Received: 12 December 2018 | Revised: 14 July 2019 | Accepted: 15 July 2019

DOI: 10.1111/mec.15189

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Genomic data reveal deep genetic structure but no support for current taxonomic designation in a grasshopper species complex

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

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

[2]Department of Biological Sciences, University of Cyprus, Nicosia, Cyprus

Correspondence
Vanina Tonzo, Estación Biológica de Doñana, EBD-CSIC, Seville, Spain.
Email: vaninatonzo@gmail.com

Funding information
Ministerio de Economía y Competitividad, Grant/Award Number: BES-2015-73159; European Regional Development Fund (ERDF), Grant/Award Number: CGL2014-54671-P and CGL2017-83433-P; Severo Ochoa, Grant/Award Number: SEV-2012-0262; Ramon y Cajal, Grant/Award Number: RYC-2013-12501

Abstract

Taxonomy has traditionally relied on morphological and ecological traits to interpret and classify biological diversity. Over the last decade, technological advances and conceptual developments in the field of molecular ecology and systematics have eased the generation of genomic data and changed the paradigm of biodiversity analysis. Here we illustrate how traditional taxonomy has led to species designations that are supported neither by high throughput sequencing data nor by the quantitative integration of genomic information with other sources of evidence. Specifically, we focus on *Omocestus antigai* and *Omocestus navasi*, two montane grasshoppers from the Pyrenean region that were originally described based on quantitative phenotypic differences and distinct habitat associations (alpine vs. Mediterranean-montane habitats). To validate current taxonomic designations, test species boundaries, and understand the factors that have contributed to genetic divergence, we obtained phenotypic (geometric morphometrics) and genome-wide SNP data (ddRADSeq) from populations covering the entire known distribution of the two taxa. Coalescent-based phylogenetic reconstructions, integrative Bayesian model-based species delimitation, and landscape genetic analyses revealed that populations assigned to the two taxa show a spatial distribution of genetic variation that do not match with current taxonomic designations and is incompatible with ecological/environmental speciation. Our results support little phenotypic variation among populations and a marked genetic structure that is mostly explained by geographic distances and limited population connectivity across the abrupt landscapes characterizing the study region. Overall, this study highlights the importance of integrative approaches to identify taxonomic units and elucidate the evolutionary history of species.

KEYWORDS

ddRADseq, geometric morphometrics, integrative species delimitation, landscape genetics, phylogenomic inference, species delimitation

1 | INTRODUCTION

Describing biological diversity and understanding the ecological and evolutionary processes that generate it is at the core of molecular ecology and allied disciplines (Avise et al., 1987; Dobzhansky, 1937; Simpson, 1944). In the last decade, this field has experienced a revolution as a result of new technological and conceptual advances that allow biologists to discover new taxa at an unprecedented resolution and gain new insights into the processes leading to species formation (Jackson, Carstens, Morales, & O'Meara, 2017; Sites & Marshall, 2003; Wiens, 2007). These advances are timely, as the current biodiversity crisis—i.e., the accelerated loss of species and the rapid degradation of ecosystems—requires the establishment of efficient conservation policies that often put the focus on red lists of formally described taxa (Agapow et al., 2004; Mace, 2004). Yet, the definition of the fundamental biological unit that conforms a species remains controversial due to many alternative species concepts and the different properties/criteria used to define them (see de Queiroz, 2007). Given the inherent difficulties to test for reproductive isolation (i.e., the biological species concept; Dobzhansky, 1970; Mayr, 1942), especially in allopatric populations, most species descriptions have been traditionally based on qualitative or quantitative phenotypic differences (diagnosable or phenetic species concept; e.g., Nelson & Platnick, 1981; Sokal & Crovello, 1970) or distinct ecological traits (e.g., Jones & Weisrock, 2018 and references therein). However, speciation may not necessarily involve shifts in niche preferences or diagnosable morphological changes and adhering to these two sources of evidence could result in an underestimation of the number of biological species.

The wide use of DNA sequence data in taxonomy (e.g., Hebert, Cywinska, Ball, & DeWaard, 2003; Tautz, Arctander, Minelli, Thomas, & Vogler, 2003) and the application of phylogenetic/coalescent-based species delimitation approaches (e.g., Fujita, Leaché, Burbrink, McGuire, & Moritz, 2012; Pons et al., 2006; Yang & Rannala, 2010) have allowed the discovery of cryptic species and the resolution of many taxonomic uncertainties (e.g., Blair et al., 2019; Burbrink et al., 2011; Leavitt, Johnson, Goward, & St Clair, 2011; Niemiller, Near, & Fitzpatrick, 2012). In particular, the incorporation of coalescent theory under an integrative taxonomic framework has increased the statistical rigor of species delimitation and has helped to move away from subjective decisions on the degree of differentiation (e.g., phenotypic, genetic, ecological, etc) that is needed for considering different lineages or populations as distinct taxa (Fujita et al., 2012; Jones & Weisrock, 2018; Wiens & Servedio, 2000). In recent years, the advent of high throughput DNA sequencing technology has exponentially increased our capability to obtain large scale genomic data from non-model organisms, removing previous data-related constraints, while promoting further development of new phylogenetic inference methods and model-based approaches of species delimitation (van Dijk, Auger, Jaszczyszyn, & Thermes, 2014; Ekblom & Galindo, 2011; Fujita et al., 2012). In particular, the multispecies coalescent model (MCM) is considered an excellent approach to test alternative hypotheses of lineage divergence (Knowles & Carstens, 2007; Yang & Rannala, 2010) and identify boundaries among recently diverged species using multilocus data (Domingos, Colli, Lemmon, Lemmon, & Beheregaray, 2017; Rannala & Yang, 2013; Yang, 2015). More recent developments have incorporated the possibility of combining diverse sources of information (quantitative traits and genomic data) to objectively identify independently evolving lineages into an integrative analytical framework (Edwards & Knowles, 2014; Fujita et al., 2012; Solis-Lemus, Knowles, & Ane, 2015). However, integrative species delimitation approaches are not exempt of limitations (Huang, 2018; Sukumaran & Knowles, 2017). One of these limitations is the challenge to deal with sexually dimorphic traits (Solis-Lemus et al., 2015), which can have a considerable impact on species delimitation inferences (Chan et al., 2017; Noguerales, Cordero, & Ortego, 2018). More importantly, the finest detection of genetic differentiation brought by the highest resolution of genomic data can lead to a potential confusion of population genetic structure with species boundaries (Carstens, Pelletier, Reid, & Satler, 2013; O'Meara, 2010; Sukumaran & Knowles, 2017). That is, in some situations, the higher power of genomic data to detect population structuring may drive to an artefactual increase in the number of inferred species (Fujita et al., 2012; Huang, 2018; Isaac, Mallet, & Mace, 2004). Taxonomic inflation, in turn, can have a negative impact on subsequent ecological and evolutionary studies, compromise our ability to reach solid conclusions on the origin and dynamics of biodiversity and, ultimately, lead to misguided conservation policies and inadequate management practices (Carstens et al., 2013; O'Meara, 2010; Sukumaran & Knowles, 2017).

As mentioned above, a common source of taxonomic controversy is the historical description of species based on the concurrence of ecological and phenotypic distinctiveness (e.g., Dowle, Morgan-Richards, & Trewick, 2014; Jones & Weisrock, 2018). These are often presumed to represent cases of ecological speciation, i.e., the formation of new species when divergent natural selection under contrasting environmental conditions (e.g., elevation, salinity, etc) leads to reproductive isolation (Rundle & Nosil, 2005; Schluter, 2001). Genetic data have often supported classic taxonomic studies separating species based on ecological and phenotypic differences (e.g., Lamichhaney et al., 2015). However, in many other cases the recognized taxonomic units are not consistent with patterns of genome-wide differentiation (e.g., Dowle et al., 2014; Jones & Weisrock, 2018), suggesting that the link between environmental and phenotypic divergence is a consequence of plasticity or divergent selection at a few genes or genomic islands involved in local adaptation processes that have not led to the reduction of gene flow (e.g., Mason & Taylor, 2015; Soria-Carrasco et al., 2014). An intermediate situation along the speciation continuum is the frequently reported association between environmental dissimilarity and phenotypic and genetic differentiation found in many studies (Sexton, Hangartner, & Hoffmann, 2014; Shafer & Wolf, 2013; Wang & Bradburd, 2014). This pattern of progressive genetic and phenotypic differentiation among populations experiencing contrasting environmental conditions has been termed isolation-by-ecology (IBE) and is generally interpreted as a signal of ongoing local

adaptation processes and incipient speciation (Shafer & Wolf, 2013). Thus, the study of these evolutionary phenomena is not only relevant in the context of resolving taxonomic uncertainties, but it also allows to identify important ecological and evolutionary processes along the speciation continuum that might deserve to be protected (Moritz, 2002).

The grasshopper subgenus *Dreuxius* Defaut, 1988 (genus *Omocestus*, Bolívar, 1878) is a complex of nine recently diversified species distributed in the Iberian Peninsula (six species) and northwestern Africa (three species; Cigliano, Braun, Eades, & Otte, 2019; García-Navas, Noguerales, Cordero, & Ortego, 2017). Most taxa have allopatric distributions and are often isolated at high elevations in different mountain systems (Cigliano et al., 2019). The putative sister species pair *Omocestus antigai* (Bolívar, 1897) and *Omocestus navasi* (Bolívar, 1908) are distributed through the Pyrenees, including some populations in the pre-Pyrenees and Catalan Pre-Coastal ranges (Poniatowski, Defaut, Llucia-Pomares, & Fartmann, 2009). These two species were originally described based on distinct habitat associations and certain quantitative phenotypic traits such as body size and forewing shape (Clemente, García, Arnaldos, Romera, & Presa, 1999; Olmo-Vidal, 2002; Puissant, 2008). The current known ranges of both species follow the west to east orientation of the Pyrenees, but their populations are mostly allopatric. *Omocestus antigai* is distributed at high elevations (1,450–2,350 m) associated with alpine or subalpine grasslands above the tree line, whereas *O. navasi* is circumscribed to Mediterranean scrubby habitats at lower elevations (700–1,600 m; Llucia-Pomares, 2002; Olmo-Vidal, 2002; Poniatowski et al., 2009). These strong ecological differences and subtle morphological variation have motivated different authors to synonymize the two taxa (Ragge & Reynolds, 1998; Reynolds, 1986), revalidate their distinct species status (Clemente et al., 1999) and even describe a new subspecific taxon (*O. navasi bellmanni*; Puissant, 2008) in the last decades (Cigliano et al., 2019). Resolving the taxonomy of the complex also has important implications for conservation and management, as the two putative species are included in the IUCN red list of threatened species but with different assessment categories (Braud, Hochkirch, Fartmann, et al., 2016; Braud, Hochkirch, Presa, et al., 2016). *Omocestus navasi* is classified as "endangered" due to its extremely small estimated area of occupancy (60–200 km^2) and the continuous decline in the extend and quality of habitats (Braud, Hochkirch, Fartmann, et al., 2016). In contrast, the high-elevation grasslands occupied by *O. antigai* are supposed to be less affected by direct human impacts and this taxon was assessed as "vulnerable", with the main threats identified being the small size and high degree of fragmentation of its populations, overgrazing in some areas, and the decline of the quality of its specific alpine and subalpine habitats due to the progressive abandonment of traditional grassland management practices (Braud, Hochkirch, Presa, et al., 2016). Overall, this makes this complex an excellent model system to test species boundaries and understand the factors that have contributed to genomic, phenotypic and ecological divergence, which will in turn provide the necessary baseline information to determine the conservation status of the different taxonomic units within the complex and design proper on-ground management practices.

In this study, we generated genomic data using restriction-site-associated DNA sequencing (ddRADseq; Peterson, Weber, Kay, Fisher, & Hoekstra, 2012) and obtained phenotypic information through geometric morphometric analyses (Bookstein, 1992) to shed light on the taxonomy and evolutionary history of *O. antigai* and *O. navasi*. First, we tested the hypothesis that the current taxonomic designation of the complex is supported by genomic and phenotypic data. In particular, we used genome-wide single nucleotide polymorphisms (SNP) data and alternative coalescent-based methods to infer the phylogenetic relationships among all sampled populations of the two taxa. Then, we compared the taxonomic status of the studied populations with phylogenomic reconstructions and inferred patterns of spatial genetic structure (Dayrat, 2005; Jones & Weisrock, 2018), while we employed a suite of Bayesian model-based species delimitation approaches, one of them integrating genomic data and quantitative traits, to identify independently-evolving lineages (Leaché, Fujita, Minin, & Bouckaert, 2014; Solis-Lemus et al., 2015; Yang & Rannala, 2010). Second, we tested the hypothesis that ecological-driven divergence is the main process underlying spatial patterns of genomic and phenotypic variation in the study system. In particular, we analyzed whether environmental dissimilarity among populations and processes of local adaptation to different elevational and climatic gradients are the main factors shaping genetic differentiation and phenotypic trait variation in the study system or if, on the contrary, genetic and phenotypic structure is mostly explained by the geographical distance among populations and resistance distances defined by the complex topography of the study area.

2 | MATERIALS AND METHODS

2.1 | Population sampling

Between 2012 and 2015, we collected specimens of *Omocestus antigai* (seven populations), *O. navasi navasi* (eight populations) and *O. navasi bellmanni* (one population) from 16 sampling localities (hereafter referred to as populations; Table 1 and Figure 1). Samples represent populations across the entire known distribution range of the three taxa based on the literature (Clemente et al., 1999; Llucia-Pomares, 2002; Poniatowski et al., 2009; Puissant, 2008). We stored specimens in 2 ml vials with 96% ethanol and preserved them at −20°C until needed for geometric morphometric and genomic analyses.

2.2 | Geometric morphometric analyses

We selected a maximum of 10 adult individuals from each population (five males and five females if available) and used landmark-based geometric morphometric analyses to characterize phenotypic variation in the complex. We excluded from geometric

TABLE 1 Geographical location, elevation and number of genotyped individuals (*n*) for each sampled population (locality) of the three putative taxa and the outgroup

Species	Locality	Code	Latitude	Longitude	Elevation (m)	*n*
O. antigai	Boí Taüll	BOI	42.47945	0.86732	2,046	5
O. antigai	Llessui	LLE	42.44554	1.03668	1,979	6
O. antigai	Coll de la Creueta	CRE	42.30131	1.99350	1,928	6
O. antigai	Err	ERR	42.38830	2.08812	2,150	3
O. antigai	Queralbs	QUE	42.36627	2.14902	2,070	6
O. antigai	Turó de l'Home	TUR	41.77167	2.44335	1,622	7
O. antigai	Setcases	SET	42.42756	2.26630	2,145	7
O. navasi navasi	Borau	BOR	42.67388	−0.57932	1,357	7
O. navasi navasi	Puerto de Monrepós	MON	42.35067	−0.39762	1,267	5
O. navasi navasi	Nerín	NER	42.59047	0.01062	1,623	7
O. navasi navasi	Saravillo	SAR	42.56089	0.23046	1,313	2
O. navasi navasi	Chía	CHI	42.55418	0.43703	1,691	7
O. navasi navasi	Espés	ESP	42.44307	0.58389	1,361	6
O. navasi navasi	Perves	PER	42.35250	0.83709	1,370	6
O. navasi navasi	Montan de Tost	TOS	42.23790	1.38052	1,182	7
O. navasi bellmanni	Tour de Batère	BAT	42.50785	2.57641	1,430	7
O. femoralis	Sierra de Espuña	ESP	37.86515	−1.57115	1,514	6

morphometric analyses five populations (MON, SAR, ERR, QUE, and TUR) for which only a few individuals for one or the two sexes (<3) could be sampled due to low population densities and/or because most specimens presented damaged structures (e.g., broken forewings or ovipositor valves). Populations with available geometric morphometric data included all taxa within the complex (i.e., all species/subspecies) and were representative of all genetic clusters inferred by Bayesian clustering analyses (see Section 3). We took digital images of head, pronotum, forewing and ovipositor valves (in females) with a ZEISS stereomicroscope (STEREO DISCOVERY version 8; Carl Zeiss Microscopy GmbH). These traits correspond to those originally used to distinguish the two putative species (Clemente et al., 1999). The coordinates of landmarks (9–14 landmarks per trait) were mapped on the images using TPSDIG version 2.2 (Rohlf, 2015) and analyzed as implemented in the R version 3.3.2 (R Core Team, 2018) package *Geomorph* (Adams & Otarola-Castillo, 2013). Specifically, we performed generalized Procrustes analyses (GPA) separately for each sex to remove the effects of location, size, and rotation of the relative positions of landmarks among specimens using the function *gpagen*. This superimposition method minimizes the sum-of-squared distances between landmarks across samples (Rohlf & Slice, 1990). To examine shape differences between taxa and among populations, we performed principal component analyses (PCA) on the covariance matrix of aligned Procrustes shape coordinates using the function *plotTangentSpace*. We retained the scores from the two first principal components (PC1 and PC2) for each trait and used multivariate analyses of variance (MANOVA) to investigate if trait shapes were significantly different between taxa and among sampled populations (e.g., Adams & Otarola-Castillo, 2013; Friedline et al., 2019). Statistical significance of MANOVA models was determined using Wilk's λ as the test statistic and $\alpha = 0.05$ as a significance threshold. Then, we used one-way analyses of variance (ANOVA) to assess differences between taxa and among sampled populations separately for each phenotypic trait (e.g., Friedline et al., 2019). The first two PCs for each trait were also used for integrative species delimitation analyses using IBPP (Solis-Lemus et al., 2015) and to build matrices of phenotypic differentiation (P_{ST}; see below for details).

2.3 | Genomic library preparation and processing

We obtained genomic data for a subset of 94 specimens representative from the 16 sampled populations of *O. antigai*, *O. navasi navasi*, and *O. navasi bellmanni* (~six individuals per population, three males and three females when available; Table 1). Additionally, we genotyped six individuals of *O. femoralis* Bolívar, 1908 (Table 1), a species also belonging to the subgenus *Dreuxius* (Cigliano et al., 2019) and that was used as an outgroup in some phylogenetic and species delimitation analyses (see below for details). Details on the preparation of ddRADseq libraries (Peterson et al., 2012) are presented in Methods S1. Raw sequences were demultiplexed and preprocessed using STACKS version 1.35 (Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011; Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) and assembled using PYRAD version 3.0.66 (Eaton, 2014). Methods S2 provides all details on sequence assembling and data filtering.

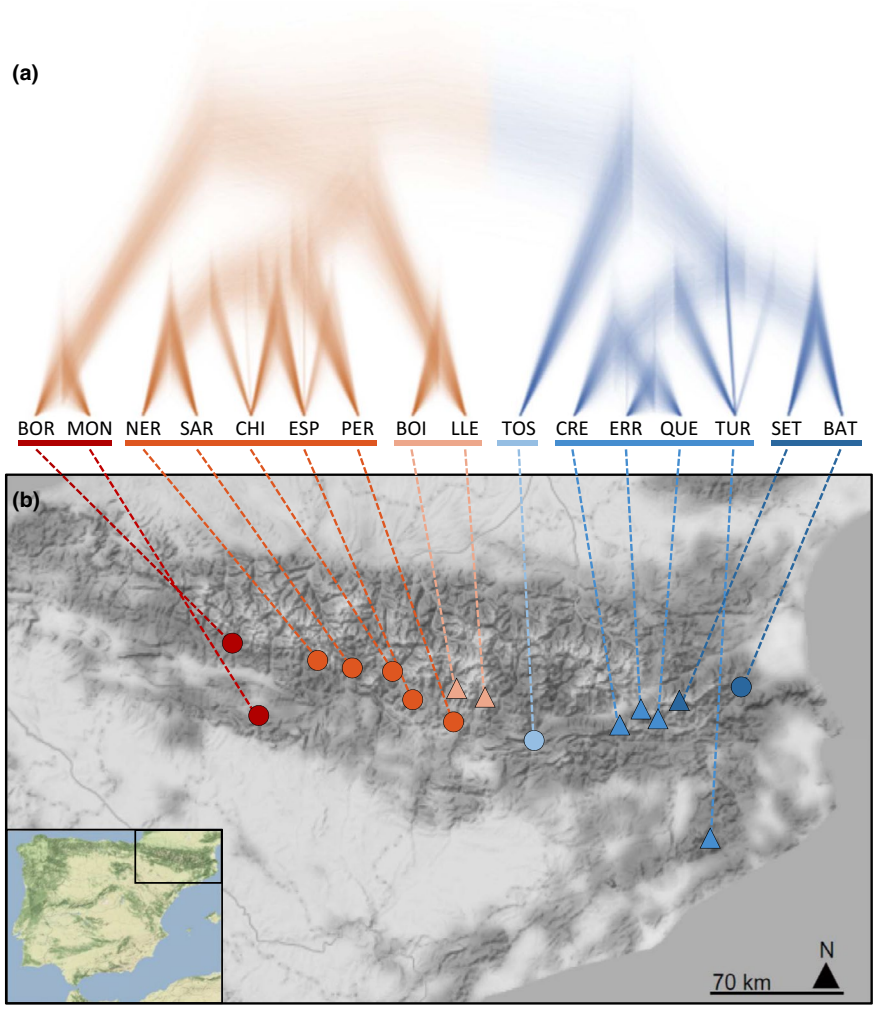

FIGURE 1 (a) Bayesian phylogenetic tree reconstructed with SNAPP (node support presented in Figure 5) and (b) map showing the distribution of the sampled populations of *Omocestus antigai* (triangles) and *Omocestus navasi* (dots) in the Pyrenees. Colours of triangles and dots on the map indicate the main genetic cluster at which populations were assigned according to STRUCTURE analyses for $K = 6$. Population codes as in Table 1 [Colour figure can be viewed at wileyonlinelibrary.com]

2.4 | Population genetic structure

We employed four complementary approaches to exhaustively explore spatial patterns of genetic structure in our study system, including: (a) classic STRUCTURE analyses (Pritchard, Stephens, & Donnelly, 2000); (b) FASTSTRUCTURE (Raj, Stephens, & Pritchard, 2014); (c) the recently developed spatial method implemented in the R program CONSTRUCT to infer patterns of genetic structure after controlling for the geographic distance separating the sampled populations (Bradburd, Coop, & Ralph, 2018), given the strong signal of isolation by distance in our study system (see Section 3.6); and (d) a principal component analysis (PCA; Jombart, 2008). Further details on these analyses are presented in Methods S3.

2.5 | Phylogenetic inference

We estimated species trees using two coalescent-based methods, SNAPP version 1.3 (Bryant, Bouckaert, Felsenstein, Rosenberg, & RoyChoudhury, 2012) as implemented in BEAST2 version 2.4.3 (Bouckaert et al., 2014) and SVDQUARTETS (Chifman & Kubatko, 2014) as implemented in PAUP* v. 4.0a152 (Swofford, 2002). SNAPP analyses are computationally highly demanding and, for this reason, we only selected one individual (the one with the highest number of retained reads; Figure S1) from each of the 16 sampled populations. The resulting data set contained 16 individuals and retained 858 polymorphic sites. We ran SNAPP analyses for 1,000,000 Markov chain Monte Carlo (MCMC) generations, sampling every 1,000 steps, and using different gamma prior distributions for *alpha* and *beta* (2, 200; 2, 2,000; 2, 20,000). The forward (u) and reverse (v) mutation rates were set to be calculated by SNAPP and we left the remaining parameters at default values. We conducted two independent runs and evaluated convergence with TRACER version 1.6. We removed 10% of trees as burnin and merged tree and log files from the different runs using LOGCOMBINER version 2.4.1. We used TREEANNOTATOR version 1.8.3 to obtain maximum credibility trees, TREESETANALYSER version 2.4.1 to identify species trees that were contained in the 95% highest posterior density (HPD), and DENSITREE version 2.2.1 (Bouckaert, 2010) to visualize the posterior distribution of trees.

SVDQUARTETS uses SNP data to infer phylogenetic relationships between quartets of taxa under the multispecies coalescent (MSC) model and then assembles these quartets into a species tree (Chifman & Kubatko, 2014). We performed SVDQUARTETS analyses including all individuals and populations and setting *O. femoralis* as an outgroup. We constructed species trees by exhaustively evaluating

Species delimitation hypothesis (H_i)	Species	MLE	2 × ln BF	Rank
H_1: (BAT + PER + SAR + CHI+TOS + NER+ESP + MON+BOR + CRE + ERR + BOI+LLE + QUE + TUR + SET)	1	−4,794.85	8.21×10^2	2
H_2: (BOI + LLE + CRE + ERR + QUE + TUR+SET) (PER + SAR + CHI + NER + ESP + MON + BOR + BAT + TOS)	2	−5,601.22	2.43×10^3	5
H_3: (BAT) (PER + SAR + CHI + TOS + NER + ESP + MON + BOR) (CRE + ERR + BOI + LLE + QUE + TUR + SET)	3	−5,384.07	2.00×10^3	4
H_4: (BAT + TOS + CRE + ERR + QUE + TUR + SET) (PER + SAR + CHI + NER + ESP + MON + BOR + BOI + LLE)	2	−5,221.06	1.67×10^3	3
H_5: (BAT + SET) (TOS) (CRE + ERR + QUE + TUR) (PER + SAR + CHI + NER + ESP) (MON + BOR) (BOI + LLE)	6	−4,384.48	—	1

TABLE 2 Results of BFD* analyses testing the support of competing species delimitation hypotheses. The table shows the clustering scheme defining each alternative species delimitation hypothesis (H_i). For each hypothesis, we show marginal likelihood estimates (MLE), their Bayes factors (calculated as 2 × ln BF) and their rank. Population codes as in Table 1

all possible quartets (i.e., every combination of four tips was examined) from the entire genomic data set (63,860 unlinked SNPs) and used 100 nonparametric bootstrapping replicates to assess branch support (Chifman & Kubatko, 2014).

2.6 | Species delimitation

We applied three Bayesian coalescent-based species delimitation approaches to determine the number of independently evolving lineages, two of them based only on molecular data (BPP; Yang & Rannala, 2010; BFD*; Leaché et al., 2014) and the third one integrating molecular and phenotypic data (IBPP; Solis-Lemus et al., 2015).

We used the BFD* method implemented in SNAPP (Leaché et al., 2014) to contrast five competing species delimitation hypotheses, including a single-species model and four alternative multi-species models informed by the current taxonomy as well as the phylogenetic and Bayesian clustering analyses (Table 2). This method allows the comparison of alternative species delimitation scenarios in an explicit MSC framework by calculating and comparing marginal likelihood estimates (MLE) for each model. Specifically, our species delimitation hypotheses included: (a) the hypothesis of a single species (H_1); (b) the current taxonomy of two species (H_2: *O. antigai* and *O. navasi*); (c) the current taxonomy but considering the subspecies *O. navasi bellmanni* as a separate species (H_3: three species); (d) a model considering as distinct species the two main genetic clusters (east-west split) revealed by phylogenetic and Bayesian clustering analyses (H_4) and (e) a model considering six species, corresponding to the six genetic clusters inferred by Bayesian clustering analyses in STRUCTURE (H_5; Table 2). We conducted a path sampling analysis of eight steps each consisting of 100,000 MCMC iterations (after a preburnin of 10,000 iterations), sampling each 100 steps, and using an α value of 0.3. These settings were sufficient to ensure convergence and obtain ESS > 200. The Bayes factor (BF) test statistic (2 × ln BF) was calculated, where BF is the difference in MLE between two competing models (base scenario-alternative scenario). These analyses were performed using a matrix with no missing data (393 SNPs; see Leaché, McElroy, & Trinh, 2018 for a similar approach) and including two individuals from each sampled population plus two individuals from the outgroup *O. femoralis*, to be able to test for the single-species model.

BPP version 3.3 can perform species delimitation analyses using a fixed input guide tree specified by the user (option A10, guided analysis) or estimating a species tree (option A11, unguided analysis; Yang, 2015; Yang & Rannala, 2014). We used both options and for guided analyses we fitted as input trees the topologies yielded by SNAPP and SVDQUARTETS (see Section 3). This allowed us to assess the potential impact that alternative tree topologies had on our species delimitation inferences (Leaché & Fujita, 2010). To make BPP analyses computationally tractable, we only included two individuals from each sampled population (same as used for BFD* analyses) and used a subset of loci from our total genomic data set (e.g., Huang, 2018; Noguerales et al., 2018; Rancilhac et al., 2019). Specifically, we selected loci with five or more variable sites, as these provide more power in species delimitation analyses than less variable loci (Huang, 2018). To evaluate the impact that the number of employed loci had on our species delimitation inferences, we used three random subsets of loci (200, 500 and 1,000 loci). According to previous studies, these numbers of loci provide considerable power for species delimitation analyses (e.g., Huang, 2018; Noguerales et al., 2018). We created BPP input files from the.loci output file from PYRAD and randomly selected the different subsets of loci with the specified minimum number of variable sites (≥5) using two R scripts (*bpp_convert_Ama_sp.r* and *numvarfrombppfile.r*, respectively) written by Huang (2018). We analyzed the impact of different demographic scenarios corresponding to different prior combinations of gamma distribution as in Huang and Knowles (2016). We considered four parameter settings: $\theta = G(1, 10)$ and $\tau = G(1, 10)$, $\theta = G(1, 10)$ and $\tau = G(2, 2{,}000)$, $\theta = G(2, 2{,}000)$ and $\tau = G(1, 10)$, and $\theta = G(2, 2{,}000)$ and $\tau = G(2, 2{,}000)$, where θ and τ refer to the ancestral population

sizes and divergence times, respectively. We ran four replicates for each combination. We used an automatic adjustment of the fine-tune parameters, allowing swapping rates to range between 0.30 and 0.70 (Yang, 2015). We ran each analysis for 100,000 generations, sampling every 10 generations (10,000 samples), after a burnin of 100,000 generations.

The integrative IBPP approach was built upon the architecture of the early version of BPP version 2.1.2 (Rannala & Yang, 2013; Yang & Rannala, 2010) and incorporates models of evolution for continuous quantitative traits under a Brownian motion process (Solis-Lemus et al., 2015). We ran IBPP analyses separately for each sex, incorporating genetic data as well as geometric morphometric data (PC1 and PC2) of three traits for males and four traits for females (see Section 2.2). As in BPP, we ran three independent analyses using as guide trees the topologies yielded by BPP (obtained using option A01; Yang, 2015), SNAPP and SVDQUARTETS, used four prior combinations for gamma distribution and same settings for all other parameter values, and ran four independent replicates for each scenario. These analyses were performed only for populations with available phenotypic data (i.e., excluding MON, SAR, ERR, QUE, and TUR), using a random subset of 500 loci with at least five variable sites, and including five individuals/population (rather than two individuals/population as in BPP analyses) in order to increase the accuracy of estimates of the phenotypic variation of populations. We ran IBPP analyses for 250,000 generations, sampling every 10 generations (25,000 samples), after a conservative burnin of 300,000 generations. Finally, we repeated these analyses considering only phenotypic data for each sex. Posterior probability (PP) of both BPP and IBPP models was considered well supported when PP > 0.95.

2.7 | Landscape genetic and phenotypic analyses

We analyzed three potential factors that could have shaped genetic (F_{ST}) and phenotypic (P_{ST}) differentiation among populations: (a) geographical distance; (b) environmental heterogeneity; (c) and isolation-by-resistance defined by topographic complexity. We estimated genetic differentiation between each pair of sampled populations calculating pairwise F_{ST} values in ARLEQUIN version 3.5 (Excoffier, Laval, & Schneider, 2005). Similarly, we estimated phenotypic differentiation between each pair of sampled populations calculating pairwise P_{ST} values for the first two PCs of each trait using the R package *Pstat* (da Silva & da Silva, 2018). We calculated the geographical distance between sampled populations using GEOGRAPHIC DISTANCE MATRIX GENERATOR version 1.2.3 (http://biodiversityinformatics.amnh.org/open_source/gdmg). To estimate environmental dissimilarity between each pair of sampled populations, we extracted for each one the values of the 19 present-day bioclimatic variables available in WORLDCLIM, downloaded at 30 arc-s (c. 1 km) resolution. Then, we used the "*rda*" function in the R package *vegan* version 2.4-4 (Oksanen et al., 2017) to perform a principal component analysis (PCA) and obtain for each population the PC scores of the first two PCs, which explained the 80% and 16% of the environmental variance, respectively. Afterward, we calculated environmental dissimilarity between each pair of populations using Euclidean distances for the obtained PC scores using the "*dist*" function in R. To investigate the role of topographic complexity, we calculated the slope for each cell from a 90 m resolution digital elevation model from NASA Shuttle Radar Topographic Mission (SRTM Digital Elevation Data) and the final raster layer was transformed to 30 arc-s (c. 1 km) resolution using QGIS version 2.8. Then, based on this layer, we used CIRCUITSCAPE version 4.0 (McRae, 2006; McRae & Beier, 2007) to calculate a matrix of resistance distances between all pairs of populations considering an eight-neighbour cell connection scheme.

We used multiple matrix regressions with randomization (MMRR) as implemented in R (Wang, 2013) to analyze: (a) pairwise population genetic differentiation (F_{ST}) in relation with geographical, resistance and environmental distances, and (b) pairwise population phenotypic differentiation (P_{ST}) in relation with genetic differentiation (F_{ST}) and geographical, resistance and environmental distances. Models for all dependent variables were initially constructed with all explanatory terms fitted (i.e., full models) and final models were selected using a backward-stepwise procedure, by progressively removing nonsignificant variables (starting with the least significant ones) until all retained terms within the model were significant. Then, we tested the significance of the rejected terms against this model to ensure that no additional variable reached significance. The result was the minimal most adequate model for explaining the variability in the response variable, where only the significant explanatory terms were retained. As we tested a high number of phenotypic traits (eight traits for females and six for males) against the same subset of independent variables, we applied a Benjamini-Hochberg false discovery rate correction to adjust *p*-values for multiple statistical tests and calculate the corresponding *q*-values using the *p.adjust* function in R.

3 | RESULTS

3.1 | Geometric morphometric analyses

The two first principal component scores, PC1 and PC2, for each trait explained most variation in shapes for pronotum (♀: 87%; ♂: 76%), head (♀: 90%; ♂: 91%), forewing (♀: 65%; ♂: 77%) and ovipositor valves (♀: 95%). Principal component plots for each trait and sex are shown in Figures S2 and S3. MANOVAs based on PC scores for all traits showed significant differences among sampled populations (♀: Wilk's $\lambda = 0.035$, $F_{80, 192.50} = 1.674$, $p = .002$; ♂: Wilk's $\lambda = 0.047$, $F_{54, 142.27} = 2.16$, $p < .001$) but not between the two taxa (♀: Wilk's $\lambda = 0.940$, $F_{8, 38} = 0.305$, $p = .959$; ♂: Wilk's $\lambda = 0.745$, $F_{6, 35} = 2.00$, $p = .092$). One-way ANOVAs revealed that multivariate differentiation among populations was driven by shape variation of forewing (PC2: $F_{10, 40} = 4.86$; $p < .001$) and ovipositor valves (PC1: $F_{10, 42} = 2.25$; $p = .03$) in females and head (PC2: $F_{10, 37} = 2.22$; $p = .039$), pronotum (PC1: $F_{10, 40} = 3.08$; $p = .005$; PC2: $F_{10, 40} = 2.47$; $p = .021$) and forewing (PC1: $F_{10, 36} = 2.75$; $p = .013$; PC2: $F_{10, 36} = 2.33$; $p = .031$) in males.

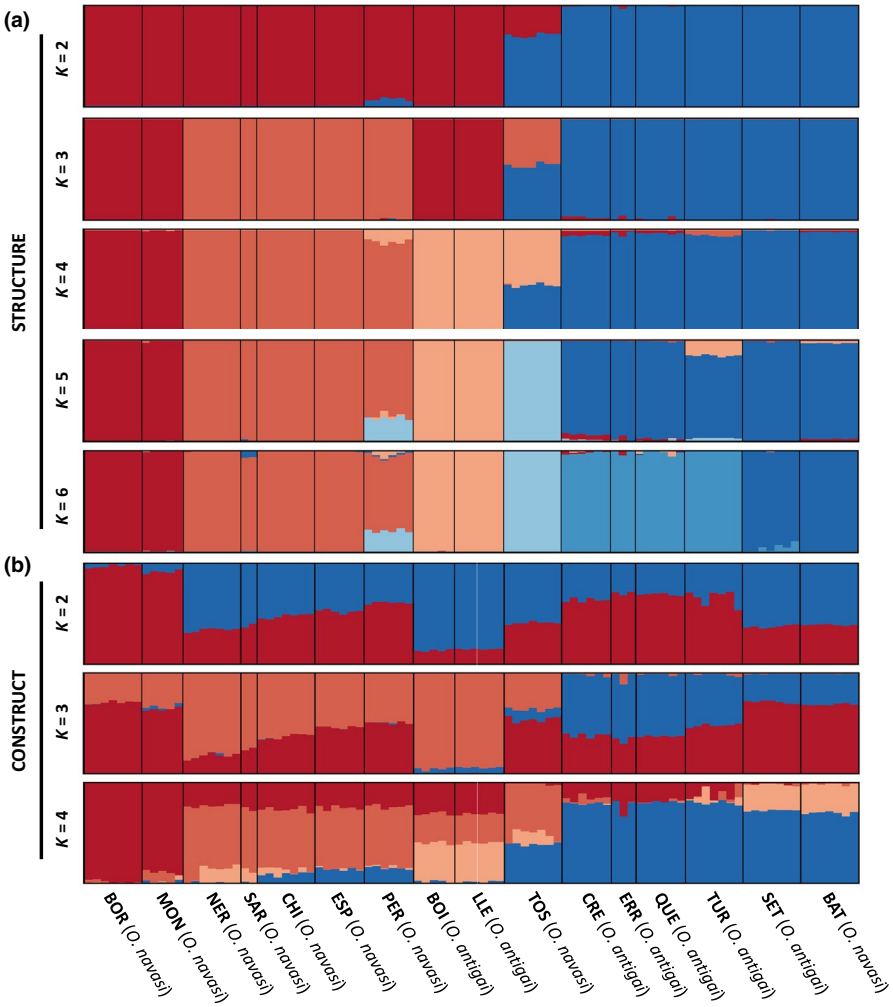

FIGURE 2 Genetic assignment of individuals based on the results of (a) STRUCTURE (from $K = 2$ to $K = 6$) and (b) CONSTRUCT (from $K = 2$ to $K = 4$). Individuals are partitioned into K coloured segments representing the probability of belonging to the cluster with that colour. Thin vertical black lines separate individuals from different populations. Population codes as in Table 1 [Colour figure can be viewed at wileyonlinelibrary.com]

3.2 | Genomic data

Illumina sequencing returned an average of 2.19×10^6 reads per sample. After quality control, an average of 2.11×10^6 reads were retained per sample (Figure S1). The total number of unlinked SNPs retained in the final data set obtained with PYRAD for all individuals and *minCov* = 25% was 64,365 SNPs (63,860 SNPs when the outgroup *O. femoralis* was included).

3.3 | Population genetic structure

STRUCTURE analyses yielded an "optimal" clustering value for $K = 2$ according to the ΔK criterion, but log probabilities of the data (LnPr(X|K)) steadily increased up to $K = 6$ (Figure S4a). The inferred genetic groups were consistent with the geographic distribution of the populations but not with their taxonomic assignment (Figure 2 and Figure S5). STRUCTURE analyses showed a hierarchical genetic structure, with a first split between eastern and western populations for $K = 2$ and subsequent subdivision of these two main population groups into other genetic clusters at smaller geographical scales from $K = 3$ to $K = 6$. Most individuals and populations were assigned to a single genetic cluster, with admixture limited to populations in putative contact zones between genetic clusters.

Maximum-likelihood scores from FASTSTRUCTURE analyses peaked for $K = 3$ and, according to the ΔK criterion the "optimal" clustering solution was $K = 2$ (Figure S4b). Model complexity that maximizes marginal likelihood was equal to three in 22 replicates, equal to four in two replicates and equal to five in one replicate, and the number of model components used to explain structure in the data was equal to four in 18 replicates and equal to seven in five replicates. FASTSTRUCTURE results mostly mirrored those obtained with STRUCTURE, but they also presented some differences (Figure S6): (a) For $K = 3$, STRUCTURE included in the same genetic cluster the populations BOR-MON and BOI-LLE. However, these two pairs of populations were assigned to different genetic clusters by FASTSTRUCTURE, which makes more geographical sense; (b) FASTSTRUCTURE analyses for $K > 5$ did not split populations CRE-ERR-QUE from TUR-SET-BAT as done by STRUCTURE for $K = 6$; (c) With the exception of TOS, no other population showed any signature of genetic admixture in FASTSTRUCTURE analyses (Figure S6). Assignment values to additional genetic clusters in FASTSTRUCTURE for analyses with $K > 5$ were extremely low in all cases ($q < 0.0001$) and, thus, their respective bar plots were virtually identical to those obtained for $K = 5$ (for a similar result, see Baiz, Tucker, & Cortés-Ortiz, 2019).

The spatial model of CONSTRUCT showed better model fit at every value of *K* than the non-spatial model (Figure S7a). Spatial models had strong statistical support up to *K* = 3–5 (Figure S7a), but layers beyond *K* = 3 contributed relatively little to total covariance (Figure S7b). CONSTRUCT analyses showed strong genetic admixture for all *K* values (e.g., Whelan et al., 2019) and only genetic assignments for *K* = 3–4 were partially compatible with the presence of the two main genetic clusters revealed by STRUCTURE and FASTSTRUCTURE (Figure 2 and Figure S8).

In agreement with previous analyses, PCA grouped populations according to their geographic location and main genetic clusters rather than by described species (Figure 3). PC1 separated the westernmost populations BOR-MON from the rest of the populations, whereas PC2 separated the rest of western populations from eastern populations and placed the admixed population of TOS at an intermediate position.

3.4 | Phylogenetic inference

Population trees reconstructed in BPP, SNAPP, and SVDQUARTETS yielded similar topologies that only differed in the inferred relationships among some nearby populations belonging to the same genetic clusters (Figures 1a and 4). For the west genetic group the three methods inferred different relationships for the populations NER-SAR-CHI-ESP-PER (Figure 4), which were grouped within the same genetic cluster at the lower hierarchical level by STRUCTURE and FASTSTRUCTURE analyses (Figure 2) and clumped together in PCA (Figure 3). For the east genetic group, SNAPP and BPP analyses produced the same topology and placed TUR as a sister clade of CRE-ERR-QUE, whereas SVDQUARTETS placed TUR as external group with respect to the five easternmost populations CRE-ERR-QUE-SET-BAT. As expected, the relationship among these populations also showed considerable uncertainty in terms of low node support (i.e. posterior probabilities for SNAPP and bootstrapping values for SVDQUARTETS; Figure 4). This is also evidenced by the SNAPP analyses, which showed considerable fuzziness in some parts of the tree probably due to gene flow among nearby populations belonging to the same genetic clusters (Figure 1a). In general, the relationships among populations inferred by all phylogenetic analyses are in good agreement with the hierarchical genetic structure yielded by STRUCTURE, FASTSTRUCTURE, and PCA and indicate no support (i.e., polyphyly) for the separation of the named taxa (Figure 1).

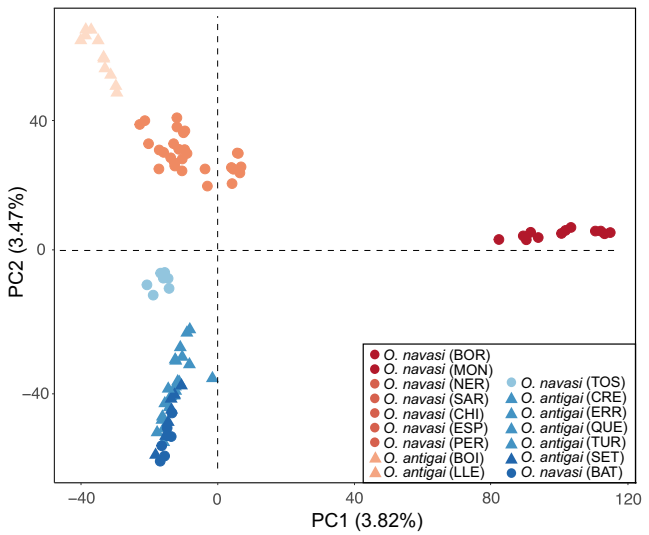

FIGURE 3 Principal component analysis (PCAs) of genetic variation for *Omocestus antigai* (triangles) and *Omocestus navasi* (dots). Colours of triangles and dots indicate the main genetic cluster at which individuals were assigned according to STRUCTURE analyses for *K* = 6. Population codes as in Table 1 [Colour figure can be viewed at wileyonlinelibrary.com]

3.5 | Species delimitation

Species delimitation analyses with the BFD* method strongly supported a six species model (H_5), which considered populations assigned to each genetic cluster identified by Bayesian clustering analyses in STRUCTURE as a distinct species (Table 2). The second most-supported scenario (H_1) was the one considering one species, but received much lower support (Table 2). BPP analyses fully supported the presence of 13 species across all tested topologies (BPP, SNAPP and SVDQUARTETS), prior combinations for gamma distribution, and number of loci (Figure 4). All sampled populations were consistently supported as distinct species when using the subset of 1,000 loci and only some nearby populations belonging to the same genetic clusters (CHI-ESP and CRE-ERR-QUE) were not supported as distinct taxa under a few prior combinations and specific guide trees for the subsets of 200 and 500 loci (Figure 4). IBPP analyses supported all populations as different species using either female or male phenotypic data (Figure 5). IBPP analyses based only on morphological data showed inconsistent results depending on sex and the tested demographic scenario, but tended to support the split into different species for the 3–4 most external nodes (Figure 5). A validation test randomizing phenotypic data across individuals yielded almost an identical result (Figure S9), indicating that the influence of phenotypic traits in IBPP analyses is probably overridden by the high amount of genomic data. Remarkably, IBPP analyses only based on randomized morphological data sets (i.e., without genomic data) also tended to support the split into different species for some external nodes, although the results strongly varied between sexes and across tested topologies and demographic priors (Figure S9).

3.6 | Landscape genetic and phenotypic analyses

MMRR analyses indicated that genetic differentiation (F_{ST}) was significantly correlated with both geographic distances (p = .001) and resistance distances defined by topographic complexity (p = .004; Table 3). However, environmental dissimilarity did not show a significant relationship with genetic differentiation and this variable was excluded from the final model (p = .374; Table 3). MMRR analyses for phenotypic differentiation (P_{ST}) showed that no trait was significantly correlated in any sex with geographical distances, resistance distances defined by topographic complexity,

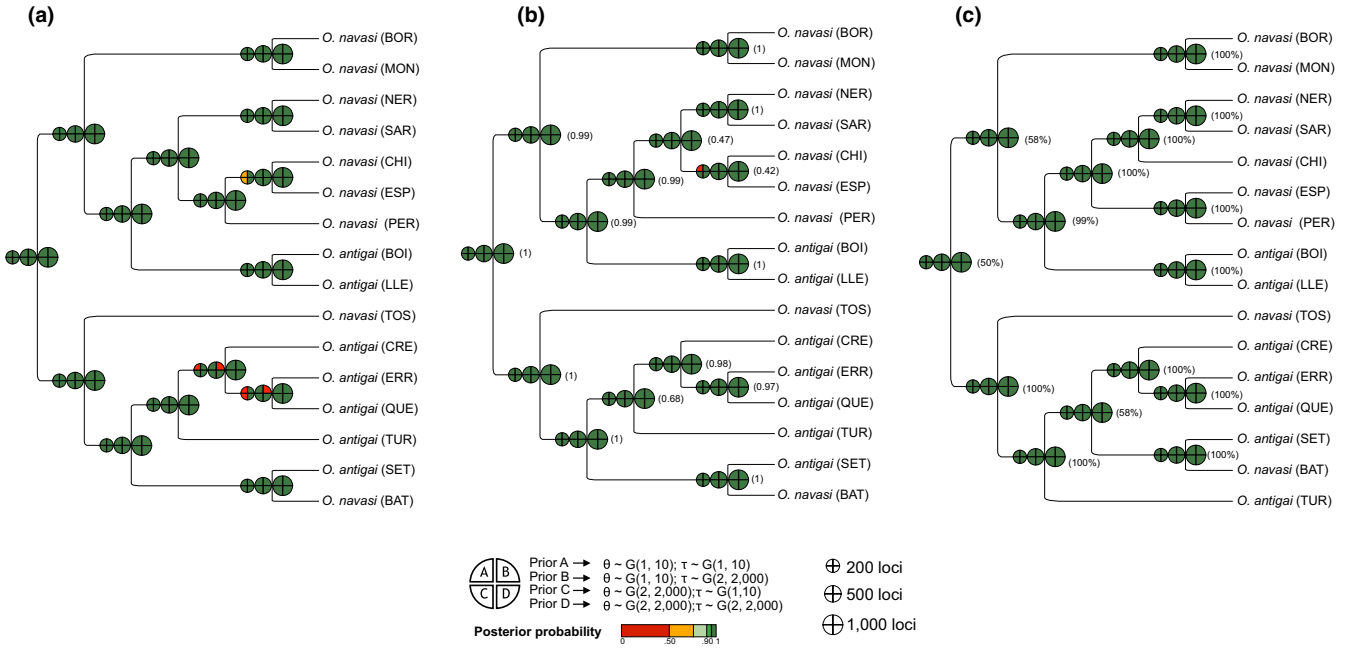

FIGURE 4 Results of species delimitation analyses in BPP. Analyses were performed using three alternative topologies (a: BPP; b: SNAPP; c: SVDQUARTETS), three different subsets of loci (200, 500, and 1,000 loci; two individuals/population), and four gamma prior combinations (gamma, α, β) for ancestral population size (θ) and root age (τ). Colored boxes at each node represent the mean posterior probability (PP) for different combinations of demographic priors (legend at bottom). Node support in terms of posterior probabilities (for SNAPP tree) and bootstrapping values (for SVDQUARTETS tree) is indicated in parentheses for each node. Population codes as in Table 1 [Colour figure can be viewed at wileyonlinelibrary.com]

environmental dissimilarity or genetic differentiation after applying Benjamini-Hochberg false discovery rate corrections for multiple testing (all q-values > 0.05; Table 4).

4 | DISCUSSION

Collectively, our analyses rejected the current hypothesis of two species and indicated that populations assigned to *O. antigai* and *O. navasi* show a distribution of genetic variation that (a) does not match with their respective taxonomic designation and (b) is incompatible with ecological/environmental speciation. Our results show the presence of two main genetic groups corresponding to an east–west split analogous to that found in many other Pyrenean taxa (Wallis, Waters, Upton, & Craw, 2016) and a marked genetic differentiation at local spatial scales reflecting limited population connectivity across the abrupt landscapes characterizing the study region (e.g., Noguerales, Cordero, & Ortego, 2016).

4.1 | Integrative species delimitation

As in many other species complexes (e.g., Leaché et al., 2018) the taxonomy of *O. antigai* and *O. navasi* has changed through time in conjunction with the discussion about the validity of the different criteria considered to establish the boundaries between the two putative taxa (Clemente et al., 1999; Ragge & Reynolds, 1998). After the two species were described by Bolívar in the transition between the XIX and XX centuries (Bolívar, 1897, 1908) and Clemente, García, and Presa (1990) confirmed their taxonomic distinction, Reynolds (1986) and Ragge and Reynolds (1998) attributed their subtle morphological diversity to the effect of population geographic isolation, denied differences in acoustic communication (i.e., courtship songs), and recommended the synonymy of *O. navasi* to *O. antigai*. However, new acoustic and biometrical analyses by Clemente et al. (1999) resurrected the species *O. navasi* and the two taxa are now considered valid species (Cigliano et al., 2019) and treated as such in conservation assessment programs (Braud, Hochkirch, Fartmann, et al., 2016; Braud, Hochkirch, Fartmann, et al., 2016; Hochkirch et al., 2016). Up to now, no molecular-based analysis had been performed to test the monophyly of the two taxa and determine whether their specific habitat associations and slight phenotypic differences have a genetic basis in line with a scenario of ecological speciation. Using genomic and phenotypic data and a comprehensive suite of phylogenomic and Bayesian clustering analyses, our results indicate no support for the current taxonomy. Phylogenomic and genetic structure analyses indicate that populations split in two main genetic groups (east/west) that do not match current taxonomic designations and, accordingly, populations sampled in the two distinct habitats (alpine vs. Mediterranean/montane) supposedly occupied by the two putative species are phylogenetically interspersed (Figure 1).

Despite phylogenomic reconstructions not recovering the monophyly of the two nominated taxa, molecular-based species delimitation analyses using BPP and BFD* identified the presence of several plausible species. Independently of the number of loci used, BPP analyses fully supported the presence of 13 species across all tested topologies

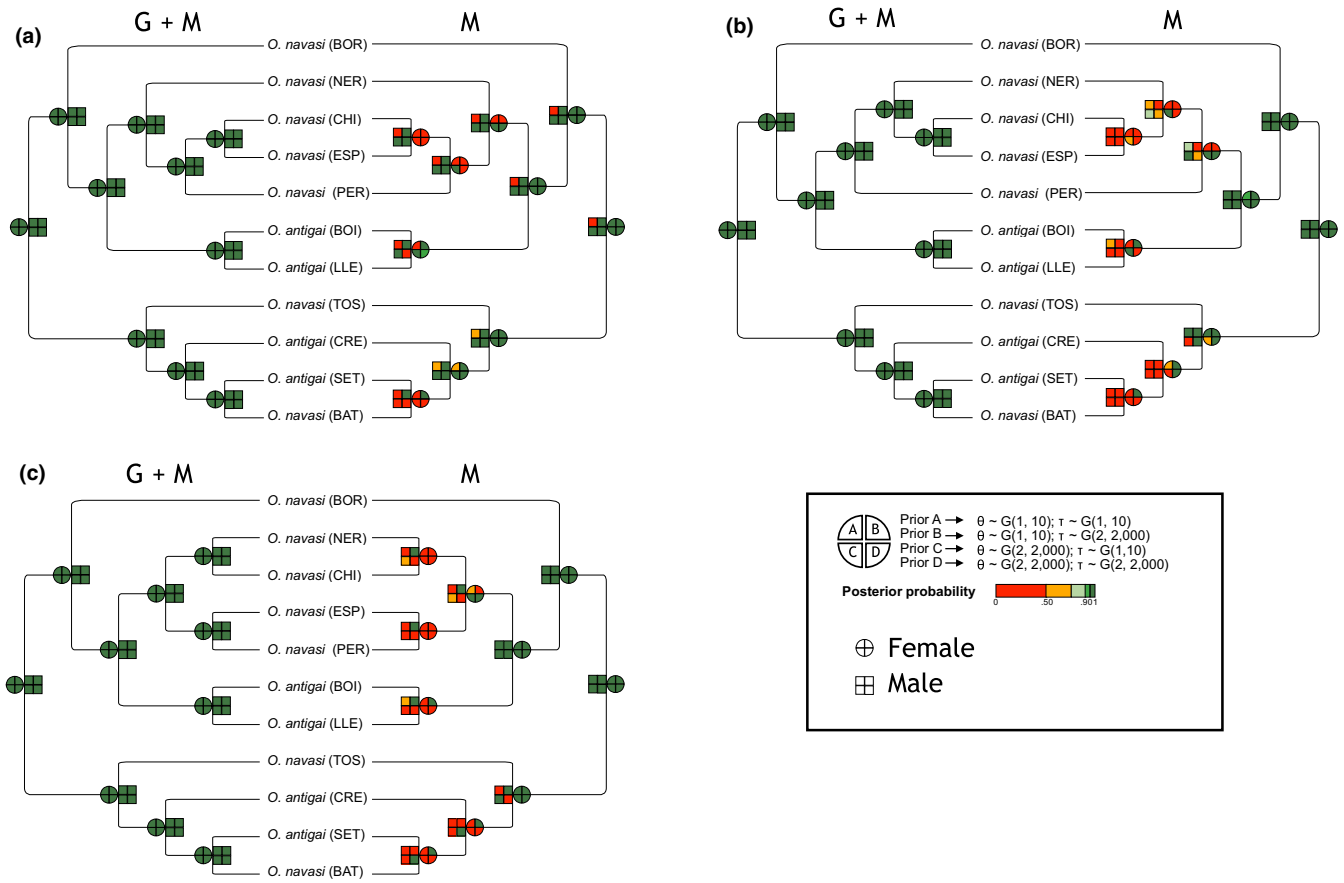

FIGURE 5 Results of species delimitation analyses in IBPP. Analyses were performed separately for each sex both combining genomic (500 loci and five individuals/population) and phenotypic data (trees on the left) and only using phenotypic data (trees on the right). All analyses were performed using three alternative topologies (a: BPP; b: SNAPP; c: SVDQUARTETS) and four gamma prior combinations (gamma, α, β) for ancestral population size (θ) and root age (τ). Coloured boxes at each node represent the mean posterior probability (PP) for different combinations of demographic priors (legend at bottom right). Population codes as in Table 1 [Colour figure can be viewed at wileyonlinelibrary.com]

and prior combinations for gamma distribution. In most cases, each of the 17 sampled populations was supported as a distinct species and only some nearby populations separated by less than 20 km and belonging to the same genetic clusters (CHI-ESP and CRE-ERR-QUE) were not supported as distinct taxa under a few prior combinations and specific topologies. It is also remarkable that many nearby populations (BOR-MON, NER-SAR, BOI-LLE and SET-BAT) assigned to the same genetic groups by PCA and Bayesian clustering analyses (STRUCTURE, FASTSTRUCTURE and CONSTRUCT) were consistently delimited as distinct species (Figure 4). Accordingly, the species delimitation model most supported by BFD* analyses was the one considering the highest number of species, one per genetic cluster inferred by STRUCTURE (see also figure 2b in Leaché et al., 2018). Recent empirical and theoretical studies have shown the limitations of Bayesian species delimitation approaches based on the multi-species coalescent (MSC) model (Rannala & Yang, 2013; Yang & Rannala, 2010), pointing out that such methods are not able to statistically distinguish genetic structure due to population isolation from true species boundaries and tend to over-split genetically differentiated populations rather than capturing species divergence (Huang, 2018; Leaché et al., 2018; Leaché, Zhu, Rannala, & Yang, 2019; Sukumaran & Knowles, 2017). The problem of species overestimation can be particularly exacerbated when using vast genome-wide data due to the high power of a large number of markers (hundreds to tens of thousands) to detect fine-grain population structure (Hime et al., 2016; Noguerales et al., 2018; Sukumaran & Knowles, 2017). In this line, recent empirical studies have found that geographically continuous intraspecific populations presenting genotypic differences simply resulted from isolation by distance can be delimited as distinct species when using hundreds of loci (Huang, 2018; Pyron, Hsieh, Lemmon, Lemmon, & Hendry, 2016).

TABLE 3 Multiple matrix regression with randomization (MMRR) for pairwise population genetic differentiation (F_{ST}) in relation with geographical, resistance and environmental distances

Variable	Coefficient	*t*	*p*-Value
Retained terms			
Intercept	−0.108	−3.079	.991
Geographical distance	0.255	6.841	.001
Resistance distance (slope)	0.007	5.278	.004
Rejected term			
Environmental distance	—	−1.061	.374

TABLE 4 Multiple matrix regressions with randomization (MMRR) for pairwise population phenotypic differentiation (P_{ST}) in relation with genetic differentiation (F_{ST}) and geographical, resistance and environmental distances

	Females		Males	
	t	*q*	*t*	*q*
Head—PC1				
Geographical distance	0.448	0.919	−0.048	0.980
Resistance distance (slope)	0.989	0.791	−1.106	0.853
Environmental distance	−1.699	0.791	−2.993	0.288
F_{ST}	0.125	0.950	−0.052	0.980
Head—PC2				
Geographical distance	1.882	0.686	1.223	0.864
Resistance distance (slope)	−1.047	0.791	−0.429	0.973
Environmental distance	−1.423	0.791	0.135	0.980
F_{ST}	2.781	0.400	1.735	0.840
Pronotum—PC1				
Geographical distance	−1.264	0.791	−0.858	0.973
Resistance distance (slope)	−0.416	0.919	0.376	0.973
Environmental distance	0.963	0.791	1.713	0.853
F_{ST}	−0.960	0.791	0.587	0.973
Pronotum—PC2				
Geographical distance	0.868	0.919	0.715	0.973
Resistance distance (slope)	−0.267	0.919	0.318	0.973
Environmental distance	−1.240	0.791	−0.054	0.980
F_{ST}	−0.735	0.919	1.152	0.853
Forewing—PC1				
Geographical distance	2.775	0.405	0.695	0.973
Resistance distance (slope)	−0.190	0.919	−0.386	0.973
Environmental distance	−2.197	0.686	−2.455	0.288
F_{ST}	2.695	0.520	−0.031	0.980
Forewing—PC2				
Geographical distance	−0.222	0.919	−1.471	0.840
Resistance distance (slope)	0.388	0.919	2.185	0.496
Environmental distance	3.159	0.400	0.782	0.973
F_{ST}	1.110	0.791	2.045	0.720
Ovopositor valve—PC1				
Geographical distance	−0.282	0.919	–	–
Resistance distance (slope)	0.210	0.919	–	–
Environmental distance	1.614	0.686	–	–
F_{ST}	−0.408	0.919	–	–
Ovopositor valve—PC2				
Geographical distance	−0.257	0.919	–	–
Resistance distance (slope)	0.637	0.919	–	–
Environmental distance	1.401	0.791	–	–
F_{ST}	0.617	0.919	–	–

Note: No independent variable was significant and retained into final models. Table shows *q*-values after applying Benjamini-Hochberg false discovery rate corrections of *p*-values to adjust for multiple statistical tests.

Recent studies have suggested that species delimitation approaches integrating genomic data with other sources of information (ecological, morphological, ethological, etc) could help to mitigate the problem of confounding population structure with species limits (Pyron et al., 2016; Sukumaran & Knowles, 2017). Although most phenotypic traits analyzed only presented very subtle differences

among some pairs of populations and did not differ between currently recognized taxa (see Figures S2 and S3), our IBPP analyses supported most populations as distinct species and a validation test randomizing phenotypic data across individuals yielded an almost identical result. Thus, incorporating morphological traits has no or very little influence on the outcome of IBPP analyses and the impact of phenotypic traits seems to be overridden by the high amount of genetic data. Our analyses suggest that the species delimitation approach implemented in IBPP does not really help to reduce the number of delineated species, indicating that integrating phenotypic and genetic data into this framework cannot prevent species overestimation (Noguerales et al., 2018; Sukumaran & Knowles, 2017). In sum, the strong genetic structure in our study system and the high number of loci employed are likely to have extraordinarily inflated the number of species identified by BPP and BFD* and integrative analyses implemented in IBPP does not seem to help to ameliorate this problem.

Beyond the taxonomic inflation issue linked with the fact that the multispecies coalescent model frequently diagnoses genetic structure rather than species, different intersecting lines of evidence suggest that the multiple delineated entities do not meet the requirements to be considered distinct taxa according to alternative contemporary species concepts (see table 1 in de Queiroz, 2007). These lines of evidence include: (a) The presence of largerly overlapping phenotypes across populations, taxa, and genetic clusters for all studied traits (Figure S2 and S3; phenetic species concept); (b) No evidence for environmental-driven divergence and the presence of phylogenetically interspersed high and low elevation populations (ecological + phylogenetic species concepts) and (c) Genetic clusters are always allopatric and admixture at contact zones happens even between the two most diverging eastern and western genetic groups (genotypic cluster species concept), which is indicative of lack of intrinsic reproductive isolation (biological species concept; de Queiroz, 2007). Although the strong signal for isolation-by-distance might be potentially compatible with a rapid stepping-stone speciation process (e.g., VanderWerf, Young, Yeung, & Carlon, 2010), such pattern most plausibly reflects population connectivity (Hutchison & Templeton, 1999; Slatkin, 1993). Thus, in our opinion, the studied populations should be treated—at least by now—as a single species taxon with strong genetic structure, which is not incompatible with the possibility of designating the main genetic clusters (e.g., east–west genetic groups) as intraspecific evolutionary significant units (Moritz, 2002) that should be considered in future conservation and management strategies (Braud, Hochkirch, Fartmann, et al., 2016; Braud, Hochkirch, Fartmann, et al., 2016).

4.2 | Factors structuring genomic and phenotypic variation

Phylogenomic reconstructions and spatial patterns of genetic structure inferred from classic Bayesian clustering analyses supported that genetic variation within the complex is hierarchically structured in congruence with the geographical distribution of the sampled populations. However, the strong signal of isolation-by-distance (r = .71, p < .0001; Table 3) probably explains why the spatial analyses in CONSTRUCT blur most genetic structure and do not even clearly recover the west/east split revealed by phylogenetic reconstructions and all other analyses. Our study system presents a high degree of genetic differentiation (mean F_{ST} = 0.35) and extraordinarily high estimates of genetic drift after divergence for all genetic clusters inferred by STRUCTURE analyses (F-value > 0.5; Pritchard et al., 2000). Also, there is a remarkable congruence between the hierarchical genetic structure inferred by nonspatial clustering analyses (STRUCTURE, FASTSTRUCTURE) and both phylogenetic reconstructions (SNAPP, SVDQUARTETS, and BPP) and PCA, the latter being a method free of the assumptions made by classic clustering analyses (Jombart, Devillard, & Balloux, 2010). All these lines of evidence suggest that results from CONSTRUCT make little biological sense in our specific study system and indicate that the strong genetic structure revealed by all other analyses is genuine and resulted from historical processes (i.e., strong isolation) rather than being an artefactual consequence of gradual genetic differentiation (i.e., migration and genetic drift equilibrium) merely driven by geographical distance (Bradburd et al., 2018).

In agreement with STRUCTURE and FASTSTRUCTURE results for K = 2, all phylogenetic analyses showed a basal split in two clades corresponding to populations located east and west of an imaginary line located around the Segre river (~1°E). Similar east-west genetic splits have been reported for many other plant and animal species from the Pyrenean region (e.g., Alvarez-Presas, Mateos, Vila-Farre, Sluys, & Riutort, 2012; Bidegaray-Batista et al., 2016; Charrier, Dupont, Pornon, & Escaravage, 2014; Milá, Carranza, Guillaume, & Clobert, 2010; Valbuena-Ureña et al., 2018). This east-west split might be well explained by detailed geological reconstructions documenting heavily glaciated areas in the central part of the range and the likely presence of ice-free refugia at its longitudinal extremes (Ehlers, Gibbard, & Hugues, 2011) that might have allowed montane/alpine species to survive glaciations and recolonize high elevations during interglacial periods (Charrier et al., 2014; Valbuena-Ureña et al., 2018). Overall, these results are consistent with the transverse breaks proposed by Wallis et al. (2016) for numerous alpine taxa and support the idea that the fragmentation of ancestral distributions during glacial periods and the presence of Pleistocene refugia along mountain ranges has played a key role on the diversification not only of lowland species but also of montane and alpine organisms.

Genomic data also revealed strong genetic differentiation and a deep genetic structure at fine spatial scales, with up to three genetic clusters comprised within each of the two main west and east genetic groups (Figure 2). The fact that almost all sampled populations have been assigned to a unique genetic cluster, with only a few cases of genetic admixture between nearby populations, suggest a strong effect of geographic isolation (Wang, 2013). Only populations TOS and PER (and into a lesser extent SAR), located in the contact zone between the west and east genetic groups presented signatures of genetic admixture, evidencing ongoing gene flow between them (Figures 1 and 2). It is noteworthy that other populations located at a

similar longitude but at higher elevations showed no evidence of genetic admixture (e.g., BOI, LLE), which suggests a higher isolation of alpine populations and increased gene flow through the less abrupt landscapes characterizing the Pyrenean foothills. Accordingly, landscape genetic analyses indicated that genetic differentiation was explained by both the geographical distance among populations and resistance distances defined by topographic roughness (Table 3). These results might reflect the low dispersal capability of the studied taxon, males being brachypterous and females micropterous, and are comparable to those obtained by previous studies showing the impact of steep slopes and complex landscapes on structuring genetic variation in montane/alpine grasshoppers (Noguerales, Cordero, et al., 2016). Genetic structure analyses showed that the split of the different populations at local/regional scales followed a longitudinal cline rather than a segregation of alpine and Mediterranean-montane populations (e.g., SET and BAT), indicating no support for either ecologically driven divergence or the taxonomic separation between the supposedly Mediterranean *O. navasi* and the alpine *O. antigai*. This was corroborated by our landscape genetic analyses, which revealed no effect of environmental dissimilarity on structuring genetic variation in the complex (i.e., isolation-by-environment; Shafer & Wolf, 2013).

Our geometric morphometric analyses showed that the subtle phenotypic differences found among populations were not explained by genetic differentiation, geographical distances or environmental dissimilarity. Overall, this suggests that the weak phenotypic differences found among populations are not a consequence of genetic drift or environmental-driven selection (Keller, Alexander, Holderegger, & Edwards, 2013; Leinonen, Cano, Makinen, & Merila, 2006; Leinonen, O'Hara, Cano, & Merila, 2008) and might be explained by ecological and evolutionary aspects not considered in this study such as sexual selection, predation risk, microhabitat structure or adaptations to different feeding resources (e.g., Ingley, Billman, Belk, & Johnson, 2014; Laiolo, Illera, & Obeso, 2013; Noguerales, García-Navas, Cordero, & Ortego, 2016).

5 | CONCLUSIONS

This study exemplifies the problems associated with species validation tests involving recently diverged allopatric taxa and highlights the importance of integrating different sources of information to delimit species that have been described solely by morphology or ecological distinctiveness (e.g., Jones & Weisrock, 2018). Phylogenetic inferences, Bayesian species delimitation analyses, and phenotypic and ecological data did not support the current taxonomic status of the complex and indicate that *O. navasi* and *O. antigai* must be synonymized into a unique taxon: *O. antigai* (Bolívar, 1897; for a list of synonyms, see Table S1). Our study also illustrates the implications that incorrect taxonomic designations can have for species conservation and management. The recent assessment of the conservation status of European grasshoppers has assigned the two taxa to different IUCN Red List categories (Hochkirch et al., 2016), considering *O. navasi* as a "endangered" species (Braud, Hochkirch, Fartmann, et al., 2016) and *O. antigai* as "vulnerable" (Braud, Hochkirch, Fartmann, et al., 2016). In the view of our results, the conservation status of the complex requires total reconsideration in future IUCN Red List assessments. Given that the range and population sizes of *O. antigai* are larger and the ecological and habitat requirements much wider than previously thought, the conservation status of the taxon should be probably downlisted to "near threatened" (Braud, Hochkirch, Fartmann, et al., 2016; Braud, Hochkirch, Fartmann, et al., 2016). However, the presence of two marked genetic groups and many genetically substructured populations should be also considered in future conservation plans as potential evolutionary significant units that probably deserve to be protected and managed independently (Moritz, 2002). Overall, our analyses point to the presence of a single species characterized by a strong genetic structure, little phenotypic variation, and a wide environmental niche. Future studies performing demographic reconstructions and spatiotemporally explicit landscape genetic analyses that integrate available information on past glacier extent (Ehlers et al., 2011) and inferred distributional shifts of the species linked to Pleistocene glacial cycles could greatly help to further understand the historical processes that have shaped the spatial distribution of genetic variation within and among the current populations (e.g., Massatti & Knowles, 2016).

ACKNOWLEDGEMENTS

We are grateful to Amparo Hidalgo-Galiana for her support during the preparation of genomic libraries, Víctor Noguerales and Pedro J. Cordero for their help during fieldwork, Francisco Rodríguez-Sánchez for providing valuable advice in spatial data analyses and Sergio Pereira (The Centre for Applied Genomics) for Illumina sequencing. We also thank three anonymous referees for helpful and 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 work 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. All authors collected the samples. V.T. performed the laboratory

work and analysed the data guided by A.P., and 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. Morphometric data and input files for all analyses are available for download from the Dryad Digital Repository (https://doi.org/10.5061/dryad.249v096).

ORCID

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

REFERENCES

Adams, D. C., & Otarola-Castillo, E. (2013). *Geomorph*: An R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*, *4*(4), 393–399. https://doi.org/10.1111/2041-210x.12035

Agapow, P. M., Bininda-Emonds, O. R. P., Crandall, K. A., Gittleman, J. L., Mace, G. M., Marshall, J. C., & Purvis, A. (2004). The impact of species concept on biodiversity studies. *Quarterly Review of Biology*, *79*(2), 161–179. https://doi.org/10.1086/383542

Alvarez-Presas, M., Mateos, E., Vila-Farre, M., Sluys, R., & Riutort, M. (2012). Evidence for the persistence of the land planarian species *Microplana terrestris* (Muller, 1774) (Platyhelminthes, Tricladida) in microrefugia during the Last Glacial Maximum in the northern section of the Iberian Peninsula. *Molecular Phylogenetics and Evolution*, *64*(3), 491–499. https://doi.org/10.1016/j.ympev.2012.05.001

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., … Saunders, N. C. (1987). Intraspecific phylogeography – The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, *18*, 489–522. https://doi.org/10.1146/annurev.ecolsys.18.1.489

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

Bidegaray-Batista, L., Sanchez-Gracia, A., Santulli, G., Maiorano, L., Guisan, A., Vogler, A. P., & Arnedo, M. A. (2016). Imprints of multiple glacial refugia in the Pyrenees revealed by phylogeography and palaeodistribution modelling of an endemic spider. *Molecular Ecology*, *25*(9), 2046–2064. https://doi.org/10.1111/mec.13585

Blair, C., Bryson, R. W., Linkem, C. W., Lazcano, D., Klicka, J., & McCormack, J. E. (2019). Cryptic diversity in the Mexican highlands: Thousands of UCE loci help illuminate phylogenetic relationships, species limits and divergence times of montane rattlesnakes (Viperidae: Crotalus). *Molecular Ecology Resources*, *19*(2), 349–365. https://doi.org/10.1111/1755-0998.12970

Bolívar, I. (1897). Catálogo sinóptico de los ortópteros de la fauna ibérica. *Annaes de Sciencias Naturaes*, *4*(4), 203–232.

Bolívar, I. (1908). Algunos ortopteros nuevos de España, Marruecos y Canarias. *Boletin de La Real Sociedad Española de Historia Natural*, *8*, 317–334.

Bookstein, F. L. (1992). *Morphometric tools for landmark data*. Cambridge, UK: Cambridge University Press.

Bouckaert, R. R. (2010). DENSITREE: Making sense of sets of phylogenetic trees. *Bioinformatics*, *26*(10), 1372–1373. https://doi.org/10.1093/bioinformatics/btq110

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., … Drummond, A. J. (2014). BEAST2: A Software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, *10*(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bradburd, G. S., Coop, G. M., & Ralph, P. L. (2018). Inferring continuous and discrete population genetic structure across space. *Genetics*, *210*(1), 33–52. https://doi.org/10.1534/genetics.118.301333

Braud, Y., Hochkirch, A., Fartmann, T., Presa, J. J., Monnerat, C., Roesti, C., & Dusoulier, F. (2016). *Omocestus antigai: The IUCN Red List of Threatened Species 2016 (Publication no. e.T16084614A70647955. doi: 10.2305/IUCN.UK.2016-3.RLTS.T16084614A70647955.en)*. Retrieved from https://www.iucnredlist.org/species/16084614/70647955

Braud, Y., Hochkirch, A., Presa, J. J., Roesti, C., Rutschmann, F., Monnerat, C., & Dusoulier, F. (2016). *Omocestus navasi: The IUCN Red List of Threatened Species 2016 (Publication no. e.T69641425A69641429. doi: 10.2305/IUCN.UK.2016-3.RLTS.T69641425A69641429.en)*. Retrieved from https://www.iucnredlist.org/species/69641425/69641429

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, *29*(8), 1917–1932. https://doi.org/10.1093/molbev/mss086

Burbrink, F. T., Yao, H., Ingrasci, M., Bryson, R. W., Guiher, T. J., & Ruane, S. (2011). Speciation at the Mogollon Rim in the Arizona Mountain Kingsnake (*Lampropeltis pyromelana*). *Molecular Phylogenetics and Evolution*, *60*(3), 445–454. https://doi.org/10.1016/j.ympev.2011.05.009

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology*, *22*(17), 4369–4383. https://doi.org/10.1111/mec.12413

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). STACKS: Building and genotyping loci *de novo* from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*(3), 171–182. https://doi.org/10.1534/g3.111.000240

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

Chan, K. O., Alexander, A. M., Grismer, L. L., Su, Y. C., Grismer, J. L., Quah, E. S. H., & Brown, R. M. (2017). Species delimitation with gene flow: A methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian Torrent Frogs. *Molecular Ecology*, *26*(20), 5435–5450. https://doi.org/10.1111/mec.14296

Charrier, O., Dupont, P., Pornon, A., & Escaravage, N. (2014). Microsatellite marker analysis reveals the complex phylogeographic history of *Rhododendron ferrugineum* (Ericaceae) in the Pyrenees. *PLoS ONE*, *9*(3), e92976. https://doi.org/10.1371/journal.pone.0092976

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics*, *30*(23), 3317–3324. https://doi.org/10.1093/bioinformatics/btu530

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., García, M. D., Arnaldos, M. I., Romera, E., & Presa, J. J. (1999). Corroboration of the specific status of *Omocestus antigai* (Bolívar, 1897) and *O. navasi* Bolívar, 1908 (Orthoptera, Acrididae). [Confirmación de las posiciones taxonómicas específicas de Omocestus antigai (Bolívar, 1897) y O. navasi Bolívar, 1908 (Orthoptera, Acrididae).]. *Boletin De La Real Sociedad Española De Historia Natural Sección Biológica*, *95*(3–4), 27–50.

Clemente, M. E., García, M. D., & Presa, J. J. (1990). Nuevos datos sobre los Acridoidea (Insecta: Orthoptera) del Pirineo y prepirineo Catalano-Aragonés. *Butlleti de la Institucio Catalana d'historia Natural*, *58*, 37–44.

da Silva, S. B., & da Silva, A. (2018). *Pstat*: An R Package to assess population differentiation in phenotypic traits. *The R Journal*, *10*(1), 447–454. https://doi.org/10.32614/RJ-2018-010

Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society*, *85*(3), 407–415. https://doi.org/10.1111/j.1095-8312.2005.00503.x

de Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*(6), 879–886. https://doi.org/10.1080/10635150701701083

Dobzhansky, T. (1937). Genetic nature of species differences. *American Naturalist*, *71*, 404–420. https://doi.org/10.1086/280726

Dobzhansky, T. (1970). *Genetics of the evolutionary process*. New York, NY: Columbia University Press.

Domingos, F., Colli, G. R., Lemmon, A., Lemmon, E. M., & Beheregaray, L. B. (2017). In the shadows: Phylogenomics and coalescent species delimitation unveil cryptic diversity in a Cerrado endemic lizard (Squamata: Tropidurus). *Molecular Phylogenetics and Evolution*, *107*, 455–465. https://doi.org/10.1016/j.ympev.2016.12.009

Dowle, E. J., Morgan-Richards, M., & Trewick, S. A. (2014). Morphological differentiation despite gene flow in an endangered grasshopper. *BMC Evolutionary Biology*, *14*, 216. https://doi.org/10.1186/s12862-014-0216-x

Eaton, D. A. R. (2014). PYRAD: Assembly of *de novo* RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*(13), 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Edwards, D. L., & Knowles, L. L. (2014). Species detection and individual assignment in species delimitation: Can integrative data increase efficacy? *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20132765. https://doi.org/10.1098/rspb.2013.2765

Ehlers, J., Gibbard, P. L., & Hugues, P. D. (2011). *Quaternary glaciations – Extent and chronology: A closer look* (1st ed.). Amsterdam, the Netherlands: Elsevier.

Ekblom, R., & Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, *107*(1), 1–15. https://doi.org/10.1038/hdy.2010.152

Excoffier, L., Laval, G., & Schneider, S. (2005). ARLEQUIN ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, *1*, 47–50.

Friedline, C. J., Faske, T. M., Lind, B. M., Hobson, E. M., Parry, D., Dyer, R. J., … Eckert, A. J. (2019). Evolutionary genomics of gypsy moth populations sampled along a latitudinal gradient. *Molecular Ecology*, *28*(9), 2206–2223. https://doi.org/10.1111/mec.15069

Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, *27*(9), 480–488. https://doi.org/10.1016/j.tree.2012.04.012

García-Navas, V., Noguerales, V., Cordero, P. J., & Ortego, J. (2017). Ecological drivers of body size evolution and sexual size dimorphism in short-horned grasshoppers (Orthoptera: Acrididae). *Journal of Evolutionary Biology*, *30*(8), 1592–1608. https://doi.org/10.1111/jeb.13131

Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, *270*(1512), 313–321. https://doi.org/10.1098/rspb.2002.2218

Hime, P. M., Hotaling, S., Grewelle, R. E., O'Neill, E. M., Voss, S. R., Shaffer, H. B., & Weisrock, D. W. (2016). The influence of locus number and information content on species delimitation: An empirical test case in an endangered Mexican salamander. *Molecular Ecology*, *25*(23), 5959–5974. https://doi.org/10.1111/mec.13883

Hochkirch, A., Nieto, A., García-Criado, M., Cálix, M., Braud, Y., Buzzetti, F. M., … Tumbrinck, J. (2016). *European red list of grasshoppers, crickets and bush-crickets*. Luxembourg: Publications Office of the European Union.

Huang, J. P. (2018). What have been and what can be delimited as species using molecular data under the multi-species coalescent model? A case study using *Hercules* beetles (Dynastes: Dynastidae). *Insect Systematics and Diversity*, *2*(2), 1–10. https://doi.org/10.1093/isd/ixy003

Huang, J. P., & Knowles, L. L. (2016). The species versus subspecies conundrum: Quantitative delimitation from integrating multiple data types within a single Bayesian approach in hercules beetles. *Systematic Biology*, *65*(4), 685–699. https://doi.org/10.1093/sysbio/syv119

Hutchison, D. W., & Templeton, A. R. (1999). Correlation of pairwise genetic and geographic distance measures: Inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, *53*(6), 1898–1914. https://doi.org/10.1111/j.1558-5646.1999.tb04571.x

Ingley, S. J., Billman, E. J., Belk, M. C., & Johnson, J. B. (2014). Morphological divergence driven by predation environment within and between species of *Brachyrhaphis* fishes. *PLoS ONE*, *9*(2), e90274. https://doi.org/10.1371/journal.pone.0090274

Isaac, N. J. B., Mallet, J., & Mace, G. M. (2004). Taxonomic inflation: Its influence on macroecology and conservation. *Trends in Ecology & Evolution*, *19*(9), 464–469. https://doi.org/10.1016/j.tree.2004.06.004

Jackson, N. D., Carstens, B. C., Morales, A. E., & O'Meara, B. C. (2017). Species delimitation with gene flow. *Systematic Biology*, *66*(5), 799–812. https://doi.org/10.1093/sysbio/syw117

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

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 94. https://doi.org/10.1186/1471-2156-11-94

Jones, K. S., & Weisrock, D. W. (2018). Genomic data reject the hypothesis of sympatric ecological speciation in a clade of *Desmognathus* salamanders. *Evolution*, *72*(11), 2378–2393. https://doi.org/10.1111/evo.13606

Keller, I., Alexander, J. M., Holderegger, R., & Edwards, P. J. (2013). Widespread phenotypic and genetic divergence along altitudinal gradients in animals. *Journal of Evolutionary Biology*, *26*(12), 2527–2543. https://doi.org/10.1111/jeb.12255

Knowles, L. L., & Carstens, B. C. (2007). Delimiting species without monophyletic gene trees. *Systematic Biology*, *56*(6), 887–895. https://doi.org/10.1080/10635150701701091

Laiolo, P., Illera, J. C., & Obeso, J. R. (2013). Local climate determines intra- and interspecific variation in sexual size dimorphism in mountain grasshopper communities. *Journal of Evolutionary Biology*, *26*(10), 2171–2183. https://doi.org/10.1111/jeb.12213

Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., … Andersson, L. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature*, *518*(7539), 371–375. https://doi.org/10.1038/nature14181

Leaché, A. D., & Fujita, M. K. (2010). Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proceedings of the Royal Society B: Biological Sciences*, *277*(1697), 3071–3077. https://doi.org/10.1098/rspb.2010.0662

Leaché, A. D., Fujita, M. K., Minin, V. N., & Bouckaert, R. R. (2014). Species delimitation using genome-wide SNP data. *Systematic Biology*, *63*(4), 534–542. https://doi.org/10.1093/sysbio/syu018

Leaché, A. D., McElroy, M. T., & Trinh, A. (2018). A genomic evaluation of taxonomic trends through time in coast horned lizards (genus *Phrynosoma*). *Molecular Ecology*, *27*(13), 2884–2895. https://doi.org/10.1111/mec.14715

Leaché, A. D., Zhu, T. Q., Rannala, B., & Yang, Z. H. (2019). The spectre of too many species. *Systematic Biology*, *68*(1), 168–181. https://doi.org/10.1093/sysbio/syy051

Leavitt, S. D., Johnson, L. A., Goward, T., & St Clair, L. L. (2011). Species delimitation in taxonomically difficult lichen-forming fungi: An

example from morphologically and chemically diverse *Xanthoparmelia* (Parmeliaceae) in North America. *Molecular Phylogenetics and Evolution*, *60*(3), 317–332. https://doi.org/10.1016/j.ympev.2011.05.012

Leinonen, T., Cano, J. M., Makinen, H., & Merila, J. (2006). Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *Journal of Evolutionary Biology*, *19*(6), 1803–1812. https://doi.org/10.1111/j.1420-9101.2006.01182.x

Leinonen, T., O'Hara, R. B., Cano, J. M., & Merila, J. (2008). Comparative studies of quantitative trait and neutral marker divergence: A meta-analysis. *Journal of Evolutionary Biology*, *21*(1), 1–17. https://doi.org/10.1111/j.1420-9101.2007.01445.x

Llucia-Pomares, D. (2002). Revision of the Orthoptera (Insecta) of Catalonia (Spain). [Revisión de los ortópteros (Insecta: Orthoptera) de Cataluña (España).]. *Monografias de la Sociedad Entomológica Aragonesa*, *7*, 1–226.

Mace, G. M. (2004). The role of taxonomy in species conservation. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, *359*(1444), 711–719. https://doi.org/10.1098/rstb.2003.1454

Mason, N. A., & Taylor, S. A. (2015). Differentially expressed genes match bill morphology and plumage despite largely undifferentiated genomes in a Holarctic songbird. *Molecular Ecology*, *24*(12), 3009–3025. https://doi.org/10.1111/mec.13140

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

Mayr, E. (1942). *Systematics and the origin of species*. New York, NY: Columbia University Press.

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

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

Milá, B., Carranza, S., Guillaume, O., & Clobert, J. (2010). Marked genetic structuring and extreme dispersal limitation in the Pyrenean brook newt *Calotriton asper* (Amphibia: Salamandridae) revealed by genome-wide AFLP but not mtDNA. *Molecular Ecology*, *19*(1), 108–120. https://doi.org/10.1111/j.1365-294X.2009.04441.x

Moritz, C. (2002). Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, *51*(2), 238–254. https://doi.org/10.1080/10635150252899752

Nelson, G., & Platnick, N. (1981). *Systematics and biogeography*. New York, NY: Columbia University Press.

Niemiller, M. L., Near, T. J., & Fitzpatrick, B. M. (2012). Delimiting species using multilocus data: Diagnosing cryptic diversity in the southern cavefish, *Typhlichthys subterraneus* (Teleostei: Amblyopsidae). *Evolution*, *66*(3), 846–866. https://doi.org/10.1111/j.1558-5646.2011.01480.x

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

Noguerales, V., Cordero, P. J., & Ortego, J. (2018). Integrating genomic and phenotypic data to evaluate alternative phylogenetic and species delimitation hypotheses in a recent evolutionary radiation of grasshoppers. *Molecular Ecology*, *27*(5), 1229–1244. https://doi.org/10.1111/mec.14504

Noguerales, V., García-Navas, V., Cordero, P. J., & Ortego, J. (2016). The role of environment and core-margin effects on range-wide phenotypic variation in a montane grasshopper. *Journal of Evolutionary Biology*, *29*(11), 2129–2142. https://doi.org/10.1111/jeb.12915

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2017). *vegan: Community Ecology Package. R package version 2.4-4*. Retrieved from http://cran.r-project.org/package=vegan

Olmo-Vidal, J. M. (2002). *Atlas Dels Ortòpters de Catalunya i Llibre Vermell 2002/Llagostes*, *Saltamartins*, *Grills*, *Someretes /Atlas de Los Ortópteros de Cataluña y Libro Rojo 2002/Atlas of the Otthoptera of Catalonia and Read Data Book 2002*. Barcelona, Spain: Generalitat de Catalunya, Departament de Medi Ambient i Habitatge.

O'Meara, B. C. (2010). New heuristic methods for joint species delimitation and species tree inference. *Systematic Biology*, *59*(1), 59–73. https://doi.org/10.1093/sysbio/syp077

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

Poniatowski, D., Defaut, B., Llucia-Pomares, D., & Fartmann, T. (2009). The Orthoptera fauna of the Pyrenean region – A field guide. *Articulata Beiheft*, *14*, 1–143.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., … Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595–609. https://doi.org/10.1080/10635150600852011

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

Puissant, S. (2008). The presence of *Omocestus navasi* Bolívar, 1908 in France with description of a new subspecies (Caelifera, Acrididae, Gomphocerinae). [Sur la presence en France d'*Omocestus navasi* Bolívar, 1908, avec description d'une nouvelle sous-espece (Caelifera, Acrididae, Gomphocerinae).]. *Materiaux Orthopteriques et Entomocenotiques*, *13*, 69–73.

Pyron, R. A., Hsieh, F. W., Lemmon, A. R., Lemmon, E. M., & Hendry, C. R. (2016). Integrating phylogenomic and morphological data to assess candidate species-delimitation models in brown and red-bellied snakes (*Storeria*). *Zoological Journal of the Linnean Society*, *177*(4), 937–949. https://doi.org/10.1111/zoj.12392

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ragge, D. R., & Reynolds, W. J. (1998). *The songs of the grasshoppers and crickets of western Europe*. Colchester, UK: Harley Books.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). FASTSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics*, *197*(2), 573–589. https://doi.org/10.1534/genetics.114.164350

Rancilhac, L., Goudarzi, F., Gehara, M., Hemami, M. R., Elmer, K. R., Vences, M., & Steinfarz, S. (2019). Phylogeny and species delimitation of near Eastern *Neurergus* newts (Salamandridae) based on genome-wide RADseq data analysis. *Molecular Phylogenetics and Evolution*, *133*, 189–197. https://doi.org/10.1016/j.ympev.2019.01.003

Rannala, B., & Yang, Z. H. (2013). Improved reversible jump algorithms for Bayesian species delimitation. *Genetics*, *194*(1), 245–253. https://doi.org/10.1534/genetics.112.149039

Reynolds, W. J. (1986). A description of the song of *Omocestus broelemanni* (Orthoptera: Acrididae) with notes on its taxonomic position. *Journal of Natural History*, *20*(1), 111–116. https://doi.org/10.1080/00222938600770101

Rohlf, F. J. (2015). The TPS series of software. *Hystrix-Italian Journal of Mammalogy*, *26*(1), 9–12. https://doi.org/10.4404/hystrix-26.1-11264

Rohlf, F. J., & Slice, D. (1990). Extensions of the procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, *39*(1), 40–59. https://doi.org/10.2307/2992207

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*(3), 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Schluter, D. (2001). Ecology and the origin of species. *Trends in Ecology & Evolution*, *16*(7), 372–380. https://doi.org/10.1016/S0169-5347(01)02198-X

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

Shafer, A. B. A., & Wolf, J. B. W. (2013). Widespread evidence for incipient ecological speciation: A meta-analysis of isolation-by-ecology. *Ecology Letters*, *16*(7), 940–950. https://doi.org/10.1111/ele.12120

Simpson, G. G. (1944). *Tempo and mode in evolution*. New York, NY: Columbia University Press.

Sites, J. W., & Marshall, J. C. (2003). Delimiting species: A Renaissance issue in systematic biology. *Trends in Ecology & Evolution*, *18*(9), 462–470. https://doi.org/10.1016/s0169-5347(03)00184-8

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

Sokal, R. R., & Crovello, T. J. (1970). Biological species concept – A critical evaluation. *American Naturalist*, *104*(936), 127–153. https://doi.org/10.1086/282646

Solis-Lemus, C., Knowles, L. L., & Ane, C. (2015). Bayesian species delimitation combining multiple genes and traits in a unified framework. *Evolution*, *69*(2), 492–507. https://doi.org/10.1111/evo.12582

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

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(7), 1607–1612. https://doi.org/10.1073/pnas.1607921114

Swofford, D. L. (2002). PAUP*. *Phylogenetic analysis using parsimony (*and other methods)*. Version 4. Sunderland, MA: Sinauer Associates.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H., & Vogler, A. P. (2003). A plea for DNA taxonomy. *Trends in Ecology & Evolution*, *18*(2), 70–74. https://doi.org/10.1016/s0169-5347(02)00041-1

Valbuena-Ureña, E., Oromi, N., Soler-Membrives, A., Carranza, S., Amat, F., Camarasa, S., … Steinfartz, S. (2018). Jailed in the mountains: Genetic diversity and structure of an endemic newt species across the Pyrenees. *PLoS ONE*, *13*(8), e0200214. https://doi.org/10.1371/journal.pone.0200214

van Dijk, E. L., Auger, H., Jaszczyszyn, Y., & Thermes, C. (2014). Ten years of next-generation sequencing technology. *Trends in Genetics*, *30*(9), 418–426. https://doi.org/10.1016/j.tig.2014.07.001

VanderWerf, E. A., Young, L. C., Yeung, N. W., & Carlon, D. B. (2010). Stepping stone speciation in Hawaii's flycatchers: Molecular divergence supports new island endemics within the elepaio. *Conservation Genetics*, *11*(4), 1283–1298. https://doi.org/10.1007/s10592-009-9958-1

Wallis, G. P., Waters, J. M., Upton, P., & Craw, D. (2016). Transverse alpine speciation driven by glaciation. *Trends in Ecology & Evolution*, *31*(12), 916–926. https://doi.org/10.1016/j.tree.2016.08.009

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

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

Whelan, N. V., Galaska, M. P., Sipley, B. N., Weber, J. M., Johnson, P. D., Halanych, K. M., & Helms, B. S. (2019). Riverscape genetic variation, migration patterns, and morphological variation of the threatened round rocksnail, *Leptoxis ampla*. *Molecular Ecology*, *28*(7), 1593–1610. https://doi.org/10.1111/mec.15032

Wiens, J. J. (2007). Species delimitation: New approaches for discovering diversity. *Systematic Biology*, *56*(6), 875–878. https://doi.org/10.1080/10635150701748506

Wiens, J. J., & Servedio, M. R. (2000). Species delimitation in systematics: Inferring diagnostic differences between species. *Proceedings of the Royal Society B: Biological Sciences*, *267*(1444), 631–636. https://doi.org/10.1098/rspb.2000.1049

Yang, Z. H. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, *61*(5), 854–865. https://doi.org/10.1093/czoolo/61.5.854

Yang, Z. H., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(20), 9264–9269. https://doi.org/10.1073/pnas.0913022107

Yang, Z., & Rannala, B. (2014). Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution*, *31*(12), 3125–3135.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Tonzo V, Papadopoulou A, Ortego J. Genomic data reveal deep genetic structure but no support for current taxonomic designation in a grasshopper species complex. *Mol Ecol*. 2019;28:3869–3886. https://doi.org/10.1111/mec.15189