Evol Biol (2017) 44:505–521
DOI 10.1007/s11692-017-9423-x

RESEARCH ARTICLE

Geographical and Ecological Drivers of Mitonuclear Genetic Divergence in a Mediterranean Grasshopper

J. Ortego[1] · V. Noguerales[2] · P. J. Cordero[2]

Received: 21 January 2017 / Accepted: 28 June 2017 / Published online: 5 July 2017
© Springer Science+Business Media, LLC 2017

Abstract The study of the neutral and/or selective processes driving genetic variation in natural populations is central to determine the evolutionary history of species and lineages and understand how they interact with different historical and contemporary components of landscape heterogeneity. Here, we combine nuclear and mitochondrial data to study the processes shaping genetic divergence in the Mediterranean esparto grasshopper (*Ramburiella hispanica*). Our analyses revealed the presence of three main lineages, two in Europe that split in the Early-Middle Pleistocene and one in North Africa that diverged from the two European ones after the Messinian. Lineage-specific potential distribution models and tests of environmental niche differentiation suggest that the phylogeographic structure of the species was driven by allopatric divergence due to the re-opening of the Gibraltar strait at the end of the Messinian (Europe–Africa split) and population fragmentation in geographically isolated Pleistocene climatic refugia (European split). Although we found no evidence for environment as an important driver of genetic divergence at the onset of lineage formation, our analyses considering the spatial distribution of populations and different aspects of landscape composition suggest that genetic differentiation at mitochondrial loci was largely explained by environmental dissimilarity, whereas resistance-based estimates of geographical distance were the only predictors of genetic differentiation at nuclear markers. Overall, our study shows that although historical factors have largely shaped concordant range-wide patterns of mitonuclear genetic structure in the esparto grasshopper, different contemporary processes (neutral gene flow vs. environmental-based selection) seem to be governing the spatial distribution of genetic variation in the two genomes.

Keywords Environmental niche modelling · Isolation-by-environment · Isolation-by-resistance · Mitonuclear discordance

Introduction

The study of the neutral and/or selective processes underlying genetic divergence has major implications for understating and preserving the idiosyncratic evolutionary trajectories of species, lineages and populations (Moritz and Potter 2013). The continuum of genetic divergence ranges from gradual differentiation among populations shaped by geographical distance (Wright 1943; Slatkin 1993; Hutchison and Templeton 1999) or environmental gradients (Shafer and Wolf 2013; Sexton et al. 2014; Wang and Bradburd 2014) to complete disruption of gene flow and species formation as a result of total spatial or ecological isolation (Graham et al. 2004; Martin and Mendelson 2012; Nosil 2012). The relative role of environment and geographic isolation in shaping genetic differentiation of natural populations has received much attention in the last years and a growing number

Electronic supplementary material The online version of this article (doi:10.1007/s11692-017-9423-x) contains supplementary material, which is available to authorized users.

✉ J. Ortego
joaquin.ortego@csic.es

[1] Department of Integrative Ecology, Estación Biológica de Doñana, EBD-CSIC, Avda. Américo Vespucio 26, 41092 Seville, Spain

[2] Grupo de Investigación de la Biodiversidad Genética y Cultural, Instituto de Investigación en Recursos Cinegéticos - IREC (CSIC, UCLM, JCCM), Ronda de Toledo 12, 13071 Ciudad Real, Spain

of empirical and experimental studies have revealed the importance of the former on the evolution of spatial genetic structure (e.g. Lee and Mitchell-Olds 2011; Wang et al. 2013) and incipient speciation processes (e.g. Soria-Carrasco et al. 2014). Accordingly, recent meta-analyses have shown that genetic divergence induced by environmental heterogeneity is pervasive across timescales and taxa and that their effects often override those of geographical isolation (Shafer and Wolf 2013; Sexton et al. 2014). However, the analysis of environmental heterogeneity as a force of evolutionary change still remains elusive due to the strong spatial autocorrelation of ecological variables and the challenge of teasing apart their effects from those of geographic distance (Bradburd et al. 2013; Wang et al. 2013; Noguerales et al. 2016).

The study of the factors structuring genetic variation in natural populations has been approached with different kinds of nuclear or mitochondrial genetic markers, with an increasingly high number of studies employing loci from both genomes (Avise 1994; Toews and Brelsford 2012). Beyond the notion that more markers provide more robust phylogenetic and demographic inferences (Brito and Edwards 2009; Edwards and Bensch 2009; Gaspari et al. 2015; Kumar et al. 2017), the integration of data from nuclear and mitochondrial genomes has yielded important information on the evolutionary history of organisms and how they interact with different aspects of landscape heterogeneity (e.g. Thorpe et al. 2008; Pavlova et al. 2013; Bar-Yaacov et al. 2015; Rosetti and Remis 2017). Several studies have reported the presence of mitonuclear discordances in spatial patterns of genetic structure, an intriguing phenomenon that has attracted the attention of evolutionary biologists in the last two decades (Avise 1994; Zink and Barrowclough 2008; Toews and Brelsford 2012). Different causes have been proposed to explain mitonuclear discordances, including shorter coalescence times of mitochondrial genes, sex-biased dispersal, and asymmetric introgression, among others (Toews and Brelsford 2012). Variation in mitochondrial genes has been long-held assumed to be primarily neutral (Ballard and Whitlock 2004; Meiklejohn et al. 2007; Galtier et al. 2009), but many studies have identified different forms of selection on mitochondrial genes linked with their important metabolic functions (e.g. Pichaud et al. 2012; Morales et al. 2015; Latorre-Pellicer et al. 2016). However, despite accumulating experimental evidence on the importance of the mitochondrial genome on different aspects of adaptive change, most phylogeographic studies on natural populations have focused on reporting spatial incongruence of genetic structure between the two genomes whereas less attention has been paid on analyzing the specific drivers of such disparities (Toews and Brelsford 2012; for some exceptions see; Cheviron and Brumfield 2009; Ribeiro et al. 2011; Pavlova et al. 2013).

Here, we study the processes underlying range-wide patterns of genetic structure in the esparto grasshopper (*Ramburiella hispanica*, Rambur 1838), a specialist insect distributed in Mediterranean environments from France, Spain and northwest Africa. This grasshopper is generally linked to esparto grass formations and previous landscape genetic studies have shown that this taxon is highly sensitive to the fragmentation of its particular habitats (Ortego et al. 2015a, b). This suggests that climate changes during the Pleistocene may have contributed to fragment its populations (Hewitt 2000, 2004) whereas the opening of the Gibraltar strait at the end of the Messinian (5.3 mya BP; Krijgsman et al. 1999) probably resulted in allopatric divergence of European and North African populations (Sanmartín 2003; Pinho et al. 2006). For these reasons, the esparto grasshopper is a good candidate for analyzing patterns of genetic divergence at different time scales spanning from landscape-level differentiation to long-term isolation driven by the past geological and climatic history of the Mediterranean region (Blondel and Aronson 1999). First, we (i) employed nuclear and mitochondrial markers to determine range-wide patterns of genetic structure and infer the demographic history of the species. In a second step, to elucidate the potential mechanisms of genetic divergence (allopatric vs. environment-mediated divergence), we (ii) linked observed patterns of genetic structure with current and past species distribution models, (iii) tested for environmental niche differentiation among inferred lineages (Warren et al. 2008; Gotelli and Stanton-Geddes 2015), and (iv) analyzed the relative contribution of geography (isolation-by-distance, IBD; Wright 1943; Slatkin 1993), the spatial distribution of suitable habitats (isolation-by-resistance, IBR; McRae and Beier 2007; McRae et al. 2008) and environmental dissimilarity (isolation-by-environment, IBE; Wang and Bradburd 2014) to explain observed patterns of genetic divergence. Finally, (v) we tested for mitonuclear concordant/discordant evolutionary trajectories by analyzing the processes (geography vs. environment) driving genetic differentiation in both genomes at different spatiotemporal scales (i.e. within and across lineages) (e.g. Thorpe et al. 2008; Pavlova et al. 2013; Sun et al. 2015).

Methods

Sampling

Between 2010 and 2014, we sampled a total of 707 esparto grasshoppers from 49 localities expanding the entire distribution range of the species in France, Spain and Morocco (Supplementary Table S1). We collected 1–22 adult individuals per population and specimens were preserved whole in 1500 µL ethanol 96% at −20 °C until needed for

genetic analyses. Population code description and further information on sampling locations are given in Supplementary Table S1. All population-based analyses were performed only considering those localities with five or more genotyped individuals (see below).

Microsatellite Data

We used a salt extraction protocol to purify genomic DNA from a hind leg of each individual (Aljanabi and Martinez 1997). We used 10 highly polymorphic microsatellite markers previously developed for esparto grasshopper (RhA105, RhA108, RhA112, RhB107, RhA2, RhC2, RhC112, RhC113, RhD2, and RhB2) to genotype each sampled individual (Aguirre et al. 2014; Ortego et al. 2015a). Amplifications were conducted in 10-µL reaction volumes containing approximately 20 ng of template DNA, 1× reaction buffer (67 mM Tris–HCL, pH 8.3, 16 mM $(NH_4)_2SO_4$, 0.01% Tween-20, EcoStart Reaction Buffer, Ecogen, Madrid, Spain), 2 mM $MgCl_2$, 0.2 mM of each dNTP, 0.15 µM of each dye-labelled primer (FAM, PET, VIC or NED) and 0.1 U of *Taq* DNA EcoStart Polymerase (Ecogen). The PCR programme used was 9 min denaturing at 95 °C followed by 40 cycles of 30 s at 94 °C, 45 s at the annealing temperature (Aguirre et al. 2014) and 45 s at 72 °C, ending with a 10 min final elongation stage at 72 °C. Amplification products were electrophoresed using an ABI 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) and genotypes were scored using GeneMapper 3.7 (Applied Biosystems).

Microsatellite loci were tested for departure from Hardy–Weinberg equilibrium within each sampling population using an exact test (Guo and Thompson 1992) based on 900,000 Markov chain iterations as implemented in the program Arlequin 3.1 (Excoffier et al. 2005). We also used Arlequin 3.1 to test for linkage disequilibrium between each pair of loci for each sampling population using a likelihood-ratio statistic, whose distribution was obtained by a permutation procedure (Excoffier et al. 2005). We applied sequential Bonferroni corrections to account for multiple comparisons.

Mitochondrial DNA Data

We used universal primers LCO1490 and HCO2198 (Folmer et al. 1994) and 16sar and 16sbr (Palumbi et al. 1991) to amplify and sequence a fragment of the mitochondrial cytochrome oxidase subunit I (COI) and 16S ribosomal RNA (16S) genes, respectively. Reagents and PCR program were the same than used for microsatellite amplifications, but reactions were performed in 15 µL volumes and with annealing temperatures at 50 °C for COI and 60 °C for 16S. Amplified products were commercially purified and sequenced (Macrogen, South Korea). Sequences were edited and aligned using Sequencher 5.0 (GeneCodes Corporation). There were no insertions or deletions and all sequences were trimmed to the same length (639 bp for COI and 491 bp for 16S). We also used Sequencher 5.0 to determine the reading frame for COI gene fragment and confirm the absence of internal stop codons that could suggest the amplification of pseudo-genes (e.g. nuclear mitochondrial sequences or NUMTs). All sequences were deposited in GenBank with accession numbers KX425019-KX425270.

For each mtDNA gene fragment sequenced, we determined the number of haplotypes, calculated the number of polymorphic sites (S), haplotype diversity (Hd) and nucleotide diversity (π), and performed Tajima's D (Tajima 1989) and Fu's F_S (Fu 1997) tests for selective neutrality using DnaSP 5.10 (Librado and Rozas 2009). We also used DnaSP 5.10 to test for selection on COI gene using the McDonald–Kreitman (MK) test (McDonald and Kreitman 1991). In order to make fully comparable the analyses for COI and 16S, these tests were performed only considering individuals with sequences available for both gene fragments ($n = 242$).

Analyses of Population Genetic Structure

For microsatellite markers, we analysed patterns of genetic structure using the Bayesian Markov chain Monte Carlo clustering analysis implemented in the program Structure 2.3.3 (Pritchard et al. 2000; Hubisz et al. 2009). We ran Structure (707 individuals collected from 49 sampling localities; Table S1) assuming correlated allele frequencies and admixture and both using and not using prior population information in two independent sets of analyses (Hubisz et al. 2009). We conducted ten independent runs for each value of $K = 1$–15 to estimate the number of clusters that best fit the data with 200,000 MCMC cycles, following a burn-in step of 100,000 iterations. These settings resulted in a good convergence among runs for all summary statistics provided by the program (e.g. α, F, D and the likelihood) (Pritchard et al. 2000). We run Structure for a maximum of $K = 15$ because (i) many of our 49 sampling localities are very closely located and genetically unstructured according to clustering analyses from previous studies on the species at the landscape scale (Ortego et al. 2015a, b) and (ii) log probabilities [Pr($X|K$)] reached an asymptote around $K = 4$–8, supporting that the number of genetic clusters is much lower than the number of sampling localities (see "Results" section). The number of populations best fitting the dataset was defined using log probabilities [Pr($X|K$)] (Pritchard et al. 2000) and the ΔK method (Evanno et al. 2005), as implemented in Structure Harvester (Earl and vonHoldt 2012). We used Clumpp 1.1.2 and the Greedy algorithm to align multiple runs of

Structure for the same K value (Jakobsson and Rosenberg 2007) and Distruct 1.1 (Rosenberg 2004) to visualize as bar plots the individual's probabilities of population membership.

To visualize the phylogenetic relationship among all populations based on microsatellite data, we built a population-based neighbour-joining (NJ) tree in Populations 1.2.30 (Langella 1999) using Cavalli-Sforza and Edwards (1967) chord distances and allele frequency data. This genetic distance has been suggested to be the most effective in recovering the correct tree topology for microsatellite markers under a variety of evolutionary scenarios without making assumptions regarding constant population size or mutation rates among loci (Takezaki and Nei 1996).

Population genetic structure for the two mtDNA gene fragments was analysed using a spatial analysis of molecular variance as implemented by Samova 2.0 (Dupanloup et al. 2002). This method uses a simulated annealing procedure to identify the optimal grouping option (K) for the data by maximizing the among-group component (F_{CT}) of the overall genetic variance. The analysis was run with default parameters and 500 simulated annealing processes for different number of populations ($K=2$–15). F_{CT} values were used to determine the most likely population clustering solution (Dupanloup et al. 2002).

The distribution of genetic variation in nuclear microsatellite markers (F_{ST}) and mtDNA sequences (Φ_{ST}) was assessed and compared using Amovas as implemented in Arlequin 3.1 (Excoffier et al. 2005). Genetic variation was hierarchically partitioned into among groups of populations defined by their geographical location (see Supplementary Table S1), among populations within groups, and among individuals within populations. Significance was tested using 10,000 permutations of the original data.

Phylogenetic Analyses, Divergence Times and Demography

We inferred an ultrametric tree and estimated divergence times (i.e. time to the most recent common ancestor, MRCA) for mtDNA sequences (COI and 16S gene fragments) using Beast 1.8.3 (Drummond et al. 2012). We used a GTR + I + G model of sequence evolution, which was determined as the best-fitting nucleotide substitution model for both gene fragments in jModelTest2 (Darriba et al. 2012). We ran several analyses considering different clock and demographic models (Supplementary Table S2), each with two independent Markov chains of 100 million generations sampled every 10,000 generations (i.e. 10,000 retained genealogies). We calibrated molecular clocks using mutation rates previously obtained for other insects (mean ± S.D.; COI = 0.0169 ± 0.0019 substitutions/site/My; 16S = 0.0049 ± 0.0008 substitutions/site/My; Papadopoulou et al. 2010). We used Tracer 1.4 to examine log files and check stationarity and convergence of the chains and confirm that effective sampling sizes (ESS) for all parameters were greater than 200. We removed 10% of trees as burn-in and combined tree and log files for replicated runs using LogCombiner 1.8.3. We used TreeAnnotator 1.8.3 to obtain maximum credibility trees and FigTree 1.4.2 to visualize final trees. Finally, we determined the best-fitting clock and demographic model using the Akaike's information criterion through Markov chain Monte Carlo (AICM; Baele et al. 2012) with 100 bootstraps as implemented in Tracer 1.6 (e.g. Magalhaes et al. 2014). We modelled demographic changes within each of the three main mtDNA clades inferred (see "Results" section) using Bayesian Skyline Plots (BSP) in Beast 1.8.3 (Drummond et al. 2012). We assumed a strict molecular clock, as exploratory runs did not support a relaxed clock.

Environmental Niche Modelling

We used environmental niche modelling (ENM) to predict the geographic distribution of suitable habitats for esparto grasshopper both in the present and during the last glacial maximum (LGM). We modelled current distributions using Maxent 3.3.3 (Phillips et al. 2006; Phillips and Dudik 2008). Species occurrence data were obtained from sampling localities ($n=49$; Table S1) as well as from records available from the literature ($n=56$) and at the Global Biodiversity Information Facility (http://www.gbif.org/) ($n=56$). Prior to modelling, all records were mapped and examined to identify and exclude those having obvious georeferencing errors. For models, a single record among those falling within the same grid cell was used, resulting in a final dataset of 153 entries. To construct the models, we used 19 bioclimatic variables from the WorldClim dataset (http://www.worldclim.org/) interpolated to 30-arcsec (c. 1-km) resolution (Hijmans et al. 2005). Then, we estimated the distribution of the species at the Last Glacial Maximum (LGM; c. 21,000 years BP) projecting contemporary species-climate relationships to this period. We used the same bioclimatic layers available at WorldClim for two palaeoclimate models of the last glacial period: the Community Climate System Model (CCSM; Kiehl and Gent 2004) and the Model for Interdisciplinary Research on Climate (MIROC; Hasumi and Emori 2004). To avoid problems due to extrapolating distributions into novel climates (Elith et al. 2011), we followed the approach described in detail in Massatti and Knowles (2014). We selected a final set of five layers to construct the models: maximum temperature of warmest month (Bio5), mean temperature of the wettest quarter (Bio8), mean temperature of warmest quarter (Bio10), annual precipitation (Bio12), and precipitation of the driest month (Bio14). Model evaluation statistics were produced

from 10 cross-validation replicate model runs and overall model performance was assessed using the area under the receiving operator characteristics curve (AUC).

Niche Divergence

We used identity and background tests in ENMTOOLS to analyze whether environmental niches are differentiated or conserved between each pair of the three main mtDNA clades recovered for the species (see "Results" section for details) (Warren et al. 2008; e.g.; Nakazato et al. 2010). For this purpose, we first built an ENM for each mtDNA clade as described above for the entire species range (e.g. Magalhaes et al. 2014; Gotelli and Stanton-Geddes 2015). Clade-specific ENMs were built using both our own records of georeferenced individuals with available genetic information and occurrence data from other sources (see above) falling within the geographic area delimited for each clade. Some records ($n=41$) were excluded from the analyses because they fell in areas close to contact zones and could not be unequivocally assigned to a particular clade. We calculated actual niche overlap between each pair of clades using two alternative statistics: I (Warren et al. 2008) and D (Schoener 1968). Then, we performed niche identity tests to compare the overlap of a clade pair's actual niches to a distribution of niche overlaps obtained from pairs of pseudoniches ($n=100$ pseudoreplicates) constructed based on randomly reshuffled occurrence points of the two clades (Warren et al. 2008). Thus, this test examines the null hypothesis that a given pair of clades is distributed in an identical environmental space. We also used background tests to analyze whether the ecological niches of a given pair of clades overlap more or less than would be expected from the differences in the environmental backgrounds of the regions where they occur (Warren et al. 2008). This test compares the observed niche overlap of a given pair of clades to a null distribution ($n=100$ random samplings) of overlap values generated by comparing the ecological niche model of one clade to an ecological niche model created from random points drawn from the geographic range of the other clade (Warren et al. 2008). This process is repeated for both clades in the comparison so that two null distributions are generated. The background area should include accessible areas for the organism, not just the observed niche or an area tightly delimited by the occurrence of the lineage/species (McCormack et al. 2010; Nakazato et al. 2010). Given that the delimitation of the background area can influence the results of the analyses, we considered different backgrounds defined by buffer zones of 1, 5, 10, and 25 km around the actual distribution of each lineage delimited by occurrence points. Background areas were obtained using ARCMAP 10.2.1.

Geographical and Environmental Drivers of Genetic Differentiation

For nuclear microsatellite markers, we estimated genetic differentiation between populations calculating pair-wise F_{ST}-values and testing their significance with Fisher's exact tests after 10,000 permutations as implemented in ARLEQUIN 3.1 (Excoffier et al. 2005). Due to the frequent presence of null alleles in grasshoppers (e.g. Chapuis et al. 2008), we used the program FREENA to estimate null allele frequencies and calculate pair-wise F_{ST} values corrected for null alleles using the so-called ENA method (Chapuis and Estoup 2007). For mtDNA gene fragments, we calculated pair-wise Φ_{ST} values (an analogue of F_{ST} for haplotype sequence similarity) (Kimura 1980) of genetic differentiation using ARLEQUIN 3.1 (Excoffier et al. 2005). Only populations with five or more analysed individuals for both nuclear microsatellite markers and mtDNA sequences were used in these and subsequent analyses.

We applied circuit theory to model gene flow across spatially heterogeneous landscapes and determine the impact of isolation-by-distance (IBD), isolation-by-resistance (IBR) and isolation-by-environment (IBE) on observed patterns of genetic differentiation among the studied populations (McRae 2006; McRae and Beier 2007). We used CIRCUITSCAPE 3.5.8 to calculate resistance distance matrices between all pairs of sampling sites considering an eight-neighbour cell connection scheme (McRae 2006). We used habitat suitability data obtained from ecological niche models (ENM) to generate three IBR scenarios based on (1) current habitat suitability (IBR-Current); (2) LGM habitat suitability based on the CCSM model (IBR-LGM_{CCSM}); and (3) LGM habitat suitability based on the MIROC model (IBR-LGM_{MIROC}). Cell size for all raster layers was 30-arcsec (c. 1-km) and habitat suitability surfaces were used as conductance grids. To test the effect of IBD we generated a matrix of resistances in CIRCUITSCAPE considering an entirely "flat" landscape, i.e. based on a raster layer in which all cells have equal resistance (resistance = 1). This matrix of flat resistance distances is expected to yield similar results than the matrix of Euclidean geographical distances, but the former has been suggested to be more appropriate for comparison with models of IBR generated with CIRCUITSCAPE (Noguerales et al. 2016).

IBE was analysed estimating environmental dissimilarity between each pair of populations using the 19 bioclimatic layers from the WorldClim dataset (Hijmans et al. 2005). First, we characterized the environmental space of the studied species extracting bioclimatic data from all available occurrence points at two spatial scales (circular areas of 1000 and 10,000 m^2 around occurrence points) in order to test the potential impact of spatial scale on the obtained inferences. We obtained qualitatively identical results when

environmental dissimilarity was estimated at either spatial scale, and only analyses for the smallest scale (1000 m^2) are presented. Then, we performed a principal component analysis (PCA) on the 19 bioclimatic variables with a varimax rotation in SPSS 22.0 (IBM Coorp., Armonk, NY, USA). For each studied population, we extracted the PC scores of the first four PCs, which had eigenvalues higher than one and explained >90% of the variance at the two studied spatial scales (Supplementary Table S3). Finally, we calculated environmental dissimilarity between each pair of populations using Euclidean distances for the obtained PC scores with the 'dist' function in R 3.2.2 (R Core Team 2016).

IBD, IBR and IBE matrices were tested against matrices of genetic differentiation for nuclear microsatellite markers (F_{ST}) and mtDNA sequences (Φ_{ST}) using multiple matrix regressions with randomization (MMRR) (Wang 2013). We used the "MMRR" function script implemented in R 3.2.2 (Wang 2013). We used a backward procedure to select final models, removing nonsignificant variables from an initial full model including all explanatory predictors. We tested the significance of the remaining variables again until no additional term reached significance. We performed these analyses on the full dataset and considering each clade separately.

Results

Microsatellite Data

All nuclear microsatellite markers were highly polymorphic and observed heterozygosity at each locus ranged from 0.44 to 0.85, with 15–47 alleles per locus. After applying sequential Bonferroni corrections to compensate for multiple statistical tests, any microsatellite locus consistently deviated from HWE in all the studied populations. We did not find any evidence of genotypic linkage disequilibrium at any pair of microsatellite loci in any population (exact tests; all *P*-values >0.05).

Mitochondrial DNA Genetic Diversity and Tests of Selection

Considering all sequences, we found 139 and 48 unique haplotypes for COI ($n = 321$ sequences; Table S1) and 16S ($n = 252$ sequences; Table S1) gene fragments, respectively. All genetic diversity estimates revealed that populations from Morocco are more variable than their European counterparts (Table 1). Significantly negative Fu's F_S, indicative of population expansion or purifying selection, were found in the two loci for the entire dataset or when each clade was analyzed separately (Table 1). Tajima's *D* for COI were significantly negative for all sites and synonymous and non-synonymous sites for Clade II (Table 1). No Tajima's test of selective neutrality was significant in COI for Clade III and only non-synonymous sites showed significantly negative Tajima's *D* values for the entire dataset or analyses focused on European populations (i.e. Clades I–II) (Table 1). No Tajima's test of selective neutrality was significant for the 16S gene (Table 1). McDonald–Kreitman (MK) tests comparing all possible pairs of haplogroups were not significant ($P > 0.2$ in all cases).

Table 1 Descriptive statistics and neutrality tests for mtDNA data (COI and 16S gene fragments)

	COI					16S				
	All	Clade I + II	Clade I	Clade II	Clade III	All	Clade I + II	Clade I	Clade II	Clade III
N individuals	242	192	93	99	50	242	192	93	99	50
H	98	65	33	32	33	47	28	10	18	19
S	101	63	35	31	44	41	23	8	15	16
Hd	0.941	0.908	0.947	0.700	0.968	0.897	0.843	0.666	0.703	0.881
π	0.038	0.021	0.008	0.003	0.015	0.011	0.007	0.002	0.002	0.005
Fu's F_S (*P*)	−26.21*	−20.22*	−16.15*	−32.04*	−14.92*	−19.39*	−10.02*	−3.28*	−15.38*	−11.34*
Tajima's *D*—all sites	0.46ns	0.01ns	−1.12ns	−2.07*	−0.61ns	−0.83ns	−0.62ns	−0.84ns	−1.78ns	−1.32ns
Tajima's *D*—S sites	0.71ns	0.29ns	−1.07ns	−1.90*	−0.45ns	–	–	–	–	–
Tajima's *D*—NS sites	−2.05*	−1.92*	−1.04ns	−1.91*	−1.70ns	–	–	–	–	–

N individuals = number of analyzed individuals, *H* = number of haplotypes, *S* = number of polymorphic sites, Hd = haplotype diversity, π = nucleotide diversity. Tajima's *D* was calculated for all sites and separately for synonymous (S) and non-synonymous (NS) sites (only for COI). Only individuals ($n = 242$) with available sequences for both gene fragments were considered in these analyses

ns not significant

*Significant

Genetic Structure

STRUCTURE analyses and the statistic ΔK indicated an "optimal" clustering for $K=2$ both considering and not considering prior population information (Supplementary Fig. S1a, b). For $K=2$, STRUCTURE grouped populations from central Iberia in one cluster and the rest of the populations in another cluster (Fig. 1e). However, Pr(X|K) steadily increased until $K=4$–8 (Supplementary Fig. S1a, b), roughly grouping populations from Ebro Valley and the Mediterranean coast from Catalonia and France, central Iberia, southeast Iberia, and Morocco (Fig. 1e). Neighbor-joining trees confirmed the patterns of genetic differentiation revealed by STRUCTURE analyses (Fig. 1d).

SAMOVA analyses on mtDNA gene fragments indicated that F_{CT} values greatly increased from $K=2$ to $K=3$ and reached a plateau for $K>4$ (Supplementary Fig. S1c). Considering $K=3$, the first group was composed of populations from the Ebro Valley and French and Spanish Mediterranean coast, the second group included populations from central and south Iberia, and the third group included all populations from Morocco (Supplementary Table 1). These groups corresponded with the three main mtDNA clades inferred in phylogenetic analyses (see next section and Fig. 1b). Haplotypes from the two European clades (Clades I and II) only co-occurred in three populations from southeast Iberia (El Bonillo, Sierra del Carché, and Berja; Fig. 1a).

AMOVA analyses indicated some differences between nuclear and mitochondrial genomes in their respective distribution of genetic variation (Table 2; Supplementary Table S1). When all populations or only European populations were considered, the percentage of total variation attributed to differences among groups and among populations within groups was high for mtDNA but low for nuclear microsatellite markers (Table 2). Conversely, variation among individuals within populations was low for mtDNA (<20%) and very high for nuclear microsatellite markers (>90%) (Table 2). This pattern differed for Moroccan populations, which did not show significant differences among groups and most variation was explained by differences among individuals within populations for both nuclear markers and mtDNA (Table 2).

Phylogenetic Analyses and Divergence Times

BEAST analyses indicate the presence of three very well-supported clades congruent with the population clustering obtained with SAMOVA analyses (Fig. 1b). Our phylogenetic analyses indicate that the split of European and Moroccan populations took place ca. 3.4 million years ago, in the transition between the Late Pliocene and the Early Pleistocene (Fig. 1b). However, confidence intervals are large (95% highest posterior density, HPD: 2.35–4.59 Ma) and the split may also correspond well with opening of the Gibraltar strait at the end of the Messinian (about 5.3 Ma) (Fig. 1b). The split between the two main European clades probably took place during the Middle-Early Pleistocene (1.37 Ma; HPD: 0.87–1.92). Most recently diverged and well supported haplogroups within each of the three main clades appeared during the Late Pleistocene and most of these haplogroups are represented in multiple populations geographically dispersed within the distribution range of each clade (Fig. 1a). Historical demographic reconstructions for each clade are shown in Fig. 2. European populations have experienced a recent demographic expansion (Fig. 2), which was particularly abrupt in Clade II (Fig. 2). In contrast, populations from North Africa (Clade III) showed a demographic expansion between ~200,000–50,000 years BP followed by a recent decline (Fig. 2).

Environmental Niche Modelling

All ENM had high AUC values (range 0.885–0.928; Supplementary Table S4) and predicted moderately well current species distribution (Fig. 2; Supplementary Fig. S2). ENMs built for a given mtDNA clade often predicted as suitable extensive areas occupied by another clade or over-predicted the current known distribution of the species (Fig. 2). Over-prediction of distribution ranges involved particularly models for Clade II and III. Model for Clade II predicted as suitable some areas where the species is not present (northwest Spain, Mediterranean islands) or where only another clade occurs (Clade I: Ebro Valley, Mediterranean coast; Clade III: North Africa). Model for Clade III over-predicted areas where the species is either not present (Mediterranean islands) or where only populations of the Iberian Clade II occur (Fig. 2). The projections of clade-specific ENM into the LGM suggest that the distribution ranges of the two European clades contracted during the LGM whereas populations from North Africa (Clade III) expanded during the LGM (Fig. 2; Supplementary Fig. S2). Remarkably, populations from Clade I are expected to have persisted during the LGM only in small suitable areas located in northeast Spain (Fig. 2; Supplementary Fig. S2).

Niche Divergence

All identity tests were significant for the two estimates of niche overlap (I and D), indicating that the three clades are not distributed in an identical environmental space (all Ps <0.01). However, most background tests were not significant or yielded mixed results depending on the direction of the test (Table 3). The environmental niches of Clades I and II were more similar than expected given the environmental background surrounding known occurrences of

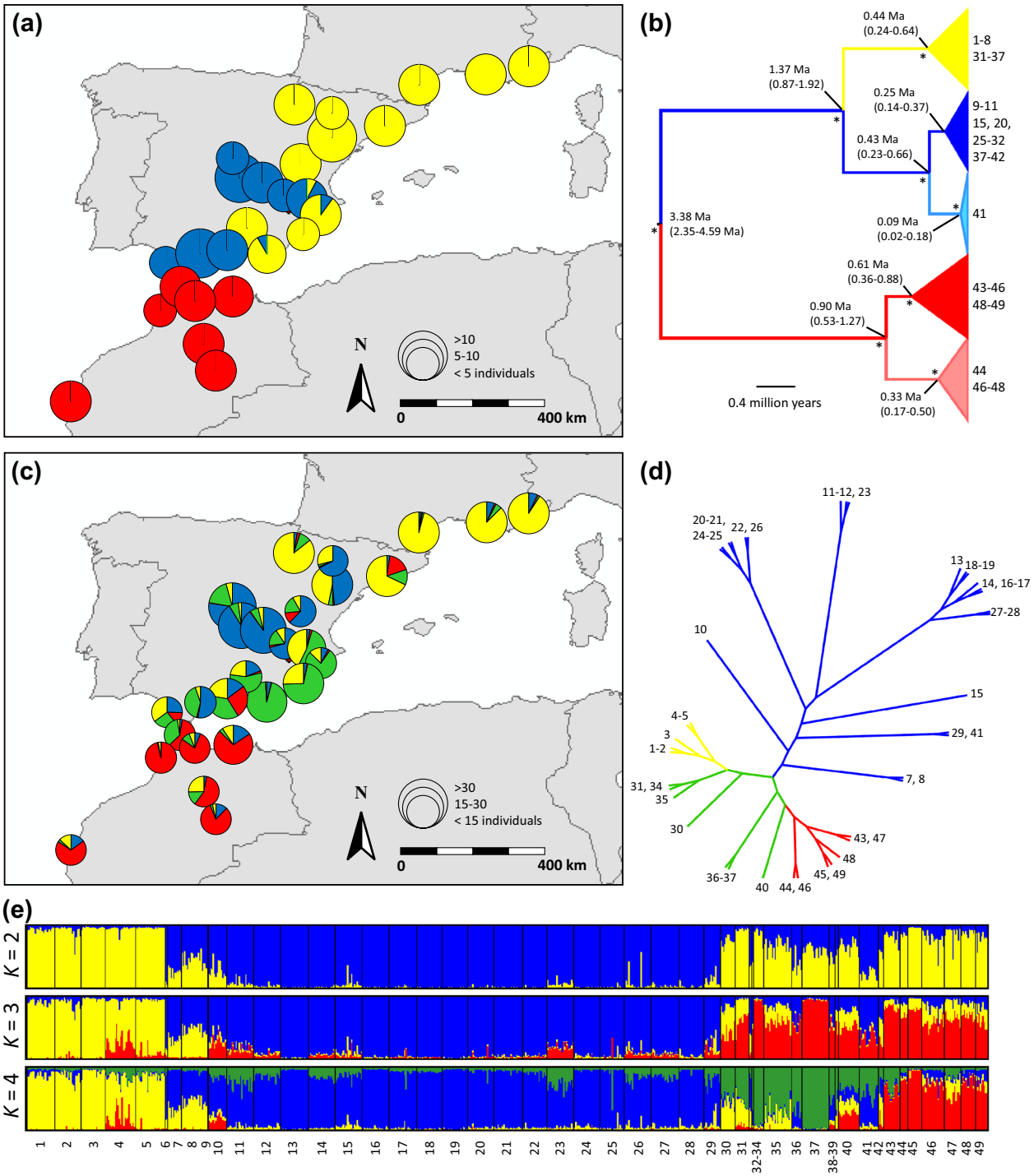

Fig. 1 **a** Geographic distribution of the three main mtDNA haplogroups for the esparto grasshopper (*Ramburiella hispanica*). **b** Maximum clade credibility tree for COI and 16S gene fragments with estimated ages (mean and lower and upper 95% highest posterior density, HPD) for each node. *Asterisks* indicate clades with Bayesian posterior probability >90%. *Tip labels* indicate the code of those populations represented in each collapsed clade. **c** Genetic assignment of populations based on nuclear microsatellite markers and the Bayesian method implemented in the program Structure. Admixture proportions generated by Structure for $K=4$ were represented using pie charts, with *each colour* indicating a different genotypic cluster. **d** Unrooted neighbour-joining tree based on Cavalli-Sforza's distance for nuclear microsatellite markers. Only populations with five or more genotyped individuals have been included in this analysis. *Tip labels* indicate population codes. **e** Genetic assignment of individuals based on Structure analyses for $K=2$–4. Each individual is represented by a *vertical bar*, which is partitioned into K coloured segments showing the individual's probability of belonging to the cluster with that colour. *Thin vertical black lines* separate individuals from different sampling localities. For the sake of presentation, some nearby populations were grouped in the same pie chart in **a** and **c**. Population codes are presented in Supplementary Table S1. (Color figure online)

Table 2 Results of AMOVAS for genetic differentiation at ten nuclear microsatellite markers and sequences of two mtDNA gene fragments (COI and 16S)

	Nuclear			Mitochondrial		
	Sum of squares	Variance components	Percentage of variation	Sum of squares	Variance components	Percentage of variation
All populations						
Among groups	41.07	0.04	2.48***	2091.11	11.01	70.18***
Among populations within groups	69.32	0.05	3.15***	465.33	2.59	16.52***
Among individuals within populations	1334.32	1.55	94.37***	394.15	2.08	13.30***
Europe						
Among groups	34.63	0.045	2.71***	630.83	4.42	49.94***
Among populations within groups	59.40	0.048	2.89***	416.76	2.78	31.43***
Among individuals within populations	1153.88	1.57	94.40***	245.66	1.65	18.63***
Morocco						
Among groups	5.29	−0.01	-0.96^{NS}	44.05	0.48	8.23^{NS}
Among populations within groups	9.91	0.08	5.65***	48.57	1.61	27.76***
Among individuals within populations	180.44	1.43	95.21***	148.49	3.71	64.01***

NS not significant

*$P<0.1$; **$P<0.05$; ***$P<0.01$

Clade I, but the reciprocal comparison was not significant (Table 3). The niches of Clades II and III were less similar than expected given the environmental background of Clade II, but the reciprocal comparison was not significant (Table 3). These counterintuitive results may be driven by differences in the heterogeneity of the environmental background for the two clades (Nakazato et al. 2010; McCormack et al. 2010).

Geographical and Environmental Drivers of Genetic Differentiation

Genetic differentiation estimated at nuclear and mtDNA markers (see Supplementary Table S5) were significantly correlated for the datasets including all populations (Mantel tests; $r=0.37$, $P<0.001$) and populations from Clades I–II ($r=0.32$, $P<0.001$) and Clade I ($r=0.28$, $P=0.017$), but not for the datasets only including populations from Clade II ($r=-0.23$, $P=0.964$) and Clade III ($r=0.01$, $P=0.467$). Univariate matrix regressions with randomization considering the entire dataset, populations from Clade I, or populations from Clades I–II (i.e. European populations) showed that genetic differentiation at nuclear markers was significantly associated with IBR distances based on a completely flat landscape (i.e. IBD) and on current and LGM environmental suitability models, but not with environmental dissimilarity (i.e. IBE) (Table 4; Fig. 3; Supplementary Fig. S3). Analyses focussed on populations from Clade II and Clade III showed no significant association between genetic differentiation at nuclear markers and any distance matrix (Table 4; Fig. 3; Supplementary Fig. S3). Genetic differentiation estimated at mtDNA markers was significantly associated with IBD, IBE and all IBR distance matrices in univariate analyses performed for datasets including all populations, and populations from Clades I–II and Clade II (Table 4; Fig. 3; Supplementary Fig. S3). Analyses focussed on populations from Clade I and Clade III showed no significant association between genetic differentiation at mitochondrial markers and any distance matrix (Table 4). Final models for nuclear microsatellite markers only included IBR based on LGM_{MIROC} habitat suitability, either including all populations ($\beta=0.140$, $t=8.11$, $P=0.002$), populations from Clades I–II ($\beta=0.240$, $t=8.90$, $P=0.001$) or populations from Clade I ($\beta=0.580$, $t=4.29$, $P=0.007$) and no other variable remained significant (all P-values >0.1). Final multivariate models considering all populations or populations from Clades I–II showed that genetic differentiation at mtDNA markers was jointly explained by a significant effect of both geographical distance (IBD, all populations: $\beta=0.453$, $t=10.53$, $P=0.001$; Clades I–II: $\beta=0.240$, $t=3.56$, $P=0.010$) and environmental dissimilarity (IBE, all populations: $\beta=0.184$, $t=4.22$, $P=0.004$; Clades I–II: $\beta=0.282$, $t=4.49$, $P=0.003$) whereas only environmental dissimilarity was retained into the final model for Clade II ($\beta=0.727$, $t=8.35$, $P=0.003$). Overall, these results pointed to geographic distance and landscape composition as the major driver of nuclear genetic differentiation, whereas both environment and geography seemed to play a similar role in shaping genetic differentiation at mtDNA.

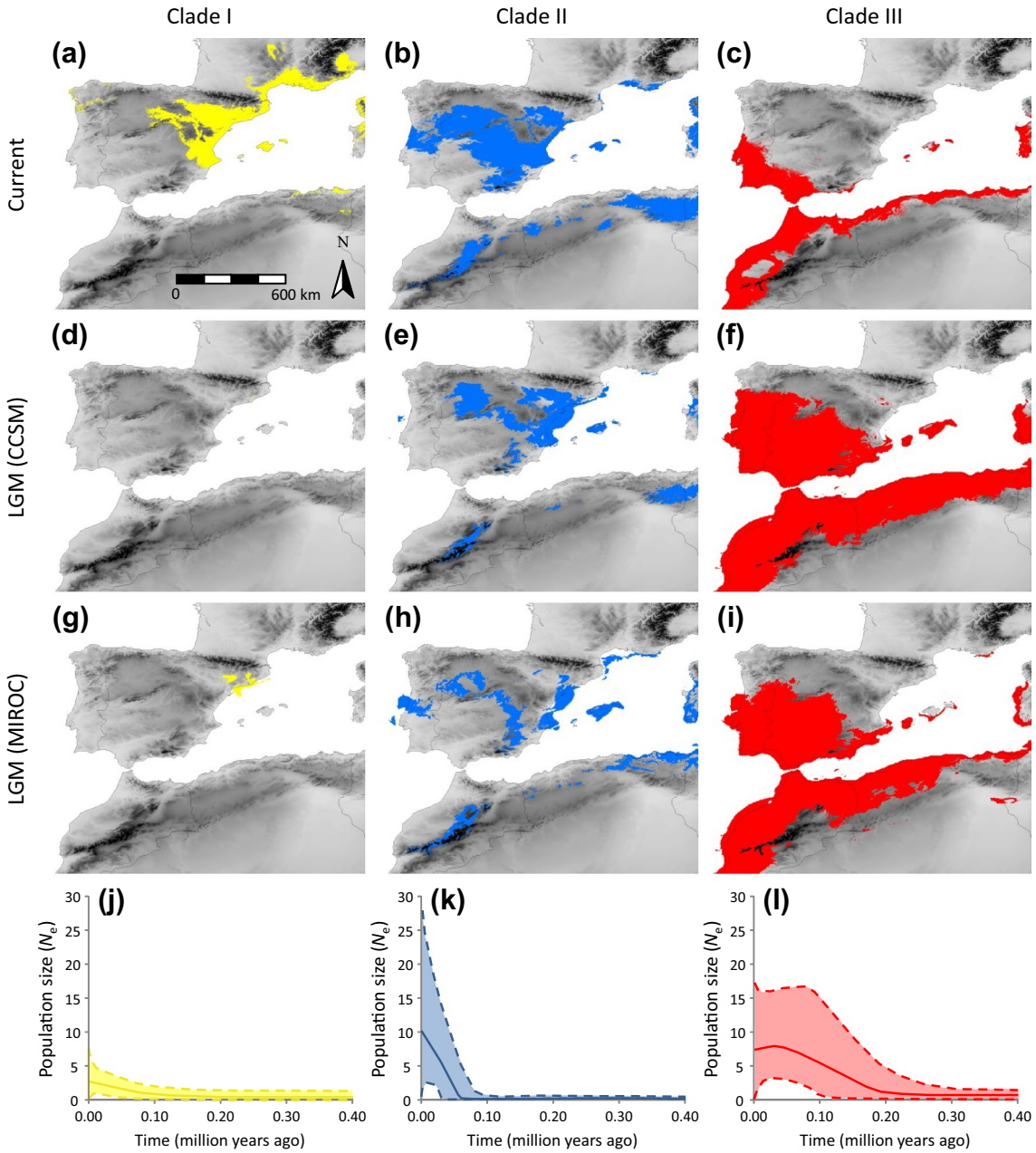

Fig. 2 **a–i** Environmental niche modelling and **j–l** past demographic history for each mtDNA clade of the esparto grasshopper (*Ramburiella hispanica*). **a–i** Maps show climatically suitable areas during the present and projections to the last glacial maximum based on two global circulation models (CCSM and MIROC). Suitable areas inferred for each clade (*yellow, blue*, and *red*) were identified based on grid cells with environmental suitability scores above the maximum training sensitivity plus specificity (MTSS) logistic threshold of MAXENT. *Grey background* represents elevation, with darker areas corresponding to higher altitude. **j–l** Bayesian Skyline Plots (BSP) representing historical demographic trends for each clade during the last 0.4 mya. *Solid line* median estimated population size, *dashed lines* upper and lower bounds of 95% highest posterior density. The *y*-axis represents the effective population size (N_e) considering the generation time of the species (=1 year). (Color figure online)

Discussion

The esparto grasshopper presents a remarkable phylogeographic structure across its distribution range in Europe and North Africa (Fig. 1). Our analyses performed at contrasting spatiotemporal scales (i.e. within and across lineages) indicate that different factors are responsible for genetic divergence, including environmental dissimilarity and geographical separation of populations probably driven by past geological events and Pleistocene climate

Table 3 Results of background tests considering two indexes of niche overlap (Warren's *I*: left; Schoener's *D*: right) and background areas obtained using different distance buffers around occurrence points

Clade for observed distribution	Clade for background	*I*	*D*	1-km buffer	5-km buffer	10-km buffer	25-km buffer
Clade I	Clade II	0.827	0.544	NS/NS	NS/NS	NS/NS	NS/NS
Clade I	Clade III	0.533	0.307	NS/NS	NS/NS	NS/NS	NS/NS
Clade II	Clade I	0.827	0.544	More***/more***	More***/more***	More**/more**	More***/more***
Clade II	Clade III	0.664	0.376	NS/NS	NS/NS	NS/NS	NS/NS
Clade III	Clade I	0.533	0.307	More*/NS	NS/NS	NS/NS	NS/NS
Clade III	Clade II	0.664	0.376	Less*/less*	Less*/less**	Less*/less**	Less*/less**

Table indicates whether actual values of niche overlap of two clades are more or less similar than expected based on the differences in the environmental background in which they occur

NS not significant

*$P < 0.1$; **$P < 0.05$; ***$P < 0.01$

Table 4 Univariate matrix regressions with randomization for genetic differentiation at ten nuclear microsatellite markers (F_{ST}-values corrected for null alleles) and sequences of two mtDNA gene fragments (Φ_{ST}) in relation with isolation by distance (IBD, based on resistance distances calculated for a completely flat landscape, i.e. with equal resistance to all pixel values), isolation-by-resistance (IBR) based on habitat suitability in the present (IBR-Current) and during the last glacial maximum obtained considering two global circulation models (IBR-LGM_{CCSM} and IBR-LGM_{MIROC}), and isolation by environment (IBE) estimated on the basis of bioclimatic dissimilarity between each pair of populations

	Nuclear				Mitochondrial			
	R^2	β	t	P	R^2	β	t	P
All populations (clades I–II–III)								
IBD	0.137	0.357	8.04	0.002	0.295	0.524	13.02	0.001
IBR-current	0.096	0.291	6.55	0.004	0.273	0.492	12.33	0.001
IBR-LGM_{CCSM}	0.130	0.346	7.76	0.002	0.295	0.523	13.00	0.001
IBR-LGM_{MIROC}	0.140	0.360	8.11	0.002	0.294	0.523	12.98	0.001
IBE	0.018	0.132	2.73	0.224	0.139	0.366	8.09	0.001
European populations (clades I–II)								
IBD	0.222	0.499	8.47	0.002	0.111	0.360	5.61	0.001
IBR-current	0.189	0.430	7.64	0.008	0.099	0.318	5.27	0.001
IBR-LGM_{CCSM}	0.212	0.497	8.23	0.004	0.057	0.234	3.91	0.006
IBR-LGM_{MIROC}	0.240	0.521	8.90	0.001	0.108	0.036	0.63	0.001
IBE	0.020	0.139	2.26	0.250	0.136	0.371	6.28	0.001
Clade I								
IBD	0.248	0.566	4.18	0.006	0.115	0.412	2.62	0.093
IBR-current	0.241	0.494	4.10	0.012	0.138	0.401	2.92	0.105
IBR-LGM_{CCSM}	0.237	0.564	4.06	0.017	0.122	0.434	2.72	0.088
IBR-LGM_{MIROC}	0.257	0.580	4.29	0.007	0.107	0.401	2.52	0.120
IBE	0.002	0.047	0.35	0.834	0.045	0.224	1.58	0.400
Clade II								
IBD	0.045	0.341	1.73	0.349	0.376	1.028	6.21	0.005
IBR-current	0.009	0.104	0.76	0.720	0.418	0.744	6.78	0.014
IBR-LGM_{CCSM}	0.060	0.414	2.02	0.280	0.338	1.023	5.72	0.010
IBR-LGM_{MIROC}	0.080	0.506	2.35	0.193	0.286	0.998	5.07	0.009
IBE	0.001	−0.014	−0.11	0.966	0.521	0.727	8.35	0.003
Clade III								
IBD	0.001	−0.018	−0.05	0.961	0.065	0.418	0.95	0.398
IBR-current	0.001	−0.008	−0.02	0.969	0.039	0.288	0.73	0.522
IBR-LGM_{CCSM}	0.005	−0.115	−0.25	0.838	0.041	0.400	0.75	0.528
IBR-LGM_{MIROC}	0.001	−0.027	−0.07	0.961	0.046	0.368	0.80	0.491
IBE	0.001	0.047	0.16	0.880	0.209	0.579	1.85	0.083

β is the regression coefficient and R^2 is the coefficient of determination

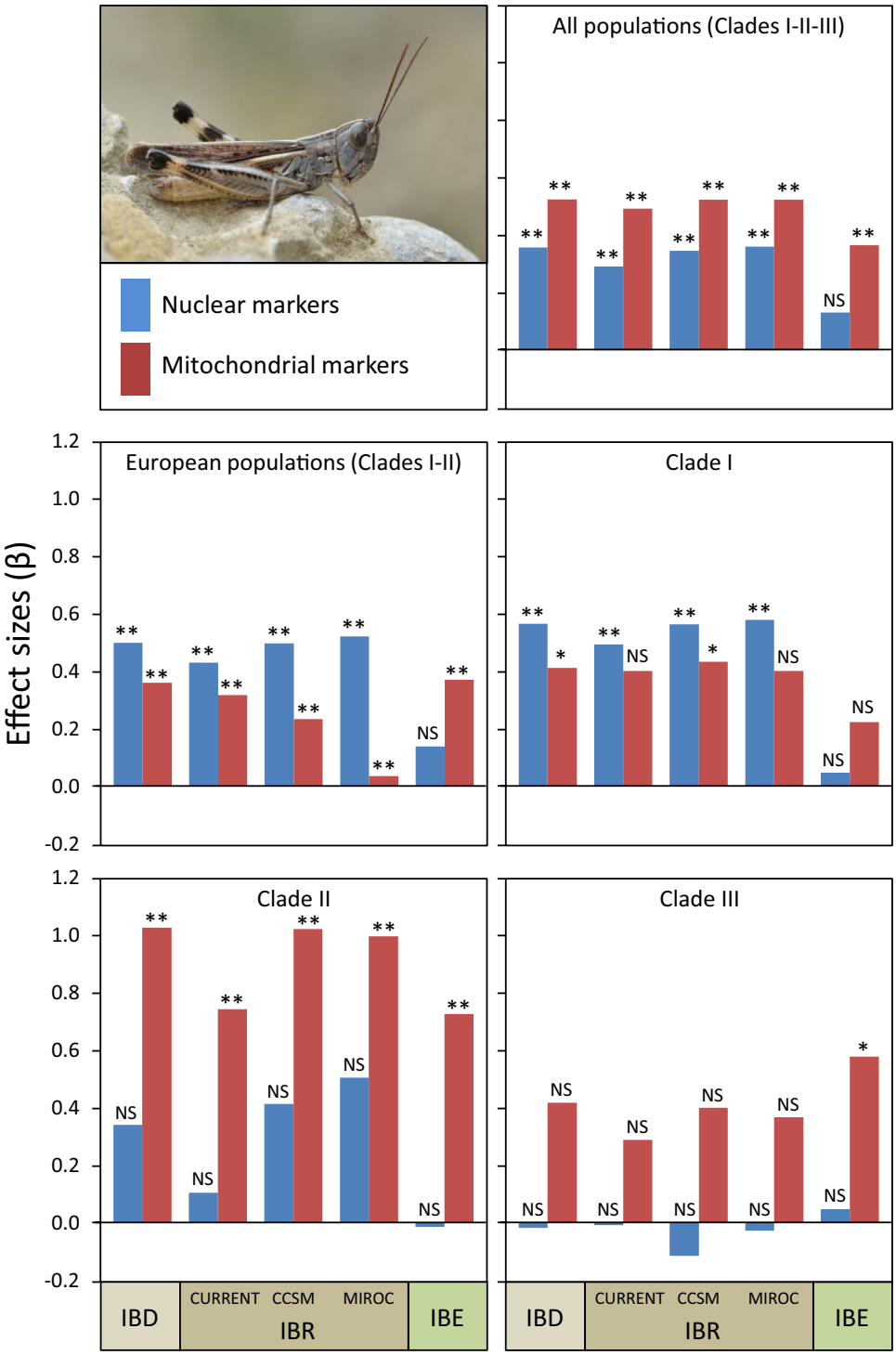

Fig. 3 Effect sizes (β) of isolation-by-distance (IBD), isolation-by-resistance (IBR) and isolation-by-environment (IBE) for nuclear (*blue*) and mitochondrial (*red*) markers obtained from univariate matrix regressions with randomization. Analyses were performed considering different subsets of populations. Isolation-by-resistance was calculated on the basis of climatic suitability maps obtained using MAXENT for the present (CURRENT) and the last glacial maximum according to two global circulation models (CCSM and MIROC). $^{**}P < 0.05$, $^{*}P < 0.1$, *NS* non-significant. *Top left* a male esparto grasshopper (*Ramburiella hispanica*) (picture by Gilles San Martin). (Color figure online)

oscillations. Despite the phylogeographic structure at both nuclear and mitochondrial markers was largely concordant, the processes underlying genetic divergence in both genomes seem to be remarkably different: geography was the major driver of nuclear genetic differentiation, whereas both environment and geography played a similar role in shaping genetic differentiation at mtDNA (Fig. 3).

Phylogeographic Genetic Structure

Our analyses revealed the presence of three major mitochondrial clades with a spatial distribution that roughly matched the global genetic structure found at nuclear microsatellite markers (Fig. 1). North African (Clade III) and European mtDNA clades (Clades I–II) were

estimated to diverge ~3.4 mya (HPD: 2.35–4.59 mya), which is consistent with post-Messinian divergence after the re-opening of the Gibraltar strait 5.33 mya (Sanmartín 2003; Pinho et al. 2006; Faille et al. 2014). The separation of the two European clades was inferred to take place much more recently, probably during the Middle-Early Pleistocene (~1.4 mya; HPD: 0.87–1.92 mya). The fact that no major geographic barrier separates populations from the two European linages together with the presence of a contact zone in southeast Iberia (Fig. 1a), suggest that their origin is probably associated with allopatric divergence in geographically isolated climate refugia (Hewitt 2004). Although the divergence time of the two European lineages predates the LGM, our global ecological niche model for the present and projections into the LGM suggests that European populations of esparto grasshopper probably were much more fragmented in glacial than in interglacial periods (Fig. 2). The esparto grasshopper is a thermophilic species distributed in habitats dominated by Mediterranean shrubby plant communities (Llucià-Pomares 2002; Ortego et al. 2015a). Thus, its particular habitat and environmental requirements together with the high sensitivity of the species to habitat fragmentation revealed by previous landscape-level genetic studies (Ortego et al. 2015a, b) are likely to have contributed to lineage split associated with population subdivision during Pleistocene glacial cycles.

Clade-specific niche models showed considerable differences among linages in their respective responses to climate change since the LGM (Gotelli and Stanton-Geddes 2015). The two European clades experienced strong range contractions during the LGM followed by a considerable expansion in the present, whereas much higher stability and a northward range shift after the LGM was inferred for the North African lineage. In particular, European mtDNA Clade-I seems to have experienced a dramatic range-size reduction during the LGM, when its populations may have persisted in small pockets of suitable habitat in northeast Iberia (Fig. 2). Our linage distribution models are in general agreement with demographic trends inferred by Bayesian Skyline analyses, which suggest recent demographic expansions in the two European clades, much smaller effective population sizes in European Clade I than in the two other lineages, and higher effective population sizes in the North African lineage during the last 100K years followed by a more recent shallow demographic decline (Fig. 2). Note, however, that the link between lineage range shifts since the LGM and the inferred demographic trends must be interpreted with extreme caution given the considerable uncertainty around estimates of effective population sizes (N_e) (i.e. large 95% highest posterior densities) and the timing of their changes over time (Fig. 2).

Isolation by Environment and Geography

Analyses of niche divergence showed that although the main lineages of esparto grasshopper are not distributed in an identical environmental space, their niches are not significantly different when considering the environmental backgrounds of the regions where they occur (Warren et al. 2008; Nakazato et al. 2010). This is consistent with the hypothesis of allopatric divergence and indicates that the most important driver of the species phylogeographic structure probably was population fragmentation resulted from Pleistocene climate changes and the opening of the Gibraltar strait at the end of the Messinian (Warren et al. 2008). Although we found no evidence for environment as an important driver of genetic divergence at the onset of lineage split, our analyses of genetic differentiation considering the spatial distribution of populations and different aspects of landscape composition suggest that environmental dissimilarity may be playing an important role in maintaining lineage boundaries and promoting genetic differentiation at finer spatiotemporal scales (Graham et al. 2004; Thorpe et al. 2008). In particular, we found that genetic differentiation at mitochondrial loci was largely explained by environmental dissimilarity, whereas geographical distance was the only predictor of genetic differentiation at nuclear markers (Fig. 3). A previous study did not find evidence for male-biased dispersal in this species (Ortego et al. 2015a), suggesting that the higher degree of genetic admixture at nuclear than at mtDNA markers and the overrepresentation of a single mtDNA linage in admixed populations located in contact zones may have resulted from environmental based-selection as inferred from our IBE analyses (Fig. 1a).

Different factors could explain the discrepancies on the inferred drivers of genetic differentiation for the two genomes. Experimental studies have revealed the important role of mitochondrial genes on key physiological traits with major consequences on fitness (e.g. Pichaud et al. 2012; Novicic et al. 2015; Latorre-Pellicer et al. 2016), an aspect that has been considered to be responsible of observed correlations between environment and mitochondrial variation in natural populations of different vertebrate (Cheviron and Brumfield 2009; Ribeiro et al. 2011; Pavlova et al. 2013; Morales et al. 2015) and invertebrate organisms (Sun et al. 2015). This, together with the very low recombination rates of the mitochondrial genome, is expected to increase the chance of detecting signals of environment-based selection that may be hard to capture with a random subset of nuclear markers (Thorpe et al. 2008; Soria-Carrasco et al. 2014; Ferrer et al. 2016). Most employed microsatellite markers are very unlikely to fall within functional genomic regions involved in local adaptation and, in absence of very strong environmental-based selection with global effects across the entire genome, they are expected to primarily

reflect background levels of neutral genetic differentiation shaped by the spatial configuration of corridors and barriers to dispersal (Shafer and Wolf 2013; Wang and Bradburd 2014). In contrast with previous studies finding sharp eco-geographical transitions of mitochondrial variants, we did not find strong genetic signatures of selection on our analysed mtDNA genes (e.g. Ribeiro et al. 2011; Pavlova et al. 2013). This suggests that the employed gene fragments are not themselves under selection and their association with climate may be a by-product of selection on other mitochondrial genes through genetic hitchhiking (Ballard and Kreitman 1994; Meiklejohn et al. 2007). Alternatively, adaptation along climate gradients (i.e. not mediated by abrupt environmental clines) suggested by our isolation-by-environment analyses may not leave strong genetic signatures traceable by traditional tests for selection. The fact that our analyses suggest a similar contribution of geography and environment on explaining genetic differentiation at mitochondrial loci indicate that genetic variation at mtDNA is also capturing neutral gene flow, which may weaken the genetic signal of selection exerted by environment (Fig. 3). The only exception is offered by analyses performed within Clade II, in which environmental dissimilarity overrides the effects of geographic distance on explaining genetic differentiation. This clade showed the strongest signatures of purifying selection or genetic hitchhiking, particularly for gene COI (Table 1), which may be the result of environmental-driven selective sweeps (Pavlova et al. 2013; Morales et al. 2015).

Conclusions

Our study shows that although historical factors have shaped largely concordant range-wide patterns of mitonuclear genetic structure in the esparto grasshopper, different contemporary processes (neutral gene flow vs. environmental-based selection) seem to be governing the spatial distribution of genetic variation in both genomes. Future studies considering genome-wide nuclear data and entire mitogenomes (e.g. Morales et al. 2015; Soria-Carrasco et al. 2014), fine-scale sampling along transects across contact zones (e.g. Cheviron and Brumfield 2009; Singhal and Moritz 2012), and laboratory tests of metabolic differences among functional genetic variants (e.g. Fontanillas et al. 2005; Pichaud et al. 2012) may help to determine the proximate mechanisms underlying observed mitonuclear discordances in the balance between local selection and neutral gene flow (Pavlova et al. 2013; Morales et al. 2015).

Acknowledgements We wish to thank to Conchi Cáliz for her valuable help in sample collection and genotyping. Two anonymous referees provided valuable comments on an earlier draft of this manuscript. JO was supported by "Ramón y Cajal" (RYC-2013-12501) and "Severo Ochoa" (SEV-2012-0262) research fellowships. VN was supported by a FPI pre-doctoral fellowship (BES-2012-053741). This work received financial support from Ministerio de Economía y Competitividad (Grants CGL2011-25053 and CGL2014-54671-P), Junta de Comunidades de Castilla-La Mancha and European Social Fund (Grants PCI08-0130-3954, POII10-0197-0167 and PEII-2014-023-P), and European Regional Development Fund (Grant UNCM08-1E-018).

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Aguirre, M. P., Noguerales, V., Cordero, P. J., & Ortego, J. (2014). Isolation and characterization of polymorphic microsatellites in the specialist grasshopper *Ramburiella hispanica* (Orthoptera: Acrididae). *Conservation Genetics Resources, 6*(3), 723–724. doi:10.1007/s12686-014-0198-4.

Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research, 25*(22), 4692–4693. doi:10.1093/nar/25.22.4692.

Avise, J. (1994). *Molecular markers, natural history and evolution*. New York: Chapman and Hall.

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M. A., & Alekseyenko, A. V. (2012). Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution, 29*(9), 2157–2167. doi:10.1093/molbev/mss084.

Ballard, J. W. O., & Kreitman, M. (1994). Unraveling selection in the mitochondrial genome of *Drosophila*. *Genetics, 138*(3), 757–772.

Ballard, J. W. O., & Whitlock, M. C. (2004). The incomplete natural history of mitochondria. *Molecular Ecology, 13*(4), 729–744. doi:10.1046/j.1365-294X.2003.02063.x.

Bar-Yaacov, D., Hadjivasiliou, Z., Levin, L., Barshad, G., Zarivach, R., Bouskila, A., et al. (2015). Mitochondrial involvement in vertebrate speciation? The case of mito-nuclear genetic divergence in chameleons. *Genome Biology and Evolution, 7*(12), 3322–3336. doi:10.1093/gbe/evv226.

Blondel, J., & Aronson, J. (1999). *Biology and wildlife of the Mediterranean region*. Oxford: Oxford University Press.

Bradburd, G. S., Ralph, P. L., & Coop, G. M. (2013). Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution, 67*(11), 3258–3273. doi:10.1111/evo.12193.

Brito, P., & Edwards, S. V. (2009). Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica, 135*(3), 439–455. doi:10.1007/s10709-008-9293-3.

Cavalli-Sforza, L. L., & Edwards, A. W. F. (1967). Phylogenetic analysis models and estimation procedures. *American Journal of Human Genetics, 19*(3P1), 233–257.

Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution, 24*(3), 621–631.

Chapuis, M. P., Lecoq, M., Michalakis, Y., Loiseau, A., Sword, G. A., Piry, S., et al. (2008). Do outbreaks affect genetic population structure? A worldwide survey in *Locusta migratoria*, a

pest plagued by microsatellite null alleles. *Molecular Ecology, 17*(16), 3640–3653.

Cheviron, Z. A., & Brumfield, R. T. (2009). Migration-selection balance and local adaptation of mitochondrial haplotypes in rofous-collared sparrows (*Zonotrichia capensis*) along an elevational gradient. *Evolution, 63*(6), 1593–1605. doi:10.1111/j.1558-5646.2009.00644.x.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods, 9*(8), 772–772.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution, 29*(8), 1969–1973. doi:10.1093/molbev/mss075.

Dupanloup, I., Schneider, S., & Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology, 11*(12), 2571–2581.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources, 4*(2), 359–361. doi:10.1007/s12686-011-9548-7.

Edwards, S., & Bensch, S. (2009). Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. *Molecular Ecology, 18*(14), 2930–2933. doi:10.1111/j.1365-294X.2009.04270.x.

Elith, J., Phillips, S. J., Hastie, T., Dudik, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions, 17*(1), 43–57. doi:10.1111/j.1472-4642.2010.00725.x.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology, 14*(8), 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x.

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.

Faille, A., Andujar, C., Fadrique, F., & Ribera, I. (2014). Late Miocene origin of an Ibero-Maghrebian clade of ground beetles with multiple colonizations of the subterranean environment. *Journal of Biogeography, 41*(10), 1979–1990. doi:10.1111/jbi.12349.

Ferrer, E. S., Garcia-Navas, V., Bueno-Enciso, J., Barrientos, R., Serrano-Davies, E., Caliz-Campal, C., et al. (2016). The influence of landscape configuration and environment on population genetic structure in a sedentary passerine: Insights from loci located in different genomic regions. *Journal of Evolutionary Biology, 29*(1), 205–219. doi:10.1111/jeb.12776.

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology, 3*, 294–299.

Fontanillas, P., Depraz, A., Giorgi, M. S., & Perrin, N. (2005). Nonshivering thermogenesis capacity associated to mitochondrial DNA haplotypes and gender in the greater white-toothed shrew, *Crocidura russula. Molecular Ecology, 14*(2), 661–670. doi:10.1111/j.1365-294X.2004.02414.x.

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics, 147*(2), 915–925.

Galtier, N., Nabholz, B., Glemin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: A reappraisal. *Molecular Ecology, 18*(22), 4541–4550. doi:10.1111/j.1365-294X.2009.04380.x.

Gaspari, S., Scheinin, A., Holcer, D., Fortuna, C., Natali, C., Genov, T., et al. (2015). Drivers of population structure of the bottlenose dolphin (*Tursiops truncatus*) in the Eastern Mediterranean sea. *Evolutionary Biology, 42*(2), 177–190. doi:10.1007/s11692-015-9309-8.

Gotelli, N. J., & Stanton-Geddes, J. (2015). Climate change, genetic markers and species distribution modelling. *Journal of Biogeography, 42*(9), 1577–1585. doi:10.1111/jbi.12562.

Graham, C. H., Ron, S. R., Santos, J. C., Schneider, C. J., & Moritz, C. (2004). Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution, 58*(8), 1781–1793. doi:10.1554/03-274.

Guo, S. W., & Thompson, E. A. (1992). A monte-carlo method for combined segregation and linkage analysis. *American Journal of Human Genetics, 51*(5), 1111–1126.

Hasumi, H., & Emori, S. (2004). *K-1 coupled GCM (MIROC) description*. Center for Climate System Research, University of Tokyo, National Institute for Environmental Studies, Frontier Research Center for Global Change, Tokyo.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature, 405*(6789), 907–913.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 359*(1442), 183–195.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology, 25*(15), 1965–1978. doi:10.1002/joc.1276.

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources, 9*(5), 1322–1332. doi:10.1111/j.1755-0998.2009.02591.x.

Hutchison, D. W., & Templeton, A. R. (1999). Correlation of pairwise genetic and geographic distance measures: Inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution, 53*(6), 1898–1914.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics, 23*(14), 1801–1806. doi:10.1093/bioinformatics/btm233.

Kiehl, J. T., & Gent, P. R. (2004). The community climate system model, version 2. *Journal of Climate, 17*(19), 3666–3682.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. *Journal of Molecular Evolution, 16*(2), 111–120. doi:10.1007/bf01731581.

Krijgsman, W., Hilgen, F. J., Raffi, I., Sierro, F. J., & Wilson, D. S. (1999). Chronology, causes and progression of the Messinian salinity crisis. *Nature, 400*(6745), 652–655.

Kumar, A., Ghazi, M. G. U., Hussain, S. A., Bhatt, D., & Gupta, S. K. (2017). Mitochondrial and nuclear DNA based genetic assessment indicated distinct variation and low genetic exchange among the three subspecies of swamp deer (*Rucervus duvaucelii*). *Evolutionary Biology, 44*(1), 31–42. doi:10.1007/s11692-016-9387-2.

Langella, O. (1999). Populations 1.2.31 software. http://bioinformatics.org/populations/. Accessed 2 Oct 2016.

Latorre-Pellicer, A., Moreno-Loshuertos, R., Lechuga-Vieco, A. V., Sanchez-Cabo, F., Torroja, C., Acin-Perez, R., et al. (2016). Mitochondrial and nuclear DNA matching shapes metabolism and healthy ageing. *Nature, 535*(7613), 561–565. doi:10.1038/nature18618.

Lee, C. R., & Mitchell-Olds, T. (2011). Quantifying effects of environmental and geographical factors on patterns of genetic differentiation. *Molecular Ecology, 20*(22), 4631–4642. doi:10.1111/j.1365-294X.2011.05310.x.

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics, 25*(11), 1451–1452. doi:10.1093/bioinformatics/btp187.

Llucià-Pomares, D. (2002). Revision of the Orthoptera (Insecta) of Catalonia (Spain). *Monografias SEA, 7*, 1–226.

Magalhaes, I. L. F., Oliveira, U., Santos, F. R., Vidigal, T., Brescovit, A. D., & Santos, A. J. (2014). Strong spatial structure, Pliocene diversification and cryptic diversity in the Neotropical dry forest spider *Sicarius cariri. Molecular Ecology, 23*(21), 5323–5336. doi:10.1111/mec.12937.

Martin, M. D., & Mendelson, T. C. (2012). Signal divergence is correlated with genetic distance and not environmental differences in darters (Percidae: Etheostoma). *Evolutionary Biology, 39*(2), 231–241. doi:10.1007/s11692-012-9179-2.

Massatti, R., & Knowles, L. L. (2014). Microhabitat differences impact phylogeographic concordance of codistributed species: Genomic evidence in montane sedges (*Carex* L.) from the Rocky Mountains. *Evolution, 68*(10), 2833–2846. doi:10.1111/evo.12491.

McCormack, J. E., Zellmer, A. J., & Knowles, L. L. (2010). Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation?: Insights from tests with niche models. *Evolution, 64*(5), 1231–1244. doi:10.1111/j.1558-5646.2009.00900.x.

McDonald, J. H., & Kreitman, M. (1991). Adaptive protein evolution at the ADH locus in *Drosophila. Nature, 351*(6328), 652–654. doi:10.1038/351652a0.

McRae, B. H. (2006). Isolation by resistance. *Evolution, 60*(8), 1551–1561. doi:10.1111/j.0014-3820.2006.tb00500.x.

McRae, B. H., & Beier, P. (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America, 104*(50), 19885–19890. doi:10.1073/pnas.0706568104.

McRae, B. H., Dickson, B. G., Keitt, T. H., & Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology, 89*(10), 2712–2724. doi:10.1890/07-1861.1.

Meiklejohn, C. D., Montooth, K. L., & Rand, D. M. (2007). Positive and negative selection on the mitochondrial genome. *Trends in Genetics, 23*(6), 259–263. doi:10.1016/j.tig.2007.03.008.

Morales, H. E., Pavlova, A., Joseph, L., & Sunnucks, P. (2015). Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. *Molecular Ecology, 24*(11), 2820–2837. doi:10.1111/mec.13203.

Moritz, C. C., & Potter, S. (2013). The importance of an evolutionary perspective in conservation policy planning. *Molecular Ecology, 22*(24), 5969–5971. doi:10.1111/mec.12565.

Nakazato, T., Warren, D. L., & Moyle, L. C. (2010). Ecological and geographic models of species divergence in wild tomatoes. *American Journal of Botany, 97*(4), 680–693. doi:10.3732/ajb.0900216.

Noguerales, V., Cordero, P. J., & Ortego, J. (2016). Hierarchical genetic structure shaped by topography in a narrow-endemic montane grasshopper. *BMC Evolutionary Biology, 16*, 96. doi:10.1186/s12862-016-0663-7.

Nosil, P. (2012). *Ecological speciation*. New York: Oxford University Press.

Novicic, Z. K., Immonen, E., Jelic, M., Andelkovic, M., Stamenkovic-Radak, M., & Arnqvist, G. (2015). Within-population genetic effects of mtDNA on metabolic rate in *Drosophila subobscura. Journal of Evolutionary Biology, 28*(2), 338–346. doi:10.1111/jeb.12565.

Ortego, J., Aguirre, M. P., Noguerales, V., & Cordero, P. J. (2015a). Consequences of extensive habitat fragmentation in landscape-level patterns of genetic diversity and structure in the Mediterranean esparto grasshopper. *Evolutionary Applications, 8*(6), 621–632. doi:10.1111/eva.12273.

Ortego, J., Garcia-Navas, V., Noguerales, V., & Cordero, P. J. (2015b). Discordant patterns of genetic and phenotypic differentiation in five grasshopper species codistributed across a microreserve network. *Molecular Ecology, 24*(23), 5796–5812. doi:10.1111/mec.13426.

Palumbi, S. R., Martin, A., Romano, S. L., McMillian, W. O., Stice, L., & Grabowski, G. (1991). *The simple Fool's guide to PCR, version 2.0*. Honolulu: University of Hawaii

Papadopoulou, A., Anastasiou, I., & Vogler, A. P. (2010). Revisiting the insect mitochondrial molecular clock: The Mid-Aegean trench calibration. *Molecular Biology and Evolution, 27*(7), 1659–1672. doi:10.1093/molbev/msq051.

Pavlova, A., Amos, J. N., Joseph, L., Loynes, K., Austin, J. J., Keogh, J. S., et al. (2013). Perched at the mito-nuclear crossroads: Divergent mitochondrial lineages correlate with environment in the face of ongoing nuclear gene flow in an Australian bird. *Evolution, 67*(12), 3412–3428. doi:10.1111/evo.12107.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling, 190*(3–4), 231–259. doi:10.1016/j.ecolmodel.2005.03.026.

Phillips, S. J., & Dudik, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography, 31*(2), 161–175. doi:10.1111/j.0906-7590.2008.5203.x.

Pichaud, N., Ballard, J. W. O., Tanguay, R. M., & Blier, P. U. (2012). Naturally occurring mitochondrial DNA haplotypes exhibit metabolic differences: Insight into functional properties of mitochondria. *Evolution, 66*(10), 3189–3197. doi:10.1111/j.1558-5646.2012.01683.x.

Pinho, C., Ferrand, N., & Harris, D. J. (2006). Reexamination of the Iberian and North African *Podarcis* (Squamata: Lacertidae) phylogeny based on increased mitochondrial DNA sequencing. *Molecular Phylogenetics and Evolution, 38*(1), 266–273. doi:10.1016/j.ympev.2005.06.012.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945–959.

R Core Team (2016). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Ribeiro, A. M., Lloyd, P., & Bowie, R. C. K. (2011). A tight balance between natural selection and gene flow in a souhern African arid-zone endemic bird. *Evolution, 65*(12), 3499–3514. doi:10.1111/j.1558-5646.2011.01397.x.

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes, 4*(1), 137–138. doi:10.1046/j.1471-8286.2003.00566.x.

Rosetti, N., & Remis, M. I. (2017). Variability of minisatellite loci and mtDNA in individuals with and without B chromosomes from populations of the grasshopper *Dichroplus elongatus. Evolutionary Biology, 44*(2), 273–283. doi:10.1007/s11692-016-9406-3.

Sanmartín, I. (2003). Dispersal vs. vicariance in the Mediterranean: Historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). *Journal of Biogeography, 30*(12), 1883–1897.

Schoener, T. W. (1968). Anolis lizards of Bimini—Resource partitioning in a complex fauna. *Ecology, 49*(4), 704–726. doi:10.2307/1935534.

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by environment or distance: Which patterns of gene flow is most common? *Evolution, 68*(1), 1–15. doi:10.1111/evo.12258.

Shafer, A. B. A., & Wolf, J. B. W. (2013). Widespread evidence for incipient ecological speciation: A meta-analysis

of isolation-by-ecology. *Ecology Letters, 16*(7), 940–950. doi:10.1111/ele.12120.

Singhal, S., & Moritz, C. (2012). Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard. *Evolution, 66*(5), 1474–1489, doi:10.1111/j.1558-5646.2011.01539.x.

Slatkin, M. (1993). Isolation by distance in equilibrium and nonequilibrium populations. *Evolution, 47*(1), 264–279.

Soria-Carrasco, V., Gompert, Z., Comeault, A. A., Farkas, T. E., Parchman, T. L., Johnston, J. S., et al. (2014). Stick insect genomes reveal natural selection's role in parallel speciation. *Science, 344*(6185), 738–742. doi:10.1126/science.1252136.

Sun, J. T., Wang, M. M., Zhang, Y. K., Chapuis, M. P., Jiang, X. Y., Hu, G., et al. (2015). Evidence for high dispersal ability and mito-nuclear discordance in the small brown planthopper, *Laodelphax striatellus. Scientific Reports*. doi:10.1038/srep08045.

Tajima, F. (1989). Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics, 123*(3), 585–595.

Takezaki, N., & Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics, 144*(1), 389–399.

Thorpe, R. S., Surget-Groba, Y., & Johansson, H. (2008). The relative importance of ecology and geographic isolation for speciation in anoles. *Philosophical Transactions of the Royal Society B-Biological Sciences, 363*(1506), 3071–3081. doi:10.1098/rstb.2008.0077.

Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology, 21*(16), 3907–3930. doi:10.1111/j.1365-294X.2012.05664.x.

Wang, I. J. (2013). Examining the full effects of landscape heterogeneity on spatial genetic variation: A multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution, 67*(12), 3403–3411. doi:10.1111/evo.12134.

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology, 23*(23), 5649–5662. doi:10.1111/mec.12938.

Wang, I. J., Glor, R. E., & Losos, J. B. (2013). Quantifying the roles of ecology and geography in spatial genetic divergence. *Ecology Letters, 16*(2), 175–182. doi:10.1111/ele.12025.

Warren, D. L., Glor, R. E., & Turelli, M. (2008). Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution, 62*(11), 2868–2883. doi:10.1111/j.1558-5646.2008.00482.x.

Wright, S. (1943). Isolation by distance. *Genetics, 28*(2), 114–138.

Zink, R. M., & Barrowclough, G. F. (2008). Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology, 17*(9), 2107–2121. doi:10.1111/j.1365-294X.2008.03737.x.