MOLECULAR ECOLOGY

Genomic footprints of an old affair: SNP data reveal historical hybridization and the subsequent evolution of reproductive barriers in two recently diverged grasshoppers with partly overlapping distributions

Journal:	*Molecular Ecology*
Manuscript ID	MEC-19-1360.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	10-May-2020
Complete List of Authors:	Tonzo, Vanina; Estacion Biologica de Donana CSIC, Integrative Ecology Papadopoulou, Anna; University of Cyprus, Department of Biological Sciences Ortego, Joaquin; Estacion Biologica de Donana CSIC, Integrative Ecology
Keywords:	coalescent-based simulations, ddRAD-seq, introgression, reproductive isolation, Hybridization

Genomic footprints of an old affair: SNP data reveal historical hybridization and the subsequent evolution of reproductive barriers in two recently diverged grasshoppers with partly overlapping distributions

Vanina Tonzo[1], Anna Papadopoulou[2], Joaquín Ortego[1]

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

[2]Department of Biological Sciences, University of Cyprus; 1 Panepistimiou Av. – 2109; Nicosia, Cyprus

Author for correspondence:

Vanina Tonzo

Estación Biológica de Doñana, EBD-CSIC,

Avda. Américo Vespucio 26, E-41092 Seville, Spain

E-mail: vaninatonzo@gmail.com

Phone: +34 954 232 340

Abstract

Secondary contact in close relatives can result in hybridization and the admixture of previously isolated gene pools. However, after an initial period of hybridization, reproductive isolation can evolve through different processes and lead to the interruption of gene flow and the completion of the speciation process. *Omocestus minutissimus* and *O. uhagonii* are two closely related grasshoppers with partially overlapping distributions in the Central System mountains of the Iberian Peninsula. To analyse spatial patterns of historical and/or contemporary hybridization between these two taxa and understand how species boundaries are maintained in the region of secondary contact, we sampled sympatric and allopatric populations of the two species and obtained genome-wide SNP data using a restriction site-associated DNA sequencing approach. We used Bayesian clustering analyses to test the hypothesis of contemporary hybridization in sympatric populations and employed a suite of phylogenomic approaches and a coalescent-based simulation framework to evaluate alternative hypothetical scenarios of interspecific gene flow. Our analyses rejected the hypothesis of contemporary hybridization but revealed past introgression in the area where the distributions of the two species overlap. Overall, these results point to a scenario of historical gene flow after secondary contact followed by the evolution of reproductive isolation that currently prevents hybridization among sympatric populations.

KEYWORDS: coalescent-based simulations, ddRAD-seq, hybridization, introgression, reproductive isolation

1. INTRODUCTION

Elucidating the processes that generate and maintain species diversity is a main ambition of evolutionary research (Graham, Ron, Santos, Schneider, & Moritz, 2004; Grant, Grant, Markert, Keller, & Petren, 2004; Fitzpatrick, Fordyce, & Gavrilets, 2009). The formation and persistence of species often depends on the evolution of reproductive isolation mechanisms that prevent interbreeding with other recently diverged taxa or closely related lineages (Marques, Draper, Riofrio, & Naranjo, 2014; Soltis & Soltis, 2009). In the context of allopatric populations, long-term geographic isolation can facilitate the development of reproductive barriers through genetic drift or divergent selection pressures, which can ultimately lead to population divergence and speciation (Hoskin, Higgie, McDonald, & Moritz, 2005; Maguilla, Escudero, Hipp, & Luceno, 2017; Schenk, Kontur, Wilson, Noble, & Derryberry, 2018). However, lineages or species often come into secondary contact and, if barriers to gene flow are lacking or incomplete, hybridization will take place. This phenomenon can have different outcomes with important evolutionary consequences (Abbott et al., 2013; Mallet, 2005; Mayr, 1963). At one extreme, barriers to gene exchange may break down upon secondary contact and lead to the collapse of formerly distinct species in a hybrid swarm characterized by extensive admixture of parental genotypes (i.e., speciation reversal or lineage fusion; Kearns et al., 2018; Taylor et al., 2006). At the opposite extreme, reduced fitness of hybrids and strong selection against them will favour the rapid evolution of barriers to gene flow (i.e., prezygotic isolation), which can ultimately lead to total reproductive isolation and culminate in the completion of the speciation process (i.e., reinforcement of isolation; Butlin, 1995; Dobzhansky, 1937; Lemmon & Juenger, 2017; Servedio & Kirkpatrick, 1997). An intermediate scenario is the

formation of tension zones with a variable geographical width determined by the equilibrium between dispersal and selection against hybrids (Barton & Hewitt, 1985; Key, 1968).

Hybrid zones have been defined as ‘‘windows on the evolutionary processes’’ (Harrison, 1990; Hewitt, 1988) and their study through space (Barton & Hewitt, 1989; Howard, Waring, Tibbets, & Gregory, 1993) and time (Britch, Cain, & Howard, 2001; Buggs, 2007) have provided important insights into the evolution of reproductive isolation and the formation of new species. Past climate changes, such as Pleistocene glacial cycles, have resulted in recurrent range expansions and contractions in many organisms, often putting into geographic contact closely related species and linages that had remained geographically isolated for large periods of time. Alpine and montane species represent a paradigmatic example of species experiencing dramatic distributional shifts in response to Pleistocene glacial cycles, descending to lower altitudes and expanding their distributions in cold periods and shrinking their ranges to high elevation areas during interglacials (Schmitt, 2009; Seddon, Santucci, Reeve, & Hewitt, 2001; Tzedakis, Emerson, & Hewitt, 2013). Under these conditions, closely related species that evolved in isolation during interglacial periods can recurrently come into secondary contact and hybridize. In other cases, interbreeding species share a large portion of their respective ranges but only hybridize in certain areas with specific environmental conditions or scattered patches where they meet, forming a mosaic hybrid structure rather than a well-defined cline limited to narrow contact zones (Barton & Hewitt, 1985). Independently of their nature and origin, contact zones offer the opportunity to study in real time the process of reproductive isolation or, if completed, to obtain indirect evidence about when (tempo) and how (modes) it might have evolved. However, testing alternative hypotheses about the evolution of reproductive isolation

can be extremely challenging, making necessary the integration of multiple sources of analytical evidence (e.g., phylogenetics and population genetics) and the development of model-based approaches considering the biogeographical context of gene flow (Payseur & Rieseberg, 2016). The high power of genomic data to resolve historical events of hybridization and test complex scenarios of gene flow in virtually any organismal model and biogeographical setting has exponentially increased our capacity to quantify the magnitude and timing of interspecific gene flow and distinguish among alternative demographic scenarios (e.g., de Manuel et al., 2016; Lohse, Clarke, Ritchie, & Etges, 2015; Ortego, Gugger, & Sork, 2018).

Here, we use as a study system two recently diverged grasshopper species with partly overlapping distributions to illustrate the potential of integrating different analytical approaches for inferring the tempo and mode of evolution of reproductive isolation (or its lack thereof) and gain insights into the speciation process. Grasshoppers (Orthoptera: Caelifera) are an interesting system to study hybridization and its evolutionary consequences, as many species have very recently evolved in allopatry (Mayer, Berger, Gottsberger, & Schulze, 2010; Ragge & Reynolds, 1998) and present incomplete barriers to gene flow (e.g., Saldamando, Tatsuta, & Butlin, 2005). In turn, their distributions often overlap across large geographical areas (Hill, 2015), show a mosaic distribution (Rohde et al., 2017) or have recurrently come into secondary contact as a consequence of range shifts driven by past climate changes (Bridle, Baird, & Butlin, 2001), providing ideal biogeographic scenarios for the study of hybridization and the evolution of reproductive isolation (Butlin, Ritchie, & Hewitt, 1991; Virdee & Hewitt, 1994). In this study we focus on two grasshopper species of the subgenus *Dreuxius* (genus *Omocestus*), a species complex comprised of nine taxa distributed in the Iberian Peninsula and Northwestern Africa (Cigliano, Braun, Eades, & Otte, 2019;

García-Navas, Noguerales, Cordero, & Ortego, 2017). Most species of this complex are distributed in allopatry, isolated at high elevation in different mountain ranges. One exception are the taxa *Omocestus minutissimus* (Brullé 1832) and *O. uhagonii* (Bolivar 1876), which show partially overlapping distributions in the Central System Mountains from the Iberian Peninsula. As the rest of species of the subgenus, both taxa are brachypterous, present a low dispersal capacity, and have a similar annual life cycle, with an adult breeding phase from the end of July to the beginning of October (Clemente, Garcia, & Presa, 1991). The two species are predominantly graminivorous and occupy open habitats tightly linked to cushion and thorny shrub formations that they use as refuge (Clemente et al., 1991). However, both species differ on the extent of their distributions and elevational ranges. While *O. minutissimus* presents a wider distribution, with patchy populations distributed in eastern and central Iberia from sea level to 2,500 m of elevation, *O. uhagonii* is restricted to the Central System and altitudes over 1,600-1,800 m (Figure 1). The two taxa partially co-occur across the distribution range of *O. uhagonii*, with several sympatric populations at high elevations in which adult individuals of the two species co-exist at high numbers in the same microhabitats (J. Ortego, personal observation). Therefore, this system provides an interesting case study to analyse the presence of contemporary and past hybridization and understand the maintenance of species boundaries in two closely related taxa that might have weak or recently evolved reproductive isolation mechanisms.

We extensively sampled sympatric and allopatric populations of *O. minutissimus* and *O. uhagonii* across their respective distribution ranges and genotyped them via restriction-site-associated DNA sequencing (ddRAD-seq; (Peterson, Weber, Kay, Fisher, & Hoekstra, 2012) to infer historical and contemporary interspecific gene flow and elucidate the evolutionary outcomes of such processes. Specifically, we first used

Bayesian clustering analyses to determine the genetic ancestry of individuals and test the hypothesis of contemporary hybridization in sympatric populations of the two taxa. Second, we employed a suite of phylogenomic approaches and a coalescent-based simulation framework to evaluate alternative scenarios of historical hybridization and estimate the timing, magnitude and directionality of interspecific gene flow. Our genomic data rejected the hypothesis of contemporary hybridization but revealed the presence of past genetic introgression, pointing to a scenario of historical hybridization after secondary contact followed by the evolution of reproductive barriers that nowadays prevent gene flow among sympatric populations of the two species.

2 MATERIALS AND METHODS

2.1. Population sampling

Between 2011 and 2015, we sampled *Omocestus minutissimus* (88 individuals, 15 populations) and *O. uhagonii* (64 individuals, 10 populations) from a total of 25 populations that cover the entire distribution range of the two taxa (Table S1; Figure 1). Eleven sampling populations of *O. minutissimus* are located in central Iberia and partly overlap with the distribution range of *O. uhagonii* (hereafter referred as sympatric populations). Among these populations of *O. minutissimus*, five were strictly sympatric with *O. uhagonii* and the two species were collected from the same localities (Table S1; Figure 1). The rest of the sampled populations of *O. minutissimus* are located in eastern Iberia and separated >300 km from the nearest population of *O. uhagonii* (hereafter referred as allopatric populations). We stored specimens in 2 ml vials with 96% ethanol and preserved them at −20° C until needed for DNA extraction. Detailed information on sampling populations is presented in Table S1.

169 **2.2. Genomic library preparation and data processing**

170 We used NucleoSpin Tissue kits (Macherey-Nagel, Düren, Germany) to extract and
171 purify total genomic DNA from a hind leg of each individual. We processed genomic
172 DNA in house following the double-digest restriction-site associated DNA procedure
173 (ddRADseq) described in Peterson et al. (2012), with some minor modifications
174 detailed in Lanier, Massatti, He, Olson, & Knowles (2015). Briefly, we digested DNA
175 with the restriction enzymes EcoRI and MseI (New England Biolabs, Ipswich, MA,
176 USA) and ligated Illumina adaptors including unique 7-base-pair barcodes to the
177 digested fragments of each individual. We pooled ligation products into four different
178 libraries, size selected for fragments between 475 and 580 bp using a Pippin Prep
179 machine (Sage Science, Beverly, MA, USA), and amplified them by PCR with 10-12
180 cycles using the iProof™ High-Fidelity DNA Polymerase (BIO-RAD, Veenendaal, The
181 Netherlands). We sequenced the libraries in single-read 151-bp lanes on an Illumina
182 HiSeq2500 platform at The Centre for Applied Genomics (Hospital for Sick Children,
183 Toronto, ON, Canada).

184 We demultiplexed raw sequences using *process_radtags*, a program distributed
185 as part of the STACKS pipeline (Catchen, Hohenlohe, Bassham, Amores, & Cresko,
186 2013). We only retained reads with Phred scores ≥ 10 (using a sliding window of 15%),
187 no adaptor contamination, and unambiguous barcode and restriction cut sites. We
188 checked read quality in FASTQC v.0.11.5
189 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and trimmed sequences to
190 130 bp using SEQTK script (Heng Li, https://github.com/lh3/seqtk) in order to remove
191 low-quality reads near the 3' ends. As an additional quality-filtering step, we used
192 PYRAD v.3.0.66 (Eaton, 2014) to convert base calls with a Phred score <20 into Ns and
193 discard reads with >2 Ns. We assembled retained reads into *de novo* loci using PYRAD,

considering parameter values for clustering threshold of sequence similarity (W_{CLUST} = 0.85), minimum coverage depth (d = 5), maximum number of individuals with shared heterozygous sites (*maxSH* = p.10), and maximum number of polymorphic sites in a final locus (*maxSNPs* = 20) based on suggestions from the literature (Eaton, 2014; Eaton & Ree, 2013; Takahashi, Nagata, & Sota, 2014). Finally, we generated final datasets for subsequent analyses discarding loci that were not present in at least ~25 % of the samples (*minCov* = ~25 %).

2.3. Genetic structure and hybrid identification

We identified hybrids and introgressed individuals between *O. minutissimus* and *O. uhagonii* using the Bayesian clustering methods implemented in the programs FASTSTRUCTURE v.1.0 (Raj, Stephens, & Pritchard, 2014) and STRUCTURE v.2.3.4 (Pritchard, Stephens, & Donnelly, 2000). First, we used the highly efficient algorithm implemented in FASTSTRUCTURE to analyse the entire dataset including all populations. Second, we used both FASTSTRUCTURE and classic STRUCTURE to perform more detailed analyses focused on the Central System, the region where the ranges of the two taxa partly overlap, with some populations living in sympatry and, thus, the two species currently have the opportunity to hybridize. We ran FASTSTRUCTURE analyses using a simple prior, considering a convergence criterion of 1×10^{-7} and conducting 25 independent runs for each value of K (from K = 1 to K = 10). Following Raj et al. (2014), we used the *chooseK.py* script to assess model complexity by estimating the metrics $K^{*}_{\varnothing}{}^{c}$, the value of K that maximizes log-marginal likelihood lower bound (LLBO) of the data, and K^{*}_{ε}, the smallest number of model components explaining at least 99% of cumulative ancestry contribution. We plotted individual co-ancestry coefficients for the most likely K value using DISTRUCT v.1.1 (Rosenberg, 2004). We

ran STRUCTURE assuming correlated allele frequencies and admixture, and without using prior population information (Hubisz, Falush, Stephens, & Pritchard, 2009; Pritchard et al., 2000). We conducted 15 independent runs for each value of K (from $K = 1$ to $K = 10$) to estimate the optimal number of genetic clusters with 200,000 MCMC cycles, following a burn-in step of 100,000 iterations. We used STRUCTURE HARVESTER (Earl & vonHoldt, 2012) to assess the number of genetic clusters that best describes our data according to log probabilities of the data (LnPr(X|K) for each value of K (Pritchard et al., 2000) and the ΔK method (Evanno, Regnaut, & Goudet, 2005). We used CLUMPP v.1.1.2 and the Greedy algorithm to align multiple runs of STRUCTURE for the same K value (Jakobsson & Rosenberg, 2007) and DISTRUCT to visualize as bar plots the individual's probabilities of population membership.

2.4. Phylogenomic analyses and inference of historical hybridization

To determine the presence of historical hybridization (i.e., introgression), we employed four-taxon ABBA/BABA tests based on the D-statistic (Durand, Patterson, Reich, & Slatkin, 2011) and TREEMIX analyses (Pickrell & Pritchard, 2012). Assuming that the sister taxa 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 ($D > 0$). We assigned sympatric populations of *O. minutissimus* to P1 (64 individuals, from localities COVA, PESQ, NEGR, AVIL, PARA, GRED, PICO, SERR, MIJA, CASI, and MALA; hereafter, MS), allopatric populations of *O. minutissimus* to P2 (24 individuals, from localities MONT, PORT, ESPU, and TEJE; hereafter, MA), and populations of *O. uhagonii* to P3 (64 individuals from all sampling localities for this taxon; hereafter, US). Note that the sympatric and allopatric

populations of *O. minutissimus* are located in the central and west portions of the species distribution range, respectively, and correspond to the two well-defined genetic clusters identified by Bayesian clustering analyses for this taxon (see Results section and Figure 1). We used as the outgroup (O) the taxon *O. antigai* (Bolívar, 1897), a species also belonging to the subgenus *Dreuxius* (Cigliano et al., 2019; García-Navas et al., 2017). Specifically, we used sequences from 94 individuals of this species available at NCBI Sequence Read Archive (SRA) under BioProject PRJNA543714 (Tonzo, Papadopoulou, & Ortego, 2019). We performed ABBA/BABA tests in PYRAD and used 1,000 bootstrap replicates to obtain the standard deviation of the *D*-statistic (Eaton & Ree, 2013; e.g., Huang, 2016).

We also analysed the potential presence of introgression and determined the direction of gene flow with TREEMIX v.1.12 (Pickrell & Pritchard, 2012). We used TREEMIX to construct a tree-based model of population genetic relationships and infer events of genetic admixture using SNP frequency data. TREEMIX fits a population graph (i.e., a phylogenetic tree that incorporates admixture) on the basis of allele frequencies and a Gaussian approximation to genetic drift, allowing patterns of splits and mixtures in multiple populations to be inferred. To perform TREEMIX analyses, we pooled populations into the same three groups used to run ABBA/BABA tests. In a first step, we estimated a maximum-likelihood tree rooted with *O. antigai*. Then, we tested the existence of a range of migration events (m = 0 to 5, with three replicated runs each) and calculated the proportion of the variance in population covariances explained by the population graph with different numbers of admixture events to determine the model best fitting the data (e.g., Gompert et al., 2014). We assumed the independence of all SNPs and used a window size of one SNP (k = 1; e.g., Vera, Díez-del-Molino, & García-Marín, 2016).

2.5. Testing alternative models of gene flow

As TREEMIX only models migration as discrete events and does not consider continuous gene flow (Pickrell & Pritchard, 2012) we applied the coalescent-based modelling approach implemented in FASTSIMCOAL2 (Excoffier, Dupanloup, Huerta-Sanchez, Sousa, & Foll, 2013) to statistically test the relative fit of more complex historical demographic models to our genomic data. We used FASTSIMCOAL2 and the site frequency spectrum (SFS) (Excoffier et al., 2013) to test seven alternative models of gene flow (Figure 2). These models considered the same three population groups used in ABBA/BABA tests and TREEMIX analyses. All scenarios considered an early split between the two species followed by the divergence between sympatric (MS) and allopatric (MA) populations of *O. minutissimus*. The tested scenarios considered (i) total absence of post-divergence gene flow (Model 0, not shown in Figure 2); (ii) gene flow only between the two population groups of *O. minutissimus* (i.e., absence of interspecific gene flow) (Models 1-2); (iii) historical gene flow between sympatric populations of *O. uhagonii* and *O. minutissimus* (i.e., interspecific gene flow) but absence of post-divergence gene flow between the two population groups of *O. minutissimus* (Models 3-4); (iv) gene flow between the two population groups of *O. minutissimus* and historical gene flow between sympatric populations of *O. uhagonii* and *O. minutissimus* (Models 5-6) (Figure 2). These scenarios were tested considering both symmetric (Models 1, 3 and 5) and asymmetric (Models 2, 4 and 6) gene flow (Figure 2). It must be noted that detailed Bayesian clustering analyses performed across multiple populations of the two taxa in the region where their distribution ranges overlap did not show any evidence of ongoing interspecific gene flow (see Results

section) and, for this reason, we did not consider models incorporating contemporary interspecific gene flow.

We calculated a folded joint SFS considering a single SNP per locus to avoid the effects of linkage disequilibrium. Because we did not include invariable sites in the SFS, we fixed the effective population size for one population group (US) to enable the estimation of other parameters in FASTSIMCOAL2 (e.g., Lanier et al., 2015; Papadopoulou & Knowles, 2015). The effective population size (N_e) fixed in the models was calculated from the level of nucleotide diversity (π) and estimates of mutation rate per site per generation (μ), according to the equation $N_e = \pi/4\mu$ (Lynch & Conery, 2003). Nucleotide diversity for US ($\pi = 0.005$) was estimated from polymorphic and non-polymorphic loci using DNASP v.6.11.01 (Librado & Rozas, 2009) and the *.allele* file generated by PYRAD. We considered the average mutation rate per site per generation of 2.80×10^{-9} estimated for *Drosophila melanogaster* (Keightley, Ness, Halligan, & Haddrill, 2014). To remove all missing data for the calculation of the joint SFS and minimize errors with allele frequency estimates, each population group was down sampled to 25% of individuals (32, 12 and 32 genes for MS, MA and US, respectively) using a custom Python script written by Isaac Overcast and available at GitHub (https://github.com/isaacovercast/easySFS). The final SFS contained 2,071 variable SNPs. Each of the seven models was run 100 replicated times using the computing resources provided by CESGA (Galician Supercomputer Center, Spain) considering 100,000–250,000 simulations for the calculation of the composite likelihood, 10–40 expectation-conditional maximization (ECM) cycles, and a stopping criterion of 0.001 (Papadopoulou & Knowles, 2015). We used an information-theoretic model selection approach based on the Akaike's information criterion (AIC) to determine the probability of each model given the observed data (Burnham &

Anderson, 2002; e.g., Abascal et al., 2016; Thome & Carstens, 2016). After the maximum likelihood was estimated for each model in every replicate, we calculated the AIC scores as detailed in Thome & Carstens (2016). AIC values for each model were rescaled (ΔAIC) calculating the difference between the AIC value of each model and the minimum AIC obtained among all competing models (i.e., the best model has ΔAIC = 0). Point estimates of the different demographic parameters for the best-supported model were selected from the run with the highest maximum composite likelihood. Finally, we calculated confidence intervals of parameter estimates from 100 parametric bootstrap replicates by simulating SFS from the maximum composite likelihood estimates and re-estimating parameters each time (Excoffier et al., 2013; e.g., Papadopoulou & Knowles, 2015).

3. RESULTS

3.1. Genomic data

Illumina sequencing returned an average of 2.74×10^6 reads per sample. After quality control, an average of 2.33×10^6 reads per sample was retained (Figure S1). The data sets obtained with PYRAD for all populations and only those from the Central System retained 15,219 and 20,350 unlinked SNPs, respectively.

3.2. Genetic structure and hybrid identification

The model complexity value that maximized the marginal likelihood in FASTSTRUCTURE analyses for the dataset including all populations was $K = 3$ across all replicates and the number of model components used to explain structure in the data was $K = 3$ in 15 replicates and $K = 4$ in 10 replicates. FASTSTRUCTURE analyses for $K = 3$ split

populations of *O. uhagonii* and sympatric and allopatric populations of *O. minutissimus* in different genetic clusters in which all individuals showed high probabilities of population membership ($q > 0.99$; Figure 1C). Assignment values to additional genetic clusters in FASTSTRUCTURE for analyses with $K > 4$ were extremely low in all cases ($q < 0.001$) and, thus, their respective bar plots were virtually identical to those obtained for $K = 3$ (for a similar result, see Baiz, Tucker, & Cortes-Ortiz, 2019; Tonzo et al., 2019).

The model complexity value that maximized the marginal likelihood in FASTSTRUCTURE analyses of the Central System populations was equal to $K = 2$ in all replicates and the number of model components used to explain structure in the data was $K = 2$ in 16 replicates and $K = 3$ in nine replicates. FASTSTRUCTURE analyses for $K = 2$ split populations of the two species in different genetic clusters and all individuals showed high probabilities of population membership ($q > 0.99$; Figure 3A). Again, assignment values to additional genetic clusters in FASTSTRUCTURE for analyses with $K > 2$ were extremely low in all cases ($q < 0.001$) and, thus, their respective bar plots were virtually identical to those obtained for $K = 2$. Classic STRUCTURE analyses also yielded an 'optimal' clustering value for $K = 2$ according to the ΔK criterion (Figure S2). The two inferred genetic groups supported a clear separation of the two species (Figure 3B). STRUCTURE analyses showed that populations of *O. uhagonii* from the eastern Central System present no signal of introgression from *O. minutissimus*. However, several populations of *O. minutissimus* and *O. uhagonii* from the western Central System, where the distribution of the two taxa overlap and five sampling localities present sympatric populations, showed signals of reciprocal genetic introgression (Figure 3B). These results suggest that FASTSTRUCTURE is less likely to reveal small proportions of admixed ancestry in comparison with STRUCTURE, which has been also suggested in previous studies (Stift, Kolar, & Meirmans, 2019; Tonzo et al., 2019).

STRUCTURE analyses indicate that populations of *O. uhagonii* and *O. minutissimus* present significant differences in the degree of introgression from the other species (one-way ANOVAs, introgression of *O. minutissimus* into *O. uhagonii*: $F_{9, 54} = 48.86$, $P < 0.001$; introgression of *O. uhagonii* into *O. minutissimus*: $F_{10, 53} = 96.68$, $P < 0.001$; Figure 3B). Although visually imperceptible in the FASTSTRUCTURE bar plot (Figure 3A), probabilities of population membership inferred by this software also revealed that populations of both *O. uhagonii* and *O. minutissimus* present significant differences in the degree of introgression from the other species (one-way ANOVAs, introgression of *O. minutissimus* into *O. uhagonii*: $F_{9, 54} = 14.01$, $P < 0.001$; introgression of *O. uhagonii* into *O. minutissimus*: $F_{10, 53} = 3.19$, $P = 0.003$; Figure 3A). The degree of introgression from *O. minutissimus* into *O. uhagonii* estimated by either FASTSTRUCTURE or STRUCTURE significantly decreased with longitude (FASTSTRUCTURE: $F_{1,8} = 6.99$, $P = 0.030$; STRUCTURE: $F_{1,8} = 16.22$, $P = 0.004$; see maps in Figure 3), but did not significantly differ between currently sympatric and allopatric populations (FASTSTRUCTURE: $F_{1,8} = 0.06$, $P = 0.81$; STRUCTURE: $F_{1,8} = 1.64$, $P = 0.236$). In contrast, the degree of introgression from *O. uhagonii* into *O. minutissimus* estimated by either FASTSTRUCTURE or STRUCTURE did not significantly decrease with longitude (FASTSTRUCTURE: $F_{1,9} = 1.12$, $P = 0.317$; STRUCTURE: $F_{1,9} = 0.41$, $P = 0.538$) or differ between currently sympatric and allopatric populations (FASTSTRUCTURE: $F_{1,9} = 0.39$, $P = 0.550$; STRUCTURE: $F_{1,9} = 0.03$, $P = 0.863$). For illustrative purposes, we displayed on a map the probabilities of assignment of the populations of *O. uhagonii* to the genetic cluster of *O. minutissimus* (i.e., the degree of introgression from *O. minutissimus* into *O. uhagonii*) by conducting a spatial interpolation using the Inverse Distance Weight (IDW) function available in ARCGIS v.10.5 (ESRI, Redlands, CA, USA) (Figure 3).

3.3. Phylogenomic analyses and inference of historical hybridization

The test of introgression based on the *D*-statistic supported post-divergence gene flow between sympatric populations of *O. minutissimus* and *O. uhagonii* (BABA = 220.59; ABBA = 162.38; *D*-statistic = -0.152; Z = 3.39; $P < 0.001$). Accordingly, TREEMIX analyses supported a single migration event (Figure S3) corresponding to admixture between sympatric populations of *O. minutissimus* and *O. uhagonii* (Figure 4).

3.4. Testing alternative models of gene flow

FASTSIMCOAL2 analyses showed that the most supported model (Model 5; ΔAIC = 0; Table 1) was the one considering symmetric interspecific gene flow between sympatric populations of *O. uhagonii* and *O. minutissimus* during a given period of time (T_{ADM1} and T_{ADM2}) and symmetric gene flow between sympatric (central) and allopatric (eastern) populations of *O. minutissimus*. Remarkably, models only considering interspecific gene-flow during a given period of time tended to be slightly more supported (ΔAIC > 0.6) than models only considering intraspecific gene flow between sympatric and allopatric populations of *O. minutissimus* (Table 1). Point estimates of demographic parameters under the best fitting model are presented in Table 2. Considering that the studied taxa are univoltine (i.e., 1-year generation time), FASTSIMCOAL2 analyses inferred that the separation of the two species (T_{DIV2}) and the split of eastern and central populations of *O. minutissimus* (T_{DIV1}) took place during the Early Pleistocene (Calabrian age) (Table 2). Gene flow between sympatric populations of *O. minutissimus* and *O. uhagonii* was estimated to happen during a period of time expanding ~15 ka in the Late Pleistocene (Tarantian age) (Table 2). Note, however, that although confidence intervals around point estimates of most parameters were reasonably tight, there was considerable uncertainty around estimates for the two-time

parameters delimiting the period of interspecific gene flow (i.e., T_{ADM1} and T_{ADM2}; Table 2). Finally, migration rates between sympatric populations of *O. uhagonii* and *O. minutissimus* (m_2) did not significantly differ from those estimated between sympatric (western) and allopatric (eastern) populations of *O. minutissimus* (m_1) (i.e., 95% CIs of m_1 and m_2 overlapped; Table 2). This indicates that historical interspecific gene flow was of the same order of magnitude as intraspecific gene flow between the two allopatric genetic clusters of *O. minutissimus*.

4. DISCUSSION

Hybridization is a phenomenon that has been extensively documented in contact zones where the distributions of closely related species with weak reproductive barriers meet (e.g., Folk, Soltis, Soltis, & Guralnick, 2018; Gugger & Cavender-Bares, 2013; Nadeau et al., 2013; Ortego, Gugger, Riordan, & Sork, 2014). Inferring events of past interspecific gene flow has important implications to understand the evolutionary history of organisms (e.g., humans; Prüfer et al., 2014; Wall et al., 2013), yet, detecting the footprints of such processes is challenging (Payseur & Rieseberg, 2016; e.g., Eaton, Hipp, González-Rodríguez, & Cavender-Bares, 2015; Ortego et al., 2018). Here, by combining a suite of phylogenomic and population genetic tools and extensive population sampling across the entire distribution of our two focal species, including currently sympatric and allopatric populations, we found no evidence for contemporary hybridization. However, we did detect signals of past introgression in the geographical region where the distribution range of the two taxa currently overlap.

4.1. Absence of contemporary interspecific gene flow

Bayesian clustering analyses showed a clear genotypic differentiation of the two species and further revealed the presence of two well-defined genetic clusters within *O. minutissimus,* corresponding with the populations of this taxon located in eastern (allopatric) and central (sympatric) Iberia (Figure 1). Detailed analyses across 21 populations from central Iberia where the distribution of the two species partially overlap and, thus, they currently have the opportunity to hybridize, showed no evidence of ongoing interspecific gene flow (i.e., F1 or first generation backcrosses). However, these analyses also revealed footprints of reciprocal introgression in the westernmost portion of the Central System, the area where the two species present overlapping distributions and some populations even co-occur. Although the degree of introgression did not statistically differ between currently sympatric and allopatric populations of the two species in the Central System, it was not spatially homogeneous. On the one hand, the proportion of genetic introgression significantly differed across populations of the two species. On the other hand, the degree of introgression from *O. minutissimus* into *O. uhagonii* increased westwards and the populations from the easternmost portion of the distribution range of this species, where *O. minutissimus* is not currently present, showed negligible signals of past hybridization (Figure 3). These results indicate spatial heterogeneity in the levels of introgression, suggesting that the magnitude and/or timing of historical hybridization differed among populations of the two species in the Central System (e.g., de Manuel et al., 2016; Ortego et al., 2018; Wall et al., 2013). The degree of introgression was consistently small in all populations of both species (STRUCTURE: <9 %; FASTRUCTURE: <0.02 %) and similar across individuals within populations. Thus, although our sample sizes are modest (128 individuals) and we cannot categorically discard that the two species sporadically hybridize, the observed patterns of introgression indicate that contemporary populations are at genotypic equilibrium (i.e.,

backgrounds of introgression are similar across all individuals within a given population) and suggest that hypothetical contemporary hybridization, if it even happens, is unlikely to have transcended F1 hybrids at least in the last generations. This end is also supported by the fact that the five populations where the two species currently co-occur do not show higher levels of introgression than nearby allopatric populations, which points to the fact that the observed patterns of genetic introgression reflect historical rather than contemporary interspecific gene flow.

4.2. Inferring historical hybridization

Both the phylogenomic analyses in TREEMIX and the *D*-statistic test yielded results compatible with those inferred by Bayesian clustering analyses and supported the hypothesis of historical hybridization between *O. uhagonii* and geographically overlapping populations of *O. minutissimus*. TREEMIX identifies gene flow in the context of competing hypotheses, taking into account the full sampled phylogeny when inferring admixture and introgression events (Pickrell & Pritchard, 2012). Even when we allowed several admixture edges for the two species, TREEMIX only inferred one event of introgression involving the sympatric populations of the two species. In agreement with the TREEMIX results, the *D*-statistic analysis detected the same pattern of interspecific gene flow. Bringing the TREEMIX and the *D*-statistic results together support a scenario of historical hybridization and enhance the interpretation of the low levels of genetic introgression revealed by clustering analyses in currently coexisting populations.

Coalescent-based analyses in FASTSIMCOAL2, which provides detailed estimation of demographic parameters, further supported a scenario of past interspecific gene flow

494 over a strictly bifurcating evolutionary history. Specifically, the most supported model
495 was the one considering interspecific gene flow between sympatric populations of *O.*
496 *uhagonii* and *O. minutissimus* during a given period of time and gene flow between
497 sympatric (central) and allopatric (eastern) populations of *O. minutissimus*. The
498 preferred model traced back the divergence of *O. uhagonii* and *O. minutissimus* to the
499 Late Pleistocene (~1.4 Ma) and this event was followed shortly after by the split of
500 eastern and central lineages of *O. minutissimus* (~1.2 Ma). These divergence times
501 agree with the Pleistocene diversification observed across most clades within the highly
502 speciose acridid subfamily Gomphocerinae (Song at al., 2015). Additionally, the split
503 time estimates are congruent with the crown age (< 3 Ma) inferred for the recent
504 radiation of the subgenus *Dreixius* based on mitochondrial DNA (García-Navas et al.,
505 2017). The results of coalescent analyses also indicate that gene flow between *O.*
506 *uhagonii* and *O. minutissimus* in the area of geographical overlap happened during a
507 limited amount of time (~15,000 years) around the last glacial maximum (14-29 ka).
508 During this period the two species engaged in gene flow at the same rate as the one
509 estimated between the two lineages of *O. minutissimus*, indicating that interspecific
510 gene flow, albeit low, was comparatively remarkable. The estimated timing of
511 interspecific gene flow is compatible with the likely expansion of *O. uhagonii* to lower
512 elevations during glacial periods. The shift to lower elevations might have put this
513 species into extensive geographic contact with the more ubiquitous *O. minutissimus*,
514 which presents many populations in foothills and valley bottoms where *O. uhagonii* is
515 not present nowadays. It must be noted, however, that confidence intervals around point
516 estimates for the time of interspecific gene flow are ample (Table 2), particularly for the
517 onset of this period (T_{ADM2}), and thus these results must be interpreted with extreme
518 caution. Uncertainty in these estimates could in part be driven by heterogeneity in the

519 timing and extent of hybridization among the different populations in the sympatric
520 area, as suggested by the significant differences in the signal of genetic introgression
521 observed among populations (see previous section).
522
523 **4.4. Inferred evolutionary scenario**
524
525 Our genomic analyses point to a scenario in which *O. uhagonii* and ancestral
526 populations of *O. minutissimus* likely diverged in allopatry followed by the split of *O.*
527 *minutissimus* into two lineages, one of which came into secondary contact and
528 hybridized with *O. uhagonii* during a limited period of time (Figure 5). Despite the fact
529 that nowadays the two species have ample opportunity to hybridize (i.e., a large
530 proportion their respective ranges currently overlap and several populations even co-
531 occur in the Central System), our analyses did not find any evidence of contemporary
532 hybridization, which suggests that reproductive isolation likely evolved after secondary
533 contact and historical gene flow (Figure 5). Speciation of Gomphocerinae and other
534 grasshoppers has been generally linked to allopatric divergence (Mayer et al. 2010), a
535 process that in the specific case of montane/alpine species of Pleistocene origin was
536 probably caused by the extensive fragmentation of ancestral populations driven by
537 Quaternary climatic oscillations (e.g., Huang, Hill, Ortego, & Knowles, 2020; Knowles,
538 2000; Scattolini, Confalonieri, Lira-Noriega, Pietrokovsky, & Cigliano, 2018). Different
539 lines of evidence also point to allopatric divergence as the most plausible mode of
540 speciation for *O. uhagonii* and *O. minutissimus*. First, the allopatric lineage of *O.*
541 *minutissimus* from eastern Iberia present a much larger distribution (Figure 1) and
542 significantly higher levels of genomic diversity (one-way ANOVA: $F_{1,13} = 12.05$, $P =$
543 0.004; Table S1) and effective population sizes (non-overlapping 95% CIs for N_e

estimates in FASTSIMCOAL2 analyses; Table 2) than the sympatric lineage of *O. minutissimus* from the Central System. This supports that *O. minutissimus* most likely originated in eastern Iberia and subsequently colonized the Central System, where it came into secondary contact with *O. uhagonii*. Second, the ecological niches of *O. minutissimus* and *O. uhagonii* are very similar and the two species present graminivorous feeding habits and same microhabitat preferences (Clemente et al., 1991; J. Ortego, personal observation). This points to considerable niche conservatism, typical of allopatric speciation, and rejects sympatric speciation via disruptive ecological selection (e.g., Grace, Wisely, Brown, Dowell, & Joern, 2010). Finally, our analyses showed no evidence of gene flow between *O. uhagonii* and ancestral populations of *O. minutissimus* (Figure 4), indicating that the two species likely evolved in allopatry and only exchanged gene flow after secondary contact in the region where their distribution ranges overlap (see also Sankararaman et al., 2014; Sankararaman, Patterson, Li, Paabo, & Reich, 2012).

The footprints of historical introgression among currently sympatric populations of *O. uhagonii* and *O. minutissimus* and the lack of evidence for contemporary gene flow lead us to hypothesize an evolutionary scenario in which reproductive isolation evolved after historical hybridization in the area where the distribution ranges of the two taxa currently overlap. The observed differences in the levels of introgression among sympatric populations suggest that barriers to gene flow might have evolved multiple times or, alternatively, could reflect heterogeneity in the proportion of the genome of the other species retained after the interruption of interspecific gene flow due to differences among the studied populations in their demographic histories (e.g., bottlenecks; Amorim et al., 2017; Lawson, Van Dorp, & Falush, 2018; Quilodran, Nussberger, Montoya-Burgos, & Currat, 2019) or spatial variation in the strength of

hypothetical purifying selection acting against introgressed alleles (Juric, Aeschbacher, & Coop, 2016; Petr, Paabo, Kelso, & Vernot, 2019). Alternative processes might have led to the evolution of reproductive isolation after secondary contact and hybridization. It has been frequently documented that interspecific gene flow can increase phenotypic and genomic divergence via the evolution of reproductive isolation and character displacement (Garner, Goulet, Farnitano, Molina-Henao, & Hopkins, 2018; Hopkins, Levin, & Rausher, 2012; Pfennig & Pfennig, 2009). One of the many potential costs of hybridization are the ecological and genetic dysfunctions of hybrid offspring, which can reduce their fitness and drive to reinforcement (Ortiz-Barrientos, Counterman, & Noor, 2004). In the reinforcement process, enhanced prezygotic isolation is favoured in sympatry in response to postzygotic isolation due to a strong selection against hybrids (Butlin, 1995; Coyne & Orr, 2004; Servedio & Noor, 2003). Selection for prezygotic isolation leads, in turn, to more divergent phenotypes and reproductive behaviours between species in sympatry than in allopatry (Moran, Zhou, Catchen, & Fuller, 2018). Thus, one possibility is that secondary contact and hybridization promoted the evolution of reproductive isolation via reinforcement, probably after an initial balance between dispersal and selection against hybrids in historical tension zones during which the two species experienced genetic exchange and introgression (Barton & Hewitt, 1985). An alternative explanation is that reproductive isolation evolved in geographical isolation as a consequence of genetic drift or as a fortuitous by-product of divergent selection on other traits (Coyne & Orr, 1989; Fitzpatrick, 2002; Sasa, Chippindale, & Johnson, 1998). The mosaic distribution of the two species in the Central System, with the presence of several sympatric populations but also large areas where the two species do not occur (e.g., eastern Central System and foothills; Figure 1), might have also

593 provided ample opportunity for the evolution of reproductive isolation in geographically
594 separated populations (Fitzpatrick, 2002).
595
596 **4.5. Conclusions and future directions**
597
598 The results of this study add to the growing body of evidence supporting that speciation-
599 with-gene-flow is more prevalent in nature than formerly acknowledged (Nosil, 2008;
600 Pinho & Hey, 2010; Roux et al., 2016). Our study system is very well-suited to study
601 the proximate mechanisms (e.g., reinforcement *vs*. genetic drift) that might have led to
602 the evolution of reproductive isolation. Future research should focus on analysing
603 mating preferences, phenotypic differentiation and reproductive character displacement
604 (song, courtship behaviour, genitalic structures, etc.) between currently sympatric and
605 allopatric populations of the two species (e.g., Butlin et al., 1991; Hollander, Smadja,
606 Butlin, & Reid, 2013), determining mating success and the viability of offspring
607 through experimental hybridization attempts in the laboratory (e.g., Coyne & Orr, 1989;
608 Hoskin et al., 2005; Saldamando et al., 2005) and, applying whole genome or
609 transcriptome sequencing data to detect potential genomic signatures of reinforcement
610 and/or identify loci that might be involved in reproductive isolation (Garner et al., 2018;
611 Hopkins et al., 2012; Roda, Mendes, Hahn, & Hopkins, 2017).
612
613 **ACKNOWLEDGEMENTS**
614
615 We are grateful to Amparo Hidalgo-Galiana for her valuable help during laboratory
616 work and the grasshopper illustration. We also would like to thank to Víctor
617 Noguerales, Conchi Cáliz, and Pedro J. Cordero for their help during field and

laboratory work, Sergio Pereira (The Centre for Applied Genomics) for Illumina sequencing, and two anonymous referees for constructive comments on an earlier version of this article. 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 to 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 Ministerio de Economía y Competitividad. During this work, AP and JO were supported by a Severo Ochoa (SEV-2012-0262) and a Ramón y Cajal (RYC-2013-12501) research fellowship, respectively.

AUTHOR CONTRIBUTIONS

V.T., A.P., and J.O. conceived and designed the study and analyses. J.O. collected the samples. V.T. performed the laboratory work and analysed the data guided by J.O. V.T. wrote the manuscript with help of J.O., and inputs from A.P.

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.12251600).

ORCID

Vanina Tonzo https://orcid.org/0000-0002-5062-1070

Anna Papadopoulou https://orcid.org/0000-0002-4656-4894

Joaquín Ortego https://orcid.org/0000-0003-2709-429X

REFERENCES

Abascal, F., Corvelo, A., Cruz, F., Villanueva-Canas, J. L., Vlasova, A., Marcet-Houben, M., . . . Godoy, J. A. (2016). Extreme genomic erosion after recurrent demographic bottlenecks in the highly endangered Iberian lynx. *Genome Biology, 17*, 251. doi:10.1186/s13059-016-1090-1

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., . . . Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology, 26*(2), 229-246. doi:10.1111/j.1420-9101.2012.02599.x

Amorim, C. E. G., Hofer, T., Ray, N., Foll, M., Ruiz-Linares, A., & Excoffier, L. (2017). Long-distance dispersal suppresses introgression of local alleles during range expansions. *Heredity, 118*(2), 135-142. doi:10.1038/hdy.2016.68

Baiz, M. D., Tucker, P. K., & Cortes-Ortiz, L. (2019). Multiple forms of selection shape reproductive isolation in a primate hybrid zone. *Molecular Ecology, 28*(5), 1056-1069. doi:10.1111/mec.14966

Barton, N. H., & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics, 16*, 113-148. doi:10.1146/annurev.ecolsys.16.1.113

Barton, N. H., & Hewitt, G. M. (1989). Adaptation, speciation and hybrid zones. *Nature, 341*(6242), 497-503. doi:10.1038/341497a0

Bridle, J. R., Baird, S. J. E., & Butlin, R. K. (2001). Spatial structure and habitat variation in a grasshopper hybrid zone. *Evolution, 55*(9), 1832-1843. doi:10.1554/0014-3820(2001)055[1832:ssahvi]2.0.co;2

Britch, S. C., Cain, M. L., & Howard, D. J. (2001). Spatio-temporal dynamics of the *Allonemobius fasciatus-A. socius* mosaic hybrid zone: a 14-year perspective. *Molecular Ecology, 10*(3), 627-638. doi:10.1046/j.1365-294x.2001.01215.x

Buggs, R. J. A. (2007). Empirical study of hybrid zone movement. *Heredity, 99*(3), 301-312. doi:10.1038/sj.hdy.6800997

Burnham, K. P., & Anderson, D. R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. New York, USA: Springer.

Butlin, R. K. (1995). Reinforcement - An idea evolving. *Trends in Ecology & Evolution, 10*(11), 432-434. doi:10.1016/s0169-5347(00)89173-9

Butlin, R. K., Ritchie, M. G., & Hewitt, G. M. (1991). Comparisons among morphological characters and between localities in the *Chorthippus parallelus* hybrid zone (Orthoptera, Acrididae). *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 334*(1271), 297-308. doi:10.1098/rstb.1991.0119

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. doi:10.1111/mec.12354

Cigliano, M. M., Braun, H., Eades, D. C., & Otte, D. (2019). *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* Bolivar, 1878. (Insecta, Orthoptera, Caelifera). *Graellsia, 46*, 191-246.

691 Coyne, J. A., & Orr, H. A. (1989). Patterns of speciation in *Drosophila*. *Evolution,*
692 *43*(2), 362-381. doi:10.2307/2409213
693 Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Association.
694 de Manuel, M., Kuhlwilm, M., Frandsen, P., Sousa, V. C., Desai, T., Prado-Martinez, J.,
695 . . . Marques-Bonet, T. (2016). Chimpanzee genomic diversity reveals ancient
696 admixture with bonobos. *Science, 354*(6311), 477-481.
697 doi:10.1126/science.aag2602
698 Dobzhansky, T. (1937). Genetic nature of species differences. *American Naturalist, 71*,
699 404-420. doi:10.1086/280726
700 Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for ancient
701 admixture between closely related populations. *Molecular Biology and*
702 *Evolution, 28*(8), 2239-2252. doi:10.1093/molbev/msr048
703 Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program
704 for visualizing STRUCTURE output and implementing the Evanno method.
705 *Conservation Genetics Resources, 4*(2), 359-361. doi:10.1007/s12686-011-
706 9548-7
707 Eaton, D. A. R. (2014). PYRAD: assembly of *de novo* RADseq loci for phylogenetic
708 analyses. *Bioinformatics, 30*(13), 1844-1849. doi:10.1093/bioinformatics/btu121
709 Eaton, D. A. R., Hipp, A. L., González-Rodríguez, A., & Cavender-Bares, J. (2015).
710 Historical introgression among the American live oaks and the comparative
711 nature of tests for introgression. *Evolution, 69*(10), 2587-2601.
712 doi:10.1111/evo.12758
713 Eaton, D. A. R., & Ree, R. H. (2013). Inferring phylogeny and introgression using
714 RADseq data: An example from flowering plants (*Pedicularis*: *Orobanchaceae*).
715 *Systematic Biology, 62*(5), 689-706. doi:10.1093/sysbio/syt032

716 Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of
717 individuals using the software STRUCTURE: a simulation study. *Molecular*
718 *Ecology, 14*(8), 2611-2620. doi:10.1111/j.1365-294X.2005.02553.x
719 Excoffier, L., Dupanloup, I., Huerta-Sanchez, E., Sousa, V. C., & Foll, M. (2013).
720 Robust demographic inference from genomic and SNP data. *PLoS Genetics,*
721 *9*(10), e1003905. doi:10.1371/journal.pgen.1003905
722 Fitzpatrick, B. M. (2002). Molecular correlates of reproductive isolation. *Evolution,*
723 *56*(1), 191-198. doi:10.1111/j.0014-3820.2002.tb00860.x
724 Fitzpatrick, B. M., Fordyce, J. A., & Gavrilets, S. (2009). Pattern, process and
725 geographic modes of speciation. *Journal of Evolutionary Biology, 22*(11), 2342-
726 2347. doi:10.1111/j.1420-9101.2009.01833.x
727 Folk, R. A., Soltis, P. S., Soltis, D. E., & Guralnick, R. (2018). New prospects in the
728 detection and comparative analysis of hybridization in the tree of life. *American*
729 *Journal of Botany, 105*(3), 364-375. doi:10.1002/ajb2.1018
730 García-Navas, V., Noguerales, V., Cordero, P. J., & Ortego, J. (2017). Ecological
731 drivers of body size evolution and sexual size dimorphism in short-horned
732 grasshoppers (Orthoptera: Acrididae). *Journal of Evolutionary Biology, 30*(8),
733 1592-1608. doi:10.1111/jeb.13131
734 Garner, A. G., Goulet, B. E., Farnitano, M. C., Molina-Henao, Y. F., & Hopkins, R.
735 (2018). Genomic signatures of reinforcement. *Genes, 9*(4), 191.
736 doi:10.3390/genes9040191
737 Gompert, Z., Lucas, L. K., Buerkle, C. A., Forister, M. L., Fordyce, J. A., & Nice, C. C.
738 (2014). Admixture and the organization of genetic diversity in a butterfly species
739 complex revealed through common and rare genetic variants. *Molecular*
740 *Ecology, 23*(18), 4555-4573. doi:10.1111/mec.12811

741 Grace, T., Wisely, S. M., Brown, S. J., Dowell, F. E., & Joern, A. (2010). Divergent
742 host plant adaptation drives the evolution of sexual isolation in the grasshopper
743 *Hesperotettix viridis* (Orthoptera: Acrididae) in the absence of reinforcement.
744 *Biological Journal of the Linnean Society, 100*(4), 866-878.
745 Graham, C. H., Ron, S. R., Santos, J. C., Schneider, C. J., & Moritz, C. (2004).
746 Integrating phylogenetics and environmental niche models to explore speciation
747 mechanisms in dendrobatid frogs. *Evolution, 58*(8), 1781-1793. doi:10.1554/03-
748 274
749 Grant, P. R., Grant, B. R., Markert, J. A., Keller, L. F., & Petren, K. (2004). Convergent
750 evolution of Darwin's finches caused by introgressive hybridization and
751 selection. *Evolution, 58*(7), 1588-1599.
752 Gugger, P. F., & Cavender-Bares, J. (2013). Molecular and morphological support for a
753 Florida origin of the Cuban oak. *Journal of Biogeography, 40*(4), 632-645.
754 doi:10.1111/j.1365-2699.2011.02610.x
755 Harrison, R. G. (1990). Hybrid zones: windows on evolutionary process. *Oxford*
756 *Surveys of Evolutionary Biology, 7*, 69-128.
757 Hewitt, G. M. (1988). Hybrid zones - Natural laboratories for evolutionary studies.
758 *Trends in Ecology & Evolution, 3*(7), 158-167. doi:10.1016/0169-
759 5347(88)90033-x
760 Hill, J. G. (2015). Revision and biogeography of the *Melanoplus scudderi* species group
761 (Orthoptera: Acrididae: Melanoplinae) with a description of 21 new species and
762 establishment of the *Carnegiei* and *Davisi* species groups. *Transactions of the*
763 *American Entomological Society, 141*(2), 252-350. doi:10.3157/061.141.0201

764 Hollander, J., Smadja, C. M., Butlin, R. K., & Reid, D. G. (2013). Genital divergence in
765 sympatric sister snails. *Journal of Evolutionary Biology, 26*(1), 210-215.
766 doi:10.1111/jeb.12029
767 Hopkins, R., Levin, D. A., & Rausher, M. D. (2012). Molecular signatures of selection
768 on reproductive character displacement of flower color in *Phlox drummondii*.
769 *Evolution, 66*(2), 469-485. doi:10.1111/j.1558-5646.2011.01452.x
770 Hoskin, C. J., Higgie, M., McDonald, K. R., & Moritz, C. (2005). Reinforcement drives
771 rapid allopatric speciation. *Nature, 437*(7063), 1353-1356.
772 doi:10.1038/nature04004
773 Howard, D. J., Waring, G. L., Tibbets, C. A., & Gregory, P. G. (1993). Survival of
774 hybrids in a mosaic hybrid zone. *Evolution, 47*(3), 789-800.
775 doi:10.2307/2410184
776 Huang, J. P. (2016). Parapatric genetic introgression and phenotypic assimilation:
777 testing conditions for introgression between Hercules beetles (Dynastes,
778 Dynastinae). *Molecular Ecology, 25*(21), 5513-5526. doi:10.1111/mec.13849
779 Huang, J. P., Hill, J. G., Ortego, J., & Knowles, L. L. (2020) Paraphyletic species no
780 more - genomic data resolve a Pleistocene radiation and validate morphological
781 species of the *Melanoplus scudderi* complex (Insecta: Orthoptera). *Systematic*
782 *Entomology*, in press. doi:10.1111/syen.12415
783 Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak
784 population structure with the assistance of sample group information. *Molecular*
785 *Ecology Resources, 9*(5), 1322-1332. doi:10.1111/j.1755-0998.2009.02591.x
786 Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and
787 permutation program for dealing with label switching and multimodality in

788 analysis of population structure. *Bioinformatics, 23*(14), 1801-1806.
789 doi:10.1093/bioinformatics/btm233
790 Juric, I., Aeschbacher, S., & Coop, G. (2016). The strength of selection against
791 Neanderthal introgression. *PLoS Genetics, 12*(11), e1006340.
792 doi:10.1371/journal.pgen.1006340
793 Kearns, A. M., Restani, M., Szabo, I., Schroder-Nielsen, A., Kim, J. A., Richardson, H.
794 M., . . . Omland, K. E. (2018). Genomic evidence of speciation reversal in
795 ravens. *Nature Communications, 9*, 906. doi:10.1038/s41467-018-03294-w
796 Keightley, P. D., Ness, R. W., Halligan, D. L., & Haddrill, P. R. (2014). Estimation of
797 the spontaneous mutation rate per nucleotide site in a *Drosophila melanogaster*
798 full-sib family. *Genetics, 196*(1), 313-320. doi:10.1534/genetics.113.158758
799 Key, K. H. L. (1968). Concept of stasipatric speciation. *Systematic Zoology, 17*(1), 14-
800 22. doi:10.2307/2412391
801 Knowles, L. L. (2000). Tests of Pleistocene speciation in montane grasshoppers (genus
802 *Melanoplus*) from the sky islands of western North America. *Evolution, 54*(4),
803 1337-1348.
804 Lanier, H. C., Massatti, R., He, Q., Olson, L. E., & Knowles, L. L. (2015). Colonization
805 from divergent ancestors: glaciation signatures on contemporary patterns of
806 genomic variation in collared pikas (*Ochotona collaris*). *Molecular Ecology,*
807 *24*(14), 3688-3705. doi:10.1111/mec.13270
808 Lawson, D. J., Van Dorp, L., & Falush, D. (2018). A tutorial on how not to over-
809 interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications, 9*,
810 3258. doi:10.1038/s41467-018-05257-7

Lemmon, E. M., & Juenger, T. E. (2017). Geographic variation in hybridization across a reinforcement contact zone of chorus frogs (*Pseudacris*). *Ecology and Evolution, 7*(22), 9485-9502. doi:10.1002/ece3.3443

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

Lohse, K., Clarke, M., Ritchie, M. G., & Etges, W. J. (2015). Genome-wide tests for introgression between cactophilic *Drosophila* implicate a role of inversions during speciation. *Evolution, 69*(5), 1178-1190. doi:10.1111/evo.12650

Lynch, M., & Conery, J. S. (2003). The origins of genome complexity. *Science, 302*(5649), 1401-1404. doi:10.1126/science.1089370

Maguilla, E., Escudero, M., Hipp, A. L., & Luceno, M. (2017). Allopatric speciation despite historical gene flow: Divergence and hybridization in *Carex furva* and *C. lucennoiberica* (Cyperaceae) inferred from plastid and nuclear RAD-seq data. *Molecular Ecology, 26*(20), 5646-5662. doi:10.1111/mec.14253

Mallet, J. (2005). Hybridization as an invasion of the genorne. *Trends in Ecology & Evolution, 20*(5), 229-237.

Marques, I., Draper, D., Riofrio, L., & Naranjo, C. (2014). Multiple hybridization events, polyploidy and low postmating isolation entangle the evolution of neotropical species of *Epidendrum* (Orchidaceae). *BMC Evolutionary Biology, 14*, 20. doi:10.1186/1471-2148-14-20

Mayer, F., Berger, D., Gottsberger, B., & Schulze, W. (2010). Non-ecological radiations in acoustically communicating grasshoppers? In M. Glaubrecht (Ed.), *Evolution in Action* (pp. 451-464). Berlin, Germany: Springer-Verlag.

Mayr, E. (1963). *Animal Species and Evolution*. Cambridge, USA: Harvard University Press.

Moran, R. L., Zhou, M. C., Catchen, J. M., & Fuller, R. C. (2018). Hybridization and postzygotic isolation promote reinforcement of male mating preferences in a diverse group of fishes with traditional sex roles. *Ecology and Evolution, 8*(18), 9282-9294. doi:10.1002/ece3.4434

Nadeau, N. J., Martin, S. H., Kozak, K. M., Salazar, C., Dasmahapatra, K. K., Davey, J. W., . . . Jiggins, C. D. (2013). Genome-wide patterns of divergence and gene flow across a butterfly radiation. *Molecular Ecology, 22*(3), 814-826. doi:10.1111/j.1365-294X.2012.05730.x

Nosil, P. (2008). Speciation with gene flow could be common. *Molecular Ecology, 17*(9), 2103-2106. doi:10.1111/j.1365-294X.2008.03715.x

Ortego, J., Gugger, P. F., Riordan, E. C., & Sork, V. L. (2014). Influence of climatic niche suitability and geographical overlap on hybridization patterns among southern Californian oaks. *Journal of Biogeography, 41*(10), 1895-1908. doi:10.1111/jbi.12334

Ortego, J., Gugger, P. F., & Sork, V. L. (2018). Genomic data reveal cryptic lineage diversification and introgression in Californian golden cup oaks (section *Protobalanus*). *New Phytologist, 218*(2), 804-818. doi:10.1111/nph.14951

Ortiz-Barrientos, D., Counterman, B. A., & Noor, M. A. F. (2004). The genetics of speciation by reinforcement. *PLoS Biology, 2*(12), 2256-2263. doi:10.1371/journal.pbio.0020416

Papadopoulou, A., & Knowles, L. L. (2015). Genomic tests of the species-pump hypothesis: Recent island connectivity cycles drive population divergence but

not speciation in Caribbean crickets across the Virgin Islands. *Evolution, 69*(6), 1501-1517. doi:10.1111/evo.12667

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Molecular Ecology, 25*(11), 2337-2360. doi:10.1111/mec.13557

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. doi:10.1371/journal.pone.0037135

Petr, M., Paabo, S., Kelso, J., & Vernot, B. (2019). Limits of long-term selection against Neandertal introgression. *Proceedings of the National Academy of Sciences of the United States of America, 116*(5), 1639-1644. doi:10.1073/pnas.1814338116

Pfennig, K. S., & Pfennig, D. W. (2009). Character displacement: Ecological and reproductive responses to a common evolutionary problem. *Quarterly Review of Biology, 84*(3), 253-276. doi:10.1086/605079

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics, 8*(11), e1002967. doi:10.1371/journal.pgen.1002967

Pinho, C., & Hey, J. (2010). Divergence with gene flow: Models and data. *Annual Review of Ecology, Evolution, and Systematics, 41*, 215-230.

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

Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., . . . Paabo, S. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature, 505*(7481), 43-49. doi:10.1038/nature12886

883 Quilodran, C. S., Nussberger, B., Montoya-Burgos, J. I., & Currat, M. (2019).
884 Hybridization and introgression during density-dependent range expansion:
885 European wildcats as a case study. *Evolution, 73*(4), 750-761.
886 doi:10.1111/evo.13704
887 Ragge, D. R., & Reynolds, W. J. (1998). *The Songs of the Grasshoppers and Crickets of*
888 *Western Europe*. Essex, UK: Harley Books.
889 Raj, A., Stephens, M., & Pritchard, J. K. (2014). FASTSTRUCTURE: Variational inference
890 of population structure in large SNP data sets. *Genetics, 197*(2), 573-589.
891 doi:10.1534/genetics.114.164350
892 Roda, F., Mendes, F. K., Hahn, M. W., & Hopkins, R. (2017). Genomic evidence of
893 gene flow during reinforcement in Texas Phlox. *Molecular Ecology, 26*(8),
894 2317-2330. doi:10.1111/mec.14041
895 Rohde, K., Hau, Y., Kranz, N., Weinberger, J., Elle, O., & Hochkirch, A. (2017).
896 Climatic effects on population declines of a rare wetland species and the role of
897 spatial and temporal isolation as barriers to hybridization. *Functional Ecology,*
898 *31*(6), 1262-1274. doi:10.1111/1365-2435.12834
899 Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population
900 structure. *Molecular Ecology Notes, 4*(1), 137-138. doi:10.1046/j.1471-
901 8286.2003.00566.x
902 Roux, C., Fraisse, C., Romiguier, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016).
903 Shedding light on the grey zone of speciation along a continuum of genomic
904 divergence. *PLoS Biology, 14*(12), e2000234. doi:10.1371/journal.pbio.2000234
905 Saldamando, C. I., Tatsuta, H., & Butlin, R. K. (2005). Hybrids between *Chorthippus*
906 *brunneus* and *C. jacobsi* (Orthoptera: Acrididae) do not show endogenous

907 postzygotic isolation. *Biological Journal of the Linnean Society, 84*(2), 195-203.
908 doi:10.1111/j.1095-8312.2005.000424.x
909 Sankararaman, S., Mallick, S., Dannemann, M., Prüfer, K., Kelso, J., Paabo, S., . . .
910 Reich, D. (2014). The genomic landscape of Neanderthal ancestry in present-day
911 humans. *Nature, 507*(7492), 354-357. doi:10.1038/nature12961
912 Sankararaman, S., Patterson, N., Li, H., Paabo, S., & Reich, D. (2012). The date of
913 interbreeding between Neandertals and modern humans. *PLoS Genetics, 8*(10),
914 e1002947. doi:10.1371/journal.pgen.1002947
915 Sasa, M. M., Chippindale, P. T., & Johnson, N. A. (1998). Patterns of postzygotic
916 isolation in frogs. *Evolution, 52*(6), 1811-1820. doi:10.1111/j.1558-
917 5646.1998.tb02258.x
918 Scattolini, M. C., Confalonieri, V., Lira-Noriega, A., Pietrokovsky, S., & Cigliano, M.
919 M. (2018). Diversification mechanisms in the Andean grasshopper genus
920 *Orotettix* (Orthoptera: Acrididae): ecological niches and evolutionary history.
921 *Biological Journal of the Linnean Society, 123*(4), 697-711.
922 doi:10.1093/biolinnean/bly008
923 Schenk, J. J., Kontur, S., Wilson, H., Noble, M., & Derryberry, E. (2018). Allopatric
924 speciation drives diversification of ecological specialists on sandhills.
925 *International Journal of Plant Sciences, 179*(4), 325-339. doi:10.1086/697073
926 Schmitt, T. (2009). Biogeographical and evolutionary importance of the European high
927 mountain systems. *Frontiers in Zoology, 6*, 9. doi:10.1186/1742-9994-6-9
928 Seddon, J. M., Santucci, F., Reeve, N. J., & Hewitt, G. M. (2001). DNA footprints of
929 European hedgehogs, *Erinaceus europaeus* and *E. concolor*. Pleistocene refugia,
930 postglacial expansion and colonization routes. *Molecular Ecology, 10*(9), 2187-
931 2198. doi:10.1046/j.0962-1083.2001.01357.x

Servedio, M. R., & Kirkpatrick, M. (1997). The effects of gene flow on reinforcement. *Evolution, 51*(6), 1764-1772. doi:10.2307/2410999

Servedio, M. R., & Noor, M. A. F. (2003). The role of reinforcement in speciation: Theory and data. *Annual Review of Ecology Evolution and Systematics, 34*, 339-364. doi:10.1146/annurev.ecolsys.34.011802.132412

Soltis, P. S., & Soltis, D. E. (2009). The role of hybridization in plant speciation. *Annual Review of Plant Biology, 60*, 561-588. doi:10.1146/annurev.arplant.043008.092039

Song, H. J., Amedegnato, C., Cigliano, M. M., Desutter-Grandcolas, L., Heads, S. W., Huang, Y., . . . Whiting, M. F. (2015). 300 million years of diversification: elucidating the patterns of orthopteran evolution based on comprehensive taxon and gene sampling. *Cladistics, 31*(6), 621-651. doi:10.1111/cla.12116

Stift, M., Kolar, F., & Meirmans, P. G. (2019). STRUCTURE is more robust than other clustering methods in simulated mixed-ploidy populations. *Heredity, 123*(4), 429-441. doi:10.1038/s41437-019-0247-6

Takahashi, T., Nagata, N., & Sota, T. (2014). Application of RAD-based phylogenetics to complex relationships among variously related taxa in a species flock. *Molecular Phylogenetics and Evolution, 80*, 137-144. doi:10.1016/j.ympev.2014.07.016

Taylor, E. B., Boughman, J. W., Groenenboom, M., Sniatynski, M., Schluter, D., & Gow, J. L. (2006). Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology, 15*(2), 343-355. doi:10.1111/j.1365-294X.2005.02794.x

Thome, M. T. C., & Carstens, B. C. (2016). Phylogeographic model selection leads to insight into the evolutionary history of four-eyed frogs. *Proceedings of the*

National Academy of Sciences of the United States of America, 113(29), 8010-8017. doi:10.1073/pnas.1601064113

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. doi:10.1111/mec.15189

Tzedakis, P. C., Emerson, B. C., & Hewitt, G. M. (2013). Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution, 28*(12), 696-704. doi:10.1016/j.tree.2013.09.001

Vera, M., Díez-del-Molino, D., & García-Marín, J. L. (2016). Genomic survey provides insights into the evolutionary changes that occurred during European expansion of the invasive mosquitofish (*Gambusia holbrooki*). *Molecular Ecology, 25*(5), 1089-1105. doi:10.1111/mec.13545

Virdee, S. R., & Hewitt, G. M. (1994). Clines for hybrid dysfunction in a grasshopper hybrid zone. *Evolution, 48*(2), 392-407. doi:10.2307/2410100

Wachter, G. A., Papadopoulou, A., Muster, C., Arthofer, W., Knowles, L. L., Steiner, F. M., & Schlick-Steiner, B. C. (2016). Glacial refugia, recolonization patterns and diversification forces in Alpine-endemic *Megabunus* harvestmen. *Molecular Ecology, 25*(12), 2904-2919. doi:10.1111/mec.13634

Wall, J. D., Yang, M. A., Jay, F., Kim, S. K., Durand, E. Y., Stevison, L. S., . . . Slatkin, M. (2013). Higher levels of Neanderthal ancestry in east Asians than in Europeans. *Genetics, 194*(1), 199-209. doi:10.1534/genetics.112.148213

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information

982 section at the end of the article.

983 Tables

984

985 **TABLE 1** Comparison of alternative migration models (detailed in Figure 2) tested
986 using FASTSIMCOAL2. For each model, the table shows the maximum likelihood
987 estimate of the model ($\log_{10}$L), the number of parameters (k), the Akaike's information
988 criterion score (AIC), the difference in AIC value of each model from that of the
989 strongest model (ΔAIC), and AIC weight (ω_i). Best-supported model (ΔAIC < 2) is
990 indicated in bold.

991

Model	k	$\log_{10}$L	AIC	ΔAIC	ω_i
Model 0	6	-3539.32	7090.64	90.01	0.00
Model 1	7	-3518.85	7051.71	51.08	0.00
Model 2	8	-3518.42	7052.84	52.21	0.00
Model 3	9	-3515.49	7048.98	48.34	0.00
Model 4	10	-3515.54	7051.09	50.45	0.00
Model 5	**10**	**-3490.32**	**7000.63**	**0.00**	**0.86**
Model 6	12	-3490.09	7004.19	3.56	0.14

992

993

994

995

TABLE 2 Parameters inferred from coalescent simulations with FASTSIMCOAL2 under the most supported demographic model (Model 5). Table shows point estimates and lower and upper 95% confidence intervals. Note that the effective population size of *O. uhagonii* is not presented in this table because it was fixed in FASTSIMCOAL2 analyses to enable the estimation of other parameters (see the Materials and Methods section for further details). θ, mutation-scaled effective population sizes; T_{DIV} and T_{ADM}, timing of population divergence and admixture, respectively (given in number of generations); m, migration rates per generation. Each specific parameter is illustrated in Figure 2.

Parameter	Point estimate	Lower Bound	Upper Bound
θ_{ANC}	702,036	376,562	832,461
θ_{MS}	813,263	686,748	890,000
θ_{MA}	1,556,235	1,293,951	1,675,243
T_{DIV1}	1,199,191	922,762	1,380,088
T_{DIV2}	1,382,958	1,278,828	1,670,203
T_{ADM1}	14,090	1632	23,626
T_{ADM2}	29,040	50,260	750,959
m_1	1.28×10^{-07}	9.02×10^{-08}	1.82×10^{-07}
m_2	2.41×10^{-07}	2.14×10^{-08}	1.37×10^{-07}

Legends to figures

FIGURE 1 Biogeographical setting of the study system. (A-B) Maps show the sampled populations and the distribution range of the two studied taxa based on our own species records (*O. minutissimus*: purple areas and squares; *O. uhagonii*: light orange areas and dots). *Omocestus minutissimus* presents a partially overlapping distribution with *O. uhagonii* in the Central System (deep purple) and allopatric populations (light purple) in eastern Iberia. Triangles indicate sampling localities where the two species were found living in sympatry. (C) Genetic assignment of individuals based on the results of FASTSTRUCTURE. Individuals are partitioned into *K* coloured segments representing the probability of belonging to the cluster with that colour and thin vertical black lines separate individuals from different populations. Population codes as in Table S1.

FIGURE 2 Alternative migration models tested using FASTSIMCOAL2. Parameters include ancestral (θ_{ANC}) and contemporary (θ_{US}, θ_{MS}, θ_{MA}) effective population sizes, timing of population split (T_{DIV}) and admixture (T_{ADM}), and migration rates (m) between different pairs of populations. Grey background highlights the most supported model.

FIGURE 3 Genetic assignment of *O. minutissimus* and *O. uhagonii* from the Central System based on the results of (A) FASTSTRUCTURE and (B) STRUCTURE. Individuals are partitioned into *K* coloured segments representing the probability of belonging to the cluster with that colour and thin vertical black lines separate individuals from different populations. Maps display the distribution of *O. uhagonii* and the probability of assignment of the populations of this species (spatial interpolation) to the genetic cluster

of *O. minutissimus* (i.e., the degree of introgression from *O. minutissimus* to *O. uhagonii*). Population codes as in Table S1.

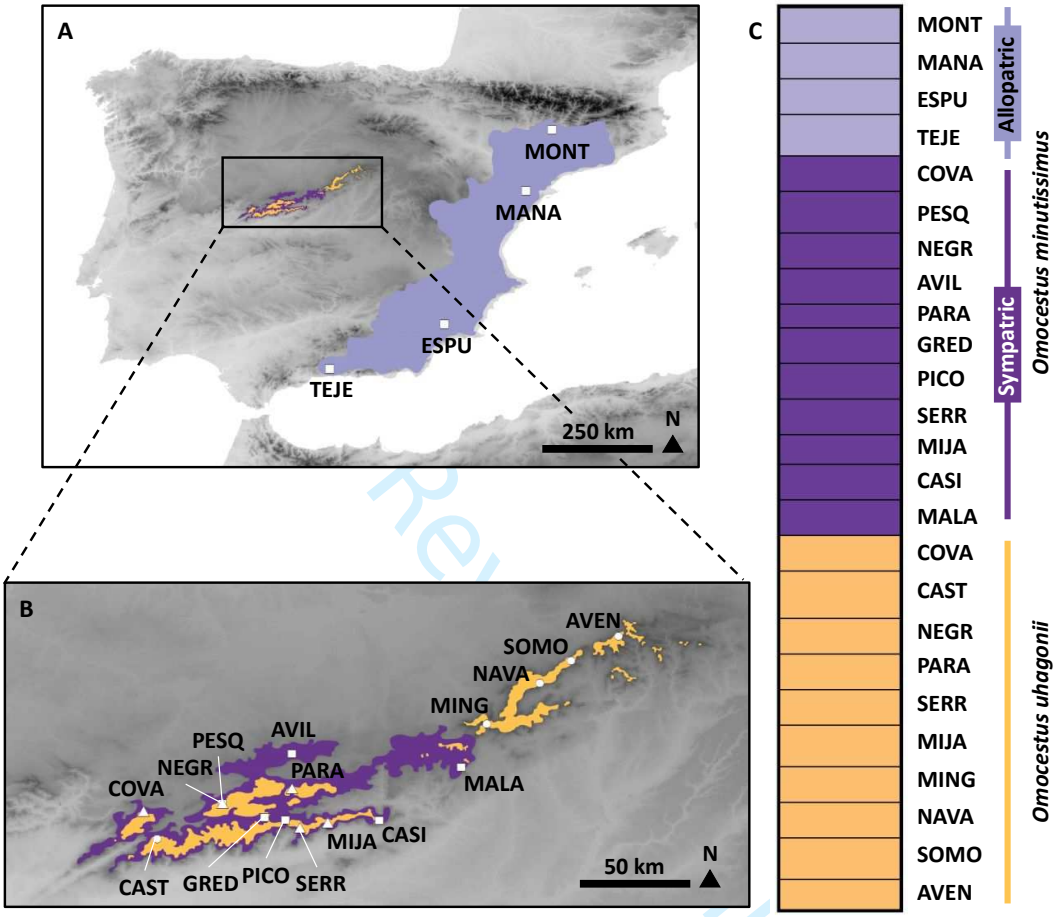
FIGURE 4 Maximum-likelihood tree inferred with TREEMIX for *O. uhagonii* (US) and sympatric (MS) and allopatric (MA) populations of *O. minutissimus*. The direction of gene flow (from MS to US) for the most likely migration event ($m = 1$) inferred is represented with an arrow colored according to the percentage of alleles (weight) originating from the source.

FIGURE 5 Schematic representation of the events documented in this study and the inferred biological processes. These correspond to the best-fit demographic model (Model 5) for *O. uhagonii* (US) and sympatric (MS) and allopatric (MA) populations of *O. minutissimus*. Vertical bars connecting MS and US represent historical gene flow. Note that geological reference time is not scaled and only point estimates inferred by FASTSIMCOAL2 are presented to simplify visualization.

1049 Figure 1

1050

1051

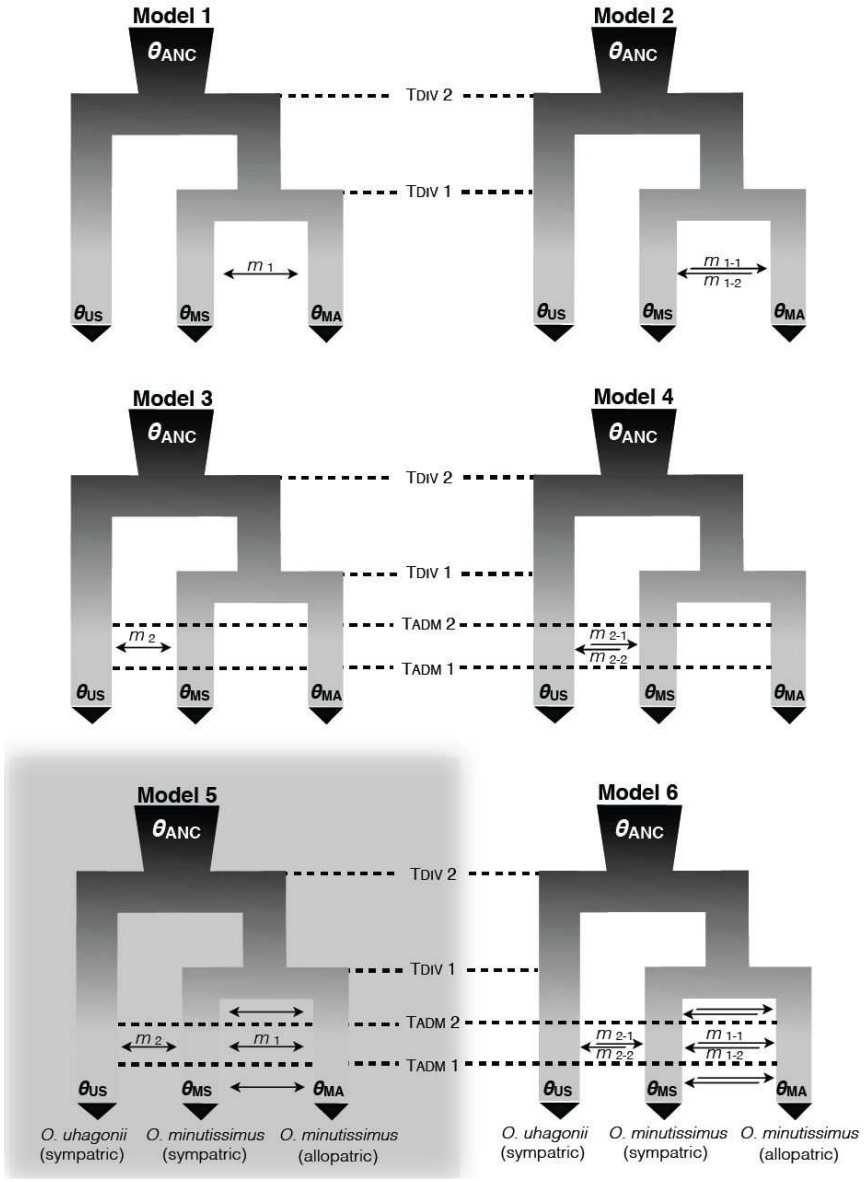


1052

1053

1054 Figure 2

1055

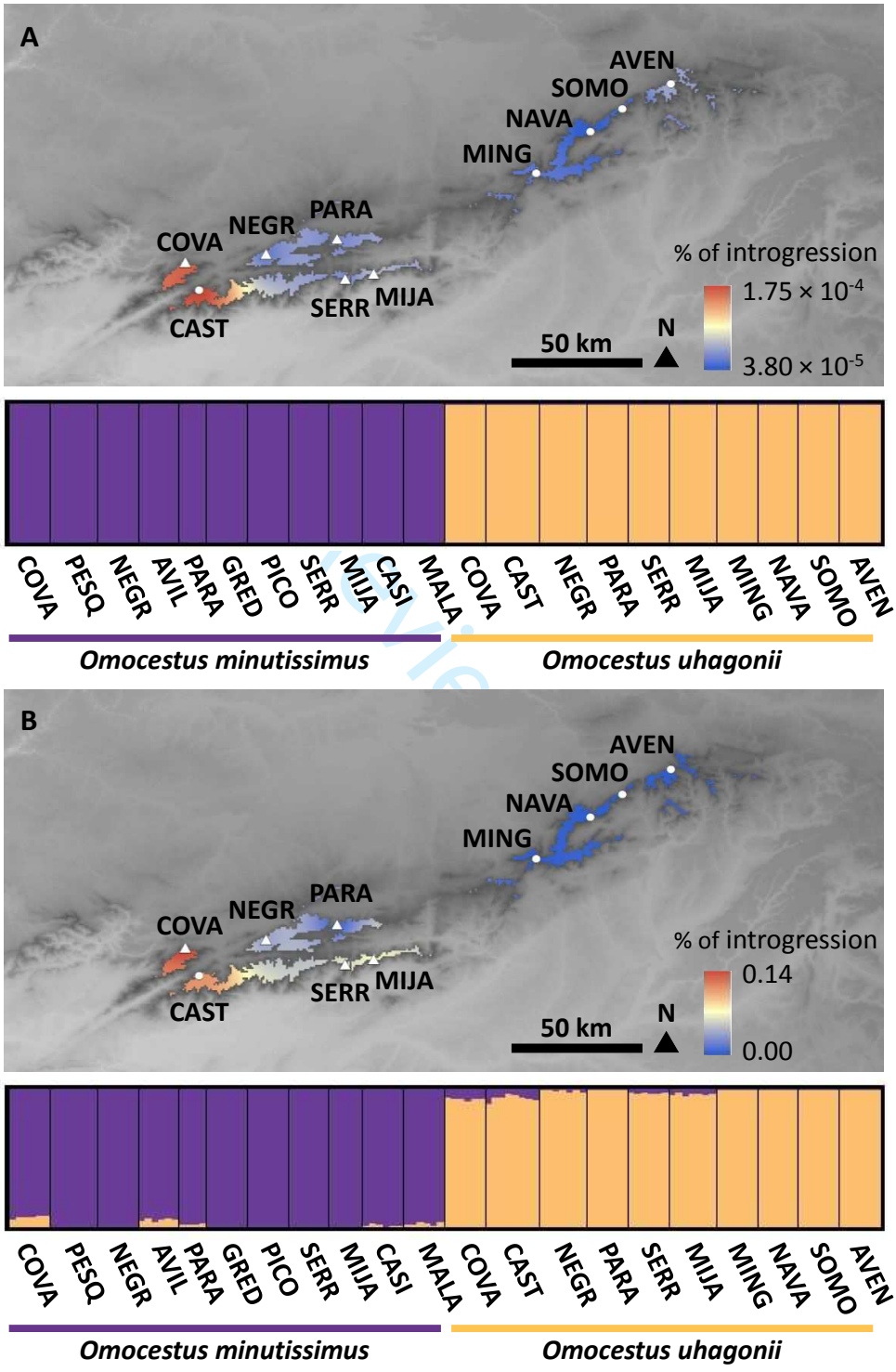


1056

1057

1058 Figure 3

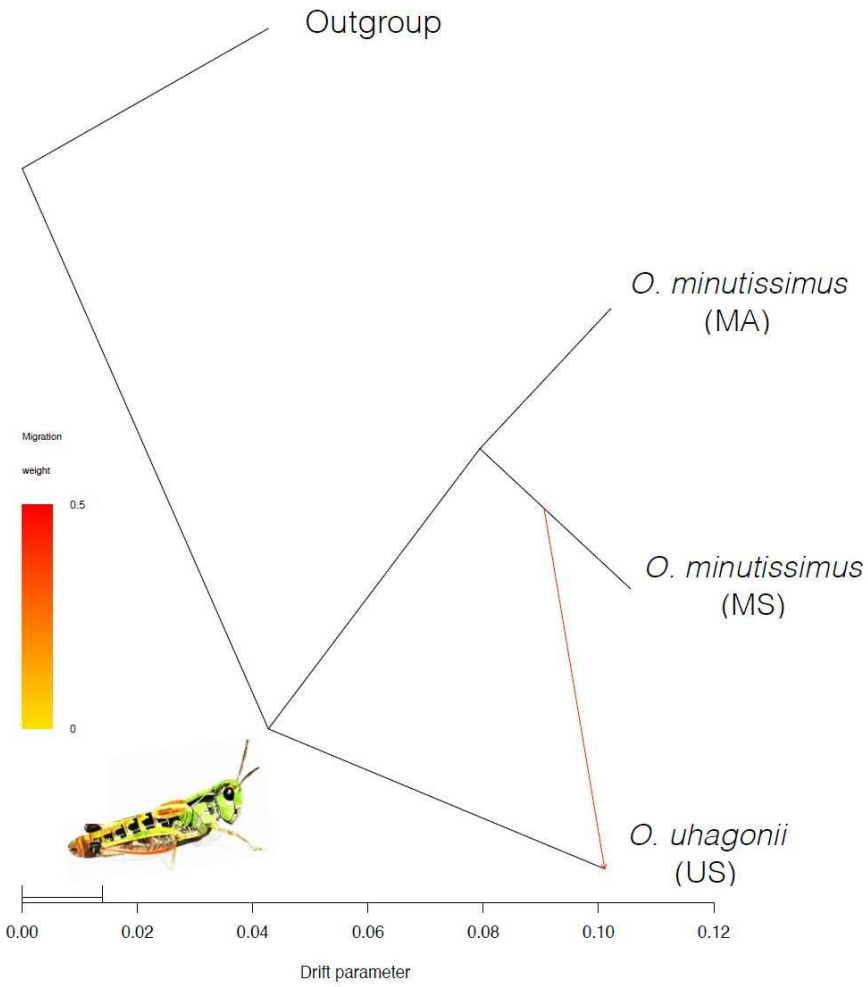
1059



1060

1061 Figure 4

1062

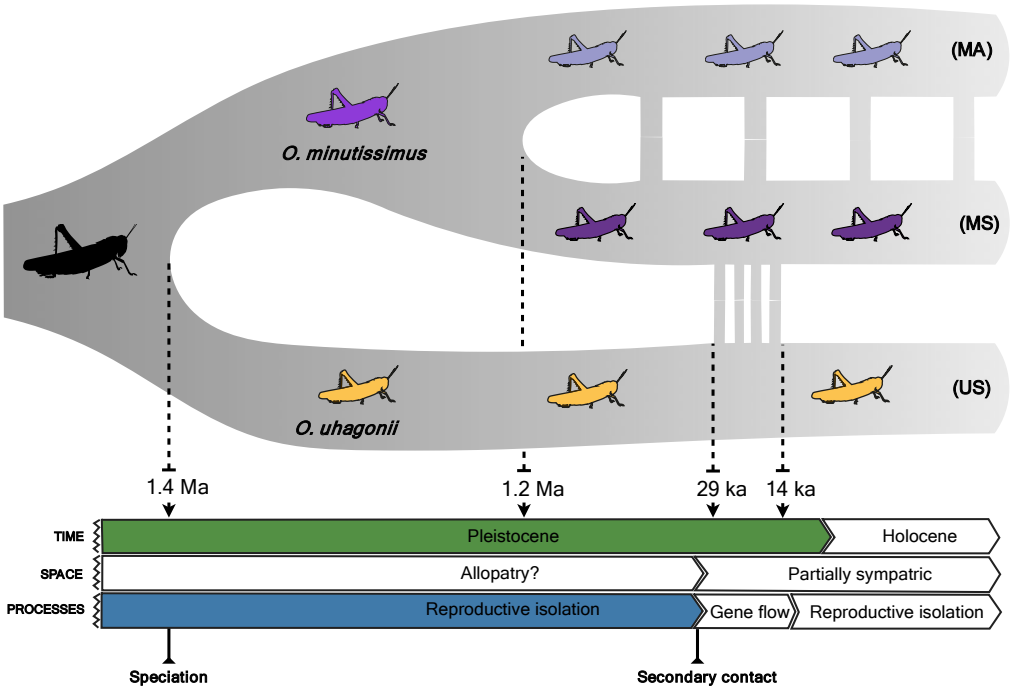


1063

1064

1065 Figure 5

1066



1067

1068

1069

1070