MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 2046–2064 doi: 10.1111/mec.13585

Imprints of multiple glacial refugia in the Pyrenees revealed by phylogeography and palaeodistribution modelling of an endemic spider

LETICIA BIDEGARAY-BATISTA,*† ‡ ALEJANDRO SÁNCHEZ-GRACIA,*§ GIULIA SANTULLI,*† LUIGI MAIORANO,¶** ANTOINE GUISAN,¶†† ALFRIED P. VOGLER‡ ‡ §§ and MIQUEL A. ARNEDO*†

*Institut de Recerca de la Biodiversitat, Universitat de Barcelona, Av. Diagonal 643, 08028 Barcelona, Spain, †Departament de Biologia Animal, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 643, 08028 Barcelona, Spain, ‡Laboratorio de Etología, Ecología y Evolución, Instituto de Investigaciones Biológicas Clemente Estable, Avenida Italia 3318, 11600 Montevideo, Uruguay, §Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 643, 08028 Barcelona, Spain, ¶Department of Ecology and Evolution, University of Lausanne, Biophore Building, CH-1015 Lausanne, Switzerland, **Department of Biology and Biotechnologies 'Charles Darwin', University of Rome 'La Sapienza', viale dell'Università 32, 00185 Rome, Italy, ††Institute of Earth Surface Dynamics, University of Lausanne, Geopolis Building, CH-1015 Lausanne, Switzerland, ‡‡Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK, §§Department of Life Sciences, Imperial College London, Silwood Park Campus, Ascot SL5 7PY, UK*

Abstract

Mediterranean mountain ranges harbour highly endemic biota in islandlike habitats. Their topographic diversity offered the opportunity for mountain species to persist in refugial areas during episodes of major climatic change. We investigate the role of Quaternary climatic oscillations in shaping the demographic history and distribution ranges in the spider *Harpactocrates ravastellus*, endemic to the Pyrenees. Gene trees and multispecies coalescent analyses on mitochondrial and nuclear DNA sequences unveiled two distinct lineages with a hybrid zone around the northwestern area of the Catalan Pyrenees. The lineages were further supported by morphological differences. Climatic niche-based species distribution models (SDMs) identified two lowland refugia at the western and eastern extremes of the mountain range, which would suggest secondary contact following postglacial expansion of populations from both refugia. Neutrality test and approximate Bayesian computation (ABC) analyses indicated that several local populations underwent severe bottlenecks followed by population expansions, which in combination with the deep population differentiation provided evidence for population survival during glacial periods in microrefugia across the mountain range, in addition to the main Atlantic and Mediterranean (western and eastern) refugia. This study sheds light on the complexities of Quaternary climatic oscillations in building up genetic diversity and local endemicity in the southern Europe mountain ranges.

Keywords: Climatic oscillations, glacial refugia, mitochondrial genes, mountain species, nuclear intron

Received 13 November 2014; revision received 24 January 2016; accepted 26 January 2016

Introduction

Quaternary climatic oscillations had a major impact on biodiversity globally (Hewitt 2004). Glaciations reshaped species distributions and changed population

Correspondence: Leticia Bidegaray-Batista,
Fax: +598 2 4875461; E-mail: letigaray@yahoo.com

sizes, leaving an enduring imprint on the genetic structure of populations and species (Miraldo *et al.* 2011; Espíndola *et al.* 2012; Garcia-Porta *et al.* 2012). The magnitude of range and demographic shifts largely depended on the species' environmental requirements, dispersal capabilities and generation times (Stewart *et al.* 2010; Arenas *et al.* 2012), and in some organisms eventually led to speciation (Knowles 2000; Barraclough & Vogler 2002; Ribera & Vogler 2004; Muster & Berendonk 2006; Gante *et al.* 2009).

Research on the effect of glaciations on European diversity has mostly focused on postglacial colonization pathways of northern Europe from the southern refugia, or the characterization of postglacial contact zones (Taberlet *et al.* 1998; Hewitt 2001; Schmitt 2007). More recently, the interest has shifted towards identifying the geographic location of putative refugia where species could have persisted during glacial maxima, for instance in the Alps (Schorr *et al.* 2013) or in the southern European Peninsulas (Hewitt 1996; Gómez & Lunt 2007). The Mediterranean mountain ranges have been identified as 'biological islands' because of their high level of endemism (Médail & Quezel 1997; Médail & Diadema 2009). Glaciation-driven altitudinal distribution shifts, facilitated by topographic gradients, probably played a key role in promoting species diversification in mountain ranges (Gutierrez Larena *et al.* 2002; Vila *et al.* 2005; Schmitt 2009; Alarcón *et al.* 2012; Milá *et al.* 2013). Glaciations offered the opportunity for mountain species to migrate among high-elevation mountain islands by establishing lowland corridors, while interglacials promoted diversification of geographically isolated lineages (Schoville & Roderick 2009). However, not all mountain species responded to glaciations in the same manner, and hence, different demographic and genetic diversity patterns should be expected in mountain species depending on where species survived the glaciations. Holderegger & Thiel-Egenter (2009) differentiated three main kinds of glacial refugia in mountain species: small areas free of ice on mountain tops (nunatak refugia); areas at the edge of mountain systems (peripheral refugia); and in areas located outside mountain systems (lowland refugia). The effect of glacial cycles on the distributions and demographic history of mountain species within the Mediterranean mountain ranges, however, remains largely unstudied, especially at the regional scale.

The Pyrenees are a major mountain range in southern Europe that stretch 400 km between the Mediterranean coast to the east and the Atlantic Bay of Biscay to the west. The Pyrenees exhibit physiographic asymmetry: the northern slope is more pronounced than the southern slope, and likewise, the eastern is more pronounced than the western. Moreover, the northern and western slopes receive abundant precipitation throughout the year, whereas the southern and eastern slopes have a more Mediterranean climate (Calvet 2004). The extent of Quaternary glaciations was affected by these asymmetries resulting in particularly pronounced differences in the snowline, which was located at higher elevations on the southeastern slopes. Small glaciers remain today only in the central Pyrenees (e.g. the Maladeta mountain) and further west (e.g. peaks of Balaïtous and Vignemale), while no glaciers are left in the eastern Pyrenees (Calvet 2004; Hughes *et al.* 2006).

Due to their west–east orientation reaching from the Mediterranean to the Atlantic Ocean, the Pyrenees acted as a barrier to dispersal for many temperate species during their postglacial northward expansions (Taberlet *et al.* 1998). The existence of hybrid zones, on the other hand, provides evidence for the postglacial expansion and secondary contact of Iberian genotypes, for example in the grasshopper, *Chorthippus parallelus* (Hewitt 1996, 2000).

Yet, little is known about how the Pyrenean species endured Quaternary glacial cycles and specifically about their persistence in refugia within or outside the Pyrenees. A recent study on *Antirrhinum* (Liberal *et al.* 2014) revealed separate evolutionary histories for populations in the western, central and eastern sections of the mountain range and found evidence for the long-term persistence of local populations since the late Pleistocene within the Pyrenees, although peripheral areas played an important role in maintaining local genetic diversity. In the rock lizard *Iberolacerta bonnali,* peripheral refugia have also been identified as important reservoirs of genetic diversity (Mouret *et al.* 2011).

Dispersal ability is a crucial trait affecting the evolutionary history of species (McPeek & Holt 1992; Nathan *et al.* 2003; Gillespie *et al.* 2012). Variation in vagility frequently leads to striking differences in phylogeographic patterns (McPeek & Holt 1992; Hodges *et al.* 2007; Papadopoulou *et al.* 2009). The chances of a species persisting through glacial cycles will depend on both the availability and location of refugia and the ability of the species to track them (Peterson 2009). Low vagility species might be less prone to track refugia outside their distribution range, and thus will depend on local refugia, while limited gene flow will result in genetic differentiation among adjacent populations. Conversely, good dispersers are more likely to track distant refugia, experiencing large-scale range shifts and exhibiting weaker local geographic structure due to high levels of gene flow.

The ground-dwelling spider *Harpactocrates ravastellus* Simon 1914 (Araneae, Dysderidae) is endemic to the Pyrenees. The species shows a clear preference for cool and humid environments: they are usually found at elevations (above 1000 m) in temperate and moist forests

(dominated by white oaks, beech and pine trees). This nocturnal wandering hunter builds its silk nest under stones and dead logs to spend daylight hours and to protect its eggs. The species has never been reported to use ballooning (i.e. aerial dispersal by means of silk threads, a dispersal mechanism that allows long-distance dispersal). This observation together with the habitat preferences of the species suggests a restricted vagility. The genus includes about a dozen additional species across the main mountain ranges in the western Mediterranean, all sharing similar ecological preferences. The species' distributions are mostly allopatric and are usually restricted to single mountain ranges or systems. The closest relatives of *H. ravastellus* are found in the Cantabrian Mountains, the western Alps and northernmost Apennines (Bidegaray-Batista *et al.* 2014). Because of its restricted distribution, habitat preferences and likely low dispersal ability, *H. ravastellus* is a promising model to test scenarios of climatic change in shaping the southern European mountain fauna.

The integration of phylogeographic analyses and palaeodistribution models has been used for inferences on the evolutionary process over time and space (Graham *et al.* 2004; Carstens & Richards 2007; Richards *et al.* 2007; Carnaval *et al.* 2009; Galbreath *et al.* 2011; Marske *et al.* 2011; Schorr *et al.* 2012, 2013). Here, we combine molecular tools, morphological analyses and species distribution modelling to investigate the phylogeographic patterns, demographic history and morphological differentiation of *H. ravastellus* in order to unravel the effects of Quaternary climatic oscillations. Specifically, we test whether *H. ravastellus* experienced range shifts towards lowland refugia outside the Pyrenees or alternatively withstood the glaciations *in situ*, and what was the effect of climate change on the evolutionary diversification of lineages.

Materials and methods

Morphological identification

Specimens were examined with a Leica MZ16A dissection microscope equipped with a Nikon DXM1200 digital camera. Female vulvas were removed with the aid of needles, and muscle tissue was digested with a 35% KOH solution before observation. Incident light photography digital images were obtained of male copulatory bulb and female vulva and focal planes composed with Helicon Focus 4.62 Pro (www.heliconsoft.com).

Molecular data collection

Harpactocrates ravastellus specimens were collected from 26 localities along an east–west transect across the Pyrenees including localities from both northern and southern slopes (Fig. 1, Table 1). Specimens were stored in absolute ethanol at −20 °C. Samples of three close relatives, *Harpactocrates apennicola* from the northern Apennines, *Harpactocrates intermedius* from the western Alps and a new species from the Cantabric mountain (hereafter referred as *Harpactocrates* n. sp. 'Cantabrian'), were also included in the analysis along with the species *Harpactocrates radulifer*, which was used to root the trees (Bidegaray-Batista *et al.* 2014).

Genomic DNA was extracted from two legs of adult specimens or whole juveniles, after piercing the carapace, using the Promega 96-well plate kit. A mitochondrial fragment (*rrnL-nad1*) spanning the 3′ half of the 16S rRNA ribosomal subunit (*rrnL*), the complete tRNA ^{leu}(*L1*), the 5′ half of the NADH dehydrogenase subunit I (*nad1*) and the 5′ half of cytochrome c oxidase subunit I (*cox1*) were amplified and sequenced using the primer pairs LR-N-13398 (Simon *et al.* 1994) and N1-J-12350 (Bidegaray-Batista & Arnedo 2011), and C1-J-1546 (Bidegaray-Batista *et al.* 2007) and C1-N-2770 (Srapepis: 5′-CAGAATAACGTCGAGGCATCCC-3′), respectively. The nuclear intron of the gene encoding signal recognition particle 54-kDA subunit (*srp54*) was additionally sequenced for the specimens of *H. ravastellus* with the primer pairs SRP54f1 and SRP52r1 (Jarman *et al.* 2002). PCR conditions were as follows: 94 °C for 2 min; 35× (94 °C for 35 s; 45 °C for 45 s for *rrnL-nad1* and *cox1*, and from 45 °C to 50 °C for 35 s for *srp54*; 72 °C for 45 s for *rrnL-nad1* and *cox1*, and 35 s for *srp54*; 72 °C for 5 min). PCR products were purified using MultiScreen 96-well plates from Millipore and sequenced in both directions with ABI BigDye technology and an ABI PRISM 3700. DNA sequences were edited using GENEIOUS PRO v.5.3.6 (Drummond *et al.* 2010).

Alleles in heterozygous individuals for the *srp54* intron were separated using the PHASE algorithm (Stephens *et al.* 2001; Stephens & Donnelly 2003), as implemented in DNASP v5 (Librado & Rozas 2009).

Genetic diversity and geographic structure of genetic variation

Diversity indices, including nucleotide (π_n) and haplotype (h) diversities, and neutrality tests including Tajima's D (Tajima 1989b) and Fu's F_S (Fu 1997) were calculated using ARLEQUIN v3.01 (Excoffier *et al.* 2005). All statistics were calculated by locality only for the mitochondrial genes, as the *srp54* intron showed no variation within most sampled localities (see Results).

Genetic differentiation among localities was tested using ϕ_{ST} pairwise genetic distances and their statistical significance assessed by performing 1000 permutations in ARLEQUIN. Localities represented by a single

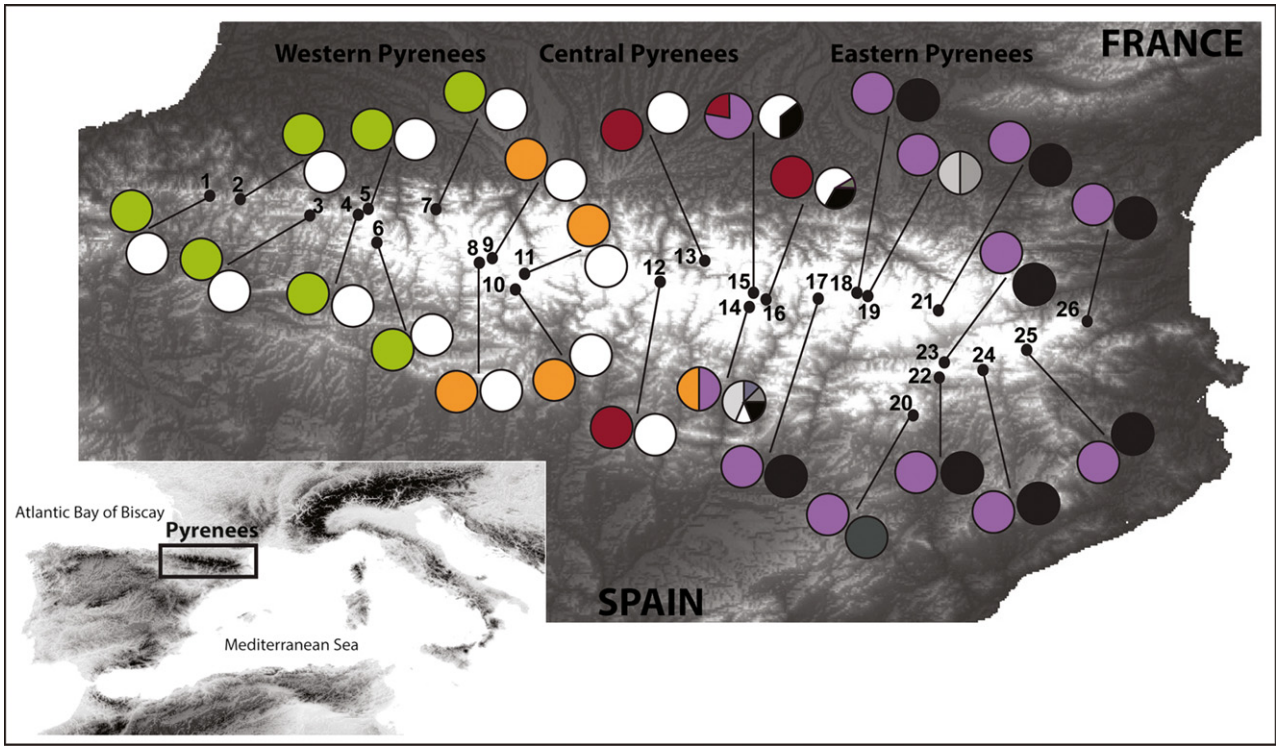

Fig. 1 Map showing the distribution of genetic variation across the sampled locations of *Harpactocrates ravastellus*. Sampled locations are indicated by black dots and numbered as in Table 1. Each location includes two pie diagrams; the left one represents the frequency of the main mitochondrial lineages as identified and coloured in the haplotype tree (see Fig. 2), and the right one represents the frequency of nuclear alleles as identified and coloured in the allele network (see Fig. 3).

individual were excluded to estimate diversity indices and ϕ_{ST} (localities 5, 18, 19 and 20, see Appendix S1, Supporting information). Isolation by distance was tested including populations from the entire distribution range of *H. ravastellus*, as well as for the SAMOVA groupings, using Mantel tests implemented in ARLEQUIN. Matrix correlation between populations ϕ_{ST} or log-ϕ_{ST} and log-transformed geographic distances were assessed for significance using 1000 permutations.

The geographic structure of the mitochondrial gene variation was assessed using spatial analysis of molecular variance implemented in SAMOVA v1.0 (Dupanloup *et al.* 2002), which optimizes the partitioning of total genetic variance for *k* population groups.

Phylogenetic analyses and haplotype networks

The alignment of the *cox1* and *nad1* sequences of *H. ravastellus* and its congeneric relatives was trivial as no insertions/deletions (indels) were observed. Likewise, no indels were observed in the *srp54* sequences of *H. ravastellus*. Conversely, sequences of the *rrnL* and *trnL1* gene fragments showed length polymorphism and were aligned using the online version of MAFFT v7 (Katoh *et al.* 2002) using the Q-ins-I algorithm, which takes into account the RNA secondary structure of RNA. Gaps were scored as presence/absence data using the method of Simmons & Ochoterena (2000) as implemented in GapCoder (Young & Healy 2002), and the resulting new characters were included in parsimony and Bayesian analyses. The aligned mitochondrial genes were concatenated using WINCLADA v.1.00.08 (Nixon 2002).

Parsimony, maximum likelihood (ML) and Bayesian inference (BI) analyses were conducted on the mitochondrial data set including only unique haplotypes of *H. ravastellus* and sequences of its congeneric relatives. Parsimony analyses were conducted with TNT v1.0 (Goloboff *et al.* 2008) using a heuristic search with 1000 replicates of random sequence addition, followed by TBR branch swapping, holding five trees per iteration and a final round of branch swapping on all retained trees. Clade support was assessed by means of 1000 jackknife replicates (Farris *et al.* 1996). The best partition scheme and the best-fitting model for each partition were selected with PartitionFinder (Lanfear *et al.* 2012). Independent GTR + G substitution models were specified for each partition in ML analyses using RAXML (Stamatakis 2006) and run remotely at the CIPRES portal (Miller *et al.* 2009). The best likelihood tree was selected

Table 1 Code: locality number that correspond to localities in Fig. 1. N: number of individual sampled per locality. See Appendix S1 (Supporting information) for information on coordinates and haplotype/allele frequencies

Locality	Code	N
Pto. de Urkiaga, Navarra, Spain	1	6
Pto. de Ibañeta, Navarra, Spain	2	3
Pto. de Larrau, Navarra, Spain	3	4
San Martín Harria, Navarra, Spain	4	8
Arette, Pyrénées-Atlantiques, Aquitaine, France	5	1
Selva de Oza, Huesca, Aragón, Spain	6	4
Forêt Communale de Laruns, Pyrénées-Atlantiques, Aquitaine, France	7	10
Balneario de Panticosa, Huesca, Aragón, Spain	8	6
Piau Engaly, Hautes-Pyrénées, Midi-Pyrénées, France	9	9
Parque Nacional Ordesa, Huesca, Aragón, Spain	10	7
Valle de Bujaruelo, Huesca, Aragón, Spain	11	3
Ball de Remuñe, Huesca, Aragón, Spain	12	5
Saut deth Pish, Lleida, Catalunya, Spain	13	12
Parc Nacional d'Aigüestortes i Estany de Sant Maurici, Lleida, Catalunya, Spain	14	8
Estany de Gerber, Lleida, Catalunya, Spain	15	14
Les Planes de Son, Lleida, Catalunya, Spain	16	6
Font de Punta Nou, Lleida, Catalunya, Spain	17	3
Pleta del Castellar, La Massana, Principat d'Andorra	18	1
Canya de la Rabassa, Ordino, Principat d'Andorra	19	1
Rasos de Peguera, Barcelona, Catalunya, Spain	20	1
Porté, Pyrénées-Orientales, Languedoc-Roussillon, France	21	4
Font de la Doble Ona, Bagà, Berguedà, Barcelona, Catalunya, Spain	22	6
La Masella, Cerdanya, Lleida, Catalunya, Spain	23	2
Navà, Ripollès, Girona, Catalunya, Spain	24	5
Pto. Vallter 2000, Ripollès, Girona, Catalunya, Spain	25	7
Mont Canigou, Pyrénées-Orientales, Languedoc-Roussillon, France	26	7
Total		143

out of 50 iterations of random addition of taxa and clade support assessed by 1000 bootstrapped matrices. Bayesian inference was conducted using MRBAYES v.3.1.2 (Ronquist & Huelsenbeck 2003). A standard discrete model was defined for the gaps scored as absence/presence. Two independent runs of 30 million generations were carried out simultaneously, with six simultaneous MCMC chains, each starting from random trees, and sampling trees every 10^3 generation. TRACER v1.5 (Drummond & Rambaut 2007) was used to ensure that the Markov chains had reached stationarity by examining the effective sample size (ESS) and to determine the number of generations of burn-in.

The haplotype network of the nuclear intron *srp54* was estimated using TCS v. 1.21 (Clement *et al.* 2000). A minimum number of recombination events (R_M) of Hudson & Kaplan (1985) test was performed in DNASP v5 (Librado & Rozas 2009) and its significance assessed by 1000 coalescent simulations.

Estimation of lineage divergence times

Lineage ages were estimated from the mitochondrial data set of *H. ravastellus* using a partitioning scheme by gene in BEAST v1.6.2 (Drummond & Rambaut 2007), under an uncorrelated lognormal relaxed clock and Yule speciation process as tree prior. The best substitution model for each partition was chosen based on the Bayesian information criterion (BIC). To avoid mixing different lineage branching rates (see Papadopoulou *et al.* 2010), lineage ages were estimated including only one individual per cluster identified by the generalized mixed Yule coalescent (GMYC) method (Pons *et al.* 2006; Fontaneto *et al.* 2007), under the single-threshold option. The GMYC analysis was conducted with the R package SPLITS (Species Limits by Threshold Statistic, http://r-forge.r-project.org/projects/splits/). The ultrametric tree required for the analyses was first estimated in BEAST by including all individuals of *H. ravastellus* under a strict clock and mean rate fixed at 1. The coalescent constant size demographic model was used as tree prior, as it provides a more conservative alternative to the identification of lineage splits (i.e. minimizes type I error) (Monaghan *et al.* 2009).

Specific *Harpactocrates* mitochondrial substitution rates estimated in Bidegaray-Batista *et al.* (2014), based on fossil and biogeographical calibration points, were used to infer absolute lineage ages. Individual substitution rates for each mitochondrial gene were incorporated as normal prior distribution of the ucld.mean rate as follows: mean = 0.0255 substitutions/my and Stdev = 0.002 for *cox1*; mean = 0.0226 substitutions/my and Stdev = 0.002 for *nad1*; mean = 0.0083 substitutions/my and Stdev = 0.0008 for *rrnL-L1*. Two independent runs of 10^7 generations were performed, sampling every 10^3 generations. Convergence and mixing of MCMC chains was assessed with TRACER v.1.5. Independent runs were combined with LogCombiner (10% burn-in), and TreeAnnotator was used to summarize the information from the sampled trees.

Additional divergence time estimates were obtained using the multispecies coalescent approach implemented in * BEAST (Heled & Drummond 2010), including both mitochondrial and nuclear genes. For these analyses, we included all individuals from the six sampling localities that were used to compare the different demographic models in the ABC-based analyses. As the method requires *a priori* definition of populations (or

species), we defined each sampling locality as an independently evolving entity. The multilocus * BEAST analyses were carried out using a partitioned data approach with independent model parameters for each gene partition. We set a strict molecular clock model and a Yule speciation prior and incorporated the same substitution rates priors as above. Two independent runs of 10^8 generations, sampling trees every 10^4 generations, were carried out and convergence and mixing of each MCMC chain assessed and combined as described above.

Testing for demographic scenarios using ABC

To investigate the effect of the Pleistocene glacial cycles in shaping the genetic diversity, we contrasted different coalescent-based evolutionary models using the ABC-GLM method (Leuenberger & Wegmann 2010). We evaluated three different demographic scenarios: (i) a standard neutral model (SNM), representing a long-term stable population, that is not affected by recent climatic changes, (ii) a bottleneck model (BTM), including the population size reduction phase and the posterior expansion, and (iii) a two-refugia model, in which two previously isolated populations underwent a recent admixture (IAM). If *H. ravastellus* populations were affected by Pleistocene glacial cycles, we would expect that genetic data supported more strongly either the BTM or the IAM than the SNM model. As our sample showed strong genetic differentiation among localities (see Results), each locality was treated as an independent population for the ABC analysis for testing the three competing evolutionary models. Analyses were restricted to localities with sample size $n \geq 7$ (localities 4, 7, 9, 10, 13, 14, 15, 25 and 26). We used only *cox1* and *nad1* genes. The *rrnL-L1* marker was not included because of its different substitution rate (see Bidegaray-Batista & Arnedo 2011) that could not be accommodated using the rate heterogeneity models available in the coalescent simulation program used. We also excluded from the analysis the *srp54* intron because of its low genetic variation within populations. Priors for the coalescent times, t_E (time where populations start the growth phase) and t_B (duration of the bottleneck phase) in BTM, were uniformly distributed as the total time from the bottleneck was always ≤5 ky [using the substitution rates estimated for *Harpactocrates* mitochondrial genes and a generation time, g = 1.5 years (Cooke 1965)]. Priors for coalescent times under the IAM model were uniformly distributed as t_S (the time when ancestral population split to generate the two isolated population) and t_A (the time of recent admixture) were <5 Ma and <2.5 Ma, respectively.

We calculated a vector of 4 summary statistics describing the intraspecific variation (Tajima's *D*, Fu's *Fs*, ZnA, and the Watterson parameter, θ) in both 2 million simulated data set (for each model) and in the observed data, using the program MLCOALSIM v.1.98 (Ramos-Onsins & Mitchell-Olds 2007) and mstatpop (http://bioinformatics.cragenomica.es/numgenomics/people/sebas/software/software.html), respectively. We retained the 0.5% of simulated data sets with the small Euclidian distance to the vector of the observed summary statistic to estimate the general linear model (GLM) and to calculate approximate posterior probabilities (PP). Model selection was performed using Bayes factors (BI) and was validated as described in Peter *et al.* (2010). Briefly, we simulated 1000 replicates under each of the three competing models from the same prior distributions used in the ABC analysis, that is pseudo-observed data sets (PODS). Then, we applied to each pseudo-observation the same model selection procedure applied to the observed data. We assessed the power and accuracy of our ABC model choice by estimating the confusion matrix (i.e. the proportion of PODS from each model that are correctly classified by the model choice procedure), the percentage of misclassifications (i.e. false allocations) and the misclassification probabilities of the model supported by the *H. ravastellus* empirical data. We also checked for biases in ABC posterior probabilities as in Chu *et al.* (2013). All ABC analyses were performed using the ABCTOOLBOX package (Wegmann *et al.* 2010) and in-house R and Shell language scripts.

We also used the hierarchical approximate Bayesian computation (hABC) method implemented in HBAYESSC (Chan *et al.* 2014) to detect putative concerted demographic histories across populations with the same demography inferred in local ABC analyses (populations from localities 4, 7, 9, 10 and 13, see Results). We used the same prior distributions of model parameters, the same set of summary statistics and the same hABC procedure as in Chan *et al.* (2014). Briefly, multispecies coalescent simulations (250 000 replicates for each of the six possible coexpansion models for five contemporary populations; a total of 1.5 simulated random draws from the priors) and summary statistics (the first four sample moments of the number of haplotypes, haplotype diversity, nucleotide diversity and Tajima's *D*) were obtained with Bayes Serial Simcoal (Anderson *et al.* 2005) and the Phyton scripts included in the hBayesSSC package (https://github.com/UH-Bioinformatics/hBayeSSC). A local linear regression model was applied to the 1000 closest Euclidean distances between simulated and observed data to obtain the adjusted estimate of the joint posterior distribution of the proportion

of coexpanding populations (ζ) and the coexpansion time (τ_e).

Species distribution modelling and projections into the Late Pleistocene

Niche-based species distribution models (SDMs, Guisan & Thuiller 2005) were constructed in order to (i) investigate the potential effects of Pleistocene climatic oscillations on the geographic distribution patterns of *H. ravastellus* since 40 000 years ago to the present and (ii) assess whether its potential past distribution explains the current distribution of genetic diversity. Climate data for 52 sites with presence of *H. ravastellus* were used to construct the SDMs (Table S1, Supporting information). These sites included the sampling localities used for this study together with occurrence data from additional sources (De Mas-Castroverde 2007; and J. Moya-Laraña pers. comm). To calibrate the climatic niche of the species, we selected a set of 4 uncorrelated bioclimatic variables (Pearson $r \leq 0.8$) assumed to be important in shaping the species' climatic niche: annual temperature range (tave-range), winter mean temperature (tave-winter), annual precipitation range (prc-range) and summer mean precipitation (prc-summer). All bioclimatic variables for the current climate were obtained from the Climatic Research Unit (Mitchell *et al.* 2004). The same variables for the past (at intervals of 4000 years for the past 40 000 years to the present, with a spatial resolution 30 arc-seconds) were obtained from a global ocean–atmosphere climate model (HadCM3) with a temporal resolution of 1000 years and a spatial resolution of 3.75° by 2.5° (Singarayer & Valdes 2010). The past climate was downscaled to a spatial resolution of 30 arc-seconds following the methods outlined in Maiorano *et al.* (2013) (as also used in Espíndola *et al.* 2012). In particular, we first calculated climate anomalies by contrasting the past monthly temperature and precipitations values against the pre-industrial climate as obtained from Singarayer & Valdes (2010). Then, given the existing offset in temperature and precipitation between the pre-industrial and the current climate, we calculated a second set of climate anomalies by contrasting climate means for the period 1901–1920 against the current climate (both obtained from Mitchell *et al.* 2004), assuming that the former provides an approximation closer to the pre-industrial climate than the available simulations. The combination of these two anomalies allowed us to avoid the usual inconsistency when analysing time series that involve both measured (current) and projected (past or future) climates. Anomalies were calculated in both cases as absolute temperature difference (Δ°C) and relative precipitation differences (% change) per coarse resolution pixel measured directly at the model output. The study area considered for the SDM was defined considering the Pyrenees and surrounding areas, considering the Ebro river in Spain and the Garonne river in France as outer boundaries.

Before calibrating the SDMs, a niche similarity test was performed to assess whether there were differences between the climatic niches of two main groups from the west-central and eastern Pyrenees (see Results section). The similarity test was performed in R using the scripts developed by Broennimann *et al.* (2012). The test examines whether the climatic niche of one group is more similar to the niche occupied by the other group than expected by chance, and the overlap of niches is measured using Schoener's D index. To conduct the test, the climatic niche was calibrated considering the first two axes of the PCA performed using the 4 variables described above and considering a background area, drawn from the surrounding species occurrence points, as environment available for the species.

To avoid spatial autocorrelation among occurrence points during the construction of SDMs, 50 subsamples were generated at random from the original presences using the 'Average Nearest Neighbour Distance' tool in ARCGIS 9.3.1 to calculate the mean expected distance for unclustered points. For each subsample, a Maximum Entropy modelling algorithm, as implemented in MAXENT v3.3.3e (Phillips *et al.* 2006), was used to calibrate a set of SDMs considering a random selection of 80% of presence points from each subsample. The remaining 20% of points were used to assess model performance measuring the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. Each model calibration was repeated 10 times for each of the 50 subsamples, giving a total of 500 models. The AUC values and variable importance were averaged for all SDMs. Finally, all the SDMs were projected onto the current and past climate surfaces at 4000-year intervals (4–40 ka BP) and averaged for each time frame.

For each time frame, the models were projected, converted into binary maps (presence/absence maps) and summed across all time frames to predict areas of stability. In the latter, areas with highest counts were then assumed to be refugial areas (predicted areas of occurrence for the species independent of the time period) for the species throughout the Pyrenees since the Late Pleistocene. Binary maps were created by applying two different threshold criteria: a default threshold of 0.5 and one that minimizes the mean error rate for positive and negative observations ['Default' and 'MaxSens + Spec' criterion, respectively, in PresenceAbsence R package (see Freeman & Moisen 2008)].

Results

Sequence variation

We collected 143 specimens of *Harpactocrates ravastellus* from 26 localities throughout the Pyrenees (Fig. 1, Table 1), of which 142 were sequenced for the *srp54* intron and 141 for the mitochondrial genes *cox1*, *rrnL*, *L1* and *nad1*. The alignments of the protein coding genes and the nuclear intron yielded 1030 characters (*cox1*), 347 characters (*nad1*) and 143 characters (*srp54*). The MAFFT alignment of the nonprotein genes (*rrnL-L1*) yielded 552 characters. The inclusion of 5 additional outgroup taxa increased the aligned matrix by 4 characters. The mitochondrial genes were concatenated resulting in a combined matrix of 1929 characters (1933 characters with outgroups included). Coding informative gaps as absence/presence characters added 5 additional characters to the matrix. The mitochondrial data set included 106 haplotypes (see Accession numbers in Data Accessibility). Of 142 individuals sequenced for the nuclear intron, six were heterozygous (2 with a single heterozygous site and 4 with two or more). Individual alleles could be resolved using the 0.99 probability threshold in PHASE. The nuclear intron data set yielded a total of 284 phased sequences representing 9 alleles (see Accession numbers in Data Accessibility). The R_M test revealed nonsignificant values of recombination ($P = 0.967$).

The average genetic divergence (p-distance) among the mitochondrial haplotypes and nuclear alleles were 4.9% and 3.1%, respectively. The maximum sequence divergence was 7.2% for the mitochondrial genes and 5.6% for the nuclear intron. The average genetic distances among the main mitochondrial lineages were around 6%, except for a value of 4.5% between the Western and Central 2 clades.

Genetic diversity and spatial genetic variation

The nucleotide diversity in *H. ravastellus* was 0.048 for the mitochondrial genes and 0.023 for the nuclear intron *srp54*, and the haplotype/allelic diversity 0.99 and 0.52, for the mitochondrial genes and the nuclear intron, respectively. Mitochondrial nucleotide diversity within each locality ranged from 0.00035 (locality 2) to 0.041 (locality 14) and the haplotype diversity from 0.66 (localities 2 and 7) to 1 (localities 3, 8, 9, 11, 14, 17, 22, 23, 24 and 26) (see Fig. 1 and Table 1 for localities names and locations). The distribution of mitochondrial haplotypes was geographically structured. If present at multiple localities, haplotypes were always confined to adjacent localities. Tajima's *D* and Fu's F_S were both negative and significantly different from neutral expectations for localities 4, 7, 8 and 9, whereas for localities 1 and 22, only Tajima's *D* or Fu's F_S were negative and significant, respectively. The ϕ_{ST} pairwise genetic distance values among localities were overall high and significant ($P < 0.05$).

The spatial analysis of molecular variance (SAMOVA) showed different groupings from $K = 2$ to 9 that maximize the proportion of total genetic variance due to differences between local groups. The ϕ_{CT} indices (fixation indices that measure the genetic differentiation among groups) were significant and showed higher values as the *K* groups increased, but a decrease in the rate of increase of the ϕ_{CT} was meaningful after $K = 4$. For grouping $K = 4$, the genetic variance among groups was 71.0%, among localities within groups 18.29% and within localities 10.69% with $\phi_{CT} = 0.71$, $\phi_{SC} = 0.63$ and $\phi_{ST} = 0.89$ significant ($P < 0.01$). Each of the four groups identified by SAMOVA included the following localities: (i) all localities from western Pyrenees: 1, 2, 3, 4, 6 and 7; (ii) the localities from central Pyrenees: 8, 9, 10 and 11; (iii) the localities from eastern and central Pyrenees: 12, 13 and 16; (iv) all localities from eastern Pyrenees: 14, 15, 17, 21, 22, 23, 24, 25 and 26.

Mantel tests showed a significant correlation between genetic and geographic distances among localities ($r = 0.64$, $P < 0.01$), and also for the SAMOVA groups (i) and (iv) ($r = 0.71$ and 0.47, $P < 0.01$, respectively), as referred above, but not significant for the groups (ii) and (iii).

Phylogenetic inferences and mitochondrial lineage ages

Results of the different phylogenetic analyses conducted on the mitochondrial data set are summarized in Fig. 2. All analyses supported the monophyly of *H. ravastellus*. Alternative tree inference methods found different sister-group relationships of *H. ravastellus* with the outgroup taxa, although with low support, except for parsimony, which showed *Harpactocrates* n. sp. 'Cantabrian' as sister group to *H. ravastellus* with 83% jackknife support. All inference methods converged on almost identical topologies for the *H. ravastellus* haplotypes. Four well-supported clades (Fig. 2) with strong geographic signal (Fig. 1) were recognized: the Western clade included haplotypes from the westernmost localities (1 to 7), the Eastern clade included easternmost localities (14, 15 and 17 to 26), the Central-1 clade included haplotypes from central localities 8 to 11 and 14, while localities 12, 13, 15 and 16 were included in Central-2 clade. Different haplotype clades coexisted in two localities (14 and 15, see Fig. 1). The Eastern and Central-1 clades were sister groups, with high ML (>70%) and BI (>0.95) support, except for the parsimony analyses that yielded an unresolved polytomy at the

base of *H. ravastellus*. ML and BI analyses recovered the Western clade as sister to the Eastern + Central 1 clade. Time-calibrated trees obtained with BEAST and * BEAST without outgroups supported the same major clades but differed in the tree root location, which was assigned to the branch between the Eastern and Central-1 clades (Figs 4 and S1, Supporting information).

The statistical parsimony allele network from *srp54* sequences (Fig. 3) connected alleles by between 1 and 3 mutational steps. The A1 and A3 alleles were the most frequent. A1 was found in localities 21 to 26, 17 and 19, which correspond to the Eastern mitochondrial clade and A3 was found in all individuals from localities 1 to 13, which belong to the Western, Central-1 and Central-2 mitochondrial clades. Both alleles coexisted in the same localities that share more than one mitochondrial clade (14 and 15), as well as in locality 16. In two of these localities (15 and 16), we found individuals (qrav833241, qrav833242, qravLB341 and qravLB336) harbouring one of the two common alleles that are usually associated with another mitochondrial clade (see black and white arrows in Fig. 1). Heterozygous individuals were only present in localities 14, 16 and 19 (see Appendix S1, Supporting information).

The GMYC model showed a significantly better fit to the data than the null model of a single coalescent branching process ($P < 0.01$) and identified 16 coalescent groups (confidence interval 15–18), comprised of 12 clusters (confidence interval 12–14) and 4 singletons (see Fig. S2, Supporting information). GMYC clusters included individuals from a single or multiple (up to four) immediately adjacent localities. The only exceptions were localities 14 and 15, which include individuals from different haplotype clades and thus belong to divergent GMYC clusters.

The lineage age estimates based on an uncorrelated lognormal relaxed clock and Yule speciation process including one representative of each GMYC cluster are shown in Fig. 4. The time of the most recent common ancestor (TMRCA) of *H. ravastellus* haplotypes was dated at 2.88 Ma (3.47–2.31 Ma) and the TMRCA of each of the four main mitochondrial lineages ranged from ~0.75 to 1.5 Ma. The estimates of the TMRCA of *H. ravastellus* obtained from * BEAST were slightly older, 3.43 Ma (4.30–2.60 Ma). The TMRCA of populations 4 and 7 and populations 9 and 10 were dated at ~22 ky and ~48 ky, respectively (Fig. S1, Supporting information).

Testing for demographic hypotheses using ABC

The results of ABC-based analyses show that the BTM was generally favoured over the SNM in four western localities (4, 7, 9 and 10) and in the central Pyrenees population (13) (Fig. 5 and Table S2, Supporting information). Conversely, the IAM scenario was strongly supported for easternmost population 26. In population 25, the SNM had the highest *PP*, although Bayes factors do not rule out the IAM. Noticeably, for the other central populations 14 and 15 we were unable to test our demographic hypothesis because postsampling adjustment failed to generate the GLM. We found that for these sampling localities some of the statistics calculated from the observed data fell far outside the observed distribution range obtained in the simulated data (data not shown).

The approximate posterior estimates of model parameters indicate a severe bottleneck for populations 4, 7, 9, 10 and 13, which were reduced to <10% of their ancestral population sizes. The ABC analyses, however, did not allow a precise estimate of time parameters in the range of values used as priors, although the shape of the posterior distributions suggests that the bottleneck time should be considerably younger than the prior upper limit of 500 ky (data not shown). Regardless of the lack of precision in the estimation of some model parameters, our model choice procedure is powerful and accurate enough to be confident about the best-supported model for each of the populations. First, the power to distinguish among models is reasonably high, as BTM and SNM were correctly allocated in more than 70% of the cases in all populations (Figs S4A and S5A, Supporting information). Second, the models that were highly supported by the empirical data, BTM and IAM, but were incorrectly allocated by the ABC model choice, showed lower posterior probabilities supporting the wrong model substantially over those estimated for the studied populations (i.e. BTM and IAM were rarely selected with high probabilities when they were not the true model). Finally, although the ABC posterior probabilities were highly biased, the results of the ABC model choice for *H. ravastellus* data remained unaffected for all localities.

Given that bottleneck times could not be accurately estimated in ABC analyses of individual localities, we performed a prospective analysis in HBAYESSC (Chan *et al.* 2014) using the data from the five localities with strong bottleneck signals (each one putatively expanded from a single refugium) to test for putative synchronous expansions. The results of this analysis clearly point to the synchronous co-expansion (τ_e = 26 771 years; 95% quantiles = 3018–312 178) of the majority of these populations (the 95% quantiles of the posterior density of ζ are between 0.4 and 1 (e. g. between two and five co-expanding populations) while the mode of its posterior density is close to 1 (e.g. all five populations co-expanded at the same time), suggesting a clear effect of glaciation cycles on the observed local demographic patterns.

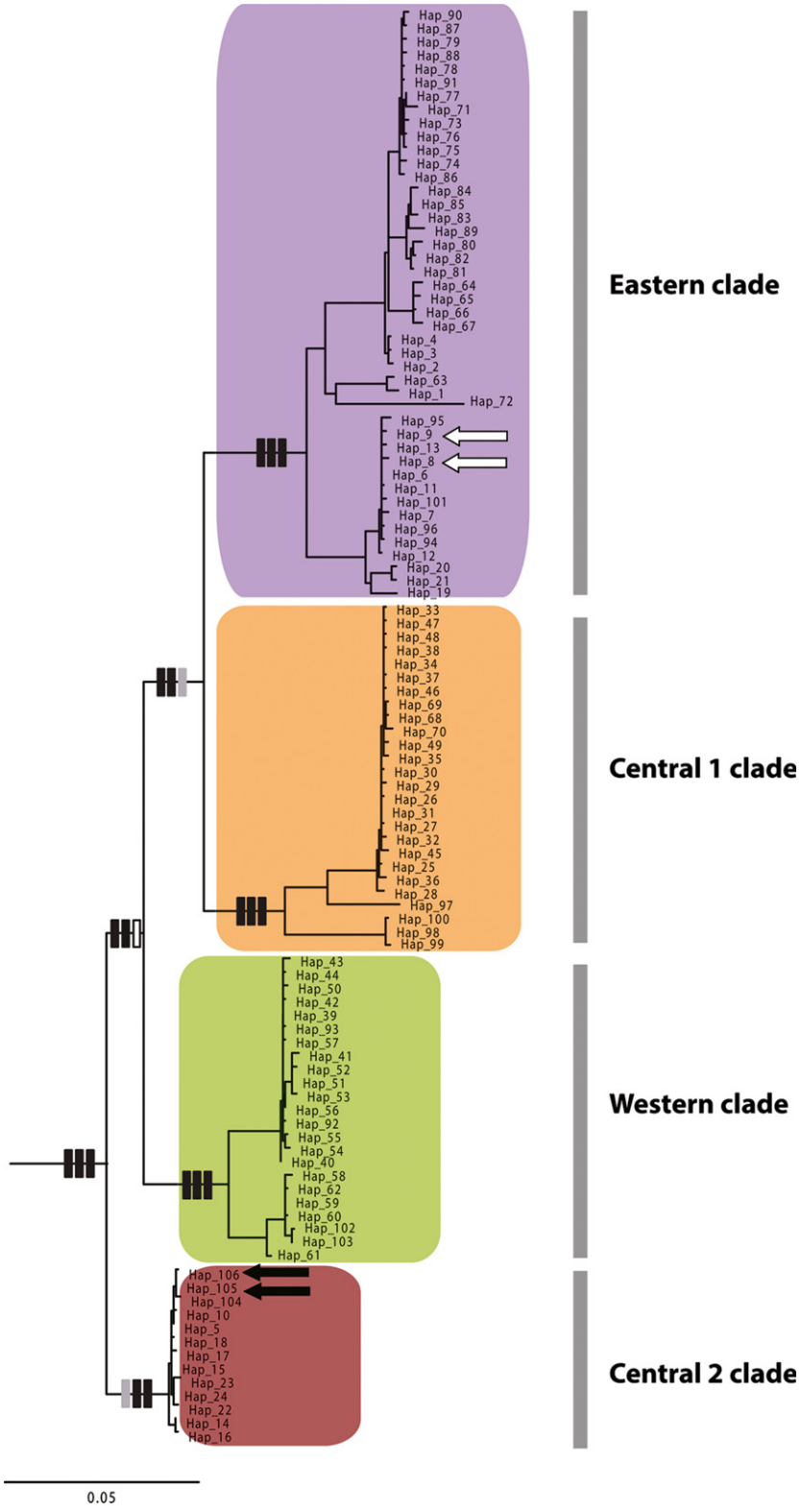

Fig. 2 Maximum-likelihood tree of *Harpactocrates ravastellus* mitochondrial haplotypes, obtained from the preferred partition scheme (full codon by gene partition scheme). Bars on branches indicate support from alternative analyses (from left to right): maximum likelihood (ML), Bayesian inference and parsimony (P). Black bar: clade supported by ML bootstrap and P jackknife >70% and Bayesian PP >0.95. Grey bar: clade recovered but with support below the threshold values above. White bar: clade not recovered in the analyses. The main mitochondrial lineage colours correspond to geographic distribution as shown in Fig. 1. Arrows point to haplotypes that were present in individuals bearing alleles more common in other haplotype lineages (A1: black arrow and A3: white arrow, see Fig. 1 and Table 1). The outgroups *Harpactocrates radulifer*, and *Harpactocrates apennicola*, *Harpactocrates intermedius* and *Harpactocrates* n. sp. have been trimmed from the tree.

Climatic niche and distribution modelling

Schoener's D index, which measures the overlap of climatic niches of the eastern and western-central Pyrenees, was 0.491. For both groups, the niche similarity test was significant (P = 0.0198), indicating that both niches are reciprocally more similar than expected by chance (Fig. S5, Supporting information).

Five hundred SDMs were generated, which were projected and averaged for each time frame. The averaged AUC was 0.93, indicating an excellent model performance (Swets 1988). The winter mean temperature was the most important variable in determining the species climate range (89.4%), followed by annual precipitation range (4%), summer mean precipitation (3.7%) and annual temperature range (2.9%). The winter mean temperature for the present species data was 2.89 °C, whereas for the defined study area it was 5.6 °C.

Summary maps obtained by summing binary maps across all predicted SDMs projected at eleven time frames (4 Holocene and 7 Pleistocene time frames) showed almost no stable area for the species throughout the Pyrenees since the Late Pleistocene. The predicted stable areas in the Pyrenees were located in small areas near the westernmost Pyrenees. Only two presence records fall in the stable area located in Puerto de Urkiaga and Puerto de Ibañeta (localities 1 and 2, Fig. 1). Other predicted stable areas were located in the Natural Park of Montseny (a mountain southeast of the Pyrenees and parallel to the Mediterranean Sea coast in Catalonia), and in Picos de Europa and País Vasco mountains, which form part of the Cantabrian mountain ranges. *H. ravastellus* has never been collected in the Natural Park of Montseny in spite of sampling effort, whereas for the Cantabrian mountain ranges the sister species of *H. ravastellus* has been collected. Fig. 6 shows the summary maps for four Holocene (present, 4, 8 and 12 ky BP) and for seven Pleistocene time frames (16, 12, 24, 28, 32, 36 and 40 ky BP) according to both criteria of threshold selection. The summary maps show a striking difference between the predicted species occurrences during the Holocene and Pleistocene. These differences were recovered for both thresholds used to create the binary maps.

Fig. 3 Statistical parsimony network of the *srp54* alleles of *Harpactocrates ravastellus*. The area of the pie plots is proportional to the number of individuals found for each allele. The number of individuals bearing each allele is included in brackets. Alleles are coloured in grey scale according to geographic distribution, as shown in Fig. 1.

Morphological analysis

Morphological examination of 49 adult individuals used for molecular analyses revealed subtle but constant differences in both male and female genitalia. Two male copulatory bulb (a modification of the palpal tarsus) morphologies were observed (hereafter referred as type A and type B), which mostly differed in the length and shape of the embolus tip (Fig. 7). Type A corresponds to the original drawings by Simon (1914) as well as those appearing in a subsequent revision by Ferrández (1986). Similarly, two different vulva shapes were also recognized among the studied material (Fig. 7, hereafter referred as type A and type B). The female of *H. ravastellus* was not illustrated in the original description of the species, but was subsequently included in a revision of Iberian *Harpactocrates* (Ferrández 1986). This illustration corresponds to our type B. However, the

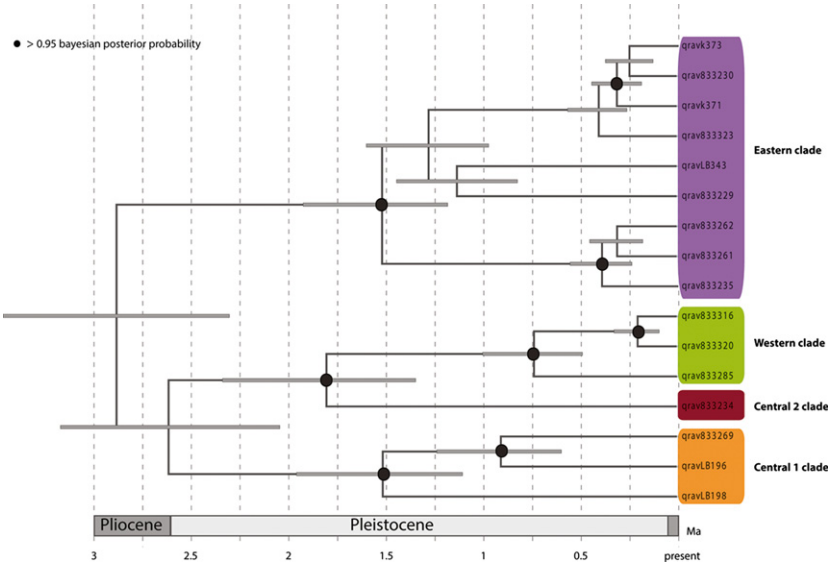

Fig. 4 Chronogram inferred from the mitochondrial genes with BEAST under partition scheme by gene and relaxed lognormal clock, including only one individual from each of the 16 GMYC clusters found. Bars at nodes indicate 95% HPD interval. Colours denote the main mitochondrial lineages as in Figs. 1 and 2.

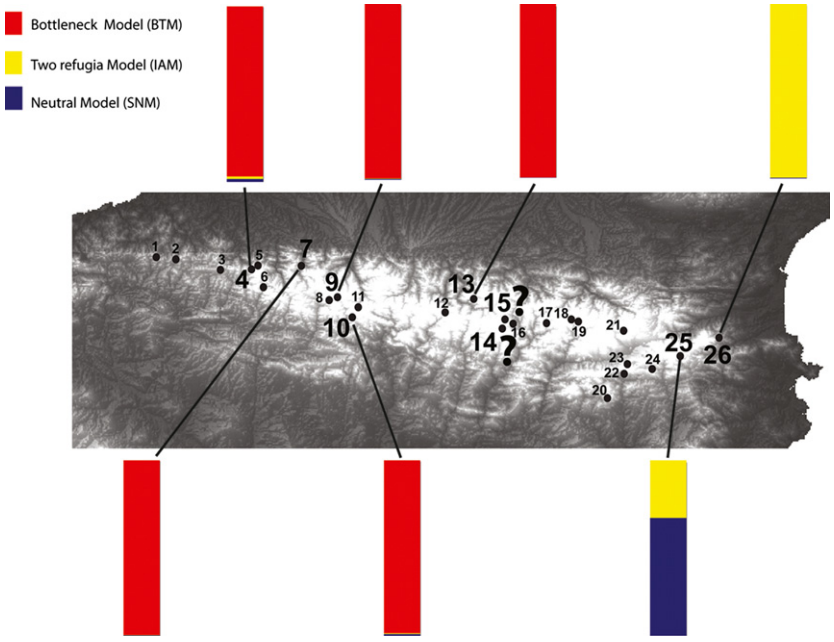

Fig. 5 Map showing the study area and sampled locations. The diagram represents the support of the empirical data for the demographic models in the ABC analysis. Bars of different colours denote the posterior probability estimated for each of the three models in a particular population.

published drawings of the male bulb and of the female vulva do not seem to match the distribution we found of both types. Males exhibiting the bulb type illustrated in the original description (type A) co-occurred with females with a vulva shape that differs from the one illustrated in the literature. Conversely, males with the new bulb type (type B) coexisted with the females with the vulva illustrated in Ferrández (1986) (Appendix S1, Supporting information).

We found a strong correlation between genitalic type and the nuclear genotype (Appendix S1, Supporting information): all individuals with genitalic type A (either male or female) bore allele A1, while all individuals with genitalic type B bore allele A3. Interestingly, the only exceptions corresponded to three individuals collected in localities from the contact zone, namely Les Planes de Son and Estany de Gerber, which combined a genitalic type B with an allele A1 or, in one of the specimens, a heterozygotic A3/A7.

Discussion

Mountain systems have played a key role in generating arthropod diversity, both at the intra- and interspecific levels (Garrick 2011). In this study, the integration of phylogeographic analyses, morphological evidence and species distribution modelling techniques allowed us to propose that *Harpactocrates ravastellus* may include two different species with a narrow contact zone, and that some of their populations survived glacial cycles *in situ*, probably by retreating to deep valleys within the mountain range.

Harpactocrates ravastellus: not one but two species with a contact zone?

Phylogenetic and population analyses of the mitochondrial genes revealed four well-supported groups in *H. ravastellus* distributed along a longitudinal gradient. Although inferences of relationships among the groups were sensitive to the inclusion of outgroups, the mitochondrial and the multilocus Bayesian analyses (i.e. * BEAST) identified the separation between the Eastern and the remaining groups as the deepest split within the lineage (see Figs 4 and S1, Supporting information). Similar western/eastern splits in the Pyrenees have been detected in the European beech, *Fagus sylvatica* (Magri *et al.* 2006; Magri 2008), the mountain ringlet butterfly *Erebia epiphoron* (Schmitt *et al.* 2006) and the snapdragon *Antirrhinum* (Liberal *et al.* 2014).

Further morphological analyses of male and female genitalia provided diagnostic differences for these two main lineages (see Fig. 7, hereafter referred as the Eastern and Central-Western lineages). It is worth noting that one of the vulva types found (type A) was illustrated in the revision of the Iberian *Harpactocrates* (Ferrández 1986), but was wrongly assigned to a different species, *H. radulifer*. We suspect that this misidentification explains why *H. radulifer* had been formerly reported from the Pyrenees, as we have not been able to find any specimens of this species in our thorough sampling of this mountain range. The combination of genetic and morphological evidence, along with the geographic separation (see Figs 1 and 7), points towards

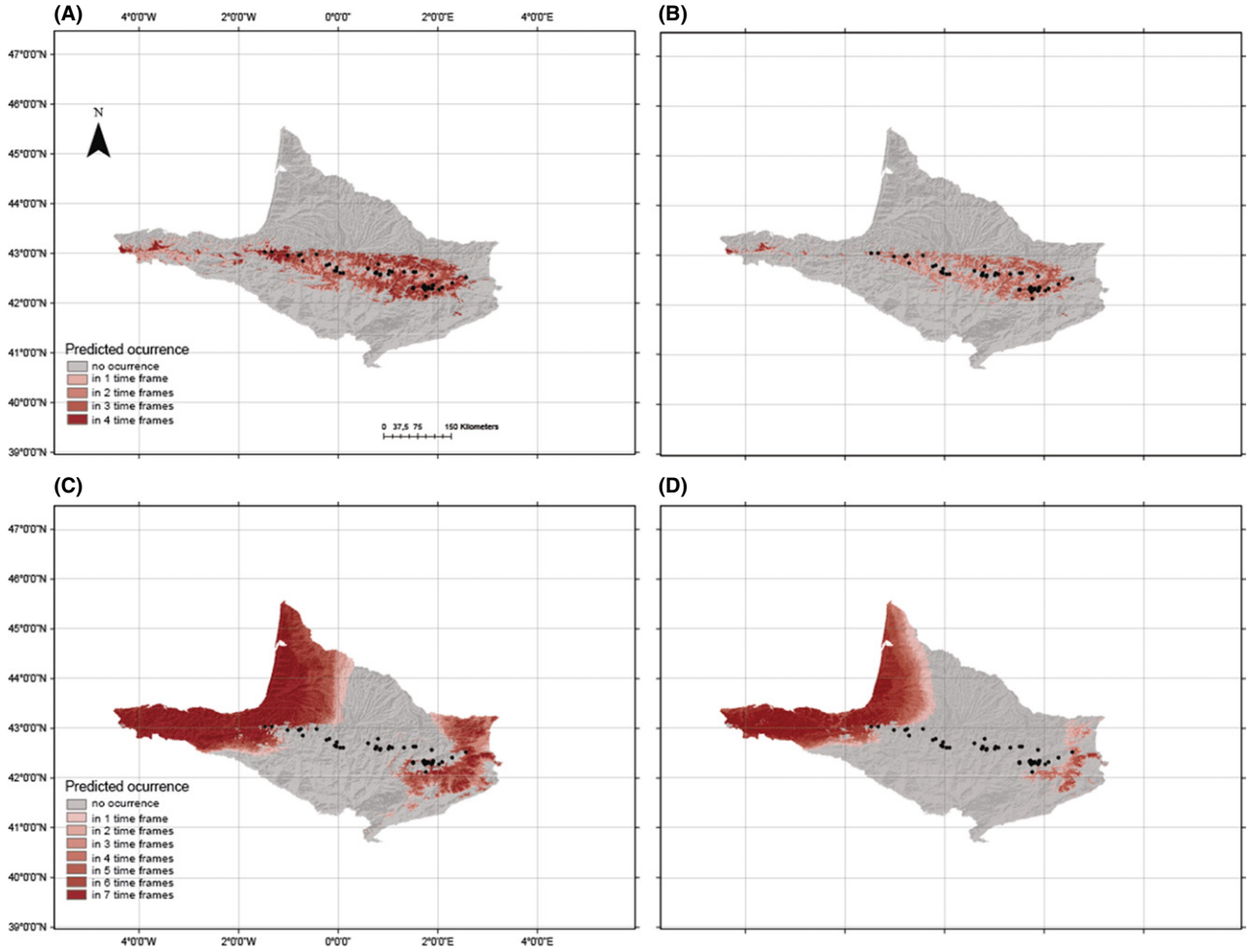

Fig. 6 Summary maps of predicted species occurrences (binary maps) during the Holocene (present, 4, 8 and 12 ky BP) and Pleistocene time frames (16, 20, 24, 28, 32, 36 and 40 ky BP) according to two thresholds for probability of occurrence (MaxSens + Spec threshold and the default threshold set at 0.5). (A) and (B) correspond to summary maps for the Holocene and Pleistocene for the MaxSens + Spec threshold, respectively, and (C) and (D) for Holocene and Pleistocene for the default threshold, respectively. Reference colours correspond to agreement areas for occurrences across the different time frames. Black dots denote all localities used to construct the models.

the existence of two independently evolving lineages that may deserve formal species description.

The two lineages seem to maintain a narrow contact zone. In localities 14, 15 and 16, which approximately correspond to the limits of the Aigüestortes i Estany de Sant Maurici National Park in northwestern Catalonia, we detected the coexistence of divergent mitochondrial haplotypes and nuclear alleles, with the presence of heterozygotic individuals and others showing evidence of mitochondrial introgression. In theses localities, we also find the only exceptions to the strong correlations observed between the genitalic type and the nuclear genotype. Sampling at finer geographic scales will be required to fully delimit the extent of the contact zone.

Although the diverse topography of the Pyrenean mountain range offered ample opportunities for allopatric isolation, it is difficult to identify any particular geographic barrier to pinpoint a vicariance event between the Eastern and Central-Western lineages (see Figs 1 and 4). Similarly, niche similarity tests based on climatic variables did not reveal any differences between the Eastern and Central-Western lineages. However, the projection of species occurrences into the past revealed that most of the Pyrenees was a region of great instability for *H. ravastellus* during cooling periods. The SDMs suggest that the lineage underwent a major range shift towards both extremes of the Pyrenees to track suitable climatic habitats, defining a Mediterranean and an Atlantic refugium, possibly outside the mountain range. Although our demographic history inferences support the persistence of some populations throughout the glaciations within the mountain

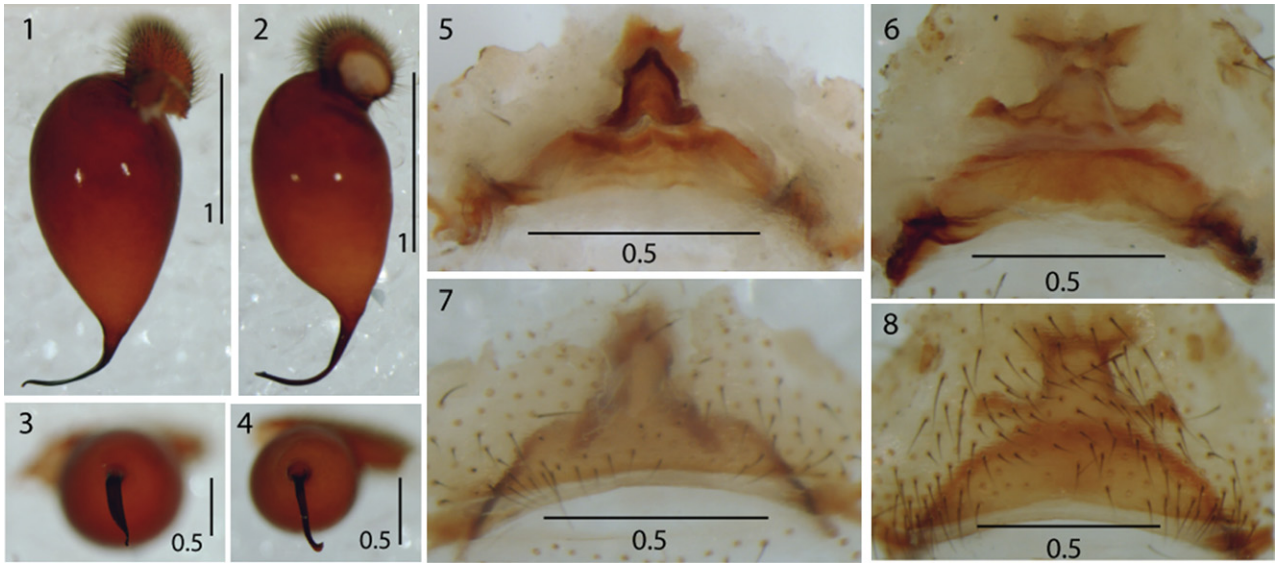

Fig. 7 Type A and B of male and female genitalia. Left male bulb: (1) and (2) correspond to retrolateral view of genitalic type A and B, respectively; (3) and (4) to ventral view of genitalic type A and B, respectively. Female vulva: (5) and (6) correspond to dorsal view of genitalic type A and B, respectively; (7) and (8) to ventral view of genitalic type A and B, respectively. Scale bar is in millimetres.

range (see next section), optimal conditions during glacial maxima existed mostly at both longitudinal extremes of the range which supports a vicariant origin of the two lineages and the subsequent formation of a hybrid zone resulting from postglacial recolonization of the formerly glaciated areas.

Although little is known about the exact dating, and the size of the glaciers and ice sheets during the Early Pleistocene glaciations (2.6–0.78 Ma), there is ample evidence that major cooling events in the Mediterranean region started between 2.8 and 2.5 Ma (Gibbard *et al.* 2010). The time of the split between the eastern and central-western lineages approximately corresponded to the onset of the Early Pleistocene glaciations (3.47–2.31 Ma and 4.30–2.60 Ma, for the mitochondrial and multispecies coalescent approaches, respectively). Therefore, to explain such divergence times between the lineages, we should assume that similar range shifts as inferred from our SDMs in the Late Pleistocene occurred repeatedly throughout the Quaternary glacial cycles.

Harpactocrates ravastellus endured glaciations within the Pyrenees

Genetic data revealed that most localities are strongly differentiated as reflected by the high values and significance of almost all ϕ_{ST} comparisons. The few comparisons that did not show significant differentiation were among neighbouring localities. The high level of haplotype diversity (0.99 for the entire lineage) and the 16 GMYC clusters inferred confirm the strong spatial structuring.

Neutrality tests and ABC-based analyses provide some hints on the depth of population differentiation. In several cases (three populations from western Pyrenees, two from central Pyrenees and one from eastern Pyrenees), neutrality tests revealed significant departures of the standard neutral model compatible, among others, with past reductions in the effective population size followed by population growth (Tajima 1989a; Fu 1997). Although some of these signatures could be explained by positive selection, the former demographic hypothesis seems a more plausible explanation under a glacial cycle scenario. The ABC analyses further supported a scenario of local bottlenecks. The ABC results showed that in at least five of the nine populations analysed the bottleneck model (see Fig. 5) is clearly the best model with very high posterior probabilities. In addition, the high fraction of the retained simulations with a smaller or equal likelihood than the observed data under the estimated GLM (*f* in Table S2, Supporting information) demonstrate that this scenario is not only the most probable among competing models but also is capable of reproducing the observed data with high probability in these populations. Lastly, and most importantly, validation results demonstrate that ABC model choice is powerful enough (and unbiased) to distinguish among alternative models, even with current limited marker sampling (Figs S3 and S4, Supporting information). Although similar bottlenecks could also be generated from a dispersal limited species in the

absence of glaciation cycles, the indications of concerted demographic history across the five localities obtained in the hABC analysis would be more likely under a glacial cycle's survival scenario (i.e. concerted response to climate changes). Although further loci need to be sequenced to corroborate the impact of dispersal, the survival scenario seems to be the most plausible explanation for the observed patterns in the available data.

For the two populations (14 and 15) that could not be accommodated by the ABC analyses, the particular distribution of variation is difficult to explain without invoking a succession of more complex scenarios of diversity reduction and admixture, probably as the result of their location at the contact zone. Interestingly, the only population that was found to be in standard neutral model (25) is located in the southeastern of the Pyrenees, an area that was weakly glaciated, while the remaining populations are in areas that were glaciated extensively (Calvet 2004). Similar patterns of strong genetic differentiation among Pyrenean populations (i.e. high pairwise F_{ST} values between localities and/or between main clusters, and high frequencies of fixed alleles) have also been reported by Charrier *et al.* (2014) in the alpine shrub *Rhododendron ferrugineum*, by Milá *et al.* (2010) in the newt *Calotriton asper*, and by Lihová *et al.* (2009) in the alpine plant *Cardamine alpina*. These studies also suggested that populations would have survived glacial periods in high-elevation refuges. The concordance of these phylogeographic patterns further supports the survival scenario of *H. ravastellus* in microrefugia within the Pyrenees.

These results seem to be at odds with the scenario derived from the projection into the past of the species distribution models, which suggest that the distribution range was shifted towards suitable habitats near the Atlantic and Mediterranean coastal ranges from which the central part of the Pyrenees was recolonized during the Holocene. The severity of population bottlenecks and the strong population differentiation, which clearly pre-date the last glacial maximum (LGM), is more compatible with the glacial survival of some populations in small refugia within the mountain range. The coarse scale (1 km grain size) used to model the species distribution may explain the lack of congruence between demographic histories and SDMs, because fine details of climatic diversity in the Pyrenees would have been insufficiently resolved. Alternatively, the inability of SDMs to detect narrow refugia within the Pyrenees could be the result of not considering potential changes in the species' niches through time (Pearman *et al.* 2008). However, this explanation seems unlikely because SDMs predicted the distribution of the species in an area inhabited by a close relative with similar habitat preferences, from which it split about 8 Ma (Bidegaray-Batista *et al.* 2014). Although SDMs failed to detect microrefugia within the range, it provides evidence for the great instability of most of the Pyrenees during the climatic oscillations, which ultimately led to nonequilibrium populations and lineages as indicated by demographic indices.

Changes in vegetation recorded from pollen data in the Pyrenees during Pleistocene climatic oscillations (Jalut *et al.* 1992) provide further support for the *in situ* survival of *H. ravastellus* populations. Pollen data show a dominance of steppe vegetation (*Artemisia*) during glacial maxima, while tree pollen of species such as *Abies*, *Fagus*, *Quercus* and *Pinus*, plant communities favoured by *H. ravastellus*, was also present. Glacial regressions were characterized by an increase of tree pollen while brief cooler periods witnessed a regression of *Pinus* and the re-advance of steppe vegetation. This pattern was more pronounced in areas near the Mediterranean basin than near the Atlantic Ocean (Elenga *et al.* 2000; Jiménez-Moreno *et al.* 2010). It is possible that during cold and dry periods *H. ravastellus* populations experienced range contractions into small refugia where suitable forest habitat remained, probably on southern valley slopes, while during interglacials, populations would expand uphill and towards the central Pyrenees.

The combination of molecular tools with species distribution modelling techniques revealed high genetic structure correlated with geographic distance, at different hierarchical levels (i.e. major clades, GMYC groups and allele frequencies) in *H. ravastellus*. At first sight, the demographic results seem to stand at odds with those inferred from the niche-based species distribution modelling analysis, although both suggest limited expansion and great instability for *H. ravastellus* during cooling periods. While the climate niche models suggest two major refugia outside the mountain range, the demographic analyses indicate survival of population in microrefugia within the mountain range. As a solution to this apparent conundrum, we propose an evolutionary scenario in which geographically close populations took refuge in the main Atlantic and Mediterranean lowland refugia, while populations within the mountain range sheltered in local microrefugia. The recurrent episodes of population contractions, local extinctions and limited dispersal driven by cyclic glacial events would have eventually shaped the complex present day pattern recovered in *H. ravastellus*.

In conclusion, our study suggests the existence and location of lowland and peripheral refugia where a Pyrenean species endured glacial periods. Future research coupling genetic data and species distribution models in other terrestrial organisms with contrasting biological traits will shed light on the generality of the

observed patterns and provide a more complete understanding of the role of climatic changes in shaping the highly endemic and diverse biota of the southern European mountain ranges.

Acknowledgements

We are grateful to N. Macías-Hernández and the late M. Mejía-Chang for their assistance in the field. We also thank S. de la Cruz, E. de Mas and E. Mateos, for providing us with additional samples, and S. Ramos-Onsins for his assistance with the coalescent simulation software. This project was funded by the Spanish Ministry of Science and Innovation (MICINN) grant CGL2006-08617 (MA) and by a SYNTHESYS project funded by the European Community Research Infrastructure Action under the FP6 'Structuring the European Research Area' Programme (to LBB). LBB was supported by a graduate fellowship (FI-DGR 2009) and a short stay grant (BE 2010) from the Generalitat de Catalunya. Further support was provided by Sistema Nacional de Investigadores (Agencia Nacional de Investigación e Innovación) and Programa de Desarrollo de las Ciencias Básicas to LBB, and by an ICREA Academia award for excellence in research from the Generalitat de Catalunya to MA.

References

Anderson CN, Ramakrishnan U, Chan YL, Hadly EA (2005) Serial SimCoal: a population genetics model for data from multiple populations and points in time. *Bioinformatics*, **21**, 1733–1734.

Alarcón M, Vargas P, Sáez L, Molero J, Aldasoro JJ (2012) Genetic diversity of mountain plants: two migration episodes of Mediterranean *Erodium* (Geraniaceae). *Molecular Phylogenetics and Evolution*, **63**, 866–876.

Arenas M, Ray N, Currat M, Excoffier L (2012) Consequences of range contractions and range shifts on molecular diversity. *Molecular Biology and Evolution*, **29**, 207–218.

Barraclough TG, Vogler AP (2002) Recent diversification rates in North American tiger beetles estimated from a dated mtDNA phylogenetic tree. *Molecular Biology and Evolution*, **19**, 1706–1716.

Bidegaray-Batista L, Arnedo MA (2011) Gone with the plate: the opening of the Western Mediterranean basin drove the diversification of ground-dweller spiders. *BMC Evolutionary Biology*, **11**, 317.

Bidegaray-Batista L, Macías-Hernández N, Oromí P, Arnedo MA (2007) Living on the edge: demographic and phylogeographical patterns in the woodlouse-hunter spider *Dysdera lancerotensis* Simon, 1907 on the eastern volcanic ridge of the Canary Islands. *Molecular Ecology*, **16**, 3198–3214.

Bidegaray-Batista L, Ferrández MA, Arnedo MA (2014) Winter is coming: Miocene and Quaternary climatic shifts shaped the diversification of Western-Mediterranean *Harpactocrates* (Araneae, Dysderidae) spiders. *Cladistics*, **30**, 428–446.

Broennimann O, Fitzpatrick MC, Pearman PB *et al.* (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481–497.

Calvet M (2004) The Quaternary glaciation of the Pyrenees. *Developments in Quaternary Science*, **2**, 119–128.

Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C (2009) Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, **323**, 785–789.

Carstens BC, Richards CL (2007) Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution*, **61**, 1439–1454.

Chan YL, Schanzenbach D, Hickerson MJ (2014) Detecting concerted demographic response across community assemblages using hierarchical approximate Bayesian computation. *Molecular Biology and Evolution*, **31**, 2501–2515.

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.

Chu JH, Wegmann D, Yeh CF *et al.* (2013) Inferring the geographic mode of speciation by contrasting autosomal and sex-linked genetic diversity. *Molecular Biology and Evolution*, **30**, 2519–2530.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Cooke JAL (1965) Spider Genus *Dysdera* (Araneae, Dysderidae). *Nature*, **205**, 1027–1028.

De Mas-Castroverde E (2007) Evaluación y predicción de la biodiversidad. Un modelo con araneidos en el Parque Natural del Cadí-Moixeró. Ph.D. dissertation, Universitat de Barcelona.

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

Drummond AJ, Ashton B, Buxton S *et al.* (2010) Geneious v.5.3. URL:http://www.geneious.com.

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

Elenga H, Peyron O, Bonnefille R *et al.* (2000) Pollen-based biome reconstruction for southern Europe and Africa 18,000 yr BP. *Journal of Biogeography*, **27**, 621–634.

Espíndola A, Pellissier L, Maiorano L *et al.* (2012) Predicting present and future intra-specific genetic structure through niche hindcasting across 24 millennia. *Ecology Letters*, **15**, 649–657.

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

Farris JS, Albert VA, Kallersjo M, Lipscomb D, Kluge AG (1996) Parsimony jackknifing outperforms neighbor-joining. *Cladistics*, **12**, 99–124.

Ferrández MA (1986) Las especies ibéricas del género *Harpactocrates* Simon 1914 (Araneida: Dysderidae). *X International Congress of Arachnology*, **1**, 337–348.

Fontaneto D, Herniou EA, Boschetti C *et al.* (2007) Independently evolving species in asexual bdelloid rotifers. *PLoS Biology*, **5**, e87.

Freeman EA, Moisen G (2008) PresenceAbsence: an R package for presence absence analysis. *Journal of Statistical Software*, **23**, 31.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Galbreath KE, Cook JA, Eddingsaas AA, DeChaine EG (2011) Diversity and demography in Beringia: multilocus test of

paleodistribution models reveal the complex history of arctic ground squirrels. *Evolution*, **65**, 1879–1896.

Gante HF, Micael J, Oliva-Paterna FJ *et al.* (2009) Diversification within glacial refugia: tempo and mode of evolution of the polytypic fish *Barbus sclateri*. *Molecular Ecology*, **18**, 3240–3255.

Garcia-Porta J, Litvinchuk SN, Crochet PA *et al.* (2012) Molecular phylogenetics and historical biogeography of the west-palearctic common toads (*Bufo bufo* species complex). *Molecular Phylogenetics and Evolution*, **63**, 113–130.

Garrick RC (2011) Montane refuges and topographic complexity generate and maintain invertebrate biodiversity: recurring themes across space and time. *Journal of Insect Conservation*, **15**, 469–478.

Gibbard PL, Head MJ, Walker MJC (2010) Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. *Journal of Quaternary Science*, **25**, 96–102.

Gillespie RG, Baldwin BG, Waters JM *et al.* (2012) Long-distance dispersal: a framework for hypothesis testing. *Trends in Ecology & Evolution*, **27**, 47–56.

Goloboff PA, Farris JS, Nixon KC (2008) TNT, a free program for phylogenetic analysis. *Cladistics*, **24**, 774–786.

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography in Southern European Refugia* (eds Weiss S, Ferrand N), pp. 155–188. Springer, Dordrecht.

Graham CH, Ron SR, Santos JC, Schneider CJ, Moritz C (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution*, **58**, 1781–1793.

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

Gutierrez Larena B, Fuertes Aguilar J, Nieto Feliner G (2002) Glacial-induced altitudinal migrations in *Armeria* (Plumbaginaceae) inferred from patterns of chloroplast DNA haplotype sharing. *Molecular Ecology*, **11**, 1965–1974.

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

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

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology*, **10**, 537–550.

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.

Hodges KM, Rowell DM, Keogh JS (2007) Remarkably different phylogeographic structure in two closely related lizard species in a zone of sympatry in south-eastern Australia. *Journal Of Zoology*, **272**, 64–72.

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.

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

Hughes PD, Woodward JC, Gibbard PL (2006) Quaternary glacial history of the Mediterranean mountains. *Progress in Physical Geography*, **30**, 334–364.

Jalut G, Marti JM, Fontugne M *et al.* (1992) Glacial to interglacial vegetation changes in the northern and southern Pyrenees: deglaciation, vegetation cover and chronology. *Quaternary Science Reviews*, **11**, 449–480.

Jarman SN, Ward RD, Elliott NG (2002) Oligonucleotide primers for PCR amplification of coelomate introns. *Marine Biotechnology*, **4**, 347–355.

Jiménez-Moreno G, Fauquette S, Suc JP (2010) Miocene to Pliocene vegetation reconstruction and climate estimates in the Iberian Peninsula from pollen data. *Review of Palaeobotany and Palynology*, **162**, 403–415.

Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**, 3059–3066.

Knowles LL (2000) Tests of pleistocene speciation in montane grasshoppers (Genus *Melanoplus*) from the sky islands of western north America. *Evolution*, **54**, 1337–1348.

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

Leuenberger C, Wegmann D (2010) Bayesian computation and model selection without likelihoods. *Genetics*, **184**, 243–252.

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

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

Lihová J, Carlsen T, Brochmann C, Marhold K (2009) Contrasting phylogeographies inferred for the two alpine sister species *Cardamine resedifolia* and *C. alpina* (Brassicaceae). *Journal of Biogeography*, **36**, 104–120.

Magri D (2008) Patterns of post-glacial spread and the extent of glacial refugia of European beech (*Fagus sylvatica*). *Journal of Biogeography*, **35**, 450–463.

Magri D, Vendramin GG, Comps B *et al.* (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, **171**, 199–221.

Maiorano L, Cheddadi R, Zimmermann NE *et al.* (2013) Building the niche through time: using 13,000 years of data to predict the effects of climate change on three tree species in Europe. *Global Ecology and Biogeography*, **22**, 302–317.

Marske KA, Leschen RAB, Buckley TR (2011) Reconciling phylogeography and ecological niche models for New Zealand beetles: Looking beyond glacial refugia. *Molecular Phylogenetics and Evolution*, **59**, 89–112.

McPeek MA, Holt RD (1992) The evolution of dispersal in spatially and temporally varying environments. *American Naturalist*, **140**, 1010–1027.

Médail F, Diadema K (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, **36**, 1333–1345.

Médail F, Quezel P (1997) Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin. *Annals of the Missouri Botanical Garden*, **84**, 112–127.

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.

Milá B, Surget-Groba Y, Heulin B, Gosá A, Fitze PS (2013) Multilocus phylogeography of the common lizard *Zootoca vivipara* at the Ibero-Pyrenean suture zone reveals lowland barriers and high-elevation introgression. *BMC Evolutionary Biology*, **13**, 192.

Miller MA, Holder MT, Vos R *et al.* (2009) The CIPRES Portals. URL:http://www.phylo.org/sub_sections/portal.

Miraldo A, Hewitt GM, Paulo OS, Emerson BC (2011) Phylogeography and demographic history of *Lacerta lepida* in the Iberian Peninsula: multiple refugia, range expansions and secondary contact zones. *BMC Evolutionary Biology*, **11**, 170.

Mitchell TD, Carter TR, Jones PD, Hulme M, New M (2004) A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100). *Tyndall Centre for Climate Change Research Working Paper*, **55**, 1–25.

Monaghan MT, Wild R, Elliot M *et al.* (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology*, **58**, 298–311.

Mouret V, Guillaumet A, Cheylan M *et al.* (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.

Muster C, Berendonk TU (2006) Divergence and diversity: lessons from an arcticalpine distribution (*Pardosa saltuaria* group, Lycosidae). *Molecular Ecology*, **15**, 2921–2933.

Nathan R, Perry G, Cronin JT, Strand AE, Cain ML (2003) Methods for estimating long-distance dispersal. *Oikos*, **103**, 261–273.

Nixon KC (2002) *WinClada*. Published by the Author, Ithaca, New York.

Papadopoulou A, Anastasiou I, Keskin B, Vogler AP (2009) Comparative phylogeography of tenebrionid beetles in the Aegean archipelago: the effect of dispersal ability and habitat preference. *Molecular Ecology*, **18**, 2503–2517.

Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Molecular Biology and Evolution*, **27**, 1659–1672.

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

Peter BM, Wegmann D, Excoffier L (2010) Distinguishing between population bottleneck and population subdivision by a Bayesian model choice procedure. *Molecular Ecology*, **19**, 4648–4660.

Peterson AT (2009) Phylogeography is not enough: The need for multiple lines of evidence. *Frontiers of Biogeography*, **1**, 19–25.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Pons J, Barraclough T, Gomez-Zurita J *et al.* (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, **55**, 595–609.

Ramos-Onsins SE, Mitchell-Olds T (2007) Mlcoalsim: multilocus coalescent simulations. *Evolutionary Bioinformatics Online*, **3**, 41–44.

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

Richards CL, Carstens BC, Knowles LL (2007) Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *Journal of Biogeography*, **34**, 1833–1845.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Schmitt T (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, **4**, 1742–9994.

Schmitt T (2009) Biogeographical and evolutionary importance of the European high mountain systems. *Frontiers in Zoology*, **6**, 1–10.

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.

Schorr G, Holstein N, Pearman PB, Guisan A, Kadereit JW (2012) Integrating species distribution models (SDMs) and phylogeography for two species of Alpine *Primula*. *Ecology and Evolution*, **2**, 1260–1277.

Schorr G, PrB P, Guisan A, Kadereit JW (2013) Combining palaeodistribution modelling and phylogeographical approaches for identifying glacial refugia in Alpine Primula. *Journal of Biogeography*, **40**, 1947–1960.

Schoville SD, Roderick GK (2009) Alpine biogeography of Parnassian butterflies during Quaternary climate cycles in North America. *Molecular Ecology*, **18**, 3471–3485.

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology*, **49**, 369–381.

Simon E (1914) Les Arachnides de France. Tome VI, 1ère partie, París.

Simon C, Frati F, Beckenbach A *et al.* (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, **87**, 651–701.

Singarayer JS, Valdes PJ (2010) High-latitude climate sensitivity to ice-sheet forcing over the last 120kyr. *Quaternary Science Reviews*, **29**, 43–55.

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

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

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

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

Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285–1293.

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

Tajima F (1989a) The effect of change in population size on DNA polymorphism. *Genetics*, **123**, 597–601.

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

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

Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010) ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*, **11**, 116.

Young ND, Healy J (2002) GapCoder. http://www.trinity.edu/nyoung/GapCoder/Download.html.

M.A.A. and L.B.B. designed the study. Specimens were collected by L.B.B. and M.A. or kindly provided by colleagues. L.B.B. did the molecular work and morphological examination. L.B.B. conducted the phylogenetic and population genetic analyses, and A.S.G. conducted the ABC-based analyses. M.A.A. and A.P.V. supervised the laboratory work and molecular genetic analyses. L.B.B., G.S. and L.M. conducted the S.D.Ms, and A.G. supervised the analyses. L.B.B. wrote the first draft, and M.A.A., A.P.V., A.G., A.S.G., G.S. and L.M. improved successive versions. All authors read and approved the final manuscript.

Data accessibility

DNA sequences: deposited in GenBank (Accession nos. KU552358-KU552898); updated in the GenBank (Accession nos. KM219492-KM219493, KM219495-KM219500); downloaded from GenBank (Accession nos. AF244235/JN689146, AF244150/ EU139682, KM219441-KM219433, KM219505, KM219513 KM219514, JN689145, JN705762). Original sequence alignments, MAXENT input files and specimens photographs were deposited in the Dryad data repository: doi:10.5061/dryad.k3v8j.

Supporting information

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

Appendix S1 Sampled localities of *Harpactocrates ravastellus* and haplotype/allele frequencies.

Table S1 Localities used for the SDM analyses.

Table S2 ABC-GLM results.

Fig. S1 Chronogram obtained from both mitochondrial and nuclear genes with *BEAST.

Fig. S2 Ultrametric tree obtained from the mitochondrial genes.

Fig. S3 Results of the ABC model choice validation for locality 4.

Fig. S4 Results of the ABC model choice validation for locality 26.

Fig. S5 Niche of the *Harpactocrates ravastellus* specimens in climatic space.