UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS

Linking phylogeography and recent dispersal in high mountains: Insights from two Iberian amphibians

"Documento Definitivo"

Doutoramento em Biodiversidade, Genética e Evolução

Federica Lucati

Tese orientada por:
Doutor Marc Ventura Oller
Professor Doutor Rui Miguel Borges Sampaio e Rebelo

Documento especialmente elaborado para a obtenção do grau de doutor
2020

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS

Linking phylogeography and recent dispersal in high mountains: Insights from two Iberian amphibians

Doutoramento em Biodiversidade, Genética e Evolução

Federica Lucati

Tese orientada por:
Doutor Marc Ventura Oller
Professor Doutor Rui Miguel Borges Sampaio e Rebelo

Júri:
Presidente:

- Doutora Sólveig Thorsteinsdottir, Professora Associada com Agregação e Presidente do Departamento de Biologia Animal da Faculdade de Ciências da Universidade de Lisboa

Vogais:

- Doutor Marc Ventura Oller, Investigador, Centro de Estudios Avanzados de Blanes do Consejo Superior de Investigaciones Científicas (CSIC), Espanha (Orientador)
- Doutor Joaquín Ortego Lozano, Investigador, Estacion Biologica de Doñana do Consejo Superior de Investigaciones Científicas (CSIC), Espanha
- Doutor Salvador Carranza Gil Dolz Del Castellar, Investigador Principal, Institut de Biologia Evolutiva (CSIC – UPF), Espanha
- Doutora Angelica Crottini, Investigadora Auxiliar, CIBIO-InBIO da Universidade do Porto
- Doutor Carlos Alberto Rodrigues Fernandes, Investigador Auxiliar, Faculdade de Ciências da Universidade de Lisboa

Trabalho apoiado pela Fundação para a Ciência e a Tecnologia (FCT) através da bolsa de doutoramento PD/BD/52598/2014, pela Comissão Europeia através do projeto LIFE+ LimnoPirineus (LIFE13 NAT/ES/001210) e pela Societas Europaea Herpetologica (SEH)

Documento especialmente elaborado para a obtenção do grau de doutor
2020

This dissertation should be cited as:

Lucati F (2020) Linking phylogeography and recent dispersal in high mountains: Insights from two Iberian amphibians. PhD Thesis, Universidade de Lisboa, Portugal.

Nota Prévia

A presente tese apresenta resultados de trabalhos já publicados (capítulo 2), submetidos para publicação (capítulo 3), ou em preparação para publicação (capítulo 4), de acordo com o previsto no nº 2 do artigo 25º do regulamento de Estudos de Pós-Graduação da Universidade de Lisboa, publicado no Diário de República, 2.ª série — N.º 60 — 26 de março de 2018. Tendo os trabalhos sido realizados em colaboração, o candidato esclarece que participou integralmente na conceção dos trabalhos, obtenção dos dados, análise e discussão dos resultados, bem como na redação dos manuscritos.

Lisboa, Dezembro de 2020

Federica Lucati

Acknowledgements

16/12/2020, 17:08 pm: in a few hours I will finally submit my PhD thesis, and the one thing left is the acknowledgement page. My eyes get wet every time I think about the journey that led me here. 6 years ago, I would have never imagined myself concluding and eventually submitting my PhD thesis. It seemed to be so far away and unreachable. What a roller coaster these past years have been! 6 years ago, following a pretty bad car accident, I was at the point to leave the PhD program and quit science, which has been my greatest passion since I was a kid. I went back to Italy to recover and undergo an intensive physical therapy program, sick and tired of that situation, which was hard for me to understand and digest at the time. I swore I would quit science and start working as a waitress/saleswoman etc. (whatever but science), and I was anxious about the idea of going back to Lisbon and restart the PhD. I had been told that I could do a big part (2 years) of the PhD abroad, but that idea didn't thrill me at the beginning. I was stuck inside a sort of black hole, and didn't know what to do with my life and where to go. After nearly one and a half year of surgeries and physical therapy, and with my morale dropping down more and more, I decided to look for a place abroad where I could continue my PhD and give it a try. I arrived in Barcelona with nothing but a backpack, a little hope and a list of potential advisors and research centres written by hand. I did some interviews and eventually met who would have been my advisor, Marc. I didn't know he would be the best PhD advisor I could ask for.

Now is the moment to say thank you to all those who helped me achieve this terrific goal:

Thanks to Paolo for always being there and dragging me out of my downward spiral. I wouldn't have made it without you.

Thanks to my supervisors: Marc, for welcoming me with open arms when I had nothing but sorrow to offer: I'll never be grateful enough; and Rui, for being always kind and supportive to me. And thanks to my former supervisor, Margarida Santos-Reis, for being so understanding and caring to me and for always believing in me.

Thanks to the Portuguese crew (Joana, Sara, Adrià, Duygu, Yasaman, Paula, William, Silvia) for the good times we spent together during the first year of my PhD.

Thanks to the Spanish crew (Marc, Teresa, Jenny, Victor, Ibor, Jongmo, Nuria, Alex, Jan, Alba, Albert, Blanca, Eloi, Isma) for welcoming me into the CEAB, for the adventures we had together in the Pyrenees and the endless nights in search of newts. And how can I forget the GuaGua (Nuria, Xavier, Mac, Rafel, Anna, Erik, Jen, Tim)? Thank you guys for the laugh, I truly had a great time with you in the car.

Thanks to my parents and sister, for always being supportive no matter what, although I'm pretty sure that they see me as the "weirdo" of the family and still have no clue what I am doing and why :)

Thanks to all my friends, especially those who were forced to listen to my PhD-related complaints :)

The amazing illustrations on chapter opening pages were kindly designed by Ottone Studio (thank you Morris and Sara). Also thanks Azzurra for helping me with the maps.

Finally, thanks to Salvatore Aranzulla, for always having a quick solution to my computer issues.

I still can't believe I made it.

Resumo

As montanhas constituem geralmente "hotspots" de biodiversidade, albergando comunidades únicas que incluem geralmente *taxa* endémicos e/ou ameaçados. Na montanha, os organismos apresentam frequentemente adaptações específicas ao frio e às condições adversas em que vivem, especialmente importantes nos habitats de alta montanha, onde vive uma flora e fauna diversificada e adaptada localmente. Embora sejam geralmente considerados como intocados, os habitats de alta montanha estão muitas vezes expostos a várias ameaças de origem antrópica, tais como espécies invasoras e doenças emergentes.

A Península Ibérica tem uma fisiografia complexa, com várias cadeias montanhosas que permitiram a sobrevivência de populações de várias espécies durante as flutuações climáticas que caracterizaram o Período Quaternário, muitas vezes isoladas durante períodos longos em refúgios glaciares geograficamente separados. Houve padrões complexos na sobrevivência às eras glaciares em diferentes refúgios da Península Ibérica, como por ex. os Pirinéus, os Picos de Europa e o Sistema Central (por exemplo, os cenários "refúgio-dentro-de-refúgio"). A Península Ibérica é especialmente importante para a conservação dos anfíbios, com muitas espécies endémicas e/ ou espacialmente estruturadas, incluindo aquelas nas quais esta tese se centra: o tritão-dos-ribeiros dos Pirenéus, *Calotriton asper* e os sapos-parteiros *Alytes obstetricans*/*almogavarii*.

A estrutura populacional de uma espécie resulta da interação de processos contemporâneos e históricos. Recentemente, o uso de ferramentas de análise genética aumentou consideravelmente a compreensão do efeito destes processos, atuais ou históricos. Os dados genéticos são cada vez mais frequentemente usados para estimar a história demográfica de populações ou espécies, assim como os processos e vias de colonização pós-glaciares, incluindo a localização de refúgios, o tamanho das populações ancestrais, as datas aproximadas de divergência populacional e a ocorrência de efeitos de gargalo ou de expansão das populações. A genética populacional permite também a avaliação da importância de processos atuais, tais como a conectividade e as taxas de dispersão, assim como dos tamanhos populacionais, oferecendo uma alternativa à tradicional obtenção de dados no campo.

O objetivo geral desta tese foi elucidar os principais fatores, históricos e contemporâneos, que explicam a atual diversidade genética e a estrutura populacional de anfíbios de alta montanha.

Para tal, abordei uma variedade de tópicos usando uma combinação de abordagens indiretas (genética) e diretas (trabalho de campo). Usei métodos genéticos que revelaram novas pistas sobre a filogeografia das espécies estudadas, informaram sobre a conectividade e a dispersão atuais de *Calotriton asper*, e sobre as consequências genéticas dos surtos de quitridiomicose em sapos-parteiros. Quanto a *C. asper*, também combinei análises genéticas e trabalho de campo (captura-recaptura de animais fotoidentificados) para definir o papel da dispersão a uma escala espacial pequena e caracterizar o processo de recolonização de lagos após a erradicação de peixes invasores.

Os capítulos desta tese estão divididos em duas partes: a parte 1, composta pelos capítulos 2 e 3, aborda o estudo da filogeografia das espécies estudadas, e a parte 2, composta pelo capítulo 4, os padrões de dispersão recente e a dinâmica de colonização de novos habitats por *C. asper*.

Capítulo 2: Refúgios glaciares múltiplos e dispersão contemporânea definem a estrutura genética de um anfíbio endémico dos Pirinéus

O objetivo principal deste estudo foi desvendar a filogeografia e determinar o grau de conectividade das populações atuais de *C. asper* e dos habitats usados por esta espécie à escala dos Pirinéus, usando uma abordagem multilocus. A análise de microssatélites revelou cinco linhagens principais com indicações de fluxo nas zonas limite. A aplicação de análises Bayesianas aproximadas e de modelos lineares indicou que as cinco linhagens provavelmente têm histórias evolutivas separadas e podem ser rastreadas até cinco refúgios glaciares distintos. A diferenciação das linhagens começou por altura do Último Máximo Glaciar em três áreas principais (Pirinéus Ocidentais, Centrais e Orientais) e continuou até ao final do Último Período Glaciar nos Pirenéus centrais, onde divergiram mais duas linhagens. Não foi encontrada nenhuma evidência de dispersão recente entre as linhagens, e as fronteiras atuais representam provavelmente zonas de contato secundário após a expansão a partir dos vários refúgios. Em cada linhagem, e embora a maioria dos eventos de dispersão envolva populações geograficamente próximas, na sua maioria pertencentes ao mesmo vale, também encontrei potencial para fenómenos ocasionais de dispersão a longa distância.

Capítulo 3: Padrões contrastantes de mistura genética explicam a filogeografia das populações ibéricas de sapos-parteiros de alta montanha

Este capítulo elucidou os promotores da diferenciação geográfica das populações de alta montanha dos sapos-parteiros *A. obstetricans*/*almogavarii* dos Pirinéus, Picos da Europa e Guadarrama, usando um conjunto de marcadores mitocondriais e nucleares. Nas três regiões montanhosas analisadas encontrámos discordâncias mitonucleares marcadas e evidências de mistura entre *taxa*. Com análises de agrupamento, identificámos três linhagens principais divergentes nos Pirinéus (correspondendo aos Pirinéus orientais, centrais e centro-ocidentais), que se expandiram e misturaram recorrentemente durante a sucessão de períodos glaciares-interglaciares que caracterizaram o Pleistocénico Superior. As populações dos Picos de Europa (NO da Península Ibérica) mostraram uma afinidade mitocondrial com as populações do centro-oeste dos Pirenéus e uma afinidade nuclear com as populações da Península Ibérica central, sugestão de uma provável origem mista das populações dos Picos de Europa. Finalmente, as populações da cordilheira de Guadarrama (centro da Península Ibérica) têm uma diversidade genética reduzida, possivelmente em consequência de uma recente epidemia de quitridiomicose.

Capítulo 4: Avaliação genética da conectividade populacional regional do anfíbio endémico *Calotriton asper* e padrões de recolonização após a erradicação de peixes

Este capítulo resulta de uma bolsa que recebi em 2018 da Societas Europaea Herpetologica (SEH). O objetivo deste capítulo foi analisar, a uma escala local, a conectividade das populações e os padrões de recolonização de *C. asper* após a erradicação de peixes invasores, usando uma combinação de análises genéticas e captura-recaptura de animais foto-identificados. Os resultados mostraram que a análise genética foi mais eficiente na caracterização dos padrões de dispersão. A validade da foto-identificação para o reconhecimento individual de *C. asper* foi testada pela primeira vez. Os resultados indicaram que o padrão ventral da espécie é altamente padronizado e suficientemente estável ao longo do tempo para permitir uma identificação individual confiável e fornecer uma alternativa eficaz a abordagens mais intrusivas. No entanto, o número de recapturas foi baixo, indicando que é necessário um esforço de amostragem muito maior para estudar a dinâmica de dispersão desta espécie usando esta abordagem. Os microssatélites revelaram níveis altos de diferenciação genética e estruturação populacional na área de estudo. No entanto, inferimos níveis modestos de dispersão e identificámos um pequeno número de migrantes e de indivíduos de ascendência mista, sugerindo uma possível dispersão entre vales e conectividade reprodutora recente entre várias populações, com exceção das mais isoladas. A análise das rotas de recolonização de lagos

previamente com peixe e atualmente sem peixe revelou a importância das áreas vizinhas onde os peixes nunca foram introduzidos como fonte de indivíduos. Também foram calculadas as primeiras estimativas do tamanho da população de *C. asper* em dois lagos de alta montanha.

Palavras-chave: *Alytes obstetricans*/*almogavarii*, *Calotriton asper*, fluxo genético, refúgios glaciares, Península Ibérica

Abstract

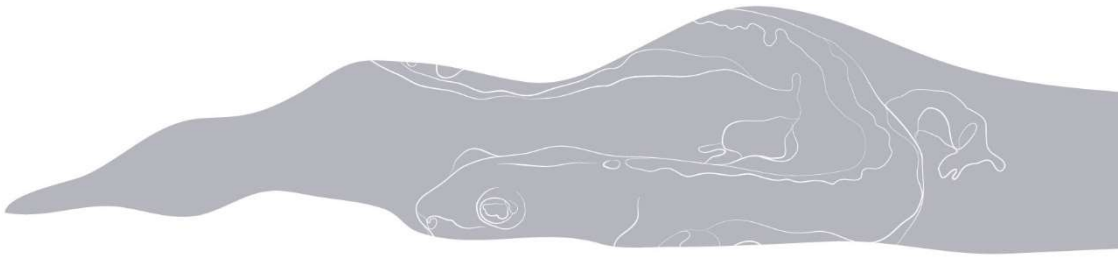
Unveiling how genetic diversity and structure of high mountain populations evolved and what factors are currently shaping it is crucial to predict how species will respond to threats such as climate change, and this will ultimately help to design species-specific conservation measures. Therefore, an integrative approach that combines the study of phylogeographic processes and contemporary dispersal dynamics is required to shed light on the mechanisms underlying spatial patterns of present-day genetic diversity and population structure, which can help to predict species responses to ongoing or future environmental changes. The general aim of this thesis was to uncover major historical and contemporary factors explaining the current genetic diversity and structure of selected Iberian amphibians living in high mountains, namely the Pyrenean brook newt, *Calotriton asper*, and the midwife toads of the *Alytes obstetricans* complex. I integrated various techniques, including genetic and mark-recapture analyses, to provide new insights into the phylogeographic history of the study species, describe the genetic consequences of chytridiomycosis outbreaks in midwife toads, inform on current connectivity in *C. asper*, and characterize its process of lake recolonization following invasive fish removal. Specifically, I showed that *C. asper* likely endured the last glaciation in five distinct glacial refugia across the Pyrenees, and I detected evidence of extensive mito-nuclear discordances and admixture between taxa of the *A. obstetricans* complex across three Iberian mountain ranges. *A. obstetricans* populations hit by a relatively recent chytridiomycosis outbreak had a reduced effective size and were depleted of genetic diversity. Finally, although I found that the majority of *C. asper* dispersal events involved geographically close populations, I also found potential for occasional long-distance dispersal and revealed the importance of neighbouring fishless areas as source of individuals for restored lakes. Overall, the results of this thesis contribute to improve our understanding of the factors responsible for species structuring in high mountains and may be useful to guide future management decisions for *C. asper* and the *A. obstetricans* complex.

Keywords: *Alytes obstetricans*/*almogavarii*, *Calotriton asper*, gene flow, glacial refugia, Iberian Peninsula

Table of Contents

Chapter 1

General introduction



High mountain areas at a glance

From ages, high mountain areas have captivated people around the globe. Such areas were once filled with fascination and mystery, evoking myths and legends in local inhabitants that contributed to the creation of a unique cultural heritage closely linked with landscapes (Speed et al. 2012). In addition to their cultural significance, high mountain systems provide key ecosystem services, the primary ones being provision of water and climate regulation (Foggin 2016). Mountains constitute sources of water for lowlands, especially in regions with relatively dry climate such as the Mediterranean, and can help to mitigate the impact of extreme climatic events (Foggin 2016; Spehn et al. 2010). Furthermore, mountains are generally recognised as hotspots of global biodiversity with unique assemblages of species, including endemic or endangered taxa (Foggin 2016; Spehn et al. 2010).

High mountain lentic habitats and associated threats

Most of the existing high mountain lakes originated during the last glaciation by erosion of glaciers upon crystalline bedrocks (Wetzel 2001; Yao et al. 2018). The peculiar location of high mountain aquatic ecosystems makes them exposed to harsh environmental and climatic conditions that are traditionally considered extreme for life (Catalan et al. 2006). Such ecosystems are usually oligotrophic, due to the low rock weathering of hard crystalline bedrocks, limited soil development, sparse vegetation cover and small catchment size (Skjelkvåle and Wright 1998). Furthermore, lakes and ponds are characterized by the formation and development of an ice and snow cover during several months of the year. This, together with the effects of high altitude and thus thinner atmosphere, makes the aquatic biota typically exposed to high UV radiation during the ice free period and to light-limited conditions in winter (Sommaruga 2001). As a consequence, organisms living in high mountain aquatic habitats face several challenges related to e.g. osmoregulation, nutrient limitation, low prey encounter rates and colonisation of new suitable habitats (Catalan et al. 2006).

Although high-altitude aquatic habitats are generally located in remote areas, they are frequently exposed to several anthropogenic threats, making it hard to find a truly pristine water body in some mountain regions. Major threats are related to either global stressors, such as climate change (Rogora et al. 2003) and long-range atmospheric transport of air pollutants (including acid deposition, persistent organic pollutants and heavy metals) (Camarero and Catalan 1993; Davidson and Knapp 2007; Ørbæk et al. 2007), and local stressors, such as water

exploitation (Godlewska et al. 2003; Toro et al. 2006) and introduction of invasive fish species (Eby et al. 2006).

The introduction of invasive alien fish is posing the greatest threat to high mountain aquatic ecosystems (Ventura et al. 2017). Northern hemisphere high mountain water bodies are naturally fishless, due to physical barriers that have prevented upstream colonisation by fish. Nevertheless, several high mountain areas in the world have been stocked with fish in an effort to create recreational fishing (e.g. Miró and Ventura 2013; Pister 2001). For example, it is estimated that 60% of mountain lakes in the western United States and 72% of the lentic surface area in the Pyrenees contain non-native fish (Bahls 1992; Ventura et al. 2017). When introduced in fishless lakes, fish occupy a higher trophic level that was previously inexistent, leading to profound alterations of food-webs and ecosystem processes (Eby et al. 2006), which in turn can have indirect cascading effects on the surrounding terrestrial habitats (Epanchin et al. 2010) (Figure 1.1). Fish introductions shift community composition and generally result in the extirpation or reduction of native aquatic species, especially amphibians, conspicuous macroinvertebrates and large zooplankton (Knapp et al. 2005; Miró et al. 2018).

The detrimental effects caused by introduced fish have raised concern for the conservation of high mountain lentic habitats and species. To counteract invasive species pressure, numerous restoration projects have taken place in mountain areas around the globe, demonstrating that many of the impacts caused by introduced fish can be reversed over a short period of time (e.g. Bosch et al. 2019; Knapp et al. 2007; Tiberti et al. 2019; Vredenburg 2004). Restoration activities generally consist in the intensive control and, if feasible, eradication of non-native fish through the use of either nets, electrofishing and/or piscicides (e.g. rotenone and antimycin), to allow the recovery of native fauna and habitats (e.g. Gresswell 1991; Tiberti et al. 2019). In the Pyrenees, the conservation project LIFE+ LIMNOPIRINEUS (2014-2019; www.lifelimnopirineus.eu) successfully returned eight high mountain lakes and surrounding streams to their fish-free natural state by means of eradication or control of introduced fish (Figure 1.2). This project opened up a wide array of opportunities to advance the knowledge on the restoration of fish invaded landscapes, including the opportunity to deepen understudied topics such as the process of recolonization of lentic habitats following fish removal and the identification and characterization of amphibian recolonization routes, which form part of this PhD thesis.

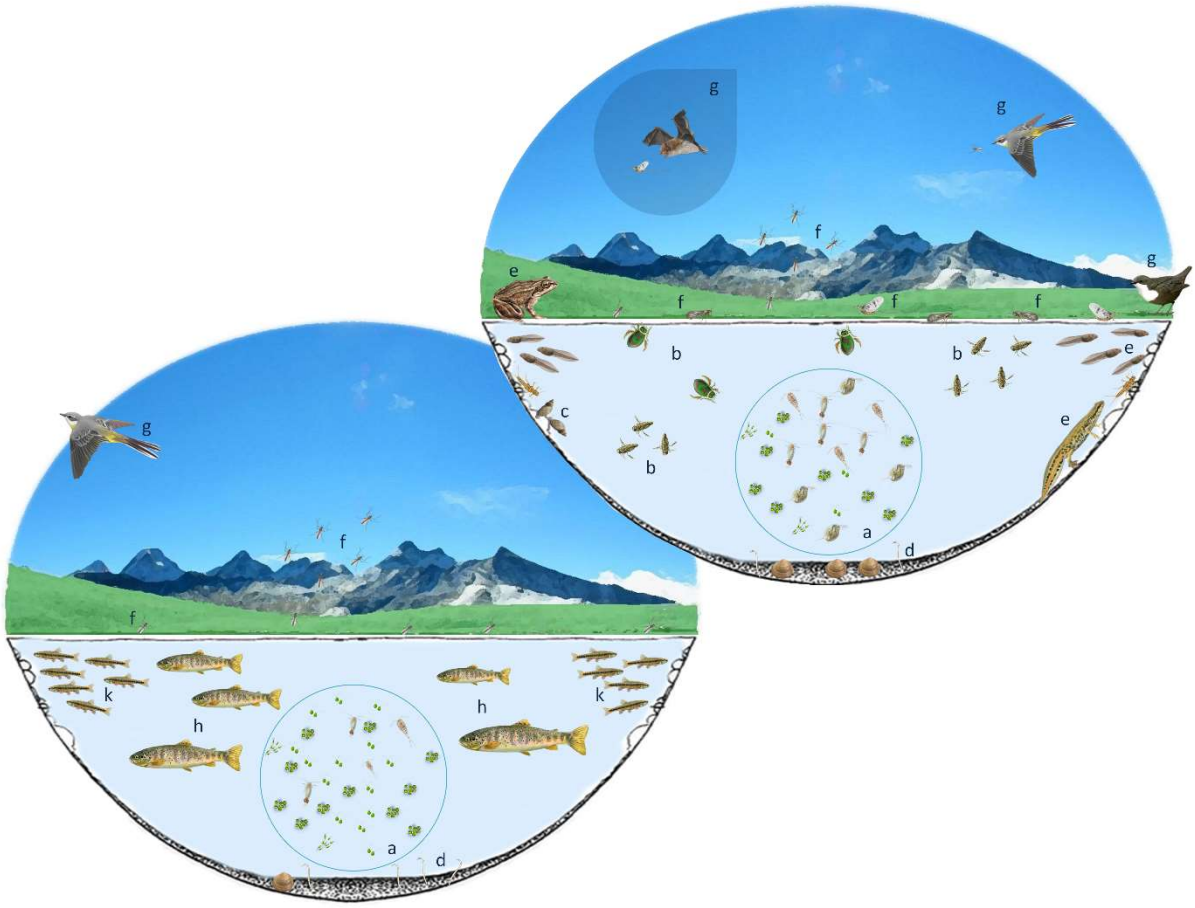


Figure 1.1 Schematic representation of the cascading effects induced by the introduction of fish into originally fishless lakes (taken from Ventura et al. 2017): (a) direct predation can affect large planktonic crustaceans and generate cascading effects altering small zooplankton and phytoplankton communities; (b) nektonic and (c) benthonic macroinvertebrates often experience local extinction following fish introduction, while (d) fossorial macroinvertebrates are generally unaffected; (e) fish introductions often lead to amphibian declines; (f) introduced fish can alter the aquatic nutrient subsidy (in the form of emerging insects and amphibians) entering the terrestrial ecosystem, and indirectly affect (g) terrestrial consumers, e.g. birds, reptiles, amphibians and bats; (h) salmonid fish species are usually introduced in high mountain lakes to sustain recreational fishing; (k) small fish species, such as *Phoxinus* sp., are used as live baits by anglers and accidentally or voluntary released in high mountain lakes with important additive ecological impacts.

Amphibians in high mountains and target species

Nearly half of the world's biodiversity hotspots are located in mountain areas (Foggin 2016). Mountain organisms often present specific adaptations to cope with the harsh and cold conditions in which they live in. This is especially true in high mountain habitats, where a highly diversified and locally adapted flora and fauna are found (Spehn et al. 2010).

Amphibians are among the most conspicuous animal groups in high mountain water bodies. They are highly diverse and have been widely advocated as indicators of environmental health because of their relatively narrow environmental tolerance that makes them very sensitive to habitat alterations (Collins and Storfer 2003; Cooke 1981). Nevertheless, their global status is

of particular concern, as they represent one of the most threatened animal groups (Beebee and Griffiths 2005). Indeed, it is estimated that a third of the world's amphibians have undergone severe decline or extinction and are currently threatened (Stuart et al. 2004).

In high mountains, some of the biggest threats to amphibians include emergent diseases, introduced fish and pesticides (Davidson and Knapp 2007; Joseph and Knapp 2018; Miró et al. 2018). As for emergent diseases, the chytrid fungus *Batrachochytrium dendrobatidis* (*Bd*), which is responsible for causing the disease chytridiomycosis, is a major cause of amphibian population declines and species extinctions in montane habitats worldwide (Beebee and Griffiths 2005; O'Hanlon et al. 2018; Skerratt et al. 2007). With regard to introduced fish, several studies have reported strong negative effects on amphibians, being direct predation and competition for prey the most obvious (Beebee and Griffiths 2005). Nevertheless, studies have also shown that reduction or elimination of invasive fish pressure by means of restoration projects is effective and allows for rapid recovery of high mountain amphibian communities when nearby resilient populations still remain (Knapp et al. 2005; Miró et al. 2020).

Figure 1.2 Three of the lakes targeted by the project LIFE+ LIMNOPIRINEUS in the eastern Pyrenees (Naorte, Rovinets and Closell), before (2014) and after (2018) fish eradication. Fish removal led to a marked increase in the abundance of crustaceans, a decrease in phytoplankton biomass, and a consequent increase in the transparency of the water column.

The Iberian Peninsula is part of the Mediterranean biodiversity hotspot (Myers et al. 2000). This area, which features several large mountain ranges, is especially important for the conservation of amphibians, as it comprises numerous endemic and/or spatially structured species. This leads us to the amphibian species on which this thesis is focused (Figure 1.3): the Pyrenean brook newt *Calotriton asper* (Dugès 1852) and the common midwife toad *Alytes obstetricans* (Laurenti 1768). Both species are protected by EU conservation legislation, as they are listed in Annex IV of the European Directive 92/43/EEC. The Pyrenean brook newt is endemic to the Pyrenees and nearby areas and it is also listed in Appendix II of the Bern Convention and categorized as Near Threatened by the International Union for Conservation of Nature (IUCN). The common midwife toad is a small anuran widely distributed in central and western Europe, which shows strong genetic subdivision in the Iberian Peninsula where four subspecies are found (Gonçalves et al. 2015). Very recently, the subspecies *almogavarii*, which is endemic to Catalonia and adjacent areas in north-eastern Spain and southern France, was recognized as an incipient species (i.e. Catalonian midwife toad, *A. almogavarii*; Dufresnes and Martínez-Solano 2019; Speybroeck et al. 2020). In order to facilitate the flow of the thesis, and given that the suggestion of *A. almogavarii* as a different species is quite recent, I will hereafter refer to these two *Alytes* species either as *A. obstetricans*/*almogavarii* or as *A. obstetricans* complex.

Figure 1.3 Juvenile Pyrenean brook newt (left, photo by Claudine Delmas) and adult male common midwife toad (right, photo by Jaime Bosch).

Iberian high mountain ranges as hotspots of speciation

Due to their exceptional altitudinal gradients leading to high levels of environmental heterogeneity, mountains have provided conditions suitable for speciation and have acted as evolutionary engines due to isolation (Foggin 2016; Nagy and Grabherr 2009). This, coupled

with the climatic fluctuations that characterized the Quaternary Period, strongly contributed to shaping the current distribution and genetic constitution of montane species (Hewitt 1996, 2000). The Quaternary Period, which began ca. 2.6 million years ago and extends into the present, featured alternating series of glacial and interglacial periods that caused repeated changes in species' distributions, leading to periodic events of range contraction and expansion (Hofreiter and Stewart 2009; Schmitt 2007). More specifically, during periods of adverse climatic conditions populations took refuge in isolated and unglaciated glacial refugia, where genetic differentiation and divergence between populations was favoured. As climatic conditions ameliorated and the ice retreated, species expanded their ranges out of the refugia and across formerly glaciated areas, and reclaimed their former distribution. Several species still bear the hallmarks of these past dynamics, with regions of high genetic diversity indicating former refugia and little genetic variation expected in areas claimed postglacially (Excoffier et al. 2009).

Within the Quaternary, the last full glacial cycle from ca. 120,000 years ago to the present is the best understood and is of particular interest when it comes to analysing the evolution and current distribution of species (Hewitt 1999). The Last Glacial Period (ca. 120,000–12,000 years ago) is characterized by alternating episodes of glacier advance and retreat, being the Last Glacial Maximum (LGM) the time when ice sheets were at their greatest extent (Allen et al. 1999; Dansgaard et al. 1993). In Europe, the LGM is estimated to have occurred ca. 30,000–16,000 years ago. After the Last Glacial Period temperatures increased, marking the beginning of the current warm interglacial, i.e. the Holocene.

During the last glaciation, southern Europe hosted three major refugial areas that provided shelter for a variety of species, i.e. Iberia, Italy and the Balkans (Schmitt 2007). The Iberian Peninsula has high physiographic complexity, with several large mountain ranges that allowed survival of populations through long-term isolation in geographically separate refugia (Gómez and Lunt 2007). In fact, growing evidence has revealed complex patterns of glacial survival within Iberia, with multiple refugial areas present within the peninsula, e.g. the Pyrenees, Picos de Europa and the Central System (the so called "refugia-within-refugia" scenario; Figure 1.4) (Abellán and Svenning 2014; Gómez and Lunt 2007). As a consequence, many European species with widespread distribution show strong genetic subdivision in the Iberian Peninsula indicative of past population isolation, such as the common midwife toad (Gonçalves et al.

2015; Maia-Carvalho et al. 2018), the palmate newt *Lissotriton helveticus* (Recuero and Garcia-Paris 2011) and the European beech *Fagus sylvatica* (Magri et al. 2006). At the same time, the complex climatic and topographic features of the Iberian Peninsula allowed for speciation within each mountain range, giving rise to hotspot areas of endemism that served as glacial refugia (Gómez and Lunt 2007). This is the case of the Pyrenees, where many endemic species are found, such as the Pyrenean brook newt (Valbuena-Ureña et al. 2018), the Pyrenean rock lizard *Iberolacerta bonnali* (Mouret et al. 2011), the ground-dwelling spider *Harpactocrates ravastellus* (Bidegaray-Batista et al. 2016) and the dwarf yams *Borderea chouardii* and *B. pyrenaica* (Segarra-Moragues and Catalán 2008). Nevertheless, there are still gaps that need to be filled, and how some of these species endured Quaternary glaciations is still uncertain.

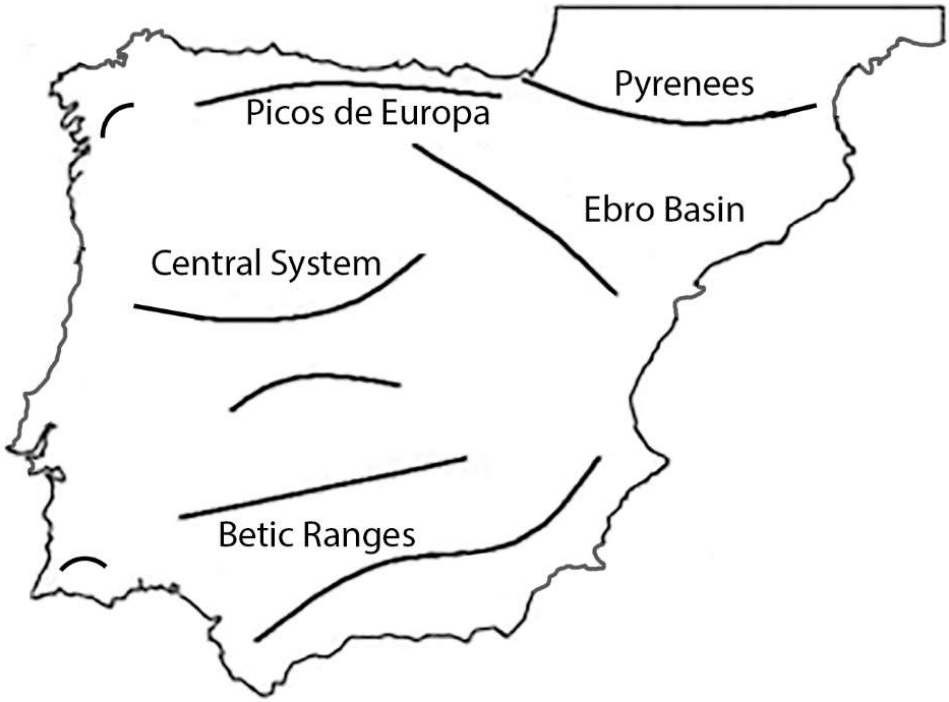


Figure 1.4 Map of the Iberian Peninsula showing the inferred location of major Pleistocene refugia (modified from Gómez and Lunt 2007).

Genetic tools to enhance the understanding of present and past processes

Population structure results primarily from the interplay of present processes and past history (Hewitt and Butlin 1997). In recent times, the use of genetic analytical tools considerably increased our understanding of both present and historical dynamics (Broquet and Petit 2009; Emerson and Hewitt 2005; Hewitt 2004). Indeed, the resolution provided by genetic analyses often outperforms that of classical methods, and cases are known in which molecular methods

revealed misinterpretations and flaws of classical analyses, based on e.g. fossil records and chorological analyses (Schmitt 2007).

Phylogeography, i.e. the study of the historical processes responsible for the contemporary geographic distribution of genetic lineages, is crucial to deduce the evolutionary history of species and subspecies (Avise 2000; Emerson and Hewitt 2005). Recently, a wide array of DNA techniques, jointly with new analytical methods and palaeoclimatic and geological studies are providing crucial insights into the distribution of genetic diversity around the globe, and how it evolved. Genetic data are being increasingly used to estimate the demographic history of populations or species and the processes and paths of postglacial colonisation, including the location of glacial refugia, the size of ancestral populations, the approximate dates of population divergence and the occurrence of bottlenecks or population expansions (Emerson and Hewitt 2005). For the estimation of ancient historical events, more slowly evolving markers are needed, such as mitochondrial or nuclear gene fragments, whereas for the most recent dynamics, in the order of tens of thousands of years, fast evolving or more variable markers are required, such as microsatellites (Cornuet et al. 2010; Emerson and Hewitt 2005). The importance of using a range of markers (i.e. a multilocus approach) is clear if a complete and more reliable picture of the population genetic structure and demographic history of the study species is to be achieved.

Population genetic methods also allow for the rapid assessment of present processes, such as connectivity, dispersal rates and population size (Broquet and Petit 2009; Cayuela et al. 2018). These genetic estimates are particularly valuable for species difficult to see and capture in the wild (Wang et al. 2009) or that are of conservation concern (Valbuena-Ureña et al. 2017). Population genetic estimates have been shown to generally match well with field-ecological estimates (Coulon et al. 2008; Wang and Shaffer 2017; but see Yu et al. 2010; Lowe and Allendorf 2010), offering an efficient alternative or replacement to traditional field data. Indeed, for many species it is probably easier, cheaper and less intrusive to capture and sample individuals in order to perform fast genetic analyses than to carry out multiyear mark-recapture studies. Furthermore, the field of population genetics has witnessed a quick growth that has led to a proliferation of analytical methods and molecular approaches that may provide deeper insights into patterns of movement, habitat preference and gene flow in an accurate and effective way (Bergl and Vigilant 2007; Cayuela et al. 2018).

General overview of the thesis and chapters outline

The general aim of this thesis was to uncover major historical and contemporary factors explaining the current genetic diversity and structure of species living in high mountains. To this end, I focused on two amphibian species (the Pyrenean brook newt *C. asper* and the midwife toads of the *A. obstetricans* complex) to draw broad conclusions from specific instances, with the ultimate goal of aiding in the description of past and present processes in high mountain areas. Using this approach, this work describes some fundamental aspects of the phylogeographic history of the target species, while also analysing present-day dynamics. This thesis is based on extensive fieldwork carried out during the course of this thesis between 2017 and 2018 and relies also on previous field campaigns, which made it possible to collect ca. 1,300 samples for *C. asper* and ca. 900 samples for *A. obstetricans*/*almogavarii* from a significant portion of the species ranges. This was possible thanks to national and international collaborators, whose precious contribution helped achieving a more complete picture of the population structure and history of the studied species.

The chapters of this thesis can be divided into two parts: part 1, consisting of chapters 2 and 3, is concerned with the study of the phylogeographic history of the target species, and part 2, consisting of chapter 4, where I investigate patterns of recent dispersal and colonization dynamics in *C. asper*. Below I provide a brief description of each chapter:

- Chapter 2: Multiple glacial refugia and contemporary dispersal shape the genetic structure of an endemic amphibian from the Pyrenees (Lucati et al. 2020b).
  The main goal of this study was to unravel the phylogeographic history and determine the degree of connectivity of present-day *C. asper* populations and habitats at the Pyrenean scale.

- Chapter 3: Contrasting patterns of genetic admixture explain the phylogeographic history of Iberian high mountain populations of midwife toads (under review, submitted to *Heredity*).
  This chapter shed light into the drivers of geographical differentiation in Iberian high mountain populations of the midwife toads *A. obstetricans*/*almogavarii*, using a comprehensive set of mitochondrial and nuclear markers.

- Chapter 4: A genetic appraisal of regional population connectivity in the endemic Pyrenean amphibian *Calotriton asper* and patterns of recolonization after fish removal (in preparation).

  This chapter stems from a grant I was awarded in 2018 by the Societas Europaea Herpetologica (SEH) (Appendix 1; Lucati et al. 2020a), with the purpose of studying the factors affecting population connectivity and the patterns of recolonization following fish removal in *C. asper* on a local scale, using a combination of genetic and photographic mark-recapture analyses. The validity of photographic mark-recapture for *C. asper* individual identification was also tested.

References

Abellán P, Svenning J-C (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.

Allen JR, Brandt U, Brauer A, Hubberten H-W, Huntley B, Keller J et al. (1999) Rapid environmental changes in southern Europe during the last glacial period. *Nature* **400:** 740-743.

Avise JC (2000) *Phylogeography: The history and formation of species.* Harvard University Press: Cambridge, MA, USA.

Bahls P (1992) The status of fish populations and management of high mountain lakes in the western United States. *Northwest Science* **66:** 183-193.

Beebee TJ, Griffiths RA (2005) The amphibian decline crisis: A watershed for conservation biology? *Biological Conservation* **125:** 271-285.

Bergl RA, Vigilant L (2007) Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*). *Molecular Ecology* **16:** 501-516.

Bidegaray-Batista L, Sánchez-Gracia A, Santulli G, Maiorano L, Guisan A, Vogler AP et al. (2016) Imprints of multiple glacial refugia in the Pyrenees revealed by phylogeography and palaeodistribution modelling of an endemic spider. *Molecular Ecology* **25:** 2046-2064.

Bosch J, Bielby J, Martin-Beyer B, Rincón P, Correa-Araneda F, Boyero L (2019) Eradication of introduced fish allows successful recovery of a stream-dwelling amphibian. *PLoS One* **14:** e0216204.

Broquet T, Petit EJ (2009) Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics* **40:** 193-216.

Camarero L, Catalan J (1993) Chemistry of bulk precipitation in the central and eastern Pyrenees, northeast Spain. *Atmospheric Environment Part A General Topics* **27:** 83-94.

Catalan J, Camarero L, Felip M, Pla S, Ventura M, Buchaca T et al. (2006) High mountain lakes: Extreme habitats and witnesses of environmental changes. *Limnetica* **25:** 551-584.

Cayuela H, Rougemont Q, Prunier JG, Moore JS, Clobert J, Besnard A et al. (2018) Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Molecular Ecology* **27:** 3976-4010.

Collins JP, Storfer A (2003) Global amphibian declines: Sorting the hypotheses. *Diversity and Distributions* **9:** 89-98.

Cooke A (1981) Tadpoles as indicators of harmful levels of pollution in the field. *Environmental Pollution Series A, Ecological and Biological* **25:** 123-133.

Cornuet JM, Ravigne V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* **11:** 401.

Coulon A, Fitzpatrick JW, Bowman R, Stith BM, Makarewich CA, Stenzler LM et al. (2008) Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (*Aphelocoma coerulescens*). *Molecular Ecology* **17:** 1685-1701.

Dansgaard W, Johnsen SJ, Clausen HB, Dahl-Jensen D, Gundestrup N, Hammer C et al. (1993) Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature* **364:** 218-220.

Davidson C, Knapp RA (2007) Multiple stressors and amphibian declines: Dual impacts of pesticides and fish on yellow-legged frogs. *Ecological Applications* **17:** 587-597.

Dufresnes C, Martínez-Solano Í (2019) Hybrid zone genomics supports candidate species in Iberian *Alytes obstetricans*. *Amphibia-Reptilia* **41:** 105-112.

Eby LA, Roach WJ, Crowder LB, Stanford JA (2006) Effects of stocking-up freshwater food webs. *Trends in Ecology & Evolution* **21:** 576-584.

Emerson BC, Hewitt GM (2005) Phylogeography. *Current Biology* **15:** R367-R371.

Epanchin PN, Knapp RA, Lawler SP (2010) Nonnative trout impact an alpine-nesting bird by altering aquatic-insect subsidies. *Ecology* **91:** 2406-2415.

Excoffier L, Foll M, Petit R (2009) Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, Systematics* **40:** 481-501.

Foggin J (2016) *Conservation issues: Mountain ecosystems. Reference module in Earth systems and environmental sciences*. Elsevier: Toronto, Canada.

Godlewska M, Mazurkiewicz-Boroń G, Pociecha A, Wilk-Woźniak E, Jelonek M (2003) Effects of flood on the functioning of the Dobczyce reservoir ecosystem. *Hydrobiologia* **504:** 305-313.

Gómez A, Lunt DH (2007) Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S and Ferrand N (eds) *Phylogeography of southern European refugia*. Springer: Amsterdam, Netherlands, pp 155-188.

Gonçalves H, Maia-Carvalho B, Sousa-Neves T, Garcia-Paris M, Sequeira F, Ferrand N et al. (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.

Gresswell RE (1991) Use of antimycin for removal of brook trout from a tributary of Yellowstone Lake. *North American Journal of Fisheries Management* **11:** 83-90.

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 (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68:** 87-112.

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; discussion 195.

Hewitt GM, Butlin RK (1997) Causes and consequences of population structure. In: Krebs JR and Davies N (eds) *Behavioral Ecology, 4th edn.* Blackwell: Oxford, pp 350-372.

Hofreiter M, Stewart J (2009) Ecological change, range fluctuations and population dynamics during the Pleistocene. *Current Biology* **19:** R584-R594.

Joseph MB, Knapp RA (2018) Disease and climate effects on individuals drive post-reintroduction population dynamics of an endangered amphibian. *Ecosphere* **9:** e02499.

Knapp RA, Boiano DM, Vredenburg VT (2007) Removal of nonnative fish results in population expansion of a declining amphibian (mountain yellow-legged frog, *Rana muscosa*). *Biological Conservation* **135:** 11-20.

Knapp RA, Hawkins CP, Ladau J, McClory JG (2005) Fauna of Yosemite National Park lakes has low resistance but high resilience to fish introductions. *Ecological Applications* **15:** 835-847.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology* **19:** 3038-3051.

Lucati F, Miró A, Ventura M (2020a) Conservation of the endemic Pyrenean newt (*Calotriton asper*) in the age of invasive species: Interlake dispersal and colonisation dynamics. *Amphibia-Reptilia* **1:** 1-2.

Lucati F, Poignet M, Miró A, Trochet A, Aubret F, Barthe L et al. (2020b) Multiple glacial refugia and contemporary dispersal shape the genetic structure of an endemic amphibian from the Pyrenees. *Molecular Ecology* **29:** 2904-2921.

Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D et al. (2006) A new scenario for the Quaternary history of European beech populations: Palaeobotanical evidence and genetic consequences. *New Phytologist* **171:** 199-221.

Maia-Carvalho B, Vale CG, Sequeira F, Ferrand N, Martínez-Solano I, Gonçalves H (2018) The roles of allopatric fragmentation and niche divergence in intraspecific lineage diversification in the common midwife toad (*Alytes obstetricans*). *Journal of Biogeography* **45:** 2146-2158.

Miró A, O'Brien D, Tomàs J, Buchaca T, Sabás I, Osorio V et al. (2020) Rapid amphibian community recovery following removal of non-native fish from high mountain lakes. *Biological Conservation* **251:** 108783.

Miró A, Sabás I, Ventura M (2018) Large negative effect of non-native trout and minnows on Pyrenean lake amphibians. *Biological Conservation* **218:** 144-153.

Miró A, Ventura M (2013) Historical use, fishing management and lake characteristics explain the presence of non-native trout in Pyrenean lakes: Implications for conservation. *Biological Conservation* **167:** 17-24.

Mouret V, Guillaumet A, Cheylan M, Pottier G, Ferchaud AL, Crochet PA (2011) The legacy of ice ages in mountain species: Post-glacial colonization of mountain tops rather than current range fragmentation determines mitochondrial genetic diversity in an endemic Pyrenean rock lizard. *Journal of Biogeography* **38:** 1717-1731.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* **403:** 853-858.

Nagy L, Grabherr G (2009) *The biology of alpine habitats*. Oxford University Press: Oxford, UK.

O'Hanlon SJ, Rieux A, Farrer RA, Rosa GM, Waldman B, Bataille A et al. (2018) Recent Asian origin of chytrid fungi causing global amphibian declines. *Science* **360:** 621-627.

Ørbæk JB, Kallenborn R, Tombre I, Hegseth EN, Falk-Petersen S, Hoel AH (2007) *Arctic alpine ecosystems and people in a changing environment*. Springer: Berlin, Germany.

Pister EP (2001) Wilderness fish stocking: History and perspective. *Ecosystems* **4:** 279-286.

Recuero E, Garcia-Paris 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.

Rogora M, Mosello R, Arisci S (2003) The effect of climate warming on the hydrochemistry of alpine lakes. *Water, Air, and Soil Pollution* **148:** 347-361.

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

Segarra-Moragues JG, Catalán P (2008) Glacial survival, phylogeography, and a comparison of microsatellite evolution models for resolving population structure in two species of dwarf yams (*Borderea*, Dioscoreaceae) endemic to the central Pyrenees. *Plant Ecology & Diversity* **1:** 229-243.

Skerratt LF, Berger L, Speare R, Cashins S, McDonald KR, Phillott AD et al. (2007) Spread of chytridiomycosis has caused the rapid global decline and extinction of frogs. *EcoHealth* **4:** 125-134.

Skjelkvåle BL, Wright RF (1998) Mountain lakes; sensitivity to acid deposition and global climate change. *Ambio* **27:** 280-286.

Sommaruga R (2001) The role of solar UV radiation in the ecology of alpine lakes. *Journal of Photochemistry and Photobiology B: Biology* **62:** 35-42.

Speed JD, Austrheim G, Birks HJB, Johnson S, Kvamme M, Nagy L et al. (2012) Natural and cultural heritage in mountain landscapes: Towards an integrated valuation. *International Journal of Biodiversity Science, Ecosystem Services & Management* **8:** 313-320.

Spehn EM, Rudmann-Maurer K, Körner C, Maselli D (2010) *Mountain biodiversity and global change*. GMBA-DIVERSITAS: Basel, Switzerland.

Speybroeck J, Beukema W, Dufresnes C, Fritz U, Jablonski D, Lymberakis P et al. (2020) Species list of the European herpetofauna–2020 update by the Taxonomic Committee of the Societas Europaea Herpetologica. *Amphibia-Reptilia* **1:** 1-51.

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

Tiberti R, Bogliani G, Brighenti S, Iacobuzio R, Liautaud K, Rolla M et al. (2019) Recovery of high mountain Alpine lakes after the eradication of introduced brook trout *Salvelinus fontinalis* using non-chemical methods. *Biological Invasions* **21:** 875-894.

Toro M, Granados I, Robles S, Montes C (2006) High mountain lakes of the Central Range (Iberian Peninsula): Regional limnology & environmental changes. *Limnetica* **25:** 217-252.

Valbuena-Ureña E, Oromi N, Soler-Membrives A, Carranza S, Amat F, Camarasa S et al. (2018) Jailed in the mountains: Genetic diversity and structure of an endemic newt species across the Pyrenees. *PLoS One* **13:** e0200214.

Valbuena-Ureña E, Soler-Membrives A, Steinfartz S, Orozco-terWengel P, Carranza S (2017) No signs of inbreeding despite long-term isolation and habitat fragmentation in the critically endangered Montseny brook newt (*Calotriton arnoldi*). *Heredity* **118:** 424-435.

Ventura M, Tiberti R, Buchaca T, Buñay D, Sabás I, Miró A (2017) Why should we preserve fishless high mountain lakes? In: Catalan J, Ninot J and Aniz M (eds) *Advances in Global Change Research*. Springer: Cham, ZG. Vol. 62: High mountain conservation in a changing world, pp 181-205.

Vredenburg VT (2004) Reversing introduced species effects: Experimental removal of introduced fish leads to rapid recovery of a declining frog. *Proceedings of the National Academy of Sciences* **101:** 7646-7650.

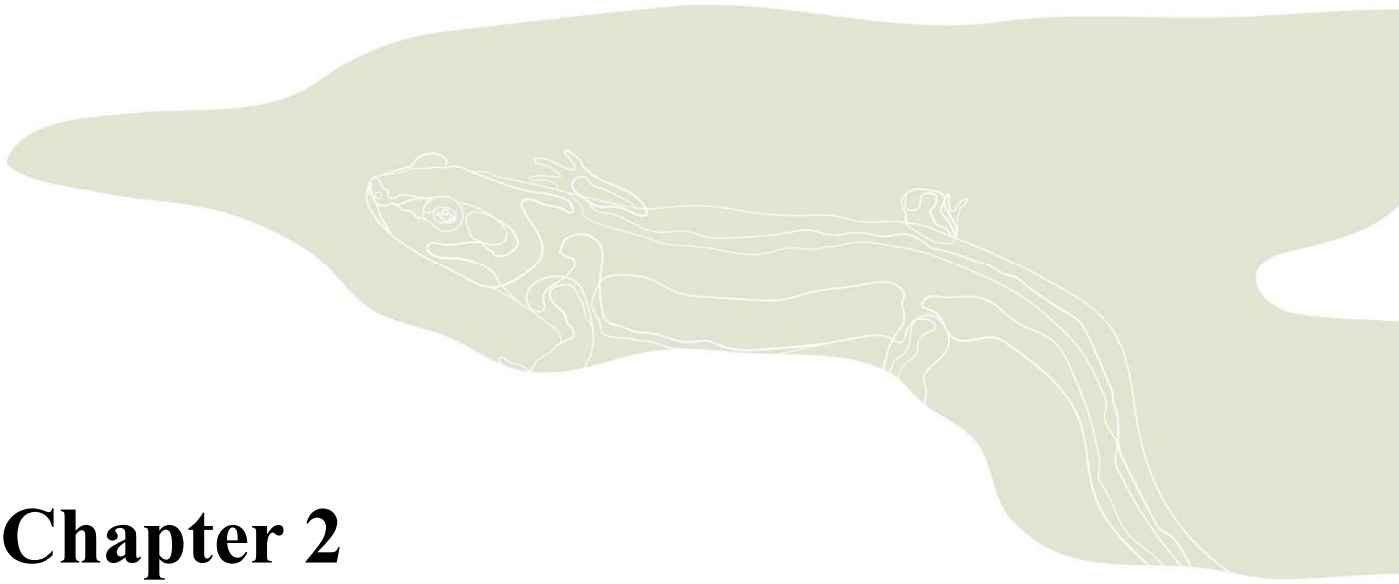
Wang IJ, Savage WK, Bradley Shaffer H (2009) Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (*Ambystoma californiense*). *Molecular Ecology* **18:** 1365-1374.

Wang IJ, Shaffer HB (2017) Population genetic and field-ecological analyses return similar estimates of dispersal over space and time in an endangered amphibian. *Evolutionary Applications* **10:** 630-639.

Wetzel RG (2001) *Limnology: Lake and river ecosystems*, 3rd edn. Academic Press: San Diego, USA.

Yao X, Liu S, Han L, Sun M, Zhao L (2018) Definition and classification system of glacial lake for inventory and hazards study. *Journal of Geographical Sciences* **28:** 193-205.

Yu H, Nason JD, Ge X, Zeng J (2010) Slatkin's Paradox: When direct observation and realized gene flow disagree. A case study in *Ficus*. *Molecular Ecology* **19:** 4441-4453.



Chapter 2

Multiple glacial refugia and contemporary dispersal shape the genetic structure of an endemic amphibian from the Pyrenees

Federica Lucati, Manon Poignet, Alexandre Miró, Audrey Trochet, Fabien Aubret, Laurent Barthe, Romain Bertrand, Teresa Buchaca, Jenny Caner, Elodie Darnet, Mathieu Denoël, Olivier Guillaume, Hugo Le Chevalier, Albert Martínez-Silvestre, Marc Mossoll-Torres, David O'Brien, Calvez Olivier, Víctor Osorio, Gilles Pottier, Murielle Richard, Ibor Sabás, Jérémie Souchet, Jan Tomàs, Marc Ventura

Lucati F, Poignet M, Miró A, Trochet A, Aubret F, Barthe L, Bertrand R, Buchaca T, Calvez O, Caner J, Darnet E, Denoël M, Guillaume O, Le Chevalier H, Martínez-Silvestre A, Mossoll-Torres M, O'Brien D, Osorio V, Pottier G, Richard M, Sabás I, Souchet J, Tomàs J, Ventura M (2020) Multiple glacial refugia and contemporary dispersal shape the genetic structure of an endemic amphibian from the Pyrenees. *Molecular Ecology* **29:** 2904–2921. https://doi.org/10.1111/mec.15521

Abstract

Historical factors (colonization scenarios, demographic oscillations) and contemporary processes (population connectivity, current population size) largely contribute to shaping species' present-day genetic diversity and structure. In this study, we use a combination of mitochondrial and nuclear DNA markers to understand the role of Quaternary climatic oscillations and present-day gene flow dynamics in determining the genetic diversity and structure of the newt *Calotriton asper* (Dugès, 1852), endemic to the Pyrenees. Mitochondrial DNA did not show a clear phylogeographic pattern and presented low levels of variation. In contrast, microsatellites revealed five major genetic lineages with admixture patterns at their boundaries. Approximate Bayesian computation analyses and linear models indicated that the five lineages likely underwent separate evolutionary histories and can be tracked back to distinct glacial refugia. Lineage differentiation started around the Last Glacial Maximum at three focal areas (western, central and eastern Pyrenees) and extended through the end of the Last Glacial Period in the central Pyrenees, where it led to the formation of two more lineages. Our data revealed no evidence of recent dispersal between lineages, whereas borders likely represent zones of secondary contact following expansion from multiple refugia. Finally, we did not find genetic evidence of sex-biased dispersal. This work highlights the importance of integrating past evolutionary processes and present-day gene flow and dispersal dynamics, together with multilocus approaches, to gain insights into what shaped the current genetic attributes of amphibians living in montane habitats.

Keywords: *Calotriton*, genetic structure, phylogeographic history, Pyrenean brook newt, recent dispersal, Pyrenees

Introduction

Unveiling the mechanisms driving species genetic diversity and structure is of crucial interest in phylogeography (Avise 2000). The extent of genetic structure of a species is regarded to result primarily from the interplay of historical factors (e.g. colonization scenarios, demographic oscillations) and current population connectivity, namely gene flow (Hewitt and Butlin 1997; Nichols and Beaumont 1996). Unravelling the phylogeographic history of species and populations is important to understand their present-day and future distribution, genetic structure and adaptations (Hewitt 2004). Historical processes are largely dependent on past climatic conditions and geological events. Such climatic and geological changes have significantly contributed to laying the genetic foundations of contemporary populations, which can be used to make inferences on their past dynamics (Cabrera and Palsbøll 2017; Hewitt and Butlin 1997). In addition, dispersal, which can include gene flow, is a significant component of metapopulation structure and dynamics and can counteract both neutral and selective processes (Johnson and Gaines 1990; Ronce 2007; Tallmon et al. 2004). A reduction in connectivity will ultimately result in a lack of dispersal among populations, increasing the risk of genetic variability loss (Ronce, 2007; but see Orsini et al. 2013). For this reason, dispersal is deemed crucial for the long-term survival of populations under changing conditions (Saccheri et al. 1998). In some circumstances, other processes might explain the genetic variability of populations, such as isolation by environment (reduction in gene flow among ecologically divergent habitats as a result of local adaptation) and by colonisation (reduction in gene flow among all populations in the landscape caused by local genetic adaptation following colonisation; Orsini et al. 2013). The literature on historic vs. contemporary mechanisms shaping the genetic attributes of species is mostly focused on either landscape genetics or dispersal processes alone, or tackles temporal dynamics dealing with the relatively recent past (Chiucchi and Gibbs 2010; Epps and Keyghobadi 2015; Noguerales et al. 2017; Zellmer and Knowles 2009). An integrative approach that combines the study of past evolutionary and phylogeographic processes and present-day gene flow and dispersal dynamics is required to shed light on the mechanisms underlying spatial patterns of contemporary genetic diversity and population structure, which can ultimately help to predict their responses to ongoing or future environmental changes.

In Europe, Quaternary climatic oscillations played a major role in shaping the geographic distribution and genetic constitution of species (Hewitt 2000, 2004). Glacial and interglacial

periods caused repeated changes in species' distributions, leading to events of contraction and expansion and, consequently, to periodic waves of colonization or recolonization. Mountain ranges across Europe are regarded as biodiversity cradles, where diversification is promoted during periods when species' ranges are restricted to geographically isolated glacial refugia (Hewitt 2000; Schmitt 2009). As glaciers repeatedly advance and retreat, species are displaced outside or to the margin of mountain systems into lowland and peripheral areas, respectively, or survive in nunataks, namely areas above glaciers not covered with ice (Holderegger and Thiel-Egenter 2009). Mountain ecosystems are home to many endemisms that still carry genetic imprints of these past dynamics, and thus represent excellent models with which to study the influence of climatic fluctuations on the diversification and postglacial colonization of species (Schmitt 2009).

As one of the major European mountain ranges and separating the Iberian Peninsula from the rest of continental Europe, the Pyrenees played a considerable role in limiting postglacial dispersal routes of numerous temperate species (Taberlet et al. 1998). During glacial periods, the Pyrenees were largely covered with ice (Calvet 2004; González-Sampériz et al. 2006). Nevertheless, it is suggested that some species could have survived glaciations in ice-free areas along the chain, such as nunataks and peripheral lower areas that served as glacial refugia (Bidegaray-Batista et al. 2016; Charrier et al. 2014; Liberal et al. 2014; Mouret et al. 2011). Following the end of glacial periods, deglaciation allowed recolonization along routes spreading from these refugia and this ultimately sculptured a complex genetic structure in the Pyrenees (Hewitt 1999; Taberlet et al. 1998). However, there has been little attempt to identify the geographic location of putative refugia where Pyrenean endemics survived glaciations, and to trace back their postglacial recolonization routes.

Dispersal capability is a crucial trait affecting the genetic composition of species and populations (Clobert et al. 2009; Ronce 2007; Tallmon et al. 2004), implying that variation in vagility generally leads to clear differences in genetic patterns. Good dispersers are likely to present less structured metapopulations than low vagility organisms (Allentoft et al. 2009; Burns et al. 2004; Kraaijeveld-Smit et al. 2005; Vos et al. 2001). Amphibians are generally regarded as low vagility and philopatric species (Gill 1978) but this is being confuted in a number of studies (Denoël et al. 2018; Smith and Green 2005, 2006). Selective pressures favouring or restraining dispersal may act differently on males and females and result in sex-

specific dispersal strategies (Li and Kokko 2019). Accordingly, sex-biased dispersal has been identified in a number of species, including newts (Denoël et al. 2018; Trochet et al. 2016). Furthermore, orographic features such as ridges and valleys can act as either barriers or bridges to dispersal and thus drive genetic structuring (Caplat et al. 2016; Noguerales et al. 2016). Although it is deemed important to better understand the processes underlying genetic differentiation in natural populations, the combined influence of sex differences and orographic features on dispersal has rarely been studied (Roffler et al. 2014; Tucker et al. 2017). Indeed, males and females may have different dispersal abilities and therefore orographic features may differently affect them, resulting in contrasting patterns of gene flow between sexes in mountain regions (see Cayuela et al. 2020 for a review).

The genus *Calotriton* (Gray, 1858) includes two species restricted to north-eastern Iberian Peninsula (Carranza and Amat 2005). Speciation within the genus has been dated to the beginning of the Pleistocene (Carranza and Amat 2005) but how these species endured Quaternary glaciations is still uncertain. The Pyrenean brook newt *C. asper* (Dugès, 1852) is a small-bodied amphibian endemic to the Pyrenees (Bosch et al. 2009). It is a largely aquatic montane species that inhabits brooks, alpine lakes and caves between 250 and 2,500 m a.s.l. (Clergue-Gazeau and Martínez-Rica 1978; Martínez-Rica and Clergue-Gazeau 1977). As expected for many amphibian species, *C. asper* is believed to have low dispersal ability (Milá et al. 2010; Montori et al. 2008b), although little attention has been paid to this aspect. Following metamorphosis, a juvenile dispersal phase of at least 2 years is described before reaching the adult stage (Montori and Llorente 2014), but it remains unclear how far individuals can disperse.

So far, few studies have analysed the genetic differentiation of the Pyrenean brook newt in a geographic context. Analysis of allozymes (Montori et al. 2008a) and mitochondrial DNA (mtDNA; Milá et al. 2010; Valbuena-Ureña et al. 2013) revealed low levels of genetic variation. Higher levels of genetic differentiation and population structuring were detected using genome-wide amplified fragment length polymorphism (AFLP; Milá et al. 2010) and microsatellite markers (Valbuena-Ureña et al. 2018). However, these studies were either based on small numbers of populations and markers with low variability (Montori et al. 2008a; Valbuena-Ureña et al. 2013), did not characterize the entire range and habitat types of the species (Milá et al. 2010), or addressed specific questions targeting the role of geographic gradients and

habitat type in shaping the current genetic attributes of the species (Valbuena-Ureña et al. 2018). Furthermore, the timing of lineage divergence and the relative importance of phylogeographic processes versus contemporary dispersal have not been studied in *C. asper*.

Here, we employ a multilocus approach aimed to disentangle major historical and contemporary processes that contributed to shaping the present genetic constitution of *C. asper* over most of its distribution range. We combine comprehensive sample collection across all habitat types with coalescent model frameworks and dispersal analyses to shed light on the evolutionary history of the species and determine the degree of connectivity of present-day populations and habitats. Specifically, we explore the effect of Quaternary climatic oscillations on the evolutionary diversification of lineages and the formation of postglacial colonization routes. Furthermore, we describe contemporary patterns of dispersal and investigate whether sex-specific dispersal strategies, orography or geography played a role in determining the species' current genetic structure.

Materials and Methods

Sampling and DNA extraction

Sampling was conducted in the period 2004-2017 across the whole Pyrenees (Figure 2.1; Table 2.S1), encompassing most of the species range. DNA was sampled via buccal swab or toe clipping of metamorphosed individuals. Samples were preserved in EDTA or absolute ethanol and stored at -20°C until DNA extraction. The collection of samples was approved by the corresponding authorities: as for the French sampling, by the Conseil Scientifique Régional du Patrimoine Naturel (CSRPN, DREAL) of the Region of Occitanie; as for the Andorran sampling, by the Principality of Andorra; as for the Spanish sampling, by the Departament d'Agricultura, Ramaderia, Pesca, Alimentació i Medi Natural of the Catalan Government and the Instituto Aragonés de Gestión Ambiental of the Aragonese Government. Procedures followed guidelines established by the Association for the Study of Animal Behaviour and complied with current French, Andorran and Spanish regulations.

Genomic DNA was extracted using QIAGEN DNeasy Blood and Tissue Kit (Qiagen^{TM}, Hilden, Germany) according to the manufacturer's protocol, or following the HotSHOT method (Montero-Pau et al. 2008), in a total volume of 100 µl.

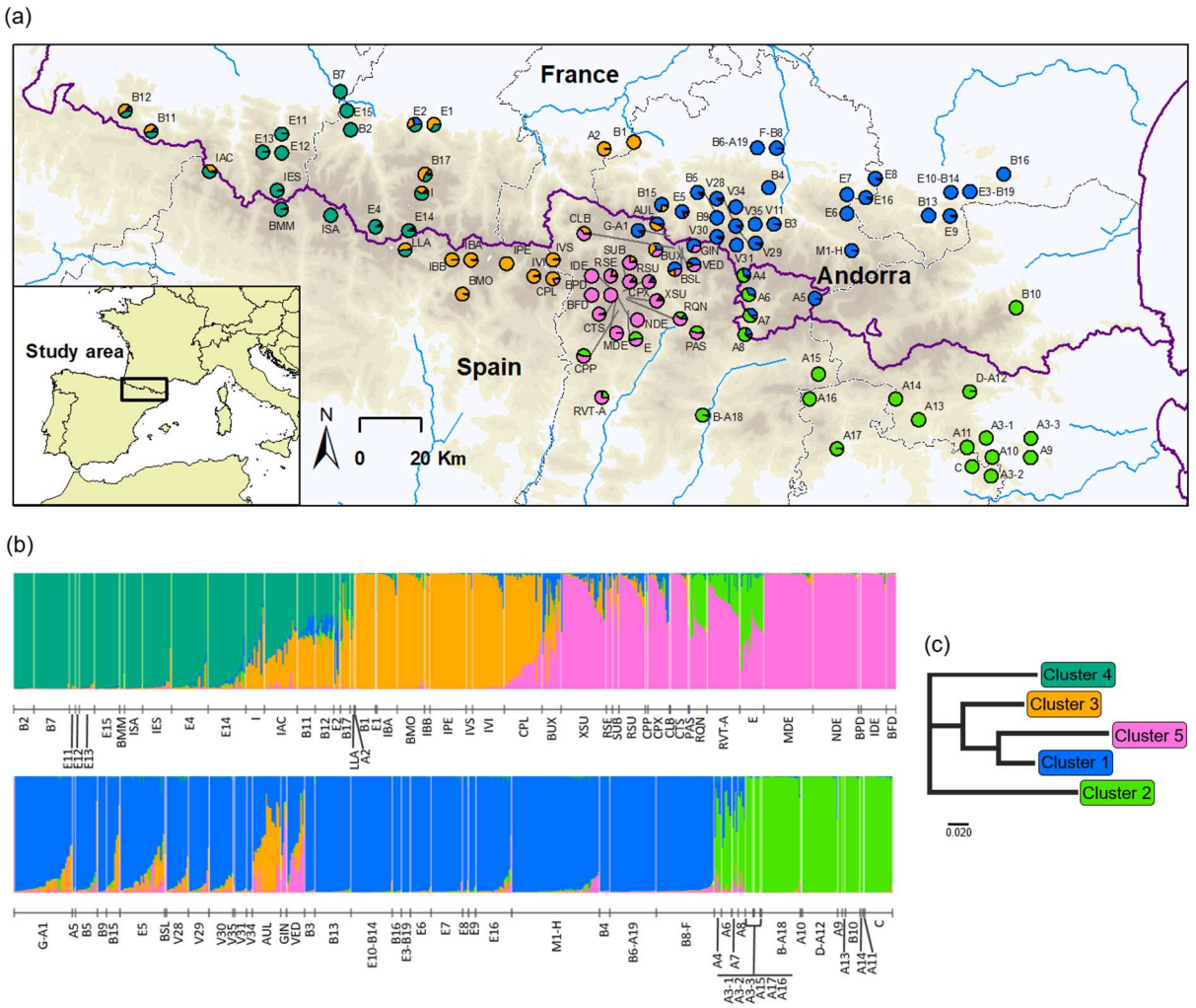


Figure 2.1 Results of the Bayesian clustering analysis across *Calotriton asper* distribution range. Panel (a) shows the geographic distribution of the five genetic clusters identified by STRUCTURE. Sampled populations are represented by pie charts highlighting the population cluster membership obtained in STRUCTURE. Panel (b) shows STRUCTURE barplot of membership assignment for K = 5. Each individual is represented by a vertical bar corresponding to the sum of assignment probabilities to the K cluster. White lines separate populations. Panel (c) represents a neighbour-joining tree based on net nucleotide distances among clusters inferred by STRUCTURE. For population codes see Table 2.S1.

Mitochondrial DNA sequencing and microsatellite screening

A fragment of the cytochrome *b* (cyt-*b*) gene was sequenced from 258 individuals from 59 sampling sites (Table 2.S1). We amplified a fragment of 374 bp using primers Cytb1EuprF and Cytb2EuprR (Carranza and Amat 2005). Amplification conditions were those described in Carranza et al. (2000). Sequences were aligned using the ClustalW algorithm in MEGA 7 (Kumar et al. 2016).

A total of 1,299 individuals from 96 sampling sites were genotyped for a set of 17 microsatellite loci combined in three multiplexes (Table 2.S1; Drechsler et al. 2013). Fragments were sized

with LIZ-500 size standard and binned using either GeneMapper 4.0 (Applied Biosystems) or Geneious 11.0.5 (Kearse et al. 2012). Only individuals that could be scored in a reliable manner for at least 15 loci were included in the analyses.

Mitochondrial DNA analysis

Gene genealogy networks were generated using Haploviewer (Salzburger et al. 2011). jModelTest 2.1.3 (Darriba et al. 2012) was run to determine the appropriate nucleotide-substitution model, under the Akaike Information Criterion (AIC). Phylogenetic reconstructions among haplotypes were estimated using a maximum likelihood approach as implemented in RAxML 7.7.1 (Stamatakis 2006), and the best generated tree was used to estimate the haplotype network. The program was run with a GTRCAT model of rate heterogeneity and no invariant sites, applying 1,000 bootstrap replicates. Haplotype network reconstruction was implemented in Haploviewer, based on all sequences available from GenBank and this study. Overall number of haplotypes (H) and polymorphic sites (S), as well as haplotype (Hd) and nucleotide (Π) diversity indices were calculated in DnaSP 6.11.01 (Rozas et al. 2017).

Microsatellite analysis

The presence of potential scoring errors, stuttering, large allele dropout and null alleles was tested using MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004). The frequency of null alleles for each locus and population was further investigated using the expectation maximization algorithm implemented in FreeNA (Chapuis and Estoup 2006). The same program was used to calculate global F_{ST} values corrected for null alleles following the Excluding Null Alleles (ENA) correction method. Bootstrap 95% confidence intervals (CI) were calculated using 1,000 replicates over loci. We tested for linkage disequilibrium between loci and for deviations from Hardy-Weinberg equilibrium (HWE) in each population and for each locus in GENEPOP 4.2 (Rousset 2008). Significance levels for multiple comparisons were adjusted using the Bonferroni correction ($\alpha = 0.05$; Rice 1989).

Parameters of genetic diversity were estimated for populations with five or more genotyped individuals and for the genetic clusters inferred by STRUCTURE. Calculation of diversity estimates only in populations with larger sample size ($\geq$ 10 individuals) yielded very similar results in terms of mean genetic diversity and in the spatial interpolation analysis. We calculated

observed (H_O) and expected heterozygosity (H_E) using the PopGenKit R package (Rioux Paquette 2011) in R 3.5.1 (R Core Team 2018). Allelic richness (Ar) standardized for sample size and rarefied private allelic richness (PAAr; calculated only at the cluster level) were calculated in HP-RARE 1.1 (Kalinowski 2005). Inbreeding coefficients (F_{IS}) were estimated in FSTAT 2.9.3.2 (Goudet 2002). We visualised geographic patterns of genetic diversity by computing a spatial interpolation of H_E and Ar values using the Inverse Distance Weighting tool implemented in ArcGIS 10.1 (ESRI, Redlands, CA, USA).

Population structure was investigated using a Bayesian approach implemented in STRUCTURE 2.3.4 (Pritchard et al. 2000). We conducted 20 independent simulations for each K value from one to 50, with 100K burn-in steps followed by 500K Markov chain Monte Carlo (MCMC) repetitions. It is highly unlikely that *C. asper* would reveal more than 50 genetic units, given that previous studies conducted at the Pyrenean scale returned a much smaller number of nuclear partitions (Milá et al. 2010; Valbuena-Ureña et al. 2018). The program was run using the admixture model with correlated allele frequencies. The analysis was conducted for the whole dataset and for each cluster separately. The optimal number of genetic clusters was determined using both the original method of Pritchard et al. (2000) and the ΔK method of Evanno et al. (2005), as implemented in the R package pophelper (Francis 2017). The same package was used to average replicate runs of the optimal K (Jakobsson and Rosenberg 2007) and plot the final output. In addition, to visualise genetic divergence between populations, we constructed a neighbour-joining (NJ) tree using the program POPTREEW (Takezaki et al. 2014). We used Nei's genetic distance (D_A; Nei et al. 1983) and performed 1,000 bootstraps. Genetic relationships between STRUCTURE clusters for the optimal K were visualised by drawing a NJ tree based on net nucleotide distances (Pritchard et al. 2010) using the program NEIGHBOR in the PHYLIP package 3.695 (Felsenstein 2005).

Isolation by distance was calculated via a Mantel test (Mantel and Valand 1970) using the R package ade4 (Dray and Dufour 2007), to explore the relationship between genetic and geographic distances among populations. We used standardized values of F_{ST} ($F_{ST}/(1-F_{ST})$) and log-transformed values of geographic distance as dependent and independent variables, respectively (Rousset 1997). Significance was estimated with 10,000 permutations. Analyses were performed between all populations and by grouping sampling localities as indicated by STRUCTURE.

The estimation of recent dispersal was conducted using a twofold approach. An assignment test was performed in GeneClass2 (Piry et al. 2004) to assign or exclude reference populations as possible origin of individuals (Paetkau et al. 2004). The test was run only for populations with 10 or more genotyped individuals (49 populations). The same program was used to detect first generation migrants, i.e. individuals born in a population other than that where they were collected. Details on parameters used in these analyses are presented in the supplement.

The sibship assignment method implemented in Colony 2.0.6.4 (Jones and Wang 2010) was used to infer the effective size (N_e) of populations with more than 15 genotyped individuals (35 populations) under the hypothesis of random mating. Details on parameters used in the analysis are presented in the supplement.

We tested for sex-biased dispersal by calculating F_{ST}, F_{IS} and assignment values (AI_C) within each sex (Goudet et al. 2002) using the hierfstat R package (Goudet and Jombart 2015). We performed 1,000 permutations using the "two sided" alternative method (Helfer et al. 2012). F_{ST} and F_{IS} are expected to be lower and higher for the dispersing sex compared to the philopatric sex, respectively (Goudet et al. 2002). AI_C values determine the probability that an individual genotype originated from the population from which it was sampled, correcting for differences in population genetic diversity (Favre et al. 1997). The distribution of AI_C values is centred around a mean (mAI_C) of zero, with lower values expected for the dispersing sex. In contrast, the variance of AI_C (vAI_C) is expected to be higher for the dispersing sex.

To examine whether the genetic structure revealed by the Bayesian clustering analysis could be explained by orographic features such as tributary valleys (i.e. valleys whose brooks or rivers flow into greater ones) and ridges (i.e. a chain of mountains or hills that form a continuous elevated crest), we conducted analyses of molecular variance (AMOVA) using a nested design (Excoffier et al. 1992). We implemented a four-level hierarchical approach and ran two separate AMOVA analyses: in the first analysis we estimated variance components among genetic clusters identified by STRUCTURE and among tributary valleys nested within clusters; next, in the second analysis we tested for evidence of structuring among valleys and among populations within valleys, without taking genetic clusters into account (see Werth et al. 2007

for a similar approach). Further details and parameters used in the analysis are described in the supplement.

We investigated how habitat type (lakes, streams and caves) and geographic variables (latitude, longitude and altitude) explained genetic diversity estimates using multiple linear regression models. Model selection was performed in R using backward stepwise selection, where variables were dropped iteratively from the full model minimizing AIC values. Violation of the assumptions of normality, homogeneity in variance, multicollinearity and autocorrelation were checked by examining the residuals. Analyses were performed between all populations and by grouping sampling localities as indicated by STRUCTURE.

To investigate *C. asper* evolutionary history and estimate divergence times among STRUCTURE-defined genetic lineages, we employed an approximate Bayesian computation (ABC) approach, as implemented in the software DIYABC 2.1.0 (Cornuet et al. 2014). We performed the computations both combining microsatellites and mtDNA data, and separately for microsatellites to assess the impact of using different types of markers on scenario choice and posterior parameter estimation. To reduce computational demands, we selected 50 individuals from each of the five genetic groups defined by STRUCTURE. Pilot runs confirmed that varying the sample size for microsatellites (from 30 individuals per cluster to all 1,299 individuals) did not substantially affect the final outcome in terms of best supported scenario and estimated parameters (Table 2.S2). Within each group, we selected populations representative of all habitat types, choosing among individuals with STRUCTURE ancestry coefficient $Q \geq 0.9$ to exclude potentially confounding effects of contemporary gene flow (see Ortego et al. 2015). Following the recommendations of Cabrera and Palsbøll (2017) to improve DIYABC ability to reveal the true demographic model, we focused on simple contrasting models and reduced the number of candidate scenarios to three (Figures 2.2 and 2.S1). The first type of scenario is a null model with all five lineages diverging at the same time from a common ancestor (general scenario 1). The second type is a model of initial divergence between two eastern and western ancestral lineages, keeping the eastern as ancestral, and subsequent formation of the five current genetic lineages, as suggested by Valbuena-Ureña et al. (2018) (general scenario 2). Finally, the third type is a hierarchical split model directly following results from STRUCTURE analysis, where clusters 1 and 5 were generated from cluster 3, after

an initial split between clusters 2, 3 and 4 (general scenario 3). Further details on model specifications and run parameters are outlined in the supplement (Table 2.S3).

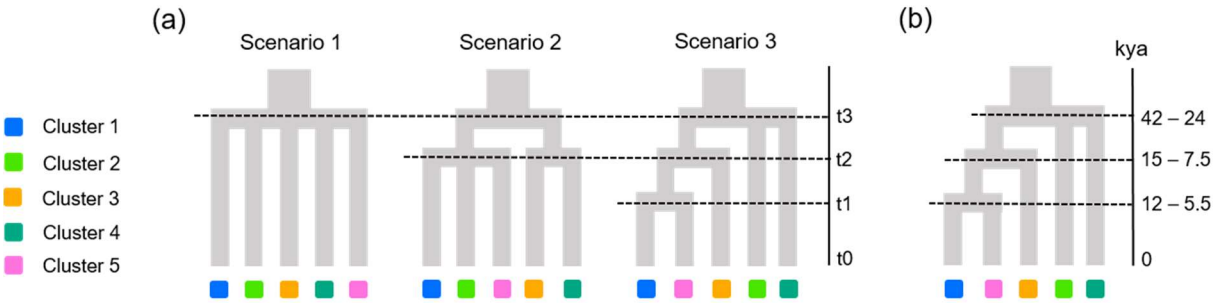


Figure 2.2 Phylogeographic scenarios tested in DIYABC during phase 2 (a). The most likely scenario, namely number 3, with the estimated time points (t1 – t3) of each split is shown in panel (b). More information on tested scenarios, estimated parameters and respective priors is given in Table 2.4 and 2.S2 and in Figure 2.S1.

Results

Multilocus genetic diversity

From the 258 individuals analysed for the cyt-*b* gene, we identified a total of 11 haplotypes. The haplotype network showed that adjacent haplotypes were separated by a single mutational step and confirmed the presence of two main central haplotypes separated from each other by two mutational steps (haplotype codes H5 and H9; Figure 2.S2). The overall mean haplotype (Hd) and nucleotide (Π) diversities were 0.570 ± 0.031 and 0.003 ± 0.0002, respectively.

Regarding microsatellites, we did not find evidence of stuttering or large allele dropout. Mean null allele frequency across all loci was 0.037, ranging from 0.018 to 0.069. Global F_{ST} values with and without correcting for null alleles were 0.377 and 0.383, respectively, and had overlapping 95% CI (0.342–0.433 for F_{ST} using ENA and 0.350–0.443 for F_{ST} not using ENA), indicating that the impact of null alleles is negligible. After applying the Bonferroni correction ($P < 0.0004$), significant linkage disequilibrium was found only in two populations between a total of three pairs of loci (in population NDE between locus pairs Ca1-Us3 and Us7-Ca16, and in population RVT-A between locus pair Ca22-Ca29). Significant deviations from HWE were observed in 18 (19%) localities after Bonferroni correction. 13 loci indicated significant departures from HWE in one to eight populations: Ca32, Ca25, Ca 23, Us7, Ca24 and Ca22 in one population, Us3, Ca8 and Ca1 in two populations, Ca30 in three populations and Ca16 in eight populations. However, this is probably the result of genetic structure in the populations,

as most of the loci showed occasional departures from HWE in three or more populations that were not consistent across populations or loci.

We recorded variable levels of nuclear genetic diversity across the study area (Table 2.S1). Mean values were 0.445 for H_O (0.162–0.698), 0.457 for H_E (0.171–0.626) and 2.659 for Ar (1.380–3.050). Westernmost populations exhibited the highest values, together with a group of central-eastern populations (Figure 2.3). F_{IS} values were generally low (mean F_{IS} = 0.069), ranging from -0.210 to 0.367.

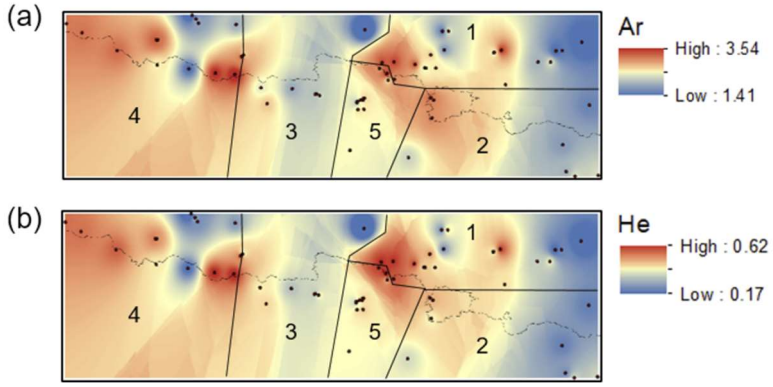


Figure 2.3 Spatial interpolation of allelic richness (Ar; a) and expected heterozygosity (H_E; b) among populations of *Calotriton asper*. Black dots denote sampling localities and black lines delimit the five genetic clusters inferred by STRUCTURE. Each cluster is identified with its corresponding number. Only populations with five or more genotyped individuals were considered in the analysis. Population codes are given in Figure 2.1.

Population structure analyses

STRUCTURE analysis revealed five well-supported groups (Figure 2.1). Log-likelihood values showed a steady increase from K = 2 to K = 5 before slowing down and eventually reaching a plateau (Figure 2.S3). Although ΔK values showed several peaks at different values of K, the peak at K = 5 was markedly higher and corresponded to the smallest variance. This chaotic behaviour has been reported when analysing data displaying strong isolation by distance with STRUCTURE (Ferchaud et al. 2015). Therefore, we assumed K = 5 as the clustering solution that best explained the spatial genetic structure of the species at the Pyrenean scale.

The five clusters were spatially distributed over the Pyrenean chain along a longitudinal gradient: the first cluster included the north-eastern (French) localities and four central-southern (Spanish) localities; the second cluster grouped together all Andorran localities, the south-

eastern (Spanish) sites and the north-eastern population Valmanya (B10); the third cluster included the central-western localities from both sides of the Pyrenees; the fourth cluster comprised all localities at both sides of the western Pyrenees; finally, sites located on the southern (Spanish) side of the central Pyrenees in-between the first three clusters formed a fifth group (see colour codes in Figure 2.1: cluster 1, blue; cluster 2, light green; cluster 3, orange; cluster 4, dark green; cluster 5, pink). Relatively high levels of admixture were detected where the genetic clusters met (Figure 2.1). When analysing each cluster separately, further substructure emerged from clusters 1 and 2 (i.e. the easternmost clusters; Figure 2.S4): sampling localities in cluster 1 grouped into three subclusters and those included in cluster 2 grouped into four subclusters. The NJ tree for the five clusters indicated that clusters 2 and 4, corresponding to the clusters at the eastern and western edges of the species range, respectively, were the most genetically differentiated (Figure 2.1). In addition, cluster 4 was the richest in terms of genetic diversity (Table 2.1). The NJ tree inferred from D_A distances over all populations revealed the five groups identified by STRUCTURE, with geographically close populations usually grouped together (Figure 2.4).

Table 2.1 Genetic diversity parameters for each genetic cluster identified by STRUCTURE analysis in *Calotriton asper*. Abbreviations: N, sample size; Ar, allelic richness standardized for sample size; PAAr, rarefied private allelic richness standardized for sample size; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, inbreeding coefficient.

Cluster	N	Ar	PAAr	H_O	H_E	F_{IS}
1	470	8.120	0.530	0.484	0.647	0.253
2	129	7.540	0.320	0.389	0.552	0.298
3	160	7.680	0.220	0.369	0.626	0.414
4	259	10.240	1.440	0.460	0.734	0.375
5	281	7.690	0.410	0.422	0.633	0.335

A significant isolation by distance (IBD) was found between all pairs of populations ($R = 0.499$, $P < 0.001$; Figure 2.S5). Similar but generally stronger IBD patterns were revealed when analysing each cluster separately (cluster 1: $R = 0.469$, $P < 0.001$; cluster 2: $R = 0.702$, $P < 0.001$; cluster 3: $R = 0.687$, $P < 0.001$; cluster 5: $R = 0.764$, $P < 0.001$; Figure 2.S5), with the exception of cluster 4 that did not show a significant IBD signal ($P = 0.053$).

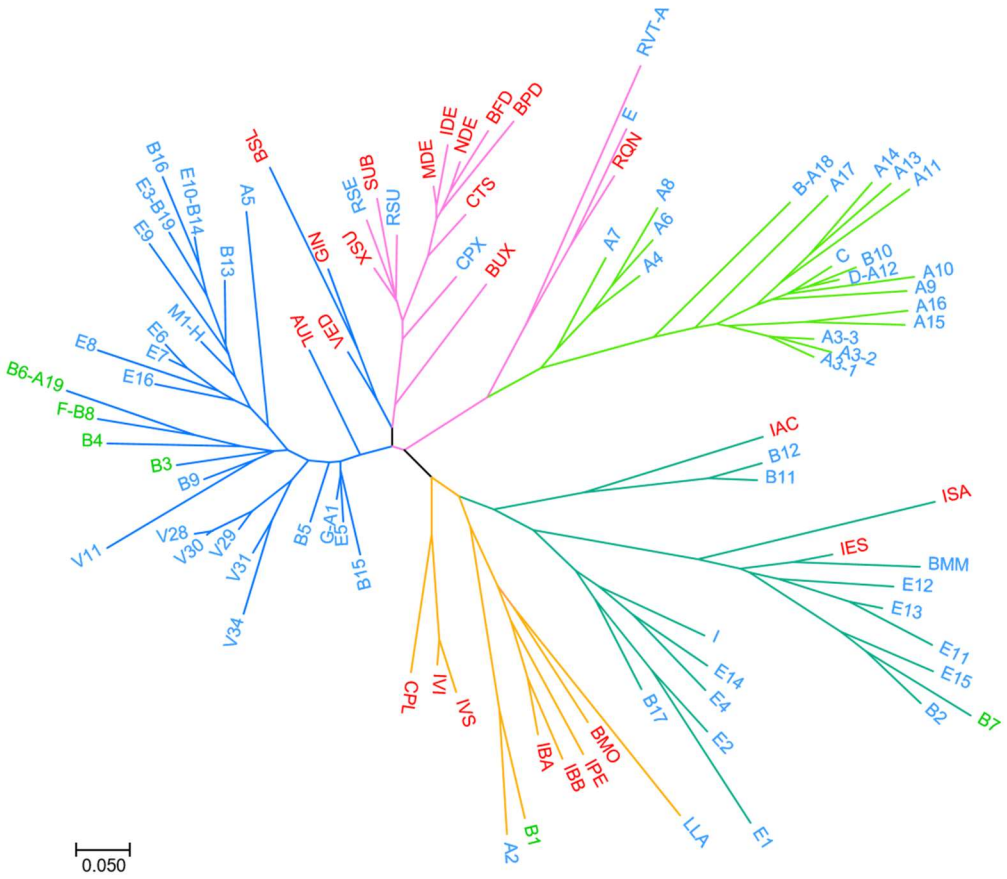


Figure 2.4 Neighbour-joining tree over all *Calotriton asper* populations based on D_A distances. Branch colours delineate the five genetic clusters identified by STRUCTURE analysis (blue: cluster 1, light green: cluster 2, orange: cluster 3, dark green: cluster 4, pink: cluster 5), while population code colours correspond to the distinct habitat types (blue: streams, red: lakes, green: caves). See Table 2.S1 for population codes.

Contemporary dispersal, effective population size and sex-biased dispersal

The assignment test conducted in GeneClass2 returned an assignment rate of 82.7%, meaning that 922 individuals out of 1,115 were assigned to the localities where they were sampled (Table 2.S4). Although the majority of misassignments were to localities belonging to the same cluster, three populations from cluster 5 and one population from cluster 3 showed ancestry to cluster 1. A total of 63 (4.9%) individuals were identified as first generation migrants: 14 and 28 individuals were selected using the L_{home} and L_{home}/L_{max} approaches, respectively, and 21 were selected by both likelihood methods. Of the 63 individuals, 27 had similar migration probabilities for several localities, indicating that these samples represented individuals whose source locality could not be determined due to the presence of unsampled populations in the study area. Among the 36 migration events with estimated origin, 19 involved stream populations only, 9 involved lake populations, 7 occurred between lake and stream populations and one between cave and stream populations. In all but one instance (one individual sampled

in population E2 and detected to be coming from E1, which are separated by only 1.7 km), migration was limited within groups detected with STRUCTURE and usually involved geographically close populations (Figures 2.5 and 2.S6). Indeed, most individuals migrated less than 1 km (17 individuals), or between 1 and 10 km (12 individuals). However, for four individuals we found potential for recent migration between localities separated by a Euclidean distance between 24 and 33 km. Dispersal between these localities would have implied either downstream migration or migration between adjacent glacial cirques, but no data are available from some intermediate localities. The remaining putative long dispersal events were below 12 km Euclidean distance and were all amongst adjacent glacial cirques.

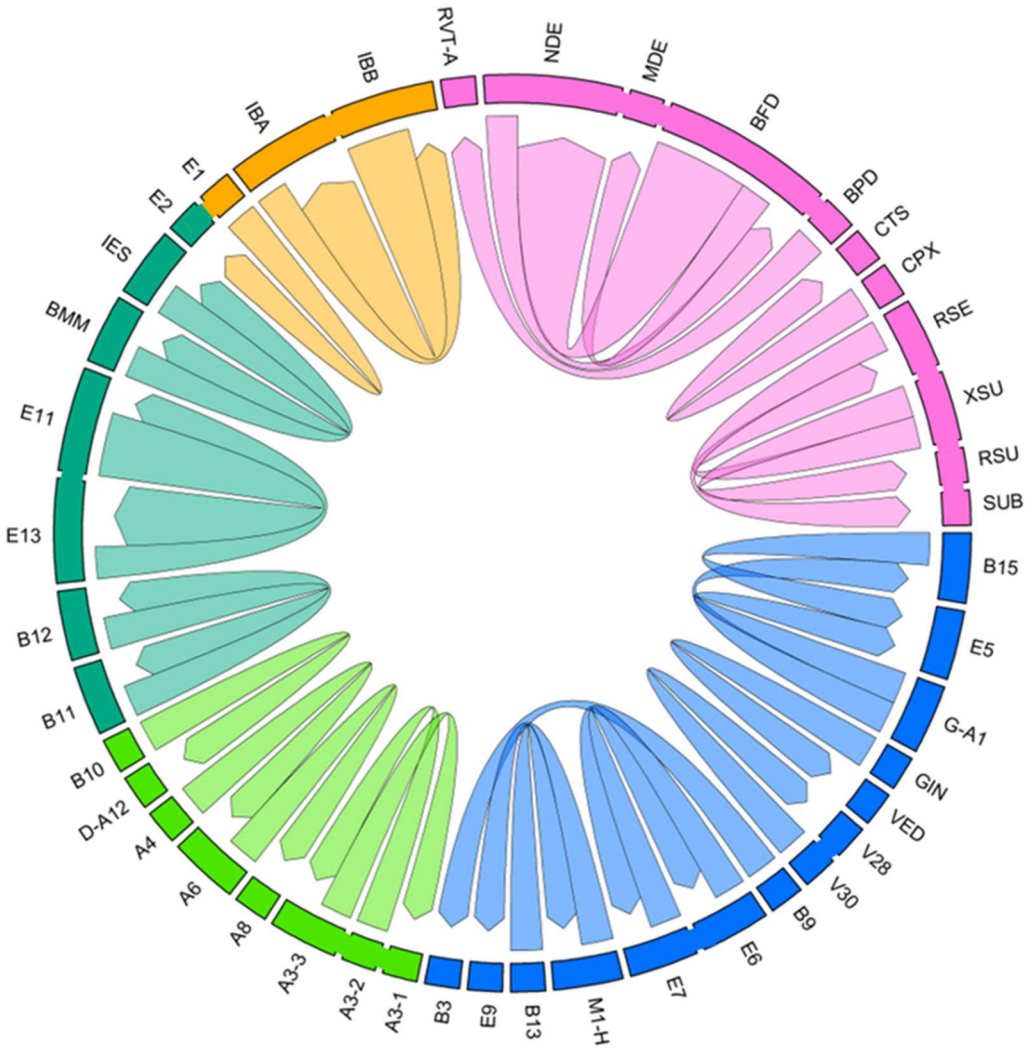


Figure 2.5 Chord diagram tracking first generation migrants flows between *Calotriton asper* sampled populations as inferred by GeneClass2. Chord size is proportional to the number of migrants detected and arrows indicate the direction of migration. Colours delineate the five genetic clusters identified by STRUCTURE analysis (blue: cluster 1, light green: cluster 2, orange: cluster 3, dark green: cluster 4, pink: cluster 5). In the outer ring, populations belonging to the same glacial cirque or valley are connected together. Only populations where first generation migrants with known source locality were detected are shown. For population codes see Table 2.S1.

Colony returned low values of effective population sizes (Table 2.S1). Values ranged from nine in the cave population Pas du Loup (B1) to 46 breeding individuals in the stream population Ruisseau de Peyrenère (E4), with a mean N_e of 26.

Results from sex-biased dispersal analysis showed that F_{ST} and F_{IS} values were not significantly different between sexes (males: $F_{ST} = 0.377$, $F_{IS} = 0.088$; females: $F_{ST} = 0.367$, $F_{IS} = 0.101$; $P_{Fst} = 0.610$, $P_{Fis} = 0.300$). Similarly, there was no significant difference in either the mean or the variance of AIc between sexes (males: mAIc = 0.052, vAIc = 16.332; females: mAIc = -0.051, vAIc = 17.771; $P_{mAIc} = 0.711$, $P_{vAIc} = 0.375$).

Influence of orography, geography and habitat

AMOVA analyses suggested significant structure at all tested levels (Table 2.2). When partitioning molecular variance between genetic clusters and tributary valleys, most molecular variance was found within valleys, followed by the among clusters component. Results did not differ substantially whether including in the analysis either all valleys or only those featuring a unique genetic cluster (data not shown). Within valleys, most variation was found among individuals, as expected for polymorphic loci such as microsatellites.

Table 2.2 Analysis of molecular variance (AMOVA) for *Calotriton asper* at the Pyrenean scale. Two hierarchical structures were tested: (1) among clusters identified by STRUCTURE analysis and among tributary valleys within clusters, (2) among tributary valleys and among populations within valleys. Abbreviations: d.f., degrees of freedom; SS, sum of squares; F_{CT}, fixation index among groups; F_{SC}, fixation index among populations within groups; F_{ST}, fixation index within populations; F_{IS}, fixation index among individuals within populations; F_{IT}, fixation index within individuals. *** $P < 0.001$

Source of variation	d.f.	SS	Variance component	% Variation	Fixation indices
(1)					
Among clusters	4	2723.166	1.368	19.195	$F_{CT} = 0.192$***
Among valleys within clusters	17	998.617	0.961	13.490	$F_{SC} = 0.167$***
Within valleys	1928	9040.858	4.797	67.315	$F_{ST} = 0.327$***
(2)					
Among valleys	20	3875.541	1.266	20.480	$F_{CT} = 0.205$***
Among populations within valleys	48	2372.971	1.295	20.960	$F_{SC} = 0.264$***
Among individuals within populations	1169	4479.044	0.212	3.430	$F_{IS} = 0.059$***
Within individuals	1238	4218.500	3.408	55.130	$F_{IT} = 0.449$***

At the Pyrenean scale, model selection indicated that altitude had a significant positive effect on H_E and Ar, whereas longitude had a significant negative effect on Ar (Figures 2.6 and 2.S7). Regarding habitat types, streams showed significantly higher levels of genetic diversity compared to lakes and caves, although this pattern was lost when performing the analysis at the genetic cluster level. Indeed, only clusters 1, 2 and 4 showed significant effects. In cluster 1,

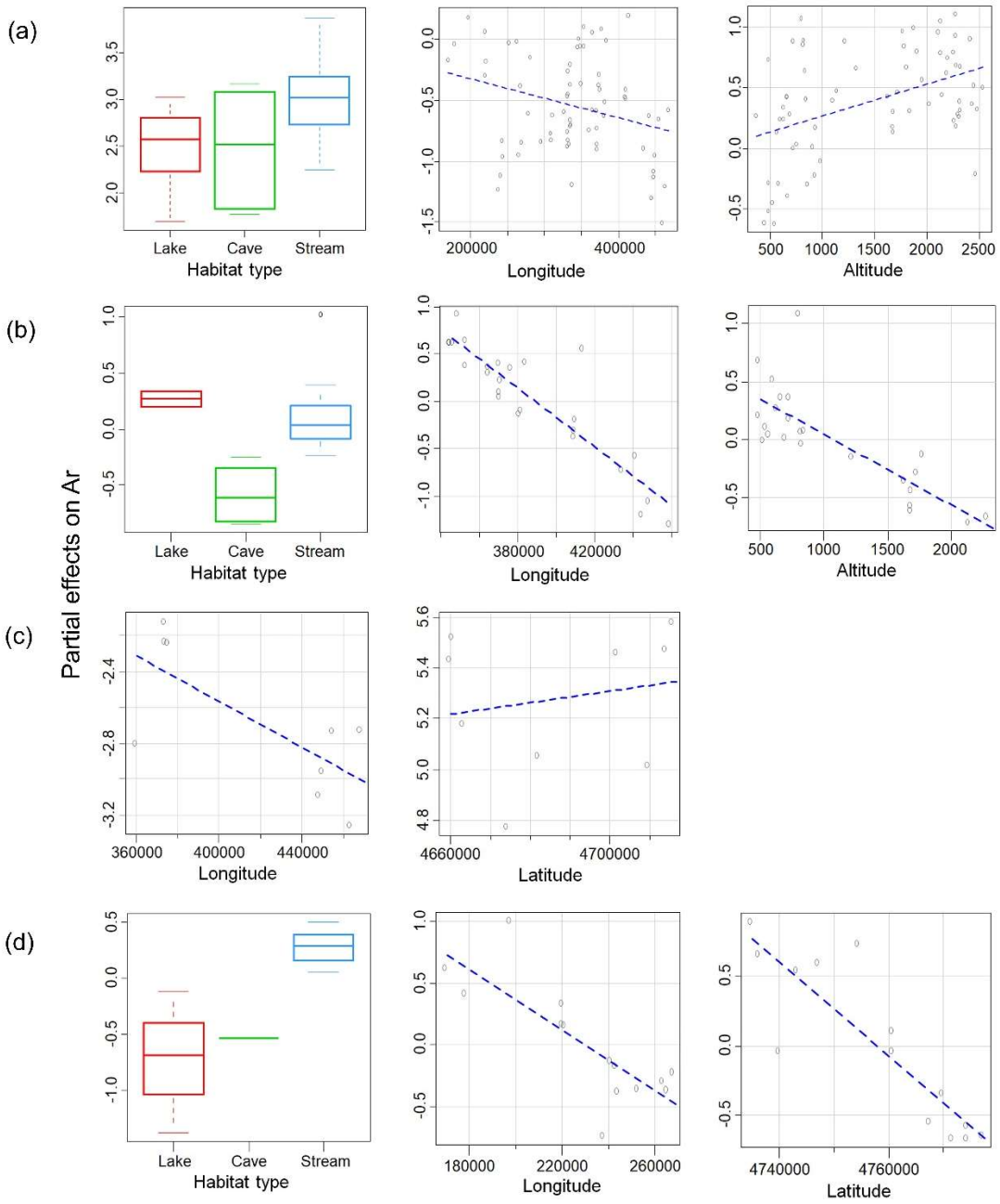


Figure 2.6 Partial effects of environmental (habitat type) and geographic (latitude, longitude and altitude) variables on allelic richness (Ar). (a), all populations; (b), cluster 1; (c), cluster 2; (d), cluster 4. Only variables that had a significant effect on Ar as determined by linear model selection are drawn. Latitude and longitude are in UTM coordinates and altitude is expressed in meters.

altitude was negatively associated with Ar and longitude was negatively associated with both H_E and Ar, and lakes were the most diverse habitat. In cluster 2, longitude had a negative effect and latitude a positive effect on both estimates; comparison between habitats was not possible because only streams were sampled. Finally, in cluster 4, longitude and latitude were negatively associated with both estimates and streams were the most diverse habitat.

Colonization history

The pre-evaluation step confirmed that the chosen priors ensured a good fit between simulated and observed data sets for all tested scenarios (Figure 2.S8). Analyses suggested highest support for scenario 3 (the multiple-refugia population model directly following Bayesian clustering analysis results) regardless of the genetic markers used (microsatellites or microsatellites + mtDNA; Figure 2.2). This scenario had the highest posterior probability (PP) and its 95% CI did not overlap with those for the other scenarios (Table 2.3). Type I and type II errors for scenario 3 were low, denoting high confidence in scenario choice (Table 2.3). RMedAD values were relatively small (< 0.25 in most cases), indicating precise parameter estimations (Table 2.4). Finally, model checking revealed that the observed dataset fell within the cloud of points of the simulated datasets obtained from the parameter posterior distribution (Figure 2.S8).

Analyses based on either microsatellites or microsatellites + mtDNA returned similar parameter estimates (Table 2.4). Results suggested that peripheral genetic lineages (clusters 2 and 4), together with the central group, diverged from a common ancestor around the Last Glacial Maximum (LGM), approximately 42,000–24,000 years ago (t3). Subsequently, the central-western lineage (cluster 3) split from the central clade ~15,000–7,500 years ago (t2), whereas the most recent divergence occurred ~12,000–5,400 years ago (t1) between the central-southern and central-eastern lineages (clusters 5 and 1; Figure 2.2).

Table 2.3 Posterior probability of tested scenarios and 95% confidence intervals (CI) estimated with DIYABC analysis when considering only microsatellites and when including both mtDNA (cyt-*b*) and microsatellite markers. Type I and II errors for the best supported scenario are indicated. See Figure 2.2 for more information on the tested scenarios.

	Microsatellites				Microsatellites + cyt-*b*			
Scenario	Posterior probability	95% CI	Type I error	Type II error	Posterior probability	95% CI	Type I error	Type II error
1	0.002	0.002-0.003			0.007	0.006-0.008		
2	0.002	0.002-0.003			0.010	0.009-0.012		
3	0.996	0.995-0.996	0.034	0.041	0.983	0.980-0.985	0.039	0.029

Table 2.4 Posterior parameters (median and 95% confidence intervals) and RMedAD (Relative Median Absolute Deviation) estimated with DIYABC analysis for the best supported scenario (scenario 3) when considering only microsatellites (simple sequence repeats – SSRs) and when including both mtDNA (cyt-*b*) and microsatellite markers. See Figures 2.2 and 2.S1 for more information on the tested scenarios. Abbreviations: N, effective population size for each analysed deme (1 – cluster 1; 2 – cluster 2; 3 – cluster 3; 4 – cluster 4, 5 – cluster 5; 135 – central clusters; 241 – three oldest glacial refugia: eastern, western and central); t, time of events in generations (t_1 – time to the most recent split; t_2 – time to the intermediate split; t_3 – time to the most ancient split); mean μ, mean mutation rate; mean *P*, mean coefficient *P*; mean *k1*, mean coefficient *k1*; $Q_{2.5}$, quantile 2.5%; $Q_{97.5}$, quantile 97.5%.

	Microsatellites				Microsatellites + cyt-*b*			
Parameter	Median	$Q_{2.5}$	$Q_{97.5}$	RMedAD	Median	$Q_{2.5}$	$Q_{97.5}$	RMedAD
N_1	3 460	1 310	9 940	0.197	3 000	962	9 930	0.225
N_2	4 600	2 470	7 950	0.186	3 790	1 550	8 690	0.195
N_3	6 380	2 790	12 600	0.175	5 940	2 160	12 800	0.195
N_4	10 300	5 680	14 200	0.135	9 620	4 730	14 200	0.154
N_5	6 930	3 110	12 900	0.183	7 100	2 400	13 700	0.207
N_{135}	7 590	997	14 400	0.324	6 490	667	14 200	0.314
N_{241}	14 400	2 680	19 700	0.269	13 200	1 980	19 500	0.307
t_1	4 050	1 470	7 770	0.245	2 700	636	6 470	0.288
t_2	5 020	1 510	9 560	0.209	3 680	860	9 200	0.255
t_3	14 200	6 730	19 600	0.169	12 000	4 620	19 300	0.219
Mean $\mu_{(SSRs)}$	$1.32x10^{-4}$	$1.01x10^{-4}$	$2.45x10^{-4}$	0.278	$1.59x10^{-4}$	$1.06x10^{-4}$	$3.37x10^{-4}$	0.253
Mean $P_{(SSRs)}$	0.229	0.122	0.300	0.188	0.195	0.110	0.291	0.180
Mean $\mu_{(cyt\text{-}b)}$	-	-	-	-	$1.69x10^{-7}$	$6.08x10^{-8}$	$4.00x10^{-7}$	0.276
Mean $k1_{(cyt\text{-}b)}$	-	-	-	-	7.920	0.410	18.800	0.442

Discussion

Refugia within refugia: the Pyrenees

Mountain systems played a crucial role in determining species diversity, and the origin of intraspecific genetic structuring has been frequently tracked back to putative glacial refugia where populations survived Quaternary ice ages (Wallis et al. 2016). In Europe, the Iberian Peninsula served as one of the most important Pleistocene glacial refugia (Gómez and Lunt 2007). The complex climatic and topographic features of this region allowed for lineage persistence in "refugia within refugia", the Pyrenees being one of them (Abellán and Svenning 2014; Gómez and Lunt 2007). For this reason, the Pyrenees are considered as a biodiversity hotspot with a rich endemic flora and fauna (Wallis et al. 2016). Here, ABC-based analyses revealed that *C. asper* microsatellite lineage differentiation started either during or slightly before the LGM (~42,000–24,000 years ago) at three main focal centres (western –cluster 4–, central and eastern –cluster 2– Pyrenees) and continued within the central group through the end of the Last Glacial Period, until ~12,000–5,500 years ago (Figure 2.2; Table 2.4). Indeed, the second and third splits straddled the Pleistocene-Holocene boundary and involved the central group only, with a first divergence event consisting of the separation of the central-

western lineage (cluster 3) ~15,000–7,500 years ago, followed by a split between the central Spanish and the central-eastern French lineages (clusters 5 and 1; ~12,000–5,500 years ago).

Our study describes the existence of five main genetic lineages in *C. asper*, which are distributed longitudinally along the Pyrenees. Previous studies mainly reported two or three major longitudinal splits in the Pyrenees in a number of species, such as the mountain ringlet butterfly *Erebia epiphron* (Schmitt et al. 2006), the European beech *Fagus sylvatica* (Magri et al. 2006), the snapdragon *Antirrhinum* (Liberal et al. 2014), the rusty-leaved alpenrose *Rhododendron ferrugineum* (Charrier et al. 2014) and the ground-dwelling spider *Harpactocrates ravastellus* (Bidegaray-Batista et al. 2016). However, most of these studies either dealt with species complexes and therefore evolutionary time lags of millions of years (Bidegaray-Batista et al. 2016; Liberal et al. 2014), or did not attempt to date back the phylogeographic history of the study species across the Pyrenees (Charrier et al. 2014; Magri et al. 2006; Schmitt et al. 2006). Other studies have only focussed on the post-glacial colonisation history (e.g. from 15,000 years ago to the present), such as in the case of the Pyrenean rock lizard *Iberolacerta bonnali* (Ferchaud et al. 2015) or the water flea *Daphnia longispina* (Ventura et al. 2014).

Phylogeography of C. asper

The times of the splits approximately correspond to major cooling events in the Pyrenees. The LGM in the Pyrenees is estimated to have occurred ~22,500–18,000 years ago (González-Sampériz et al. 2006); glacial advance likely promoted species retreat to isolated refugial areas (three main refugial areas: western, central and eastern) and subsequent genetic differentiation. After the LGM, a period of increase in temperature (the Bølling-Allerød period, ~15,000–13,000 years ago) may have created favourable conditions for dispersion outside the refugia and colonization of suitable areas in the Pyrenees. This was followed by a cold period (the Younger Dryas, ~13,000–11,500 years ago; González-Sampériz et al. 2006) that likely prompted species retreat to refugial areas where further genetic differentiation was favoured (divergence of cluster 3 from the central group). The Younger Dryas marked the end of the Pleistocene and, with the beginning of the Holocene, temperatures increased again, favouring species expansion uphill and towards the central Pyrenees. An additional abrupt cooling episode took place ~8,400–8,000 years ago (8,200–yr event; Alley et al. 1997; González-Sampériz et al. 2006), which likely promoted the last split between clusters 1 and 5.

During the last glaciation, Pyrenean glaciers reached their maximum extent earlier than the LGM at > 30,000 years ago, though a later glaciar re-advance occurred during the LGM (García-Ruiz et al. 2003). During these periods, most of the Pyrenees was extensively covered with ice and likely represented an unsuitable region for *C. asper*. Although we cannot rule out that some *C. asper* populations survived glacial events in microrefugia *in situ* (e.g. in deep valleys or on southern valley slopes), optimal conditions during glacial maxima existed mostly in peripheral areas outside the mountain range, unlike other species that likely survived in nunataks along the chain (e.g. Charrier et al. 2014). We thus hypothesise that, at the time of the first split, populations took refuge in three major refugial areas (corresponding to the western, central and eastern genetic lineages) located outside the mountain range. The long branches defining these three lineages in the NJ tree and their geographic consistency support a scenario of allopatric divergence and long-term lineage persistence in separated refugia (Figure 2.1). After the LGM, temperatures increased and created favourable conditions for the species to recolonise suitable habitats inside the chain. The following splits were likely prompted by cooling events occurring over shorter intervals and characterized by a lesser glacier extent (i.e. the Younger Dryas and the 8,200-yr event; González-Sampériz et al. 2006), leading to a wide availability of habitats inside the Pyrenees even during cold periods. It is reasonable to assume that *C. asper* endured these cooling periods in refugia located within the Pyrenees, where differentiation of the central group was favoured.

We would like to stress that ABC modelling has some uncertainty. Firstly, the tested models do not represent a comprehensive range of all possible scenarios, but are instead based on a selection of hypotheses that we consider are most likely to reflect our data. We focused our analysis on three simple contrasting models aimed at capturing the key demographic events, avoiding overcomplex and similar models. This approach has proven useful to increase the ability of DIYABC to reveal the true model, as well as to better estimate the error and accuracy of parameter estimates (Cabrera and Palsbøll 2017). Secondly, ABC modelling is based on scenarios where no gene flow is permitted between populations after they initially diverge. Only single events of admixture between populations are considered, whereas recurrent gene flow due to dispersal cannot be incorporated. However, we believe that not incorporating gene flow had only a marginal effect on our ABC results, as ABC analyses run using all 1,299 individuals (and thus including admixed populations located at cluster borders) yielded parameter estimates

similar to those from computations based on 50 individuals per cluster (Table 2.S2). Thirdly, it is important to note that the time estimates presented for *C. asper* have relatively large confidence intervals, although they still embrace values broadly referred to the time of the last glaciation.

Mito-nuclear discordance

Population analyses of nuclear microsatellites revealed that the Pyrenean brook newt is subdivided into five well-supported genetic groups mainly distributed along a longitudinal gradient (Figure 2.1), with eastern genetic groups displaying finer substructure (Figure 2.S4). This is in agreement with previous studies investigating the nuclear genetic structure of the species (Milá et al. 2010; Valbuena-Ureña et al. 2018). However, mitochondrial DNA did not show a clear phylogeographic pattern coinciding with the five microsatellite lineages (Figure 2.S2). Haplotype H9 partly corresponds to cluster 2 (eastern Pyrenees; but see Valbuena-Ureña et al. 2013) and haplotype H7 shows some affinity to cluster 3 (central-western Pyrenees); the remaining area is dominated by haplotype H5, which is the most widespread haplotype. The almost perfect match between ABC analyses based on either microsatellites or microsatellites + mtDNA was possibly due to the lack of mtDNA variation. In *C. asper*, a similar mito-nuclear discordance was detected by Milá et al. (2010): variation at several mtDNA regions (2,040 bp) was low, whereas differentiation at AFLP loci was high and consistent with the structure here identified with microsatellites (see also Valbuena-Ureña et al. 2018). Milá et al. (2010) suggested that variation at AFLP loci could have been abnormally high because of the high amount of satellite DNA in *C. asper* genome, which possibly interfered in the amplification. However, the marked genetic structuring detected with microsatellites, which is consistent with the genetic units revealed by AFLP, indicates that AFLP loci variation was not an artefact but the product of real population structuring in the species. Divergence times estimated with microsatellites approximately correspond to major cooling events that likely impacted and shaped the genetic constitution of *C. asper*. Furthermore, the high differentiation at AFLP and microsatellite markers is consistent with the high morphological diversification reported among *C. asper* populations (Montori et al. 2008a). An alternative possibility is that the observed mtDNA variation could be due to female-biased dispersal, with female-mediated gene flow and phylopatric males leading to a pattern of mito-nuclear discordance (Prugnolle and De Meeus 2002). However, our results do not support a sex-biased dispersal scenario. A more plausible explanation for the observed discordance would be a selective sweep on mtDNA, bringing

haplotypes H5 and H9 close to fixation in most populations over most of the species range (see also Valbuena-Ureña et al. 2013). Empirical evidence of selection on mtDNA is accumulating in the literature and possible cases of selective sweep have been reported in a number of taxa (Bazin et al. 2006; Bensch et al. 2006; Ferchaud et al. 2015; Rato et al. 2010). As for *C. asper*, a selective sweep of favourable mtDNA variants was previously suggested by Milá et al. (2010) to explain the lack of mtDNA diversity. A selective sweep could account for the low variation in mtDNA compared to nuclear DNA and for the geographic distribution of haplotypes. However, further studies are needed to confirm this hypothesis.

Contemporary dispersal and influence of environmental and geographic variables

Our analyses revealed restricted contemporary gene flow and dispersal between populations of *C. asper* across the five genetic lineages (Figure 2.5; Table 2.S4). This is supported by the clear pattern of isolation by distance (Figure 2.S5) and by 19% of the observed genetic variation being explained by differences between major genetic clusters (Table 2.2). However, population structure analysis revealed admixture patterns at boundaries between genetic clusters, implying potential recent gene flow across all clusters' borders (Figure 2.1). Molecular estimates of dispersal corroborated this finding: genetic signs of contemporary dispersal, albeit weak, were detected between a number of populations located at clusters' borders. This holds especially true for cluster 5, with three populations showing ancestry to cluster 1 (Table 2.S4). According to ABC analyses, clusters 1 and 5 were the last to diverge and may have retained a higher degree of connectivity (Figure 2.2).

Moderate levels of dispersal and connectivity between habitat types were detected within genetic clusters (Figure 2.5; Table 2.S4). Nevertheless, migration preferentially involved geographically close populations (0–4 km Euclidean distance; Figure 2.S6) and it was mostly restricted within valleys. This is in agreement with Montori et al. (2008b), which mainly recorded short-range movements in *C. asper* using a capture-recapture framework. The short mean dispersal distances, coupled with low effective population sizes ($N_e < 50$), may explain the high levels of genetic structuring and differentiation for *C. asper* populations across the entire species range. On the other hand, our estimations suggested potential for rare long-distance dispersal (up to 33 km). This might include both movements along the stream network and overland dispersal (Grant et al. 2010). Some individuals could have also been carried downstream during floods (Montori et al. 2012). However, although long-distance dispersal of

few individuals per population remains possible in amphibians (Cayuela et al. 2020), a plausible alternative scenario is that potential unsampled source populations located in between the study sites may have been at the origin of migrants if they shared alleles with the putative sites of origin. This is possible given the high availability of suitable habitats for *C. asper* in the study area. Nevertheless, long-distance dispersal, possibly over a few successive generations (Saura et al. 2014), is in line with our estimates of genetic diversity, as shown by most populations presenting low inbreeding coefficients (mean F_{IS} = 0.069) and levels of genetic variability within the range of other urodeles and temperate amphibians (Chan and Zamudio 2009).

The high overall F_{ST} value, together with the clear pattern of isolation by distance (especially at the genetic cluster level), indicate that divergence between populations is spatially structured. The strong spatial structuring, even across contrasting habitats, suggests no support for isolation by environment (Orsini et al. 2013). Indeed, populations from different habitats clustered together in four of the five lineages, and neighbour-joining analysis showed that populations are mainly grouped by valleys rather than habitats (Figure 2.4). Marked genetic differentiation exists at the scale of tributary valleys, as suggested by 20.5% of the molecular variance being attributable to differences between valleys (Table 2.2). Furthermore, we detected recent dispersal (as inferred by microsatellites) among populations inhabiting different habitats. In accordance with Valbuena-Ureña et al. (2018), we found evidence for a negative longitudinal and positive altitudinal gradient of genetic diversity over all *C. asper* populations, and streams showed higher values of genetic diversity compared to lakes and caves (Figures 2.6 and 2.S7). This trend has been previously interpreted as evidence of preference for cooler and wetter environments, typical of the western sector of the Pyrenees and high altitudes, by *C. asper* (Valbuena-Ureña et al. 2018). However, linear models conducted at the genetic cluster level revealed contrasting patterns of genetic diversity that do not conform with the general trend. This, together with the strong isolation by distance revealed at the cluster level, suggests that the pattern detected at the Pyrenean scale is likely the result of independent drivers acting within clusters. Clusters may thus be considered as independent units as a result of independent phylogeographic histories, each being the product of separate post-glacial colonisation routes. In light of the above, isolation by colonisation remains a plausible explanation for the resulting pattern of isolation by distance (Orsini et al. 2013), but further studies focussing on local adaptation might be necessary to confirm this point (see also Oromi et al. 2018). An alternative possibility is that the contrasting patterns at the cluster level could have arisen through the

combined effects of latitude, longitude and habitat type. Habitat type might have an influence on the level of genetic variation in the residing populations and the contrasting patterns among clusters could be caused by the differential availability of these habitats in different areas.

Concluding remarks

This study highlights the importance of integrating past evolutionary processes and present-day gene flow and dispersal dynamics to shed light onto what shaped (and is currently shaping) the observed genetic composition and structure of endemic species. Here, we demonstrate that the endemic newt *C. asper* probably recolonized the Pyrenees from at least five distinct glacial refugia. Differentiation started before the LGM and continued through the end of the Last Glacial Period, leading to the formation of five well-supported genetic lineages that likely underwent separate evolutionary histories. There is currently limited gene flow between lineages, although borders represent zones of admixture resulting from postglacial recolonization of formerly glaciated areas. Within lineages, dispersal distances are relatively short, although long-distance dispersal may be accomplished by a few individuals. The incongruence between the high variation in nuclear DNA and low variation in mtDNA could be interpreted as evidence of selective sweep in mtDNA and underscores the importance of using a multilocus approach to achieve a complete picture of the population structure and history of the study species. Given the age of the studied lineages and the restricted present-day gene flow, we suggest that these broad areas should be regarded as separate management units worthy of independent conservation consideration. At smaller spatial scales, specific lake populations of *C. asper* have been also found to merit special conservation focus (i.e. the paedomorphic populations described in Oromi et al. 2018).

Acknowledgements

We thank Meritxell Cases, Alba Castrillón, Eloi Cruset, Blanca Font, Ismael Jurado and Quim Pou-Rovira for field assistance. Economic support was provided by the European Commission LIFE+ project LimnoPirineus (LIFE13 NAT/ES/001210), by the Spanish Government project Funbio (RTI2018-096217-B-I00), by the Interreg POCTEFA ECTOPYR project (EFA031/15) and by the Societas Europaea Herpetologica (SEH, research grant awarded to F.L.). F.L. had a doctoral grant funded by Fundação para a Ciência e a Tecnologia (FCT**,** grant number PD/BD/52598/2014). M.D. is research Director at Fonds de la Recherche Scientifique – FNRS.

References

Abellán P, Svenning J-C (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.

Allentoft ME, Siegismund HR, Briggs L, Andersen LW (2009) Microsatellite analysis of the natterjack toad (*Bufo calamita*) in Denmark: Populations are islands in a fragmented landscape. *Conservation Genetics* **10:** 15-28.

Alley RB, Mayewski PA, Sowers T, Stuiver M, Taylor KC, Clark PU (1997) Holocene climatic instability: A prominent, widespread event 8200 yr ago. *Geology* **25:** 483-486.

Avise JC (2000) *Phylogeography: The history and formation of species.* Harvard University Press: Cambridge, MA, USA.

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

Bensch S, Irwin DE, Irwin JH, Kvist L, Åkesson S (2006) Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Molecular Ecology* **15:** 161-171.

Bidegaray-Batista L, Sánchez-Gracia A, Santulli G, Maiorano L, Guisan A, Vogler AP et al. (2016) Imprints of multiple glacial refugia in the Pyrenees revealed by phylogeography and palaeodistribution modelling of an endemic spider. *Molecular Ecology* **25:** 2046-2064.

Bosch J, Tejedo M, Lecis R, Miaud C, Lizana M, Edgar P et al. (2009) *Calotriton asper*. The IUCN Red List of Threatened Species 2009: e.T59448A11943040. doi:10.2305/IUCN.UK.2009.RLTS.T59448A11943040.en.

Burns EL, Eldridge MD, Houlden BA (2004) Microsatellite variation and population structure in a declining Australian Hylid *Litoria aurea*. *Molecular Ecology* **13:** 1745-1757.

Cabrera AA, Palsbøll PJ (2017) Inferring past demographic changes from contemporary genetic data: A simulation-based evaluation of the ABC methods implemented in DIYABC. *Molecular Ecology Resources* **17:** e94-e110.

Calvet M (2004) The Quaternary glaciation of the Pyrenees. *Developments in Quaternary Sciences* **2:** 119-128.

Caplat P, Edelaar P, Dudaniec RY, Green AJ, Okamura B, Cote J et al. (2016) Looking beyond the mountain: Dispersal barriers in a changing world. *Frontiers in Ecology and the Environment* **14:** 262-269.

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

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

Cayuela H, Valenzuela-Sanchez A, Teulier L, Martínez-Solano Í, Léna J-P, Merilä J et al. (2020) Determinants and consequences of dispersal in vertebrates with complex life

cycles: A review of pond-breeding amphibians. *The Quarterly Review of Biology* **95:** 1-36.

Chan LM, Zamudio KR (2009) Population differentiation of temperate amphibians in unpredictable environments. *Molecular Ecology* **18:** 3185-3200.

Chapuis M-P, Estoup A (2006) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* **24:** 621-631.

Charrier O, Dupont P, Pornon A, Escaravage N (2014) Microsatellite marker analysis reveals the complex phylogeographic history of *Rhododendron ferrugineum* (Ericaceae) in the Pyrenees. *PLoS One* **9:** e92976.

Chiucchi JE, Gibbs H (2010) Similarity of contemporary and historical gene flow among highly fragmented populations of an endangered rattlesnake. *Molecular Ecology* **19:** 5345-5358.

Clergue-Gazeau M, Martínez-Rica J (1978) Les différents biotopes de l'urodèle pyrénéen, *Euproctus asper*. *Bulletin de la Société d'Histoire Naturelle de Toulouse* **114:** 461-471.

Clobert J, Le Galliard JF, Cote J, Meylan S, Massot M (2009) Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecology Letters* **12:** 197-209.

Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R et al. (2014) DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30:** 1187-1189.

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

Denoël M, Dalleur S, Langrand E, Besnard A, Cayuela H (2018) Dispersal and alternative breeding site fidelity strategies in an amphibian. *Ecography* **41:** 1543-1555.

Dray S, Dufour A-B (2007) The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software* **22:** 1-20.

Drechsler A, Geller D, Freund K, Schmeller DS, Kuenzel S, Rupp O et al. (2013) What remains from a 454 run: estimation of success rates of microsatellite loci development in selected newt species (*Calotriton asper*, *Lissotriton helveticus*, and *Triturus cristatus*) and comparison with Illumina-based approaches. *Ecology and Evolution* **3:** 3947-3957.

Epps CW, Keyghobadi N (2015) Landscape genetics in a changing world: Disentangling historical and contemporary influences and inferring change. *Molecular Ecology* **24:** 6021-6040.

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, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131:** 479-491.

Favre L, Balloux F, Goudet J, Perrin N (1997) Female-biased dispersal in the monogamous mammal *Crocidura russula*: Evidence from field data and microsatellite patterns. *Proceedings of the Royal Society of London Series B: Biological Sciences* **264:** 127-132.

Felsenstein J (2005) PHYLIP (phylogeny inference package) version 3.6. Distributed by the author. Seattle (WA): Department of Genome Sciences, University of Washington.

Ferchaud AL, Eudeline R, Arnal V, Cheylan M, Pottier G, Leblois R et al. (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.

Francis RM (2017) pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* **17:** 27-32.

García-Ruiz JM, Valero-Garcés BL, Martí-Bono C, González-Sampériz P (2003) Asynchroneity of maximum glacier advances in the central Spanish Pyrenees. *Journal of Quaternary Science* **18:** 61-72.

Gill DE (1978) The metapopulation ecology of the red-spotted newt, *Notophthalmus viridescens* (Rafinesque). *Ecological Monographs* **48:** 145-166.

Gómez A, Lunt DH (2007) Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S and Ferrand N (eds) *Phylogeography of southern European refugia*. Springer: Amsterdam, Netherlands, pp 155-188.

González-Sampériz P, Valero-Garcés BL, Moreno A, Jalut G, García-Ruiz JM, Martí-Bono C et al. (2006) Climate variability in the Spanish Pyrenees during the last 30,000 yr revealed by the El Portalet sequence. *Quaternary Research* **66:** 38-52.

Goudet J (2002) FSTAT version 2.9. 3.2, a program to estimate and test gene diversities and fixation indices. Lausanne, Switzerland: Institute of Ecology. http://www2.unil.ch/popgen/softwares/fstat.htm.

Goudet J, Jombart T (2015) hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-22.

Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology* **11:** 1103-1114.

Grant EHC, Nichols JD, Lowe WH, Fagan WF (2010) Use of multiple dispersal pathways facilitates amphibian persistence in stream networks. *Proceedings of the National Academy of Sciences* **107:** 6936-6940.

Helfer V, Broquet T, Fumagalli L (2012) Sex-specific estimates of dispersal show female philopatry and male dispersal in a promiscuous amphibian, the alpine salamander (*Salamandra atra*). *Molecular Ecology* **21:** 4706-4720.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68:** 87-112.

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; discussion 195.

Hewitt GM, Butlin RK (1997) Causes and consequences of population structure. In: Krebs JR and Davies N (eds) *Behavioral Ecology, 4th edn.* Blackwell: Oxford, pp 350-372.

Holderegger R, Thiel-Egenter C (2009) A discussion of different types of glacial refugia used in mountain biogeography and phylogeography. *Journal of Biogeography* **36:** 476-480.

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.

Johnson ML, Gaines MS (1990) Evolution of dispersal: Theoretical models and empirical tests using birds and mammals. *Annual Review of Ecology and Systematics* **21:** 449-480.

Jones OR, Wang J (2010) COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10:** 551-555.

Kalinowski ST (2005) HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5:** 187-189.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S et al. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28:** 1647-1649.

Kraaijeveld-Smit FJ, Beebee TJ, Griffiths RA, Moore RD, Schley L (2005) Low gene flow but high genetic diversity in the threatened Mallorcan midwife toad *Alytes muletensis*. *Molecular Ecology* **14:** 3307-3315.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33:** 1870-1874.

Li XY, Kokko H (2019) Sex-biased dispersal: A review of the theory. *Biological Reviews* **94:** 721-736.

Liberal IM, Burrus M, Suchet C, Thebaud C, Vargas P (2014) The evolutionary history of *Antirrhinum* in the Pyrenees inferred from phylogeographic analyses. *BMC Evolutionary Biology* **14:** 146.

Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D et al. (2006) A new scenario for the Quaternary history of European beech populations: Palaeobotanical evidence and genetic consequences. *New Phytologist* **171:** 199-221.

Mantel N, Valand RS (1970) A technique of nonparametric multivariate analysis. *Biometrics* **26:** 547-558.

Martínez-Rica J, Clergue-Gazeau M (1977) Données nouvelles sur la répartition géographique de l'espèce *Euproctus asper* Dugès, Urodèle, Salamandridae. *Bulletin de la Société d'Histoire Naturelle de Tolouse* **113:** 318-330.

Milá B, Carranza S, Guillaume O, Clobert J (2010) Marked genetic structuring and extreme dispersal limitation in the Pyrenean brook newt *Calotriton asper* (Amphibia: Salamandridae) revealed by genome-wide AFLP but not mtDNA. *Molecular Ecology* **19:** 108-120.

Montero-Pau J, Gómez A, Muñoz J (2008) Application of an inexpensive and high-throughput genomic DNA extraction method for the molecular ecology of zooplanktonic diapausing eggs. *Limnology and Oceanography: Methods* **6:** 218-222.

Montori A, Llorente GA (2014) Tritón pirenaico–*Calotriton asper* (Dugès, 1852). In: Salvador A and Martínez-Solano I (eds) *Enciclopedia Virtual de los Vertebrados*. Museo Nacional de Ciencias Naturales: Madrid, Spain

Montori A, Llorente GA, García-París M (2008a) Allozyme differentiation among populations of the Pyrenean newt *Calotriton asper* (Amphibia: Caudata) does not mirror their morphological diversification. *Zootaxa* **1945:** 39-50.

Montori A, Llorente GA, Richter-Boix A (2008b) Habitat features affecting the small-scale distribution and longitudinal migration patterns of *Calotriton asper* in a Pre-Pyrenean population. *Amphibia-Reptilia* **29:** 371-381.

Montori A, Richter-Boix A, Franch M, Santos X, Garriga N, Llorente GA (2012) Natural fluctuations in a stream dwelling newt as a result of extreme rainfall: A 21-year survey of a *Calotriton asper* population. *Basic and Applied Herpetology* **26:** 43-56.

Mouret V, Guillaumet A, Cheylan M, Pottier G, Ferchaud AL, Crochet PA (2011) The legacy of ice ages in mountain species: Post-glacial colonization of mountain tops rather than current range fragmentation determines mitochondrial genetic diversity in an endemic Pyrenean rock lizard. *Journal of Biogeography* **38:** 1717-1731.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19:** 153-170.

Nichols RA, Beaumont MA (1996) Is it ancient or modern history that we can read in the genes? In: Hochberg ME, Clobert J and Barbault R (eds) *Aspects of the Genesis and Maintenance of Biological Diversity*. Oxford University Press: Oxford, pp 69-87.

Noguerales V, Cordero PJ, Ortego J (2016) Hierarchical genetic structure shaped by topography in a narrow-endemic montane grasshopper. *BMC Evolutionary Biology* **16:** 96.

Noguerales V, Cordero PJ, Ortego J (2017) Testing the role of ancient and contemporary landscapes on structuring genetic variation in a specialist grasshopper. *Ecology and Evolution* **7:** 3110-3122.

Oromi N, Valbuena-Ureña E, Soler-Membrives A, Amat F, Camarasa S, Carranza S et al. (2018) Genetic structure of lake and stream populations in a Pyrenean amphibian (*Calotriton asper*) reveals evolutionary significant units associated with paedomorphosis. *Journal of Zoological Systematics and Evolutionary Research* **57:** 418-430.

Orsini L, Vanoverbeke J, Swillen I, Mergeay J, De Meester L (2013) Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology* **22:** 5983-5999.

Ortego J, Noguerales V, Gugger PF, Sork VL (2015) Evolutionary and demographic history of the Californian scrub white oak species complex: An integrative approach. *Molecular Ecology* **24:** 6188-6208.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: A simulation-based exploration of accuracy and power. *Molecular Ecology* **13:** 55-65.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95:** 536-539.

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

Pritchard JK, Wen X, Falush D (2010) Documentation for STRUCTURE software, version 2.3. Department of Human Genetics University of Chicago, Department of Statistics University of Oxford.

Prugnolle F, De Meeus T (2002) Inferring sex-biased dispersal from population genetic tools: A review. *Heredity* **88:** 161-165.

R Core Team (2018) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Rato C, Carranza S, Perera A, Carretero MA, Harris DJ (2010) Conflicting patterns of nucleotide diversity between mtDNA and nDNA in the Moorish gecko, *Tarentola mauritanica*. *Molecular Phylogenetics and Evolution* **56:** 962-971.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution* **43:** 223-225.

Rioux Paquette S (2011) PopGenKit: Useful functions for (batch) file conversion and data resampling in microsatellite datasets. R package version 1.0.

Roffler GH, Talbot SL, Luikart G, Sage GK, Pilgrim KL, Adams LG et al. (2014) Lack of sex-biased dispersal promotes fine-scale genetic structure in alpine ungulates. *Conservation Genetics* **15:** 837-851.

Ronce O (2007) How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics* **38:** 231-253.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145:** 1219-1228.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8:** 103-106.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE et al. (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34:** 3299-3302.

Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature* **392:** 491-494.

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

Saura S, Bodin Ö, Fortin MJ (2014) Stepping stones are crucial for species' long-distance dispersal and range expansion through habitat networks. *Journal of Applied Ecology* **51:** 171-182.

Schmitt T (2009) Biogeographical and evolutionary importance of the European high mountain systems. *Frontiers in Zoology* **6:** 9.

Schmitt T, Hewitt GM, Muller P (2006) Disjunct distributions during glacial and interglacial periods in mountain butterflies: *Erebia epiphron* as an example. *Journal of Evolutionary Biology* **19:** 108-113.

Smith MA, Green DM (2005) Dispersal and the metapopulation paradigm in amphibian ecology and conservation: Are all amphibian populations metapopulations? *Ecography* **28:** 110-128.

Smith MA, Green DM (2006) Sex, isolation and fidelity: Unbiased long-distance dispersal in a terrestrial amphibian. *Ecography* **29:** 649-658.

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

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

Takezaki N, Nei M, Tamura K (2014) POPTREEW: web version of POPTREE for constructing population trees from allele frequency data and computing some other quantities. *Molecular Biology and Evolution* **31:** 1622-1624.

Tallmon DA, Luikart G, Waples RS (2004) The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology & Evolution* **19:** 489-496.

Trochet A, Courtois EA, Stevens VM, Baguette M, Chaine A, Schmeller DS et al. (2016) Evolution of sex-biased dispersal. *The Quarterly Review of Biology* **91:** 297-320.

Tucker JM, Allendorf FW, Truex RL, Schwartz MK (2017) Sex-biased dispersal and spatial heterogeneity affect landscape resistance to gene flow in fisher. *Ecosphere* **8:** e01839.

Valbuena-Ureña E, Amat F, Carranza S (2013) Integrative phylogeography of *Calotriton* newts (Amphibia, Salamandridae), with special remarks on the conservation of the endangered Montseny brook newt (*Calotriton arnoldi*). *PLoS One* **8:** e62542.

Valbuena-Ureña E, Oromi N, Soler-Membrives A, Carranza S, Amat F, Camarasa S et al. (2018) Jailed in the mountains: Genetic diversity and structure of an endemic newt species across the Pyrenees. *PLoS One* **13:** e0200214.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4:** 535-538.

Ventura M, Petrusek A, Miró A, Hamrová E, Buñay D, De Meester L et al. (2014) Local and regional founder effects in lake zooplankton persist after thousands of years despite high dispersal potential. *Molecular Ecology* **23:** 1014-1027.

Vos CC, Antonisse-De Jong AG, Goedhart PW, Smulders MJM (2001) Genetic similarity as a measure for connectivity between fragmented populations of the moor frog (*Rana arvalis*). *Heredity* **86:** 598-608.

Wallis GP, Waters JM, Upton P, Craw D (2016) Transverse alpine speciation driven by glaciation. *Trends in Ecology & Evolution* **31:** 916-926.

Werth S, Gugerli F, Holderegger R, Wagner HH, Csencsics D, Scheidegger C (2007) Landscape-level gene flow in *Lobaria pulmonaria*, an epiphytic lichen. *Molecular Ecology* **16:** 2807-2815.

Zellmer A, Knowles LL (2009) Disentangling the effects of historic vs. contemporary landscape structure on population genetic divergence. *Molecular Ecology* **18:** 3593-3602.

Supplementary material

Supplementary methods

Recent dispersal (*GeneClass2)*

We employed two different likelihood computation criteria for migrant detection: L_{home}, that is the likelihood of finding an individual in the population where it was sampled, and L_{home}/L_{max}, the ratio of L_{home} to the highest likelihood among all sampled populations. L_{home} is more appropriate when, as in our case, not all source populations for migrants have been sampled. However, it is less powerful than L_{home}/L_{max}. We used the Bayesian method of Rannala and Mountain (1997) coupled with the Monte-Carlo resampling method of Paetkau et al. (2004). We used 10,000 simulated individuals for assignment and L_{home} computations, while 1,000 simulated individuals were employed for L_{home}/L_{max} computations, because of RAM limitations. Type I error (alpha level) was set to 0.01. A number of individuals identified by GeneClass2 as first generation migrants had similar migration probabilities for several areas and therefore could not be unambiguously assigned to a source population. This was especially evident in areas where sampling effort was low and, as a result, some of the sampled populations were apparently isolated. The individuals in question were thus classified as first generation migrants whose source locality could not be determined (Bergl and Vigilant 2007). To visualize first generation migrants' trajectories, we generated a chord diagram using the circlize R package (Gu et al. 2014).

Effective population size (Colony)

This software uses a maximum likelihood method to conduct parentage and sibship inference to estimate N_e and can accommodate null alleles and other genotyping errors. Because we could not reliably discriminate between adult and subadult individuals (and thus discriminate between putative parental and offspring genotypes), we opted to pool all sampled individuals in the offspring sample only (J. Wang, personal communication). The software was run with the null allele frequencies computed with FreeNA as allelic dropout rate and the rate of other kinds of genotyping errors was set to 0.01 at all loci (see Mokhtar-Jamai et al. 2013 for a similar approach). Mating system was assumed to be both-sex polygamy, with no sibship prior. Three medium length runs were conducted to ensure convergence of the annealing procedure, and other parameters were used as default.

Analysis of molecular variance (AMOVA)

Because we had a few instances of valleys where two different genetic clusters coexisted, we repeated the first AMOVA twice: first retaining only valleys harbouring a unique genetic cluster (17 out of 22), and then including all valleys and assigning the "admixed" ones to the most representative cluster in terms of number of populations. Analyses were performed in Arlequin 3.5.2.2 (Excoffier and Lischer 2010) with 10,000 permutations to assess statistical significance of fixation indices. Only populations with five or more genotyped individuals were considered in the analyses.

Approximate Bayesian Computation (DIYABC)

Candidate scenarios were built on the basis of STRUCTURE and NJ analyses and previously published information on the population structure and putative past distribution of the species (Carranza and Amat 2005; Valbuena-Ureña et al. 2013; Valbuena-Ureña et al. 2018). Three types of scenarios of historical divergence were tested (Figure 2.S1). Model selection was conducted in a hierarchical manner, in a similar fashion to Barbosa et al. (2017): each general scenario was first tested independently to examine alternative hypotheses about the timeframe of divergence among clusters (Figure 2.S1); the configuration with the highest posterior probability for each general scenario was kept for the final analysis, where we compared the best scenarios against each other (Figure 2.2). We performed preliminary runs with varying priors for effective population size and divergence times (t) to adjust them to the most appropriate values. The final parameter setting is shown in Table 2.S3. We assumed a generation time of 2 to 3 years (Montori 1988; Montori and Llorente 2014), so we multiplied time estimates by two and by three to convert demographic parameters into absolute times. We generated 10^6 simulated datasets per scenario, assuming a 1:1 female to male sex ratio (Montori and Llorente 2014). We used the following summary statistics (SS) for microsatellites: mean number of alleles, mean genetic diversity and mean allele size variance as one sample SS, and mean number of alleles, Fst and $(d\mu)^2$ distance as two sample SS. As for cyt-*b*, the following SS were used: number of haplotypes, number of segregating sites and Tajima's D as one sample SS, and number of haplotypes and Fst as two sample SS. The "Pre-evaluation of scenarios and prior distributions" option was employed to ensure that our observed dataset was positioned well within the cloud of simulated datasets for all competing scenarios, through the computation of a principal component analysis (PCA) on summary statistics (Cornuet et al. 2015; Cornuet et al. 2010). Selection of the most supported scenario, confidence in scenario choice (type I and

II errors), model checking and estimation of the posterior distribution of parameters for the most supported scenario followed Ortego et al. (2015). Bias and precision on parameters estimation were evaluated by calculating the median of the Relative Median Absolute Deviation (RMedAD) of each parameter based on 5,000 pseudo-observed datasets (Cornuet et al. 2015).

Supplementary tables

Table 2.S1 Geographic information and standard genetic statistics of *Calotriton asper* sampling localities. Abbreviations: Lat., latitude; Long., longitude; Alt., altitude in meters; N, sample size for microsatellites; Ar, allelic richness standardized for sample size; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, inbreeding coefficient; N_e, effective population size; N mtDNA, sample size for mtDNA; mtDNA haps, occurrence and code (in parentheses) of mitochondrial haplotypes identified in each population.

Population	Code	Lat.	Long.	Alt.	Habitat	N	Ar	H_O	H_E	F_{IS}	N_e	N mtDNA	mtDNA haps
Irati	B12	43.05	-1.06	1100	Stream	15	3.129	0.583	0.560	-0.007	-	-	-
Olhadoko	B11	42.99	-0.95	662	Stream	14	3.152	0.597	0.558	-0.032	-	-	-
Ibón Acherito	IAC	42.88	-0.71	1872	Lake	26	3.184	0.500	0.571	0.145	27	3	3(H7)
Ruisseau de Leignières 2	E13	42.95	-0.44	834	Stream	12	3.297	0.583	0.561	0.004	-	-	-
Ibón Espelunciecha	IES	42.79	-0.43	1951	Lake	23	2.726	0.380	0.470	0.214	28	5	5(H5)
Ruisseau de Gourzy	E11	42.95	-0.43	1137	Stream	5	3.132	0.518	0.518	0.111	-	5	5(H5)
Balsa Pertacua	BPE	42.71	-0.42	1913	Lake	-	-	-	-	-	-	1	1(H5)
Ruisseau de Leignières 1	E12	42.94	-0.42	1356	Stream	3	-	-	-	-	-	3	2(H5), 1(H18)
Río Barranco Mina Millor	BMM	42.78	-0.41	1800	Stream	4	-	-	-	-	-	4	4(H5)
Ibón Serrato Superior	ISA	42.76	-0.21	2459	Lake	14	1.925	0.219	0.295	0.291	-	4	4(H5)
Betharram	B7	43.10	-0.19	446	Cave	28	1.478	0.202	0.185	-0.072	11	4	4(H5)
Ruisseau Coume Rège	E15	43.08	-0.16	581	Stream	20	2.332	0.394	0.391	0.018	22	8	8(H5)
Genie Longue	B2	43.05	-0.15	669	Stream	16	2.229	0.349	0.339	0.003	24	5	5(H5)
Ruisseau de Peyrenère	E4	42.74	-0.03	1782	Stream	29	3.446	0.557	0.617	0.115	46	-	-
Río Llanos de Larri	LLA	42.70	0.09	1595	Stream	3	-	-	-	-	-	4	2(H7), 2(H8)
Gavarnie	E14	42.73	0.10	2107	Stream	30	3.549	0.584	0.626	0.084	40	10	8(H5), 2(H7)
Sarramea et Las Carraous	E2	43.04	0.11	855	Stream	5	2.290	0.259	0.353	0.367	-	-	-
Castelmouly	E1	43.04	0.13	806	Stream	1	-	-	-	-	-	-	-
Néouvielle	I	42.84	0.15	2183	Stream	15	3.202	0.587	0.574	0.013	-	-	-
Rau d'Estibère	B17	42.85	0.16	2280	Stream	8	2.764	0.419	0.437	0.106	-	-	-
Ibón Barleto Inferior	IBB	42.65	0.28	2483	Lake	4	-	-	-	-	-	5	5(H7)
Ibón Barleto Superior	IBA	42.64	0.28	2531	Lake	16	2.571	0.393	0.467	0.194	19	5	5(H7)

Ibón Bassa de la Mora	BMO	42.54	0.33	1903	Lake	22	2.862	0.356	0.483	0.293	28	4	4(H7)
Ibón Perramó	IPE	42.64	0.50	2254	Lake	29	2.264	0.341	0.394	0.155	20	4	4(H5)
Ibón Bajo de Vallibierna	IVI	42.60	0.66	2428	Lake	25	2.379	0.386	0.422	0.106	19	4	4(H15)
Ibón Alto de Vallibierna	IVS	42.60	0.66	2474	Lake	5	2.336	0.335	0.370	0.206	-	4	2(H5), 2(H15)
Estany Cap de Llauset	CPL	42.60	0.69	2443	Lake	30	2.529	0.444	0.472	0.076	28	4	4(H5)
Arbas	A2	42.99	0.88	526	Stream	1	-	-	-	-	-	1	1(H5)
Rivert	RVT-A	42.25	0.90	932	Stream	26	2.661	0.397	0.455	0.147	23	3	2(H5), 1(H12)
Estanyet Dellui 17	IDE	42.55	0.94	2306	Lake	20	2.238	0.293	0.362	0.218	19	4	3(H5), 1(H14)
Estanyet Dellui 15	MDE	42.55	0.94	2314	Lake	39	2.651	0.355	0.448	0.221	24	5	5(H5)
Bassa Dellui	BPD	42.55	0.94	2307	Lake	2	-	-	-	-	-	2	2(H5)
Estanyet Dellui 13	NDE	42.55	0.94	2313	Lake	36	2.364	0.381	0.419	0.105	17	5	5(H5)
Estanyet Dellui 12	BFD	42.55	0.95	2314	Lake	8	2.292	0.378	0.383	0.084	-	5	5(H5)
Estany Corticelles	CTS	42.56	0.95	2278	Lake	14	2.661	0.420	0.470	0.148	-	3	3(H5)
Estany Gran Dellui	GDE	42.55	0.95	2349	Lake	-	-	-	-	-	-	1	1(H5)
Estany Redó	RQN	42.51	0.95	2411	Lake	14	2.882	0.460	0.501	0.124	-	-	-
Riu Coma de Peixerani	CPX	42.57	0.97	2297	Stream	17	2.758	0.471	0.506	0.102	32	3	3(H5)
Estany Xic de Subenuix	XSU	42.57	0.99	2272	Lake	36	2.901	0.461	0.533	0.150	33	5	4(H5), 1(H14)
Estany Subenuix	SUB	42.57	0.99	2194	Lake	5	2.718	0.459	0.413	0.007	-	3	3(H5)
Riu de Subenuix Esquerre	RSE	42.58	0.99	2142	Stream	5	2.912	0.424	0.481	0.231	-	3	3(H5)
Vall Fosca	E	42.51	0.99	1830	Stream	19	2.782	0.480	0.505	0.078	25	-	-
Riu de Subenuix	RSU	42.58	0.99	2021	Stream	21	2.839	0.504	0.513	0.046	23	3	3(H5)
Pas du Loup	B1	43.01	1.00	487	Cave	16	1.414	0.162	0.171	0.087	9	6	6(H5)
Barranc del Cap del Port de Peguera	CPP	42.54	1.03	2510	Stream	2	-	-	-	-	-	2	1(H9), 1(H13)
Estany Port d'Aulà	AUL	42.77	1.10	2128	Lake	19	3.004	0.484	0.592	0.217	21	5	5(H5)
Arcouzan	B15	42.80	1.12	1214	Stream	9	3.334	0.698	0.614	-0.077	-	-	-
Salau	G-A1	42.73	1.15	1764	Stream	39	3.415	0.586	0.621	0.068	40	-	-
Estany Buixasse Nord	BUX	42.69	1.16	2250	Lake	15	2.739	0.545	0.503	-0.047	-	3	1(H5), 2(H9)
Estany Calberante	CLB	42.69	1.17	2418	Lake	1	-	-	-	-	-	1	1(H9)
Estany Inferior de la Gallina	GIN	42.71	1.19	2268	Lake	4	-	-	-	-	-	4	4(H9)

Bassa Llavera	BSL	42.71	1.19	2193	Lake	1	-	-	-	-	-	2	2(H9)
Hoque de Fustès	E5	42.81	1.20	820	Stream	30	3.333	0.621	0.606	-0.008	35	-	-
Estanyet de Vedos	VED	42.70	1.20	2270	Lake	12	3.047	0.549	0.585	0.105	-	4	2(H5), 2(H9)
Barranc Romadriu	PAS	42.45	1.27	1684	Stream	1	-	-	-	-	-	1	1(H9)
Organyà	B-A18	42.20	1.30	758	Stream	26	2.467	0.391	0.420	0.088	26	16	15(H17), 1(H19)
Ribaui	B5	42.79	1.34	814	Stream	15	3.282	0.611	0.589	-0.003	-	5	5(H5)
Vicdessos ruisseau 3	V29	42.77	1.41	1718	Stream	14	2.879	0.514	0.521	0.051	-	-	-
Vicdessos ruisseau 4	V30	42.77	1.41	1671	Stream	16	2.594	0.498	0.477	-0.012	24	-	-
Vicdessos ruisseau 2	V28	42.77	1.41	1674	Stream	15	2.551	0.496	0.477	-0.005	-	-	-
Vicdessos ruisseau 7	V35	42.77	1.42	1667	Stream	1	-	-	-	-	-	-	-
Vicdessos ruisseau 6	V34	42.77	1.42	1668	Stream	4	-	-	-	-	-	-	-
Vicdessos ruisseau 5	V31	42.77	1.42	1677	Stream	8	2.721	0.538	0.492	-0.027	-	-	-
Courbiere	B9	42.85	1.45	1622	Stream	6	2.854	0.569	0.524	0.007	-	-	-
Comapedrosa	A4	42.58	1.45	2127	Stream	5	3.194	0.494	0.513	0.147	-	-	-
Vicdessos ruisseau 1	V11	42.77	1.45	1040	Stream	1	-	-	-	-	-	-	-
Riu de Mossers	A8	42.45	1.46	1322	Stream	5	3.069	0.518	0.459	-0.017	-	-	-
Riu de Turer	A6	42.57	1.47	1803	Stream	7	3.077	0.528	0.505	0.034	-	-	-
Riu enclar	A7	42.50	1.48	1435	Stream	4	-	-	-	-	-	-	-
"B17"	B3	42.77	1.49	725	Cave	7	2.753	0.543	0.469	-0.081	-	-	-
Grotte Bernard	B6-A19	43.00	1.53	564	Cave	31	1.998	0.347	0.344	0.006	18	5	5(H5)
Siech	B4	42.88	1.55	690	Cave	7	2.149	0.370	0.343	0.000	-	4	4(H5)
Labouiche	F-B8	43.00	1.57	485	Cave	39	2.593	0.497	0.478	-0.028	32	10	10(H5)
Gisclareny	A16	42.26	1.72	1200	Stream	1	-	-	-	-	-	1	1(H9)
Salana Pas	A5	42.56	1.74	1948	Stream	2	-	-	-	-	-	-	-
Bellver Cerdanya	A15	42.33	1.76	1302	Stream	1	-	-	-	-	-	1	1(H9)
Berga	A17	42.11	1.83	882	Stream	4	-	-	-	-	-	4	4(H9)
Fontestorbes 1	E6	42.87	1.89	627	Stream	14	2.692	0.465	0.460	0.026	-	5	5(H5)
Fontestorbes 2	E7	42.87	1.89	663	Stream	21	2.773	0.455	0.484	0.086	25	-	-
Rieufourcant	E16	42.88	1.94	797	Stream	24	3.415	0.507	0.602	0.179	35	-	-

Orlu	M1-H	42.66	1.97	841	Stream	58	2.760	0.428	0.474	0.106	38	5	5(H5)
Belesta	E8	42.91	1.97	506	Stream	4	-	-	-	-	-	-	-
Gombren	A14	42.26	2.07	1055	Stream	2	-	-	-	-	-	2	2(H9)
Ripoll	A13	42.20	2.16	801	Stream	2	-	-	-	-	-	2	2(H9)
Cailla	B13	42.81	2.19	725	Stream	24	2.312	0.428	0.408	-0.026	24	-	-
Ruisseau de Fabournet Puilaurens	E9	42.81	2.28	595	Stream	5	2.544	0.576	0.439	-0.210	-	-	-
Cass-Rats	E10-B14	42.88	2.32	522	Stream	28	1.851	0.303	0.286	-0.038	20	10	9(H5), 1(H7)
Vidra	A11	42.12	2.35	1046	Stream	1	-	-	-	-	-	-	-
Font de Dotz	E3-B19	42.88	2.36	484	Stream	6	2.011	0.392	0.332	-0.090	-	-	-
Camprodon	D-A12	42.28	2.36	980	Stream	24	2.186	0.305	0.323	0.077	17	10	10(H9)
Rassa de l'Areny	A3-1	42.12	2.38	906	Stream	5	2.300	0.318	0.340	0.176	-	-	-
Hostalets d'en Bas	C	42.09	2.45	626	Stream	19	2.522	0.406	0.416	0.052	25	9	9(H9)
Riu Tec	A3-2	42.09	2.45	582	Stream	3	-	-	-	-	-	-	-
Sant Esteve d'en bas	A10	42.09	2.45	511	Stream	1	-	-	-	-	-	-	-
Auriac	B16	42.93	2.49	541	Stream	6	1.651	0.294	0.253	-0.071	-	5	5(H5)
Valmanya	B10	42.53	2.54	924	Stream	10	2.048	0.363	0.320	-0.082	-	4	4(H9)
Sant Aniol Finestres	A9	42.09	2.61	367	Stream	3	-	-	-	-	-	-	-
Rassa Mosquera	A3-3	42.09	2.61	372	Stream	6	2.528	0.451	0.416	0.006	-	-	-
Total						1299						258	

Table 2.S2 Posterior parameters (median and 95% confidence intervals) and RMedAD (Relative Median Absolute Deviation) estimated in the pilot runs with DIYABC for the best supported scenario (scenario 3) when considering all 1,299 individuals screened for microsatellites (simple sequence repeats – SSRs). See Figures 2.2 and 2.S1 for more information on the tested scenarios. Abbreviations: N, effective population size for each analysed deme (1 – cluster 1; 2 – cluster 2; 3 – cluster 3; 4 – cluster 4, 5 – cluster 5; 135 – central clusters; 241 – three oldest glacial refugia: eastern, western and central); t, time of events in generations (t_1 – time to the most recent split; t_2 – time to the intermediate split; t_3 – time to the most ancient split); mean μ, mean mutation rate; mean *P*, mean coefficient *P*; $Q_{2.5}$, quantile 2.5%; $Q_{97.5}$, quantile 97.5%.

	Microsatellites			
Parameter	Median	$Q_{2.5}$	$Q_{97.5}$	RMedAD
N_1	8 170	5 190	9 800	0.126
N_2	4 640	2 740	6 920	0.158
N_3	5 750	2 780	9 150	0.146
N_4	8 950	6 610	9 920	0.130
N_5	5 290	2 470	9 040	0.165
N_{135}	3 400	361	9 310	0.410
N_{241}	2 210	154	11 600	0.560
t_1	2 570	907	6 020	0.278
t_2	4 850	1 660	9 410	0.238
t_3	8 160	3 150	14 400	0.229
Mean $\mu_{(SSRs)}$	2.80×10^{-4}	1.56×10^{-4}	5.63×10^{-4}	0.394
Mean $P_{(SSRs)}$	0.275	0.159	0.300	0.209

Table 2.S3 Parameters used in DIYABC analysis and respective priors for the best supported scenario (scenario 3). See Figures 2.2 and 2.S1 for more information on tested scenarios. Abbreviations: Parameters: N, effective population size for each analysed deme (1 – cluster 1; 2 – cluster 2; 3 – cluster 3; 4 – cluster 4, 5 – cluster 5; 135 – central clusters; 241 – three oldest glacial refugia: eastern, western and central); t, time of events in generations (t_1 – time to the most recent split; t_2 – time to the intermediate split; t_3 – time to the most ancient split); Microsatellite ($_{SSRs}$) and mitochondrial ($_{cyt\text{-}b}$) parameters: mean μ, mean mutation rate; individual locus μ, individual locus mutation rate; mean *P*, mean coefficient *P*; individual locus *P*, individual locus coefficient *P*; SNI, Single Nucleotide Insertion rate; mean *k1*, mean coefficient *k1*; individual locus *k1*, individual locus coefficient *k1*. Conditions: sequence data were simulated under a Kimura two-parameter (K2P) mutation model. Distribution: parameter distributions were left as default, with the exception of maximum N values, which were set to 15,000 or 20,000, maximum t value for the most ancient split, which was set to 20,000, SNI rate that was set to zero, and minimum and maximum mutation rates for the cyt-*b* gene that were set to 10^{-10} and 10^{-6}, respectively.

Parameter	Conditions	Distribution [min-max]
N_1		Uniform [10 - 15 000]
N_2		Uniform [10 - 15 000]
N_3		Uniform [10 - 15 000]
N_4		Uniform [10 - 15 000]
N_5		Uniform [10 - 15 000]
N_{135}		Uniform [10 - 15 000]
N_{241}		Uniform [10 - 20 000]
t_1		Uniform [10 - 10 000]
t_2	$t_2>t_1$	Uniform [10 - 10 000]
t_3	$t_3>t_2$	Uniform [10 - 20 000]
Mean $\mu_{(SSRs)}$		Uniform [10^{-4} - 10^{-3}]
Individual locus $\mu_{(SSRs)}$		Gamma [10^{-5} - 10^{-2}]
Mean $P_{(SSRs)}$		Uniform [10^{-1} - $3x10^{-1}$]
Individual locus $P_{(SSRs)}$		Gamma [10^{-2} - $9x10^{-1}$]
$SNI_{(SSRs)}$		Log-u [0]
Mean $\mu_{(cyt\text{-}b)}$	K2P	Uniform [10^{-10} - 10^{-6}]
Individual locus $\mu_{(cyt\text{-}b)}$	K2P	Gamma [10^{-10} - 10^{-6}]
Mean $k1_{(cyt\text{-}b)}$	K2P	Uniform [0.05 - 20]
Individual locus $k1_{(cyt\text{-}b)}$	K2P	Gamma [0.05 - 20]

Table 2.S4 Assignment of individuals to populations of origin and admixture proportions to other sampling localities in *Calotriton asper* as estimated with GeneClass2. Numbers in parentheses indicate minor assignment proportions (only assignment probabilities ≥ 0.1 were considered). For population codes see Table 2.S1. Abbreviations: AProb, probability of assignment to the population of collection.

					Cluster 1															
		Highest AProb	Lowest AProb	Mean AProb	G-A1	B6-A19	B5	F-B8	B13	E10-B14	E5	E6	E7	E16	M1-H	V28	V29	V30	AUL	VED
Cluster 1	G-A1	0.94	0.01	0.38	39	-	(10)	-	-	-	(17)	-	-	(1)	-	-	-	-	-	(1)
	B6-A19	0.99	0	0.45	(5)	30	-	-	-	-	-	-	-	(2)	-	-	-	-	-	-
	B5	0.87	0.01	0.38	4(5)	-	11(3)	-	-	-	(9)	-	-	-	-	-	(1)	-	(2)	(1)
	F-B8	0.99	0.05	0.50	-	-	(2)	38	-	-	-	-	-	-	-	-	-	-	-	(2)
	B13	0.93	0.01	0.45	(4)	-	(1)	-	16(4)	-	(1)	(5)	(4)	4(18)	3(10)	-	-	-	-	-
	E10-B14	0.92	0	0.51	(6)	-	-	-	(11)	11(15)	-	(5)	(15)	14(12)	1(19)	-	(1)	-	-	-
	E5	0.97	0.01	0.40	6(9)	-	1(4)	-	-	-	22(7)	-	-	-	-	-	-	-	(1)	-
	E6	0.94	0	0.40	(2)	-	-	-	-	-	-	3(9)	6(6)	4(8)	(4)	-	-	-	-	-
	E7	0.95	0	0.45	(1)	-	-	-	-	-	(1)	2(12)	10(7)	8(12)	(4)	-	-	-	-	-
	E16	0.97	0	0.41	(2)	-	(2)	-	-	-	-	-	(5)	24	(3)	-	-	-	-	-
	M1-H	0.99	0.01	0.55	(10)	-	(4)	-	-	-	(2)	(4)	(3)	1(21)	57	-	-	-	-	-
	V28	0.95	0.01	0.48	(8)	-	(2)	-	-	-	(1)	-	-	-	-	3(9)	5(8)	6(8)	-	-
	V29	0.90	0.02	0.37	3(3)	-	(6)	-	-	-	(4)	-	-	(1)	-	(3)	10(3)	(5)	-	-
	V30	0.93	0.01	0.46	1(9)	-	(1)	-	-	-	(4)	-	-	-	-	8(7)	5(10)	2(12)	-	-
	AUL	0.97	0.06	0.42	(1)	-	-	-	-	-	(1)	-	-	-	-	-	-	-	18	-
	VED	0.78	0.04	0.38	(1)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	12
Cluster 2	D-A12	0.98	0	0.46	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B-A18	0.94	0	0.49	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B10	0.76	0.01	0.38	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	C	0.92	0	0.51	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cluster 3	B1	1	0	0.61	-	-	-	-	-	-	-	-	-	-	-	-	-	-	(1)	-
	IBA	0.93	0.04	0.43	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	BMO	0.98	0.01	0.45	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	IPE	1	0.05	0.61	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	IVI	0.96	0.10	0.43	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CPL	1	0.01	0.55	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Cluster 4	B2	0.91	0.02	0.48	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B7	1	0	0.70	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B11	0.98	0.01	0.32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	B12	0.78	0	0.34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	E4	0.98	0	0.38	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	E13	0.76	0.02	0.27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	E14	0.93	0.01	0.41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	E15	1	0.01	0.40	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	I	0.83	0.03	0.41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	ISA	0.99	0.01	0.48	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	IES	0.99	0.01	0.60	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	IAC	0.98	0.01	0.54	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cluster 5	XSU	0.99	0.01	0.54	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	(3)
	MDE	1	0	0.57	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	(1)
	NDE	1	0.02	0.57	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	IDE	1	0	0.56	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	BUX	0.99	0.04	0.61	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CPX	1	0.01	0.50	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	(1)	(1)
	CTS	0.85	0.01	0.43	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	RQN	0.99	0	0.46	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	RSU	1	0.07	0.63	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	RVT-A	0.99	0	0.54	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	E	0.93	0	0.39	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

					Cluster 2				Cluster 3					
		Highest AProb	Lowest AProb	Mean AProb	D-A12	B-A18	B10	C	B1	IBA	BMO	IPE	IVI	CPL
Cluster 1	G-A1	0.94	0.01	0.38	-	-	-	-	-	-	-	-	-	-
	B6-A19	0.99	0	0.45	-	-	-	-	-	-	-	-	-	-
	B5	0.87	0.01	0.38	-	-	-	-	-	-	-	-	-	-
	F-B8	0.99	0.05	0.50	-	-	-	-	-	-	-	-	-	-
	B13	0.93	0.01	0.45	-	-	-	-	-	-	-	-	-	-
	E10-B14	0.92	0	0.51	-	-	-	-	-	-	-	-	-	-
	E5	0.97	0.01	0.40	-	-	-	-	-	-	-	-	-	-
	E6	0.94	0	0.40	-	-	-	-	-	-	-	-	-	-
	E7	0.95	0	0.45	-	-	-	-	-	-	-	-	-	-
	E16	0.97	0	0.41	-	-	-	-	-	-	-	-	-	-
	M1-H	0.99	0.01	0.55	-	-	-	-	-	-	-	-	-	-
	V28	0.95	0.01	0.48	-	-	-	-	-	-	-	-	-	-
	V29	0.90	0.02	0.37	-	-	-	-	-	-	-	-	-	-
	V30	0.93	0.01	0.46	-	-	-	-	-	-	-	-	-	-
	AUL	0.97	0.06	0.42	-	-	-	-	-	-	-	-	-	-
	VED	0.78	0.04	0.38	-	-	-	-	-	-	-	-	-	-
Cluster 2	D-A12	0.98	0	0.46	15(8)	-	1(6)	7(14)	-	-	-	-	-	-
	B-A18	0.94	0	0.49	-	25	-	-	-	-	-	-	-	-
	B10	0.76	0.01	0.38	(7)	-	3(6)	7(3)	-	-	-	-	-	-
	C	0.92	0	0.51	(3)	-	(3)	19	-	-	-	-	-	-
Cluster 3	B1	1	0	0.61	-	-	-	-	15	-	-	-	-	-
	IBA	0.93	0.04	0.43	-	-	-	-	-	16	(1)	-	-	-
	BMO	0.98	0.01	0.45	-	-	-	-	-	(1)	22	-	-	-
	IPE	1	0.05	0.61	-	-	-	-	-	-	-	29	-	-
	IVI	0.96	0.10	0.43	-	-	-	-	-	-	-	-	25	-
	CPL	1	0.01	0.55	-	-	-	-	-	-	-	-	-	30
Cluster 4	B2	0.91	0.02	0.48	-	-	-	-	-	-	-	-	-	-
	B7	1	0	0.70	-	-	-	-	-	-	-	-	-	-
	B11	0.98	0.01	0.32	-	-	-	-	-	-	-	-	-	-
	B12	0.78	0	0.34	-	-	-	-	-	-	-	-	-	-
	E4	0.98	0	0.38	-	-	-	-	-	-	-	-	-	-
	E13	0.76	0.02	0.27	-	-	-	-	-	-	-	-	-	-
	E14	0.93	0.01	0.41	-	-	-	-	-	-	-	-	-	-
	E15	1	0.01	0.40	-	-	-	-	-	-	-	-	-	-
	I	0.83	0.03	0.41	-	-	-	-	-	-	-	-	-	-
	ISA	0.99	0.01	0.48	-	-	-	-	-	-	-	-	-	-
	IES	0.99	0.01	0.60	-	-	-	-	-	-	-	-	-	-
	IAC	0.98	0.01	0.54	-	-	-	-	-	-	-	-	-	-
Cluster 5	XSU	0.99	0.01	0.54	-	-	-	-	-	-	-	-	-	-
	MDE	1	0	0.57	-	-	-	-	-	-	-	-	-	-
	NDE	1	0.02	0.57	-	-	-	-	-	-	-	-	-	-
	IDE	1	0	0.56	-	-	-	-	-	-	-	-	-	-
	BUX	0.99	0.04	0.61	-	-	-	-	-	-	-	-	-	-
	CPX	1	0.01	0.50	-	-	-	-	-	-	-	-	-	-
	CTS	0.85	0.01	0.43	-	-	-	-	-	-	-	-	-	-
	RQN	0.99	0	0.46	-	-	-	-	-	-	-	-	-	-
	RSU	1	0.07	0.63	-	-	-	-	-	-	-	-	-	-
	RVT-A	0.99	0	0.54	-	-	-	-	-	-	-	-	-	-
	E	0.93	0	0.39	-	-	-	-	-	-	-	-	-	-

					Cluster 4											
		Highest AProb	Lowest AProb	Mean AProb	B2	B7	B11	B12	E4	E13	E14	E15	I	ISA	IES	IAC
Cluster 1	G-A1	0.94	0.01	0.38	-	-	-	-	-	-	-	-	-	-	-	-
	B6-A19	0.99	0	0.45	-	-	-	-	-	-	-	-	-	-	-	-
	B5	0.87	0.01	0.38	-	-	-	-	-	-	-	-	-	-	-	-
	F-B8	0.99	0.05	0.50	-	-	-	-	-	-	-	-	-	-	-	-
	B13	0.93	0.01	0.45	-	-	-	-	-	-	-	-	-	-	-	-
	E10-B14	0.92	0	0.51	-	-	-	-	-	-	-	-	-	-	-	-
	E5	0.97	0.01	0.40	-	-	-	-	-	-	-	-	-	-	-	-
	E6	0.94	0	0.40	-	-	-	-	-	-	-	-	-	-	-	-
	E7	0.95	0	0.45	-	-	-	-	-	-	-	-	-	-	-	-
	E16	0.97	0	0.41	-	-	-	-	-	-	-	-	-	-	-	-
	M1-H	0.99	0.01	0.55	-	-	-	-	-	-	-	-	-	-	-	-
	V28	0.95	0.01	0.48	-	-	-	-	-	-	-	-	-	-	-	-
	V29	0.90	0.02	0.37	-	-	-	-	-	-	-	-	-	-	-	-
	V30	0.93	0.01	0.46	-	-	-	-	-	-	-	-	-	-	-	-
	AUL	0.97	0.06	0.42	-	-	-	-	-	-	-	-	-	-	-	-
	VED	0.78	0.04	0.38	-	-	-	-	-	-	-	-	-	-	-	-
Cluster 2	D-A12	0.98	0	0.46	-	-	-	-	-	-	-	-	-	-	-	-
	B-A18	0.94	0	0.49	-	-	-	-	-	-	-	-	-	-	-	-
	B10	0.76	0.01	0.38	-	-	-	-	-	-	-	-	-	-	-	-
	C	0.92	0	0.51	-	-	-	-	-	-	-	-	-	-	-	-
Cluster 3	B1	1	0	0.61	-	-	-	-	-	-	-	-	-	-	-	-
	IBA	0.93	0.04	0.43	-	-	-	-	-	-	-	-	-	-	-	-
	BMO	0.98	0.01	0.45	-	-	-	-	-	-	-	-	-	-	-	-
	IPE	1	0.05	0.61	-	-	-	-	-	-	-	-	-	-	-	-
	IVI	0.96	0.10	0.43	-	-	-	-	-	-	-	-	-	-	-	-
	CPL	1	0.01	0.55	-	-	-	-	-	-	-	-	-	-	-	-
Cluster 4	B2	0.91	0.02	0.48	16	-	-	-	-	-	-	-	-	-	(1)	-
	B7	1	0	0.70	2(4)	26(1)	-	-	-	-	-	-	-	-	(1)	-
	B11	0.98	0.01	0.32	-	-	11	1(4)	-	-	-	-	-	-	-	-
	B12	0.78	0	0.34	-	-	1(8)	13(1)	-	-	-	-	-	-	-	-
	E4	0.98	0	0.38	-	-	-	-	29	-	(2)	-	-	-	-	-
	E13	0.76	0.02	0.27	-	-	-	-	-	12	-	-	-	-	-	-
	E14	0.93	0.01	0.41	-	-	-	-	(1)	-	30	-	-	-	-	-
	E15	1	0.01	0.40	(1)	-	-	-	-	(3)	-	18	-	-	-	-
	I	0.83	0.03	0.41	-	-	-	-	-	-	-	-	14	-	-	-
	ISA	0.99	0.01	0.48	-	-	-	-	-	-	-	-	-	14	-	-
	IES	0.99	0.01	0.60	-	-	-	-	-	1	-	-	-	-	22	-
	IAC	0.98	0.01	0.54	-	-	-	-	-	-	-	-	-	-	-	26
Cluster 5	XSU	0.99	0.01	0.54	-	-	-	-	-	-	-	-	-	-	-	-
	MDE	1	0	0.57	-	-	-	-	-	-	-	-	-	-	-	-
	NDE	1	0.02	0.57	-	-	-	-	-	-	-	-	-	-	-	-
	IDE	1	0	0.56	-	-	-	-	-	-	-	-	-	-	-	-
	BUX	0.99	0.04	0.61	-	-	-	-	-	-	-	-	-	-	-	-
	CPX	1	0.01	0.50	-	-	-	-	-	-	-	-	-	-	-	-
	CTS	0.85	0.01	0.43	-	-	-	-	-	-	-	-	-	-	-	-
	RQN	0.99	0	0.46	-	-	-	-	-	-	-	-	-	-	-	-
	RSU	1	0.07	0.63	-	-	-	-	-	-	-	-	-	-	-	-
	RVT-A	0.99	0	0.54	-	-	-	-	-	-	-	-	-	-	-	-
	E	0.93	0	0.39	-	-	-	-	-	-	-	-	-	-	-	-

					Cluster 5										
		Highest AProb	Lowest AProb	Mean AProb	XSU	MDE	NDE	IDE	BUX	CPX	CTS	RQN	RSU	RVT-A	E
Cluster 1	G-A1	0.94	0.01	0.38	-	-	-	-	-	-	-	-	-	-	-
	B6-A19	0.99	0	0.45	-	-	-	-	-	-	-	-	-	-	-
	B5	0.87	0.01	0.38	-	-	-	-	-	-	-	-	-	-	-
	F-B8	0.99	0.05	0.50	-	-	-	-	-	-	-	-	-	-	-
	B13	0.93	0.01	0.45	-	-	-	-	-	-	-	-	-	-	-
	E10-B14	0.92	0	0.51	-	-	-	-	-	-	-	-	-	-	-
	E5	0.97	0.01	0.40	-	-	-	-	-	-	-	-	-	-	-
	E6	0.94	0	0.40	-	-	-	-	-	-	-	-	-	-	-
	E7	0.95	0	0.45	-	-	-	-	-	-	-	-	-	-	-
	E16	0.97	0	0.41	-	-	-	-	-	-	-	-	-	-	-
	M1-H	0.99	0.01	0.55	-	-	-	-	-	-	-	-	-	-	-
	V28	0.95	0.01	0.48	-	-	-	-	-	-	-	-	-	-	-
	V29	0.90	0.02	0.37	-	-	-	-	-	-	-	-	-	-	-
	V30	0.93	0.01	0.46	-	-	-	-	-	-	-	-	-	-	-
	AUL	0.97	0.06	0.42	-	-	-	-	-	-	-	-	-	-	-
	VED	0.78	0.04	0.38	-	-	-	-	-	-	-	-	-	-	-
Cluster 2	D-A12	0.98	0	0.46	-	-	-	-	-	-	-	-	-	-	-
	B-A18	0.94	0	0.49	-	-	-	-	-	-	-	-	-	-	-
	B10	0.76	0.01	0.38	-	-	-	-	-	-	-	-	-	-	-
	C	0.92	0	0.51	-	-	-	-	-	-	-	-	-	-	-
Cluster 3	B1	1	0	0.61	-	-	-	-	-	-	-	-	-	-	-
	IBA	0.93	0.04	0.43	-	-	-	-	-	-	-	-	-	-	-
	BMO	0.98	0.01	0.45	-	-	-	-	-	-	-	-	-	-	-
	IPE	1	0.05	0.61	-	-	-	-	-	-	-	-	-	-	-
	IVI	0.96	0.10	0.43	-	-	-	-	-	-	-	-	-	-	-
	CPL	1	0.01	0.55	-	-	-	-	-	-	-	-	-	-	-
Cluster 4	B2	0.91	0.02	0.48	-	-	-	-	-	-	-	-	-	-	-
	B7	1	0	0.70	-	-	-	-	-	-	-	-	-	-	-
	B11	0.98	0.01	0.32	-	-	-	-	-	-	-	-	-	-	-
	B12	0.78	0	0.34	-	-	-	-	-	-	-	-	-	-	-
	E4	0.98	0	0.38	-	-	-	-	-	-	-	-	-	-	-
	E13	0.76	0.02	0.27	-	-	-	-	-	-	-	-	-	-	-
	E14	0.93	0.01	0.41	-	-	-	-	-	-	-	-	-	-	-
	E15	1	0.01	0.40	-	-	-	-	-	-	-	-	-	-	-
	I	0.83	0.03	0.41	-	-	-	-	-	-	-	-	-	-	-
	ISA	0.99	0.01	0.48	-	-	-	-	-	-	-	-	-	-	-
	IES	0.99	0.01	0.60	-	-	-	-	-	-	-	-	-	-	-
	IAC	0.98	0.01	0.54	-	-	-	-	-	-	-	-	-	-	-
Cluster 5	XSU	0.99	0.01	0.54	31(3)	-	-	-	-	(1)	-	-	4(19)	-	-
	MDE	1	0	0.57	(1)	30(7)	1(14)	1(16)	-	(3)	3(21)	-	3(9)	-	-
	NDE	1	0.02	0.57	(3)	18(15)	14(20)	3(12)	-	(3)	(11)	-	(6)	-	-
	IDE	1	0	0.56	(2)	11(7)	(11)	6(12)	-	-	1(13)	-	(7)	-	-
	BUX	0.99	0.04	0.61	-	-	-	-	15	-	-	-	-	-	-
	CPX	1	0.01	0.50	(2)	-	-	-	-	16	(1)	-	(4)	-	-
	CTS	0.85	0.01	0.43	1(1)	(9)	(2)	(3)	-	(3)	8(3)	-	3(6)	-	-
	RQN	0.99	0	0.46	-	-	-	-	-	-	-	14	-	-	(1)
	RSU	1	0.07	0.63	3(6)	-	-	-	(1)	(2)	(1)	-	18	-	-
	RVT-A	0.99	0	0.54	-	-	-	-	-	-	-	-	-	26	-
	E	0.93	0	0.39	-	-	-	-	-	-	-	(2)	-	-	19

Supplementary figures

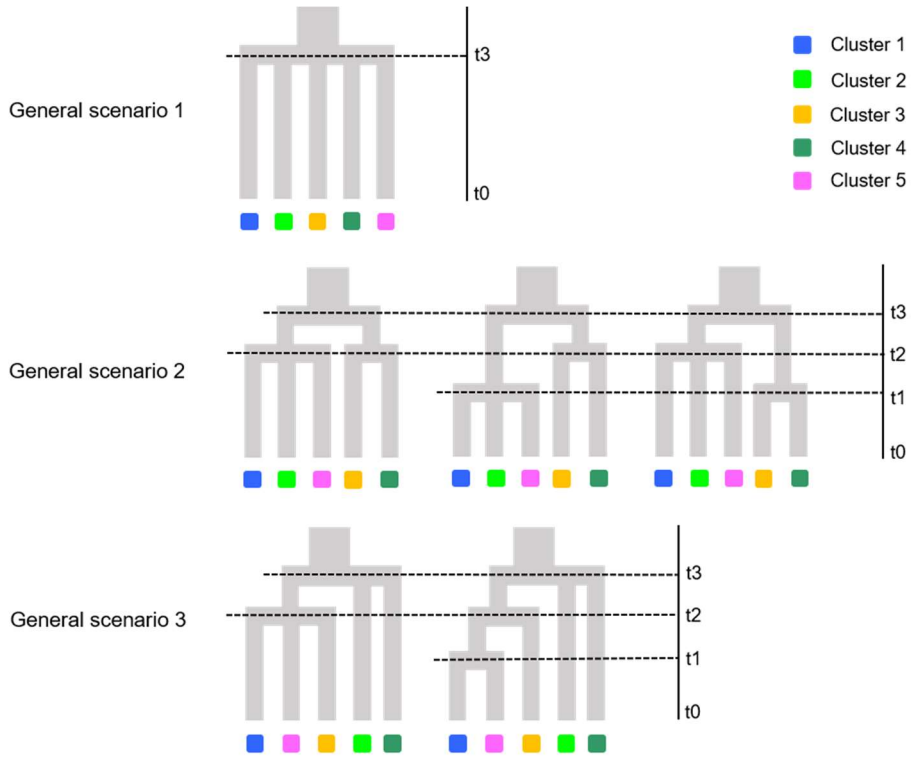


Figure 2.S1 Phylogeographic models of *Calotriton asper* evolution tested with DIYABC during phase 1. The best scenario for each general topology was selected for phase 2, where we compared the best trees for each general scenario against each other. See Table 2.S3 for more information on parameters for effective population size and time of events (t), as well as prior distributions.

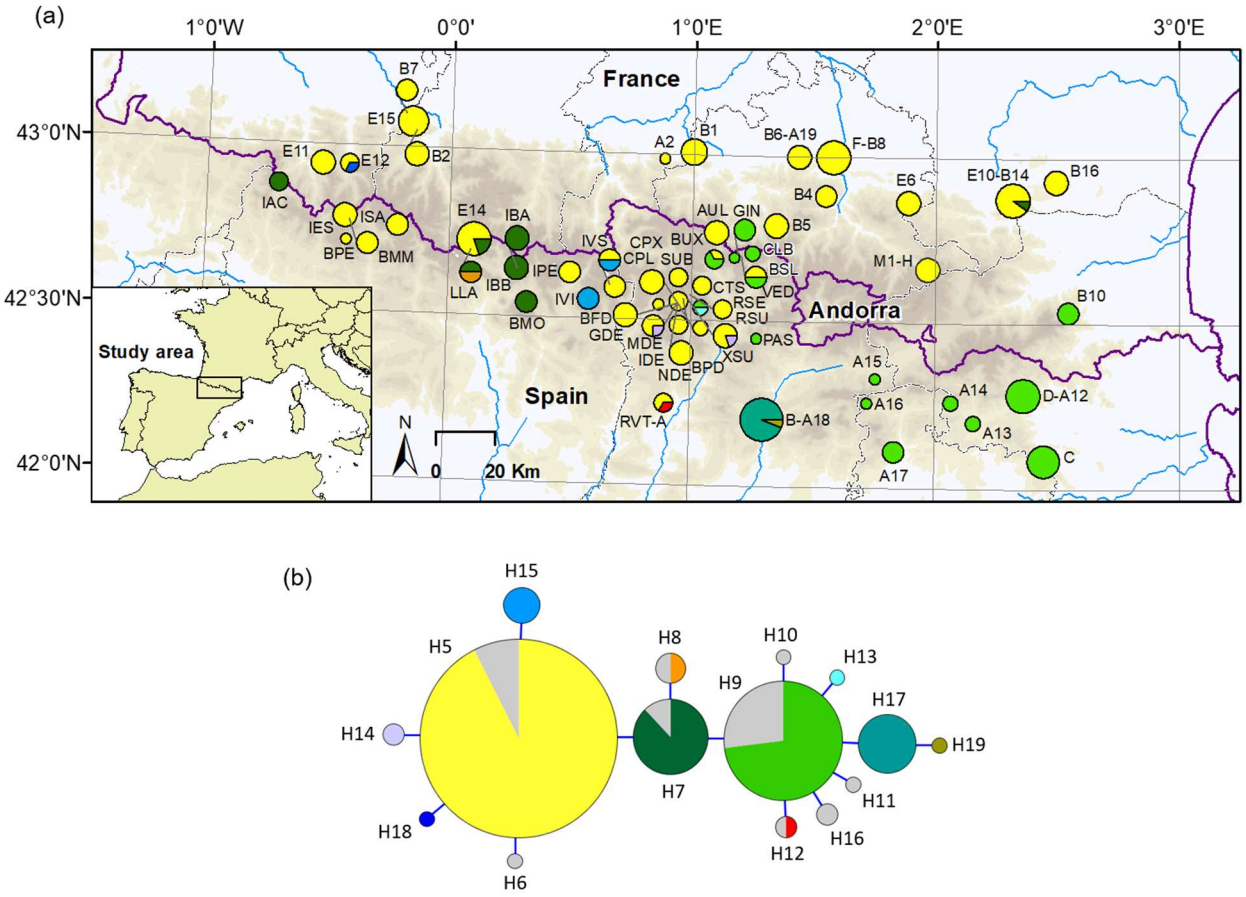


Figure 2.S2 Results of the network analysis of the cytochrome *b* (cyt-*b*) mtDNA across *Calotriton asper* distribution range. Panel (a) shows the geographic distribution of cyt-*b* haplotypes identified from 258 individuals. The size of circles is proportional to the number of analysed sequences. For population codes, as well as number of analysed sequences in each population, see Table 2.S1. Panel (b) shows the haplotype network inferred from cyt-*b* sequences. Each circle represents a unique haplotype and the circle area is proportional to the number of sequences of a given haplotype. Connecting lines correspond to one mutational step. Colours and codes of haplotypes follow Valbuena-Ureña et al. (2013). Sequences depicted in grey were retrieved from GenBank.

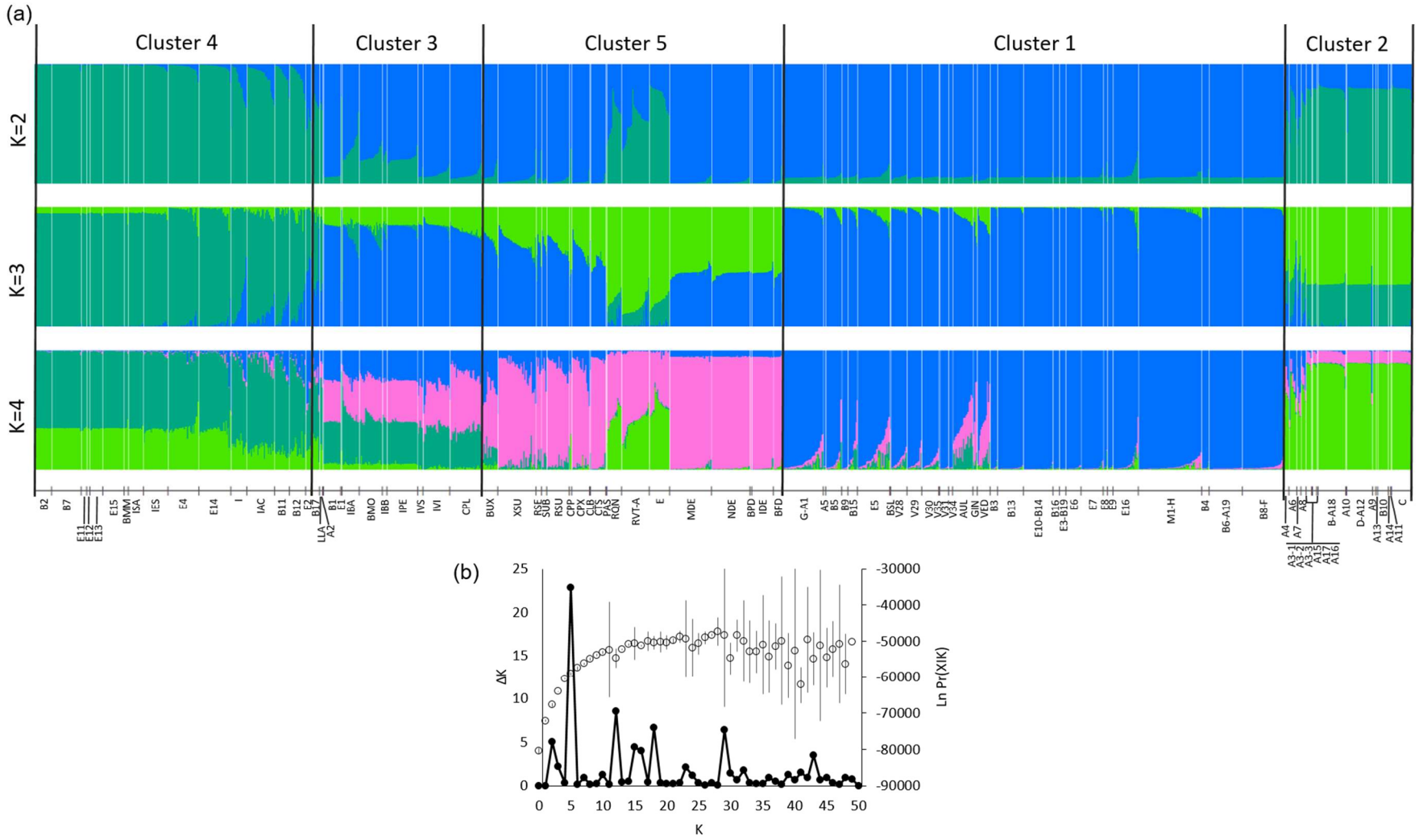


Figure 2.S3 Resulting plots from STRUCTURE analysis for the complete dataset. Panel (a) shows barplots of membership assignment for K = 2-4 genetic clusters: each individual is represented by a vertical bar corresponding to the sum of assignment probabilities to the K cluster. White lines separate populations and black lines separate clusters. Further information on population codes and sampling sites are given in Table 2.S1. In panel (b), the right axis (open dots with error bars) displays mean (± SD) log probability of the data [Ln Pr(XIK)] over 20 runs, for each value of K; the left axis (black dots) shows ΔK values as a function of K, calculated according to Evanno et al. (2005).

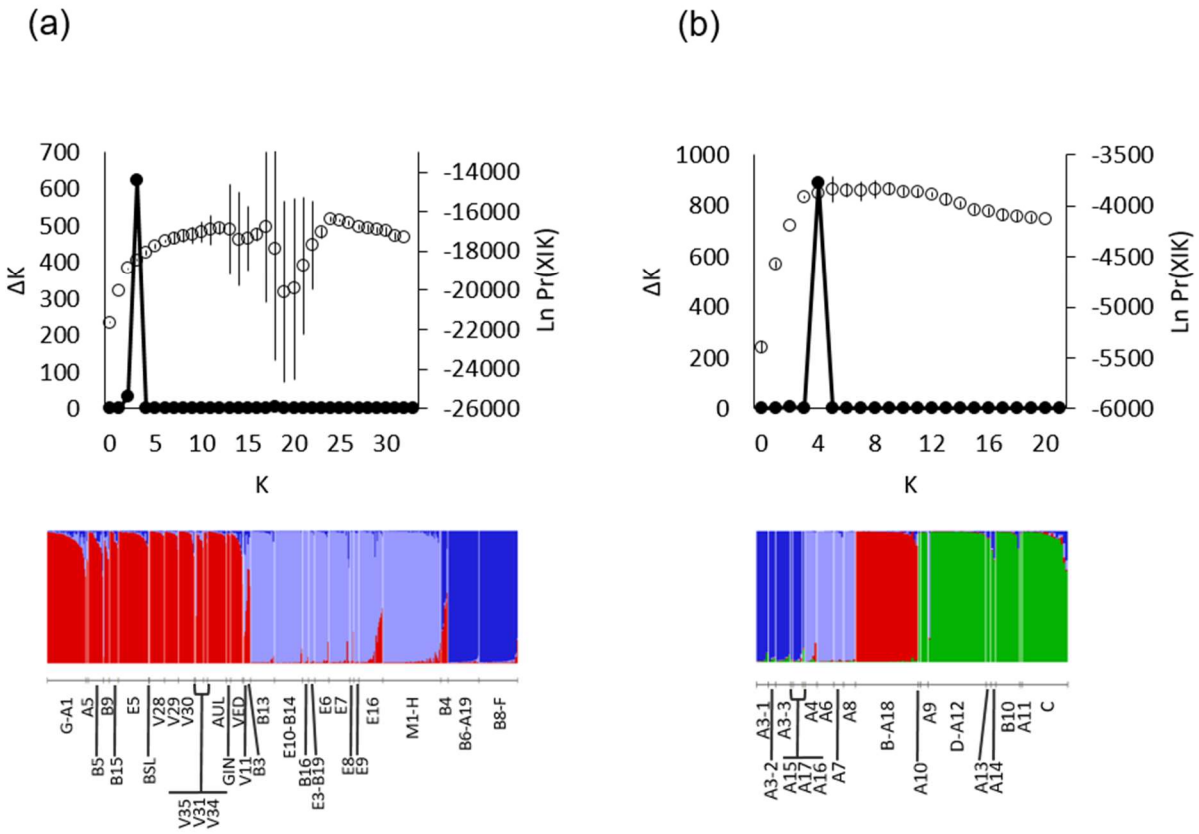


Figure 2.S4 Resulting plots from STRUCTURE analysis for cluster 1 (eastern French populations; a) and cluster 2 (eastern Spanish populations; b). In the upper plots, the right axis (open dots with error bars) displays mean (± SD) log probability of the data [Ln Pr(XIK)] over 20 runs, for each value of K; the left axis (black dots) shows ΔK values as a function of K, calculated according to Evanno et al. (2005). Lower plots represent barplots of membership assignment for K = 3 (cluster 1; a) and K = 4 (cluster 2; b): each individual is represented by a vertical bar corresponding to the sum of assignment probabilities to the K cluster. White lines separate populations. Further information on population codes and sampling sites are given in Table 2.S1.

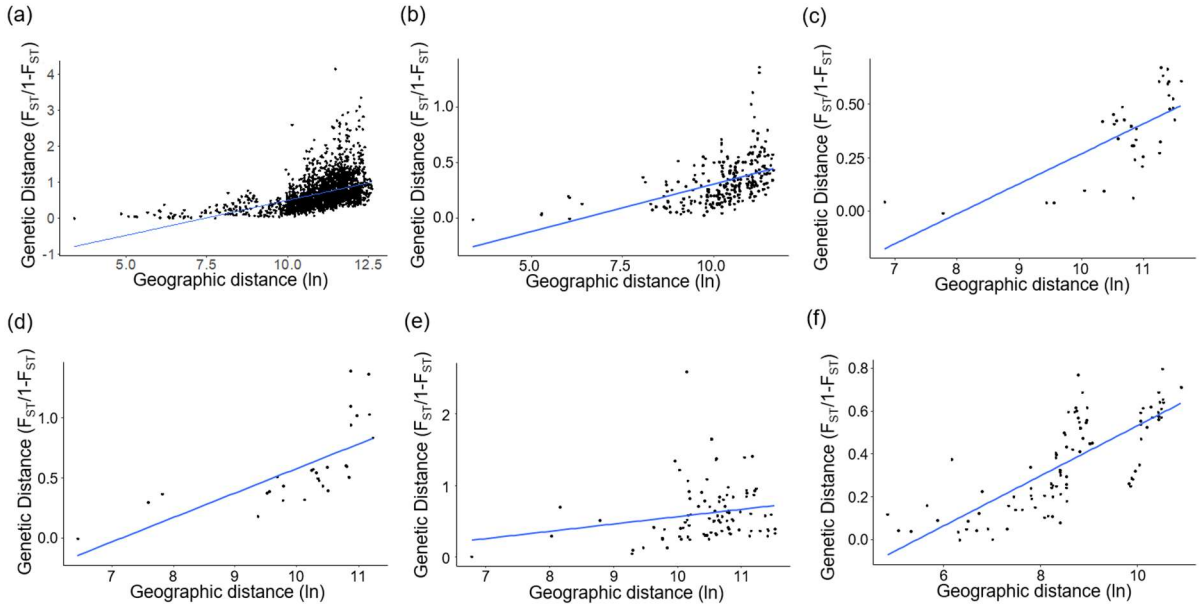


Figure 2.S5 IBD (isolation by distance) analysis over all *Calotriton asper* population pairs (a) and for each genetic cluster as inferred by STRUCTURE (b, cluster 1; c, cluster 2; d, cluster 3; e, cluster 4; f, cluster 5).

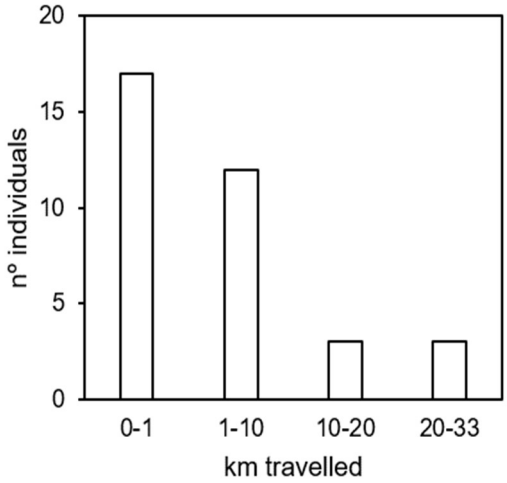


Figure 2.S6 Histogram showing the Euclidean distance covered by *Calotriton asper* first generation migrants as inferred by GeneClass2. Only individuals whose source locality could be determined are shown.

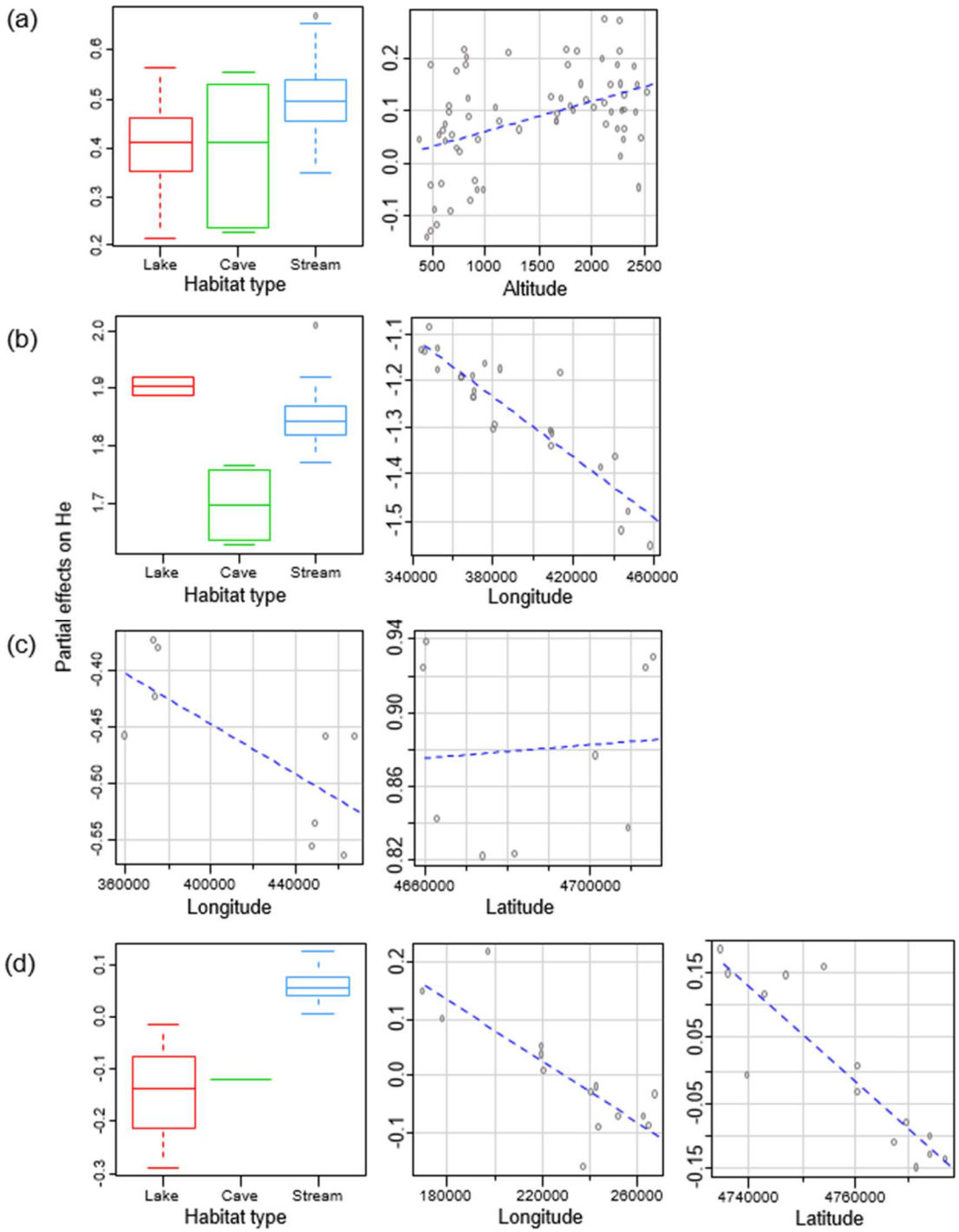


Figure 2.S7 Partial effects of environmental (habitat type) and geographic (latitude, longitude, altitude) variables on expected heterozygosity (H_E). (a), all populations; (b), cluster 1; (c), cluster 2; (d), cluster 4. Only variables that had a significant effect on H_E as determined by linear model selection are drawn. Latitude and longitude are in UTM coordinates and altitude is expressed in meters.

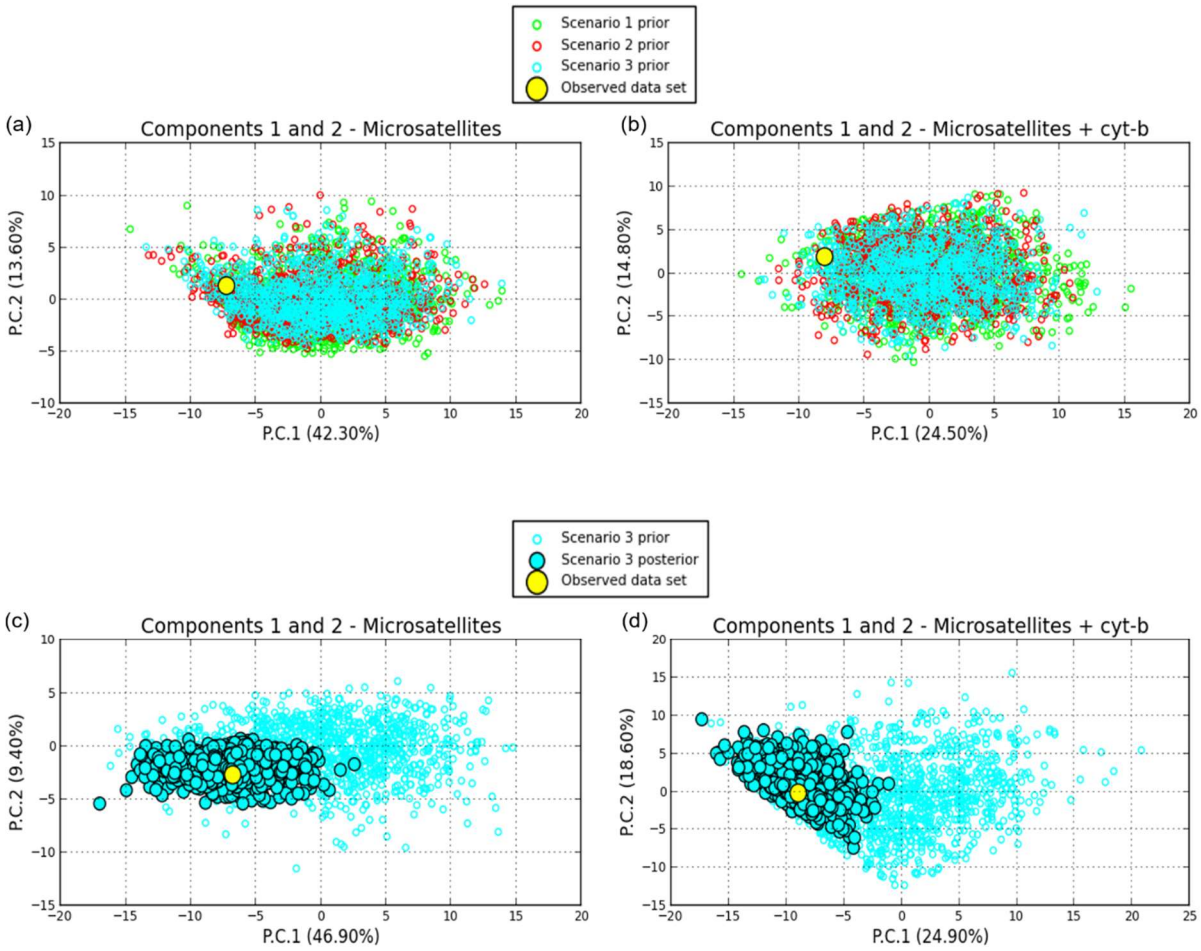


Figure 2.S8 Resulting plots from DIYABC analysis. Panels (a) and (b) show the pre-evaluation step of scenarios tested during phase 2, conducted through a PCA on summary statistics of simulated and observed datasets. Panels (c) and (d) show the model checking step for the most probable scenario (scenario 3), through performance of a PCA on the simulated datasets generated from posterior parameter distributions and the observed dataset. Simulations were performed only including microsatellites (panels a and c) and considering both mtDNA and microsatellite markers (panels b and d). The observed dataset (yellow dot) falls within the cloud of simulated points, indicating a good fit between observed and simulated datasets under the chosen priors. See Figure 2.S1 for details on the tested scenarios.

References

Barbosa S, Pauperio J, Herman JS, Ferreira CM, Pita R, Vale-Goncalves HM et al. (2017) Endemic species may have complex histories: Within-refugium phylogeography of an endangered Iberian vole. *Molecular Ecology* **26:** 951-967.

Bergl RA, Vigilant L (2007) Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*). *Molecular Ecology* **16:** 501-516.

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

Cornuet JM, Pudlo P, Veyssier J, Loire E, Santos F, Dehne-Garcia A et al. (2015) DIYABC version 2.1. A user-friendly software for inferring population history through

Approximate Bayesian Computations using microsatellite, DNA sequence and SNP data. Montferrier-sur-Lez, France: Institut National de la Recherche Agronomique.

Cornuet JM, Ravigne V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* **11:** 401.

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10:** 564-567.

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.

Gu Z, Gu L, Eils R, Schlesner M, Brors B (2014) *circlize* implements and enhances circular visualization in R. *Bioinformatics* **30:** 2811-2812.

Mokhtar-Jamai K, Coma R, Wang J, Zuberer F, Feral JP, Aurelle D (2013) Role of evolutionary and ecological factors in the reproductive success and the spatial genetic structure of the temperate gorgonian *Paramuricea clavata*. *Ecology and Evolution* **3:** 1765-1779.

Montori A (1988) Estudio sobre la biología y ecología del tritón pirenaico *Euproctus asper* (Dugès, 1852) en la Cerdanya. PhD thesis, University of Barcelona, Spain.

Montori A, Llorente GA (2014) Tritón pirenaico–*Calotriton asper* (Dugès, 1852). In: Salvador A and Martínez-Solano I (eds) *Enciclopedia Virtual de los Vertebrados*. Museo Nacional de Ciencias Naturales: Madrid, Spain p28.

Ortego J, Noguerales V, Gugger PF, Sork VL (2015) Evolutionary and demographic history of the Californian scrub white oak species complex: An integrative approach. *Molecular Ecology* **24:** 6188-6208.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: A simulation-based exploration of accuracy and power. *Molecular Ecology* **13:** 55-65.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 9197-9201.

Valbuena-Ureña E, Amat F, Carranza S (2013) Integrative phylogeography of *Calotriton* newts (Amphibia, Salamandridae), with special remarks on the conservation of the endangered Montseny brook newt (*Calotriton arnoldi*). *PLoS One* **8:** e62542.

Valbuena-Ureña E, Oromi N, Soler-Membrives A, Carranza S, Amat F, Camarasa S et al. (2018) Jailed in the mountains: Genetic diversity and structure of an endemic newt species across the Pyrenees. *PLoS One* **13:** e0200214.

Chapter 3

Contrasting patterns of genetic admixture explain the phylogeographic history of Iberian high mountain populations of midwife toads

Federica Lucati, Alexandre Miró, Jaime Bosch, Jenny Caner, Michael Joseph Jowers, Xavier Rivera, David Donaire-Barroso, Marc Ventura

Abstract

Multiple Quaternary glacial refugia in the Iberian Peninsula, commonly known as "refugia within refugia", allowed diverging populations to come into contact and admix following range shifts tracking the climatic fluctuations, potentially boosting substantial mito-nuclear discordances. In this study, we employ a comprehensive set of mitochondrial and nuclear markers to shed light onto the drivers of geographical differentiation in Iberian high mountain populations of the midwife toads *Alytes obstetricans* and *A. almogavarii* from the Pyrenees, Picos de Europa and Guadarrama Mountains. In the three analysed mountain regions, we detected evidence of extensive mito-nuclear discordances and admixture between taxa. Clustering analyses identified three major divergent lineages in the Pyrenees (corresponding to the eastern, central and central-western Pyrenees), which recurrently expanded and admixed during the succession of glacial-interglacial periods that characterised the Late Pleistocene. On the other hand, populations from the Picos de Europa mountains (NW Iberian Peninsula) showed a mitochondrial affinity to central-western Pyrenean populations and a nuclear affinity to populations from the central Iberian Peninsula, suggesting a likely admixed origin for Picos de Europa populations. Finally, populations from the Guadarrama Mountain Range (central Iberian Peninsula) were depleted of genetic diversity, possibly as a consequence of a recent epidemic of chytridiomycosis. This work highlights the complex evolutionary history that shaped the current genetic composition of high mountain populations, and underscores the importance of using a multilocus approach to capture true dynamics of population divergence.

Keywords: *Alytes obstetricans*, *Alytes almogavarii*, chytrid fungus, contact zones, glacial refugia, admixture

Introduction

Patterns of population structure and genetic divergence within species primarily result from their evolutionary history and contemporary dispersal capability, and these processes are in turn responsible of generating specific phylogeographic signatures (Avise 2000). The description of these signatures return valuable information on the mechanisms underlying spatial patterns of contemporary genetic diversity and structure, and this knowledge can ultimately help in the conservation and management of species and populations under global change (Fordham et al. 2014).

Organisms with limited dispersal capacity and a typically allopatric type of speciation, such as many amphibians, generally present complex genetic signals at lineage borders resulting in reticulate patterns (Denton et al. 2014; Gonçalves et al. 2007; Pereira et al. 2016). A number of mechanisms are involved in the generation of reticulate patterns in phylogeography, among them secondary contact, hybridisation and introgression (Edwards et al. 2016). Such mechanisms, among others, may result in the formation of discordant patterns of variation among genetic markers (i.e. mito-nuclear discordances; e.g. Bisconti et al. 2018; Pereira et al. 2016; Toews and Brelsford 2012). As a consequence, these discordances represent a powerful source of insights into the evolutionary history of species.

Quaternary climatic fluctuations had a major impact on worldwide biotas and organisms, reshaping the distribution of species and modelling their genetic structure (Hewitt 2000, 2004). In Europe, it is acknowledged that the Iberian Peninsula served as one of the most important refugia for temperate species during periods of climatic instability (Gómez and Lunt 2007). Growing evidence has revealed that in this region, characterized by a great variety of climatic and ecological conditions, multiple isolated refugia were present, which promoted genetic diversification and ultimately determined complex phylogeographic patterns in a number of taxa (the "refugia within refugia" scenario; Abellán and Svenning 2014; Gómez and Lunt 2007). In this context, Iberian mountain ranges played a major role in favouring survival throughout the Pleistocene. The Iberian Peninsula has several mountain ranges that permitted flexibility and survival of populations through altitudinal shifts, allowing for movements up or down the mountains in search of suitable microclimates as the temperatures worsened or ameliorated (Gómez and Lunt 2007; Hewitt 1996). With the beginning of the Holocene, temperatures

increased substantially, favouring species expansion outside the refugia and consequent formation of contact zones at lineage borders (Gómez and Lunt 2007).

The common midwife toad (*Alytes obstetricans*) is a small anuran widely distributed in central and western Europe. It is a common species that inhabits a wide variety of habitats between sea level and 2,400 meters above sea level (m.a.s.l.) in the Pyrenees (Bosch et al. 2009; Miró et al. 2018). Major threats to the species are related to habitat loss and fragmentation through pollution, commercial and agricultural development, introduction of non-native species, and more recently emergent diseases such as chytridiomycosis, which is caused by the chytrid fungus *Batrachochytrium dendrobatidis* (*Bd*) (Bosch et al. 2009; Bosch et al. 2001; Miró et al. 2018). *Bd* has led to episodes of mass mortality in *A. obstetricans*, which is known to be highly susceptible to the pathogen (Albert et al. 2014; Bosch et al. 2001; Tobler and Schmidt 2010).

A. obstetricans represents an excellent biological system to study phylogeographic patterns derived from post-glacial expansion, such as contact zones and hybridisation. The species shows strong genetic subdivisions in the Iberian Peninsula indicative of past population isolation (Gonçalves et al. 2015). Until recently, *A. obstetricans* was defined by six divergent and geographically structured mtDNA haplogroups (named A to F; Gonçalves et al. 2015; Maia-Carvalho et al. 2014), delineating as many genetic lineages that probably interbreed in zones of secondary contact. Each of the six mitochondrial lineages corresponded to a unique nuclear microsatellite clade, except for lineage B that harboured two distinct microsatellite clusters (Maia-Carvalho et al. 2018). More recently, the subspecies *almogavarii* (mtDNA lineages E-F) was recognized as an incipient species (i.e. *A. almogavarii*; Dufresnes and Martínez-Solano 2019) by RAD-sequencing analysis.

Previous studies revealed a complex scenario of relationships and admixture between lineages of *A. obstetricans*/*almogavarii*, evidencing the existence of contact and hybrid zones at some lineage borders (Arntzen and Garcia-Paris 1995; Dufresnes and Martínez-Solano 2019; García-París 1995; Gonçalves et al. 2015). Yet, there has been little attempt to identify the causes and consequences of such admixture. Furthermore, none of these studies was specifically focused on high mountains, which are well-known hotspots of genetic diversity and have been shown to play a considerable role in separating well-differentiated intraspecific clades in numerous species (e.g. Pereira et al. 2016; Pöschel et al. 2018). Here, we employ a multilocus approach

combined with comprehensive sample collection across four of the defined mtDNA genetic lineages to analyse the geographical differentiation in *A. obstetricans*/*almogavarii*, placing special focus on three high mountain regions in the Iberian Peninsula: Pyrenees, Picos de Europa and Guadarrama Mountain Range. Specifically, we aimed to (1) describe finer-scale geographical patterns of genetic diversity and structure in high mountain populations of *A. obstetricans*/*almogavarii* and compare different scenarios of population divergence; (2) identify putative contact zones and assess patterns of introgression/hybridisation; and (3) better delineate the geographic distribution of major genetic lineages. We also provide deeper analyses on the genetic status of populations that have been hit by a recent chytridiomycosis outbreak.

Materials and Methods

Sampling and DNA extraction

Sampling was conducted across northern and central Iberian Peninsula, as it is known to host most of the genetic diversity of the species (Gonçalves et al. 2015; Maia-Carvalho et al. 2018). We paid special emphasis on three mountain regions: Pyrenees, Picos de Europa mountains (Cantabrian Mountains, NW Iberian Peninsula) and Guadarrama Mountain Range (Central System, central Iberian Peninsula), although we also took some additional samples from neighbouring lowland localities (Figure 3.1, Tables 3.1 and 3.S1). We sampled 51 sites in the Pyrenees (1,010–2,447 m.a.s.l.), 13 in Picos de Europa (1,120–2,079 m.a.s.l.), nine in Guadarrama Mountains (1,527–1,980 m.a.s.l.), and 32 sites in lowland areas (52–992 m.a.s.l.), totalling 890 individuals from 105 sampling sites (Table 3.1). Lowland populations were grouped according to their geographic proximity to the analysed mountain regions (Table 3.1). A subset of samples was used in a previous study (Albert et al. 2014).

Tissue samples were collected via tail clipping in the case of larvae and toe clipping in the case of adults. Samples were stored in absolute ethanol and maintained at -20 °C. Genomic DNA was extracted using QIAGEN DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol, or following the HotSHOT method (Montero-Pau et al. 2008), in a final volume of 100 µl for toe clips and 250 µl for tail clips.

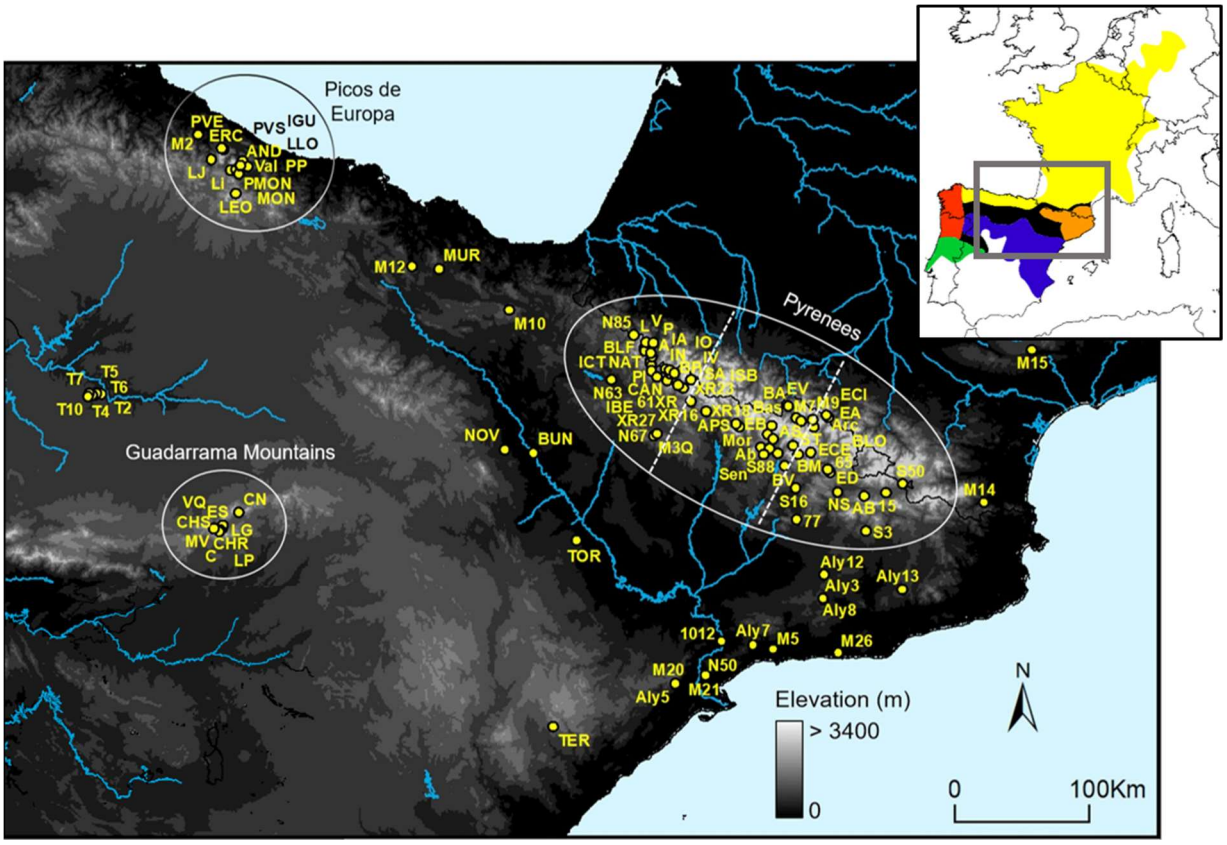


Figure 3.1 Geographic location of sampling sites for *Alytes obstetricans*/*almogavarii*. The three analysed mountain regions are circled in white. In the Pyrenees, dashed lines delimit the three geographic sections (eastern, central and western Pyrenees). For population codes and further information on sampling sites see Table 3.S1. The inset map shows the distribution of the main lineages: orange - ND4 haplogroups E-F (*A. almogavarii*), yellow - ND4 haplogroup B (*A. o. obstetricans*), blue - ND4 haplogroup A (*A. o. pertinax*), red - ND4 haplogroup C (*A. o. boscai*), green - ND4 haplogroup D (*A. o. boscai*), black - unclear (adapted from Dufresnes and Martínez-Solano 2019).

Table 3.1 Genetic diversity parameters as estimated from microsatellites for each analysed mountain region and corresponding lowland areas. Alt., altitudinal range in meters; N_{pops}, number of populations; N, sample size for microsatellites; Na, mean number of alleles; Ar, allelic richness standardized for sample size; PAAr, rarefied private allelic richness standardized for sample size; H_O, observed heterozygosity; H_E, expected heterozygosity; N_e, range of effective population size; ND4 haps, occurrence and code (in parentheses) of mitochondrial ND4 haplogroups identified in each deme.

	Alt.	N_{pops}	N	Na	Ar	PAAr	H_O	H_E	N_e	ND4 haps
Eastern Pyrenees	1010-2406	17	144	13.471	5.310	0.780	0.480	0.701	14-51	28(F)
Eastern Pyrenees – lowland	62-948	9	21	8.588	5.310	0.830	0.460	0.691	-	11(F)
Central Pyrenees	1211-2003	9	52	10.824	5.130	0.920	0.401	0.651	68	15(E), 4(B)
Central Pyrenees – lowland	567	1	2	-	-	-	-	-	-	2(E)
Central-western Pyrenees	1043-2447	25	284	14.118	5.240	0.990	0.377	0.708	12-84	61(B)
Central-western Pyrenees – lowland	992	1	1	-	-	-	-	-	-	1(B)
Picos de Europa	1120-2079	13	202	12.706	5.740	1.080	0.595	0.789	19-160	31(B)
Picos de Europa – lowland	464-928	3	3	-	-	-	-	-	-	3(B)
Guadarrama	1527-1980	9	112	7.529	3.940	0.580	0.442	0.613	5-15	25(A)
Guadarrama – lowland	52-830	18	57	14.529	6.580	1.080	0.620	0.805	70	36(A), 2(B)

Nuclear and mitochondrial genes sequencing

We amplified five gene regions, including four mitochondrial fragments (cytochrome *b* gene – cyt-*b*; 12S rRNA gene – 12S; 16S rRNA gene – 16S; and NADH dehydrogenase subunit 4 gene and adjacent tRNAs – ND4) and one nuclear gene (β-fibrinogen intron 7 – β-fibint7). Primers used for amplification and sequencing were: for cyt-*b*, a shortened version of primers Ptacek1-L (5'-TGAGGACAAATATCATTCTGAGG-3'; Ptacek et al. 1994) and CB3Xen-H (5'-GGCGAATAGGAARTATCATTC-3'; Goebel et al. 1999); for 12S, primers 12Sa (Kocher et al. 1989) and H1557mod (Zaher et al. 2009); for 16S, primers 16sar and 16Sbr (Palumbi et al. 1991). Amplification conditions followed standard PCR amplifications with annealing temperatures ranging between 48-56 ºC. As for ND4, primers ND4 and Leu, described by Arevalo et al. (1994), were initially used for amplification and sequencing. PCR conditions followed Gonçalves et al. (2007). However, we obtained non-resolvable electropherograms and/or very short sequences for a high number of individuals. We thus designed a new set of primers specific for *A. obstetricans*/*almogavarii*, ND4-F (5'-TACCCCTTTATTGGTCTCGC-3') and Leu-F (5'-GGTTCCTAAGACCAACGGAT-3'), which allowed us to recover high quality sequences. PCR conditions were the same as described in Gonçalves et al. (2007), except for the annealing temperature that was set to 57 ºC. Similarly, we designed specific primers for β-fibint7 based on sequences that amplified using β-fibint7 primers BFxF and BFxR (Sequeira et al. 2006). These primers were AlyFibF (5'-GACCTTATGATAATCATCTTTATTGGCT-3') and AlyFibR (5'-TGTTTGTTTGGATCTGCATGTAGCCTGG-3'), with annealing temperatures ranging between 50-56 ºC. Primer combinations were: BFxF-BFxR, BFxF-AlyFibR and AlyFibF-BFxR.

Resulting sequences were aligned using the ClustalW algorithm in MEGA 7 (Kumar et al. 2016).

Microsatellite screening and estimation of genetic diversity

878 individuals from 102 localities were screened for 17 previously characterised microsatellite loci combined in five multiplexes (Maia-Carvalho et al. 2014). Fragments were sized with LIZ-500 size standard and binned using Geneious 11.0.5 (Kearse et al. 2012). Individuals that could not be scored for a minimum of 15 loci were excluded from further analyses.

The presence of potential scoring errors, large allele dropout and null alleles was tested using MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004). We tested for linkage disequilibrium between loci and for departures from Hardy-Weinberg equilibrium (HWE) in each population and for each locus using GENEPOP 4.2 (Rousset 2008). The Bonferroni correction was applied to adjust for multiple comparisons ($\alpha = 0.05$; Rice 1989).

Genetic diversity parameters were calculated for populations with ≥ five genotyped individuals and for each analysed mountain region and corresponding lowland areas. Observed (H_O) and expected heterozygosity (H_E) and mean number of alleles (Na) were obtained with GenAlEx 6.5 (Peakall and Smouse 2012). Allelic richness (Ar) standardized for sample size and rarefied private allelic richness (PAAr – calculated only at the genetic cluster level) were calculated in HP-RARE 1.1 (Kalinowski 2005). We estimated the inbreeding coefficient (F_{IS}) within each population with GENETIX 4.05.2 (Belkhir et al. 1996-2004).

Phylogenetic analyses

For ND4, we estimated genealogic relationships among haplotypes using Haploviewer (Salzburger et al. 2011). The optimal nucleotide-substitution model was determined by jModelTest 2.1.3 (Darriba et al. 2012), under the Akaike Information Criterion (AIC). The phylogeny was estimated with RAxML 7.7.1 (Stamatakis 2006) using a Maximum Likelihood (ML) approach. The program was run with a gamma model of rate heterogeneity and no invariant sites (GTRGAMMA), applying 1,000 bootstrap replicates. The best tree was selected for haplotype network construction in Haploviewer, based on all sequences retrieved from GenBank and this study. Furthermore, genetic diversity parameters, namely number of haplotypes (H) and polymorphic sites (S), as well as haplotype (Hd) and nucleotide (Π) diversity indices, were estimated for the whole dataset and for each analysed mountain region and corresponding lowland areas using DnaSP 6.11.01 (Rozas et al. 2017).

In order to describe the affinities of the different genetic lineages described in *A. obstetricans*/*almogavarii*, we constructed a species tree with the five partially sequenced genes using the multispecies coalescent approach implemented in *BEAST (Heled and Drummond 2009). The substitution model for each marker was determined by jModelTest. We set a strict molecular clock model and a Yule speciation prior. The clock and tree models for mtDNA markers were linked. We defined five groups in *A. obstetricans*/*almogavarii*, corresponding to

the four main population lineages identified in the haplotype network (A, B, E, F) and further subdividing lineage B into two groups (central-western Pyrenees and Picos de Europa; see Results). Samples of *A. maurus* were used as outgroup. Chains were run for 200 million generations, sampling trees every 5,000. Tracer 1.6 (Drummond and Rambaut 2007) was used to check for stationarity and convergence of MCMC chains. A maximum clade credibility tree was constructed in TreeAnnotator 2.6.2 and visualized using FigTree 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree), discarding the first 10% generations as burn-in.

Genetic structure

For microsatellites, population structure was inferred with a Discriminant Analysis of Principal Components (DAPC), using the ADEGENET package 2.1.1 (Jombart 2008; Jombart et al. 2010) in R 3.5.1 (R Core Team 2018). The optimal number of clusters was assessed using the *find.clusters* function and determined as the K value above which BIC (Bayesian Information Criterion) values decreased substantially. We covered a range of possible clusters from 1 to 100 and retained all principal components (PCs), as suggested by Jombart and Collins (2015). We also used DAPC to investigate the hierarchical genetic structure of our data through the examination of various values of K and the sequence of differentiation when K increases. The same package was also used to perform a Principal Component Analysis (PCA), for comparison purposes. For DAPC analysis, 150 PCs (80% of variance) and all discriminant functions were retained. The number of PCs retained for DAPC was estimated using cross-validation through the *xvalDapc* command. Furthermore, the Bayesian cluster analysis implemented in STRUCTURE 2.3.4 (Pritchard et al. 2000) was performed to corroborate the optimal clustering solution inferred by DAPC and PCA analyses. All runs were repeated 10 times for each K, set between 1 and 15, with 100K burn-in steps followed by 100K MCMC repetitions. We used the admixture model with correlated allele frequencies. The optimal number of genetic clusters was determined using both the original method of Pritchard et al. (2000) and the ΔK method of Evanno et al. (2005), as implemented in STRUCTURE HARVESTER (Earl and vonHoldt 2012). The R package pophelper (Francis 2017) was used to average replicate runs of the optimal K (Jakobsson and Rosenberg 2007) and plot the final output. Genetic relationships between STRUCTURE clusters were visualised by constructing a neighbour-joining (NJ) tree based on net nucleotide distances (Pritchard et al. 2010) using the program NEIGHBOR in the PHYLIP package 3.695 (Felsenstein 2005). In addition, to visualise genetic divergence between sampled sites and check for consistency with ND4-inferred lineages, we drew a NJ tree using

the program POPTREEW (Takezaki et al. 2014). We used Nei's genetic distance (D_A, Nei et al. 1983) and performed 1,000 bootstraps. As POPTREEW does not allow loci to have no data for an entire locality, we excluded 22 localities with data missing for at least one locus.

To further explore the genetic differentiation between the defined genetic units, we calculated pairwise F_{ST} for both ND4 and microsatellites. Computations were performed in Arlequin 3.5.2.2 (Excoffier and Lischer 2010) in the case of ND4 and in GenAlEx in the case of microsatellites, with 1,000 permutations to assess statistical significance. Similarly, to partition genetic variability at different hierarchical levels (among genetic clades, among populations within clades and within populations), we conducted an analysis of molecular variance (AMOVA) as implemented in Arlequin, using 10,000 permutations to assess significance of variance components (Excoffier et al. 1992).

Effective population size

The effective population size (N_e) of populations with ≥ 15 genotyped individuals was calculated in Colony 2.0.6.5 (Jones and Wang 2010). The program was run with the parameters specified in Lucati et al. (2020).

Demographic history

The approximate Bayesian computation (ABC) approach, as implemented in the software DIYABC 2.1.0 (Cornuet et al. 2014), was used to reconstruct the history of divergence among the genetic clades identified in *A. obstetricans*/*almogavarii*. Only high mountain populations were included in the analysis, as we were interested in understanding the evolutionary history of the three high mountain areas described above. We grouped populations into five groups according to their geographic location. We also incorporated the information from previous studies dealing with the phylogenetics and population structure of the different subspecies, which point to a shared origin between populations from the central and eastern Pyrenees (mitochondrial ND4 haplogroups E and F in Gonçalves et al. 2015; and microsatellite clusters F1n and F2n in Maia-Carvalho et al. 2018): populations from the Picos de Europa mountains (PEU) and the central-western Pyrenees (CWPY; corresponding to subspecies *A. o. obstetricans*) formed two separate groups, whereas populations from the central Pyrenees (CPY), eastern Pyrenees (EPY; corresponding to the putative incipient species *A. almogavarii*) and Guadarrama mountains (GUA; corresponding to subspecies *A. o. pertinax*) were subdivided

into three groups. To reduce computational demands, we selected 50 individuals from each of the five population groups, maximizing the number of samples with available ND4 sequences and minimizing missing data for microsatellites. We generated five different scenarios (Figure 3.2a): (1) null model with all groups diverging simultaneously from a common ancestor, except for lineage CPY that originates from EPY; (2) sequential splitting model directly following results from STRUCTURE and DAPC analyses, where on the one hand populations from the Pyrenees, and on the other GUA and PEU populations share a common origin; (3) similar to the second one, but predicts a common origin of PEU and CWPY populations, as suggested by ND4 haplotype network; (4) same as the previous scenario, with a common origin between populations belonging to ND4 haplogroups A and B, as suggested by *BEAST analysis; and (5) where the PEU clade was created by admixture of CWPY and GUA populations. We performed the computations both combining microsatellites and ND4 data, and separately for microsatellites. The mutation rate prior distribution assumed for ND4 included a range of values comprehensive of the mutation rate calculated in a previous phylogenetic study (0.85×10^{-8} substitutions/site/year; Gonçalves et al. 2015). We also performed a preliminary analysis run using a fixed prior for mutation rate that appeared unsuccessful, as the estimation of prior distribution of parameters showed a lack of correspondence between simulated and observed data sets (data not shown). The final parameter setting is shown in Table 3.S2. We generated 10^6 simulated datasets per scenario, assuming a 1:1 female to male sex ratio and a generation time of 1 to 2 years (Tobler et al. 2013). The following summary statistics were used for microsatellites: mean number of alleles, mean genetic diversity and mean allele size variance as one sample statistics, and Fst and $(d\mu)^2$ distance as pairwise statistics. As for ND4, the following summary statistics were used: number of segregating sites, mean of pairwise differences, variance of pairwise differences, Tajima's D and private segregating sites as one sample statistics, and number of haplotypes and Fst as pairwise statistics. Pre-evaluation of scenarios, selection of the most supported scenario, confidence in scenario choice (type I and II errors), model checking, estimation of the posterior distribution of parameters and evaluation of bias and precision on parameters estimation for the most supported scenario followed Lucati et al. (2020).

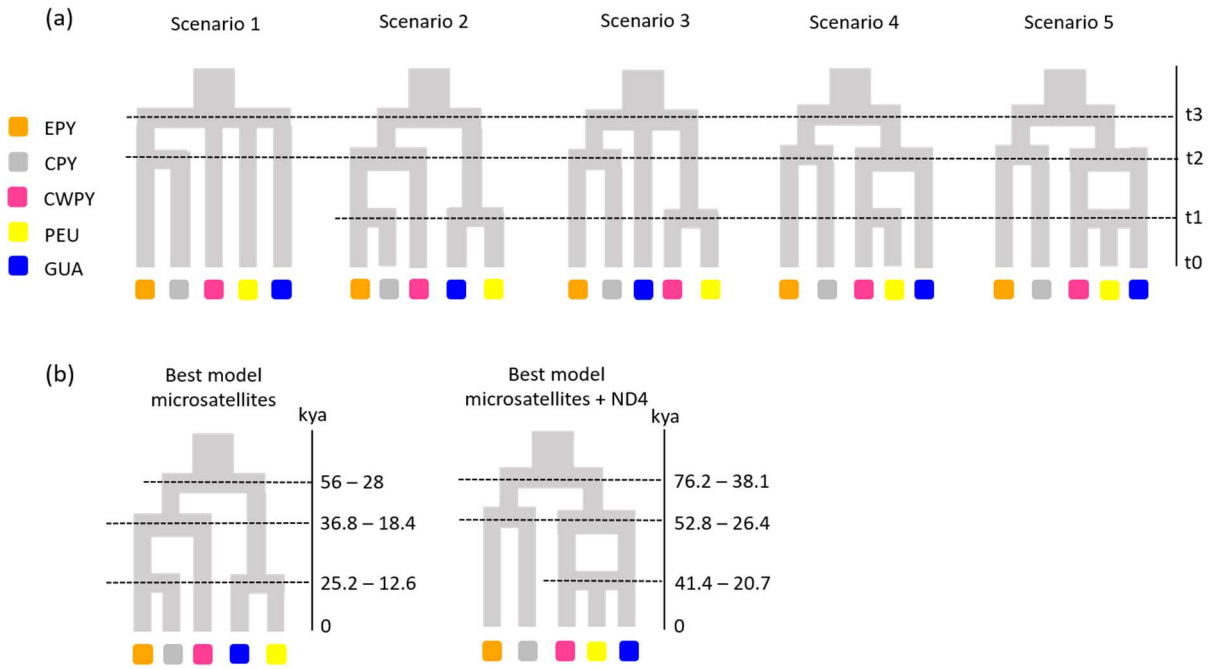


Figure 3.2 Phylogeographic scenarios tested in DIYABC considering the three analysed high mountain areas (a). The best supported scenarios, namely number 2 when considering only microsatellites and number 5 when considering both mtDNA (ND4) and microsatellite markers, with the estimated divergence times (t1 – t3) of each split are shown in panel (b). Population groups were defined on the basis of their geographic distribution and the results from clustering analyses: eastern Pyrenees (EPY, orange), central Pyrenees (CPY, grey), central-western Pyrenees (CWPY, pink), Picos de Europa mountains (PEU, yellow), and Guadarrama Mountain Range (GUA, blue).

Results

Sequence variation and genetic diversity

The nuclear DNA alignment (β-fibint7) included 20 sequences of 605 bp, while for the mitochondrial dataset we obtained 219 sequences of 654 bp for ND4, 40 sequences of 325 bp for cyt-*b*, 42 sequences of 341 bp for 12S, and 40 sequences of 579 bp for 16S (Table 3.S1). With regard to ND4, overall haplotype (Hd) and nucleotide (Π) diversities were 0.930 ± 0.008 and 0.020 ± 0.0008, respectively. Across the analysed mountain regions, we detected comparable levels of ND4 genetic diversity (eastern Pyrenees: Hd = 0.770 ± 0.039, Π = 0.002 ± 0.0003; central Pyrenees: Hd = 0.724 ± 0.101, Π = 0.014 ± 0.003; central-western Pyrenees: Hd = 0.805 ± 0.027, Π = 0.002 ± 0.0002; Picos de Europa: Hd = 0.602 ± 0.091, Π = 0.003 ± 0.001), except for Guadarrama Mountains that showed the lowest values (Hd = 0.290 ± 0.109, Π = 0.001 ± 0.0002).

For microsatellite loci, we did not find evidence of large allele dropout, stuttering or null allele artefacts. Similarly, no significant linkage disequilibrium or departures from HWE across

populations and loci were detected after applying the Bonferroni correction. Observed (H_O) and expected heterozygosity (H_E) ranged from 0.197 to 0.794 and from 0.222 to 0.720, respectively (mean H_O = 0.473, mean H_E = 0.530; Table 3.S1). Mean number of alleles (Na) and allelic richness (Ar) varied from 2.059 to 7.824 and from 1.270 to 5.620, respectively (mean Na = 4.440, mean Ar = 1.873), and the inbreeding coefficient (F_{IS}) ranged from -0.110 to 0.403 (mean = 0.157). Among the analysed mountain regions, Picos de Europa and the eastern Pyrenees were the richest regions in terms of genetic diversity, whereas Guadarrama Mountains was the poorest, with the central and central-western Pyrenees presenting intermediate values (Table 3.1). With regard to lowland areas, populations located in the Duero and Ebro basins (herein referred to as Guadarrama lowland localities) were the most diverse, while eastern Pyrenean localities were the least diverse.

Phylogenetic analyses

From the 219 individuals analysed for the ND4 gene we identified 34 haplotypes, of which 22 are newly described (Figure 3.3, Table 3.S1). Haplotypes were defined by 61 polymorphic sites, of which 49 were parsimony informative. The majority of newly described haplotypes were found in the Pyrenees (13), whereas only three and one haplotypes were detected in Picos de Europa and Guadarrama Mountains, respectively; the remaining five haplotypes were found in lowland localities. The haplotype network showed that haplotypes clustered into four well-differentiated haplogroups with a strong association with geography (Figures 3.3 and 3.4a): haplogroup A included sequences from Guadarrama and corresponding lowland areas (corresponding to populations of *A. o. pertinax*), haplogroup F included sequences from the eastern Pyrenees (corresponding to populations of *A. almogavarii*), haplogroup E corresponded to populations from the central Pyrenees, and haplogroup B (corresponding to populations of *A. o. obstetricans*) included sequences from the central-western Pyrenees and Picos de Europa. Within haplogroup B, sequences corresponding to central-western Pyrenean populations were clearly separated from those from Picos de Europa. In addition, in four localities we detected the presence of more than one haplogroup (Figures 3.3 and 3.4a, Table 3.S1): haplogroups A and B were found to co-occur in population Fte. Nueva de Bardales (T2), whereas both haplogroups B and E were found in populations Plano de Igüer (PI), Balsa Pertacua (BP) and Ibón de los Asnos (61XR).

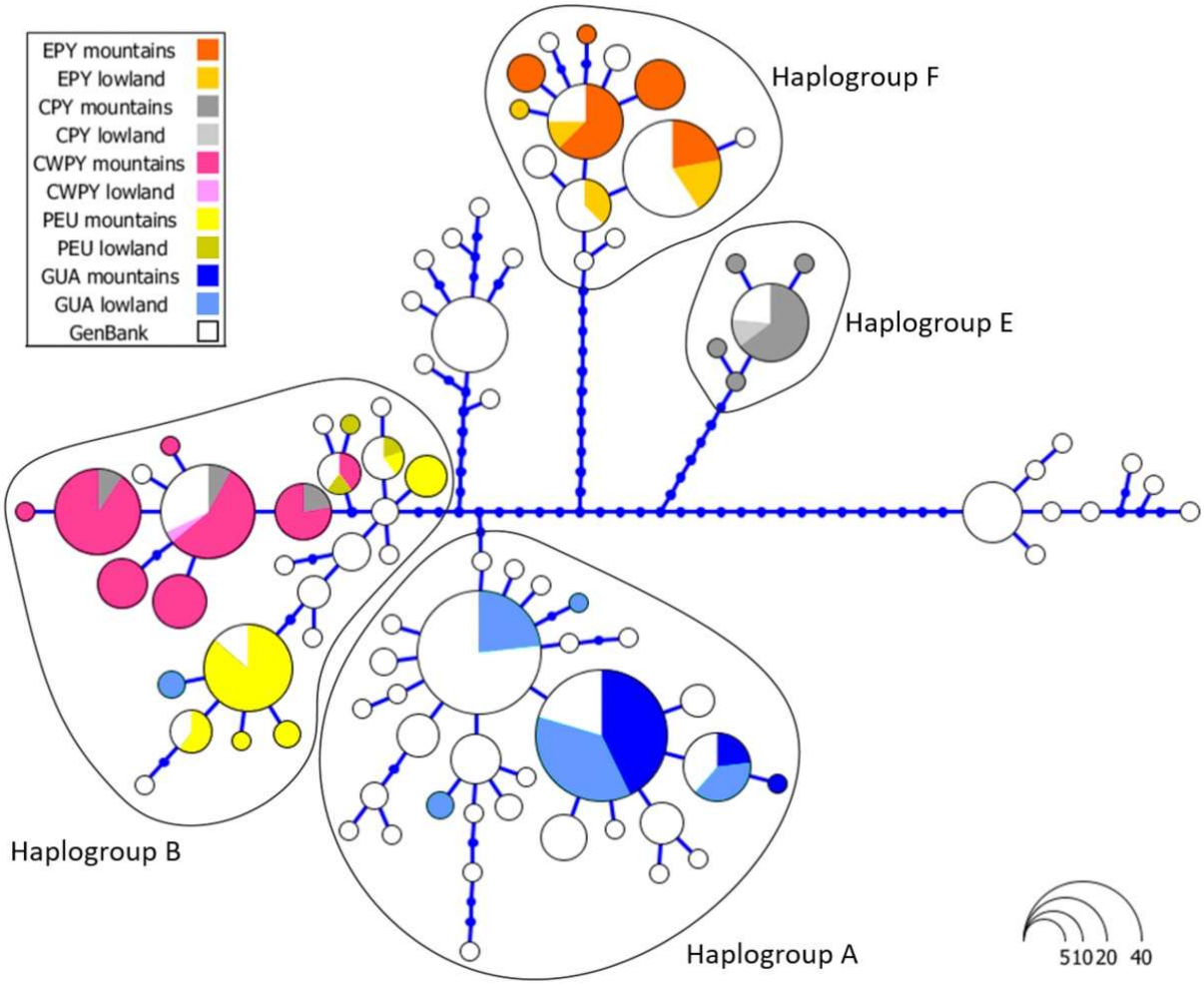


Figure 3.3 Haplotype network of ND4 sequences analysed in *Alytes obstetricans*/*almogavarii*. Each circle represents a unique haplotype and the circle area is proportional to the number of sequences of a given haplotype. Blue dots correspond to inferred unsampled haplotypes. Sequences depicted in white were retrieved from GenBank. The analysed high mountain regions and corresponding lowland areas are indicated by different colours: eastern Pyrenees (EPY, orange), central Pyrenees (CPY, grey), central-western Pyrenees (CWPY, pink), Picos de Europa mountains (PEU, yellow), and Guadarrama Mountain Range (GUA, blue).

In order to describe the affinities between the different haplogroups, we built a consensus tree with the five sequenced gene fragments using a multispecies coalescent approach (species tree; *BEAST). The tree placed lineage E as sister to the rest of the clades, with lineage A as sister to lineage B. Finally, lineage B split into central-western Pyrenees and Picos de Europa groups (Figure 3.5).

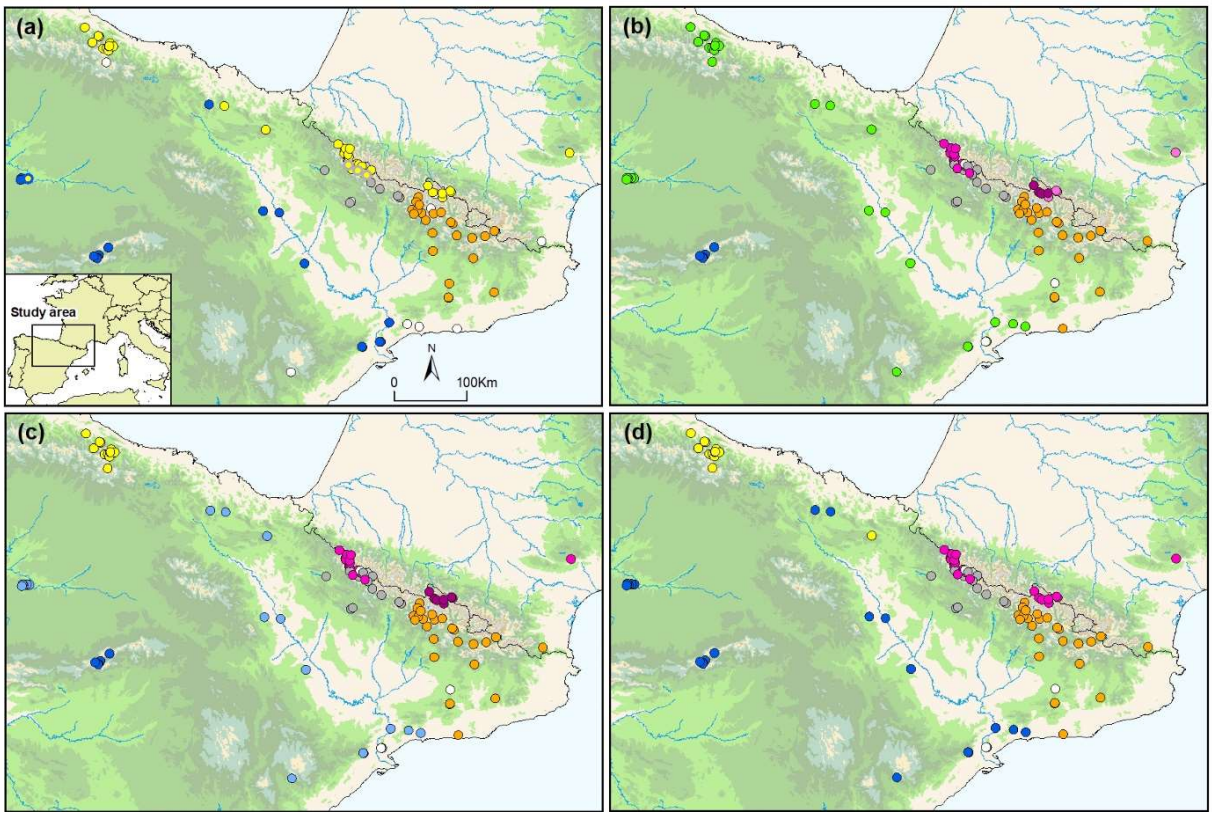


Figure 3.4 Results of phylogenetic and clustering analyses for *Alytes obstetricans*/*almogavarii*. Panel (a) shows the geographic distribution of the four mtDNA (ND4) haplogroups recovered in the analysis (blue: mtDNA haplogroup A, yellow: mtDNA haplogroup B, grey: mtDNA haplogroup E, orange: mtDNA haplogroup F). In four populations (T2, 61XR, BP and PI) we detected the presence of more than one haplogroup. Panels (b) and (c) represent, respectively, the results from Discriminant Analysis of Principal Components (DAPC) and STRUCTURE for K = 7 microsatellite groups (see Figure 3.6). Panel (d) shows the geographic distribution of the five genetic groups identified by microsatellite-based neighbour-joining analysis (see Figure 3.7). White circles indicate populations with no data available for either ND4 or microsatellites. For population codes see Figure 3.1.

Genetic structure and effective population sizes

All the clustering analyses performed on microsatellites recovered congruent results (Figures 3.4 and 3.6). Regarding DAPC analysis, the optimum number of clusters was inferred to be 7 (Figures 3.4b, 3.6c and 3.S1b). The inferred clusters were consistent with the geographic location of populations. A closer look from K = 2 to K = 7 revealed a spatial hierarchical genetic structure (Figure 3.S1a), with a first split between the central-western Pyrenees and all other populations, and a subsequent subdivision of this second group into central Pyrenees–eastern Pyrenees and Guadarrama and corresponding lowland populations–Picos de Europa. At K = 4, populations from the Guadarrama Mountain Range split from lowland populations located in the Duero and Ebro basins and Picos de Europa Mountains, whereas at K = 5 there was a split between the central and eastern Pyrenees. The following splits occurred within the central-western Pyrenean group. According with DAPC analysis, the PCA grouped populations

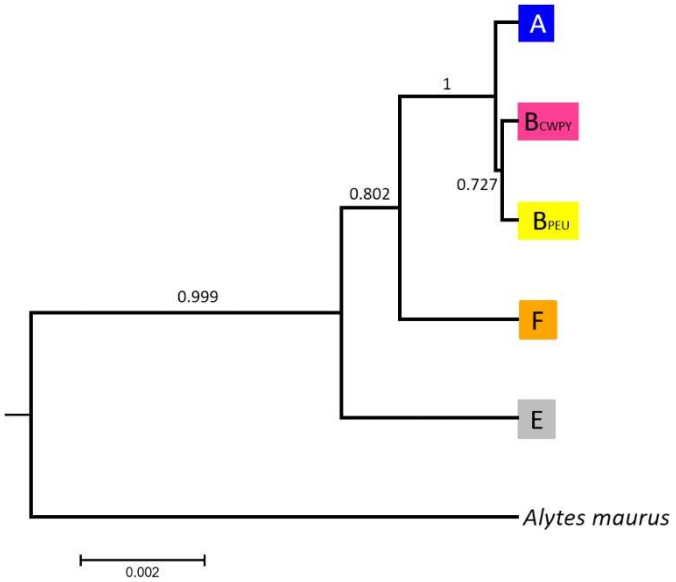


Figure 3.5 *Alytes* species tree produced in *BEAST based on one nuclear gene (β-fibint7) and four mitochondrial fragments (ND4, cyt-*b*, 12S and 16S). Labels on branch tips correspond to the distinct ND4 haplogroups identified (blue: mtDNA haplogroup A, pink: mtDNA haplogroup B (central-western Pyrenean populations), yellow: mtDNA haplogroup B (Picos de Europa populations), grey: mtDNA haplogroup E, orange: mtDNA haplogroup F). Posterior probabilities of lineage divergence are indicated on branch labels.

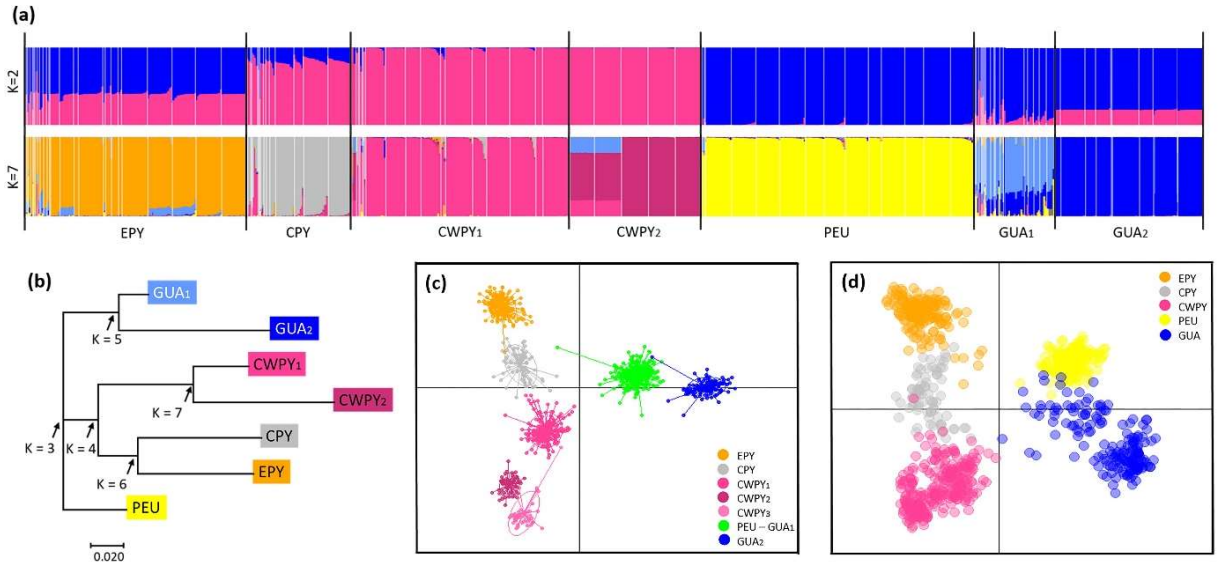


Figure 3.6 Results of clustering analyses of *Alytes obstetricans*/*almogavarii* populations based on microsatellite data. Panel (a) shows STRUCTURE barplots of membership assignment for K = 2 and K = 7. Each individual is represented by a vertical bar corresponding to the sum of assignment probabilities to the K cluster. White lines separate populations and black lines separate clusters. Panel (b) represents a neighbour-joining tree based on net nucleotide distances among the seven clusters inferred by STRUCTURE. Arrows indicate the sequence of differentiation when K increases. Panel (c) shows a summary plot from Discriminant Analysis of Principal Components (DAPC) for K = 7 genetic clusters. Dots represent individuals and genetic clusters are shown as inertia ellipses. Panel (d) shows the plot obtained in the Principal Component Analysis (PCA). Each dot represents one individual. Labels indicate the different genetic clusters: eastern Pyrenees (EPY, orange), central Pyrenees (CPY, grey), central-western Pyrenees (CWPY, pink), Picos de Europa mountains (PEU, yellow), and Guadarrama Mountain Range (GUA, blue).

according to their geographic distribution rather than by described subspecies (Figure 3.6d). PC1 separated the Pyrenees from the Picos de Europa and Guadarrama, whereas PC2 distributed Pyrenean populations along the Pyrenean chain and contributed to the separation of the Picos de Europa and Guadarrama. STRUCTURE analysis showed patterns of genetic differentiation similar to those inferred by PCA, and identified K = 2 as the best clustering solution (Figures 3.6a and 3.S2): the first cluster grouped together all Pyrenean populations and the second cluster included peninsular ones. A smaller peak was detected at K = 7 (Figure 3.4c), in agreement with DAPC-inferred clusters, but segregating populations located in the Duero and Ebro basins from Guadarrama Mountains, while splitting central-western Pyrenean populations into two groups. In order to describe the segregation history amongst the seven groups identified by STRUCTURE, we built a NJ tree based on net nucleotide distances: the tree started at K = 3 with the separation of Guadarrama and corresponding lowland populations from Picos de Europa and the Pyrenees, then at K = 4 populations from the central and eastern Pyrenees split from central-western Pyrenean populations, whereas the following splits occurred within the different major groups (Figure 3.6b). We also built a second NJ tree inferred from microsatellite-based D_A distances over all populations that identified five discrete clades, which match with ND4-inferred lineages, except that populations included in haplogroup B formed two separate groups (Figures 3.4d and 3.7). Nevertheless, the tree suggested that the pairs of clades central Pyrenees–eastern Pyrenees and Guadarrama–Picos de Europa were more closely related with each other than with the clade of central-western Pyrenees, which in turn appeared more distant from the rest of the clades.

In all cases, pairwise F_{ST} values were significantly different from zero ($P < 0.01$; Table 3.S3), indicating significant genetic differences between the seven clusters defined by STRUCTURE. ND4-based pairwise F_{ST} values ranged from 0.951 for eastern Pyrenees–Guadarrama Mountains to 0.089 for Guadarrama Mountains–Guadarrama lowlands, whereas microsatellite-based pairwise F_{ST} values ranged from 0.222 for Guadarrama Mountains–central-western Pyrenees to 0.060 for Guadarrama lowlands–Picos de Europa. Accordingly, AMOVA analyses suggested significant structure among the seven genetic clades ($P < 0.001$; Table 3.S4). The proportion of variation attributable to differences among clades was lower in microsatellites (26.171%) than in ND4 (84.945%). Conversely, variation among individuals within populations was low for ND4 (8.736%) and high for microsatellite (56.425%) markers, as expected for polymorphic loci such as microsatellites.

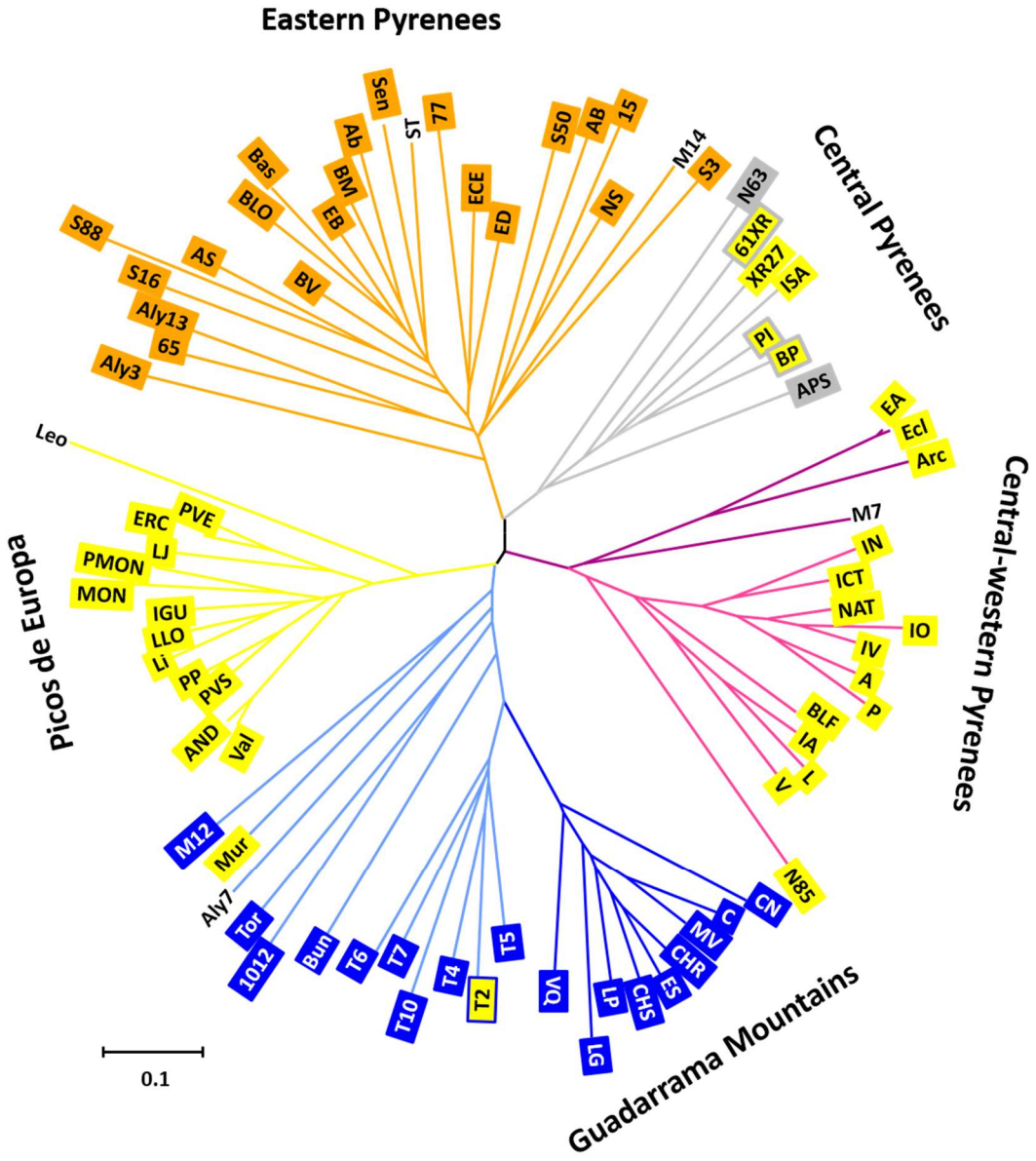


Figure 3.7 Neighbour-joining tree based on D_A distances for microsatellite markers. Branch colours delineate the seven genetic clusters inferred by STRUCTURE (see Figure 3.6a), while colour shades around population codes correspond to the distinct mtDNA (ND4) haplogroups (blue: mtDNA haplogroup A, yellow: mtDNA haplogroup B, grey: mtDNA haplogroup E, orange: mtDNA haplogroup F; see Figure 3.3). In four populations (T2, 61XR, BP and PI) we detected the presence of more than one haplogroup. For population codes see Table 3.S1.

Colony returned, in general, low values of effective population sizes (Table 3.S1). Values ranged from five in the population Laguna de Pájaros (LP) to 160 breeding individuals in the population Lago Ercina (ERC), with a mean N_e of 43. The lowest values were found in populations from the Guadarrama Mountains (N_e range = 5–15), while higher values were found in the Pyrenees (eastern Pyrenees: 14–51, central Pyrenees: 68, central-western Pyrenees: 12–84) and Picos de Europa (19–160).

Demographic history

DIYABC analyses suggested highest support for two different but compatible scenarios of population divergence depending on the genetic markers used (Figure 3.2b, Table 3.S5). Both scenarios had non-overlapping 95% confidence intervals and type I and II errors for the best supported models were low (Table 3.S5), indicating high confidence in scenario choice. Model checking confirmed that the best supported scenarios provided a good fit to the observed data (data not shown). Furthermore, the mean mutation rate estimated for ND4 was found to be an order of magnitude higher than the literature value (1.96×10^{-7} substitutions/site/year; Table 3.S6), but still much lower than that estimated for microsatellites (ranging from 1.28×10^{-4} to 3.89×10^{-4} and from 4.59×10^{-5} to 3.31×10^{-4} for analyses conducted using only microsatellites and including both ND4 and microsatellite markers, respectively).

The analysis focused on microsatellites indicated that scenario 2 (the sequential splitting model based on results from clustering analyses) was the best supported model (Figure 3.2b, Table 3.S5). According to this scenario, the first split led to the separation of Pyrenean populations from peninsular ones. During the second split there was the divergence of, on the one hand, central-western and central-eastern Pyrenean populations, and on the other Picos de Europa and Guadarrama populations. Finally, populations from the central Pyrenees originated from eastern Pyrenean populations during the third split. The first split occurred 56,000–28,000 years ago, and the second and third splits 36,800–18,400 and 25,200–12,600 years ago, respectively (Table 3.S6). The analysis based on the combination of microsatellites and ND4, however, suggested that Picos de Europa populations were generated by admixture of populations from the central-western Pyrenees and Guadarrama (scenario 5; Figure 3.2b, Table 3.S5). Results indicated that the first split occurred 76,200–38,100 years ago, the second split 52,800–26,400 years ago and the admixture event 41,400–20,700 years ago (Table 3.S6).

Discussion

Overall, our multilocus analysis of genetic divergence in high mountain populations of *A. obstetricans*/*almogavarii* from the Pyrenees, Picos de Europa and Guadarrama mountains revealed a partial concordance between the two main marker types employed (mtDNA –ND4– and microsatellites) and a strong association of the defined genetic lineages with geography. Analyses of mtDNA showed that the four ND4 haplogroups detected (namely A, B, E and F) were geographically structured and spatially segregated in specific mountain regions, except

for lineage B that was found to be highly diversified and was recovered from two different high mountain areas (central-western Pyrenees and Picos de Europa; Figure 3.4a). Our study also describes the presence of seven different nuclear microsatellite clusters (Figure 3.6). Most of the genetic differentiation was found in Pyrenean populations, which in turn were genetically more related to each other than with the other two high mountain areas. Based on ND4 sequence data, Picos de Europa populations showed a close affinity to the subspecies *A. o. obstetricans*, whereas microsatellites revealed a close relationship between Picos de Europa and *A. o. pertinax*, suggesting a possible admixed origin for Picos de Europa. Finally, Guadarrama mountain populations were genetically distinct from the corresponding lowland *A. o. pertinax* populations.

Differences in microsatellites and mitochondrial DNA evolutionary times

Recently, Gonçalves et al. (2015) presented a comprehensive species tree of all *A. obstetricans*/*almogavarii* extant clades based on five gene regions, where ND4 lineage F was recovered as sister taxon to the rest of the clades, from which it was estimated to diverge around 2.5 Mya. Our results confirm the tree topology recovered by Gonçalves et al. (2015), with the only difference that lineage E was estimated to be the basal taxon (Figure 3.5). The simplest explanation for this discordance might be the use of different sets of markers in the two studies, given that in our analyses we included four mitochondrial and one nuclear gene fragments, while one mitochondrial and four nuclear genes were used by Gonçalves et al. (2015). Furthermore, our species tree goes a step further by incorporating information on the genetic distinctiveness between central-western Pyrenean and Picos de Europa populations, both belonging to lineage B (subspecies *A. o. obstetricans*). The split between these two population groups was estimated to be the most recent divergence event.

Using the mutation rate estimated by Gonçalves et al. (2015) for ND4 (0.85×10^{-8} substitutions/site/year) and assuming constant evolutionary rates, this last split would be estimated to have occurred approximately 1.25 Mya, whereas when applying the mutation rate estimated by DIYABC (Table 3.S6) we obtain a divergence time of ~50,000 years ago (see below). Indeed, results from ABC modelling based on either microsatellites or microsatellites + mtDNA resulted in time estimates extremely different from those of the species tree (e.g. ranging from ca. 13,000 to 76,000 years ago; Figure 3.2b), suggesting that the two analyses are depicting different evolutionary events. In fact, while the species tree has proven useful to

estimate divergence times associated with the formation of species/subspecies, ABC modelling was here employed to gain insights into the colonisation of three high mountain regions in the Iberian Peninsula by the different *A. obstetricans*/*almogavarii* taxa. Microsatellite markers, with their fast mutation rate, are known to perform poorly for the estimation of ancient historical events (Cornuet et al. 2010). On the contrary, microsatellites provide substantially better estimations than mtDNA for the most recent dynamics. However, the combination of both marker types in our ABC modelling resulted in similar divergence times than those estimated using microsatellites alone. Despite introducing the mutation rate obtained by Gonçalves et al. (2015) for ND4, DIYABC did not stabilise until using a wider mutation rate prior distribution (see Results; Table 3.S6). Consequently, the mean mutation rate estimated for ND4 by ABC analyses was higher than the literature value. It is thus possible that combining the two marker types together may have caused a bias toward microsatellites and the estimation of recent historical events. The divergence times estimated with DIYABC may thus mirror the properties of microsatellites, which are better suited for disentangling recent evolutionary dynamics (Cornuet et al. 2010). An alternative explanation for the observed discordance is that ND4 mutation rate has not been constant along the evolutionary history of the species, and that some bottleneck events might have accelerated the segregation of populations (Lynch 2010). According to this scenario, ND4 mutation rate estimated by DIYABC might reflect the last evolutionary events, and therefore the separation of Picos de Europa populations from the central-western Pyrenees depicted by the species tree might be matching the changes described by microsatellites.

Are Pyrenean populations shaped by glacial isolation?

In accordance with previous studies (Gonçalves et al. 2015; Maia-Carvalho et al. 2018), we recovered three different lineages within the Pyrenees (Figure 3.4): *A. o. obstetricans* (ND4 lineage B, central-western Pyrenees), *A. almogavarii* (ND4 lineage F, eastern Pyrenees) and ND4 lineage E (central Pyrenees), which is currently considered as part of *A. almogavarii* (based on microsatellites; Maia-Carvalho et al., 2018), although it has not yet been described morphologically. Furthermore, our results show that the current distribution of lineage E is more extended than previously considered. A large majority of the populations bearing this haplogroup is found in the Spanish central Pyrenees (Aragón region), between the Aragón River (province of Zaragoza) in the West and the Gistaín Valley in the East. The southern distribution limit falls in the Aragonese Pre-Pyrenees, at the foothills of the Sierra de Gratal.

Clustering analyses based on microsatellites indicated that high mountain Pyrenean populations were more genetically related to each other than with any of the other analysed high mountain regions (i.e. Guadarrama Mountains and Picos de Europa; Figure 3.6), which is in apparent contradiction with the species tree based on functional genes, where the two most recent clades (ND4 clades A and B, corresponding to subspecies *A. o. obstetricans* and *A. o. pertinax*, respectively) show the highest affinity (Figure 3.5). Following the splits between major population lineages in *A. obstetricans*/*almogavarii*, which date back to the Early Pleistocene (starting from 2.5 Mya; Gonçalves et al. 2015), there have been several glacial-interglacial episodes that likely provided opportunities for diverging taxa to come into contact and interbreed following range shifts tracking the climatic fluctuations (Abellán and Svenning 2014; Wallis et al. 2016). In this regard, putative episodes of admixture within *A. obstetricans*/*almogavarii* in the Pyrenees were suggested by previous phylogenetic analyses (Gonçalves et al. 2007). Here, microsatellite-based clustering analyses, together with ABC modelling, support the hypothesis that Pyrenean high mountain populations have gone through relatively recent events of admixture or hybridisation, likely favoured by the different glacial-interglacial episodes that characterised the Late Pleistocene. Specifically, according to ABC analyses, the two clades ascribed to *A. almogavarii* (E-F, central and eastern Pyrenees) seem to have undergone a certain degree of contact and admixture until divergence took place ~52,000–12,000 years ago, when the two clades likely remained confined in two distinct glacial refugia (Figure 3.2b). The older divergence event involving *A. o. obstetricans* (central-western Pyrenean group) suggests that this clade remained isolated from *A. almogavarii* at a slightly earlier time (~52,000–18,000 years ago; Figure 3.2b). Contemporary signs of connectivity between the two species were detected at lineage borders, as indicated by the occurrence of a contact zone between ND4 lineage E and *A. o. obstetricans* in the West (Figures 3.3 and 3.7). Similarly, signs of introgression and extensive gene flow between *A. o. obstetricans* and *A. almogavarii* in the western Pyrenees were previously pointed out by allozyme markers (Arntzen and Garcia-Paris 1995; García-París 1995). Furthermore, within the central-western Pyrenean group, clustering analyses revealed a finer genetic structure at regional scales that partitioned the microsatellite lineage into smaller genetic units. More into detail, this group was further subdivided by STRUCTURE and DAPC analyses into two or three different genetic clusters, respectively (Figures 3.4 and 3.6). This is consistent with a scenario of allopatric divergence in different glacial refugia separated by major geographic barriers, with consequent reduction or

disruption of gene flow (e.g. Charrier et al. 2014). From a conservation point of view, we suggest that these areas should be treated as separate conservation and management units.

Colonization of Picos de Europa and admixture between subspecies

All population structure analyses performed on the microsatellite dataset provided evidence that the two nuclear counterparts of ND4 lineage B (i.e. central-western Pyrenees and Picos de Europa groups), corresponding to the subspecies *A. o. obstetricans*, are extremely divergent (Figures 3.4 and 3.6). Microsatellite data also suggested that Picos de Europa clade was more genetically related to populations from the Guadarrama Mountains and lowland areas in the Duero and Ebro basins, ascribed to the subspecies *A. o. pertinax*, than to the central-western Pyrenean clade. Indeed, in DAPC analysis the pair of clades Picos de Europa–Guadarrama (mainly lowland populations located in the Duero and Ebro basins) was recovered as a single genetic unit, and in PCA and NJ analyses the pair appeared more closely related than to any of the other clades. In support of this, ABC analyses solely based on microsatellites pointed to a shared origin of populations from Picos de Europa and Guadarrama, and nuclear-based pairwise F_{ST} values between the cluster pair were the lowest (Figure 3.2b, Table 3.S3). In contrast, the mtDNA (ND4) of Picos de Europa populations was derived from *A. o. obstetricans* (Figures 3.3 and 3.5). Therefore, as suggested by ABC analyses based on both microsatellites and ND4, the extant Picos de Europa populations may have originated from an extensive admixture event involving *A. o. obstetricans* (the central-western Pyrenean group) as the maternal donor and *A. o. pertinax* as the paternal donor (Figure 3.2b). A similar pattern was observed in the fire salamander *Salamandra salamandra* (Donaire-Barroso and Rivera 2016) and in crustaceans of the genus *Daphnia* (Thielsch et al. 2017). A plausible scenario could be that, during Late Pleistocene glacial periods, some *A. o. obstetricans* populations remained confined in isolated refugia where they coincided with *A. o. pertinax*. Subdivided glacial refugia could have experienced events of admixture during the succession of glacial and interglacial periods, with consequent fusion between refugial lineages (Canestrelli et al. 2014; Dufresnes et al. 2020). Mito-nuclear discordances with evidence for admixture or hybridisation are not uncommon in amphibians (e.g. Bisconti et al. 2018; Pereira et al. 2016). Furthermore, the Cantabrian Mountains represent a peculiar biogeographic region and a recognised hotspot of genetic diversity in amphibians, being home to endemic refugial clades in a number of species with broad European distribution (Dufresnes et al. 2020; Recuero and Garcia-Paris 2011). Finally, more recently, Picos de Europa and *A. o. pertinax* populations may have come into secondary

contact and interbred following an expansion phase, creating zones of admixture at lineage borders. Indeed, both ND4 and microsatellites identified a contact zone between *A. o. obstetricans* and *A. o. pertinax* in the South (Duero basin) and in the East (Cantabrian Mountains; Figures 3.4 and 3.7).

Genetic distinctiveness of Guadarrama mountain populations

Within the range occupied by *A. o. pertinax*, DAPC and STRUCTURE analyses revealed a pattern of high genetic differentiation between populations from the Guadarrama Mountain Range and lowland localities from the Duero and Ebro basins (Figures 3.4 and 3.6). Furthermore, both nuclear and mtDNA-based pairwise F_{ST} values indicated significant genetic differences between the two groups (Table 3.S3). It should be noted that our sampling was performed approximately 10 years after an epidemic of the disease chytridiomycosis, which hugely impacted this area causing several *A. obstetricans* populations to decrease or even disappear (Bosch et al. 2018; Bosch et al. 2001). The genetic distinctiveness and limited genetic diversity of populations from the Guadarrama Mountains herein detected (Table 3.1) are likely early signs of inbreeding depression induced by an emerging pathogen, i.e. the chytrid fungus *Batrachochytrium dendrobatidis*. Accordingly, the only sampled population in the affected area that did not show signs of chytridiomycosis was the one to present the highest values of genetic diversity (Montes de Valsain, MV; Table 3.S1). Our findings complement those of Albert et al. (2014) that, using a different set of 10 microsatellite loci, detected evidence of low genetic variability and population bottleneck in *A. obstetricans* from the Guadarrama Mountains. In addition to this, we provide the first estimates of effective population sizes of Guadarrama populations, which were among the lowest overall (N_e = 5–15; Table 3.1), raising concern for the long-term persistence of these populations, which are small and isolated. Alarming signs of reduced genetic diversity were also pointed out by our mtDNA analyses, as indicated by the low values of ND4 genetic diversity of Guadarrama Mountains compared to the other lineages. In light of the above, our results support previous findings on the catastrophic consequences of chytridiomycosis outbreaks and provide new insights into the understanding of emergent diseases.

Concluding remarks

Our multilocus phylogeography across the Iberian Peninsula revealed high genetic structure correlated with geography and a complex pattern of lineage admixture in high mountain

populations of *A. obstetricans*/*almogavarii*. Our study evidenced how each analysed mountain region underwent a peculiar phylogeographic history through the Late Pleistocene, which is consistent with the "refugia within refugia" model (Abellán and Svenning 2014; Gómez and Lunt 2007) and confirms previous studies on a number of Iberian amphibian species (e.g. Maia-Carvalho et al. 2018; Recuero and Garcia-Paris 2011; Vences et al. 2013; Vences et al. 2017). Results also support the assumption that refugia within refugia may be hotspots of extensive mito-nuclear discordances (Dufresnes et al. 2020), highlighting the importance of multilocus approaches to capture true dynamics of population divergence.

Acknowledgements

We thank Saioa Fernández-Beaskoetxea, Amparo Mora, Susana Marquínez, Fernando Rivada, René A. Priego, Abel Bermejo and Matthew C. Fisher for field assistance, and Eva Albert, Laura Méndez, Annie Machordom and Andrés Fernández-Loras for laboratory assistance. This work was supported by the European Commission LIFE+ project LimnoPirineus (LIFE13 NAT/ES/001210), the MINECO grant (CGL2015-70070-R; PI: Jaime Bosch) and Picos de Europa National Park grants for 2015 and 2016 (PI: Jaime Bosch). F.L. had a doctoral grant funded by Fundação para a Ciência e a Tecnologia (FCT, grant number PD/BD/52598/2014). M.J.J. had an International collaboration grant (National Institute of Ecology).

References

Abellán P, Svenning J-C (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.

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.

Arevalo 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, Garcia-Paris 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, MA, USA.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Bisconti R, Porretta D, Arduino P, Nascetti G, Canestrelli D (2018) Hybridization and extensive mitochondrial introgression among fire salamanders in peninsular Italy. *Scientific Reports* **8:** 13187.

Bosch J, Beebee T, Schmidt B, Tejedo M, Martinez Solano I, Salvador A et al. (2009) *Alytes obstetricans* (errata version published in 2016). The IUCN Red List of Threatened Species 2009: e.T55268A87541047. doi:10.2305/IUCN.UK.2009.RLTS.T55268A11283700.en.

Bosch J, Fernández-Beaskoetxea S, Garner TW, Carrascal LM (2018) Long-term monitoring of an amphibian community after a climate change- and infectious disease-driven species extirpation. *Global Change Biology* **24:** 2622-2632.

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.

Canestrelli D, Bisconti R, Sacco F, Nascetti G (2014) What triggers the rising of an intraspecific biodiversity hotspot? Hints from the agile frog. *Scientific Reports* **4:** 5042.

Charrier O, Dupont P, Pornon A, Escaravage N (2014) Microsatellite marker analysis reveals the complex phylogeographic history of *Rhododendron ferrugineum* (Ericaceae) in the Pyrenees. *PLoS One* **9:** e92976.

Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R et al. (2014) DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30:** 1187-1189.

Cornuet JM, Ravigne V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* **11:** 401.

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

Denton RD, Kenyon LJ, Greenwald KR, Gibbs HL (2014) Evolutionary basis of mitonuclear discordance between sister species of mole salamanders (*Ambystoma* sp.). *Molecular Ecology* **23:** 2811-2824.

Donaire-Barroso D, Rivera X (2016) La salamandra común *Salamandra salamandra* (Linnaeus, 1758) en el subcantábrico: Origen, dispersión, subespecies y zonas de introgresión. *Butlletí Societat Catalana de Herpetología* **23:** 7-38.

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

Dufresnes C, Martínez-Solano Í (2019) Hybrid zone genomics supports candidate species in Iberian *Alytes obstetricans*. *Amphibia-Reptilia* **41:** 105-112.

Dufresnes C, Nicieza AG, Litvinchuk SN, Rodrigues N, Jeffries DL, Vences M et al. (2020) Are glacial refugia hotspots of speciation and cyto-nuclear discordances? Answers from the genomic phylogeography of Spanish common frogs. *Molecular Ecology* **29:** 986-1000.

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

Edwards SV, Potter S, Schmitt CJ, Bragg JG, Moritz C (2016) Reticulation, divergence, and the phylogeography–phylogenetics continuum. *Proceedings of the National Academy of Sciences* **113:** 8025-8032.

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, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10:** 564-567.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131:** 479-491.

Felsenstein J (2005) PHYLIP (phylogeny inference package) version 3.6. Distributed by the author. Seattle (WA): Department of Genome Sciences, University of Washington.

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.

Francis RM (2017) pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* **17:** 27-32.

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

Goebel A, Donnelly J, Atz M (1999) PCR primers and amplification methods for 12S ribosomal DNA, the control region, cytochrome oxidase I, and cytochrome *b* in bufonids and other frogs, and an overview of PCR primers which have amplified DNA in amphibians successfully. *Molecular Phylogenetics and Evolution* **11:** 163-199.

Gómez A, Lunt DH (2007) Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S and Ferrand N (eds) *Phylogeography of southern European refugia*. Springer: Amsterdam, Netherlands, pp 155-188.

Gonçalves H, Maia-Carvalho B, Sousa-Neves T, Garcia-Paris M, Sequeira F, Ferrand N et al. (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.

Heled J, Drummond AJ (2009) 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 (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **359:** 183-195; discussion 195.

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.

Jombart T, Collins C (2015) A tutorial for discriminant analysis of principal components (DAPC) using *adegenet* 2.0.0. Imperial College London, MRC Centre for Outbreak Analysis and Modelling.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics* **11:** 94.

Jones OR, Wang J (2010) COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10:** 551-555.

Kalinowski ST (2005) HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5:** 187-189.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S et al. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28:** 1647-1649.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX et al. (1989) Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences* **86:** 6196-6200.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33:** 1870-1874.

Lucati F, Poignet M, Miró A, Trochet A, Aubret F, Barthe L et al. (2020) Multiple glacial refugia and contemporary dispersal shape the genetic structure of an endemic amphibian from the Pyrenees. *Molecular Ecology* **29:** 2904-2921.

Lynch M (2010) Evolution of the mutation rate. *Trends in Genetics* **26:** 345-352.

Maia-Carvalho B, Gonçalves H, Martínez-Solano I, Gutiérrez-Rodríguez J, Lopes S, Ferrand N et al. (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.

Maia-Carvalho B, Vale CG, Sequeira F, Ferrand N, Martínez-Solano I, Gonçalves H (2018) The roles of allopatric fragmentation and niche divergence in intraspecific lineage diversification in the common midwife toad (*Alytes obstetricans*). *Journal of Biogeography* **45:** 2146-2158.

Miró A, Sabás I, Ventura M (2018) Large negative effect of non-native trout and minnows on Pyrenean lake amphibians. *Biological Conservation* **218:** 144-153.

Montero-Pau J, Gómez A, Muñoz J (2008) Application of an inexpensive and high-throughput genomic DNA extraction method for the molecular ecology of zooplanktonic diapausing eggs. *Limnology and Oceanography: Methods* **6:** 218-222.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19:** 153-170.

Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowski G (1991) *Simple fool's guide to PCR*. University of Hawaii Press: Honolulu, Hawaii.

Peakall R, Smouse P (2012) GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28:** 2537-2539.

Pereira RJ, Martinez-Solano I, Buckley D (2016) Hybridization during altitudinal range shifts: Nuclear introgression leads to extensive cyto-nuclear discordance in the fire salamander. *Molecular Ecology* **25:** 1551-1565.

Pöschel J, Heltai B, Graciá E, Quintana MF, Velo-Antón G, Arribas O et al. (2018) Complex hybridization patterns in European pond turtles (*Emys orbicularis*) in the Pyrenean Region. *Scientific Reports* **8:** 1-13.

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

Pritchard JK, Wen X, Falush D (2010) Documentation for STRUCTURE software, version 2.3. Department of Human Genetics University of Chicago, Department of Statistics University of Oxford.

Ptacek MB, Gerhardt HC, Sage RD (1994) Speciation by polyploidy in treefrogs: Multiple origins of the tetraploid, *Hyla versicolor*. *Evolution* **48:** 898-908.

R Core Team (2018) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Recuero E, Garcia-Paris 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.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution* **43:** 223-225.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8:** 103-106.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE et al. (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34:** 3299-3302.

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

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

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

Takezaki N, Nei M, Tamura K (2014) POPTREEW: web version of POPTREE for constructing population trees from allele frequency data and computing some other quantities. *Molecular Biology and Evolution* **31:** 1622-1624.

Thielsch A, Knell A, Mohammadyari A, Petrusek A, Schwenk K (2017) Divergent clades or cryptic species? Mito-nuclear discordance in a *Daphnia* species complex. *BMC Evolutionary Biology* **17:** 227.

Tobler U, Garner TW, Schmidt BR (2013) Genetic attributes of midwife toad (*Alytes obstetricans*) populations do not correlate with degree of species decline. *Ecology and Evolution* **3:** 2806-2819.

Tobler U, Schmidt BR (2010) Within- and among-population variation in chytridiomycosis-induced mortality in the toad *Alytes obstetricans*. *PLoS One* **5:** e10927.

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21:** 3907-3930.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4:** 535-538.

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

Vences M, Sarasola-Puente V, Sanchez E, Amat F, Hauswaldt JS (2017) Diversity and distribution of deep mitochondrial lineages of the common frog, *Rana temporaria*, in northern Spain. *Salamandra* **53:** 25-33.

Wallis GP, Waters JM, Upton P, Craw D (2016) Transverse alpine speciation driven by glaciation. *Trends in Ecology & Evolution* **31:** 916-926.

Zaher H, Grazziotin FG, Cadle JE, Murphy RW, Moura-Leite JCd, Bonatto SL (2009) Molecular phylogeny of advanced snakes (Serpentes, Caenophidia) with an emphasis on South American Xenodontines: A revised classification and descriptions of new taxa. *Papéis Avulsos de Zoologia* **49:** 115-153.

Supplementary material

Supplementary tables

Table 3.S1 Geographic information and standard genetic statistics of *Alytes obstetricans*/*almogavarii* sampling localities. Populations are grouped according to the genetic group of interest (EPY: eastern Pyrenees, CPY: central Pyrenees, CWPY: central-western Pyrenees, PEU: Picos de Europa mountains, GUA: Guadarrama Mountain Range). Lat., latitude; Long., longitude; Alt., altitude in meters; N, sample size for microsatellites; Na, mean number of alleles; Ar, allelic richness standardized for sample size; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, inbreeding coefficient; N_e, effective population size; N ND4, sample size for ND4; ND4 haps, occurrence and code (in parentheses) of mitochondrial ND4 haplogroups identified in each population (see Figure 3.3); N cyt-*b*, sample size for cyt-*b*; N 12S, sample size for 12S; N 16S, sample size for 16S; N β-fibint7, sample size for β-fibint7.

Genetic group	Population	Code	Lat.	Long.	Alt.	N	Na	Ar	H_O	H_E	Fis	N_e	N ND4	ND4 haps	N cyt-*b*	N 12S	N 16S	N β-fibint7
$EPY_{mountains}$	Abeurador Abella	Ab	42.452	0.564	1360	3	-	-	-	-	-	-	1	1(F)	-	-	-	-
	Abeurador Boscalt	AB	42.314	1.587	1450	1	-	-	-	-	-	-	2	2(F)	-	-	-	-
	Abeurador Sarroqueta	AS	42.444	0.724	1106	3	-	-	-	-	-	-	2	2(F)	-	-	-	-
	Abeurador Señuy	Sen	42.460	0.641	1227	5	2.412	1.440	0.388	0.392	0.126	-	2	2(F)	-	-	-	-
	Barranc de Viu	BV	42.373	0.812	1125	19	6.353	1.630	0.471	0.606	0.279	14	2	2(F)	-	-	-	-
	Barranco Llana de Obarra	BLO	42.528	0.653	1421	5	3.647	1.620	0.535	0.547	0.158	-	2	2(F)	-	-	-	-
	Bassa Manyanet	BM	42.470	0.908	2277	19	4.765	1.590	0.53	0.571	0.104	33	1	1(F)	-	-	-	-
	Bassa Vallibierna	Bas	42.613	0.610	1912	19	3.294	1.480	0.432	0.468	0.104	26	3	3(F)	-	-	-	-
	Bellver de Cerdanya	15	42.372	1.777	1087	1	-	-	-	-	-	-	1	1(F)	-	-	-	-
	Enveitg	S50	42.459	1.906	1537	3	-	-	-	-	-	-	3	3(F)	-	-	-	-
	Estany Basibé	EB	42.548	0.596	2254	19	5.412	2.690	0.494	0.595	0.198	53	2	2(F)	-	-	-	-
	Estanyet Coma d'Espós	ECE	42.509	1.017	2406	18	5.235	1.60	0.508	0.573	0.149	51	1	1(F)	-	-	-	-
	Estanyet de Davall	ED	42.415	1.232	2058	18	6.059	2.740	0.494	0.647	0.265	51	3	3(F)	-	-	-	-
	Les Paüls	S88	42.408	0.606	1242	2	-	-	-	-	-	-	1	1(F)	-	-	-	-
	Navès	S3	42.088	1.677	1010	1	-	-	-	-	-	-	1	1(F)	-	-	-	-
	Safareig de Taüll	ST	42.522	0.846	1515	7	2.706	2.550	0.366	0.425	0.230	-	-	-	-	-	-	-
	Serra de l'Orri	65	42.428	1.205	1592	1	-	-	-	-	-	-	1	1(F)	2	2	2	1
$EPY_{lowland}$	Castellterçol	Aly13	41.763	2.118	678	1	-	-	-	-	-	-	1	1(F)	1	1	1	-
	Els Banys i Palaldà	M14	42.473	2.670	305	1	-	-	-	-	-	-	-	-	1	1	1	1

	Ferran	Aly12	41.730	1.401	672	-	-	-	-	-	-	-	1	1(F)	1	1	1	-
	Gavet de la Conca	77	42.045	1.037	948	2	-	-	-	-	-	-	2	2(F)	1	1	1	-
	La Goda	Aly8	41.563	1.446	691	1	-	-	-	-	-	-	2	2(F)	4	4	4	-
	La Pobla de Segur	S16	42.250	0.963	812	1	-	-	-	-	-	-	1	1(F)	-	-	-	-
	Noves de Segre	NS	42.295	1.342	713	11	4.294	2.680	0.447	0.564	0.258	-	2	2(F)	-	-	-	-
	Pobla de Carivenys	Aly3	41.570	1.441	691	3	-	-	-	-	-	-	2	2(F)	1	1	1	1
	Vilanova i la Geltrú	M26	41.239	1.683	62	1	-	-	-	-	-	-	-	-	3	3	3	1
$CPY_{mountains}$	Aljibe Pino Simón	APS	42.561	0.291	1426	4	-	-	-	-	-	-	2	2(E)	-	-	-	-
	Arguis	N67	42.326	0.413	1211	2	-	-	-	-	-	-	4	4(E)	2	2	2	2
	Balsa Pertacua	BP	42.714	0.423	1913	7	3.588	1.540	0.479	0.504	0.132	-	2	1(B), 1(E)	-	-	-	-
	Bassa de la Mora	Mor	42.545	0.326	1903	14	4.059	5.350	0.272	0.407	0.403	-	3	3(E)	-	-	-	-
	Buisán	XR18	42.585	0.010	1482	1	-	-	-	-	-	-	1	1(E)	-	-	-	-
	Ibón de los Asnos	61XR	42.693	0.267	2003	4	-	-	-	-	-	-	4	2(B), 2(E)	-	-	-	-
	Ibón Serrato Bajo	ISB	42.765	0.215	2447	3	-	-	-	-	-	-	2	2(B)	-	-	-	-
	Linás de Broto	XR16	42.623	0.168	1504	1	-	-	-	-	-	-	1	1(E)	-	-	-	-
	Monrepos	M3Q	42.347	0.391	1214	1	-	-	-	-	-	-	-	-	1	1	1	1
	Plano de Igüer	PI	42.745	0.588	1585	18	6.588	1.650	0.447	0.629	0.318	68	2	1(B), 1(E)	-	-	-	-
$CPY_{lowland}$	Asó-Veral	N63	42.610	0.925	567	2	-	-	-	-	-	-	2	2(E)	1	1	1	1
$CWPY_{mountains}$	Arties	M7	42.699	0.862	1548	3	-	-	-	-	-	-	-	-	1	2	-	-
	Barranco Las Foyas	BLF	42.862	0.696	1372	4	-	-	-	-	-	-	2	2(B)	-	-	-	-
	Bassa d'Arres	BA	42.769	0.715	1562	19	2.765	5.620	0.338	0.336	0.023	21	4	4(B)	-	-	-	-
	Canfranc	CAN	42.717	0.523	1501	2	-	-	-	-	-	-	1	1(B)	2	2	2	1
	Estanhet d'Arcoïls	Arc	42.680	0.988	2392	20	2.235	1.270	0.271	0.262	0.036	20	3	3(B)	-	-	-	-
	Estanho Vilac	EV	42.709	0.814	1638	20	2.059	5.510	0.197	0.222	0.140	21	2	2(B)	-	-	-	-
	Estany d'Aulà	EA	42.769	1.099	2128	20	2.471	1.360	0.292	0.352	0.197	17	3	3(B)	-	-	-	-
	Estany de Clavera	ECl	42.778	1.077	2230	20	2.353	1.340	0.308	0.335	0.106	25	3	3(B)	-	-	-	-
	Formigal	XR27	42.775	0.376	1823	3	-	-	-	-	-	-	3	3(B)	-	-	-	-
	Ibón Acherito	IA	42.880	0.707	1872	21	5.588	1.560	0.493	0.548	0.127	51	5	5(B)	2	1	1	1
	Ibón Campo de Troya Inferior	ICT	42.766	0.409	2071	10	4.294	1.530	0.400	0.499	0.249	-	3	3(B)	-	-	-	-

	Ibón Espelunciecha	IBE	42.787	0.431	1951	-	-	-	-	-	-	-	1	1(B)	-	-	-	-
	Ibón Negras	IN	42.787	0.465	2079	18	3.941	1.450	0.378	0.433	0.155	44	4	4(B)	-	-	-	-
	Ibón Orná	IO	42.798	0.614	1851	5	2.941	1.470	0.379	0.422	0.213	-	3	3(B)	-	-	-	-
	Ibón Serrato Alto	ISA	42.765	0.213	2459	17	3.529	1.460	0.332	0.443	0.286	27	4	4(B)	-	-	-	-
	Ibón Viejo	IV	42.781	0.602	2111	20	5.588	1.560	0.400	0.550	0.296	22	2	2(B)	-	-	-	-
	Isaba	N85	42.946	0.835	1227	2	-	-	-	-	-	-	2	2(B)	1	1	1	1
	Lac d'Arlet	A	42.839	0.615	1998	15	5.176	1.550	0.379	0.525	0.316	84	2	2(B)	-	-	-	-
	Lac de Lhurs	L	42.922	0.704	1698	15	4.176	1.530	0.422	0.509	0.209	52	2	2(B)	-	-	-	-
	Lescun	V	42.934	0.637	903	15	3.353	1.480	0.345	0.459	0.286	12	3	3(B)	-	-	-	-
	Naval Aguas Tuertas	NAT	42.812	0.621	1623	19	6.118	1.630	0.501	0.613	0.213	49	3	3(B)	-	-	-	-
	Pla de Beret	M9	42.725	0.964	2070	1	-	-	-	-	-	-	1	1(B)	1	1	1	-
	Puit d'Arious	P	42.864	0.633	1874	11	4.588	1.570	0.437	0.53	0.249	-	2	2(B)	-	-	-	-
	Tramacastilla de Tena	XR23	42.705	0.320	1267	1	-	-	-	-	-	-	1	1(B)	-	-	-	-
$CWPY_{lowland}$	Lac du Saut de Vésoles	M15	43.555	2.794	992	1	-	-	-	-	-	-	1	1(B)	1	1	1	1
$PEU_{mountains}$	Ándara	AND	43.213	4.716	1346	14	4.941	1.670	0.519	0.648	0.235	-	1	1(B)	-	-	-	-
	Ftes. Carrionas	Leo	43.011	4.744	2079	2	-	-	-	-	-	-	-	-	-	-	-	-
	Lago de Valdominguero	Val	43.207	4.727	1834	19	5.647	1.650	0.592	0.635	0.097	26	3	3(B)	-	-	-	-
	Lago Ercina	ERC	43.267	4.979	1120	16	7.824	1.720	0.652	0.697	0.098	160	2	2(B)	-	-	-	-
	Liordes	Li	43.149	4.857	1885	14	4.588	1.620	0.562	0.598	0.100	-	3	3(B)	-	-	-	-
	Llagos de Jesús	LJ	43.175	5.056	1199	19	5.824	1.670	0.602	0.654	0.108	98	2	2(B)	-	-	-	-
	Pilón de Igüedri	IGU	43.146	4.774	1386	10	4.706	1.730	0.587	0.687	0.200	-	3	3(B)	-	-	-	-
	Pilón de Moñetas	PMON	43.202	4.783	1530	20	4.647	1.670	0.602	0.654	0.106	19	3	3(B)	-	-	-	-
	Pilón de Pandébano	PP	43.235	4.781	1249	18	6.471	1.740	0.602	0.720	0.195	87	3	3(B)	-	-	-	-
	Pilón Vegas de Sotres	PVS	43.208	4.767	1530	18	6.471	1.730	0.635	0.711	0.136	20	3	3(B)	-	-	-	-
	Pilón Vegas del Enol	PVE	43.269	4.998	1120	18	7.824	1.710	0.682	0.688	0.038	77	3	3(B)	-	-	-	-
	Pozo de Moñetas	MON	43.197	4.786	1530	17	4.353	1.440	0.419	0.430	0.058	32	3	3(B)	-	-	-	-
	Pozos de Lloroza	LLO	43.165	4.811	1482	17	5.824	1.710	0.653	0.692	0.087	68	2	2(B)	-	-	-	-
$PEU_{lowland}$	Murua	Mur	42.976	2.736	638	1	-	-	-	-	-	-	1	1(B)	-	-	-	-
	Puerto de Lizarraga	M10	42.860	2.005	928	1	-	-	-	-	-	-	1	1(B)	1	1	1	-
	Valle del Tendi	M2	43.302	5.249	464	1	-	-	-	-	-	-	1	1(B)	1	1	1	1

$GUA_{mountains}$	Charcas de la Rubia	CHR	40.846	3.949	1898	18	3.471	1.490	0.475	0.474	0.029	11	3	3(A)	-	-	-	-
	Charcas del Salto	ES	40.846	3.948	1898	11	3.176	1.480	0.463	0.451	0.030	-	3	3(A)	-	-	-	-
	Charcas Secas	CHS	40.848	3.948	1898	8	2.412	1.380	0.385	0.357	0.007	-	3	3(A)	-	-	-	-
	Circo del Nevero	CN	40.979	3.844	1980	20	2.941	1.440	0.377	0.425	0.152	15	3	3(A)	-	-	-	-
	Laguna de Pájaros	LP	40.860	3.948	1898	20	3.471	1.490	0.430	0.476	0.124	5	3	3(A)	-	-	-	-
	Laguna Grande	LG	40.840	3.957	1898	1	-	-	-	-	-	-	1	1(A)	-	-	-	-
	Montes de Valsaín	MV	40.824	4.017	1527	11	3.765	1.580	0.514	0.558	0.125	-	3	3(A)	-	-	-	-
	Puerto de Cotos	C	40.823	3.962	1829	5	3.000	1.560	0.485	0.501	0.144	-	3	3(A)	-	-	-	-
	Valdesqui	VQ	40.815	3.961	1829	18	4.588	1.530	0.431	0.510	0.188	20	3	3(A)	-	-	-	-
$GUA_{lowland}$	Barranc de la Galera	N50	40.746	0.332	723	1	-	-	-	-	-	-	-	-	1	1	1	-
	Bassa de Saranou	M21	40.864	0.581	243	1	-	-	-	-	-	-	1	1(A)	1	1	1	-
	Bassa Gran	M20	40.860	0.571	243	-	-	-	-	-	-	-	1	1(A)	1	1	1	-
	Buñuel	Bun	41.982	1.438	240	3	-	-	-	-	-	-	5	5(A)	3	3	3	3
	Fte. de la Marlota	T10	41.377	5.452	751	2	-	-	-	-	-	-	4	4(A)	-	-	-	-
	Fte. de los Perros	T6	41.395	5.423	693	5	4.000	1.690	0.659	0.615	0.056	-	4	4(A)	-	-	-	-
	Fte. de Picarico	T7	41.400	5.451	751	6	4.471	1.720	0.794	0.651	0.110	-	4	4(A)	-	-	-	-
	Fte. Los Billares	T5	41.409	5.393	772	15	6.824	1.750	0.646	0.719	0.138	70	3	3(A)	-	-	-	-
	Fte. Nueva de Bardales	T2	41.424	5.344	706	5	4.471	1.660	0.562	0.590	0.160	-	4	2(A), 2(B)	-	-	-	-
	Fte. Valdespino	T4	41.416	5.387	751	4	-	-	-	-	-	-	4	4(A)	-	-	-	-
	L'Argentera	Aly7	41.144	0.921	341	2	-	-	-	-	-	-	-	-	-	1	1	-
	Mas de Barberans	Aly5	40.748	0.325	723	1	-	-	-	-	-	-	2	2(A)	2	2	2	-
	Mora la Nova	1012	41.109	0.642	52	3	-	-	-	-	-	-	1	1(A)	-	1	1	-
	Nacimiento Nervión	M12	42.939	2.982	830	1	-	-	-	-	-	-	1	1(A)	2	2	2	2
	Nogueruelas	TER	40.235	0.633	1203	1	-	-	-	-	-	-	-	-	-	-	-	-
	Novallas	NOV	41.947	1.950	435	3	-	-	-	-	-	-	2	2(A)	-	-	-	-
	Reus	M5	41.150	1.106	108	1	-	-	-	-	-	-	-	-	-	-	-	-
	Torrecilla de Valmadrid	Tor	41.501	0.854	396	3	-	-	-	-	-	-	2	2(A)	2	2	2	2
	Total					878							219		40	42	40	20

Table 3.S2 Parameters used in DIYABC analysis and respective priors for the best supported scenarios, namely scenario 2 when considering only microsatellites (simple sequence repeats – SSRs) and scenario 5 when including both mtDNA (ND4) and microsatellite markers. See Figure 3.2 for more information on tested scenarios. N, effective population size for each analysed deme (EPY – eastern Pyrenees; CPY – central Pyrenees; CWPY – central-western Pyrenees; PEU – Picos de Europa Mountains; GUA – Guadarrama Mountains); ra, admixture rate; t, time of events in generations (t_1 – time to the most recent split; t_2 – time to the intermediate split; t_3 – time to the most ancient split). Microsatellite (SSRs) and mitochondrial (ND4) parameters: mean μ, mean mutation rate; individual locus μ, individual locus mutation rate; mean *P*, mean coefficient *P*; individual locus *P*, individual locus coefficient *P*; SNI, Single Nucleotide Insertion rate; mean *k*, mean coefficient *k*; individual locus *k*, individual locus coefficient *k*. Microsatellite loci were divided in two groups depending on the motif length (tri- and tetranucleotide loci). Conditions: sequence data were simulated under a Tamura Nei (TN93) mutation model.

	Microsatellites		Microsatellites + ND4	
Parameter	Conditions	Distribution [min-max]	Conditions	Distribution [min-max]
N_{EPY}		Uniform [10 - 20 000]		Uniform [10 - 20 000]
N_{CPY}		Uniform [10 - 20 000]		Uniform [10 - 20 000]
N_{CWPY}		Uniform [10 - 20 000]		Uniform [10 - 20 000]
N_{PEU}		Uniform [10 - 20 000]		Uniform [10 - 20 000]
N_{GUA}		Uniform [10 - 20 000]		Uniform [10 - 20 000]
$N_{EPY\text{-}GUA}$		Uniform [10 - 40 000]		-
$N_{EPY\text{-}CWPY\text{-}GUA}$		-		Uniform [10 - 40 000]
ra		-		0.001 – 0.999
t_1		Uniform [10 - 30 000]		Uniform [10 - 30 000]
t_2	$t_2>t_1$	Uniform [10 - 40 000]	$t_2>t_1$	Uniform [10 - 40 000]
t_3	$t_3>t_2$	Uniform [10 - 50 000]	$t_3>t_2$	Uniform [10 - 50 000]
Mean $\mu 1_{(SSRs)}$		Uniform [10^{-5} - 10^{-3}]		Uniform [10^{-5} - 10^{-3}]
Individual locus $\mu 1_{(SSRs)}$		Gamma [10^{-5} - 10^{-2}]		Gamma [10^{-5} - 10^{-2}]
Mean $P1_{(SSRs)}$		Uniform [10^{-1} - $3x10^{-1}$]		Uniform [10^{-1} - $3x10^{-1}$]
Individual locus $P1_{(SSRs)}$		Gamma [10^{-2} - $9x10^{-1}$]		Gamma [10^{-2} - $9x10^{-1}$]
$SNI_{(SSRs)}$		Log-u [0]		Log-u [0]
Mean $\mu 2_{(SSRs)}$		Uniform [10^{-5} - 10^{-3}]		Uniform [10^{-5} - 10^{-3}]
Individual locus $\mu 2_{(SSRs)}$		Gamma [10^{-5} - 10^{-2}]		Gamma [10^{-5} - 10^{-2}]
Mean $P2_{(SSRs)}$		Uniform [10^{-1} - $3x10^{-1}$]		Uniform [10^{-1} - $3x10^{-1}$]
Individual locus $P2_{(SSRs)}$		Gamma [10^{-2} - $9x10^{-1}$]		Gamma [10^{-2} - $9x10^{-1}$]
$SNI_{(SSRs)}$		Log-u [0]		Log-u [0]
Mean $\mu_{(ND4)}$	-	-	TN93	Uniform [10^{-10} - 10^{-6}]
Individual locus $\mu_{(ND4)}$	-	-	TN93	Gamma [10^{-10} - 10^{-6}]
Mean $k1_{(ND4)}$	-	-	TN93	Uniform [0.05 - 20]
Individual locus $k1_{(ND4)}$	-	-	TN93	Gamma [0.05 - 20]
Mean $k2_{(ND4)}$	-	-	TN93	Uniform [0.05 - 20]
Individual locus $k2_{(ND4)}$	-	-	TN93	Gamma [0.05 - 20]

Table 3.S3 Microsatellite-based (below diagonal) and ND4-based (above diagonal) pairwise estimates of F_{ST} between the seven genetic groups identified by STRUCTURE in *Alytes obstetricans*/*almogavarii* (see Figure 3.6). All P values < 0.01.

	Eastern Pyrenees	Central Pyrenees	Central-western Pyrenees 1	Central-western Pyrenees 2	Picos de Europa	Guadarrama 1 – lowland	Guadarrama 2
Eastern Pyrenees		0.831	0.941	0.937	0.910	0.913	0.951
Central Pyrenees	0.098		0.744	0.642	0.681	0.710	0.746
Central-western Pyrenees 1	0.122	0.097		0.302	0.727	0.820	0.906
Central-western Pyrenees 2	0.176	0.192	0.093		0.688	0.781	0.914
Picos de Europa	0.092	0.107	0.095	0.155		0.776	0.858
Guadarrama 1 – lowland	0.106	0.132	0.102	0.151	0.060		0.089
Guadarrama 2	0.175	0.195	0.157	0.222	0.115	0.084	

Table 3.S4 Analysis of molecular variance (AMOVA) for mitochondrial (ND4) and nuclear (microsatellites) markers based on the seven genetic groups identified by STRUCTURE in *Alytes obstetricans*/*almogavarii* (see Figure 3.6). All P values < 0.001.

	ND4			Microsatellites		
Source of variation	SS	Variance component	% Variation	SS	Variance component	% Variation
Among clusters	1149.248	6.259	84.945	1639.308	1.042	26.171
Among populations within clusters	148.165	0.466	6.319	1275.452	0.693	17.405
Among individuals within populations	77.883	0.644	8.736	3640.199	2.247	56.425

Table 3.S5 Posterior probability of tested scenarios and 95% confidence intervals (CI) estimated with DIYABC analysis when considering only microsatellites and when including both mtDNA (ND4) and microsatellite markers. Type I and II errors for the best supported scenarios are indicated. See Figure 3.2 for more information on tested scenarios.

	Microsatellites				Microsatellites + ND4			
Scenario	Posterior probability	95% CI	Type I error	Type II error	Posterior probability	95% CI	Type I error	Type II error
1	0.101	0.038-0.163			0.063	0.032-0.095		
2	0.849	0.837-0.860	0.114	0.075	0.292	0.274-0.310		
3	0.005	0.000-0.073			0.056	0.023-0.089		
4	0.007	0.000-0.074			0.109	0.058-0.159		
5	0.039	0.000-0.112			0.480	0.459-0.502	0.078	0.089

Table 3.S6 Posterior parameters (median and 95% confidence intervals) and RMedAD (Relative Median Absolute Deviation) estimated with DIYABC analysis for the best supported scenarios when considering only microsatellites (simple sequence repeats – SSRs; scenario 2) and when including both mtDNA (ND4) and microsatellite markers (scenario 5). See Figure 3.2 for more information on tested scenarios. N, effective population size for each analysed deme (EPY – eastern Pyrenees; CPY – central Pyrenees; CWPY – central-western Pyrenees; PEU – Picos de Europa Mountains; GUA – Guadarrama Mountains); ra, admixture rate; t, time of events in generations (t_1 – time to the most recent split; t_2 – time to the intermediate split; t_3 – time to the most ancient split); mean μ, mean mutation rate; mean P, mean coefficient P; mean k, mean coefficient k; $Q_{2.5}$, quantile 2.5%; $Q_{97.5}$, quantile 97.5%.

	Microsatellites				Microsatellites + ND4			
Parameter	Median	$Q_{2.5}$	$Q_{97.5}$	RMedAD	Median	$Q_{2.5}$	$Q_{97.5}$	RMedAD
N_{EPY}	12 100	6 110	18 500	0.159	15 800	10 500	19 200	0.201
N_{CPY}	8 830	3 820	17 300	0.195	18 200	14 300	19 800	0.202
N_{CWPY}	10 500	4 580	18 100	0.189	13 700	8 200	18 500	0.196
N_{PEU}	11 800	5 220	19 000	0.184	14 600	8 370	19 200	0.189
N_{GUA}	5 500	2 460	12 100	0.193	5 590	2 620	11 600	0.278
$N_{EPY-GUA}$	18 200	1 330	38 400	0.376	-	-	-	-
$N_{EPY-CWPY-GUA}$	-	-	-	-	11 700	742	35 900	0.486
ra	-	-	-	-	0.350	0.043	0.874	0.334
t_1	12 600	5 370	24 000	0.223	20 700	11 200	28 100	0.360
t_2	18 400	6 770	36 900	0.228	26 400	12 900	38 000	0.247
t_3	28 000	9 600	48 500	0.194	38 100	19 100	49 200	0.249
Mean $\mu 1_{(SSRs)}$	$1.28x10^{-4}$	$4.36x10^{-5}$	$4.05x10^{-4}$	0.414	$4.59x10^{-5}$	$1.93x10^{-5}$	$1.70x10^{-4}$	0.614
Mean $P1_{(SSRs)}$	0.216	0.111	0.294	0.258	0.177	0.104	0.287	0.222
Mean $\mu 2_{(SSRs)}$	$3.89x10^{-4}$	$1.86x10^{-4}$	$8.07x10^{-4}$	0.328	$3.31x10^{-4}$	$1.49x10^{-4}$	$7.41x10^{-4}$	0.536
Mean $P2_{(SSRs)}$	0.268	0.144	0.300	0.262	0.192	0.107	0.292	0.228
Mean $\mu_{(ND4)}$	-	-	-	-	$1.96x10^{-7}$	$1.06x10^{-7}$	$4.42x10^{-7}$	0.217
Mean $k1_{(ND4)}$	-	-	-	-	17.400	2.620	20.000	0.421
Mean $k2_{(ND4)}$	-	-	-	-	11.500	0.724	19.800	0.410

Supplementary figures

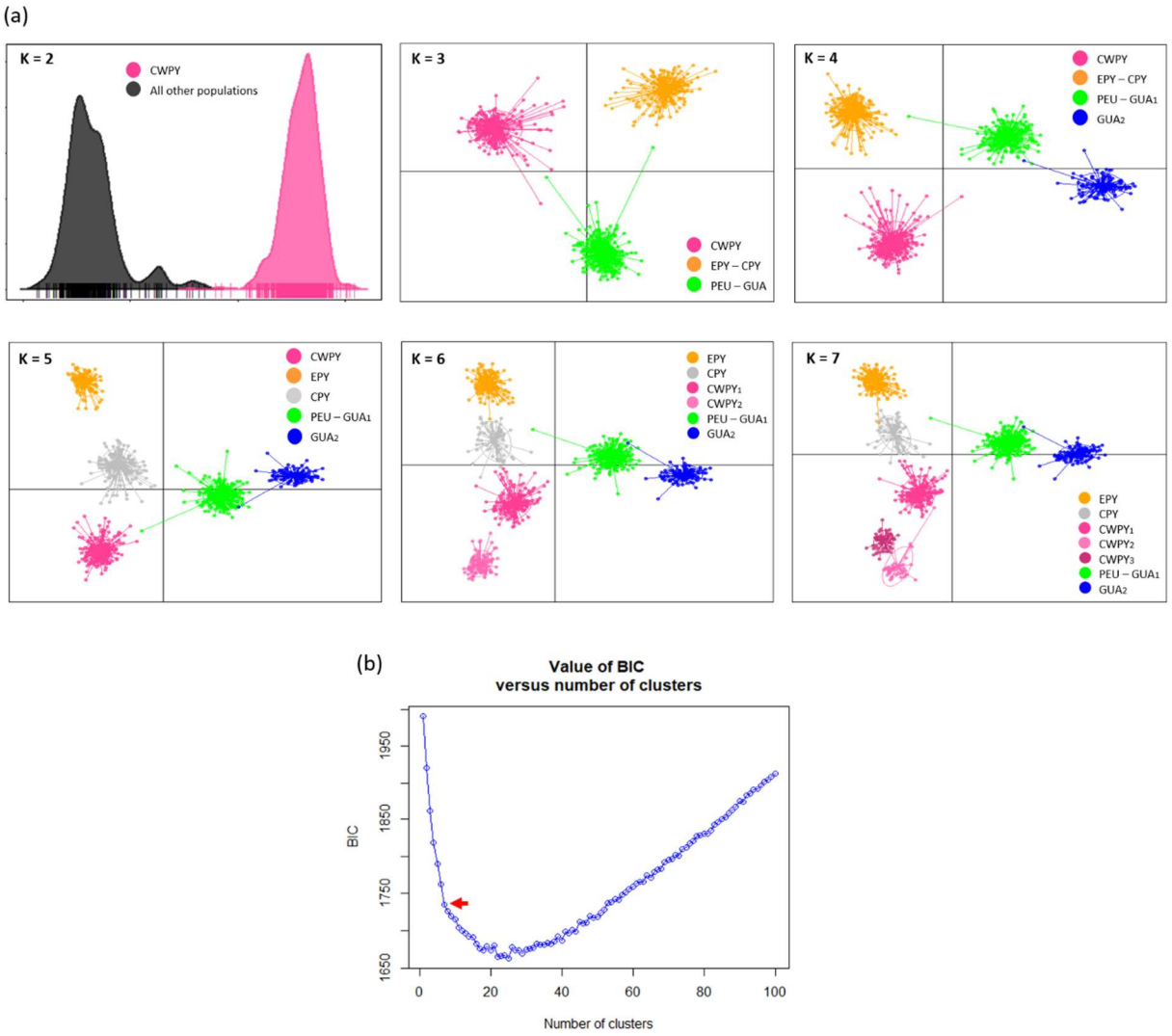


Figure 3.S1 Resulting plots from Discriminant Analysis of Principal Components (DAPC) across all *Alytes obstetricans*/*almogavarii* populations. Panel (a) shows summary plots for K = 2-7 genetic clusters. At K = 2, genetic clusters are represented as density curves. At K = 3-7, dots represent individuals and genetic clusters are shown as inertia ellipses. Legend labels indicate the different genetic clusters: eastern Pyrenees (EPY, orange), central Pyrenees (CPY, grey), central-western Pyrenees (CWPY, pink), Picos de Europa mountains (PEU, yellow), and Guadarrama Mountain Range (GUA, blue). Panel (b) shows the distribution of BIC (Bayesian Information Criterion) values according to the number of clusters. The red arrow indicates the number of clusters chosen for DAPC analysis.

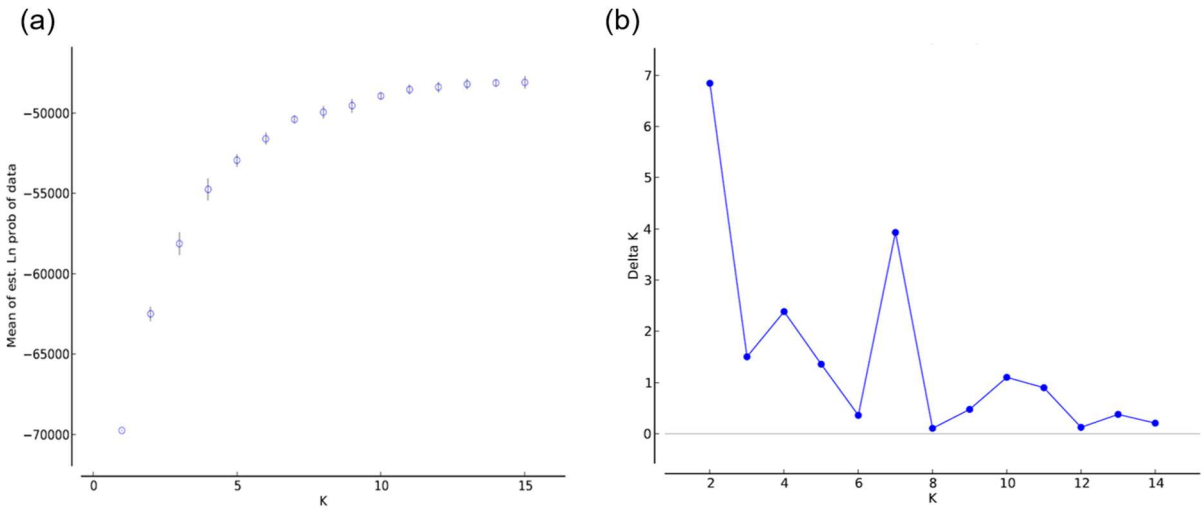
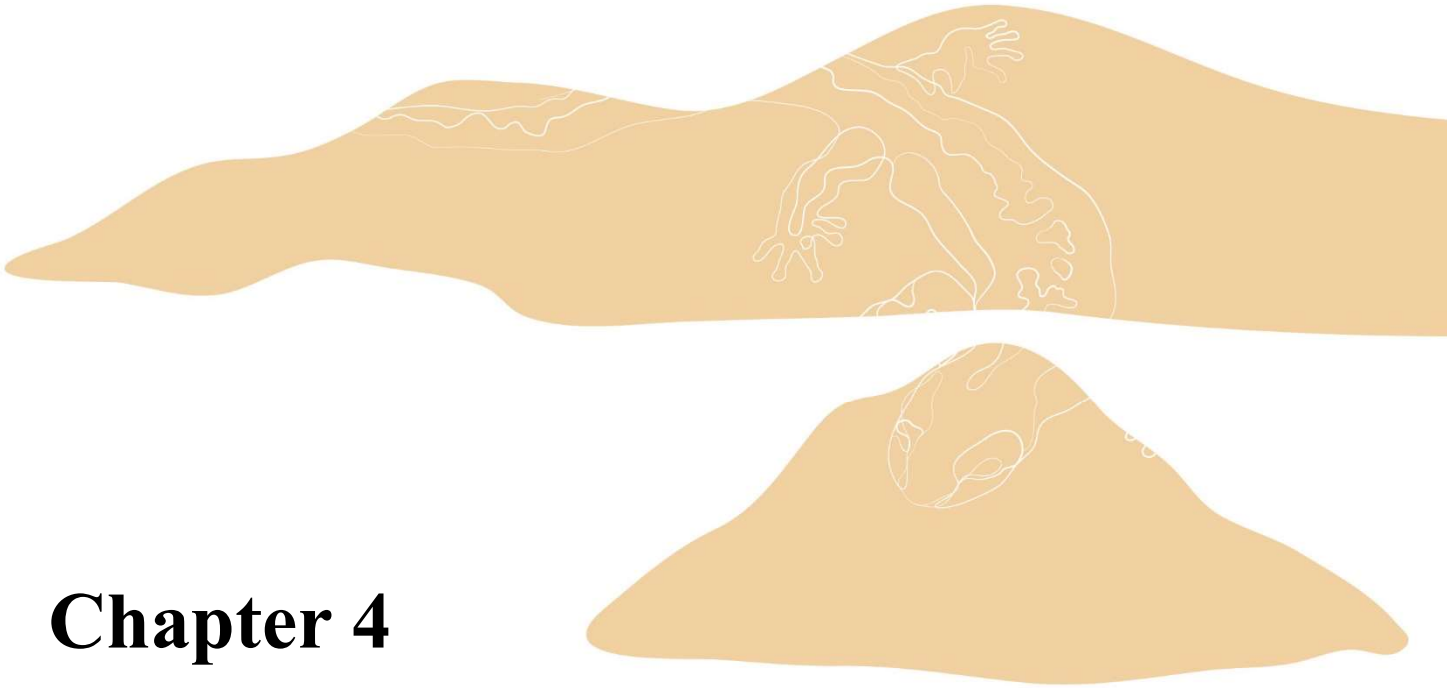


Figure 3.S2 Results of Bayesian clustering analyses in STRUCTURE for microsatellites. Panel (a) displays mean (± SD) log probability of the data [Ln Pr(XIK)] over 10 runs, for each value of K. Panel (b) shows ΔK values as a function of K, calculated according to Evanno et al. (2005).

References

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.



Chapter 4

A genetic appraisal of regional population connectivity in the endemic Pyrenean amphibian *Calotriton asper* and patterns of recolonization after fish removal

Federica Lucati, Alba Castrillón, Alexandre Miró, Ibor Sabás, Jan Tomàs, Blanca Font, David O'Brien, Teresa Buchaca, Daniel Oro, Marc Ventura

Abstract

Dispersal plays a key role in maintaining connectivity among populations and, if limited, is one of the main processes to promote regional genetic structuring. Here we made use of a combination of genetic and photographic mark-recapture analyses to study the factors affecting population connectivity in the endemic newt *Calotriton asper* at a small spatial scale, as well as the process of amphibian recolonization following non-native fish removal, in an interconnected system of 14 high mountain water bodies located in a protected area in the central-eastern Pyrenees. Microsatellite loci revealed high levels of genetic differentiation and population structuring in the study area. Nevertheless, we inferred modest levels of dispersal and identified a number of migrants and individuals of admixed ancestry, suggesting possible intervalley movement and recent reproductive connectivity between several populations, except those from the most isolated localities. Analysis of recolonization routes of stocked-now-fishless lakes revealed the importance of neighbouring never-stocked areas as source of individuals. Finally, we provide the first estimates of census population size for *C. asper* in two high mountain lakes. Conservation efforts should focus on preserving connectivity between populations that are still in migratory contact and attempt to increase or restore connectivity in isolated localities.

Keywords: gene flow, introduced fish, pattern recognition, photo-identification, Pyrenean brook newt

Introduction

Dispersal, intended as the movement of individuals that can sustain gene flow, plays a key role in maintaining connectivity among populations and, if limited, is one of the main processes to drive and promote regional genetic structuring (Ronce 2007). Dispersal patterns among populations can be strongly affected by landscape features, which can be perceived as a structuring force (Baguette et al. 2013; Zarnetske et al. 2017). Therefore, unveiling how movement and dispersal among populations are influenced by landscape complexity is of crucial interest in biology.

Introduced species can have profound effects on landscape structure, as they constitute an important stressor that can alter the habitat structure of native species and ultimately promote habitat fragmentation and degradation (Turner 2005; Zarnetske et al. 2017), and this in turn can hamper dispersal of native species among suitable habitat patches (Didham 2010). The impact of introduced species on the landscape is especially evident in high mountain water bodies, where non-native fish are causing profound ecological changes in several areas around the world (Davidson and Knapp 2007; Miró et al. 2018; Ventura et al. 2017). Northern hemisphere high mountain water bodies are naturally fishless, due to natural barriers that have kept them isolated from lower altitude streams. When introduced, fish constitute a threat to native species, through direct predation on native fauna or competition for prey (Epanchin et al. 2010; Miró et al. 2018). To reverse this biodiversity loss, several regions of the world have implemented restoration projects aiming to return high mountain water bodies to their natural fish-free status (Ventura et al. 2017), showing how fish removal allows for rapid natural recovery of native fauna and habitats (Knapp et al. 2005; Miró et al. 2020; Tiberti et al. 2019).

High mountain ecosystems host high levels of amphibian endemism and thus represent important reservoirs of genetic diversity for this animal class (Sandel et al. 2011). The succession of glacial and interglacial periods during Quaternary climatic oscillations caused repeated range shifts in species' distributions, and thus contributed significantly to shaping the current genetic structure of montane species (Hewitt 2000, 2004). During periods of adverse climatic conditions, populations remained confined in glacial refugia, where genetic differentiation and divergence among local populations was promoted. Such genetic isolation ultimately contributed to laying the genetic foundations of contemporary populations, which still bear hallmarks of these past dynamics (Schmitt 2009). Nowadays, amphibians are among

the most conspicuous native species in high mountain lentic habitats. Nevertheless, several mountain regions are exhibiting high levels of amphibian population declines, being non-native fish, habitat loss and fragmentation and pesticides among the main threats (e.g. Davidson and Knapp 2007; Miró et al. 2018).

The Pyrenean brook newt (*Calotriton asper*) is a small-sized urodele endemic to the Pyrenees. It is a highly aquatic species that can be found in mountain lakes, ponds, caves and streams at mid to high elevations (Clergue-Gazeau and Martínez-Rica 1978; Martínez-Rica and Clergue-Gazeau 1977). During its life cycle, the species interacts with both freshwater and terrestrial habitats: upon metamorphosis, a 2-years-long juvenile terrestrial phase precedes the adult stage, where individuals return to the water for the remainder of their life cycle (Montori 1988; Montori and Llorente 2014). Evidence suggests that *C. asper* experienced dramatic population declines during the 20th century (Bosch et al. 2009), and the negative effect of introduced fish on the species is documented (Miró et al. 2018). Its listing as *Near Threatened* by the IUCN due to habitat loss and fragmentation shows that non-native fish can seriously hamper the conservation of the species (Bosch et al. 2009). Previous studies on *C. asper* revealed five major genetic lineages at the Pyrenean scale (Milá et al. 2010; Valbuena-Ureña et al. 2018), which correspond to distinct glacial refugia where the species survived Quaternary glaciations (Lucati et al. 2020b). Both genetic and field-based approaches evidenced the species' fairly low vagility (Lucati et al. 2020b; Montori et al. 2008a; Montori et al. 2008b), while denoting a complex pattern of migration between neighbouring lakes and/or streams (Lucati et al. 2020a; Oromi et al. 2018).

We focused on the National Park of Aigüestortes i Estany de Sant Maurici (PNAESM) in the central-eastern Pyrenees. In this region, lake restoration activities via fish removal have been ongoing since 2015 within the context of the conservation project LIFE+ LIMNOPIRINEUS (www.lifelimnopirineus.eu), with the aim of returning selected high mountain lakes to their fish-free natural state and allow the recovery of native species, such as amphibians (Miró et al. 2020). This, coupled with the potentially high habitat availability for *C. asper* in PNAESM, makes it a natural laboratory to study, at a regional scale, the factors affecting gene flow and population connectivity, as well as the process of amphibian recolonization following fish removal. Within PNAESM, we sampled populations at both broad and restricted spatial scales, in order to gain insights into *C. asper* short- to medium-distance movements and evaluate the

importance of the landscape in influencing its distribution and dispersal patterns. To do this, we used a combination of both genetic and mark-recapture (via photo-identification) analyses. This also allowed us to directly compare the congruency of genetic estimates of effective population size with more traditional population size estimates based on mark-recapture methods.

Materials and Methods

Study area and fieldwork

Assessment of population structure and recent dispersal was conducted at 14 permanent high mountain lakes, ponds and rivers across six glacial cirques of PNAESM (Figure 4.1, Table 4.1). All sites are of glacial origin and lie at elevations ranging from 2,021 to 2,510 m a.s.l. Distance between individual sites within glacial cirques is less than 1 km. Of the 14 water bodies, 8 are fishless and sustain natural populations of *C. asper*, 5 are undergoing complete fish eradication or intensive control and are being colonized by *C. asper*, and one still harbours non-native fish and almost no newts were reported (Figure 4.1, Table 4.1). Population size estimation took place at two of the 14 water bodies, which are fishless and are known to host stable *C. asper* populations. The lake (Estany Xic de Subenuix, XSU) covers an area of 0.58 ha and has a maximum depth of 4.3 m. The pond (Bassa Natural de Dellui, NDE) is smaller in size (0.32 ha) and has a maximum depth of 1.95 m.

Sampling was mostly carried out during the summers of 2016-2017. Four populations were resampled in the two years to investigate the dynamics of recolonization following fish removal (Table 4.1). In the case of population size estimation, fieldwork was conducted for three consecutive years (2016-2018), totalling six sampling sessions for each study site, with an interval of at least two days between each consecutive visit, which lasted for 2-3 hours. Newts were captured manually or using hand nets, sexed, weighed and measured (snout-vent length – SVL) before having their ventral side photographed. Tissue samples were collected via toe clipping and stored in absolute ethanol at -20 °C. All individuals were released at their capture location.

Genetic analyses

Whole genomic DNA was extracted in a total volume of 100 µl following the HotSHOT extraction method (Montero-Pau et al. 2008). Samples were genotyped for a set of 17 microsatellite loci multiplexed in three mixes (Drechsler et al. 2013). Multiplex combinations

and amplification conditions follow Drechsler et al. (2013). PCR products were electrophoresed on an ABI 3730 capillary sequencer (Secugen, Madrid, Spain), fragments were sized with LIZ-500 size standard and binned using Geneious 11.0.5 (Kearse et al. 2012). A subset of samples was used in previous studies (Lucati et al. 2020a; Lucati et al. 2020b). We used MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) to identify potential scoring errors, large allele dropout and null alleles. We tested for linkage disequilibrium and for departures from Hardy-Weinberg equilibrium (HWE) using GENEPOP 4.2 (Rousset 2008). The Bonferroni correction was applied to adjust for multiple comparisons ($\alpha = 0.05$; Rice 1989).

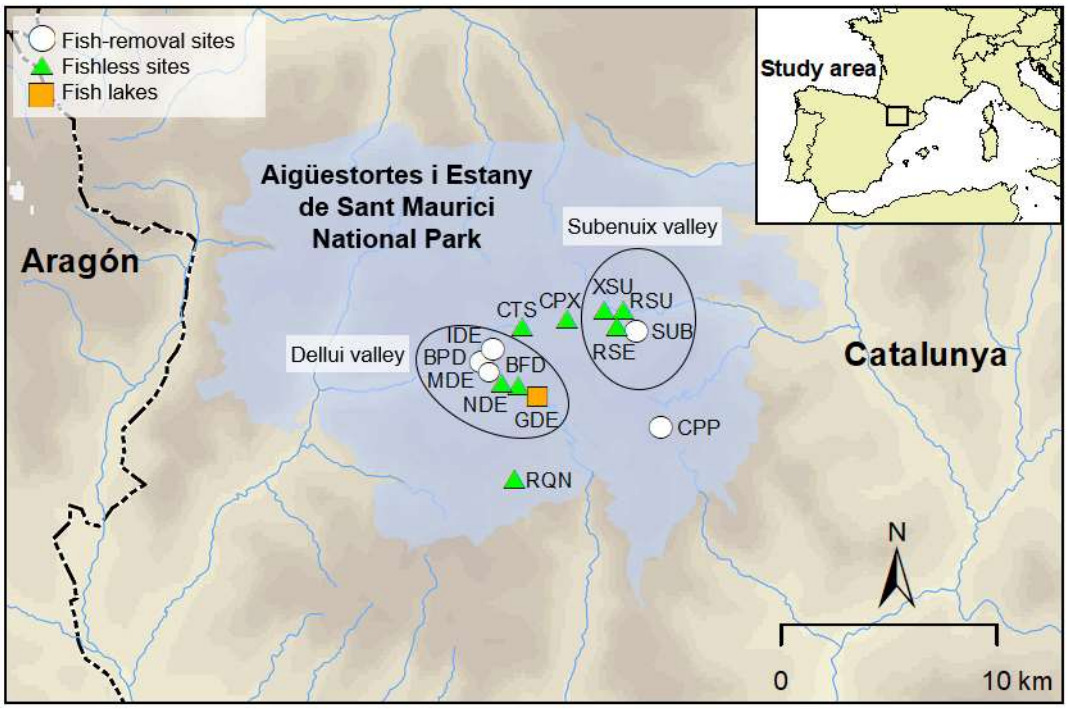


Figure 4.1 Map of the study area showing the sampling sites located in the Aigüestortes i Estany de Sant Maurici National Park. See Table 4.1 for details on each site.

Population structure was inferred using a Bayesian cluster analysis implemented in STRUCTURE 2.3.4 (Pritchard et al. 2000). All runs were repeated 20 times for each K, set between 1 and 14 (i.e. the number of sampling locations plus one), with 100K burn-in steps followed by 500K MCMC repetitions. We used the admixture model, assuming independent allele frequencies and an alpha value of 1/n (where n is the number of populations), individually inferred for each assumed population. This parameter set is recommended to improve individual

Table 4.1 Descriptive characteristics and geographical information for each analysed population. Lat., latitude; Long., longitude; Alt., altitude in meters; N microsat., sample size for microsatellites (asterisks indicate the four resampled populations); N photos, total number of photographed individuals; Recaptures, number of individuals recaptured at each sampling site; N_c, mean census population size; N_e, effective population size as calculated by Lucati et al. (2020b).

Population	Code	Lat.	Long.	Alt.	Type	Fish status	N microsat.	N photos	Recaptures	N_c	N_e
Dellui Nord	IDE	42.55	0.94	2306	Pond	Eradication	20*	24	0	-	19
Dellui Mig	MDE	42.55	0.94	2314	Pond	Eradication	39*	44	2	-	24
Dellui Freda	BFD	42.55	0.95	2314	Pond	Fishless	8	8	0	-	-
Dellui Natural	NDE	42.55	0.94	2313	Pond	Fishless	37*	62	10	80 ± 18.92	17
Dellui Gran	GDE	42.55	0.95	2349	Lake	Fish	-	1	0	-	-
Dellui Phoxinus	BPD	42.55	0.94	2307	Pond	Eradication	2	-	-	-	-
Riu de Subenuix Esquerre	RSE	42.58	0.99	2142	Stream	Fishless	5	5	0	-	-
Riu de Subenuix	RSU	42.58	0.99	2021	Stream	Fishless	21	21	1	-	23
Subenuix	SUB	42.57	0.99	2194	Lake	Eradication	5	11	0	-	-
Xic de Subenuix	XSU	42.57	0.99	2272	Lake	Fishless	36*	191	11	859 ± 196.96	33
Coma de Peixerani	CPX	42.57	0.97	2297	Stream	Fishless	15	21	0	-	32
Redó de Riquerna	RQN	42.51	0.95	2411	Lake	Fishless	14	21	0	-	-
Corticelles	CTS	42.56	0.95	2278	Lake	Fishless	15	16	0	-	-
Cap del Port de Peguera	CPP	42.54	1.03	2510	Stream	Eradication	2	-	-	-	-
Total							234	425	24		

assignment to populations when sampling is uneven (Wang 2017). The software was run also for each cluster separately. To estimate the optimal number of clusters we used both the original method of Pritchard et al. (2000) and the ΔK method of Evanno et al. (2005), as implemented in STRUCTURE HARVESTER (Earl and vonHoldt 2012). We averaged replicate runs of the optimal K using CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007) and plotted the output with DISTRUCT 1.1 (Rosenberg 2004). In addition, to further investigate the population structure of the studied populations, pairwise F_{ST} among populations was calculated with GenAlEx 6.5 (Peakall and Smouse 2012), applying 1,000 permutations to assess statistical significance.

GeneClass2 was used to detect first generation migrants, i.e. individuals born in a population other than that where they were collected (Piry et al. 2004). We employed two different likelihood computation criteria for migrant detection: L_{home}, that is the likelihood of finding an individual in the population where it was sampled, and L_{home}/L_{max}, the ratio of L_{home} to the highest likelihood among all sampled populations. L_{home} is more appropriate when, as in our case, not all source populations for migrants have been sampled. However, it is less powerful than L_{home}/L_{max}. We used the Bayesian method of Rannala and Mountain (1997) coupled with the Monte-Carlo resampling method of Paetkau et al. (2004), and applied 10,000 simulated individuals. Type I error (alpha level) was set to 0.01. To visualize first generation migrants' trajectories, we generated a chord diagram using the circlize R package (Gu et al. 2014).

We resampled four populations in 2 years to examine between-year reproductive patterns and the dynamics of colonization of new ecosystems. Of the four sites, two are fishless (population codes NDE and XSU) and two are undergoing complete fish eradication (codes IDE and MDE). ML-RELATE (Kalinowski et al. 2006) was employed to calculate maximum likelihood estimates of relatedness between individuals, setting confidence intervals at 0.95 and applying 10,000 randomizations. The program assigns relatedness according to the following categories: parent-offspring (PO), full-siblings (FS), half-siblings (HS) and unrelated (U). If U had equal probability as HS and FS, we grouped these relationships as a new category (third-order, e.g. cousins) (see Savage et al. 2010 for a similar approach). For each sampled site, we calculated relatedness separately for each year, as well as by pooling samples from both years. Furthermore, we used POPTREEW (Takezaki et al. 2014) to build a neighbour-joining (NJ) tree, using Nei's genetic distance (D_A, Nei et al. 1983) with 1,000 bootstraps. NJ analysis was performed among all sampling sites, including the four resampled populations.

We used Linkage Mapper 2.0.0 (McRae and Kavanagh 2011) to identify dispersal corridors and calculate least-cost paths between sampled areas. The least-cost path, i.e. the most likely and cost-effective route an individual would take to move between two points, represents a more realistic measure than Euclidean distance, as it takes into account the landscape configuration and topographic complexity (Adriaensen et al. 2003). Core areas were represented by the 14 target water bodies. We created a resistance surface that considered the following landscape features: slope, elevation and fish presence. Slope and elevation were derived from a digital elevation model, whereas fish data were obtained from field surveys and previously published literature (Miró and Ventura 2013). For each landscape feature, we produced a resistance surface in which values ranged from 1 to 1,000, where values of 1 represent ideal habitats and higher values indicate habitats of increasingly lower suitability (Table 4.S1; Beier et al. 2011; WHCWG 2010). The final resistance surface raster was generated using Raster Calculator in ArcMap 10.1 (ESRI, Redlands, CA, USA). Linkage Mapper was run in ArcMap using the Linkage Pathways tool with default parameters.

I^3S individual identification

Individual photographs were compared and matched using the open-source pattern identification software I^3S Pattern$^+$ 4.1 (https://reijns.com/i3s/; Van Tienhoven et al. 2007). I^3S Pattern$^+$ requires the user to specify three reference points on the animal's body to be used as reference system; we thus selected the right and left base of the head and the base of the tail. The user is asked to annotate the region of interest for individual identification and to segment the region into foreground (the coloured pattern) and background (dark spots and marks) by marking a number of representative points. The program compares the annotated image to the ones already included in the dataset and presents the user with a list of possible matches, ranked according to a dissimilarity score calculated for each pair of images. The user can then select whether to accept or reject a match.

The software's performance was assessed using I^3S evaluation tool ("Expert" menu). We determined whether I^3S Pattern$^+$ can be reliably used for *C. asper* individual identification by performing an intra-individual comparison between images of the same individual and counting the total number of recaptures that were ranked within the top 10 images. The stability of the

ventral pattern was evaluated by visually inspecting photographs of newts recaptured ≥ 1 year apart.

Population size estimation

The program MARK 9.0 was used to estimate recapture (p) and survival (ϕ) probabilities of individuals within each of the two populations under the Cormack-Jolly-Seber (CJS) model (White and Burnham 1999). We fitted two models for each population: recapture probability was set as fixed (.) while survival probability was defined as either fixed or variable (t) over time. The model with the lowest adjusted Akaike's Information Criterion (AIC_c) was selected as the model that best fit the data (Hurvich and Tsai 1989; Sugiura 1978). Population size (N) estimates for each capture event i were calculated using the recapture probabilities retrieved from the models, according to the formula: $N_i = n_i/p_i$ (Pollock et al. 1990). The standard error (SE) was calculated using the Delta method (Seber 1982): $SE(N_i) \approx ((p_i)(n_i/p_i) + (n_i(1-p_i)/p_i))^{1/2}$.

Results

Overall genetic structure and dispersal

MICRO-CHECKER did not detect the presence of scoring errors, null alleles or large allele dropout. Similarly, there was no sign of linkage disequilibrium or departures from HWE after applying the Bonferroni correction.

STRUCTURE analysis revealed that the 13 populations were subdivided into three genetic clusters with a strong association with geography (Figures 4.2 and 4.S1). Indeed, the three clusters were restricted to the eastern, western and southern part of PNAESM, respectively. Modest levels of admixture were detected between clusters 1 (west: Dellui valley + CTS) and 2 (east: Subenuix valley + CPX), indicating possible intervalley movement, while cluster 3 (south) was highly distinct. When analysing clusters 1 and 2 separately, both grouped into two subclusters (Figures 4.2 and 4.S1): clustering of western populations resulted in a north-south gradient of genetic differentiation, whereas eastern populations showed a sharp genetic break between the two valleys under consideration (Subenuix and Coma de Peixerani valleys). Pairwise F_{ST} values confirmed these findings, denoting genetic similarity in populations belonging to the valley of Subenuix, and higher differentiation in the valley of Dellui (Table 4.2).

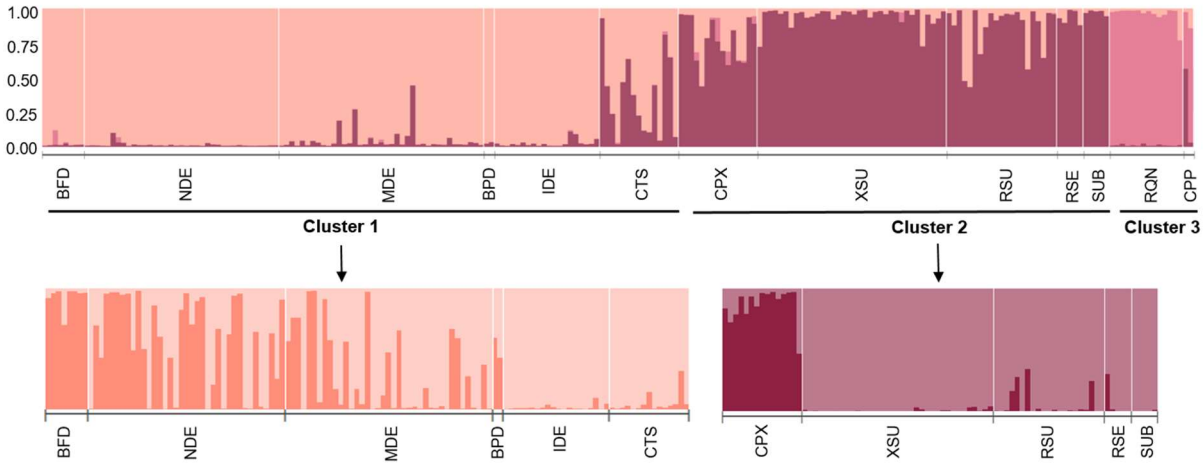


Figure 4.2 Resulting bar plots of STRUCTURE analyses. The top bar plot is for K = 3 across all populations, and below are the bar plots for K = 2 for clusters 1 and 2 analysed separately. Each individual is represented by a single vertical bar, and the length of each bar represents the probability of membership in each cluster. White lines separate populations. For population codes see Table 4.1.

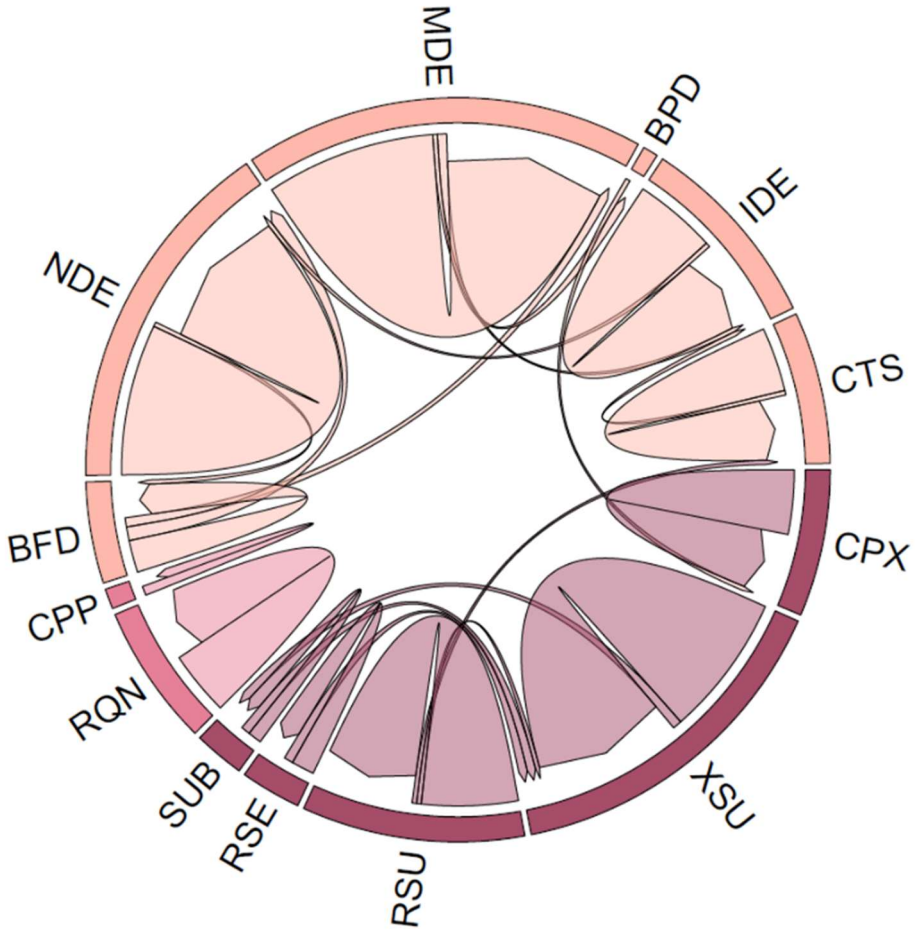


Figure 4.3 Chord diagram tracking first generation migrants' flows between sampled populations as inferred by GeneClass2. Chord size is proportional to the number of migrants detected and arrows indicate the direction of migration. Colours delineate the three genetic clusters identified by STRUCTURE (light pink: cluster 1, burgundy: cluster 2, dark pink: cluster 3). For population codes see Table 4.1.

Table 4.2 Pairwise F_{ST} estimated for the 13 sites sampled for *C. asper*. Bold values denote statistical significance ($P < 0.05$). NA = not available. For population codes see Table 4.1.

	SUB	XSU	RSE	RSU	CPX	CTS	IDE	BPD	MDE	NDE	BFD	CPP	RQN
SUB	0												
XSU	**0.033**	0											
RSE	0.030	**0.016**	0										
RSU	**0.052**	**0.059**	**0.076**	0									
CPX	**0.186**	**0.146**	**0.134**	**0.142**	0								
CTS	**0.093**	**0.096**	**0.125**	**0.063**	**0.152**	0							
IDE	**0.316**	**0.226**	**0.273**	**0.204**	**0.281**	**0.134**	0						
BPD	0.226^{NA}	0.136^{NA}	0.148^{NA}	0.124^{NA}	0.027^{NA}	0.047^{NA}	0.106^{NA}	0					
MDE	**0.205**	**0.173**	**0.181**	**0.145**	**0.158**	**0.072**	**0.049**	-0.038^{NA}	0				
NDE	**0.293**	**0.210**	**0.246**	**0.210**	**0.201**	**0.149**	**0.108**	-0.024^{NA}	**0.044**	0			
BFD	**0.309**	**0.230**	**0.266**	**0.243**	**0.180**	**0.218**	**0.305**	0.051^{NA}	**0.138**	**0.089**	0		
CPP	0.258^{NA}	0.162^{NA}	0.144^{NA}	0.208^{NA}	0.205^{NA}	0.222^{NA}	0.348^{NA}	0.122^{NA}	0.232^{NA}	0.249^{NA}	0.221^{NA}	0	
RQN	**0.362**	**0.323**	**0.309**	**0.313**	**0.295**	**0.316**	**0.393**	0.257^{NA}	**0.307**	**0.335**	**0.331**	0.118^{NA}	0

Estimation of recent dispersal in GeneClass2 indicated that, for clusters 1 and 2, migration between populations belonging to the same genetic cluster was relatively common (cluster 1: 12 cases, cluster 2: 8 cases; Figure 4.3). This was especially true for populations located in the same glacial cirque. Nevertheless, consistently with the results of STRUCTURE, we also detected two cases of recent migration between clusters 1 and 2. Least-cost path analysis indicated that the length of dispersal corridors used by first generation migrants was below 4 km (Figure 4.4, Table 4.S2). STRUCTURE also identified 26 individuals with admixed ancestry, i.e. individuals having an intermediate genotype between two clusters, suggesting that they are the product of admixture between genetic lineages (Figure 4.5). Such individuals have Q-values between 0.2 and 0.8, indicating nontrivial membership in more than one cluster (Bergl and Vigilant 2007; Lecis et al. 2006; Vähä and Primmer 2006). Of the 26 newts, two were previously identified as migrants by GeneClass2. All but two individuals were inferred to be admixed between clusters 1 and 2, the majority belonging to populations located at the border between the two clusters (i.e. CTS and CPX).

Temporally resampled populations and restored lakes

Our relatedness analysis showed that very few individuals, both within and between years, were half or full siblings, with most individuals being unrelated or only weakly related, and that the relative proportion of relationship categories was relatively stable between years (Figure 4.6). Furthermore, we carried out STRUCTURE analyses for the two restored populations pooled across the 2 years (Figure 4.6). Bar plots showed that while in population MDE K = 2 genetic units exist, with relatively high probabilities of individual assignment to the respective year of sampling, no apparent genetic structure was detected in population IDE. Similarly, the NJ tree revealed that individuals collected in population MDE in 2016 were more genetically related to populations located in the southern sector of the Dellui valley, whereas individuals collected in 2017 had higher affinity to the northern Dellui lakes (Figure 4.7).

I^3S individual identification

A total of 425 individual Pyrenean brook newts were captured and identified during the 3-year period (Table 4.1). The ventral pattern of each captured individual was unique and sufficiently varied to allow for the reliable individual identification in I^3S Pattern$^+$ (Figure 4.8). Of the 425 individuals, 24 were recaptured at least once. All recaptures were ranked between the first 10

listed images in I[3]S Pattern[+]. Furthermore, the visual inspection of recaptures revealed no detectable changes in the ventral pattern of individual newts over the study period.

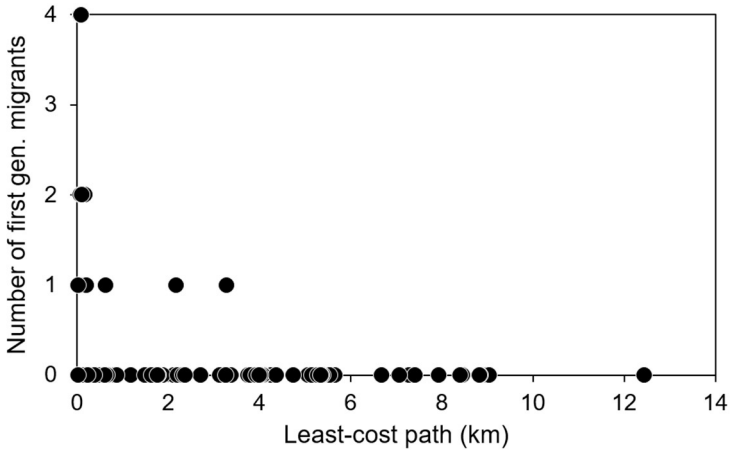


Figure 4.4 Scatter plot showing the least-cost distance covered by first generation migrants as inferred by GeneClass2.

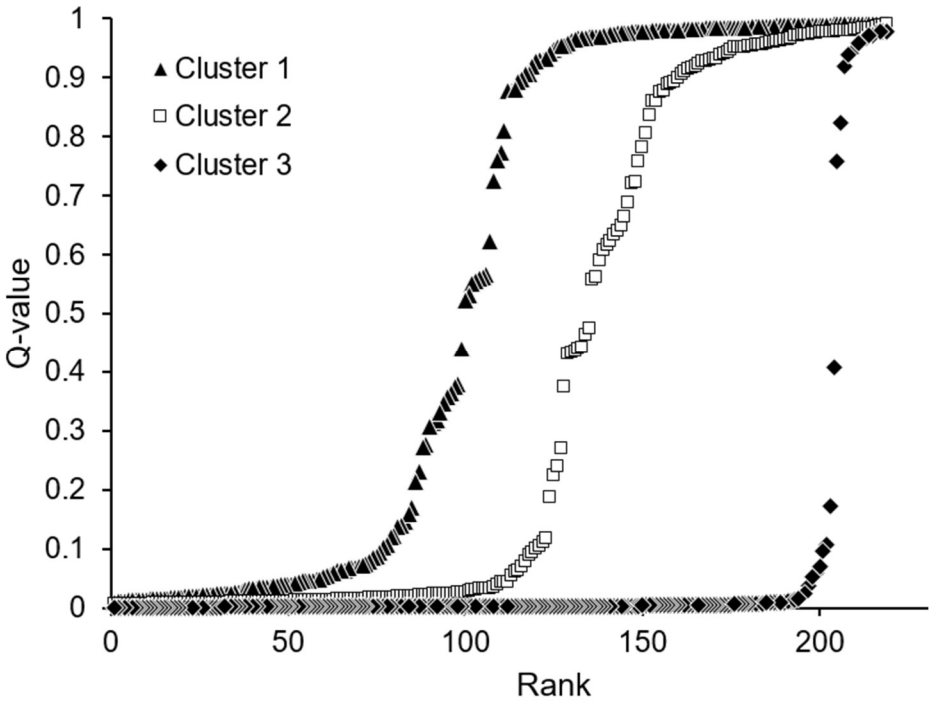


Figure 4.5 Ranked Q-values (probability of membership in each cluster) for each individual in each cluster as inferred by STRUCTURE.

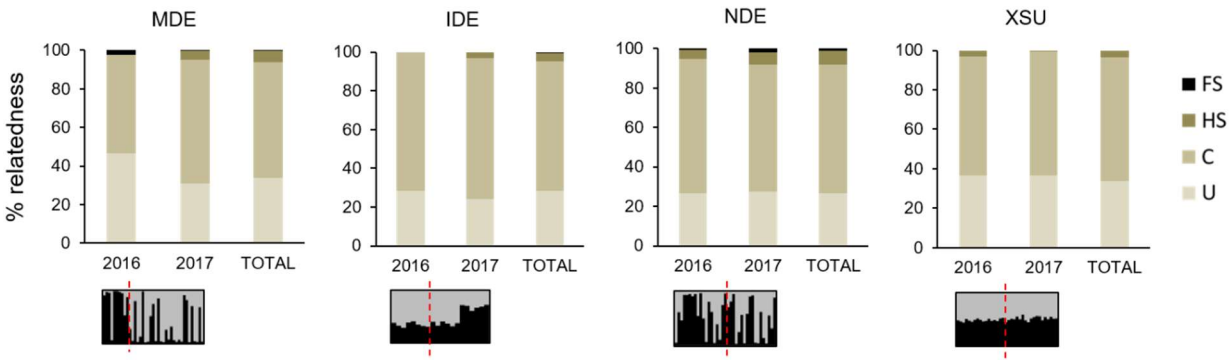


Figure 4.6 Percentage of individual relatedness within each of the four resampled populations. Relatedness was calculated per sample year (2016 and 2017) and by pooling samples from both years (TOTAL). Relationship categories, as estimated by ML-RELATE, were as follows: U (unrelated), C (third-order, e.g. cousins), HS (half-siblings) and FS (full-siblings). No parent-offspring relationships were detected. STRUCTURE bar plots (K = 2) showing between-year sample assignment proportion are presented for the two restored lakes (MDE and IDE) and the two fishless lakes (NDE and XSU).

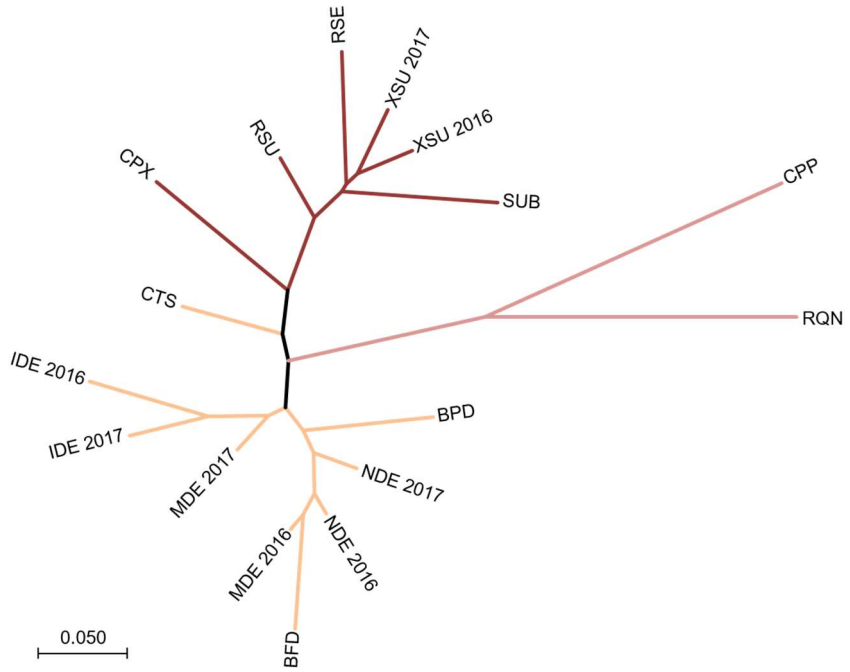


Figure 4.7 Neighbour-joining tree based on D_A distances showing the relationship between the three genetic clusters inferred by STRUCTURE (light pink: cluster 1, burgundy: cluster 2, dark pink: cluster 3; see Figure 4.2). The four resampled populations (MDE, IDE, NDE and XSU) are labelled by year of sampling. See Table 4.1 for population codes.

Of the 24 recaptures, only three were found in a site other than the place of first collection. In all cases, migration involved neighbouring populations separated by a least-cost distance of 15-84 m (Table 4.S2): the first individual was first captured in population SUB (a restored lake) and recaptured one year later in RSU (a fishless stream), while the second and third individuals

were first captured in NDE (fishless lake) and recaptured two years later in MDE (restored lake).

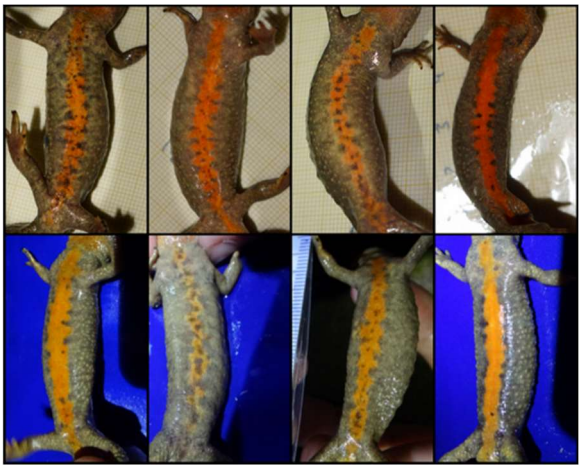


Figure 4.8 Example images of captured newts showing the variation in ventral pattern.

Population size estimation

Overall, 191 (87 males, 104 females) and 62 (45 males, 17 females) Pyrenean brook newts were captured and photographed in populations XSU and NDE, respectively (Table 4.1). The total number of recaptures in both sites was low, namely 11 and 10 in XSU and NDE, respectively.

The model with fixed recapture probability and variable survival probability best fitted the data for both populations ($p.\phi_t$). Population sizes, as predicted from the model, varied considerably between sites and also within sites between successive capture occasions (Table 4.S3). Mean population size estimates (mean ± SE) over a period of three years were higher in XSU (859 ± 196.96) than in NDE (80 ± 18.92). As expected, the low proportion of recaptured individuals led to estimated population sizes with large variation.

Discussion

Genetic structure and dispersal on a small spatial scale

Extreme habitats such as high mountains, which are typically characterized by severe climatic, topographical and ecological conditions, may limit population sizes and dispersal, leading to high levels of genetic subdivision (Giordano et al. 2007; Sánchez-Montes et al. 2018). Indeed, the intensity of cold temperatures, the duration of the snow cover and the lake freezing/drying

cycles pose several challenges to amphibians living in high mountain habitats. Here, we found evidence of fine-scale genetic structuring in the endemic newt *C. asper* in a restricted area of the central-eastern Pyrenees (Figure 4.2). More into detail, we detected three genetically and geographically distinct lineages located in the eastern, western and southern parts of PNAESM. Furthermore, further substructure emerged from the eastern and western clusters. These findings complement and extend previous work on the spatial genetic structure of *C. asper*, which was mostly conducted at broad spatial scales (e.g. Lucati et al. 2020b; Milá et al. 2010; Oromi et al. 2018; Valbuena-Ureña et al. 2018), by providing a smaller-scale perspective on the species' dynamics.

In spite of the high levels of genetic subdivision, our data suggested that inter-cluster movements continue or have recently occurred between several localities. In fact, results showed that a few individuals migrated between glacial cirques or genetic clusters within the current generation (Figure 4.3), whereas several other individuals showed recent ancestry in a genetic cluster other than their source cluster (Figure 4.5). As expected, signs of recent migration and/or admixture mainly involved populations situated at clusters' borders and tended to fade away with distance. Nevertheless, this is suggesting that Pyrenean brook newts are not only able to occasionally engage in inter-valley and/or inter-cluster movements, but that they are also able to reproduce in the new area. Such reproductive connectivity is deemed crucial for the long-term maintenance of genetic diversity in small fragmented populations (Bijlsma et al. 2000; Tallmon et al. 2004). In newts, the terrestrial juvenile phase is the primary dispersal stage in the life cycle (Gill 1978; Pittman et al. 2014; Semlitsch 2008). Juveniles can migrate relatively long distances, and they will be responsible for most genetic exchanges.

The southern cluster (cluster 3) remained distinct from the others and showed little evidence of genetic connectivity with other areas (Figures 4.3 and 4.5). The long branches separating this lineage from the others in the NJ tree are a clear sign of genetic divergence, possibly as a result of the lack of exchange of regular dispersers (Figure 4.7). Although the southern cluster is formed by the two most distant and isolated populations (least-cost distance to populations belonging to clusters 1 and 2 was between 5 and 12 km), this does not appear to be the whole story, because previous work on *C. asper* showed that the species is capable of dispersing over such distance on occasion (Lucati et al. 2020b). One possible explanation for this lack of connectivity is habitat fragmentation induced by introduced fish (Bradford et al. 1993; Miró et

al. 2018). Populations from the southern cluster are largely surrounded by water bodies stocked with fish, which may interfere strongly with *C. asper* dispersal movements. Under this scenario, opportunities for dispersing individuals to find new suitable habitats will be limited, thus promoting isolation and genetic divergence of populations (Gong et al. 2010). Taken together, our data suggest that the combination of distance and non-native species may be driving the local population structure seen in *C. asper*.

Temporal sampling and recolonization of restored lakes: where do newts come from?

Given the high levels of genetic structuring and differentiation observed in this study, we expected that the small local breeding populations at each site would be relatively inbred (Frankham et al. 2002; Madsen et al. 1999). Instead, we found that the four resampled populations were formed by largely unrelated individuals (Figure 4.6). Furthermore, no difference was found between fishless and restored lakes in terms of relatedness between individuals. This is in line with previous studies on *C. asper* showing low inbreeding coefficients and levels of genetic variability within the range of other temperate amphibians (Chan and Zamudio 2009; Lucati et al. 2020b; Valbuena-Ureña et al. 2018). In light of this, most of the water bodies included in this study seem to represent an interconnected set of populations, where productive sites supply neighbouring sites with regular dispersers.

With regard to the two temporally resampled restored lakes, we found differential patterns of year-to-year genetic differentiation and lake colonization (Figures 4.6 and 4.7). Both lakes (codes MDE and IDE) are located in the valley of Dellui, where fish eradication activities have been ongoing since 2015 (Figure 4.S2). At the beginning of fish eradication work, these lakes had no detectable amphibian populations. Species richness recovered completely only one year after the eradication began without the need for translocation or reintroduction, being the Pyrenean brook newt among the first colonists (Miró et al. 2020). Here, our molecular approach allowed us to uncover two genetic subclusters in the valley of Dellui, whose prevalence shifts in the central lakes (Figure 4.2). Lake MDE, situated in the central part of the valley (Figure 4.S2), exhibited high year-to-year genetic differentiation. According to the NJ tree, samples collected in 2016 were more genetically related to the southern fishless lakes, whereas in samples from 2017 there was higher ancestry in the northern lakes. The most plausible explanation is that, at the beginning of fish eradication, the first colonists came from the closest fishless water bodies (namely NDE and BFD), and, as fish removal work progressed, the

northern fish-removal lake IDE, which was being recolonized by newts itself, may have started to serve as source of individuals for MDE. On the other hand, lake IDE, situated at the northern edge of the valley, showed genetic consistency from year to year and exhibited genetic affinity to CTS (Figure 4.2), a fishless lake located in the adjacent glacial cirque. It is reasonable to think that CTS, despite not being the closest lake, acted as main donor for IDE at least at the initial stages of fish removal, given that the central Dellui lakes were all stocked with fish, which made it arduous for newts dispersing from the southern lakes to reach IDE (Figure 4.S2). In support of this, we found evidence of recent migration from CTS to IDE (Figure 4.3).

Genetic data vs. mark-recapture data

To provide insights into the spatial population structure and patterns of movement of species, both direct (mark-recapture) and indirect (genetic) estimates are useful (McCall et al. 2013; Trenham et al. 2001; Wang and Shaffer 2017). Direct data from marked individuals would be preferable, but the time and effort needed to characterize individual movements is often prohibitive. Indeed, for many species it is probably easier, cheaper and less intrusive to capture and sample individuals in order to perform fast genetic analyses than to carry out multiyear mark-recapture studies. This is especially true for endangered and secretive species such as *C. asper*, making molecular approaches an efficient way to collect data for conservation and management planning (Wang and Shaffer 2017).

Here, we confirm previous findings on the reliability and efficacy of photo-identification for *C. asper* individual recognition based on its ventral pattern (Lucati et al. 2020a). However, the low number of recaptures did not allow us to deeply characterize the species' recent movement patterns, as only three instances of interlake dispersal were detected. We therefore suggest caution should be exercised when inferring temporal dynamics from mark-recapture data over a short-term study such as ours, as reliable estimates of the dynamics of *C. asper* populations probably require multiple sampling sessions and long-term field monitoring. Nevertheless, we provide the first estimates of census population size for *C. asper* in two high mountain lakes. Census size was much higher for population XSU than NDE; lake XSU is twice as big and deeper than NDE and field observations indicate that XSU sustains a large *C. asper* population. Comparison with genetic-based effective population sizes of the same two populations (Lucati et al. 2020b; Table 4.1) gives a ratio of effective to census size (N_e/N_c) of 0.04 and 0.19 for XSU and NDE, respectively, which is within the range expected for animal populations

(Frankham 1995). In amphibians, small N_e/N_c ratios have been frequently reported (see Schmeller and Merilä 2007 for a review), indicating that the difference between effective and census size can be sometimes high. It has been suggested that estimates of N_e/N_c may be so low due to fluctuations in population size and variance in family size, as a result of e.g. environmental fluctuations, extreme winters and disease (Frankham 1995).

Implications for conservation

The genetic population structure and colonization dynamics we detected in our Pyrenean brook newt populations have important implications for the conservation and management of this endemic and near threatened amphibian at a local scale. Overall, although we found high levels of genetic structuring, we also documented considerable reproductive connectivity during the current generation and genetic admixture between several populations, except the most peripheral and isolated localities. Furthermore, this study reports the first evidence of amphibian recolonization of habitats formerly occupied by introduced fish in high mountain lakes using a genetic approach. Several studies exist on the recovery of amphibian species following invasive fish removal in high mountains (e.g. Knapp et al. 2005; Miró et al. 2020; Tiberti et al. 2019; Vredenburg 2004), but none had specifically focused on the characterization of recolonization routes using a combination of genetic and mark-recapture approaches. In our case, genetic tools proved more efficient and less time-consuming than mark-recapture in accurately characterizing *C. asper* patterns of dispersal, highlighting the higher sampling effort and dedication needed with more conventional approaches. Results also confirm the high capacity of *C. asper* to disperse and naturally colonize new suitable environments following reduction or elimination of introduced species pressure (Lucati et al. 2020a; Miró et al. 2020).

In light of the above, conservation measures should focus on the preservation of the genetic identity of high mountain populations and on the maintenance and if possible increase of connectivity between localities, especially the most isolated and fragmented. Fish eradication or intensive control have proved effective for this purpose, demonstrating that negative effects caused by non-native fish can be reversed (Bosch et al. 2019; Knapp et al. 2007; Ventura et al. 2017). However, eradication programmes should also undertake long-term monitoring to ensure native species recovery and protect restored lakes from future fish introductions, or maintain fish numbers low enough for amphibian survival.

Acknowledgements

We thank Meritxell Cases, Eloi Cruset, Ismael Jurado and Quim Pou-Rovira for field assistance, and Jenny Caner for laboratory assistance. Economic support was provided by the European Commission LIFE+ project LimnoPirineus (LIFE13 NAT/ES/001210) and by the Societas Europaea Herpetologica (SEH, research grant awarded to F.L.). F.L. had a doctoral grant funded by Fundação para a Ciência e a Tecnologia (FCT, grant number PD/BD/52598/2014).

References

Adriaensen F, Chardon J, De Blust G, Swinnen E, Villalba S, Gulinck H et al. (2003) The application of 'least-cost' modelling as a functional landscape model. *Landscape and Urban Planning* **64:** 233-247.

Baguette M, Blanchet S, Legrand D, Stevens VM, Turlure C (2013) Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews* **88:** 310-326.

Beier P, Spencer W, Baldwin RF, McRae BH (2011) Toward best practices for developing regional connectivity maps. *Conservation Biology* **25:** 879-892.

Bergl RA, Vigilant L (2007) Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*). *Molecular Ecology* **16:** 501-516.

Bijlsma R, Bundgaard J, Boerema A (2000) Does inbreeding affect the extinction risk of small populations?: Predictions from *Drosophila*. *Journal of Evolutionary Biology* **13:** 502-514.

Bosch J, Bielby J, Martin-Beyer B, Rincón P, Correa-Araneda F, Boyero L (2019) Eradication of introduced fish allows successful recovery of a stream-dwelling amphibian. *PloS One* **14:** e0216204.

Bosch J, Tejedo M, Lecis R, Miaud C, Lizana M, Edgar P et al. (2009) *Calotriton asper*. The IUCN Red List of Threatened Species 2009: e.T59448A11943040. doi:10.2305/IUCN.UK.2009.RLTS.T59448A11943040.en.

Bradford DF, Tabatabai F, Graber DM (1993) Isolation of remaining populations of the native frog, *Rana muscosa*, by introduced fishes in Sequoia and Kings Canyon National Parks, California. *Conservation Biology* **7:** 882-888.

Chan LM, Zamudio KR (2009) Population differentiation of temperate amphibians in unpredictable environments. *Molecular Ecology* **18:** 3185-3200.

Clergue-Gazeau M, Martínez-Rica J (1978) Les différents biotopes de l'urodèle pyrénéen, *Euproctus asper*. *Bulletin de la Société d'Histoire Naturelle de Toulouse* **114:** 461-471.

Davidson C, Knapp RA (2007) Multiple stressors and amphibian declines: Dual impacts of pesticides and fish on yellow-legged frogs. *Ecological Applications* **17:** 587-597.

Didham RK (2010) *Ecological consequences of habitat fragmentation.* Wiley: London, UK.

Drechsler A, Geller D, Freund K, Schmeller DS, Kuenzel S, Rupp O et al. (2013) What remains from a 454 run: Estimation of success rates of microsatellite loci development in selected

newt species (*Calotriton asper*, *Lissotriton helveticus*, and *Triturus cristatus*) and comparison with Illumina-based approaches. *Ecology and Evolution* **3:** 3947-3957.

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

Epanchin PN, Knapp RA, Lawler SP (2010) Nonnative trout impact an alpine-nesting bird by altering aquatic-insect subsidies. *Ecology* **91:** 2406-2415.

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.

Frankham R (1995) Effective population size/adult population size ratios in wildlife: A review. *Genetics Research* **66:** 95-107.

Frankham R, Ballou SEJD, Briscoe DA, Ballou JD (2002) *Introduction to conservation genetics.* Cambridge University Press: Cambridge, MA.

Gill DE (1978) The metapopulation ecology of the red-spotted newt, *Notophthalmus viridescens* (Rafinesque). *Ecological Monographs* **48:** 145-166.

Giordano AR, Ridenhour BJ, Storfer A (2007) The influence of altitude and topography on genetic structure in the long-toed salamander (*Ambystoma macrodactulym*). *Molecular Ecology* **16:** 1625-1637.

Gong W, Gu L, Zhang D (2010) Low genetic diversity and high genetic divergence caused by inbreeding and geographical isolation in the populations of endangered species *Loropetalum subcordatum* (Hamamelidaceae) endemic to China. *Conservation Genetics* **11:** 2281-2288.

Gu Z, Gu L, Eils R, Schlesner M, Brors B (2014) *circlize* implements and enhances circular visualization in R. *Bioinformatics* **30:** 2811-2812.

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; discussion 195.

Hurvich CM, Tsai CL (1989) Regression and time-series model selection in small samples. *Biometrika* **76:** 297-307.

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.

Kalinowski ST, Wagner AP, Taper ML (2006) ML-RELATE: A computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes* **6:** 576-579.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S et al. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28:** 1647-1649.

Knapp RA, Boiano DM, Vredenburg VT (2007) Removal of nonnative fish results in population expansion of a declining amphibian (mountain yellow-legged frog, *Rana muscosa*). *Biological Conservation* **135:** 11-20.

Knapp RA, Hawkins CP, Ladau J, McClory JG (2005) Fauna of Yosemite National Park lakes has low resistance but high resilience to fish introductions. *Ecological Applications* **15:** 835-847.

Lecis R, Pierpaoli M, Biro Z, Szemethy L, Ragni B, Vercillo F et al. (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology* **15:** 119-131.

Lucati F, Miró A, Ventura M (2020a) Conservation of the endemic Pyrenean newt (*Calotriton asper*) in the age of invasive species: Interlake dispersal and colonisation dynamics. *Amphibia-Reptilia* **1:** 1-2.

Lucati F, Poignet M, Miró A, Trochet A, Aubret F, Barthe L et al. (2020b) Multiple glacial refugia and contemporary dispersal shape the genetic structure of an endemic amphibian from the Pyrenees. *Molecular Ecology* **29:** 2904-2921.

Madsen T, Shine R, Olsson M, Wittzell H (1999) Restoration of an inbred adder population. *Nature* **402:** 34-35.

Martínez-Rica J, Clergue-Gazeau M (1977) Données nouvelles sur la répartition géographique de l'espèce *Euproctus asper* Dugès, Urodèle, Salamandridae. *Bulletin de la Société d'Histoire Naturelle de Tolouse* **113:** 318-330.

McCall BS, Mitchell MS, Schwartz MK, Hayden J, Cushman SA, Zager P et al. (2013) Combined use of mark-recapture and genetic analyses reveals response of a black bear population to changes in food productivity. *The Journal of Wildlife Management* **77:** 1572-1582.

McRae BH, Kavanagh DM (2011) Linkage mapper connectivity analysis software. The Nature Conservancy: Seattle, WA.

Milá B, Carranza S, Guillaume O, Clobert J (2010) Marked genetic structuring and extreme dispersal limitation in the Pyrenean brook newt *Calotriton asper* (Amphibia: Salamandridae) revealed by genome-wide AFLP but not mtDNA. *Molecular Ecology* **19:** 108-120.

Miró A, O'Brien D, Tomàs J, Buchaca T, Sabás I, Osorio V et al. (2020) Rapid amphibian community recovery following removal of non-native fish from high mountain lakes. *Biological Conservation* **251:** 108783.

Miró A, Sabás I, Ventura M (2018) Large negative effect of non-native trout and minnows on Pyrenean lake amphibians. *Biological Conservation* **218:** 144-153.

Miró A, Ventura M (2013) Historical use, fishing management and lake characteristics explain the presence of non-native trout in Pyrenean lakes: Implications for conservation. *Biological Conservation* **167:** 17-24.

Montero-Pau J, Gómez A, Muñoz J (2008) Application of an inexpensive and high-throughput genomic DNA extraction method for the molecular ecology of zooplanktonic diapausing eggs. *Limnology and Oceanography: Methods* **6:** 218-222.

Montori A (1988) Estudio sobre la biología y ecología del tritón pirenaico *Euproctus asper* (Dugès, 1852) en la Cerdanya. PhD thesis, University of Barcelona, Spain.

Montori A, Llorente GA (2014) Tritón pirenaico–*Calotriton asper* (Dugès, 1852). In: Salvador A and Martínez-Solano I (eds) *Enciclopedia Virtual de los Vertebrados.* Museo Nacional de Ciencias Naturales: Madrid, Spain.

Montori A, Llorente GA, García-París M (2008a) Allozyme differentiation among populations of the Pyrenean newt *Calotriton asper* (Amphibia: Caudata) does not mirror their morphological diversification. *Zootaxa* **1945:** 39-50.

Montori A, Llorente GA, Richter-Boix A (2008b) Habitat features affecting the small-scale distribution and longitudinal migration patterns of *Calotriton asper* in a Pre-Pyrenean population. *Amphibia-Reptilia* **29:** 371-381.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19:** 153-170.

Oromi N, Valbuena-Ureña E, Soler-Membrives A, Amat F, Camarasa S, Carranza S et al. (2018) Genetic structure of lake and stream populations in a Pyrenean amphibian (*Calotriton asper*) reveals evolutionary significant units associated with paedomorphosis. *Journal of Zoological Systematics and Evolutionary Research* **57:** 418-430.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: A simulation-based exploration of accuracy and power. *Molecular Ecology* **13:** 55-65.

Peakall R, Smouse P (2012) GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28:** 2537-2539.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95:** 536-539.

Pittman SE, Osbourn MS, Semlitsch RD (2014) Movement ecology of amphibians: A missing component for understanding population declines. *Biological Conservation* **169:** 44-53.

Pollock KH, Nichols JD, Brownie C, Hines JE (1990) Statistical inference for capture-recapture experiments. *Wildlife Monographs* **107:** 3-97.

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

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 9197-9201.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution* **43:** 223-225.

Ronce O (2007) How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics* **38:** 231-253.

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

Rousset F (2008) GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8:** 103-106.

Sánchez-Montes G, Wang J, Ariño AH, Martínez-Solano Í (2018) Mountains as barriers to gene flow in amphibians: Quantifying the differential effect of a major mountain ridge on the genetic structure of four sympatric species with different life history traits. *Journal of Biogeography* **45:** 318-331.

Sandel B, Arge L, Dalsgaard B, Davies R, Gaston K, Sutherland W et al. (2011) The influence of Late Quaternary climate-change velocity on species endemism. *Science* **334:** 660-664.

Savage WK, Fremier AK, Bradley Shaffer H (2010) Landscape genetics of alpine Sierra Nevada salamanders reveal extreme population subdivision in space and time. *Molecular Ecology* **19:** 3301-3314.

Schmeller DS, Merilä J (2007) Demographic and genetic estimates of effective population and breeding size in the amphibian *Rana temporaria*. *Conservation Biology* **21:** 142-151.

Schmitt T (2009) Biogeographical and evolutionary importance of the European high mountain systems. *Frontiers in Zoology* **6:** 9.

Seber G (1982) *The estimation of animal abundance and related parameters.* Griffin: London, UK.

Semlitsch RD (2008) Differentiating migration and dispersal processes for pond-breeding amphibians. *The Journal of Wildlife Management* **72:** 260-267.

Sugiura N (1978) Further analysts of the data by akaike's information criterion and the finite corrections: Further analysts of the data by akaike's. *Communications in Statistics-Theory and Methods* **7:** 13-26.

Takezaki N, Nei M, Tamura K (2014) POPTREEW: Web version of POPTREE for constructing population trees from allele frequency data and computing some other quantities. *Molecular Biology and Evolution* **31:** 1622-1624.

Tallmon DA, Luikart G, Waples RS (2004) The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology & Evolution* **19:** 489-496.

Tiberti R, Bogliani G, Brighenti S, Iacobuzio R, Liautaud K, Rolla M et al. (2019) Recovery of high mountain Alpine lakes after the eradication of introduced brook trout *Salvelinus fontinalis* using non-chemical methods. *Biological Invasions* **21:** 875-894.

Trenham PC, Koenig WD, Shaffer HB (2001) Spatially autocorrelated demography and interpond dispersal in the salamander *Ambystoma californiense*. *Ecology* **82:** 3519-3530.

Turner MG (2005) Landscape ecology: What is the state of the science? *Annual Review of Ecology, Evolution, and Systematics* **36:** 319-344.

Vähä JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15:** 63-72.

Valbuena-Ureña E, Oromi N, Soler-Membrives A, Carranza S, Amat F, Camarasa S et al. (2018) Jailed in the mountains: Genetic diversity and structure of an endemic newt species across the Pyrenees. *PLoS One* **13:** e0200214.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4:** 535-538.

Van Tienhoven AM, Den Hartog JE, Reijns RA, Peddemors VM (2007) A computer-aided program for pattern-matching of natural marks on the spotted raggedtooth shark *Carcharias taurus*. *Journal of Applied Ecology* **44:** 273-280.

Ventura M, Tiberti R, Buchaca T, Buñay D, Sabás I, Miró A (2017) Why should we preserve fishless high mountain lakes? In: Catalan J, Ninot J and Aniz M (eds) *Advances in Global Change Research.* Springer: Cham, ZG. Vol. 62: High mountain conservation in a changing world, pp 181-205.

Vredenburg VT (2004) Reversing introduced species effects: Experimental removal of introduced fish leads to rapid recovery of a declining frog. *Proceedings of the National Academy of Sciences* **101:** 7646-7650.

Wang IJ, Shaffer HB (2017) Population genetic and field-ecological analyses return similar estimates of dispersal over space and time in an endangered amphibian. *Evolutionary Applications* **10:** 630-639.

Wang J (2017) The computer program STRUCTURE for assigning individuals to populations: Easy to use but easier to misuse. *Molecular Ecology Resources* **17:** 981-990.

Washington Wildlife Habitat Connectivity Working Group (WHCWG) (2010) *Washington Connected Landscapes Project: Statewide analysis.* Washington Departments of Fish and Wildlife, and Transportation: Olympia, WA, USA.

White GC, Burnham KP (1999) Program MARK: Survival estimation from populations of marked animals. *Bird Study* **46:** 120-139.

Zarnetske PL, Baiser B, Strecker A, Record S, Belmaker J, Tuanmu M-N (2017) The interplay between landscape structure and biotic interactions. *Current Landscape Ecology Reports* **2:** 12-29.

Supplementary material

Supplementary tables

Table 4.S1 Resistance values of elevation, slope and fish presence used in Linkage Mapper for least-cost path analysis.

Feature	Category	Resistance values
Elevation	841-3027 m	1-1000
Slope	40º-60º	400
	> 60º	600
Fish presence	No fish	1
	Land	400
	Fish	800

Table 4.S2 Least-cost path lengths between sampled sites as inferred by Linkage Mapper. See Table 4.1 for population codes.

From	To	Least-cost path (m)
CPP	CTS	8445
CPP	IDE	7935
CPP	SUB	5657
CPP	MDE	8876
CPP	NDE	8980
CPP	GDE	7274
CPP	BPD	8840
CPP	BFD	9039
CPP	XSU	5041
CPP	RQN	7399
CPP	CPX	6671
CPP	RSU	5071
CPP	RSE	5147
CTS	IDE	626
CTS	SUB	3371
CTS	MDE	712
CTS	NDE	816
CTS	GDE	1171
CTS	BPD	676
CTS	BFD	875
CTS	XSU	3123
CTS	RQN	5530
CTS	CPX	1493
CTS	RSU	3273

CTS	RSE	3252
IDE	SUB	3997
IDE	MDE	120
IDE	NDE	204
IDE	GDE	661
IDE	BPD	50
IDE	BFD	288
IDE	XSU	3749
IDE	RQN	5393
IDE	CPX	2119
IDE	RSU	3899
IDE	RSE	3878
MDE	NDE	84
MDE	GDE	541
MDE	BPD	70
MDE	BFD	168
MDE	XSU	3835
MDE	RQN	5273
MDE	CPX	2205
MDE	RSU	3985
MDE	RSE	3964
NDE	GDE	457
NDE	BPD	154
NDE	BFD	84
NDE	XSU	3939
NDE	RQN	5357
NDE	CPX	2309
NDE	RSU	4089
NDE	RSE	4068
GDE	BPD	611
GDE	BFD	373
GDE	XSU	3979
GDE	RQN	4359
GDE	CPX	2701
GDE	RSU	4009
GDE	RSE	4030
BPD	BFD	238
BPD	XSU	3799
BPD	RQN	5343
BPD	CPX	2169
BPD	RSU	3949
BPD	RSE	3928
BFD	XSU	3998
BFD	RQN	4732
BFD	CPX	2368
BFD	RSU	1780
BFD	RSE	1759

SUB	MDE	4083
SUB	NDE	4187
SUB	GDE	4045
SUB	BPD	4047
SUB	BFD	4246
SUB	XSU	66
SUB	RQN	8404
SUB	CPX	1878
SUB	RSU	15
SUB	RSE	36
XSU	RQN	12440
XSU	CPX	1630
XSU	RSU	30
XSU	RSE	106
RSU	RSE	21
RQN	CPX	7060
RQN	RSU	8840
RQN	RSE	8819
CPX	RSU	1780
CPX	RSE	1759

Table 4.S3 Estimates of adult census population size of *Calotriton asper* at two high mountain lakes (XSU and NDE) for the different capture occasions. Standard errors are given in parentheses. See Table 4.1 for population codes.

Capture occasion	Census population size	
	XSU	NDE
1	513 (22.20)	120 (12.28)
2	692 (25.90)	140 (13.56)
3	230 (14.90)	87 (10.04)
4	1487 (24.50)	73 (9.06)
5	1333 (36.00)	47 (6.83)
6	897 (29.95)	14 (3.65)
Mean	859 (196.96)	80 (18.92)

Supplementary figures

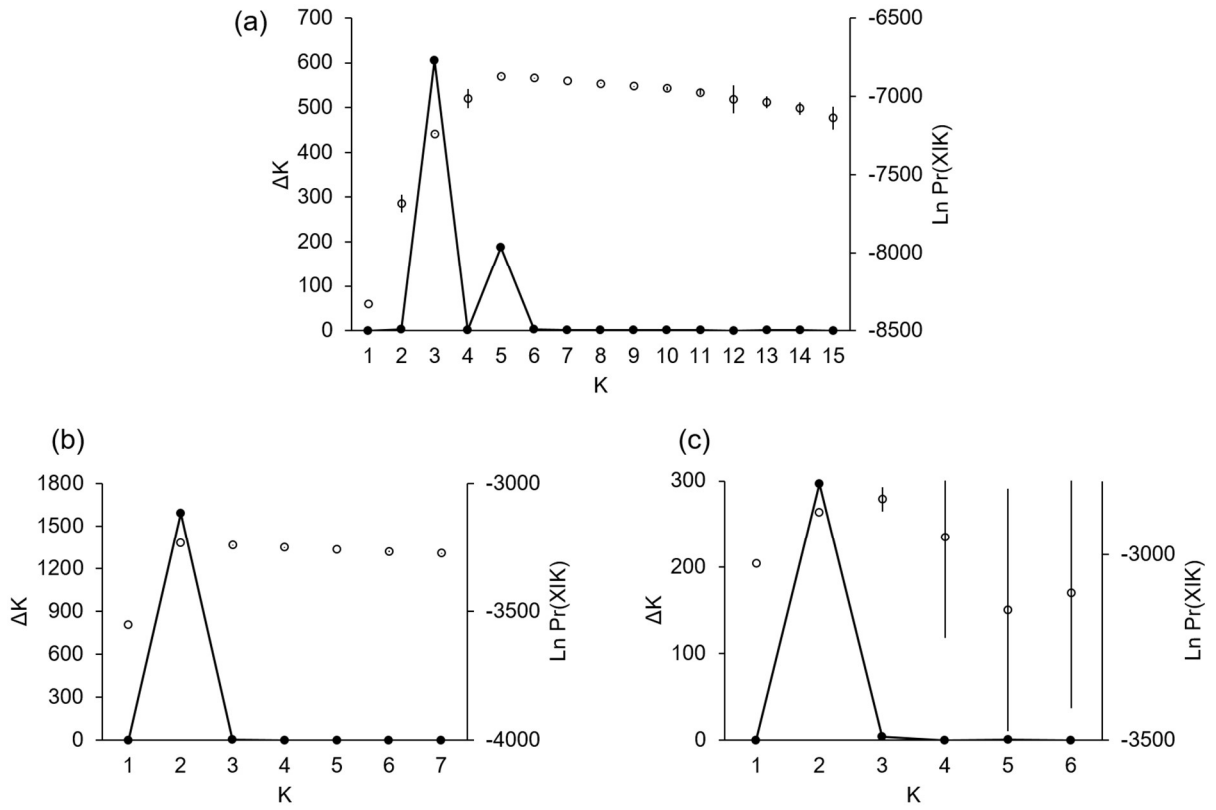


Figure 4.S1 Results of Bayesian clustering analyses in STRUCTURE for the total dataset (a), cluster 1 (western populations, b) and cluster 2 (eastern populations, c). The right axis (open dots with error bars) displays mean (± SD) log probability of the data [Ln Pr(XIK)] over 20 runs, for each value of K; the left axis (black dots) shows ΔK values as a function of K, calculated according to Evanno et al. (2005).

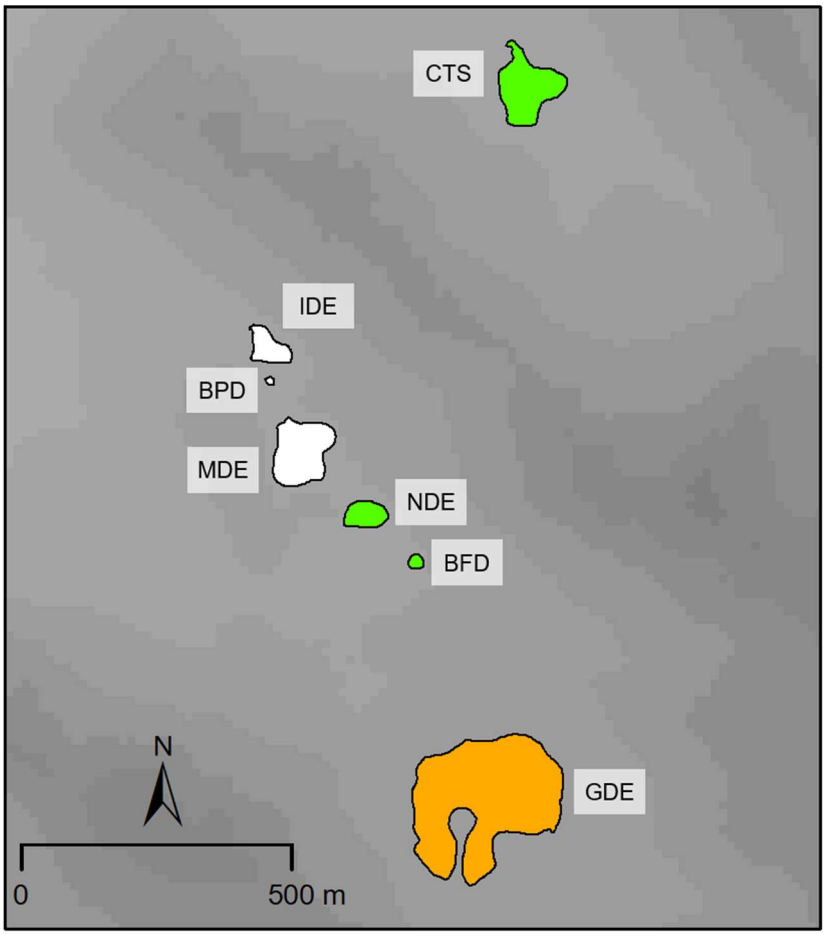


Figure 4.S2 Map showing the lakes located in the valley of Dellui and the adjacent valley of Corticelles (CTS). Altitude (meters) is based on a gray scale, where darker colours indicate higher elevations. Fishless lakes are coloured in green, fish-removal lakes in white and the only stocked lake in orange. For population codes see Table 4.1.

References

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.

Chapter 5

General discussion

Unveiling how genetic diversity and structure of high mountain populations evolved and what factors are currently shaping it is crucial to predict how species will respond to threats such as global change, and this will ultimately help to design species-specific conservation measures. Mountains are among the regions most affected by climate change, as testified by the shrinkage of glaciers, the change in precipitation patterns and the increased frequency and magnitude of extreme events (Foggin 2016; Kohler et al. 2010). For several montane populations, this will translate into either upslope migrations in search for new suitable habitats, range contractions or local extinctions (Spehn et al. 2010). Therefore, an integrative approach that combines the study of phylogeographic processes and contemporary dispersal dynamics is required to shed light on the mechanisms underlying spatial patterns of present-day genetic diversity and population structure, which can help to predict species responses to ongoing or future environmental changes.

This thesis contributes to improving our understanding of the factors responsible for species structuring in high mountains (Figure 5.1). To achieve this, I addressed fundamental research questions about the contribution of Pleistocene glacial cycles and contemporary dispersal on shaping the genetic composition of montane populations, and the factors and threats that may be currently responsible for promoting regional genetic structuring, such as invasive species and emergent diseases. I tackled a variety of topics using a combination of indirect (genetic) and direct (field-based) approaches. I used genetic methods to provide new insights into the phylogeographic history of the study species (the Pyrenean brook newt *Calotriton asper*, and the midwife toads *Alytes obstetricans*/*almogavarii*), inform on current connectivity and dispersal in the Pyrenean brook newt and describe the genetic consequences of chytridiomycosis outbreaks in midwife toads. As for the Pyrenean brook newt, I also combined genetic and field-based (mark-recapture) analyses to target the role of dispersal at small spatial scales and characterize the process of lake recolonization following invasive fish removal.

In this chapter, I summarise the most important findings of this thesis and incorporate my results into a wider perspective, while also analysing the conservation implications for amphibians living in high mountains. Finally, I discuss study limitations and suggest new avenues for future research.

Figure 5.1 Some of the study sites. From top-left to bottom-right: Pozo Llau (Picos de Europa National Park, photo by Jaime Bosch), Laguna de los Pájaros and Laguna Grande de Peñalara (Guadarrama National Park - Central System, photos by Jaime Bosch), Dellui Mig and Tres Estanys (eastern Pyrenees, photos by Jenny Caner and Víctor Osorio).

Phylogeography and past population dynamics

Genetic data convey information about the evolutionary history of populations from which they were obtained. In fact, genetic analyses have been crucial for elucidating the historical processes that shaped the geographic distribution of species and revealing their population structure (e.g. Bisconti et al. 2018; Ferchaud et al. 2015; Lourenço et al. 2018; Mouret et al. 2011). Specifically, phylogeographic studies have as main focus the analysis and interpretation of historical factors that led to the current geographic distribution of species and genetic lineages (Avise 2000). In the last few years, new computational methods such as Approximate Bayesian Computation (ABC) have developed as a new tool for inference on the past history of populations and species (Beaumont and Rannala 2004; Cornuet et al. 2010; Cornuet et al. 2008). This has led to great advances in the field of phylogeography, making it possible to better delineate the complex demographic processes that have acted on populations by comparing competing evolutionary scenarios and quantifying their relative support. Consequently, we are now able to infer the size of ancestral populations, the approximate dates of population divergence and the occurrence of bottlenecks or population expansions with greater precision, by combining DNA sequence and microsatellite data.

As for *C. asper*, previous studies based on morphological, allozymic and genetic data showed contrasting levels of genetic structuring across the Pyrenees, with allozymic and mtDNA data revealing low levels of genetic variation and morphological, AFLP and microsatellite data denoting higher levels of differentiation (Milá et al. 2010; Montori et al. 2008; Valbuena-Ureña et al. 2013; Valbuena-Ureña et al. 2018). Putative glacial refugia where the species survived Pleistocene glaciation events in the Pyrenees were previously postulated (Carranza and Amat 2005; Valbuena-Ureña et al. 2018), but several basic questions, such as the location and number of refugia and the process of postglacial colonization, remained poorly understood. In chapter 2, I corroborated previous studies as for the strong spatial structure revealed by microsatellites, while also reconstructing the demographic history of the species. ABC analyses supported the presence of five different glacial refugia in the Pyrenees: differentiation started shortly before the Last Glacial Maximum (~42,000–24,000 years ago) at three focal areas (western, central and eastern Pyrenees) and continued through the end of the Last Glacial Period in the central Pyrenees until ~12,000–5,500 years ago, leading to the formation of a total of five well-supported genetic lineages. Cyclic cooling periods characterized by recurrent events of population contraction, downhill retreat and limited dispersal prompted episodes of genetic differentiation that eventually determined the present-day structure revealed in *C. asper*.

With regard to *A. obstetricans*/*almogavarii*, high levels of intraspecific diversity and population structure were reported in prior studies from the Iberian Peninsula using a range of morphological, allozymic and genetic data, which led to the description of four parapatric subspecies: *A. o. obstetricans*, distributed across western Europe and northern Spain, *A. o. pertinax*, present in the central and eastern areas of the Iberian Peninsula, *A. o. boscai*, present in northern and central Portugal and adjacent areas in Spain, and *A. o. almogavarii*, endemic to north-eastern Spain and south-eastern France (Arntzen and Garcia-Paris 1995; Gonçalves et al. 2015; Gonçalves et al. 2007; Martínez-Solano et al. 2004). Recently, the latter taxon was recognized as a distinct species (*A. almogavarii*; Dufresnes and Martínez-Solano 2019). Nevertheless, doubts remain regarding how the different Iberian *A. obstetricans*/*almogavarii* lineages responded to Pleistocene glacial cycles. Particularly, no prior study was specifically focused on high mountains, which are well-known hotspots of genetic diversity and have been shown to play a considerable role in separating well-differentiated intraspecific clades in numerous species (e.g. Pereira et al. 2016; Sánchez-Montes et al. 2018). My results, provided in chapter 3, revealed higher levels of intraspecific genetic structuring than previously detected,

which were correlated with geography, and complex patterns of lineage admixture, mito-nuclear discordance and secondary contact in Iberian high mountain populations of *A. obstetricans*/*almogavarii*. Nevertheless, we cannot discard other factors, such as stochastic rather than geographic processes, as responsible for some of the patterns we found. However, results indicate that each analysed mountain region (i.e. the Pyrenees, Picos de Europa and Guadarrama Mountains) likely underwent a peculiar phylogeographic history through the Late Pleistocene, which is consistent with the "refugia within refugia" model (Gómez and Lunt 2007).

Results provided in chapters 2 and 3 add to a growing body of literature highlighting the role of the Iberian Peninsula as one of the most important Pleistocene glacial refugia and hotspot of biological diversity and endemism (Gómez and Lunt 2007; Schmitt 2009; Taberlet et al. 1998; Wallis et al. 2016). It emerges the importance of this region as evolutionary hotspot of Iberian amphibians, where multiple intra-specific lineages and contact zones co-occur (Dufresnes et al. 2020; Gonçalves et al. 2015; Recuero and Garcia-Paris 2011; Vences et al. 2017).

***A. obstetricans*/*almogavarii*: how many species are there?**

The contact zones between different lineages of a species complex may reflect different processes. The outcome along different contact zones may vary, with gene flow being high in some contact zones and absent in others. Previous studies have evidenced that the subspecies *A. o. almogavarii*, which is distributed in Catalonia and adjacent areas in north-eastern Spain and southern France, is highly differentiated in allozyme (Arntzen and Garcia-Paris 1995) and genetic markers (Gonçalves et al. 2015; Maia-Carvalho et al. 2018), and presents peculiar osteological (Martínez-Solano et al. 2004) and bioacoustic characters (Márquez and Bosch 1995). In the course of this PhD project, *A. o. almogavarii*, which is represented by lineages E and F, was recognized as a distinct species (*A. almogavarii*) based on the analysis of the hybrid zone between *A. o. almogavarii* and *A. o. pertinax* along the Catalonian coast in north-eastern Spain, using RAD-sequencing analysis (Dufresnes and Martínez-Solano 2019; Speybroeck et al. 2020).

My results, provided in chapter 3, allow for discussion on the validity of *A. almogavarii* as a different species. At lineage borders, I detected signs of contemporary connectivity between *A. almogavarii* lineage E and *A. o. obstetricans* in the western Pyrenees using both microsatellites

and mitochondrial DNA, which confirmed previous results based on allozymes (Arntzen and Garcia-Paris 1995; García-París 1995). In contrast, *A. almogavarii* lineage F showed no signs of admixture at lineage borders with *A. o. obstetricans* in the North or *A. o. pertinax* in the South, in line with previous studies indicating no admixture between this lineage and *A. o. obstetricans* (Gonçalves et al. 2015), and weak connectivity with *A. o. pertinax* (Dufresnes and Martínez-Solano 2019; Gonçalves et al. 2015). Furthermore, my results highlight the genetic distinctiveness of Pyrenean populations, which are genetically more related to each other than with any of the other analysed high mountain regions (based on microsatellites).

This scenario of different contact zones telling different stories is exemplified in ring species, i.e. a system formed by a region of interconnected populations with both ends of the ring coming into contact without apparent admixture (Alcaide et al. 2014; Irwin et al. 2001a; Irwin et al. 2001b). In the Pyrenees, the *A. obstetricans* complex displays the characteristics of a ring species, where taxa at the western part of the ring likely interbreed (*A. almogavarii* lineage E and *A. o. obstetricans*) and those at the eastern side don't (*A. almogavarii* lineage F and *A. o. obstetricans*), displaying a continuum from slightly divergent neighbouring populations to substantially reproductively isolated taxa. Ring species represent cases of speciation in action and have been cited as evidence of evolution (Irwin et al. 2001b), with some of the most well-known cases being identified in amphibians, such as the *Ensatina eschscholtzii* salamander complex of western North America (Moritz et al. 1992). In light of this, my results are not entirely in line with the distinction of *A. almogavarii* as a different species, although lineage F shows evidence for speciation and reproductive isolation, and further studies are needed to complete our understanding of the *A. obstetricans* complex in the Pyrenees.

Gene flow and dispersal

Understanding a species' distribution across the landscape and the connectivity among geographically discrete populations is fundamental for its conservation and management. Amphibians are generally considered to be poor dispersers with philopatric tendencies (Gamble et al. 2007; Gill 1978), leading to phylogeographic breaks and deeply subdivided populations (Vences and Wake 2007). Furthermore, the harsh terrestrial conditions and dramatic seasonal climate changes typical of high mountains are factors affecting dispersal (Sánchez-Montes et al. 2018; Savage et al. 2010). However, not all amphibians are strictly philopatric, as testified

by some species being fairly quick to colonize newly available habitat patches (Pechmann et al. 2001; Semlitsch 2008).

In this thesis, I addressed gene flow and dispersal in the Pyrenean brook newt at both broad (chapter 2) and small (chapter 4) spatial scales. In chapter 2, although I found that the majority of dispersal events involved geographically close populations (0-4 km) mostly belonging to the same valley, I also found potential for occasional long-distance dispersal (up to 33 km) using microsatellites. In chapter 4, I detected high levels of genetic differentiation and population structuring in a protected area in the central-eastern Pyrenees, with modest levels of short-distance dispersal and admixture between individuals suggesting reproductive connectivity between several populations, except the most isolated localities. The short mean dispersal distances, coupled with low effective population sizes ($N_e < 50$), may explain the high levels of genetic structuring for *C. asper* populations. On the other hand, rare long-distance dispersal and intervalley movements are in line with my estimates of genetic diversity, as shown by most populations presenting low inbreeding coefficients and levels of genetic variability within the range of other urodeles and temperate amphibians (Chan and Zamudio 2009). Furthermore, *C. asper* proved to be a quick colonizer when new fishless lakes within close proximity of existing ponds become available (chapter 4, see also Miró et al. 2020). Taken together, results suggest that *C. asper* is moderately vagile and not strictly philopatric. It has been suggested that colonizing newly available ponds may be associated with high fitness of adults changing ponds or for successfully dispersing juveniles breeding for the first time, through enhanced offspring performance (Semlitsch 2008). In *C. asper*, paedomorphic (when larval traits, such as gills, are retained in the adult stage) and cave-dwelling populations have also been occasionally reported (e.g. Guillaume et al. 2020; Oromi et al. 2018). In such cases, dispersal may be extremely limited, if not absent.

Habitat fragmentation in high mountains

Species are generally negatively affected by habitat fragmentation, when suitable habitat patches became too narrow and isolated (Fahrig 2003; Fischer and Lindenmayer 2007). Such fragmentation, which can have either natural or anthropogenic causes, may lead to loss of fitness through inbreeding or local extinction through stochastic effects. Negative effects caused by fragmentation may be stronger in species with restricted distribution ranges (Sodhi et al. 2008), such as *C. asper* and the different lineages of *A. obstetricans*/*almogavarii*. In high

mountains, in addition to natural spatial patterns, the degree of isolation among populations is markedly increased by introduced fish and emergent diseases (Bradford et al. 1993; Walker et al. 2010).

In general, high mountain lake fauna has low resistance but high resilience to fish introductions (Knapp et al. 2001). In fact, a rapidly accumulating literature indicates that amphibian populations are capable of recovering quickly after invasive fish eradication or intensive control (e.g. Bosch et al. 2019; Knapp et al. 2007; Miró et al. 2020). In chapter 4, I showed how removal of invasive fish allowed not only for the recovery of declining amphibians, but also for local recolonization of sites where *C. asper* was considered extinct. In the same chapter, using a combination of genetic and mark-recapture analyses, I also showed that fragments of non-invaded habitats such as satellite or neighbouring ponds function as reservoirs in maintaining newt populations, and serve as sources of individuals as invasive fish pressure is reduced. Previous research on this topic was scarce and relied exclusively on field-based data, such as visual counts (Denoël et al. 2016; Tiberti 2018). Another interesting finding is that, according to genetic and least-cost path analyses, colonizers did not necessarily all come from the closest fishless lake. Taken together, our data suggest that geographic distance may be driving *C. asper* recolonization routes, but that other factors are also contributing to the patterns of lake colonization, such as the structure of the invaded landscape. Indeed, fishless lakes surrounded by stocked water bodies are probably less likely to function as reservoirs for restored lakes than fishless areas located at a greater distance, where newts are free to wander widely over the landscape without being eaten by fish.

The chytrid fungus *Batrachochytrium dendrobatidis* (*Bd*), which is responsible for causing the disease chytridiomycosis, is widely recognized as a proximate driver of amphibian declines across the globe (O'Hanlon et al. 2018; Skerratt et al. 2007). *A. obstetricans*/*almogavarii* is known to be highly susceptible to the pathogen, which represents a potential risk to the species across the entire Iberian Peninsula (Walker et al. 2010). In chapter 3, in line with previous findings (Albert et al. 2014), I found that *A. obstetricans* populations hit by a relatively recent chytridiomycosis outbreak have a reduced effective size and may be genetically depleted to face threatening events such as introduction of invasive species or climate change. More than 10 years have passed since the outbreak in the Guadarrama Mountain Range and affected populations still bear all the hallmarks of the disease, as evidenced by both mtDNA and

microsatellite markers. In high mountains, which are associated with severe topographical, climatic and ecological conditions, suitable habitat for amphibians is usually patchy and populations are generally more isolated and fragmented than in lower elevation areas (Giordano et al. 2007). As a consequence, breeding demes may often be small and dispersal may be limited. Under these conditions, recovery of amphibian populations hit by disease outbreaks may take several decades, and leave populations exposed to inbreeding and genetic drift. It is therefore crucial to improve connectivity between populations that are still in migratory contact and attempt to restore connectivity where it has been lost.

Implications for conservation

Due to their peculiar climatic and geological history, mountains are a recognised hotspot of genetic diversity, being home to endemic refugial clades in a number of amphibian species (e.g. Dufresnes et al. 2020; Recuero and Garcia-Paris 2011; Vences et al. 2017). Conservation efforts in high mountains should thus focus on the preservation of such genetic variation, especially in the case of endemic and/or highly structured species, such as *C. asper* and *A. obstetricans*/*almogavarii*, which are generally more vulnerable to a wide variety of threats.

In chapters 3 and 4 I showed that, even at small spatial scales, amphibian populations can be highly structured. This has important implications when planning translocations of individuals or reintroductions of captive bred animals for conservation purposes, as it is extremely important to avoid genetic risks concerned with e.g. inbreeding depression (reduced biological fitness as a result of breeding of related individuals) or outbreeding depression (when crosses between two genetically distant groups result in a reduction of fitness). Translocations are being increasingly proposed as a way of maintaining biodiversity and ecosystem function, as habitat loss, introduced species, disease epidemics and climate change are threatening species worldwide (Joseph and Knapp 2018; Weeks et al. 2011). Within the context of the conservation project LIFE+ LIMNOPIRINEUS, from which this thesis stems, translocations of adult Pyrenean brook newts were initially proposed to facilitate the colonization of lakes following fish removal, because it was unclear whether newts were capable of recolonizing restored lakes alone. It has been argued that the appropriate selection of the source of founder individuals for future translocations or reintroductions may determine the success of the program (Seddon et al. 2007). For this reason, genetic studies prior to the management of target populations have increased in number, including for the conservation of threatened amphibians (Albert et al.

2014; Kraaijeveld-Smit et al. 2005). Therefore, the information presented in this thesis may be useful to guide future management decisions for *C. asper* and *A. obstetricans*/*almogavarii*, and in marking paths for future investigation.

Study limitations and future research directions

In this thesis I used genetic markers such as microsatellites and mitochondrial and nuclear gene fragments to unravel the evolutionary history of the Pyrenean brook newt and midwife toads of the *A. obstetricans* complex; I also used genetic markers in combination with photographic mark recapture data to inform on connectivity and dispersal in the Pyrenean brook newt at both broad and small spatial scales. However, some specific points remain to be addressed. Here I discuss directions for future research that could lead to an even more effective understanding of the topics addressed in this dissertation.

In ABC modelling, gene flow was not taken into account, and this may have limited inference power. This is an intrinsic limitation of the software I used, i.e. DIYABC 2.1.0, where no gene flow is permitted between populations after they have diverged (Cornuet et al. 2010). Only single events of admixture between populations are considered, whereas recurrent gene flow due to dispersal cannot be incorporated. However, these limitations may be overcome by calculating past migration rates and incorporating them to the models, which can be done using programs such as MIGRATE-N (Beerli 2009; Beerli and Felsenstein 1999) or the software series IM/IMa (e.g. IMa3; Hey et al. 2018). Furthermore, the tested models do not represent a comprehensive range of all possible scenarios, but are instead based on a selection of hypotheses that I considered were most likely to reflect my data. I focused on simple contrasting models aimed at capturing the key demographic events, avoiding overcomplex and similar models, as suggested by Cabrera and Palsbøll (2017).

Some of the markers used have limited sample size. This is especially true for three of the four mitochondrial fragments (cyt-*b*, 12S and 16S) and for the nuclear gene (β-fibint7) analysed in chapter 3. Sequencing more samples, as well as incorporating more nuclear genes, would strengthen the study and its conclusions. Furthermore, in the last few years the application of next-generation sequencing (NGS) approaches has revolutionised sequencing capabilities, providing higher throughput data with lower cost. In future studies, the reassessment of population structure and evolutionary history of the target species through the use of molecular

markers with greater variability, such as single nucleotide polymorphisms (SNPs), could provide more detailed and deeper insights into the description of evolutionary lineages, mito-nuclear discordances and introgression.

Further morphological and genetic analyses should be carried out to integrate and support my genetic results, especially in the case of *A. obstetricans*/*almogavarii*. ND4 lineage E is currently considered as part of *A. almogavarii*, although it has not yet been described morphologically. This lineage is genetically distinct and divergent from *A. almogavarii* lineage F, as evidenced by results provided in chapter 3 and by previous studies using both mitochondrial and microsatellite markers (Gonçalves et al. 2015; Maia-Carvalho et al. 2018). Furthermore, in chapter 3 I detected contemporary signs of connectivity between lineage E and *A. o. obstetricans* in the western Pyrenees, as indicated by the occurrence of a contact zone at lineage borders. The suggestion of *A. almogavarii* as a different species is quite recent and is mainly based on the analysis of the hybrid zone between lineage F and *A. o. pertinax* (Dufresnes and Martínez-Solano 2019). Morphological characterization of populations of lineage E and comparison with lineage F, as well as additional surveys across hybrid zones between lineage E and *A. o. obstetricans* would allow testing whether signs of incipient speciation are also found in lineage E.

To address some of the questions raised in chapter 4 of this thesis more fully would require additional sampling, during more years and from more populations. The low number of recaptures did not allow me to deeply characterize *C. asper* recent movement patterns using photographic mark-recapture. Also, the low proportion of recaptured individuals led to estimated population sizes with large variation. This is an indication that a much higher sampling effort is required to disentangle the species' dispersal dynamics and return precise estimates of census population size using this approach. With regard to genetic data, it would be interesting to examine the dynamics of colonization of new ecosystems over a longer period of time (i.e. for three or more consecutive years), targeting more restored lakes and amphibian species, such as the common frog *Rana temporaria* and the palmate newt *Lissotriton helveticus*.

Side projects

In the course of this PhD project, I had the opportunity to collaborate on a number of side projects related to the conservation of high mountain species, which resulted in scientific

publications and book chapters that I have authored or co-authored. Chapter 4 stemmed from the SEH Grant in Herpetology for 2018 (Appendix 1; Lucati et al. 2020), an annual award granted by the Societas Europaea Herpetologica to support a project oriented to the conservation of amphibians and reptiles in Europe. Other scientific contributions include: peer-reviewed articles on the recovery of amphibian communities following removal of non-native fish (Miró et al. 2020) and on the detection of chytrid fungi in *C. asper* (Martínez-Silvestre et al. 2020), and book chapters on the changes in lakes following reduction of fish pressure (Buchaca et al. 2019) and on the conservation status of the Pyrenean sculpin (*Cottus hispaniolensis*) in the central Pyrenees (Rocaspana et al. 2019).

References

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.

Alcaide M, Scordato ES, Price TD, Irwin DE (2014) Genomic divergence in a ring species complex. *Nature* **511:** 83-85.

Arntzen JW, Garcia-Paris 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, MA, USA.

Beaumont MA, Rannala B (2004) The Bayesian revolution in genetics. *Nature Reviews Genetics* **5:** 251-261.

Beerli P (2009) How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use. *Population Genetics for Animal Conservation* **17:** 42-79.

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* **152:** 763-773.

Bisconti R, Porretta D, Arduino P, Nascetti G, Canestrelli D (2018) Hybridization and extensive mitochondrial introgression among fire salamanders in peninsular Italy. *Scientific Reports* **8:** 13187.

Bosch J, Bielby J, Martin-Beyer B, Rincón P, Correa-Araneda F, Boyero L (2019) Eradication of introduced fish allows successful recovery of a stream-dwelling amphibian. *PLoS One* **14:** e0216204.

Bradford DF, Tabatabai F, Graber DM (1993) Isolation of remaining populations of the native frog, *Rana muscosa*, by introduced fishes in Sequoia and Kings Canyon National Parks, California. *Conservation Biology* **7:** 882-888.

Buchaca T, Sabas I, Osorio V, Pou-Rovira Q, Miró A, Puig MA et al. (2019) Changes in lakes after the reduction of fish densities. In: Carrillo E, Ninot J, Buchaca T and Ventura M

(eds) *Life+ Limnopirineus: Conservation of aquatic habitats and species in high mountains of the Pyrenees. Technical report. Blanes, Life+ Limnopirineus Technical Office: 29-40.* ISBN: 978-84-18199-17-2.

Cabrera AA, Palsbøll PJ (2017) Inferring past demographic changes from contemporary genetic data: A simulation-based evaluation of the ABC methods implemented in DIYABC. *Molecular Ecology Resources* **17:** e94-e110.

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

Chan LM, Zamudio KR (2009) Population differentiation of temperate amphibians in unpredictable environments. *Molecular Ecology* **18:** 3185-3200.

Cornuet JM, Ravigne V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* **11:** 401.

Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin JM, Balding DJ et al. (2008) Inferring population history with DIYABC: A user-friendly approach to approximate Bayesian computation. *Bioinformatics* **24:** 2713-2719.

Denoël M, Scimè P, Zambelli N (2016) Newt life after fish introduction: Extirpation of paedomorphosis in a mountain fish lake and newt use of satellite pools. *Current Zoology* **62:** 61-69.

Dufresnes C, Martínez-Solano Í (2019) Hybrid zone genomics supports candidate species in Iberian *Alytes obstetricans*. *Amphibia-Reptilia* **41:** 105-112.

Dufresnes C, Nicieza AG, Litvinchuk SN, Rodrigues N, Jeffries DL, Vences M et al. (2020) Are glacial refugia hotspots of speciation and cyto-nuclear discordances? Answers from the genomic phylogeography of Spanish common frogs. *Molecular Ecology* **29:** 986-1000.

Fahrig L (2003) Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics* **34:** 487-515.

Ferchaud AL, Eudeline R, Arnal V, Cheylan M, Pottier G, Leblois R et al. (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.

Fischer J, Lindenmayer DB (2007) Landscape modification and habitat fragmentation: A synthesis. *Global Ecology and Biogeography* **16:** 265-280.

Foggin J (2016) *Conservation issues: Mountain ecosystems. Reference module in Earth systems and environmental sciences*. Elsevier: Toronto, Canada.

Gamble LR, McGarigal K, Compton BW (2007) Fidelity and dispersal in the pond-breeding amphibian, *Ambystoma opacum*: Implications for spatio-temporal population dynamics and conservation. *Biological Conservation* **139:** 247-257.

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

Gill DE (1978) The metapopulation ecology of the red-spotted newt, *Notophthalmus viridescens* (Rafinesque). *Ecological Monographs* **48:** 145-166.

Giordano AR, Ridenhour BJ, Storfer A (2007) The influence of altitude and topography on genetic structure in the long-toed salamander (*Ambystoma macrodactulym*). *Molecular Ecology* **16:** 1625-1637.

Gómez A, Lunt DH (2007) Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S and Ferrand N (eds) *Phylogeography of southern European refugia*. Springer: Amsterdam, Netherlands, pp 155-188.

Gonçalves H, Maia-Carvalho B, Sousa-Neves T, Garcia-Paris M, Sequeira F, Ferrand N et al. (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.

Guillaume O, Deluen M, Raffard A, Calvez O, Trochet A (2020) Reduction in the metabolic levels due to phenotypic plasticity in the Pyrenean newt, *Calotriton asper*, during cave colonization. *Ecology and Evolution* **10:** 12983–12989.

Hey J, Chung Y, Sethuraman A, Lachance J, Tishkoff S, Sousa VC et al. (2018) Phylogeny estimation by integration over isolation with migration models. *Molecular Biology and Evolution* **35:** 2805-2818.

Irwin DE, Bensch S, Price TD (2001a) Speciation in a ring. *Nature* **409:** 333-337.

Irwin DE, Irwin JH, Price TD (2001b) Ring species as bridges between microevolution and speciation. *Genetica* **112-113:** 223-243.

Joseph MB, Knapp RA (2018) Disease and climate effects on individuals drive post-reintroduction population dynamics of an endangered amphibian. *Ecosphere* **9:** e02499.

Knapp RA, Boiano DM, Vredenburg VT (2007) Removal of nonnative fish results in population expansion of a declining amphibian (mountain yellow-legged frog, *Rana muscosa*). *Biological Conservation* **135:** 11-20.

Knapp RA, Matthews KR, Sarnelle O (2001) Resistance and resilience of alpine lake fauna to fish introductions. *Ecological Monographs* **71:** 401-421.

Kohler T, Giger M, Hurni H, Ott C, Wiesmann U, von Dach SW et al. (2010) Mountains and climate change: A global concern. *Mountain Research and Development* **30:** 53-55.

Kraaijeveld-Smit FJ, Beebee TJ, Griffiths RA, Moore RD, Schley L (2005) Low gene flow but high genetic diversity in the threatened Mallorcan midwife toad *Alytes muletensis*. *Molecular Ecology* **14:** 3307-3315.

Lourenço A, Sequeira F, Buckley D, Velo-Antón G (2018) Role of colonization history and species-specific traits on contemporary genetic variation of two salamander species in a Holocene island-mainland system. *Journal of Biogeography* **45:** 1054-1066.

Lucati F, Miró A, Ventura M (2020) Conservation of the endemic Pyrenean newt (*Calotriton asper*) in the age of invasive species: Interlake dispersal and colonisation dynamics. *Amphibia-Reptilia* **1:** 1-2.

Maia-Carvalho B, Vale CG, Sequeira F, Ferrand N, Martínez-Solano I, Gonçalves H (2018) The roles of allopatric fragmentation and niche divergence in intraspecific lineage

diversification in the common midwife toad (*Alytes obstetricans*). *Journal of Biogeography* **45:** 2146-2158.

Márquez R, Bosch J (1995) Advertisement calls of the midwife toads *Alytes* (Amphibia, Anura, Discoglossidae) in continental Spain. *Journal of Zoological Systematics and Evolutionary Research* **33:** 185-192.

Martínez-Silvestre A, Trochet A, Calvez O, Poignet M, Le Chevalier H, Souchet J et al. (2020) Presence of the fungus *Batrachochytrium dendrobatidis*, but not *Batrachochytrium salamandrivorans*, in wild Pyrenean brook newts (*Calotriton asper*) in Spain and France. *Herpetological Review* **51:** 738–743.

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.

Milá B, Carranza S, Guillaume O, Clobert J (2010) Marked genetic structuring and extreme dispersal limitation in the Pyrenean brook newt *Calotriton asper* (Amphibia: Salamandridae) revealed by genome-wide AFLP but not mtDNA. *Molecular Ecology* **19:** 108-120.

Miró A, O'Brien D, Tomàs J, Buchaca T, Sabás I, Osorio V et al. (2020) Rapid amphibian community recovery following removal of non-native fish from high mountain lakes. *Biological Conservation* **251:** 108783.

Montori A, Llorente GA, García-París M (2008) Allozyme differentiation among populations of the Pyrenean newt *Calotriton asper* (Amphibia: Caudata) does not mirror their morphological diversification. *Zootaxa* **1945:** 39-50.

Moritz C, Schneider CJ, Wake DB (1992) Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. *Systematic Biology* **41:** 273-291.

Mouret V, Guillaumet A, Cheylan M, Pottier G, Ferchaud AL, Crochet PA (2011) The legacy of ice ages in mountain species: Post-glacial colonization of mountain tops rather than current range fragmentation determines mitochondrial genetic diversity in an endemic Pyrenean rock lizard. *Journal of Biogeography* **38:** 1717-1731.

O'Hanlon SJ, Rieux A, Farrer RA, Rosa GM, Waldman B, Bataille A et al. (2018) Recent Asian origin of chytrid fungi causing global amphibian declines. *Science* **360:** 621-627.

Oromi N, Valbuena-Ureña E, Soler-Membrives A, Amat F, Camarasa S, Carranza S et al. (2018) Genetic structure of lake and stream populations in a Pyrenean amphibian (*Calotriton asper*) reveals evolutionary significant units associated with paedomorphosis. *Journal of Zoological Systematics and Evolutionary Research* **57:** 418-430.

Pechmann JH, Estes RA, Scott DE, Gibbons JW (2001) Amphibian colonization and use of ponds created for trial mitigation of wetland loss. *Wetlands* **21:** 93-111.

Pereira RJ, Martinez-Solano I, Buckley D (2016) Hybridization during altitudinal range shifts: Nuclear introgression leads to extensive cyto-nuclear discordance in the fire salamander. *Molecular Ecology* **25:** 1551-1565.

Recuero E, Garcia-Paris 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.

Rocaspana R, Aparicio E, Pou-Rovira Q, Cruset E, Caner J, Lucati F et al. (2019) Conservation status of the Pyrenean sculpin (*Cottus hispaniolensis*) in the Aran Valley. In: Carrillo E, Ninot J, Buchaca T and Ventura M (eds) *Life+ Limnopirineus: Conservation of aquatic habitats and species in high mountains of the Pyrenees. Technical report. Blanes, Life+ Limnopirineus Technical Office: 109-113.* ISBN: 978-84-18199-17-2.

Sánchez-Montes G, Wang J, Ariño AH, Martínez-Solano Í (2018) Mountains as barriers to gene flow in amphibians: Quantifying the differential effect of a major mountain ridge on the genetic structure of four sympatric species with different life history traits. *Journal of Biogeography* **45:** 318-331.

Savage WK, Fremier AK, Shaffer HB (2010) Landscape genetics of alpine Sierra Nevada salamanders reveal extreme population subdivision in space and time. *Molecular Ecology* **19:** 3301-3314.

Schmitt T (2009) Biogeographical and evolutionary importance of the European high mountain systems. *Frontiers in Zoology* **6:** 9.

Seddon PJ, Armstrong DP, Maloney RF (2007) Developing the science of reintroduction biology. *Conservation Biology* **21:** 303-312.

Semlitsch RD (2008) Differentiating migration and dispersal processes for pond-breeding amphibians. *The Journal of Wildlife Management* **72:** 260-267.

Skerratt LF, Berger L, Speare R, Cashins S, McDonald KR, Phillott AD et al. (2007) Spread of chytridiomycosis has caused the rapid global decline and extinction of frogs. *EcoHealth* **4:** 125-134.

Sodhi NS, Bickford D, Diesmos AC, Lee TM, Koh LP, Brook BW et al. (2008) Measuring the meltdown: Drivers of global amphibian extinction and decline. *PLoS One* **3:** e1636.

Spehn EM, Rudmann-Maurer K, Körner C, Maselli D (2010) *Mountain biodiversity and global change*. GMBA-DIVERSITAS: Basel, Switzerland.

Speybroeck J, Beukema W, Dufresnes C, Fritz U, Jablonski D, Lymberakis P et al. (2020) Species list of the European herpetofauna–2020 update by the Taxonomic Committee of the Societas Europaea Herpetologica. *Amphibia-Reptilia* **1:** 1-51.

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

Tiberti R (2018) Can satellite ponds buffer the impact of introduced fish on newts in a mountain pond network? *Aquatic Conservation: Marine and Freshwater Ecosystems* **28:** 457-465.

Valbuena-Ureña E, Amat F, Carranza S (2013) Integrative phylogeography of *Calotriton* newts (Amphibia, Salamandridae), with special remarks on the conservation of the endangered Montseny brook newt (*Calotriton arnoldi*). *PLoS One* **8:** e62542.

Valbuena-Ureña E, Oromi N, Soler-Membrives A, Carranza S, Amat F, Camarasa S et al. (2018) Jailed in the mountains: Genetic diversity and structure of an endemic newt species across the Pyrenees. *PLoS One* **13:** e0200214.

Vences M, Sarasola-Puente V, Sanchez E, Amat F, Hauswaldt JS (2017) Diversity and distribution of deep mitochondrial lineages of the common frog, *Rana temporaria*, in northern Spain. *Salamandra* **53:** 25-33.

Vences M, Wake DB (2007) Speciation, species boundaries and phylogeography of amphibians. *Amphibian Biology* **7:** 2613-2671.

Walker SF, Bosch J, Gomez V, Garner TW, Cunningham AA, Schmeller DS et al. (2010) Factors driving pathogenicity vs. prevalence of amphibian panzootic chytridiomycosis in Iberia. *Ecology Letters* **13:** 372-382.

Wallis GP, Waters JM, Upton P, Craw D (2016) Transverse alpine speciation driven by glaciation. *Trends in Ecology & Evolution* **31:** 916-926.

Weeks AR, Sgro CM, Young AG, Frankham R, Mitchell NJ, Miller KA et al. (2011) Assessing the benefits and risks of translocations in changing environments: A genetic perspective. *Evolutionary Applications* **4:** 709-725.

Appendix 1

BRILL

Amphibia-Reptilia (2020) DOI:10.1163/15685381-2020SEH1

brill.com/amre

SEH Grant in Herpetology for 2018: Conservation – Project report

Conservation of the endemic Pyrenean newt (*Calotriton asper*) in the age of invasive species: interlake dispersal and colonisation dynamics

Federica Lucati[1,2,*], Alexandre Miró[1], Marc Ventura[1]

The aim of the project was to investigate patterns of colonisation of new ecosystems following invasive fish removal and the process of dispersal from nearby populations in the Pyrenean newt (*Calotriton asper*) (fig. 1), a Near Threatened (NT-IUCN) amphibian endemic to the Pyrenees. We used a multidisciplinary approach that combined genetic (population structure and dispersal analyses) and field-based (mark-recapture using photo-identification) approaches. The study site was located in the cirque of Dellui, in the National Park of Aigüestortes i Estany de Sant Maurici (central-eastern Pyrenees). This glacial cirque features six high mountain lakes, of which two are fishless and sustain natural populations of *C. asper*, three are undergoing complete fish eradication and are being colonised by *C. asper*, and one still harbours invasive fish and no newts were reported. Fish eradication activities have been ongoing since 2014, with the purpose of returning selected high mountain lakes to their fish-free natural state and allow the recovery of native species such as amphibians (LIFE+ LIMNOPIRINEUS; www.lifelimnopirineus.eu).

Figure 1. An adult individual of Pyrenean newt (*Calotriton asper*). Photo by Marc Ventura.

SEH funds were the only financial support available for this project in 2018 and were used to cover sampling and laboratory costs. Data collected in 2018 were complemented with those obtained from previous field campaigns in 2016 and 2017.

Assessment of population structure and recent dispersal was conducted through the analysis of 17 nuclear microsatellite markers (Drechsler et al., 2013). In total, 106 individuals from five of the six target lakes were sampled (we did not detect newts in the fish-containing lake). Seven populations (82 individuals) from nearby cirques were also sampled to gain further insights into the dispersal capability of the species. *C. asper* individual photo-identification was performed through the pattern-recognition

1 - Centre for Advanced Studies of Blanes (CEAB-CSIC), Accés Cala St. Francesc, 14, 17300, Blanes, Girona, Spain

2 - Centre for Ecology, Evolution and Environmental Changes (cE3c), Faculty of Sciences, University of Lisbon, Campo Grande, 1749-016 Lisbon, Portugal

*Corresponding author;
e-mail: federicalucati@hotmail.com

software I^3S Pattern$^+$ 4.1 (Van Tienhoven et al., 2007), based on the species spotted ventral pattern. Three to six sampling sessions were carried out at four target lakes during the 3-year period.

Population structure analysis revealed three well-supported groups (genetic lineages), corresponding to the geographic areas under study. Dellui lakes clustered together with the adjacent cirque of Corticelles. We found a relatively high degree of migration within the cirque of Dellui that involved 34 individuals across all sampled lakes. Furthermore, we detected two migration events from Corticelles to Dellui. Nonetheless, migration was limited within lineages. The photographic dataset included 125 individuals, 20 of which were recaptured at least once. The species ventral pattern was sufficiently varied to allow for the reliable identification of all individuals and it did not substantially change over the study period. Of the 20 recaptures, only one was found in a lake other than the lake of first collection. The individual in question was first captured at a fishless lake in 2016 and recaptured two years later at one of the restored lakes.

This work emphasizes the high resilience of *C. asper* high-mountain populations after reducing or eliminating the source of perturbation. The species showed a remarkable capability to colonize, establish and spread into new suitable environments when invasive fish predation pressure was reduced, even without complete fish eradication. *C. asper* was detected within one year of starting fish removal at the restored lakes, and both reproduction and survival of larvae were confirmed. To our knowledge, this is the first study attempting to use computer-assisted photo-identification in a capture-recapture study targeting the Pyrenean newt. Results indicate that the species ventral pattern is highly patterned and stable enough over time to allow for reliable individual identification, and provide an effective alternative to more intrusive approaches. The number of recaptures was low, indicating that a much higher sampling effort is required to disentangle the species' dispersal dynamics using this approach. Nevertheless, the only case of inter-lake dispersal shows the importance of fishless lakes as reservoirs for restored lakes. On the other hand, genetics proved more efficient in characterising patterns of dispersal. The studied populations are genetically and geographically structured and migration usually involved neighbouring lakes. However, we detected evidence of dispersal between adjacent cirques. Finally, this study highlights the importance of combining genetic and field-based approaches to shed light on dispersal dynamics.

Acknowledgements. We thank Jenny Caner for laboratory assistance, and Jan Tomàs, Blanca Font, Ismael Jurado, Meritxell Cases, Eloi Cruset, David O'Brien and Quim Pou-Rovira for field assistance. We are especially grateful to Alba Castrillón for her help during field work and for photo-identification data processing.

References

Drechsler, A., Geller, D., Freund, K., Schmeller, D.S., Kuenzel, S., Rupp, O., Loyau, A., Denoël, M., Valbuena-Ureña, E., Steinfartz, S. (2013): What remains from a 454 run: estimation of success rates of microsatellite loci development in selected newt species (*Calotriton asper*, *Lissotriton helveticus*, and *Triturus cristatus*) and comparison with illumina-based approaches. Ecol. Evol. **3**: 3947-3957.

Van Tienhoven, A.M., Den Hartog, J.E., Reijns, R.A., Peddemors, V.M. (2007): A computer-aided program for pattern-matching of natural marks on the spotted raggedtooth shark *Carcharias taurus*. J. Appl. Ecol. **44**: 273-280.