Biogeografia dos sapos-parteiros (*Alytes* spp.): análise multilocus das relações filogenéticas e da diversificação intraespecífica

Bruno Maia-Carvalho
Doutoramento em Biodiversidade, Genética e Evolução
InBIO
2016

Orientador
Nuno Ferrand de Almeida, Professor Catedrático, Faculdade de Ciências da Universidade do Porto

Coorientadores
Helena Gonçalves, Investigadora de Pós-doutoramento, CIBIO-UP
Iñigo Martínez-Solano, Científico Titular, MNCN-CSIC

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

CIBIO
Centro de Investigação em Biodiversidade e Recursos Genéticos

Biogeografia dos sapos-parteiros (*Alytes* spp.): análise multilocus das relações filogenéticas e da diversificação intraespecífica

Bruno Maia-Carvalho

Orientadores:

Nuno Ferrand de Almeida

Maria Helena Aguiar Gonçalves

Iñigo Martínez-Solano

Porto – Portugal

2016

Nota prévia

Na elaboração desta dissertação, e nos termos do número 2 do Artigo 4º do Regulamento Geral dos Terceiros Ciclos de Estudos da Universidade do Porto e do Artigo 31º do D.L. 74/2006, de 24 de Março, com a nova redação introduzida pelo D.L. 230/2009, de 14 de Setembro, foi efetuado o aproveitamento total de um conjunto coerente de trabalhos de investigação já publicados ou submetidos para publicação em revistas internacionais indexadas e com arbitragem científica, os quais integram alguns dos capítulos da presente tese. Tendo em conta que os referidos trabalhos foram realizados com a colaboração de outros autores, o candidato esclarece que, em todos eles, participou ativamente na conceção, obtenção, análise e discussão de resultados, bem como na elaboração da sua forma publicada.

A presente investigação foi apoiada pelo Ministério da Ciência Tecnologia e Ensino Superior do Governo de Portugal e pela Fundação para a Ciência e Tecnologia (FCT), através da atribuição de uma bolsa de doutoramento SFRH/BD/60305/2009.

Este trabalho foi desenvolvido em:

Esta tese deve ser citada como:

Maia-Carvalho, B (2016) Biogeografia dos sapos-parteiros (*Alytes* spp.): análise multilocus das relações filogenéticas e da diversificação intraespecífica. Tese de Doutoramento, Universidade do Porto, Porto, Portugal.

À minha amada família do Brasil e
à minha família lusa de coração.

Agradecimentos

Ciclos encerrados e muitas emoções vividas ... agora, uma delas será registada: a gratidão. Independente do caminho percorrido, tenho a consciência de que o desenvolvimento e término deste estudo só foi possível graças ao suporte, ajuda e incentivo de muitas pessoas e instituições. Uma somatória de esforços e energias. Se calhar, não é por um acaso que os agradecimentos vêm na frente de qualquer outra secção, mesmo antes do objetivo principal (razão) desta tese. Sou grato pelas oportunidades, grato pelas trocas de conhecimento, por toda ajuda recebida, por todas as experiências vivenciadas. Para todas as pessoas envolvidas neste projeto profissional e de vida: muito obrigado!

Neste sentido, agradeço ao Prof. Dr. Nuno Ferrand de Almeida por ter-me aceitado como aluno de doutoramento. O seu apoio e suporte, durante estes anos, foram fundamentais para a conclusão deste trabalho. Obrigado pela oportunidade.

À Dra. Helena Gonçalves, meu muito obrigado. Agradeço a atenção e disponibilidade em ajudar-me, incluindo a composição do tema da tese, e a possibilidade de dar continuidade na investigação sobre os *Alytes*. Eu acredito que em todos os processos, sejam eles evolutivos ou não, a adaptação sempre é um fator desafiante, uma barreira a ser ultrapassada. Por isso agradeço a sua exigência, preocupação e a sua paciência. Chegamos ao final. E não sou o mesmo. Obrigado por ter contribuído para o meu crescimento pessoal e científico!

Agradeço também ao Dr. Iñigo Martínez-Solano. Mesmo que o tempo de trabalho não tenha sido regular e contínuo, acredite, a sua participação foi fundamental para a minha formação. Obrigado, sempre, pela disponibilidade e vontade em ajudar-me. A estada no *Museu Nacional de Ciências Naturales* (MNCN), em Madrid, ficará para sempre na minha memória.

Agradeço também ao Dr. Fernando Sequeira e ao Dr. Marcelo Vallinoto. Certamente vocês foram pessoas fundamentais para que eu conseguisse chegar até aqui. Para além do conhecimento científico, agradeço as conversas, o incentivo e suporte pessoal.

Aos muitos colegas que coletaram as amostras de *Alytes*, aos colaboradores dos artigos publicados e também aos investigadores que cederam tempo para discussões académicas, muito obrigado. O apoio de vocês foi determinante para o sucesso dessa investigação.

Ao Centro de Testagem Molecular, foi uma grata surpresa trabalhar neste laboratório. Sou infinitamente agradecido por tudo, principalmente pelo acolhimento e o apoio nesta reta final. Foram muitas horas a trabalhar juntos, verificando a leitura dos micros então ... Vacas? Penas? Repetiria tudo novamente! Em nome de Susana Lopes, deixo um abraço carinhoso e saudoso para toda a equipa.

Agradeço também as instituições que permitiram o desenvolvimento deste trabalho. Ao CIBIO, minha instituição de acolhimento, pela oportunidade de trabalhar em um centro composto por pessoas de excelência. Particularmente agradeço à Sara Ferreira, Sandra Rodrigues e Sr. Bernardino pela ajuda, paciência, e, principalmente, nas resoluções de alguns imprevistos. Sarinha, não vou esquecer teu carinho durante a minha chegada em 2008. Obrigado por todo apoio.

O que seria da razão sem a emoção? Ao meu ver, algo sem sentido. A minha razão cresce, à medida que vou vivenciando novas emoções. Fazer um doutoramento é desafiador. Desenvolver esta etapa fora de casa, foi um desafio à parte. Sim, eu senti muitas emoções: novas (e foram muitas ...), antigas, algumas meio escondidas ... E para entender, analisar e tirar um resultado disso tudo, contei com a ajuda de uma equipa muito especial: meus familiares e amigos, estes últimos com uma categoria muito nova: os amigos lusitanos. Agradecer nunca será suficiente. Obrigado por tudo! Aos amigos que estão no Brasil, a distância nunca foi uma barreira. Nossa convivência diminuiu, é um fato. Mas o carinho ultrapassa qualquer mar, qualquer temporalidade. Aos amigos *de novo*, percebi o quanto são caros e especiais.

Luciana Castro, obrigado por ter estado tão perto, sempre presente. *Mi casa* es *tu casa*. Dividir contigo esta fase foi algo surreal. As lembranças são tão ativas, e tantas, que bloqueiam a ação da escrita. Simplesmente obrigado! Liane Finotelo, obrigado pela tua forma simples de ser, tua força interior. A tríade seria perfeita se um dia viesses visitar-nos. Vocês me ajudaram a plantar a semente do meu amanhã.

Tuliana Brunes, obrigado pela oportunidade de conhecer-te! Aprendi muito contigo, não imaginas o quanto. Cláudia Almendra, Ricardo Lopes, Mafalda Barata, Sónia Ferreira, Ricardo Severino, Pedro Souza e Ana Torres, a vida em Vila do Conde fez mais sentido por ter vocês sempre por perto. Obrigado pelas conversas, os muitos momentos de

descontração. Obrigado por terem mostrado um pouco de Portugal e fazer com que me sentisse em casa. Sara João, Pamela Puppo e Susana Lopes, o CTM foi só meio. Obrigado mesmo pela parceria e amizade destes anos. E que venham muitos outros. Sérgio Filipe e Olímpia Ribeiro, obrigado pela vossa atenção, paciência e todas as formas de ajuda. Os convívios, muitas vezes intensos, são memórias impressas para sempre. Às famílias Fernandes Ferreira e Vila Real Ribeiro, obrigado pelo acolhimento, pelo carinho. Serei eternamente grato.

E por último e, sim, muito importante, a minha família de raiz: Márcia, Hamilton e Thaís. Nós constatamos que o amor ultrapassa barreiras, é imensurável. Não tem tempo, não tem território. A vida terminou por nos mostrar caminhos um pouco diferentes, onde a nossa noção de convívio foi ressignificada. Entretanto, estes caminhos nos ensinaram muito. Agradeço todo apoio, incentivo e empenho. Amo vocês!

Resumo

Os sapos-parteiros (*Alytes* spp.) englobam um conjunto de cinco espécies amplamente distribuídas por toda a Bacia Mediterrânica Ocidental. Três subgéneros são atualmente reconhecidos: *Ammoryctys* (*Alytes cisternasii*), *Baleaphryne* (*Alytes dickhilleni*, *Alytes maurus* e *Alytes muletensis*) e *Alytes sensu stricto* (*Alytes obstetricans*). Nas duas últimas décadas, os estudos evolutivos do género *Alytes,* ao nível inter e intraespecífico, demonstraram: *i*) complexas relações filogenéticas entre as espécies, resultantes de retenção de polimorfismo ancestral e contato secundário, e *ii*) significativa diferenciação genética ao nível intraespecífico para *A. cisternasii* e *A. obstetricans*. A observação de padrões genéticos estruturados associados a diferenciações morfológicas permitiu ainda o reconhecimento de quatro subespécies dentro de *A. obstetricans*: *A. o. obstetricans*, *A. o. boscai*, *A. o. pertinax* e *A. o. almogavarii*. Contudo, mesmo com a aplicação de diferentes abordagens metodológicas, a biogeografia do género *Alytes* ainda não se encontra completamente esclarecida. Neste sentido, o presente trabalho teve como objetivo central elucidar os fatores históricos e ambientais utilizando dados moleculares, provenientes de diferentes marcadores genéticos, assim como dados ecológicos e variados métodos de análise, nomeadamente bayesianos e modelação de nicho ecológico.

Um aspecto interessante na história evolutiva dos sapos-parteiros é ilustrado pela rápida radiação das três espécies incluídas no clado *Baleaphryne* e a possibilidade de fluxo génico interespecífico envolvendo *A. o. almogavarii*. Assim, o primeiro objetivo desta tese consistiu na reanálise das relações filogenéticas entre as espécies do género *Alytes* (incluindo todas as subespécies reconhecidas de *A. obstetricans*), utilizando dados multilocus (fragmentos do DNA mitocondrial – mtDNA - e quatro genealogias nucleares - nDNA) e métodos de análise coalescente (árvore de espécies). Os resultados dos dados concatenados, assim como da árvore de espécies, corroboram estudos anteriores em que *Alytes cisternasii* se posiciona como grupo irmão de *A. obstetricans* + *Baleaphryne*. Dentro do clado *Baleaphryne*, foi possível demonstrar que *A. muletensis* (endémico da ilha de Maiorca) é grupo irmão de *A. dickhilleni* (endémico do sudeste Ibérico) e *A. maurus* (endémico do norte de África), sendo a radiação deste clado estimada à cerca de 5,3 milhões de anos atrás. Este cenário é consistente com evidências paleogeológicas relacionadas com a fragmentação do Maciço Bético-Rifenho e posterior abertura do Estreito de Gibraltar no final da crise Messiniense. Ao nível intraespecífico, duas linhagens divergentes de *A. o. boscai* estão bem suportadas (uma localizada a norte e outra a sul do Rio Douro) e as linhagens de *A. o. pertinax* e *A. o. obstetricans* são recuperadas como grupos monofiléticos. No entanto a posição filogenética de *A. o. almogavarii* permanece inconclusiva.

Estudos recentes demonstraram elevados níveis de diversidade e diferenciação genética entre as subespécies de *A. obstetricans* e sugerem que as mudanças climáticas e ambientais do Pleistoceno moldaram a história evolutiva desta espécie. Diante deste contexto, desenvolveu-se uma abordagem filogeográfica para analisar o padrão de distribuição da diversidade genética da espécie e, mais especificamente, entender as dinâmicas populacionais ocorridas durante o Pleistoceno (fragmentação e diferenciação) e no período pós-glaciar (expansão e estabelecimento de zonas de contacto secundário entre linhagens). Com uma amostragem abrangente, incluindo todas as subespécies descritas, a análise multilocus (mtDNA e nDNA) revelou relações históricas mais complexas do que inicialmente se hipotetizava. Os resultados evidenciaram seis haplogrupos mitocondriais divergentes e geograficamente estruturados (de A até F), incluindo um haplogrupo "fantasma" sem representação ao nível nuclear. Por outro lado, os marcadores nucleares não recuperaram nenhum padrão monofilético, provavelmente reflexo da retenção de polimorfismo ancestral. As linhagens correspondentes às subspécies de *A. o. almogavarii*, *A. o. pertinax*, *A. o. obstetricans* e *A. o. boscai* evoluíram independentemente (corroborando o modelo de "refúgios dentro dos refúgios") e, de facto, as oscilações climáticas do Pleistoceno podem estar relacionadas com esta diversificação, pois o intervalo de tempo obtido para a separação destas linhagens está compreendido entre 2,6 milhões e 450 mil anos atrás. As linhagens de *A. o. pertinax* e *A. o. boscai* são as mais amplamente distribuídas na Península Ibérica. Dois grupos divergentes foram descritos para *A. o. boscai*, estando a sua origem relacionada com a existência de dois refúgios: um a norte e outro a sul do rio Douro. Por outro lado, um único refúgio, no norte Ibérico, é proposto para *A. o. obstetricans*, do qual expandiu posteriormente um subgrupo para o norte da Europa, atravessando o extremo oeste dos Pirinéus. As populações francesas e alemãs analisadas revelaram um único haplótipo de mtDNA - sugestivo de forte evento de *bottleneck*. Entretanto, são as populações de *A. o. almogavarii* que possuem a distribuição mais restrita: ao norte do Rio Ebro, nas Cordilheiras dos Pirinéus, e uma pequena região no sudeste de França. Este *taxon* apresenta níveis singulares de diversidade genética e partilha poucos haplótipos com as outras linhagens descritas, cenário sugestivo de especiação incipiente.

Após a deteção de um forte padrão filogeográfico com base no mtDNA, uma valiação detalhada ao nível nuclear tornou-se crucial. Assim, delimitaram-se os últimos objetivos desta tese: descrição da estrutura populacional de *A. obstetricans* e das áreas de distribuição das linhagens, assim como a identificação das áreas de contato secundário e deteção dos fatores exógenos, nomeadamente ambientais, potencialmente envolvidos no processo de diversificação do sapo-parteiro-comum. Para este fim, um extensivo número de amostras (965 indivíduos), recolhidas principalmente na Península Ibérica, foi genotipado, usando-se um painel de 17 microssatélites, e analisado por métodos Bayesianos de

agrupamento. Estas análises revelaram sete grupos distintos, concordantes com os dados mitocondriais. No geral, estes grupos apresentam elevada diversidade genética apesar da deteção de eventos históricos de redução populacional. A combinação dos resultados individuais dos testes de atribuição (*assignment test*) com análises de modelação de nicho ecológico e sobreposição de nicho demonstraram: i) fortes evidências de divergência de nicho e ii) áreas de sobreposições estreitas entre todos os grupos de *A. obstetricans*, corroborando os mesmos limites de distribuição indicados somente do ponto de vista genético. A significativa divergência de nichos ecológicos entre as diferentes linhagens, numa área geográfica relativamente pequena, destaca o potencial da heterogeneidade topográfica e/ou climática no estabelecimento de padrões de diferenciação genética entre as populações e o potencial papel da seleção na diversificação intraespecífica de *A. obstetricans*.

Os resultados obtidos no presente trabalho ressaltam a importância da aplicação de abordagens integradas no sentido de reconstituir cenários biogeográficos complexos, incluindo não somente o uso de diferentes marcadores genéticos, como também, diferentes métodos de análise. Os fatores históricos e ambientais envolvidos no processo de diversificação das linhagens do género *Alytes* podem também ter influenciado a biogeografia de diversas outras espécies da Bacia Mediterrânica Ocidental. Neste sentido, a condução de novos estudos ao nível intraespecífico pode contribuir para o esclarecimento, dentro dos processos microevolutivos, o papel dos factores ambientais na diversificação biológica e na manutenção dos padrões de variabilidade genética ao nível intraespecífico, ou seja, o papel da seleção nas fases iniciais de divergência.

Palavras chave: *Alytes*, *Baleaphryne*, *A. obstetricans*, Biogeografia, Península Ibérica, Dados multilocus, Filogenia, Árvore de espécies, Rápida radiação, Filogeografia, Estrutura populacional, Diversidade genética, Modelação de nicho ecológico, Divergência de nicho, Diferenciação de linhagens e Especiação.

Summary

The midwife toads (*Alytes* spp.) comprise five species widely distributed throughout the Western Mediterranean Basin. Three subgenera are currently recognized: *Ammoryctys* (*Alytes cisternasii*), *Baleaphryne* (*Alytes dickhilleni, Alytes maurus* and *Alytes muletensis*) and *Alytes sensu stricto* (*Alytes obstetricans*). During the last decades, inter- and intraspecific evolutionary studies have described: i) complex phylogenetic relationships between species due to the presence of ancestral polymorphism, and the establishment of secondary contact between them, and ii) high genetic differentiation within populations of *A. cisternasii* and *A. obstetricans*. Particularly, the observation of structured patterns of genetic diversity associated with morphological differentiation has allowed the recognition of four subspecies within *A. obstetricans*: *A. o. obstetricans*, *A. o. boscai*, *A. o. pertinax* and *A. o. almogavarii*. However, the biogeographic scenario for this genus is not yet consensual. Thus, this study investigates the historical and environmental factors that may have shaped the diversification of the genus *Alytes* through an integrated analysis that combines the use of different molecular markers under phylogenetic, phylogeographical and ecological niche modeling approaches.

An interesting aspect of the evolution of midwife toads is the disjunct and highly restricted geographical ranges of the three species included in the *Baleaphryne* clade, and the possibility of interspecific gene flow with *A. o. almogavarii*. Thus, the first aim of this thesis was to re-address the phylogeny of *Alytes,* including all subspecies of *A. obstetricans*, using a multilocus dataset, mitochondrial DNA (mtDNA) and four nuclear genealogies (nDNA), and coalescent analysis. Both concatenation and species tree analyses suggest that *A. muletensis*, endemic to the Balearic island of Mallorca, is the sister taxon to a clade comprising the southeastern Iberian endemic, *A. dickhilleni*, and the North African, *A. maurus*. The *Baleaphryne* radiation is estimated to have occurred 5.3 Ma. This scenario is consistent with paleogeological evidence associated with the fragmentation of the Betic-Rifean Massif, followed by the opening of the Strait of Gibraltar. On the other hand, *A. o. almogavarii* was placed within *A. obstetricans*, recovering monophyly of this clade with high support. Furthermore, two divergent lineages of *A. o. boscai* were distinguished, geographically separated by the Douro River.

Regarding intraspecific diversification of *A. obstetricans*, recent studies have shown high levels of genetic diversity and differentiation between lineages. Given this context, a phylogeographic approach was developed in order to infer the evolutionary processes that shaped current patterns of genetic diversity and population subdivision. Coalescent multilocus analysis revealed six geographically structured mtDNA haplogroups depicting population lineages and varying levels of admixture in the four nuclear genealogies. The

major lineages evolved independently after diverging from a common ancestor, about 2.6 Ma. In fact, climatic and environmental changes during the Pleistocene seem to have shaped the diversification history of *A. obstetricans*. Survival of populations in allopatric refugia through the Ice Ages supports the ‘‘refugia-within-refugia” scenario for the Iberian Peninsula. Both *A. o. pertinax* and *A. o. boscai* are widespread and genetically diverse in Iberia. The latter comprises two divergent lineages located north and south of the Douro river, and showing a long independent history. *Alytes o. obstetricans* expanded northward from a northern Iberian refugium through the western Pyrenees, leaving a bottleneck signal with a single mtDNA haplotype in the populations from SW France to Germany. Finally, *A. o. almogavarii* populations have the most restricted distribution in the north-eastern corner of Iberia, north of the Ebro River, with additional populations in south-eastern France. This taxon exhibits unparalleled levels of genetic diversity, sharing few haplotypes with other lineages, suggesting a process of incipient speciation.

After detecting a strong mtDNA phylogeographic pattern, a detailed assessment of the genetic variability under a nuclear perspective becomes crucial. In order to delimitate subspecies ranges, identify secondary contact zones, describe population structure, and detecting exogenous factors, such as environmental factors, potentially involved in *A. obstetricans* lineage diversification process 965 individuals (mainly collected in the Iberian Peninsula) were genotyped using 17 microsatellites and analyzed with Bayesian clustering analysis. The results reveal seven distinct clusters, clearly related to the mitochondrial haplogroups. In general, these clusters have high genetic diversity even though there is evidence of historical events of effective population size reduction. Moreover, the combination of the results from the structure assignment together with tests of niche overlap, equivalency and similarity, as well as ecological niche-based models, showed strong evidence of niche divergence and limited overlap areas of suitability where the genetic range limits are coincident with niche limits. These results suggest that the heterogeneous topography of the Iberian Peninsula and/or climate can be associated to the establishment of genetic diversity patterns. Also, the adaptation to different ecological niches across current lineage distributions can be an important force driving intraspecific lineage diversification in *A. obstetricans*.

In a broader sense, the results obtained in the present study highlight the importance of an integrated approach in order to recover complex biogeographic scenarios, including the use of different genetic markers as well as methods of analysis. The historical and environmental factors involved in the process of lineage diversification in genus *Alytes* may also have influenced the biogeography of several other species in the Western Mediterranean Basin. Therefore, other investigations at an intraspecific level are crucial to

understand the microevolutionary processes such as the potential role of environmental factors in biological diversification (selection) during the initial stages of divergence.

Keywords: *Alytes*, *Baleaphryne*, *A. obstetricans*, Biogeography, Iberian Peninsula, Multilocus data, Phylogeny, Species tree, Rapid radiation, Phylogeography, Population structure, Genetic diversity, Ecological niche modeling, Niche divergence, Lineage divergence and Speciation.

Índice

Lista de Tabelas

CAPÍTULO 2

Artigo I

CAPÍTULO 3

Artigo II

CAPÍTULO 4

Artigo III

Lista de Figuras

CAPÍTULO 1

CAPÍTULO 2

Artigo I

CAPÍTULO 5

Capítulo 1

Introdução Geral

A história evolutiva dos organismos é diretamente influenciada por processos fisiográficos e geológicos. A compreensão destes processos e do modo como afetam a distribuição geográfica e a diversificação biológica, tanto no passado quanto no presente, são alguns dos objetivos da Biogeografia (Lomolino *et al.* 2010). Com o aparecimento e o desenvolvimento da biologia molecular, mudanças no campo da biogeografia histórica e evolutiva foram suscitadas com base em análises de relações filogenéticas, padrões de estrutura genética e mecanismos de dinâmica populacional entre as espécies, proporcionando descrições detalhadas de processos evolutivos e microevolutivos com suporte estatístico (por exemplo: Zeisset & Beebee 2008, San Mauro 2010, Hewitt 2011). O conhecimento ao nível molecular tem-se revelado também importante para os campos da Sistemática e Taxonomia, auxiliando na revisão e reclassificação de grupos de organismos, principalmente naqueles em que a diferenciação morfológica é difícil (ver Pyron 2015). Particularmente em vertebrados, a análise de dados moleculares tem possibilitado a elucidação de processos de extinção em massa em mamíferos (Koch & Barnosky 2006), rotas de migração em aves (Pérez-Tris *et al.* 2004, Wink 2006), processos de radiação adaptativa em peixes (Muschick *et al.* 2012) e questões mais atuais de instabilidade populacional em répteis e anfíbios devido às alterações climáticas, fragmentação e perda de habitats (Shaffer *et al.* 2015).

Em relação aos estudos evolutivos, características como baixa capacidade de dispersão, isolamento populacional e especificidades ecológicas tornam alguns organismos bons modelos de investigação. Observadas na herpetofauna, estas características, quando associadas a fortes padrões de estrutura populacional, refletem-se em importantes sinais filogeográficos, auxiliando na reconstituição de cenários biogeográficos e histórias evolutivas. Neste contexto, geralmente os anfíbios são organismos adequados para a investigação dos fatores históricos envolvidos na distribuição da diversidade biológica e sua diversificação (Zeisset & Beebee 2008, Reilly & Wake 2015). Adicionalmente, as populações de anfíbios são facilmente afetadas por alterações nos seus habitats, sejam estas de caráter antropogénico ou natural (Stuart *et al.* 2004, Allentoft & O'Brien 2010). Na sequência de um cenário global de alterações climáticas, tem-se verificado um declínio sistemático e acentuado das populações de anfíbios, estimando-se que 40% destas espécies estejam classificadas como ameaçadas ou em risco de extinção (Stuart *et al.* 2004, Hoffmann *et al.* 2010). Eventos vicariantes ocorridos no passado podem também influenciar a diversidade das populações contemporâneas, prejudicando a viabilidade de uma determinada espécie (Dufresnes *et al.* 2013, Dufresnes & Perrin 2015). Por exemplo, as glaciações do Pleistoceno em regiões de clima temperado têm sido associadas a processos de diversificação e extinção de diversas espécies de anfíbios (ver Gómez & Lunt 2007, Hewitt 2011). Assim, estudos integrados que descrevam os padrões de diversidade biológica, bem

como os processos de diversificação e a história evolutiva das espécies de anfíbios, são fundamentais, não somente para o resgate destes processos como também para a elaboração de estratégias de conservação que visam a preservação deste grupo de vertebrados.

1.1 Diversificação biológica

O processo evolutivo, por ser dinâmico e contínuo, deixa evidências em diversas características biológicas, nomeadamente fenotípicas, moleculares e comportamentais. Se, por um lado, os eventos de mutação e recombinação são responsáveis por gerar variabilidade dentro de uma população, por outro, a seleção natural, a deriva genética e o fluxo génico moldam a diferenciação desta variabilidade. Quando acumulada e transmitida à descendência, esta diferenciação promove divergência entre as linhagens que descendem de um ancestral comum, processo denominado de especiação (Coyne & Orr 2004). A compreensão dos fatores que podem conduzir à diversificação das espécies sempre foi tema de grande interesse da biologia evolutiva, destacando-se dentre eles isolamento reprodutivo, seleção sexual, *reinforcement* e divergência ecológica (ver The Marie Curie Speciation Network 2012, Etges 2014).

Os estudos de biologia molecular, baseados em níveis de divergência genética, têm possibilitado o esclarecimento de processos evolutivos e auxiliado na identificação de espécies e delimitação de linhagens evolutivas. Contudo, a análise do grau de divergência pode ser complexa, dificultando a sua utilização como critério objetivo para o diagnóstico de espécie, principalmente quando as populações divergentes cumprem funções ecológicas semelhantes e mantêm contacto secundário com a presença de fluxo génico e migração (ver Pinho & Hey 2010, Hey & Pinho 2012, The Marie Curie Speciation Network 2012). Apesar da complexidade, o surgimento de novos e melhorados métodos de análise têm conduzido a mudanças de paradigma nos campos da sistemática e taxonomia, promovendo um intenso debate sobre o conceito de espécie (Mayden 1997, Avise 2000, de Queiroz 2007).

O conceito biológico clássico define espécie como um grupo de organismos que se reproduzem entre si (ou em que existe potencial de reprodução) e geram descendentes viáveis, estando este grupo isolado reprodutivamente de outros grupos (Wright 1940, Mayr 1942, Dobzhansky 1950). Esta definição perdura até hoje, mantendo-se associada a uma escola clássica de identificação e descrição de espécies fundamentada em evidências morfológicas (Cronquist 1978, Dayrat 2005). Entretanto o conceito clássico tem sido alvo de

inúmeras críticas, dentre elas a sua inaplicabilidade a organismos de reprodução assexuada. O desenvolvimento de diversas áreas do conhecimento biológico (morfologia, fisiologia, genética, comportamento, reprodução, entre outras) culminou na elaboração de diferentes conceitos de espécie, perpetuando discussões e conflitos entre as várias definições (de Queiroz 2007). Existem mais de 22 conceitos, mas a maioria não explicita claramente em que ponto da trajetória evolutiva podem ser ancorados (Wheller & Meier 2002, de Queiroz 2007), provavelmente porque no decorrer do processo de especiação as evidências evolutivas não se estabelecem de forma ordenada e temporal regular.

De acordo com Mayden (1997), apenas o conceito evolutivo seria adequado para a definição de espécie (*Evolutionary Species Concept* – ESC), uma vez que faz referência a um(a) estado (condição) mais basilar. A definição atual de ESC consiste num grupo de organismos que mantém a sua identidade em relação a outros, através do tempo e espaço, tendo sua própria independência evolutiva e tendências históricas (Wiley & Mayden 2000). Mayden (1997) sugere que todos os outros conceitos de espécie serviriam como critérios secundários para a definição e, segundo de Queiroz (2007), funcionariam como linhas de corte dentro do processo de especiação (ver o esquema da Figura 1). Assim, quando aplicados de forma isolada sobre o processo contínuo, dinâmico e complexo de evolução, os distintos conceitos de espécie podem tornar-se insuficientes ou inadequados, gerando discordâncias.

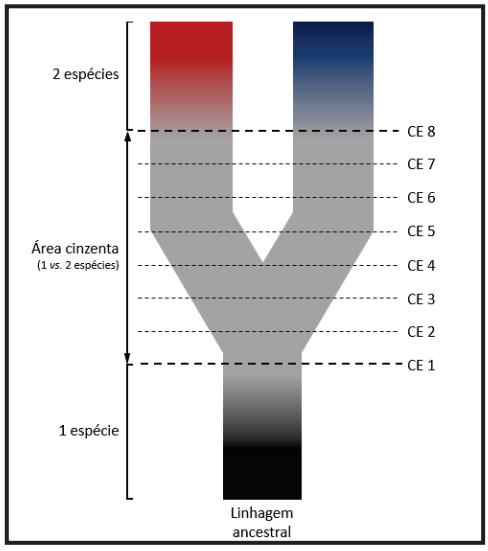


Figura 1. Esquema do processo de separação e divergência de linhagens evolutivas (especiação) baseado em de Queiroz (2007). Os conceitos de espécie (CE), fundamentados em diferentes evidências biológicas (por exemplo, diferenciação ecológica, morfológica, isolamento reprodutivo, monofilia recíproca ...), são assinalados no decorrer do processo de separação da linhagem ancestral. Enquanto que na área a cinzento espera-se a identificação de diferentes linhagens evolutivas e o uso de classificações taxonómicas ao nível de subespécies (provavelmente desde o início da fase de divergência); antes e depois desta área espera-se que os diversos CE concordem no reconhecimento de uma ou duas espécies.

A identificação de linhagens divergentes pode ser efetuada em qualquer ponto de um contínuo evolutivo, variando desde a panmixia, passando por vários níveis crescentes de distância e isolamento genético e chegando, por fim, ao isolamento reprodutivo (Dufresnes *et al.* 2014). A diferenciação destas linhagens é subtil nos processos de divergência recente, pois grande parte das características biológicas ainda são fortemente partilhadas, dentre

elas as morfológicas e ecológicas. Por isso, é na fase inicial do processo de especiação que os conflitos conceptuais são mais evidentes. Em contrapartida, evidências de divergência genética significativas podem ser detetadas com base em dados moleculares, revelando, por exemplo, estruturação ao nível intraespecífico (exemplos em Díaz-Rodríguez *et al.* 2015, Reilly & Wake 2015). Segundo Wood *et al.* (2014), a perceção de que as espécies reúnem linhagens evolutivas em diferentes estados de especiação é central para a discussão de diversidade intraespecífica. Neste sentido, níveis subespecíficos de classificação podem ser usados tanto para denotar uma possível variação fenotípica numa determinada espécie e/ou contemplar níveis hierárquicos de diferenciação genética entre linhagens (ver Wilson & Brown 1953, Dufresnes *et al.* 2014, Wood *et al.* 2014, Schield *et al.* 2015). O conceito de subespécie surgiu para a distinção de grupos de indivíduos com base na morfologia, genética e distribuição geográfica, possibilitando um refinamento da classificação taxonómica (Mayr 1942, 1982; Wilson & Brown 1953).

Mayr (1942) define subespécie como um grupo de populações fenotipicamente semelhantes que habitam uma subdivisão geográfica do intervalo de distribuição de uma espécie, sendo este grupo diferente de outras subdivisões. No entanto, diversos estudos evidenciaram desacordos entre dados genéticos e morfológicos em vários grupos taxonómicos (ver discussões em Zink 2004 – aves, Haig *et al.* 2006 – anfíbios, Torstrom *et al.* 2014 - répteis), fazendo com que dúvidas e críticas fossem levantadas sobre a validade deste nível de classificação (Phillimore & Owens 2006). Os estudos moleculares contribuíram para que algumas subespécies fossem excluídas, associadas, criadas e até mesmo elevadas ao patamar de espécies (por exemplo: Leaché & Reeder 2002, Podnar *et al.* 2004, Bryson *et al.* 2007, Guicking *et al.* 2008, Daniels *et al.* 2010). O criticismo à classificação de subespécies impulsionou o aperfeiçoamento das técnicas de obtenção de dados moleculares e das análises das relações filogenéticas, de forma a validar e suportar estes grupos como entidades naturais (Wood *et al.* 2014). Embora a aplicabilidade da classificação, significado evolutivo e importância taxonómica sejam tópicos de controvérsia, alguns autores argumentam que as subespécies podem representar unidades úteis para políticas de conservação e definição de estratégias de gestão (ver Phillimore & Owens 2006, Allentoft & O'Brien 2010, Braby *et al.* 2012, Dufresnes *et al.* 2013, Wood *et al.* 2014). Apesar de espécie ser a unidade fundamental de análise na ecologia, evolução e biologia da conservação (de Queiroz 2007), o conhecimento e a manutenção da diversidade intraespecífica são cruciais para a persistência das populações e do potencial adaptativo das espécies frente a novas exigências ambientais (Frankham 2002, 2005). Sob esta perspetiva, a classificação de um determinado grupo populacional como subespécie pode ser substituído por unidade evolutiva significativa (*Evolutionarily Significant Unit* - ESU, Moritz, 1994).

No decorrer do processo de divergência, alguns fatores podem alterar os níveis de diferenciação genética, nomeadamente o fluxo génico (no caso de especiações simpátricas e parapátricas, em que as populações não estão isoladas geograficamente), a deriva genética e o tamanho do efetivo populacional (no caso de especiações alopátricas, em que as populações se encontram isoladas geograficamente) (ver Frankham *et al.* 2010, Lomolino *et al.* 2010). Neste contexto, é esperado que populações parapátricas reflitam divergência adaptativa em que as diferenças fenotípicas são resultado da seleção natural, enquanto populações alopátricas tenham tendência a apresentar tanto divergência adaptativa quanto neutra, sendo a estrutura filogeográfica impulsionada principalmente pelos efeitos da deriva genética (Braby *et al.* 2012). Assim, populações alopátricas podem representar mais facilmente processos de especiação independentes. Neste caso, em particular, se características morfológicas e ecológicas ainda não evoluíram ou não se tornaram evidentes, as populações (linhagens evolutivas) podem ser classificadas como crípticas na existência de significativa diferenciação genética e/ou isolamento reprodutivo. Organismos que são especialistas ecológicos ou que têm baixa capacidade de dispersão, são particularmente bons candidatos para apresentarem níveis crípticos de biodiversidade devido à sua tendência para formar populações geneticamente isoladas (Reilly & Wake 2015). Isto pode ser confirmado através da análise combinada de diferentes marcadores moleculares neutros, tais como aloenzimas, DNA mitocondrial (mtDNA) e microssatélites, com métodos coalescentes e Bayesianos de análise. As análises combinadas possibilitam também a deteção dos fatores e da forma como estes influenciam os processos microevolutivos, assim como a identificação e delimitação da diversidade biológica desconhecida ou sub-representada (O'Meara 2010, Yang & Rannala 2010, Ence & Carstens 2011).

1.1.1 Filogenia

O conhecimento das relações ancestrais entre um conjunto de espécies, linhagens evolutivas ou genes é condição *sine qua non* para a elucidação dos processos históricos envolvidos na sua origem e diversificação, assim como para descrição de cenários biogeográficos. A Filogenia, área da ciência que estuda estas relações ancestrais, surgiu em termos conceptuais com os estudos de Darwin em concomitância com o conceito de ancestralidade, o qual tem como pressuposto que todas as espécies possuem um ancestral comum, mesmo que este não possa ser representado diretamente ou através do registo fóssil (Yang & Rannala 2012). Foi somente em 1866 que Haeckel implementou o termo

"filogenia" e, desde então, a representação das relações evolutivas entre grupos de organismos tem passado por diversos desdobramentos (Haeckel 1866), dentre eles a sistemática filogenética proposta por Willi Henning em 1966. Segundo Henning, os organismos que partilham condições derivadas de caracteres (ou apomórficas) poderiam ser considerados descendentes de uma espécie ancestral, em que a condição primitiva (ou plesiomórfica) passou a condição derivada (Dayrat 2003).

Desde a implementação da Filogenia, diferentes informações biológicas têm vindo a ser incorporadas nas análises filogenéticas (características morfológicas, fisiológicas, ecológicas e, mais recentemente, moleculares), bem como diferentes métodos de análise (por exemplo, distâncias genéticas e caracteres discretos). A combinação de dados moleculares e análises coalescentes, através de inferências Bayesianas, possibilitaram um rápido avanço na reconstituição das relações evolutivas complexas, conferindo à Filogenia um papel de destaque no campo da sistemática e taxonomia, assim como noutras áreas de investigação: estudo de famílias de genes, história demográfica, evolução de organismos patogénicos e determinação de unidades evolutivas significativas (ver Webb *et al.* 2002, Pybus & Rambaut 2009, Boussau & Daubin 2010).

1.1.1.1 Coalescência e árvore de espécies

A teoria matemática da coalescência foi originalmente desenvolvida no início de 1980 por John Kingman. Esta teoria permite estimar a probabilidade (incluindo variação temporal) de uma ancestralidade comum para todos os alelos de uma determinada genealogia, partindo-se de uma amostragem reduzida que representa conceptualmente uma população maior (Kingman 1982, 2000). Nesta análise é possível estimar o ancestral comum mais recente (*Most Recent Common Ancestor* - MRCA), uma vez que todos os alelos coalescem no passado (Figura 2). Na sua forma mais simplista, esta teoria assume ausência de recombinação, seleção natural, e fluxo génico ou estrutura populacional. No entanto, avanços na sua formulação já incluem praticamente qualquer modelo evolutivo ou demográfico de análise genética populacional.

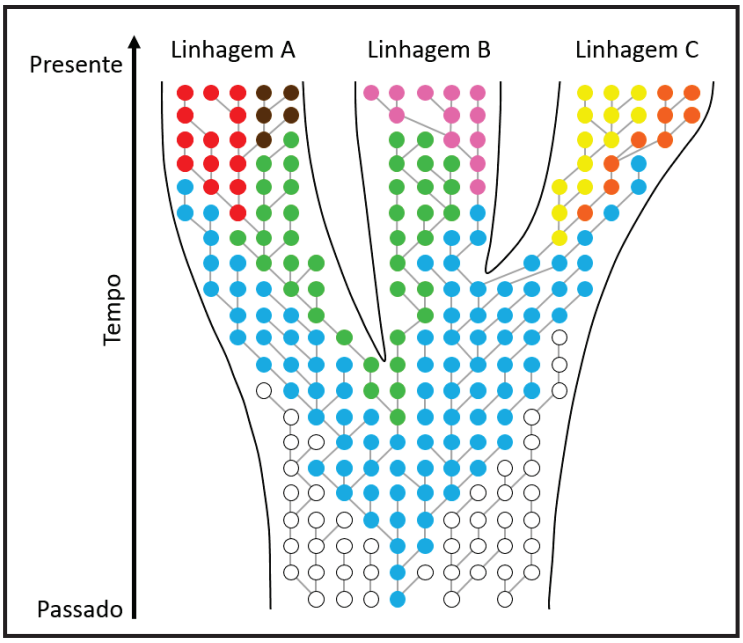


Figura 2. Genealogia Ilustrativa do fundamento da teoria da coalescência. Neste exemplo é possível observar, em três linhagens diferentes, a presença de cinco alelos (tempo presente) que coalescem no passado num único alelo.

Um aspeto importante a destacar é que o processo de coalescência é estocástico. Assim, genealogias analisadas individualmente podem gerar árvores com sinais filogenéticos distintos (árvores de genes – Figura 3). Eventos de incongruências/discordâncias entre diferentes árvores de genes têm sido amplamente documentados em muitos grupos de organismos e as interpretações evolutivas alternativas dificultam, por vezes, a reconstrução da história filogenética entre linhagens (Degnan & Rosenberg 2009, Nakhleh 2013, Leaché *et al.* 2014). Entre as causas mais comuns de discordância estão a diferenciação incompleta de linhagens (*Incomplete Lineage Sorting* - ILS), hibridação recente/ancestral, fluxo génico e politomias (Belfiore *et al.* 2008, Kubatko *et al.* 2011, Stanley *et al.* 2011, Corl & Ellegren 2013).

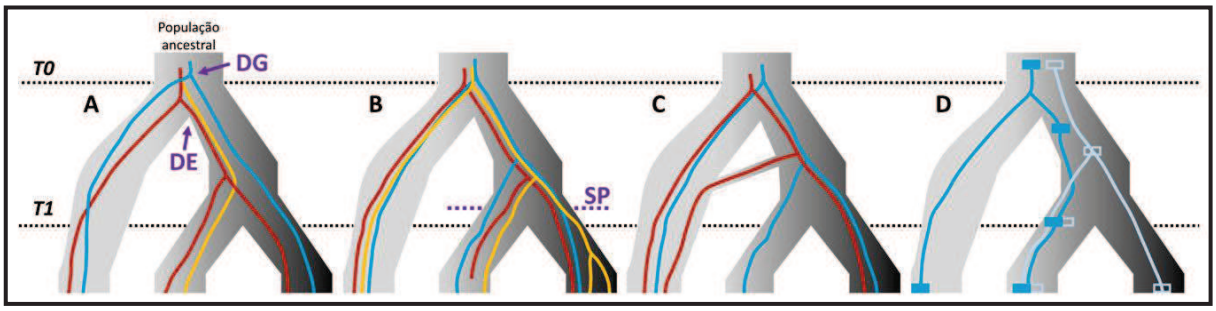


Figura 3. Incongruências entre árvores de genes e árvore de espécies (adaptado de Degnan & Rosenberg 2009). T0: tempo 0 (inicial), T1: tempo 1 (decorrido). A: as genealogias podem ter taxas de evolução desiguais, DE: divergência de espécies, DG: divergência genética. B: a perda de alelos e a duplicação de genes são eventos comuns, SP: separação populacional. C: fluxo génico pode ocorrer entre linhagens mesmo depois da separação. D: eventos de recombinação entre regiões de um mesmo *locus* podem também conduzir a incongruências nas relações filogenéticas entre populações e na história evolutiva das genealogias.

No sentido de reduzir o efeito estocástico deste processo, foram desenvolvidas abordagens de análise que combinam dados moleculares multilocus (Brito & Edwards 2009) e inferências Bayesianas de coalescência (Yang & Rannala 2010). Este fato conduziu a uma mudança de foco nos estudos filogenéticos, inicialmente pautados pela análise individual das árvores de genes, para uma abordagem mais global utilizando-se árvores de espécies. Este método de vanguarda, que avalia as incertezas provenientes das árvores de genes, tem sido eficaz na delimitação de linhagens evolutivas, na estimativa do tempo de divergência entre linhagens recentemente diferenciadas, assim como contribuiu para o desenvolvimento das análises filogeográficas (ver Edwards 2009, O'Meara 2010, 2012; Yang & Rannala 2010, Ence & Carstens 2011, Carstens *et al.* 2013, Edwards & Knowles 2014).

1.1.2 Filogeografia

Os padrões de divergência genética observados ao nível intraespecífico e entre espécies próximas são o resultado da combinação das histórias evolutivas (tempo) com os fatores que limitam a dispersão das populações (espaço). A deteção e descrição desses padrões são importantes na inferência de processos microevolutivos. Com esta finalidade, surgiu em finais da década de 80 uma nova área multidisciplinar, a Filogeografia (Avise *et al.* 1987), que incorpora conhecimentos científicos de diversos domínios: Filogenia, Genética Populacional, Ecologia, Paleontologia, entre outros (Avise 2000, 2009). Originalmente concebida como elo de ligação entre a Filogenia e a Genética Populacional, a Filogeografia cresceu como disciplina fazendo rápidos avanços nos seus métodos de análise através da incorporação da teoria da coalescência (ver Hickerson *et al.* 2010).

Devido à natureza estocástica do processo genealógico e à possível ocorrência de fenómenos seletivos, diferentes marcadores (DNA nuclear, DNA mitocondrial e microssatélites) devem ser combinados em estudos filogeográficos e, sempre que possível, aumentado o número de *loci* investigados (Edwards & Beerli 2000, Garrick *et al.* 2011). Desta forma, possíveis problemas relacionados com as estimativas do tempo de divergência e presença de polimorfismo ancestral são minimizados ou completamente resolvidos (Knowles 2009, Hickerson *et al.* 2010). Os estudos filogeográficos são fortemente apoiados em análises de dados de diversidade genética, traduzidos em árvores e redes de haplótipos. Entretanto, na última década, observou-se a introdução de outros tipos de informação nas análises filogeográficas, nomeadamente tamanho do efetivo populacional, crescimento ou declínio populacional, taxas de migração (Watterson 1975, Tajima 1989), utilização de

árvores de espécies, estimativa dos tempos de divergência (Heled & Drummond 2010, Yang & Rannala 2010), assim como, modelação de nicho ecológico (Knowles 2009, Chan *et al.* 2011, Alvarado-Serrano & Knowles 2014). Estes dados são essenciais para a determinação dos processos históricos e das dinâmicas populacionais envolvidas nos padrões de distribuição da variabilidade genética intraespecífica, principalmente em dois cenários: i) quando os eventos de divergência são muito recentes ou ii) na identificação de isolamento populacional por alopatria na presença de eventos de expansão populacional recentes ou de fluxo génico, o que pode gerar associações com o primeiro cenário.

A crescente complexidade das questões filogeográficas fez com que surgisse a necessidade de análises determinísticas e significativas, culminando com o surgimento da Filogeografia Estatística (Knowles 2009, Chan *et al.* 2011). Atualmente, a aplicação de métodos estatísticos rigorosos, utilizando complexos programas computacionais, permite análises refinadas de conectividade genética entre as populações, com a descrição de corredores de dispersão, estimativas de taxas de migração e sentido do fluxo génico (ver Kuhner 2009, Nielsen & Beaumont 2009), assim como, a avaliação de modelos de divergência populacional através de testes probabilísticos, em função da incerteza sobre os parâmetros populacionais históricos, tais como tamanho efetivo da população ancestral e tempo de divergência entre populações (Amaral *et al.* 2013, Tsai & Carstens 2013).

1.1.2.1 Filogeografia comparada

O número de trabalhos publicados sobre a Filogeografia de diversas espécies tem aumentado de forma significativa desde o final do século passado (Avise 2000, Hewitt 2000, 2004). Este aumento permitiu a comparação e posterior observação de concordância entre alguns padrões filogeográficos numa mesma área de distribuição (Filogeografia comparada – Taberlet *et al.* 1998). Tanto em ambientes terrestres como aquáticos é possível destacar mecanismos geológicos e/ou climáticos envolvidos na história evolutiva, ou mesmo ecológica, de um determinado grupo de espécies através da Filogeografia comparada. Um exemplo da aplicação deste tipo de estudos está relacionado com os ciclos glaciares do Pleistoceno, uma época caracterizada por intensa diversificação biológica em consequência de drásticas flutuações climáticas (Taberlet *et al.* 1998, Hewitt 2004, 2011; Gómez & Lunt 2007). Nestes estudos foi também possível a determinação dos efeitos dos processos de diversificação, a avaliação dos possíveis cenários biogeográficos sob uma perspetiva probabilística, a elucidação do papel dos fenómenos climáticos e históricos na subestruturação populacional e a identificação de refúgios glaciares, rotas de colonização pós-glaciar e regiões de contacto secundário (ver Carstens *et al.* 2005, Gómez & Lunt 2007,

Hewitt 2011). Este tipo de abordagem comparada tem contribuído também de forma relevante para a avaliação das respostas bióticas atuais e para a projeção destas para o futuro, visando a conservação e gestão das espécies e populações dentro de um contexto global de alterações climáticas (Avise 2009, Hickerson *et al.* 2010, Fordham *et al.* 2014).

1.1.2.2 Distribuição espacial e modelação de nicho ecológico

Um importante objetivo presente em diversas investigações nestes últimos anos é determinar quais as consequências das alterações climáticas globais para as populações e espécies. A modelação de nicho ecológico (MNE) (Haffer 1969), associada a estudos filogeográficos, tem sido uma abordagem cada vez mais aplicada para responder a estas questões, demonstrando por exemplo que as mudanças climáticas afetam a distribuição geográfica das espécies, os padrões de estrutura genética e ainda as dinâmicas de fluxo génico (ver Stewart 2009, Stewart *et al.* 2010; Tocchio *et al.* 2015). Se, por um lado, as barreiras vicariantes podem conduzir a processos de divergência alopátrica, limitando o fluxo génico e reforçando padrões de estrutura populacional (Hewitt 2004), por outro, os fatores ambientais/ecológicos têm um papel importante na diversificação, viabilidade e manutenção de populações e espécies (Thomassen *et al.* 2010, Gotelli & Stanton-Gueddes 2015). As análises de MNE utilizam dados fisiográficos georeferenciados (por exemplo, vegetação, clima, paleoclima, geologia) aplicados ao Sistemas de Informação Geográfica - SIG (do inglês *Geographic Information System*) (ver exemplos em Swenson 2008, Alvarado-Serrano & Knowles 2014). Os modelos resultantes geram mapas indicativos de áreas de alta e baixa aptidão abiótica (persistência), baseados na tolerância ecológica de uma determinada espécie ou unidade evolutiva (ver Soberón & Nakamura 2009, Warren *et al.* 2008). A previsão da distribuição geográfica de uma determinada espécie pode ser obtida em diferentes escalas temporais (passado, presente e, até mesmo, futuro).

A combinação de dados genéticos e ambientais é uma ferramenta poderosa para responder a perguntas biogeográficas mais complexas (Gotelli & Stanton-Gueddes 2015, Serra-Varela *et al.* 2015) Vários têm sido os trabalhos científicos a utilizar esta metodologia, nomeadamente na identificação de potenciais áreas de estabilidade climática nas regiões temperadas durante o Pleistoceno (Waltari *et al.* 2007, Abellán & Svenning 2014), na determinação dos requisitos ambientais responsáveis pelo estabelecimento e manutenção de zonas híbridas (Martínez-Freiría *et al.* 2008, 2009, 2015; Tarroso *et al.* 2014) e na avaliação da existência de diferenças de nicho entre linhagens evolutivas divergentes (McCormack *et al.* 2010, Alvarado-Serrano & Knowles 2014, Rato *et al.* 2015).

1.1.3 Marcadores moleculares

A relação entre linhagens ancestrais e descendentes passa necessariamente pela compreensão dos processos genealógicos (Filogenia e Filogeografia) e estrutura populacional (Genética Populacional) (Avise *et al.* 1987). Esta premissa evidencia a necessidade de integração dos diferentes métodos de análise para que a junção das diferentes partes permita a elucidação de um todo, ou seja, a recuperação da história evolutiva das espécies. Isto reflete-se ao nível dos marcadores moleculares utilizados neste tipo de estudos (Avise 2000). Se por um lado a utilização de apenas um *locus* pode conduzir a inferências incompletas acerca dos padrões evolutivos das espécies (Corl & Ellegren 2013, Leaché *et al.* 2014), as abordagens multilocus têm revelado a descrição de processos evolutivos consistentes, principalmente quando integrados com outros tipos de dados, tais como ecológicos e demográficos.

Um marcador molecular pode ser definido como qualquer *locus* génico, ou o seu produto, que apresenta variabilidade adequada, denominada de polimorfismo, para estudar uma determinada questão biológica (Snustad & Simmons 2011). Apesar dos marcadores nucleares e mitocondriais estarem sujeitos às mesmas forças evolutivas, diferenças quanto à taxa de mutação e hereditariedade podem levar à identificação de padrões de variabilidade distintos (Avise 2000) e de processos populacionais em diferentes tempos de uma história evolutiva (Thomson *et al.* 2010).

O marcador mais utilizado em estudos filogenéticos e filogeográficos é o DNA mitocondrial (mtDNA), tendo sido a principal fonte de informação biológica nas últimas duas décadas (Avise 2000, Garrick *et al.* 2011). Este fato deve-se à sua facilidade de extração; rápida taxa de evolução e tempo de coalescência mais curto do que o nDNA (Griffiths *et al.* 2010, Hedrick 2011); ausência de recombinação (no entanto ver Tsaousis *et al.* 2005 e Sammler *et al.* 2011); e modo de transmissão uniparental (Snustad & Simmons 2011). Esta última característica reflete-se no tamanho do efetivo populacional que é, em média, quatro vezes inferior ao das genealogias nucleares. Apesar da capacidade de revelar eventos de divergência mais recentes do que o nDNA (em última análise monofilia recíproca) o mtDNA é mais suscetível a fenómenos estocásticos e de deriva genética (Moore 1995). Por isso, torna-se essencial investigar marcadores moleculares nucleares, uma vez que as informações contidas nestes marcadores são qualitativamente diferentes, evidenciando, por exemplo, a ocorrência de fluxo génico e a retenção de polimorfismo ancestral (Sousa-Neves *et al.* 2013, Días-Rodríguez *et al.* 2015).

Ao nível nuclear diferentes marcadores moleculares neutrais, que não sofrem pressões seletivas, são amplamente utilizados: polimorfismos de nucleotídeo único (*Single*

Nucleotide Polymorphism – SNP) e os microssatélites. Os eventos de recombinação presentes nestes marcadores possibilitam inferências sobre níveis de miscigenação entre grupos populacionais diferenciados e, por consequência, o estudo de zonas de contato secundário com ou sem a ocorrência de hibridação (ver exemplos: Brunes 2015, Dufresnes *et al.* 2014). Os SNPs podem ser estudados através das genealogias nucleares (genes autossómicos) e permitem a descrição de histórias evolutivas em diferentes níveis de profundidade temporal, pois evoluem com taxas de mutação distintas. Em função disto, o estudo de genes nucleares tornou-se importante no resgate de informações sobre os processos históricos, demográficos e seletivos que determinam a estrutura genética das espécies (Hare 2001). No entanto, a combinação de taxas de mutação baixas com valores elevados de tamanho do efetivo populacional pode conduzir à retenção de polimorfismo ancestral o que, juntamente com a ocorrência de recombinação, pode dificultar as inferências evolutivas (Zhang & Hewitt 2003).

Por outro lado, os microssatélites revelam-se uma poderosa ferramenta para estudos de genética populacional, podendo ser aplicados em estudos microevolutivos e mais especificamente em inferência de *bottlenecks* ou de expansões populacionais. Estes marcadores consistem na repetição em *tandem* de dois a seis pares de bases. O número de repetições de cada motivo pode ser altamente variável tanto entre diferentes microssatélites, assim como dentro de um mesmo *locus* (Putman & Carbone 2014). A principal questão na utilização de microssatélites está relacionada com o modelo mutacional assumido durante as análises. O modelo mais aceite para explicar a variabilidade dos microssatélites é o modelo passo-a-passo (*Stepwise Mutation Model* – SMM – Kimura & Otha 1978), o qual assume que os eventos de mutação ocorrem, com a mesma probabilidade, por perda ou ganho de uma unidade de repetição em relação ao estado original da molécula. Em completa oposição, o modelo dos alelos infinitos (*Infinite Allele Model* – IAM – Kimura & Crow 1964) propõe que o número possível de estados alélicos de um determinado *locus* pode ser muito grande. Neste modelo, cada novo alelo (mutante) pode surgir na população com a mesma probabilidade. Por fim, o modelo de mutação de duas fases (*Two-Phase Model* – TPM – Di Rienzo *et al.* 1994) assume uma maior probabilidade para o surgimento de novos alelos pelo SMM e uma menor probabilidade para o IAM. Em relação às análises, a variabilidade polimórfica dos microssatélites pode ser investigada com base na frequência e no tamanho dos alelos, os quais são parâmetros intimamente relacionados com o fluxo génico (Griffiths *et al.* 2010, Putman & Carbone 2014). Os microssatélites têm sido amplamente utilizados em estudos de genética populacional, principalmente em análises de agrupamento Bayesiano (Pritchard 2000). Com esta metodologia, é possível fazer-se a categorização de um indivíduo, população ou linhagem e ainda, no caso de miscigenação, estimar-se a composição ancestral de um determinado indivíduo (Putman & Carbone 2014).

1.2 Península Ibérica: refúgio do Pleistoceno e *hotspot* de biodiversidade

As oscilações climáticas do Quaternário são caracterizadas por alternância de períodos glaciares e interglaciares. Essa alternância teve um papel crucial na biogeografia das espécies localizadas principalmente nas regiões temperadas, influenciando as suas histórias demográficas, processos de diversificação e distribuição geográfica das suas populações (Hewitt 2000, 2004, 2011; Taberlet *et al.* 1998). Vários estudos genéticos demonstraram que durante o Pleistoceno (época que está compreendida entre 2,6 milhões e 12 mil anos atrás) ocorreram sucessivos eventos de fragmentação populacional. Estes eventos, quando combinados com reduzidos tamanhos de efetivo populacional e a ação estocástica da deriva genética, podem ter influenciado os padrões de estruturação e diferenciação populacional, assim como os processos de diversificação intraespecífica de diversas espécies de vertebrados (ver Branco *et al.* 2000, 2002; Yang *et al.* 2006, Pinho *et al.* 2009, Hofmann 2012). Os estudos filogeográficos evidenciaram movimentos de expansão e contração populacional durante o Pleistoceno, que promoveram o aparecimento de áreas de contato secundário. Estas áreas abrangem desde zonas de completa simpatria, sem fluxo génico, até zonas com ampla introgressão. Contudo, parece que os eventos vicariantes do Pleistoceno não contribuíram decisivamente para processos de completa especiação (Hewitt 2004; Avise *et al.* 1998).

Na Europa, dados paleoecológicos e filogeográficos definem quatro áreas que mantiveram certa estabilidade climática durante o Pleistoceno: as penínsulas do sul do Mediterrâneo (Ibéria, Itália, Balcãs) e as montanhas dos Cárpatos (Figura 4) (Taberlet *et al.* 1998, Hewitt 2004). Estas regiões, definidas como refúgios, permitiram a sobrevivência de diversas espécies de plantas e animais, as quais recolonizaram o continente Europeu durante a fase de reaquecimento e equilíbrio de temperaturas no final do Pleistoceno (para exemplos e revisão ver Michaux *et al.* 2003, 2005; Hewitt 2004, Brito 2005, Rowe *et al.* 2006, Weiss & Ferrand 2007, Abellán & Svenning 2014). Este cenário caracteriza-se pela elevada variabilidade genética nos refúgios peninsulares (sul da Europa) *versus* a perda de diversidade biológica nas regiões neocolonizadas (centro e norte da Europa), resultante dos sucessivos eventos de expansão populacional e *bottlenecks* (Hewitt 2004). Entretanto, evidências com base em dados moleculares sugerem padrões mais complexos de sobrevivência durante o Quaternário. Algumas espécies temperadas e boreais apresentam claramente refúgios setentrionais na Europa (Schmitt 2007) e os refúgios do sul europeu constituem verdadeiros mosaicos heterogéneos quanto à diversificação de climas e habitats, reunindo, assim, múltiplos refúgios glaciares (Gómez & Lunt 2007, Weiss & Ferrand 2007).

Figura 4. Principais refúgios do Pleistoceno no continente Europeu (contornados em amarelo). 1: Península Ibérica; 2: Península Itálica; 3: Península dos Balcãs e 4: Montes Cárpatos. As áreas esbranquiçadas estiveram cobertas por gelo durante os períodos glaciares do Pleistoceno (Levin 2005). Durante os períodos de glaciação máxima, sugere-se que o gelo tenha coberto até um terço das superfícies de terra do planeta, estendendo-se desde o Ártico até aproximadamente 45° N de latitude. Mesmo nas regiões que não foram afetadas diretamente pela presença de espessas camadas de gelo e neve, como o hemisfério sul, estima-se que as temperaturas tenham oscilado, em média, entre os 4° e 8° C (Lomolino *et al.* 2010).

Trabalhos de filogeografia comparada e de modelação de climas com espécies endémicas da região ibérica, demonstraram a existência de vários microrefúgios dentro do refúgio ibérico com concordância de padrões de distribuição da diversidade genética e de zonas de estabilidade climática (modelo do "refúgio-dentro-do-refúgio" - Gómez & Lunt 2007). A ocorrência de fenómenos vicariantes na Península Ibérica pode ter conduzido à diferenciação de populações que, embora geograficamente próximas, sobreviveram aos períodos glaciares do Quaternário em refúgios isolados, fator este que pode ter contribuído para a diversificação de linhagens (Alexandrino *et al.* 2000, Gonçalves *et al.* 2009, Díaz-Rodríguez *et al.* 2015) (Figura 5). A existência nesta região de uma enorme diversidade biológica e importantes níveis de endemismo, designadamente no domínio herpetofaunístico, fazem da Península Ibérica uma área prioritária para a conservação da biodiversidade e um cenário ideal para o estudo da história evolutiva das espécies (Barbadillo *et al.* 1997, Abellán & Svenning 2014). Vários fatores fisiográficos e ambientais

(Figura 6) contribuíram para a importância da Península Ibérica como *hotspot* de biodiversidade: a sua localização no extremo sudoeste do continente europeu, estando separada do restante continente pelos Pirinéus; a relativa proximidade ao continente africano (aproximadamente 14 Km de distância) apenas separada pelo Estreito de Gibraltar; a apresentação de uma complexa fisiografia, refletida em cadeias montanhosas com orientação Este-Oeste, um amplo planalto central e bacias hidrográficas de grandes dimensões; e a elevada heterogeneidade climatérica, com influências climáticas do Atlântico e do Mediterrâneo (ver Riberon *et al.* 2001, Taberlet *et al.* 1998, Michaux *et al.* 2003, 2005; Ribera & Vogler 2004, Brito 2005, Rowe *et al.* 2006).

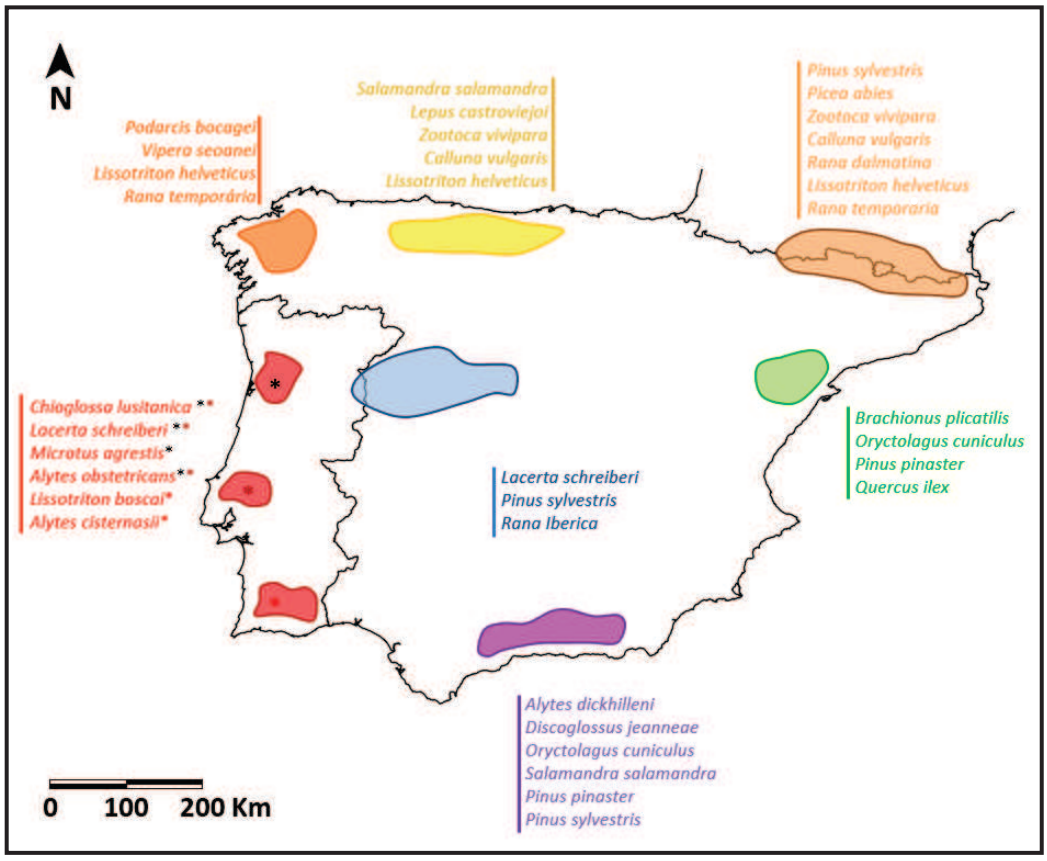


Figura 5. Principais refúgios na Península Ibérica durante o Pleistoceno (adaptado de Gómez & Lunt 2007). Outras referências: Martínez-Solano *et al.* (2006), Gonçalves *et al.* (2009), Pinho *et al.* (2011), Recuero & García-París (2011), Veith *et al.* (2012), Vences *et al.* (2013), Martínez-Freiría *et al.* (2015).

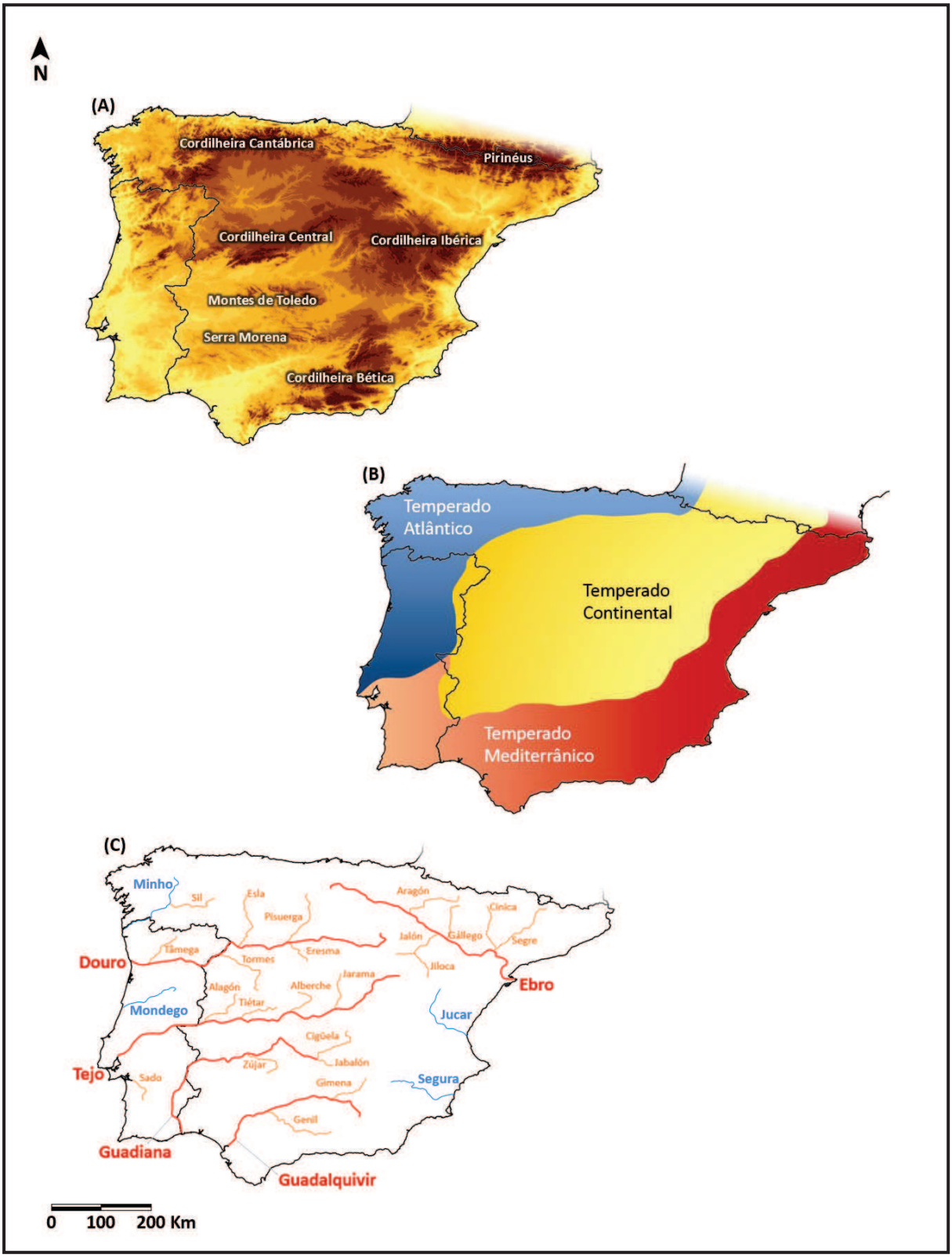


Figura 6. Características fisiográficas da Península Ibérica. (A) relevo, (B) clima, (C) bacias hidrográficas. Em vermelho destacam-se os rios principais e em laranja os rios adjacentes.

A implementação de trabalhos de investigação integrados, que combinem diferentes tipos de dados (genéticos, geográficos, ecológicos e ambientais), é fundamental para a compreensão dos processos de diversificação biológica e cenários biogeográficos. Neste contexto, organismos que apresentam limitada capacidade de dispersão, elevados níveis de estrutura populacional e vulnerabilidade ecológica, como é o caso dos anfíbios, tornam-se excelentes objetos de estudo (Zeisset & Beebee 2008, Reilly & Wake 2015).

1.3 Amphibia, Anura, Alytidae

A origem deste grupo de organismos aponta para o período Devoniano, há cerca de 350 milhões de anos atrás (ver Anderson 2008). Atualmente, estão descritas cerca de 7,5 mil espécies de anfíbios, incluindo sapos, rãs, salamandras, tritões e cecílias, distribuídas por todo globo terrestre (com exceção da Antártida) (Frost 2015). Todas as espécies vivas pertencem à subclasse Lissamphibia e estão divididas em três ordens: Anura, Caudata e Gymnophiona (San Mauro 2010). Atualmente, os anfíbios encontram-se num acentuado processo de declínio populacional (Stuart *et al.* 2004, Hoffmann *et al.* 2010). Aproximadamente 40% das espécies estão ameaçadas ou em risco de extinção (*International Union for Conservation of Nature* - IUCN). As alterações climáticas, o aparecimento de doenças infeciosas (por exemplo, a quitridiomicose e ranaviroses), contaminações químicas e a alteração, fragmentação e até mesmo a completa destruição dos seus habitats de ocorrência, são alguns dos fatores de ameaça que contribuem em última análise para a perda de diversidade genética (ver Allentoft & O'Brien 2010, Cheng *et al.* 2011, Savage & Zamudio 2011, Duarte *et al.* 2012).

Os Alytidae são uma família de sapos primitivos da ordem Anura. São conhecidos pelos nomes comuns de sapos-pintados ou sapos-parteiros. A maior parte das espécies atuais é endémica da Europa, mas existem três que ocorrem no noroeste africano. A família é também referenciada como Discoglossidae, todavia o nome mais antigo, Alytidae, tem sido o mais usado na maioria dos trabalhos recentes de investigação. A família é composta por três géneros: *Alytes*, *Discoglossus* e *Latonia* (AmphibiaWeb 2016).

1.4 Sapos-parteiros (*Alytes* spp.): aspetos biológicos e evolutivos

A origem do processo de divergência dos "sapos-parteiros", *Alytes* Wagler 1829, em relação aos demais Discoglossídeos está relacionada com o período do Cretáceo (Biton *et al.* 2013), suportada através do registo de fósseis na Ásia Central que foram atribuídos ao ancestral do sapo-parteiro (Sanchiz 1998). O nome comum atribuído a este grupo está relacionado com uma característica peculiar de cuidado parental: os machos transportam os ovos fertilizados nos membros posteriores durante todo o período de incubação. No momento da eclosão dos ovos, os machos deslocam-se para junto de massas de água onde os depositam. São sapos de pequeno porte (entre 32 e 55 milímetros) e hábitos crepusculares e noturnos (García-París *et al.* 2004).

Atualmente estão descritas cinco espécies de sapos-parteiros incluídas em três subgéneros (*Alytes*, *Ammoryctis* e *Baleaphryne* – Martínez-Solano *et al.* 2004*,* Gonçalves *et al.* 2007). Estas espécies estão distribuídas ao longo da região mediterrânica ocidental e são denominadas da seguinte forma: *Alytes* (*Alytes*) *obstetricans* (Laurenti 1768), *Alytes* (*Ammoryctis*) *cisternasii* (Boscá 1879), *Alytes* (*Baleaphryne*) *muletensis* (Sanchiz & Adrover 1979), *Alytes* (*Baleaphryne*) *dickhilleni* (Arntzen & García-París 1995) e *Alytes* (*Baleaphryne*) *maurus* (Pasteur & Bons 1962). Os subgéneros *Alytes* e *Ammoryctis* são claramente monofiléticos do ponto de vista morfológico e genético (Arntzen & García-París 1995, Fromhage *et al.* 2004), enquanto a monofilia do clado formado por *A. maurus* + *A. dickhilleni* + *A. muletensis* continua em discussão. Filogeneticamente este clado é bem suportado (Fromhage *et al.* 2004, Gonçalves *et al.* 2007), porém adaptações morfológicas de sapo trepador observadas em *A. muletensis* separam-no morfologicamente das outras duas espécies.

1.4.1 Subgénero *Ammoryctis*

O sapo-parteiro-ibérico (*A. cisternasii* – Figura 7) é uma espécie endémica do sudoeste e centro da Península Ibérica, ocorrendo desde o nível do mar até aos 1210 m de altitude (AmphibiaWeb 2016). Esta espécie fossorial exibe adaptações morfológicas aos habitats das planícies áridas mediterrânicas, ocorrendo na vizinhança de rios e ribeiros. Quanto à sua distribuição, *Alytes cisternasii* ocorre em Portugal, prevalentemente a sul do sistema montanhoso Lousã-Estrela, estendendo-se para norte através das regiões orientais das províncias da Beira-Baixa, Beira-Alta e Trás-os-Montes (Loureiro *et al.* 2008), e também em Espanha, distribuindo-se ao longo das bacias hidrográficas do Guadiana, Tejo e Douro.

A norte do planalto central, ocorre nas províncias de Zamora, Salamanca, Ávila, Valladolid, Segóvia e, a sul, nas províncias de Madrid, Guadalajara, Toledo, Cáceres, Badajoz, Ciudad Real, Huelva, Sevilha, Córdoba e Jaén (limitado a sudeste pelo rio Guadalquivir) (García-París *et al.* 2004, AmphibiaWeb 2016).

Os indivíduos de *Alytes cisternasii* são morfologicamente uniformes (Crespo, 1982) e suas populações apresentaram baixa variabilidade genética segundo dados de aloenzimas (Rosa *et al.* 1990, Arntzen & García-París 1995), sugerindo a persistência das populações de *A. cisternasii* num único refúgio mediterrânico durante o Pleistoceno. Contudo, Gonçalves *et al.* (2009) demonstraram, com base em dados mitocondriais e de microssatélites, um cenário de populações fragmentadas em variados refúgios durante as glaciações, seguido de subsequentes eventos de expansão populacional com o estabelecimento de contato secundário, além de elevada diversidade genética.

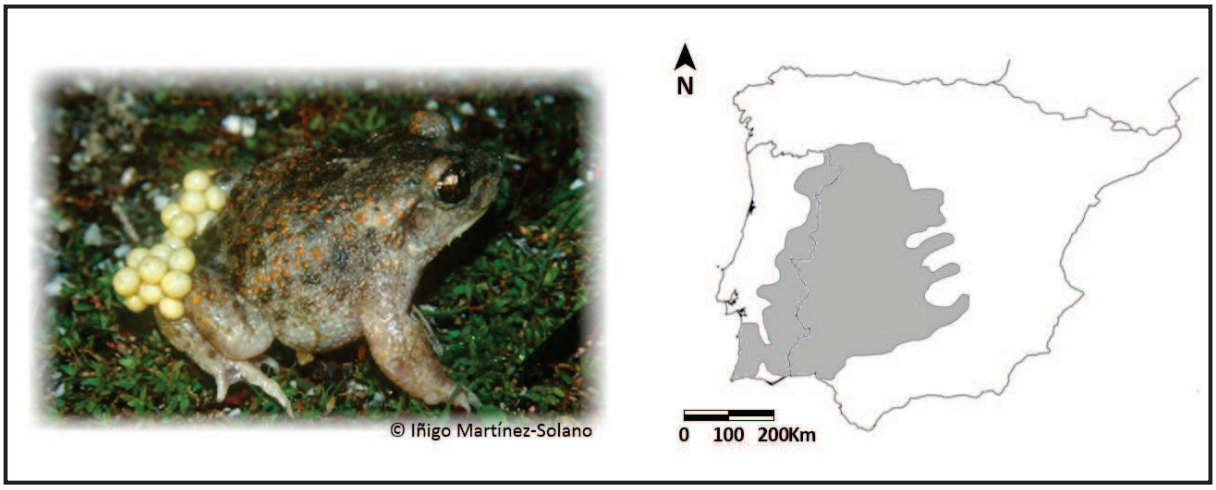


Figura 7. Sapo-parteiro-ibérico (*Alytes cisternasii*) e respetiva área de ocorrência destacada em cinzento.

1.4.2 Subgénero *Baleaphryne*

O clado *Baleaphryne* é composto por três espécies de sapo-parteiro. O sapo-parteiro-balear (*A. muletensis* – Figura 8) que, depois de se pensar que estava definitivamente extinto, foi redescoberto em 1980 (Sanchiz & Adrover 1979, Alcover & Mayol 1980, Sanchiz & Alcover 1982), tendo sido a sua osteologia descrita e comparada com a de *A. cisternasii* e *A. obstetricans* por Clarke (1984). Endémico da ilha de Maiorca, *A. muletensis* é considerado como uma espécie Vulnerável pela IUCN - *Red List of Threatened Species* (www.iucnredlist.org) e está também listado na *Evolutionarily Distinct and Globally Endangered* - EDGE (www.edgeofexistence.org), uma vez que possui uma área de

distribuição muito restrita (Serra Tramuntana) com reduzido número de adultos (entre 500 a 1500). Os indivíduos de *A. muletensis* distinguem-se claramente das outras espécies de *Alytes* por possuírem características morfológicas de sapo trepador, estando perfeitamente adaptados às paredes rochosas que caracterizam o seu habitat (Sanchiz 1984, Maxson & Szymura 1984).

O sapo-parteiro-bético (*A. dickhilleni* – Figura 8) foi descrito como uma espécie distinta em 1995 por Arntzen & García-París após estudo eletroforético de proteínas. Estes sapos possuem uma distribuição geográfica fragmentada e restrita às montanhas do sudeste ibérico, incluindo as Sierras Tejeda e Almijara (províncias de Málaga e Granada), a Sierra de Gádor (província de Almería), a Sierra de Baza (província de Granada), a Sierra Mágina (província de Jaén), e a Sierra de Alcaraz (província de Albacete) (García-París & Arntzen 2002, AmphibiaWeb 2016). A espécie é também considerada vulnerável (www.iucnredlist.org) e citada na EDGE (www.edgeofexistence.org), pois é muito sensível às mudanças climáticas e as populações fragmentadas têm sido alvo de infeções por quitridiomicose, causadas pelo fungo *Batrachochytrium dendrobatidis* (Bosch *et al.* 2013, Dias *et al.* 2014).

O sapo-parteiro-marroquino (*A. maurus* – Figura 8) ocorre geralmente em locais húmidos em áreas de montanha e escarpas, desde os 200 até 2050 metros de altitude nas regiões oeste e central das Montanhas do Rif e Atlas de Marrocos e Argélia (Donaire-Barroso & Bogaerts 2003). Ao longo da sua área de distribuição apresenta populações muito fragmentadas distribuídas por cerca de vinte localidades (AmphibiaWeb 2016). Inicialmente, esta espécie foi considerada como sendo uma subespécie de *Alytes obstetricans*. Contudo, Llorente *et al.* (1995), com base na análise biogeográfica do género *Alytes* na Península Ibérica, sugerem *A. maurus* como uma nova espécie. Esta denominação foi adotada e consolidada após os trabalhos de Donaire-Barroso & Bogaerts (2003), Fromhage *et al.* (2004), Martínez-Solano *et al.* (2004), Gonçalves *et al.* (2007) e Donaire-Barroso *et al.* (2008).

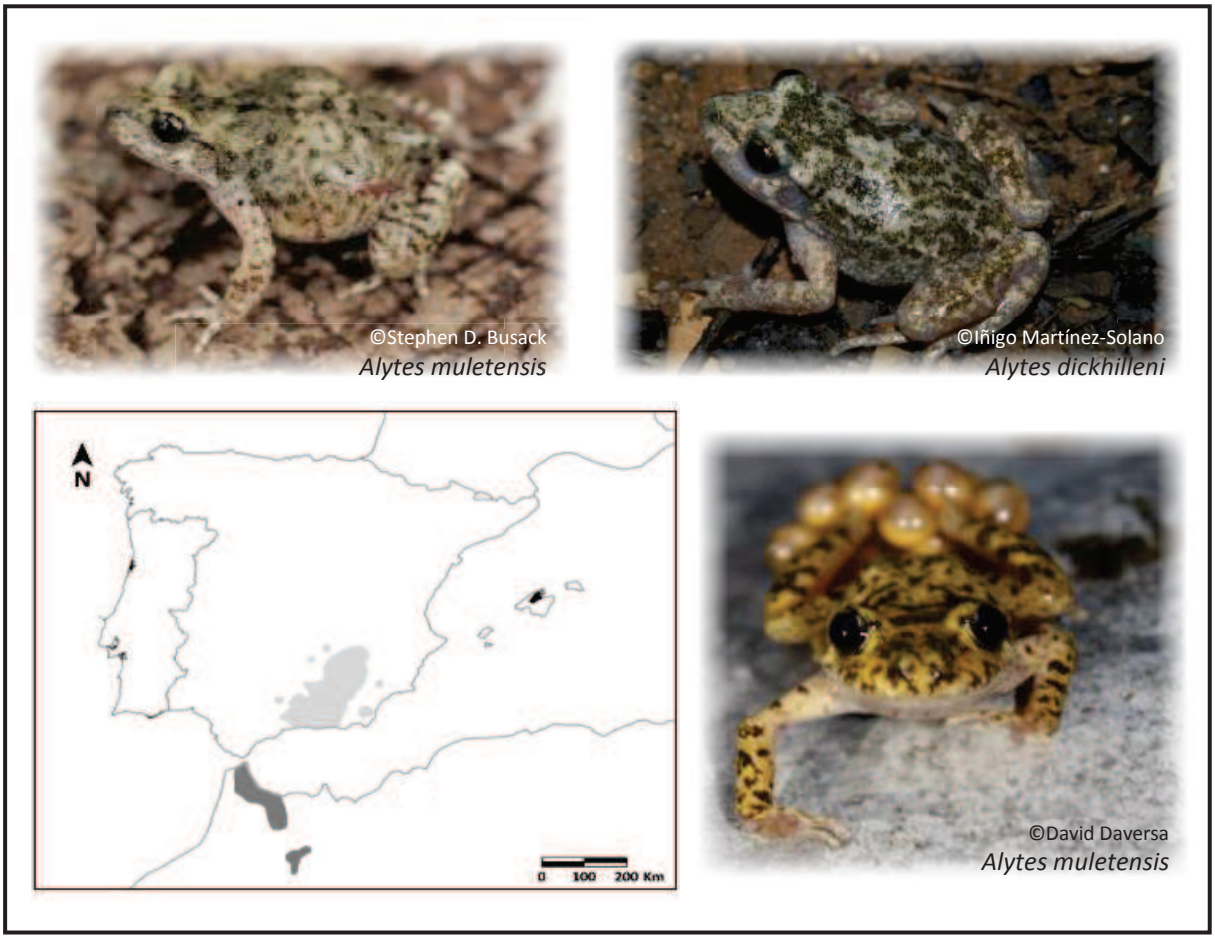


Figura 8. Espécies de sapo-parteiro do subgénero *Baleaphryne* e respetivas áreas de ocorrência. Cinzento claro: *Alytes dickhilleni*; cinzento escuro: *Alytes maurus*; preto: *Alytes muletensis*.

1.4.3 Subgénero *Alytes*

O sapo-parteiro-comum, *Alytes obstetricans*, é a espécie mais amplamente distribuída na Europa Ocidental, desde a metade norte da Península Ibérica até à Alemanha (Grossenbacher 1997). Esta espécie está adaptada a uma grande diversidade de habitats, tais como zonas rochosas de montanha, campos agrícolas e até áreas urbanas, ocorrendo desde o nível do mar até cerca dos 2400 m nos Pirinéus (Márquez & Rosa 1997). Devido ao seu prolongado desenvolvimento larvar, esta espécie é encontrada prevalentemente em massas de água permanente (Arntzen & García-París 1995). Diversos estudos evidenciaram diferenciações morfológicas e genéticas das populações de *A. obstetricans* (Crespo 1979, Crespo 1982, Arntzen & Szymura 1984, Viegas & Crespo 1985, Rosa *et al.* 1990, Arntzen & García-París 1995, García-París 1995, García-París & Martínez-Solano 2001, Fonseca *et al.* 2003, Fromhage *et al.* 2004), sendo a maior parte da diversidade genética desta espécie encontrada na Península Ibérica (Gonçalves 2007).

Atualmente são reconhecidas quatro subespécies (Figura 9): i) *A. o. obstetricans*, distribuída por toda a Europa ocidental e no norte da Península Ibérica (Navarra, País Basco e Cordilheira Cantábrica); ii) *A. o. boscai*, nas regiões norte e centro de Portugal, assim como Galiza, Castilla-León ocidental, bem como ao longo do Sistema Central de Espanha; iii) *A. o. pertinax*, presente nas regiões central e oriental da Península Ibérica (García-París & Martínez-Solano 2001); e iv) *A. o. almogavarii*, distribuída desde os Pirinéus orientais (Geniez & Crochet 2003) até à Serra de Guadarrama, a norte de Madrid (García-París 1995). No entanto, os limites da distribuição geográfica destas subespécies, potenciais zonas de simpatria e respetivas relações filogenéticas ainda não estão bem esclarecidos. Fonseca *et al.* (2003), com base na análise do mtDNA, sugerem que *A. o. obstetricans* pode ter uma área de distribuição estendida até ao rio Douro. Por outro lado, dados aloenzimáticos mostram a ocorrência de uma zona de hibridação nos Pirinéus entre *A. o. almogavarii* e *A. o. obstetricans*, especulando-se ainda potenciais áreas de sobreposição no Norte e Centro de Espanha (Arntzen & García-París 1995, García-París 1995). Outro aspeto que ainda não se encontra completamente esclarecido é qual terá sido o evento vicariante que conduziu à diferenciação das subespécies descritas, ou seja, se esta diferenciação está relacionada com a formação das principais bacias fluviais na Península Ibérica (há cerca de 4 milhões de anos atrás) ou associada à fragmentação populacional e isolamento em diferentes refúgios durante o Pleistoceno, como já foi relatado para as quatro linhagens genéticas descritas para A. *cisternasii* (Gonçalves *et al.* 2009).

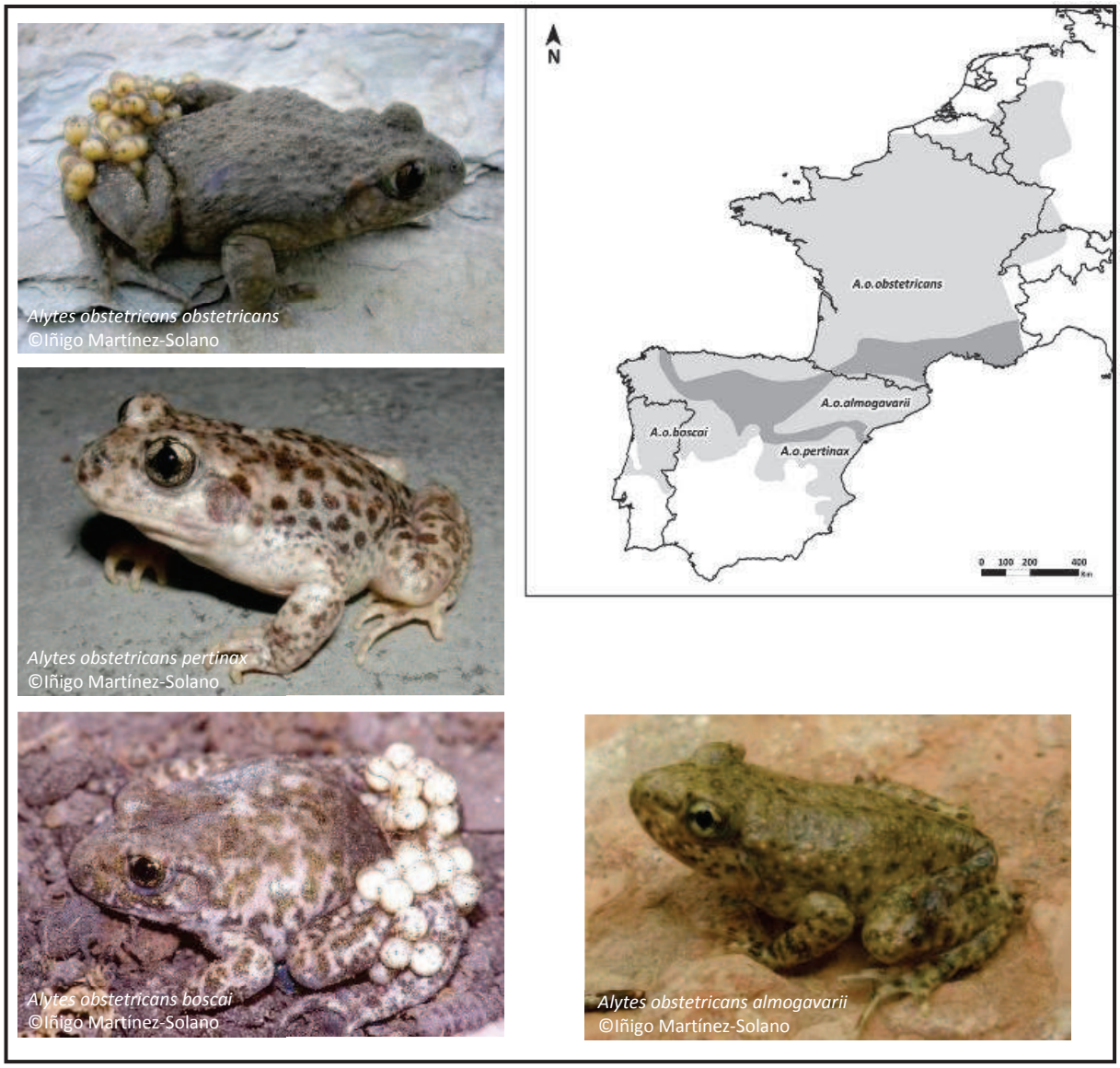


Figura 9. Subespécies de sapo-parteiro-comum (*Alytes obstetricans*) e respetivas áreas de ocorrência. Em cinzento escuro, está assinalada a área onde a determinação da subespécie não foi efetuada.

1.4.4 História evolutiva do género *Alytes*

A história evolutiva dos sapos-parteiros ainda não está bem esclarecida, refletindo-se na ausência de um cenário biogeográfico consensual. Os resultados filogenéticos obtidos até o momento mostram-se inconclusivos e inconsistentes em alguns aspetos, nomeadamente na determinação dos eventos vicariantes que levaram à diferenciação de algumas linhagens. Este fato pode ser decorrente de restrições nas análises realizadas, seja pela ausência de alguma linhagem evolutiva, a utilização de apenas um marcador genético ou a calibração do relógio molecular.

A linhagem de *A. cisternasii* terá sido a primeira a divergir, há cerca de 10 a 16 milhões de anos atrás, durante o Mioceno (Martínez-Solano *et al.* 2004). Dois eventos vicariantes podem estar relacionados com esta divergência: i) o isolamento do Maciço Bético-Rifenho (Arntzen & García-París 1995, Fromhage *et al.* 2004) que conduziria à especiação em alopatria do sapo-parteiro-ibérico (continente) e um proto-*A. obstetricans* (maciço); ou ii) a formação de grandes lagos salinos no centro da Península Ibérica, com uma posterior ocupação do Maciço Bético-Rifenho por linhagens de proto-*A. obstetricans*, proposto por Altaba (1997).

Após a publicação dos trabalhos de Arntzen & García-París (1995) e Fromhage *et al.* (2004), propõe-se que a fragmentação do Maciço Bético-Rifenho tenha sido o evento responsável pela separação de linhagens ancestrais de *Balephryne* e *A. obstetricans*. Durante a crise Messiniense, há cerca de 5,9–5,3 Ma, *A. obstetricans* terá colonizado a Península Ibérica enquanto as linhagens de *Baleaphryne* terão permanecido isoladas nas montanhas Béticas. O posterior enchimento da bacia do Mediterrâneo terá isolado as ilhas Baleares, conduzido à especiação de *A. dickhilleni*, *A. muletensis e A. maurus*.

Em 2004, Martínez-Solano *et al.* publicaram uma filogenia multilocus para o género *Alytes*, incluindo análises morfológicas, propondo um cenário biogeográfico mais consensual que se descreve seguidamente (Figuras 10 e 11). A Península Ibérica terá sido colonizada por populações ancestrais do género *Alytes*. A formação dos grandes lagos salinos no interior da Península Ibérica, há cerca de 16 milhões de anos atrás, seria o evento vicariante responsável pela separação de *A. cisternasii*, no sudoeste ibérico, das populações restantes localizadas nordeste ibérico. Este stock remanescente expandiu-se e, posteriormente, o surgimento do Estreito Bético terá isolado a linhagem ancestral dos *Baleaphryne;* permanecendo um segundo grupo na Península, ancestral das demais subespécies de *A. obstetricans*. A fragmentação do Maciço Bético-Rifenho poderá ter iniciado a divergência dos *Baleaphryne*, todavia seria a abertura do estreito de Gibraltar, com o final da crise Messiniense (em torno de 5,3 milhões de anos atrás), que possibilitaria a divergência entre

as três espécies deste clado. A rápida radiação dificultou a resolução da politomia referenciada por Martínez-Solano *et al.* (2004) para este clado. Segundo os mesmos autores, uma migração trans-marítima poderia estar relacionada com a divergência de A. *muletensis*. Em 2007, Gonçalves *et al.*, com base em dados de genealogia nuclear *(β-fibrinogénio*), observaram conflitos no sinal filogenético do clado *Baleaphryne*, evidenciando um agrupamento com a linhagem de *A. o. almogavarii* (Figura 12). Os autores levantaram a hipótese de possíveis eventos de hibridação ou retenção de polimorfismo ancestral para justificar a deteção de uma coalescência mais profunda do que se supunha, colocando não só em discussão a monofilia do clado, mas também o nível de divergência de *A. o. almogavarii* em relação às demais subespécies. Segundo Gonçalves (2007) a diversificação intraespecífica de *A. obstetricans* poderia estar relacionada com a formação das bacias hidrográficas na Península ou alternativamente associada às glaciações do Pleistoceno. A autora descreve a presença de, pelo menos, duas linhagens altamente divergentes e geograficamente estruturadas com base nas análises do mtDNA, sugerindo processos de divergência em diferentes refúgios ibéricos: um localizado a sul do rio Douro e outro nos Pirinéus. A análise comparada de dados filogenéticos e filogeográficos também forneceu fortes indicações da possível existência, durante as glaciações do Pleistoceno, de um refúgio no nordeste ibérico (Gonçalves 2007).

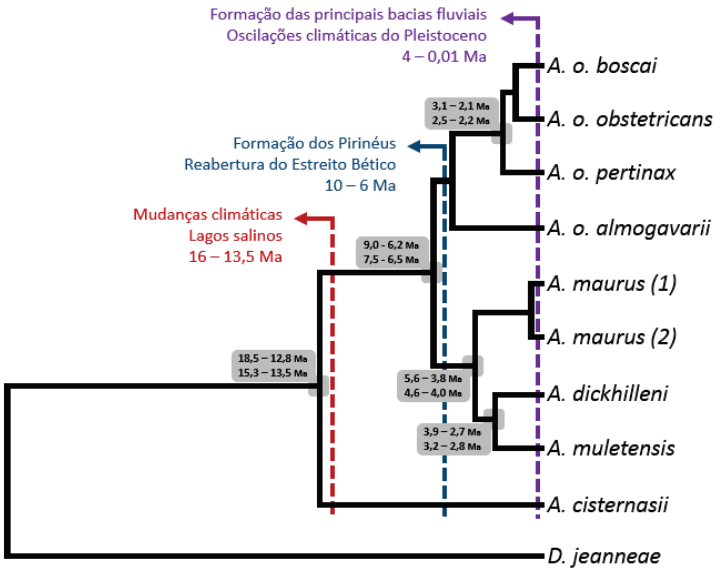


Figura 10. Árvore ultramérica de *Alytes* spp. adaptada de Martínez-Solano *et al.* (2004). Os tempos de divergência, destacados junto aos nós da árvore, foram calculados com base em duas calibrações do relógio molecular: superior - dados proteícos (Beerli *et al.* 1996); inferior - dados imunológicos (Maxson & Szymura 1984). Ma: milhões de anos atrás. Os principais eventos vicariantes são destacados em linhas tracejadas.

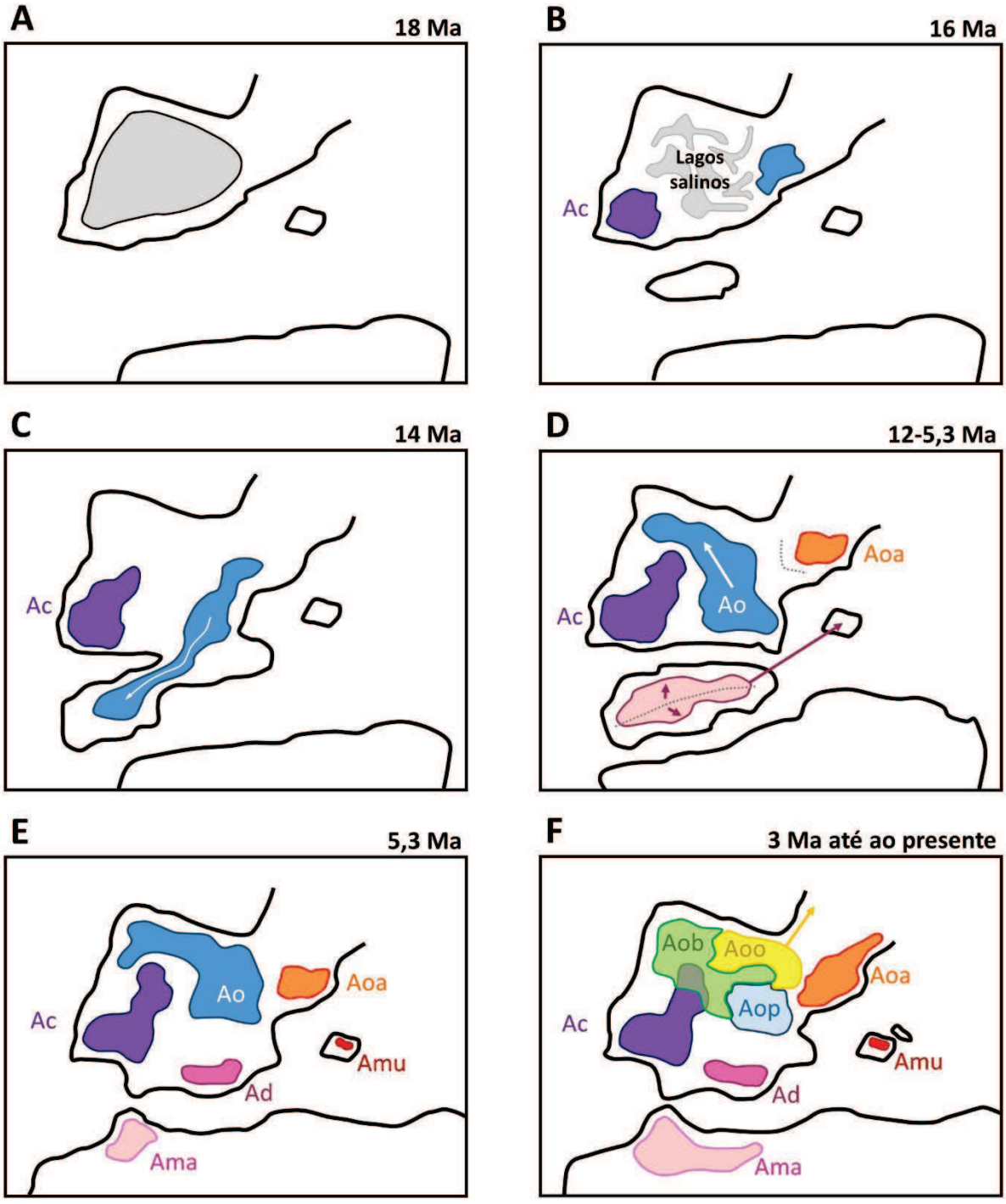


Figura 11. Cenário biogeográfico com base em Martínez-Solano *et al.* (2004) e Gonçalves *et al.* (2007). Ma: milhões de anos atrás. A: aparecimento da linhagem ancestral dos *Alytes* e ocupação da Península Ibérica. B: Formação dos grandes lagos salinos na região central da Península e/ou drásticas mudanças climáticas promoveram a diferenciação da linhagem ancestral de *A. cisternasii* (Ac), no sudoeste ibérico, vs. o stock remanescente de *Alytes*. C: grupo formado pela linhagem ancestral do clado *Baleaphryne* expande-se em direção ao Maciço Bético-Rifenho. D: Isolamento de *Baleaphryne* no Maciço devido a reabertura do Estreito Bético. Divergência da linhagem ancestral de *A.o.almogavarii* (Aoa) *vs.* ancestral de *A. obstetricans* (Ao). A crise Messiniense poderia ter facilitado a colonização das Ilhas Baleares pela linhagem ancestral de *A. muletensis* (Amu). E: a abertura do Estreito de Gibraltar e enchimento do Mar Mediterrâneo, isolando as linhagens ancestrais de *A. maurus* (Ama), *A. dickhilleni* (Ad) e *A. muletensis* (Amu). F: Diversificação intraespecífica de *A. obstetricans* e expansão populacional (Aoo - *A. o. obstetricans*; Aop - *A. o. pertinax*; Aob. - *A. o. boscai*).

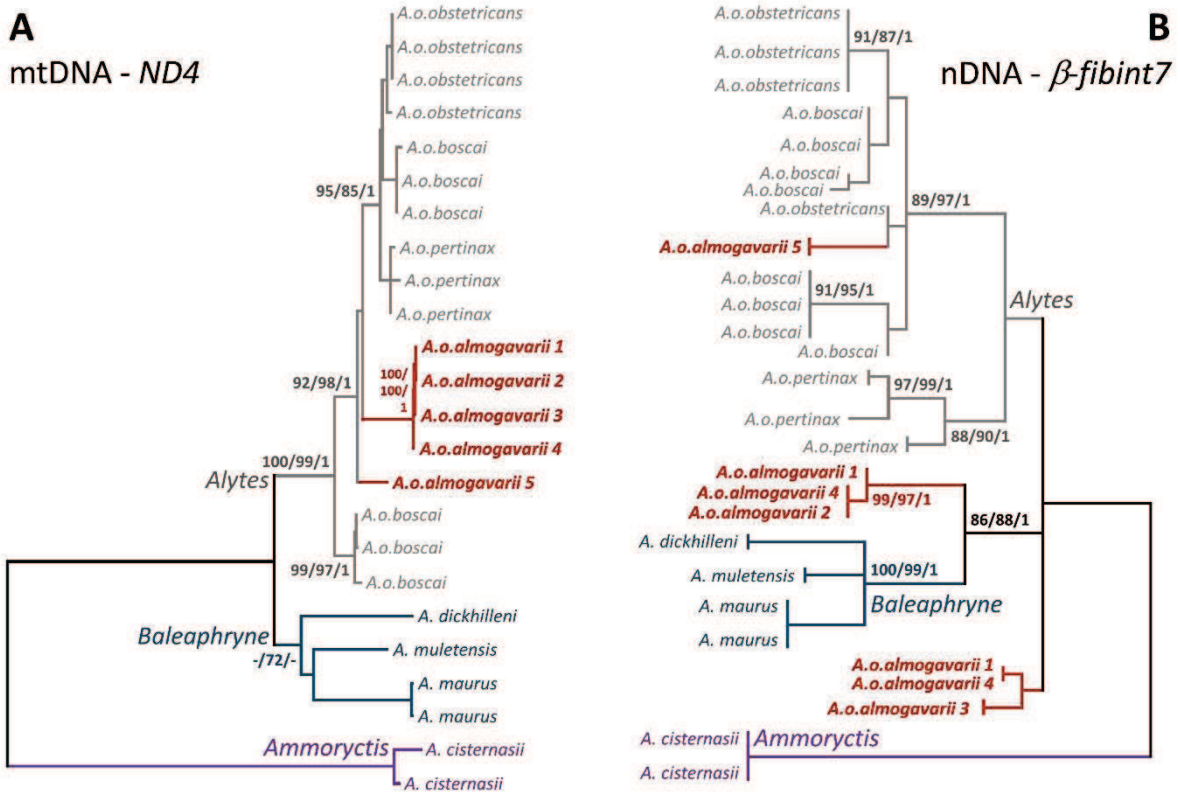


Figura 12. Árvores filogenéticas ilustrando as relações entre as várias espécies e subespécies de *Alytes* adaptadas de Gonçalves *et al.* (2007). A: gene ND4. B: gene *Beta fibrinogénio intron 7*. Os números junto aos ramos representam os valores de *bootstrap* (Máxima Parcimónia/Máxima Verossimilhança) e probabilidades posteriores.

1.5 Objetivos e estrutura da tese

O objetivo central desta tese é a investigação dos processos evolutivos que determinaram a diversificação dos sapos-parteiros (*Alytes* spp.) na Bacia Mediterrânica Ocidental através de uma abordagem integrada utilizando diferentes marcadores genéticos (mtDNA, genealogias nucleares e microssatélites) e métodos de análise (por exemplo, Bayesianos e modelação de nicho ecológico). A região da Península Ibérica foi a principal área de estudo, pois, segundo os estudos prévios, detém a maior diversidade biológica deste género. Assim, os seguintes objetivos específicos foram definidos:

- Confirmar a datação preliminar da fragmentação histórica das populações de *A. muletensis*, *A. dickhilleni* e *A. maurus* e a sua relação com as modificações geomorfológicas da Bacia Mediterrânica Ocidental;
- Descrever os padrões filogeográficos das populações de sapo-parteiro-comum (*A. obstetricans)* na Península Ibérica;
- Avaliar a existência de diversidade críptica e especiação incipiente sugeridas para algumas linhagens de *A. obstetricans*;
- Detetar a ocorrência de divergência de nichos ecológicos entre as diferentes linhagens de *A. obstetricans* e, por consequência, a associação de fatores ambientais com os processos de diversificação.

O conteúdo desta tese está organizado em cinco capítulos. No capítulo 1 faz-se uma introdução geral ao tema, destacando-se a informação teórica relevante para justificar os objetivos desta tese. Neste capítulo encontram-se revistos os conhecimentos existentes nas áreas da Biogeografia e Biologia Evolutiva, nomeadamente filogenéticos, filogeográficos e métodos de análise. Adicionalmente, faz-se uma descrição da área de estudo e dos possíveis eventos vicariantes relacionados com a diversificação dos sapos-parteiros (*Alytes* spp.). Os aspetos biológicos e a história evolutiva do objeto de estudo são também descritos neste capítulo. Em consonância com os objetivos delineados, os capítulos 2, 3 e 4 são constituídos pelos resultados obtidos no presente trabalho, sendo apresentados sob a forma de quatro artigos científicos, três dos quais já publicados em revistas científicas de circulação internacional e um submetido para publicação.

No capítulo 2 é apresentada a análise filogenética multilocus de todas as espécies e subespécies do género *Alytes* (Artigo I). Nesta secção, as relações filogenéticas entre todas as linhagens do género *Alytes* foram avaliadas com base em métodos de máxima verossimilhança e Bayesianos. Após a análise das árvores de genes e da árvore de espécies, destacam-se como resultados principais a resolução da politomia do clado

Baleaphryne e a descrição dos eventos vicariantes potencialmente envolvidos na diferenciação deste clado.

No capítulo 3 apresenta-se uma análise multilocus ao nível intraespecífico de populações de *Alytes obstetricans*. No Artigo II, foram realizadas inferências sobre os processos de divergência genética, delimitação de unidades evolutivas significativas e designação de espécies crípticas. Os resultados demonstraram que os padrões filogeográficos das populações de *A. obstetricans* são mais complexos do que se supunha, sendo influenciados pelas oscilações climáticas do Pleistoceno.

No capítulo 4, composto pelos Artigos III e IV, fez-se uma análise estrutural da diversidade genética intraespecífica das várias populações de *A. obstetricans* com base em microssatélites e a definição de zonas de persistência a partir de técnicas de modelação de nicho ecológico. A genotipagem de 17 microssatélites em 965 indivíduos revelou um claro padrão de estrutura populacional, com a caracterização de sete grupos genéticos. Estes grupos possuem diferentes requisitos ecológicos os quais são refletidos em áreas de persistência distintas com estreitas zonas de sobreposição.

Por fim, o capítulo 5 consiste na discussão integrada dos principais resultados obtidos, mais especificamente das relações filogenéticas entre as diferentes linhagens de *Alytes*, da descrição dos padrões filogeográficos e dos fatores históricos e ambientais que moldaram a estrutura da diversidade genética em *A. obstetricans*, da diversificação intraespecífica e possíveis implicações taxonómicas e, por fim, na reconstituição do cenário biogeográfico para o género *Alytes*. Neste capítulo, apresentam-se também as considerações finais sobre o trabalho realizado e a indicação de perspetivas futuras.

1.6 Lista dos artigos que integram a tese

Os artigos científicos apresentados nos Capítulos 2, 3 e 4 foram uniformizados quanto à formatação sem qualquer alteração dos conteúdos, uma vez que as diferentes revistas científicas publicaram os respetivos trabalhos com arte final diferenciada. Os artigos que compõem esta tese são:

Artigo 1. Maia-Carvalho B, Gonçalves H, Ferrand N, Martínez-Solano I (2014) Multilocus assessment of phylogenetic relationships in *Alytes* (Anura, Alytidae). *Molecular Phylogenetics and Evolution*, 72: 270–278 (doi: 10.1016/j.ympev.2014.05.033).

Artigo 2. Gonçalves H, Maia-Carvalho B, Sousa-Neves T, García-París M, Sequeira F, Ferrand N, Martínez-Solano (2015) Multilocus phylogeography of the common midwife toad, *Alytes obstetricans* (Anura, Alytidae): Contrasting patterns of lineage diversification and genetic structure in the Iberian refugium. *Molecular Phylogenetics and Evolution*, 93: 363-379 (doi: 10.1016/j.ympev.2015.08.009).

Artigo 3. Maia-Carvalho B, Gonçalves H, Martínez-Solano I, Gutiérrez-Rodríguez J, Lopes S, Ferrand N, Sequeira F (2014) Intraspecific genetic variation in the common midwife toad (*Alytes obstetricans*): subspecies assignment using mitochondrial and microsatellite markers. *Journal of Zoological Systematics and Evolutionary Research*, 52: 170–175 (doi: 10.1111/jzs.12048).

Artigo 4. Maia-Carvalho B, Vale C, Sequeira F, Ferrand N, Martínez-Solano I, Gonçalves H (*submitted*). When the environment favours lineage diversification: ecological niche divergence mirrors intraspecific patterns of genetic diversity in the Common Midwife Toad (*Alytes obstetricans*).

1.7 Referências Bibliográficas

Abellán P, Svenning JC (2014) Refugia within refugia–patterns in endemism and genetic divergence are linked to Late Quaternary climate stability in the Iberian Peninsula. *Biological Journal of the Linnean Society,* 113: 13-28.

Alcover JA, Mayol J (1980) Noticia del hallazgo de *Baleaphryne* (Amphibia: Anura: Discoglossidae) viviente en Mallorca. Doñana, *Acta Vertebrata*, 7: 266-269.

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.

Allentoft ME, O'Brien J (2010) Global amphibian declines, loss of genetic diversity and fitness: a review. *Diversity*, 2: 47-71.

Altaba C (1997) Phylogeny and biogeography of midwife toads: a reappraisal. *Contributions to Zoology*, 66: 257–262.

Alvarado-Serrano DF, Knowles LL (2014) Ecological niche models in phylogeographic studies: applications, advances and precautions. *Molecular Ecology Resources*, 14: 233-48.

Amaral FR, Albers PK, Edwards SV, Miyaki CY (2013) Multilocus tests of Pleistocene refugia and ancient divergence in a pair of Atlantic Forest antbirds (*Myrmeciza*). *Molecular Ecology*, 22: 3996-4013.

AmphibiaWeb (2017) AmphibiaWeb: Information on amphibian biology and conservation. Available from: http://amphibiaweb.org/ Berkeley, CA (acessado em: janeiro de 2017).

Anderson JS (2008) Focal review: The origin(s) of modern amphibians. *Evolutionary Biology*, 35: 231-247.

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.

Arntzen JW, Szymura JM (1984) Genetic differentiation between African and European midwife toads (*Alytes*, Discoglossidae). *Bijdragen tot de Dierkunde*, 54: 157-162.

Avise JC (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, 36: 3–15.

Avise JC (2000) Phylogeography. The History and Formation of Species. Harvard University Press. Cambridge, Massachusetts, USA, 464 pp.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, 18: 489–522.

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.

Barbadillo LJ, García-París M, Sanchiz B (1997) Orígenes y relaciones evolutivas de la herpetofauna iberica. Distribución y biogeografía de los anfibios y reptiles en España y Portugal (ed. by JM Pleguezuelos), pp. 47–100. Universidad de Granada, Asociación Herpetológica Española, Granada.

Belfiore NM, Liu L, Moritz C (2008) Multilocus phylogenetics of a rapid radiation in the genus *Thomomys* (Rodentia: Geomyidae). *Systematic Biology*, 57: 294-310.

Beerli P, Hotz H, Uzzell T (1996) Geologically dated sea barriers calibrate a protein clock for Aegean water frogs. *Evolution*, 50: 1676–1687.

Biton R, Geffen E, Vences M, Cohen O, Bailon S, Rabinovich R, Gafny S (2013) The rediscovered Hula painted frog is a living fossil. *Nature Communications*, 4: 1959.

Bosch J, García-Alonso D, Fernández-Beaskoetxea S, Fisher MC, Garner TW (2013) Evidence for the introduction of lethal Chytridiomycosis affecting wild betic midwife toads (*Alytes dickhilleni*). *EcoHealth*, 10: 82-91.

Boussau B, Daubin V (2010) Genomes as documents of evolutionary history. *Trends in Ecology and Evolution*, 25: 224-232.

Braby MF, Eastwood R, Murrau N (2012) The subspecies concept in butterflies: has its application in taxonomy and conservation biology outlived its usefulness? *Biological Journal of the Linnean Society*, 106: 699-716.

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.

Brito PH, Edwards SV (2009) Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica,* 135: 439-455.

Brunes T (2015) Biogeografia e diversificação de um grupo de anuros Neotropicais, *Phyllomedusa burmeisteri*. Abordagem integrativa através de análises moleculares e modelos de nicho ecológico. Tese de Doutoramento, Universidade do Porto, Portugal.

Bryson Jr. RW, Pastorini J, Burbrink FT, Forstner MRJ (2007) A phylogeny of the *Lampropeltis mexicana* complex (Serpentes: Colubridae) based on mitochondrial DNA sequences suggests evidence for species-level polyphyly within *Lampropeltis*. *Molecular Phylogenetic and Evolution*, 43: 674–684.

Carstens BC, Brunsfeld SJ, Demboski JR, Good JM, Sullivan J (2005) Investigating the evolutionary history of the Pacific Northwest mesic forest ecosystem: hypothesis testing within a comparative phylogeographic framework. *Evolution*, 59: 1639-1652.

Carstens BC, Pelletier TA, Reid NM, Satler JD (2013) How to fail at species delimitation. *Molecular Ecology*, 22: 4369–83.

Chan LM, Brown JL, Yoder AD (2011). Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Molecular Phylogenetics and Evolution*, 59: 523–537.

Cheng TL, Rovito SM, Wake DB, Vredenburg VT (2011) Coincident mass extirpation of neotropical amphibians with the emergence of the infectious fungal pathogen *Batrachochytrium dendrobatidis*. *Proceedings of the National Academy of Sciences*, 108: 9502-9507.

Clarke BT (1984). General skeletal morphology. In: Historia Biológica del Ferreret (eds. Hemmer H, Alcover JA), pp. 45–59. Moll, Mallorca, España.

Corl A, Ellegren H (2013) Sampling strategies for species trees: the effects on phylogenetic inference of the number of genes, number of individuals, and whether loci are mitochondrial, sex-linked, or autosomal. *Molecular Phylogenetics and Evolution*, 67: 358-366.

Coyne JA, Orr HA (2004) Speciation. Sinauer Associates, Inc. Sunderland, Massachusetts, USA, 545 pp.

Crespo EG (1979) Contribuição para o conhecimento da biologia dos *Alytes* ibéricos, *Alytes obstetricans boscai* Lataste, 1879 e *Alytes cisternasii* Boscá, 1879 (Amphibia, Salientia): A problemática da especiação de *Alytes cisternasii*. Tese de Doutoramento. Universidade de Lisboa, Portugal.

Crespo EG (1982) Contribuição para o conhecimento da biologia das espécies ibéricas de *Alytes*, *A. obstetricans boscai*, Lataste 1879 e *A. cisternasii*, Boscá 1879 (Amphibia, Discoglossidae): morfologia dos adultos e dos girinos. *Arquivos do Museu Bocage (C)*, 1: 255–316.

Cronquist A (1978) Once again, what is a species? (Ed. Knutson LV), pp. 3-20. Montclair, NJ: Alleheld Osmun.

Daniels SR, Hofmeyr MD, Henen BT, Baard EH (2010) Systematics and phylogeography of a threatened tortoise, the speckled padloper. *Animal Conservation*, 13: 237–246.

Dayrat B (2003) The roots of phylogeny: how did Haeckel build his trees? *Systematic Biology*, 52: 515–527.

Dayrat B (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society*, 85: 407-415.

De Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology*, 56: 879-86.

Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24: 332-340.

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994). Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America*, 91: 3166–3170.

Dias G, Beltrán JF, Tejedo M, Benítez M, Miras EG, Ferrand N, Gonçalves H (2014) Limited gene flow and high genetic diversity in the threatened Betic midwife toad (*Alytes dickhilleni*): evolutionary and conservation implications. *Conservation Genetics*, 16: 459-476.

Díaz-Rodríguez J, Gonçalves H, Sequeira F, Sousa-Neves T, Tejedo M, Ferrand N, Martínez-Solano I (2015) Molecular evidence for cryptic candidate species in Iberian *Pelodytes* (Anura, Pelodytidae). *Molecular Phylogenetics and Evolution*, 83: 224-241.

Dobzhansky T (1950) Mendelian populations and their evolution. *The American Naturalist*, 84: 401–418.

Donaire-Barroso D, Bogaerts S (2003) Datos sobre taxonomia, ecologia y biológia de *Alytes maurus* (Pasteur & Bons, 1962) (Anura; Discoglossidae). *Butlletí de la Societat Catalana d'Herpetologia*, 16: 25-42.

Donaire-Barroso D, Salvador A, Slimani T, Mouden EHE, Martínez-Solano I (2008) *Alytes maurus*. IUCN 2010. IUCN Red List of Threatened Species. Version 2010.1.

Duarte H, Tejedo M, Katzenberger M, Marangoni F, Baldo D, Beltrán JF, Martí DA, Richter-Boix A, Gonzalez-Voyer A (2012) Can amphibians take the heat? Vulnerability to climate warming in subtropical and temperate larval amphibian communities. *Global Change Biology*, 18: 412-421.

Dufresnes C, Bonato L, Novarini N, Betto-Colliard C, Perrin N, Stöck M (2014) Inferring the degree of incipient speciation in secondary contact zones of closely related lineages of Palearctic green toads (*Bufo viridis* subgroup). *Heredity*, 113: 9- 20.

Dufresnes C, Perrin N (2015) Effect of biogeographic history on population vulnerability in European amphibians. *Conservation Biology*, 29: 1235-1241.

Dufresnes C, Wassef J, Ghali K, Brelsford A, Stöck M, Lymberakis P, Perrin N (2013) Conservation phylogeography: Does historical diversity contribute to regional vulnerability in European tree frogs (*Hyla arborea*)? *Molecular Ecology*, 22: 5669-5684.

Edwards DL, Knowles LL (2014) Species detection and individual assignment in species delimitation: can integrative data increase efficacy? *Proceedings of the Royal Society B: Biological Sciences*, 281: 20132765.

Edwards SV (2009) Is a new and general theory of molecular systematics emerging? *Evolution*, 63: 1–19.

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

Ence DD, Carstens BC (2011) SpedeSTEM: A rapid and accurate method for species delimitation. *Molecular Ecology Resources*, 11: 473–480.

Etges WJ (2014) No Boundaries: Genomes, Organisms, and Ecological Interactions Responsible for Divergence and Reproductive Isolation. *Journal of Heredity*, 105: 756–770.

Fonseca A, Arntzen JW, Crespo EG, Ferrand N (2003) Regional differentiation in the common midwife toad (*Alytes obstetricans*) in Portugal: a picture from mitochondrial DNA. *Zeitschrift für Feldherpetologie*, 10: 83–89.

Fordham DA, Brook BW, Moritz C, Nogués-Bravo D (2014) Better forecasts of range dynamics using genetic data. *Trends in Ecology & Evolution*, 29: 436–443.

Frankham R (2002) Population viability analysis. *Nature*, 419: 18–19.

Frankham R (2005) Genetics and extinction. *Biological Conservation*, 126: 131–140.

Frankham R, Ballou JD, Briscoe DA (2010) Introduction to Conservation Genetics. 2nd Edition, Cambridge University press. Cambridge, United Kingdom, 644 pp.

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

Frost DR (2015) Amphibian Species of the World: an Online Reference. Version 6.0 (April, 2016). Electronic Database accessible at: http://research.amnh.org/herpetology/amphibia/index.html. American Museum of Natural History, New York, USA.

García-París M (1995) Variabilidad genética y distribución geográfica de *Alytes obstetricans almogavarii* en España. *Revista Española de Herpetología*, 9: 133–138.

García-París M, Arntzen JW (2002) *Alytes dickhilleni* Arntzen & García-París 1995. In: Atlas y Libro Rojo de los Anfibios y Reptiles de España (eds. Pleguezuelos JM, Márquez R, Lizana M), pp. 76–78. Dirección General de la Conservación de la Naturaleza, Asociación Herpetológica Española, Madrid.

García-París M, Martínez-Solano I (2001) Nuevo estatus taxonómico para las poblaciones iberomediterráneas de *Alytes obstetricans* (Anura: Discoglossidae). *Revista Española de Herpetología*, 15: 99–113.

García-París M, Montori A, Herrero P (2004) *Amphibia, Lissamphibian*. In: Fauna Ibérica, vol. 24. (Eds. Ramos MA *et al.*) Museo Nacional de Ciencias Naturales. CSIC. Madrid, España.

Garrick RC, Bonatelli IA, Hyseni C, Morales A, Pelletier TA, Perez MF, Rice E, Satler JD, Symula RE, Thomé MT, Carstens BC (2011) The evolution of phylogeographic data sets. *Molecular Ecology*, 24: 1164-71.

Geniez P, Crochet PA (2003) Confirmation de l'existence, en France, de trois taxons méconnus: *Alytes obstetricans almogavarii* Arntzen & García-París, 1995 (Amphibia, Discoglossidae), *Podarcis hispanica sebastiani* (Klemmer, 1964) (Reptilia, Lacertidae) et *Natrix natrix astreptophora* (Seoane, 1884) (Reptilia, Colubridae). *Bulletin de la Société Herpétologique de France*, 105: 41-53.

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Phylogeography of Southern European Refugia: Evolutionary Perspectives on the Origin and Conservation of European Biodiversity (Eds. Weiss S, Ferrand N), pp. 155-187. Springer, Dordrecht, The Netherlands.

Gonçalves H (2007) História evolutiva dos sapos-parteiros (*Alytes* spp.) na Península Ibérica. Análise filogenética e filogeográfica, reconstrução de um cenário biogeográfico e implicações taxonómicas. Tese de Doutoramento. Universidade do Porto, Portugal.

Gonçalves H, Martínez-Solano I, Ferrand N, García-París M (2007) Conflicting phylogenetic signal of nuclear vs mitochondrial DNA markers in midwife toads (Anura, Discoglossidae, *Alytes*): Deep coalescence or ancestral hybridization? *Molecular Phylogenetics and Evolution*, 44: 494-500.

Gonçalves H, Martínez-Solano I, Pereira RJ, Carvalho B, García-París M, Ferrand N (2009) High levels of population subdivision in a morphologically conserved Mediterranean toad (*Alytes cisternasii*) result from recent, multiple refugia: evidence from mtDNA, microsatellites and nuclear genealogies. *Molecular Ecology*, 18: 5143-5160.

Gotelli NJ, Stanton-Geddes J (2015) Climate change, genetic markers and species distribution modelling. *Journal of Biogeography*, 42: 1577–1585.

Griffiths AJF, Wessler SR, Carroll SB, Doebley J (2010) Introduction to Genetic Analysis. 10th Edition, W. H. Freeman and Company,USA, 800 pp.

Grossenbacher K (1997) *Alytes obstetricans* (Laurenti, 1768). In: Atlas of amphibians and reptiles in Europe (eds. Gasc JP *et al.*), pp. 94–95. Societas Europaea Herpetologica, Paris.

Guicking D, Joger U, Wink M (2008) Molecular phylogeography of the viperine snake *Natrix maura* (Serpentes: Colubridae): evidence for strong intraspecific differentiation. *Organisms Diversity & Evolution*, 8: 130–145.

Haeckel E (1866) Generelle Morphologie der Organismen. Reimer, Berlin.

Haffer J (1969) Speciation in Amazonian forest birds. *Science*, 165: 131–137.

Haig SM, Beever EA, Chambers SM, Draheim HM, Dugger BD, Dunham S, Elliott-Smith E, Fontaine JB, Kesler DC, Knaus BJ, Lopes IF, Loschl P, Mullins TD, Sheffield LM (2006) Taxonomic considerations in listing subspecies under the U.S. Endangered Species Act. *Conservation Biology*, 20: 1584–1594.

Hare MP (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology & Evolution*, 16: 700-706.

Hedrick P (2011) Genetics of Population. 4th Edition, Jones & Bartlett Learning, Burlington, Massachusetts, USA, 675 pp.

Heled J, Drummond AJ (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, 27: 570–80.

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 B: Biological Sciences*, 359: 183–195.

Hewitt GM (2011) Quaternary phylogeography: The roots of hybrid zones. *Genetica*, 139: 617-638.

Hey J, Pinho C (2012) Population genetics and objectivity in species diagnosis. *Evolution*, 66: 1413-1429.

Hickerson MJ, Carstens BC, Cavender-bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, 54: 291–301.

Hofmann S (2012) Population genetic structure and geographic differentiation in the the hot spring snake *Thermophis baileyi* (Serpentes, Colubridae): indications for glacial refuges in southern-central Tibet. *Molecular Phylogenetics and Evolution*, 63: 396–406.

Hoffmann M, Hilton-Taylor C, Angulo A, *et al.* (2010) The impact of conservation on the status of the world's vertebrates. *Science*, 330: 1503-1509.

IUCN (2016) The IUCN red list of threatened species. Versão 2014.1. Disponível em: http://www.iucnredlist.org (acessado em Abril de 2016).

Kimura M, Crow JF (1964) The Number of Alleles That Can Be Maintained in a Finite Population. *Genetics*, 49: 725–738.

Kimura M, Ohta T (1978) Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences of the United States of America*, 75: 2868–2872.

Kingman JFC (1982) The coalescent. *Stochastic Processes and their Applications*, 13: 235-248.

Kingman JFC (2000) Origins of the coalescent 1974–1982. *Genetics*, 156: 1461–1463.

Knowles LL (2009) Statistical phylogeography. *Annual Review of Ecology, Evolution, and Systematics*, 40: 593–612.

Koch PL, Barnosky AD (2006) Late Quaternary Extinctions: State of the Debate. *Annual Review of Ecology, Evolution and Systematics*, 37: 215–50.

Kubatko LS, Gibbs HL, Bloomquist EW (2011) Inferring Species-Level Phylogenies and Taxonomic Distinctiveness Using Multilocus Data in *Sistrurus rattlesnakes*. *Systematic Biology*, 60: 393-409.

Kuhner MK (2009) Coalescent genealogy samplers: windows into population history. *Trends in Ecology & Evolution*, 24: 86–93.

Leaché AD, Harris RB, Rannala B, Yang Z (2014) The influence of gene flow on species tree estimation: a simulation study. *Systematic Biology*, 63: 17-30.

Leaché AD, Reeder TW (2002) Molecular systematics of the eastern fence lizard (*Sceloporus undulatus*): a comparison of parsimony, likelihood and Bayesian approaches. *Systematic Biology*, 51: 44–68.

Levin HL (2005) The Earth Through Time. 8th Edition, John Wiley & Sons, Inc. Hoboken, New Jersey, USA, 616pp.

Llorente GA, Montori A, Santos X, Carretero MA (1995) Atlas dels amfibis i rèptils de Catalunya i Andorra. El Brau, Figueres, 192pp.

Lomolino MV, Riddle BT, Whittaker RJ, Brown JH (2010) Biogeography. 4th Edition, Sinauer Associates, Inc. Sunderland. Massachusetts, USA, 764 pp.

Loureiro A, Ferrand de Almeida N, Carretero MA, Paulo OS (2008) Atlas dos Anfíbios e Répteis de Portugal. 1ª edição, Instituto da Conservação da Natureza e da Biodiversidade. Lisboa, Portugal, 257 pp.

Márquez R, Rosa HD (1997) *Alytes obstetricans* (Laurenti, 1768). In: Distribución y Biogeografía de los Anfibios y Reptiles en España y Portugal (ed. Pleguezuelos JM), pp. 131–133. Universidad de Granada, Asociación Herpetológica Española, Granada.

Martínez-Freiría F, Santos X, Pleguezuelos JM, Lizana M, Brito JC (2009). Geographical patterns of morphological variation and environmental correlates in contact zones: a multi-scale approach using two Mediterranean vipers (Serpentes). *Journal of Zoological Systematics and Evolutionary Research*, 47: 357–367.

Martínez-Freiría F, Sillero N, Lizana M, Brito JC (2008) GIS-based niche models identify environmental correlates sustaining a contact zone between three species of European vipers. *Diversity and Distributions*, 14: 452–461.

Martínez-Freiría F, Velo-Antón G, Brito JC (2015) Trapped by climate: interglacial refuge and recent population expansion in the endemic Iberian adder *Vipera seoanei*. *Diversity and Distributions*, 21: 331–344.

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, García-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.

Maxson LR, Szymura JM (1984) Relationships among Discoglossid frogs: an albumin perspective. *Amphibia–Reptilia*, 5: 245–252.

Mayden RL (1997) A hierarchy of species concepts: the denoument in the saga of the species problem, In: Species: The units of diversity (Eds. Claridge MF, Dawah HA, Wilson MR), pp. 381-423. Springer Netherlands.

Mayr E. (1942) Systematics and the origin of species. Columbia University Press, New York, NY, 372 pp.

Mayr E (1982) Of what use are subspecies? *Auk*, 99: 593– 595.

McCormack JE, Zellmer AJ, Knowles LL (2010) Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation? Insights from tests with niche models. *Evolution*, 64: 1231-1244.

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.

Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution*, 49: 718–726.

Moritz C (1994) Defining "Evolutionarily Significant Units" for conservation. *Trends in Ecology & Evolution*, 9: 373-375.

Muschick M, Indermaur A, Salzburger W (2012) Convergent Evolution within an Adaptive Radiation of Cichlid Fishes. *Current Biology*, 22: 2362-2368.

Nakhleh L (2013) Computational approaches to species phylogeny inference and gene tree reconciliation. *Trends in Ecology & Evolution*, 28: 719-728.

Nielsen R, Beaumont MA (2009) Statistical inferences in phylogeography. *Molecular Ecology*, 18: 1034–1047.

O'Meara BC (2010) New heuristic methods for joint species delimitation and species tree inference. *Systematic Biology*, 59: 59–73.

O'Meara BC (2012) Evolutionary Inferences from Phylogenies: A Review of Methods. *Annual Review of Ecology, Evolution and Systematics*, 43: 267-285.

Pérez-Tris J, Bensch S, Carbonell R, Helbig AJ, Tellería JL (2004) Historical Diversification of Migration Patterns in a Passerine Bird. *Evolution*, 58: 1819-1832.

Phillimore AB, Owens IPF (2006) Are subspecies useful in evolutionary and conservation biology? *Proceedings of the Royal Society B*, 273: 1049-1053.

Pinho C, Hey J (2010) Divergence with Gene Flow: Models and Data. *Annual Review of Ecology, Evolution and Systematics*, 41: 215-230.

Pinho C, Kaliontzopoulou A, Carretero MA, Harris DJ, Ferrand N (2009). Genetic admixture between the Iberian endemic lizards *Podarcis bocagei* and *Podarcis carbonelli*: evidence for limited natural hybridization and a bimodal hybrid zone. *Journal of Zoological Systematics and Evolutionary Research*, 47: 368–377.

Pinho C, Kaliontzopoulou A, Harris DJ, Ferrand N (2011) Recent evolutionary history of the Iberian endemic lizards *Podarcis bocagei* (Seoane 1884) and *P. carbonelli* Pérez-Mellado 1981 (Squamata: Lacertidae) revealed by allozyme and microsatellite markers. *Zoological Journal of the Linnean Society*, 162: 184-200.

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

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155: 945–59.

Putman AI, Carbone I (2014) Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution*, 22: 4399-4428.

Pybus OG, Rambaut A (2009) Evolutionary analysis of the dynamics of viral infectious disease. *Nature Reviews Genetics*, 10: 540-550.

Pyron RA (2015) Post-molecular systematics and the future of phylogenetics. *Trends in Ecology & Evolution*, 30: 384–389.

Rato C, Harris DJ, Perera A, Carvalho SB, Carretero MA, Rödder D (2015) A Combination of Divergence and Conservatism in the Niche Evolution of the Moorish Gecko, *Tarentola mauritanica* (Gekkota: Phyllodactylidae). *PLoS One*, 10: e0127980.

Recuero E, García-París M (2011) Evolutionary history of *Lissotriton helveticus*: Multilocus assessment of ancestral vs. recent colonization of the Iberian Peninsula. *Molecular Phylogenetics and Evolution*, 60: 170–182.

Reilly SB, Wake DB (2015) Cryptic diversity and biogeographical patterns within the black salamander (*Aneides flavipunctatus*) complex. *Journal of Biogeography*, 42: 280–291.

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

Riberon A, Miaud C, Grossenbacher K, Taberlet P (2001) Phylogeography of the Alpine salamander, *Salamandra atra* (Salamandridae) and the influence of the Pleistocene climatic oscillations on population divergence. *Molecular Ecology*, 10: 2555-2560.

Rosa HD, Viegas AM, Crespo EG (1990) Genetic structure of Portuguese populations of midwife toads, with special reference to an isolate of *Alytes obstetricans*. *Portugaliae Zoologica*, 1: 15-25.

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.

Sammler S, Bleidorn C, Tiedemann R (2011) Full mitochondrial genome sequences of two endemic Philippine hornbill species (Aves: Bucerotidae) provide evidence for pervasive mitochondrial DNA recombination. *BMC Genomics*, 12: 35.

San Mauro D (2010) A multilocus timescale for the origin of extant amphibians. *Molecular Phylogenetics and Evolution*, 56: 554-61.

Sanchiz B (1984) Análisis filogenético de la tribu Alytini (Anura, Discoglossidae) mediante el estudio de su morfoestructura ósea. In: Historia Biológica del Ferreret (eds. Hemmer H, Alcover JA), pp. 61–108. Moll, Mallorca.

Sanchiz B (1998). Encyclopedia of Paleoherpetology, Part IV (Salientia). Friedrich Pfeil, München.

Sanchiz B, Adrover R (1979) Anfibios fósiles del Pleistoceno de Mallorca. *Doñana, Acta Vertebrata*, 4: 5-25.

Sanchiz B, Alcover JA (1982) Un nou discoglossid (Amphibia: Anura) de l'Holocé de Menorca. *Bulleti de la Institució Catalana d'Historia Natural*, 48: 99-105.

Savage AE, Zamudio KR (2011) MHC genotypes associate with resistance to a frog-killing fungus. *Proceedings of the National Academy of Sciences*, 108: 16705–16710.

Schmitt T (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, 4: 11.

Schield DR, Card DC, Adams RH, Jezkova T, Reyes-Velasco J, Proctor FN, Spencer CL, Herrmann HW, Mackessy SP, Castoe TA (2015) Incipient speciation with biased gene flow between two lineages of the Western Diamondback Rattlesnake (*Crotalus atrox*). *Molecular Phylogenetic and Evolution*, 83: 213-223.

Serra-Varela MJ, Grivet D, Vincenot L, Broennimann O, Gonzalo-Jiménez J, Zimmermann NE (2015) Does phylogeographical structure relate to climatic niche divergence? A test using maritime pine (*Pinus pinaster* Ait.). *Global Ecology and Biogeography*, 24: 1302–1313.

Shaffer HB, Gidis M, McCartney-Melstad E, Neal KM, Oyamaguchi HM, Tellez M, Toffelmier EM (2015) Conservation Genetics and Genomics of Amphibians and Reptiles. *The Annual Review of Animal Biosciences*, 3: 113 – 38.

Snustad & Simmons (2011) Principles of Genetics. 6th Edition, John Wiley & Sons, Inc. Hoboken, New Jersey, USA, 784pp.

Soberón J, Nakamura M (2009) Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America*, 106: 19644–19650.

Sousa-Neves T, Aleixo A, Sequeira F (2013) Cryptic patterns of diversification of a widespread Amazonian woodcreeper species complex (Aves: Dendrocolaptidae) inferred from multilocus phylogenetic analysis: implications for historical biogeography and taxonomy. *Molecular Phylogenetics and Evolution*, 68: 410–24.

Stanley EL, Bauer AM, Jackman TR, Branch WR, Mouton PLFN (2011). Between a rock and a hard polytomy: rapid radiation in the rupicolous girdled lizards (Squamata: Cordylidae). *Molecular Phylogenetics and Evolution*, 58: 53–70.

Stewart JR (2009) The evolutionary consequence of the individualistic response to climate change. *Journal of Evolutionary Biology*, 22: 2363–2375.

Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia revisited: individualistic responses of species in space and time. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 277: 661–671.

Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues ASL, Fischman DL, Waller RW (2004) Status and trends of amphibian declines and extinctions worldwide. *Science*, 306: 1783-1786.

Swenson NG (2008) The past and future influence of geographic information systems on hybrid zone, phylogeographic and speciation research. *Journal of Evolutionary Biology*, 21: 421–434.

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

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

Tarroso P, Pereira RJ, Martínez-Freiría F, Godinho R, Brito JC (2014) Hybridization at an ecotone: ecological and genetic barriers between three Iberian vipers. *Molecular Ecology*, 23: 1108–1123.

The Marie Curie Speciation Network (2012) What do we need to know about speciation? *Trends in Ecology & Evolution*, 27: 27-39.

Thomassen HA, Cheviron ZA, Freedman AH, Harrigan RJ, Wayne RK, Smith TB (2010) Spatial modeling and landscape level approaches for visualizing intraspecific variation. *Molecular Ecology*, 19: 3532–3548.

Thomson RC, Wang IJ, Johnson JR (2010) Genome-enabled development of molecular markers for ecology, evolution and conservation. *Molecular Ecology*, 19: 2184–2195.

Tocchio LJ, Gurgel-Gonçalves R, Escobar LE, Peterson AT (2015) Niche similarities among white-eared opossums (Mammalia, Didelphidae): Is ecological niche modelling relevant to setting species limits? *Zoologica Scripta*, 44: 1–10.

Torstrom SM, Pangle KL, Swanson BJ (2014) Shedding subspecies: The influence of genetics on reptile subspecies taxonomy. *Molecular Phylogenetics and Evolution*, 76: 134–143.

Tsai YE, Carstens BC (2013). Assessing model fit in phylogeographical investigations: an example from the North American sandbar willow *Salix melanopsis*. *Journal of Biogeography*, 40: 131–141.

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.

Veith M, Baumgart A, Dubois A, Ohler A, Galán P, Vieites DR, Nieto-Román S, Vences M (2012) Discordant patterns of nuclear and mitochondrial introgression in Iberian populations of the European common frog (*Rana temporaria*).*The Journal of Heredity*, 103: 240-249.

Vences M, Hauswaldt JS, Steinfartz S, Rupp O, Goesmann A, Künzel S, Smirnov NA (2013) Radically different phylogeographies and patterns of genetic variation in two European brown frogs, genus *Rana*. *Molecular Phylogenetics and Evolution*, 68: 657-670.

Viegas AM, Crespo EG (1985) Sur la structure génétique de deux “populations” allopatriques d’*Alytes obstetricans boscai* et d’*Alytes cisternasii* (Amphibia, Discoglossidae) du Portugal. Alytes, 4: 1-11.

Waltari E, Hijmans RJ, Peterson AT, Nyári ÁS, Perkins SL, Guralnick RP (2007) Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *Plos One*, 2: e563.

Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, 62: 2868– 2883

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

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ (2002) Phylogenies and Community Ecology. *Annual Review of Ecology and Systematics*, 33: 475-505.

Weiss S, Ferrand N (2007) Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origin and Conservation of European Biodiversity. Springer, Dordrecht, The Netherlands, 377pp.

Wheeler QD and Meier R (2000) Species concepts and phylogenetic Theory: A Debate. Columbia University Press. New York, USA 230 pp.

Wiley EO and Mayden RL (2000) The evolutionay species concept. In: Species concepts and phylogenetic Theory: A Debate (Eds. Wheeler QD, Meier R), pp70-89. Columbia University Press. New York, USA.

Wilson EO, Brown WL (1953) The subspecies concept and its taxonomic application. *Systematic Zoology*, 2: 97–111.

Wink M (2006) Use of DNA markers to study bird migration. *Journal of Ornithology*, 147: 234-244.

Wood DA, Fisher RN, Vandergast AG (2014) Fuzzy boundaries: Color and gene flow patterns among parapatric lineages of the Western shovel-nosed snake and taxonomic implication. *PLoS ONE*, 9: 16–21.

Wright S. (1940) The statistical consequences of Mendelian heredity in relation to speciation. In: The New Systematics (Ed. Huxley J), pp. 161–183. Oxford University Press. London.

Yang SJ, Yin ZH, Ma XM, Lei FM (2006) Phylogeography of ground tit (*Pseudopodoces humilis*) based on mtDNA: Evidence of past fragmentation on the Tibetan Plateau. *Molecular Phylogenetics and Evolution*, 41: 257–265.

Yang Z, Rannala B (2012) Molecular phylogenetics: principles and practice. *Nature Reviews Genetics*, 13: 303-14.

Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America*, 107: 9264–9269.

Zeisset I, Beebee TJC (2008) Amphibian phylogeography: a model for understanding historical aspects of species distributions. *Heredity*, 101: 109-119.

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

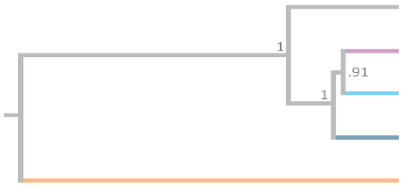
Zink RM (2004) The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proceedings of the Royal Society of London B*, 271: 561–564.

Capítulo 2

Relações filogenéticas dos sapos-parteiros (*Alytes* spp.)

Artigo I

Maia-Carvalho B, Gonçalves H, Ferrand N and Martínez-Solano I (2014) Multilocus assessment of phylogenetic relationships in *Alytes* (Anura, Alytidae). *Molecular Phylogenetics and Evolution* 72, 270 – 278. doi: 10.1016/j.ympev.2014.05.033



Artigo I

Original Article
Accepted 28 May 2014

Multilocus assessment of phylogenetic relationships in *Alytes* (Anura, Alytidae)

Bruno Maia-Carvalho[1,2], Helena Gonçalves[1], Nuno Ferrand[1,2], and Iñigo Martínez-Solano[1,3]

Molecular Phylogenetics and Evolution, 72: 270 – 278
doi: 10.1016/j.ympev.2014.05.033

[1]*CIBIO, Centro de Investigacão em Biodiversidade e Recursos Genéticos da Universidade do Porto, InBIO 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]*Instituto de Investigación en Recursos Cinegéticos (CSIC-UCLM-JCCM), Ronda de Toledo s/n, 13005 Ciudad Real, Spain*

2.1 Multilocus assessment of phylogenetic relationships in *Alytes* (Anura, Alytidae)

2.1.1 Abstract

With the advent of large multilocus datasets, molecular systematics is experiencing very rapid progress, but important challenges remain regarding data analysis and interpretation. Midwife toads (genus *Alytes*) exemplify two of the most widespread problems for accurate phylogenetic reconstruction: discerning the causes of discordance between gene trees, and resolving short internodes produced during rapid, successive splitting events. The three species in subgenus *Baleaphryne* (*A. maurus*, *A. dickhilleni* and *A. muletensis*), the sister group to *A. obstetricans*, have disjunct and highly restricted geographical ranges, which are thought to result from old vicariant events affecting their common ancestor, but their phylogenetic relationships are still unresolved. In this study we re-address the phylogeny of *Alytes* with a special focus on the relationships in *Baleaphryne* with a multilocus dataset including >9000 base pairs of mitochondrial DNA and four nuclear markers (3142 bp) in all recognized taxa, including all subspecies of *A. obstetricans*. Both concatenation and species tree analyses suggest that *A. muletensis*, endemic to the Balearic island of Mallorca, is the sister taxon to a clade comprising the southeastern Iberian endemic *A. dickhilleni* and the North African *A. maurus*. This scenario is consistent with palaeogeological evidence associated with the fragmentation of the Betic-Rifean Massif, followed by the opening of the Strait of Gibraltar. On the other hand, analyses of intraspecific variation in *A. obstetricans* are inconclusive regarding relationships between major clades and conflict with current subspecific taxonomy.

Key words: Alytes, Baleaphryne, mtDNA, Multispecies coalescent, Species trees, Concatenation

2.1.2 Introduction

Molecular systematics is experiencing rapid advances in the reconstruction of historical relationships among organisms. New developments are the consequence of the increasing feasibility of compiling large multilocus datasets and the incorporation of new analytical tools in phylogenetic inference (O'Meara 2012, Yang and Rannala 2012). This has led to a shift in interest from gene tree to species tree inference (Edwards 2009, Corl and Ellegren 2013, Reid *et al.* 2014). However, multilocus datasets present new theoretical and computational challenges for the inference of species trees. For instance, events of gene tree incongruence have been widely reported in many taxa, and their alternative interpretations in terms of evolutionary processes complicate phylogenetic reconstruction (Degnan and Rosenberg 2009, Nakhleh 2013). Among the most common causes of discordance between gene trees are incomplete lineage sorting and ancient or recent hybridization or gene flow (Belfiore *et al.* 2008, Corl and Ellegren 2013). These processes are usually hard to discriminate because they produce similar molecular signatures, but their consequences for the interpretation of speciation processes are very different (Leaché *et al.* 2014). Another challenge is the difficulty in distinguishing ''soft'' (where lack of support is typically related to lack of informative characters) versus ''hard'' polytomies, (in which population lineages are not strictly bifurcating, as in rapid radiations, see Kubatko *et al.* 2011, Stanley *et al.* 2011).

In order to solve or minimize the impact of these problems, different alternatives have been proposed (Camargo *et al.* 2012). For instance, enlarged taxon and data sampling (genes or alleles – Belfiore *et al.* 2008, Chung and Ané, 2011 -, as well as the use of larger gene fragments, e.g. mitogenomes – Pabijan *et al.* 2013, Zhang *et al.* 2013) have been used to resolve conflicting nodes in many groups of organisms (Whitield and Lockhart 2007, Pacheco *et al.* 2011, Kapralov *et al.* 2013, Sanders *et al.* 2013), including amphibians (Espregueira-Themudo *et al.* 2009, Garcia-Porta *et al.* 2012, Recuero *et al.* 2012, Williams *et al.* 2013). While amphibians are excellent models for studies of speciation due to their low vagility and geographically structured patterns of genetic variation (Vences and Wake 2007, San Mauro 2010), inference of phylogenetic relationships has been often complicated by findings of strong discordance across datasets, sometimes related to rapid diversification (Blackburn *et al.* 2013). Midwife toads (*Alytes* spp.), a species-poor genus including representatives of an ancient anuran clade illustrate these problems.

The genus *Alytes* is divided into five species, which occur throughout the Western Mediterranean Basin (Fig. 1). These species are classified in three subgenera, each of which is in principle a monophyletic group (but see Gonçalves *et al.* 2007). Subgenus *Alytes* includes *Alytes obstetricans* (Laurenti, 1768), widely distributed in Western Europe, with four currently

recognized subspecies (Fig. 1). Subgenus *Ammoryctis* includes *Alytes cisternasii* Boscá, 1879, endemic to the center and southwest of the Iberian Peninsula. Finally, subgenus *Baleaphryne* comprises three species: *Alytes muletensis* (Sanchiz and Adrover, 1979) from the island of Mallorca in the Balearic archipelago; *Alytes dickhilleni* Arntzen and García-París, 1995, endemic to the Betic Mountains in southeastern Spain, and *Alytes maurus* Pasteur and Bons 1962, which is distributed in northern Africa in Morocco (Rif and Middle Atlas Mountains) and Algeria (Bons and Geniez 1996, Grossenbacher 1997, Crespo 1997, García-París and Arntzen 2002, Román 2002, Beukema *et al.* 2013, Mateo *et al.* 2013, de Pous *et al.* 2013).

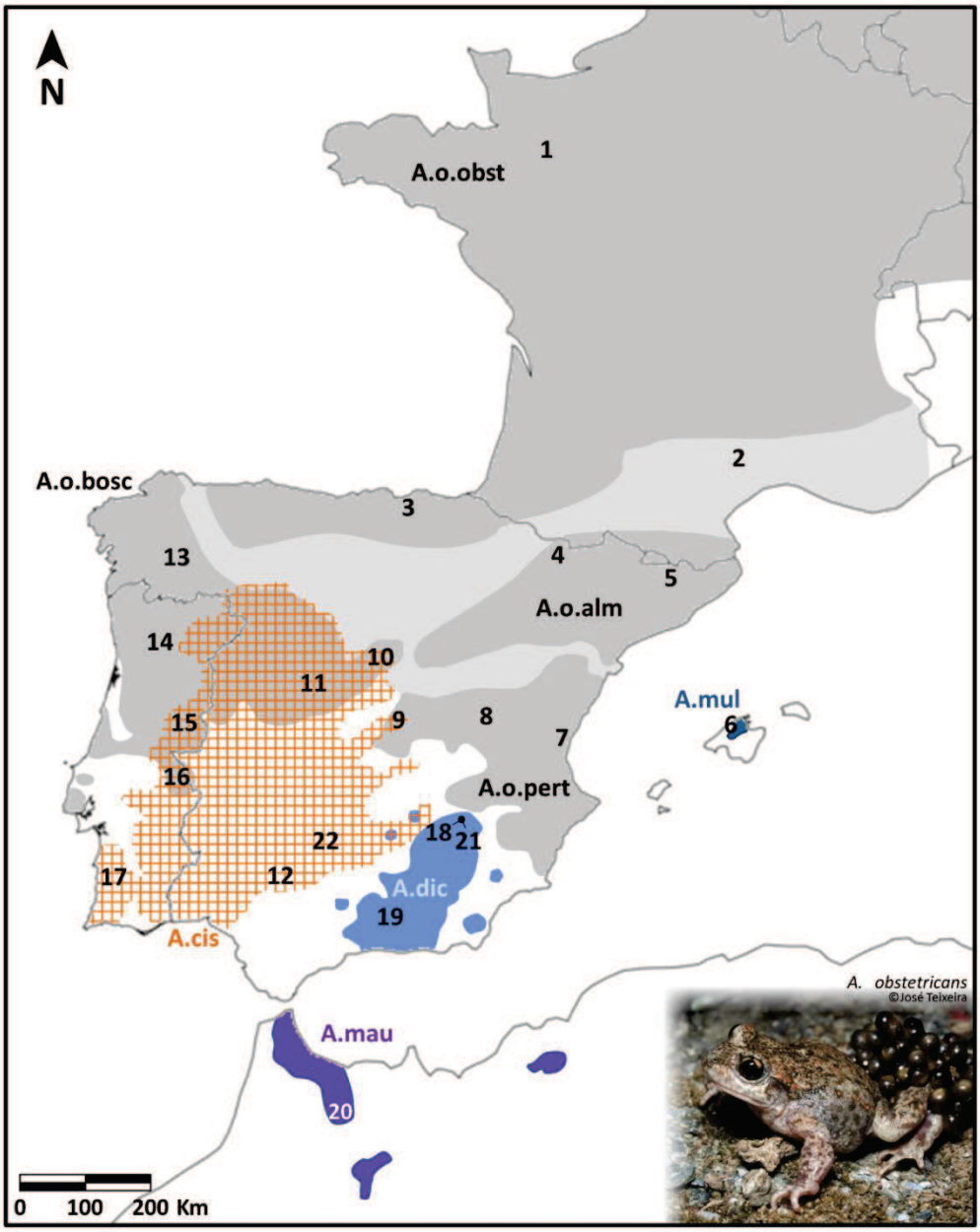


Fig. 1. Distribution map of *Alytes* species and subspecies with indication of the geographical origin of all samples analyzed in the present study. Light gray shading represents areas where the assignment to the different *Alytes obstetricans* subspecies is doubtful (see for instance Maia-Carvalho *et al.* 2014). Sample codes are described in Table 1.

Phylogenetic relationships in *Alytes* have been addressed in several studies based on morphology, genetic data, and both types of data combined (Arntzen and García-París 1995, Fromhage *et al.* 2004, Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007, Biton *et al.* 2013), but some questions remain open. In particular, phylogenetic relationships between the three members of the subgenus *Baleaphryne* are still unresolved. The origin and diversification of this clade dates back to the Miocene and seems to be related to fragmentation of the Betic-Rifean Massif and the Messinian Salinity crisis ending with the opening of the Strait of Gibraltar about 5.33 mya, but the relative splitting order between species is apparently very close in time, precluding full resolution of the trichotomy in previous studies (Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007). Moreover, Gonçalves *et al.* (2007) questioned the monophyly of subgenus *Alytes* based on the finding of *Beta-fibrinogen intron 7* (*β-fibint7*) haplotypes in the most divergent lineage within *A. obstetricans* (subspecies *A. o. almogavarii*) that grouped with *Baleaphryne*, suggesting the possibility of ancient interspecific gene flow or, alternatively, the persistence of shared ancestral polymorphisms across subgenera. In this study we re-address phylogenetic relationships in *Alytes* with a special focus on the relationships in *Baleaphryne* with a multilocus dataset including >9000 base pairs of mtDNA and four nuclear markers (3142 bp) in all recognized taxa. Different analyses suggest a resolution of the polytomy with *A. muletensis* being the sister taxon to *A. dickhilleni* + *A. maurus* and incomplete lineage sorting as the most likely source of discordance in *β-fibint7*. We discuss the implications of this new finding for elucidation of the evolutionary history of *Alytes*.

2.1.3 Material and methods

2.1.3.1 Sample collection

A total of 32 individuals was analyzed in the present study, representing all the recognized species and subspecies of the genus *Alytes* (Fig. 1 and Table 1). Some of these samples were previously used by Gonçalves *et al.* (2007, 2009) and Pinho *et al.* (2010), and we incorporated some sequences from those studies in our analyses (see Supplementary Table S1). Newly collected tissue samples were obtained from toe tips of adults or tail tips of larvae, and preserved in 95% ethanol. All individuals were released back in their place of capture after tissue collection.

Genomic DNA was extracted using EasySpin Genomic DNA Minipreps Tissue Kit (SP-DT-250, Qiagen, Hilden, Germany) following the fabricant's protocol. We amplified by polymerase chain reaction (PCR) several fragments of mitochondrial DNA (mtDNA) (see

below) and fragments of four nuclear genes: Protein phosphatase 3, catalytic subunit, alpha isoform (*PPP3CAint4*), Ribosomal protein L9 intron 4 (*RPL9int4*), Cellular myelocytomatosis (*C-myc*) and Beta-fibrinogen intron 7 (*β-fibint7*). PCRs were performed in 10 μL reaction volumes containing 5 μL of Phusion Master Mix (Thermo Scientific, Waltham, Massachusetts, USA), 0.2 mM each primer and 50 ng of genomic DNA.

Table 1. Locality information of *Alytes* samples used in the study. The ''captivity'' samples (*A. muletensis*) are from a captive breeding programme at Jersey Zoo. Taxon code as in Figs. 2, 3B, S1 and S2. Map code as in Fig. 1.

Taxa	Taxon code	Voucher	Population	Country	Map code	Latitude	Longitude
A. o. pertinax	A.o.pert	MAD4	Belmonte de Tajo, Madrid	Spain	9	40.13	-3.33
		CUE1	Buenache de la Sierra, Cuenca	Spain	8	40.13	-2.00
		VLC02	Algar de Palancia, Valencia	Spain	7	39.78	-0.36
A. o. obstetricans	A.o.obst	FRAN14	St Pierre-de-la-Fage	France	2	43.79	3.42
		FRAN01	Jublains	France	1	48.25	-0.50
		SAN01	Puerto de la Magdalena, Cantabria	Spain	3	43.35	-3.33
A. o. boscai	A.o.boscN	OUR03	Penalba, Ourense	Spain	13	43.42	-7.73
	A.o.boscS	MTM02	Cinfães, Serra de Montemuro	Portugal	14	41.07	-8.02
		MTM05	Cinfães, Serra de Montemuro	Portugal	14	41.07	-8.02
		SMA25	Rib. S. Bento, Serra S. Mamede	Portugal	16	39.31	-7.41
A. o. almogavarii	A.o.alm	HUE1	Ibón de Piedrafita, Huesca	Spain	4	42.70	-0.33
		BER1	Rasos de Peguera, Berga, Barcelona	Spain	5	42.13	1.76
		BER3	Rasos de Peguera, Berga, Barcelona	Spain	5	42.13	1.76
Alytes maurus	A.mau	MAR03	Bab Bou Idir, Taza	Morocco	20	34.06	-4.11
		MAR04	Bab Bou Idir, Taza	Morocco	20	34.06	-4.11
		MAR05	Bab Bou Idir, Taza	Morocco	20	34.06	-4.11
		MAR06	Bab Bou Idir, Taza	Morocco	20	34.06	-4.11
Alytes dickhilleni	A.dic	HG103	Puerto de las Crucetillas, Sierra de Alcaraz, Albacete	Spain	18	38.54	-2.38
		HG104	Puerto de las Crucetillas, Sierra de Alcaraz, Albacete	Spain	18	38.54	-2.38
		Aly120D	Fuente Alta, Illora, Granada	Spain	19	37.31	-3.86
		IMS4189	Riopar Viejo, Albacete	Spain	21	38.50	-2.44
Alytes muletensis	A.mul	CAP01	Captivity	Spain	6	-	-
		AM02	Captivity	Spain	6	-	-
		MAI06	Sierra de Tramuntana, Mallorca	Spain	6	39.58	2.50

Table 1. Continued

Taxa	Taxon code	Voucher	Population	Country	Map code	Latitude	Longitude
Alytes cisternasii	A.cis	CER17	Cercal	Portugal	17	37.75	-8.65
		HG112	Río Adaja, Ávila	Spain	11	40.65	-4.70
		MAD01	Montejo de la Sierra, Madrid	Spain	10	41.06	-3.53
		IDN03	Idanha-a-Nova	Portugal	15	40.00	-7.23
		NAV22	Las Navas de la Concepción, Constantina	Spain	12	37.86	-5.61
		IMS2004	Alcaracejos, Córdoba	Spain	22	38.38	-4.96

mtDNA. We amplified and sequenced the fragment *ND4-tRNA* using the primers described by Arévalo *et al.* (1994), and following the amplification conditions of Gonçalves *et al.* (2007). Additionally, we designed several primers along the mitochondrial genome with melting temperatures (Tm) around 56°C to avoid nonspecific amplifications. Primers were combined in different ways to amplify templates between 1000 and 2000 base pairs. A touchdown PCR program was applied in all cases, consisting of: 98°C for 3 min (pre-denaturing step); 1st round (9 cycles) of 98°C for 30 s, 60–56°C (decreasing 0.5°C in each cycle) for 40 s, and 72°C for 40 s; 2nd round (31 cycles) of 98°C for 30 s, 56°C for 30 s, and 72°C for 40 s, with a final extension at 72°C for 5 min. Successful amplification products of the same fragment for all taxa (Table 2) were sequenced in both forward and reverse directions.

Table 2. Primers used for the PCR amplification of ten mtDNA and four nuclear gene fragments in *Alytes*.

Fragment amplified	Estimated size (bp)	Primer Name	Sequence (5' - 3')	Reference
mtDNA				
RNA1	1000	Aly406F	GTCGCACGCTTCAGTTG	Present study
		Aly1405R	TAGGCTTGTCACCTCTACTC	
RNA2	1017	Aly1184F	GTAAGGGAAAGATGAAATAGC	Present study
		Aly2200R	CTTGCTTGTGGTTAGATGAG	
ND1	1220	Aly2623F	GACAAAAATTGCAACTAAGC	Present study
		Aly3842R	ACTTTACTAGGAAGTGGCATA	
ND2	1343	Aly3825F	GCCACTTCCTAGTAAAGTCA	Present study
		Aly5167R	GGAGAAGTAGAATGAAGCTC	
COX1	1811	Aly5129F	GCTAAACGCTCAATCCAG	Present study
		Aly6940R	AATGACAGAGTGGTGATGTG	
COX2-APT6	1915	Aly6867F	GCCACTAACGAGAAAAGAG	Present study
		Aly8781R	TCGTTAGAAGTATGGTGATTAG	
COX3-ND4L	1611	Aly8617F	GTCTAATGGCACACCAAG	Present study
		Aly10227R	GTCATAGGGCTGGAATAAG	
ND4-tRNALEU	880	ND4	CACCTATGACTACCAAAAGCTCATGTAGAAGC	Arévalo *et al.* (1994)
		Leu	CATTACTTTTACTTGGATTTGCACCA	
ND5	1185	Aly11732F	CAAAGCCTCTTGGTGCAACT	Present study
		Aly12916R	GGTTCCGGTTAGGGCTAGG	
ND6	718	Aly13407F	AGCTCTAGTACCACAAACAAA	Present study

Table 2. (Continued)

Fragment amplified	Estimated size (bp)	Primer Name	Sequence (5' - 3')	Reference
nDNA				
β-fibint7	900	FIBX7	GGAGANAACAGNACNATGACAATNCAC	Sequeira *et al.* (2006)
		FIBX8	ATCTNCCATTAGGNTTGGCTGCATGGC	
	700	BFXF	CAGYACTTTYGAYAGAGACAAYGATGG	
		BFXR	TTGTACCACCAKCCACCACCRTCTTC	
PPP3CAint4	700	PPP3CA4F1	CTGTAYTTGTGGGCCTTGAAAATTC	Pinho *et al.* (2010)
		PPP3CA5R1	AAGGCATCCATGCAGGCATCATATA	
C-myc	1200	Cmyc1U	GAGGACATCTGGAARAARTT	Crawford (2003)
		Cmyc3cat	GTTGYTGCTGATCTGTTTGAG	Brunes *et al.* (2010)
RPL9int4	500	RPL94F	CGTGTKGACAAATGGTGGGGTAA	Pinho *et al.* (2010)
		RPL95R	ATGGGAAAGTGAGCRTACACAGA	

nDNA. For *PPP3CAint4* and *RPL9int4*, amplification conditions followed Pinho *et al.* (2010), with some modifications: 98°C for 3 min (pre-denaturing step); 40 cycles of 98°C for 30 s, 60°C for 20 s, and 72°C for 20 s; and a final extension at 72°C for 5 min. For *C-myc*, we used primers *Cmyc1U* (Crawford 2003) and *Cmyc3cat* (Brunes *et al.* 2010). Amplification was performed as follows: 98°C for 3 min (pre-denaturing step); 40 cycles of 98°C for 30 s, 62.5°C for 40 s, and 72°C for 40 s; and a final extension at 72°C for 5 min. For *β-fibint7* a two-step amplification procedure, with the primers described by Sequeira *et al.* (2006), was followed. Amplification conditions were performed as described in Gonçalves *et al.* (2007). Purified products of each reaction were sequenced with the ABI Prism BigDye Terminator v3.1 Sequencing Kit protocol on an ABI3130xl DNA analyzer (Applied Biosystems, Foster City, California, USA). GenBank Accession Numbers of the new sequences are: KJ858769-KJ859062 (Table S1).

Some DNA fragments (mtDNA and nDNA) could not be obtained for a small group of samples. In these cases, to avoid biases in the analyses, we replaced the sample in question by another one of the same taxon and, when possible, of the same population of origin, as indicated in Table S1. This is equivalent to the concept of ‘‘composite taxa’’ (Campbell and Lapointe 2009, Alonso *et al.* 2012); we used this approach to increase phylogenetic accuracy. As mentioned above, some sequences were available from GenBank (*16S* and *ND4* mtDNA sequences and some *PPP3CAint4*, *RPL9int4*, *β-fibint7* nuclear sequences, see Table S1).

2.1.3.2 Sequence analyses and genetic variation

Sequences were assembled with the software CHROMASPRO v1.5 (www.technelysium.com.au/ChromasPro.html). The assembled sequences were then edited and aligned using the program BIOEDIT v7.1.3.0 (Hall 1999); this preliminary alignment was subsequently refined by eye. Since at the relatively deep phylogenetic levels investigated some saturation might be expected to occur in faster-evolving markers, we performed saturation tests in different mtDNA fragments with DAMBE v5.3.70 (Xia 2013). On the other hand, single and multiple-base insertions or deletions (indels) were observed at nuclear markers. These events were used to solve the haplotype phase in all cases (see below). Decoding was made with the direct interpretation of the mixed trace formed by the two allele peaks superimposed onto each other downstream of the indel (Sousa-Neves *et al.* 2013). Other polymorphic sites in heterozygous individuals were coded with IUPAC ambiguity codes. In these cases, the probabilistic Bayesian algorithm implemented in PHASE v2.1.1 (Stephens *et al.* 2001, Stephens and Donnelly 2003) was used to phase haplotypes. Input files were formatted with SEQPHASE (Flot 2010). Due to the low haplotype variability in our data set, we

incorporated additional sequences from a larger *A. obstetricans* dataset (Gonçalves *et al.* unpublished) into each input file to increase the precision of phased haplotypes. We ran PHASE three times with different random seeds to check the consistency of haplotype reconstructions across runs. All ambiguous positions could be resolved with posterior probabilities of 0.90 or higher. Finally, we tested for recombination in the nuclear markers through calculation of the *R* statistic (minimum number of recombination events – Hudson and Kaplan 1985), as implemented in DNASP v5.10 (Librado and Rozas 2009).

We calculated several summary statistics for each marker to describe their levels of polymorphism, using the software DNASP: number of segregating sites (*S*), nucleotide diversity (π) (Nei 1987), and Theta (θ per site, Watterson 1975). We also calculated genetic distances (*p*-uncorrected) within and between species with MEGA5 (Tamura *et al.* 2011).

2.1.3.3 Phylogenetic analyses

We performed phylogenetic analyses based on two optimality criteria: Maximum Likelihood (ML) and Bayesian Inference (BI). Different analyses were performed on: (1) the concatenated mtDNA dataset; (2) each nuclear gene analyzed separately; (3) the four nuclear genes concatenated; and (4) a concatenated mtDNA + nDNA dataset. For the concatenated mtDNA and nDNA datasets, the optimal partitioning strategies and the respective best-fit models of evolution for each partition were estimated under the Bayesian Information Criterion (Schwarz 1978) as implemented in PARTITIONFINDER v1.1.0 (Lanfear *et al.* 2012) (see Supplementary Table S2). For this, the mtDNA alignment was divided based on functional categories (genes, tRNAs and rRNAs) and into 1st, 2nd and 3rd codon position for protein-coding genes (see Brandley *et al.* 2005, Pabijan *et al.* 2013). In nDNA markers, partitioning was between intron and exon regions (see Wiens *et al.* 2010). For estimation of individual gene trees of the four nuclear DNA markers, we selected the best-fit substitution model for each gene with JMODELTEST 2.1.1, based on the Bayesian Information Criterion (Darriba *et al.* 2012)(Table S2). For the combined mtDNA + nDNA dataset, only one of the two phased alleles per individual was included, and the optimal partitioning strategies for each data type (eight for mtDNA and three for nDNA, see Table S2) were specified based on PARTITIONFINDER results.

ML analyses of both concatenated (partitioned) and single-gene datasets were conducted using GARLI v.2.0 (Zwickl 2006). In all cases, we performed 100 bootstrap replicates (BS) and a 50% majority-rule consensus tree was subsequently computed with PAUP (Swofford 2003). BI analyses were run in MRBAYES v.3.2.2 (Huelsenbeck and Ronquist 2001, Ronquist *et al.* 2012). Two replicate runs with four independent chains were run for 10 x 10^7 generations, sampling every 10,000 generations. A majority-rule consensus tree was

computed after discarding the first 25% generations. TRACER v1.6 (Rambaut and Drummond 2007) was used to check convergence of results by plotting the log-likelihood values versus generation number. Branch support was based on Bayesian Posterior Probabilities (BPPs).

In addition to concatenation approaches, we analyzed our multilocus dataset under the multispecies coalescent as implemented in *BEAST v1.8 (Drummond and Rambaut 2007; Heled and Drummond 2010). We defined 12 independent partitions: eight in the mitochondrial dataset, as previously selected by PARTITIONFINDER (see Table S2 for details), and each of the four nuclear markers independently. Substitution models were unlinked across partitions, and molecular clock and tree priors were linked in the mtDNA partitions and unlinked in the nuclear genes. As species-tree prior, we used the Yule speciation model. Preliminary runs indicated deviations from the strict clock model in three partitions: mtDNA, *β-fibint7* and *RPL9int4*, and therefore we used in subsequent analyses the uncorrelated (lognormal) relaxed clock instead of the strict clock used in the remaining partitions (*PPP3CAint4* and *C-myc*).

The fossil record of *Alytes* includes Middle Miocene remains that have been assigned to a new species predating the diversification of extant species, with a minimum age of 15.8 million years (Bastir *et al.* 2014). In order to accommodate this prior information in the analyses, we specified a lognormal prior for the root of the tree with a mean of 20 million years and a log (standard deviation) of 0.1, yielding dates from 16.36 (2.5% quantile) to 24.21 (97.5% quantile) million years. For substitution rates, we used a wide, largely uninformative prior with a lognormal distribution (mean: 0.01 substitutions per lineage per million years; log (st. dev.): 1;95% range: $1.171E^{-3}$ -$3.142E^{-2}$).

Two independent analyses were run for 300 million generations in order to check for convergence of results across runs. Logfiles were inspected in TRACER v1.6 to assess convergence. All effective sample sizes (ESS) of parameters estimated were >200, as recommended by the authors. A burn-in period of 10% of the total running time was specified and a Maximum Clade Credibility consensus tree was constructed with the program TREEANNOTATOR, which is distributed as part of the BEAST package.

Since previous studies suggested the possibility of hybridization between ancestors of *Baleaphryne* and *A. obstetricans* based on *β-fibint7* data (Gonçalves *et al.* 2007), we used the posterior predictive checking approach implemented in JML v1.02 (Joly *et al.* 2009, Joly 2012) to discriminate between this and the alternative hypothesis of incomplete lineage sorting in this marker. JML takes as input the posterior distribution of species trees from *BEAST and simulates gene trees under the coalescent with no migration. Minimum genetic distances between species in a simulated *β-fibint7* dataset were used to generate a posterior predictive distribution, which was then compared with empirical (observed) values.

Finally, in order to check for consistence of results across different species-tree inference methods, we used another species-tree reconstruction approach, the pseudo-

likelihood method implemented in MP-EST v.1.3 (Liu *et al.* 2010). To estimate support, 100 replicate gene trees for each marker were subjected to MP-EST analyses to derive bootstrap proportions of inferred clades in the species tree. Since we found variation in clade support in analyses of the same datasets under different optimization criteria, we used as alternative inputs for these analyses: (1) 100 ML trees from a bootstrap analysis in Garli; (2) 100 post burn-in trees from a MrBayes run; and (3) 100 post burn-in trees from a *BEAST analysis.

2.1.4 Results

2.1.4.1 Sequence variation and genetic distances

The final mtDNA alignment totaled 9045 bp of which 7091 bp are related with protein-coding genes (Table S1). A gap of 3 bp (nucleotide positions 763–765, according to the reference sequence of *A. obstetricans* from GenBank, Accession Number: NC_006688.1, San Mauro *et al.* 2004) was observed at the *ND1* gene (mt03 fragment, see Table S1) in the two samples of *A. cisternasii*. This codon specifies serine in *A. obstetricans* and threonine in *Baleaphryne*. For the nuclear markers, we obtained a concatenated alignment of 3142 bp, including indels in all genes. No evidence of recombination or saturation was detected in any marker (results not shown).

Supplementary Table S3 presents an overview of genetic diversity. All markers and partitions were polymorphic, with a range of 36–322 segregating sites in mtDNA and 74–85 in nDNA. Mitochondrial and nuclear genetic distances (*p*-uncorrected) between species and subgenera are in Table 3. Within *Baleaphryne*, species differed between 7.2% (*A. maurus* - *A. muletensis*) and 8.2% (*A. dickhilleni* - *A. muletensis*) in mtDNA and 1.0% (*A. maurus* - *A. muletensis*) to 1.2% (*A. maurus* - *A. dickhilleni*) in nDNA (4-gene concatenated alignment). Within *A. obstetricans*, genetic distances ranged from 0.9% (*A. o. obstetricans* and *A. o. pertinax*) to 3.4% (*A. o. boscai* and *A. o. almogavarii*) in mtDNA and from 0.5% (*A. o. obstetricans* and *A. o. pertinax*) to 1.2% (*A. o. boscai* and *A. o. pertinax*) in nDNA. *Alytes cisternasii* differed by >15% in mtDNA and >6% in nDNA from other *Alytes* species, whereas *Baleaphryne* and *A. obstetricans* differed by over 9% in mtDNA and 2% in nDNA.

Table 3. Average number of pairwise sequence differences (*p*-uncorrected distance) across all *Alytes* taxa. Mitochondrial estimates are above diagonal and nuclear estimates (four markers concatenated) are below diagonal. Taxon "*A.o.bosc*" includes samples from the Southern and Northern lineages in *A. o. boscai* (see Maia-Carvalho *et al.* 2014).

	A.obstetricans	*A.o.pertinax*	*A.o.obstetricans*	*A.o.boscai*	*A.o.almogavarii*	*Baleaphryne*	*A.maurus*	*A.dickhilleni*	*A.muletensis*	*A.cisternasii*
A.obst	0	*	*	*	*	0.091	0.091	0.096	0.088	0.153
A.o.pert	*	0	0.009	0.024	0.029	0.092	0.092	0.096	0.088	0.154
A.o.obst	*	0.005	0	0.027	0.031	0.093	0.094	0.096	0.090	0.153
A.o.bosc	*	0.012	0.010	0	0.034	0.091	0.091	0.095	0.088	0.152
A.o.alm	*	0.010	0.010	0.011	0	0.090	0.088	0.096	0.085	0.153
Baleap	0.021	0.022	0.021	0.022	0.021	0	*	*	*	0.158
A.mau	0.021	0.022	0.021	0.021	0.021	*	0	0.081	0.072	0.162
A.dic	0.022	0.023	0.022	0.023	0.022	*	0.012	0	0.082	0.156
A.mul	0.021	0.021	0.021	0.021	0.020	*	0.010	0.011	0	0.155
A.cist	0.067	0.066	0.066	0.068	0.067	0.066	0.067	0.070	0.068	0

* Non-applicable.

2.1.4.2 Phylogenetic analyses

The ML and BI trees based on the concatenated mtDNA dataset are highly concordant, including a fully resolved tree (Fig. 2a). Three main clades were recovered with high support (BPP ≥ 0.9, BS ≥ 75), corresponding to the three subgenera: *Ammoryctis* (*A. cisternasii*), *Alytes* (*A. obstetricans*), and *Baleaphryne* (*A. muletensis*, *A. dickhilleni* and *A. maurus*). *Alytes cisternasii* is the sister taxon of *A. obstetricans* + *Baleaphryne*, as in previous studies. Within *Baleaphryne*, *A. muletensis* is recovered as the sister group of *A. maurus* + *A. dickhilleni* (BS = 76; BPP = 1.0). Within *A. obstetricans*, *A. o. pertinax* and *A. o. obstetricans* are recovered as monophyletic and grouped as sister taxa. On the other hand, subspecies boscai and almogavarii are not recovered as monophyletic groups (Fig. 2a). *Alytes o. boscai* is divided into a well-supported, highly divergent southern group (samples 02S and 03S) and a northern group (sample 01N); however, in *A. o. almogavarii* the sample HUE from the Pyrenees in Huesca does not group with the other two samples (alm01 and alm02), which form a well-supported clade.

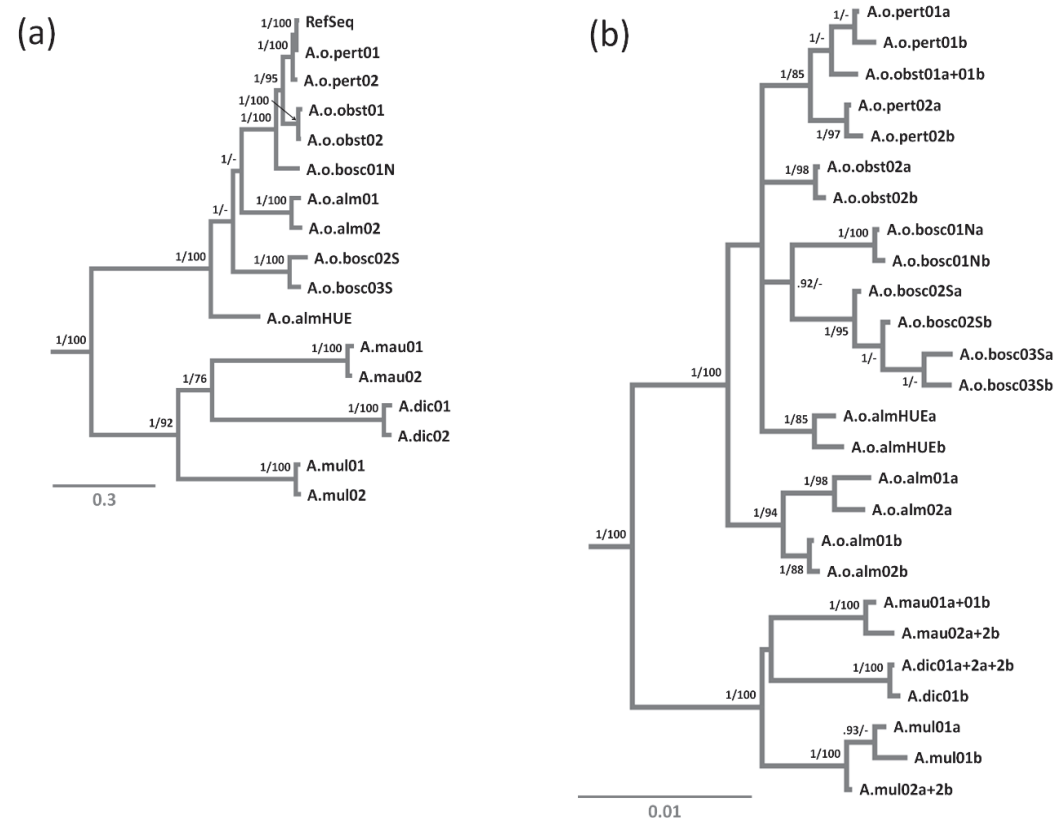


Fig. 2. Phylogenetic relationships of all recognized *Alytes* species and subspecies based on (a) mtDNA and (b) nDNA concatenated sequenced data under BI and ML analyses. Phylograms are Bayesian consensus trees; the outgroup (*A. cisternasii*) was omitted for clarity. Posterior probability and bootstrap values of well-supported nodes (BPP ≥ 0.9 / BS ≥ 75) are shown at nodes. Sample codes as in Table 1. RefSeq is the reference mitogenome of *A. o. pertinax* (GenBank Accession: NC_006688.1, San Mauro *et al.* 2004), which was included for reference.

Regarding each nuclear marker, individual gene trees had similar topologies based on ML or BI. Support for some of the clades recovered in analyses of mtDNA, based on BS and BPPs, was lower (Supplementary Figs. S1 and S2). For instance, *A. obstetricans* was recovered as a monophyletic group in *PPP3CAint4* and *C-myc* gene trees, but not in *β-fibint7* or *RPL9int4*. Monophyly of *Baleaphryne* was supported by all markers in both BI and ML analyses. However, alternative resolutions of the trichotomy in *Baleaphryne* were recovered by different genes. For instance, *C-myc* recovers *A. muletensis* as the sister group of *A. maurus* + *A. dickhilleni* with moderate (BS: 64) to high (BPP: 0.94) support, as in mtDNA, whereas *RPL9int4* groups *A. muletensis* and *A. dickhilleni* instead, although with low support (BPP < 0.9, BS < 75, Figs. S1 and S2). On the other hand, some *β -fibint7* alleles of *A. o. almogavarii* cluster with *Baleaphryne* in a well-supported clade (BPP: 1.0; BS: 79). According to JML results, this is consistent with the coalescent with no migration model, thus supporting the hypothesis of incomplete lineage sorting as the source of discordance. *PPP3CAint4*, on the other hand, does not resolve the relationships within *Baleaphryne*. The differences across nuclear gene trees are reflected in the lack of resolution of the concatenated nDNA tree, where the *Baleaphryne* trichotomy remains unresolved (Fig. 2b).

Regarding subspecies of *A. obstetricans*, there is some discordance between the different nuclear gene trees. Some clades are recovered with strong support in both ML and BI inference, although there is no correspondence with described subspecies: no subspecies is recovered as a monophyletic group across all markers (Figs. S1 and S2). When nDNA fragments are concatenated into a single alignment, BPPs and BS values for some clades increase. Samples of *A. o. pertinax* group with part of *A. o. obstetricans* in a well-supported clade (BPP: 1.0; BS: 85), the monophyly of *A. o. boscai* is well supported by BI (BPP: 0.92) but not in ML analyses, and *A. o. almogavarii* (excluding alleles from Huesca) is recovered as a monophyletic group (BPP: 1.0; BS: 94).

The concatenated mtDNA + nDNA tree is well-resolved (Fig. 3b) and similar to the mtDNA tree. Both *Baleaphryne* and *A. obstetricans* are recovered as monophyletic, and relationships within *Baleaphryne* are well-resolved, with *A. maurus* being the sister taxon to *A. dickhilleni*. Within *A. obstetricans*, *A. o. pertinax* is more closely related to *A. o. obstetricans*, whereas the other two subspecies (*A. o. boscai* and *A. o. almogavarii*) are not recovered as monophyletic, as in the mtDNA tree.

Results of *BEAST analyses are shown in Fig. 3a. The monophyly of subgenera *Alytes* and *Baleaphryne* is supported by BPPs of 1.0, and there is also support for a sister-group relationship between *A. dickhilleni* and *A. maurus* (BPP = 0.91). Time estimates (time to most recent common ancestor, or TMRCA) for major splits include median values of 3.39 million years for *Baleaphryne* (95% highest posterior density interval, HPD: 2.04–5.11 million years) and 5.57 for *Baleaphryne* + *A. obstetricans* (95% HPD: 3.43–8.16 million years). MP-EST

recovered the same species tree, including a sister-group relationship between *A. maurus* and *A. dickhilleni*, although the level of support for this clade varied depending on the sources of gene trees used in bootstrapping analyses. When ML trees were used, this clade was recovered only 41% of the time, but this value increased to 52% with MRBAYES trees and up to 84% with trees from the *BEAST run used as input.

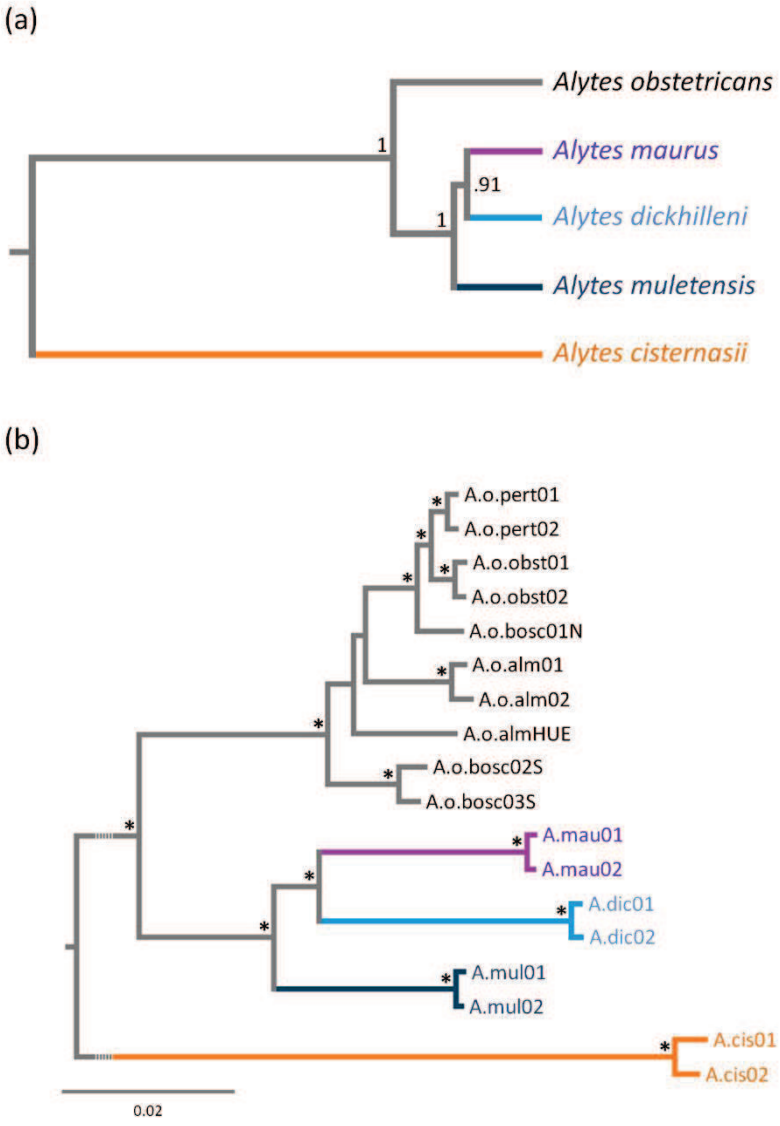


Fig. 3. (a) Species tree of *Alytes* based on *BEAST results of the analysis of mtDNA and four nuclear genes (*β-fibint7*; *C-myc*; *PPP3CAint4* and *RPL9int4*). Values at nodes indicate Bayesian posterior probabilities. (b) Bayesian phylogram based on analysis of concatenated mtDNA and nDNA data. * = BPP: ≥ 0.95 and BS: ≥ 85.

2.1.5 Discussion

Recently Biton *et al.* (2013) presented a comprehensive molecular phylogeny for Alytidae based on six genes (2500 bp), which recovered a well-supported *A. maurus* + *A. dickhilleni* clade. Our mtDNA data confirm this result in an enlarged dataset including a good representation of intraspecific variation in *A. obstetricans*, and it appears to be robust to different methods of analysis. On the other hand, our four nuclear markers, despite presenting significant variation, provided less resolution, although all species, including the three species in *Baleaphryne*, were recovered as monophyletic groups in most nuclear-gene trees (Figs. S1 and S2), with the only exception of *β-fibint7*, as previously reported by Gonçalves *et al.* (2007). Lack of reciprocal monophyly in this marker can be explained, based on JML results, which suggest absence of hybridization, by a deep coalescence event.

Increased data sampling, especially in mtDNA, has been shown to increase phylogenetic accuracy, even in situations of rapid, simultaneous divergence that make inference of population history extremely challenging (see Pabijan *et al.* 2013, Williams *et al.* 2013). In our case, it allowed full resolution of the *Baleaphryne* radiation with very high support. This contrasts with the lower support derived from nuclear data, which may stem from several factors, apart from the higher number of base pairs sequenced for mtDNA. It has been long known that the higher mutation rate and lower effective population size of mitochondrial DNA permit more precise phylogenetic resolution compared to nuclear markers (Avise 2000, Zhang *et al.* 2008, Belfiore *et al.* 2008). Conflicting gene trees are more common among nuclear markers due to incomplete lineage sorting, as in *Alytes*. Even with informative markers, as in our study, many additional independent nuclear markers may be required to resolve short internodes, such as those in the *Baleaphryne* radiation.

Despite the lack of resolution in nuclear DNA markers (both individually and concatenated), implementation of the multispecies coalescent in *BEAST produced a species tree with good support for the maurus + dickhilleni clade. In general, methods that estimate a species tree from independently estimated gene trees (*BEAST), including those that account for the stochastic nature of genetic drift in the lineage-sorting process, may be more likely to reconstruct the true species phylogeny, even where there is strong discordance among gene trees (Edwards *et al.* 2007, Williams *et al.* 2013). Leaché and Rannala (2011) suggested that Bayesian implementations of the multispecies coalescent could produce very accurate estimates of the species tree. Whether support for a maurus-dickhilleni clade in our species tree results from dominance of the mitochondrial marker in *BEAST analyses (which simultaneously estimates gene trees and the species tree) is unclear, although simulation studies suggest that this should not be the case (Heled and Drummond, 2010). In fact, results

of MP-EST, which does not infer gene trees and thus is expected to weight them equally when inferring the species tree, also reveal a sistergroup relationship between *A. maurus* and *A. dickhilleni*, although with weaker support. Support for this clade increased significantly when using as input gene trees from a *BEAST analysis, highlighting the potential for biases from the faster-evolving marker in the *BEAST approach during co-estimation of gene and species trees. In any event, one of the nuclear markers (*C-myc*) also provides robust support for that clade (Figs. S1c and S2c), and the other markers do not provide strong evidence for conflicting relationships. Therefore, in the absence of evidence for introgression or any other potentially confounding factor in either mtDNA or *C-myc*, we consider our results well supported. This is in agreement with previous analyses of rapid radiations using mitogenomes, in which mtDNA recovered the same pattern of phylogenetic relationships than that estimated by multilocus coalescence analyses (see Steinfartz *et al.* 2007, Zhang *et al.* 2008).

The new phylogenetic hypothesis is consistent with paleogeological evidence regarding the fragmentation of the Betic-Rifean Massif (see for instance, Martín *et al.* 2009), with the species endemic to the Betic (*A. dickhilleni*) and Rifean (*A. maurus*) mountains sharing a more recent common ancestor with respect to the other species, endemic to the Balearic islands (*A. muletensis*). The opening of the Strait of Gibraltar might be associated with the split of *A. maurus* and *A. dickhilleni*. Our time estimates have wide confidence intervals that encompass the extended timeframe of this geological event, although median values tend to be somewhat recent and probably underestimate the actual divergence time. In any case, the overseas dispersal event postulated by Martínez-Solano *et al.* (2004) to explain the differentiation of *A. muletensis* is no longer required.

Phylogenetic relationships between subspecies of *A. obstetricans* were not resolved in our study. Subspecies pertinax and obstetricans are recovered as monophyletic groups and form a clade in most gene trees. On the contrary, subspecies boscai and almogavarii are not recovered as monophyletic in any of the gene trees. Both taxa include very divergent lineages whose affinities are uncertain (see also Maia-Carvalho *et al.* 2014). Delineating major clades within *A. obstetricans* and resolving their evolutionary history requires extensive sampling and is thus beyond the scope of this paper; however, the nuclear markers used here are informative at the intraspecific level. Phylogenetic structure of lineages within *A. obstetricans* is nonetheless expected to be complicated by demographic fluctuations, secondary contact and admixture across lineages during the Quaternary.

2.1.6 Acknowledgments

We thank J. Teixeira for providing the picture of *A. obstetricans*. The editor and two anonymous reviewers provided comments that improved the manuscript. This work was supported through Project ‘‘Genomics and Evolutionary Biology’’ co-financed by North Portugal Regional Operational Programme 2007/2013 (ON.2 – O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF); the Program Operacional Factores de Competitividade (COMPETE), and by national funds from Fundação para a Ciência e a Tecnologia (FCT), through the research Project PTDC/BIA-BEC/099915/2008 to HG; and partial funds provided by grants CGL2008-04271-C02-01/ BOS and CGL2011-28300 (Ministerio de Ciencia e Innovación, Ministerio de Economía y Competitividad, Spain, and FEDER) and PPII10-0097-4200 (Junta de Comunidades de Castilla la Mancha and FEDER) to IMS, who is currently funded by Project ‘‘Biodiversity, Ecology and Global Change’’, co-financed by North Portugal Regional Operational Programme 2007/2013 (ON.2 – O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF). BMC is supported by a PhD grant from FCT (SFRH/BD/60305/2009), and HG is supported by a postdoctoral grant from FCT (SFRH/BPD/ 26555/2006).

2.1.7 References

Alonso R, Crawford AJ, Bermingham E (2012) Molecular phylogeny of an endemic radiation of Cuban toads (Bufonidae: *Peltophryne*) based on mitochondrial and nuclear genes. *Journal of Biogeography*, 39: 434–451.

Arévalo E, Davis SK, Sites 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.

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

Bastir M, Böhme M, Sanchiz B (2014) Middle Miocene remains of *Alytes* (Anura, Alytidae) as an example of the unrecognized value of fossil fragments for evolutionary morphology studies. *Journal of Vertebrate Paleontology*, 34: 69–79.

Belfiore N, Liu L, Moritz C (2008) Multilocus phylogenetics of a rapid radiation in the genus *Thomomys* (Rodentia: Geomyidae). *Systematic Biology*, 57: 294–310.

Beukema W, de Pous P, Donaire-Barroso D, Bogaerts S, Garcia-Porta J, Escoriza D, Arribas OJ, Mouden EHE, Carranza S (2013) Review of the systematics, distribution, biogeography and natural history of Moroccan amphibians. *Zootaxa*, 3661: 1–60.

Biton R, Geffen E, Vences M, Cohen O, Bailon S, Rabinovich R, Malka Y, Oron T, Boistel R, Brumfeld V, Gafny S (2013) The rediscovered Hula painted frog is a living fossil. *Nature Communications*, 4: 1959.

Blackburn DC, Siler CD, Diesmos AC, McGuire JA, Cannatella DC, Brown RM (2013) An adaptive radiation of frogs in a Southeast Asian island archipelago. *Evolution*, 67: 2631–2646.

Bons J, Geniez P (1996) Amphibiens et Reptiles du Maroc (Sahara Occidental compris), Atlas biogéographique. Asociación Herpetológica Española, Barcelona, Spain.

Brandley MC, Schmitz A, Reeder TW (2005) Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. *Systematic Biology*, 54: 373–390.

Brunes TO, Sequeira F, Haddad CFB, Alexandrino J (2010) Gene and species trees of a Neotropical group of treefrogs: genetic diversification in the Brazilian Atlantic Forest and the origin of a polyploid species. *Molecular Phylogenetic and Evolution*, 57: 1120–1133.

Camargo A, Avila LJ, Morando M, Sites JW (2012) Accuracy and precision of species trees: effects of locus, individual, and base pair sampling on inference of species trees in lizards of the *Liolaemus darwinii* group (Squamata, Liolaemidae). *Systematic Biology*, 61: 272–288.

Campbell V, Lapointe FJ (2009) The use and validity of composite taxa in phylogenetic analysis. *Systematic Biology*, 58: 560–572.

Chung Y, Ané C (2011) Comparing two Bayesian methods for gene tree/species tree reconstruction: simulations with incomplete lineage sorting and horizontal gene transfer. *Systematic Biology*, 60: 261–275.

Corl A, Ellegren H (2013) Sampling strategies for species trees: the effects on phylogenetic inference of the number of genes, number of individuals, and whether loci are mitochondrial, sex-linked, or autosomal. *Molecular Phylogenetic and Evolution*, 67: 358–366.

Crawford A (2003) Huge populations and old species of Costa Rican and Panamanian dirt frogs inferred from mitochondrial and nuclear gene sequences. *Molecular Ecology*, 12: 2525–2540.

Crespo EG (1997) *Alytes cisternasii* Boscá, 1879 In: Pleguezuelos, J.M. (Ed.), Distribución y Biogeografía de los Anfibios y Reptiles en España y Portugal. Universidad de Granada - Asociación Herpetológica Española, Granada, España.

Darriba D, Taboada GL, Doallo R, Posada D (2012) JModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9: 772.

de Pous P, Metallinou M, Donaire-Barroso D, Carranza S, Sanuy D (2013) Integrating mtDNA analyses and ecological niche modelling to infer the evolutionary history of *Alytes maurus* (Amphibia; Alytidae) from Morocco. *Herpetological Journal*, 23: 153–160.

Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24: 332–340.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7: 214.

Edwards SV (2009) Is a new and general theory of molecular systematic emerging? *Evolution*, 63: 1–19.

Edwards SV, Liu L, Pearl DK (2007) High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences USA*, 104: 5936–5941.

Themudo EG, Wielstra B, Arntzen JW (2009) Multiple nuclear and mitochondrial genes resolve the branching order of a rapid radiation of crested newts (*Triturus*, Salamandridae). *Molecular Phylogenetic and Evolution*, 52: 321–328.

Flot JF (2010) SeqPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. *Molecular Ecology Resources*, 10: 162–166.

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

García-París M, Arntzen JW (2002) *Alytes dickhilleni* Arntzen and García-París, 1995. In: Pleguezuelos, J.M., Márquez, R., Lizana, M. (Eds.), Atlas y Libro Rojo de los Anfibios y Reptiles de España. Dirección General de la Conservación de la Naturaleza, Asociación Herpetológica Española, Madrid, España.

Garcia-Porta J, Litvinchuk SN, Crochet PA, Romano A, Geniez PH, Lo-Valvo M, Lymberakis P, Carranza S (2012) Molecular phylogenetics and historical biogeography of the west-palearctic common toads (*Bufo bufo* species complex). *Molecular Phylogenetic and Evolution*, 63: 113–130.

Gonçalves H, Martínez-Solano I, Ferrand N, García-París M (2007) Conflicting phylogenetic signal of nuclear vs mitochondrial DNA markers in midwife toads (Anura, Discoglossidae, *Alytes*): deep coalescence or ancestral hybridization? *Molecular Phylogenetic and Evolution*, 44: 494–500.

Gonçalves H, Martínez-Solano I, Pereira RJ, Carvalho B, García-París M, Ferrand N (2009) High levels of population subdivision in a morphologically conserved Mediterranean toad (*Alytes cisternasii*) result from recent, multiple refugia: evidence from mtDNA, microsatellites and nuclear genealogies. *Molecular Ecology*, 18: 5143–5160.

Grossenbacher K (1997) *Alytes obstetricans* (Laurenti, 1768). In: Gasc *et al.* (Eds.), Atlas of Amphibians and Reptiles in Europe. Societas Europaea Herpetologica, Paris, France. pp. 94–95.

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.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, 27: 570–580.

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.

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

Joly S (2012) JML: testing hybridization from species trees. *Molecular Ecology Resources*, 12: 179–184.

Joly S, McLenachan PA, Lockhart PJ (2009) A statistical approach for distinguishing hybridization and incomplete lineage sorting. *The American Naturalist*, 175: 621–622.

Kapralov MV, Votintseva AA, Filatov DA (2013) Molecular adaptation during a rapid adaptive radiation. *Molecular Biology and Evolution*, 30: 1051–1059.

Kubatko LS, Gibbs HL, Bloomquist EW (2011) Inferring species-level phylogenies and taxonomic distinctiveness using multilocus data in *Sistrurus rattlesnakes*. *Systematic Biology*, 60: 393–409.

Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29: 1695–1701.

Leaché AD, Rannala B (2011) The accuracy of species tree estimation under simulation: a comparison of methods. *Systematic Biology*, 60: 126–137.

Leaché AD, Harris RB, Rannala B, Yang Z (2014) The influence of gene flow on species tree estimation: a simulation study. *Systematic Biology*, 63: 17–30.

Librado P, Rozas J (2009) DNAsp v5: a software for comprehensive analysis of polymorphism DNA data. *Bioinformatics*, 25: 1451–1452.

Liu L, Yu L, Edwards SV (2010) A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. *BMC Evolutionary Biology*, 10: 302.

Maia-Carvalho B, Gonçalves H, Martínez-Solano I, Gutiérrez-Rodriguez J, Lopes S, Ferrand N, Sequeira F (2014) Intraspecific genetic variation in the common midwife toad (*Alytes obstetricans*): subspecies assignment using mitochondrial and microsatellite markers. *Journal of Zoological Systematics and Evolutionary Research*, 52: 170–175.

Martín JM, Braga J, Aguirre J, Puga-Bernabéu A (2009) History and evolution of the North-Betic Strait (Prebetic Zone, Betic Cordillera): a narrow, early Tortonian, tidal-dominated, Atlantic-Mediterranean marine passage. *Sedimentary Geology*, 216: 80–90.

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

Mateo JA, Geniez P, Pether J (2013) Diversity and conservation of Algerian amphibian assemblages. *Basic and Applied Herpetology*, 27: 51–83.

Nakhleh L (2013) Computational approaches to species phylogeny inference and gene tree reconciliation. *Trends in Ecology & Evolution*, 28: 719–728.

Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press, New York, USA.

O'Meara BC (2012) Evolutionary inferences from phylogenies: a review of methods. *Annual Review of Ecology, Evolution, and Systematics*, 43: 267–285.

Pabijan M, Wandycz A, Hofman S, Wecek K, Piwczynski M, Szymura JM (2013) Complete mitochondrial genomes resolve phylogenetic relationships within *Bombina* (Anura Bombinatoridae). *Molecular Phylogenetic and Evolution*, 69: 63–74.

Pacheco MA, Battistuzzi FU, Lentino M, Aguilar R, Kumar S, Escalante AA (2011) Evolution of modern birds revealed by mitogenomics: timing the radiation and origin of major orders. *Molecular Biology and Evolution*, 28: 1927–1942.

Pinho C, Rocha S, Carvalho BM, Lopes S, Mourao S, Vallinoto M, Brunes TO, Haddad CFB, Gonçalves H, Sequeira F (2010) New primers for the amplification and sequencing of nuclear loci in a taxonomically wide set of reptiles and amphibians. *Conservation Genetics Resources*, 2: 181–185.

Rambaut A, Drummond AJ (2007) Tracer v1.4, <http://beast.bio.ed.ac.uk/Tracer>. (accessed 28.02.14).

Recuero E, Canestrelli D, Vörös J, Szabó K, Poyarkov NA, Arntzen JW, Crnobrnja-Isailovic J, Kidov AA, Coga Iniceanu D, Caputo FP, Nascetti G, Martínez-Solano I (2012) Multilocus species tree analyses resolve the radiation of the widespread *Bufo bufo* species group (Anura, Bufonidae). *Molecular Phylogenetic and Evolution*, 62: 71–86.

Reid NM, Hird SM, Brown JM, Pelletier TA, McVay JD, Satler JD, Carstens BC (2014) Poor fit to the multispecies coalescent is widely detectable in empirical data. *Systematic Biology*, 63: 322–333.

Román A (2002) *Alytes muletensis* (Sanchiz y Adrover, 1979). In: Pleguezuelos, J.M., Márquez, R., Lizana, M. (Eds.), Atlas y Libro Rojo de los Anfibios y Reptiles de España. Dirección General de la Conservación de la Naturaleza, Asociación Herpetológica Española, Madrid, pp. 79–81.

Ronquist F, Teslenko M, Van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61: 539–542.

San Mauro D (2010) A multilocus timescale for the origin of extant amphibians. *Molecular Phylogenetic and Evolution*, 56: 554–561.

San Mauro D, García-París M, Zardoya R (2004) Phylogenetic relationships of discoglossid frogs (Amphibia: Anura: Discoglossidae) based on complete mitochondrial genomes and nuclear genes. *Gene*, 343: 357–366.

Sanchiz B, Adrover R (1979) Anfibios fósiles del Pleistoceno de Mallorca. *Doñana, Acta Vertebrata*, 4: 5–25.

Sanders KL, Lee MSY, Mumpuni, Bertozzi T, Rasmussen AR (2013) Multilocus phylogeny and recent rapid radiation of the viviparous sea snakes (Elapidae: Hydrophiinae). *Molecular Phylogenetic and Evolution*, 66: 575–591.

Schwarz G (1978) Estimating the dimension of a model. Annals of Statistics 6, 461–464.

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

Sousa-Neves T, Aleixo A, Sequeira F (2013) Cryptic patterns of diversification of a widespread Amazonian Woodcreeper species complex (Aves: Dendrocolaptidae) inferred from multilocus phylogenetic analysis: implications for historical biogeography and taxonomy. *Molecular Phylogenetic and Evolution*, 68: 410–424.

Stanley EL, Bauer AM, Jackman TR, Branch WR, Mouton PLN (2011) Between a rock and a hard polytomy: rapid radiation in the rupicolous girdled lizards (Squamata: Cordylidae). *Molecular Phylogenetic and Evolution*, 58: 53–70.

Steinfartz S, Vicario S, Arntzen JW, Caccone A (2007) A Bayesian approach on molecules and behavior: reconsidering phylogenetic and evolutionary patterns of the Salamandridae with emphasis on *Triturus newts*. *Journal of experimental zoology. Part B, Molecular and developmental evolution*, 308B: 139–162.

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, 73: 1162–1169.

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

Swofford DL (2003) PAUP. Phylogenetic Analysis Using Parsimony (and Other Methods), v4. Sunderland: Sinauer Associates.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28: 2731–2739.

Vences M, Wake D (2007) Speciation, species boundaries and phylogeography of amphibians. In: Heatwhole, H., Tyler, M. (Eds.), Amphibian Biology, Systematics, vol. 7. *Surrey Beatty & Sons, Australia*, pp. 2613–2671.

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

Whitfield JB, Lockhart PJ (2007) Deciphering ancient rapid radiations. *Trends in Ecology & Evolution*, 22: 258–265.

Wiens JJ, Kuczynski CA, Arif S, Reeder TW (2010) Phylogenetic relationships of phrynosomatid lizards based on nuclear and mitochondrial data, and a revised phylogeny for Sceloporus. *Molecular Phylogenetic and Evolution*, 54: 150–161.

Williams JS, Niedzwiecki JH, Weisrock DW (2013) Species tree reconstruction of a poorly resolved clade of salamanders (Ambystomatidae) using multiple nuclear loci. *Molecular Phylogenetic and Evolution*, 68: 671–682.

Xia X (2013) DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Molecular Biology and Evolution*, 30: 1720–1728.

Yang Z, Rannala B (2012) Molecular phylogenetics: principles and practice. *Nature Reviews Genetics*, 13: 303–314.

Zhang P, Papenfuss TJ, Wake MH, Qu L, Wake DB (2008) Phylogeny and biogeography of the family Salamandridae (Amphibia: Caudata) inferred from complete mitochondrial genomes. *Molecular Phylogenetic and Evolution*, 49: 586–597.

Zhang P, Liang D, Mao RL, Hillis DM, Wake DB, Cannatella DC (2013) Efficient sequencing of anuran mtDNAs and a mitogenomic exploration of the phylogeny and evolution of frogs. *Molecular Biology and Evolution*, 30: 1899–1915.

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. USA.

2.1.8 Supplementary material

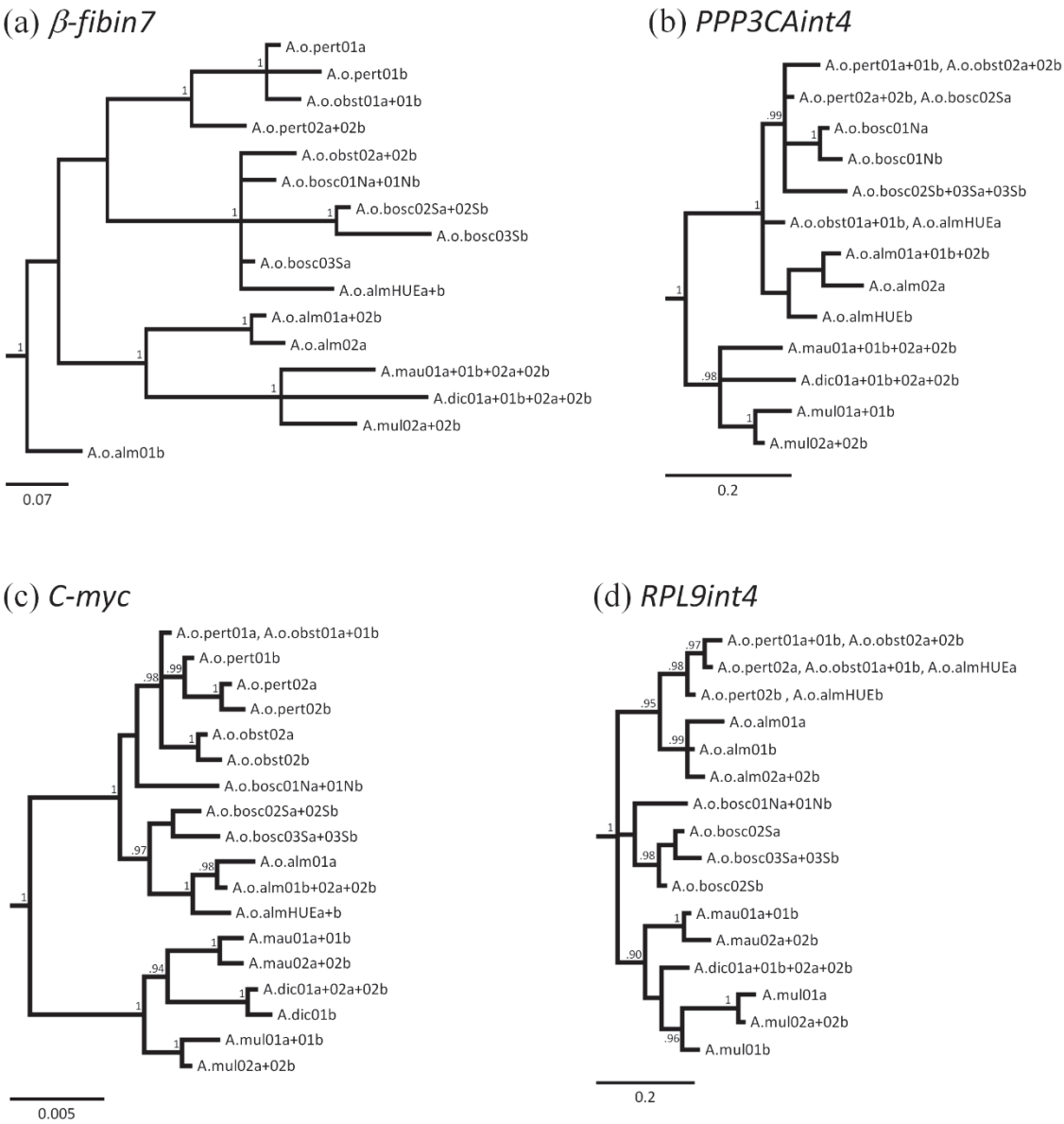


Fig. S1. Bayesian nuclear gene trees of 16 samples representing all species and subspecies of *Alytes* (a: *β-fibint7*; b: *PPP3CAint4*; c: *C-myc*; and d: *RPL9int4*). The outgroup *A. cisternasii* was omitted for clarity. Numbers at nodes represent Bayesian posterior probabilities (only values ≥ 0.9 are shown).

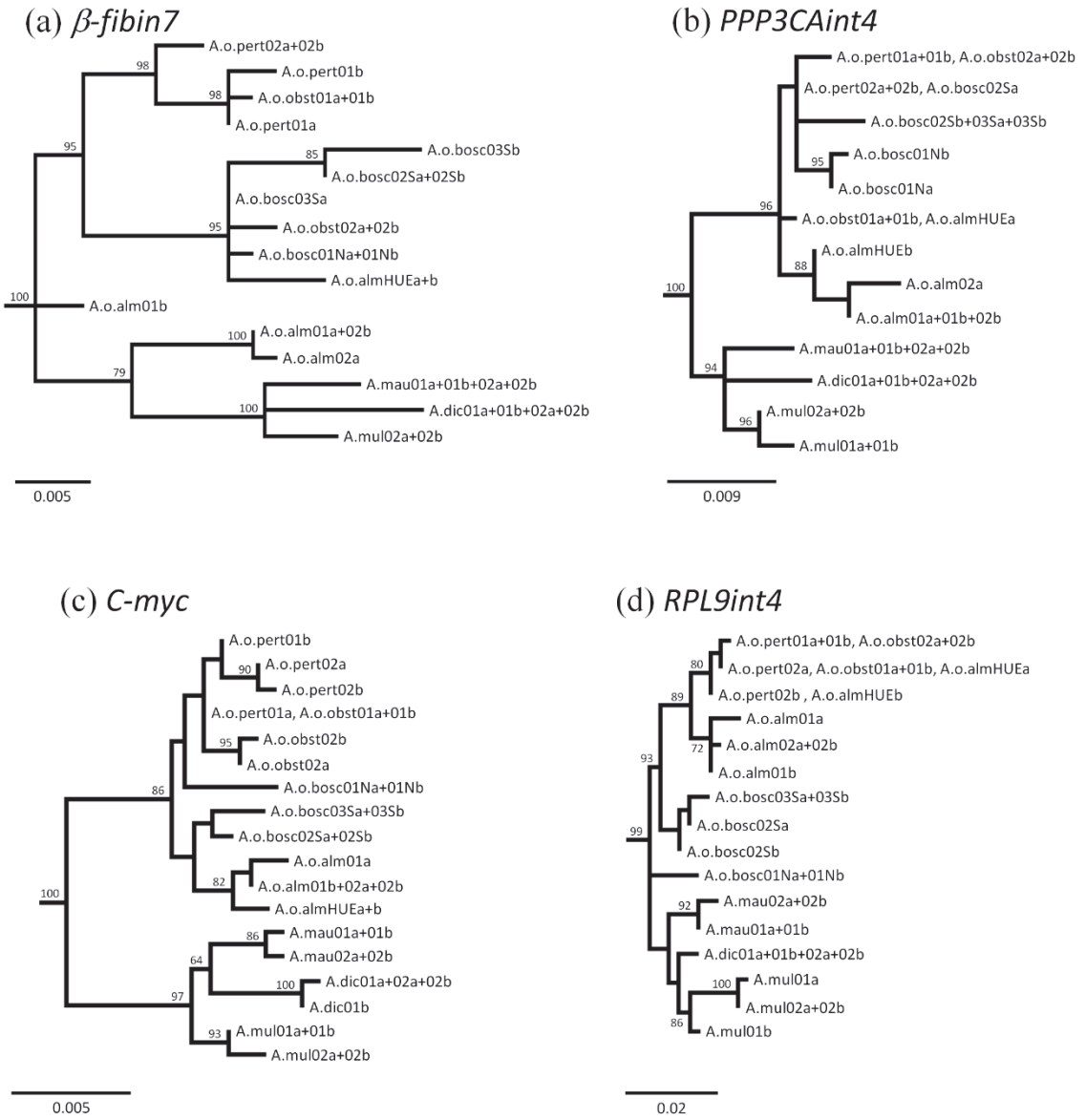


Fig. S2. Maximum likelihood gene trees of 16 samples representing all species and subspecies of *Alytes* based on Garli analyses (a: *β-fibint7*; b: *PPP3CAint4*; c: *C-myc*; and d: *RPL9int4*). The outgroup *A. cisternasii* was omitted for clarity. RAxML results were very similar in topology and branch support (not shown). The phylogram represents the best tree obtained by Garli for each nuclear marker. Numbers at nodes represent bootstrap values (only values ≥ 75 are shown).

Table S1. GenBank accession numbers and sequence information.

Taxa	Taxon code	mtDNA fragments								
		FA	*RNA1*	*RNA2*	*ND1*	*ND2*		*COX1*		*COX2-APT6*	
		FS	mt01(F+R)	mt02(F+R)	mt03(F+R)	mt04(F)	mt05(R)	mt06(F)	mt07(R)	mt08(F)	mt09(R)
		NP	482-1387	1952-2156	2688-3780	3878-4477	4855-5115	5181-5913	6516-6898	6920-7607	8202-8703
A. o. pertinax	A.o.pert01		+	+	+	+	+	+	+	+	+
	(MAD04)		KJ858769	KJ858787	KJ858803	KJ858821	KJ858839	KJ858857	KJ858875	KJ858893	KJ858911
	A.o.pert02		+	+	+	+	+	+	+	+	+
	(CUE01)		KJ858770	KJ858788	KJ858804	KJ858822	KJ858840	KJ858858	KJ858876	KJ858894	KJ858912
A. o. obstetricans	A.o.obst01		+	+	+	+	+	+	+	+	+
	(FRAN14)		KJ858771	KJ858789	KJ858805	KJ858823	KJ858841	KJ858859	KJ858877	KJ858895	KJ858913
	A.o.obst02		+	+	+	+	+	+	+	+	+
	(FRAN01)		KJ858772	KJ858790	KJ858806	KJ858824	KJ858842	KJ858860	KJ858878	KJ858896	KJ858914
A. o. boscai	A.o.bosc01N		+	+	+	+	+	+	+	+	+
	(OUR03)		KJ858773	KJ858791	KJ858807	KJ858825	KJ858843	KJ858861	KJ858879	KJ858897	KJ858915
	A.o.bosc02S		+	+	+	+	+	+	+	+	+
	(SMA25)		KJ858774	KJ858792	KJ858808	KJ858826	KJ858844	KJ858862	KJ858880	KJ858898	KJ858916
	A.o.bosc03S		+	+	+	+	+	+	+	+	+
	(MTM05)		KJ858775	KJ858793	KJ858809	KJ858827	KJ858845	KJ858863	KJ858881	KJ858899	KJ858917
A. o. almogavarii	A.o.alm01		+	+	+	+	+	+	+	+	+
	(BER01)		KJ858776	KJ858794	KJ858810	KJ858828	KJ858846	KJ858864	KJ858882	KJ858900	KJ858918
	A.o.alm02		+	+	+	+	+	+	+	+	+
	(BER03)		KJ858777	KJ858795	KJ858811	KJ858829	KJ858847	KJ858865	KJ858883	KJ858901	KJ858919
	A.o.almHUE		+	+	+	+	+	+	+	+	+
	(HUE01)		KJ858778	KJ858796	KJ858812	KJ858830	KJ858848	KJ858866	KJ858884	KJ858902	KJ858920
Alytes maurus	A.mau01		+	+	+	+	+	+	+	+	+
	(MAR04)		KJ858779	KJ858797	KJ858813	KJ858831	KJ858849	KJ858867	KJ858885	KJ858903	KJ858921
	A.mau02		+	+	+	+	+	+	+	+	+
	(MAR05)		KJ858780	KJ858798	KJ858814	KJ858832	KJ858850	KJ858868	KJ858886	KJ858904	KJ858922
Alytes dickhilleni	A.dic01		+	+	+	+	+	+	+	+	+
	(HG104)		KJ858781	KJ858799	KJ858815	KJ858833	KJ858851	KJ858869	KJ858887	KJ858905	KJ858923
	A.dic_02		+	+	+	+	+	+	+	+	+
	(Aly120D)		KJ858782	KJ858800	KJ858816	KJ858834	KJ858852	KJ858870	KJ858888	KJ858906	KJ858924
Alytes muletensis	A.mul01		+	+	+	+	+	+	+	+	+
	(CAP01)		KJ858783	KJ858801	KJ858817	KJ858835	KJ858853	KJ858871	KJ858889	KJ858907	KJ858925
	A.mul02		+	MAI06	+	+	+	+	MAI06	+	MAI06
	(AMD2)		KJ858784	KJ858802	KJ858818	KJ858836	KJ858854	KJ858872	KJ858890	KJ858908	KJ858926
Alytes cisternasii	A.cis01		+	IMS2004 (a)	+	+	+	+	+	+	+
	(MAD01)		KJ858785	AY442027	KJ858819	KJ858837	KJ858855	KJ858873	KJ858891	KJ858909	KJ858927
	A.cis02		+		+	+	+	+	+	+	+
	(IDN03)		KJ858786	***	KJ858820	KJ858838	KJ858856	KJ858874	KJ858892	KJ858910	KJ858928

Taxon code as in Table 1. Voucher numbers (in parentheses, see Table 1) indicate the main samples representing each lineage. +: indicates that the sequence was obtained from the main voucher; other codes indicate additional samples used to complement the dataset. FA: fragment amplified; FS: Fragment sequenced. Capital letters in parentheses indicate sequencing direction (F: Forward, R: Reverse). NP: nucleotide position based on the mtDNA genome (reviewed reference sequence, San Mauro *et al.* 2004) of *A. o. pertinax* (GI: NC_006688.1). Lowercase letters in parentheses refer to sequences obtained from GenBank: (a) Martínez-Solano *et al.* (2004); (b) Gonçalves *et al.* (2007); (c) Gonçalves *et al.* (2009); (d) Pinho *et al.* (2010); ***: no data.

					nDNA - fragments			
COX3-ND4L		*ND4*	*ND5*	*ND6*	*β-fibint7*	*PPP3CAint4*	*RPL9int4*	*C-myc*
mt10(F)	mt11(R)	mt12(F+R)	mt13(F+R)	mt14(F+R)				
8731-9313	9624-10170	10880-11693	11788-12861	13468-14119				
+	+	(b)	+	+	(b)	+	+	VLC22
KJ858929	KJ858947	EF441298	KJ858970	KJ858988	EF441322/EF441323	KJ859033	KJ859047	KJ859009/KJ859010
+	+	(b)	+	+	(b)	(d)	(d)	+
KJ858930	KJ858948	EF441299	KJ858971	KJ858989	EF441324	GU181146	GU181179	KJ859011/KJ859012
+	+	+	+	+	SAN01	+	+	SAN01
KJ858931	KJ858949	KJ858965	KJ858972	KJ858990	KJ859005	KJ859034	KJ859048	KJ859013
+	+	(b)	+	+	(b)	(d)	(d)	+
KJ858932	KJ858950	EF441292	KJ858973	KJ858991	EF441316	GU181145	GU181178	KJ859014/KJ859015
+	+	(b)	+	+	(b)	(d)	(d)	+
KJ858933	KJ858951	EF441295	KJ858974	KJ858992	EF441319	GU181147/KJ859035	GU181180	KJ859016
+	+	(b)	+	+	(b)	+	+	+
KJ858934	KJ858952	EF441303	KJ858975	KJ858993	EF441330	KJ859036/KJ859037	KJ859049/KJ859050	KJ859017
+	+	+	MTM02	+	MTM02	+	+	+
KJ858935	KJ858953	KJ858966	KJ858976	KJ858994	KJ859006/KJ859007	KJ859038	KJ859051	KJ859018
+	+	(b)	+	+	(b)	(d)	(d)	+
KJ858936	KJ858954	EF441305	KJ858977	KJ858995	EF441332/EF441333	GU181144	KJ859052/GU181177	KJ859019/KJ859020
+	+	(b)	+	+	(b)	+	+	+
KJ858937	KJ858955	EF441306	KJ858978	KJ858996	EF441334/EF441335	KJ859039/KJ859040	KJ859053	KJ859021
+	+	(b)	+	+	(b)	+	+	+
KJ858938	KJ858956	EF441304	KJ858979	KJ858997	EF441331	KJ859041/KJ859042	KJ859054/KJ859055	KJ859022
+	+	(b)	+	+	(b)	+	+	MAR06
KJ858939	KJ858957	EF441312	KJ858980	KJ858998	EF441341	KJ859043	KJ859056	KJ859023
+	+	MAR03 (b)	+	+	MAR03 (b)	MAR03 (d)	+	+
KJ858940	KJ858958	EF441311	KJ858981	KJ858999	EF441340	GU181141	KJ859057	KJ859024
+	+	+	+	+	HG103 (b)	+	+	+
KJ858941	KJ858959	KJ858967	KJ858982	KJ859000	EF441338	KJ859044	KJ859058	KJ859025/KJ859026
+	+	+	+	+	IMS4189	+	+	+
KJ858942	KJ858960	KJ858968	KJ858983	KJ859001	KJ859008	KJ859045	KJ859059	KJ859027
+	+	+	+	+		+	+	+
KJ858943	KJ858961	KJ858969	KJ858984	KJ859002	***	KJ859046	KJ859060/KJ859061	KJ859028
+	+	MAI06 (b)	+	+	MAI06 (b)	MAI06 (d)	MAI06 (d)	MAI06
KJ858944	KJ858962	EF441310	KJ858985	KJ859003	EF441339	GU181142	GU181175	KJ859029
+	+	HG112 (b)	+	+	HG112 (b)	HG112 (d)	HG112 (d)	HG112
KJ858945	KJ858963	EF441313	KJ858986	KJ859004	EF441342	GU181140	GU181173	KJ859030
+	+	CER01 (b)	+		CER01 (b)	CER01 (c)	CER01	NAV22
KJ858946	KJ858964	EF441314	KJ858987	***	EF441342	GU086791	KJ859062	KJ859031/KJ859032

Table S2. Optimal partition schemes (eight for mtDNA and three for nDNA) and corresponding nucleotide substitution models selected by PARTITIONFINDER for MRBAYES and GARLI analyses, based on the Bayesian Information Criterion.The ordinal numbers 1st, 2nd, 3rd correspond to the codon position in the respective gene. FS: Fragment sequenced, FC: Functional category.

Marker	FS	FC	Input file	Partition	Best fit model for partition (PFinder)	Best fit model (jModeltest)
mtDNA						
	mt01	*RNA*	mt01	mtDNA1	HKY + I + G	
	mt02	*RNA*	mt02	mtDNA2	TrNef + I	
	mt03	*RNA*	mt03a	mtDNA3	JC	
		ND1	mt03b 1st	mtDNA1		
			mt03b 2nd	mtDNA4	F81	
			mt03b 3rd	mtDNA5	TrN +I + G	
		RNA	mt03c	mtDNA1		
	mt04	*RNA*	mt04a	mtDNA6	TrN +I	
		ND2	mt04b 1st	mtDNA1		
			mt04b 2nd	mtDNA6		
			mt04b 3rd	mtDNA5		
	mt05	*ND2*	mt05a 1st	mtDNA5		
			mt05a 2nd	mtDNA1		
			mt05a 3rd	mtDNA6		
		RNA	mt05b	mtDNA1		
	mt06	*RNA*	mt06a	mtDNA1		
		COX1	mt06b 1st	mtDNA2		
			mt06b 2nd	mtDNA4		
			mt06b 3rd	mtDNA5		
	mt07	*COX1*	mt07a 1st	mtDNA4		
			mt07a 2nd	mtDNA5		
			mt07a 3rd	mtDNA2		
		RNA	mt07b	mtDNA1		
	mt08	*RNA*	mt08a	mtDNA1		
		COX2	mt08b 1st	mtDNA2		
			mt08b 2nd	mtDNA4		
			mt08b 3rd	mtDNA5		
	mt09	*APT6*	mt09a 1st	mtDNA5		
			mt09a 2nd	mtDNA1		
			mt09a 3rd	mtDNA4		
		COX3	mt09b 1st	mtDNA3		
			mt09b 2nd	mtDNA4		
			mt09b 3rd	mtDNA1		
	mt10	*COX3*	mt10 1st	mtDNA6		
			mt10 2nd	mtDNA5		
			mt10 3rd	mtDNA2		

Table S2. Continued.

Marker	FS	*FC*	Input file	Partition	Best fit model for partition (PFinder)	Best fit model (jModeltest)
	mt11	*ND3*	mt11a 1st	mtDNA1		
			mt11a 2nd	mtDNA6		
			mt11a 3rd	mtDNA5		
		RNA	mt11b	mtDNA1		
		ND4L	mt11c 1st	mtDNA1		
			mt11c 2nd	mtDNA4		
			mt11c 3rd	mtDNA5		
	mt12	*ND4*	mt12a 1st	mtDNA5		
			mt12a 2nd	mtDNA1		
			mt12a 3rd	mtDNA6		
		RNA	mt12b	mtDNA1		
	mt13	*ND5*	mt13 1st	mtDNA5		
			mt13 2nd	mtDNA1		
			mt13 3rd	mtDNA6		
	mt14	*ND5*	mt14a 1st	mtDNA5		
			mt14a 2nd	mtDNA1		
			mt14a 3rd	mtDNA4		
		ND6	mt14b 1st	mtDNA7	TIM + G	
			mt14b 2nd	mtDNA6		
			mt14b 3rd	mtDNA8	TrN + G	
		RNA	mt14c	mtDNA1		
nDNA						
	β-fibint7			nDNA1	HKY+G	F81
	PPP3CAint4			nDNA1		
	RPL9int4			nDNA1		
	C-myc		Exon2 1st	nDNA2	JC	HKY
			Exon2 2nd	nDNA2		
			Exon2 3rd	nDNA2		
			Intron2	nDNA3	F81	F81+G
			Exon3 1st	nDNA2		
			Exon3 2nd	nDNA2		
			Exon3 3rd	nDNA2		

Table S3. Summary results of genetic diversity in mtDNA and nuclear gene sequences.

Marker	Functional Categories	Clade	Length	Polymorphism						
				N	*S*	*Eta*	*h*	*π*	*SD(π)*	*θ*
mtDNA										
	ND1	*Alytes*	960-963	18	254	285	16	0.0849	0.0133	0.0769
	(mt03)	*Alytes obstetricans*	963	9	57	57	8	0.0234	0.0032	0.0218
		Baleaphryne	963	6	122	127	5	0.0686	0.0109	0.0555
	ND2	*Alytes*	665	18	36	42	16	0.0270	0.0133	0.0843
	(mt04+ mt05)	*Alytes obstetricans*	665	9	50	50	8	0.0282	0.0045	0.0277
		Baleaphryne	665	6	91	93	5	0.0714	0.0113	0.0599
	COX1	*Alytes*	949	18	228	259	18	0.0805	0.0114	0.0699
	mt06+07	*Alytes obstetricans*	949	9	53	53	9	0.0210	0.0031	0.0206
		Baleaphryne	949	6	94	95	6	0.0516	0.0077	0.0539
	COX2	*Alytes*	590	18	142	164	17	0.0772	0.0120	0.0700
	(mt08)	*Alytes obstetricans*	590	9	39	39	9	0.0261	0.0034	0.0243
		Baleaphryne	590	6	65	66	5	0.0596	0.0090	0.0483
	ATP6	*Alytes*	420	18	114	130	16	0.0865	0.0135	0.0789
	(mt09)	*Alytes obstetricans*	420	9	22	22	8	0.0194	0.0030	0.0193
		Baleaphryne	420	6	51	53	5	0.0641	0.0096	0.0532
	COX3	*Alytes*	665	18	140	164	16	0.0665	0.0096	0.0612
	mt09+10	*Alytes obstetricans*	665	9	38	38	8	0.0221	0.0031	0.0210
		Baleaphryne	665	6	61	64	5	0.0491	0.0073	0.0402
	ND3	*Alytes*	195	18	43	52	14	0.1246	0.0179	0.1069
	(mt11)	*Alytes obstetricans*	195	9	18	18	8	0.0362	0.0057	0.0340
		Baleaphryne	195	6	27	27	4	0.1214	0.0188	0.1011
	ND4L	*Alytes*	282	18	76	87	15	0.0859	0.0128	0.0784
	(mt11)	*Alytes obstetricans*	282	9	17	17	8	0.0219	0.0044	0.0222
		Baleaphryne	282	6	40	42	4	0.0768	0.0115	0.0621
	ND4	*Alytes*	677	18	206	241	17	0.1036	0.0149	0.0958
	(mt12)	*Alytes obstetricans*	677	9	70	74	8	0.0395	0.0054	0.0381
		Baleaphryne	677	6	90	93	6	0.0754	0.0111	0.0630
	ND5	*Alytes*	1174	18	322	374	18	0.1022	0.0152	0.0872
	(mt13)	*Alytes obstetricans*	1174	9	64	64	9	0.0235	0.0038	0.0219
		Baleaphryne	1174	6	141	145	6	0.0692	0.0104	0.0575
	ND6	*Alytes*	510	17	157	187	16	0.0978	0.0142	0.0911
	(mt14)	*Alytes obstetricans*	510	9	39	39	8	0.0301	0.0045	0.0281
		Baleaphryne	510	6	83	87	6	0.0873	0.0129	0.0713
	Non coding	*Alytes*	1954	17	245	267	17	0.0310	0.0062	0.0373
		Alytes obstetricans	1954	9	61	63	9	0.0112	0.0015	0.0115
		Baleaphryne	1954	6	66	68	6	0.0178	0.0026	0.0148
	Coding	*Alytes*	7091	17	1898	2185	17	0.0816	0.0125	0.0806
		Alytes obstetricans	7091	9	478	482	9	0.0259	0.0036	0.0248
		Baleaphryne	7091	6	878	905	6	0.0669	0.0099	0.0552

Table S3. Continued.

Marker	Functional Categories	Clade	Length	Polymorphism						
				N	*S*	*Eta*	*h*	*π*	*SD(π)*	*θ*
nDNA										
	β-fibint7	*Alytes*	634	36	76	77	17	0.0297	0.0026	0.0299
		Alytes obstetricans	634	18	37	37	12	0.0174	0.0013	0.0170
		Baleaphryne	634	12	12	12	3	0.0093	0.0008	0.0063
	PPP3CAint4	*Alytes*	712	36	74	77	15	0.0252	0.0051	0.0262
		Alytes obstetricans	712	18	14	15	8	0.0059	0.0006	0.0059
		Baleaphryne	712	12	14	14	5	0.0088	0.0008	0.0067
	C-myc	*Alytes*	1317	36	85	86	20	0.0150	0.0024	0.0160
		Alytes obstetricans	1317	18	21	21	12	0.0045	0.0003	0.0047
		Baleaphryne	1317	12	15	15	6	0.0047	0.0004	0.0039
	RPL9int4	*Alytes*	479	36	76	81	18	0.0480	0.0116	0.0513
		Alytes obstetricans	479	18	19	19	10	0.0141	0.0017	0.0123
		Baleaphryne	479	12	15	15	6	0.0134	0.0015	0.0111

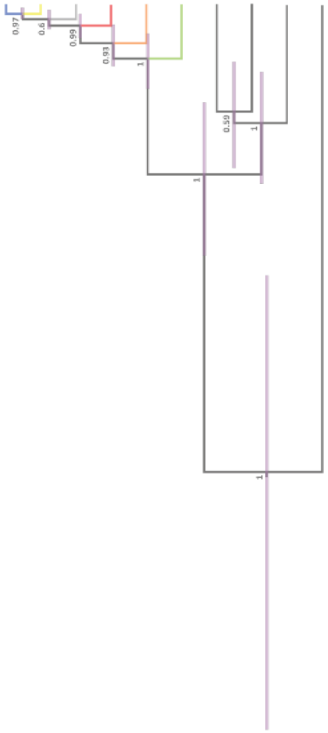
N: Number of samples, *S*: Number of polymorphic sites; *Eta*: Total number of mutations; *h*: Total number of haplotypes; *π*: Nucleotide diversity; *SD*: Standard deviation and *θ*: Theta. Text refers to the fragment sequenced – FS (see Table S1).

Capítulo 3

Padrões intraespecíficos de variabilidade genética e diversificação de linhagens em *Alytes obstetricans*

Artigo II

Gonçalves H, Maia-Carvalho B, Souza-Neves AC, García-París M, Sequeira F, Ferrand N, Martínez-Solano I (2015). Multilocus phylogeography of the common midwife toad, *Alytes obstetricans* (Anura, Alytidae): Contrasting patterns of lineage diversification and genetic structure in the Iberian refugium. *Molecular Phylogenetics and Evolution* 93, 363 – 379. doi: 10.1016/j.ympev.2015.08.009



Artigo II

Original Article
Accepted 7 August 2015

Multilocus phylogeography of the common midwife toad, *Alytes obstetricans* (Anura, Alytidae): Contrasting patterns of lineage diversification and genetic structure in the Iberian refugium

H. Gonçalves [a], B. Maia-Carvalho [a,b], T. Souza-Neves [a,c], M. García-París [d], F. Sequeira [a], N Ferrand [a,b,e], I. Martínez-Solano [a,f,g]

Molecular Phylogenetics and Evolution, 93: 363 – 379
doi: 10.1016/j.ympev.2015.08.009

[a] *CIBIO Research Centre in Biodiversity and Genetic Resources, InBIO, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal*
[b] *Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, 4099-002 Porto, Portugal*
[c] *Museu Paraense Emílio Goeldi, Caixa Postal 399, Belém, PA 66040-170, Brazil*
[d] *Museo Nacional de Ciencias Naturales, MNCN-CSIC, c/José Gutiérrez Abascal 2, 28006 Madrid, Spain*
[e] *Department of Zoology, University of Johannesburg, Johannesburg, South Africa*
[f] *Instituto de Investigación en Recursos Cinegéticos (CSIC-UCLM-JCCM), Ronda de Toledo, s/n, 13005 Ciudad Real, Spain*
[g] *Ecology, Evolution and Development Group, Department of Wetland Ecology, Estación Biológica de Doñana (EBD-CSIC), Avenida Américo Vespucio, s/n, 41092 Sevilla, Spain*

3.1 Multilocus phylogeography of the common midwife toad, *Alytes obstetricans* (Anura, Alytidae): Contrasting patterns of lineage diversification and genetic structure in the Iberian refugium

3.1.1 Abstract

Recent investigations on the evolutionary history of the common midwife toad (*Alytes obstetricans*) revealed high levels of geographically structured genetic diversity but also a situation where delineation of major historical lineages and resolution of their relationships are much more complex than previously thought. We studied sequence variation in one mitochondrial and four nuclear genes throughout the entire distribution range of all recognized *A. obstetricans* subspecies to infer the evolutionary processes that shaped current patterns of genetic diversity and population subdivision. We found six divergent, geographically structured mtDNA haplogroups diagnosing population lineages, and varying levels of admixture in nuclear markers. Given the timeframe inferred for the splits between major lineages, the climatic and environmental changes that occurred during the Pleistocene seem to have shaped the diversification history of *A. obstetricans*. Survival of populations in allopatric refugia through the Ice Ages supports the generality of the ‘‘refugia-within-refugia” scenario for the Iberian Peninsula. However, lineages corresponding to subspecies *A. o. almogavarii*, *A. o. pertinax*, *A. o. obstetricans*, and *A. o. boscai* responded differently to Pleistocene climatic oscillations after diverging from a common ancestor. *Alytes o. obstetricans* expanded northward from a northern Iberian refugium through the western Pyrenees, leaving a signal of contrasting patterns of genetic diversity, with a single mtDNA haplotype north of the Pyrenees from SW France to Germany. Both *A. o. pertinax* and *A. o. boscai* are widespread and genetically diverse in Iberia, the latter comprising two divergent lineages with a long independent history. Finally, *A. o. almogavarii* is mostly restricted to the north-eastern corner of Iberia north of the Ebro river, with additional populations in a small region in south-eastern France. This taxon exhibits unparalleled levels of genetic diversity and little haplotype sharing with other lineages, suggesting a process of incipient speciation.

Key words: Amphibia, Historical biogeography, Demography, Diversification, Speciation

3.1.2 Introduction

Patterns of genetic diversity within species are a consequence of their evolutionary history and of contemporary constraints to dispersal, and these processes are expected to give rise to specific phylogeographic patterns (Avise 2000, 2009). The detection and description of these patterns are the key to infer the evolutionary history of a species, and this knowledge has important implications for predicting biotic responses to current and future periods of global climate change, and ultimately, to provide valuable information for the conservation and management of species and populations (Moritz and Agudo 2013, Velo-Antón *et al.* 2013, Fordham *et al.* 2014).

In the last three decades, numerous phylogeographic studies have shown that Quaternary climatic oscillations caused severe range shifts in species distributions and have been among the most important historical factors in shaping the genetic structure of temperate species (Taberlet *et al.* 1998, Hewitt 2000, 2004). It is well known that Southern Europe, including the peninsulas of Iberia, Italy, and the Balkans, has functioned as a refugial area for temperate species survival during periods of adverse climatic conditions (Taberlet *et al.* 1998, Hewitt 2001, Stewart *et al.* 2010). Geographical separation and long-lasting persistence in these refugia caused divergence of local populations and the evolution of genetic lineages or even speciation (Hewitt 1996). High genetic diversity can often be detected in populations located in former refugia (Hewitt 2000, 2004), while populations at the leading edge of a species' distribution are mostly characterized by lower genetic diversity (Hampe and Petit 2005, Recuero and García-París 2011). Within the Iberian Peninsula, a geologically and ecologically complex and heterogeneous region, multiple population-divergence processes across different geographical and temporal scales have been described for a variety of taxa, showing the existence of ''refugia within refugia" (Gómez and Lunt 2007, Abellán and Svenning 2014). Characterizing these refugial areas is of the utmost importance for the long-term conservation of species and populations, since they frequently harbor most of the species' genetic diversity and thus their potential to cope with environmental changes (Hampe and Petit 2005, Schoville *et al.* 2012, Dufresnes *et al.* 2013).

The common midwife toad, *Alytes obstetricans* (Laurenti, 1768) is widely distributed in Western Europe, from Germany to the northern half of the Iberian Peninsula (Fig. 1; Grossenbacher 1997). However, most of the species' genetic diversity is found in Iberia (Maia-Carvalho *et al.* 2014a). Currently, four subspecies are recognized within *A. obstetricans* (Fig. 1): (i) *A. o. obstetricans* (Laurenti, 1768), distributed across western Europe and the northern Iberian Peninsula (Navarra, Basque Country and Cantabrian Mountains); (ii) *A. o. boscai* Lataste, 1879, in northern and central Portugal, Galicia, western Castilla-León as well as along

the Sistema Central mountains; (iii) *A. o. pertinax* García-París and Martínez-Solano, 2001, present in the central and eastern regions of the Iberian Peninsula (García-París and Martínez-Solano 2001); and (iv) *A. o. almogavarii* Arntzen and García-París, 1995, distributed from the eastern Pyrenees (Geniez and Crochet 2003) up to the Sierra de Guadarrama, north of Madrid (García-París 1995). According to Arntzen and García-París (1995) and Martínez-Solano *et al.* (2004), the differentiation between subspecies may be related to the formation of the main fluvial drainages in the Iberian Peninsula, about 4 Mya (millions of years ago) or, alternatively, be associated with isolation in different Pleistocene glacial refugia. While several studies have addressed patterns of variation in morphology, allozymes, and mitochondrial and nuclear DNA (Arntzen and García-París 1995, Fonseca *et al.* 2003, Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007), some aspects of the evolutionary history of this species remain unsolved. Recently, a study by Maia-Carvalho *et al.* (2014a), based on mitochondrial and microsatellite data in a set of populations representative of all currently recognized subspecies, revealed discordances with previous studies. First, the location of the northern boundary between *A. o. boscai* and *A. o. obstetricans* had been inferred to be in the Pyrenees, based on allozyme data (Arntzen and Szymura 1984), but was shown to be more likely located in the Iberian northwest, somewhere between northern Portugal and the Cantabrian Mountains. Second, Maia-Carvalho *et al.* (2014a) revealed the existence of two well-differentiated groups within *A. o. boscai*, separated by the Douro River, as suggested by Fonseca *et al.* (2003). Additionally, the recent multilocus assessment of phylogenetic relationships in *Alytes* by Maia-Carvalho *et al.* (2014b) produced inconclusive results about the relationships between major clades in *A. obstetricans* and indicated conflict with the current morphology-based subspecific taxonomy.

In this study, we use a range-wide multilocus dataset (one mitochondrial and four nuclear genes) to investigate further the evolutionary history of *A. obstetricans*, with special interest in patterns of population subdivision and postglacial expansion. Specifically, we assess the extent of mutual compatibility between nuclear and mitochondrial markers and the relationship with previously described morphological subspecies, and the dynamics of population fragmentation, differentiation and post-glacial expansion. We discuss the evolutionary processes that may have acted as drivers of diversification and the taxonomic and conservation implications of our results.

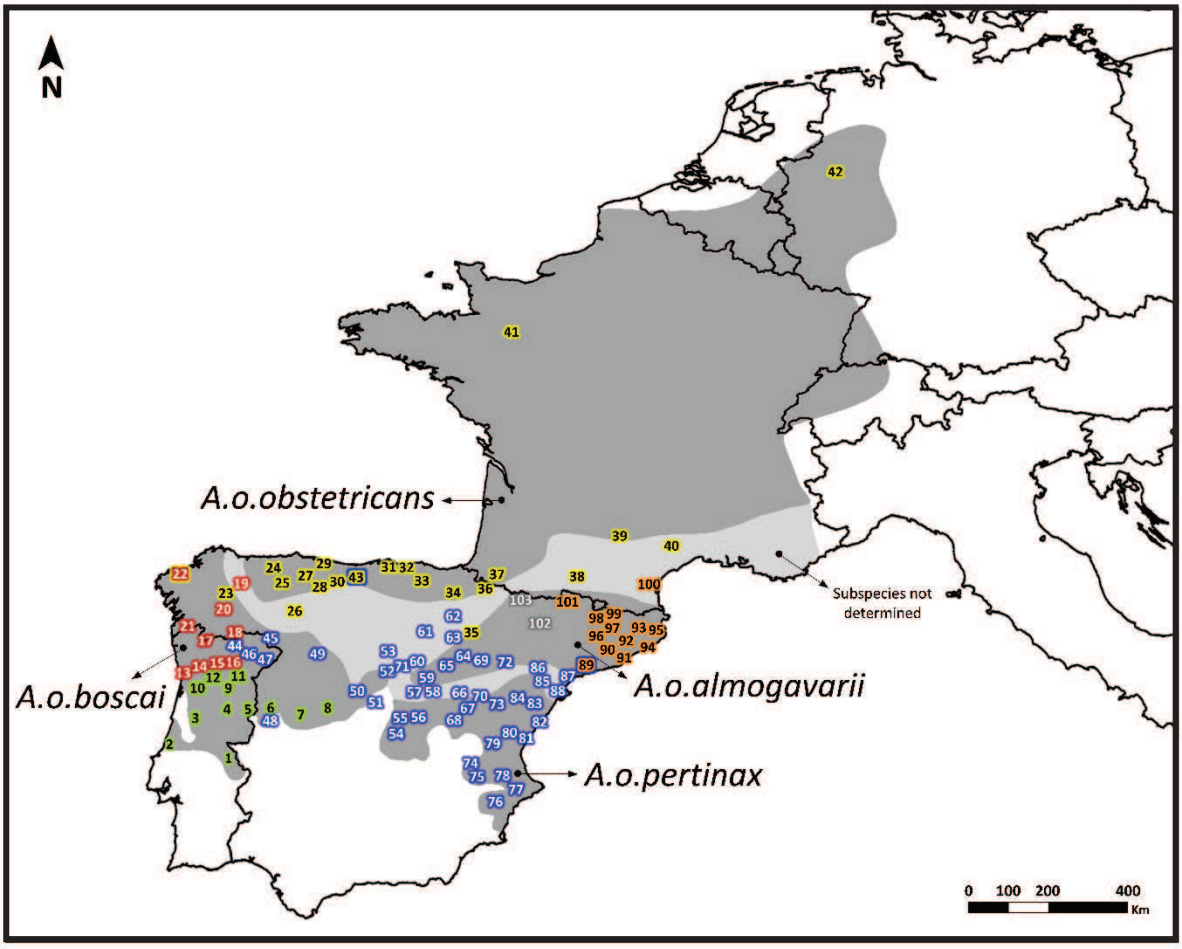


Fig. 1. Distribution map of Alytes obstetricans subspecies (dark shading, adapted from García-París and Martínez-Solano 2001) with indication of the geographical origin of all samples analyzed in the present study (population numbers as in Table 1). Note the broad areas where subspecific assignment of populations is undetermined (light shading). Colored circles around numbers represent the six mtDNA haplogroups recovered in the analyses. In three populations (22, 43 and 89), we found haplotypes of different haplogroups co-occurring. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

2.1.3 Material and methods

2.1.3.1 Sampling and DNA extraction

We analyzed 227 common midwife toads from 103 sampling sites spread across the species' entire geographical range (Fig. 1 and Table 1). Our sampling scheme was designed to incorporate individuals from all recognized subspecies, including their type localities when possible (Fig. 1). Some mitochondrial and nuclear sequences (identified in Table 1) were generated in previous genetic studies on *Alytes* (Gonçalves *et al.* 2007, 2009; Pinho *et al.* 2010, Maia-Carvalho *et al.* 2014a,b). Newly collected tissue samples were obtained from toe

tips of adults or tail tips of larvae, and preserved in 95% ethanol. All individuals were released in situ. Whole genomic DNA was extracted following the standard high-salt protocol of Sambrook *et al.* (1989) or using EasySpin Genomic DNA Minipreps Tissue Kit (SP-DT-250, Qiagen, Hilden, Germany) following the manufacturer's protocol.

2.1.3.2 Amplification, sequencing and haplotype determination

Five gene regions including one mitochondrial fragment of the NADH dehydrogenase subunit 4 gene and adjacent tRNAs (hereafter referred to as *ND4*) and four nuclear genes (β-fibrinogen intron 7 – *β-fibint7*; Protein phosphatase 3, catalytic subunit, alpha isoform intron 4 – *PPP3CAint4*; Ribosomal protein L9 intron 4 – *RPL9int4*; and a segment of exon 2 and intron 2 of the cellular myelocytomatosis proto-oncogene – *C-myc*) were amplified via polymerase chain reaction (PCR). The following primers were used for amplification and sequencing: for *ND4* – primers *ND4* and *Leu* (Arévalo *et al.* 1994); for *β-fibint7* a two-step amplification procedure was used, with a combination of two primer pairs – PCR1: *FIBX7* and *FIBX8*; PCR2: *BFXF* and *BFXR* (Sequeira *et al.* 2006); for *PPP3CAint4* – primers *PPP3CA4F1* and *PPP3CA5R1* (Pinho *et al.* 2010); for *RPL9int4* – *RPL94F* and *RPL95R* (Pinho *et al.* 2010); and for *C-myc* – primers *Cmyc1U* (Crawford 2003) and *Cmyc3cat* (Brunes *et al.* 2010). PCRs were carried out in 10 μL volume containing 1 X PCR buffer (50mmTris–HCl, 50mmNaCl, pH 8.5); 3mMMgCl2; 0.4mM each dNTPs, 0.5U of Phusion High-Fidelity PCR Master Mix (Thermo Scientific), 0.3 μM each primer and approximately 50 ng of genomic DNA. For *ND4* and *β-fibint7* amplification conditions were those described in Gonçalves *et al.* (2007). For *PPP3CAint4*, *RPL9int4* and *C-myc* amplification conditions followed Maia-Carvalho *et al.* (2014b). Purified products of each reaction were sequenced with the ABI Prism BigDye Terminator v3.1 Sequencing Kit protocol on an ABI3130xl DNA analyzer (Applied Biosystems, Foster City, California, USA). All sequences generated for this study are deposited in GenBank under Accession Numbers KT363119–KT363648 (Table 1).

Sequences were assembled with the software CHROMASPRO v1.5 (www.technelysium.com.au/ChromasPro.html). The assembled sequences were then edited and aligned manually using the program BIOEDIT v7.1.3.0 (Hall 1999).

Table 1. Samples analyzed in this study, with mtDNA haplogroups (see text), sample code, locality, country, population number (as in Fig. 1), latitude and longitude, and GenBank accession numbers for newly generated sequences. Sample numbers as in Figs. 2 and 3. --- : no data.

Species	mtDNA haplogroup	Sample Code	Locality	Country	Pop No.	Sample No.	Latitude	Longitude	GeneBank Accession No.				
									ND4	*ß-fibint7*	*C-myc*	*RPL9int4*	*PPP3CAint4*
A. obstetricans	D	SMA25	Rib. S. Bento, Serra S. Mamede	Portugal	1	1.1	39.317	-7.417	EF441303(a)	EF441330(a)	KJ859017(e)	KJ859049/ KJ859050(e)	KJ859036/ KJ859037(e)
		SMA26				1.2			KT363268	---	---	KT363640/ KT363641	KT363589/ KT363590
		NAZ1	Valado dos Frades, Nazaré	Portugal	2	2.1	39.596	-9.020	KF626441(d)	KT363445/ KT363446	---	---	---
		NAZ2				2.2			KF626442(d)	KT363447/ KT363448	KT363549	KT363633	KT363582/ KT363583
		IMS1460	Coimbra a Miranda do Corvo	Portugal	3	3.1	40.111	-8.379	KT363270	KT363449	---	---	---
		IMS1461				3.2			KT363271	KT363450	---	---	---
		EST26	Serra da Estrela	Portugal	4	4.1	40.333	-7.617	KT363266	---	---	---	---
		EST27				4.2			KT363267	---	---	---	---
		MAL4	Serra da Malcata	Portugal	5	5.1	40.306	-7.078	KT363264	KT363441/ KT363442	---	---	---
		MAL5				5.2			KT363265	KT363443/ KT363444	---	---	---
		GAT1	Puerto Nuevo, Sierra de Gata, Cáceres	Spain	6	6.1	40.361	-6.536	EF441301(a)	EF441327(a)	KT363541	KT363614/ KT363615	---
		GAT2				6.2			KT363260	KT363383/ KT363384	---	KT363616/ KT363617	KT363570/ KT363571
		GAT3				6.3			---	---	KT363542	---	---
		IMS2928	Cabezuela del Valle, Cáceres	Spain	7	7	40.211	-5.778	KT363272	KT363515	---	---	---
		IMS3640	Navarredonda de Gredos, Ávila	Spain	8	8	40.330	-5.116	KT363273	KT363516/ KT363517	---	---	---
		MAN1	Freixiosa, Mangualde	Portugal	9	9	40.600	-7.683	KT363263	KT363439/ KT363440	---	---	---
		VOU8	Rio Mau, Sever do Vouga	Portugal	10	10.1	40.733	-8.400	---	KT363438	---	---	---
		VOU11				10.2			KT363262	---	---	---	---
		FCO2	Penedono, V.N. Foz Côa	Portugal	11	11	40.983	-7.400	KT363269	---	---	---	---
		MTM1	Cinfães, Serra de Montemuro	Portugal	12	12.1	41.078	-8.024	EF441302(a)	EF441328/ EF441329(a)	KT363547	---	KT363579

	MTM2				12.2			KT363261	KJ859006/ KJ859007(e)	KT363548	KT363631/ KT363632	KT363580/ KT363581
	MTM5				12.3			---	---	KJ859018(e)	KJ859051(e)	KJ859038(e)
C	VAL22	Alfena, Valongo	Portugal	13	13.1	41.240	-8.525	EF441300(a)	EF441325/ EF441326(a)	KT363553/ KT363554	KT363645	KT363594
	VAL24				13.2			KT363251	---	---	---	---
	LOU2	Lousada	Portugal	14	14.1	41.303	-8.253	KT363252	KT363435	---	---	---
	LOU6				14.2			KT363253	KT363436/ KT363437	---	---	---
	VPA5	Cerva, Serra do Alvão	Portugal	15	15.1	41.467	-7.833	---	KT363424/ KT363425	---	---	---
	VPA6				15.2			---	KT363426	---	---	---
	VPA9				15.3			KT363248	---	---	---	---
	VPA13				15.4			KT363249	---	---	---	---
	MUR1	Jou, Murça	Portugal	16	16	41.483	-7.433	KT363250	KT363427/ KT363428	---	---	---
	GER1	Carris, Serra do Gerês	Portugal	17	17.1	41.817	-8.050	KT363246	KT363416	KT363543	KT363618	KT363572
	GER2				17.2			KT363247	KT363417/ KT363418	---	---	---
	VER1	Castrelo del Valle, Verín, Pontevedra	Spain	18	18.1	41.983	-7.417	KT363243	KT363370/ KT363371	---	---	---
	VER2				18.2			KT363244	KT363372/ KT363373	---	KT363646/ KT363647	KT363595/ KT363596
	LUG1	Fontaneira, Lugo	Spain	19	19.1	43.034	-7.200	KT363240	KT363366	KT363546	KT363630	KT363578
	LUG2				19.2			KT363241	---	---	---	---
	OUR2	Penalba, Ourense	Spain	20	20.1	43.425	-7.733	KT363245	KT363369	---	---	---
	OUR3				20.2			EF441295(a)	EF441319(a)	KJ859016(e)	GU181180(c)	GU181147(c)/ KJ859035(e)
	PON1	Monte Aloia, Tuy, Pontevedra	Spain	21	21.1	42.050	-8.633	EF441296(a)	EF441320(a)	KT363551/ KT363552	KT363636/ KT363637	KT363586
	MNCN8590				21.2			KT363259	KT363528	---	---	---
	IMS3624				21.3			KT363254	---	---	---	---
	IMS3625				21.4			KT363255	---	---	---	---
	IMS3626				21.5			KT363256	---	---	---	---

Species	mtDNA haplogroup	Sample Code	Locality	Country	Pop No.	Sample No.	Latitude	Longitude	GeneBank Accession No.				
									ND4	*ß-fibint7*	*C-myc*	*RPL9int4*	*PPP3CAint4*
		IMS3627				21.6			KT363257	---	---	---	---
		IMS3628				21.7			KT363258	---	---	---	---
		COR1	Arteixo, A Coruña	Spain	22	22.1	43.375	-8.433	KT363242	KT363367	---	---	---
	B	COR2				22.2			KT363221	KT363368	---	---	---
		IMS2630	Alfonxe, Lugo	Spain	23	23.1	42.924	-7.389	KT363229	KT363497/ KT363498	---	---	---
		IMS2631				23.2			KT363230	KT363499	---	---	---
		TIN1	Tineo, Asturias	Spain	24	24	43.333	-6.417	KT363219	KT363365	---	KT363642/ KT363643	KT363591/ KT363592
		MNCN4785	Somiedo, Asturias	Spain	25	25.1	43.070	-6.306	KT363235	KT363522	---	---	---
		MNCN4786				25.2			KT363236	KT363523/ KT363524	---	---	---
		LEO3	Veguellina de Órbigo, León	Spain	26	26	42.450	-5.885	KT363224	KT363364	---	---	---
		TLV1	Tolivia, Asturias	Spain	27	27	43.200	-5.583	EF441299(a)	EF441318(a)	---	---	---
		LEO1	Lago de Isoba, León	Spain	28	28.1	43.046	-5.315	KT363222	KT363362	---	KT363626/ KT363627	---
		LEO2				28.2			KT363223	KT363363	---	KT363628/ KT363629	---
		LEO4				28.3			---	---	KT363545	---	---
		MNCN8444				28.4			KT363237	KT363527	---	---	---
		OVI1	Puerto de San Isidro, Asturias			28.5	43.050	-5.325	KT363218	KT363360/ KT363361	---	---	---
		RIB1	El Fito, Asturias	Spain	29	29	43.433	-5.148	KT363220	KT363359	---	---	---
		IMS3621	Casavegas, La Pernía, Palencia	Spain	30	30	43.022	-4.513	KT363233	---	---	---	---
		SAN4	Fresnedo, Cantabria	Spain	31	31	43.367	-3.567	KT363214	KT363353/ KT363354	---	---	---
		SAN1		Spain	32	32.1	43.350	-3.333	KT363211	KJ859005(e)	KJ859013(e)	KT363638/ KT363639	KT363587/ KT363588
		SAN2				32.2			KT363212	KT363350	---	---	---
		SAN3				32.3			KT363213	KT363351/ KT363352	---	---	---
		IMS3009	P.N. Gorbeia, Vizcaya	Spain	33	33.1	43.088	-2.792	KT363231	---	---	---	---

	IMS3010				33.2			KT363232	KT363516/ KT363517	---	---	---
	MNCN4445	Sierra de Andía, Navarra	Spain	34	34	42.818	-1.970	KT363234	---	---	---	---
	IMS4096	Ribaforada, Navarra	Spain	35	35	42.006	-1.536	KT363239	---	---	---	---
	PAM1	Irati, Navarra	Spain	36	36.1	42.993	-1.100	KT363228	KT363348/ KT363349	KT363550	KT363634/ KT363635	KT363584/ KT363585
	PAM2				36.2			EF441293(a)	EF441317(a)	---	---	---
	IMS3959	Iraty Lake	France	37	37	43.046	-1.074	KT363238	---	---	---	---
	FRAN10	Saint-Michel	France	38	38	43.168	1.084	KT363227	---	---	---	---
	FRAN3	La Sallète	France	39	39	43.993	2.167	KT363226	KT363311	---	---	---
	FRAN14	St Pierre-de-la-Fage	France	40	40	43.794	3.420	---	---	---	KJ859048(e)	KJ859034(e)
	FRAN1	Jublains	France	41	41.1	48.251	-0.501	EF441292(a)	EF441316(a)	KJ859014/ KJ859015(e)	GU181178(c)	GU181145(c)
	FRAN2				41.2			KT363225	KT363310	---	---	---
	FRAN4				41.3			---	---	KT363540	---	---
	GERM1	Rhür Valley	Germany	42	42	51.517	7.450	EF441291(a)	EF441315(a)	---	---	---
	TUD2	Tudanca, Cantabria	Spain	43	43.1	43.150	-4.367	KT363215	KT363357/ KT363358	---	---	---
	TUD3				43.2			KT363216	---	---	---	---
	TUD4				43.3			KT363217	---	---	---	---
A	TUD1				43.4			KT363129	KT363355/ KT363356	---	KT363644	KT363593
	IMS3507	Feces de Abaixo	Portugal	44	44	41.810	-7.422	KT363204	---	---	---	---
	MON1	Refega, Serra de Montesinho	Portugal	45	45.1	41.783	-6.583	KT363142	KT363419	---	---	---
	MON12				45.2			KT363143	KT363420/ KT363421	---	---	---
	MON15				45.3			KT363144	KT363422/ KT363423	---	---	---
	MIR1	Mirandela	Portugal	46	46	41.483	-7.167	KT363145	KT363429/ KT363430	---	---	---
	MOG5	Mogadouro	Portugal	47	47.1	41.333	-6.717	KT363146	KT363431/ KT363432	---	---	---
	MOG6				47.2			KT363147	KT363433/ KT363434	---	---	---
	CAC1	Arroyo San Blas, Sierra de Gata, Cáceres	Spain	48	48.1	40.233	-6.600	KT363135	KT363385	---	---	---

Species	mtDNA haplogroup	Sample Code	Locality	Country	Pop No.	Sample No.	Latitude	Longitude	GeneBank Accession No.				
									ND4	*ß-fibint7*	*C-myc*	*RPL9int4*	*PPP3CAint4*
		CAC2				48.2			---	KT363386/ KT363387	---	---	---
		ZAM1	Valdefinjas, Zamora	Spain	49	49.1	41.451	-5.451	KT363130	KT363376/ KT363377	---	---	---
		ZAM2				49.2			EF441297(a)	EF441321(a)	---	---	---
		ZAM3				49.3			---	KT363378/ KT363379	---	---	---
		ZAM4				49.4			---	KT363380	---	---	---
		ZAM5				49.5			---	KT363381/ KT363382	---	---	---
		IMS3244				49.6			KT363202	KT363519	---	---	---
		IMS3245				49.7			KT363203	KT363520	---	---	---
		IMS4301	Vegas de Matute, Segovia	Spain	50	50.1	40.801	-4.258	KT363207	---	---	---	---
		IMS4302				50.2			KT363208	---	---	---	---
		IMS2749	Valsaín, Segovia	Spain	51	51.1	40.804	-4.023	KT363199	KT363513	---	---	---
		IMS2750				51.2			KT363200	KT363514	---	---	---
		GUA1	Pto. Cotos, Sierra de Guadarrama, Madrid			51.3	40.824	-3.962	KT363121	KT363388	---	KT363621/ KT363622	---
		GUA2				51.4			KT363122	KT363389/ KT363390	---	KT363623/ KT363624	KT363575
		IMS2669	Circo del Nevero, Madrid	Spain	52	52.1	40.979	-3.844	KT363197	KT363511	---	---	---
		IMS2670				52.2			KT363198	KT363512	---	---	---
		IMS4055	Villaverde de Montejo, Segovia	Spain	53	53	41.504	-3.666	KT363206	---	---	---	---
		TOL1	La Guardia, Toledo	Spain	54	54.1	39.783	-3.483	KT363133	KT363403	---	---	---
		TOL2				54.2			KT363134	KT363404	---	---	---
		MAD1	Colmenar de Oreja, Madrid	Spain	55	55.1	40.108	-3.387	KT363123	KT363397/ KT363398	---	---	---
		MAD4	Belmonte de Tajo, Madrid			55.2	40.133	-3.339	EF441298(a)	EF441322/ EF441323(a)	---	KJ859047(e)	KJ859033(e)
		MAD2	Valdelaguna, Madrid			55.3	40.162	-3.368	KT363124	KT363399/ KT363400	---	---	---
		MAD3				55.4			KT363125	KT363401/ KT363402	---	---	---

IMS2538	Valdaracete, Madrid	Spain	56	56	40.160	-3.179	KT363182	KT363492	---	---	---
GLJ2	Torija, Guadalajara	Spain	57	57.1	40.744	-3.030	KT363127	KT363393/ KT363394	---	KT363620	KT363573
GLJ3				57.2			KT363128	KT363395/ KT363396	KT363544	KT363619	KT363574
GLJ1	Ciruelas, Guadalajara			57.3	40.752	-3.090	KT363126	KT363391/ KT363392	---	---	---
IMS2620	Gárgoles, Guadalajara	Spain	58	58.1	40.757	-2.631	KT363185	---	---	---	---
IMS2621				58.2			KT363186	KT363495/ KT363496	---	---	---
IMS2571	Canredondo, Guadalajara			58.3	40.784	-2.544	KT363183	KT363493	---	---	---
IMS2572				58.4			KT363184	KT363494	---	---	---
IMS4358	Mirabueno, Guadalajara	Spain	59	59.1	40.948	-2.729	KT363209	---	---	---	---
IMS4359				59.2			KT363210	---	---	---	---
IMS3022	Cañamares, Guadalajara	Spain	60	60	41.227	-2.956	KT363201	KT363518	---	---	---
SOR1	Vinuesa, Soria	Spain	61	61	41.912	-2.763	KT363131	KT363374	---	---	---
IMS3866	Arnedo, La Rioja	Spain	62	62	42.224	-2.076	KT363205	---	---	---	---
OLV1	Ólvega, Soria	Spain	63	63	41.781	-1.985	KT363132	KT363375	---	---	---
IMS2267	Moros, Zaragoza	Spain	64	64.1	41.402	-1.820	KT363166	KT363467	---	---	---
IMS2268				64.2			KT363167	KT363468	---	---	---
IMS2247	Montuenga, Soria	Spain	65	65.1	41.225	-2.210	KT363164	---	---	---	---
IMS2248				65.2			KT363165	---	---	---	---
IMS2665	Terzaga a Peralejos de las Truchas, Guadalajara	Spain	66	66.1	40.647	-1.907	KT363195	KT363508	---	---	---
IMS2666				66.2			KT363196	KT363509/ KT363510	---	---	---
IMS2643	Nacimiento río Tajo, Teruel	Spain	67	67.1	40.321	-1.698	KT363191	KT363504	---	---	---
IMS2644				67.2			KT363192	KT363505/ KT363506	---	---	---
IMS2641	Frías de Albarracín, Teruel			67.3	40.362	-1.615	KT363189	KT363502	---	---	---
IMS2642				67.4			KT363190	KT363503	---	---	---
IMS2645	Guadalaviar, Teruel			67.5	40.395	-1.728	KT363193	KT363507	---	---	---

Species	mtDNA haplogroup	Sample Code	Locality	Country	Pop No.	Sample No.	Latitude	Longitude	GeneBank Accession No.				
									ND4	*ß-fibint7*	*C-myc*	*RPL9int4*	*PPP3CAint4*
		IMS2646				67.6			KT363194	---	---	---	---
		CUE1	Buenache de la Sierra, Cuenca	Spain	68	68	40.134	-2.000	EF441299(a)	EF441324(a)	---	GU181179(c)	GU181146(c)
		IMS2288	Tobed, Zaragoza	Spain	69	69.1	41.340	-1.399	KT363168	KT363469	---	---	---
		IMS2289				69.2			KT363169	KT363470	---	---	---
		IMS2633	Albarracín, Teruel	Spain	70	70.1	40.394	-1.416	KT363187	KT363500	---	---	---
		IMS2634				70.2			KT363188	KT363501	---	---	---
		IMS2522	Hayedo de Tejera Negra, Guadalajara	Spain	71	71	41.237	-3.350	KT363181	KT363490/ KT363491	---	---	---
		IMS2299	Belchite, Zaragoza	Spain	72	72.1	41.285	-0.776	KT363170	KT363471	---	---	---
		IMS2300				72.2			KT363171	KT363472	---	---	---
		IMS2126	Corbalán, Teruel	Spain	73	73.1	40.406	-0.986	KT363162	---	---	---	---
		IMS2127				73.2			KT363163	KT363466	---	---	---
		ALB1	Higueruela, Albacete	Spain	74	74.1	38.950	-1.450	KT363136	KT363411/ KT363412	---	---	KT363558/ KT363559
		ALB2				74.2			KT363137	---	---	KT363598	---
		IMS1851	Las Fuentes, Albacete	Spain	75	75.1	39.003	-1.304	KT363148	KT363451	---	---	---
		IMS1852				75.2			KT363149	KT363452/ KT363453	---	---	---
		ALI1	El Pinoso, Alicante	Spain	76	76.1	38.400	-1.033	KT363119	---	---	KT363599	---
		ALI2				76.2			KT363120	---	---	---	---
		ALI3				76.3			---	KT363413	---	KT363600	KT363560
		ALI4				76.4			---	KT363414/ KT363415	---	---	---
		IMS1932	Mas de Celedons, Alicante	Spain	77	77.1	38.669	-0.525	KT363152	KT363455	---	---	---
		IMS1933				77.2			KT363153	---	---	---	---
		IMS1901	Enguera, Valencia	Spain	78	78.1	38.929	-0.855	KT363150	KT363454	---	---	---

	IMS1902				78.2			KT363151	---	---	---	---
	IMS2006	Las Nogueras, Valencia	Spain	79	79.1	39.588	-1.082	KT363154	KT363456	---	---	---
	IMS2007				79.2			KT363155	KT363457	---	---	---
	IMS2021	Alcublas, Valencia	Spain	80	80.1	39.815	-0.700	KT363156	KT363458/ KT363459	---	---	---
	IMS2022				80.2			KT363157	KT363460	---	---	---
	VLC1	Algar de Palancia, Valencia	Spain	81	81.1	39.782	-0.367	KT363138	---	KT363555	---	---
	VLC2				81.2			KT363139	KT363409/ KT363410	KJ859009/ KJ859010(e)	KT363648	KT363597
	CAS1	Barranco de Santa Agueda, Benicassim, Castellón	Spain	82	82.1	40.050	0.066	KT363140	KT363405/ KT363406	---	KT363612/ KT363613	---
	CAS2				82.2			KT363141	KT363407/ KT363408	---	---	---
	IMS2071	Barranc dels Horts, Castellón	Spain	83	83.1	40.414	-0.066	KT363158	KT363461/ KT363462	---	---	---
	IMS2072				83.2			KT363159	KT363463/ KT363464	---	---	---
	IMS2091	Cantavieja, Teruel	Spain	84	84.1	40.516	-0.459	KT363160	---	---	---	---
	IMS2092				84.2			KT363161	KT363465	---	---	---
	IMS2317	Mazaleón, Teruel	Spain	85	85.1	41.061	0.110	KT363174	---	---	---	---
	IMS2318				85.2			KT363175	---	---	---	---
	IMS2314	Caspe, Zaragoza	Spain	86	86.1	41.148	0.013	KT363172	KT363473	---	---	---
	IMS2315				86.2			KT363173	KT363474	---	---	---
	IMS2367	Riudoms, Tarragona	Spain	87	87.1	41.148	1.043	KT363178	KT363478	---	---	---
	IMS2368				87.2			KT363179	KT363479	---	---	---
	IMS2337	El Perelló, Tarragona	Spain	88	88.1	40.858	0.670	KT363176	KT363475/ KT363476	---	---	---
	IMS2338				88.2			KT363177	KT363477	---	---	---
	IMS2387	El Port d'Armentera, Tarragona	Spain	89	89.1	41.387	1.359	KT363180	KT363480/ KT363481	---	---	---
F	IMS2388				89.2			KT363294	---	---	---	---
	IMS2407	Piera, Barcelona	Spain	90	90.1	41.511	1.737	KT363295	KT363482	---	---	---
	IMS2408				90.2			KT363296	---	---	---	---

Species	mtDNA haplogroup	Sample Code	Locality	Country	Pop No.	Sample No.	Latitude	Longitude	GeneBank Accession No.				
									ND4	*ß-fibint7*	*C-myc*	*RPL9int4*	*PPP3CAint4*
		BAR1	Collserola, Barcelona	Spain	91	91	41.419	2.099	---	---	---	KT363601/ KT363602	KT363561
		IMS2427	Sant Miquel del Fai, Barcelona	Spain	92	92.1	41.716	2.193	KT363297	KT363483	---	---	---
		IMS2428				92.2			KT363298	---	---	---	---
		IMS2447	Sant Hilari Sacalm, Girona	Spain	93	93.1	41.903	2.506	KT363299	KT363484/ KT363485	---	---	---
		IMS2448				93.2			KT363300	---	---	---	---
		MNCN4334	Riudarenes, Girona	Spain	94	94	41.822	2.717	KT363305	KT363521	---	---	---
		IMS2467	La Bisbal, Girona	Spain	95	95.1	41.939	2.949	KT363301	KT363486	---	---	---
		IMS2468				95.2			KT363302	KT363487	---	---	---
		IMS2492	Prades, Lleida	Spain	96	96.1	41.795	1.579	KT363303	---	---	---	---
		IMS2493				96.2			KT363304	KT363488	---	---	---
		BER5	Llinars, Lleida	Spain	97	97.1	42.131	1.709	KT363279	KT363314/ KT363315	KT363532	KT363605	KT363564
		BER10				97.2			KT363283	KT363319/ KT363320	---	---	---
		BER11				97.3			KT363284	KT363321/ KT363322	---	---	---
		BER12				97.4			KT363285	KT363323/ KT363324	---	---	---
		BER13				97.5			KT363286	KT363325/ KT363326	---	KT363606/ KT363607	KT363566
		BER14				97.6			KF626443(d)	KT363327/ KT363328	KT363534/ KT363535	---	---
		BER15				97.7			KF626444(d)	KT363329/ KT363330	KT363536/ KT363537	---	---
		BER16				97.8			KT363287	KT363331/ KT363332	---	---	---
		BER6				97.9			EF441308(a)	EF441337(a)	---	---	---
		BER7				97.10			KT363280	---	---	---	---
		BER1	Rasos de Peguera, Berga, Barcelona			97.11	42.136	1.761	EF441305(a)	EF441332/ EF441333(a)	KJ859019/ KJ859020(e)	GU181177(c)/ KJ859052(e)	GU181144(c)
		BER2				97.12			KT363278	KT363312/ KT363313	---	---	---
		BER3				97.13			EF441306(a)	EF441334/ EF441335(a)	KJ859021(e)	KJ859053(e)	KJ859039/ KJ859040(e)

		BER4				97.14			EF441307(a)	EF441336(a)	KT363531	KT363603/ KT363604	KT363562/ KT363563
		BER8				97.15			KT363281	KT363316	---	---	---
		BER9				97.16			KT363282	KT363317/ KT363318	KT363533	---	KT363565
		BER20	La Coma, Berga, Barcelona	Spain	98	98.1	42.188	1.581	KT363290	KT363339/ KT363340	KT363539	KT363610/ KT363611	KT363568/ KT363569
		BER21				98.2			KT363291	KT363341/ KT363342	---	---	---
		BER22				98.3			KT363292	KT363343/ KT363344	---	---	---
		BER23				98.4			KT363293	KT363345	---	---	---
		BER17	Soldes, Túnel del Cadí, Berga, Barcelona	Spain	99	99.1	42.350	1.700	---	KT363333/ KT363334	---	---	---
		BER18				99.2			KT363288	KT363335/ KT363336	KT363538	KT3636087 KT363609	KT363567
		BER19				99.3			KT363289	KT363337/ KT363338	---	---	---
		IMS4107	Tuchan	France	100	100	42.885	2.733	KT363306	---	---	---	---
		BEN1	Fonchanina, Huesca	Spain	101	101	42.522	0.652	KT363277	KT363347	---	---	---
	E	IMS2501	Naval, Huesca	Spain	102	102.1	42.188	0.154	KT363275	KT363489	---	---	---
		IMS2502				102.2			KT363276	---	---	---	---
		HUE1	Ibón de Piedrafita, Huesca	Spain	103	103.1	42.700	-0.333	EF441304(a)	EF441331(a)	KJ859022(e)	---	---
		HUE2				103.2			KT363274	KT363346	---	---	KT363576
		HUE4				103.3			---	---	---	KT363625	KT363577
A. cisternasii		CER17	Cercal	Portugal	-	-	37.750	-8.650	EF441314(a)	---	---	---	---
A. cisternasii		CER01	Cercal	Portugal	-	-	37.750	-8.650	---	EF441343(a)	---	KJ859062(e)	---
A. cisternasii		HG113	Río Adaja, Ávila	Spain	-	-	40.650	-4.700	---	---	KT363557	---	---
A. dickhilleni		IMS3489	Puerto de las Crucetillas, Albacete	Spain	-	-	38.545	-2.383	KT363307	---	---	---	---
A. dickhilleni		HG103	Puerto de las Crucetillas, Sierra de Alcaraz, Albacete	Spain	-	-	38.545	-2.383	---	EF441338(a)	---	GU181176(c)	GU181143(c)
A. dickhilleni		HG104	Puerto de las Crucetillas, Sierra de Alcaraz, Albacete	Spain	-	-	38.545	-2.383	---	---	KJ859026/ KJ859025(e)	---	---
A. muletensis		IMS3517	Captivity	Spain	-	-	-	-	KT363309	---	---	---	---

Species	mtDNA haplogroup	Sample Code	Locality	Country	Pop No.	Sample No.	Latitude	Longitude	GeneBank Accession No.				
									ND4	*ß-fibint7*	*C-myc*	*RPL9int4*	*PPP3CAint4*
A. maurus		MAR01	Bab Bou Idir, Taza	Morocco	-	-	34.060	-4.110	KT363308	---	---	---	---
A. maurus		MAR04	Bab Bou Idir, Taza	Morocco	-	-	34.060	-4.110	---	EF441341(a)	---	KJ859056(e)	KJ859043(e)
A. maurus		MAR06	Bab Bou Idir, Taza	Morocco	-	-	34.060	-4.110		---	KT363556		
A. muletensis		MAI06	Sierra de Tramuntana, Mallorca	Spain	-	-	39.580	2.500	---	EF441339(a)	KJ859029(e)	GU181175(c)	GU181142(c)

(a) Gonçalves *et al.* (2007); (b) Gonçalves *et al.* (2009); (c) Pinho *et al.* (2010); (d) Maia-Carvalho *et al.* (2014a); (e) Maia-Carvalho *et al.* (2014b)

Polymorphic positions of *β-fibint7*, *PPP3CAint4*, *RPL9int4* and *C-myc* corresponding to heterozygous individuals were coded with IUPAC ambiguity codes. For each input genotype sequence of heterozygous individuals, we inferred phased haplotypes probabilistically through the Bayesian algorithm implemented in PHASE v2.1.1 (Stephens *et al.* 2001; Stephens and Donnelly 2003), using SEQPHASE (Flot 2010) to format the input files. In the four nuclear loci we detected single and multiple-base insertions or deletions (indels). For phasing analyses we pruned the data assuming that indels likely resulted from a single evolutionary step. We left only the first base of the indel (in the case of an insertion) or reduced them to one single step (deletion). We choose this approach rather than completely removing indels because this would significantly reduce the number of polymorphic sites and disregard some of the information contained in the data sets. Heterozygous indels, which resulted from the amplification of alleles of different sizes in a single individual, were decoded interpreting directly the mixed trace formed by the two allelic peaks superimposed onto each other downstream of the indel (Sousa-Neves *et al.* 2013). All known haplotypes were incorporated for subsequent haplotype inference. We ran PHASE three times with different random seeds and checked if haplotype estimation was consistent across runs. Each run was conducted using default values. We used a threshold of 0.90 posterior probability to accept a given haplotype phase reconstruction.

2.1.3.3 Data analysis

A gene tree for the mtDNA marker *ND4* was reconstructed using maximum likelihood (ML) and Bayesian inference (BI) analyses. The most appropriate model of nucleotide evolution was selected using PARTITIONFINDER v1.1.0 (Lanfear *et al.*, 2012) under the Akaike information criterion (AIC; Akaike 1973). ML analyses were performe with the software RAXML v7.2.8 (Silvestro and Michalak 2010), using the graphical front-end RAX-ML GUI v1.1 (Randomized Axelerated Maximum Likelihood; Stamatakis 2006). Through the bootstrap option, ML analyses were run ten times from starting random seeds to generate 1000 nonparametric bootstrap replicates. Bayesian analyses were performed in the version of MRBAYES v3.1.2 hosted at the CIPRES Science Gateway Portal v3.1 (San Diego Supercomputer Center; Miller *et al.* 2010; http://www.phylo.org/portal/), using two replicate searches with 10 x 10^7 generations each and sampling every 10,000th generations. Four MCMC (Markov chain Monte Carlo) were run simultaneously in each analysis. Stationarity for each run was detected through three different ways. First, we assessed the convergence among chains by plotting the log-likelihood values against generation number using TRACER v1.6 (Rambaut *et al.* 2014). Secondly, we used the online program AWTY (Nylander *et al.*

2008) to analyze the trace plot of the log-likelihood and the cumulative split frequencies across all post burn-in generations within each analysis. Finally, we checked the standard deviation of split frequencies as a convergence index (<0.001). After assessing chain convergence, we discarded all samples obtained during the first ten million generations as burn-in. Post-burn-in trees from all replicates were combined estimating a 50% majority-rule consensus tree. The frequency of any particular clade in the consensus tree represents the posterior probability of that clade (Huelsenbeck and Ronquist 2001).

To test for recombination in nuclear loci we used the difference in sums of squares (DSS) method as implemented in software TOPALi v2.5 (Milne *et al.* 2004), with a sliding window of 100-bp and 10-bp step size.

Genealogical relationships among haplotypes (haplotype networks) for each locus were estimated using phylogenetic algorithms with proper models of sequence evolution, as implemented in HAPLOVIEWER (Salzburger *et al.* 2011). Phylogenetic reconstructions among haplotypes for each locus were estimated using a Maximum Likelihood (ML) approach, as implemented in the software RAxML. Using default options, we ran the program with the best-fit model for each locus as selected by PARTITIONFINDER, and the generated trees were used to estimate each haplotype network.

For both mtDNA and nuclear fragments genetic diversity parameters were estimated using DNASP v5.10 (Librado and Rozas 2009). We also used DNASP to perform a pairwise mismatch-distribution analysis (Rogers and Harpending 1992), and to calculate Fu's Fs (Fu 1997), Tajima's D (Tajima 1989) and Ramos-Onsins & Rozas' R2 (Ramos-Onsins and Rozas 2002) statistics, in order to test for molecular signatures of demographic expansion. Genetic distances (*p*-uncorrected) within and between lineages were calculated with MEGA v5 (Tamura *et al.* 2011). Additionally, we performed an Extended Bayesian Skyline Plot (Heled and Drummond 2008) analysis using all available sequences for *A. obstetricans* (*ND4*: 210; *β-fibint7*: 169; *C-myc*: 30; *PPP3CAint4*: 38; and *RPL9int4*: 41; numbers refer to individuals, and for nuclear markers the number of sequences is twice that number because there are two copies for each individual). This coalescent-based analysis estimates the effective population size (*Ne*) through time and is implemented in BEAST v.1.8 (Drummond *et al.* 2012). Optimal nucleotide-substitution models were selected by JMODELTEST v2.1 (Darriba *et al.* 2012). For models incorporating both gamma-distributed rate variation (+G) and proportion of invariant sites (+I) we implemented only + G with 10 rate categories due to the computational difficulty of simultaneously estimating both + G and + I.

We analyzed our multilocus dataset under the multispecies coalescent as implemented in *BEAST v1.8 (Heled and Drummond 2010). The alignments included *ND4* sequences of 174 individuals, and phased nuclear sequences (two alleles per individual) for 27 (*C-myc*), 35 (*PPP3CAint4*), 35 (*RPL9int4*), and 140 (*β-fibint7*) individuals. We defined five independent

partitions, one for each marker. The optimal substitution models for each partition were selected by JMODELTEST. Molecular clocks (strict for all partitions) and tree priors were unlinked across partitions. As species-tree prior, we used the Yule speciation model. We used samples of all extant *Alytes* species, including *A. cisternasii*, *A. muletensis*, *A. dickhilleni* and *A. maurus* as outgroups in the analyses. We also defined six groups in *A. obstetricans* corresponding to the six major population lineages identified in previous analyses (A – F, see Section 3). Samples from the three localities where two mtDNA haplotype clades were found in sympatry (22, 43, and 89) were excluded from the analyses since *BEAST does not take into account the possibility of gene flow across lineages. Additionally, we also excluded individuals from localities where gene flow was suspected because of geographical proximity and/or patterns of nuclear haplotype sharing. Localities excluded from this analysis for this reason included 6, 16, 19, 23, 28, 31–36, 44–48, 62–63, and 102. In the absence of prior information about clock rates for the markers used, we performed an additional phylogenetic analysis on a reduced *ND4* dataset including representatives of all major mtDNA haplogroups within *A. obstetricans* (sample codes: ALB1, BER6, HUE1, MNCN8590, SMA25, and TIN1, see Table 1), all other species of *Alytes* (*A. cisternasii* CER17; *A. muletensis* IMS3517; *A. maurus* MAR01; and *A. dickhilleni* IMS3489), and sequences of the related genera *Discoglossus* and *Bombina* downloaded from GenBank (*Discoglossus galganoi* AY442088; *D. jeanneae* AY442115; *D. pictus* AY442137; *D. scovazzi* AY442139; *Bombina bombina* JX893173; *B. variegata* AY971143; *B. variegata* JX893176; *B. variegata* JX893179; *B. orientalis* AY585338; *B. microdeladigitora* JX893182; and *B. maxima* JX893181) (Martínez-Solano 2004, San Mauro *et al.* 2004, Pabijan *et al.* 2013). This reduced dataset was analyzed as a single partition in EAST, with the optimal substitution model selected by JMODELTEST. We used the Yule speciation prior with a strict clock and calibrated the node separating eastern and western Palearctic Bombina species following Pabijan *et al.* (2013), with a logNormalPrior encompassing values between 13.4 and 32.2 Mya (mean = 0.994, st. dev. = 1, offset = 13.0, meanInRealSpace = false). The mean and 95% Highest Posterior Density interval (95% HPD) estimated in this analysis for the parameter ‘‘clock rate” were subsequently used to specify a normal prior for the *ND4* clock rate in the *BEAST analysis (mean: 0.0085, st. dev. = 0.002), with diffuse Gamma priors (shape: 0.01, scale: 100) for the clock rates of nuclear markers. Species-tree analyses were run for 200 million generations, and the logfile was inspected in TRACER to assess adequate mixing and convergence. All effective sample sizes (ESS) of parameters estimated were well > 200, as recommended by the authors of the software. A burn-in period of 10% of the total running time was specified after trace inspection, and a Maximum Clade Credibility consensus tree was constructed with the program TREEANNOTATOR, which is distributed as part of the BEAST package. A similar analysis, but

excluding the mtDNA partition was also run to test for the effect of mtDNA on the species tree (see for instance Jockusch *et al.* 2015).

2.1.4 Results

The mitochondrial DNA alignment consisted of 210 sequences of 813 base pairs (bp) yielding 88 haplotypes defined by 138 polymorphic sites, of which 127 were parsimony informative (Table 2). For the nuclear dataset we obtained 338 sequences with 630– 635 bp for *β-fibint7*, 76 sequences with 597–604 bp for *PPP3CAint4*, 82 sequences with 442–460 bp for *RPL9int4*, and 60 sequences with 1257–1261 bp for *C-myc* (Table 2). In the *β-fibint7* alignment, 57 out of 59 polymorphic sites were parsimony informative. For *PPP3CAint4*, 22 out of 26 polymorphic sites were parsimony informative. For *RPL9int4* and *C-myc*, 27 and 24 out of 29 and 26 polymorphic sites were parsimony informative, respectively. We did not find significant evidence of recombination in the nuclear markers.

Phylogenetic analyses of mtDNA sequences recovered a well-resolved tree with six major haplotype clades (labeled A–F, Fig. 2), with a strong association with geography (Fig. 1), diagnosing six population lineages. Briefly, haplogroup A (blue in the figures) corresponds to populations of *A. o. pertinax* but extends much further than previously thought, along the Iberian northern plateau up to the Cantabrian Mountains and in the east end of the Central System mountains (Sierras de Guadarrama and Ayllón), filling most of the region not ascribed to any subspecies in previous studies; haplogroup B (yellow) corresponds to populations of *A. o. obstetricans* and extends from northern Spain (Cantabrian Mountains) to central Europe and also to the west, up to Galicia; haplogroup C (red) corresponds to populations of *A. o. boscai* and occurs in Galicia and northwest Portugal; haplogroup D (green) includes all Portuguese populations south of the Douro River and the Spanish populations from the western part of the Central System (Sierras de Gata and Gredos, traditionally considered as part of *A. o. boscai*); haplogroup E (gray) was only detected in populations of the Spanish central Pyrenees, while haplogroup F (orange) corresponds to populations of *A. o. almogavarii*. Haplogroup D is the sister taxon to a clade comprising the other five haplogroups (BPP: 1.0), and haplogroups A, B and C form a monophyletic group (BPP: 1.0). In three populations, we found haplotypes of different haplogroups co-occurring (Table 2 and Fig. 1). These include localities 22 (haplogroups B + C), 43 (haplogroups A + B) and 89 (haplogroups A + F).

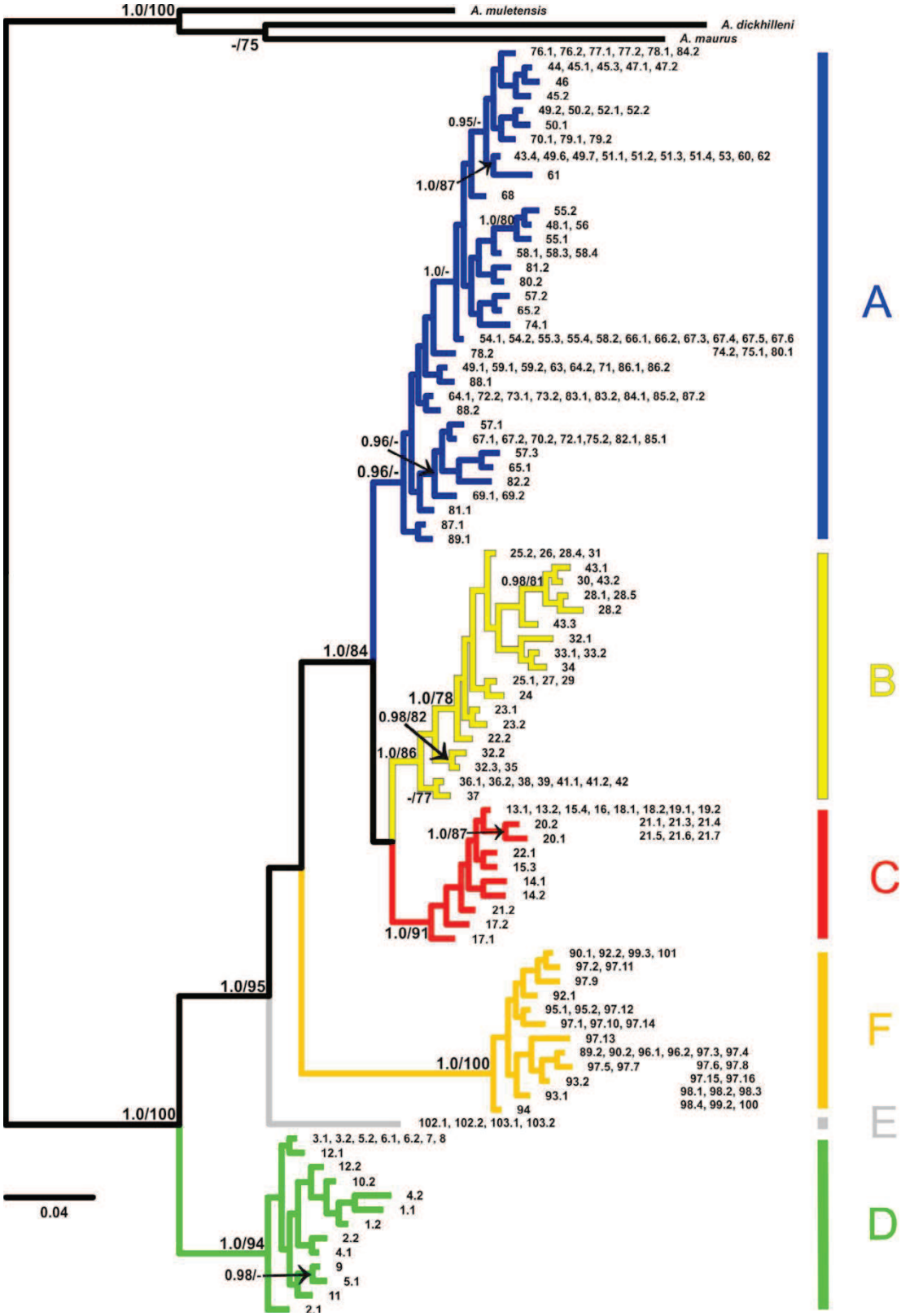


Fig. 2. Phylogenetic tree inferred from a Bayesian analysis of sequences of the mitochondrial *ND4* gene (813 bp) in *Alytes obstetricans*. Posterior probabilities and maximum likelihood bootstrap values of well-supported nodes (BPP ≥ 0.9 / BS ≥ 75) are shown at nodes. Sample codes as in Table 1.

Table 2. Summary statistics for the five molecular markers sequenced in *A. obstetricans* in this study. (N) number of sequences; (S) number of segregating sites, in parentheses, number of parsimony-informative sites; (H) number of haplotypes; (Hd) haplotype diversity; (p) nucleotide diversity. High levels of genetic diversity were found in *A. obstetricans*. Significant values of Fu's Fs and Ramos-Onsins & Rozas' R2 statistics were obtained for the population lineages corresponding to mtDNA haplogroups A, C and D (*P < 0.05; **P < 0.01), suggesting demographic expansions.

Marker	Haplogroups	N	S	H	Hd	π	Tajima's D	Fu's Fs	R_2
ND4 (813bp)		210	138 (127)	88	0.975±0.004	0.0243±0.0012	-0.475	-24.076**	0.0765
	A	95	37	34	0.942±0.010	0.0042±0.0003	-1.657	-23.384**	0.0433*
	B	33	24	18	0.936±0.027	0.0063±0.0005	-0.467	-5.691*	0.0960
	C	23	19	10	0.640±0.116	0.0024±0.0007	-2.262**	-4.111*	0.0512**
	D	19	15	13	0.877±0.074	0.0029±0.0007	-1.681	-8.883**	0.0665**
	E	4	0	1	-	-	-	-	-
	F	36	15	12	0.787±0.064	0.0029±0.0004	-1.119	-3.788*	0.0778
b-fibint7 (630–635 bp)		338	59 (57)	108	0.971±0.004	0.0155±0.0003	0.175	-73.961**	0.0857
C-myc (1257–1261 bp)		60	26 (24)	18	0.904±0.024	0.0035±0.0002	-0.713	-3.446	0.0892
PPP3CAint4 (597–604 bp)		76	26 (22)	21	0.900±0.017	0.0075±0.0004	-0.486	-4.508	0.1003
RPL9int4 (442–460 bp)		82	29 (27)	26	0.908±0.023	0.0124±0.0007	-0.188	-6.392*	0.1014

Haplotype networks based on mtDNA and nuclear sequences are presented in Fig. 3. The mtDNA haplotype genealogy showed the same six groups as the phylogenetic tree. Within each haplogroup, the most frequent haplotypes generally show an interior position, whereas the remaining are in most cases closely connected by one-step mutations. Haplogroups C and D exhibit star-like topologies, with one high frequency, central haplotype and several additional haplotypes connected by few-step mutations, suggesting recent demographic expansion. The most frequent haplotype in haplogroup C is present in 14 out of 23 sequences and is widely distributed, with 9 haplotypes occurring only once. The most frequent haplotype in haplogroup D occurs in 7 out of 19 individuals from the Iberian Central System, with 12 haplotypes with a frequency of 1. In haplogroup B the most frequent haplotype (frequency = 7) is almost exclusive of the French and German populations. We found a total of 18 haplotypes in this haplogroup, with frequencies ranging from 1–7. Haplogroup A has several frequent haplotypes that are widely distributed. We found 34 haplotypes in this haplogroup, with a highest frequency of 14, and 22 haplotypes observed only once. The most frequent haplotype in haplogroup F occurs in 16 out of 36 individuals and is widely distributed through Catalonia. Finally, haplogroup E is represented by a single haplotype. The nuclear networks were generally compatible with mtDNA in the diagnosis of six major population lineages, with similar results for each of the four genes (Fig. 3). In the *C-myc* network, all population lineages had non-overlapping sets of haplotypes, whereas the other three nuclear networks (*β-fibint7*, *PPP3CAint4* and *RPL9int4*) showed the presence of heterozygous individuals with alleles characteristic of different inferred population lineages (Figs. 1 and 3).

Measures of genetic diversity for the mitochondrial gene and the four nuclear loci are summarized in Table 2. We found high levels of genetic diversity in *A. obstetricans* (π *ND4* = 0.0243; π *β-fibint7* = 0.0155; π *RPL9int4* = 0.0124; π *PPP3CAint4* = 0.0075; π *C-myc* = 0.0035). The mitochondrial average genetic distance (*p*-uncorrected) ranged from 5.2% between haplotype clades D and F to 1.4% between haplotype clades B and C (Table 3). Within haplogroups, *p*-distances were higher in haplogroups B and C (0.70%) than in the other haplogroups (0.00–0.40%, Table 3). Significant values of Fu's Fs and Ramos-Onsins & Rozas' R2 statistics were obtained for the population lineages corresponding to mtDNA haplogroups A, C and D (Table 2). Tajima's D statistic was also significant for the population lineage corresponding to mtDNA haplogroup C. Mismatch distributions showed similar patterns, supporting scenarios of demographic expansion in the population lineages A, C and D (Fig. 4). The EBSP showed a sustained increase in *Ne* through the Pleistocene and up to the present (Fig. 5). Species-tree analyses recovered largely unresolved topologies (Fig. 6), with lineage F as the sister taxon to a clade comprising the other five lineages but with low support (BPP < 0.9), with only lineage A being recovered as the sister taxon to lineage B (BPP: 0.99). Estimates of split times were calculated for several nodes in the species-tree, with median

values for the root of 21.1 Mya (95% HPDi: 10.1–37.9), 5.6 Mya (95% HPDi: 3.0–9.8) for the basal split in *Baleaphryne* (the clade that includes *A. dickhilleni*, *A. maurus* and *A. muletensis*), and 2.5 Mya for the basal split in *A. obstetricans*, separating lineage F from the rest (95% HPDi: 1.4 – 4.3). Thus, the splits between the six major populations lineages in *A. obstetricans* date back to the Pleistocene, although their relative splitting order is unresolved (Fig. 6). The same largely unresolved topology was recovered in analyses based on the four nuclear markers, with a weakly supported basal split separating lineage F from the rest (BPP: 0.87) and strong support for a sister-group relationship between lineages A and B (BPP: 0.99).

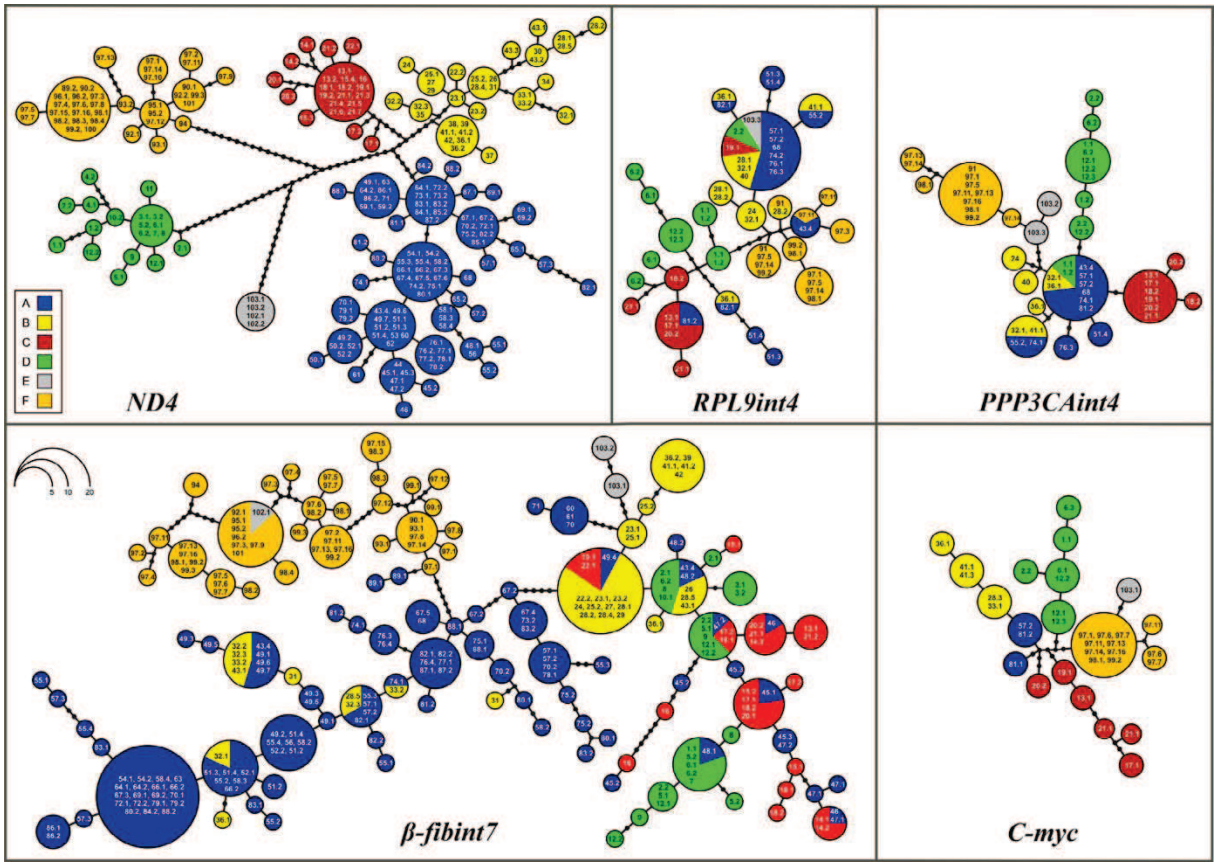


Fig. 3. Haplotype genealogies from a maximum likelihood analysis of mtDNA *ND4*, and nuclear *RPL9int4*, *PPP3CAint3*, *β-fibint7* and *C-myc* genes performed with the software HAPLOVIEWER. Colours represent the different mtDNA haplogroups (as in Figs. 1 and 2, see text for details). Each circle represents a different haplotype and its size is proportional to its relative frequency (see scale). Dots represent inferred unsampled or extinct haplotypes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Table 3. Average genetic distances (*p*-uncorrected) within (on the diagonal, including average and range) and between the main mtDNA haplogroups in *Alytes obstetricans*.

mtDNA Haplogroup	A	B	C	D	E	F
A	0.004 (0.001 - 0.012)	0.015	0.016	0.040	0.033	0.045
B		0.007 (0.001 - 0.012)	0.014	0.043	0.033	0.045
C			0.007 (0.001 - 0.023)	0.044	0.036	0.045
D				0.003 (0.001 - 0.009)	0.039	0.052
E					0.00	0.046
F						0.003 (0.001 - 0.01)

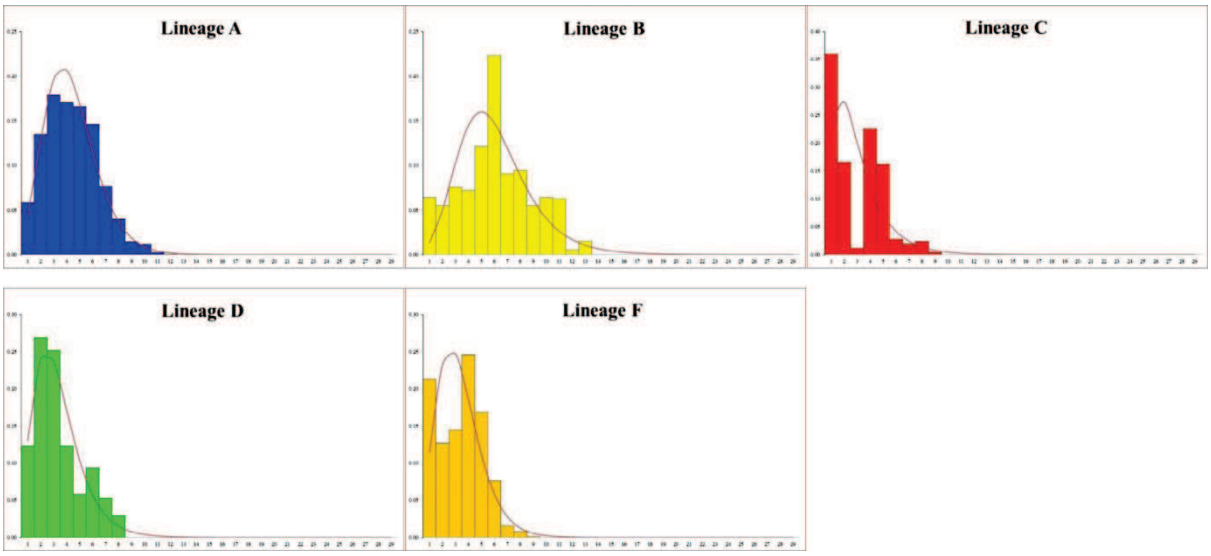


Fig. 4. Mismatch-distributions for each population lineage. Red curves show the expected distribution of mutations according to the null hypothesis of demographic expansion. The number of pairwise differences and their frequencies are shown on the horizontal and vertical axes, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

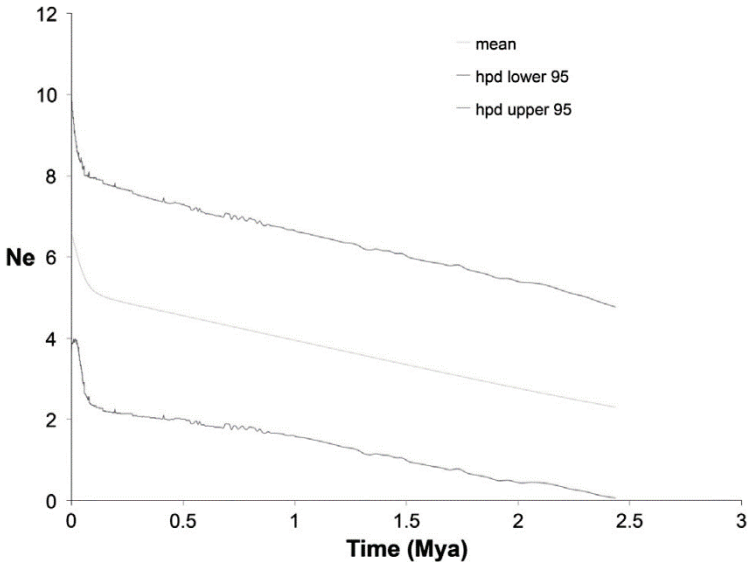


Fig. 5. Extended Bayesian Skyline Plot showing a sustained increase in effective population size (*Ne*) in *A. obstetricans* since the Pleistocene (horizontal axis, scale in millions of years). Lines represent the mean and upper and lower limits of the 95% HPDi.

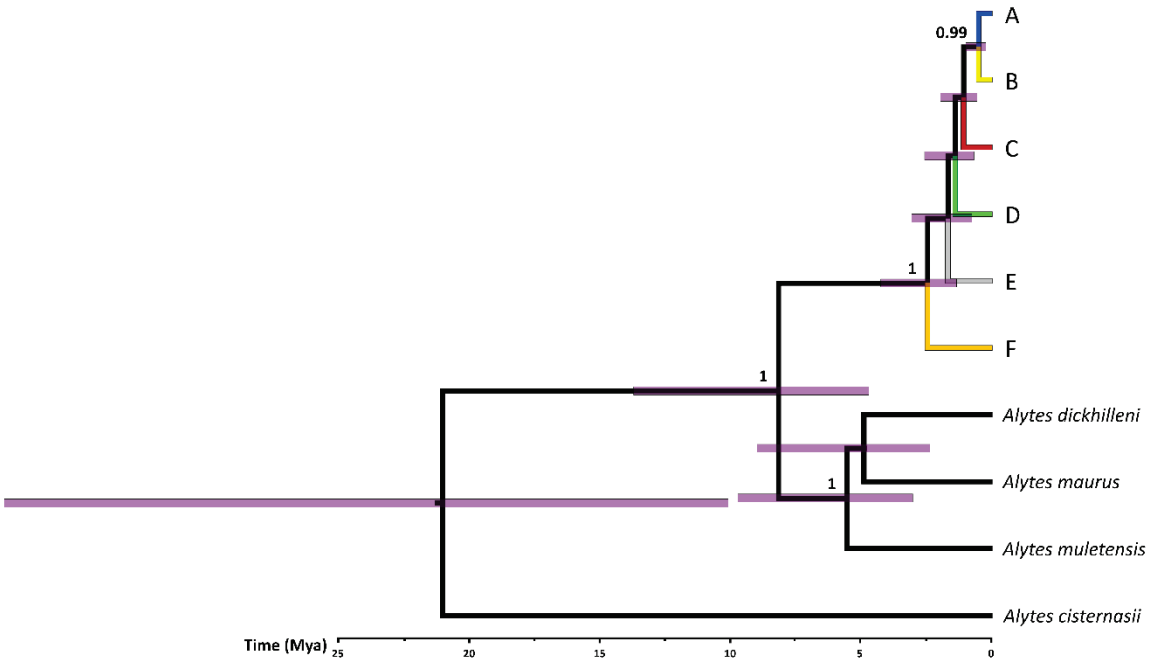


Fig. 6. *Alytes* species tree including all major population lineages in *Alytes obstetricans* based on the coalescent-based analysis of sequences of mtDNA *ND4* and four nuclear genes (*RPL9int4*, *PPP3CAint3*, *β-fibint7* and *C-myc*) in *BEAST. Values at nodes indicate Bayesian posterior probabilities. Bars show 95% highest posterior density intervals for split times. Scale (bottom) in millions of years.

2.1.5 Discussion

Our analysis of genetic diversity in *A. obstetricans* revealed very high levels of intraspecific variation in Iberia, in agreement with previous results based on allozymes (Arntzen and García-París 1995), morphology (Martínez-Solano *et al.* 2004), and mtDNA and nuclear markers (Fonseca *et al.* 2003, Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007, Maia-Carvalho *et al.* 2014a). The common midwife toad has been previously considered as including four essentially parapatric subspecies in the Iberian Peninsula (*A. o. obstetricans*, *A. o. boscai*, *A. o. pertinax*, and *A. o. almogavarii*; see Fig. 1 for their respective geographical distributions), but our study shows additional and different patterns of genetic subdivision and allows finer delineation of their ranges. Based on mtDNA data, both the phylogenetic tree and the haplotype genealogy clearly recovered six different haplogroups, three off which are highly divergent (D, E and F, with average *p*-uncorrected distances with other haplogroups of 4.4%, 3.7%, and 4.7%, respectively, see Table 3). These mtDNA haplogroups diagnose six population lineages with very strong geographical concordance: only in three locations we documented the presence of more than one lineage (Fig. 1). Although the nuclear genealogies are in general mutually compatible with mtDNA data, they do not recover monophyly of the six mtDNA-defined haplogroups. Probably this result reflects the retention of ancestral polymorphisms in the slowest-evolving markers.

Our species-tree analysis recovered largely unresolved trees (Fig. 6), which may be a consequence of insufficient resolution in the markers used or of nearly simultaneous divergence across population lineages. Nevertheless, the species tree recovered lineage F as the sister taxon to a clade comprising the other five extant lineages, albeit with low support (BPP: 0.78 in the combined vs BPP: 0.87 in the nuclear species tree). Lineage F includes populations previously ascribed to *A. o. almogavarii*, and was estimated to diverge from the other lineages around 2.5 Mya (1.4 – 4.3). The combined evidence from mtDNA and nuclear markers suggests that *A. o. almogavarii* is an ancient form restricted to the northeastern corner of the Iberian Peninsula, where it persisted probably since the Pleistocene. Its presence in SE France, anticipated by Geniez and Crochet (2003), was confirmed in our study, but the contact zone with *A. obstetricans* in France remains to be described in detail in future studies. Apparently, this lineage does not cross the Ebro valley, in contradiction with previous results that suggested an extended geographical distribution reaching the mountains of Guadarrama, close to Madrid (García-París 1995). Our results also contradict the initial hypothesis of García-París (1995), later expanded by Martínez-Solano *et al.* (2004) and Gonçalves *et al.* (2007), suggesting that *A. o. almogavarii* might have been a highly divergent lineage that progressively lost its genetic identity due to extensive hybridization and introgression with its neighboring

lineages. On the contrary, the high diversity, overall concordance across markers, with little allele sharing with other lineages, and geographical confinement of *A. o. almogavarii* suggests that it may represent an incipient species. However, potential changes in its specific taxonomic status must await further studies focusing on admixture patterns in the putative contact zones with both *A. o. obstetricans* and *A. o. pertinax* that have been revealed in this study, as well as in other as yet understudied areas, in order to assess the extent of reproductive isolation across lineages. While the range of *A. o. almogavarii* apparently remained stable for a long period of time within a restricted area, the other extant *A. obstetricans* subspecies may have expanded and contracted multiple times. Given the inferred timeframe for the split between lineage F and these lineages, it is likely that these demographic fluctuations were associated with the climatic oscillations that characterized the Pleistocene. Lineage D includes populations south of the Douro River in Portugal and along the western part of the Sistema Central mountains in central Iberia, previously assigned to subspecies *A. o. boscai*. This finding confirms previous results based on mtDNA and microsatellite data (Fonseca *et al.* 2003; Maia-Carvalho *et al.* 2014a), which highlighted the genetic distinctiveness of these populations. The mountains south of the Douro are well-known hotspots of genetic diversity and correspond to important glacial refugia where many different species have persisted across the Ice Ages (e.g. Alexandrino *et al.* 2000; Paulo *et al.* 2001; Martínez-Solano *et al.* 2006). Lineage C is restricted to north-western Iberia and includes the type locality of *A. o. boscai* (Tuy, in the province of Pontevedra, Spain; see García-París and Martínez-Solano 2001). The pattern of mtDNA haplotype diversity and significant values of Tajima's D and Ramos-Onsins & Rozas' R_2 statistics suggest a rapid and recent expansion, consistent with the results of the mismatch distribution analysis. Further research is needed to delineate the contact zone with lineage B, although the finding of haplotypes of both lineages (B + C) co-occurring in locality 22 (Arteixo, A Coruña – Table 1 and Fig. 1) is in accordance with previous descriptions of the geographical border between the two subspecies, which was located in the western slopes of the Cantabrian Mountains in Galicia, based on allozymes and coloration patterns (Arntzen and García-París 1995). The results from both mtDNA and nuclear genealogies show that lineage A corresponds to subspecies *A. o. pertinax* (GarcíaParís and Martínez-Solano 2001). However, its current range is much more extensive than previously considered, occupying, to the north, the southern margin of the Ebro River up to the Cantabrian Mountains and, to the west, the Spanish-Portuguese border. Both mtDNA and nuclear data suggest the occurrence of a contact zone between this subspecies and *A. o. almogavarii* in the northeast, with *A. o. obstetricans* in the north and north-west, and with *A. o. boscai* in the west (Sierra de Gata), respectively (Fig. 1). Thus, high-elevation habitats in the Estrela and Gredos mountains (western Central System) and Guadarrama (eastern Central System) seem to have been colonized independently by lineages corresponding to part of subspecies *A. o. boscai* (the

highly divergent lineage D) and subspecies *A. o. pertinax*, respectively. Similarly, the high-elevation mountain habitats in the north have been independently colonized by subspecies *A. o. obstetricans* (Cantabrian mountains) and *A. o. almogavarii* (Pyrenees). Since these populations are those most heavily affected by chytridiomycosis related mass mortality events, taking into account the phylogenetic affinities of extirpated and potential source populations with no record of massive die-offs will be critical to plan successful captive breeding and reintroduction. The susceptibility to disease (bacteria, chytrid fungi, ranaviruses) of high-elevation populations in four of the six major population lineages identified in our study (Márquez *et al.* 1995, Bosch *et al.* 2001, Walker *et al.* 2010, Rosa *et al.* 2012, Price *et al.* 2014) favors environmental correlates as explanatory variables in population die-offs and dismisses the role of potential historical factors derived from a shared evolutionary history. However, genetic bottlenecks associated with the colonization of high-elevation habitats probably also represent a contributing factor (Albert *et al.* 2014). Finally, lineage B (*A. o. obstetricans*) extends from Northern Spain to Central Europe. North of the Pyrenees, the large geographical area occupied by this lineage (comprising France, Luxembourg and parts of Germany, Netherlands, Switzerland and Belgium) is almost depleted of genetic variation, as shown by the reduced haplotypic diversity, suggesting a rapid and recent expansion from the southern slopes of the western Pyrenees, probably after the Last Glacial Maximum. In contrast, populations south of the Pyrenees show considerable levels of both mtDNA and nuclear diversity. Contrasting patterns of genetic diversity cannot be attributed to biases in our sampling, where populations north of the Pyrenees are under-represented, but sampling effects may explain lack of significant results in neutrality tests. This pattern of contrasting diversity south and north of the Pyrenees and the inferred colonization route resemble those described for other amphibian species, such as *Lissotriton helveticus* (Recuero and García-París 2011). Taken together with phylogenetic results, both genetic diversity and phylogeographical data provide compelling evidence that the Iberian Peninsula has served as a long-term refugium for *A. obstetricans*. In fact, our results indicate that ancestral *A. obstetricans* populations may have persisted and survived in several refugia within the Iberian Peninsula during climatic oscillations in the Pleistocene, in agreement with the ‘‘refugia-within-refugia” scenario (Gómez and Lunt 2007, Abellán and Svenning 2014). Our results also show that lineage F has had a long independent evolutionary history (see also Arntzen and García-París 1995), splitting probably early in the Lower Pleistocene from other population lineages in *A. obstetricans*, and it might constitute a cryptic species. More detailed analyses of contact zones are needed to assess patterns of reproductive isolation and to clarify the taxonomic status of this divergent lineage, but its recognition as an evolutionary significant unit for conservation purposes is warranted.

In conclusion, our results evidenced a complex pattern of lineage isolation and admixture in the common midwife toad, suggesting that a mosaic of habitats in heterogeneous landscapes could be of major importance for this species to persist through changing environmental conditions. In general, the biogeographic pattern here uncovered, derived from an intense diversification process spanning the late Tertiary and Quaternary periods, supports the model of refugia-within-refugia (Gómez and Lunt 2007), corroborating previous findings on a variety of Iberian amphibian species (Martínez-Solano *et al.* 2005, 2006; Sequeira *et al.* 2008, Gonçalves *et al.* 2009, Vences *et al.* 2013, Díaz-Rodríguez *et al.* 2015). Despite the influence of common environmental factors, the accumulated data have shown that temporal and spatial patterns of genetic diversity in this hotspot cannot be explained by any general model, being more consistent with idiosyncratic and organism-specific responses to driving factors of diversification. However, these idiosyncratic patterns call for meta-analyses based on hypothesis testing approaches that may clarify the importance of exogenous vs endogenous (species natural history) factors in shaping lineage and species diversity.

2.1.6 Acknowledgments

We thank Agentes Forestales de la Comunidad de Madrid, F. Álvarez, J. Álvarez, J.W. Arntzen, E. Ayllón, L. Arregui, A. Bermejo, M. Berroneau, D. Buckley, M.A. Carretero, F.A. Fernández-Álvarez, A. Fonseca, C. Grande, J. Gutiérrez, H. Iglesias, A. Loureiro, R. Márquez, L. Martínez-Solano, P. Mas, J.C. Monzó, J. Oliver, E. Pascual, P. Pavón, J.M. Quiñones, R. Pereira, E. Recuero, J. Rubio, G. Sánchez, V. Sancho, N. Sillero, C. Soares, J. Teixeira, G. Themudo, A. Valdeón and G. Velo-Antón for sharing samples and/or help during fieldwork. We also thank B. Álvarez and I. Rey (Colección de ADN, MNCN-CSIC) for loaning tissues under their care, and MJ Fernández Benéitez (IREC-CSIC-UCLM-JCCM) for her work as a lab technician. This work was supported through Project ‘‘Genomics and Evolutionary Biology” co-financed by North Portugal Regional Operational Programme 2007/2013 (ON. 2 – O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF); by FEDER funds through the Operational Programme for Competitiveness Factors – COMPETE and by National Funds through FCT – Foundation for Science and Technology PTDC/BIABEC/099915/2008 and FCOMP-01-0124-FEDER-008915 to HG; and by partial funds provided by Grants CGL2008-04271-C02-01/ BOS and CGL2011-28300 (Ministerio de Ciencia e Innovación, Ministerio de Economía y Competitividad, Spain, and FEDER) and PPII10-0097-4200 (Junta de Comunidades de Castilla la Mancha and FEDER) to IMS. IMS was funded by Project ‘‘Biodiversity, Ecology and

Global Change", co-financed by ON. 2–O Novo Norte, under the NSRF, through the ERDF, and is currently supported by funding from the Spanish Severo Ochoa Program (SEV-20120262). HG was supported by a postdoctoral Grant from FCT (SFRH/BPD/26555/2006) and by funding from the Projecto Incentivo (INCENTIVO/BIA/LA0027/2013), and is currently supported by a postdoctoral Grant from FCT (SFRH/BPD/102966/2014). BMC was supported by a PhD Grant from FCT (SFRH/BD/60305/2009). FS is supported by a postdoctoral Grant from FCT (SFRH/BPD/87721/2012) and TSN by a research Grant (BI) from ICETA/UP.

2.1.7 References

Abellán P, Svenning JC (2014) Refugia within refugia – patterns in endemism and genetic divergence are linked to Late Quaternary climate stability in the Iberian Peninsula. *Biological Journal of the Linnean Society*, 113: 13–28.

Akaike H (1973) Information theory as an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory. Akademiai Kiado, Budapest.

Albert EM, Fernández-Beaskoetxea S, Godoy JA, Tobler U, Schmidt BR, Bosch J (2014) Genetic management of an amphibian population after a chytridiomycosis outbreak. *Conservation Genetics*, 16: 103–111.

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.

Arévalo, E., Davis, S.K., Sites Jr., J.W., 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.

Arntzen JW, Szymura JM (1984) Genetic differentiation between African and European midwife toads (*Alytes*, Discoglossidae). *Bijdragen tot de Dierkunde*, 54: 157–162.

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 (2000) Phylogeography: the history and formation of species. Cambridge, Massachusetts, Harvard University Press.

Avise JC (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, 36: 3–15.

Bosch J, Martínez-Solano I, García-París M (2001) Evidence of a chytrid fungus infection involved in the decline of the common midwife toad (*Alytes obstetricans*) in protected areas of central Spain. *Biological Conservation*, 97: 331–337.

Brunes T, Sequeira F, Haddad CFB, Alexandrino J (2010) Gene and species trees of a Neotropical group of treefrogs: genetic diversification in the Brazilian Atlantic Forest and the origin of a polyploid species. *Molecular Phylogenetics and Evolution*, 57: 1120-1133.

Crawford AJ (2003) Huge populations and old species of Costa Rican and Panamanian dirt frogs inferred from mitochondrial and nuclear gene sequences. *Molecular Ecology*, 12: 2525-2540.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9: 772.

Díaz-Rodríguez J, Gonçalves H, Sequeira F, Sousa-Neves T, Tejedo M, Ferrand N, Martínez-Solano I (2015) Molecular evidence for cryptic candidate species in Iberian *Pelodytes* (Anura, Pelodytidae). *Molecular Phylogenetics and Evolution*, 83: 224-241.

Drummond AJ, Rambaut A (2007) *BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7: 214.

Dufresnes C, Wassef J, Ghali K, Brelsford A, Stöck M, Lymberakis P, Crnobrnja-Isailović J, Perrin N (2013) Conservation phylogeography: does historical diversity contribute to regional vulnerability in European tree frogs (*Hyla arborea*)? *Molecular Ecology*, 22: 5669–5684.

Flot JF (2010) SeqPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. *Molecular Ecology Resources*, 10: 162-166.

Fonseca A, Arntzen JW, Crespo EG, Ferrand N (2003) Regional differentiation in the common midwife toad (*Alytes obstetricans*) in Portugal: a picture from mitochondrial DNA. *Zeitschrift für Feldherpetologie*, 10: 83–89.

Fordham DA, Brook BW, Moritz C, Nogués-Bravo D (2014) Better forecasts of range dynamics using genetic data. *Trends in Ecology & Evolution*, 29: 436–443. Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147: 915-925.

García-París M (1995) Variabilidad genética y distribución geográfica de *Alytes obstetricans almogavarii* en España. *Revista Española de Herpetología*, 9: 133–138.

García-París M, Martínez-Solano I (2001) Nuevo estatus taxonómico para las poblaciones iberomediterráneas de *Alytes obstetricans* (Anura: Discoglossidae). *Revista Española de Herpetología*, 15: 99–113.

Geniez P, Crochet P-A (2003) Confirmation de l'existence, en France, de trois taxons méconnus: *Alytes obstetricans almogavarii* Arntzen & García-París, 1995 (Amphibia,

Discoglossidae), *Podarcis hispanica sebastiani* (Klemmer, 1964) (Reptilia, Lacertidae) et *Natrix natrix astreptophora* (Seoane, 1884) (Reptilia, Colubridae). *Bulletin de la Société Herpétologique de France*, 105: 41-53.

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

Gonçalves H, Martínez-Solano I, Ferrand N, García-París M (2007) Conflicting phylogenetic signal of nuclear *vs* mitochondrial DNA markers in midwife toads (Anura, Discoglossidae, *Alytes*): deep coalescence or ancestral hybridization? *Molecular Phylogenetics and Evolution*, 44: 494-500.

Gonçalves H, Martínez-Solano I, Pereira RJ, Carvalho B, García-París M, Ferrand N (2009) High levels of population subdivision in a morphologically conserved Mediterranean toad (*Alytes cisternasii*) result from recent, multiple refugia: evidence from mtDNA, microsatellites and nuclear genealogies. *Molecular Ecology*, 18: 5143–5160.

Grossenbacher K (1997) *Alytes obstetricans* (Laurenti, 1768). Atlas of amphibians and reptiles in Europe (eds. J-P Gasc *et al.*), pp. 94–95. Societas Europaea Herpetologica, Paris.

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.

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, 8: 461–467.

Heled J, Drummond AJ (2008) Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology,* 8: 289.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, 27: 570–580.

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

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.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B, Biological Sciences*, 359: 183–195.

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17: 754–755.

Jockusch, E.L., Martínez-Solano, Í., Timpe, E.K., 2015. The effects of inference method, population sampling, and gene sampling on species tree inferences: an empirical study in slender salamanders (Plethodontidae: Batrachoseps). *Systematic Biology*, 64: 66–83.

Lanfear R, Calcott B, Ho SYW, Guindon S (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29: 1695-1701.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of polymorphism DNA data. *Bioinformatics*, 25: 1451-1452.

Maia-Carvalho B, Gonçalves H, Martínez-Solano I, Gutiérrez-Rodríguez J, Lopes S, Ferrand N, Sequeira F (2014a) Intraspecific genetic variation in the common midwife toad (*Alytes obstetricans*): subspecies assignment using mitochondrial and microsatellite markers. *Journal of Zoological Systematics and Evolutionary Research*, 52: 170-175.

Maia-Carvalho B, Gonçalves H, Ferrand N, Martínez-Solano I (2014b) Multilocus assessment of phylogenetic relationships in *Alytes* (Anura, Alytidae). *Molecular Phylogenetics and Evolution*, 79: 270-278.

Márquez R, Olmo JL, Bosch J (1995) Recurrent mass mortality of larval midwife toads *Alytes obstetricans* in a lake in the Pyrenean mountains. *Herpetological Journal*, 5: 287-289.

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, Alcobendas M, Buckley D, García-París, M (2005) Molecular characterisation of the endangered *Salamandra salamandra almanzoris* (Caudata, Salamandridae). *Annales Zoologici Fennici,* 42: 57–68.

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

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), 14 November 2010, New Orleans, LA, pp. 1–8.

Milne I, Wrigh F, Rowe G, Marshall DF, Husmeier D, McGuire G (2004) TOPALi: software for automatic identification of recombinant sequences within DNA multiple alignments. *Bioinformatics*, 20: 1806-1807.

Moritz C, Agudo R (2013) The Future of Species Under Climate Change: Resilience or Decline? *Science*, 341: 504–508.

Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): a system 803 for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*, 24: 581–583.

Pabijan, M., Wandycz, A., Hofman, S.,Wecek, K., Piwczyn´ ski, M., Szymura, J.M., 2013. Complete mitochondrial genomes resolve phylogenetic relationships within Bombina (Anura: Bombinatoridae). *Molecular Phylogenetic and Evolution*, 69: 63–74.

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. P*roceedings of the Royal Society of London B, Biological Sciences*, 268: 1625–1630.

Pinho C, Rocha S, Carvalho BM, Lopes S, Mourão S, Vallinoto M, Brunes TO, Haddad CFB, Gonçalves H, Sequeira F (2010). New primers for the amplification and sequencing of nuclear loci in a taxonomically wide set of reptiles and amphibians. *Conservation Genetics Resources*, 2: 181-185.

Price SJ, Garner TWJ, Nichols RA, Balloux F, Ayres C, de Alba AM-C, Bosch J (2014) Collapse of amphibian communities due to an introduced Ranavirus. *Curbio*, 24: 2586–2591.

Rambaut A, Suchard MA, Xie D, Drummond AJ (2014) Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, 19: 2092-2100.

Recuero E, García-París M (2011) Evolutionary history of *Lissotriton helveticus*: Multilocus assessment of ancestral vs. recent colonization of the Iberian Peninsula. *Molecular Phylogenetics and Evolution,* 60: 170–182.

Rogers AR, Harpending HC (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, 9: 552–569.

Rosa GM, Anza I, Moreira PL, Conde J, Martins F, Fisher MC, Bosch J (2013). Evidence of chytrid mediated population declines in common midwife toad in Serra da Estrela, Portugal. *Animal Conservation*, 16: 306–315.

Salzburger W, Ewing GB, Von Haeseler A (2011) The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Molecular Ecology*, 20: 1952-1963.

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

San Mauro D, García-París M, Zardoya R (2004) Phylogenetic relationships of discoglossid frogs (Amphibia: Anura: Discoglossidae) based on complete mitochondrial genomes and nuclear genes. *Gene*, 343: 357–366.

Schoville SD, Bonin A, Francois O, Lobreaux S, Melodelima C, Manel S (2012) Adaptive Genetic Variation on the Landscape: Methods and Cases. *Annual Review of Ecology, Evolution, and Systematics*, 43: 23–43.

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

Sequeira F, Alexandrino J, Weiss S, Ferrand N (2008) Documenting the advantages and limitations of different classes of molecular markers in a well-established phylogeographic context: lessons from the Iberian endemic Golden-striped salamander, *Chioglossa lusitanica* (Caudata: Salamandridae). *Biological Journal of the Linnean Society*, 95: 371–387.

Silvestro D, Michalak I (2010) RAxML GUI: A Graphical Front-end for RAML. <http://sourceforge.net/projects/raxmlgui/>.

Sousa-Neves T, Aleixo A, Sequeira F (2013). Cryptic patterns of diversification of a widespread Amazonian Woodcreeper species complex (Aves: Dendrocolaptidae) inferred from multilocus phylogenetic analysis: Implications for historical biogeography and taxonomy. *Molecular Phylogenetics and Evolution*, 68: 410-424.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixedmodels. *Bioinformatics*, 22: 2688–2690.

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

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, 73: 1162-1169.

Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B*, 277: 661–671.

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

Tajima F (1989) Statistical-Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics*, 123: 585-595.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*, 28: 2731-2739.

Velo-Antón G, Parra JL, Parra-Olea G, Zamudio KR (2013) Tracking climate change in a dispersal-limited species: reduced spatial and genetic connectivity in a montane salamander. *Molecular Ecology*, 22: 3261–3278. doi:10.1111/mec.12310

Vences M, Hauswaldt JS, Steinfartz S, Rupp O, Goesmann A, Künzel S, Orozco-terWengel P, Vieites DR, Nieto-Roman S, Haas S, Laugsch C, Gehara M, Bruchmann S, Pabijan M, Ludewig A-K, Rudert D, Angelini C, Borkin LJ, Crochet P-A, Crottini A, Dubois A, Ficetola GF, Galán P, Geniez P, Hachtel M, Jovanovic O, Litvinchuk SN, Lymberakis P, Ohler A, Smirnov NA (2013): Radically different phylogeographies and patterns of genetic variation in two European brown frogs, genus *Rana*. *Molecular Phylogenetics and Evolution*, 68: 657-670.

Walker SF, Bosch J, Gomez V, Garner TWJ, Cunningham AA, Schmeller DS, Ninyerola M, Henk DA, Ginestet C, Arthur C-P, Fisher MC (2010) Factors driving pathogenicity vs. prevalence of amphibian panzootic chytridiomycosis in Iberia. *Ecology Letters*, 13: 372–382. doi:10.1111/j.1461-0248.2009.01434.x

Capítulo 4

Fatores ambientais e a diferenciação de linhagens em *Alytes obstetricans*

Artigo III

Maia-Carvalho B, Gonçalves H, Martínez-Solano I, Lopes S, Gutiérrez-Rodríguez J, Ferrand N and Sequeira F (2014) Intraspecific genetic variation in the common midwife toad (*Alytes obstetricans*): subspecies assignment using mitochondrial and microsatellite markers. *Journal of Zoological Systematics and Evolutionary Research* 52, 170 – 175. doi: 10.1111/jzs.12048

Artigo IV

Maia-Carvalho B, Vale CG, Sequeira F, Martínez-Solano I, Ferrand N, Gonçalves H (*submitted*) When the environment favours lineage diversification: ecological niche divergence mirrors intraspecific patterns of genetic diversity in the Common Midwife Toad (*Alytes obstetricans*).

Artigo III

Short Communication
Accepted 27 September 2013

Intraspecific genetic variation in the common midwife toad (Alytes obstetricans): subspecies assignment using mitochondrial and microsatellite markers

Bruno Maia-Carvalho[1,2], Helena Gonçalves[1], Iñigo Martínez-Solano[1,3], Jorge Gutiérrez-Rodríguez[4], Susana Lopes[1], Nuno Ferrand[1,2], and Fernando Sequeira[1]

Journal of Zoological Systematics and Evolutionary Research (2014) 52, 170 - 175
doi: 10.1111/jzs.12048

[1] CIBIO/UP, Centro de Investigação em Biodiversidade e Recursos Genéticos da Universidade do Porto, InBIO, Vairão, Portugal;
[2] Departamento de Zoologia e Antropologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal;
[3] nstituto de Investigación en Recursos Cinegéticos (CSIC-UCLM-JCCM), Ciudad Real, Spain;
[4] Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain

4.1 Intraspecific genetic variation in the common midwife toad (*Alytes obstetricans*): subspecies assignment using mitochondrial and microsatellite markers

4.1.1 Abstract

The common midwife toad (*Alytes obstetricans*), widely distributed in the northern half of the Iberian Peninsula and part of Western Europe, is currently subdivided into four subspecies: *A. o. obstetricans*, *A. o. boscai*, *A. o. pertinax* and *A. o. almogavarii*. However, the delimitation of these subspecies and their ranges are still under discussion because strong discordances have been found between morphological and molecular data, and especially among different genetic markers. Here, we screen a set of novel microsatellite loci and mtDNA sequences of *A. obstetricans* populations representative of all currently recognized subspecies to investigate the correspondence between genetic groupings inferred from clustering analysis of microsatellite genotypes and the described subspecies and test whether patterns of mtDNA variation are concordant with those genetic clusters. Our results confirm previous expectations of extremely high intraspecific diversity in *A. obstetricans* in Iberia. Analyses of microsatellite and mtDNA data were concordant in recovering five well-defined groups, of which three correspond to previously defined subspecies, while the two additional clusters correspond to populations of subspecies *A. o. boscai* separated by the Douro River. Our results suggest the occurrence of two distinct genetic units within *A. o. boscai* that likely result from a long independent evolutionary history, thus deserving special attention from a conservation point of view.

Key words: STRs – mtDNA – amphibians – Iberian Peninsula – Alytes.

4.1.2 Introduction

Subspecies are diagnosable units that correspond to discontinuities in the geographical distribution of some morphological and/or genetic characteristics within a species. Whereas the evolutionary meaning and taxonomic importance of subspecies are still highly controversial topics (e.g. Hawlitschek *et al.* 2012), some authors argue that they may represent useful units for conservation because the maintenance of genetic diversity is crucial for population viability and the adaptive potential of species (e.g. Phillimore and Owens 2006, Allentoft and O'Brien

2010). This seems to be particularly relevant in the case of species that present a high degree of population genetic differentiation and restricted distribution ranges, often associated with very low effective population sizes. In many cases, these differentiated populations represent geographically structured, long-term diverging evolutionary units that are maintained due to reduced dispersal capabilities and/or local adaptation promoted by different ecological requirements (Köhler *et al.* 2005, Martínez-Solano *et al.* 2006, Gómez and Lunt 2007). Nonetheless, one of the major problems with the acceptance and delimitation of subspecific categories stems from the finding that in many cases, inferred subspecies boundaries are not congruent among different genetic markers and between these and morphological traits (e.g. Arntzen *et al.* 2007, Glor and Laport 2012).

One such case of discordance between the geographical distribution of genetic and morphological characteristics and their relationship with described subspecies is the common midwife toad, *Alytes obstetricans* (Laurenti 1768). This species, widely distributed in Western Europe, ranging from Germany to the northern half of the Iberian Peninsula (Grossenbacher 1997), is currently subdivided into four subspecies (Fig. 1a): *A. o. obstetricans* distributed from northern Iberia to Germany, *A. o. boscai* distributed in north-western and central Iberia, *A. o. pertinax* distributed in central and eastern Iberia, and *A. o. almogavarii* distributed in north-eastern Iberia (García-París 1995, García-París and Martínez-Solano 2001). However, the delimitation of these subspecies and their ranges is still under discussion because strong discordances have been found between morphological and molecular data, and especially among different genetic markers, including allozymes and mtDNA (see Arntzen and García-París 1995, Fonseca *et al.* 2003, Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007).

The hypervariability and simple inheritance mode of microsatellites provide a powerful tool for population genetic studies and coupled to Bayesian model-based clustering procedures (Pritchard *et al.* 2000) they have been widely used to assign individuals to subspecies/populations or to estimate the most likely ancestral composition of admixed individuals, especially when the phenotypic differentiation between the subspecies/populations in question is weak. In this study, we used 14 microsatellite loci (here reported and characterized for the first time) and extended mtDNA sequencing analysis (*ND4* gene) to assess intraspecific levels of genetic variation over the *A. obstetricans* range. Using a set of *A. obstetricans* populations representative of all currently recognized subspecies, we aimed to contrast patterns of mtDNA variation and genetic groupings as revealed by clustering analysis of microsatellite markers in the light of previously inferred subspecies delimitation.

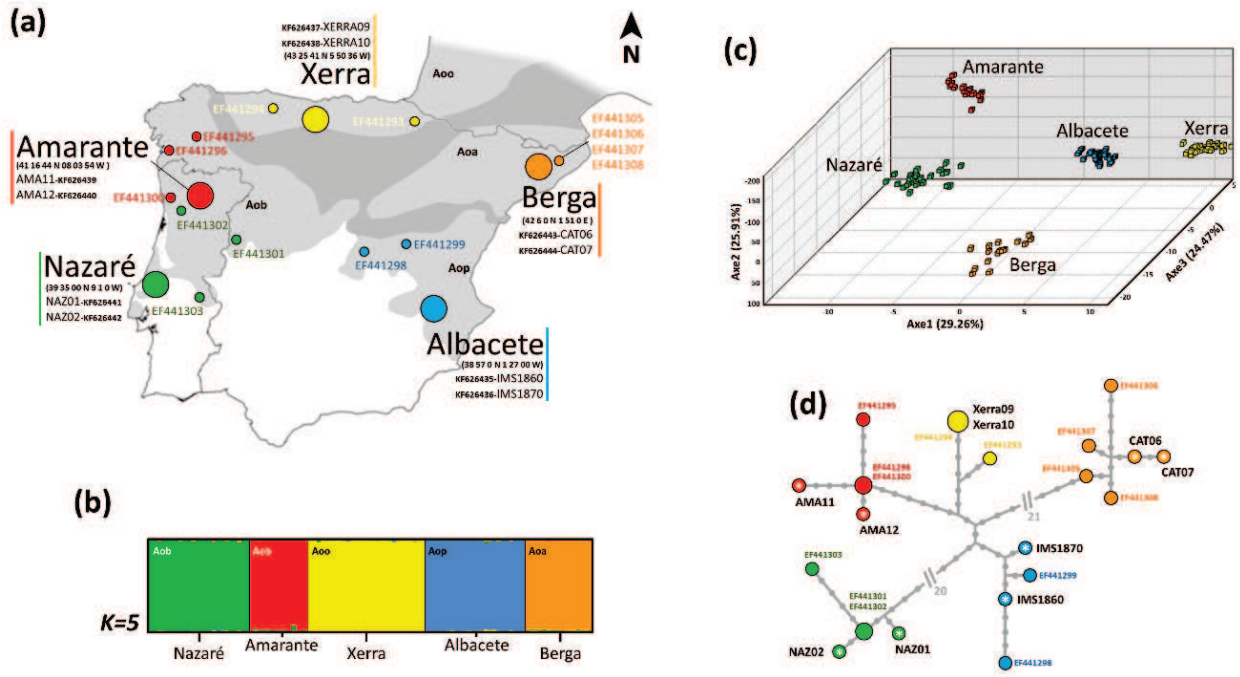


Fig. 1. (a) Distribution map of *Alytes obstetricans* subspecies (García-París and Martínez-Solano 2001) in the Iberian Peninsula (shaded) with indication of the geographical origin of the samples analysed in this study (*A. o. boscai* – Aob, *A. o. obstetricans* – Aoo, *A. o. pertinax* – Aop and *A. o. almogavarii* – Aoa). Dark grey area corresponds to zones where the assignment to the different *Alytes* subspecies is doubtful. Big coloured circles represent the populations genotyped for microsatellites and additionally sequenced for mtDNA analysis, with indication of the respective voucher number of each sample. Small coloured circles represent the sample localization of *ND4* sequences (EF numbers, GenBank) from Gonçalves *et al.* (2007). (b) Bayesian clustering results of STRUCTURE analysis for microsatellites (*K* = 5). Each individual is represented as a vertical line partitioned into *K* coloured segments, whose length is proportional to the individual's estimated membership coefficient. A black line separates individuals of different populations. (c) Factorial correspondence analysis (FCA) based on microsatellite genotyping of 106 individuals from five *A. obstetricans* populations. (d) Median-joining network for *ND4* haplotypes observed in *Alytes obstetricans* showing the five haplogroups recovered. Each circle represents a different haplotype with size proportional to its relative frequency (with indication of the respective GenBank accession number). Asterisks indicate the new haplotypes described in this study. Grey dots along branches represent hypothetical unsampled haplotypes.

4.1.3 Material and methods

The microsatellite loci were developed from a partial enriched genomic library prepared from one individual of *A. obstetricans pertinax* (voucher no. IMS1239) collected in Villanueva de Alcorón, Guadalajara, Spain (N 40.67°; W 2.25°). We extracted genomic DNA from toe or tail clips using EasySpin Genomic DNA Minipreps Tissue Kit (SP-DT-250, Qiagen, Hilden, Germany) following the fabricant protocol. The same methodology was applied for all 106 samples used for genotyping. The microsatellite library was constructed at the Evolutionary Genetics Core Facility (EGCF) from Cornell University Life Sciences Core Laboratories Center (CLC), following the protocol described by Andrés and Bogdanowicz (2011). Briefly, genomic DNA was completely digested with a restriction enzyme (five-base cutter). Linkers were ligated

to the digested DNA, and the resulting fragments were enriched for microsatellites by hybridization and magnetic capture of biotinylated repeat probes of two dimers, five trimers and four tetramers. Enriched genomic fragments were amplified by PCR, ligated to Roche/454 Titanium Multiplex Identifier (MID) adapters and size separated in an agarose gel. Sequences were generated on a Roche/454 sequencer (Titanium chemistry and adapters).

We designed one hundred primer pairs with PRIMER3PLUS software (http://bioinfo.ebc.ee/mprimer3/). Resulting primers were tested for potential interactions with each other, including primer–dimer and intramolecular hairpin formation using AUTODIMER software (Vallone and Butler 2004). After initial tests for PCR amplification, we discarded 66 loci due to unsatisfactory amplification or PCR products > 500 base pairs. From the remaining 34 loci, we selected a set of 17 loci that exhibited clear and reliable amplification for most of the target subspecies at population level. A total of 106 individuals from five populations representative of all currently recognized *A. obstetricans* subspecies (Fig. 1a) were genotyped using those 17 loci in five multiplex reactions and the Qiagen Multiplex PCR Master Mix 2X (Appendix 1). For each locus, the forward primer was 5'-labelled with a fluorescent dye (VIC, PET, 6-FAM or NED, Appendix 1). Multiplex PCR amplifications were performed in a 10 μL reaction volume containing 5 μL of the same Master Mix, 0.5 μL primer mix and 1.0 μL (50 ng) genomic DNA. Touchdown PCR programmes were carried out as follows: 95°C for 15 min; 1st round (11 cycles) of 95°C for 30 s, 61°C to 56°C (for Plex 1 to Plex 3) and 56°C to 51°C (for Plex 4 and Plex 5) decreasing 0.5°C in each cycle for 1 min and 72°C for 35 s; 2nd round (20 cycles) of 95°C for 30 s, 56°C (for Plex 1 to Plex 3) and 51°C (for Plex 4 and Plex 5) for 40 s and 72°C for 35 s; 3rd round (eight cycles) of 95°C for 30 s, 53°C for 30 s and 72°C for 35 s, with a final extension at 60°C for 30 min.

We used 1 μL of PCR diluted product in combination with 10 μL of deionized formamide and 0.2 μL of internal size standard (Genescan-500 LIZ, ABI, Foster City, California, USA) for genotyping. Fragment size was determined on an ABI Prism 3130XL capillary sequencer. Fragments were scored and binned using GENEMAPPER v4.0 software (Applied Biosystems, Foster City, California, USA). To determine possible genotyping errors, null alleles and allelic dropout, we used MICROCHECKER v2.2.3 (Van Oosterhout *et al.* 2004). Tests for Hardy–Weinberg equilibrium (HWE), linkage disequilibrium (LD) and standard genetic diversity measures, such as the mean of allele numbers, observed (Ho) and expected (He) heterozygosities, allele frequencies and allelic richness were estimated using FSTAT v2.9.3.2 (Goudet 2001). Critical probability for each test was adjusted with a sequential Bonferroni correction (Rice 1989). Population genetic differentiation was analysed using Fisher's exact test and by estimating F_{st} (Weir and Cockerham 1984) in FSTAT. Population structure was investigated with a factorial correspondence analysis (FCA), as implemented in GENETIX

v4.05.2 (Belkhir *et al.* 2000), and also using the Bayesian multilocus clustering analysis implemented in STRUCTURE v2.3.1 (Pritchard *et al.* 2000). To determine the number of clusters (*K*), ten independent simulations for *K* = 1–10 were carried out with 100,000 burn-in iterations and 500,000 data iterations. Analyses were performed using the admixture model of population structure with allele frequencies correlated among populations. We selected the optimal *K* value based on the rate of change in the log probability of data between successive *K* values Δ*K* (Evanno *et al.* 2005), as implemented in STRUCTURE HARVESTER (Earl and vonHoldt 2012). We combined values from different run replicates under the optimal K value with CLUMPP v1.1.2 (Jakobsson and Rosenberg 2007).

A mitochondrial fragment (630 bp) of the NADH dehydrogenase subunit four gene (*ND4*) and adjacent tRNAs was amplified using the primers *ND4* and *Leu* previously described by Arévalo *et al.* (1994). Two samples of each population (Fig. 1a) were sequenced. PCR was performed with Phusion Master Mix (Thermo Scientific, Waltham, Massachusetts, USA), and products of each reaction were sequenced following the ABI Prism BigDye Terminator v3.1 Sequencing Kit protocol on an ABI3130xl DNA Analyzer (Applied Biosystems, Foster City, California, USA) as in Gonçalves *et al.* (2007). Sequences were edited and aligned using the program BIOEDIT v7.1.3.0 (Hall 1999). We added 14 published *ND4* sequences from GenBank (Gonçalves *et al.* 2007) to the 10 sequences obtained (Fig 1a). A median-joining network was constructed using NETWORK v4.6.1.0 (http://www.fluxus-engineering.com). Pairwise population *p*-uncorrected genetic distances were calculated with MEGA v5.1 (Tamura *et al.* 2011). All sequences generated in this study are deposited in GenBank under accession numbers KF626401-KF626434.

4.1.4 Results and discussion

From the initially selected 17 microsatellite loci, only 14 were considered for the analyses of the 106 individuals, because three loci amplified inconsistently in some subspecies (Table 1). All loci were found to be in HWE and linkage equilibrium after applying the Bonferroni correction for multiple tests. Evidences for genotyping errors, null alleles or allelic dropout were not found. The number of alleles per locus varied between three and 20. Values of allelic richness and expected heterozygosity were relatively high for all populations, ranging from 4.61 and 0.65 to 7.44 and 0.78, respectively (Table 1). All populations were significantly differentiated from each other (Fisher's exact test; $p < 0.001$). For all populations, the pairwise F_{st} values were highly significant ($p < 0.001$), ranging from 0.174 to 0.305 (Table 2). Bayesian clustering analysis using STRUCTURE indicated that the most likely value for *K* was five (Fig.

1b). Three clusters correspond to previously defined subspecies (*A. o. obstetricans*, *A. o. pertinax* and *A. o. almogavarii*), while the two additional clusters correspond to the populations Amarante and Nazaré (subspecies *A. o. boscai*, Fig. 1b). For all individuals, assignment probability to its respective cluster was > 95%. Results from FCA were concordant with Bayesian clustering analysis showing the same five clusters, corresponding to the *a priori* defined populations (Fig. 1c).

Table 1. Polymorphism statistics of 17 microsatellite markers in five populations representative of all recognized *Alytes obstetricans* subspecies

Locus	Populations				
	Nazaré (Aob)	Amarante (Aob)	Xerra (Aoo)	Albacete (Aop)	Berga (Aoa)
Aobstμ01					
N	24	14	28	24	16
n	14(7)	5(1)	5	15(10)	12(4)
R	9.82	4.29	3.89	9.74	9.28
Aobstμ02					
N	24	14	28	24	16
n	13(5)	9(1)	6(1)	11(5)	9(6)
R	8.88	7.66	5.19	7.68	6.01
Aobstμ03					
N	0	14	28	24	12
n	0	3(3)	4(1)	5(2)	6(4)
R	0	2.86	3.33	4.38	6.00
Aobstμ04					
N	24	14	28	23	0
n	6(3)	7(1)	5(1)	4	0
R	5.65	7.00	4.82	4.00	0
Aobstμ05					
N	0	0	28.00	24.00	16.00
n	0	0	8(5)	12(10)	3(2)
R	0	0	7.65	10.83	3.00
Aobstμ06					
N	22	12	28	24	16
n	19(9)	5(1)	8(2)	10(1)	19(9)
R	11.28	4.49	6.27	7.64	12.91
Aobstμ07					
N	22.00	12.00	28.00	24.00	16.00
n	13(11)	14(5)	6	11(1)	13(4)
R	9.45	11.63	5.31	8.30	10.91
Aobstμ08					
N	24	12	28	24	13
n	4(3)	4	11(3)	10(1)	13(5)
R	2.91	4.00	8.00	7.74	11.00

Aobstμ09					
N	23.00	14.00	28.00	24.00	16.00
n	3(2)	5(1)	4(1)	4(1)	5(2)
R	2.78	4.88	3.64	3.62	4.69
Aobstμ10					
N	24	14	28	24	16
n	5(1)	2(1)	2	5(1)	2
R	4.52	2.00	2.00	4.83	2.00
Aobstμ11					
N	24	14	28	24	16
n	11(2)	9(1)	7(1)	8(1)	13(1)
R	8.48	7.62	5.88	6.35	9.87
Aobstμ12					
N	24	14	28	24	16
n	3(1)	1	2	4	5(2)
R	2.36	1.00	2.00	3.32	3.94
Aobstμ13					
N	24	14	28	24	16
n	10(6)	3	3	4(1)	3(2)
R	7.32	2.64	2.99	3.53	2.13
Aobstμ14					
N	24	14	28	24	16
n	7(4)	5(5)	4	8(3)	9(5)
R	5.37	4.94	3.50	6.37	8.04
Aobstμ15					
N	19	9	28	23	16
n	14(4)	6(3)	9(6)	12(8)	20(8)
R	10.01	6.00	6.43	8.21	13.66
Aobstμ16					
N	24	14	28	24	16
n	15(6)	14(5)	5(1)	11(2)	12(5)
R	10.12	11.00	4.25	8.14	9.35
Aobstμ17					
N	22	14	27	24	15
n	8(4)	3	4(2)	8(3)	7(2)
R	5.70	2.96	3.15	6.34	6.22
AR	6.98	5.31	4.61	6.53	7.44
A	9.67	5.94	5.47	8.35	9.44
Ho	0.55	0.54	0.69	0.74	0.54
He	0.66	0.65	0.66	0.78	0.69

N, sample size; *n*, number of alleles detected (private alleles are indicated in parentheses); *R*, allelic richness per locus; *AR*, allelic richness per population; *A*, mean number of alleles. *Ho* and *He* are the average observed and expected heterozygosity, respectively, across all loci. *, loci that were not consistently amplified in all currently recognized subspecies: Aob – *Alytes obstetricans boscai*, Aoo – *Alytes obstetricans obstetricans*, Aop – *Alytes obstetricans pertinax*, Aoa – *Alytes obstetricans almogavarii* (García-París and Martínez-Solano 2001).

The mitochondrial DNA alignment (*ND4*) consisted of 24 sequences of 630 bp yielding 20 haplotypes. Eight haplotypes are newly described, whereas all other haplotypes were already reported in Gonçalves *et al.* (2007). The median-joining network for *ND4* haplotypes is shown in Fig. 1d. The haplotypes were clustered into five well-differentiated haplogroups, which are in full agreement with the microsatellite-defined clusters. The average number of sequence pairwise differences (*p*-uncorrected distance) between haplogroups was relatively high (Table 2), ranging from 1.3% between Albacete (*A. o. pertinax*) and Xerra (*A. o. obstetricans*) to 5.7% between Berga (*A. o. almogavarii*) and Nazaré (*A. o. boscai*).

Table 2. Average number of pairwise sequence differences (*p*-uncorrected distance, above diagonal) and microsatellite pairwise F_{st} estimates (below diagonal) for the five populations of *Alytes obstetricans* studied.

	Nazaré (Aob)	Amarante (Aob)	Xerra (Aoo)	Albacete (Aop)	Berga (Aoa)
Nazaré(Aob)	-	0.052	0.046	0.044	0.057
Amarante(Aob)	0.19499	-	0.019	0.020	0.052
Xerra(Aoo)	0.27935	0.30463	-	0.013	0.050
Albacete(Aop)	0.18514	0.17374	0.19747	-	0.050
Berga(Aoa)	0.24914	0.27094	0.25271	0.17656	-

Our results clearly suggest that microsatellites developed herein are useful markers for subspecies assignment and population substructure analysis. Indeed, all microsatellite analyses confirm previous expectations of high intraspecific genetic diversity in *A. obstetricans* in Iberia and are fully concordant with mtDNA in identifying groups corresponding to the four currently recognized subspecies of *A. obstetricans*. Although applied on a limited number of populations, our microsatellite results do not agree with the previously inferred northern boundary between *A. o. boscai* and *A. o. obstetricans* based on allozyme data (Pyrenees; Arntzen and Szymura 1984), suggesting that this is likely placed somewhere between northern Portugal and the Cantabrian Mountains.

Microsatellite data were also congruent with mtDNA results in revealing the existence of two well-differentiated groups within *A. o. boscai*, separated by the Douro River. The pronounced genetic differentiation at the nuclear level seems to provide additional support for the previously hypothesized existence of two lineages within *A. o. boscai* based on mtDNA analysis (Fonseca *et al.* 2003). The level of genetic differentiation between northern and southern populations of *A. o. boscai* as measured by F_{st} and sequence divergence (*p*-uncorrected distance) was similar or, in some cases, higher than the level of differentiation between the other subspecies of *A. obstetricans* (Table 2). Remarkably, sequence divergence between southern and northern populations of *A. o. boscai* and the other subspecies is within the same order of magnitude of divergence estimates between *A. o. almogavarii*, the most

differentiated taxon within *A. obstetricans*, and the other subspecies, which may have diverged from each other more than 5 million years ago (Martínez-Solano *et al.* 2004). Multicharacter analyses involving molecular as well as morphological and bioacoustical data based on a more comprehensive sampling scheme should be undertaken before making taxonomic considerations, but our results suggest the occurrence of two genetic evolutionary units within *A. o. boscai* that likely resulted from a long independent evolutionary history, thus deserving a special attention from a conservation point of view.

Although our findings gave new insights into the patterns of population subdivision and raise some new taxonomic questions, many aspects of the biogeographic history of *A. obstetricans* remain unresolved, including the identification of putative contact zones and the geographical delimitation of the subspecies, which only could be conveniently addressed with a more comprehensive sampling scheme. Nonetheless, this study constitutes a valuable basis for stimulating further detailed investigations. In particular, these novel microsatellites, together with the 12 recently described loci for *A. o. obstetricans* populations from Central Europe (Tobler *et al.* 2013), have great potential for addressing questions related to fine-scale population structure, estimation of effective population size and parentage analysis and will thus help conservation efforts, including captive breeding programs, already underway in some areas (European Commission 2012).

4.1.5 Acknowledgements

We thank Enrique Ayllón, Alfredo Nicieza and Leticia Viesca for providing some of the samples used here. We further thank Steve Bogdanowicz (Evolutionary Genetics Core Facility, Cornell University Life Sciences Core Laboratories Center) for microsatellite library construction. This work was financed through the Programa Operacional Factores de Competitividade (COMPETE) and by national funds from Fundaçãoo para a Ciência e a Tecnologia (FCT), through the research projects PTDC/BIABEC/105083/2008 and PTDC/BIA-BEC/099915/2008 to FS and HG and grants CGL2008-04271-C02-01/BOS and CGL2011-28300 (Ministerio de Ciencia e Innovación, Ministerio de Economía y Competitividad, Spain, and FEDER) and PPII10-0097- 4200 (Junta de Comunidades de Castilla la Mancha) to IMS. FS and HG are supported by a postdoctoral grant from FCT (SFRH/BPD/87721/2012 and SFRH/BPD/26555/2006, respectively), and BMC is supported by a PhD grant from FCT (SFRH/BD/ 60305/2009). IMS was supported by a 'Ramón y Cajal' postdoctoral fellowship funded by the Spanish Ministerio de Ciencia e Innovación and the Universidad de Castilla la Mancha and is currently funded by Project "Biodiversity, Ecology and Global Change" cofinanced by North Portugal Regional Operational Programme 2007/2013 (ON.2–O Novo

Norte), under the National Strategic ReferenceFramework (NSRF), through the European Regional Development Fund (ERDF), and JGR is supported by the Consejo Superior de Investigaciones Científicas of Spain (CSIC) and the European Social Fund (ESF) (JAE-pre-PhD fellowship).

4.1.5 References

Allentoft ME, O'Brien J (2010) Global amphibian declines, loss of genetic diversity and fitness: a review. Diversity 2: 47–71.

Andrés JA, Bogdanowicz SM (2011) Isolating microsatellite loci: looking back, looking ahead. In: Orgogozo V, Rockman MV (eds), Molecular Methods for Evolutionary Genetics, Methods in Molecular Biology 772, Part 3, Humana Press, NY, USA, pp 211–232.

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

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

Arntzen JW, Szymura JM (1984) Genetic differentiation between African and European midwife toads (*Alytes*, Discoglossidae). Bijdr Dierkd 54: 157–162.

Arntzen JW, Groenenberg DSJ, Alexandrino J, Ferrand N, Sequeira F (2007) Geographical variation in the golden-striped salamander, *Chioglossa lusitanica* Bocage, 1864 and the description of a newly recognized subspecies. J Nat Hist 41: 925–936.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2000) 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.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4: 359–361.

European Commission (2012) Action Plan for the Conservation of the Common Midwife Toad (*Alytes obstetricans*) in the European Union. Final draft (April 17 2012)

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611–2620.

Fonseca A, Arntzen JW, Crespo EG, Ferrand N (2003) Regional differentiation in the common midwife toad (*Alytes obstetricans*) in Portugal: a picture from mitochondrial DNA. Z Feldherpetol 10: 83– 89.

García-París M (1995) Variabilidad genética y distribución geográfica de *Alytes obstetricans almogavarii* en España. Rev Esp Herpetol 9: 133– 138.

García-París M, Martínez-Solano I (2001) Nuevo estatus taxonómico para las poblaciones ibero-mediterráneas de *Alytes obstetricans* (Anura: Discoglossidae). Rev Esp Herpetol 15: 99–113.

Glor RE, Laport RG (2012) Are subspecies of *Anolis* lizards that differ in dewlap color and pattern also genetically distinct? A mitochondrial analysis. Mol Phylogenet Evol 64: 255–260.

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

Gonçalves H, Martínez-Solano I, Ferrand N, García-París M (2007) Conflicting phylogenetic signal of nuclear vs mitochondrial DNA markers in midwife toads (Anura, Discoglossidae, Alytes): deep coalescence or ancestral hybridization? Mol Phylogenet Evol 44: 494– 500.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www2.unil. ch/popgen/softwares/fstat.htm.

Grossenbacher K (1997) *Alytes obstetricans* (Laurenti, 1768). In: Gasc JP, Cabela A, Crnobrnja-Isailovic J, Dolmen D, Grossenbacher K, Haffner P, Lescure J, Martens H, Martínez-Rica JP, Maurin H, Oliveira ME, Sofianidou TS, Veith M and Zuiderwijk A (eds), Atlas of Amphibians and Reptiles in Europe, Societas Europaea Herpetologica, Paris, pp 94–95.

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

Hawlitschek O, Nagy ZT, Glaw F (2012) Island evolution and systematic revision of Comoran snakes: why and when subspecies still make sense. PLoS ONE 7: e42970. doi:10.1371/journal.pone.0042970.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801–1806.

Köhler J, Vieites DR, Bonett RM, García FH, Glaw F, Steinke D, Vences M (2005) New amphibians and global conservation: a boost in species discoveries in a highly endangered vertebrate group. Bioscience 55: 693–696.

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

Martínez-Solano I, Teixeira J, Buckley D, García-París M (2006) Mitochondrial DNA phylogeography of *Lissotriton boscai* (Caudata, Salamandridae): evidence for old, multiple refugia in an Iberian endemic. Mol Ecol 15: 3375–3388.

Phillimore AB, Owens IPF (2006) Are subspecies useful in evolutionary and conservation biology? Proc Biol Sci 273: 1049–1053.

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

Rice W (1989) Analyzing tables of statistical tests. Evolution 43: 223– 225.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731–2739.

Tobler U, Garner TWJ, Schmidt BR (2013) Genetic attributes of midwife toad (*Alytes obstetricans*) populations do not correlate with degree of species decline. Ecology and Evolution. doi:10.1002/ece3.677.

Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primerdimer and hairpin structures. Biotechniques 37: 226–231.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4: 535–538.

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

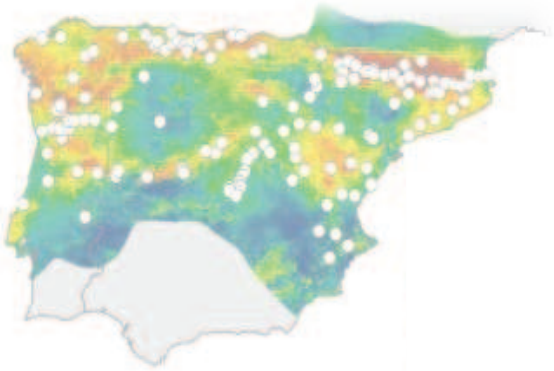
Appendix 1. Characterization of seventeen microsatellite markers developed in *Alytes obstetricans*

Locus	GenBank	Repeat motif	Primer Sequences	MP	Primer label (flourescent Dye)	Allele size range (bp)	Cross Amplification*			
							Ad	Ama	Amu	Ac
Aobstμ01	KF626401	(ACAT)n	F: *AGGCCATAGACTGCGTTTCA	Plex1	VIC	130-266	2	3	1	-
			R: TGCTCACTCACTGCTTAAATTCC							
Aobstμ02	KF626402	(ACAG)n	F: *TCAACTTAGCTTACACCAAAATGG	Plex1	NED	120-204	2	4	4	2
			R: CATTGCTCTCCTACACCACTG							
Aobstμ03	KF626403	(AAC)n	F: *TGTCCCACCTGAGTGTCTGA	Plex1	PET	239-278	-	4	1	1
			R: AGTTTCTCTGTTCGCATCATTG							
Aobstμ04	KF626404	(GTTT)n	F: *TGCAGGTTATATGTGTAATGGTGTT	Plex2	VIC	222-330	3	4	3	2
			R: TGGCATAAGGTTTGAACATGC							
Aobstμ05	KF626405	(ATCT)n	F: *CCTGCTCGCTTAACCCTATG	Plex2	NED	175-359	3	3	3	3
			R: TGACTTGGAGTTACATTGTGTTGTT							
Aobstμ06	KF626406	(CTGT)n(ATCT)n	F: *TGAGCAGGACTGGCTTTCTT	Plex2	6-FAM	299-391	3	4	3	2
			R: CCTGCAATACTCAAACTCTTACCA							
Aobsttμ07	KF626407	(ATCT)n	F: *CTGTCCTGTGACCCCTCTGT	Plex3	VIC	230-582	4	-	-	-
			R: TAATACCAGACCCGGCAAAG							
Aobsttμ08	KF626408	(ATCT)n	F: *CTTAGGCGTGTCCAGTAAAATG	Plex3	NED	218-410	-	-	-	-
			R: GGAAAGGTGATGTCCATTGTG							
Aobstμ09	KF626409	(AAAC)n	F: *GACCATAACGACTCTGCACTG	Plex3	PET	214-270	3	2	2	-
			R: CTTGCATCTGAGCTTGTGGA							
Aobstμ10	KF626410	(GTTT)n	F: *TGAAAGCACTGGCACTAAGG	Plex4	VIC	284-304	2	4	4	-
			R: CATGAACCCAAAATAAGCAGTAA							
Aobstμ11	KF626411	(GTTT)n	F: *GGTTTTCCATATTCCAACATCAA	Plex4	NED	233-305	-	2	4	2
			R: TGACCTAGAATTTACCAGCATGAA							

Appendix 1. *Continued.*

Locus	GenBank	Repeat motif	Primer Sequences	MP	Primer label (flourescent Dye)	Allele size Range (bp)	Cross Amplification			
							Ad	Ama	Amu	Ac
Aobstμ12	KF626412	(GTTT)n	F: *GGCAGCAGATTTAGTGGACA	Plex4	6-FAM	124-172	3	4	4	3
			R: AGCATTTGTTTTCCTTCTTCAG							
Aobstμ13	KF626413	(ACT)n	F: *CTGCATTGCTCGTATTCTGC	Plex4	PET	217-271	2	2	3	-
			R: TCACCTGCCACATAAGAGTCC							
Aobstμ14	KF626414	(ACT)n	F: *TGTGGGAACCTTTACATCATAA	Plex5	VIC	166-214	-	-	-	-
			R: CCCTCCTCTAAGCCGTCA							
Aobstμ15	KF626415	(AGAT)n	F: *TTGGATGGTGGGTACAATCA	Plex5	NED	268-444	-	-	-	-
			R: TGAGGACAAATGCCTGACAA							
Aobstμ16	KF626416	(AGAT)n	F: *TCAGAATAAACAAGAGCTGCAAA	Plex5	6-FAM	284-440	-	4	2	-
			R: GGAGATCCACGCTCAGGATA							
Aobstμ17	KF626417	(ACC)n	F: *CGGTGTCCCCATCTTATCAA	Plex5	PET	272-308	-	4	3	3
			R: CCCAGTGCTCAAACCTCAAT							

GenBank accession numbers, repeat motif, forward (F) and reverse (R) primer sequences, amplification multiplex panel (MP), primer label, allele size range and cross amplification results. *The cross-amplification was tested in four individuals of each other congeneric species: Ad-*Alytes dickhilleni*, Ama-*Alytes maurus*, Amu-*Alytes muletensis*, Ac-*Alytes cis*



Artigo IV

Original Article
Subimitted to Journal of Biogeography

When the environment favours lineage diversification: ecological niche divergence mirrors intraspecific patterns of genetic diversity in the Common Midwife Toad (*Alytes obstetricans*)

Bruno Maia-Carvalho[1,2], Cândida Gomes Vale[1], Fernando Sequeira[1], Nuno Ferrand[1,2,3,4], Iñigo Martinez-Solano[1,5,6*], Helena Gonçalves[1,4]

[1] CIBIO Research Centre in Biodiversity and Genetic Resources, InBIO, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal

[2] Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, 4099-002 Porto, Portugal

[3] Department of Zoology, University of Johannesburg, Johannesburg, South Africa

[4] Museu de História Natural e da Ciência da Universidade do Porto (MHNC-UP), Praça Gomes Teixeira, 4099-002 Porto, Portugal

[5] Instituto de Investigación en Recursos Cinegéticos (IREC-CSIC-UCLM), Ronda de Toledo, s/n, 13005 Ciudad Real, Spain

[6]Ecology, Evolution and Development Group, Department of Wetland Ecology, Estación Biológica de Doñana (EBD-CSIC), Avenida Américo Vespucio, s/n, 41092 Sevilla, Spain

4.2 When the environment favours lineage diversification: ecological niche divergence mirrors intraspecific patterns of genetic diversity in the Common Midwife Toad (*Alytes obstetricans*)

4.2.1 Abstract

4.2.1.1 Aim

Unravelling the relative role of geographic and climatic factors in promoting lineage diversification is central to our understanding of how new species arise and evolve. To address this topic, we apply an innovative approach that integrates multilocus datasets and species distribution models to test for ecological divergence across genetically differentiated population lineages at the intraspecific level. We use as a case study the Common Midwife Toad (*Alytes obstetricans*), for which previous phylogenetic and phylogeographic analyses of mitochondrial and nuclear markers have identified a major role for climatic and environmental changes during the Pleistocene in shaping the population history of six divergent genetic lineages.

4.2.1.2 Location

Iberian Peninsula, Europe.

4.2.1.3 Methods

We combined information from 12 nuclear microsatellite markers with tests of niche overlap, equivalency and similarity as well as ecological niche-based models in an extensive dataset (965 individuals from 142 localities) to finely delineate the ranges of genetic lineages of *A. obstetricans*, explore the extent of niche divergence between them, and determine whether their range boundaries correspond to divergent niche limits.

4.2.1.4 Results

We recovered seven differentiated genetic clusters, with high levels of geographically structured genetic variation, and a strong signal of niche divergence across most cluster pairs. We also found limited overlap in areas of environmental suitability across genetic clusters, indicating that genetic range limits are coincident with ecological niche limits.

4.2.1.5 Main conclusions

Our results suggest that adaptation to different ecological niches may have played an important role in driving intraspecific lineage diversification in *A. obstetricans*. The topographic and climatic heterogeneity in the Iberian Peninsula has probably favored processes of adaptive

genetic differentiation following older population fragmentation events. Our study highlights the importance of incorporating intraspecific genetic structure into species distribution models to explore spatial patterns of genetic diversity in terms of their underlying generating processes.

Keywords: Ecological niche-based modelling, incipient speciation, intraspecific lineage, lineage diversification, microsatellites, niche divergence.

4.2.2 Introduction

Identifying factors that promote genetic differentiation among populations is crucial for understanding how new species arise, diversify and adapt (Wiens 2004, Frankham 2005, Freedman *et al.* 2010, Dufresnes *et al.* 2014). Traditionally, biological diversification has been mainly explained in the light of classical hypotheses relying on vicariance, especially on the emergence of geophysical barriers (Avise 2000, Hewitt 2004). Vicariant events, such as the uplift of mountain ranges and the formation of river basins, are drivers of allopatric divergence by limiting or impeding genetic exchange between populations (Hewitt 2004). Notwithstanding, there is growing evidence that climatic/environmental factors also play a key role in the process of biological diversification (see Sacks *et al.* 2004, Schoville *et al.* 2012, Alberdi *et al.* 2015, Tocchio *et al.* 2015). Recent studies using ecological niche modelling approaches have shown that environmental conditions can drive the evolutionary fate of populations and/or species, particularly through effects on: i) the geographical distribution - even when there is niche conservatism –, and ii) patterns of genetic structure - even with the occurrence of gene flow (Wiens 2004, Ortego *et al.* 2012, but see McCormack *et al.* 2010 and Tarroso *et al.* 2014). When both factors (geophysical and ecological) act synergistically, they may enhance the population divergence process (e.g. “species pumps”; Bryant *et al.* 2008; Kozak and Wiens 2010).

Most organisms live in environments that vary over time and space. In facing this environmental heterogeneity, genetic differentiation among populations across species’ ranges can arise due to dispersal limitations and genetic drift (Dyer *et al.* 2010, Freeland *et al.* 2010, Stewart 2009, Stewart *et al.* 2010), or to natural selection towards local adaptation (Garant *et al.* 2005, Postma and Van Noordwijk 2005). However, disentangling the relative effects of those processes in shaping patterns of population genetic structure remains a complex task. This is because divergent populations or intraspecific lineages (i.e. subspecies and evolutionary significant units – ESUs, see Mayr 1942, Moritz 1994 and Wood *et al.* 2014) can respond similarly to environmental changes, being in effect ecologically analogous – the niche conservatism concept (Stewart *et al.* 2010, but see Hughes and Woodward 2008). In fact, there is an ongoing debate about the importance of fundamental niche retention in closely related lineages (niche conservatism) and how the environment promotes or restricts genetic divergence (Wiens and Graham 2005, Serra-Varela *et al.* 2015). Importantly, the phylogenetic niche conservatism theory considers adaptation as a continuous process at local scales. In consequence, intraspecific lineages may also exhibit either niche divergence or conservatism (Pyron *et al.* 2014, but see Pearman *et al.* 2008).

Climatic factors are often as dynamic as topography, and climatic changes through time may intensify niche divergence effects (see Abellán and Svenning 2014, Gómez-Rodríguez *et al.* 2015). It is well documented that alternating glacial and interglacial cycles during the Pleistocene have played a major role in processes of lineage diversification (Taberlet *et al.* 1998, Hewitt 2000, 2011), affecting mainly species in temperate regions. In Europe, the Iberian Peninsula was one of the most important long-term refugia due to its geographical position and high physiographic complexity, encompassing very distinct bioclimatic areas (Sillero *et al.* 2009). This topographic and climatic heterogeneity has also contributed to favor the persistence of isolated species and populations in multiple microrefugia (the "refugia-within-refugia" model), promoting high levels of intraspecific genetic diversity (Gómez and Lunt 2007) and perhaps additional processes of ecological divergence. Genetic effects of vicariant events associated with climate-mediated range expansions and contractions during glacial cycles are well documented in several phylogeographic studies highlighting the existence of cryptic diversity in many taxa (e.g. Martínez-Solano *et al.* 2006, Sequeira *et al.* 2008, Pinho *et al.* 2011, Recuero and García-París 2011, Díaz-Rodríguez *et al.* 2015, Gonçalves *et al.* 2015). However, whereas genetic data has been widely used to document the persistence of populations in refugia, few efforts have been directed towards investigating potential processes of adaptation to new environments and their role in generating current patterns of genetic structure (see Frankham 2005, Allentoft and O'Brien 2010, Dufresnes *et al.* 2013, 2014; Ferchaud *et al.* 2015). In this respect, the integration of molecular, ecological and environmental data at the intraspecific level has great potential to clarify the role of processes of local adaptation and genetic divergence (see Kozak *et al.* 2008, Pearman *et al.* 2008, McCormack *et al.* 2010). Ecological niche-based models (ENMs) and niche overlap analysis (Kozak *et al.* 2008, Tarroso *et al.* 2014, Rato *et al.* 2015) have been applied to assess the role of environmental conditions in the diversification, viability and maintenance of populations and species lineages (Guisan and Thuiller 2005, Thomassen *et al.* 2010, Gotelli and Stanton-Gueddes 2015). Although these approaches have been used in several studies during the last decade, only few of them have included both climatic and genetic data (see D´Amen *et al.* 2013, Lee-Yaw and Irwin 2015, Serra-Varela *et al.* 2015, Morales *et al.* 2016). The combination of genetic and environmental data can provide a powerful tool to evaluate population structure and answer biogeographical questions (Gotelli and Stanton-Gueddes 2015), identifying whether niche adaptation may be an effective force shaping phylogeographic structure (Serra-Varela *et al.* 2015).

In the present study we integrated multilocus genetic data and ecological niche-based models (ENMs) to explore the factors that have shaped patterns of genetic diversity and promoted the process of intraspecific diversification in the Common Midwife Toad (*Alytes obstetricans*) in the Iberian Peninsula. This species is widely distributed in Western Europe

and occurs in a wide variety of habitats (mountain, farmland and even urban areas) and bioclimatic zones (Temperate Atlantic, Continental and Mediterranean), occurring from sea level up to 2,400 m in the Pyrenees (Márquez and Rosa 1997). A recent phylogeographic study (Gonçalves *et al.* 2015) revealed divergent, geographically structured mtDNA haplogroups diagnosing six population lineages (labeled A to F), with varying but overall low levels of admixture in nuclear markers. According to this study, and given the estimated timeframe for the splits between major lineages, the climatic changes that occurred during the Pleistocene would have left a remarkable imprint on the diversification history of *A. obstetricans* in the Iberian Peninsula, where most of its genetic diversity is found (see Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007, Maia-Carvalho *et al.* 2014a,b; Gonçalves *et al.* 2015). The lineage diversification process in *A. obstetricans* could thus be the result of demographic events (cycles of population contraction/expansion), accompanied or not by processes of local adaptation. Here we build on previous studies by analysing a more extensive genetic dataset (both in terms of the number of individuals and populations analysed and the number of genetic markers) in combination with ecological niche modeling with the specific goals of: i) describing finer-scale patterns of intraspecific genetic diversity and population structure in *A. obstetricans*, including delineating the ranges of major genetic clusters; ii) evaluating the existence of niche divergence or niche conservatism among genetic clusters; iii) analysing if geographical contact zones, based on genetic data, correspond to ecological niche overlap areas; and iv) examining the relative role of environmental and geographical effects as drivers of intraspecific diversification among *A. obstetricans* lineages. We discuss the benefits of integrating genetic and environmental data in population differentiation and speciation studies.

4.2.3 Material and Methods

4.2.3.1 Genetic approach

4.2.3.1.1 Sampling, DNA extraction and microsatellite genotyping

A total of 965 individuals from 142 localities were sampled across the species' entire geographical range (Fig. 1 and Table S1 - Supporting Information, Appendix S2). Sampling effort focused in the Iberian Peninsula, where most of the genetic diversity of the species is found (Gonçalves *et al.* 2015). A limited subset of samples was used in previous studies (Gonçalves *et al.* 2007, 2015; Maia-Carvalho *et al.* 2014a,b). Tissue samples were collected from adult toe tips or larval tails and preserved in 95% ethanol. Genomic DNA was extracted

from each tissue sample using EasySpin Genomic DNA Minipreps Tissue Kit (SP-DT-250, Qiagen, Hilden, Germany) following the fabricant's protocol. We used seventeen polymorphic microsatellite markers previously characterized for *A. obstetricans* (Maia-Carvalho *et al.* 2014a) to genotype all samples. For specific annealing temperatures, PCR conditions and scoring methods see Maia-Carvalho *et al.* (2014a).

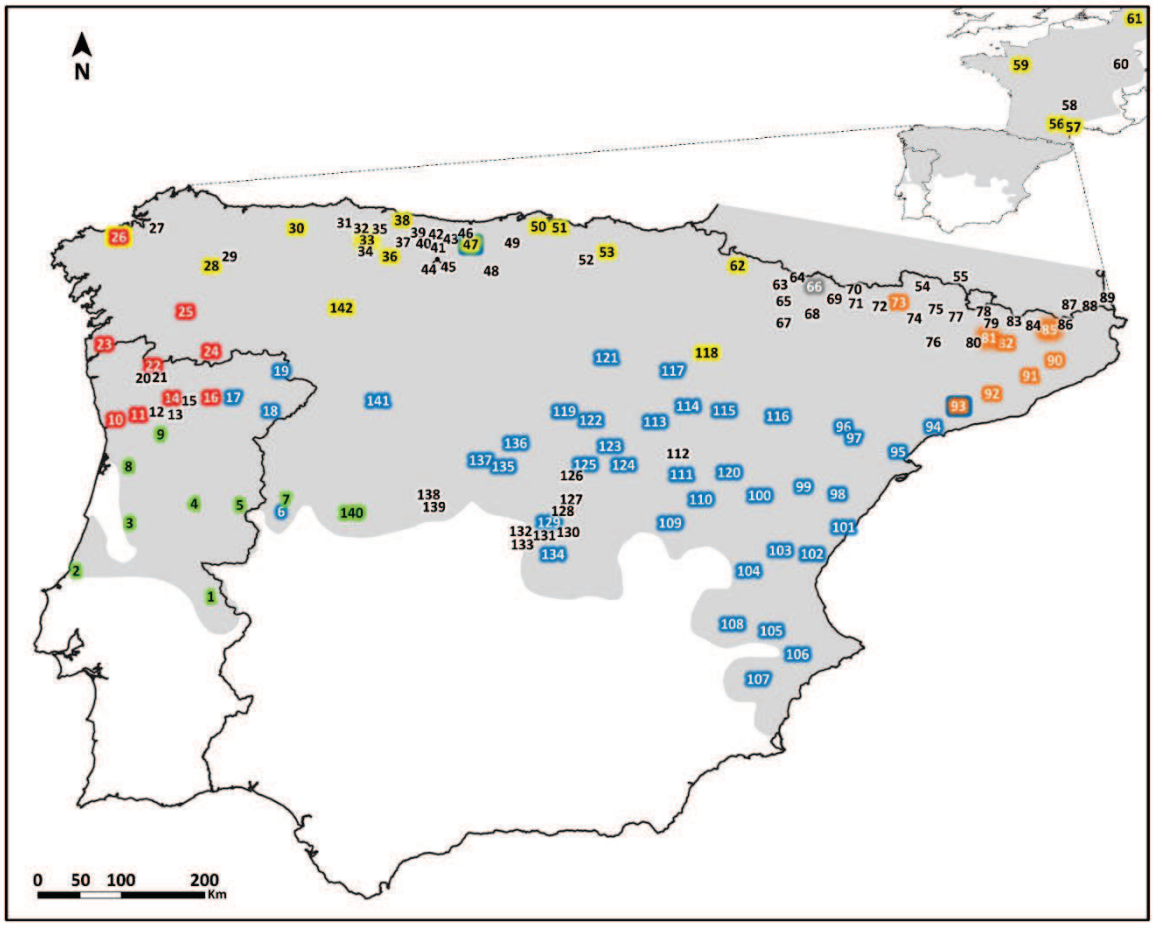


Figure 1. Distribution map of *Alytes obstetricans* populations analysed in the present study (population numbers according to Table S1, Appendix S2, Supporting Information). Coloured shades around numbers represent populations where the mtDNA haplogroup is already known from a previous study (Gonçalves *et al.* 2015). Blue: mtDNA haplogroup A; yellow: mtDNA haplogroup B; red: mtDNA haplogroup C; green: mtDNA haplogroup D; gray: mtDNA haplogroup E; and orange: mtDNA haplogroup F. Note that in three populations (26, 47 and 93) haplotypes of different haplogroups were found co-occurring (Gonçalves *et al.* 2015).

4.2.3.1.2 Genetic diversity and population structure

We assessed standard genetic diversity statistics, including observed (H_O) and expected (H_E) heterozygosities and allelic richness (A_R) and tested for deviations from Hardy-Weinberg Equilibrium (HWE) and Linkage Disequilibrium (LD) in 56 populations with a sample size ≥5 (Table S2, Supporting Information, Appendix S2). The same basic genetic diversity

statistics were estimated for each differentiated genetic cluster as defined by results of Bayesian model-based clustering and Discriminant Analysis of Principal Components (DAPC, see Results). All tests and estimates were calculated in FSTAT v2.9.3.2 (Goudet 2001), adjusting significance levels for multiple comparisons using the sequential Bonferroni correction (α = 0.05; Rice, 1989). Patterns of population structure were assessed using three different approaches: i) a non-spatial Bayesian multilocus clustering analysis implemented in STRUCTURE v2.3.4 (Pritchard *et al.* 2000); ii) a spatial-based clustering analysis implemented in TESS v2.3.1 (Chen *et al.* 2007); and iii) a discriminant analysis of principal components (Jombart 2008). Details on models and parameters used in these analyses are presented in Appendix S1.

Genetic differentiation between clusters (as defined by results of Bayesian clustering and DAPC analyses, see Results) was quantified with pairwise F_{st} values, estimated in ARLEQUIN V3.5.1.2 (Excoffier 2005). Statistical significance was assessed through 1,000 iterations and 95% bootstrapped confidence intervals (CI). We also performed a Mantel test (Mantel 1967) as implemented in software IBD (Bohonak 2002) to test the Isolation-by-distance (IBD) hypothesis in each genetic cluster (see Meirmans 2012, about the potential confounding role of IBD when analysing population structure).

4.2.3.2 Spatial approach

4.2.3.2.1 Study area

The study area used for modelling purposes was restricted to the Iberian Peninsula. It was defined with a buffer of 100km encompassing the range of the IUCN extent of the species' occurrence polygon in the Iberian Peninsula and further extended 100km to the northeast, to account for potential dispersal and include the entire Pyrenees conifer and mixed forests and Cantabrian mixed forests ecoregions (Olson *et al.* 2001). The rest of the European range of the species, north of the Pyrenees, was not included for modelling purposes, due to the lack of genotyped samples at a comparable level of sampling detail.

4.2.3.2. 2 Environmental analysis

A set of uncorrelated ecogeographical variables (EGVs; $r < 0.73$ in all cases; Table 1) was chosen based on their putative effect in shaping the range of *A. obstetricans,* including intraspecific lineages (=genetic clusters), and their representativeness in the study area. EGVs comprise three main categories: (i) topographical – slope derived from the SRTM digital elevation model (http:// srtm.usgs.gov); (ii) climatic – five grids extracted from WORLDCLIM v1.4 (Hijmans *et al.* 2005) and one extracted from CGIAR-CSI (Trabucco and Zomer 2009); and (iii) habitats – five distance to land cover grids derived for years 2004-2006 (Bicheron *et al.* 2008). To convert the original categorical habitat EGVs into continuous variables, one binary grid for each habitat type was created. The Euclidean distance of each grid cell to the closest habitat type cell was then calculated. All variables have a grid resolution of approximately 1 km^2.

Table 1. General information on the environmental variables used to predict the potential niche of *Alytes obstetricans* genetic clusters.

Code	Description	Range	Units
Bioclimatic			
Bio3	Isothermality (mean diurnal range/annual range) (* 100)	29-46	Coef. var.
Bio7	Temperature Annual Range (Max. Temp. of warmest month - Min. Temp. of coldest month)	12.2-33.6	°C
Bio11	Mean Temperature of Coldest Quarter	-8.9 - 12.3	°C
Bio15	Precipitation Seasonality	11 - 66	Coef. var.
Bio16	Precipitation of Wettest Quarter	97 - 719	mm
Pet	Annual average potential evapotranspiration	399 - 1311	mm
Topographical			
Slope	Slope (derived from altitude)	0 - 68.9	%
Habitat (distance to)			
CropVeg	Mosaic cropland (> 50%) / vegetation (grassland/shrubland/forest) (< 50%)	0 - 0.25	º
DecFor	Closed (>40%) broadleaved deciduous forest (>5m)	0 - 1.20	º
RainCrop	Rainfed croplands	0 - 0.32	º
Rivers	Hydrological basins	0 - 0.75	º
SparVeg	Sparse (<15%) vegetation	0 - 0.15	º

4.2.3.2.3 Ecological niche-based analyses

From the initial 142 geo-referenced populations that were genetically analysed, those located outside the Iberian training area were excluded (N=6). Of the remaining populations, 132 non-clustered populations were randomly selected for modelling purposes according to the Nearest Neighbour Index (NNI=0.82, *p*=0.08) estimated using ARCGIS v10.0 (ESRI 2011). Ecological niche-based models were developed using the artificial neural networks (ANN) algorithm implemented in SIMAPSE (Tarroso *et al.* 2012). ANN is a machine learning algorithm that finds the ecological niche model for a given population structure. It uses a training data set with population affinity scores and the subset of EGVs that might represent the environmental features acting on the ecological divergence between taxa (Tarroso *et al.* 2014). Compared to other presence-only methods, the advantages of this modelling technique include: (i) it uses the full continuum of membership coefficients to any given number of taxa, clusters or lineages, avoiding issues related to the assignment of individuals to discrete taxa (Tarroso *et al.* 2014); (ii) it has optimum performance with nonlinear relationships between dependent and independent variables (Olden *et al.* 2008); and (iii) it uses a flexible algorithm supporting several degrees of complexity (Olden *et al.* 2008). Cluster membership probabilities (Q-scores) as determined by STRUCTURE were directly input to SIMAPSE with the corresponding geographical coordinates of each population sample location. All parameters and conditions used in this modelling approach are described in Appendix S1.

4.2.3.2.3 Niche analyses

We tested for niche overlap in both geographical and environmental space. In geographical space, binary maps for each cluster were intersected to create a single map ("Combined") highlighting the overlap in predicted suitable areas among the different genetic clusters. Potential values range from zero (i.e. bioclimatic conditions are unsuitable for any cluster) to seven (i.e. cells where bioclimatic conditions are suitable for all clusters). Then, we quantified overlap among pairs of clusters.

In environmental space, niche overlap among clusters was assessed based on the same set of uncorrelated EGVs used for modelling purposes as well as in the same set of populations used for converting continuous to binary maps. Cluster B2 (see results) was excluded from this analysis, due to the low number of populations within the Iberian Peninsula. We applied the statistical framework termed PCA-env (Broennimann *et al.* 2012) for each cluster, and the available environmental space was defined by all conditions within a buffer of 100km enclosing the population localities, representing the potentially colonisable

environmental space accounting for both natural and human-mediated dispersal (Rato *et al.* 2015). Both the set of population localities and each set of available EGVs were projected into PCA space. The relative density of background conditions and cluster records across the first two principal components were captured by a kernel density smoother function in order to create density grids of r x r cells in environmental space. Despite using smoothed densities from a kernel density function ensured the independency of the measure of overlap from the grid's resolution, the latter was set to 100 (Broennimann *et al.* 2012). The density grids were then used to quantify niche overlap in terms of Schoener´s *D*, which ranges from 0 (no overlap) to 1 (complete overlap) (Warren *et al.* 2008).

Hypotheses of niche equivalency and niche similarity were computed from the density estimations of the clusters in environmental space (Broennimann *et al.* 2012). The niche equivalency test evaluates if the niches of two entities are significantly different from each other and if niche spaces are interchangeable. The niche similarity test whether similarity in the realized niche of two taxa can be due to available environmental space or to processes of active, divergent habitat selection. All analyses were performed in R version 3.1.1 (www.r-project.org) using the scripts provided by Broennimann *et al.* (2012) which were adapted to our data (see Appendix S1 for details).

4.2.4 Results

4.2.4.1 Genetic diversity

From the initially selected 17 loci, five were discarded due to amplification inconsistencies in some groups of samples. No significant deviations from Hardy-Weinberg or linkage equilibria across all populations and loci were identified after applying the sequential Bonferroni correction. Statistics of genetic variation at all microsatellite loci are summarized in Table S2 (Supporting Information, Appendix S2). All loci showed high levels of polymorphism, with an average of 2.42 to 11.92 sampled alleles per locus and population.

4.2.4.2 Population structure and genetic differentiation

All the clustering analyses performed recovered concordant results (Fig. 2). The Bayesian clustering analysis with STRUCTURE identified two main groups (*K*=2) as the highest hierarchical level of genetic structure in our dataset (based on both the Ln Pr(*X/K*) and the ΔK methods) (Fig. 2a): one comprising all sampled populations in northeastern Spain, corresponding to subspecies *A. o. almogavarii,* and a second group with all the remaining *A. obstetricans* populations. This distinction was also supported by the DAPC analysis, which shows the clear distinctiveness of populations of *A. o. almogavarii* (Fig. 2c). The spatial clustering analyses performed with TESS provide evidence for further population subdivision, recovering seven genetic clusters (Fig. 2b). This partitioning into seven clusters is also supported by results of the DAPC analysis (Fig. 2c).

When STRUCTURE was performed forcing *K*=7 the same population subdivision was observed. Although some admixed individuals were identified, most genotyped individuals in the dataset had an assignment probability greater than 0.80 to one genetic cluster. Concerning the population assignment results, there was a clear prevalence of one genetic cluster in most populations, although a few localities included individuals from different genetic clusters in similar proportions (Fig. 2a).

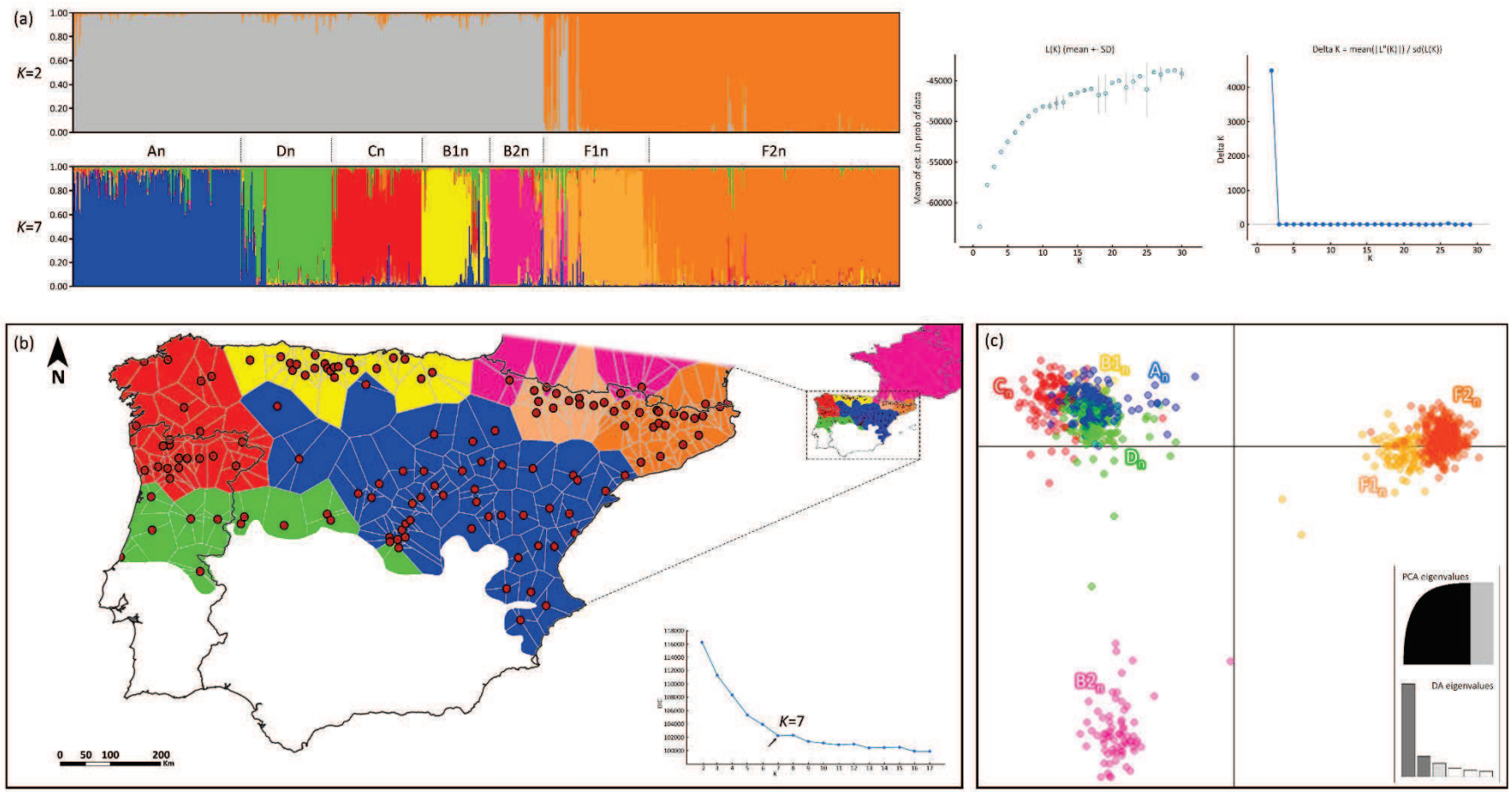


Figure 2. Population structure of *Alytes obstetricans* in the Iberian Peninsula based on 12 microsatellite loci. (a) Bayesian clustering results of STRUCTURE analyses. Individual assignment probabilities to each genetic cluster (*K*=2, top and *K*=7, bottom). Each individual is represented as a vertical line partitioned into *K* colored segments whose length is proportional to the individual's estimated membership coefficient Q. Graphics on the right show the mean likelihood scores and ΔK values estimated for different values of *K*, respectively. Colors represent the seven genetic clusters identified: blue – A_n; yellow – $B1_n$; pink – $B2_n$; red – C_n; green – D_n; light orange – $F1_n$; orange – $F2_n$. (b) Illustrative map of admixture proportions generated from the Bayesian analysis implemented in TESS. Red dots indicate the populations analyzed. The graphic on the bottom right shows the mean values of the DIC statistic (averaged over five runs) estimated for different *K* values, ranging from 1 to 17. Color codes as for STRUCTURE results (c) Plot of the first two axes obtained in the Discriminant Analysis of Principal Components (DAPC). Each dot represents one individual. Color codes as for STRUCTURE results.

Parameters of standard genetic diversity for each cluster are described in Table 2. Cluster $B2_n$ exhibited the lowest genetic diversity values, whereas cluster D_n presented the highest. Pairwise F_{st} values between clusters ranged from 0.120 to 0.307 and were all significantly different from zero, indicating strong genetic structure (Table 3). Concerning Isolation-by-distance tests, support for an IBD scenario was observed only for clusters Cn and F2n ($P < 0.01$) (Fig. S1, Appendix S3, Supporting Information).

Table 2. Summary statistics of genetic variation in each of the seven major genetic clusters in *Alytes obstetricans*. N: Number of samples; N_a: number of alleles sampled; A_r: Allelic Richness; H_o: observed heterozygosity; H_e: expected heterozygosity. IBD: p-values for the Isolation-by-distance tests in each genetic cluster.

		Genetic diversity				
Cluster	N	N_a	A_r	*Ho*	*He*	IBD
A_n	171	19.1	14.8	0.661	0.824	0.063
$B1_n$	50	9.1	9.0	0.611	0.688	**
$B2_n$	51	8.9	8.8	0.397	0.668	0.708
C_n	87	23.6	19.6	0.590	0.771	0.000
D_n	82	24.9	21.3	0.653	0.837	0.308
$F1_n$	92	15.0	13.0	0.621	0.689	0.206
$F2_n$	275	23.1	15.8	0.617	0.714	0.005

Table 3. Microsatellite-based pairwise F_{st} estimates between the seven genetic clusters in *Alytes obstetricans*. All values are significantly different from zero ($P < 0.05$).

	A	B1	B2	C	D	F1	F2
A	0						
B1	0.180	0					
B1	0.207	0.278	0				
C	0.120	0.237	0.208	0			
D	0.089	0.169	0.193	0.122	0		
F1	0.216	0.218	0.307	0.292	0.229	0	
F2	0.212	0.240	0.275	0.263	0.209	0.166	0

4.2.4.3 Ecological niche modelling

The ecological niche-based model for *Alytes obstetricans* was accurate in predicting the presence of the species ($AUC_{A.\ obstetricans} = 0.806$), closely depicting its known distribution in the Iberian Peninsula (Fig. 3). Cluster-based models exhibited limited overlap between areas of suitability, with the exception of cluster Cn, which overlapped with cluster Dn in the southwest of the study area. Accordingly, the models showed steep transitions between areas of high and low suitability for all clusters, with the exception of clusters B2n and Dn (Fig. 3).

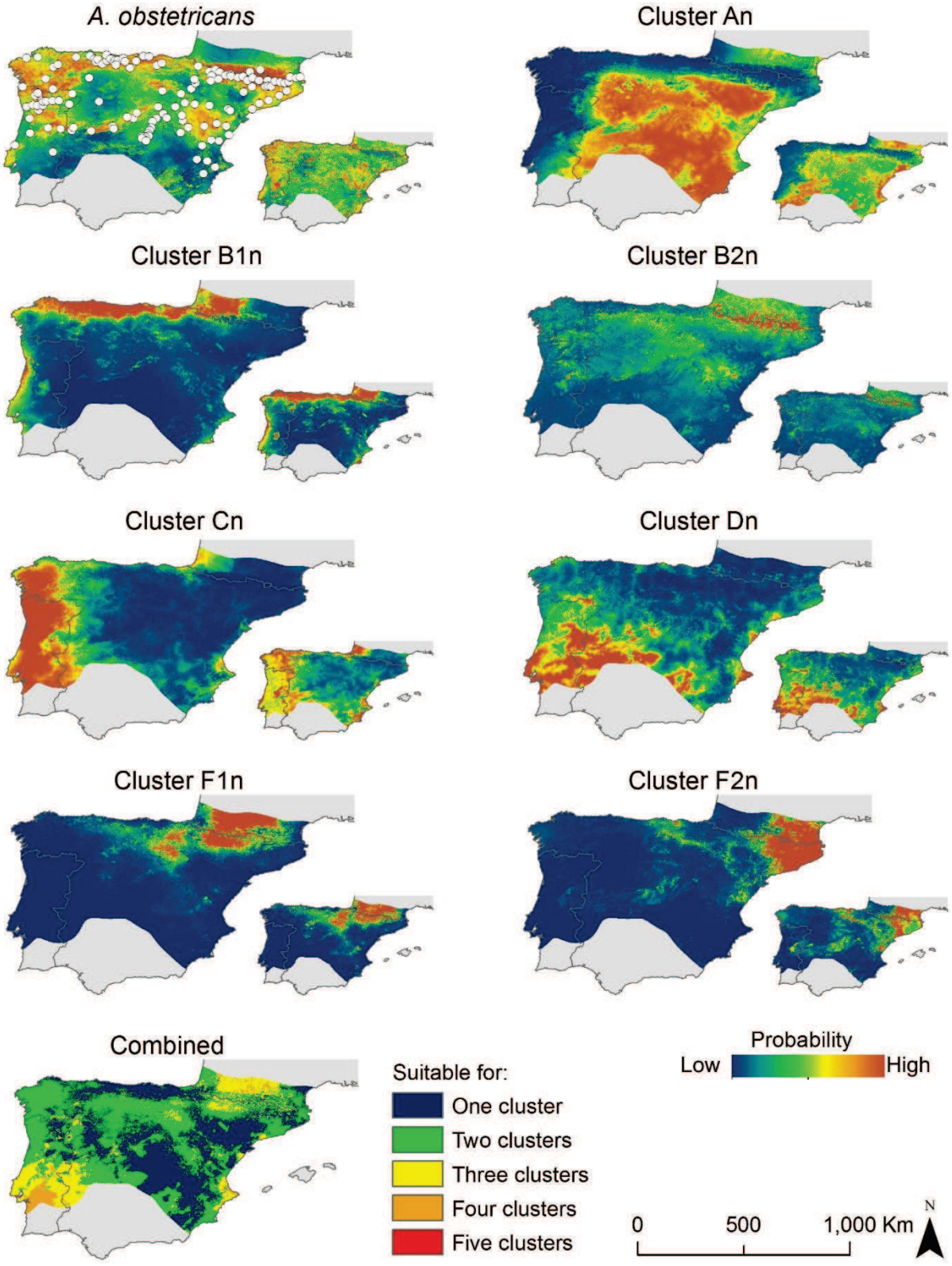


Figure 3. Ecological niche-based models of *Alytes obstetricans* using membership probabilities of sampled populations to inferred genetic clusters. Larger maps represent the probability of presence for *A. obstetricans* and for each cluster separately. Smaller maps at the lower right side of each map represent the standard deviation between replicates. The Combined map represents niche overlap among lineages in geographic space.

The accuracy of the ecological niche-based analysis for each cluster was high, since ecological model results were a significant predictor of Q-scores ($P < 0.01$ for all models; Fig. S2 in Appendix S3, Supporting Information). EGVs had different contributions for the multivariate model of *A. obstetricans* and for each cluster (Fig. S3, Appendix S3, Supporting Information). For instance, "Slope" and "Distance to mosaic cropland/vegetation" were the most important EGVs explaining the overall distribution of *A. obstetricans*. However, different EGVs contributed differently to explain the distribution of each genetic cluster. The variable "Precipitation of wettest quarter (bio16)" had high importance for the models of clusters An and Cn, followed by "Distance to closed broadleaved deciduous forest (decfor)" and "Temperature annual range (bio 7)" for cluster An, and by "Isothermality (bio3)" and "Precipitation seasonality (bio15)" for cluster Cn. Bio15 was the most important EGV for explaining the predicted distribution of clusters Dn, F1n and F2n. However, Dn models also had high factor loadings for the variable "Distance to deciduous forests and rivers", while in F2n models it is variables bio3 and "Mean temperature of coldest quarter (bio11)". Bio3 was the most important EGV in the model for cluster B1n, followed by "Annual average potential evapotranspiration (pet)". Bio11 had high importance in the model for cluster B2n, followed by "Distance to rainfed croplands (raincrop)" and "Slope" (Fig. S3, Appendix S3, Supporting Information).

4.2.4.4 Niche overlap analyses

Overall, the overlap values obtained from both geographical and environmental analyses were low (Table 4). From the total study area, only 0.1% was not suitable for any cluster, while 0.003 was predicted as suitable for five clusters. Around 37.1% and 45.15 % of the study area was inferred to be suitable for only one or two clusters, respectively. Areas with favourable conditions for three and four clusters (14.43% and 3.26%, respectively) were mostly located in the south-west and north-east of the study area (Fig. 3). For instance, clusters Cn and Dn overlapped 32.17% in geographic space in the south-west of the study area (Table 5 and Fig. 3), while clusters An and B1n overlapped 30.24% in the core of the study area and in both the north-east and south-west.

In the environmental analysis, the sum of the first two axes of the PCA-env for all pairwise comparisons among clusters explained between 44.33% and 53.78% of the total variance (Fig. S3, Appendix S3, Supporting Information). The pairwise niche overlap analysis using the D metric (Table 5) varied between no or very limited overlap (D = 0.023; cluster Cn — cluster F2n) to low overlap (D = 0.373; cluster F1n - cluster F2n), according to the metrics suggested by Rödder and Engler (2011). The null hypothesis of niche equivalence was rejected for all pairwise comparisons ($P<0.05$), suggesting that all clusters possess more

significantly distinct niches than expected by chance. For the pairs An-Dn, B1n-Cn, Cn-Dn, Cn-F2n, Dn-F1n, Dn-F2n and F1n-F2n, niche overlap is higher than expected based on the environmental conditions available to each, although for An-Dn, B1n-Cn, Cn-Dn, Cn-F2n and Dn-F1n, significant results were obtained in only one direction (Table 5).

Table4. Pairwise geographical (above the diagonal) and environmental (below the diagonal) overlap between clusters. Geographical overlap is calculated as the percentage of the total area suitable for any of the two clusters that is suitable for both at the same time. Environmental overlap is quantified by the D metric (Warren *et al.* 2008).

	A_n	$B1_n$	$B2_n$	C_n	D_n	$F1_n$	$F2_n$
A_n		30.24	2.19	10.24	19.36	5.675	4.4
$B1_n$	0.088		7.34	26.86	27.23	10.9	2.2
$B2_n$	NA	NA		0.935	0.301	13.7	9.9
C_n	0.097	0.131	NA		32.17	0.33	0
Dn	0.196	0.137	NA	0.259		0	0
$F1_n$	0.143	0.138	NA	0.034	0.165		3.4
$F2_n$	0.177	0.183	NA	0.023	0.196	0.37	

Table 5. Pairwise niche overlap values using the D metric, and P-values of niche similarity and equivalence via randomization test. Significant values (P<0.05) in the niche equivalency test and falling outside of the 95% CI of the niche similarity test are shown in bold.

Comparison on (1-2)	Niche overlap (D)	Niche Equivalency	Niche Similarity (2→1)	Niche Similarity (1→2)
Cluster An - Cluster B1n	0.088	**0.02**	0.28	0.67
Cluster An - Cluster Cn	0.097	**0.02**	0.38	0.30
Cluster An - Cluster Dn	0.196	**0.02**	**0.04**	0.20
Cluster An - Cluster F1n	0.143	**0.02**	0.36	0.28
Cluster An - Cluster F2n	0.177	**0.02**	0.59	0.89
Cluster B1n - Cluster Cn	0.131	**0.02**	0.44	**0.02**
Cluster B1n - Cluster Dn	0.137	**0.02**	0.05	0.69
Cluster B1n - Cluster F1n	0.138	**0.02**	0.79	0.85
Cluster B1n - Cluster F2n	0.183	**0.02**	0.73	0.75
Cluster Cn - Cluster Dn	0.259	**0.02**	0.26	**0.04**
Cluster Cn - Cluster F1n	0.034	**0.02**	0.83	0.10
Cluster Cn - Cluster F2n	0.023	**0.02**	0.18	**0.04**
Cluster Dn - Cluster F1n	0.165	**0.02**	1.00	**0.02**
Cluster Dn - Cluster F2n	0.196	**0.02**	**0.02**	**0.02**
Cluster F1n - Cluster F2n	0.373	**0.04**	**0.02**	**0.02**

4.2.5 Discussion

Our results show a striking degree of divergence in the environmental space between intraspecific lineages over a relatively small geographic area, highlighting the potential role of topographic and/or climatic heterogeneity in promoting or reinforcing patterns of genetic differentiation across populations. Whereas the dynamics of population expansion/contraction seem to have played a major role in lineage divergence during the Pleistocene (Gonçalves *et al.* 2015), there is an additional and clear signal of divergence in environmental space, suggesting possible processes of local adaptation in each major lineage. Furthermore, transition zones in environmental space coincide to a significant extent with genetic boundaries across lineages. By delimiting each cluster's potential distribution area based on independent yet complementary approaches, this study allows a detailed examination of patterns of gene flow across contact zones and will guide the search for signatures of adaptation at the genome level, making this an interesting emerging study system in speciation research.

Previous phylogenetic and phylogeographic analyses on the evolutionary history of *A. obstetricans* revealed high levels of population differentiation, suggesting that this species likely survived in allopatric refugia through the Pleistocene Ice Ages and responded differently to climatic oscillations after diverging from a common ancestor (Gonçalves *et al.* 2015). By extending previous molecular datasets in an increased population and genetic sample, including fast-evolving markers like microsatellite loci, we provide further insights into the relative role of historical and contemporary processes in determining current patterns of genetic structure and diversity in *A. obstetricans*. All analyses performed on the microsatellite dataset produced consistent results, supporting previous results based on mtDNA and nuclear DNA sequences, further confirming the existence of high levels of geographically structured genetic diversity in *A. obstetricans* in Iberia and highlighting the existence of at least five well-differentiated historical lineages. The most outstanding case is the *A. o. almogavarii* lineage, clearly the most differentiated in all analyses. Combined with previous mtDNA and nuclear DNA sequence data (Gonçalves *et al.* 2015), and further considering the low levels of admixture observed with geographically close populations of other lineages, the new evidence from fast-evolving markers supports the genetic distinctiveness of *A. o. almogavarii*, already apparent in the allozyme analyses of Arntzen and García-París (1995), suggesting the need to review its specific taxonomic status as a candidate species.

The spatial clustering analysis performed with TESS and the DAPC analysis provide evidence for further population subdivision, recovering up to seven genetic clusters (Fig. 2a). These results are also concordant with the mitochondrial DNA-diagnosed population lineages in the recent phylogeographic study by Gonçalves *et al.* (2015). More specifically, mtDNA

haplogroups A, B, C, D and F (Gonçalves *et al.* 2015) have a nuclear counterpart based on microsatellite genotypes, with mtDNA haplogroups B and F showing further subdivision (microsatellite clusters B1n and B2n, F1n and F2n). Indeed, this subdivision was already detected in the phylogenetic tree based on DNA sequence data of Gonçalves *et al.* (2015), but with low support. The only exception is mtDNA haplogroup E that is here recovered as part of cluster F1n. Taken together with the previous phylogeographic data, our microsatellite results provide compelling evidence for the presence of distinct genetic groups within *A. obstetricans* occupying distinct geographical regions of the contemporary landscape.

Disentangling the relative roles of geographical and ecological factors as drivers of intraspecific diversification remains a challenge, but is of central importance in hypotheses explaining biogeographic patterns and speciation processes (Nosil 2012, Pinho and Faria 2016, Warren *et al.* 2014). The concordance between different genetic markers and the low levels of gene flow detected between clusters may be a consequence of random demographic events, or may be associated with some degree of separation of the different lineages in environmental space, producing a signature of niche divergence between lineages, as detected in our study. Our results show that different sets of environmental variables are important to predict suitable areas for each genetic cluster and highlight limited geographic overlap between suitable areas (Fig. 3 and Table 4). The highest degree of observed overlap occurs between clusters Cn - Dn. Both clusters are mostly distributed in areas with similar environmental conditions characterized by an Atlantic climate. In fact, precipitation (bio15 and bio16) is the most important EGV for explaining the predicted distribution of clusters Cn and Dn. The lack of individuals in the Cn cluster in the south-west of the study area may be related with historical factors, such as the formation of the current Douro River basin, which has probably acted as a dispersal barrier, preventing further range expansion and admixture across lineages. But generally, we observed a clear signature of niche divergence between all pairs of clusters (Tables 4 and 5). Overall, patterns of environmental suitability for the different clusters have high explanatory power to locate range boundaries as defined by molecular markers. Thus, these results additionally support the role of ecological barriers in preventing gene flow and in the maintenance of lineage boundaries and their evolutionary trajectories. This observation can be further verified in genome-wide studies focusing on identifying adaptive genetic variants, and if confirmed, it would stress the importance of divergent natural selection at early stages of diversification in the speciation process (Rundell and Price 2009, Schluter and Conte 2009, Nosil 2012).

Contrasting with the niche conservatism concept, according to which intraspecific lineages present a tendency to respond in an ecologically similar way to environmental perturbations (Stewart *et al.* 2010), our results provide strong evidence for ecological niche divergence among the distinct population clusters in *A. obstetricans*. Nevertheless, in addition

to climatic variables, observed patterns of population structure probably result from the action of other factors, including demographic, genetic (genetic drift), and ecological (interspecific competition). Some studies have suggested that closely-related taxa are ecologically too similar to allow co-existence in sympatry, and thus competition is the major force responsible for the establishment of allopatric patterns, resulting, for instance, from post glacial population expansions and secondary contact (Rundle and Nosil 2005). However, this model would only explain the strong genetic structure observed in *A. obstetricans,* but not the clear signals of niche divergence.

On the contrary, some models have postulated that recently isolated populations can adapt to different habitats (Mayr 1942, 1947, 1963). In line with this view, several studies have argued that in a population expansion scenario, after consecutive periods of contraction/expansion, different lineages might come into contact at ecotones forming abrupt barriers, remaining in the habitats where they are best adapted (see Zink 2014). In this model, allopatry would result mainly from ecological divergence, possibly occurring at the beginning of the lineage diversification process (at the intraspecific level). In this sense, our results seem to corroborate the model proposed by Zink (2014). Together with the ENMs, the PCA-env analyses point towards an absence of niche overlap in areas where different genetic groups co-occur (for example haplogroups A and B in NW Iberia) and the niche similarity tests show that the populations of a given cluster are not ecologically interchangeable.

To our knowledge, apart from the work by Tarroso *et al.* (2014), who firstly incorporated cluster membership probabilities to build ecological niche-based models in order to test for the contribution of ecological divergence to the establishment of species boundaries in a contact zone, this is a pioneer study in incorporating genetic data directly into the modelling procedure to explore spatial patterns of intraspecific genetic variation. In their recent perspective, Gotelli and Stanton-Geddes (2015) argue that standard species distribution modelling methods applied to presence-only species occurrence records cannot easily accommodate the possibility of local adaptation and evolutionary change within different parts of a species' geographical range, and they suggest the incorporation of intraspecific genetic variation to build species distribution models. Our results show the benefits of this approach and provide evidence for niche evolution at the intraspecific level for *A. obstetricans,* suggesting that the different lineages may have adapted independently to different ecological niches across its current range. This study opens new lines of research focused on identifying regions of the genome associated with transitions in climatic and genetic boundaries. This information is critical to build up on our knowledge about the process of species formation. Furthermore, given the negative prospects for the long-term survival of the species, with projected contraction of its range of up to 88% in the period 2041-2070 (Araujo *et al.* 2011), it will also provide useful information to guide conservation efforts.

4.2.5 Acknowledgements

This work was supported through Project ‘‘Genomics and Evolutionary Biology” co-financed by North Portugal Regional Operational Programme 2007/2013 (ON. 2 – O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF); by FEDER funds through the Operational Programme for Competitiveness Factors – COMPETE and by National Funds through FCT – Foundation for Science and Technology PTDC/BIA-BEC/099915/2008 and FCOMP-01-0124-FEDER-008915 to HG. Partial funds were also provided by grants CGL2008-04271-C02-01/ BOS and CGL2011-28300 (Ministerio de Ciencia e Innovación, Ministerio de Economía y Competitividad, Spain, and FEDER) and PPII10-0097-4200 (Junta de Comunidades de Castilla la Mancha and FEDER) to IMS, who was funded by Project ‘‘Biodiversity, Ecology and Global Change”, co-financed by ON. 2–O Novo Norte, under the NSRF, through the ERDF and the Spanish Severo Ochoa Program (SEV-2012-0262). HG is supported by a postdoctoral Grant from FCT (SFRH/BPD/102966/2014). BMC was supported by a PhD Grant from FCT (SFRH/BD/60305/2009). FS is supported by a postdoctoral Grant from FCT (SFRH/BPD/87721/2012).

4.2.6 References

Abellán P, Svenning JC (2014) Refugia within refugia – patterns in endemism and genetic divergence are linked to Late Quaternary climate stability in the Iberian Peninsula. *Biological Journal of the Linnean Society,* 113: 13–28.

Alberdi A, Gilbert MTP, Razgour O, Aizpurua O, Aihartza J, Garin I (2015) Contrasting population-level responses to Pleistocene climatic oscillations in an alpine bat revealed by complete mitochondrial genomes and evolutionary history inference. *Journal of Biogeography*, 42: 1689–1700.

Allentoft ME, O’Brien J (2010) Global amphibian declines, loss of genetic diversity and fitness: a review. *Diversity*, 2: 47–71.

Araújo MB, Rozenfeld R, Rahbek C, Marquet PA (2011) Using species co-occurrence networks to assess the impacts of climate change. *Ecography*, 34: 897-908.

Arntzen JW, García-París M (1997) Phylogeny and biogeography of midwife toads (*Alytes*, Discoglossidae): a rebuttal. *Contributions to Zoology*, 66: 263–268.

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

Bicheron P, Defourny P, Brockmann C, Schouten L, Vancutsem C, Huc M, Bontemps S, Leroy M, Achard F, Herold M, Ranera F, Arino O (2008) GLOBCOV- ER: Products description and validation report. Medias- France and Postel, http://postel.mediasfrance.org.

Bohonak AJ (2002) IBD (Isolation by Distance): A program for analyses of isolation by distance. *Journal of Heredity*, 93: 153-154.

Broennimann O, Fitzpatrick MC, Pearman PB, Petitpierre B, Pellissier L, Yoccoz NG, Thuiller W, Fortin M, Randin C, Zimmermann NE, Graham CH, Guisan A (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, 21: 481–497.

Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL (2008) Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 105: 11505–11511.

Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes,* 7: 747–756.

D'Amen M, Zimmermann NE, Pearman PB (2013) Conservation of phylogeographic lineages under climate change. *Global Ecology and Biogeography*, 22: 93–104.

Díaz-Rodríguez J, Gonçalves H, Sequeira F, Sousa-Neves T, Tejedo M, Ferrand N, Martínez-Solano I (2015) Molecular evidence for cryptic candidate species in Iberian *Pelodytes* (Anura, Pelodytidae). *Molecular Phylogenetics and Evolution,* 83: 224-241.

Dufresnes C, Bonato L, Novarini N, Betto-Colliard C, Perrin N, Stöck M (2014) Inferring the degree of incipient speciation in secondary contact zones of closely related lineages of Palearctic green toads (*Bufo viridis* subgroup). *Heredity*, 113, 9 – 20.

Dufresnes C, Wassef J, Ghali K, Brelsford A, Stöck M, Lymberakis P, Perrin N (2013) Conservation phylogeography: Does historical diversity contribute to regional vulnerability in European tree frogs (*Hyla arborea*)? *Molecular Ecology*, 22: 5669 – 5684.

Dyer RJ, Nason JD, Garrick RC (2010) Landscape modeling of gene flow: improved power using conditional genetic distance derived from the topology of population networks. *Molecular Ecology*, 19: 3746–3759.

Earl DA, Von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4: 359–361.

ESRI (2011) ArcGIS Desktop, Release 10. Environmental Systems Research Institute Inc., Redlands, CA, USA.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14: 2611–2620.

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

Ferchaud AL, Eudeline R, Arnal V, Cheylan M, Pottier G, Leblois R, Crochet PA (2015) Congruent signals of population history but radically different patterns of genetic diversity between mitochondrial and nuclear markers in a mountain lizard. *Molecular Ecology*, 24: 192-207.

Frankham R (2005) Genetics and extinction. *Biological Conservation*, 126: 131–140.

Freedman AH, Thomassen HA, Buermann W, Smith TB (2010) Genomic signals of diversification along ecological gradients in a tropical lizard. *Molecular Ecology*, 19: 3773–88.

Freeland JR, Biss P, Conrad KF, Silvertown J (2010) Selection pressures have caused genome-wide population differentiation of *Anthoxanthum odoratum* despite the potential for high gene flow. *Journal of Evolutionary Biology*, 23: 776–782.

Garant D, Kruuk LEB, Wilkin TA, McCleery RH, Sheldon BC (2005) Evolution driven by differential dispersal within a wild bird population. *Nature*, 433: 60–65.

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss, S. & Ferrand, N. (eds) Phylogeography of Southern European Refugia: Evolutionary Perspectives on the Origin and Conservation of European Biodiversity, pp. 155-187, Springer, Dordrecht, The Netherlands.

Gómez-Rodríguez C, Baselga A, Wiens JJ (2015) Is diversification rate related to climatic niche width? *Global Ecology and Biogeography*, 24: 383–395.

Gonçalves H, Maia-Carvalho B, Sousa-Neves T, García-París M, Sequeira F, Ferrand N, Martínez-Solano I (2015) Multilocus phylogeography of the common midwife toad, *Alytes obstetricans* (Anura, Alytidae): contrasting patterns of lineage diversification and genetic structure in the Iberian refugium. *Molecular Phylogenetics and Evolution*, 93: 363–379.

Gonçalves H, Martínez-Solano I, Ferrand N, García-París M (2007) Conflicting phylogenetic signal of nuclear vs mitochondrial DNA markers in midwife toads (Anura, Discoglossidae, *Alytes*): deep coalescence or ancestral hybridization? *Molecular Phylogenetics and Evolution*, 44: 494– 500.

Gotelli NJ, Stanton-Geddes J (2015) Climate change, genetic markers and species distribution modelling. *Journal of Biogeography*, 42: 1577–1585.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www2.unil.ch/popgen/softwares/fstat.htm.

Guisan A, Thuiller W (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8: 993–1009.

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, Biological Sciences*, 359: 183–195.

Hewitt GM (2011) Quaternary phylogeography: The roots of hybrid zones. *Genetica*, 139: 617–638.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25: 1965–1978.

Hughes PD, Woodward JC (2008) Timing of glaciation in the Mediterranean mountains during the last cold stage. *Journal of Quaternary Science*, 23: 575–588.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23: 1801–1806.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24: 1403–1405.

Kozak KH, Wiens JJ (2010) Niche conservatism drives elevational diversity patterns in Appalachian salamanders. *American Naturalist*, 176: 40–54.

Kozak KH, Graham CH, Wiens JJ (2008) Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology & Evolution*, 23: 141– 148.

Lee-Yaw JA, Irwin DE (2015) The importance (or lack thereof) of niche divergence to the maintenance of a northern species complex: the case of the long-toed salamander (*Ambystoma macrodactylum* Baird). *Journal of Evolutionary Biology*, 28: 917-30.

Maia-Carvalho B, Gonçalves H, Ferrand N, Martínez-Solano I (2014b) Multilocus assessment of phylogenetic relationships in *Alytes* (Anura, Alytidae). *Molecular Phylogenetics and Evolution*, 79: 270-278.

Maia-Carvalho B, Gonçalves H, Martínez-Solano I, Gutiérrez-Rodríguez J, Lopes S, Ferrand N, Sequeira F (2014a) Intraspecific genetic variation in the common midwife toad (*Alytes obstetricans*): subspecies assignment using mitochondrial and microsatellite markers. *Journal of Zoological Systematics and Evolutionary Research*, 52: 170-175.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27: 209–220.

Márquez R, Rosa HD (1997) *Alytes obstetricans* (Laurenti, 1768). *Distribución y Biogeografía de los Anfibios y Reptiles en España y Portugal* (ed. JM Pleguezuelos), pp. 131–133. Universidad de Granada, Asociación Herpetológica Española, Granada.

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.

Mayr E (1942) Systematics and the origin of species. New York, NY: Columbia University Press.

Mayr E (1947) Ecological factors in speciation. *Evolution*, 1: 263–288.

Mayr E (1963) Animal species and evolution. Harvard University Press, Cambridge, MA.

McCormack JE, Zellmer AJ, Knowles LL (2010) Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation? Insights from tests with niche models. *Evolution*, 64: 1231-44.

Meirmans PG (2012) The trouble with isolation by distance. *Molecular Ecology,* 21**:** 2839-2846.

Morales A, Villalobos F, Velazco PM, Simmons NB, Piñero D (2016) Environmental niche drives genetic and morphometric structure in a widespread bat. *Journal of Biogeography*, 43: 1057–1068.

Moritz C (1994) Defining “Evolutionarily Significant Units” for conservation. *Trends in Ecology & Evolution*, 9: 373-375.

Nosil P (2012) Ecological Speciation. Oxford University Press, New York, NY.

Olden JD, Lawler JJ, Poff NL (2008) Machine learning methods without tears: a primer for ecologists. *The Quarterly Review of Biology*, 83: 171–193.

Olson DM, Dinerstein E, Wikramanayake E, Burgess ND, Powell GVN, Underwood EC, D’Amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR (2001). Terrestrial ecoregions of the world: a new map of life on Earth. *BioScience*, 51: 933–938.

Ortego J, Riordan EC, Gugger PF, Sork VL (2012) Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. *Molecular Ecology*, 21: 3210-23.

Pearman P, Guisan A, Broennimann O, Randin C (2008) Niche dynamics in space and time. *Trends in Ecology & Evolution*, 23: 149–158.

Pinho C, Faria R (2016) Magadi tilapia ecological specialization: filling the early gap in the speciation continuum. *Molecular Ecology*, 25: 1420–1422.

Pinho C, Kaliontzopoulou A, Harris DJ, Ferrand N (2011) Recent evolutionary history of the Iberian endemic lizards *Podarcis bocagei* (Seoane 1884) and *P. carbonelli* Pérez-Mellado 1981 (Squamata: Lacertidae) revealed by allozyme and microsatellite markers. *Zoological Journal of the Linnean Society*, 162: 184-200.

Postma E, Van Noordwijk AJ (2005) Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature*, 433: 65–68.

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

Pyron RA, Costa GC, Patten MA, Burbrink F (2014) Phylogenetic niche conservatism and the evolutionary basis of ecological speciation. *Biological reviews of the Cambridge Philosophical Society*, 90: 1248-62.

Rato C, Harris DJ, Perera A, Carvalho SB, Carretero MA, Rödder D (2015) A combination of divergence and conservatism in the niche evolution of the Moorish Gecko, *Tarentola mauritanica* (Gekkota: Phyllodactylidae). *PLoS One*, 10: e0127980.

Recuero E, García-París M (2011) Evolutionary history of *Lissotriton helveticus*: Multilocus assessment of ancestral vs. recent colonization of the Iberian Peninsula. *Molecular Phylogenetics and Evolution*, 60: 170–182.

Rödder D, Engler JO (2011) Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks. *Global Ecology and Biogeography*, 20: 915–927.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, 4: 137–138.

Rundell RJ, Price TD (2009) Adaptive radiation, nonadoptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology & Evolution*, 24: 394–399.

Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, 8: 336–352.

Sacks BN, Brown SK, Ernest HB (2004) Population structure of California coyotes corresponds to habitat-specific breaks and illuminates species history. *Molecular Ecology*, 13: 1265–1275.

Schluter D (2009) Evidence for ecological speciation and its alternative. *Science*, 323: 737–741.

Schoville SD, Bonin A, François O, Lobreaux S, Melodelima C, Manel S (2012) Adaptive genetic variation on the landscape: methods and cases. *Annual Review of Ecology, Evolution, and Systematics*, 43: 23–43.

Sequeira F, Alexandrino J, Weiss S, Ferrand N (2008) Documenting the advantages and limitations of different classes of molecular markers in a well-established phylogeographic context: lessons from the Iberian endemic Golden-striped salamander, *Chioglossa lusitanica* (Caudata: Salamandridae). *Biological Journal of the Linnean Society*, 95: 371 – 387.

Serra-Varela MJ, Grivet D, Vincenot L, Broennimann O, Gonzalo-Jiménez J, Zimmermann NE (2015) Does phylogeographical structure relate to climatic niche divergence? A test using maritime pine (*Pinus pinaster* Ait.). *Global Ecology and Biogeography*, 24: 1302–1313.

Sillero N, Brito JC, Skidmore AK, Toxopeus AG (2009) Biogeographical patterns derived from remote sensing variables: the amphibians and reptiles of the Iberian Peninsula. *Amphibia-Reptilia*, 30: 185–206.

Stewart JR (2009) The evolutionary consequence of the individualistic response to climate change. *Journal of Evolutionary Biology*, 22: 2363–2375.

Stewart JR, Lister AM, Barnes I, Dalen L (2010) Refugia revisited: individualistic responses of species in space and time. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 277: 661–671.

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

Tarroso P, Carvalho SBS, Brito JC (2012) Simapse – simulation maps for ecological niche modelling. *Methods in Ecology and Evolution*, 3, 787–791.

Tarroso P, Pereira RJ, Martínez-Freiría F, Godinho R, Brito JC (2014) Hybridization at an ecotone: ecological and genetic barriers between three Iberian vipers. *Molecular Ecology*, 23: 1108–23.

Thomassen HA, Cheviron ZA, Freedman AH, Harrigan RJ, Wayne RK, Smith TB (2010) Spatial modeling and landscape- level approaches for visualizing intra-specific variation. *Molecular Ecology*, 19: 3532–3548.

Tocchio LJ, Gurgel-Gonçalves R, Escobar LE, Peterson AT (2015) Niche similarities among white-eared opossums (Mammalia, Didelphidae): Is ecological niche modelling relevant to setting species limits? *Zoologica Scripta*, 44: 1–10.

Trabucco A, Zomer RJ (2009). Global Aridity Index (Global-Aridity) and Global Potential Evapo-Transpiration (Global-PET) Geospatial Database. CGIAR Consortium for Spatial Information. Published online, available from the CGIAR-CSI GeoPortal at: http://www.csi.cgiar.org.

Warren DL, Cardillo M, Rosauer DF, Bolnick DI (2014) Mistaking geography for biology: inferring processes from species distributions. *Trends in Ecology and Evolution*, 29: 572–580.

Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, 62: 2868-83.

Wiens J (2004) Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. *Evolution*, 58:193–197.

Wiens JJ, Graham CH (2005). Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, 36: 519–539.

Wood D, Fisher RN, Vandergast AG (2014). Fuzzy boundaries: Color and gene flow patterns among parapatric lineages of the Western shovel-nosed snake and taxonomic implication. *PLoS One*, 9: 16–21.

Zink RM (2014) Homage to Hutchinson, and the role of ecology in lineage divergence and speciation. *Journal of Biogeography*, 41: 999–1006.

4.2.6 Supporting Information

4.2.6.1 **Appendix S1**. Details of genetic and spatial analyses.

4.2.6.1.1 *Genetic approach*

Bayesian multilocus assignment methods

Analyses with STRUCTURE were performed using the admixture model with correlated allele frequencies, assuming no a priori information about the origin of individuals. Ten independent simulations were run with 1,000,000 Markov chain Monte Carlo (MCMC) iterations, discarding the first 500,000 replicates as burn-in, and assuming a range of clusters (*K*) from 1 to 30. The optimal *K* value was determined using STRUCTURE HARVESTER (Earl and von Holdt, 2012) based on graphical files representing ln Pr (*X/K*) (Pritchard *et al*. 2000) and the ΔK approach (Evanno *et al*. 2005) per *K* and per repeated run. TESS analyses were carried out by modelling admixture using the conditional autoregressive (CAR) model with the following conditions: 180,000 runs; 120,000 burn-in and 10 replicates for each *K* value tested (from 2 to 17). The spatial interaction parameter (ψ) was set to 0.6 and the option to update this parameter was activated. The best *K* was determined by plotting DIC (Deviation information Criterion) against *K* and choosing the values of *K* that correspond to a plateau of the curve (Durand *et al*. 2009). The software CLUMPP v1.1.2 (Jakobsson and Rosenberg, 2007) was used to estimate average matrices of proportion of ancestry membership (admixture *Q* matrix output) over the 10 replicates of the optimal *K* value obtained with both STRUCTURE and TESS. Finally, the program DISTRUCT (Rosenberg, 2002) was used to display individual membership coefficients for each cluster.

Discriminant analysis of Principal Components (DAPC)

ADEGENET package v1.3 (Jombart 2010) was used to plot clusters of genotypes from the DAPC analysis in *R* version 3.1.1 (www.r-project.org). This approach overlooks within-groups variation and summarizes the amount of between-groups variation, making this method an adequate option for assessing relationships between populations. Clustering solutions for different *K* values were compared by calculating Bayesian Information Criterion (BIC). Following Jombart *et al*. (2010) the optimal *K* value corresponds to the lowest value of BIC.

4.2.6.1.2 *Spatial approach*

Ecological niche-based analyses

We built 25 model replicates for each cluster to yield a consensus model with a standard deviation. Each replicate was based on a network with a single hidden layer with five nodes; the learning rate was set to 0.01 and the momentum to 0.1, with 2,500 iterations. From the 132 available samples, 25% were randomly selected for testing the network and the iteration with the best fit, while avoiding overfitting (Tarroso *et al*. 2014). Overall prediction ability of the models was tested by comparing membership probabilities with model predicted values using linear regressions, and the significance of the predictor was tested with ANOVA (Tarroso *et al*. 2014).

Average continuous model outputs were converted to binary presence–absence maps by defining thresholds that force all populations with a membership probability to a given cluster (> 0.75) to be considered as predicted. Populations were first assigned to a given cluster based on the membership probability results from STRUCTURE (> 0.75). Then, population sample locations were intersected with the average probability models for each cluster and the minimum probability value was taken as threshold. Additionally and to elucidate how the full species niche is represented in the study area, the full distribution of *Alytes obstetricans* was modeled in the Iberian Peninsula. The same SIMAPSE parameters were applied to the same set of EGVs and the 132 non-clustered population occurrences to build the model.

Niche analyses

For the niche equivalency test, all occurrences of two clusters were pooled and randomly split into two datasets. This procedure was repeated 100 times and the niche overlap statistic *D* was calculated. The simulated values were then used to construct 95% confidence interval. The null hypothesis of niche equivalency is rejected if the observed value of niche overlap falls outside this confidence interval. For the niche similarity test, niche overlaps are computed based on the environmental conditions of one cluster and records randomly generated in the second cluster environmental space. This is performed in both directions assessing whether the observed niche overlap is due to the available environmental spaces. The null hypothesis is rejected, if the observed overlap is greater or smaller than 95% of the simulated values. This indicates that niche differentiation between clusters is due to habitat selection and not an artefact related to differences of the underlying environment.

4.2.6.2 **Appendix S2**. Supplementary tables.

Table S1. Detail of sampling locations used in the present study. For Map Number see Figure 1. Coordinates are presented in WGS84. N: sample size. Cluster: based on microsatellites analysis (cut off ≥ 0.8).

Map Number	Name of Locality	Country	Code	Cordinates (X/Y)		N	Cluster
1	S.S.Mamede	Portugal	SMA	39.3167	-7.4167	17	D
2	Nazaré	Portugal	NAZ	39.5958	-9.0203	20	D
3	Coimbra	Portugal	COM	40.1106	-8.3786	2	D
4	Serra da estrela	Portugal	EST	40.3333	-7.6167	13	D
5	S.Malcata	Portugal	MAL	40.3056	-7.0778	15	D
6	Cáceres	Portugal	CAC	40.2333	-6.6000	2	D
7	Sierra Gata	Portugal	SGA	40.3614	-6.5361	3	D
8	Server do Vouga	Portugal	VOU	40.7333	-8.4000	6	**
9	Montemuro	Portugal	MON	41.0783	-8.0242	6	**
10	Valongo	Portugal	VAL	41.2400	-8.5247	14	C
11	Lousada	Portugal	LOU	41.3028	-8.2528	5	C
12	Amarante	Portugal	AMA	41.2789	-8.0650	14	C
13	Vila Real	Portugal	VRE	41.2902	-7.8391	2	C
14	Alvão	Portugal	ALV	41.4667	-7.8333	1	**
15	Pontido	Portugal	APO	41.4614	-7.6826	4	C
16	Murça	Portugal	MUR	41.4833	-7.4333	2	C
17	Mirandela	Portugal	MIR	41.4833	-7.1667	1	C
18	Mogadouro	Portugal	MOG	41.3333	-6.7167	14	C
19	Montesinho	Portugal	MTO	41.7833	-6.5833	15	C
20	Ane	Portugal	ANE	41.7130	-8.0360	1	C
21	Ventuzelo	Portugal	VEI	41.7001	-8.1675	2	C
22	Carris	Portugal	CRR	41.8167	-8.0500	2	C
23	Ponte Vedra	Spain	PON	42.0812	-8.6865	5	C
24	Verin	Spain	VER	41.9833	-7.4167	2	C
25	Ourense	Spain	COR	42.4250	-7.7333	3	C
26	Arteixo	Spain	ART	43.3062	-8.5251	6	C
27	Monfero	Spain	MOF	43.3325	-8.0570	1	**
28	Alfonxe	Spain	ALF	42.9240	-7.3890	2	C
29	Fontaneira	Spain	FTN	43.0333	-7.2000	2	C
30	Tineo	Spain	TIN	43.3333	-6.4167	1	B1
31	San Isidro	Spain	SIS	43.3826	-5.8352	3	B1
32	Rozada	Spain	ROZ	43.2684	-5.6399	1	**
33	Tolivia	Spain	TLV	43.2000	-5.5833	1	**
34	Pelúgano	Spain	PEL	43.1443	-5.5843	2	B1
35	Xerra	Spain	XER	43.2542	-5.5036	30	B1
36	Isoba	Spain	ISO	43.0460	-5.3151	4	B1
37	Llomena	Spain	LLO	43.1856	-5.1235	2	B1
38	Ribadesella	Spain	RIB	43.4333	-5.1500	1	B1
39	Muñegru	Spain	MUN	43.2583	-4.9576	2	B1
40	Cimeru	Spain	CIM	43.1722	-4.8676	2	B1

Map Number	Name of Locality	Country	Code	Cordinates (X/Y)	N	Cluster	
41	Liordes	Spain	LIO	43.1441	-4.8197	2	B1
42	VegasToro	Spain	VET	43.2164	-4.7578	1	**
43	Andara	Spain	AND	43.2130	-4.6525	2	**
44	Hoyos de Vargas S	Spain	HVS	43.0198	-4.7269	2	**
45	Hoyos de Vargas I	Spain	HVI	43.0123	-4.7438	2	**
46	Sopenilla	Spain	SOP	43.2818	-4.4261	1	**
47	Tudanca	Spain	TUD	43.1500	-4.3667	7	**
48	Palencia	Spain	PAL	42.8721	-4.1298	2	A
49	Gamonal	Spain	GAM	43.1755	-3.8857	2	B1
50	Fresnedo	Spain	FRS	43.3667	-3.5667	2	**
51	La Magdalena	Spain	MAG	43.3500	-3.3333	3	**
52	Orduña	Spain	ORD	42.9921	-3.0121	1	**
53	Vizcaya	Spain	VIZ	43.0875	-2.7922	3	**
54	Pla de Beret	France	BET	42.7114	0.9484	18	B2
55	Etang de Lers	France	ETG	42.8073	1.3826	16	B2
56	Saletes	France	SAL	43.9933	2.1669	3	**
57	St Pierre-de-la-Fage	France	PIE	43.7939	3.4200	6	B2
58	Auvergne	France	AUZ	45.4167	3.0000	3	B2
59	Alsace	France	JUB	48.2508	-0.5014	6	B2
60	La Bru	France	LBR	48.3471	6.6775	3	**
61	Germany	Germany	GER	51.5434	7.5863	4	**
62	Pamplona	Spain	PAM	42.9500	-1.2500	5	**
63	Hecho	Spain	HEC	42.7366	-0.7620	15	F1
64	Astun	Spain	AST	42.8022	-0.5182	13	**
65	Santa Cruz de la Seros	Spain	SER	42.5433	-0.6828	10	F1
66	Ibón de Piedrafita	Spain	IBN	42.7000	-0.3333	2	**
67	Sopeira	Spain	SPR	42.3333	-0.7250	3	**
68	Sabiñanigo	Spain	SBG	42.4106	-0.3896	4	F1
69	Fanlo	Spain	BRO	42.5689	-0.0769	8	F1
70	Escalona	Spain	FAN	42.5977	0.1327	9	F1
71	Meson de Puertolas	Spain	MÊS	42.5290	0.1466	21	F1
72	Seira	Spain	SEI	42.4802	0.4291	18	F1
73	Fonchanina	Spain	FON	42.5221	0.6520	13	F1
74	Perves	Spain	PER	42.3565	0.8340	19	**
75	Llesui	Spain	LLS	42.4522	1.0645	17	F2
76	Sant Salvador de Tolo	Spain	TOL	42.0808	1.0321	14	F2
77	Avellanet	Spain	AVE	42.3517	1.3288	21	F2
78	Sanilles	Spain	SAN	42.3703	1.6833	9	F2
79	Soldes	Spain	SOL	42.3500	1.7000	2	F2
80	Busa	Spain	BES	42.0859	1.6079	17	F2
81	Llinars	Spain	LLN	42.1138	1.7091	4	**
82	Barcelona	Spain	BCN	42.1000	1.8500	22	F2
83	Toses	Spain	TOS	42.3174	2.0048	19	F2
84	Sant Marti d"Ogassa	Spain	OGA	42.2692	2.2225	20	F2
85	La Coma	Spain	LAC	42.2385	2.4357	4	F2

86	Sadernes	Spain	ANI	42.2885	2.5892	21	F2
87	Taulis	France	TAU	42.4986	2.6571	21	F2
88	Sant Martin d'Albera	France	ALR	42.4624	2.9150	19	F2
89	Madeloc	France	MDL	42.4998	3.0669	1	F2
90	Sant Hilari Sacalm	Spain	HIL	41.9029	2.5058	10	F2
91	Sant Miquel del Fai	Spain	MIQ	41.7163	2.1927	20	F2
92	Piera	Spain	PIR	41.5106	1.7371	19	F2
93	El Pont d"Armentera	Spain	ARM	41.3869	1.3595	20	F2
94	Riudoms	Spain	RIU	41.1481	1.0429	20	A
95	El Perelló	Spain	PLL	40.8579	0.6697	2	A
96	Caspe	Spain	CAS	41.1483	0.0128	2	A
97	Mazaleón	Spain	MAS	41.0606	0.1103	2	A
98	Barranc dels Horts	Spain	BAR	40.4144	-0.0661	2	A
99	Cantavieja	Spain	CAT	40.5160	-0.4590	3	A
100	Corbalán	Spain	COB	40.4056	-0.9864	2	A
101	Benicassim	Spain	BEM	40.0509	0.0450	3	A
102	Algar de Palancia	Spain	ALG	39.7833	-0.3667	3	A
103	Alcublas	Spain	ALC	39.8147	-0.6997	2	A
104	Las Nogueras	Spain	NOG	39.5881	-1.0819	3	A
105	Enguera	Spain	ENG	38.9289	-0.8550	3	A
106	Mas de Celedons	Spain	CEL	38.6689	-0.5247	3	**
107	Pinoso	Spain	PIN	38.4000	-1.0333	12	A
108	La Fuente	Spain	ALB	39.0028	-1.3040	26	A
109	Buenache de la Sierra	Spain	BUE	40.1333	-2.0000	1	A
110	Guadalaviar	Spain	GUA	40.3625	-1.6702	6	A
111	Terzaga	Spain	TER	40.6472	-1.9072	2	A
112	Molina de Aragón	Spain	MOL	40.8692	-1.9403	2	A
113	Montuenga	Spain	MTG	41.2247	-2.2097	4	A
114	Moros	Spain	MRO	41.4017	-1.8200	2	A
115	Tobed	Spain	TOB	41.3403	-1.3986	3	A
116	Belchite	Spain	BEC	41.2853	-0.7756	3	A
117	Ólvega	Spain	OLV	41.7805	-1.9852	1	**
118	Ribaforada	Spain	RBF	42.0058	-1.5364	1	A
119	Hayedo de Tejera Negra	Spain	HAY	41.2370	-3.3500	2	**
120	Albarracín	Spain	ABR	40.3940	-1.4164	2	A
121	Vinuesa	Spain	VIN	41.9167	-2.7667	4	A
122	Cañamares	Spain	CAN	41.2272	-2.9558	3	A
123	Mirabueno	Spain	MIB	40.9475	-2.7292	2	**
124	Gargoles	Spain	GAG	40.7572	-2.5738	5	A
125	Torija	Spain	TRJ	40.7444	-3.0308	2	A
126	Ciruelas	Spain	CRL	40.6084	-3.1834	4	**
127	Villar del Olmo	Spain	OLM	40.3362	-3.2387	20	A
128	Tielmes	Spain	TLM	40.2333	-3.3167	2	A
129	Colmenar de Oreja	Spain	COL	40.1000	-3.3833	2	**
130	Villarrubia de Santiago	Spain	VRB	39.9895	-3.3467	5	A
131	Ocaña	Spain	OCN	39.9618	-3.5074	5	A

Map Number	Name of Locality	Country	Code	Cordinates (X/Y)	N	Cluster	
132	Noblejas	Spain	NBJ	39.9749	-3.6445	19	A
133	Yepes	Spain	YEP	39.9035	-3.6470	5	A
134	La Guardia	Spain	GAD	39.7834	-3.4696	4	**
135	Pto. Cotos	Spain	COT	40.7333	-4.0000	8	**
136	Circo del Nevero	Spain	CIR	40.9790	-3.8440	2	A
137	Vegas de Matute	Spain	VEG	40.8010	-4.2580	4	**
138	Venta del Obispo	Spain	VEM	40.4128	-4.8704	4	D
139	Mijares	Spain	MJR	40.2921	-4.8342	1	D
140	Cabezuela del Valle	Spain	CBZ	40.2115	-5.7781	1	D
141	Valdefinjas	Spain	VDJ	41.4508	-5.4511	5	**
142	Veguellina de Órbigo	Spain	VGL	42.4500	-5.8850	1	**

**: not attributed

Table S2. Statistics of average genetic measures based in 12 microsatellite loci for 56 populations of *Alytes obstetricans*. Population codes as in Table S1. Map Number as in Figure 1. *Na*: number of alleles sampled; *Ar*: Allelic Richness; *He*: expected heterozygosity.

Populations	Map Number	N_a	A_r	H_e
SMA	1	9.083	3.731	0.696
NAZ	2	9.667	3.872	0.763
EST	4	8.583	3.999	0.785
MAL	5	11.333	4.313	0.815
VOU	8	6	4.051	0.803
MON	9	5.5	3.941	0.764
VAL	10	6.167	3.236	0.646
LOU	11	5.583	3.966	0.739
AMA	12	5.583	3.168	0.649
MOG	18	8.417	3.972	0.781
MOT	19	8.583	3.843	0.743
POV	23	3.75	2.97	0.627
ART	26	4.417	2.991	0.557
XER	35	4.917	2.88	0.621
TUD	47	4.25	2.935	0.608
BET	54	2.917	1.939	0.343
ETG	55	2.917	2.071	0.427
PIE	57	3.667	2.779	0.559
ALS	59	2.917	2.302	0.355
PAM	62	4.833	3.631	0.715

HEC	63	7.917	3.547	0.682
AST	64	7.167	3.526	0.708
SER	65	5.833	3.134	0.563
FAN	69	4.167	3.034	0.597
ESC	70	6.083	3.335	0.595
MES	71	6.833	3.248	0.616
SEI	72	4.333	2.862	0.61
FON	73	5.083	3.029	0.6
PER	74	7.333	3.434	0.661
LLS	75	5.417	2.917	0.582
TOL	76	6.083	3.356	0.674
AVE	77	5.833	2.912	0.588
SAN	78	2.667	1.923	0.334
BES	80	9.917	3.677	0.701
BCN	82	11.917	3.882	0.721
TOS	83	6.25	3.146	0.619
OGA	84	6.667	3.108	0.629
SAD	86	8.25	3.307	0.627
TAU	87	7.833	3.112	0.591
ALR	88	4.583	2.806	0.562
HIL	90	7.583	3.64	0.684
SMA	1	9.083	3.731	0.696
MIQ	91	7.917	3.402	0.665
PIR	92	8.5	3.604	0.715
ARM	93	4.083	2.656	0.553
RIU	94	7.167	3.528	0.715
PIN	107	5.917	3.255	0.666
LAF	108	9	3.808	0.773
GUA	110	5.917	4.049	0.791
GAG	124	5.083	3.907	0.776
OLM	127	5.25	3.146	0.671
VRB	130	3.417	2.917	0.643
OCN	131	3.417	2.907	0.639
NBJ	132	6.833	3.304	0.683
YEP	133	2.417	2.237	0.516
COT	135	5.333	3.454	0.704
VDJ	141	4.833	3.727	0.761

4.2.6.3 **Appendix S3**. Supplementary figures.

Figure S1. Relationship between pairwise F_{st} values and geographic distances between populations in six genetic clusters in *Alytes obstetricans*.

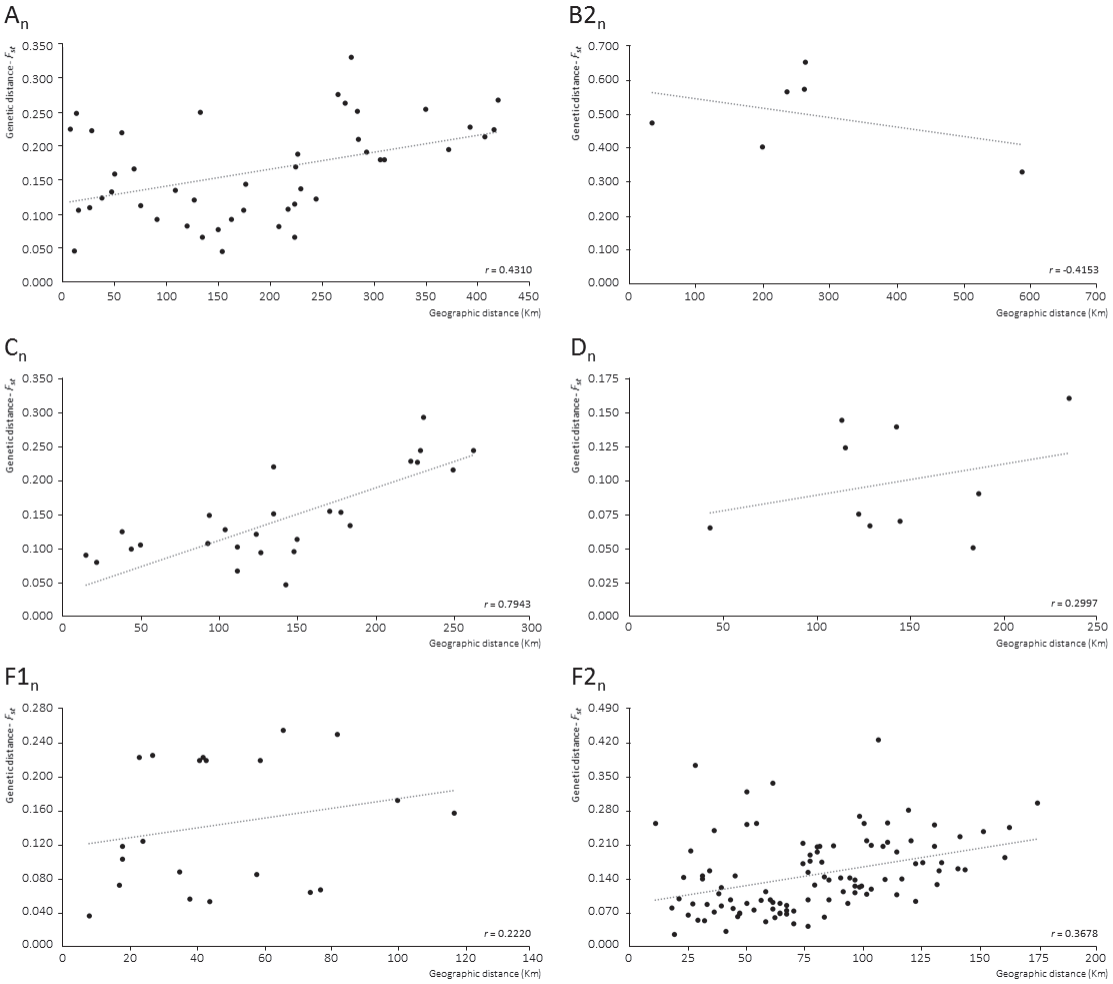


Figure S2. Linear regressions between membership probabilities and values predicted by the ecological niche-based models. Analyses of variance (ANOVA) indicate that the model is a significant predictor of membership probabilities.

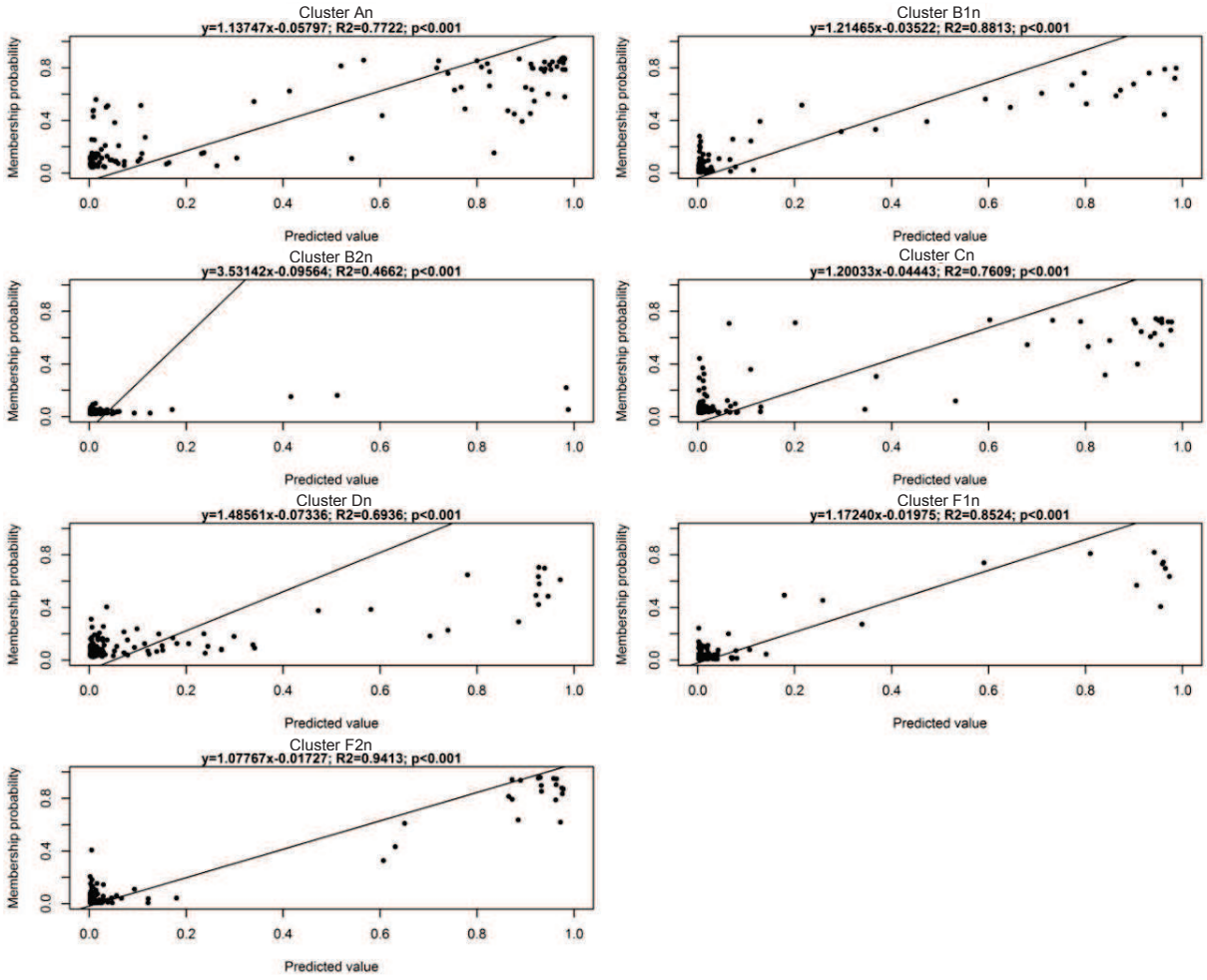


Figure S3. Relative contribution of environmental variables in niche models for *A. obstetricans* and all major genetic clusters.

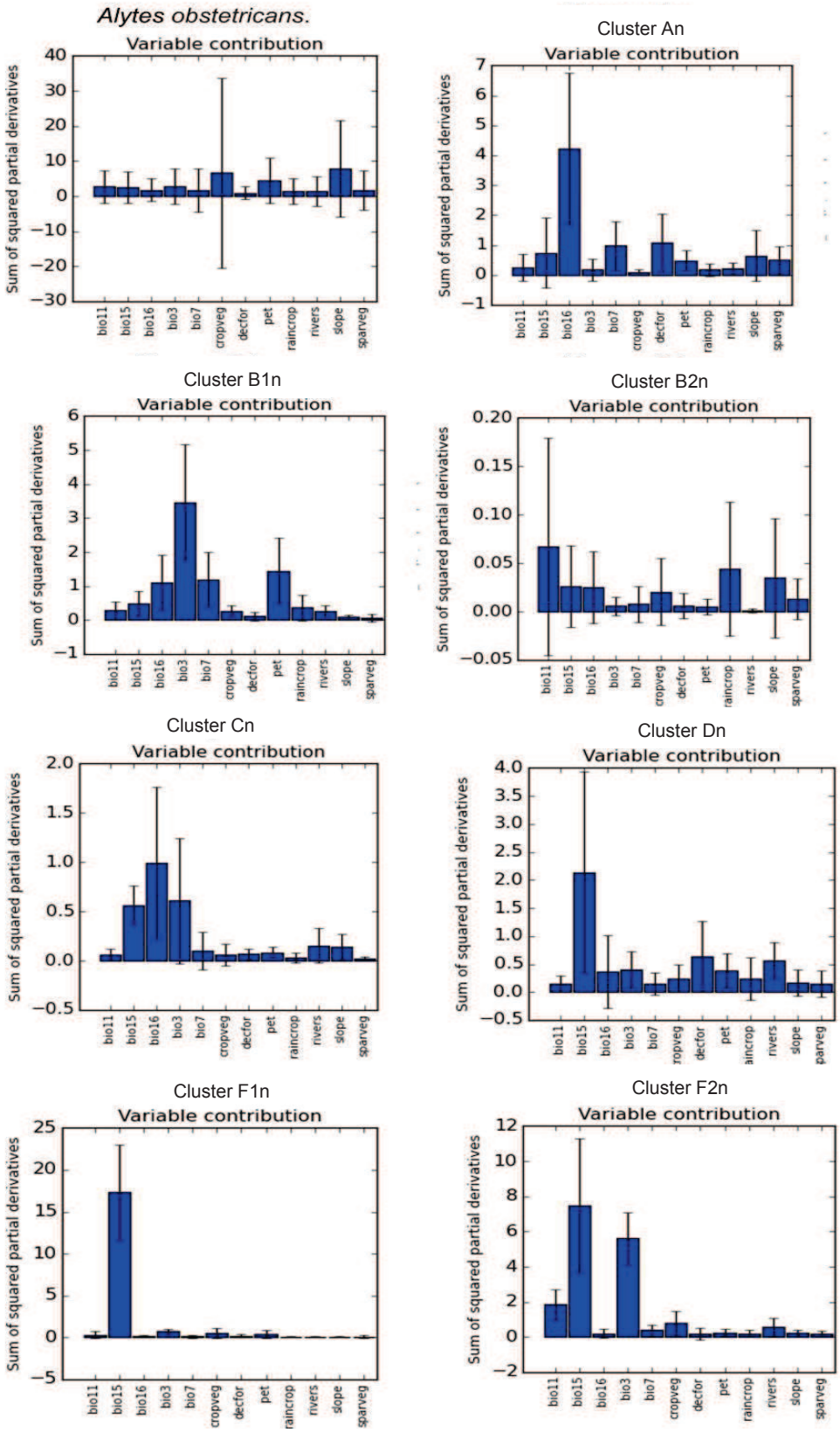
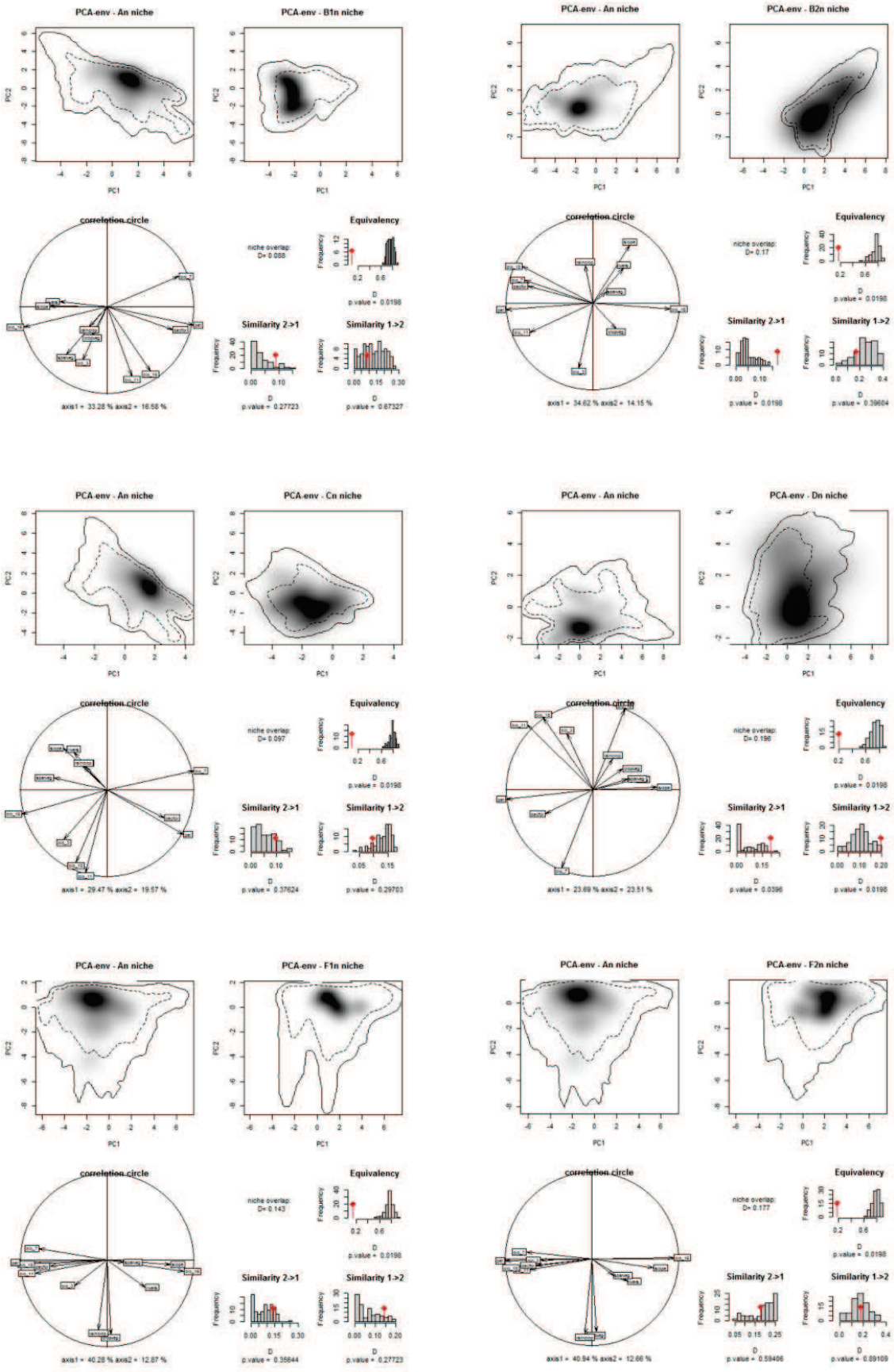
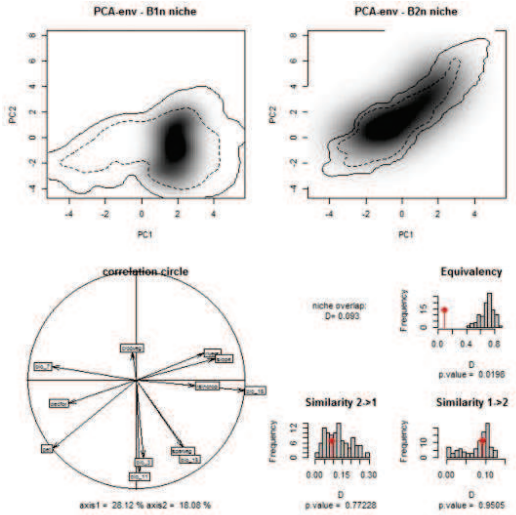
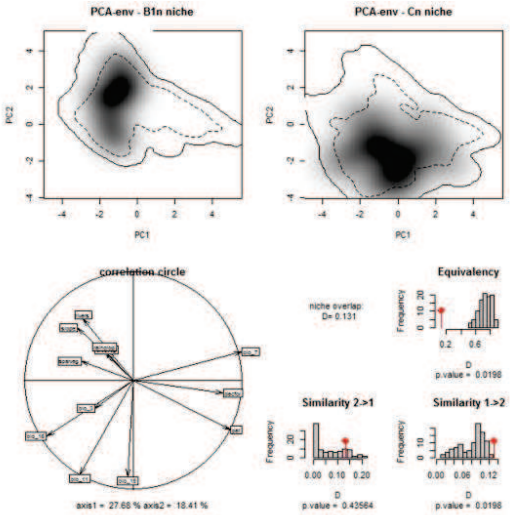
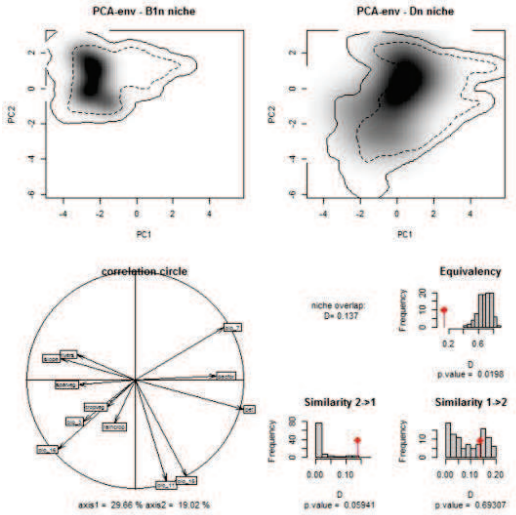


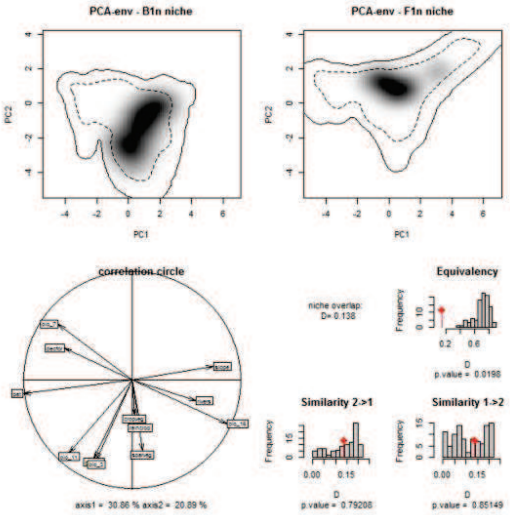
Figure S4. PCA-env and niche overlap results of genetic clusters in *Alytes obstetricans*.

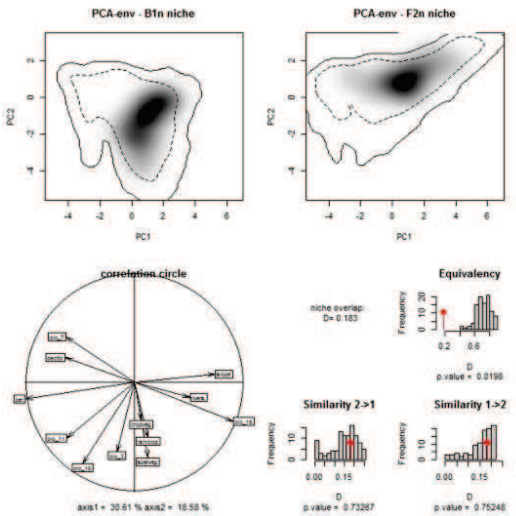


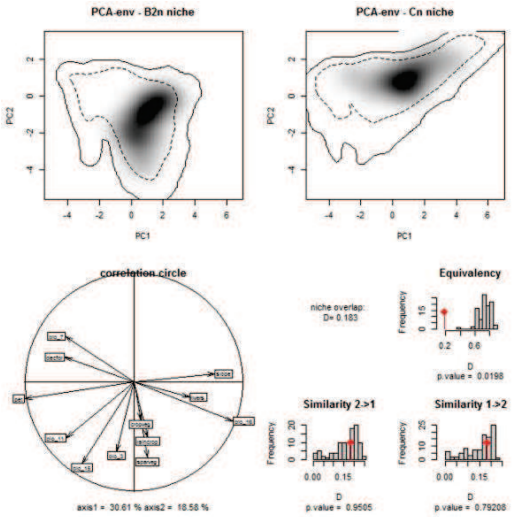

PCA-env - B1n niche
PCA-env - B2n niche
PC1
PC2
correlation circle
Equivalency
niche overlap:
D= 0.093
D
p.value = 0.0198
Similarity 2->1
Similarity 1->2
Frequency
D
p.value = 0.77228
p.value = 0.9505
axis1 = 28.12 % axis2 = 18.08 %

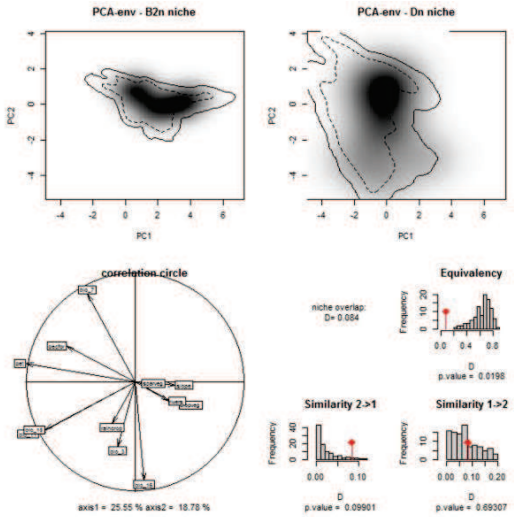

PCA-env - B1n niche
PCA-env - Cn niche
PC1
PC2
correlation circle
Equivalency
niche overlap:
D= 0.131
D
p.value = 0.0198
Similarity 2->1
Similarity 1->2
Frequency
D
p.value = 0.42564
p.value = 0.0198
axis1 = 27.66 % axis2 = 18.41 %

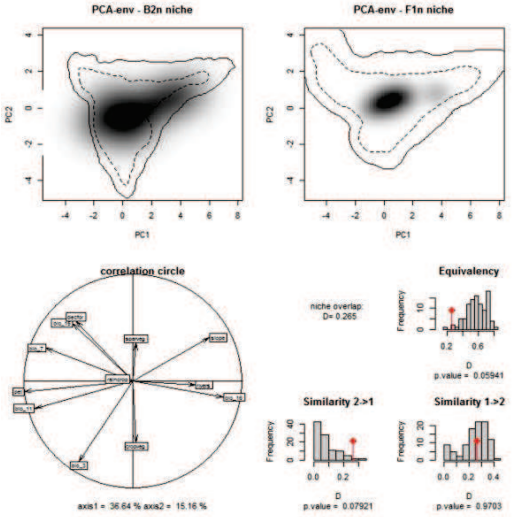

PCA-env - B1n niche
PCA-env - Dn niche
PC1
PC2
correlation circle
Equivalency
niche overlap:
D= 0.137
D
p.value = 0.0198
Similarity 2->1
Similarity 1->2
Frequency
D
p.value = 0.05941
p.value = 0.69307
axis1 = 29.66 % axis2 = 19.02 %

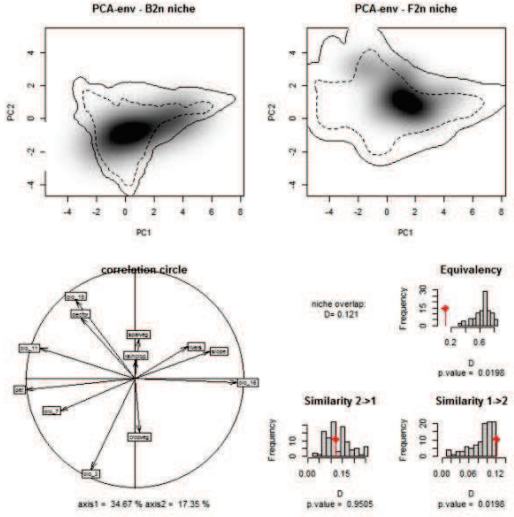

PCA-env - B1n niche
PCA-env - F1n niche
PC1
PC2
correlation circle
Equivalency
niche overlap:
D= 0.138
D
p.value = 0.0198
Similarity 2->1
Similarity 1->2
Frequency
D
p.value = 0.79208
p.value = 0.85149
axis1 = 30.86 % axis2 = 20.89 %

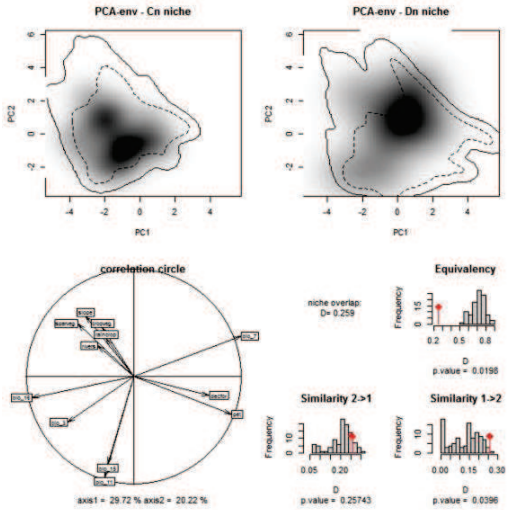

PCA-env - B1n niche
PCA-env - F2n niche
PC1
PC2
correlation circle
Equivalency
niche overlap:
D= 0.183
D
p.value = 0.0198
Similarity 2->1
Similarity 1->2
Frequency
D
p.value = 0.73267
p.value = 0.75248
axis1 = 30.61 % axis2 = 18.58 %

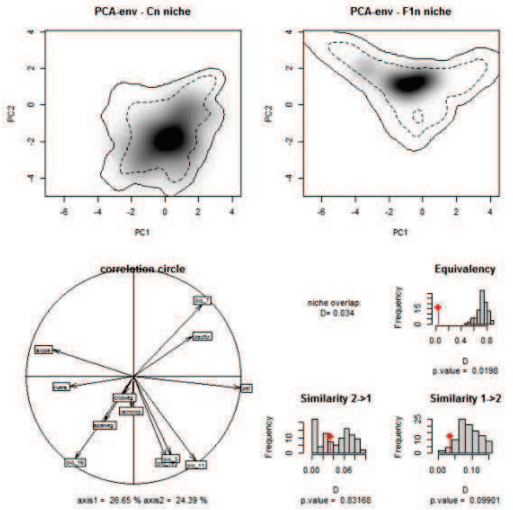

PCA-env - B2n niche
PCA-env - Cn niche
PC1
PC2
correlation circle
axis1 = 30.61 % axis2 = 18.58 %
Equivalency
niche overlap:
D= 0.183
p.value = 0.0198
Similarity 2->1
p.value = 0.9505
Similarity 1->2
p.value = 0.79208
Frequency
D

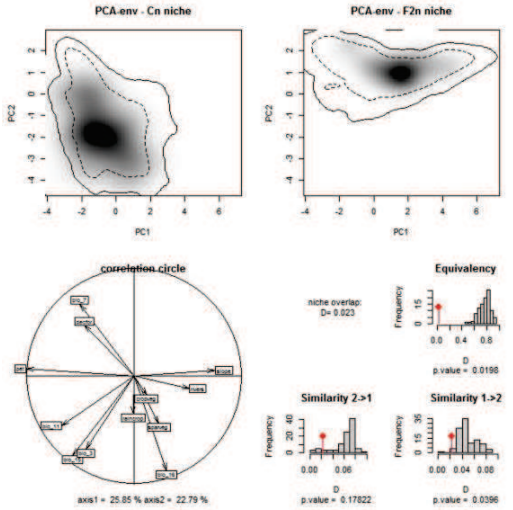

PCA-env - B2n niche
PCA-env - Dn niche
PC1
PC2
correlation circle
axis1 = 25.55 % axis2 = 18.78 %
Equivalency
niche overlap:
D= 0.084
p.value = 0.0198
Similarity 2->1
p.value = 0.09901
Similarity 1->2
p.value = 0.69307
Frequency
D

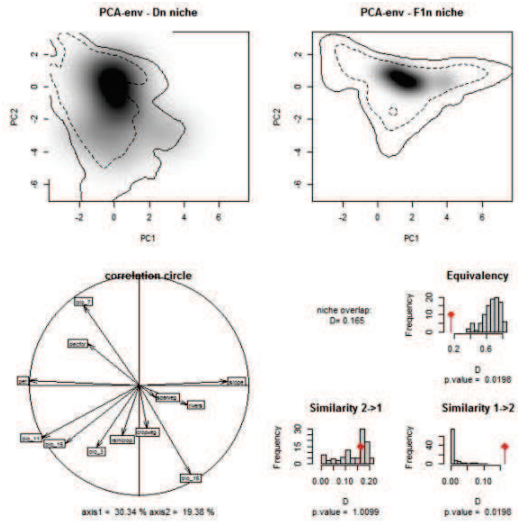

PCA-env - B2n niche
PCA-env - F1n niche
PC1
PC2
correlation circle
axis1 = 36.64 % axis2 = 15.16 %
Equivalency
niche overlap:
D= 0.265
p.value = 0.05941
Similarity 2->1
p.value = 0.07921
Similarity 1->2
p.value = 0.9703
Frequency
D

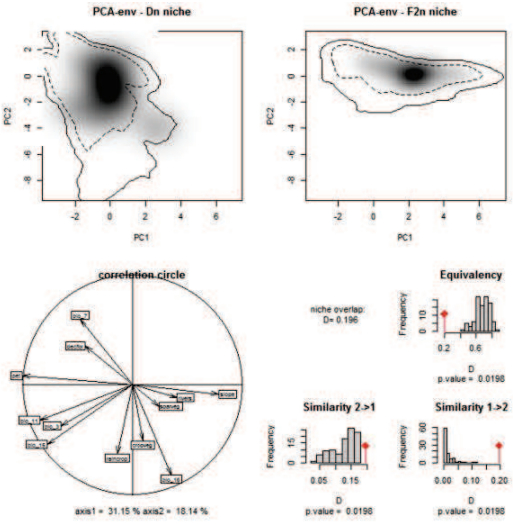

PCA-env - B2n niche
PCA-env - F2n niche
PC1
PC2
correlation circle
axis1 = 34.67 % axis2 = 17.35 %
Equivalency
niche overlap:
D= 0.121
p.value = 0.0198
Similarity 2->1
p.value = 0.9505
Similarity 1->2
p.value = 0.0198
Frequency
D

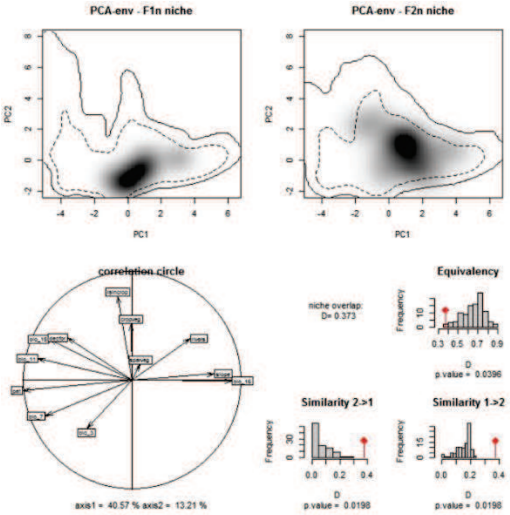

PCA-env - Cn niche
PCA-env - Dn niche
PC1
PC2
correlation circle
axis1 = 29.72 % axis2 = 20.22 %
Equivalency
niche overlap:
D= 0.259
D
p.value = 0.0198
Similarity 2->1
D
p.value = 0.25743
Similarity 1->2
D
p.value = 0.0396
Frequency


PCA-env - Cn niche
PCA-env - F1n niche
PC1
PC2
correlation circle
axis1 = 26.65 % axis2 = 24.39 %
Equivalency
niche overlap:
D= 0.034
D
p.value = 0.0198
Similarity 2->1
D
p.value = 0.83168
Similarity 1->2
D
p.value = 0.09901
Frequency


PCA-env - Cn niche
PCA-env - F2n niche
PC1
PC2
correlation circle
axis1 = 25.85 % axis2 = 22.79 %
Equivalency
niche overlap:
D= 0.823
D
p.value = 0.0198
Similarity 2->1
D
p.value = 0.17822
Similarity 1->2
D
p.value = 0.0396
Frequency


PCA-env - Dn niche
PCA-env - F1n niche
PC1
PC2
correlation circle
axis1 = 30.34 % axis2 = 19.38 %
Equivalency
niche overlap:
D= 0.165
D
p.value = 0.0198
Similarity 2->1
D
p.value = 1.0099
Similarity 1->2
D
p.value = 0.0198
Frequency


PCA-env - Dn niche
PCA-env - F2n niche
PC1
PC2
correlation circle
axis1 = 31.15 % axis2 = 18.14 %
Equivalency
niche overlap:
D= 0.196
D
p.value = 0.0198
Similarity 2->1
D
p.value = 0.0198
Similarity 1->2
D
p.value = 0.0198
Frequency


PCA-env - F1n niche
PCA-env - F2n niche
PC1
PC2

correlation circle
axis1 = 40.57 % axis2 = 13.21 %

niche overlap:
D= 0.373
Equivalency
Frequency
D
p.value = 0.0396
Similarity 2->1
p.value = 0.0198
Similarity 1->2
p.value = 0.0198

Capítulo 5

Discussão Geral

Os objetivos apresentados no capítulo 1 delinearam os temas centrais dos quatro artigos científicos que integram esta tese. Nestes artigos, os resultados foram apresentados e discutidos de forma seccionada. Assim, neste capítulo, pretendeu-se fazer uma discussão geral com o propósito de fornecer uma perspetiva global dos processos evolutivos que determinaram a diversificação dos sapos-parteiros (*Alytes* spp.) na Bacia Mediterrânica Ocidental.

5.1 Relações filogenéticas do género *Alytes* baseadas em análises multilocus

Ao longo das últimas décadas, vários estudos com base em dados morfológicos e genéticos tentaram elucidar as relações filogenéticas entre as espécies de sapo-parteiro (*Alytes* spp.) (Arntzen & García-París 1995, Fromhage *et al.* 2004, Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007, Biton *et al.* 2013). Apesar dos esforços metodológicos, alguns aspectos continuram por esclarecer, nomeadamente i) a politomia do subgénero *Baleaphryne* (*A. muletensis*, *A. dickhilleni* e *A. maurus*), cuja a rápida radiação data do Mioceno e parece estar relacionada com a fragmentação do Maciço Bético-Rifenho e com o fim da Crise Salina Messiniense, aquando da abertura do Estreito de Gibraltar (cerca de 5,3 Ma) (Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007); e ii) a monofilia do subgénero *Alytes*, uma vez que, além da observação de características osteológicas comuns partilhadas entre *A. o. almogavarii* e as espécies do clado *Baleaphryne* (Martínez-Solano *et al.* 2004), haplótipos do gene nuclear beta-fibrinogénio intrão 7 também são partilhados entre os referidos taxa (Gonçalves *et al.* 2007), sugerindo, assim, a existência de fluxo génico interespecífico no passado ou, alternativamente, a persistência de polimorfismo ancestral. No sentido de tentar esclarecer estas questões filogenéticas, recorreu-se no presente trabalho (Artigo I e II) à utilização de diferentes marcadores moleculares (DNA mitocondrial e quatro genes nucleares) e métodos coalescentes de análise. Em particular, pretendeu-se i) esclarecer a tricotomia descrita para a filogenia de *A. muletensis*, *A. dickhilleni* e *A. maurus;* ii) avaliar a datação preliminar da fragmentação histórica das populações destas espécies e a sua associação à crise Messiniense, com a indicação do cenário biogeográfico mais plausível; iii) determinar a posição filogenética de *A. o. almogavarii* relativamente às restantes subespécies de *Alytes obstetricans*; e iv) analisar a diversificação intraespecífica de *A. obstetricans* e sua relação com as oscilações climáticas do Pleistoceno.

A aplicação de métodos recentes de análise multilocus, nomeadamente árvore de espécies baseada na teoria da coalescência (*BEAST), destaca-se como um aspeto

determinante deste trabalho, pois, além de ter contribuído para a elucidação de parte da história evolutiva de *Alytes*, contribuiu também para o debate sobre as limitações dos estudos filogenéticos baseados apenas na utilização de genes mitocondriais e de métodos de análise individual de genes em inferências multilocus – árvores de genes (Degnan & Rosenberg 2009, Heled & Drummond 2010). A análise de árvores de espécies tem-se revelado de extrema importância em grupos de organismos proximamente relacionados ou de espécies crípticas, nos quais as rápidas radiações e eventos de hibridação podem conduzir a inferências evolutivas discordantes quando diferentes marcadores genéticos são analisados isoladamente. Estas discordâncias de sinal filogenético são geralmente comuns quando se analisam genealogias nucleares devido à separação incompleta de linhagens. A distinção dos fatores que promovem discordâncias entre árvores de genes é usualmente difícil e complexa, uma vez que podem produzir perfis moleculares muito semelhantes como já demonstrado, por exemplo, em estudos da história evolutiva de *Alytes* (Fromhage *et al.* 2004, Martínez-Solano *et al.* 2004 e Gonçalves *et al.* 2007). Neste sentido, o uso de métodos de análise coalescente multiespecíficos, que estimam a árvore de espécies, é mais apropriado na reconstrução das relações filogenéticas (Edwards *et al.* 2007, Leaché and Rannala 2011, Williams *et al.* 2013).

Em cenários evolutivos sugestivos de divergência rápida e simultânea, o incremento do número de organismos amostrados e de marcadores genéticos analisados tem demonstrado ser uma abordagem eficiente para o aumento da acurácia filogenética (ver Pabijan *et al.* 2013, Williams *et al.* 2013). Neste trabalho, a implementação de ambas as condições contribuiu decisivamente para a eficácia das análises filogenéticas, tendo as análises coalescentes multiespecíficas (*BEAST - Artigo I e II) recuperado uma árvore de espécies bem suportada e com incremento do sinal filogenético comparativamente com os estudos anteriores. Isto permitiu a resolução da radiação de *Baleaphryne* e a definição da posição filognética de *A.o.almogavarii*.

5.1.2 Relações filogenéticas em *Baleaphryne*

Os resultados obtidos no Artigo I, baseados na análise de dados mitocondriais (aproximadamente 9.000 pares de base), recuperaram o clado *A. maurus* + *A. dickhilleni* bem suportado, sendo *A. muletensis* grupo irmão deste clado (Figura 1a). Este resultado corrobora os dados obtidos por Biton *et al.* (2013) na filogenia apresentada para a família Alytidae baseada em seis genes (mtDNA e nucleares - 2.500bp). No entanto, os quatro genes nucleares estudados (*β-fibint7*; *PPP3CAint4; C-myc* e *RPL9int4*, num total de 3.142 pares de

bases analisados), apesar de apresentarem variação genética significativa, mostraram menor poder de resolução (Figura 1b) comparativamente com o mtDNA. Mesmo assim, os clados representados pelas linhagens de *Baleaphryne* foram recuperados como monofiléticos em todas as genealogias nucleares (ver figuras S1 e S2 do Artigo I), com exceção do *gene beta-fibrinogénio intrão 7* (corroborando os dados de Gonçalves *et al.* 2007), em que se verificou a presença de linhagens de *A. o. almogavarii*. De acordo com os nossos dados e com os de Gonçalves *et al.* (2007), a hipótese para a observação de ausência de monofilia recíproca no marcador *beta-fibrinogénio intrão 7* poderá resultar de um profundo evento coalescente. A menor resolução observada nas genealogias nucleares pode ser consequência da menor taxa de mutação e maior efetivo populacional em relação ao mtDNA (Avise 2000, Zhang *et al.* 2008, Belfiore *et al.* 2008), tornando-as menos suscetíveis à ocorrência de fenómenos estocásticos de extinção e fixação de linhagens (Hare 2001).

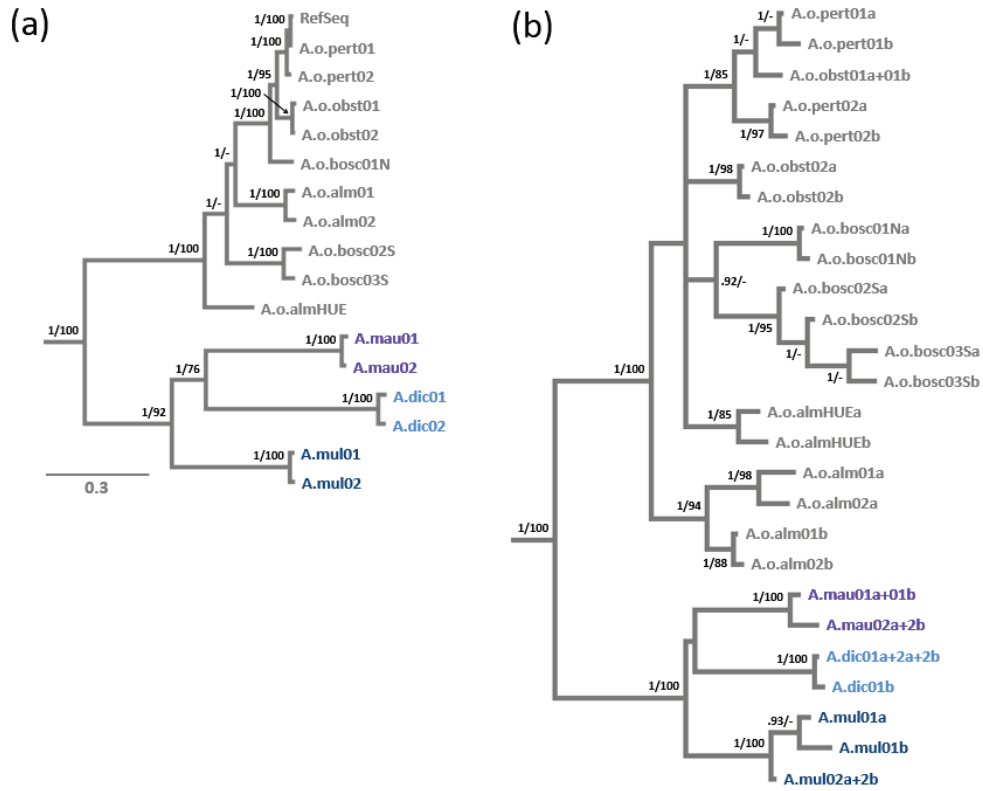


Figura 1. Relações filogenéticas entre as espécies do género *Alytes* baseadas em dados concatenados de (a) mtDNA e (b) nDNA com base em análises Bayesianas e de Máxima Verossimilhança. Representantes de todas as espécies de *Alytes* e das linhagens intraespecíficas descritas para *A. obstetricans* foram incluídos nas análises. Os filogramas apresentados consistem em árvores consenso Bayesianas onde o grupo externo (*A. cisternasii*) foi omitido. Os valores junto aos ramos indicam probabilidades posteriores bayesianas (BPP) e de *bootstrap* (BS) apenas em nós suportados (BPP ≥ 0.9/BS ≥ 75). O código das amostras está descrito na Tabela I do Artigo I. A cor cinzenta corresponde a amostras de *A. obstetricans*, a cor roxa de *A. maurus*, a cor azul claro *A. dickhilleni* e a cor azul escuro *A. muletensis*. O código RefSeq equivale à sequência referência do mtDNA de *A. o. pertinax* (Número de acesso GenBank: NC_006688.1, San Mauro *et al.* 2004).

Em relação aos resultados obtidos na árvore de espécies (Figura 2), algumas críticas podem ser levantadas quanto uma possível dominância do marcador mitocondrial nas análises do *BEAST. No entanto, estudos prévios baseados em simulações têm sugerido que este não será o caso (Heled & Drummond 2010) e os resultados obtidos com o MP-EST, embora com menor suporte (Figura 3b – Artigo I), recuperaram também o clado *A. maurus* + *A. dickhilleni* reforçando a ausência de dominância do mtDNA, uma vez que esta análise não infere árvores de genes para a obtenção da árvore de espécies e, por isso, atribuiu o mesmo peso a todos os marcadores analisados. Além disso, os resultados obtidos no Artigo I estão de acordo com estudos anteriores de radiações rápidas que utilizaram mitogenomas nas suas inferências. Nesses estudos, o mtDNA recuperou o mesmo padrão de relações filogenéticas do que o estimado pelas análises de árvores de genes (ver Steinfartz *et al.* 2007, Zhang *et al.* 2008).

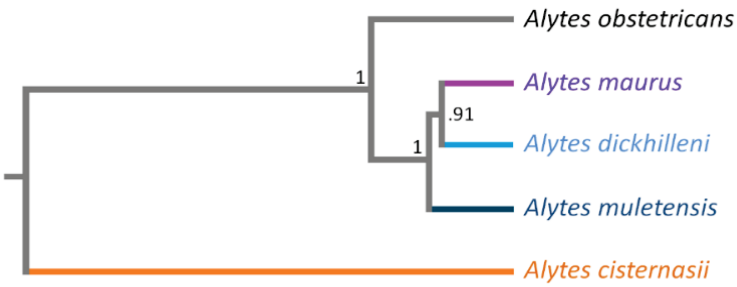


Figura 2. Árvore de espécies do género *Alytes* com base nas análises do *BEAST (mtDNA e quatro genealogias nucleares: *β-fibint7*; *PPP3CAint4; C-myc* and *RPL9int4*). Os valores junto aos ramos indicam probabilidades posteriores Bayesianas.

Um aspecto interessante da biogeografia de *Baleaphryne* está relacionado com a distribuição restrita e disjunta das três espécies incluídas neste clado (*A. muletensis*, *A. dickhilleni* e *A. maurus*). De acordo com as hipóteses filogenéticas publicadas até ao momento, estas espécies divergiram quase simultaneamente. Os eventos vicariantes relacionados com o processo de separação destas linhagens (Arntzen & García-París 1995, 1997; Altaba 1997, Fromhage *et al.* 2004, Martínez-Solano *et al.* 2004) ainda são alvo de discussão, pois os estudos anteriores basearam-se em amostragens incompletas e/ou em um único locus e recorreram a diferentes estratégias na calibração do relógio molecular em função do marcador utilizado (ver Maxson & Szymura 1984, Beerli *et al.* 1996). A abertura do Estreito Bético (López-Martínez 1989) terá sido o evento que conduziu ao isolamento da população ancestral de *Baleaphryne* na região sub-bética e, por sua vez, a abertura do Estreito de Gibraltar o responsável pela diferenciação das três espécies deste clado (Arntzen & García-París 1995, Fromage *et al.* 2004).

Com base na estimativa da árvore de espécies (Artigo I), a radiação de *Baleaphryne* terá acontecido aproximadamente há 3,4 Ma e a separação deste clado de *A. obstetricans* terá ocorrido há 5,6 Ma, contudo amplos intervalos de confiança foram observados em ambas as estimativas. Desta forma, sugere-se que o Maciço Bético-Rifenho terá sido colonizado por populações ancestrais do clado *Baleaphryne* e a fragmentação do maciço, durante o Mioceno, terá sido o evento vicariante que promoveu a divergência das espécies deste clado. O cenário proposto é corroborado por evidências palaeogeológicas descritas por Martín *et al.* (2009). Durante a fragmentação do maciço terá ocorrido a separação e o isolamento da linhagem ancestral de *A. muletensis* nas Ilhas Baleares e, posteriormente, a separação de *A. dickhilleni* + *A. maurus*, uma vez que estas espécies compartilham um ancestral comum mais recente. No final da crise Messiniense, a abertura do Estreito de Gibraltar (Krijgsman *et al.* 1999, Duggen *et al.* 2003, Crespo-Blanc *et al.* 2016) terá sido o evento vicariante que concluiu a total separação das espécies de *Baleaphryne*, aproximadamente há 5,33 Ma. O enchimento da Bacia do Mediterrâneo conduziu ao isolamento quase simultâneo das linhagens que originaram *A. maurus*, no Norte de África, *A. dickhilleni*, nas Serras Béticas, e *A. muletensis*, nas ilhas Baleares. Neste contexto, a dispersão trans-marítima postulada por Martínez-Solano *et al.* (2004) para explicar a diferenciação de *A. muletensis* parece pouco provável.

5.1.3 Relações filogenéticas em *A. obstetricans*

No estudo da história evolutiva do género *Alytes*, a posição filogenética da subespécie *A. o. almogavarii* sempre foi um tema de intenso debate, sendo descrita como divergente e independente das demais linhagens de *A. obstetricans* (Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007). Por isso, no Artigo II desenvolveu-se uma abordagem multilocus combinada com um elevado número de indivíduos analisados na tentativa de clarificar as relações evolutivas entre todas as subespécies reconhecidas de *A. obstetricans*. As análises filogenéticas de sequências de mtDNA recuperaram uma árvore bem resolvida com seis haplogrupos distintos (A até F), sendo três destes altamente divergentes (D, E e F, com uma distância média *p*–não corrigida em relação aos outros haplogrupos de 4,4%, 3,7% e 4,7%, respectivamente, ver Figura 3).

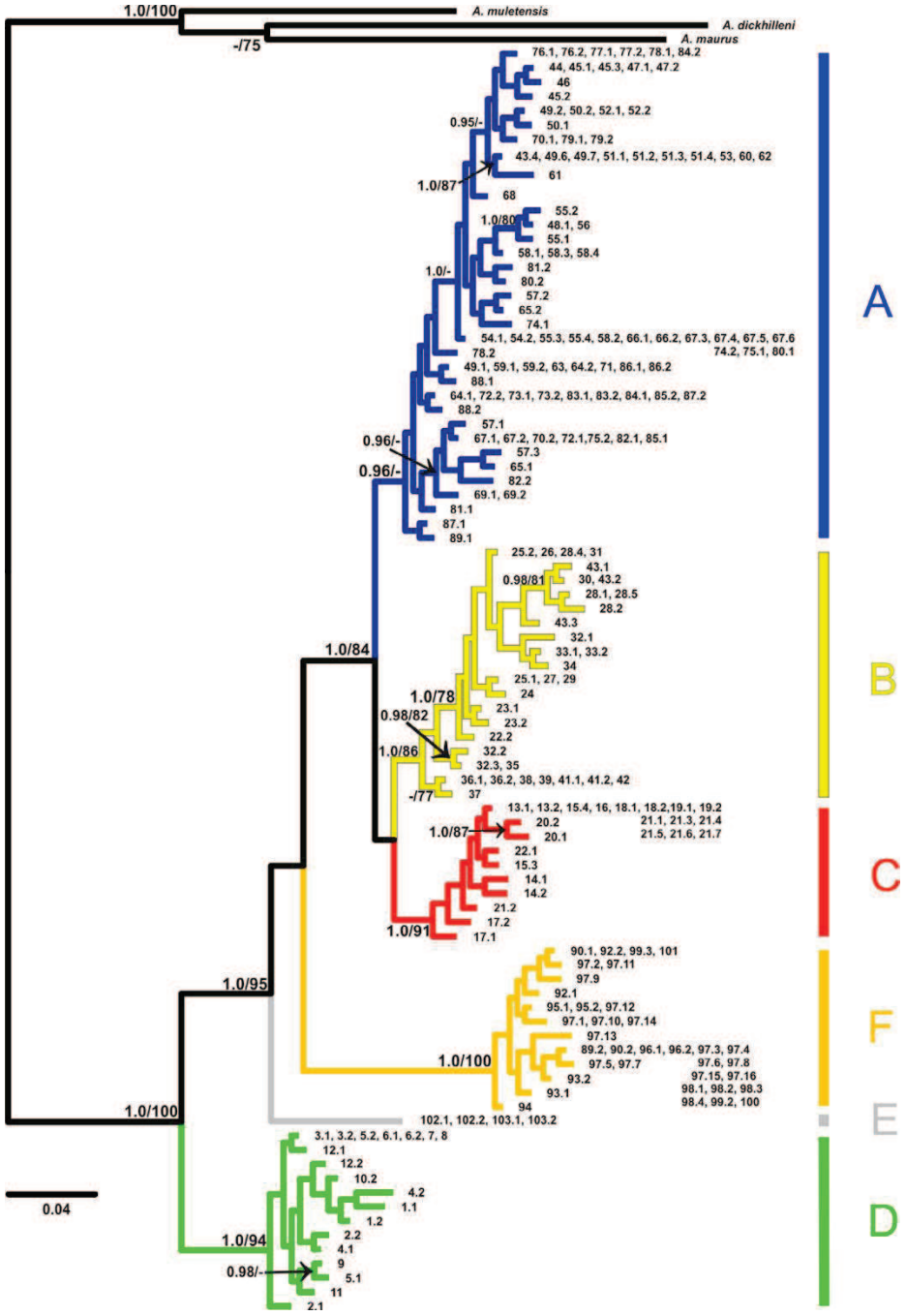


Figura 3. Relações filogenéticas entre as linhagens de *A. obstetricans* com base em análises Bayesianas e de máxima verossimilhança do gene mitocondrial *ND4* (813 bp). O filograma representa a árvore de consenso Bayesiana, em que os valores junto aos ramos indicam probabilidades posteriores bayesianas e de *bootstrap* apenas em nós suportados (BPP ≥ 0.9/BS ≥ 75). O código das amostras está descrito na Tabela I do Artigo II.

Embora as genealogias nucleares estejam tendencialmente de acordo com os dados mitocondriais, elas não recuperam a monofilia dos seis haplogrupos definidos pelo mtDNA. A falta de congruência entre as árvores de genes pode ser reflexo da ocorrência de fluxo génico entre as linhagens de *A. obstetricans*, o que seria concordante com a deteção de pelo menos três zonas de contato entre as diferentes linhagens de mtDNA. Além disso, a análise

da árvore de espécies (Figura 4), mesmo com o incremento do número de haplótipos e marcadores genéticos analisados, também não demonstrou suportes significativos entre as diferentes linhagens (A-F) de *A. obstetricans* definidas pelo mtDNA. Este fato pode ser consequência de: i) uma resolução insuficiente dos marcadores nucleares utilizados e/ou da divergência recente e em simultâneo das linhagens deste grupo, que poderia contribuir para a existência de alguns haplótipos partilhados (principalmente em marcadores com taxa de mutação mais lenta), ii) da retenção de polimorfismo ancestral (Avise 2000, Belfiore *et al.* 2008).

Por outro lado, a árvore de espécies (Figura 4) recuperou a linhagem F, representada por populações previamente identificadas como *A. o. almogavarii*, como grupo irmão do clado que contém as restantes linhagens de *A. obstetricans*, ou seja, um claro padrão de monofilia. Deste modo, foi possível a resolução da posição filogenética de *A. o. almogavarii*, excluindo-se a possibilidade deste taxon ser grupo irmão do clado *Baleaphryne*, uma das hipóteses anteriormente proposta por Gonçalves *et al.* (2007). A retenção de polimorfismo ancestral parece ser a causa mais provável para a observação de haplótipos partilhados do gene beta-fibrinogénio intrão 7 entre os referidos taxa (García-París 1995, Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007). De acordo com a árvore de espécies, a radiação deste taxon foi estimada para 2,5 Ma, no Terciário tardio (Plioceno). Além disso, os nossos resultados corroboram a sugestão de Arntzen & García-París (1995), em que terão sido as oscilações climáticas do Quaternário e a formação das principais bacias hidrográficas na Península Ibérica os eventos vicariantes responsáveis pela diversificação das demais linhagens, influenciando também os padrões de estruturação da diversidade genética (ver discussão na secção 5.2.3.1). Estas estimativas temporais coincidem com as obtidas para outros grupos de organismos amplamente distribuídos na Península Ibérica (ver por exemplo Alexandrino *et al.* 2000, Gonçalves *et al.* 2009, Recuero & García-París 2011, Vences *et al.* 2013, Díaz-Rodríguez *et al.* 2015, Martínez-Freiría *et al.* 2015).

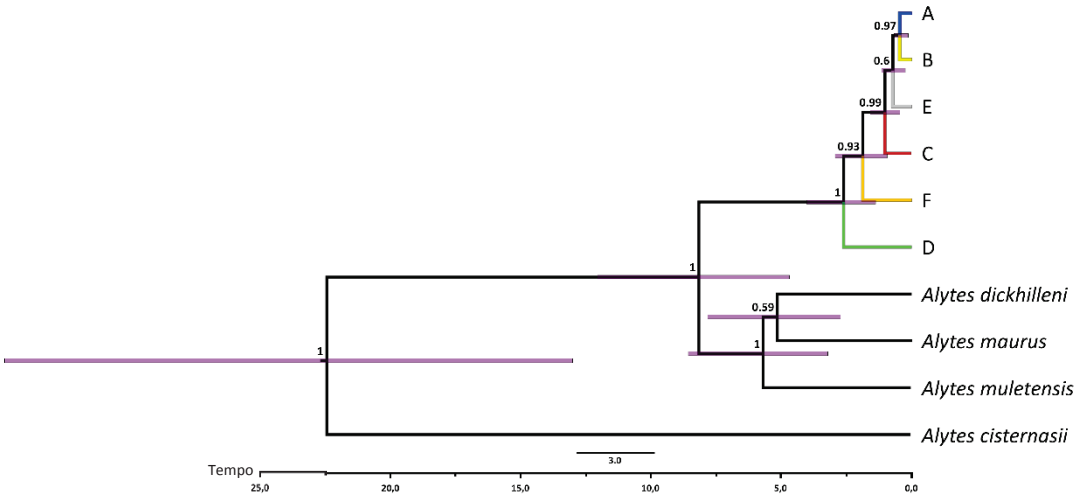


Figura 4. Árvore de espécies do género *Alytes* baseada na análise coalescente de sequências mitocondriais (ND4) e nucleares (*β-fibint7*; *PPP3CAint4; C-myc* e *RPL9int4*) obtida com o programa *BEAST. Esta análise incluiu representantes de todas as espécies de *Alytes* e das linhagens intraespecíficas descritas para *A. obstetricans*. Os valores nos nós indicam probabilidades posteriores bayesianas. As barras horizontais indicam intervalos de densidade posterior superior a 95% para os tempos de divergência. A escala é apresentada em milhões de anos.

5.2 Filogeografia e dinâmicas populacionais em *A. obstetricans*

Os estudos filogeográficos têm como objetivo a análise e interpretação dos factores históricos que conduziram à distribuição geográfica actual das unidades evolutivas proximamente relacionadas (Avise 2000). Nos últimos anos, a aplicação de análises de modelação de nicho ecológico (MNE) a estes estudos tem possibilitado a determinação do papel dos fatores ambientais na manutenção da alopatria entre linhagens que divergiram recentemente (Zink 2014). Assim, os estudos filogeográficos permitem: i) a delimitação de linhagens proximamente relacionadas e a estimativa do tempo de divergência entre elas (Avise 2000); e ii) a utilização de testes quantitativos para determinar se taxa aparentados possuem diferentes nichos ecológicos e se estes nichos são mais ou menos divergentes do que o esperado dado as variáveis ambientais para cada um (ver MacCormack *et al.* 2010). Em relação a *A. obstetricans*, análises anteriores, baseadas em dados aloenzimáticos, morfológicos e genéticos, indicaram elevados níveis de diversidade intraespecífica e estruturação populacional, conduzindo à descrição de quatro subespécies parapátricas: *A. o. obstetricans*, *A. o. boscai*, *A. o. pertinax* e *A. o. almogavarii* (Arntzen & García-Paris 1995, Fonseca *et al.* 2003, Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007). No entanto, existem ainda dúvidas sobre a origem, variabilidade e expansão geográfica destas subspécies. No sentido de elucidar algumas partes deste *puzzle* biogeográfico, o presente trabalho implementou uma abordagem integrada, combinando análises genéticas, espaciais e

ecológicas. As diferentes análises realizadas nos Artigos II, III e IV não só corroboraram os estudos anteriores quanto à elevada diversidade intraespecífica em *A. obstetricans*, como também demonstraram um forte padrão de estrutura populacional. Adicionalmente, possibilitaram a associação de eventos vicariantes com os padrões filogeográficos descritos, a identificação de processos de especiação incipientes e a reunião de evidências que demonstram discordâncias entre a atual classificação taxonómica e a variabilidade genética existente em *A. obstetricans*.

5.2.1 Padrões de estrutura e diversidade genética

As análises de diversidade genética das populações de *A. obstetricans*, apresentadas no Artigo II, revelaram elevados níveis de variação e diversidade intraespecífica (π *ND4* = 0,0243; π *β-fibint7* = 0,0155; π *RPL9int4* = 0,0124; π *PPP3CAint4* = 0,0075; π *C-myc* = 0,0035). Estes resultados corroboraram trabalhos previamente publicados (Arntzen & García-Paris 1995, Martínez-Solano *et al.* 2004, Fonseca *et al.* 2003, Martínez-Solano *et al.* 2004, Gonçalves *et al.* 2007), destacando a região da Península Ibérica como a área que concentra a maior diversidade genética, como já descrito para outros grupos de organismos (ver Hewitt 2004, Schmitt 2007, Taberlet *et al.* 1998, Recuero and García-París 2011). No entanto, o padrão de estruturação desta diversidade varia de acordo com os diferentes marcadores genéticos investigados (mtDNA e genes nucleares). Cinco dos seis haplogrupos mitocondriais descritos no Artigo II (exceção do E) refletem forte concordância geográfica. No entanto, os resultados das análises dos genes nucleares (Figura 3 do Artigo II) evidenciam apenas uma tendência para a formação de grupos equivalentes aos haplogrupos mitocondriais. Em estudo filogeográficos, a observação de padrões diferentes entre marcadores nucleares e mitocondrial é relativamente comum e esperado entre linhagens intraespecíficas e encontra-se amplamente documentada na Peninsula Ibérica para diversas espécies de anfíbios, répteis e mamíferos (García-París *et al.* 2003, Melo-Ferreira *et al.* 2005, Sequeira *et al.* 2008, Pinho *et al.* 2009, Días-Rodrígues *et al.* 2015). No caso de *A. obstetricans*, a existência de diferenças entre mtDNA e genes nucleares pode indicar separação incompleta de linhagens e/ou a presença fluxo génico intraespecífico, corroborando os pressupostos de García-París (1995). Este autor descreve a existência de contato secundário entre subespécies de *A. obstetricans* e a ocorrência de fluxo génico tanto entre *A. o. almogavarii* e *A. o. obstetricans*, como entre *A. o. almogavarii* e *A. o. pertinax.* Neste contexto, a utilização de marcadores com taxas de mutação mais elevadas, como os microssatélites, permitiu a detecção de mecanismos mais

detalhados de dinâmica populacional (por exemplo, fluxo génico e isolamento por distância) e padrões de estruturação genética ao nível nuclear.

Os dados resultantes das análises Bayesianas de agrupamento dos microssatélites (Artigos III e IV) mostraram elevados valores de diversidade genética (riqueza alélica = 46 / heterozigotia esperada = 0,88) e um padrão de estrutura genética muito semelhante ao obtido para o mtDNA, nomeadamente uma elevada concordância geográfica. Numa primeira análise dois grupos foram detectados (Figura 5): i) um formado por amostras previamente identificadas como *A. o. almogavarii* (haplogrupo F) e ii) outro incluindo as restantes amostras de *A. obstetricans*. Análises subsequentes revelaram um segundo nível de estruturação com sete grupos bem definidos (Figura 5) e diferenciados (F_{st} global: 0.208). Todos os grupos possuem clara concordância com cinco dos seis haplogrupos mitocondriais. Além disso, os haplogrupos B e F mostraram uma subdivisão adicional, grupos B1n/B2n e F1n/F2n, respectivamente, enquanto que o haplogrupo E não revelou nenhuma equivalência ao nível nuclear. Neste último caso, todas as amostras pertencentes ao haplogrupo E, de acordo com as análises Bayesianas de agrupamento, foram incluidas no grupo F1n. Uma possível explicação para este facto seria que populações de *A. o. almogavarii*, mais especificamente do grupo F1n, teriam capturado a linhagem E presente em populações ancestrais durante movimentos de expansão e colonização secundária dos Pirinéus centrais. Alternativamente, esta linhagem mitocondrial estaria representada em populações isoladas e ancestrais de *A. o. almogavarii* em frequências diminutas, e, por deriva genética, encontra-se sub-representada atualmente.

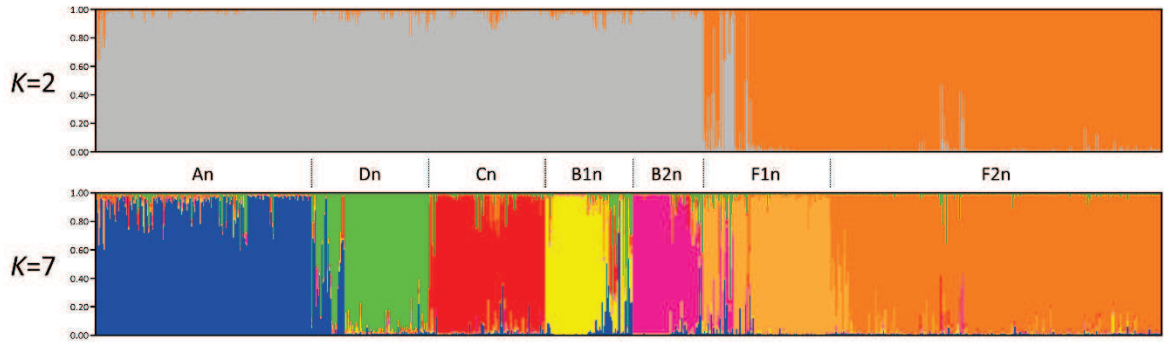


Figura 5. Estrutura populacional de *A. obstetricans*, *K*=2 e *K*=7, obtida com base em análises Bayesianas de agrupamento de 12 microssatélites. Cada indivíduo é representado por uma linha vertical, que se encontra dividida em porções correspondentes à fração do genoma desse indivíduo com origem em cada uma das K populações ancestrais.

As análises Bayesianas de agrupamento permitem a recuperação de níveis maiores ou menores de estrutura populacional de acordo com a variação do número de marcadores utilizados e a diversidade genética (polimorfismos) de cada locus (Pritchard *et al.* 2000, Porras-Hurtado *et al.* 2013). Na presente investigação, os grupos genéticos nucleares (An, B1n, B2n, Cn, Dn, F1n, F2n), descritos no Artigo IV, possuem um claro padrão filogeográfico concordante com os haplogrupos mitocondriais (A, B, C, D, F) descritos no Artigo II. De acordo com as características dos marcadores em questão, como a taxa de mutação e o tamanho do efetivo populacional, os resultados da análise do mtDNA e dos microssatélites podem refletir períodos temporais diferenciados e por isso apresentar padrões distintos: um mais histórico (mtDNA), relacionado com eventos vicariantes, e outro mais contemporâneo (microssatélites), relacionado, por exemplo, com dinâmicas populacionais (Avise 2000). Assim, mecanismos de isolamento populacional por efeito de distância, associados à baixa capacidade de dispersão da maioria das espécies de anfíbios, podem revelar uma maior definição de subestruturação populacional com base em análises de agrupamento de microssatélites (por exemplo, Martínez-Solano *et al.* 2006, Sequeira *et al.* 2008, Gonçalves *et al.* 2009, ainda, ver discussão em Christiansen & Reyer 2011). No entanto, os resultados dos testes de Mantel (Artigo IV) somente demonstraram resultados significativos de isolamento por distância em apenas dois dos sete grupos: Cn e F2n. Por outro lado, os dados provenientes das análises do mtDNA e das genealogias nucleares (Artigo II) apontam a possibilidade da ocorrência de fluxo génico entre linhagens. Neste sentido, seria de esperar que os microssatélites indicassem a existência de fluxo génico entre as subespécies, principalmente quando o contato secundário entre elas é favorecido pela ausência de barreiras geográficas em muitas regiões da Península Ibérica (ver exemplos na secção 5.2.3 deste capítulo). Contudo, os resultados da análise de fluxo génico entre os sete grupos genéticos (Artigo IV) não suportam esta hipótese.

Diante do exposto, as populações de *A. obstetricans*, aparentemente, não estão isoladas por efeito da distância geográfica, mas a existência de fluxo génico entre os diferentes grupos de linhagens parece reduzido ou praticamente inexistente. Outros fatores poderão estar envolvidos na manutenção do padrão filogeográfico revelado pelo mtDNA e microssatélites. Alguns estudos têm sugerido que fatores ambientais podem estar envolvidos no estabelecimento de padrões filogeográficos, atuando, inclusivemente, em processos de divergência intraespecífica por meio de seleção adaptativa (ver Ortego *et al.* 2012, Abellán and Svenning 2014). Assim, combinámos os resultados das análises Bayesianas de agrupamento dos microssatélites com métodos de modelação de nicho ecológico (Artigo IV) para avaliar se fatores ambientais estão envolvidos, de alguma forma, no processo microevolutivo de *A. obstetricans*.

5.2.2 Fatores ambientais, estrutura populacional e adaptação

Os estudos que combinam a diversidade biológica presente em diferentes marcadores genéticos com modelação de nicho ecológico (MNE) têm contribuído para o esclarecimento do papel dos fatores ambientais no estabelecimento de padrões de estrutura da variabilidade genética de uma espécie e para a avaliação temporal dos padrões demográficos (Carsten & Richards, 2007). Adicionalmente, o conhecimento do efeito das variáveis ecológicas nos processos de divergência genética e/ou fluxo génico intraespecífico tem tido implicações diretas na forma como reconhecemos, classificamos e conservamos a diversidade biológica (Morales *et al.* 2016). Desta forma, a abordagem integrada (microssatélites e MNE), implementada no Artigo IV, foi de extrema importância para o entendimento dos padrões de diversidade intraespecífica e estrutura populacional, assim como permitiu fazer inferências evolutivas e taxonómicas em *Alytes obstetricans*.

A modelação de nicho ecológico em associação com os testes de sobreposição de nicho (Artigo IV) possibilitaram a delimitação de zonas de contacto entre as linhagens de *A. obstetricans*, assim como a avaliação da existência de divergência ecológica entre elas. Os resultados destes testes demonstraram evidências de divergências significativas de nicho entre os sete grupos genéticos definidos pelos microssatélites (Figura 6 deste capítulo e Tabela 5 do Artigo IV). Mais do que isso, os modelos obtidos revelaram que cada nicho é definido por um conjunto de variáveis ambientais distintas. Do ponto de vista ecológico, este fato poderá ser a causa da observação de zonas de contato relativamente diminutas, ou seja, estreitas áreas de sobreposição de nicho entre os sete grupos distintos, com exceção da sobreposição de nicho observada entre os grupos Cn e Dn. Neste caso, estes grupos distribuem-se numa região de clima tipicamente atlântico e os resultados obtidos por meio da MNE mostraram que as suas áreas de persistência são potencialmente amplas, sendo a precipitação (bio15 e bio16) a variante ecogeográfica mais importante na sua definição. Isto possibilitaria a ocupação, por parte dos grupos Cn e Dn, de uma vasta área sob influência de clima atlântico e consequente sobreposição. No entanto estes grupos, representados por populações previamente identificadas como pertencentes a *A.o.boscai*, não se distribuem nas regiões localizadas a sul do rio Douro, no caso do grupo Cn, e a norte do mesmo rio, no caso de Dn. Portanto, o rio Douro tem um claro papel de barreira geográfica, pois, mesmo com características ambientais favoráveis, as populações de ambas os grupos não ultrapassam o seu limite.

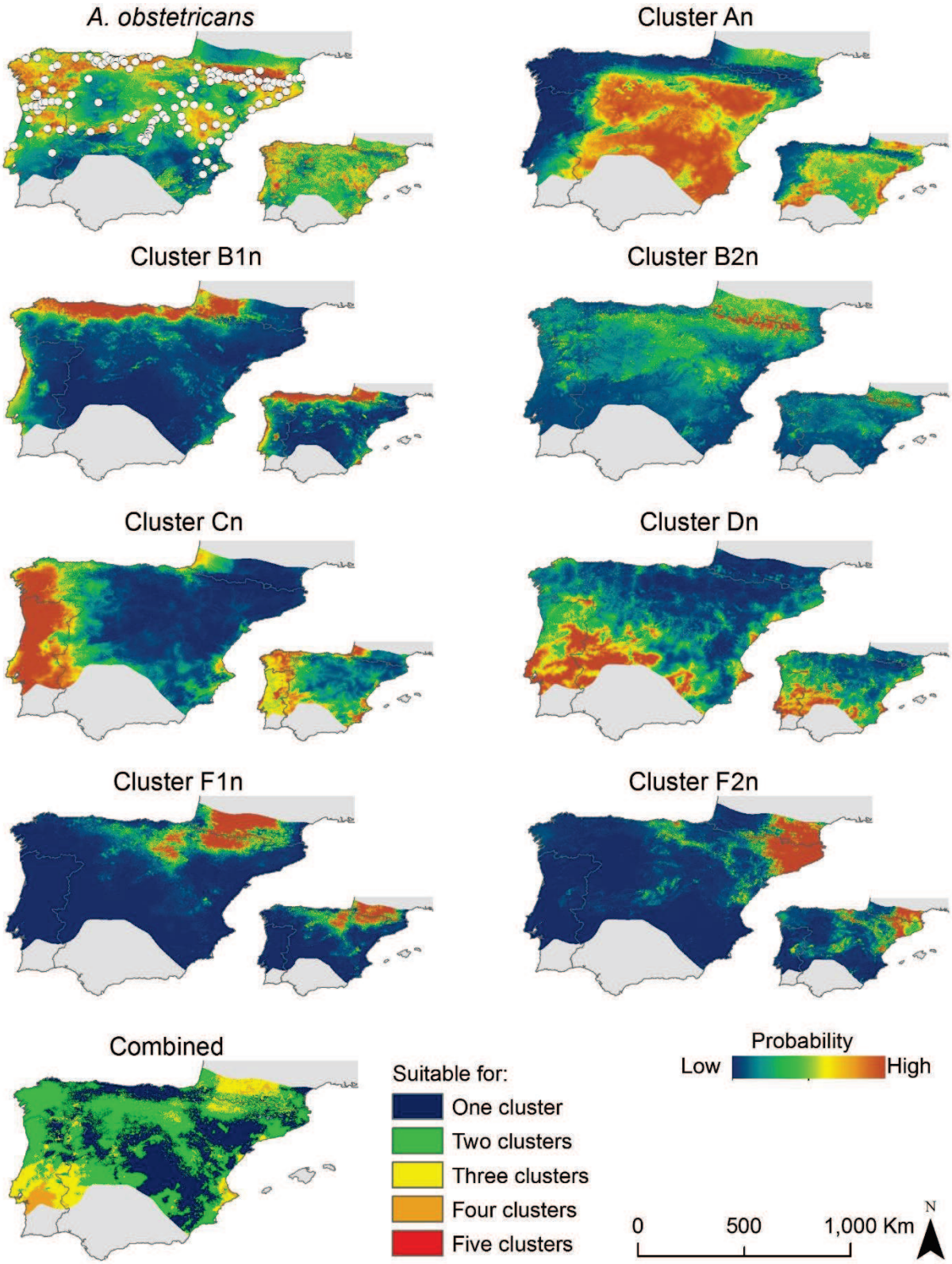


Figura 6. Modelação de nicho ecológico em *A. obstetricans* com base nas análises Bayesianas de agrupamento. Os mapas maiores representam a probabilidade da presença de *A. obstetricans* para cada um dos grupos obtidos na análise de agrupamento Bayesiano. Os mapas menores de cada modelo representam o desvio padrão entre réplicas. O mapa "Combinado" representa a sobreposição de nicho entre os diferentes grupos no espaço geográfico.

Em oposição ao conceito de conservação de nicho, o qual define que linhagens intraespecíficas têm tendência para responder de forma ecologicamente análoga às alterações ambientais (ver Stewart *et al.* 2010 e Hughes & Woodward 2008), os resultados apresentados no Artigo IV demonstram fortes indícios de divergência de nicho ecológico entre os grupos de *A. obstetricans.*

Parece evidente que os resultados da análise de divergência de nicho destacam a influência dos fatores ambientais no padrão de distribuição da diversidade intraespecífica de *A. obstetricans*. No entanto, não podemos descartar que o padrão de estruturação populacional observado possa ser resultado da influência de outros fatores, nomeadamente demográficos, genéticos (deriva genética) e ecológicos (relações de competição). Alguns estudos sugerem que *taxa* proximamente relacionados serão muito semelhantes ecologicamente para coexistirem numa mesma área, sendo as relações de competição as responsáveis pelo estabelecimento de padrões alopátricos como resultado, por exemplo, de expansões populacionais pós-glaciares e contato secundário (Rundle & Nosil, 2005). No entanto, este modelo suportaria apenas o forte padrão de estruturação populacional revelado pelas análises do mtDNA e microssatélites e não a divergência de nichos em si.

Por outro lado, alguns modelos postulam que populações recentemente isoladas podem adaptar-se a diferentes habitats (Mayr 1942, 1947, 1963). Ainda neste sentido, outros estudos defendem que num cenário de expansão populacional, após períodos consecutivos de contração, diferentes linhagens podem encontrar-se e formar ecótonos ou barreiras abruptas, permanecendo nos habitats em que se tornaram melhor adaptadas (ver Zink 2014). Neste modelo, a alopatria seria o resultado de divergência ecológica, podendo ocorrer no início de processos de diversificação de linhagens (nível intraespecífico). Assim, os nossos resultados parecem ilustrar o modelo discutido em Zink (2014). Somando-se à MNE do Artigo IV, as análises de PCA indicaram a ausência de sobreposição de nichos em áreas com indicação de co-ocorrência de diferentes grupos genéticos (por exemplo, haplogrupos A e B, na Galiza) e, adicionalmente, os testes de similaridade demonstraram que populações de um determinado grupo não têm tendência a ocupar os nichos de outros grupos. Além disso, as nossas análises incluíram variáveis ecogeográficas, bioclimáticas e os dados genéticos provenientes dos microssatélites que, diferentemente do mtDNA, revelam um padrão filogeográfico tendencialmente contemporâneo (Avise 2000). Portanto, toda a análise conduzida no Artigo IV tende a apoiar a influência dos fatores ambientais na biogeografia de *A. obstetricans*.

Com a excepção do trabalho realizado por Tarroso *et al.* (2014), que integra ferramentas moleculares e modelação de nicho ecológico para estimar o papel da divergência ecológica nos limites da distribuição entre espécies, este é o primeiro estudo a incorporar dados genéticos resultantes de métodos de agrupamento Bayesiano para criar modelos

baseados em nichos ecológicos e explorar os padrões espaciais da diversidade genética. No seu recente trabalho Gotelli e Stanton-Geddes (2015) argumentam que os métodos de modelação da distribuição das espécies aplicados apenas aos registos de presença/ausência da espécie podem não acomodar facilmente a possibilidade de adaptação local e mudança evolutiva dentro de diferentes partes de uma área de distribuição e sugerem a incorporação da diversidade genética intraespecífica para construir estes modelos. Deste modo, se os nossos resultados indicam que o padrão obtido poderá ser o resultado da influência de fatores ambientais, então seria de questionar o papel da seleção adaptativa no processo microevolutivo desta espécie, ou seja, a importância da seleção natural nos estádios iniciais do processo de especiação (Rundell & Price 2009, Schluter 2009, Nosil 2012). Contudo, mais estudos noutros organismos serão necessários para solidificar esta hipótese, uma vez que o sinal de conservação de nichos é comum entre linhagens divergentes do período pós-glaciar (Lee-Yaw & Irwin 2015).

5.2.3 Diversificação de linhagens e biogeografia

Do ponto de vista filogenético, os haplogrupos mitocondriais D e F são os mais basais e divergentes de *A. obstetricans* (Figura 4). Esta condição ancestral reflete-se também nos índices de diversidade genética e na estrutura da rede de haplótipos observados tanto para o mtDNA quanto para os microssatélites. O haplogrupo F será o mais antigo, divergindo dos demais há cerca de 2,5 Ma (ver Artigo II). Este haplogrupo, que inclui as populações previamente designadas por *A. o. almogavarii*, está aparentemente restrito ao extremo nordeste da Península Ibérica, nos Pirinéus, onde persistiu provavelmente desde o Pleistoceno. No entanto, a presença deste haplogrupo no sudeste de França, descrita anteriormente por Geniez & Crochet (2003), foi corroborada neste estudo pelos resultados de mtDNA e microssatélites. Desta forma, a existência de uma zona de contacto com o haplogrupo B (relacionada com as populações designadas previamente de *A. o. obstetricans*), em território francês, parece provável. Estudos futuros serão necessários para confirmar este contato secundário, delinear a sua extensão e as dinâmicas populacionais existentes entre os dois grupos. Quanto à distribuição a sul dos Pirinéus, o haplogrupo F aparentemente não atravessa o vale do Ebro, discordando de trabalhos anteriores (García-París 1995), que sugeriam uma ampla distribuição geográfica até as montanhas de Guadarrama, nas redondezas de Madrid. As análises da genotipagem dos microssatélites (Figura 5) demonstram uma clara e forte subestruturação deste haplogrupo com a definição de dois sub-grupos: F1n e F2n, um na região oeste dos Pirinéus e outro a leste dos Pirinéus. O sinal de

diferenciação dos microssatélites pode ser identificado na árvore de mtDNA (Figura 3), contudo sem suporte estatístico. A sobreposição geográfica destes grupos ocorre entre os Rios Segre e Cinica, os quais podem funcionar como uma barreira geográfica, mantendo estes grupos isolados, limitando o fluxo génico e favorecendo processos de diferenciação tanto ao nível genético quanto ecológico. Esta hipótese de diversificação é apoiada pelos resultados das análises de agrupamento Bayesianas e de modelação de nicho ecológico (Figura 5 e 6). No presente, a população de *A. o. almogavarii* provavelmente mantem-se estável nos Pirinéus (ver *mismatch distribution*, Figura 4 – Artigo II e ausência de fluxo génico com outras subespécies). No entanto, os testes de *bottleneck* apontam para que reduções populacionais tenham ocorrido no passado (Artigo IV). Este cenário, compatível com as oscilações climáticas do Pleistoceno, foi bem caracterizado para o grupo F2n em detrimento do F1n (menos diverso geneticamente). A população ancestral de *A. o. almogavarii* pode estar relacionada com o grupo F2n, assim como a provável área de refúgio, na porção oriental dos Pirinéus, durante os períodos glaciares.

Os ciclos de contração e expansão populacional, muito característicos durante o Pleistoceno nos organismos da Península Ibérica, podem ter contribuído para a diferenciação e flutuações populacionais das demais linhagens de *A. obstetricans*, dado que a estimativa de tempo de divergência entre elas é de cerca de 2 Ma. O haplogrupo D é o segundo mais diferenciado dentro de *A. obstetricans* (ver árvore de espécies, Figura 2 deste capítulo) e, segundo os dados da genotipagem de microssatélites, corresponde ao grupo Dn. Este haplogrupo inclui populações a sul do Rio Douro, em Portugal, e a oeste das montanhas do Sistema Central em Espanha, descritas previamente como *A. o. boscai*. Esta área de distribuição, concordante com os dados de mtDNA obtidos por Fonseca *et al.* (2003), inclui as montanhas a sul do Rio Douro que são reconhecidas como uma zona de grande diversidade genética e um importante refúgio glaciar durante o Pleistoceno, onde diferentes espécies resistiram a condições climáticas adversas (ver por exemplo, Alexandrino *et al.* 2000, Martínez–Solano *et al.* 2006). De acordo com as análises do mtDNA, este haplogrupo mostra sinais de expansão demográfica (valores significativos de *Tajima's D* e de R^2 - Ramos-Onsins & Rozas – Artigo II) e, segundo os dados de modelação de nicho ecológico, possui uma extensa área favorável à sua persistência no Sistema Central (ver Artigo IV).

Por outro lado, outras populações de sapo-parteiro-comum identificadas previamente como *A. o. boscai* (cuja a localidade-tipo é Tuy, província de Pontevedra, Espanha; ver García-París & Martínez-Solano, 2001) foram agrupadas no haplogrupo C. Segundo os dados da genotipagem de microssatélites, este haplogrupo corresponde ao grupo Cn. Aparentemente restrito ao noroeste da Península Ibérica, o haplogrupo C mostra também sinais de expansão populacional recente com base na análise do mtDNA (ver Artigo II), sendo a sua distribuição claramente limitada a sul pelo Rio Douro. Contudo, com base nas análises de MNE, habitats

muito favoráveis para a sua ocorrência são identificados a sul deste rio. Por outro lado, a leste da distribuição do haplogrupo C, foi identificada uma zona de contato secundário com o haplogrupo B. Os resultados mitocondriais mostram a ocorrência de ambos os haplogrupos nas redondezas da província de Lugo, Espanha (Figura 1 – Artigo II). A co-ocorrência destes haplogrupos é também observada na localidade da Corunha, Espanha. Estes dados corroboram os resultados de Arntzen *&* García-París (1995), obtidos após análise de aloenzimas e padrões fenotípicos de coloração. Os resultados obtidos com base nos microssatélites indicam que a área de fronteira entre as regiões da Galiza e Astúrias será o limite da distribuição dos dois grupos, Cn e B1n, atuando como zona de contato secundário. Esta área é coincidente com a região prevista pelas análises de modelação ecológica de nicho, pois, enquanto o grupo B1n parece mais restrito à região da cordilheira Cantábrica, o grupo Cn tem na região da Galiza o seu habitat mais favorável. De acordo com os dados apresentados, é possível que, no passado, as populações de *A. o. obstetricans* (haplogrupo B) pudessem ter ocupado todo o litoral norte da Península Ibérica, incluindo a região da Galiza (ver Fonseca *et al.* 2003). Porém, a expansão demográfica das populações de *A. o. boscai* (haplogrupo C), favorecida pelas condições ambientais, poderá ter provocado uma retração na distribuição geográfica dos *A. o. obstetricans* (haplogrupo B). Assim, um estudo genético mais detalhado será necessário para a definição da zona de contato entre estas duas subspécies e para avaliação da ocorrência de fluxo génico.

O haplogrupo B, descrito em populações identificadas previamente como *A. o. obstetricans*, distribui-se desde o norte de Espanha (Cantábria) até à Europa Central, incluindo França e partes do território da Áustria e Bélgica, Luxemburgo e região oeste da Alemanha (ver Figura 1 – Artigo II). A norte dos Pirinéus, toda a extensa área geográfica ocupada por esta linhagem apresenta reduzida variabilidade genética, tanto ao nível mitocondrial como nuclear (genealogias e microssatélites), sugerindo uma expansão rápida e recente, provavelmente após o Último Máximo Glaciar. Em contraste, os mesmos marcadores moleculares indicam que as populações da Península Ibérica apresentam níveis mais elevados de diversidade genética. Este contraste de diversidade permitu a caracterização, com base nos microssatélites, de dois grupos fortemente estruturados denominados de B1n, no norte da Península Ibérica, e B2n, a norte dos Pirinéus. A dicotomia nos índices de diversidade pode ser consequência da baixa representatividade das populações amostradas a norte dos Pirinéus, assim como a perda de significância nos testes de neutralidade (Artigo II). No entanto, este padrão contrastante, assim como a sugestão da existência de um corredor de colonização no sentido norte europeu por *A. o. obstetricans*, é análogo a outras espécies de anfíbios, como *Lissotriton helveticus* e outros vertebrados (Recuero & García-París, 2011). Se geneticamente os grupos B1n e B2n são bem diferenciados, os testes de modelação de nicho ecológico mostram que o habitat logo a norte dos Pirinéus, ocupado pela linhagem B2n,

é também propício para a linhagem B1n (Figura 3 – Artigo IV). Por esta razão, o estudo das populações francesas de *A. o. obstetricans* será primordial para definir a zona de contato entre os dois grupos definidos pelos microssatélites, e, ainda, avaliar se os Pirinéus estarão a funcionar como barreira geográfica ao fluxo génico.

O haplogrupo A, que inclui as populações previamente designadas de *A. o. pertinax*, mostra sinais de expansão demográfica (ver Artigo II). Este haplogrupo, com base nos dados de microssatélites, corresponde ao grupo An. De acordo com os Artigos II e IV, foi possível demonstrar que a área de distribuição desta linhagem é muito mais extensa do que anteriormente considerada. Esta subespécie ocorre nas margens do rio Ebro, percorrendo toda sua extensão desde a Catalunha até às Montanhas Cantábricas. Nas montanhas Cantábricas, no norte da Península Ibérica, esta subespécie distribui-se a sul deste sistema montanhoso. No extremo oeste cantábrico, os *A. o. pertinax* distribuem-se no sentido sul, ocupando regiões próximas da fronteira entre Portugal e Espanha. Já no centro da Península, ocorrem em todo o Sistema Central. Os marcadores genéticos (mtDNA e microssatélites) apontam para a ocorrência de zonas de contacto com: i) *A. o. almogavarii* no nordeste da Península Ibérica; ii) *A. o. obstetricans* na região da Cantábria; e iii) *A. o. boscai* (haplogrupo C) no noroeste da Península. Por último, existe ainda uma importante área de sobreposição entre as populações de *A. o. pertinax* e as populações pertencentes à linhagem D*,* desde a Serra da Gata até às redondezas de Madrid.

De acordo com a análise dos dados filogenéticos, filogeográficos e de MNE obtidos nesta investigação, reconhece-se a existência de, pelo menos, cinco linhagens distintas e muito bem definidas, sob o ponto de vista genético, de *A. obstetricans* (A, B, C, D e F). O elevado número de indivíduos amostrados, representativo de toda a área de distribuição da espécie, foi um fator determinante para a definição destas linhagens intraespecíficas que, dentro de um contínuo evolutivo, parecem apresentar trajetórias diferentes e independentes. De acordo com os dados de mtDNA, o processo de diversificação intraespecífico de *A. obstetricans* iniciou-se antes do Pleistoceno, contudo foram as oscilações climáticas desta época que atuaram mais diretamente na diferenciação das linhagens. Paralelamente, as grandes bacias hidrográficas da Península Ibérica parecem ter atuado como importantes barreiras geográficas no processo microelutivo de *A. obstetricans*, nomeadamente as bacias do Douro e Ebro. De uma forma geral, as populações de *A. obstetricans*, durante os ciclos glaciares do Pleistoceno, passaram por períodos de isolamento, contração e expansão populacional. Este fato conduziu ao estabelecimento de padrões de estruturação populacional e de divergência genética, para os quais contribuíram não só fatores estocásticos e demográficos, mas também ambientais. Segundo os dados provenientes da MNE, é possível ainda levantar a hipótese de que parte do processo de divergência intraespecífica possa ser influenciado por forças de seleção adaptativa, corroborando que a heterogeneidade

topográfica e/ou climática da Península Ibérica torna-a uma região propícia para intensa diferenciação intraespecífica (Gómez & Lunt, de 2007, Abellán & Svenning 2014).

Historicamente, as populações ancestrais do sapo-parteiro-comum terão persistido em vários refúgios durante as oscilações climáticas do Pleistoceno (Figura 7), assim como descrito para diversas outras espécies de anfíbios ibéricos (Martínez-Solano *et al.* 2004, 2006; Sequeira *et al.* 2008, Gonçalves *et al.* 2009, Vences 2013, Díaz-Rodríguez *et al.* 2015), corroborando o modelo de "refúgio-dentro-do-refúgio" (ver Gómez & Lunt 2007, Weiss & Ferrand 2007, Abellán Svenning 2014). Assim, propõem-se que os habitats de elevada altitude, como as serras da Estrela e Gredos (oeste do Sistema Central) e Guadarrama (este do Sistema Central), possam ter sido colonizados de forma independente por populações correspondentes à subespécie *A. o. boscai* (referente a linhagem D, altamente divergente) e à subespécie *A.o.pertinax*, respectivamente. Do mesmo modo que as altas montanhas do norte da Península Ibérica foram independentemente colonizadas por populações pertencentes à subespécie *A. o. obstetricans* (Montanhas Cantábricas) e por populações de *A. o. almogavarii* (Pirinéus).

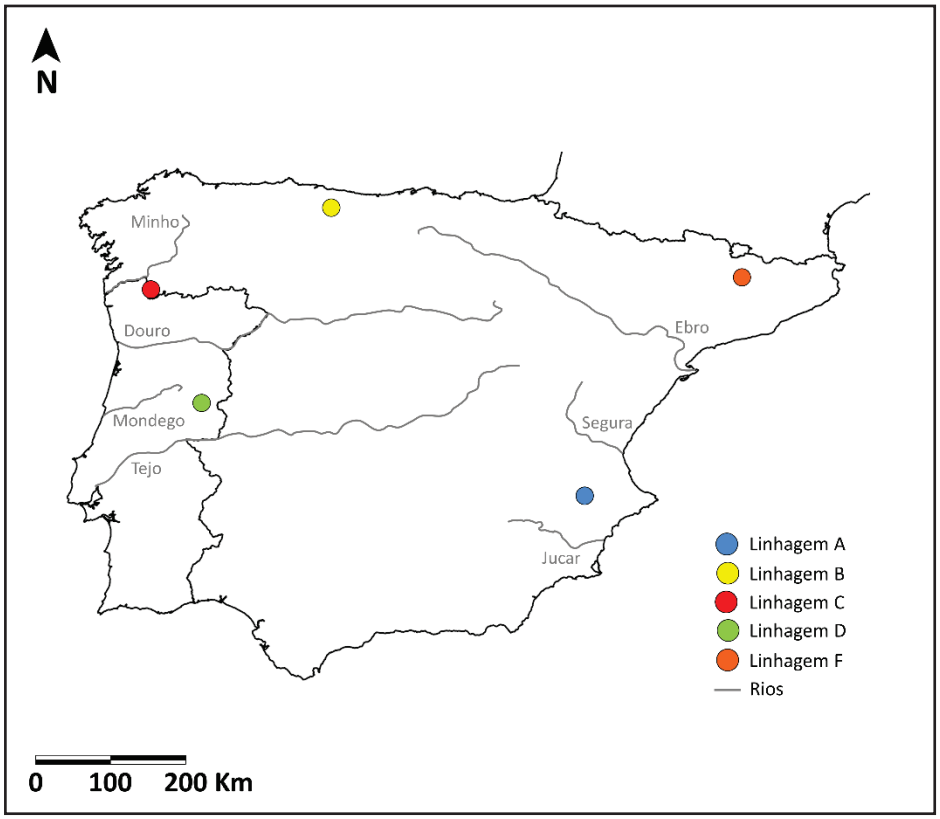


Figura 7. Provaváveis áreas de refúgio de *A. obstetricans* durante o Pleistoceno com base na análise de marcadores moleculares (mtDNA, microssatélites) e MNE.

5.3 Implicações taxonómicas

De acordo com estudos anteriores, baseados em dados aloenzimáticos, morfológicos e genéticos, as populações de *A. obstetricans* possuem elevada diversidade biológica (Arntzen & García-París 1995, Fonseca *et al.* 2003, Martínez-Solano *et al.* 2004, Gonçalves 2007), possibilitando a diferenciação de grupos populacionais e o reconhecimento de quatro subespécies: i) *A. o. obstetricans* (Laurenti, 1768), distribuída no norte da Península Ibérica e norte e centro da Europa, ii) *A. o. boscai* Lataste, 1879, presente nas regiões norte e centro de Portugal, Galiza, oeste de Castilla-León, assim como ao longo do Sistema Central; iii) *A. o. pertinax* García-París & Martínez-Solano, 2001, ocorre nas regiões centro e leste da Península Ibérica (García-París & Martínez-Solano, 2001); e iv) *A. o. almogavarii* Arntzen & García-París, 1995, distribuída a este dos Pirinéus (Geniez & Crochet, 2003) até a Serra de Guadarrama no norte de Madrid (García-París, 1995).

No Artigo II, as relações filogenéticas entre *A. obstetricans* e as demais espécies do género foram esclarecidas com a inclusão dos *A. o. almogavarii* dentro do subgénero *Alytes*. Uma vez comprovada a monofilia da espécie, foi possível diagnosticar, com base no mtDNA, o caráter ancestral e divergente das linhagens F e D. As linhagens A, B e C possuem uma estimativa de divergência mais recente, relacionada com as glaciações do Pleistoceno. Estas linhagens ao nível nuclear, segundo os dados da genotipagem de microssatélites (Artigo III e IV), possuem uma clara correspondência com os grupos An, Bn e Cn. O padrão de estrutura genética equivalente entre os marcadores refletiu-se também na distribuição geográfica. Além disso, os dados provenientes das análises de MNE (Artigo IV) mostraram diferenças significativas em relação aos habitats ocupados por cada um dos grupos mitocondriais (divergência de nicho). Portanto, a combinação de dados moleculares (mtDNA e microssatélites) com dados ambientais (testes de sobreposição de nicho e modelação de nicho ecológico) corroboraram o estatuto subespecífico das linhagens A, B e C, correspondendo às subespécies anteriormente descritas como *A. o. pertinax*, *A. o. obstetricans* e *A, o. boscai*, respectivamente. Confirma-se ainda a distribuição parapátrica destes *taxa*, destacando-se a observação de estreitas zonas de contato e limitado fluxo génico.

Em relação a linhagem F, a mais divergente de *A. obstetricans*, os dados de microssatélites evidenciaram forte subestruturação populacional refletida na caracterização de dois grupos distintos, F1n e F2n. O segundo grupo será o mais antigo e estabelece uma estreita zona de contato com F1n. As análises indicam a existência de reduzido fluxo génico entre si, assim como com os grupos An, B1n e B2n. Além disso, as populações representantes de F2n ainda mostram sinal de isolamento por distância. Corroborando esta diferenciação, as

análises de MNE mostram que ambos os grupos possuem nichos ecológicos diferentes, bem caracterizados e com estreita área de sobreposição. A linhagem F corresponde, portanto, às populações descritas como *A. o. almogavarii*. No entanto, a sua elevada diferenciação, diversidade genética e distribuição aparentemente restrita aos Pirinéus indicam uma trajetória evolutiva certamente independente e divergente das demais congéneres, justificando o estatuto de espécie, como já anteriormente sugerido por Gonçalves (2007).

Por fim, a linhagem D, representada por populações descritas como *A. o. boscai*, é o segundo taxon mais divergente de *A. obstetricans*. Ao nível nuclear, esta linhagem corresponde ao grupo Dn, apresentando uma distribuição delimitada pelos rios Douro, a norte, e Tejo, a sul. De acordo com as análises de MNE, esta área de distribuição é extremamente favorável à ocorrência e persistência da linhagem D. Os resultados obtidos corroboram o estatuto subespecífico da linhagem D, mas a sua história evolutiva, padrão filogeográfico e nicho ecológico são distintos e divergentes da linhagem C, para que ambas sejam descritas como *A. o. boscai*, sugerindo a necessidade de uma reorganização taxonómica desta subespécie.

Em resumo, é possível constatar discordância na taxonomia subespecífica atual em relação às linhagens F e D. Assim, sugerimos que estas linhagens sejam consideradas unidades taxonómicas diferenciadas, uma vez que apresentam trajetórias evolutivas longas (evidência filogenética), padrões de diversidade e estrutura genética únicos (evidência filogeográfica) e nichos ecológicos divergentes (evidência ecológica). Este reconhecimento pode ter consequência direta nos planos de avaliação de biodiversidade e políticas públicas de conservação (Bickford *et al.* 2007). No entanto, estudos adicionais das populações de *A. obstetricans* são necessários, nomeadamente análises detalhadas das zonas de contato, para identificar o nível de porosidade entre as linhagens, como também novas análises morfológicas, fisiológicas, de seleção sexual, comportamental, entre outras.

5.4 Considerações finais

Os quatro artigos que compõe esta tese contribuiram para a elucidação da história evolutiva e aspectos biogeográficos das espécies de sapo-parteiro (*Alytes* spp.). Num sentido mais amplo, este trabalho permitiu a compreensão do papel de eventos vicariantes e fatores ambientais nos processos de diversificação ao nível inter e intraespecífico, destacando-se a resolução da tricotomia do clado *Baleaphryne* e a descrição da diversificação das linhagens de *A. obstetricans*. A combinação de uma ampla e representativa amostragem, o uso diferentes marcadores moleculares (mtDNA, genealogias nucleares e microssatélites) e,

ainda, a utilização de vários métodos de análise, destacando-se as análises coalescentes, as Bayesianas de agrupamento e a modelação de nicho ecológico, foram aspectos determinantes no alcance dos objetivos propostos neste trabalho.

Do ponto de vista filogenético, os resultados obtidos nesta tese foram particularmente importantes, pois possibilitaram a resolução de algumas incongruências detectadas em trabalhos anteriores, especialmente a rápida radiação do clado *Baleaphryne* e a definição da posição filogenética de *A. o. almogavarii*. A nossa abordagem coalescente multilocus (mtDNA e quatro genealogias nucleares), através da análise da árvore de espécies (*BEAST), permitiu a definição de *A. muletensis* como grupo irmão do clado formado por *A. maurus* + *A. dickhilleni* e relacionar a divergência destas espécies com a fragmentação do Maciço Bético-Rifenho, durante o Mioceno, seguido da abertura do estreito de Gibraltar, no final da crise Messiniense. Outro aspecto filogenético interessante foi a definição de *A. o. almogavarii* como grupo irmão das demais linhagens de *A. obstetricans*, descartando-se a hipótese inicial de que esta subespécie faria parte do clado *Baleaphryne.*

Além de resolver o carater monofilético de *A. obstetricans*, as nossas análises filogenéticas, com base no mtDNA, possibilitaram a definição de seis haplogrupos mitocondriais (A-F) muito diversos e diferenciados geneticamente. De acordo com a análise da árvore de espécies, a radiação deste taxon foi estimada para 2.5 Ma aproximadamente, ou seja, desde o Terciário tardio (Plioceno). No entanto, terão sido as oscilações climáticas do Pleistoceno, os eventos vicariantes responsáveis pela diversificação das demais linhagens de *A. obstetricans*. Assim, a Península Ibérica terá sido a principal região de diversificação intraespecífica deste taxon, reforçando o modelo de "refúgio-dentro-do-refúgio" proposto por Gómez & Lunt (2007).

Em paralelo, a análise dos dados da genotipagem de microssatélites, com base análises Bayesianas de agrupamento, possibilitaram a definição de sete grupos geneticamente diferenciados. O padrão de estruturação genética destes grupos tem uma clara correspondência com aquele obtido pelo mtDNA. Dentro de cada grupo, as análises populacionais demonstraram ausência de isolamento em função da distância geográfica (exceções: Cn e F2n) e a existência de fluxo génico reduzido entre eles.

Os padrões concordantes de estrutura genética (mtDNA e microssatélites) também foram corroborados do ponto de vista ecológico. Segundo as nossas análises de MNE, fatores ambientais podem estar a influenciar diretamente a manutenção deste padrão, pois todos os grupos definidos com base nos microssatélites apresentaram divergências de nichos ecológico significativas. Isto justificaria a existência de estreitas áreas de contato secundário, revelada pela MNE, e do reduzido fluxo génico entre os grupos, sugerindo a possibilidade de seleção adaptativa no processo microevolutivo de *A. obstetricans*.

Em estudos futuros, a análise massiva de polimorfismos obtidos por meio da sequenciação de nova geração poderia não somente definir as relações filognéticas que ainda permanecem sem precisão, assim como esclarecer aspectos microevolutivos intrínsecos deste grupo, por exemplo, a evidência de seleção adaptativa levantada nesta tese. A implementaçao de abordagens como *reverse genetics* e *candidate gene,* pode confirmar e descrever o papel da selecção na diversificação intraespecífica dos *A. obstetricans*, assim como auxiliar na identificação de genes que estejam sob a influencia de factores ambientais, por exemplo aqueles que determinaram a identificação e o padrão geográfico dos diferentes grupos genéticos.

O estudo integrado implementado nesta tese, com o uso de diferentes marcadores genéticos e abordagens de análises, forneceu múltiplas evidências de que o sapo-parteiro-comum, *A. obstetricans*, possui, duas unidades evolutivas muito bem caracterizadas genética e ecologicamente, representadas pelas linhagens D e F. Por isso, informações adicionais quanto a variabilidade morfológica, zonas de contato secundário com as restantes linhagens e dados genómicos seriam cruciais para a definição do estatuto taxonómico destas linhagens.

Os resultados apresentados nesta tese, em conjunto com as propostas de investigações futuras, poderão certamente contribuir para a compreensão do processo de diversificação intraespecífico de *A. obstetricans*, incluindo o papel da seleção adaptativa, assim como para o estabelecimento de políticas mais eficazes de conservação da diversidade biológica desta espécie.

5.5 Referências Bibliográficas

Abellán P, Svenning JC (2014) Refugia within refugia–patterns in endemism and genetic divergence are linked to Late Quaternary climate stability in the Iberian Peninsula. Peninsula. *Biological Journal of the Linnean Society*, 113: 13–28.

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.

Altaba C (1997) Phylogeny and biogeography of midwife toads: a reappraisal. *Contributions to Zoology*, 66: 257–262.

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.

Arntzen JW, García-París M (1997) Phylogeny and biogeography of midwife toads (*Alytes*, Discoglossidae): a rebuttal. *Contributions to Zoology, 66*: 263–268.

Avise JC (2000) Phylogeography. The History and Formation of Species. Harvard University Press. Cambridge, Massachusetts, USA, 464 pp.

Beerli P, Hotz H, Uzzell T (1996) Geologically dated sea barriers calibrate a protein clock for Aegean water frogs. *Evolution*, 50: 1676 – 1687.

Belfiore N, Liu L, Moritz C (2008) Multilocus phylogenetics of a rapid radiation in the genus *Thomomys* (Rodentia: Geomyidae). *Systematic Biology*, 57: 294–310.

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.

Biton R, Geffen E, Vences M, Cohen O, Bailon S, Rabinovich R, Gafny S (2013) The rediscovered Hula painted frog is a living fossil. *Nature Communications*, 4: 1959.

Carstens BC, Richards CL (2007) Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution*, 61: 1439-54.

Christiansen DG, Reyer HU (2011) Effects of geographic distance, sea barriers and habitat on the genetic structure and diversity of all-hybrid water frog populations. *Heredity*, 106: 25-36.

Crespo-Blanc A, Comas M, Balanyá JC (2016) Clues for a Tortonian reconstruction of the Gibraltar Arc: Structural pattern, deformation diachronism and block rotations. *Tectonophysics*, 683: 308–324.

Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24: 332-340.

Díaz-Rodríguez J, Gonçalves H, Sequeira F, Sousa-Neves T, Tejedo M, Ferrand N, Martínez-Solano I (2015) Molecular evidence for cryptic candidate species in Iberian *Pelodytes* (Anura, Pelodytidae). *Molecular Phylogenetics and Evolution*, 83, 224-241.

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

Edwards SV, Liu L, Pearl DK (2007) High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences USA*, 104: 5936–5941.

Fonseca A, Arntzen JW, Crespo EG, Ferrand N (2003) Regional differentiation in the common midwife toad (*Alytes obstetricans*) in Portugal: a picture from mitochondrial DNA. *Zeitschrift für Feldherpetologie*, 10: 83–89.

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

García-París M (1995) Variabilidad genética y distribución geográfica de *Alytes obstetricans almogavarii* en España. *Revista Española de Herpetología*, 9: 133–138.

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

García-París M, Martínez-Solano I (2001) Nuevo estatus taxonómico para las poblaciones iberomediterráneas de *Alytes obstetricans* (Anura: Discoglossidae). *Revista Española de Herpetología*, 15: 99–113.

Geniez P, Crochet PA (2003) Confirmation de l'existence, en France, de trois taxons méconnus: *Alytes obstetricans almogavarii* Arntzen & García-París, 1995 (Amphibia, Discoglossidae), *Podarcis hispanica sebastiani* (Klemmer, 1964) (Reptilia, Lacertidae) et *Natrix natrix astreptophora* (Seoane, 1884) (Reptilia, Colubridae). *Bulletin de la Société Herpétologique de France*, 105: 41-53.

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Phylogeography of Southern European Refugia: Evolutionary Perspectives on the Origin and Conservation of European Biodiversity (Eds. Weiss S, Ferrand N), pp. 155-187. Springer, Dordrecht, The Netherlands.

Gonçalves H (2007) História evolutiva dos sapos-parteiros (*Alytes* spp.) na Península Ibérica. Análise filogenética e filogeográfica, reconstrução de um cenário biogeográfico e implicações taxonómicas. Tese de Doutoramento. Universidade do Porto, Portugal.

Gonçalves H, Martínez-Solano I, Ferrand N, García-París M (2007) Conflicting phylogenetic signal of nuclear vs mitochondrial DNA markers in midwife toads (Anura, Discoglossidae, *Alytes*): Deep coalescence or ancestral hybridization? *Molecular Phylogenetics and Evolution*, 44: 494-500.

Gonçalves H, Martínez-Solano I, Pereira RJ, Carvalho B, García-París M, Ferrand N (2009) High levels of population subdivision in a morphologically conserved Mediterranean toad (*Alytes cisternasii*) result from recent, multiple refugia: evidence from mtDNA, microsatellites and nuclear genealogies. *Molecular Ecology*, 18: 5143-5160.

Gotelli NJ, Stanton-Geddes J (2015) Climate change, genetic markers and species distribution modelling. *Journal of Biogeography*, 42: 1577–1585.

Hare MP (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology and Evolution*, 16: 700-706.

Heled J, Drummond AJ (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, 27: 570–80.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359: 183–195.

Hughes PD, Woodward JC (2008) Timing of glaciation in the Mediterranean mountains during the last cold stage. *Journal of Quaternary Science*, 23: 575–588.

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

Leaché AD, Rannala B (2011) The accuracy of species tree estimation under simulation: a comparison of methods. *Systematic Biology*, 60: 126–137.

Lee-Yaw JA, Irwin DE (2015) The importance (or lack thereof) of niche divergence to the maintenance of a northern species complex: the case of the long-toed salamander (*Ambystoma macrodactylum* Baird). *Journal of Evolutionary Biology*, 28: 917-30.

López Martínez N (1989). Tendencias en Paleobiogeografía. El futuro de la biogeografía del pasado. In: Paleontología (Ed. Aguirre E), pp. 271–296. CSIC, Madrid.

Martín JM, Braga J, Aguirre J, Puga-Bernabéu A (2009) History and evolution of the North-Betic Strait (Prebetic Zone, Betic Cordillera): a narrow, early Tortonian, tidal-dominated, Atlantic-Mediterranean marine passage. *Sedimentary Geology*, 216: 80–90.

Martínez-Freiría F, Velo-Antón G, Brito JC (2015) Trapped by climate: interglacial refuge and recent population expansion in the endemic Iberian adder *Vipera seoanei*. *Diversity and Distributions*, 21: 331–344.

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.

Maxson LR, Szymura JM (1984) Relationships among Discoglossid frogs: an albumin perspective. *Amphibia–Reptilia*, 5: 245–252.

Mayr E (1942) Systematics and the origin of species. Columbia University Press, New York, NY, 372 pp.

Mayr, E (1947) Ecological factors in speciation. *Evolution*, 1: 263–288.

Mayr, E (1963) Animal species and evolution. Harvard University Press, Cambridge, MA, 797 pp.

McCormack JE, Zellmer AJ, Knowles LL (2010) Does niche divergence accompany allopatric divergence in *Aphelocoma jays* as predicted under ecological speciation? Insights from tests with niche models. *Evolution*, 64: 1231-1244.

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.

Morales A, Villalobos F, Velazco PM, Simmons NB, Piñero D (2016) Environmental niche drives genetic and morphometric structure in a widespread bat. *Journal of Biogeography*, 43: 1057-1068.

Nosil P. 2012. Ecological Speciation. New York: Oxford Univ. Press, 304 pp.

Ortego J, Riordan EC, Gugger PF, Sork VL (2012) Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. *Molecular Ecology*, 21: 3210-3223.

Pabijan M, Wandycz A, Hofman S, Wecek K, Piwczynski M, Szymura JM (2013) Complete mitochondrial genomes resolve phylogenetic relationships within *Bombina* (Anura Bombinatoridae). *Molecular Phylogenetic and Evolution*, 69: 63–74.

Pinho C, Kaliontzopoulou A, Carretero MA, Harris DJ, Ferrand N (2009). Genetic admixture between the Iberian endemic lizards *Podarcis bocagei* and *Podarcis carbonelli* : evidence for limited natural hybridization and a bimodal hybrid zone. *Journal of Zoological Systematics and Evolutionary Research*, 47: 368–377.

Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo A, Lareu MV (2013) An overview of STRUCTURE: applications, parameter settings, and supporting software. *Frontiers in genetics*, 4: 98.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155: 945–59.

Recuero E, García-París M (2011) Evolutionary history of *Lissotriton helveticus*: Multilocus assessment of ancestral vs. recent colonization of the Iberian Peninsula. *Molecular Phylogenetics and Evolution*, 60: 170–182.

Rundell RJ, Price TD (2009) Adaptive radiation, nonadaptive radiation, ecological speciation and non-ecological speciation. *Trends in Ecology & Evolution*, 24: 394–399.

Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, 8: 336–352.

San Mauro D, García-París M, Zardoya R (2004) Phylogenetic relationships of discoglossid frogs (Amphibia: Anura: Discoglossidae) based on complete mitochondrial genomes and nuclear genes. *Gene*, 343: 357–366.

Santos NP, Ribeiro-Rodrigues EM, Ribeiro-Dos-Santos AK, Pereira R, Gusmão L, Amorim A, Guerreiro JF, Zago MA, Matte C, Hutz MH, Santos SE (2010) Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. *Human Mutation*, 31: 184-190, 2010.

Schluter D, Conte GL (2009) Genetics and ecological speciation. *Proceedings of the National Academy of Sciences of the United States of America*, 106: 9955-9962.

Schmitt T (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, 4: 11.

Sequeira F, Alexandrino J, Weiss S, Ferrand N (2008) Documenting the advantages and limitations of different classes of molecular markers in a well-established phylogeographic context: Lessons from the Iberian endemic Golden-striped salamander, *Chioglossa*

lusitanica (Caudata: Salamandridae). *Biological Journal of the Linnean Society*, 95: 371 – 387.

Steinfartz S, Vicario S, Arntzen JW, Caccone A (2007) A Bayesian approach on molecules and behavior: reconsidering phylogenetic and evolutionary patterns of the Salamandridae with emphasis on *Triturus* newts. *Journal of experimental zoology. Part B, Molecular and developmental evolution*, 308B: 139–162.

Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia revisited: individualistic responses of species in space and time. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 277: 661–671.

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

Tarroso P, Pereira RJ, Martínez-Freiría F, Godinho R, Brito JC (2014) Hybridization at an ecotone: ecological and genetic barriers between three Iberian vipers. *Molecular Ecology*, 23: 1108–23.

Vences M, Hauswaldt JS, Steinfartz S, Rupp O, Goesmann A, Künzel S, Smirnov NA (2013) Radically different phylogeographies and patterns of genetic variation in two European brown frogs, genus *Rana*. *Molecular Phylogenetics and Evolution*, 68: 657-670.

Weiss S, Ferrand N (2007) Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origin and Conservation of European Biodiversity. Springer, Dordrecht, The Netherlands. 377pp.

Williams JS, Niedzwiecki JH, Weisrock DW (2013) Species tree reconstruction of a poorly resolved clade of salamanders (Ambystomatidae) using multiple nuclear loci. *Molecular Phylogenetic and Evolution*, 68: 671–682.

Zhang P, Papenfuss TJ, Wake MH, Qu L, Wake DB (2008) Phylogeny and biogeography of the family Salamandridae (Amphibia: Caudata) inferred from complete mitochondrial genomes. *Molecular Phylogenetic and Evolution*, 49: 586–597.

Zink RM (2014) Homage to Hutchinson, and the role of ecology in lineage divergence and speciation. *Journal of Biogeography*, 41: 999–1006.