Check for updates

Received: 12 November 2020 | Revised: 2 February 2021 | Accepted: 23 April 2021

DOI: 10.1111/ddi.13306

RESEARCH ARTICLE

Diversity and Distributions WILEY

Glacial connectivity and current population fragmentation in sky islands explain the contemporary distribution of genomic variation in two narrow-endemic montane grasshoppers from a biodiversity hotspot

Vanina Tonzo | Joaquín Ortego

Department of Integrative Ecology, Estación Biológica de Doñana (EBD-CSIC), Seville, Spain

Correspondence
Vanina Tonzo, Estación Biológica de Doñana, EBD-CSIC, Avda. Américo Vespucio 26, E-41092 Seville, Spain.
Email: vaninatonzo@gmail.com

Funding information
Spanish Ministry of Economy and Competitiveness, Grant/Award Number: CGL2014-54671; European Regional Development Fund (ERDF), Grant/Award Number: CGL2017-83433-P; Spanish Ministry of Economy and Competitiveness predoctoral fellowship, Grant/Award Number: BES-2015-73159

Editor: Zhonge Hou

Abstract

Aim: Cold-adapted biotas from mid-latitudes often show small population sizes, harbour low levels of local genetic diversity and are highly vulnerable to extinction due to ongoing climate warming and the progressive shrinking of montane and alpine ecosystems. In this study, we use a suite of analytical approaches to infer the demographic processes that have shaped contemporary patterns of genomic variation in *Omocestus bolivari* and *Omocestus femoralis*, two narrow-endemic and red-listed Iberian grasshoppers forming highly fragmented populations in the sky island archipelago of the Baetic System.

Location: South-eastern Iberia.

Methods: We quantified genomic variation in the two focal taxa and coupled ecological niche models and a spatiotemporally explicit simulation approach based on coalescent theory to determine the relative statistical support of a suite of competing demographic scenarios representing contemporary population isolation (i.e. a predominant role of genetic drift) versus historical connectivity and post-glacial colonization of sky islands (i.e. pulses of gene flow and genetic drift linked to Pleistocene glacial cycles).

Results: Inference of spatial patterns of genetic structure, environmental niche modelling and statistical evaluation of alternative species-specific demographic models within an approximate Bayesian computation framework collectively supported genetic admixture during glacial periods and post-glacial colonization of sky islands, rather than long-term population isolation, as the scenario best explaining the current distribution of genomic variation in the two focal taxa. Moreover, our analyses revealed that isolation in sky islands has also led to extraordinary genetic fragmentation and contributed to reduce local levels of genetic diversity.

Main conclusions: This study exemplifies the potential of integrating genomic and environmental niche modelling data across biological and spatial replicates to determine whether organisms with similar habitat requirements have experienced

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd.

concerted/idiosyncratic responses to Quaternary climatic oscillations, which can ultimately help to reach more general conclusions about the vulnerability of mountain biodiversity hotspots to ongoing climate warming.

KEYWORDS

approximate Bayesian computation, ddRAD-seq, demographic history, environmental niche modelling, landscape genetics, Pleistocene glaciations

1 | INTRODUCTION

Distributional shifts in response to Quaternary climatic oscillations had a dramatic impact on biogeographical patterns of species diversity, abundance and local endemism (Hewitt, 1996; Sandel et al., 2011). These cycles, in particular the high-amplitude climatic changes characterizing the mid and late Pleistocene (<0.8 Ma; Jouzel et al., 2007), have also shaped the distribution and spatial patterns of genetic variation in many organisms (Hewitt, 2000). However, multiple studies have documented considerable heterogeneity across regions and taxa in the demographic consequences of Pleistocene glacial cycles (Hewitt, 2000). On the one hand, the impact of Pleistocene glaciations strongly depended on latitude and regional topography, with extinction–recolonization dynamics at higher latitudes (i.e. "southern richness to northern purity" paradigm; Hewitt, 1996) and elevational shifts and more complex processes of population fragmentation and connectivity at lower latitudes such as the tropics or temperate regions (e.g. "refugia within refugia" concept; Gómez & Lunt, 2007). On the other hand, the way organisms respond to climate changes strongly depends on species-specific niche requirements and life-history traits, which define favourable/unfavourable climatic periods (glacials or interglacials; Bennett & Provan, 2008) and their capacity to deal with population fragmentation (e.g. microhabitat preferences, dispersal capacity; Massatti & Knowles, 2016; Papadopoulou & Knowles, 2016). These aspects determined the location and extension of Pleistocene refugia, which have played a predominant role on species' persistence during unfavourable climatic periods and acted as source populations from which species expanded their ranges at the onset of more favourable conditions (Bennett & Provan, 2008). Temperate species currently inhabiting low-elevation areas generally restricted their distributions to southern refugia during glacial phases (i.e. glacial refugia) and expanded during interglacial periods whereas cold-adapted species, nowadays presenting fragmented populations at high elevations/latitudes (i.e. interglacial refugia), had much more widespread distributions in glacial stages (Hewitt, 2000).

Mid- and low-latitude alpine and montane taxa represent small-scale replicates of cold-adapted species living at high latitudes (Hewitt, 1996, 2000). Sky islands are a paradigmatic example of interglacial refugia in which alpine and montane organisms currently form highly isolated populations in mountain ranges separated from each other by intervening valleys with unsuitable environmental conditions (e.g. DeChaine & Martin, 2005; Knowles & Massatti, 2017; McCormack et al., 2008; Mouret et al., 2011). Changing climatic conditions causes suitable habitats to expand or shrink along elevational gradients, leading to alternating periods of population isolation and connectivity (DeChaine & Martin, 2005; Knowles, 2000). Periods of isolation on mountain tops are expected to lead to demographic bottlenecks, genetic drift and divergence among populations, while periods of connectivity allow for dispersal and gene flow among formerly isolated populations (DeChaine & Martin, 2005). Species restricted to sky islands often show unique patterns of population genetic structure resulting from climate-induced distributional shifts (Mouret et al., 2011), processes that in some cases have led to lineage diversification and speciation (Knowles, 2000; Knowles & Massatti, 2017). Therefore, montane habitats are often important hotspots of intraspecific genetic diversity, species richness and local endemism, providing an excellent study system for understanding how climatic changes associated with Pleistocene glacial cycles impacted species distributions, demography and genetic diversification (Mairal et al., 2017). Their study takes particular relevance if we consider that species restricted to sky islands often show small population sizes, harbour reduced levels of local genetic diversity and are highly vulnerable to extinction due to ongoing climate warming and the progressive shrink of alpine ecosystems (Rubidge et al., 2012).

South-eastern Iberia constitutes an important biodiversity hotspot due to the interplay among a complex geological history (Meulenkamp & Sissingh, 2003), an extraordinarily complex topography (Braga et al., 2003) and a limited impact of Quaternary glaciations (Hughes & Woodward, 2017). The presence of vast areas free of permanent ice during the coldest stages of the Pleistocene made this region an important glacial refugium for warm-temperate taxa (Gómez & Lunt, 2007; Hewitt, 2011). At the same time, the mountain ranges of the region (Baetic system; >3,000 m) constitute important interglacial refugia for several cold-adapted organisms that currently persist in severely fragmented populations at high elevation patches of suitable habitat and contribute disproportionally to the extraordinary rates of local endemism of the area (Mota et al., 2002). This offers an ideal biogeographical setting to jointly study the genetic consequences of long-term population fragmentation in this sky island system and assess the time-scale at which past distributional shifts have shaped contemporary patterns of genetic variation in alpine organisms from mid-latitudes.

Here, we use genomic data obtained via restriction site-associated DNA sequencing (ddRAD-seq; Peterson et al., 2012) and a suite of analytical approaches to shed light on the demographic

processes underlying contemporary patterns of genetic variation in the grasshoppers *Omocestus bolivari* Chopard, 1939, and *Omocestus femoralis* Bolívar, 1908, two narrow-endemic Iberian species inhabiting high elevation areas in the sky island archipelago of the Baetic System (Presa et al., 2016a, 2016b). The two taxa are closely related non-sister species that present adjacent but not overlapping distributions (Tonzo et al., 2021; see Figure 1), both occupying similar open habitats of thorny shrub formations (e.g. *Erinacea* sp., *Festuca* sp., *Juniperus* sp.) located at the upper (>1,500 m; see Table S1) thermoclimatic belts of the Mediterranean region (Clemente et al., 1991; Presa et al., 2016a, 2016b). They are graminivorous, univoltine (i.e. a single generation per year) and present similar phenologies, with adult population peaks in July (Clemente et al., 1991). The distribution range of the two species includes sky islands of varying size, ranging from large ones in the main massifs of the region (*O. bolivari*: Sierra Nevada and Sierra de Baza-Filabres; *O. femoralis*: Sierra de Cazorla) to tiny patches of suitable habitat located in isolated mountains with maximum elevations close to the lower altitudinal limit of the species (Figure 1). By using these two species of uniform ecological and life-history traits as independent replicates, we exemplify how the integration of genomic and environmental niche modelling (ENM) data within a comparative framework can help to elucidate the responses of species with similar habitat requirements to processes of population fragmentation/connectivity driven by Pleistocene climatic oscillations. This can ultimately help to understand the temporal scale at which such processes have shaped contemporary patterns of genetic variation in the focal species and other co-distributed organisms with analogous environmental niches (Knowles & Massatti, 2017). In a first step, (a) we tested whether observed patterns of genomic variation in the two biological replicates reflect the signals of historical population connectivity (i.e. gene flow during glacial periods) and subsequent colonization and population isolation in sky islands (i.e. genetic drift during interglacial periods) or, rather, if post-glacial population fragmentation and genetic drift after the Pleistocene-Holocene transition have entirely monopolized the genetic makeup of contemporary populations. Specifically, we quantified genetic structure and coupled ecological niche models (ENMs) and a spatio-temporally explicit simulation approach based on coalescent theory to mimic complex demographic processes under a suite of scenarios representing historical gene flow and post-glacial colonization in sky islands versus contemporary population isolation (Knowles & Massatti, 2017; Ray et al., 2010). The expectations under these alternative scenarios were tested and validated against observed genomic data within an approximate Bayesian computation (ABC) framework (e.g. He et al., 2013). In a second step, (b) we analysed the impact of sky island size and the spatial location of contemporary populations on their demographic fate and levels of genetic variation. To this end, we tested the hypothesis that genetic diversity is positively associated with the geographic centrality of populations and habitat patch size and used demographic reconstructions based on genomic data to determine whether changes in effective population size (N_e) over time differed between sky islands of contrasting size and degree of peripheral location (Lira-Noriega & Manthey, 2014).

2 | MATERIALS AND METHODS

2.1 | Population sampling

Between 2011 and 2016, we sampled eleven populations of the Iberian grasshopper *O. bolivari* Chopard, 1939 (n = 82 individuals), and five populations of the stripe-legged grasshopper *O. femoralis* Bolívar, 1908 (n = 36 individuals) (Orthoptera: Acrididae, subgenus *Dreuxius*; Cigliano et al., 2020) (Figure 1; Table S1). The two species present allopatric distributions and can be easily identified based on diagnostic morphological traits (i.e. shape and relative size of forewings and hindwings; Clemente et al., 1991). According to our extensive surveys in south-eastern Iberia and available records in the literature (Presa et al., 2016a, 2016b), the sampled populations cover the entire known distribution range of the two species (Figure 1). We stored specimens in 2-ml vials with 96% ethanol and preserved them at −20°C until needed for genomic analyses. All specimens were preserved as vouchers in our Orthoptera collections, with labels for genotyped individuals available at the NCBI Sequence Read Archive (SRA) under BioProject PRJNA543714. The two studied species are red-listed (Presa et al., 2016a, 2016b) and were collected under permits provided by the respective environmental authorities from the different Spanish Autonomous Communities (Castilla–La Mancha: DGPFEN/SEN/AVP_16_179; Murcia: AUF/2016/0064; Andalucía: SGIB/AF).

2.2 | Genomic library preparation and processing

We used Nucleo Spin Tissue kits (Macherey-Nagel, Düren, Germany) to extract and purify total genomic DNA from a hind leg of each individual. Genomic DNA was individually barcoded and processed in-house using the double-digest restriction site-associated DNA procedure (ddRAD-seq) described in Peterson et al. (2012). Briefly, DNA was doubly digested with *EcoRI* and *MseI* restriction enzymes, followed by the ligation of Illumina adaptor sequences and unique 7-base pair barcodes. Ligation products were pooled into four libraries, size-selected for fragments between 475 and 580 bp using a Pippin prep machine (Sage Science) and amplified by iProof™ High-Fidelity DNA Polymerase (Bio-Rad) with 12 cycles. Single-read 151-bp sequencing was performed on an Illumina HiSeq2500 platform at The Centre for Applied Genomics (Hospital for Sick Children). We used the different programs distributed as part of the STACKS v. 1.35 pipeline (Catchen et al., 2013) to filter and assemble our sequences into de novo loci and call genotypes. All details on sequence assembling and data filtering are presented in Methods S1.

2.3 | Population genetic structure

We analysed population genetic structure of the two species using the Bayesian Markov chain Monte Carlo clustering method implemented in the program STRUCTURE v. 2.3.3 (Pritchard et al., 2000). We

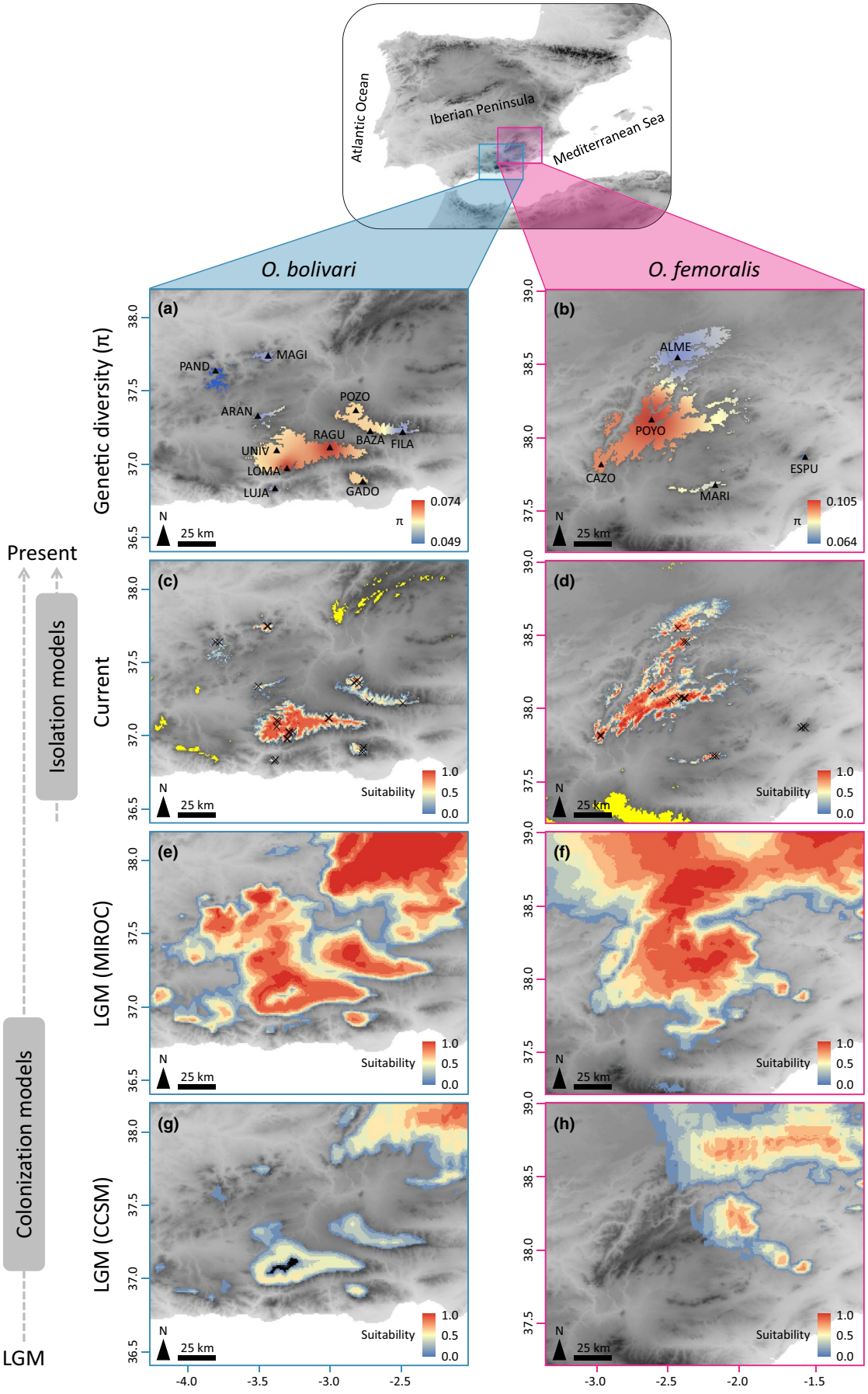
Atlantic Ocean
Iberian Peninsula
Mediterranean Sea
O. bolivari
O. femoralis
Genetic diversity (π)
Current
LGM (MIROC)
LGM (CCSM)
Present
LGM
Isolation models
Colonization models
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
PAND
MAGI
ARAN
POZO
RAGU
BAZA
FILA
UNIV
LOMA
LUJA
GADO
ALME
POYO
CAZO
ESPU
MARI
N
25 km
π
0.074
0.049
0.105
0.064
Suitability
1.0
0.5
0.0
38.0
37.5
37.0
36.5
39.0
38.5
-4.0
-3.5
-3.0
-2.5
-2.0
-1.5

FIGURE 1 (a, b) Spatial patterns of genetic diversity and (c–h) environmental niche models (ENM) for *Omocestus bolivari* and *O. femoralis*. Panels a–b show genetic diversity (π) of the studied populations (triangles, population codes as in Table S1) interpolated across the current distribution of each species as predicted by their respective ENMs. Panels C-H show the projection of ENMs for (c–d) present and Last Glacial Maximum (LGM) conditions under the (e–f) CCSM4 and (g–h) MIROC-ESM general atmospheric circulation models. Maps for current distributions show records (crosses) used to build the ENMs and the dark yellow colour indicates areas predicted as suitable out of the known limits of species' ranges (i.e. overpredictions, not shown in panels a–b). All maps show suitable areas above the maximum training sensitivity plus specificity (MTSS) logistic threshold of MAXENT. Grey background represents elevation, with darker areas corresponding to higher altitudes. Latitude and longitude in decimal degrees. The temporal scale of "isolation" (static) and "colonization" (dynamic) demographic models is indicated on the left

ran STRUCTURE assuming correlated allele frequencies and admixture without using prior population information. We conducted 15 independent runs with 200,000 MCMC cycles, following a burn-in step of 100,000 iterations, for each of the different possible *K* genetic clusters. We retained the ten runs having the highest likelihood for each value of *K* and inferred the number of populations best fitting the dataset using log probabilities [Pr(*X*|*K*)] (Pritchard et al., 2000) and the Δ*K* method (Evanno et al., 2005). Complementary to STRUCTURE analyses, we performed principal component analyses (PCA) as implemented in the R package *adegenet* (Jombart, 2008).

2.4 | Environmental niche modelling

We built environmental niche models (ENM) to predict the geographic distribution of climatically suitable habitats for *O. bolivari* and *O. femoralis* both in the present and during the Last Glacial Maximum (LGM, 21 ka). Then, we used this information to create maps of environmental suitability in these two time periods and generate alternative demographic models as detailed in the next section (e.g. He et al., 2013; Massatti & Knowles, 2016). To build ENMs, we used the maximum entropy algorithm implemented in MAXENT v.3.3.3 (Phillips et al., 2006) and the 19 bioclimatic variables from the WORLDCLIM dataset (http://www.worldclim.org/) interpolated to 30-arc-sec resolution (~1 km^2 cell size) (Hijmans et al., 2005). To generate climate suitability maps during the LGM, we projected the ENM onto LGM bioclimatic conditions derived from the CCSM4 (Community Climate System Model; Braconnot et al., 2007) and the MIROC-ESM (Model of Interdisciplinary Research on Climate; Hasumi & Emori, 2004) general atmospheric circulation models. Further details on ENM are presented in Methods S2.

2.5 | Testing alternative demographic models

We used the integrative distributional, demographic and coalescent (iDDC) approach (He et al., 2013) and an approximate Bayesian computation (ABC) framework (Beaumont et al., 2002) to test alternative scenarios representing different hypotheses about how landscape heterogeneity (or its lack thereof) and colonization from glacial refugia versus contemporary population isolation in sky islands explain the spatial distribution of genomic variation in our two focal species. This approach is described in He et al. (2013) and consists of three main steps: (a) constructing alternative demographic models representing different hypotheses about the processes shaping spatial patterns of genetic structure and diversity; (b) running demographic and genetic simulations under each model using the software SPLATCHE2 (see Ray et al., 2010); (c) evaluating the fit of observed genomic data (i.e. empirical genomic data obtained after genotyping sampled populations) to the genetic expectations under each model, identifying the most probable model/s and estimating demographic parameters (e.g. González-Serna et al., 2019; He et al., 2013). Below, we describe the most relevant aspects of this approach and present an extended version with all the details in Methods S3.

2.5.1 | Constructing demographic models

We generated two main sets of models that differ in the hypothetical demographic processes that have shaped spatial patterns of contemporary genetic variation (Table 1):

(a) Colonization of sky islands from glacial refugia (Models A, B, D and E). The first two models (Models A and B) are dynamic models (sensu He et al., 2013; Figure 1) incorporating the colonization process from hypothetical glacial refugia and distributional shifts resulted from the interaction between the species bioclimatic envelope and Pleistocene glacial cycles (e.g. He et al., 2013; Massatti & Knowles, 2016). In these scenarios, carrying capacities change over time according to climatic suitability maps obtained from projections of the ENM to the present and the LGM bioclimatic conditions under the MIROC-ESM (Model A) and CCSM (Model B) general atmospheric circulation models (see Section 2.4). These models considered landscapes from three consecutive time periods (LGM, intermediate and current) reflecting temporal shifts in the spatial distribution of environmentally suitable areas for the species in response to climate changes since the LGM (e.g. He et al., 2013; Massatti & Knowles, 2016). Forward demographic simulations under this model initialized 21 ka BP from hypothesized refugial populations (i.e. source populations, each one with an effective population size of N_{ANC}) that were located at every habitat patch predicted as suitable for each focal species during the LGM according to MIROC-ESM (Model A) and CCSM (Model B) projections. Specifically, suitable habitat patches during the LGM were identified as those cells with a probability of presence of the species above the maximum training sensitivity plus specificity (MTSS) logistic threshold for occurrence from MAXENT (Liu et al., 2005). Finally, we generated two

TABLE 1 Statistics from the ABC procedure used for evaluating the relative support of each demographic model in *Omocestus bolivari* and *Omocestus femoralis*

Model		Ancestral sources	Layer/s	Marginal density	Wegmann's *p*-value	Bayes factor	R^2		
							K_{MAX}	*m*	N_{ANC}
O. bolivari									
A	Colonization	LGM_{MIROC}	LGM_{MIROC} → Current	9.14×10^{-07}	.368	2.09	0.78	0.61	0.92
B	Colonization	LGM_{CCSM}	LGM_{CCSM} → Current	1.91×10^{-06}	.174	—	0.73	0.61	0.91
C	Isolation	Current	Current	1.52×10^{-13}	<.001	1.26×10^{07}	0.79	0.66	0.91
D	Colonization	LGM_{MIROC}	Flat landscape	2.76×10^{-09}	.006	6.93×10^{02}	0.69	0.68	0.93
E	Colonization	LGM_{CCSM}	Flat landscape	9.15×10^{-08}	.012	2.09×10^{01}	0.70	0.70	0.91
F	Isolation	Current	Flat landscape	3.67×10^{-13}	<.001	5.22×10^{06}	0.69	0.72	0.90
O. femoralis									
A	Colonization	LGM_{MIROC}	LGM_{MIROC} → Current	1.26×10^{-06}	.026	1.28×10^{01}	0.36	0.68	0.96
B	Colonization	LGM_{CCSM}	LGM_{CCSM} → Current	8.82×10^{-06}	.195	1.83	0.28	0.66	0.93
C	Isolation	Current	Current	1.00×10^{-100}	<.001	1.61×10^{95}	0.41	0.59	0.92
D	Colonization	LGM_{MIROC}	Flat landscape	6.35×10^{-09}	.001	2.54×10^{03}	0.24	0.69	0.96
E	Colonization	LGM_{CCSM}	Flat landscape	1.61×10^{-05}	.239	—	0.19	0.73	0.96
F	Isolation	Current	Flat landscape	1.00×10^{-100}	<.001	1.61×10^{95}	0.19	0.66	0.91

Note: A higher marginal density corresponds to a higher model support and a high Wegmann's *p*-value indicates that the model is able to generate data in agreement with empirical data. Bayes factors represent the degree of relative support for the model with the highest marginal density over the other models. R^2 is the coefficient of determination from a regression between each demographic parameter (K_{MAX}, *m*, N_{ANC}) and the five partial least squares (PLS) extracted from all summary statistics.

K_{MAX}, carrying capacity of the deme with highest suitability (100%).

m, migration rate per deme per generation.

N_{ANC}, ancestral population size.

more models (Models D and E) analogous to the previous ones but in which carrying capacities (*k*) are homogeneous across space and through time. These static models (sensu He et al., 2013) are equivalent to a flat landscape or an isolation-by-distance model and only differ among them in the location of ancestral populations, which were based on the patches of suitable habitat identified under LGM-MIROC (Model D) and LGM-CCSM (Model E) bioclimatic conditions.

(b) Population isolation in sky islands (Models C and F). The first model of this set (Model C) is a static model representing genetic drift associated with the isolation of populations in sky islands according to the current geographical configuration of suitable habitats (Figure 1). In this model, carrying capacities do not change over time and are defined by an environmental suitability layer obtained from the projection of the ENM to the current bioclimatic conditions. Forward demographic simulations under this model initialized 7 ka BP from hypothesized refugial populations located in every patch of habitat predicted as suitable during present time according to the MTSS logistic threshold for occurrence from MAXENT. Thus, this model hypothesizes that the spatial distribution of contemporary genetic variation reflects population fragmentation and isolation in sky islands since the mid-Holocene (e.g. Knowles & Massatti, 2017). Similar to the first set of models, we also generated a scenario (Model F) analogous to the previous one but in which carrying capacities (*k*) are homogeneous across space (i.e. equivalent to an isolation-by-distance model; He et al., 2013).

2.5.2 | Model choice and parameter estimation

We used an approximate Bayesian computation (ABC) framework to perform model selection and parameter estimation, as implemented in ABCTOOLBOX programs (TRANSFORMER and ABCESTIMATOR) and R scripts (*findPLS*) (Wegmann et al., 2010). We used the R package *pls* v.2.6-0 (Mevik & Wehrens, 2007) and the *findPLS* script to extract partial least-squares (PLS) components with Box–Cox transformation from the summary statistics of the first 10,000 simulations for each model (Wegmann et al., 2010). The first five PLSs extracted from the summary statistics were used for ABC analyses, as the root-mean-squared error (RMSE) of the three demographic parameters employed (K_{MAX}, *m*, N_{ANC}; see Methods S3 for details) for the two species did not decrease significantly with additional PLSs (Figure S1). We used the linear combinations of summary statistics obtained from the first 10,000 simulations for each model to transform all datasets (observed empirical and simulated datasets) with the program TRANSFORMER (Wegmann et al., 2010). For each model, we retained the 1,000 simulations (0.5%) closest to observed empirical data and used them to approximate marginal densities and posterior distributions of the parameters with a post-sampling regression adjustment using the ABC-GLM (general linear model) procedure detailed in Leuenberger and Wegmann (2010) and implemented in ABCESTIMATOR. We used Bayes factors (BF) for model selection (Kass & Raftery, 1995).

2.5.3 | Model validation

To evaluate the ability of each model to generate the empirical data, we calculated Wegmann's *p*-value from the 1,000 retained simulations (Wegmann et al., 2010). We also assessed the potential for a parameter to be correctly estimated by computing the proportion of parameter variance that was explained (i.e. the coefficient of determination, R^2) by the retained PLSs. For the most supported model for each species, we determined the accuracy of parameter estimation using a total of 1,000 pseudo-observation datasets (PODs) generated from prior distributions of the parameters. If the estimation of the parameters is unbiased, posterior quantiles of the parameters obtained from PODs should be uniformly distributed (Wegmann et al., 2010). As with the empirical data, we calculated the posterior quantiles of true parameters for each pseudo run based on the posterior distribution of the regression-adjusted 1,000 simulations closest to each pseudo-observation.

2.6 | Genetic diversity and historical changes in effective population size

To further explore the processes determining the spatial distribution of genetic variation in the two species, we (a) tested the association between genetic diversity and the geographic centrality and habitat patch size of the studied populations and (b) used genomic data to reconstruct changes in effective population size (N_e) over time. We employed linear regressions in SPSS v. 26.0 to analyse whether genetic diversity (nucleotide diversity, π) is associated with the distance of each population to the species' distribution centroid and contemporary suitable habitat patch size. Given that the precision of genetic diversity estimates may differ among populations due to differences in sample sizes, we used a weighted least-square (WLS) method where weight equals the number of genotyped individuals for each population (Table S1). We calculated nucleotide diversity for each population using the program *populations* from STACKS. The centroid of species distribution was calculated in ARCMAP v.10.3 on the basis of polygons (patches) of suitable habitat defined by grid cells with a probability of presence of the focal species above the maximum training sensitivity plus specificity (MTSS) logistic threshold for occurrence from MAXENT (Liu et al., 2005). Similarly, patch size for each study population was calculated considering the area (in km^2) of continuous suitable habitat defined as above on the basis of raster maps obtained from projecting ENMs to contemporary bioclimatic conditions.

We inferred changes in effective population sizes (N_e) over time for each studied population using STAIRWAY PLOT (Liu & Fu, 2015), a flexible multi-epoch model approach based on the site frequency spectrum (SFS) that does not require a predefined demographic model for estimating past demographic histories. To compute a folded SFS for each population, we ran STACKS in order to obtain a VCF file retaining only loci represented in at least 50% of the individuals of the population. To remove all missing data for the calculation of the SFS and minimize errors with allele frequency estimates, each population was down-sampled to five individuals using a custom Python script written by Qixin He and available on Dryad (http://dx.doi.org/10.5061/dryad.23hs1). We ran STAIRWAY PLOT considering a 1-year generation time, assuming the mutation rate per site per generation of 2.8×10^{-9} estimated for *Drosophila melanogaster* (Keightley et al., 2014) and performing 200 bootstrap replicates to estimate 95% confidence intervals.

3 | RESULTS

3.1 | Genomic data

After quality filtering, 182,588,301 and 74,830,407 reads were retained across all genotyped individuals of *O. bolivari* and *O. femoralis*, respectively (Figure S2). Final exported datasets obtained with STACKS after removing loci that did not meet the different filtering requirements contained 2,262 SNPs for *O. bolivari* and 14,116 SNPs for *O. femoralis*.

3.2 | Population genetic structure

STRUCTURE analyses showed that log probabilities of the data (LnPr(X|*K*)) reached a plateau for $K = 8$ in *O. bolivari* and $K = 5$ in *O. femoralis* (Figure S3). The ΔK criterion indicated an "optimal" clustering solution for $K = 5$ in *O. bolivari* and $K = 3$ in *O. femoralis* (Figure S3). The two species presented a strong genetic structure, with most (*O. bolivari*) or all (*O. femoralis*) populations being assigned to unique genetic clusters with little or no genetic admixture among them (Figures 2 and S4). In the two species, $K = 2$ separated north-western (*O. bolivari*: PAND, MAGI, ARAN, and UNIV; *O. femoralis*: ALME and CAZO) from south-eastern populations (*O. bolivari*: POZO, BAZA, FILA, RAGU, LUMA, LUJA and GADO; *O. femoralis*: POYO, MARI and ESPU). These two main clusters split hierarchically at increasingly higher *K* values, generally grouping the different populations in accordance with their geographical location. Remarkably, populations located in the same sky island did not always cluster together at different hierarchical levels. For instance, in *O. bolivari*, population UNIV grouped with north-western populations (PAND, MAGI and ARAN) rather than with other populations (LOMA and RAGU) located within its same mountain range (i.e. the main Sierra Nevada massif; Figures 1a and 2b). In the same line, population POYO from *O. femoralis* clustered together with the highly isolated south-eastern populations (MARI and ESPU) rather than with the two other populations (ALME and CAZO) located within the Sierra de Cazorla massif and with which it is expected to be currently more connected (Figures 1b and 2b). Principal component analyses for both species yielded analogous results to those obtained with STRUCTURE, separating along PC1 north-western from south-eastern populations and grouping populations according to their geographical location (Figure S5).

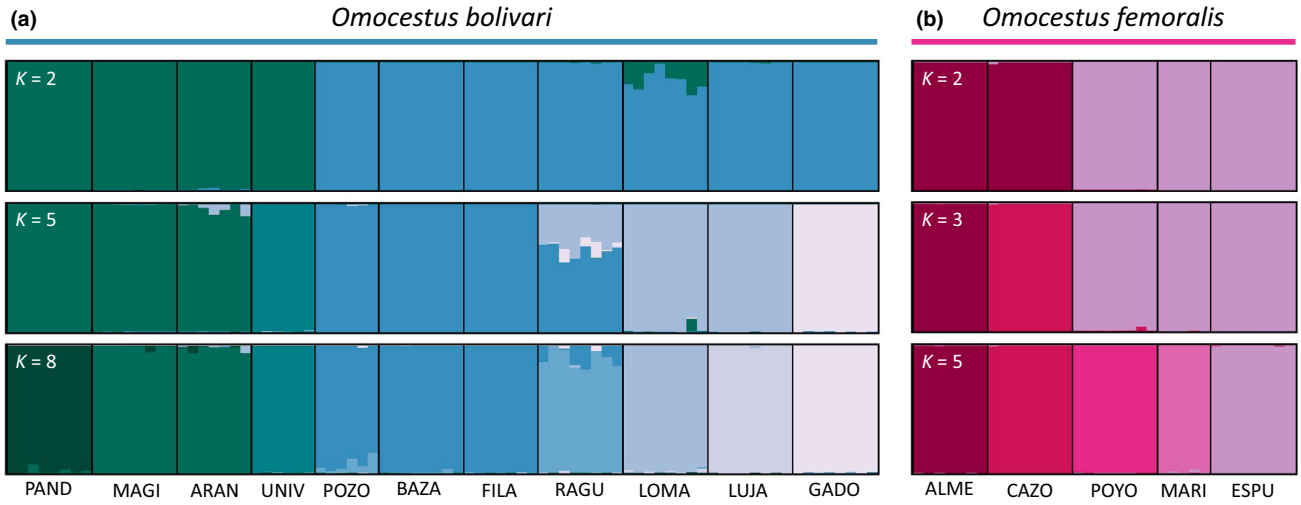

FIGURE 2 Genetic assignment of individuals based on the results of STRUCTURE for (a) *Omocestus bolivari* ($K = 2$, $K = 5$ and $K = 8$) and (b) *Omocestus femoralis* ($K = 2$, $K = 3$ and $K = 5$). Individuals are partitioned into K coloured segments representing the probability of belonging to the cluster with that colour. Thin vertical black lines separate individuals from different populations. Population codes as in Table S1

3.3 | Environmental niche modelling

We generated final ENMs using the feature class (FC) combination (*O. bolivari* = LQ; *O. femoralis* = LQ) and regularization multiplier (*O. bolivari* = 6.0; *O. femoralis* = 2.0) that minimized AICc across the set of models tested with ENMEVAL for each species. After removing highly correlated variables ($r > 0.9$; e.g. Jueterbock et al., 2016) and those with a zero percent contribution, our models retained four bioclimatic variables for *O. bolivari* (sorted by percent contribution, BIO8: 36.8%; BIO7: 34.8%; BIO11: 16.5%; BIO2: 11.9%) and eight variables for *O. femoralis* (BIO11: 37.9%; BIO2: 33.9%; BIO7: 13.6%; BIO8: 7.4%; BIO19: 4.3%; BIO15: 1.3%; BIO12: 1.2%; BIO18: 0.4%). Inspection of predicted distributions for the present confirmed that ENMs yielded distribution patterns coherent with the observed current distribution of the two species (Figure 1c,d). However, current distribution maps for the two species also identified as suitable some high elevation areas outside their respective distribution ranges. In the two species, overpredicted areas included partially (*O. bolivari*) or totally (*O. femoralis*) the distribution range of each other, which is not surprising given their similar habitat, elevation and environmental requirements. Projections of present-day climate niche envelopes to LGM climatic conditions under two general atmospheric circulation models (CCSM4 and MIROC-ESM) suggest that the two species have experienced important distributional shifts in response to Pleistocene glaciations (Figure 1e–h). As expected for montane–alpine organisms, the distribution of the two species expanded into lowland areas during the LGM (Figure 1e–h). However, projections under the two circulation models also presented important differences. Populations of the two species were predicted to be much more connected during the LGM under the MIROC-ESM than under the CCSM model. According to MIROC-ESM, most currently isolated populations of the two species (with the exception of Sierra de Lújar and Sierra de Gádor in *O. bolivari*) became connected by corridors of suitable habitat during the LGM. However, in *O. bolivari*, most contemporary sky islands were also predicted to be isolated from each other during the LGM according to the CCSM model (with the exception of Sierra de Arana and Sierra Nevada; Figure 1g). In the case of *O. femoralis*, the southwesternmost portion of its current distribution was predicted as unsuitable under the CCSM model, suggesting that this species might have experienced important distributional changes beyond elevational shifts (Figure 1h).

3.4 | Testing alternative demographic models

Based on marginal densities calculated from the 0.5% of retained simulations, the best ranked models in the two species were those incorporating the colonization process from hypothetical glacial refugia (Models A, B, D and E; Table 1). In contrast, models of isolation in contemporary sky islands had considerably lower marginal densities and a difference in Bayes factors $> 10^6$ with the most supported model (Table 1), indicating strong relative support for colonization models in the two species (Kass & Raftery, 1995). Moreover, only colonization models were able to simulate genetic data comparable with empirical data (Models A and B in *O. bolivari* and Models B and E in *O. femoralis*), unlike the isolation models in which there was a substantial difference between the likelihoods of the simulated data compared with the empirical data (Wegmann's p-values $< .001$ in all cases; Table 1). Specifically, for *O. bolivari*, dynamic colonization models incorporating changes in habitat suitability over time under both the CCSM and MIROC projections (Models A and B) were those most supported by our empirical data and small Bayes factors (BF < 3; Table 1) suggest that they are statistically indistinguishable

(Kass & Raftery, 1995). In *O. femoralis*, both dynamic and static (i.e. flat landscape) colonization models based on CCSM projections (Models B and E) were the most supported by our empirical data and statistically indistinguishable (BF < 2; Table 1). The dynamic colonization model based on MIROC bioclimatic conditions also had a low Bayes factor (<20), but its low Wegmann's *p*-value (<.03) indicates that it was not capable of generating data compatible with our empirical data (Table 1).

Posterior distributions of parameters under the most probable models were considerably distinct from the prior, indicating that the simulated data contained information relevant to estimating the parameters (Figure S6). Comparison of the posterior distributions before and after the ABC-GLM also showed the improvement that this procedure had on parameter estimates (Figure S6). The posterior distribution of K_{MAX} in *O. femoralis* was flat, indicating uncertainty in the estimation of this parameter (Figure S6b). Accordingly, the coefficients of determination (R^2) from a multiple regression between each demographic parameter and the five retained PLS indicated that the employed summary statistics had a moderately high potential to correctly estimate all the parameters except K_{MAX} in *O. femoralis* (Table 1). Posterior quantiles of *m* in *O. bolivari*, K_{MAX} in *O. femoralis* and N_{ANC} in both taxa significantly deviated from a uniform distribution, indicating a potential bias in the estimation of these parameters (Figure S7).

3.5 | Genetic diversity and historical changes in effective population sizes

Visual inspection of the spatial distribution of genetic variation in the two species evidenced that peripheral populations and those located in smaller sky islands tended to present lower levels of genetic diversity (Figure 1a,b). Accordingly, genetic diversity (π) in populations of *O. bolivari* was negatively associated with population centrality ($r = -.913$, $t = -6.72$, $p < .001$; Figure 3a) and showed a positive relationship with patch size ($r = 0.733$, $t = 3.23$, $p = .010$; Figure 3b). In the case of *O. femoralis*, genetic diversity was positively associated with patch size ($r = 0.878$, $t = 3.17$, $p = .050$; Figure 3d) but the negative relationship with population centrality did not reach statistical significance ($r = 0.743$, $t = -1.92$, $p = .150$; Figure 3c). STAIRWAY PLOT analyses for *O. bolivari* revealed similar results across most of the populations, with an abrupt demographic expansion led by the advance of the glaciation till its maximum level around 21 ka BP (Figure 4a,b). Afterwards, populations experienced a decline in N_e until recent times (Figure 4a,b). One exception was the highly isolated populations PAND, MAGI and LUJA, which presented a general demographic stability during the last 100 ka (Figure 4b). In the case of *O. femoralis*, most populations showed idiosyncratic demographic histories, which ranged from expansions during the last glacial period followed by reductions in N_e at the onset of the Holocene (ALME and MARI) to marked bottlenecks around the LGM (ESPU and CAZO) (Figure 4c,d).

4 | DISCUSSION

Our study contributed to a better understanding of the demographic processes underlying spatial patterns of genetic variation in cold-adapted biotas currently forming severely fragmented populations in sky island archipelagos from temperate regions (Hewitt, 1996; Knowles & Massatti, 2017; Perrigo et al., 2020). Specifically, the combination of population genomic data, niche modelling and spatiotemporally explicit coalescent-based simulations supported population connectivity during glacial periods followed by colonization and population isolation in sky islands as the most supported scenario explaining contemporary patterns of population genetic diversity and structure in the two focal taxa. Although our analyses indicate that post-glacial fragmentation and genetic drift have not blurred the genomic signatures left by historical patterns of population connectivity, the lower levels of genetic diversity in peripheral populations confined to small patches of suitable habitat (i.e. small-sized sky islands) evidence the genetic consequences of long-term isolation. This study exemplifies how the integration of multiple lines of evidence provided by a comprehensive suite of analytical approaches applied to two independent spatial and biological replicates can help to identify concerted or idiosyncratic evolutionary and demographic histories of species with similar habitat requirements. Ultimately, this can lead to more general conclusions about the way organisms and whole communities have responded to Quaternary climatic oscillations. These aspects are of pivotal importance to understand the dynamics of montane and alpine ecosystems from mid- and low-latitude regions, which not only harbour high levels of local endemism but have been also identified among the most vulnerable to extinction due to ongoing climate warming and the progressive shrink of these habitats (Perrigo et al., 2020).

4.1 | Marked genetic structure in sky islands

The contemporary genetic structure of *O. bolivari* and *O. femoralis* exhibited many similarities, with both taxa showing a marked north-west–south-east pattern of genetic differentiation and most genetic clusters identified at the different hierarchical levels presenting a very low degree of genetic admixture (Figure 2). The north-west–south-east pattern of differentiation likely reflects the predominant connectivity and origin of source populations that led to the colonization of sky islands from the more widespread lowland distributions presented by the two species during glacial periods, as supported by projections of environmental niche models to LGM bioclimatic conditions under alternative general circulation models (Figure 1). Remarkably, the two taxa presented populations that, being located on the same mountain range, are assigned to different genetic clusters, which, in turn, are shared with populations located in different sky islands. This is exemplified by the genetic assignment of UNIV and POYO populations from *O. bolivari* and *O. femoralis*, respectively. Population UNIV is situated in a slope facing north in Sierra Nevada massif

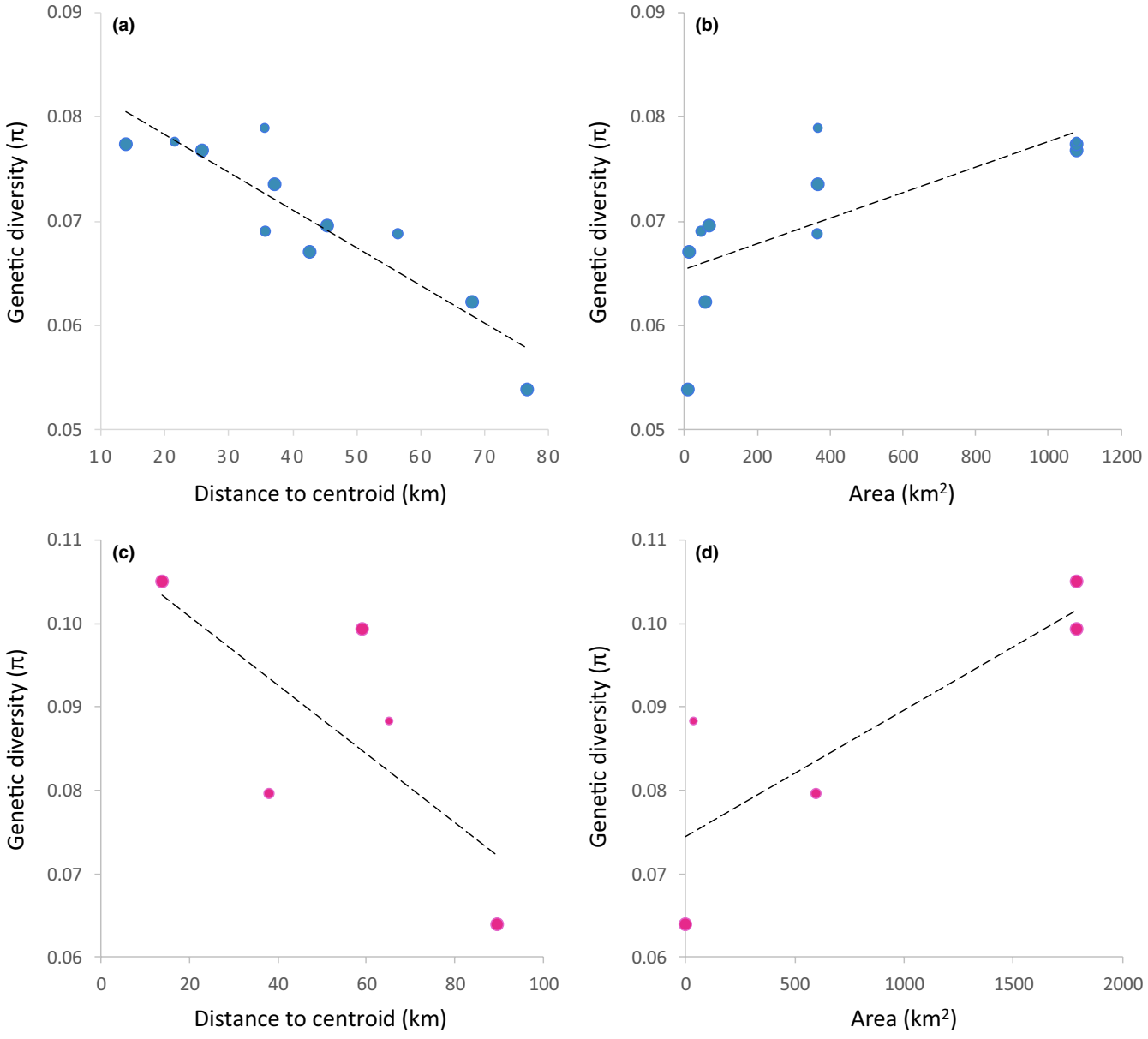

FIGURE 3 Relationship between the genetic diversity (π) of populations and (a, c) distance to centroid of species' distribution and (b, d) area of climatically suitable habitat for (a, b) *Omocestus bolivari* and (c, d) *Omocestus femoralis*. Regression lines are indicated and dot size is proportional to sample size for each studied population

but shares a common ancestry with the northernmost populations of the species rather than with the other two populations (LOMA and RAGU) located within the same mountain range (Figure 2a). Similarly, the large Sierra de Cazorla massif encompasses populations CAZO, ALME and POYO, but the latter clusters with the highly isolated and southernmost populations MARI and ESPU (Figure 2b). This indicates that patterns of genetic structure at different hierarchical levels are in some instances more congruent with the geographical arrangement of populations than with their location in a specific sky island. Projections of ENMs to LGM bioclimatic conditions point to the connection of these populations during cold periods, indicating that apparently incongruent patterns of genetic structure likely reflect the identity of ancestral glacial-source populations rather than the spatial configuration of contemporary suitable habitats and hypothetical corridors to gene flow. The low-dispersal capacity of the brachypterous focal taxa (Presa et al., 2016a, 2016b) and the considerable topographic complexity (i.e. ridges and steep slopes) within the main massifs (>3,400 m in Sierra Nevada) of the region might explain the deep genetic structure and low or absent genetic admixture among populations located within the same sky island even if these are predicted to be nowadays connected by corridors of suitable habitat according to ENMs (González-Serna et al., 2019; Tonzo et al., 2019). These inferences, coupled with those provided by our spatially explicit demographic modelling, emphasize how the genetic makeup and demographic dynamics of sky island populations

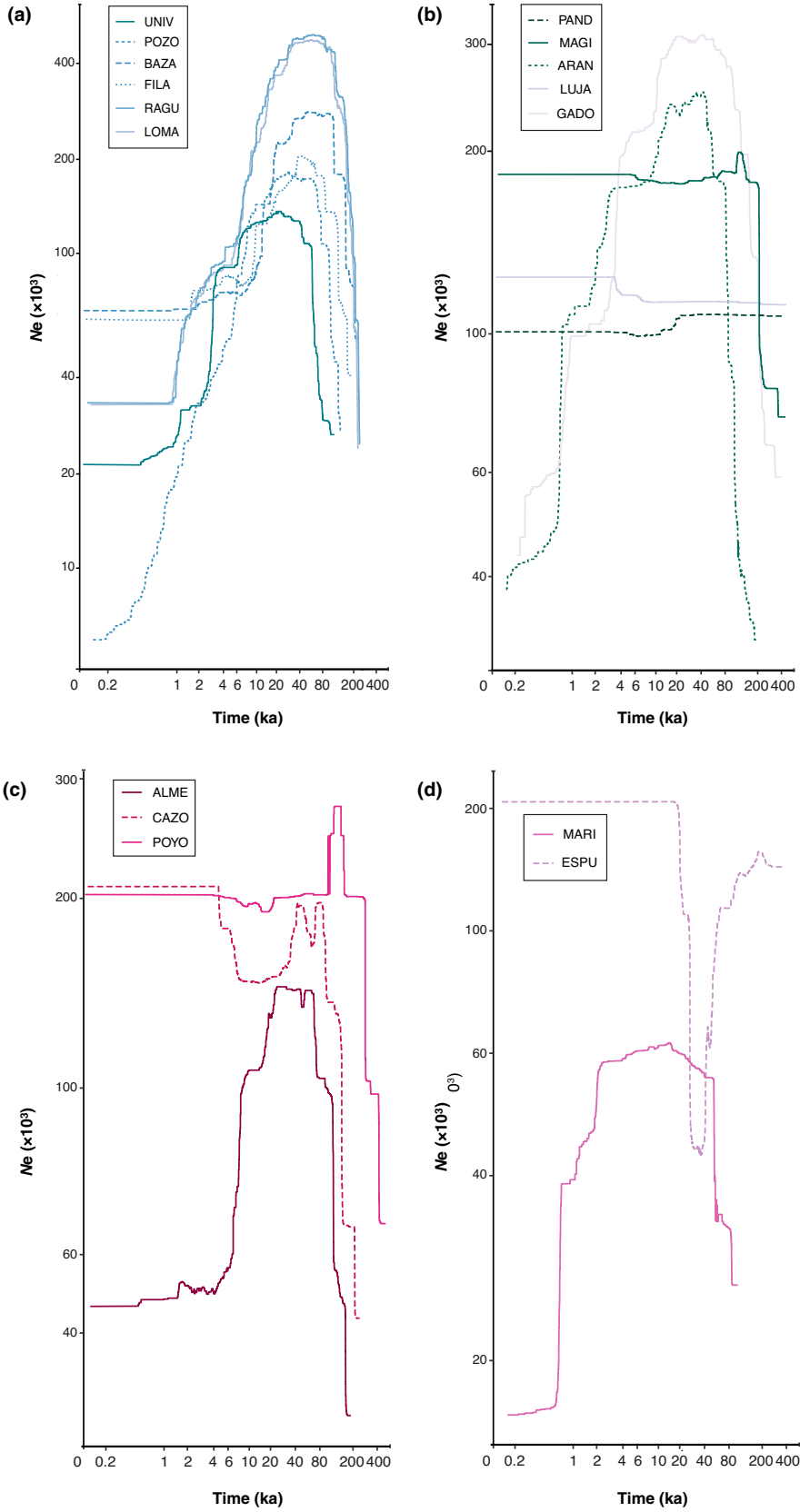

FIGURE 4 Demographic history of each studied population of (a, b) *Omocestus bolivari* and (c, d) *Omocestus femoralis* inferred using STAIRWAY PLOT. Lines show the median estimate of effective population size (N_e) over time (ka) for populations located in (a, c) large and (b, d) small sky islands. Population codes shown in the legend of each panel are described in Table S1

transcend their contemporary distributions in specific mountain ranges, which warns against oversimplified isolation models often assumed for alpine and montane organisms (see also Knowles & Alvarado-Serrano, 2010; Knowles & Massatti, 2017). Thus, the idiosyncratic nature of isolation of sky island biotas cannot be fully understood without a thoughtful assessment of the spatial

and temporal dimensions of population fragmentation and connectivity (i.e. colonization history, multiple source populations and historical gene flow), which present certain analogies but are also expected to substantially differ from the evolutionary pathways and demographic dynamics dominating true island systems (Flantua et al., 2020).

4.2 | Post-glacial colonization of sky islands

Model selection under an ABC framework revealed that the spatially explicit demographic scenarios most supported by empirical data were those in which the spatial distribution of contemporary genomic variation was shaped by processes of post-glacial colonization of sky islands from source populations defined by the spatial location of environmentally suitable areas during the last ice age (Knowles & Massatti, 2017; Mouret et al., 2011). In contrast, models of isolation in sky islands since the onset of the Holocene had a very low statistical support, indicating that genetic drift after post-glacial fragmentation cannot solely explain the genetic makeup of contemporary populations. Although scenarios based on colonization of sky islands were the only ones able to simulate genomic data compatible with observed data, our approach was not able to distinguish between dynamic models considering changes in landscape heterogeneity since the LGM and static models considering a diffusion scenario of expansion from glacial-source populations (see also Knowles & Massatti, 2017). These results are surprising considering that the two focal taxa have strict habitat preferences for high-elevation areas and unsuitable lowlands are expected to act as impassable barriers to dispersal, as evidenced by distribution models (Figure 1) and strong genetic structure of contemporary populations (Figure 2). Simple Mantel tests indicate that genetic differentiation (F_{ST}) is positively correlated with geographical distances among populations in *O. bolivari* ($r = 0.57$, $p < .001$, $n = 11$) and marginally in *O. femoralis* ($r = 0.60$, $p = .058$, $n = 5$). Processes of genetic admixture and drift mediated by population expansion and contraction during Quaternary glacial cycles are expected to have resulted in this pattern of isolation by distance (Slatkin, 1993), which is similar to the expected outcome of migration-genetic drift equilibrium under a static, flat landscape scenario (He et al., 2013; Knowles & Massatti, 2017). Thus, admixture among nearby populations is likely to have reduced the power of our analyses to discriminate between dynamic (Models A–B) and static (Models D–E) scenarios of post-glacial colonization, particularly in *O. femoralis* for which only five populations were available for analyses (Table S1).

The two general circulation palaeoclimatic models supported the expansion of the two focal taxa towards lowland areas during the LGM, but their respective populations were predicted to be much more connected under the MIROC-ESM than under the CCSM model (Figure 1). Although differences between palaeoclimatic models in predicting past species distributions are a frequent outcome of most studies (e.g. Ramírez-Barahona & Eguiarte, 2014), the relative support for contrasting predictions is rarely validated using independent sources of information (Nogués-Bravo, 2009). Such differences seemed to be captured by our model-based approach, which revealed that glacial-source populations identified for the two focal taxa on the basis of projections of ENMs to bioclimatic conditions under the CCSM general circulation model tended to provide a better fit to our genomic data than those obtained based on MIROC projections (BF > 2 in all cases; Table 1). Although differences in the relative likelihood for scenarios based on the two palaeoclimatic models are small and must be interpreted with extreme caution, our results highlight the potential of spatiotemporally explicit coalescent-based simulations to refine hypotheses about the location of glacial refugial (Bemmels et al., 2019) and validate distributional shifts inferred from ENMs that cannot usually be contrasted with independent sources of information (e.g. fossil records; Fordham et al., 2014; Metcalf et al., 2014).

4.3 | Genetic diversity and demographic reconstructions

The spatial distribution of genetic variation in the two species evidenced that peripheral populations and those located in smaller sky islands tend to present comparatively lower levels of genetic diversity (Figure 1). These findings are consistent with the centre-periphery model, which predicts that populations located on the species range edges become progressively smaller, are spatially more isolated, and harbour lower levels of genetic diversity than populations situated at the core of the distribution (e.g. Lira-Noriega & Manthey, 2014). In contrast, genomic-based demographic reconstructions indicate that, in general terms, most populations presented larger effective population sizes during the last glacial period and experienced demographic declines at the onset of the Holocene (Figure 4), in agreement with palaeoclimatic-based reconstructions of species distributions (Figure 1). Inference of changes of N_e through time did not reveal obvious differences between populations located in large versus small sky islands, suggesting that contemporary populations, and their ancestral sources, reacted in a similar fashion to Quaternary climatic oscillations independently of their current geographical location. This is not surprising given the small range of the two focal taxa (Presa et al., 2016a, 2016b), with most distant populations separated by <150 km, and the likely similar regional impact of Pleistocene glacial cycles across the entire distribution of the species. Also, it must be considered that, according to our own field observations, small habitat patches can often sustain high local densities of the species. This might contribute to maintain large effective population sizes and avoid strong genetic drift even in populations currently confined to tiny sky islands. These results indicate that although the genetic makeup of contemporary populations has been shaped in a great extent by historical processes of genetic admixture (glacial periods) and drift during the retreat to interglacial refugia (e.g. Knowles & Massatti, 2017), contemporary isolation has also contributed to erode the levels of genetic diversity of populations persisting in peripheral sky islands of small size (Lira-Noriega

& Manthey, 2014; Rubidge et al., 2012). Collectively, our results are in agreement with the "interglacial refugia model" documented for other montane and alpine organisms from temperate and tropical regions in which populations experienced expansions during glacial periods followed by severe demographic contractions and disruption of gene flow during interglacials (Bennett & Provan, 2008; Camacho-Sánchez et al., 2018). It must be noted that this scenario is very different to the dynamics of alpine organisms from temperate regions that were more extensively glaciated (e.g. North America) and for which genetic-based inferences have often supported strong isolation and divergence in glacial refugia and net demographic expansions during warm interglacials ("expanding alpine archipelago model"; e.g. Maier et al., 2019; Schoville et al., 2011). This emphasizes the importance of inferring the demographic fate of montane and alpine biotas under different biogeographical contexts, as these will largely determine, in interaction with taxon-specific traits (e.g. dispersal capacity and habitat specialization), patterns of local endemism, the extent of population fragmentation, and the magnitude of distributional shifts in response to past and future climate changes (Flantua et al., 2020; Massatti & Knowles, 2016).

5 | CONCLUSIONS

By considering two biological replicates from the sky island archipelago of the Baetic System, our study highlights the potential of integrating different sources of information to infer the evolutionary and demographic processes shaping spatial patterns of genetic variation in cold-adapted faunas from temperate regions. Our spatiotemporally explicit simulations and testing of alternative demographic scenarios supported population expansion during cold periods and subsequent post-glacial colonization and isolation in sky islands as the most likely explanation for the current distribution of genetic variation in alpine and montane biotas from the Mediterranean biodiversity hotspot. Global warming is expected to threaten the persistence of alpine and cold-adapted organisms from mid- and low-latitude regions by extraordinarily reducing the extent of suitable habitats already confined to mountain tops (Rubidge et al., 2012). The similar demographic responses inferred for the two focal taxa to past climate changes suggest that they and other co-distributed organisms with similar life-history traits will probably present concerted responses to ongoing global change. Thus, our results can help to establish unified conservation strategies aimed at preserving the extraordinary rates of local endemism of this and other mountain biodiversity hotspots, from a community-level perspective (Perrigo et al., 2020).

ACKNOWLEDGEMENTS

We are much indebted to Anna Papadopoulou for her valuable help in study design and useful comments, suggestions and corrections on a first draft of the manuscript. We also wish to thank Pedro J. Cordero and Víctor Noguerales for their help during fieldwork and providing samples from some populations, Amparo Hidalgo for support during laboratory work, Sergio Pereira (The Centre for Applied Genomics) for Illumina sequencing and three anonymous referees for their constructive and valuable comments on an earlier version of the manuscript. Logistical support was provided by Laboratorio de Ecología Molecular (LEM-EBD) and Laboratorio de Sistemas de Información Geográfica y Teledetección (LAST–EBD) from Estación Biológica de Doñana. We also thank Centro de Supercomputación de Galicia (CESGA) and Doñana's Singular Scientific-Technical Infrastructure (ICTS-RBD) for access to computer resources. This study was funded by the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (ERDF) (CGL2014-54671-P and CGL2017-83433-P). VT was supported by an FPI predoctoral fellowship (BES-2015-73159) from the Spanish Ministry of Economy and Competitiveness.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ddi.13306.

DATA AVAILABILITY STATEMENT

Raw Illumina reads have been deposited at the NCBI Sequence Read Archive (SRA) under BioProject PRJNA543714. Input files for all analyses are available for download on Figshare (https://doi.org/10.6084/m9.figshare.13643030).

ORCID

Vanina Tonzo https://orcid.org/0000-0002-5062-1070
Joaquín Ortego https://orcid.org/0000-0003-2709-429X

REFERENCES

Beaumont, M. A., Zhang, W. Y., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, *162*(4), 2025–2035. https://doi.org/10.1093/genetics/162.4.2025

Bemmels, J. B., Knowles, L. L., & Dick, C. W. (2019). Genomic evidence of survival near ice sheet margins for some, but not all, North American trees. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(17), 8431–8436. https://doi.org/10.1073/pnas.1901656116

Bennett, K. D., & Provan, J. (2008). What do we mean by refugia? *Quaternary Science Reviews*, *27*(27–28), 2449–2455. https://doi.org/10.1016/j.quascirev.2008.08.019

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J.-Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, T. H., Hewitt, C. D., Kageyama, M., Kitoh, A., Laîné, A., Loutre, M.-F., Marti, O., Merkel, U., Ramstein, G., Valdes, P., Weber, S. L., … Zhao, Y. (2007). Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum – Part 1: Experiments and large-scale features. *Climate of the Past*, *3*(2), 261–277. https://doi.org/10.5194/cp-3-261-2007

Braga, J. C., Martin, J. M., & Quesada, C. (2003). Patterns and average rates of late Neogene-Recent uplift of the Betic Cordillera, SE Spain. *Geomorphology*, *50*(1–3), 3–26. https://doi.org/10.1016/s0169-555x(02)00205-2

Camacho-Sánchez, M., Quintanilla, I., Hawkins, M. T. R., Tuh, F. Y. Y., Wells, K., Maldonado, J. E., & Leonard, J. A. (2018). Interglacial refugia on tropical mountains: Novel insights from the summit rat (*Rattus baluensis*), a Borneo mountain endemic. *Diversity and Distributions*, *24*(9), 1252–1266. https://doi.org/10.1111/ddi.12761

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). STACKS: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Cigliano, M. M., Braun, H., Eades, D. C., & Otte, D. (2020). *Orthoptera species file*. Version 5.0/5.0. Retrieved from http://orthoptera.speciesfile.org/

Clemente, M. E., Garcia, M. D., & Presa, J. J. (1991). Los Gomphocerinae de la península Iberica: 2. *Omocestus* Bolívar, 1878. (Insecta, Orthoptera, Caelifera). *Graellsia*, *46*, 191–246.

DeChaine, E. G., & Martin, A. P. (2005). Historical biogeography of two alpine butterflies in the Rocky Mountains: Broad-scale concordance and local-scale discordance. *Journal of Biogeography*, *32*(11), 1943–1956. https://doi.org/10.1111/j.1365-2699.2005.01356.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. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Flantua, S. G. A., Payne, D., Borregaard, M. K., Beierkuhnlein, C., Steinbauer, M. J., Dullinger, S., Essl, F., Irl, S. D. H., Kienle, D., Kreft, H., Lenzner, B., Norder, S. J., Rijsdijk, K. F., Rumpf, S. B., Weigelt, P., & Field, R. (2020). Snapshot isolation and isolation history challenge the analogy between mountains and islands used to understand endemism. *Global Ecology and Biogeography*, *29*(10), 1651–1673. https://doi.org/10.1111/geb.13155

Fordham, D. A., Brook, B. W., Moritz, C., & Nogués-Bravo, D. (2014). Better forecasts of range dynamics using genetic data. *Trends in Ecology & Evolution*, *29*(8), 436–443. https://doi.org/10.1016/j.tree.2014.05.007

Gómez, A., & Lunt, D. H. (2007). Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. In S. Weiss, & N. Ferrand (Eds.), *Phylogeography of southern European refugia* (pp. 155–188). Springer. https://doi.org/10.1007/1-4020-4904-8

González-Serna, M. J., Cordero, P. J., & Ortego, J. (2019). Spatiotemporally explicit demographic modelling supports a joint effect of historical barriers to dispersal and contemporary landscape composition on structuring genomic variation in a red-listed grasshopper. *Molecular Ecology*, *28*(9), 2155–2172. https://doi.org/10.1111/mec.15086

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

He, Q. X., Edwards, D. L., & Knowles, L. L. (2013). Integrative testing of how environments from the past to the present shape genetic structure across landscapes. *Evolution*, *67*(12), 3386–3402. https://doi.org/10.1111/evo.12159

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, *58*(3), 247–276. https://doi.org/10.1006/bijl.1996.0035

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

Hewitt, G. M. (2011). Mediterranean Peninsulas: The evolution of hotspots. In F. E. Zachos, & J. C. Habel (Eds.), *Biodiversity hotspots* (pp. 123–147). Springer.

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. https://doi.org/10.1002/joc.1276

Hughes, P. D., & Woodward, J. C. (2017). Quaternary glaciation in the Mediterranean mountains: A new synthesis. In P. D. Hughes, & J. C. Woodward (Eds.), *Quaternary glaciation in the Mediterranean Mountains* (Vol. *433*, pp. 1–23). Geological Soc Publishing House.

Jombart, T. (2008). *Adegenet*: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jouzel, J., Masson-Delmotte, V., Cattani, O., Dreyfus, G., Falourd, S., Hoffmann, G., Minster, B., Nouet, J., Barnola, J. M., Chappellaz, J., Fischer, H., Gallet, J. C., Johnsen, S., Leuenberger, M., Loulergue, L., Luethi, D., Oerter, H., Parrenin, F., Raisbeck, G., … Wolff, E. W. (2007). Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science*, *317*(5839), 793–796. https://doi.org/10.1126/science.1141038

Jueterbock, A., Smolina, I., Coyer, J. A., & Hoarau, G. (2016). The fate of the Arctic seaweed *Fucus distichus* under climate change: An ecological niche modeling approach. *Ecology and Evolution*, *6*(6), 1712–1724. https://doi.org/10.1002/ece3.2001

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Keightley, P. D., Ness, R. W., Halligan, D. L., & Haddrill, P. R. (2014). Estimation of the spontaneous mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib family. *Genetics*, *196*(1), 313–320. https://doi.org/10.1534/genetics.113.158758

Knowles, L. L. (2000). Tests of Pleistocene speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of western North America. *Evolution*, *54*(4), 1337–1348. https://doi.org/10.1111/j.0014-3820.2000.tb00566.x

Knowles, L. L., & Alvarado-Serrano, D. F. (2010). Exploring the population genetic consequences of the colonization process with spatio-temporally explicit models: Insights from coupled ecological, demographic and genetic models in montane grasshoppers. *Molecular Ecology*, *19*(17), 3727–3745. https://doi.org/10.1111/j.1365-294X.2010.04702.x

Knowles, L. L., & Massatti, R. (2017). Distributional shifts – Not geographic isolation – As a probable driver of montane species divergence. *Ecography*, *40*(12), 1475–1485. https://doi.org/10.1111/ecog.02893

Leuenberger, C., & Wegmann, D. (2010). Bayesian computation and model selection without likelihoods. *Genetics*, *184*(1), 243–252. https://doi.org/10.1534/genetics.109.109058

Lira-Noriega, A., & Manthey, J. D. (2014). Relationship of genetic diversity and niche centrality: A survey and analysis. *Evolution*, *68*(4), 1082–1093. https://doi.org/10.1111/evo.12343

Liu, C. R., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, *28*(3), 385–393. https://doi.org/10.1111/j.0906-7590.2005.03957.x

Liu, X. M., & Fu, Y. X. (2015). Exploring population size changes using SNP frequency spectra. *Nature Genetics*, *47*(5), 555–U172. https://doi.org/10.1038/ng.3254

Maier, P. A., Vandergast, A. G., Ostoja, S. M., Aguilar, A., & Bohonak, A. J. (2019). Pleistocene glacial cycles drove lineage diversification and fusion in the Yosemite toad (*Anaxyrus canorus*). *Evolution*, *73*(12), 2476–2496. https://doi.org/10.1111/evo.13868

Mairal, M., Sanmartín, I., Herrero, A., Pokorny, L., Vargas, P., Aldasoro, J. J., & Alarcón, M. (2017). Geographic barriers and Pleistocene climate change shaped patterns of genetic variation in the Eastern Afromontane biodiversity hotspot. *Scientific Reports*, *7*, 45749. https://doi.org/10.1038/srep45749

Massatti, R., & Knowles, L. L. (2016). Contrasting support for alternative models of genomic variation based on microhabitat preference: Species-specific effects of climate change in alpine sedges. *Molecular Ecology*, *25*(16), 3974–3986. https://doi.org/10.1111/mec.13735

McCormack, J. E., Bowen, B. S., & Smith, T. B. (2008). Integrating paleoecology and genetics of bird populations in two sky island archipelagos. *BMC Biology*, *6*, 28. https://doi.org/10.1186/1741-7007-6-28

Metcalf, J. L., Prost, S., Nogués-Bravo, D., DeChaine, E. G., Anderson, C., Batra, P., Araújo, M. B., Cooper, A., & Guralnick, R. P. (2014). Integrating multiple lines of evidence into historical biogeography hypothesis testing: A *Bison bison* case study. *Proceedings of the Royal Society B-Biological Sciences*, *281*(1777), 20132782. https://doi.org/10.1098/rspb.2013.2782

Meulenkamp, J. E., & Sissingh, W. (2003). Tertiary palaeogeography and tectonostratigraphic evolution of the Northern and Southern Peri-Tethys platforms and the intermediate domains of the

African-Eurasian convergent plate boundary zone. *Palaeogeography Palaeoclimatology Palaeoecology*, *196*(1–2), 209–228. https://doi.org/10.1016/s0031-0182(03)00319-5

Mevik, B. H., & Wehrens, R. (2007). The *pls* package: Principal component and partial least squares regression in R. *Journal of Statistical Software*, *18*(2), 1–23.

Mota, J. F., Pérez-García, F. J., Jiménez, M. L., Amate, J. J., & Peñas, J. (2002). Phytogeographical relationships among high mountain areas in the Baetic Ranges (South Spain). *Global Ecology and Biogeography*, *11*(6), 497–504. https://doi.org/10.1046/j.1466-822X.2002.00312.x

Mouret, V., Guillaumet, A., Cheylan, M., Pottier, G., Ferchaud, A. L., & Crochet, P. A. (2011). The legacy of ice ages in mountain species: Post-glacial colonization of mountain tops rather than current range fragmentation determines mitochondrial genetic diversity in an endemic Pyrenean rock lizard. *Journal of Biogeography*, *38*(9), 1717–1731. https://doi.org/10.1111/j.1365-2699.2011.02514.x

Nogués-Bravo, D. (2009). Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography*, *18*(5), 521–531. https://doi.org/10.1111/j.1466-8238.2009.00476.x

Papadopoulou, A., & Knowles, L. L. (2016). Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(29), 8018–8024. https://doi.org/10.1073/pnas.1601069113

Perrigo, A., Hoorn, C., & Antonelli, A. (2020). Why mountains matter for biodiversity. *Journal of Biogeography*, *47*(2), 315–325. https://doi.org/10.1111/jbi.13731

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS One*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

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

Presa, J. J., García, M., Clemente, M., Barranco Vega, P., Correas, J., Ferreira, S., Hochkirch, A., Lemos, P., Odé, B., & Prunier, F. (2016a). *Omocestus femoralis: The IUCN Red List of Threatened Species* (Publication no. https://dx.doi.org/ https://doi.org/10.2305/IUCN.UK.2016-3.RLTS.T16084447A75088712.en). Retrieved from https://www.iucnredlist.org/species/16084447/75088712

Presa, J. J., García, M., Clemente, M., Barranco Vega, P., Correas, J., Ferreira, S., Hochkirch, A., Lemos, P., Odé, B., & Prunier, F. (2016b). *Omocestus bolivari: The IUCN Red List of Threatened Species* (Publication no. https://dx.doi.org/ https://doi.org/10.2305/IUCN.UK.2016-3.RLTS.T16084452A75088604.en). Retrieved from https://www.iucnredlist.org/species/16084452/75088604

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Ramírez-Barahona, S., & Eguiarte, L. E. (2014). Changes in the distribution of cloud forests during the last glacial predict the patterns of genetic diversity and demographic history of the tree fern *Alsophila firma* (Cyatheaceae). *Journal of Biogeography*, *41*(12), 2396–2407. https://doi.org/10.1111/jbi.12396

Ray, N., Currat, M., Foll, M., & Excoffier, L. (2010). SPLATCHE2: A spatially explicit simulation framework for complex demography, genetic admixture and recombination. *Bioinformatics*, *26*(23), 2993–2994. https://doi.org/10.1093/bioinformatics/btq579

Rubidge, E. M., Patton, J. L., Lim, M., Burton, A. C., Brashares, J. S., & Moritz, C. (2012). Climate-induced range contraction drives genetic erosion in an alpine mammal. *Nature Climate Change*, *2*(4), 285–288. https://doi.org/10.1038/nclimate1415

Sandel, B., Arge, L., Dalsgaard, B., Davies, R. G., Gaston, K. J., Sutherland, W. J., & Svenning, J. C. (2011). The influence of Late Quaternary climate-change velocity on species endemism. *Science*, *334*(6056), 660–664. https://doi.org/10.1126/science.1210173

Schoville, S. D., Stuckey, M., & Roderick, G. K. (2011). Pleistocene origin and population history of a neoendemic alpine butterfly. *Molecular Ecology*, *20*(6), 1233–1247. https://doi.org/10.1111/j.1365-294X.2011.05003.x

Slatkin, M. (1993). Isolation by distance in equilibrium and nonequilibrium populations. *Evolution*, *47*(1), 264–279. https://doi.org/10.1111/j.1558-5646.1993.tb01215.x

Tonzo, V., Bellvert, A., & Ortego, J. (2021). Reticulate evolutionary history in a recent radiation of montane grasshoppers revealed by genomic data. *BioRxiv*. https://doi.org/10.1101/2021.01.12.426362

Tonzo, V., Papadopoulou, A., & Ortego, J. (2019). Genomic data reveal deep genetic structure but no support for current taxonomic designation in a grasshopper species complex. *Molecular Ecology*, *28*(17), 3869–3886. https://doi.org/10.1111/mec.15189

Wegmann, D., Leuenberger, C., Neuenschwander, S., & Excoffier, L. (2010). ABCTOOLBOX: A versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*, *11*, 116. https://doi.org/10.1186/1471-2105-11-116

BIOSKETCHES

Vanina Tonzo is a postdoctoral fellow at the Institute of Evolutionary Biology (IBE-CSIC, Barcelona, Spain). Her current research is primarily focused on the evolution and biogeography of arthropods, in particular exploring the potential impacts of climate change and orography on the genetic patterns of species using a population and comparative genomic approach.

Joaquín Ortego is a researcher at Estación Biológica de Doñana (EBD-CSIC, Seville, Spain) broadly interested in the study of speciation and the processes underlying spatial patterns of genetic variation in different organism groups. More information about the laboratory and ongoing research projects can be found at https://www.ortegolab.com/.

Authors contribution: V.T. and J.O. conceived and designed the study, analyses and collected the samples. V.T. performed the laboratory work and analysed the data guided by J.O. V.T. and J.O. wrote the manuscript.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Tonzo V, Ortego J. Glacial connectivity and current population fragmentation in sky islands explain the contemporary distribution of genomic variation in two narrow-endemic montane grasshoppers from a biodiversity hotspot. *Divers Distrib*. 2021;00:1–15. https://doi.org/10.1111/ddi.13306