

## Original Article

# Mountain speciation driven by high rates of lineage formation and rapid evolution of partial reproductive isolation: insights from a recent radiation of grasshoppers (Orthoptera: Gomphocerinae)

Joaquín Ortego<sup>1,\*</sup> and Víctor Nogueras<sup>2</sup>

<sup>1</sup>Department of Ecology and Evolution, Estación Biológica de Doñana, EBD-CSIC, Seville, Spain

<sup>2</sup>Departamento de Biología Animal, Edafología y Geología, Universidad de La Laguna, San Cristóbal de La Laguna, Tenerife, Canary Islands, Spain

\*Corresponding author. Department of Ecology and Evolution, Estación Biológica de Doñana, EBD-CSIC, Avda. Américo Vespucio 26, Seville E-41092, Spain. E-mail: [joaquin.ortego@csic.es](mailto:joaquin.ortego@csic.es)

## ABSTRACT

Mountains contribute disproportionately to global biodiversity, yet the ecological and evolutionary processes that make them important centres of speciation remain poorly understood. Here, we use a complex of Gomphocerinae grasshoppers distributed across the Pyrenees as a model system to investigate the process of speciation in mountain regions. Clustering analyses revealed a marked genetic fragmentation of populations and divergence time estimates indicated that all lineages split during the Pleistocene (<0.9 Mya). Although all taxa currently exhibit allopatric distributions, fine-scale temporal distribution modelling suggests that their ranges probably overlapped across extensive areas during glacial periods. Despite ample opportunities for historical hybridization, introgression tests indicate that gene flow was restricted to peripheral populations and only one population exhibited a high degree of admixed ancestry. Accordingly, our analyses reveal abrupt genetic discontinuities corresponding to the distributional boundaries of each taxon. Although taxa have probably not achieved complete reproductive isolation, semi-permeable species boundaries do not appear to have compromised their evolutionary independence. These results emphasize the interplay between high rates of lineage formation and the rapid evolution of reproductive barriers in promoting microgeographic speciation in topographically complex landscapes, shedding light on the proximate processes underlying the high levels of neo-endemism observed across mountain ranges worldwide.

**Keywords:** genetic admixture; hybridization; introgression; mountain ecosystems; Pleistocene speciation; species' delimitation; subspecies

## INTRODUCTION

Speciation hotspots—geographic areas where species of one or more taxonomic groups originate and accumulate at disproportionately high rates—represent unique natural laboratories for studying the microevolutionary processes underlying lineage formation and understanding how such lineages persist and ultimately give rise to full-fledged species (Davis *et al.* 2008). Mountains are recognized as both speciation and biodiversity hotspots, harbouring a high number of species compared to other ecosystems, many of which are narrow endemics of significant conservation concern (Wollenberg *et al.* 2008, Steinbauer *et al.*

2016, Noroozi *et al.* 2018, Rahbek *et al.* 2019). Despite their disproportionate contribution to global biodiversity, our understanding of the proximate ecological and evolutionary processes that make mountain regions key centres of speciation remains limited (Antonelli *et al.* 2018, Rahbek *et al.* 2019, Perrigo *et al.* 2020). Several abiotic factors have been proposed to contribute to mountain biodiversity, including the role of dispersal barriers in promoting population isolation and igniting lineage formation, and the presence of steep environmental gradients that provide opportunities for ecological speciation and create microclimatic refugia where species may persist under unfavourable environmental

Received 14 May 2025; revised 5 July 2025; accepted 26 August 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of The Linnean Society of London.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [reprints@oup.com](mailto:reprints@oup.com) for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com).

conditions (Sandel *et al.* 2011, Steinbauer *et al.* 2013, Steinbauer *et al.* 2016). Within this ecological context, reproductive isolation may evolve rapidly, reducing the potential for genetic exchange among lineages following secondary contact (Dynesius and Jansson 2014). Conversely, under scenarios of incomplete speciation, closely related lineages with adjacent distributions may merge (i.e. speciation reversal; Kearns *et al.* 2018) as a result of secondary contact during range expansions (Klicka and Zink 1997; e.g. Maier *et al.* 2019). Thus, the interplay between the particular physiography of mountain ecosystems, range dynamics driven by Pleistocene climate oscillations, species-specific life-history traits (e.g. dispersal capacity, generation time), and the pace at which reproductive isolation evolves is probably key to understanding the rates at which new lineages arise and how they persist through evolutionary time and progress toward speciation (Tzedakis 2009, Feliner 2014, Dynesius and Jansson 2014).

A major challenge in studying speciation in mountain systems lies in the difficulty of delineating independently evolving lineages within complexes of closely related taxa that are typically allopatric and for which reproductive isolation cannot be readily assessed in natural populations (Joshi *et al.* 2024). Numerous allopatric narrow-endemic species have been described in different groups of alpine and montane organisms based on subtle morphological differences and limited or no evidence of evolutionary independence (i.e. genetically and reproductively distinct populations that qualify as full-fledged species), raising concerns about their taxonomic validity [i.e. taxonomic inflation *sensu* Isaac *et al.* (2004), e.g. Pallarés *et al.* (2024)]. Moreover, the disruption of gene flow by topographic complexity in dispersal-limited alpine and montane species often results in marked genetic structures, leading to a pattern of population divergence that is confounded with speciation by current coalescent-based species' delimitation approaches (Sukumaran and Knowles 2017, Tonzo *et al.* 2019, Huang 2020). In this context, fine-scale population sampling can help evaluate whether proposed taxonomic units correspond to full-fledged species or instead reflect mere genetic discontinuities across the landscape shaped by historical or contemporary barriers to gene flow (e.g. Pyron *et al.* 2023, Burbrink *et al.* 2024, Pyron *et al.* 2024). This information, combined with reconstructions of past range dynamics and inferences about the timing and extent of historical hybridization (or its lack thereof) among putative species, offers a powerful framework for assessing the evolutionary independence of lineages that are currently allopatric (e.g. Ortego and Knowles 2022).

In this study, we used four closely related taxa from the complex *Chorthippus* (*Glyptobothrus*) superspecies *binotatus* (Defaut, 2011) (Orthoptera: Gomphocerinae), distributed across the Pyrenees, as a model system to investigate the processes promoting the formation and persistence of recently diverged lineages and to understand the high rates of local neo-endemism—i.e. recently formed, range-restricted species—that characterize mountain systems in the Mediterranean region (e.g. Tzedakis 2009, Feliner 2014). The complex includes two species—*C. binotatus* (Charpentier, 1825) and *C. saulcyi* (Krauss, 1888)—and three subspecific taxa—*C. saulcyi saulcyi* (Krauss, 1888), *C. saulcyi moralesi* Uvarov, 1954, and *C. saulcyi vicdessossi* Defaut, 2011—with adjacent distributions, making it an ideal system for exploring the mode and tempo of lineage formation along the 'grey zone' of the

speciation continuum (Roux *et al.* 2016). Although all taxa are currently allopatric, some peripheral populations are separated by very short geographical distances (<10 km), suggesting that secondary contact probably occurred during range shifts driven by Pleistocene climatic oscillations and provided opportunities for past hybridization. A preliminary genomic-based species' delimitation study supported the distinctiveness of all taxa within the complex (Noguerales *et al.* 2018). However, the limited number of populations analysed (one population per taxon), combined with strong genetic differentiation among populations within the complex (Noguerales *et al.* 2016), may have led the coalescent-based methods used for delineating species to conflate population genetic structure with species boundaries, potentially overestimating the number of taxa (Sukumaran and Knowles 2017). Furthermore, unresolved phylogenetic relationships among taxa suggest incomplete lineage sorting due to recent divergence and/or post-divergence gene flow, casting further doubt about their full-fledged species' status (Noguerales *et al.* 2018).

Here, we integrate fine-scale population genomic data with reconstructions of range dynamics through time to evaluate the status of the different taxa within the studied complex and to gain insights into the processes that have contributed to the formation and persistence of the lineages over evolutionary time. Specifically, we first reconstructed phylogenetic relationships among populations and performed clustering analyses to assess the status of putative taxa under both the phylogenetic and genotypic cluster species concepts, which are supported by the reciprocal monophyly of populations within taxa and by sharp genetic discontinuities corresponding with the boundaries of taxa distributions, respectively (De Queiroz 2007). Second, we tested for evidence of genetic admixture and/or introgression among the delineated taxa and related these genomic signals to reconstructions of historical range overlap based on lineage-specific distribution models projected into the past at fine temporal resolution. Historical range overlap in the absence of gene flow would indicate reproductive isolation and the completion of speciation, whereas signals of introgression combined with well-defined taxonomic boundaries (i.e. genetically and phenotypically cohesive population groups) would suggest that incomplete reproductive isolation has not compromised their evolutionary independence. Finally, we used simulation-based demographic modelling to compare alternative scenarios of species' formation, to estimate key demographic parameters, and to infer the pace and mode of speciation in the study system.

## MATERIALS AND METHODS

### Study area and sampling

The study area encompasses the Pyrenees Mountains and foothills in north-eastern Spain and southern France, where four taxa of the *Chorthippus* (*Glyptobothrus*) superspecies *binotatus* exhibit adjacent distributions (Cigliano *et al.* 2025). We sampled 28 localities (hereafter referred to as populations) within this species complex, comprising the taxa *C. binotatus* ( $N = 5$  populations, 39 individuals), *C. saulcyi moralesi* ( $N = 15$  populations, 98 individuals), *C. saulcyi saulcyi* ( $N = 5$  populations, 23 individuals), and *C. saulcyi vicdessossi* ( $N = 3$  populations, 20 individuals). Specimens of *C. reissingeri* Harz, 1972 were used as an outgroup in

phylogenomic analyses and *D*-statistic tests, whereas an extra-Pyrenean population of *C. binotatus* was used as a non-introgressed reference population in *D*-statistic tests (see below). Spatial coordinates were recorded using a Global Positioning System (GPS) device, and whole specimens were preserved at  $-20^{\circ}\text{C}$  in 1500  $\mu\text{L}$  of 96% ethanol until needed for genomic analyses. Further details on taxa, geographic location of populations, and number of genotyped individuals per population are provided in the [Supporting Information, Table S1](#).

### Genomic library preparation and processing

We extracted and purified DNA from each specimen using NucleoSpin Tissue kits (Macherey-Nagel, Dürren, Germany). Genomic libraries were prepared using the double-digest restriction-site-associated DNA sequencing (ddRAD-seq) protocol described by Peterson *et al.* (2012) and detailed in the [Supporting Information, Methods S1](#). Raw sequences were demultiplexed and preprocessed with STACKS v.1.35 (Catchen *et al.* 2013) and *de novo* assembled using PYRAD v.3.0.66 (Eaton 2014). Unless otherwise indicated, downstream analyses were performed using datasets of unlinked single nucleotide polymorphisms (SNPs) (i.e. one random SNP per RAD locus) obtained with PYRAD considering a clustering threshold of sequence similarity of 85% ( $W_{\text{clust}} = 0.85$ ) and discarding loci that were not present in at least 25%–50% of individuals ( $\text{minCov} = 0.25\text{--}0.50$ ). Further details on genomic data filtering and assembly are provided in the [Supporting Information, Methods S2](#).

### Phylogenomic analyses

We inferred phylogenetic relationships among populations using two independent approaches: the maximum-likelihood (ML) method implemented in RAXML v.8.2.12 (Stamatakis 2014) and the coalescent-based approach implemented in SVDQUARTETS (Chifman and Kubatko 2014). Eight individuals of *C. reissingeri* were used as the outgroup.

**RAXML**—We ran RAXML using a matrix of concatenated SNPs (all SNPs per locus; *.snp* file from PYRAD) edited using the R v.4.4.3 (R Core Team 2025) package ‘phrynomics’ (B. Banbury, <http://github.com/bbanbury/phrynomics>) to remove non-binary and invariant sites. This filtering step yielded an input dataset including 70789 SNPs. As in previous studies using concatenated SNP data, we considered a GTR-GAMMA model of nucleotide evolution (e.g. Kim *et al.* 2018, Barley *et al.* 2022). We applied an ascertainment bias correction using the conditional likelihood method, performed 100 rapid bootstrap replicates, and searched for the best-scoring ML tree.

**SVDQUARTETS**—We ran SVDQUARTETS analyses in PAUP\* v.4.0a169 (Swofford 2002) using a dataset of 3205 unlinked SNPs (*.usnp* file from PYRAD). To reduce computational demands, we selected the three individuals per population with the lowest proportion of missing data. We exhaustively evaluated all possible quartets and performed 100 non-parametric bootstrap replicates to assess topological support.

### Estimation of divergence time

The phylogenetic tree inferred using SVDQUARTETS was used as a fixed topology in A00 analyses in BPP v.4.7.0 (Flouri *et al.* 2018) to estimate posterior distributions of divergence times ( $\tau$ ). We

used a custom UNIX script (<https://github.com/OrtegoLab/ipyrad2bpp>) to edit and convert the *.loci* file from PYRAD into the BPP input format. Due to computational constraints, we limited the dataset to 1000 loci and one representative individual per population. We excluded from the analyses the admixed Montan de Tost (TOST) population (see Results). We set uniform rooted trees as species’ tree prior, applied an automatic adjustment of fine-tune parameters, set the diploid option to indicate that the input sequences are unphased, and adjusted inverse-gamma prior distributions for  $\theta$  ( $\alpha = 3, \beta = 0.04$ ) and  $\tau$  ( $\alpha = 3, \beta = 0.07$ ) according to empirical estimates calculated based on the number of segregating sites per site (Huang *et al.* 2020). Note that BPP analyses do not require a defining outgroup, as the program samples the root position along with the other nodes of the tree. In order to ensure the convergence of the runs ( $\text{ESS} > 200$ ), we ran two independent replicate analyses for 2000000 generations, sampling every two generations, after a burn-in of 100000 generations. We estimated divergence times using the formula  $\tau = 2\mu t$ , where  $\tau$  is the divergence in substitutions per site estimated by BPP,  $\mu$  is the per-site mutation rate per generation, and  $t$  is the absolute divergence time in years (Walsh 2001). We considered the mutation rate per site per generation of  $2.8 \times 10^{-9}$  estimated for *Drosophila melanogaster* (Keightley *et al.* 2014), which is similar to the spontaneous mutation rate estimated for the butterfly *Heliconius mel-pomene* ( $2.9 \times 10^{-9}$ ; Keightley *et al.* 2015).

### Testing for introgression

We used four-taxon ABBA/BABA tests based on the *D*-statistic to evaluate whether the different populations of a given taxon have experienced genetic introgression from other taxa (Durand *et al.* 2011). Assuming that the sister-taxon P1 and P2 diverged from P3 and an outgroup species O, the *D*-statistic is used to test the null hypothesis of no introgression ( $D = 0$ ) between P3 and P1 or P2. *D*-values significantly different from zero indicate gene flow between P1 and P3 ( $D < 0$ ) or between P2 and P3 ( $D > 0$ ). Tests were designed to evaluate introgression from one taxon (P3) into each population of another taxon (P1). To this end, for both P2 and P3 we used reference populations with no signatures of genetic introgression from other taxa based on inferences from exhaustive *D*-statistic tests. In the case of *C. binotatus*, we used as reference an extra-Pyrenean conspecific population from Sierra Espuña, south-eastern Spain ([Supporting Information, Table S1](#)). *Chorthippus reissingeri* was used as the outgroup (P4) in all analyses ([Supporting Information, Table S1](#)). Due to the limited distribution range of *C. saulcyi vicdessossi* and the difficulty in establishing a non-introgressed reference population for this taxon, *D*-statistic tests were exclusively focused on comparisons involving *C. binotatus*, *C. saulcyi moralesi*, and *C. saulcyi saulcyi* and pairs of taxa with adjacent or partially overlapping distributional ranges.

### Genetic diversity, structure, and admixture

We calculated haplotype diversity ( $H_d$ ) and nucleotide diversity ( $\pi$ ) for each population in DNASP v.6.12.03 (Rozas *et al.* 2017). We used the Bayesian Markov chain Monte Carlo clustering method implemented in the program STRUCTURE v.2.3.3 (Pritchard *et al.* 2000) to quantify genetic structure and admixture across populations of the four focal taxa ([Supporting Information,](#)



Table S1). To fully explore population genetic subdivision and admixture, we initially analysed the data from all taxa and populations jointly and, subsequently, we ran independent analyses focused on populations assigned to each taxon (e.g. Ortego *et al.* 2024). We ran STRUCTURE analyses assuming correlated allele frequencies and admixture and without using prior population information. The number of genetic clusters ( $K$ ) considered ranged from  $K = 1$  to  $K = 10$  in analyses including all taxa and from  $K = 1$  to  $K = N + 1$  for each dataset of  $N$  populations in taxon-specific analyses. We conducted 15 independent runs for each  $K$  value to estimate the most likely number of genetic clusters with 200000 MCMC cycles, following a burn-in step of 100000 iterations. We retained the 10 runs having the highest likelihood for each value of  $K$  and determined the number of genetic clusters that best describes our data according to log probabilities of the data [ $\text{LnPr}(X|K)$ ; Pritchard *et al.* 2000] and the  $\Delta K$  method (Evanno *et al.* 2005), as implemented in STRUCTURE HARVESTER (Earl and vonHoldt 2012). We used CLUMPP v.1.1.2 and the Greedy algorithm to align multiple runs of STRUCTURE for the same  $K$  value (Jakobsson and Rosenberg 2007) and DISTRUCT v.1.1 (Rosenberg 2004) to visualize the individuals' probabilities of cluster membership as bar plots. In addition, we performed principal component analyses (PCAs) as implemented in the R package 'ade4' (Jombart 2008). Before running PCAs, we replaced missing data by the mean allele frequency of the corresponding locus estimated across all samples (Jombart 2008). Both STRUCTURE and PCA analyses were performed with datasets only including SNPs with a minor allele frequency (MAF) >1% (for details on datasets, see Supporting Information, Table S2).

### Mode and timing of speciation and admixture

D-statistic tests and Bayesian clustering analyses in STRUCTURE indicated that the putative population of *C. binotatus* from Montan de Tost (TOST) was admixed, with approximately 25% of its ancestry originating from the eastern lineage of *C. saulcyi moralesi* (see Results). Based on this finding, we used FASTSIMCOAL2 (Excoffier *et al.* 2021) to infer the divergence mode and to estimate the timing of the split between *C. binotatus* and *C. saulcyi moralesi*, as well as the secondary contact event that gave rise to the admixed TOST population. We tested four competing demographic models, including scenarios of (i) divergence in strict isolation followed by an admixture event (Model A), (ii) divergence in strict isolation followed by an admixture event and contemporary gene flow (Model B), (iii) divergence with gene flow followed by an admixture event and no contemporary gene flow (Model C), and (iv) divergence with gene flow followed by an admixture event and contemporary gene flow (Model D) (Supporting Information, Fig. S1). For these analyses, we included the admixed TOST population and selected one representative population for each parental taxon: MONT for *C. binotatus* and ERRM for *C. saulcyi moralesi*. These two parental populations were selected because they were assigned to the same genetic clusters than the admixed TOST population at high-level hierarchical analyses in STRUCTURE ( $K = 7$ ) and had no signatures of genetic introgression from other taxa (see Results). Further details on FASTSIMCOAL2 analyses and model selection are presented in the Supporting Information, Methods S3.

### Comparing taxa's environmental space

We used available occurrence records and a comprehensive suite of environmental variables to test for differences in environmental space among taxa. For each record, we extracted (i) elevation from the 90-m-resolution digital elevation model from NASA Shuttle Radar Topographic Mission (SRTM Digital Elevation Data; <http://srtm.csi.cgiar.org/>), and (ii) 19 bioclimatic variables (Supporting Information, Table S3) from the CHELSA database (<http://chelsa-climate.org/bioclim/>; Karger *et al.* 2017). To summarize variation and to reduce strong redundancy among variables, we ran a principal component analysis (PCA) with VARIMAX rotation and obtained scores for the first three principal components (eigenvalues > 1), which cumulatively explained 88.75% of the variance (PC1: 43.65%; PC2: 37.32%; PC3: 10.77%; see factor loadings in Supporting Information, Table S4). Then, we performed a MANOVA to compare elevation and bioclimatic conditions (PC1, PC2, and PC3 scores) among the four taxa and used post hoc Tukey tests to examine differences in environmental variables between each pair of taxa. All statistical analyses were performed in SPSS v.26 (IBM, NY, USA).

### Environmental niche modelling

We built environmental niche models (ENMs) to reconstruct the distribution of climatically suitable areas for the four focal taxa from the last glacial maximum (LGM, c. 22 ka) to present. We used this information to infer distributional shifts experienced by each taxon in response to Quaternary climatic oscillations and (i) determine across different time-periods the extent of climatically suitable areas and patterns of population fragmentation/connectivity resulted from the spatial configuration of hypothetical barriers to dispersal and (ii) the temporal dynamics of range overlap between each pair of taxa. To build ENMs, we used the maximum entropy algorithm implemented in MAXENT v.3.4.1 (Phillips *et al.* 2006, Phillips and Dudik 2008), occurrence records available for each taxon, and the 19 bioclimatic layers (30-arcsec resolution) from the CHELSA database (<http://chelsa-climate.org/bioclim/>; Karger *et al.* 2017). To estimate environmental suitability from the LGM to the present, and dynamics of range overlap between each pair of taxa, we projected taxon-specific ENMs to bioclimatic conditions during the last 22000 years at 100-year time intervals (i.e. from CE1990 to the LGM, for a total of 220 snapshots) using bioclimatic layers available at a high resolution (30-arcsec) from the CHELSA-TraCE21k v.1.0 database (<https://chelsa-climate.org/>; Karger *et al.* 2023). Further details on ENM are presented in the Supporting Information, Methods S4.

## RESULTS

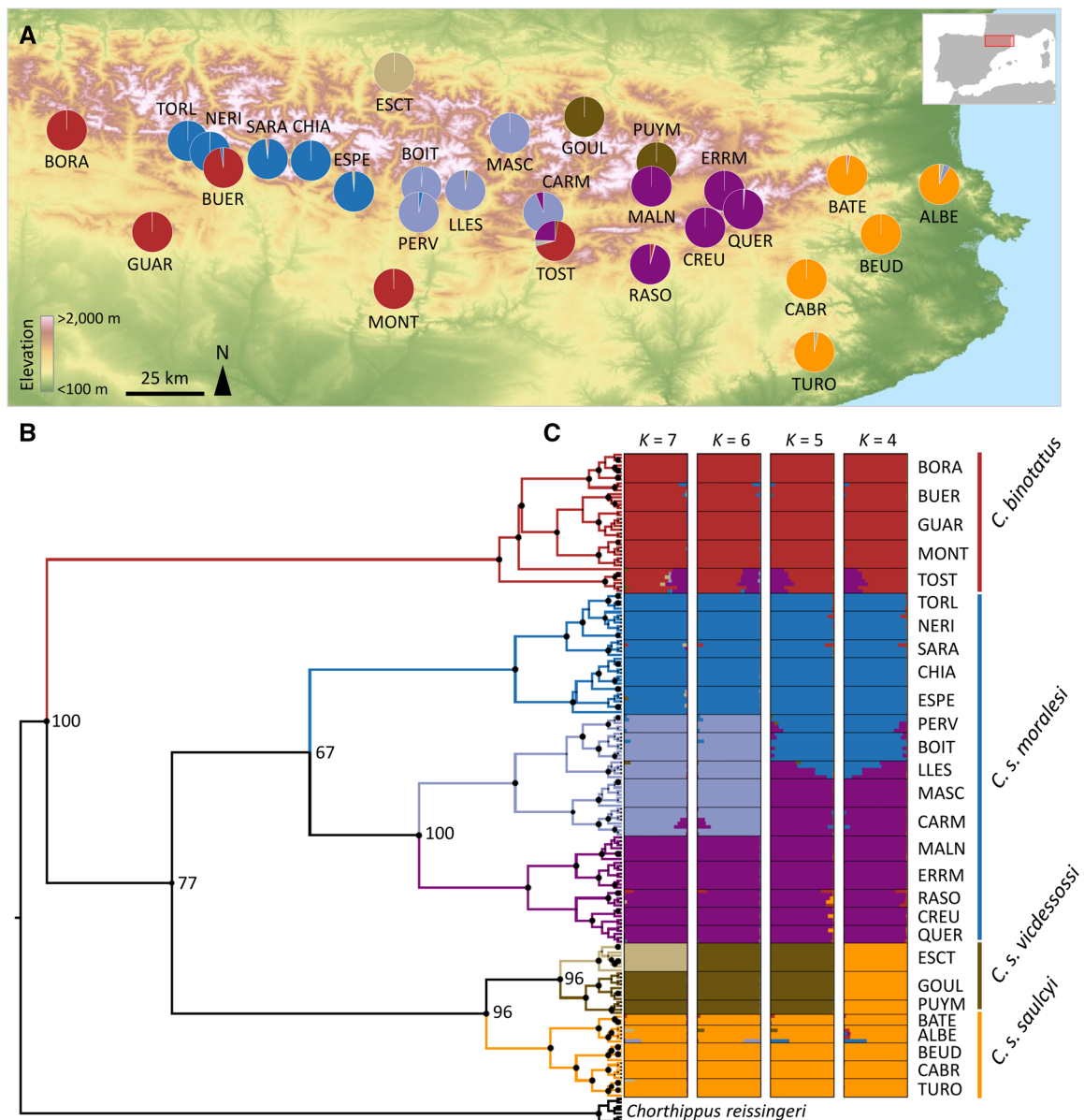
### Genomic datasets

After filtering loci, the final dataset, including all individuals ( $\text{min-Cov} = 25\%$ ), retained 8885 SNPs. Taxon-specific datasets ( $\text{min-Cov} = 50\%$ ) retained 6847 SNPs for *C. binotatus*, 3839 SNPs for *C. saulcyi moralesi*, 8100 SNPs for *C. saulcyi viduessosi*, and 7844 SNPs for *C. saulcyi moralesi*. Other attributes of the genomic datasets used in each specific analysis are presented in the Supporting Information, Table S2.

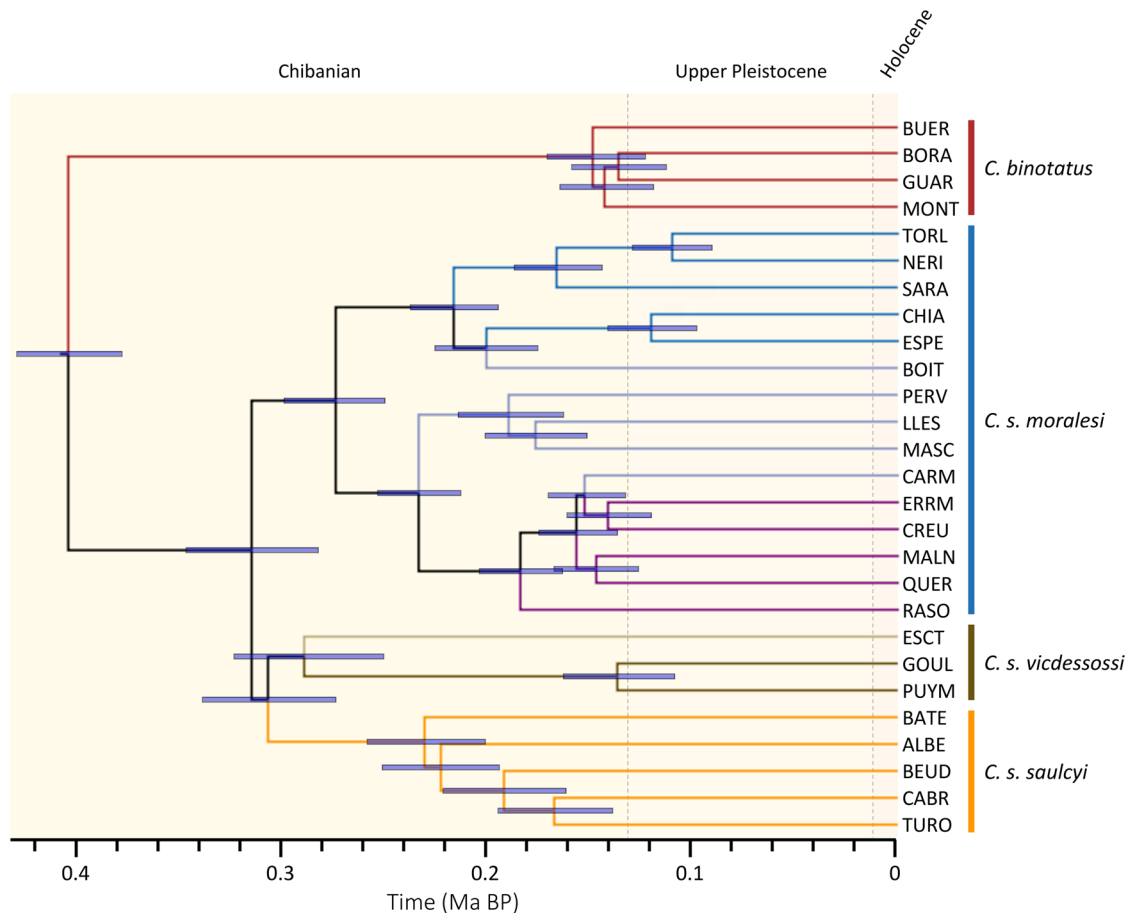
### Phylogenomic analyses and estimation of divergence time

Phylogenomic reconstructions using RAXML and SVDQUARTETS revealed two main clades: one comprising all populations of *C. binotatus* and another including the three subspecific taxa within *C. saulcyi* (Fig. 1; Supporting Information, Fig. S2). The *C. saulcyi* clade was structured into two sister-subclades: one including all populations of *C. saulcyi moralesi*, and another comprising *C. saulcyi vicdessossi* and *C. saulcyi saulcyi* (Supporting Information, Figs 1 and S2). Within *C. saulcyi moralesi*, three lineages were identified, distributed across the western, central, and eastern Pyrenees (Fig. 1; Supporting Information, Fig. S2). Nodes separating the

main clades were poorly supported in SVDQUARTETS' analyses, with bootstrap support (BSS) values <50% (Supporting Information, Fig. S2). In contrast, RAXML analyses based on concatenated SNP data showed strong support for the nodes separating *C. binotatus* from *C. saulcyi*, and *C. saulcyi vicdessossi* from *C. saulcyi saulcyi* (Fig. 1). Divergence-time estimates in BPP suggested that *C. binotatus* and *C. saulcyi* diverged c. 0.4 Mya, while the three subspecies within *C. saulcyi* split c. 0.3 Mya (Chibanian period; Fig. 2). Note that these estimates assume no post-divergence gene flow (i.e. a strictly bifurcating model), and may thus underestimate actual divergence times (see FASTSIMCOAL2 analyses below).



**Figure 1.** (A, C) results of genetic assignments based on STRUCTURE and (B) phylogenetic relationships inferred with RAXML for the four taxa of the *Chorthippus binotatus* species complex from the Pyrenees Mountains. A, pie charts on map show the geographic location of genotyped populations and their respective genetic assignments for  $K = 7$ . C, In barplots, each individual is represented by a horizontal bar partitioned into  $K$  coloured segments showing the individual's probability of belonging to the cluster with that colour; thin horizontal black lines separate individuals from different populations. B, In phylogenetic tree, black dots on nodes indicate bootstrap support (BSS) values >0.95%; estimated BSS values are indicated on the right side of nodes separating main clades. Population codes as described in Supporting Information, Table S1.



**Figure 2.** Population tree inferred with SVDQUARTETS and divergence times estimated using BPP (analysis A00). Bars on nodes indicate 95% highest posterior densities (HPD) of divergence times estimated considering a genomic mutation rate of  $2.8 \times 10^{-9}$  per site per generation and a one-year generation time. Background colours indicate geological divisions of the Quaternary. Population codes as described in [Supporting Information, Table S1](#).

### Testing for introgression

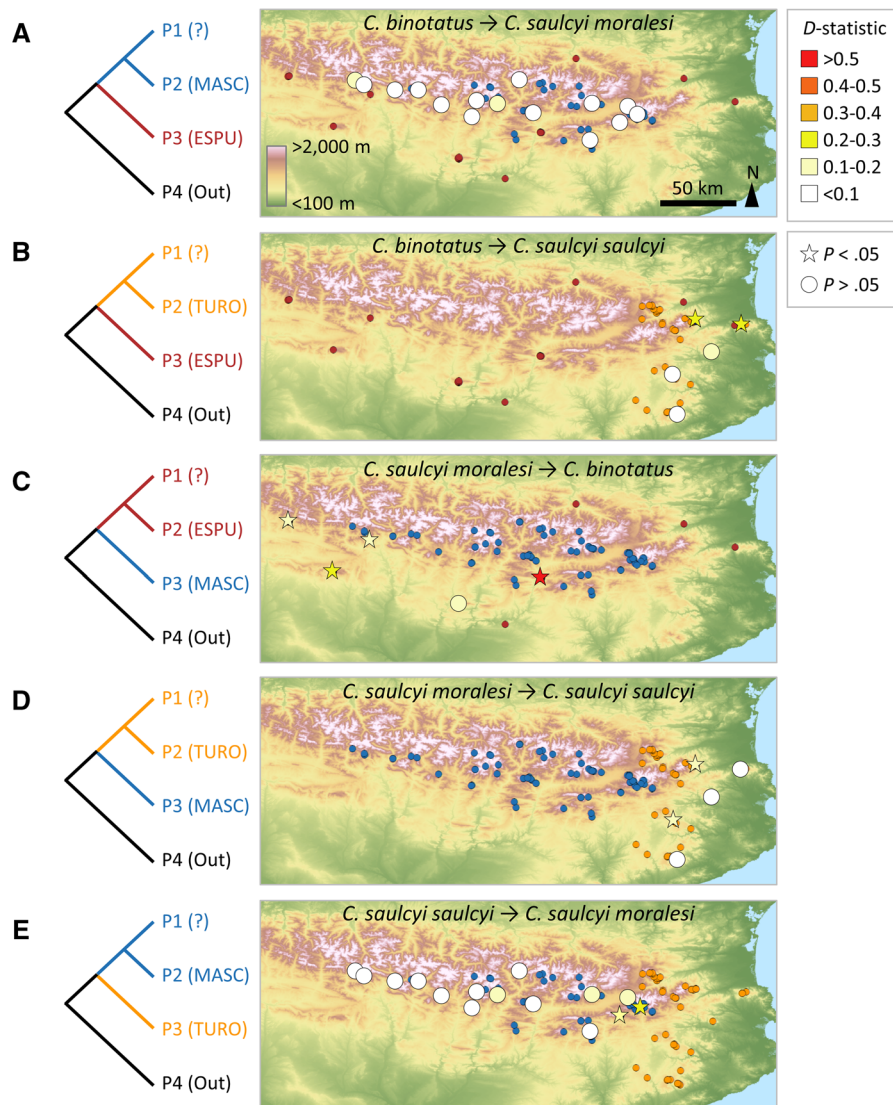
D-statistic tests revealed significant genetic introgression (i) from *C. saulcyi moralesi* into most populations of *C. binotatus* (BORA, BUER, GUAR, and TOST) and into the two westernmost populations of *C. saulcyi saulcyi* (BATE and CABR); (ii) from *C. saulcyi saulcyi* into the two easternmost populations of *C. saulcyi moralesi* (CREU and QUER); and (iii) from *C. binotatus* into the two northernmost populations of *C. saulcyi saulcyi* (BATE and ALBE) (Fig. 3; [Supporting Information, Table S5](#)). However, D-statistic tests did not show significant genetic introgression from *C. binotatus* into any population of *C. saulcyi moralesi* (Fig. 3; [Supporting Information, Table S5](#)).

### Genetic diversity, structure, and admixture

STRUCTURE analyses, including all populations and taxa, showed that  $\Delta K$  peaked at  $K = 2$  and had a secondary peak at  $K = 7$ , while  $\text{LnPr}(X|K)$  reached an asymptote at  $K = 7-8$  ([Supporting Information, Fig. S3A](#)). Genetic clusters inferred at  $K = 2-3$  were inconsistent with phylogenomic inferences and current taxonomic assignments (see [Supporting Information, Fig. S4](#)), which might be a consequence of uneven population sampling across taxa/lineages ([Puechmaille 2016](#)) and/or reflect the difficulties of STRUCTURE to identify higher level clustering schemes and to

assign individuals from complex genetic histories into too few clusters ([Kalinowski 2011](#), [Funk et al. 2020](#)). From  $K = 4$  to  $K = 7$ , STRUCTURE analyses showed a hierarchical split among taxa and within-taxon lineages consistent with phylogenomic inferences (Fig. 1). The population from TOST was identified as admixed, with approximately 25% of its ancestry deriving from the eastern lineage of *C. saulcyi moralesi* (Fig. 1). Genetic diversity of this admixed population was significantly higher than in parental populations of *C. binotatus* (one-sample  $t$ -tests, Hd:  $t = -4.43$ ,  $P = .011$ ;  $\pi$ :  $t = -4.07$ ,  $P = 0.027$ ) and *C. saulcyi moralesi* (one-sample  $t$ -tests, Hd:  $t = -9.39$ ,  $P < .001$ ;  $\pi$ :  $t = -10.74$ ,  $P < .001$ ), and was the highest among all analysed populations ([Supporting Information, Table S1](#)). STRUCTURE analyses showed that all other populations presented limited signatures of genetic admixture with other taxa (Fig. 1). Analyses restricted to individual taxa (excluding TOST for *C. binotatus*) revealed strong genetic structure within each, with most populations forming distinct clusters at higher  $K$  values ([Supporting Information, Figs S3B-E, S5](#)). Admixture among genetic clusters within taxa was absent or limited, mostly involving nearby populations of the more densely sampled taxon *C. saulcyi moralesi*. Principal component analyses (PCA) of genetic variation showed a genetic clustering of populations similar to that inferred by STRUCTURE analyses





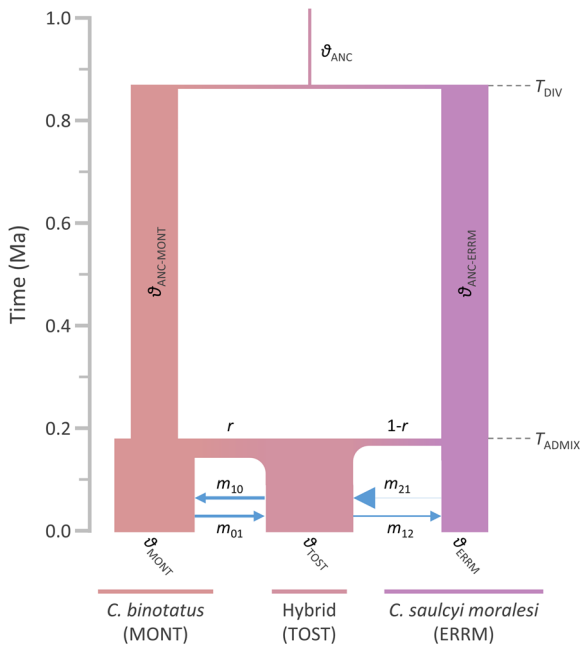
**Figure 3.** Analyses of introgression based on four-taxon  $D$ -statistic (ABBA/BABA) tests. Analyses were designed to test for introgression from one taxon (P3) into each population of the focal taxon (P1; circles and stars on maps). For both P2 and P3, we used reference populations with no signatures of genetic introgression from other taxa. *Chorthippus reissingeri* was used as an outgroup (P4) in all analyses. Stars indicate populations presenting significant ( $P < .05$ ) genetic introgression. Colours of circles and stars show the absolute value of the inferred  $D$ -statistic estimate for each focal population. Background dots represent occurrence records of taxa involved in each comparison, as a proxy of their current distribution. Statistics for all analyses are presented in [Supporting Information, Table S3](#). Population codes as described in [Supporting Information, Table S1](#).

([Supporting Information, Figs S6, S7](#)). Remarkably, in the PCA including all taxa ([Supporting Information, Fig. S6](#)), the population ESCT of *C. saulcyi vicdessossi* was largely differentiated along PC1 from the rest of populations. This population harboured the lowest level of genetic diversity across all analysed populations (one-sample  $t$ -tests, Hd:  $t = 12.03$ ,  $P < .001$ ;  $\pi$ :  $t = 10.93$ ,  $P < .001$ ; [Supporting Information, Table S1](#)), indicating that it has probably experienced strong genetic drift.

#### Mode and timing of speciation and admixture

The best-fitting model was a scenario in which *C. binotatus* and *C. saulcyi moralesi* diverged in strict isolation followed by an admixture event that led to the formation of the hybrid population from TOST and contemporary gene flow (Model B; [Table 1](#); [Fig. 4](#)).

Alternative models received lower support ( $\Delta\text{AIC} > 2$ ; [Table 1](#)). Considering one generation per year, the split between *C. binotatus* and *C. saulcyi moralesi* was estimated to take place during the early Pleistocene (c. 0.866 Mya; Calabrian age; [Table 2](#); [Fig. 4](#)). The admixture event that led to the formation of the hybrid population of TOST was estimated to have occurred during the penultimate glacial period (c. 182 kya; [Table 2](#); [Fig. 4](#)). In line with inferences from STRUCTURE (see Results), c. 73% of the ancestry of TOST was estimated to have originated from *C. binotatus* and 27% from *C. saulcyi moralesi* ([Table 2](#); [Fig. 4](#)). Contemporary effective migration rates per generation between the admixed population from TOST and *C. binotatus* was estimated to be more than five-fold higher than between TOST and *C. saulcyi moralesi* ([Table 2](#); [Fig. 4](#)).



**Figure 4.** Schematic of the best-supported model (i.e. divergence in strict isolation followed by an admixture event and contemporary gene flow; see Table 1) for divergence of *Chorthippus binotatus* and *C. saulcyi moralesi* and the event leading to the formation of the admixed population from Montan de Tost (TOST). Parameters include mutation-scaled ancestral ( $\theta_{ANC}$ ,  $\theta_{ANC-MONT}$  and  $\theta_{ANC-ERRM}$ ) and contemporary ( $\theta_{MONT}$ ,  $\theta_{TOST}$  and  $\theta_{ERRM}$ ) effective population sizes (proportional to thickness of vertical bars), effective migration rates per generation ( $m_{01}$ ,  $m_{10}$ ,  $m_{12}$  and  $m_{21}$ , proportional to arrow thickness), the proportion of lineages transferred from parental lineages to the admixed population of TOST ( $r$  and  $1-r$ , proportional to thickness of horizontal bars), and the timing (scale bar on left) of divergence ( $T_{DIV}$ ) and admixture ( $T_{ADMIX}$ ). Migration rates ( $m$ ) are presented as  $2Nm$  and estimates of time are given in years, assuming one generation per year. Contemporary effective population size of *C. binotatus* ( $\theta_{MONT}$ ) was calculated from levels of nucleotide diversity ( $\pi$ ) and fixed in FASTSIMCOAL2 analyses to enable estimation of other parameters (see Materials and methods for further details). Demographic parameter values and confidence intervals are detailed in Table 2. Population codes as described in Supporting Information, Table S1.

### Comparing taxa's environmental space

MANOVA revealed significant differences among taxa in the environmental spaces they occupy (Wilk's  $\lambda = 0.214$ ,  $F_{12,387} = 25.52$ ,  $P < .001$ ; Fig. 5). One-way ANOVAs showed significant differences among taxa in all analysed variables (elevation:  $F_{3,149} = 37.00$ ;  $P < 0.001$ ; bioclimatic PC1:  $F_{3,149} = 9.39$ ;  $P < .001$ ; bioclimatic PC2:  $F_{3,149} = 39.76$ ;  $P < .001$ ; bioclimatic PC3:  $F_{3,149} = 42.54$ ;  $P < .001$ ; Fig. 5). Post hoc tests indicated significant pairwise differences in most cases, especially between *C. binotatus* and any subspecies of *C. saulcyi* (Supporting Information, Table S6).

### Environmental niche modelling

Climatically suitable areas predicted by taxon-specific ENMs are highly congruent with their respective present-day observed distributions (Supporting Information, Figs S8–S11). Currently suitable areas correspond to the Pyrenean and pre-Pyrenean

**Table 1.** Alternative models of divergence and admixture tested using FASTSIMCOAL2.

| Model         | lnL             | k         | AIC            | $\Delta AIC$ | $\omega_i$  |
|---------------|-----------------|-----------|----------------|--------------|-------------|
| ModelA        | –1600.18        | 8         | 3216.35        | 13.30        | 0.00        |
| <b>ModelB</b> | <b>–1591.52</b> | <b>10</b> | <b>3203.05</b> | <b>0.00</b>  | <b>0.76</b> |
| ModelC        | –1595.51        | 9         | 3209.02        | 5.97         | 0.04        |
| ModelD        | –1591.87        | 11        | 3205.74        | 2.69         | 0.20        |

Models include scenarios of divergence in strict isolation followed by an admixture event (Model A), divergence in strict isolation followed by an admixture event and contemporary gene flow (Model B), divergence with gene flow followed by an admixture event and interruption of contemporary gene flow (Model C), and divergence with gene flow followed by an admixture event and contemporary gene flow (Model D) (see Supporting Information, Fig. S1). The best-supported scenario is highlighted in bold and depicted in Figure 4. The site frequency spectrum used for FASTSIMCOAL2 analyses contained 1014 variable SNPs. lnL, maximum likelihood value of the model; k, number of parameters in the model; AIC, Akaike's information criterion value;  $\Delta AIC$ , difference in AIC value from that of the strongest model;  $\omega_i$ , AIC weight.

**Table 2.** Parameters inferred from coalescent simulations with FASTSIMCOAL2 under the most-supported scenario of divergence of *Chorthippus binotatus* and *C. saulcyi moralesi* and the event leading to the formation of the admixed population from TOST (Model B).

| Parameter           | Point estimate | Lower bound | Upper bound |
|---------------------|----------------|-------------|-------------|
| $\theta_{ANC}$      | 29536          | 66853       | 98526       |
| $\theta_{ANC-MONT}$ | 377466         | 358680      | 398651      |
| $\theta_{ANC-ERRM}$ | 456830         | 435055      | 474562      |
| $\theta_{MONT}$     | 774140         |             |             |
| $\theta_{TOST}$     | 846828         | 810570      | 878188      |
| $\theta_{ERRM}$     | 448919         | 426702      | 461465      |
| $m_{01}$            | 0.799          | 0.729       | 0.795       |
| $m_{10}$            | 0.874          | 0.779       | 0.856       |
| $m_{12}$            | 0.165          | 0.128       | 0.156       |
| $m_{21}$            | 0.087          | 0.067       | 0.081       |
| $r$                 | 0.727          | 0.712       | 0.732       |
| $T_{DIV}$           | 865855         | 783967      | 849916      |
| $T_{ADMIX}$         | 181610         | 173107      | 188102      |

Table shows point estimates and lower and upper 95% confidence intervals for each parameter, which include the mutation-scaled ancestral ( $\theta_{ANC}$ ,  $\theta_{ANC-MONT}$  and  $\theta_{ANC-ERRM}$ ) and contemporary effective population sizes (*Chorthippus binotatus*:  $\theta_{MONT}$ ; *C. saulcyi moralesi*:  $\theta_{ERRM}$ ; admixed population:  $\theta_{TOST}$ ), effective migration rates per generation ( $m_{01}$ ,  $m_{10}$ ,  $m_{12}$  and  $m_{21}$ ), the proportion of lineages transferred from *C. binotatus* to the admixed population of TOST ( $r$ ), and the timing of divergence ( $T_{DIV}$ ) and admixture ( $T_{ADMIX}$ ). Migration rates ( $m$ ) are presented as  $2Nm$  and estimates of time are given in years, assuming one generation per year. The model is illustrated in Figure 4. Contemporary effective population size of *C. binotatus* ( $\theta_{BIN}$ ) was calculated from its levels of nucleotide diversity ( $\pi$ ) and fixed in FASTSIMCOAL2 analyses to enable the estimation of all other parameters (see Materials and methods for further details).

foothills for *C. binotatus* (Supporting Information, Fig. S8B), sub-alpine-montane habitats in the central Pyrenees for *C. saulcyi moralesi* (Supporting Information, Fig. S9B), the northern central Pyrenean foothills for *C. saulcyi vicedessossi* (Supporting Information, Fig. S10B), and the eastern Pyrenean foothills and pre-Pyrenean mountains for *C. saulcyi saulcyi* (Supporting Information, Fig. S11B). Details on the performance and parameters of taxon-specific models are presented in Supporting Information, Table S7. Projections of taxon-specific ENMs under bioclimatic conditions from the LGM to the present, at 100-year intervals, indicate that all taxa underwent downslope range shifts and expansions during the last glacial period. However, the inferred range



dynamics were taxon-specific. The extent and connectivity of suitable areas for the Mediterranean-montane *C. binotatus* (Supporting Information, Fig. S8A) and *C. saulcyi saulcyi* (Supporting Information, Fig. S11A) peaked around the LGM c. 22–18 kya) and then declined abruptly from the end of the last glacial period until the onset of the Holocene (c. 18–10 kya), when both taxa experienced a gradual and slight range expansion until the present day (Supporting Information, Figs S8a and S11a). In contrast, the extent of suitable areas for the subalpine-montane *C. saulcyi moralesi* (Supporting Information, Fig. S9A) and *C. saulcyi videsossi* (Supporting Information, Fig. S10A) reached a minimum around the LGM (c. 20–18 kya), expanded abruptly c. 18 kya ago, and then progressively decreased from c. 15 kya ago until the present day (Supporting Information, Figs S9A, S10A). Consistent with these predictions, the extent of range overlap between taxa peaked at the end of the last glacial period (c. 22–18 kya for *C. saulcyi moralesi*–*C. binotatus* and c. 18–15 kya for the remaining species pairs) but remained very limited from the onset of the Holocene to the present day (Fig. 6).

## DISCUSSION

Our comprehensive suite of analyses on a complex of recently diverged (<0.9 Mya) grasshoppers distributed across the Pyrenees has revealed that the formation and persistence of narrow endemic lineages in mountain systems probably results from: (i) population isolation and divergence driven by the interplay among limited dispersal capacity, topographically complex landscapes, and range dynamics driven by Pleistocene climatic oscillations (i.e. igniters of lineage formation); and (ii) long-term persistence thanks to the continuous presence of climatically suitable refugia and short speciation times that facilitate recently formed lineages became reproductively isolated (i.e. lineage persistence) (Dynesius and Jansson 2014). Despite all taxa within the studied complex currently present adjacent distributions and had ample opportunity for hybridization during historical periods of secondary contact, clustering analyses showed abrupt genetic discontinuities coinciding with the distributional limits of each putative taxon, and introgression tests revealed that historical hybridization was restricted to boundary populations. These findings support the taxonomic status of the two currently recognized species and the genotypic distinctiveness of the three narrowly distributed subspecific lineages, emphasizing the role of Mediterranean mountains as key centres of speciation and hotspots of local endemism.

### Species' delimitation under incomplete lineage sorting and introgression

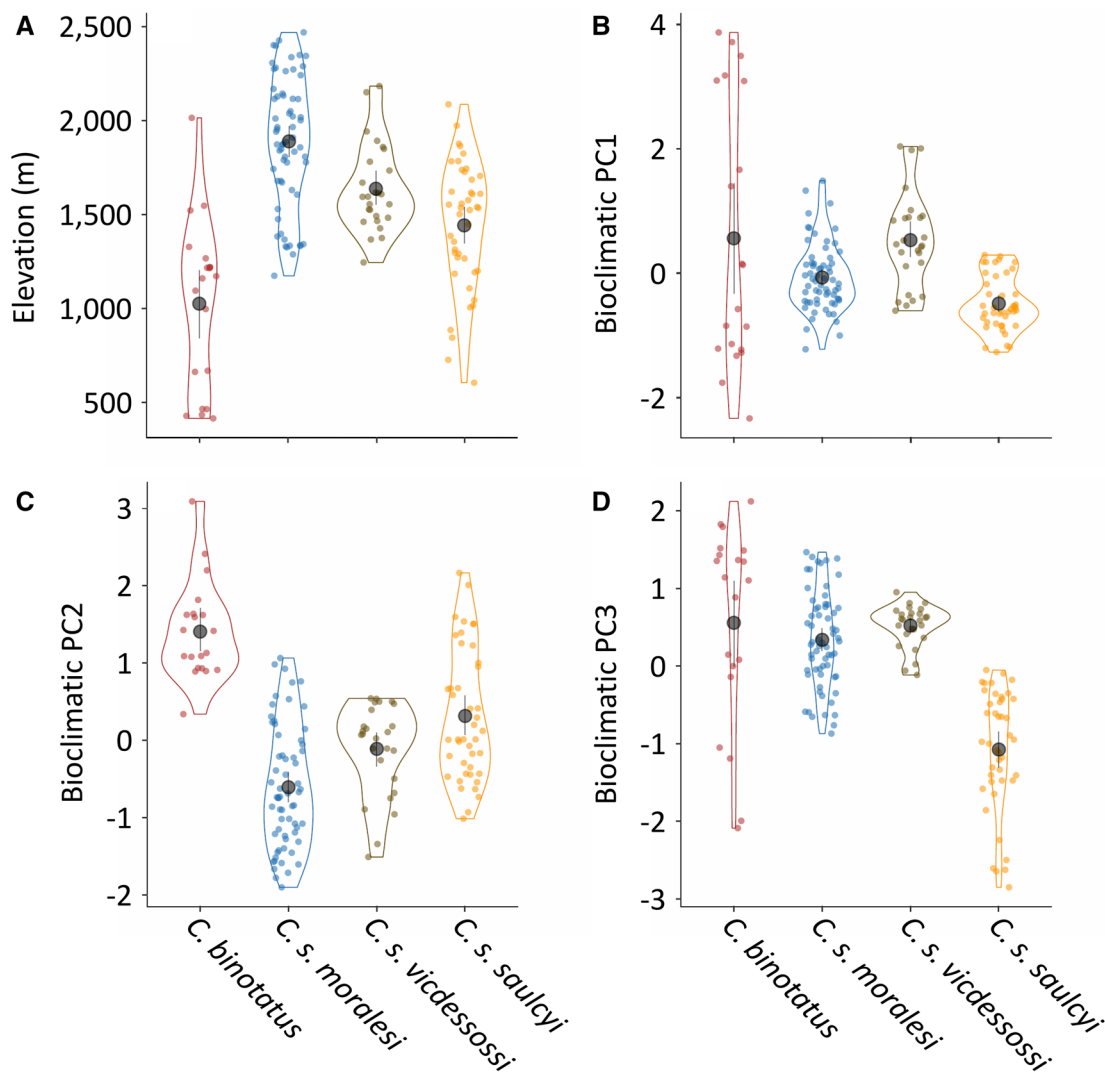
Phylogenomic analyses revealed two main clades corresponding to the currently recognized species *C. binotatus* and *C. saulcyi*, with subclades aligning with the three subspecific taxa within *C. saulcyi* and the phylogeographic breaks reported in a previous microsatellite-based study for *C. saulcyi moralesi* (Noguerales et al. 2016). However, coalescent-based phylogenetic analyses showed weak support for most internal nodes (BSS < 60%; Supporting Information, Fig. S1; see also: Noguerales et al. 2018), probably reflecting both historical hybridization (see Results) and incomplete lineage sorting due to recent divergence (Carstens and Knowles

2007). Phylogenetic dating under a strictly bifurcating tree (Fig. 2) and FASTSIMCOAL2 simulations accounting for post-divergence gene flow (Fig. 4) both estimated recent divergence of all taxa (<0.9 Mya). This may explain the weak phylogenetic support for main lineages, which is probably driven by incomplete lineage sorting (Carstens and Knowles 2007), along with other factors such as short DNA sequences and limited phylogenetically informative characters. Although low nodal support does not fulfil the reciprocal monophyly criterion required to define independently evolving lineages (Knowles and Carstens 2007), different lines of evidence support the distinctiveness of the all taxa under the genotypic cluster species concept (Mallet 1995). All putative taxa, whether specific or subspecific, showed sharp genetic discontinuities aligned with their respective distributional boundaries, supporting that the inferred high-level hierarchical genetic clusters do not correspond with genetic clines resulted from within-species differentiation (e.g. isolation by distance; e.g. Pyron et al. 2023, Ortego et al. 2024, Pyron et al. 2024). Except for the admixed population of TOST (see Results), we found little evidence of admixture among hetero(sub)specific populations, supporting that all taxa are separated by well-defined genetic boundaries.

### Admixture and genetic introgression

Introgression analyses based on the *D*-statistic revealed that historical gene flow among taxa was limited and exclusively involved boundary populations (Fig. 3). Although all taxa are currently allopatric, inferences from taxon-specific ENMs suggest that their distributional ranges largely overlapped during glacial periods, probably leading to secondary contact and allowing for hybridization. Comparisons of spatial patterns of genetic introgression and inferred range shifts of *C. binotatus* and *C. saulcyi moralesi*—the two taxa with broader distributions—provide significant insights into the proximate processes underlying historical hybridization dynamics. Despite their current adjacent ranges, *C. saulcyi moralesi* typically occupies higher elevations in the central Pyrenees, while *C. binotatus* is mainly found at lower elevations in the foothills and pre-Pyrenean ranges (Fig. 5A; Supporting Information, Figs S8, S9). Whereas *D*-statistic tests showed that most populations of *C. binotatus* are introgressed with *C. saulcyi moralesi*, there was no evidence for significant introgression in the opposite direction, indicating that historical hybridization between the two taxa was asymmetric. This pattern aligns with inferences from taxon-specific ENM, which showed that populations of *C. saulcyi moralesi* probably expanded downslope into the range of *C. binotatus* during glacial periods, while *C. binotatus* has never colonized the subalpine and montane zones currently occupied by *C. saulcyi moralesi* in the central Pyrenees (Supporting Information, Figs S8, S9).

Among all genotyped populations, only TOST showed a high admixed ancestry resulting from historical hybridization between *C. binotatus* and one of the lineages of *C. saulcyi moralesi* (Fig. 1). FASTSIMCOAL2 analyses showed that *C. binotatus* and *C. saulcyi moralesi* probably split c. 0.9 Mya (Calabrian age), whereas the admixture event that led to the formation of the hybrid population of TOST was estimated to happen during the penultimate glacial period c. 182 kya (Late Pleistocene; Fig. 4). Unlike other



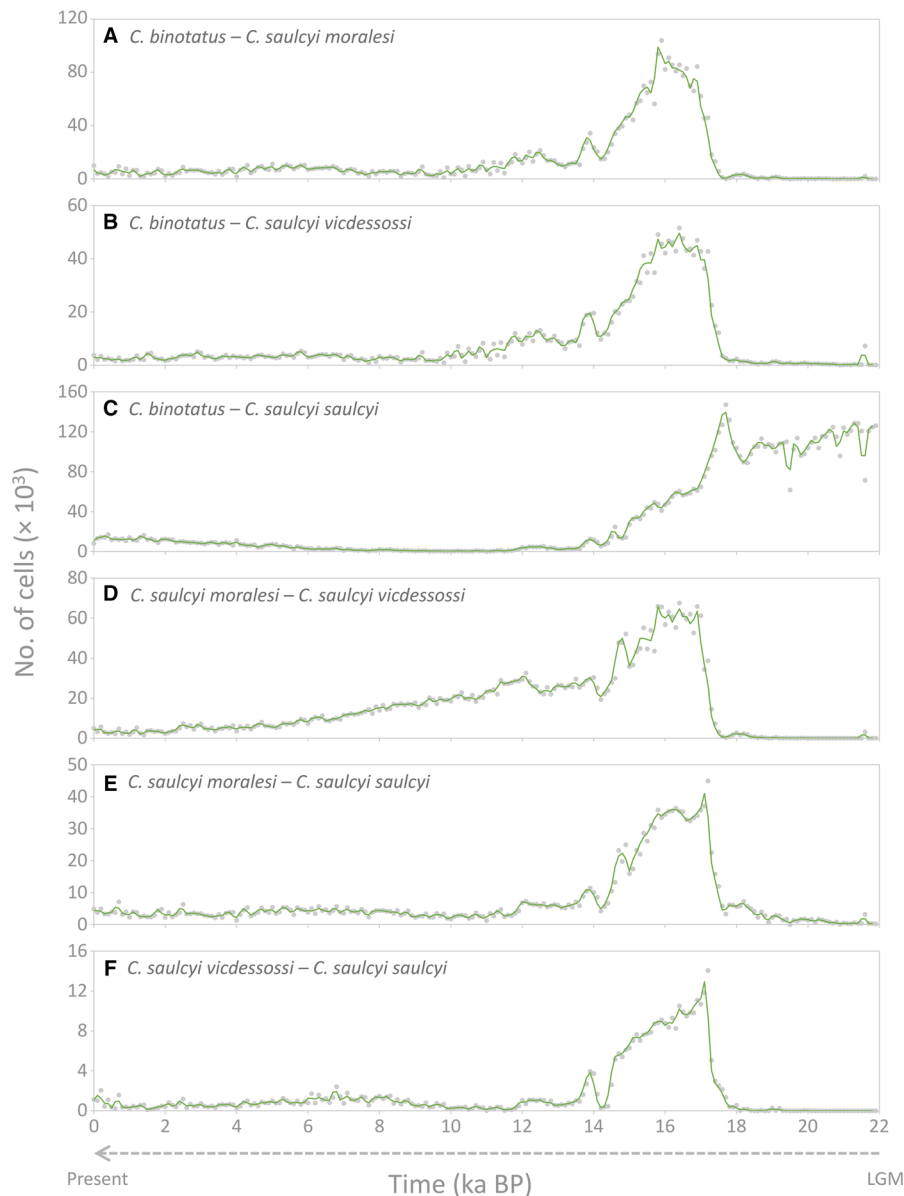
**Figure 5.** (A) elevation and (B, C) scores of a principal component analyses (PCA) on bioclimatic variables (PC1, PC2, and PC3) for occurrence sites of the four taxa of the *Chorthippus binotatus* species complex from the Pyrenees Mountains. Violin plots show values for occurrence sites (small coloured dots) and mean and confidence intervals (black dots and vertical bars, respectively) for each taxon.

populations with subtle signatures of genetic introgression—imperceptible in Bayesian clustering analyses—the high level of admixed ancestry in TOST probably reflects a stable hybrid swarm originating from parental populations that contributed substantial portions of their genomes. However, despite its high admixed ancestry, the two parental species did not contribute equally to the genome of this hybrid population. Coalescent-based analyses estimated that c. 73% and 27% of the ancestry of TOST population originated from *C. binotatus* and *C. saulcyi moralesi*, respectively (Fig. 4), which is in line with inferences from STRUCTURE analyses (Fig. 1) and consistent with the taxonomic assignment of this population to *C. binotatus* (Llucià-Pomares 2002). The population of TOST is established at 1190 m.a.s.l., in an area with a Mediterranean vegetation (*Quercus ilex*, *Buxus sempervirens*, and *Thymus* sp.), characteristic of the typical habitats occupied by *C. binotatus* in the Pyrenean foothills. Notably, gene flow between the admixed population of TOST and *C. binotatus* was estimated to be five times higher than with *C. saulcyi moralesi*, suggesting that ecological, genetic, and/or phenotypic similarities may have facilitated

subsequent gene flow with the taxon that contributed most to its genome (Fig. 4). Collectively, the phenotypic and ecological similarity of the admixed TOST population to *C. binotatus* illustrates how even substantial levels of admixture do not necessarily lead to detectable evolutionary outcomes (e.g. via phenotypic assimilation by one of the parental species; Huang 2016) and may often go unnoticed without genomic data.

#### Drivers of microgeographic speciation

Multiple lines of evidence support that the formation of narrow endemic taxa in the studied complex was probably a consequence of high rates of lineage formation coupled with rapid evolution of reproductive isolation (i.e. short-speciation times) that allowed their persistence through evolutionary time (Dynesius and Jansson 2014). Bayesian clustering analyses revealed strong genetic structure within all taxa, with populations often forming discrete clusters and showing limited admixture (Supporting Information, Figs S5, S7). This structure probably reflects differentiation at contrasting spatiotemporal scales, from the disruption of gene



**Figure 6.** Extent of range overlap between each pair of taxa as inferred from projections of taxon-specific environmental niche models (ENM) to bioclimatic conditions during the last 22000 years (i.e. from CE1990 to the LGM) at 100-year time intervals. The degree of overlap in the distribution of each pair of taxa at each time interval was calculated as the number of cells where the two focal taxa were predicted to be present according to the maximum training sensitivity plus specificity (MTSS) logistic threshold for occurrence of their respective taxon-specific ENM.

flow among nearby populations separated by dispersal barriers to the formation of phylogeographic lineages resulting from historical range fragmentation. Factors contributing to genetic isolation and lineage formation include limited dispersal capacity of most grasshoppers (e.g. *Voje et al. 2009*, *González-Serna et al. 2019*, *Ortego et al. 2021*), the high topographic complexity of the region (i.e. landscape barriers to dispersal; e.g. *Noguerales et al. 2016*, *Valbuena-Ureña et al. 2018*, *Tonzo et al. 2019*), and population fragmentation resulting from distributional shifts due to Pleistocene climatic oscillations (*Supporting Information, Figs S8–S11*). These processes are well-exemplified in *C. saulcyi moralesi*, which shows a marked population genetic structure at short geographical scales (<10 km; *Supporting Information, Fig. SSB*) and a major east–west phylogeographic break (*Fig. 1B*; see also: *Noguerales*

*et al. 2016*). This transverse break (*sensu Wallis et al. 2016*), also reported in other co-distributed organisms (*Charrier et al. 2004*, *Valbuena-Ureña et al. 2018*, *Tonzo et al. 2019*), has been attributed to range fragmentation when glaciers covered extensive areas from the central part of the Pyrenees during the coldest stages of the Pleistocene (*Ehlers et al. 2011*).

The persistence of these recently formed lineages at such fine spatial scales is probably driven by short speciation times and the rapid evolution of reproductive barriers, which is instrumental in preventing lineage fusion and speciation reversal during periods of secondary contact (*Fig. 6*; e.g. *Maier et al. 2019*). Accordingly, dating analyses suggest that divergence occurred <0.9 Mya, within the range of speciation times reported for sister-lineages of Orthoptera that are co-distributed and demonstrably reproductively



isolated (e.g. Ortego *et al.* 2021, 2024). Given the allopatric distributions of the studied taxa, speciation was most likely driven by the progressive accumulation of reproductive isolation due to strong genetic drift in small populations during extended periods of geographical isolation (e.g. Kalirad *et al.* 2024). This can include the development of both pre-zygotic (e.g. ecological isolation, gamete incompatibility, and/or divergence in internal genitalia and courtship behaviour) and post-zygotic (e.g. hybrid inviability or sterility; i.e. Dobzhansky–Muller incompatibilities) reproductive barriers, which can quickly accumulate in isolated populations (Coyne and Orr 1989, Bailey *et al.* 2004, Presgraves 2010, Langenhans *et al.* 2016, Hagberg *et al.* 2022). Although geographical isolation is the most parsimonious driver of divergence at early stages of speciation, ecological divergence might have also played an important role on the formation and/or maintenance of species boundaries in the studied complex. For instance, whereas *C. binotatus* exclusively feeds on scrub legumes, *C. saulcyi* is predominantly granivore, a contrasting feeding regime that might have triggered diversification through niche partitioning (Noguerales *et al.* 2018, Noguerales and Ortego 2022). Likewise, although the presence of impassable barriers to dispersal (i.e. Pyrenees mountains) seems to be the main driver explaining the allopatric distribution of some taxa, ecological divergence might have also played an important role in the reproductive isolation between the Mediterranean *C. binotatus* and the subalpine-montane *C. saulcyi moralesi*, which occupy different habitats and elevational ranges, despite some of their respective populations being separated by <7 km (Fig. 1A). Future genotype–environment association analyses and the investigation of lineage-specific selection signatures in particular genomic regions may help to clarify the potential role of selection in driving speciation and ecological adaptation in the studied complex (e.g. Soria-Carrasco *et al.* 2014, Chase *et al.* 2021). Altogether, although introgression and admixture analyses suggest that taxa might have not achieved complete reproductive isolation, historical gene flow has been limited and spatially restricted, indicating that semi-permeable boundaries have not compromised their genetic distinctiveness and independent evolutionary trajectories (e.g. Ortego and Knowles 2022, Ortego *et al.* 2024).

## CONCLUSIONS

Collectively, our findings emphasize the interplay between high rates of lineage formation and rapid evolution of partial reproductive isolation in promoting microgeographic speciation in topographically complex landscapes. These results highlight the potential of integrative, multidisciplinary approaches to understanding spatial heterogeneity of speciation rates and to illuminate the mechanisms that make mountain systems important hotspots of speciation and centres of endemism.

## ACKNOWLEDGEMENTS

We are grateful to P.J. Cordero, A. Papadopoulou, and V. Tonzo for their help during field work, A. Hidalgo-Galiana for genomic library preparation, and two anonymous referees for their

constructive and valuable comments on an earlier version of the manuscript. Logistic support was provided by Laboratorio de Ecología Molecular (LEM-EBD) from Estación Biológica de Doñana (CSIC). We also thank Centro de Supercomputación de Galicia (CESGA) and Doñana's Singular Scientific-Technical Infrastructure (ICTS-RBD) for access to computer resources. Sampling permits were granted by Gobierno de Aragón (2014: INAGA/500201/24/2014/02984; 2015: INAGA/500201/24/2015/3894; 2016: INAGA/500201/24/2016/2959; 2017: INAGA/500201/24/2017/03131; 2018: INAGA/500201/24/2018/759; 2019: INAGA/500201/24/2019/01203), Generalitat de Catalunya (2015: SF/590-593; 2016: SF/567-570, SF/577-578; 2017: SF/565, SF/657-662; 2018: SF/0068-0073; 2019: SF/0030/2019), Parque Nacional de Ordesa y Monte Perdido (2016: S20160232665; 2018: S20180135013; 2019: S2019 0223469), Parque Nacional de Aigüestortes y Estany de Sant Maurici (2015: PM-19/15), Parc Natural de l'Alt Pirineu (2018: 6/2018; 2019: 6/2019), Parc Natural del Cadí-Moixeró (2018: 22/2018; 2019: 19/2019), Parc Natural de les Capçaleres del Ter i del Freser (2019: SF/0030/2019), and Parc National des Pyrénées (2016: 2016-55; 2018: 2018-10; 2019: 2019-85).

## AUTHOR CONTRIBUTIONS

Joaquín Ortego (conceptualization, methodology, formal analysis, investigation, writing—original draft preparation, writing—review and editing, funding acquisition) and Víctor Noguerales (conceptualization, methodology, investigation, writing—original draft preparation, writing—review and editing)

## SUPPLEMENTARY DATA

Supplementary data is available at *Zoological Journal of the Linnean Society* online.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## FUNDING

This work was supported by grant TED2021-129328B-I00 funded by MCIN/AEI/10.13039/501100011033 and European Union NextGenerationEU/PRTR and grant PID2021-123298NB-I00 funded by MCIN/AEI/10.13039/501100011033/FEDER, UE. Víctor Noguerales was supported by a Viera y Clavijo postdoctoral fellowship funded by the Agencia Canaria de Investigación, Innovación y Sociedad and the Universidad de La Laguna.

## DATA AVAILABILITY

Raw Illumina reads have been deposited at the NCBI Sequence Read Archive (SRA) under BioProject PRJNA702631. Input files for all analyses (RAXML, SVDQUARTETS, STRUCTURE, PCAs, DNASP, FASTSIMCOAL2, and ENM) are available for download on Figshare (<https://doi.org/10.6084/m9.figshare.29046218>).

## REFERENCES

- Antonelli A, Kissling WD, Flantua SGA *et al.* Geological and climatic influences on mountain biodiversity. *Nature Geoscience* 2018;**11**:718–25. <https://doi.org/10.1038/s41561-018-0236-z>
- Bailey RI, Thomas CD, Butlin RK. Premating barriers to gene exchange and their implications for the structure of a mosaic hybrid zone between *Chorthippus brunneus* and *C. jacobsi* (Orthoptera: Acrididae). *Journal of Evolutionary Biology* 2004;**17**:108–19. <https://doi.org/10.1046/j.1420-9101.2003.00648.x>
- Barley AJ, Cordes JE, Walker JM *et al.* Genetic diversity and the origins of parthenogenesis in the teiid lizard *Aspidoscelis laredoensis*. *Molecular Ecology* 2022;**31**:266–78. <https://doi.org/10.1111/mec.16213>
- Burbrink FT, Myers EA, Pyron RA. Understanding species limits through the formation of phylogeographic lineages. *Ecology and Evolution* 2024;**14**:e70263. <https://doi.org/10.1002/ece3.70263>
- Carstens BC, Knowles LL. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. *Systematic Biology* 2007;**56**:400–11. <https://doi.org/10.1080/10635150701405560>
- Catchen J, Hohenlohe PA, Bassham S *et al.* STACKS: an analysis tool set for population genomics. *Molecular Ecology* 2013;**22**:3124–40. <https://doi.org/10.1111/mec.12354>
- Chase MA, Ellegren H, Mugal CF. Positive selection plays a major role in shaping signatures of differentiation across the genomic landscape of two independent *Ficedula* flycatcher species pairs. *Evolution; International Journal of Organic Evolution* 2021;**75**:2179–96. <https://doi.org/10.1111/evo.14234>
- Cigliano MM, Braun H, Eades DC, Otte D. *Orthoptera Species File (OSF)*, 2025. Available at <http://orthoptera.speciesfile.org>. (7 March 2025, date last accessed).
- Coyne JA, Orr HA. Patterns of speciation in *Drosophila*. *Evolution; International Journal of Organic Evolution* 1989;**43**:362–81. <https://doi.org/10.2307/2409213>
- Charrier O, Dupont P, Pornon A *et al.* Microsatellite marker analysis reveals the complex phylogeographic history of *Rhododendron ferrugineum* (Ericaceae) in the Pyrenees. *PLoS One* 2014;**9**:e92976. <https://doi.org/10.1371/journal.pone.0092976>
- Chifman J, Kubatko L. Quartet inference from snp data under the coalescent model. *Bioinformatics (Oxford, England)* 2014;**30**:3317–24. <https://doi.org/10.1093/bioinformatics/btu530>
- Davis EB, Koo MS, Conroy C *et al.* The California hotspots project: identifying regions of rapid diversification of mammals. *Molecular Ecology* 2008;**17**:120–38. <https://doi.org/10.1111/j.1365-294X.2007.03469.x>
- De Queiroz K. Species concepts and species delimitation. *Systematic Biology* 2007;**56**:879–86. <https://doi.org/10.1080/10635150701701083>
- Durand EY, Patterson N, Reich D *et al.* Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution* 2011;**28**:2239–52. <https://doi.org/10.1093/molbev/msr048>
- Dynesius M, Jansson R. Persistence of within-species lineages: a neglected control of speciation rates. *Evolution; International Journal of Organic Evolution* 2014;**68**:923–34. <https://doi.org/10.1111/evo.12316>
- Earl DA, vonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing structure output and implementing the evanno method. *Conservation Genetics Resources* 2012;**4**:359–61. <https://doi.org/10.1007/s12686-011-9548-7>
- Eaton DAR. PYRAD: assembly of *de novo* radseq loci for phylogenetic analyses. *Bioinformatics (Oxford, England)* 2014;**30**:1844–9. <https://doi.org/10.1093/bioinformatics/btu121>
- Ehlers J, Gibbard PL, Hugues PD. *Quaternary Glaciations—Extent and Chronology: A Closer Look*, 1st edn. Amsterdam, the Netherlands: Elsevier, 2011.
- Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 2005;**14**:2611–20. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Marchi N, Marques DA *et al.* FASTSIMCOAL2: demographic inference under complex evolutionary scenarios. *Bioinformatics (Oxford, England)* 2021;**37**:4882–5. <https://doi.org/10.1093/bioinformatics/btab468>
- Feliner GN. Patterns and processes in plant phylogeography in the Mediterranean basin. A review. *Perspectives in Plant Ecology Evolution and Systematics* 2014;**16**:265–78. <https://doi.org/10.1016/j.ppees.2014.07.002>
- Flouri T, Jiao XY, Rannala B *et al.* Species tree inference with bpp using genomic sequences and the multispecies coalescent. *Molecular Biology and Evolution* 2018;**35**:2585–93. <https://doi.org/10.1093/molbev/msy147>
- Funk SM, Guedaoura S, Juras R *et al.* Major inconsistencies of inferred population genetic structure estimated in a large set of domestic horse breeds using microsatellites. *Ecology and Evolution* 2020;**10**:4261–79. <https://doi.org/10.1002/ece3.6195>
- González-Serna MJ, Cordero PJ, Ortego J. 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* 2019;**28**:2155–72. <https://doi.org/10.1111/mec.15086>
- Hagberg L, Celemin E, Irisarri I *et al.* Extensive introgression at late stages of species formation: insights from grasshopper hybrid zones. *Molecular Ecology* 2022;**31**:2384–99. <https://doi.org/10.1111/mec.16406>
- Huang JP. Parapatric genetic introgression and phenotypic assimilation: testing conditions for introgression between Hercules beetles (Dynastes, Dynastinae). *Molecular Ecology* 2016;**25**:5513–26. <https://doi.org/10.1111/mec.13849>
- Huang JP. Is population subdivision different from speciation? From phylogeography to species delimitation. *Ecology and Evolution* 2020;**10**:6890–6. <https://doi.org/10.1002/ece3.6524>
- Huang JP, Hill JG, Ortego J *et al.* Paraphyletic species no more—genomic data resolve a Pleistocene radiation and validate morphological species of the *Melanoplus scudderi* complex (insecta: Orthoptera). *Systematic Entomology* 2020;**45**:594–605. <https://doi.org/10.1111/syen.12415>
- Isaac NJB, Mallet J, Mace GM. Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology & Evolution* 2004;**19**:464–9. <https://doi.org/10.1016/j.tree.2004.06.004>
- Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics (Oxford, England)* 2007;**23**:1801–6. <https://doi.org/10.1093/bioinformatics/btm233>
- Jombart T. *Adegenet*: an R package for the multivariate analysis of genetic markers. *Bioinformatics (Oxford, England)* 2008;**24**:1403–5. <https://doi.org/10.1093/bioinformatics/btn129>
- Joshi M, Espeland M, Huemer P *et al.* Species delimitation under allopatry: genomic insights within and across continents in Lepidoptera. *Insect Systematics and Diversity* 2024;**8**:ixae027. <https://doi.org/10.1093/isd/ixae027>
- Kalinowski ST. The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity* 2011;**106**:625–32. <https://doi.org/10.1038/hdy.2010.95>
- Kalirad A, Burch CL, Azevedo RBR. Genetic drift promotes and recombination hinders speciation on holey fitness landscapes. *PLoS Genetics* 2024;**20**:e1011126. <https://doi.org/10.1371/journal.pgen.1011126>
- Karger DN, Conrad O, Bohner J *et al.* Climatologies at high resolution for the earth's land surface areas. *Scientific Data* 2017;**4**:170122. <https://doi.org/10.1038/sdata.2017.122>
- Karger DN, Nobis MP, Normand S *et al.* Chelsa-trace21k-high-resolution (1 km) downscaled transient temperature and precipitation data since the last glacial maximum. *Climate of the Past* 2023;**19**:439–56. <https://doi.org/10.5194/cp-19-439-2023>
- Keightley PD, Ness RW, Halligan DL *et al.* Estimation of the spontaneous mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib family. *Genetics* 2014;**196**:313–20. <https://doi.org/10.1534/genetics.113.158758>

- Keightley PD, Pinharanda A, Ness RW *et al.* Estimation of the spontaneous mutation rate in *Heliconius melpomene*. *Molecular Biology and Evolution* 2015;**32**:239–43. <https://doi.org/10.1093/molbev/msu302>
- Kim BY, Wei XZ, Fitz-Gibbon S *et al.* RADseq data reveal ancient, but not pervasive, introgression between Californian tree and scrub oak species (*Quercus* sect. *Quercus*: Fagaceae). *Molecular Ecology* 2018;**27**:4556–71. <https://doi.org/10.1111/mec.14869>
- Klicka J, Zink RM. The importance of recent ice ages in speciation: a failed paradigm. *Science* 1997;**277**:1666–9. <https://doi.org/10.1126/science.277.5332.1666>
- Knowles LL, Carstens BC. Delimiting species without monophyletic gene trees. *Systematic Biology* 2007;**56**:887–95. <https://doi.org/10.1080/10635150701701091>
- Langerhans RB, Anderson CM, Heinen-Kay JL. Causes and consequences of genital evolution. *Integrative and Comparative Biology* 2016;**56**:741–51. <https://doi.org/10.1093/icb/icw101>
- Llucià-Pomares D. Revision of the Orthoptera (Insecta) of Catalonia (Spain). *Monografias SEA* 2002;**7**:1–226.
- Maier PA, Vandergast AG, Ostojia SM *et al.* Pleistocene glacial cycles drove lineage diversification and fusion in the Yosemite toad (*Anaxyrus canorus*). *Evolution; International Journal of Organic Evolution* 2019;**73**:2476–96. <https://doi.org/10.1111/evo.13868>
- Mallet J. A species definition for the modern synthesis. *Trends in Ecology & Evolution* 1995;**10**:294–9. [https://doi.org/10.1016/s0169-5347\(00\)89105-3](https://doi.org/10.1016/s0169-5347(00)89105-3)
- Nogueras V, Ortego J. Genomic evidence of speciation by fusion in a recent radiation of grasshoppers. *Evolution; International Journal of Organic Evolution* 2022;**76**:2618–33. <https://doi.org/10.1111/evo.14508>
- Nogueras V, Cordero PJ, Ortego J. Hierarchical genetic structure shaped by topography in a narrow-endemic montane grasshopper. *BMC Evolutionary Biology* 2016;**16**:96. <https://doi.org/10.1186/s12862-016-0663-7>
- Nogueras V, Cordero PJ, Ortego J. Integrating genomic and phenotypic data to evaluate alternative phylogenetic and species delimitation hypotheses in a recent evolutionary radiation of grasshoppers. *Molecular Ecology* 2018;**27**:1229–44. <https://doi.org/10.1111/mec.14504>
- Noroozi J, Talebi A, Doostmohammadi M *et al.* Hotspots within a global biodiversity hotspot—areas of endemism are associated with high mountain ranges. *Scientific Reports* 2018;**8**:10345. <https://doi.org/10.1038/s41598-018-28504-9>
- Ortego J, Knowles LL. Geographical isolation versus dispersal: relictual alpine grasshoppers support a model of interglacial diversification with limited hybridization. *Molecular Ecology* 2022;**31**:296–312. <https://doi.org/10.1111/mec.16225>
- Ortego J, Gutiérrez-Rodríguez J, Nogueras V. Demographic consequences of dispersal-related trait shift in two recently diverged taxa of montane grasshoppers. *Evolution; International Journal of Organic Evolution* 2021;**75**:1998–2013. <https://doi.org/10.1111/evo.14205>
- Ortego J, Kaya S, Çiplak B *et al.* Microgeographic speciation in a complex of Anatolian bush crickets facilitated by fast evolution of reproductive isolation. *Journal of Evolutionary Biology* 2024;**37**:14–27. <https://doi.org/10.1093/jeb/voad008>
- Pallarés S, Ortego J, Carbonell JA *et al.* Genomic, morphological and physiological data support fast ecotypic differentiation and incipient speciation in an alpine diving beetle. *Molecular Ecology* 2024;**33**:e17487. <https://doi.org/10.1111/mec.17487>
- Perrigo A, Hoorn C, Antonelli A. Why mountains matter for biodiversity. *Journal of Biogeography* 2020;**47**:315–25. <https://doi.org/10.1111/jbi.13731>
- Peterson BK, Weber JN, Kay EH *et al.* Double digest radseq: an inexpensive method for *de novo* snp discovery and genotyping in model and non-model species. *PLoS One* 2012;**7**: e37135. <https://doi.org/10.1371/journal.pone.0037135>
- Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 2006;**190**:231–59. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Phillips SJ, Dudik M. Modeling of species distributions with MAXENT: new extensions and a comprehensive evaluation. *Ecography* 2008;**31**:161–75. <https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- Presgraves DC. The molecular evolutionary basis of species formation. *Nature Reviews. Genetics* 2010;**11**:175–80. <https://doi.org/10.1038/nrg2718>
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000;**155**:945–59.
- Puechmaile SJ. The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular Ecology Resources* 2016;**16**:608–27. <https://doi.org/10.1111/1755-0998.12512>
- Pyron RA, O'Connell KA, Duncan SC *et al.* Speciation hypotheses from phylogeographic delimitation yield an integrative taxonomy for seal salamanders (*Desmognathus monticola*). *Systematic Biology* 2023;**72**:179–97. <https://doi.org/10.1093/sysbio/syaa065>
- Pyron RA, Kakkera A, Beamer DA *et al.* Discerning structure versus speciation in phylogeographic analysis of seepage salamanders (*Desmognathus aeneus*) using demography, environment, geography, and phenotype. *Molecular Ecology* 2024;**33**:e17219. <https://doi.org/10.1111/mec.17219>
- R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2025. Available at <https://www.R-project.org/>. (11 November 2024, date last accessed).
- Rahbek C, Borregaard MK, Antonelli A *et al.* Building mountain biodiversity: geological and evolutionary processes. *Science (New York, N.Y.)* 2019;**365**:1114–9. <https://doi.org/10.1126/science.aax0151>
- Rosenberg NA. Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* 2004;**4**:137–8. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>
- Roux C, Fraisse C, Romiguier J *et al.* Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology* 2016;**14**:e2000234. <https://doi.org/10.1371/journal.pbio.2000234>
- Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC *et al.* DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* 2017;**34**:3299–302. <https://doi.org/10.1093/molbev/msx248>
- Sandel B, Arge L, Dalsgaard B *et al.* The influence of late quaternary climate-change velocity on species endemism. *Science (New York, N.Y.)* 2011;**334**:660–4. <https://doi.org/10.1126/science.1210173>
- Soria-Carrasco V, Gompert Z, Comeault AA *et al.* Stick insect genomes reveal natural selection's role in parallel speciation. *Science (New York, N.Y.)* 2014;**344**:738–42. <https://doi.org/10.1126/science.1252136>
- Stamatakis A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)* 2014;**30**:1312–3. <https://doi.org/10.1093/bioinformatics/btu033>
- Steinbauer MJ, Irl SDH, Beierkuhnlein C. Elevation-driven ecological isolation promotes diversification on Mediterranean islands. *Acta Oecologica* 2013;**47**:52–6. <https://doi.org/10.1016/j.actao.2012.11.004>
- Steinbauer MJ, Field R, Grytnes JA *et al.* Topography-driven isolation, speciation and a global increase of endemism with elevation. *Global Ecology and Biogeography* 2016;**25**:1097–107. <https://doi.org/10.1111/geb.12469>
- Sukumaran J, Knowles LL. Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences of the United States of America* 2017;**114**:1607–12. <https://doi.org/10.1073/pnas.1607921114>
- \*Swofford D. Paup. Phylogenetic Analysis Using Parsimony (\*and Other Methods). v.4.0b10, 2002.
- Tonzo V, Papadopolou A, Ortego J. Genomic data reveal deep genetic structure but no support for current taxonomic designation in a grasshopper species complex. *Molecular Ecology* 2019;**28**:3869–86. <https://doi.org/10.1111/mec.15189>
- Tzedakis PC. Museums and cradles of Mediterranean biodiversity. *Journal of Biogeography* 2009;**36**:1033–4. <https://doi.org/10.1111/j.1365-2699.2009.02123.x>
- Valbuena-Ureña E, Oromi N, Soler-Membrives A *et al.* Jailed in the mountains: genetic diversity and structure of an endemic newt species across



- the Pyrenees. *PLoS One* 2018;**13**:e0200214. <https://doi.org/10.1371/journal.pone.0200214>
- Voje KL, Hemp C, Flagstad O *et al.* Climatic change as an engine for speciation in flightless Orthoptera species inhabiting African mountains. *Molecular Ecology* 2009;**18**:93–108.
- Walsh B. Estimating the time to the most recent common ancestor for the Y chromosome or mitochondrial DNA for a pair of individuals. *Genetics* 2001;**158**:897–912.
- Wallis GP, Waters JM, Upton P *et al.* Transverse alpine speciation driven by glaciation. *Trends in Ecology & Evolution* 2016;**31**:916–26. <https://doi.org/10.1016/j.tree.2016.08.009>
- Wollenberg KC, Vieites DR, van der Meijden A *et al.* Patterns of endemism and species richness in Malagasy cophyline frogs support a key role of mountainous areas for speciation. *Evolution; International Journal of Organic Evolution* 2008;**62**:1890–907. <https://doi.org/10.1111/j.1558-5646.2008.00420.x>