack of genomic resources for non-model vertebrates, is now a clearly feasible and promising avenue of research. In particular, we hope that in the near future we will be able to apply the study of a large number of nuclear genes to the genetic characterization of contact zones between distinct lineages, in a similar way to that which we have accomplished for the contact zone between *P. bocagei* and *P. carbonelli*. This will be of utmost importance in the evaluation of the degree of reproductive isolation that the various forms of the species complex may have achieved. It is however important that this study is carried out in a multidisciplinary framework, including a thorough morphological, ecological and physiological characterization of genetic lineages, which is already in development.

At a more general level, our results also demonstrate that, because of the diversity of distribution patterns (with forms existing in allopatry, parapatry and broad sympatry) and the probable existence of a very large number of contact zones, *Podarcis* lizards provide an excellent model system for the study of the dynamics of genetic, morphological and ecological diversification.

Besides providing data on the nature and dynamics of differentiation among species, we also addressed the subject of intraspecific variation in species of *Podarcis*. Our results suggest that the three studied species were significantly affected by the Quaternary climatic oscillations, similarly to recent descriptions of the phylogeographic structure in Iberian and, to a lesser extent, North African taxa. This reinforces what has become known as the "refugia within refugia" paradigm: southern refugia not only functioned as pockets of survival and sources of post-glacial recolonization, but also experienced drastic habitat changes that are reflected on the present genetic structure of inhabiting organisms. Nevertheless, we detected a high degree of variation in the patterns of subdivision and demography inferred for the three studied species, which most likely reflect the different latitudes on which the species are distributed.

One of those differences was the depth of observed population structure, which is correlated to the levels of persistence across different glacial cycles. The work herein presented therefore illustrates that the degree of population subdivision varies widely in both a spatial and a temporal framework even across closely-related taxa. Following this suggestion, a promising future line of work related to the Western Mediterranean Quaternary biogeography should be not only to assess concordance among patterns described for different species but also to analyse the nature and causes of discrepancy in species' responses.

Appendix

Table A1. A compilation of published and unpublished mitochondrial DNA sequencing locality data. The corresponding map is shown in figure 4.4.

	Locality	Region	Country	Gene	Reference
Pb	Esposende	Porto	Portugal	1, 3	Harris & Arnold 1999, Oliverio *et al.* 2000
	Montesinho	Bragança	Portugal	1, 2, 4	Harris & Sá-Sousa 2001, this thesis (article VI)
	Vila Pouca de Aguiar	Vila Real	Portugal	1, 2	Harris & Sá-Sousa 2001
	Serra do Gerês	Braga	Portugal	1, 2, 4	Harris & Sá-Sousa 2001, this thesis (article VI)
	Vairão	Porto	Portugal	1, 2, 3, 4	Harris & Sá-Sousa 2001, 2002, this thesis (article VI)
	Braga	Braga	Portugal	1, 2, 4	Harris & Sá-Sousa 2001, this thesis (article VI)
	Viana do Castelo	Viª do Castelo	Portugal	1, 2	Harris & Sá-Sousa 2001
	Malpica	A Coruña	Spain	1, 3, 4, 5, 6	Harris & Sá-Sousa 2002, this thesis (article I)
	Madalena	Porto	Portugal	1, 3, 4, 5, 6	Harris & Sá-Sousa 2002, this thesis (articles I, VI)
	Posadilla de la Vega	León	Spain	6	Renoult 2006
	Arosa Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Benencia Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Jidoiro Arenoso Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	San Bartolomé Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Toja Pequeña Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Praia de Louro	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Mirador de la Curota	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Mostero de la Armenteira	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Tanes	Astúrias	Spain	1, 3, 4, 5, 6	This thesis (articles I, VI)
	Valdoviño	A Coruña	Spain	4	This thesis (article VI)
	A Coruña	A Coruña	Spain	4	This thesis (article VI)
	Sarria	Lugo	Spain	4	This thesis (article VI)
	Sanxenxo	Pontevedra	Spain	4	This thesis (article VI)
	Taboadela	Ourense	Spain	4	This thesis (article VI)
	Moledo	Vª do Castelo	Portugal	4	This thesis (article VI)
	Caldas das Taipas	Braga	Portugal	4	This thesis (article VI)
	Zimão	Vila Real	Portugal	4	This thesis (article VI)
	Alvão	Vila Real	Portugal	4	This thesis (article VI)
	Marco de Canavezes	Porto	Portugal	4	This thesis (article VI)
	Penafiel	Porto	Portugal	4	This thesis (article VI)
	Espinho	Aveiro	Portugal	4	This thesis (article VI)
	Mindelo	Porto	Portugal	3	This thesis (unpublished)
Pc	Serra da Estrela	Guarda	Portugal	1, 2, 3, 4	Harris & Sá-Sousa 2001, 2002, this thesis (article VI)
	Torreira	Aveiro	Portugal	1, 2, 3	Harris & Sá-Sousa 2001, 2002
	Monte Clérigo	Faro	Portugal	1, 2, 3, 4	Harris & Sá-Sousa 2001, 2002, this thesis (article VI)
	Peniche	Leiria	Portugal	1, 2	Harris & Sá-Sousa 2001
	Berlenga Is.	Leiria	Portugal	1, 2, 4	Harris & Sá-Sousa 2001, this thesis (article VI)
	Playa del Rompeculos	Huelva	Spain	1, 3, 4, 5, 6	Harris *et al.* 2002a, this thesis (articles I, VI)
	Peña de Francia	Salamanca	Spain	6	Renoult 2006
	Aveiro	Aveiro	Portugal	1, 3, 4, 5, 6	This thesis (articles I, VI)
	La Alberca	Salamanca	Spain	1, 3, 4, 5, 6	This thesis (articles I, VI)
	Espinho	Aveiro	Portugal	3	This thesis (article V)
	Villasrúbias	Salamanca	Spain	4	This thesis (article VI)
	Sátão	Viseu	Portugal	4	This thesis (article VI)
	Pendilhe	Viseu	Portugal	4	This thesis (article VI)
	S. Pedro do Sul	Viseu	Portugal	4	This thesis (article VI)
	Tondela	Viseu	Portugal	4	This thesis (article VI)
	Esmoriz	Aveiro	Portugal	4	This thesis (article VI)
	Carriço	Leiria	Portugal	4	This thesis (article VI)
	S. Pedro de Moel	Leiria	Portugal	4	This thesis (article VI)
	Cabo Raso	Lisboa	Portugal	4	This thesis (article VI)
	Meco	Setúbal	Portugal	4	This thesis (article VI)
	Sines	Setúbal	Portugal	4	This thesis (article VI)
	S. Jacinto	Aveiro	Portugal	3	This thesis (unpublished)
	Farilhão Is,	Leiria	Portugal	3	This thesis (unpublished)
Pv	Atlas Mts.	?	Morocco	1	Castilla *et al.* 1998b
	Mid Atlas Mts.	?	Morocco	1, 2, 3, 4, 5, 6	Harris & Sá-Sousa 2001, 2002, this thesis (article I)
	10Km W Bab-Berred	Tetouan	Morocco	3	Oliverio *et al.* 2000
	Azrou	Ifrane	Morocco	1, 3, 4	Harris *et al.* 2002b, this thesis (article VI)
	Oukaïmeden	Marrakech	Morocco	1, 3, 4, 5, 6, 7	Harris *et al.* 2002b, Busack *et al.* 2005, this thesis (articles I, VI)
	Bab-Berred	Tetouan	Morocco	1, 3, 4	Harris *et al.* 2002b, this thesis (article VI)
	8Km SW Zinat	Tetouan	Morocco	1, 3, 4	Harris *et al.* 2002b, this thesis (article VI)
	15 Km SW Zinat	Tetouan	Morocco	1, 3, 4	Harris *et al.* 2002b, this thesis (article VI)
	El-Had	Tetouan	Morocco	1, 3, 4	Harris *et al.* 2002b, this thesis (article VI)
	Jebel Musa	Tetouan	Morocco	1, 3, 4	Harris *et al.* 2002b, this thesis (article VI)
	Taza	Taza	Morocco	1, 3, 4	Harris *et al.* 2002b, this thesis (article VI)
	Mischliffen	Ifrane	Morocco	1, 3, 4	Harris *et al.* 2002b, this thesis (article VI)
	Bab Taza	Chefchaouen	Morocco	1, 3, 4, 5, 6	Harris *et al.* 2002b, this thesis (articles I, VI)
	N Oukaïmeden	Marrakech	Morocco	1, 3	Harris *et al.* 2002b
	Guadalcacin	Cádiz	Spain	1, 3, 4, 5, 6	Harris *et al.* 2002b, this thesis (article I)
	Mairena del Aljarafe	Sevilla	Spain	1, 3, 4	Harris *et al.* 2002b, this thesis (article VI)
	Sevilla	Sevilla	Spain	1, 3	Harris *et al.* 2002b
	Barbate	Cádiz	Spain	1, 3	Harris *et al.* 2002b
	Huelva	Huelva	Spain	1, 3, 4, 5, 6	Harris *et al.* 2002b, this thesis (articles I, VI)
	El Ksiba	Beni-Mellal	Morocco	1, 4, 7	Busack *et al.* 2005
	Asilah	Tetouan	Morocco	1, 4, 7	Busack *et al.* 2005
	Ksar-es-Seghir	Tetouan	Morocco	1, 4, 7	Busack *et al.* 2005
	Río Hozgarganta	Málaga	Spain	1, 4, 7	Busack *et al.* 2005
	Facinas	Cádiz	Spain	1, 4, 7	Busack *et al.* 2005
	Sierra Nevada	Granada	Spain	6	Renoult 2006
	Tizi n' Tichka	Marrakech	Morocco	6	Renoult 2006
	El Rocio	Huelva	Spain	6	Renoult 2006
	Ronda	Málaga	Spain	6	Renoult 2006
	Motril	Granada	Spain	6	Renoult 2006
	La Barrosa	Cádiz	Spain	4	This thesis (article VI)
	Chefchaouen	Chefchaouen	Morocco	4	This thesis (article VI)
	Ketama	Al-Hoceima	Morocco	4	This thesis (article VI)
	Jebel Tazzeka	Taza	Morocco	4	This thesis (article VI)
	Debdou	Oujda	Morocco	4	This thesis (article VI)
	Balcon d'Ito	Ifrane	Morocco	4	This thesis (article VI)
	Foum Kheneg	Ifrane	Morocco	4	This thesis (article VI)
	Midelt	Khenifra	Morocco	4	This thesis (article VI)
	Tizi n'Test	Marrakech	Morocco	4	This thesis (article VI)

Table A1. (cont.)

	Locality	Region	Country	Gene	Reference
Pv	Imlil	Marrakech	Morocco	4	This thesis (unpublished)
	Tiaret	Tiaret	Algeria	4	This thesis (unpublished)
	M'Chedallah	Bouira	Algeria	4	This thesis (unpublished)
	Ain Harhar	Ain Defla	Algeria	4	This thesis (unpublished)
	Tahament	?	Algeria	4	This thesis (unpublished)
	Djurjura	Tizi Ouzou	Algeria	4	This thesis (unpublished)
*Ph*1A	Vila Real	Vila Real	Portugal	1, 2, 3, 4, 5, 6	Harris & Sá-Sousa 2001, 2002, this thesis (article I)
	Montesinho	Bragança	Portugal	1, 2, 3, 4, 5, 6	Harris & Sá-Sousa 2001, 2002, this thesis (articles I, IV)
	Vila de Rua	Viseu	Portugal	1, 3, 4, 5, 6	Harris & Sá-Sousa, 2002, this thesis (article I)
	Coroso Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Jidoiro Pedregoso Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Noro Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Rua Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Sálvora Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Vionta Is.	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Mirador de la Curota	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Mostero de la Armenteira	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Punta Moreiras	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Ardia	Pontevedra	Spain	1	Arntzen & Sá-Sousa in press
	Los Ancares	León	Spain	4	This thesis (article IV)
	Tua	Bragança	Portugal	4	This thesis (article IV)
	Ria de Arosa	Pontevedra	Spain	4	This thesis (article IV)
	Pendilhe	Viseu	Portugal	4	This thesis (article IV)
	Gerês	Braga	Portugal	3	This thesis (unpublished)
	Sto. Estêvão	Vila Real	Portugal	3	This thesis (unpublished)
	Oliveira do Hospital	Coimbra	Portugal	4	This thesis (unpublished)
	Vale Rossim	Serra da Estrela	Portugal	3	This thesis (unpublished)
	Zamora	Zamora	Spain	3	This thesis (unpublished)
	Celanova	Ourense	Spain	3	This thesis (unpublished)
*Ph*1B	Trujillo	Cáceres	Spain	1, 3, 4, 5, 6	Harris & Sá-Sousa 2002, this thesis (article I)
	Oropesa	Toledo	Spain	1, 3, 4, 5, 6	Harris & Sá-Sousa 2002, this thesis (article I)
	Sierra de Gredos	Ávila	Spain	1, 3	Harris *et al.* 2002b
	Maqueda	Toledo	Spain	1, 3	Harris *et al.* 2002b
	Peña de Francia	Salamanca	Spain	1, 3	Carranza *et al.* 2004
	San Martín del Pimpollar	Ávila	Spain	1, 4, 7	Busack *et al.* 2005
	Arenas de San Pedro	Ávila	Spain	1, 4, 7	Busack *et al.* 2005
	El Tiemblo	Ávila	Spain	6	Renoult 2006
	Villacastín	Segóvia	Spain	4	This thesis (article IV)
	Guadarrama	Madrid	Spain	4	This thesis (article IV)
	La Alberca	Salamanca	Spain	4	This thesis (article IV)
	Torrejón de la Calzada	Madrid	Spain	3	This thesis (unpublished)
	El Piornal	Cáceres	Spain	3	This thesis (unpublished)
	Béjar	Salamanca	Spain	3	This thesis (unpublished)
	Las Ventas c/ Peña Aguilera	Toledo	Spain	3	This thesis (unpublished)
*Ph*2	Leiria	Leiria	Portugal	1, 2, 3	Harris & Sá-Sousa 2001, 2002
	Portalegre	Portalegre	Portugal	1, 2, 3	Harris & Sá-Sousa 2001, 2002
	Beja	Beja	Portugal	1, 2, 3, 4, 5, 6	Harris & Sá-Sousa 2001, 2002, this thesis (article I)
	Marvão	Portalegre	Portugal	1, 2	Harris & Sá-Sousa 2001
	Águeda	Aveiro	Portugal	1, 2, 3	Harris & Sá-Sousa 2001, 2002
	Madrid	Madrid	Spain	1, 3, 4, 5, 6	Harris & Sá-Sousa 2002, this thesis (articles I, IV)
	Tielmes	Madrid	Spain	1, 3, 4, 5, 6	Harris *et al.* 2002b, this thesis (article I)
	Horcajada de la Torre	Cuenca	Spain	6	Renoult 2006
	Albalate de Zorita	Guadalajara	Spain	6	Renoult 2006
	Gualda	Guadalajara	Spain	6	Renoult 2006
	Cazorla	Jaén	Spain	6	Renoult 2006
	Benatae	Jaén	Spain	1, 3, 4, 5, 6	This thesis (article I)
	Castelo de Vide	Portalegre	Portugal	4	This thesis (article IV)
	S. Mamede	Portalegre	Portugal	4	This thesis (article IV)
	Saucedilla	Córdoba	Spain	4	This thesis (article IV)
	Castaño del Robledo	Huelva	Spain	4	This thesis (article IV)
	Casar de Cáceres	Cáceres	Spain	3	This thesis (unpublished)
	S. Pedro de Moel	Leiria	Portugal	3	This thesis (unpublished)
	Castanheira de Pêra	Leiria	Portugal	3	This thesis (unpublished)
	Ourém	Leiria	Portugal	4	This thesis (unpublished)
	Louriçal	Leiria	Portugal	4	This thesis (unpublished)
	Serra do Caldeirão	Faro	Portugal	3	This thesis (unpublished)
	Laminador	Albacete	Spain	3	This thesis (unpublished)
	Olmeda de Cobeta	Guadalajara	Spain	3	This thesis (unpublished)
	Évora	Évora	Portugal	3	This thesis (unpublished)
	Monte Real	Leiria	Portugal	3	This thesis (unpublished)
*Ph*ss	El Saler	Valencia	Spain	1	Castilla *et al.* 1998a,b
	El Grao	Castelló	Spain	1	Castilla *et al.* 1998a,b
	Castellón	Castelló	Spain	1	Castilla *et al.* 1998a,b
	Burjasot	Valencia	Spain	1	Castilla *et al.* 1998a
	Puerto de la Ragua	Granada	Spain	3, 1	Oliverio *et al.* 2000, Harris *et al.* 2002b
	Cuenca	Cuenca	Spain	1, 3, 4	Harris & Sá-Sousa 2002, this thesis (article IV)
	Granada	Granada	Spain	1, 3, 4	Harris & Sá-Sousa 2002, this thesis (article IV)
	Motilla del Palancar	Cuenca	Spain	1, 3, 4	Harris *et al.* 2002b, this thesis (article IV)
	Puebla de D. Fadrique	Granada	Spain	1, 3, 4, 5, 6	Renoult 2006, this thesis (article I)
	Sta. Maria de Nieva	Almería	Spain	4, 6, 3	Renoult 2006, this thesis (article IV, unpublished)
	Alcudia de Crespins	Valencia	Spain	6	Renoult 2006
	Gandia	Valencia	Spain	6	Renoult 2006
	Bétera	Valencia	Spain	6	Renoult 2006
	Nules	Castelló	Spain	6	Renoult 2006
	Catadau	Valencia	Spain	6	Renoult 2006
	Riba-roja de Turia	Valencia	Spain	6	Renoult 2006
	Ondara	Alicante	Spain	6	Renoult 2006
	Calpe	Alicante	Spain	6	Renoult 2006
	Puerto de Burriana	Castelló	Spain	6	Renoult 2006
	San Manuel	Valencia	Spain	6	Renoult 2006
	Bocairent	Valencia	Spain	6	Renoult 2006
	Ibi	Alicante	Spain	6	Renoult 2006
	Cazorla	Jaén	Spain	6	Renoult 2006
	Sierra Nevada	Granada	Spain	4	This thesis (article IV)

Table A1. (cont.)

	Locality	Region	Country	Gene	Reference
Phss	Guadix	Granada	Spain	3	This thesis (unpublished)
	Puebla del Salvador	Cuenca	Spain	3	This thesis (unpublished)
	Valencia	Valencia	Spain	3	This thesis (unpublished)
*Ph*3	Foradada Is.	Castelló	Spain	1	Castilla *et al.* 1998a,b
	El Lobo Is.	Castelló	Spain	1	Castilla *et al.* 1998a,b
	Columbrete Grande Is.	Castelló	Spain	1	Castilla *et al.* 1998a,b, Oliverio *et al.* 2000
	Mancolibre Is.	Castelló	Spain	1	Castilla *et al.* 1998a
	Barcelona	Barcelona	Spain	1, 3, 4	Harris & Sá-Sousa 2002, this thesis (article IV)
	Medinaceli	Soria	Spain	1, 3, 4, 5, 6	Harris & Sá-Sousa 2002, this thesis (article I)
	Tarragona	Tarragona	Spain	1, 3, 4	Harris & Sá-Sousa 2002, this thesis (article IV)
	Girona	Girona	Spain	1, 3	Harris & Sá-Sousa 2002
	Vall d'Alyniá	Lleida	Spain	1, 3	Harris & Sá-Sousa 2002
	Andorra	-	Andorra	1, 3	Harris *et al.* 2002b, Carranza *et al.* 2004a
	Mayorga	Valladolid	Spain	1, 4, 7	Busack *et al.* 2005
	Aguarón	Zaragoza	Spain	6	Renoult 2006
	Almenar de Soria	Soria	Spain	6	Renoult 2006
	Vinaroz	Castelló	Spain	6	Renoult 2006
*Ph*3	Farena	Tarragona	Spain	6	Renoult 2006
	Cap-de-la-Coste	Landes	France	6	Renoult 2006
	Burgos	Burgos	Spain	1, 3, 4, 5, 6	This thesis (article I)
	Getaria	Guipúzcoa	Spain	4	This thesis (article IV)
	Alcolea del Pinar	Guadalajara	Spain	3	This thesis (unpublished)
	Montpellier	Hérault	France	3	This thesis (unpublished)
*Ph*Gal	Puebla de D. Fadrique	Granada	Spain	6, 4, 3	Renoult 2006, this thesis (article IV, unpublished)
	Berja	Almería	Spain	6	Renoult 2006
	Bullas	Murcia	Spain	6, 3	Renoult 2006, this thesis (unpublished)
	Chirivel	Almería	Spain	6, 3	Renoult 2006, this thesis (unpublished)
	Vélez Blanco	Almería	Spain	6	Renoult 2006
	Tiscar	Jaén	Spain	6	Renoult 2006
	Negratin	Granada	Spain	6	Renoult 2006
	Galera	Granada	Spain	1, 3, 4, 5, 6	This thesis (articles I, IV)
*Ph*JS	Jebel Sirwah	Tiznit	Morocco	1, 3, 4, 5, 6	Harris *et al.* 2002b, this thesis (articles I, IV)
	5Km SE Jebel Sirwah	Tiznit	Morocco	4	This thesis (article IV)
*Ph*T+A	N Aïn Draham	Jendouba	Tunisia	1, 3	Harris *et al.* 2002b
	Aïn Draham	Jendouba	Tunisia	1, 3	Harris *et al.* 2002b
	S Aïn Draham	Jendouba	Tunisia	1, 3	Harris *et al.* 2002b
	10Km S Tabarca	Jendouba	Tunisia	1, 3	Harris *et al.* 2002b
	Oued Kébir	Jendouba	Tunisia	1, 3, 4, 5, 6	This thesis (articles I, IV)
	Le Kef	Le Kef	Tunisia	1, 3, 4, 5, 6	This thesis (articles I, IV)
	Jughourta Table	Kasserine	Tunisia	3, 4	This thesis (unpublished)
	Hamla	Batna	Algeria	3, 4	This thesis (unpublished)
	Azazga	Tizi Ouzou	Algeria	3, 4	This thesis (unpublished)

Lineages: *Pb*, *P. bocagei*; *Pc*, *P, carbonelli*; *Pv*, *P. vaucheri*; *Ph*1A, *P. hispanica* type 1A; *Ph*1B, *P. hispanica* type 1B; *Ph2*, *P. hispanica* type 2; *Phss, P. hispanica* sensu stricto; *Ph*3, *P. hispanica* type 3; *Ph*Gal, *P. hispanica* Galera type; *Ph*JS, *P. hispanica* J. Sirwah type; *Ph*T+A, *P. hispanica* Tunisian and Algerian types.
Genes: 1, cyt ochrome *b*; 2, Cytochrome oxidase subunit I; 3, 12S rRNA; 4, NADH dehydrogenase subunit 4; 5, 16S rRNA; 6, control region; 7, NADH dehydrogenase subunit 2.