

RESEARCH ARTICLE

Open Access

# Ecological niche and phylogeography elucidate complex biogeographic patterns in *Loxosceles rufescens* (Araneae, Sicariidae) in the Mediterranean Basin

Enric Planas<sup>1</sup>, Erin E Saupe<sup>2,3</sup>, Matheus S Lima-Ribeiro<sup>4</sup>, A Townsend Peterson<sup>3</sup> and Carles Ribera<sup>1\*</sup>

## Abstract

**Background:** Understanding the evolutionary history of morphologically cryptic species complexes is difficult, and made even more challenging when geographic distributions have been modified by human-mediated dispersal. This situation is common in the Mediterranean Basin where, aside from the environmental heterogeneity of the region, protracted human presence has obscured the biogeographic processes that shaped current diversity. *Loxosceles rufescens* (Araneae, Sicariidae) is an ideal example: native to the Mediterranean, the species has dispersed worldwide via cohabitation with humans. A previous study revealed considerable molecular diversity, suggesting cryptic species, but relationships among lineages did not correspond to geographic location.

**Results:** Delimitation analyses on cytochrome c oxidase subunit I identified 11 different evolutionary lineages, presenting two contrasting phylogeographic patterns: (1) lineages with well-structured populations in Morocco and Iberia, and (2) lineages lacking geographic structure across the Mediterranean Basin. Dating analyses placed main diversification events in the Pleistocene, and multiple Pleistocene refugia, identified using ecological niche modeling (ENM), are compatible with allopatric differentiation of lineages. Human-mediated transportation appears to have complicated the current biogeography of this medically important and synanthropic spider.

**Conclusions:** We integrated ecological niche models with phylogeographic analyses to elucidate the evolutionary history of *L. rufescens* in the Mediterranean Basin, with emphasis on the origins of mtDNA diversity. We found support for the hypothesis that northern Africa was the center of origin for *L. rufescens*, and that current genetic diversity originated in allopatry, likely promoted by successive glaciations during the Pleistocene. We corroborated the scenario of multiple refugia within the Mediterranean, principally in northern Africa, combining results from eight atmosphere–ocean general circulation models (AOGCMs) with two different refugium-delimitation methodologies. ENM results were useful for providing general views of putative refugia, with fine-scale details depending on the level of stringency applied for agreement among models.

**Keywords:** Arachnid, Evolution, Human-mediated dispersal, Pleistocene glaciations, Refugia, Spider

## Background

The Mediterranean Basin was placed among 25 world biodiversity 'hotspots for conservation priority' based on high levels of endemism and rapid loss of natural areas [1]. Humans began transforming Mediterranean ecosystems >10,000 years ago [2], such that today, only 4.7% of

primary vegetation remains unaltered in the region [3]. Despite these long-standing impacts, the Mediterranean continues to be home to a diverse flora and fauna.

Several factors (e.g. climate, geology) promoted development of this diversity at different temporal and geographic scales, such as the Messinian Salinity Crisis and the onset of a Mediterranean-type climate ~3.2 Ma [4]. Glaciations during the Pleistocene (~2.6 – 0.02 Ma) also played a role in shaping current diversity patterns [5]:

\* Correspondence: [cribera@ub.edu](mailto:cribera@ub.edu)

<sup>1</sup>Institut de Recerca de la Biodiversitat (IRBio), Departament de Biologia Animal, Universitat de Barcelona, Barcelona, Spain

Full list of author information is available at the end of the article

climatic fluctuations during this period caused regional extinctions [6,7] and promoted range shifts and diversification via allopatric speciation (e.g. [8]). The three major southern Mediterranean peninsulas (Iberia, Italy, Balkans) were long thought to have served as major refugia for European flora and fauna during Pleistocene glaciations [8], but recent studies have challenged this paradigm as too simplistic to explain observed patterns [9,10]. As a consequence, some authors have argued for refugia within refugia [11] or multiple northern refugia ([12,13]; but see [14]).

Different approaches have been used to delimit glacial refugia in the Mediterranean. Traditionally, paleoecological evidence [15-17] and concentrations of endemic taxa [18] served as primary evidence, but more recently, comparative phylogeographic studies have been used to delineate “phylogeographic hotspots”, or areas with unique genetic diversity [5], while others have used ecological niche models (ENMs) in conjunction with paleoclimate simulations (e.g. [19]). The latter methodology projects environmental requirements of species onto past conditions, thus offering an approach that is independent and complementary [20-22]. Each approach has drawbacks and merits, but identifying regions using multiple approaches offers increased confidence [21,23,24]. Although other studies have successfully integrated phylogeographic and ENM approaches to uncover putative refugial areas [25-27], few have treated the entire Mediterranean Basin [19,28,29], and none have considered the added complexity of a human commensal.

Two *Loxosceles* spider species coexist in the Mediterranean: *L. rufescens* (Dufour 1820) and the Tunisian *L. mrazigi* Ribera and Planas 2009. *Loxosceles rufescens* originated in the Mediterranean [30-34] but has been transported worldwide by humans [30-32,34,35]. Duncan et al. [32] documented diverse genetic lineages among individuals morphologically consistent with *L. rufescens*, suggesting cryptic speciation. The morphological simplicity within

*Loxosceles* makes traditional species delimitation “singularly difficult” ([36] p. 142), so genetically-based methodologies are key to illuminating the evolutionary history of this group. In addition, relationships between *L. rufescens* lineages are not predictable by geographic location [32], which contrasts with the high spatial structure of populations in other *Loxosceles* species [31,34]. Therefore, *Loxosceles rufescens* is an ideal model to unravel the role of climatic changes and human impacts on the evolutionary history of Mediterranean species.

In this contribution, we examine mtDNA diversity within *L. rufescens*, and elucidate evolutionary processes that promoted this diversity via a combination of phylogeographic and ENM approaches. Our working hypothesis is that current mtDNA diversity was generated allopatrically in glacial refugia across the Mediterranean, and that *L. rufescens* biogeography was since obfuscated by human activity. This hypothesis offers three opportunities for testing: (1) divergence times should coincide with periods of repeated glaciations (~2.6 – 0.02 My), (2) multiple putative refugia should have existed to provide areas of origin for distinct evolutionary lineages, and (3) widespread lineages should show no spatial structure within the Mediterranean Basin.

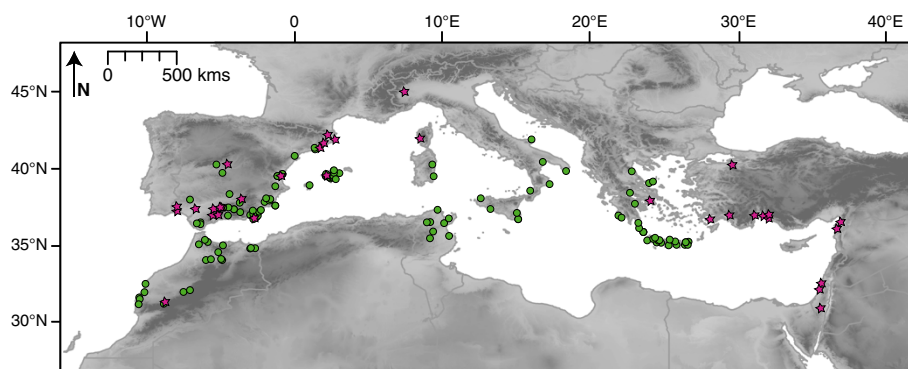
## Methods

### Taxonomic sampling

We sampled *L. rufescens* populations across the Mediterranean Basin to replicate and complement previous sampling [32] and to increase the likelihood of discovering new lineages. In all, 158 localities were sampled across eight countries (Figure 1 and Additional file 1). From these localities, 310 individuals were sequenced and included in our analyses.

### Molecular data

We included at least one individual from each locality in molecular analyses. Total genomic DNA was extracted



**Figure 1** Map of sampling localities. Green circles represent localities used in ENM analyses; pink stars indicate additional sampled localities used in the phylogenetic analyses.

using SPEEDTOOLS Tissue DNA Extraction Kit (BIO-TOOLS) following manufacturer's protocols. We amplified a portion of the cytochrome *c* oxidase subunit I (*cox1*) using LINP and GAYAR primers [34], producing 1016-bp fragments; we used different combinations of C1-J-1718, C1-J-2183 [37] and C1-N-2191 [38] internal primers when the first set failed. PCR reactions were conducted at a final volume of 25  $\mu$ L using either *Taq* polymerase (Promega) or Biotools *Pfu* DNA Polymerase (Biotools). PCR products were cycle-sequenced in both directions using the same PCR primers and the BigDye Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems). Sequences were derived from these products in an ABI 3700 automated sequencer at the Serveis Científico-Tècnics of the Universitat de Barcelona and in Macrogen, Inc. (Seoul, Korea). Raw sequences were edited and assembled with GENEIOUS v.4.6.5 [39]. To avoid amplification of pseudogenes reported by Duncan et al. [32] for *Loxosceles cox1*, we mostly used *Loxosceles*-specific primers; we also translated sequences into amino-acids, and checked for stop-codons.

Sequences were aligned unambiguously in GENEIOUS using ClustalW [40] with default parameters. We partitioned data by codon position and explored best partitioning schemes and substitution models simultaneously using PartitionFinder v.1.0.1 [41] under a Bayesian information criterion for the entire matrix. These steps were conducted independently for individuals employed in delimitation analyses, and for the reduced set used in dating analyses.

### Phylogenetic analyses

We used maximum likelihood (ML) and Bayesian inference (BI) to infer phylogenetic relationships from a dataset containing one representative of each *cox1* haplotype. ML analyses were conducted in RAxML v.7.4.2 [42] with the aid of the graphical front-end RAxML-GUI v.1.3 [43]. We applied a rapid hill-climbing search algorithm, and conducted 1000 non-parametric bootstrap replicates. BI analyses were conducted in MrBayes v.3.2 [44] with two independent runs of two million generations with four Markov chains (one cold, three heated), sampling every 1000 generations. We checked convergence of chains visually in Tracer v.1.6 [45] until effective sample sizes (EES) were above 200, and the average standard deviations of split frequencies (ASDSF) of the two runs were below 0.01. The first 25% of trees in each run were discarded as burn-in, and a majority-rule consensus tree was generated from remaining trees. BI trees were also obtained with BEAST v.1.7.4 [46] using a coalescent tree prior with a constant population size and a relaxed log-normal clock (rate fixed arbitrarily at one). Two independent runs of 20 million generations (sampling every 1000th generation) were used for each analysis. We

assessed convergence and correct mixing of chains by inspecting the trace plots and ensuring EES > 200 in Tracer. The two runs were combined using LogCombiner and TreeAnnotator [47] after removing a 10% burn-in of the samples. Position of the root of the tree was estimated implicitly in BEAST [48] and used for rooting RAxML and MrBayes trees.

Genetic p-distances between and within lineages (see *Molecular delimitation analyses*) were calculated using MEGA5 [49]. To study demographic history (only for lineages with  $N > 10$ ), we applied two neutrality tests: Fu's  $F_S$  [50] and  $R_2$  [51] in DnaSP v.5.10 [52]. We assessed statistical significance and confidence intervals using coalescence simulations in DnaSP with 1000 replicates and default parameters. We excluded six sequences from the A6 lineage that lacked the data for the 3' or 5' ends.

### Molecular delimitation analyses

Morphological traits provide few means of distinguishing lineages of *L. rufescens*, but genetic distances between lineages are high and different lineages are often sympatric at micro-scales ([32]; pers. obs.); consequently, assignment of individuals to lineages is not a simple function of geographic location. To account for this, we used two methodologies for objective delimitation of evolutionary lineages: (1) a General Mixed Yule Coalescent model (GMYC), and (2) phylogenetic network estimation using statistical parsimony (TCS). Although these methodologies have been successful at circumscribing species and often yield results congruent with alternative species delimitation methods (e.g. based on morphology, behavior; [53,54]), their utility depends on different factors (e.g. effective population size, sampling scheme; [53,55,56]), with results tending perhaps towards overestimation of species numbers [57].

### GMYC

GMYC is a species delimitation method that provides an objective way to delimit genetic clusters [56]. The method was developed to identify putative species in poorly-known groups based on single molecular markers [58]. The model seeks transition points (thresholds) between inter-specific relationships and intra-specific coalescent events, and subsequently tests the likelihood of the model against a null model that assumes a single branching process for the entire tree [58]. Since GMYC requires identical sequences to be removed [56], we included one representative of each haplotype. Because GMYC is sensitive to relative branch lengths and topology of the ultrametric tree [54], we explored effects of alternative input trees obtained from ML using RAxML, and BI using MrBayes and BEAST, as described below.

We generated a ML tree using RAxML [42], as outlined above. We converted this result to an ultrametric tree using PATHd8 [59], arbitrarily fixing the root at 100 units. Bayesian trees were derived using MrBayes [44], with two independent runs of 50 million generations with four Markov chains each (one cold, three heated), sampling every 1000 generations with a coalescent clock prior. The clock rate was arbitrarily fixed to one, and we used a lognormal distribution as a prior for population size. We checked for chain convergence as described above. The ultrametric tree was imported into R [60] and made fully dichotomous with the multi2di function in the APE package v.30.8 [61]. The ultrametric Bayesian tree from BEAST was obtained as outlined above.

### TCS

Separate haplotype networks in statistical parsimony analyses might provide a useful and objective method to delimit individuals into evolutionary significant units [62]. Networks delimited using a 95% parsimony connection limit, i.e. the probability that two DNA sequences share a parsimonious relationship without multiple substitutions underlying any single nucleotide difference [63], generally correspond to single species (78% in 663 examples in [62]). We obtained statistical parsimony networks in TCS v.1.3 [64] using our complete dataset (310 individuals), applying 95% and 99% connection limits, and treating gaps as missing data.

### Patterns of genetic diversity

We investigated how mtDNA diversity is distributed across geography within *L. rufescens* following [65]; this methodology is of particular use when localities have unequal sample sizes. Diversity statistics are computed by considering samples located within a perimeter around a grid point. We set grid points every 100 km in latitude and longitude, and computations were assessed across random sets of five individuals, bootstrapping 1000 times. We calculated three diversity indices, total diversity ( $H_T$ ), haplotype richness ( $H_R$ ), and rarity index (R) [19], in R [60] using custom scripts provided by N. Arrigo. High genetic diversity combined with rare haplotypes are characteristics of populations with a long *in situ* history, so identification of this pattern may indicate regions of origin for the different lineages [66]. We explored effects of different parameters on diversity indices, varying grid point distances 50–450 km, and numbers of individuals per analysis 3–10.

We analyzed the geographic structure of each lineage using a Mantel test implemented in the Isolation By Distance Web Service v.3.23 [67]. Geographic distances among localities were calculated using the Geographic Distance Matrix Generator v.1.2.3 [68], and genetic p-distances

between localities were calculated in MEGA. We performed Mantel tests with 999 permutations to assess significance of correlations between genetic distances and log-transformed geographic distances. We excluded lineages B1 and B3 from Mantel tests owing to low numbers of localities (< five). Lineages A1 to A5 were pooled for the analyses (see results), and remaining lineages were assessed independently.

### Dating analysis

Since we lacked reliable calibration points within the *L. rufescens* lineage, we explored two divergent rates. First, we used a *Loxosceles*-specific molecular rate obtained using fossil and island ages as calibration points [34]. Second, we applied the substitution rate obtained for the same mtDNA gene in a closely-related spider family (Dysderidae; [69]). As the two rates are fairly divergent, we suspect that actual divergence times for *L. rufescens* fall somewhere between these end points. Rates were incorporated as priors under a normal distribution with mean  $0.095 \pm 0.001$  and  $0.0199 \pm 0.001$ , respectively. Dating analyses were conducted in BEAST [46], using an uncorrelated lognormal relaxed clock [48] and a Yule tree prior. One representative of each lineage was included in analyses, and we used two independent runs of 10 million generations, sampling every 1000th generation, for each analysis. We assessed convergence and correct mixing of chains by inspecting trace plots and ensuring EES > 200 in Tracer. Runs were combined using LogCombiner and TreeAnnotator, after removing a 10% burn-in.

### Ecological Niche Modeling (ENM)

#### Study area

Niche models for *L. rufescens* were calibrated within a region that we hypothesized was 'sampled' by the species over its relevant history; in other words, a region the species had been able to deem suitable/unsuitable (**M**; *sensu* Barve et al. [70]), intersected with regions that were sampled as part of this study [71]. To calculate **M**, we buffered *L. rufescens* records by the longest distance from the sea to a documented locality (~350 km), which provided an estimate of the dispersal capability of the species. We excluded areas that we were unable to sample, or where closely-related species occur (*L. mrazig* in the southern parts of Tunisia and southeastern Morocco; an undescribed species group in the Sous Valley of Morocco). These steps left a calibration area comprising Morocco, the Iberian Peninsula, Balearic Islands, Sardinia, Sicily, continental Italy, Greece and adjacent islands, Crete, and Tunisia. After model calibration in this area, we projected results to the entirety of the Mediterranean Basin, within a bounding rectangle of 48.2–26.7° latitude and -14.8–41.2° longitude (see Figure 1).

### **Climate data**

We obtained climate data from eight coupled atmosphere–ocean general circulation model (AOGCMs) simulations: Community Climate System Model (CCSM), Centre National de Recherches Météorologiques (CNRM), Consortium for Small-scale Modelling (COSMOS), Goddard Institute for Space Studies (GISS), Institute Pierre Simon Laplace (IPSL), Model for Interdisciplinary Research on Climate (MIROC), Max-Planck Institut für Meteorologie (MPI), and Meteorological Research Institute (MRI). These AOGCMs were derived from the multi-model ensemble in the Coupled Model Intercomparison Project Phase 5 (CMIP5) [72] and the Paleoclimate Modelling Intercomparison Project Phase 3 (PMIP3) [73]; more details about the climate models are provided in Additional file 2 and in Taylor et al. [74].

We downloaded monthly simulation outputs for annual precipitation and mean, maximum, and minimum temperatures from the pre-industrial experiment, which characterized current climatic conditions. Past conditions were characterized using paleoclimate simulations for the mid-Holocene (~6 Ka) and Last Glacial Maximum (LGM, ~21 Ka) from each AOGCM, with the exception of GISS and COSMOS, which lacked mid-Holocene outputs. To produce climate scenarios at resolutions relevant to the spatial scale of species' distributions, we downscaled climate layers to 0.5° resolution using a standard change-factor approach [75]: (1) for each AOGCM, we computed the difference between the past (mid-Holocene and LGM) and current simulations (i.e. pre-industrial), and this difference (i.e. climate change trends) and the current climate were interpolated to 0.5° spatial resolution using kriging; (2) these differences were added to the interpolated current climate to obtain interpolated past conditions. We used absolute differences for temperatures and relative differences for precipitation; see [76] for further details. This procedure maintains higher-resolution topography in downscaled climates and ensures coherency of climatic patterns over time [77]. From these downscaled climatic scenarios, we computed 19 so-called 'bioclimate' variables [71], but we excluded mean temperature of the wettest and driest quarters, and precipitation of the warmest and coldest quarters, owing to the spatial artifacts that emerge in these four variables.

For each AOGCM, we performed a principal components analysis in R [60] on the 15 bioclimatic variables in the calibration area to create new axes that summarized variation in fewer, independent dimensions, and to reduce co-linearity among variables. We retained those principal components that explained cumulatively 99% of the overall variance in the dataset (i.e. the first six principal components for all AOGCMs except GISS, which required only the first five) for model calibration. These principal components were used to calculate

corresponding composite variables for mid-Holocene and LGM conditions. The PCA structure for current conditions was enforced for the past conditions using a script in R [60] written by A. Lira and N. Barve (U. Kansas).

### **Occurrence data**

We used a subset (130 records) of occurrence data associated with samples employed in the genetic analyses. These localities were obtained directly from fieldwork in natural areas by EP and others (see Acknowledgments), and have precise latitude/longitude coordinates derived from GPS measurements. To consider the potential biasing effects introduced by spatial autocorrelation, such that spatially-clumped points would over-represent certain environments, we calculated spatial lags in environmental data using Geostatistical Analyst in ArcMap v.10 (ESRI, Redlands, CA), and subsampled the records to create 10 replicate datasets using a script in R [60] written by N. Barve (U. Kansas). Based on lag calculations, we enforced a minimum distance of 50 km between localities. Subsampling occurrence data to account for environmental lag ensures that suitable conditions are evenly weighted during model calibration. Each subset included 62–65 occurrence records.

### **Modeling algorithm**

ENMs were generated using Maxent v.3.1.1 [78], which can be monitored for extrapolation errors when projecting to past climates [79,80]. Maxent minimizes the relative entropy between two probability densities—one from the distributional data and one from the background or study area—defined in covariate space [81]. We used default parameters, but specified 100 bootstrap replicates per occurrence dataset and a minimum training presence threshold rule to avoid omission error. We took the median of the 100 runs per occurrence dataset multiplied by 1000, and converted to integer grids in ArcMap. These grids were used to calculate the median of the 10 subsets for each AOGCM. The resulting models were converted to binary grids based on all 130 localities using a minimum training presence approach [82]. Use of multiple AOGCMs [83] provides a broader estimate of suitable conditions for *L. rufescens*, but we acknowledge that ensemble-modeling approaches may shed additional light on model-dependent results (see [24]).

When transferring models temporally or spatially, conditions outside the range of climatic values in the calibration region (**M**) may be encountered, leading to situations of extrapolation. To identify these regions, we used a script in R [60] written by N. Barve [84] to create Mobility Oriented Parity (MOP) maps [80]. Areas identified as both suitable and extrapolative were removed from analyses to avoid interpreting results outside of known climatic response conditions for *L. rufescens*.

### Model evaluation

To evaluate the predictive power of the models, we partitioned two of the replicate occurrence datasets at random: half of the data was used in model calibration, and the other half for model testing via partial Receiver Operator Characteristic (partial ROC) approaches [82]. Partial ROC avoids many of the problems associated with traditional ROC analyses, such as equal weighting of omission and commission errors, and consideration of model thresholds that yield irrelevant predictions. These tests were run using a Visual Basic routine developed by N. Barve [85], with an expected error rate of  $E = 1\%$  [82]. We performed 1000 bootstrap iterations by resampling 50% of test points with replacement.

### Identifying refugial areas

We identified possible refugia as areas that remained continuously suitable from the LGM to the present. Because glaciations were common throughout the Pleistocene, with glacial and interglacial conditions recurring in nearly-regular cycles with similar amplitudes (at least since 800 Ka) [86], we assume that the three time slices used here (LGM, mid-Holocene and present) capture, at least to some degree, the key climatic conditions across the entire Pleistocene [87,88]. We acknowledge, however, that these reconstructions are merely broad estimates of potential refugial conditions for *L. rufescens* and lack constraining data for the earlier half of the epoch.

We identified refugia using two approaches: approach one (M1) required AOGCMs to agree on suitable area for each time slice, applying four different thresholds (8 of 8 to AOGCMs agree, 6/8, 4/8, and 1/8). This resulted in four different suitability maps for the LGM, the mid-Holocene, and the present-day. The intersection of the three time slices was taken as the final refugial area, which created four different possible refugial scenarios (Figure 2a-d). The second approach (M2) calculated refugial area across the three time slices for each individual AOGCM (Figure 2e-h); in other words, the intersection of suitable area was taken across the three time slices (LGM, mid-Holocene, and present-day) for each AOGCM independently. From the individual AOGCM maps of refugial areas, we applied the threshold criteria of M1 to identify consensus regions (8 of 8 AOGCMs agree, 6/8, 4/8 and 1/8). In effect, these two methods explored sensitivity to threshold choice, and resulted in eight putative refugial maps. We repeated the two methods without COSMOS and IPSL, as these AOGCMs often exhibit anomalous climatic patterns compared to other AOGCMs. Using ENM to identify putative refugia can elucidate potential divergence mechanisms. Considerable caution, however, should be exercised in interpreting such analyses, particularly in light of the spatial grain of these data: coarse-resolution climate data

cannot detect fine-scale phenomena (i.e. microrefugia *sensu* Rull [89]).

## Results

### Phylogenetic analyses

New sequences obtained during this study were deposited in GenBank with accession numbers KJ560560 - KJ560863 (Additional file 1); additional sequences were downloaded from GenBank (Additional file 1). In total, 310 sequences were used, containing 63 different haplotypes. PartitionFinder suggested a non-partitioned codon scheme with a HKY + G substitution model as the best fit for these data under the Bayesian information criterion; we used this partition scheme and substitution model in all phylogenetic analyses except RAxML, where only a GTR + G model was available.

Phylogenetic results were nearly identical between BI and ML approaches (available on TreeBase S15925). In both cases, *L. rufescens* was split in two main clades: A and B (Figure 3). Clade A included six well-supported lineages (all with bootstrap support >88% and posterior probabilities >0.76). Four lineages were composed exclusively of individuals from Morocco (termed lineages A1-A4). A1 placed as sister to a well-supported clade comprising A2-A6; the clade composed of A2-A4 placed as sister to a clade composed of A5-A6. A5 included individuals from two Iberian Peninsula populations, and A6 included individuals from across the Mediterranean.

Clade B comprised five lineages (B1-B5), all well-supported (bootstrap support >97%, posterior probabilities 1.0). B5 placed as sister to the remaining lineages of clade B, and lineage B1 was sister to a well-supported clade (bootstrap support 99%, posterior probabilities 1.0) comprising B2-B4, but this latter relationship was not well supported (bootstrap support 54%, posterior probabilities 0.64). B2 placed as sister to B3 and B4 (bootstrap support 79%, posterior probabilities 0.98). B3 was composed exclusively of Iberian Peninsula individuals; the remaining lineages included individuals from different Mediterranean regions.

Genetic p-distances ranged from 1.5-7.8% between the various lineages (Additional file 3). The two major clades (A and B) were separated by a p-distance of 7.04% (Additional file 3). Neutrality tests for the lineages are presented in Table 1. Fu's  $F_s$  test for demographic expansion was negative and significant in all cases except lineage B3. The  $R_2$  test was low and significant for the two lineages with higher sample sizes (A6 and B5) in the left tail. As a whole, these results suggest a recent demographic expansion for all lineages except for B3.

### Molecular delimitation analyses

We used three methods to obtain the ultrametric trees required for GMYC analyses (Figure 3). In all three

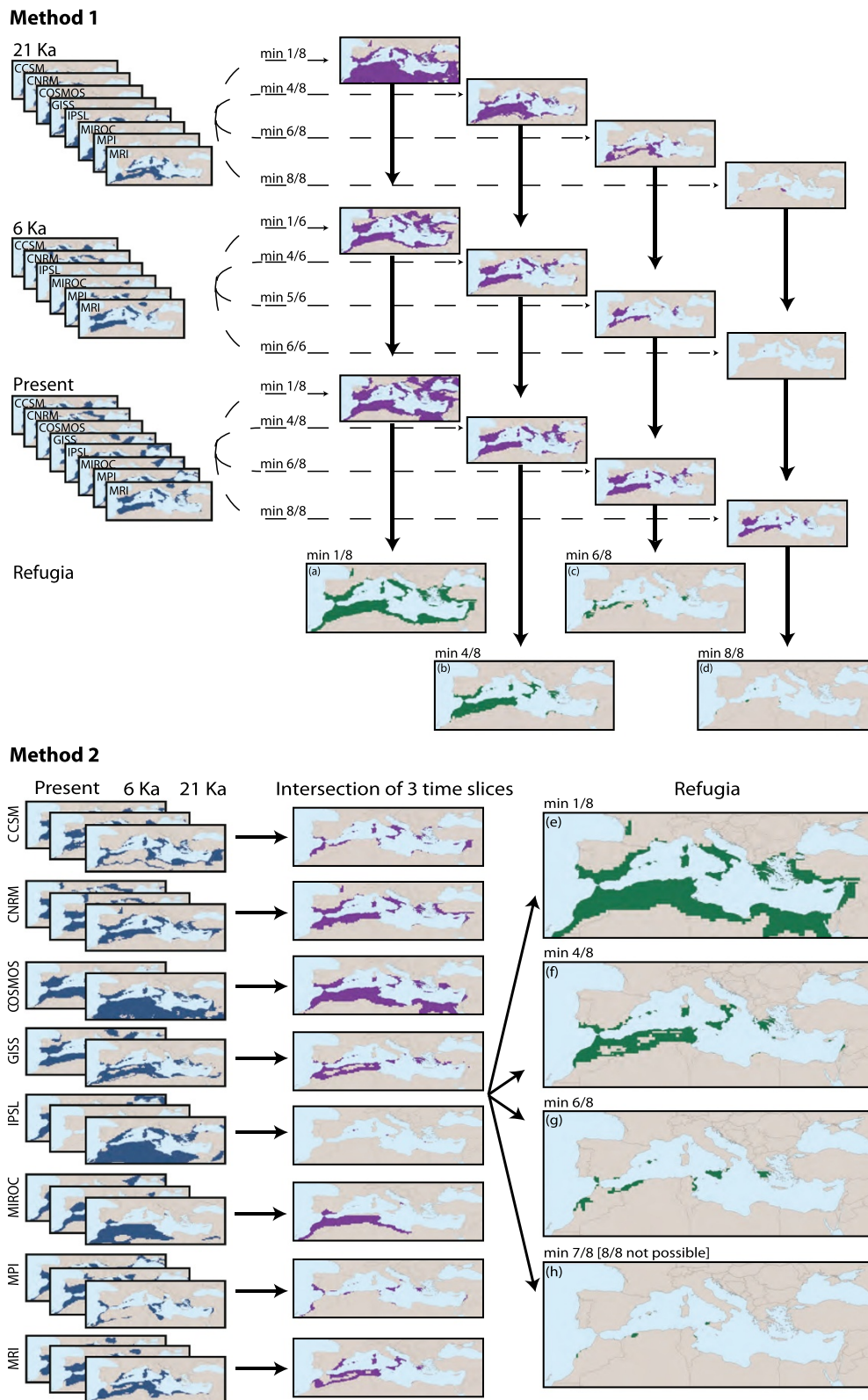


Figure 2 (See legend on next page.)

(See figure on previous page.)

**Figure 2 Refugium delimitation methods.** Refugial areas were identified as the intersection of the three time slices. Method 1 sought consensus among the eight AOGCMs in each time slice by requiring at least (a) one to agree, (b) 4/8 to agree, (c) 6/8 to agree and (d) all 8 to agree. Method 2 sought the intersection of the three time slices for each AOGCM independently, subsequently requiring the AOGCMs to agree in the same fashion as M1: at least (e) one to agree, (f) 4/8 to agree, (g) 6/8 to agree and (h) 7 to agree (8/8 was not possible).

cases, the likelihood ratio test of the Yule model was significantly better than the null hypothesis (BEAST  $p = 4.22 \times 10^{-6}$ , RAxML and PATHd8  $p = 7.42 \times 10^{-7}$ , and MrBayes  $p = 1.28 \times 10^{-6}$ , respectively). Clusters identified (i.e. GMYC groups composed of more than single individuals) were mostly congruent across methods; in all three, seven clusters were delimited. Slight differences between approaches appeared in terms of detecting singletons: in total, we recovered 14 entities (clusters plus singletons) using the ultrametric tree obtained with BEAST (confidence interval: 11–21), while the remaining two analyses delimited 11 entities (confidence interval: 11–14 with RAxML, 10–11 with MrBayes). Differences occurred in lineages B5, A4, and A6, wherein GMYC analyses using the BEAST tree split each of these lineages into two clusters (Figure 3).

TCS results varied depending on the connection limit (Figure 3): a 95% connection limit resulted in fewer independent networks compared to a 99% limit. In the former case, the maximum number of calculated steps was 13, forming seven independent networks, and in the latter case, the maximum number of calculated steps was five, with 11 independent networks. The higher connection limit mirrored the GMYC results. We found two main patterns in the haplotype networks: (1) several lineages composed of individuals restricted to one or a few localities that harbor only one or a few haplotypes, and conversely, (2) single haplotypes present in individuals from across the Mediterranean Basin, with closely related haplotypes forming a star-like network (Figure 4).

#### Patterns of genetic diversity

Analyses were conducted across the entire dataset and within lineages with wide distributions (i.e. A6, B2, B4, and B5; results presented in Additional file 4). Analyses across all lineages lacked clear geographic patterns for total diversity ( $H_T$ ) and haplotype richness ( $H_R$ ); however, they exhibited higher values for the rarity index ( $R$ ) in southern Morocco. Analyses of individual lineages showed clear geographic patterns for lineage B2 and B4, wherein higher values for all diversity statistics were obtained in the Iberian Peninsula and Balearic Islands, respectively. Lineages A6 and B5 showed no clear geographic pattern, with the highest values in various isolated regions. Different parametrizations for this test produced similar patterns (results not shown). That is, similar results were

obtained when the number of individuals was reduced to three (not recommended by the script authors, N. Arrigo pers. comm.), except within lineage B2, where high values for all statistics characterized the region around Greece. The Mantel test showed weak correlations between genetic and geographic distance matrices for lineages A6, B2, B4, and B5 ( $r < 0.24$ ), with tests not significant except for lineage B5 (Table 2). Conversely, when lineages A1–A5 were pooled, we obtained a stronger and significant correlation ( $r = 0.52$ ,  $P < 0.05$ ; Table 2), suggesting that they are geographically structured.

#### Dating analyses

Divergence time estimates differed depending on the rate used in calibration (see Additional file 5). Both analyses, however, dated major diversification events to the period of Pleistocene glaciations. The estimated split between the two main clades (A versus B, Figure 3) was dated at 0.356 Ma (95% HPD: 0.243–0.487) using the *Loxosceles*-specific rate, and at 1.968 Ma (1.322–2.698) with the *Parachtes* rate.

#### Niche models

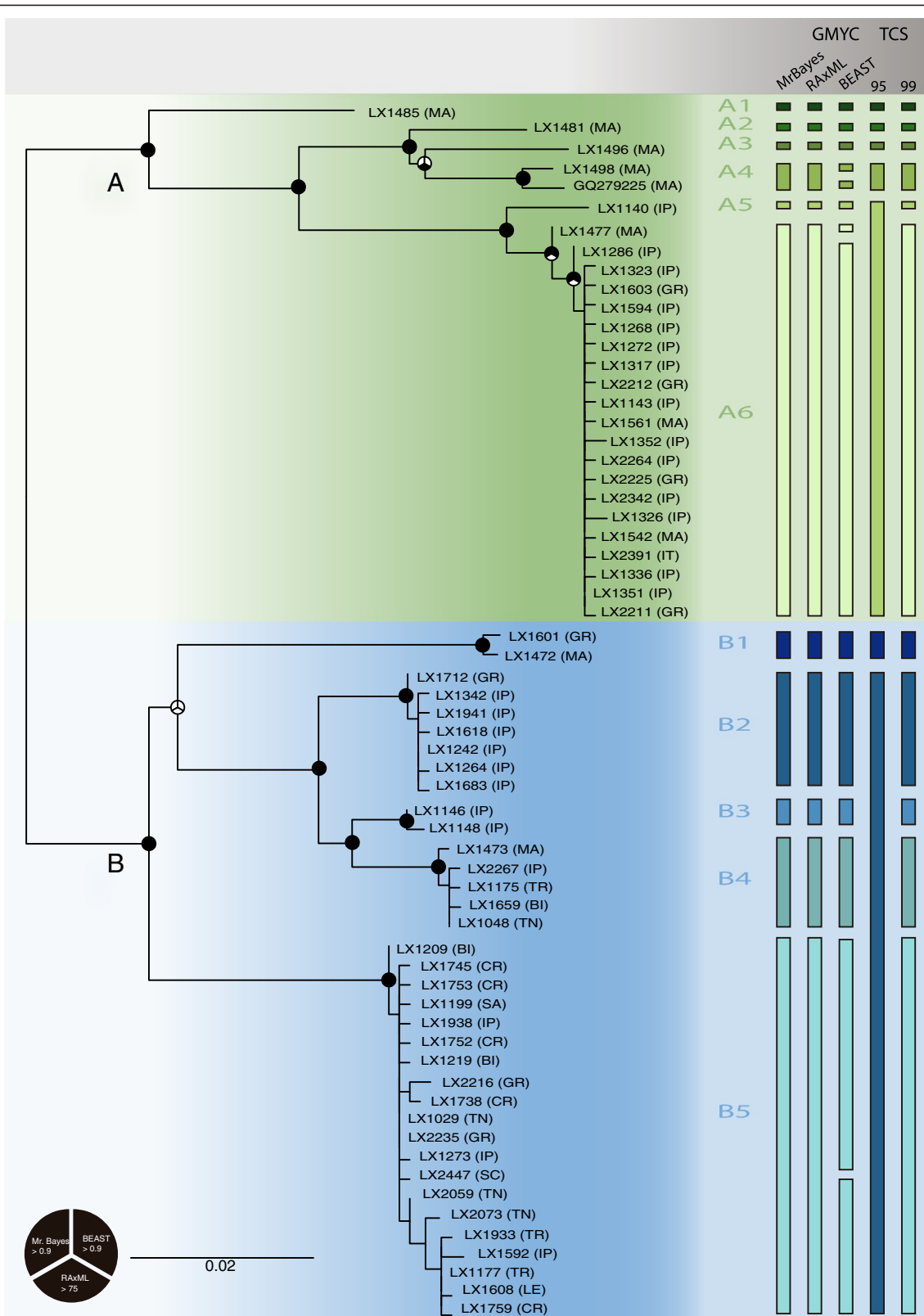
##### Present

Models from individual AOGCMs were predictive of independent suites of occurrence points, with all models statistically significant in partial ROC tests (all  $P < 0.05$ ; Additional file 6). Predictions were consistent across AOGCMs (Additional file 7), with suitable areas identified in the southern Iberian Peninsula, Italy, Greece, western Turkey, northern Africa, and various Mediterranean islands. Minor differences were noted among models; for example, less suitable area was identified in Turkey under MIROC and MRI. Regions environmentally outside those represented within M (highlighted by MOP analyses) were also consistent across AOGCMs, covering desert areas (e.g. the Sahara and Sinai Peninsula) and some regions around the Black Sea; these regions were not considered in our analyses.

##### Paleoprojections

Most regions identified as suitable in the present were also identified as suitable during the mid-Holocene (Additional file 7). MOP results were similar to those in the present, with one exception: although IPSL identified potential distributions congruent with those in other AOGCMs, MOP analyses indicated most of these regions





**Figure 3** (See legend on next page.)

(See figure on previous page.)

**Figure 3 Maximum likelihood tree based on single representatives of each *cox1* haplotype.** Node circles represent maximum likelihood bootstrap and Bayesian posterior probabilities, as shown in the legend. Green indicates the **A** clade, and blue the **B** clade. Each column on the right indicates a different delimitation method, and delimited lineages are represented with colored bars. Abbreviations: MA (Morocco), IP (Iberian Peninsula), GR (Greece), IT (Italian Peninsula), TR (Anatolian Peninsula), BI (Balearic Islands), TN (Tunisia), CR (Crete), SC (Sicily), LE (Levant).

were environmentally novel. A similar situation occurred with the LGM projection for MIROC. Refugial area delimitation was not affected by removal of these two models with odd results.

Compared to present and mid-Holocene projections, LGM potential distributions shifted southward. The main Mediterranean peninsulas (i.e. Iberia, Italy, Balkans) retained suitable conditions, although reduced in extent, but extensive regions of the Sahara, which were largely unsuitable in the present and mid-Holocene, were identified as suitable during the LGM under most AOGCMs. Regions environmentally outside those represented within **M** occurred across broad swaths of the northern and southeastern portions of our study area.

#### Identifying refugial areas

Results from individual AOGCMs were similar, with the exception of IPSL. Because IPSL uses a different vegetation model from other AOGCMs (for example, bare soil is considered a type of vegetation, whereas other AOGCMs ignore this factor), we ran refugial delimitation analyses with and without IPSL. Putative refugia were congruent across the two methodologies and with and without IPSL (Figure 5; without IPSL not shown). Depending on the level of stringency enforced for AOGCM agreement, 4–14 major, independent, and isolated refugia were identified (Figure 5a–e). When less stringent agreement thresholds were applied (Figure 5a and e), the entire Mediterranean rim was identified as refugial, except for the northern and eastern coast of the Adriatic Sea and Gulf of Genoa. With intermediate levels of stringency (Figure 5b,c,f,g), most refugia were situated in the western Mediterranean, primarily Morocco, Algeria, and parts of the Iberian Peninsula (Cabo de Gata, Cadiz/Algeciras region, Valencia region; sensu Médail and Diadema [5]). The Balearic Islands, especially Mallorca, were identified as refugial, even under the most stringent criteria (Figure 5d and h),

and some parts of Sicily were recovered under most scenarios (except the strictest criterion in M1). Unlike the western Mediterranean, few parts of the eastern half of the Mediterranean were identified as refugial. For example, only some areas of the Peloponnese (Greece) were recovered consistently as putative refugia; broad areas of Anatolia and the Levant coast were recovered as refugial only under the least stringent AOGCM agreement levels.

## Discussion

### Genetic diversity and biogeography

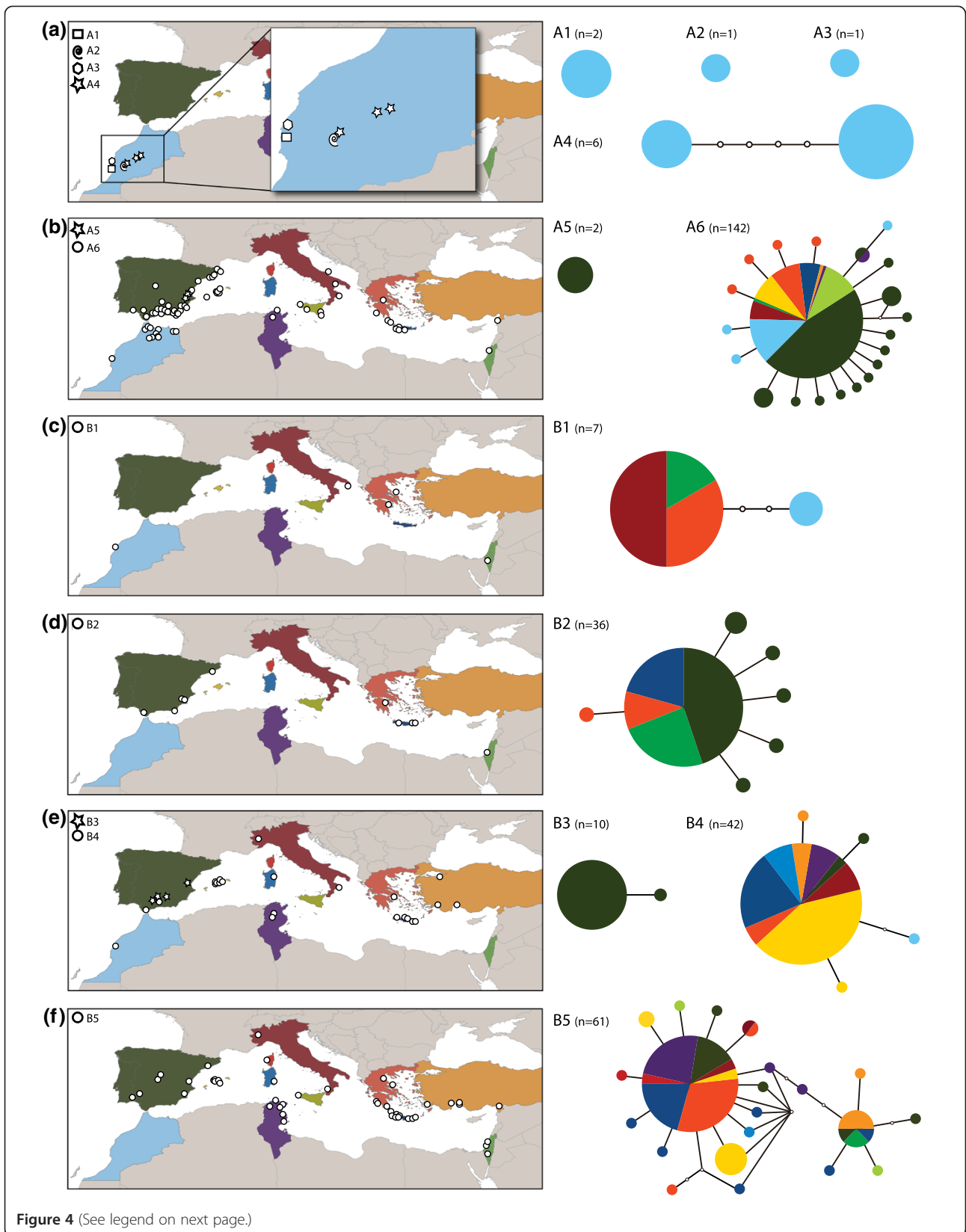
We document high mtDNA genetic diversity in *L. rufescens* across the Mediterranean Basin, as reported previously by Duncan et al. [32], underscoring the importance of broad sampling efforts for accurate representation of diversity patterns. Most of the delimitation methods recover 11 distinct evolutionary lineages (Figure 3), but molecular delimitation analyses, particularly those using only one line of evidence (in this case, mtDNA), can be prone to over-delimitation [90]. Thus, identified lineages should be taken as a basis for further studies of taxonomic status using integrative approaches based on morphology and variable nuclear markers [57], and the phylogeographic patterns uncovered herein merit reexamination with additional nuclear data. Nevertheless, even without further analyses, the existence of such divergent mitochondrial lineages deserves attention, and our main aim was to understand the factors promoting this diversity.

Genetic diversity is not distributed uniformly among lineages within *L. rufescens*, and is, in fact, highly heterogeneous (Figures 2 and 4 and Additional file 2). Broadly, we find two contrasting phylogeographic patterns: the mountainous region of Morocco harbors several lineages with well-structured populations, whereas lineages distributed across the broader Mediterranean Basin generally lack geographic structure.

**Table 1 Summary of the nucleotide diversity estimates and neutrality tests**

Lineage	n	Sites	H	H <sub>D</sub>	π	F <sub>S</sub>	95% CI	R <sub>2</sub>	95% CI
A6	136	808	20	0.298	0.00045	−34.268***	−4.435–3.293	0.019*	0.015–0.238
B2	36	480	5	0.260	0.00057	−4.111***	−1.893–2.722	0.07 NS	0.054–0.25
B3	10	565	2	0.2	0.00035	−0.339 NS	−0.594–1.523	0.3 NS	0.178–0.3
B4	42	555	5	0.184	0.00043	−4.408***	−2.089–2.631	0.078 NS	0.046–0.254
B5	61	407	12	0.572	0.00187	−9.883***	−3.394–4.148	0.042*	0.046–0.23

Abbreviations: \**p* < 0.05; \*\*\**p* < 0.01; NS non significant, *n* number of sequences, *H* number of haplotypes, *H<sub>D</sub>* Haplotype diversity, *π* nucleotide diversity, *F<sub>S</sub>* Fu's *F<sub>S</sub>*, *R<sub>2</sub>* Ramos-Onsins & Rozas *R<sub>2</sub>* test, *CI* Confidence interval.



(See figure on previous page.)

**Figure 4 Distribution map and haplotype network for each lineage.** Colors on haplotype networks correspond to colored areas on the maps to the left (a-f). Haplotype networks are not to the same scale among lineages. Each circle represents one haplotype, and colors correspond to frequencies of region of origin for the haplotype ( $n$  = number of individuals). Note the two contrasting phylogeographic patterns, with lineages A1 to A5 restricted to one or a few well-structured populations, whereas the lineages distributed across the Mediterranean Basin generally lack geographic structure, which is likely a consequence of human-mediated dispersal.

Four lineages (A1-A4) are distributed along the western slopes of the Atlas Mountains (Figure 4), including individuals from A4 referred to as the “Asni clade” by Duncan et al. [32]. Although lineages A1-A4 live in close proximity (within 50–200 km), they exhibit striking genetic divergence (>4.7%, Additional file 3). The Atlas Mountain region shows the highest rarity index values in analyses considering all lineages (Additional file 4). This pattern of deep genetic divergence and haplotype differentiation among populations in close proximity may be explained by long-term presence of this species in the region and low dispersal capacity under natural conditions. Together with A5, these lineages are spatially structured in that genetic diversity is positively correlated with distance between localities. Altogether, these patterns are consistent with spider species with similar dispersal capacities (e.g. mygalomorphs, see [57]), and with other *Loxosceles* species, such as those from the Canary Island [34], *Loxosceles mrazig*, and a related group from the Sous Valley (Morocco; Planas and Ribera unpublished data).

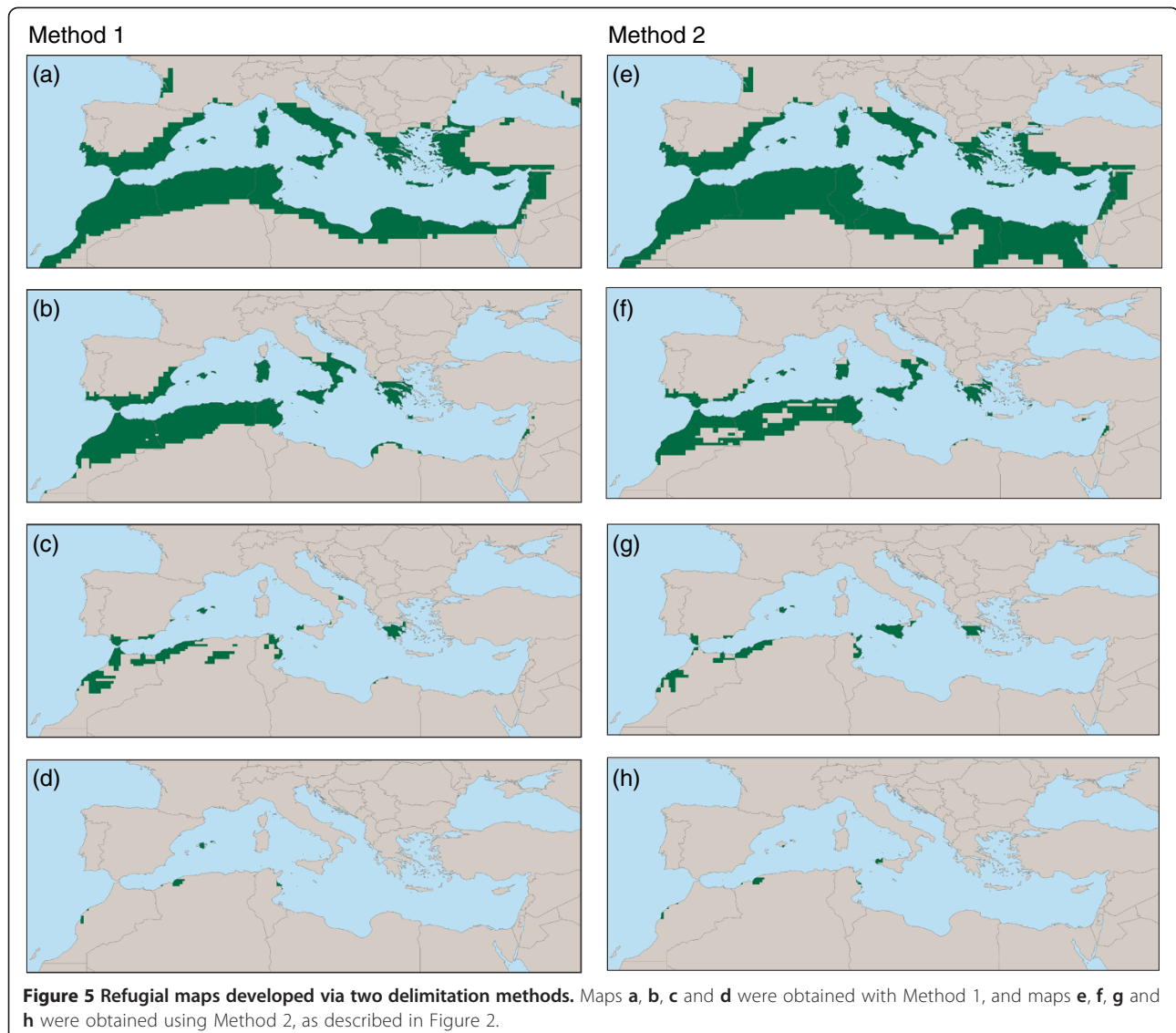
Most individuals, however, belong to lineages widespread across the Mediterranean (e.g. A6 and most B lineages, Figure 4). Haplotype networks for these lineages have star-like shapes, with one common haplotype shared among individuals distributed across the Mediterranean Basin, and fewer, less-common haplotypes, restricted to individuals from one or a few localities. For almost all lineages, no clear correspondence exists between haplotypes and geography, with weak correlations between genetic and geographic distances. For example, the most common haplotype from A6 is found across the entirety of the Mediterranean Basin (Figure 4). This lineage, named the “Iberian clade” by Duncan et al. [32], appears to be the most common in the western Mediterranean, and contains individuals from Sagunt, the type locality of *L. rufescens*.

Biogeographic patterns within clade B are more complex. The exception is lineage B3, where all individuals are found in the Iberian Peninsula. In lineages B2, B4 and B5, individuals with the most common haplotype are widespread across the Mediterranean, as in A6, although no clear pattern emerges that links genetic diversity with geography (Additional file 4 and Table 2). Lineage B1 represents the most extreme example of the complex distributional patterns found within *L. rufescens*. Even given the extensive sampling we conducted, we found individuals of this lineage at only five localities, some separated by >4000 km (although more samples from this lineage might produce the typical star-like shape of the B clade lineages).

Discerning between natural and human-mediated dispersals can be difficult in the Mediterranean ([91], and references herein). Current genetic patterns for Mediterranean lineages do not coincide with those expected as a result of secondary contact through natural processes. If naturally occurring, contact would be restricted to particular areas and/or occur between or among only a few lineages. Here, multiple lineages are distributed across the Mediterranean, including on islands, a pattern that is difficult to explain by natural processes in organisms with naturally poor dispersal abilities. The lack of geographic structure within most of the Mediterranean *Loxosceles* lineages contrasts with the highly structured patterns found for lineages A1-A4, distributed in the mountainous region of Morocco, with the former pattern a likely consequence of human-mediated dispersal. Although *L. rufescens* originated in the Mediterranean Basin ([30,32], see below), the species has been introduced to many parts of the world, including Australia, Madagascar and North America [30,32,34,92]. Human transportation seems a likely mechanism to explain how some haplotypes are distributed across the Mediterranean Basin, including on several islands and on both African and European shores. *Loxosceles rufescens* possesses two life traits that facilitate dispersal with human assistance: high starvation tolerance [93] and urban microhabitat preferences. Maritime commerce in this region has been active for >5000 years [2], and transportation of cultivated plants [94,95], domesticated animals [96], and wild animals such as reptiles [97], snails [98,99], mosquitoes [100], and freshwater triclads [101] has been documented widely throughout the Mediterranean. Thus, the expected “natural” biogeographic patterns have been blurred for this region, and the complex

**Table 2 Results of Mantel tests performed with 10000 permutations to assess the significance of the correlation between genetic distances and log-transformed geographic distances**

Lineage	r	P
A1-A5	0.524	0.012
A6	0.058	0.176
B2	0.002	0.483
B4	0.159	0.059
B5	0.242	<0.0001



phylogeographic pattern documented for *L. rufescens* represents a clear example of human influence on species' distributional dynamics.

#### Refugia and origins of genetic diversity

Although human-mediated transportation of *L. rufescens* likely impacted patterns of distribution for this species within the Mediterranean Basin, the aim of this study was to assess when and how the distinct lineages originated. Combining divergence time and refugial estimates, we marshal two distinct data streams toward answering these questions [23].

Our dating analyses place diversification events during the Pleistocene (Additional file 5). Although we are not able to link diversification events to individual climatic events (i.e. a particular glacial-interglacial cycle) with any confidence, these dates provide a coarse-resolution

estimate of diversification timing. The placement of key diversification events during the Pleistocene indicates that processes operating during this period (i.e. glacial/interglacial cycles) played an important role in shaping the current diversity of the species, as with numerous other species in the Mediterranean Basin (e.g. [98-106]). The results from our ENM analyses largely corroborate the scenario of multiple refugia within the Mediterranean for *L. rufescens*, although details depended on the level of stringency applied for agreement among the eight different AOGCM models (Figure 5). The shape and size of refugia differ markedly depending on the AOGCM used (Additional file 7). In other words, ENM is useful for providing general views of putative refugia, rather than for identifying actual borders. Combining results from different AOGCMs, we obtain a consensus view of general patterns for the latter half of the Pleistocene

epoch [107], which, when an intermediate threshold (Figure 5) is considered, agrees with refugia obtained for plants using phylogeographic approaches [5].

In both methods (Figure 5), major refugia are concentrated in the western Maghreb. Indeed, in phylogenetic terms, four evolutionary lineages (A1-A4) are found in this area, signifying a hot spot of lineage richness. This richness supports the hypothesis that northern Africa is the center of origin for *L. rufescens*, as previously hypothesized by Gertch [30] and Duncan et al. [32]. Additionally, the sister group to *L. rufescens* is found south of this area, in the High and Anti-Atlas Mountains and the Sous Valley (Morocco; Planas and Ribera unpublished data), which lends further support to a northern African origin for these lineages and for *L. rufescens* as a whole. This region has been postulated as a climatic refugium for various animals and plants in light of its complex orography (e.g. Anti-Atlas, High Atlas) and climatic stability ([90] and references therein).

More challenging is linking putative refugia to the origins of the Mediterranean lineages (A5, A6, B), with current distributional and genetic patterns most likely the result of population mixing through human-mediated transportation. Although some of these divergent lineages now occur in sympatry, they likely originated in allopatry, given the dominance of this speciation mechanism for the genus [31,34] and the geographic results summarized above (Figure 5). Lineages A5 and B3 have small distributions and are endemic to the Iberian Peninsula, which may reflect refugial areas, especially along the southern and eastern Iberian Mediterranean coast. However, five other lineages are widely distributed across the Mediterranean.

Genetic diversity patterns should help in elucidating the origins of these lineages, assuming that refugial areas harbor higher genetic diversity (but see [108]). Such is the case for lineage B4 (Additional file 4), where the highest genetic diversity is found in the Balearic Islands, an area identified consistently as a refugium (Figure 4 and Additional file 3). However, lineages A6, B1, B2, and B5 are widespread across the Mediterranean and do not show any correspondence between genetic diversity and a single putative refugial area; these lineages may have originated in one of the remaining predicted refugia (e.g. Sicily, southern Italian Peninsula, the Peloponnese), but subsequent processes (e.g. human-mediated transportation) appear to have erased ancient biogeographic signals. More extensive sampling in the central and eastern Mediterranean may help to resolve this question.

## Conclusions

In this study, we delimited 11 evolutionary lineages within *Loxosceles rufescens* in the Mediterranean Basin based on mtDNA data. Genetic diversity was not distributed uniformly, and we found two contrasting phylogeographic patterns: (1) the southern region of Morocco holds several

lineages with well-structured populations, (2) whereas lineages distributed across the broader Mediterranean Basin generally lack geographic structure. By combining results from eight AOGCMs with two different refugium-delimitation methodologies, we corroborated the scenario of multiple refugia within the Mediterranean, principally in northern Africa. ENMs were useful for providing general views of putative refugia, with fine-scale details depending on the level of stringency applied for agreement among models. Although refugial delimitation remains challenging, by combining ENM with phylogeographic approaches, we found support for the hypothesis that northern Africa was the center of origin for *L. rufescens*, that current genetic diversity probably originated in allopatry and was promoted by successive glaciations during the Pleistocene, and that protracted human activities impacted the current distributional patterns of *L. rufescens* within the Mediterranean Basin.

## Availability of supporting data

The data sets supporting the phylogenetic results of this article are available on the TreeBase repository, ID: 15925, <http://purl.org/phylo/treebase/phyloids/study/TB2:S15925>. Information on geographic location and Genbank accession number for the 310 individuals included in this study is available in Additional file 1.

## Additional files

**Additional file 1: Information on specimens and Genbank accession numbers.** Information on geographic location, inclusion in ENM analysis (Y – yes; N – no), mtDNA lineage and Genbank accession number for the 310 individuals included in this study. In bold, individuals included in the phylogenetic analyses. Abbreviations: CS – Corsica; CR – Crete; GR – Greece; IB – Balearic Islands; IT – Italy; LE – Levant; MA – Morocco; PI – Iberian Peninsula; SA – Sardinia; SC – Sicily; TN – Tunisia; TR – Turkey.

**Additional file 2: Details on climatic models (AOGCMs) used for Ecological Niche Modeling.**

**Additional file 3: Estimates of genetic p-distance between (above the diagonal) and within (in bold) lineages.** In the inset, genetic p-distance between the two clades.

**Additional file 4: Genetic diversity of *Loxosceles rufescens* in the Mediterranean.** Total diversity (HT), haplotype richness (HR) and rarity index (R) are represented for *L. rufescens* (all) and separately for lineages A6, B2, B4 and B5. Diversity statistics are computed by considering samples located within a perimeter around a grid point. We set grid points every 100 km in latitude and longitude, and computations were conducted across random sets of 5 individuals, bootstrapping 1000 times.

**Additional file 5: Dating analysis results.** Dating analysis results in Ma obtained from two different *cox1* rates: (1) using the rate for *Parachetes*, and (2) using a mean rate obtained for *Loxosceles*.

**Additional file 6: Partial ROC analyses.** Partial ROC tests ( $P < 0.05$ ) using two data subsets with 1000 replicates.

**Additional file 7: Ecological niche modelling results for each atmosphere-ocean general circulation model (AOGCM).** Ecological niche modelling results for each AOGCM in each time slice (Last Glacial Maximum, mid-Holocene, and present).

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

CR and EP conceived the study and participated in the fieldwork. EP carried out lab work and conducted the phylogenetic analyses. MLR prepared the AOGCMs models. ATP, EES and EP designed and conducted the ENM analyses. EES and EP wrote the first draft: all authors improved upon successive versions. All authors read and approved the final manuscript.

### Acknowledgments

We thank those people who diligently helped with fieldwork or provided *Loxosceles* specimens for study, particularly M.A. Arnedo, L. Bidegaray, R. Espluga, E. Gavish-Regev, M. Isaia, A. López-Pancorbo, K. Bogak, Y. Marusik, D. Martínez, M. Metallinou, J. Miñano, S. Montagud, E. Mora, V. Opatova, P. Sousa, M. Vadell and N. Yiğit. We also thank M.A. Arnedo and M. Metallinou for intellectual guidance on all aspects of this study. We are grateful to N. Barve and A. Lira for providing R scripts. Funding for this research was provided by CGL2008-03385 Project (Ministerio de Ciencia y Tecnología, Spain). E.P. was supported by a FPI grant from the Ministerio de Ciencia y Tecnología, Spain (BES-2009-015871), and E.E.S. was supported by a University of Kansas Self Graduate Fellowship and an NSF GK-12 Fellowship. We acknowledge the World Climate Research Programmer's Working Group on Coupled Modeling, which is responsible for CMIP5, and we thank the climate modeling groups (listed in Supporting Information Additional file 2 of this paper) for producing and making their model outputs available.

### Author details

<sup>1</sup>Institut de Recerca de la Biodiversitat (IRBio), Departament de Biologia Animal, Universitat de Barcelona, Barcelona, Spain. <sup>2</sup>Department of Geology, University of Kansas, Lawrence, Kansas, USA. <sup>3</sup>Biodiversity Institute, University of Kansas, Lawrence, Kansas, USA. <sup>4</sup>Departamento de Ciências Biológicas, Universidade Federal de Goiás, Campus Jataí, Jataí, GO, Brasil.

Received: 3 March 2014 Accepted: 22 August 2014

Published online: 09 October 2014

### References

- Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB, Kent J: **Biodiversity hotspots for conservation priorities.** *Nature* 2000, **403**:853–858.
- Abulafia D: *The Great Sea: A Human History of the Mediterranean.* New York: Oxford University Press; 2011.
- Geri F, Amici V, Rocchini D: **Human activity impact on the heterogeneity of a Mediterranean landscape.** *Appl Geogr* 2010, **30**:370–379.
- Blondel J, Aronson J, Bodiou JY, Boeuf G: *The Mediterranean Region: Biological Diversity in Space and Time.* New York: Oxford University Press; 2010.
- Médail F, Diadema K: **Glacial refugia influence plant diversity patterns in the Mediterranean Basin.** *J Biogeogr* 2009, **36**:1333–1345.
- Koch PL, Barnosky AD: **Late Quaternary extinctions: state of the debate.** *Annu Rev Ecol Syst* 2006, **37**:215–250.
- Postigo-Mijarra JM, Morla C, Barrón E, Morales-Molino C, García S: **Patterns of extinction and persistence of Arctotertiary flora in Iberia during the Quaternary.** *Rev Palaeobot Palynol* 2010, **162**:416–426.
- Hewitt GM: **The genetic legacy of the Quaternary ice ages.** *Nature* 2000, **405**:907–913.
- Weiss S, Ferrand N: *Phylogeography of Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of European Biodiversity.* Dordrecht: Springer; 2007.
- Feliner GN: **Southern European glacial refugia: a tale of tales.** *Taxon* 2011, **60**:365–372.
- Gómez A, Lunt DH: **Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula.** In *Phylogeography of Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of European Biodiversity.* Edited by Weiss S, Ferrand N. Dordrecht: Springer; 2007:155–188.
- Stewart JR, Lister AM: **Cryptic northern refugia and the origins of the modern biota.** *Trends Ecol Evol* 2001, **16**:608–613.
- Schmitt T, Varga Z: **Extra-Mediterranean refugia: the rule and not the exception?** *Front Zool* 2012, **9**:22.
- Tzedakis PC, Emerson BC, Hewitt GM: **Cryptic or mystic? Glacial tree refugia in northern Europe.** *Trends Ecol Evol* 2013, **28**:696–704.
- Elena H, Peyron O, Bonnefille R, Jolly D, Cheddadi R, Guiot J, Andrieu V, Bottema S, Buchet G, Hamilton AC, Maley J, Marchant R, Reille M, Riollet G, Scott L, Straka H, Taylor D, Van Campo E, Vincens A, Laarif F, Jonson H: **Pollen-based biome reconstruction for southern Europe and Africa 18,000 yr BP.** *J Biogeogr* 2000, **27**:621–634.
- Tzedakis PC, Lawson IT, Frogley MR, Hewitt GM, Preece RC: **Buffered tree population changes in a Quaternary refugium: evolutionary implications.** *Science* 2002, **297**:2044–2047.
- Carrion Y, Ntinou M, Badal E: ***Olea europaea* L. in the north Mediterranean Basin during the Pleniglacial and the Early–Middle Holocene.** *Quat Sci Rev* 2010, **29**:952–968.
- Médail F, Quezel P: **Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin.** *Ann Missouri Bot Gard* 1997, **84**:112–127.
- Besnard G, Khadari B, Navascués M, Fernández-Mazuecos M, El Bakkali A, Arrigo N, Baali-Cherif D, de Caraffa Brunini-Bronzini V, Santoni S, Vargas P, Savolainen V: **The complex history of the olive tree: from Late Quaternary diversification of Mediterranean lineages to primary domestication in the northern Levant.** *Proc Royal Soc B* 2013, **280**:20122833.
- Carstens BC, Richards CL: **Integrating coalescent and ecological niche modeling in comparative phylogeography.** *Evolution* 2007, **61**:1439–1454.
- Waltari E, Hijmans RJ, Peterson AT, Nyári Á, Perkins SL, Guralnick RP: **Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions.** *PLoS ONE* 2007, **2**:e563.
- Peterson AT, Nyári Á: **Ecological niche conservatism and pleistocene refugia in the Thrush-like Mourner, *Schiffornis* sp., in the Neotropics.** *Evolution* 2008, **62**:173–183.
- Peterson AT: **Phylogeography is not enough: the need for multiple lines of evidence.** *Front Biogeogr* 2009, **1**:19–25.
- Collevatti RG, Terribile LV, Oliveira G, Lima-Ribeiro MS, Nabout JC, Rangel TF, Diniz-Filho JAF: **Drawbacks to palaeodistribution modelling: the case of South American seasonally dry forests.** *J Biogeogr* 2013, **40**:345–358.
- Peterson AT, Martínez-Meyer E, González-Salazar C: **Reconstructing the Pleistocene geography of the *Aphelocoma* jays (Corvidae).** *Divers Distrib* 2004, **10**:237–246.
- Richards CL, Carstens BC, Knowles L: **Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses.** *J Biogeogr* 2007, **34**:1833–1845.
- Wilson JS, Pitts JP: **Identifying Pleistocene refugia in North American cold deserts using phylogeographic analyses and ecological niche modelling.** *Divers Distrib* 2012, **18**:1139–1152.
- Jakob SS, Ihlow A, Blattner FR: **Combined ecological niche modelling and molecular phylogeography revealed the evolutionary history of *Hordeum marinum* (Poaceae) niche differentiation, loss of genetic diversity, and speciation in Mediterranean Quaternary refugia.** *Mol Ecol* 2007, **16**:1713–1727.
- Lozier JD, Mills NJ: **Ecological niche models and coalescent analysis of gene flow support recent allopatric isolation of parasitoid wasp populations in the Mediterranean.** *PLoS ONE* 2009, **4**:e5901.
- Gertsch WJ: **The spider genus *Loxosceles* in South America (Araneae, Scytodidae).** *Bull Am Museum Nat Hist* 1967, **136**:117–174.
- Binford GJ, Callahan MS, Bodner MR, Rynerson MR, Núñez PB, Ellison CE, Duncan RP: **Phylogenetic relationships of *Loxosceles* and *Sicarius* spiders are consistent with western Gondwanan vicariance.** *Mol Phylogenet Evol* 2008, **49**:538–553.
- Duncan RP, Rynerson MR, Ribera C, Binford GJ: **Diversity of *Loxosceles* spiders in Northwestern Africa and molecular support for cryptic species in the *Loxosceles rufescens* lineage.** *Mol Phylogenet Evol* 2010, **55**:234–248.
- Ribera C, Planas E: **A new species of *Loxosceles* (Araneae, Sicariidae) from Tunisia.** *ZooKeys* 2009, **16**:217–225.
- Planas E, Ribera C: **Uncovering overlooked island diversity: colonization and diversification of the medically important spider genus *Loxosceles* (Arachnida: Sicariidae) on the Canary Islands.** *J Biogeogr* 2014, **41**:1255–1266.
- Platnick NI: **The world spider catalog, version 14.5.** American Museum of Natural History. In <http://research.amnh.org/iz/spiders/catalog/index.html>.
- Brignoli P: **Note sugli Scytodidae d'Italia e Malta (Araneae).** *Fragm Ent* 1969, **6**:121–166.
- Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P: **Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers.** *Ann Entomol Soc Am* 1994, **87**:651–701.

38. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R: **DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates.** *Mol Mar Biol Biotechnol* 1994, **3**:294–299.
39. Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A: **Geneious v.4.6.5.** <http://www.geneious.com>.
40. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22**:4673–4680.
41. Lanfear R, Calcott B, Ho SYW, Guindon S: **Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses.** *Mol Biol Evol* 2012, **29**:1695–1701.
42. Stamatakis A: **RaxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models.** *Bioinformatics* 2006, **22**:2688–2690.
43. Silvestro D, Michalak I: **raxMLGUI: a graphical front-end for RAXML.** *Org Divers Evol* 2011, **12**:335–337.
44. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP: **MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space.** *Syst Biol* 2012, **61**:539–542.
45. Rambaut A, Drummond AJ: **Tracer v1.4.** [<http://beast.bio.ed.ac.uk/Tracer>]
46. Drummond AJ, Suchard MA, Xie D, Rambaut A: **Bayesian phylogenetics with BEAUti and the BEAST 1.7.** *Mol Biol Evol* 2012, **29**:1969–1973.
47. Drummond AJ, Rambaut A: **BEAST: Bayesian evolutionary analysis by sampling trees.** *BMC Evol Biol* 2007, **7**:214.
48. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A: **Relaxed phylogenetics and dating with confidence.** *PLoS Biol* 2006, **4**:e88.
49. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: **MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.** *Mol Biol Evol* 2011, **28**:2731–2739.
50. Fu Y: **Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection.** *Genetics* 1997, **147**:915–925.
51. Ramos-Onsins SE, Rozas J: **Statistical properties of new neutrality tests against population growth.** *Mol Biol Evol* 2002, **19**:2092–2100.
52. Librado P, Rozas J: **DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.** *Bioinformatics* 2009, **25**:1451–1452.
53. Estelstyn JA, Evans BJ, Sedlock JL, Anwarali Khan FA, Heaney LR: **Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats.** *Proc R Soc B Biol Sci* 2012, **279**:3678–3686.
54. Papadopoulou A, Cardoso A, Gómez-Zurita J: **Diversity and diversification of Eumolpinae (Coleoptera: Chrysomelidae) in New Caledonia.** *Zool J Linn. Soc* 2013, **168**:473–495.
55. Lohse K: **Can mtDNA barcodes be used to delimit species? A response to Pons et al. (2006).** *Syst Biol* 2009, **58**:439–442.
56. Fujisawa T, Barraclough TG: **Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent (GMYC) approach: a revised method and evaluation on simulated data sets.** *Syst Biol* 2013, **65**:707–724.
57. Satler JD, Carstens BC, Hedin M: **Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*).** *Syst Biol* 2013, **62**:805–823.
58. Pons J, Barraclough T, Gomez-Zurita J, Cardoso A, Duran D, Hazell S, Kamoun S, Sumlin W, Vogler A: **Sequence-based species delimitation for the DNA taxonomy of undescribed insects.** *Syst Biol* 2006, **55**:595–609.
59. Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K: **Estimating divergence times in large phylogenetic trees.** *Syst Biol* 2007, **56**:741–752.
60. R Core Team: *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2012.
61. Paradis E, Claude J, Strimmer K: **APE: analyses of phylogenetics and evolution in R language.** *Bioinformatics* 2004, **20**:289–290.
62. Hart MW, Sunday J: **Things fall apart: biological species form unconnected parsimony networks.** *Biol Lett* 2007, **3**:509–512.
63. Templeton AR, Crandall KA, Sing F: **Cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data III. Cladogram estimation.** *Genetics* 1992, **132**:619–633.
64. Clement M, Posada D, Crandall KA: **TCS: a computer program to estimate gene genealogies.** *Mol Ecol* 2000, **9**:1657–1659.
65. Arrigo N, Felber P, Parisod C, Buerki S, Alvarez N, David J, Guadagnuolo R: **Origin and expansion of the allotetraploid *Aegilops geniculata*, a wild relative of wheat.** *New Phytol* 2010, **187**:1170–1180.
66. Ehrich D, Alsos IG, Brochmann C: **Where did the northern peatland species survive the dry glacials: cloudberry (*Rubus chamaemorus*) as an example.** *J Biogeogr* 2008, **35**:801–814.
67. Jensen JL, Bohonak AJ, Kelley ST: **Isolation by distance web service.** *BMC Genet* 2005, **6**:13.
68. Ersts PJ: **Geographic Distance Matrix Generator v.1.2.3.** [http://biodiversityinformatics.amnh.org/open\\_source/gdmg](http://biodiversityinformatics.amnh.org/open_source/gdmg).
69. Bidegaray-Batista L, Arnedo MA: **Gone with the plate: the opening of the western Mediterranean basin drove the diversification of ground-dweller spiders.** *BMC Evol Biol* 2011, **11**:317.
70. Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J, Villalobos F: **The crucial role of the accessible area in ecological niche modeling and species distribution modeling.** *Ecol Model* 2011, **222**:1810–1819.
71. Peterson AT: **Ecological niche conservatism: a time-structured review of evidence.** *J Biogeogr* 2011, **38**:817–827.
72. CMIP5. In <http://cmip-pcmdi.llnl.gov>.
73. PMIP3. In <http://pmip3.lscce.ipsl.fr>.
74. Taylor KE, Stouffer RJ, Meehl GA: **An overview of CMIP5 and the experiment design.** *Bull Am Meteorol Soc* 2012, **93**:485–498.
75. Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO: **Guidelines for use of climate scenarios developed from statistical downscaling methods.** [http://ipcc-ddc.cru.uea.ac.uk/guidelines/dgm\\_no2\\_v1\\_09\\_2004.pdf](http://ipcc-ddc.cru.uea.ac.uk/guidelines/dgm_no2_v1_09_2004.pdf).
76. WorldClim. In <http://www.worldclim.org>.
77. Hijmans RJ, Graham CH: **The ability of climate envelope models to predict the effect of climate change on species distributions.** *Glob Chang Biol* 2006, **12**:2272–2281.
78. Phillips SJ, Anderson RP, Schapire RE: **Maximum entropy modeling of species geographic distributions.** *Ecol Modell* 2006, **190**:231–259.
79. Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ: **A statistical explanation of MaxEnt for ecologists.** *Divers Distrib* 2011, **17**:43–57.
80. Owens HL, Campbell LP, Dornak LL, Saupe EE, Barve N, Soberón J, Ingenloff K, Lira-Noriega A, Hensz CM, Myers CE, Peterson AT: **Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas.** *Ecol Model* 2013, **263**:10–18.
81. Elith J, Kearney M, Phillips S: **The art of modelling range-shifting species.** *Methods Ecol Evol* 2010, **1**:330–342.
82. Peterson AT, Papeş M, Soberón J: **Rethinking receiver operating characteristic analysis applications in ecological niche modeling.** *Ecol Model* 2008, **213**:63–72.
83. Fordham DA, Wigley TML, Watts MJ, Brook BW: **Strengthening forecasts of climate change impacts with multi-model ensemble averaged projections using MAGICC/SCENGEN 5.3.** *Ecography* 2012, **35**:4–8.
84. **Constraints on Interpretation of Ecological Niche Models by Limited Environmental Ranges on Calibration Areas: Software Script Appendix.** <http://hdl.handle.net/1808/10122>.
85. **Niche Modeling: Model Evaluation.** In <http://hdl.handle.net/1808/10059>.
86. Lisiecki LE, Raymo ME: **A Pliocene-Pleistocene stack of 57 globally distributed benthic  $\delta^{18}O$  records.** *Paleoceanography* 2005, **20**, PA1003.
87. Rodríguez-Sánchez F, Arroyo J: **Reconstructing the demise of Tethyan plants: climate-driven range dynamics of *Laurus* since the Pliocene.** *Glob Ecol Biogeogr* 2008, **17**:685–695.
88. Tzedakis P, Roucoux K, Abreu L, De Shackleton N: **The duration of forest stages in southern Europe and interglacial climate variability.** *Science* 2004, **306**:2231–2235.
89. Rull V: **On microrefugia and cryptic refugia.** *J Biogeogr* 2010, **37**:1623–1625.
90. Carstens BC, Pelletier TA, Reid NM, Satler JD: **How to fail at species delimitation.** *Mol Ecol* 2013, **22**:4369–4383.
91. Husemann M, Schmitt T, Zachos FE, Ulrich W, Habel JC: **Palaeoarctic biogeography revisited: evidence for the existence of a North African refugium for Western Palaeoarctic biota.** *J Biogeogr* 2013, **41**:81–94.
92. Greene A, Breisch N, Boardman T: **The Mediterranean Recluse Spider, *Loxosceles rufescens* (Dufour): an abundant but cryptic inhabitant of deep infrastructure in the Washington, DC Area (Arachnida: Araneae: Sicariidae).** *Am Entomol* 2009, **55**:158–163.
93. Kobelt M, Nentwig W: **Alien spider introductions to Europe supported by global trade.** *Divers Distrib* 2007, **14**:273–280.
94. Khadari B, Grout C, Santoni S, Kjellberg F, Hy F: **Contrasted genetic diversity and differentiation among Mediterranean populations of *Ficus carica* L.: a study using mtDNA RFLP.** *Genet Resour Crop Ev* 2005, **52**:97–109.



95. Delplancke M, Alvarez N, Benoit L, Espindola A, Joly HJ, Neuenschwander S, Arrigo N: **Evolutionary history of almond tree domestication in the Mediterranean Basin.** *Mol Ecol* 2013, **22**:1092–1104.
96. Zeder MA: **Domestication and early agriculture in the Mediterranean Basin: origins, diffusion, and impact.** *Proc Natl Acad Sci* 2008, **105**:11597–11604.
97. Carranza S, Arnold EN: **Systematics, biogeography, and evolution of *Hemidactylus* geckos (Reptilia: Gekkonidae) elucidated using mitochondrial DNA sequences.** *Mol Phylogenet Evol* 2006, **38**:531–545.
98. Jesse R, Véla E, Pfenninger M: **Phylogeography of a land snail suggests trans-Mediterranean neolithic transport.** *PLoS ONE* 2011, **6**:e20734.
99. Guiller A, Madec L: **Historical biogeography of the land snail *Cornu aspersum*: a new scenario inferred from haplotype distribution in the Western Mediterranean basin.** *BMC Evol Biol* 2010, **10**:18.
100. Porretta D, Canestrelli D, Urbanelli S, Bellini R, Schaffner F, Petric D, Nascetti G: **Southern crossroads of the Western Palaearctic during the Late Pleistocene and their imprints on current patterns of genetic diversity: insights from the mosquito *Aedes caspius*.** *J Biogeogr* 2011, **38**:20–30.
101. Solà E, Sluys R, Gritzalis K, Riutort M: **Fluvial basin history in the northeastern Mediterranean region underlies dispersal and speciation patterns in the genus *Dugesia* (Platyhelminthes, Tricladida, Dugesiidae).** *Mol Phylogenet Evol* 2013, **66**:877–888.
102. Hewitt GM: **Mediterranean Peninsulas: The Evolution of Hotspots.** In *Biodiversity Hotspots*. Edited by Zachos FE, Habel JC. Berlin Heidelberg: Springer; 2011:123–147.
103. Rokas A, Atkinson RJ, Webster L, Csoka G, Stone GN: **Out of Anatolia: longitudinal gradients in genetic diversity support an eastern origin for a circum-Mediterranean oak gallwasp *Andricus quercustozae*.** *Mol Ecol* 2003, **12**:2153–2174.
104. Salvi D, Harris DJ, Kaliontzopoulou A, Carretero MA, Pinho C: **Persistence across Pleistocene ice ages in Mediterranean and extra-Mediterranean refugia: phylogeographic insights from the common wall lizard.** *BMC Evol Biol* 2013, **13**:147–165.
105. Kindler C, Böhme W, Corti C, Gvoždík V, Jablonski D, Jandzik D, Metallinou M, Široký P, Fritz U: **Mitochondrial phylogeography, contact zones and taxonomy of grass snakes (*Natrix natrix*, *N. megalcephala*).** *Zool Scripta* 2013, **42**:458–472.
106. Planas E, Fernández-Montraveta C, Ribera C: **Molecular systematics of the wolf spider genus *Lycosa* (Araneae: Lycosidae) in the Western Mediterranean Basin.** *Mol Phylogenet Evol* 2013, **67**:414–428.
107. Nogués-Bravo D: **Predicting the past distribution of species climatic niches.** *Glob Ecol Biogeogr* 2009, **18**:521–531.
108. Widmer A, Lexer C: **Glacial refugia: sanctuaries for allelic richness, but not for gene diversity.** *Trends Ecol Evol* 2001, **16**:267–269.

doi:10.1186/s12862-014-0195-y

**Cite this article as:** Planas et al.: Ecological niche and phylogeography elucidate complex biogeographic patterns in *Loxosceles rufescens* (Araneae, Sicariidae) in the Mediterranean Basin. *BMC Evolutionary Biology* 2014 **14**:195.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

