



# Landscape genetics of the protected Spanish Moon Moth in core, buffer, and peripheral areas of the Ordesa y Monte Perdido National Park (Central Pyrenees, Spain)

Inés González-Castellano<sup>1,2</sup> · Neus Marí-Mena<sup>1,3</sup> · Gernot Segelbacher<sup>4</sup> · Carlos Lopez-Vaamonde<sup>5,6</sup> · Jorge González-López<sup>1,3</sup> · Jaime Fagúndez<sup>1,7</sup> · Marta Vila<sup>1</sup>

Received: 10 October 2022 / Accepted: 15 May 2023 / Published online: 31 May 2023

© The Author(s) 2023

## Abstract

Conservation managers need to know the degree of connectivity showed by the populations to be preserved, especially when protected areas and/or species are involved. One of the conservation projects carried out by the Ordesa y Monte Perdido National Park (Central Pyrenees, Spain) is the monitoring of the protected Spanish Moon Moth, *Graellsia isabellae* (Lepidoptera, Saturniidae), in several sites within the actual park, buffer zone, and peripheral area. Here we studied the genetic diversity, geographical structure, and connectivity of this iconic insect in those areas with the aim of producing evidence-based recommendations that might help the National Park staff in their decision-making. For this, we non-lethally sampled 402 adult moths from 17 sites and worked at two geographic scales: Western/Central Pyrenees and the area monitored by the staff of the National Park. The multilocus genotypes obtained for nine nuclear microsatellite markers allowed us to quantify genetic variation, investigate population structure, and calculate recent migration rates. Our results revealed a large-scale (ca. 125 km) west–east cline in allele frequencies that causes low overall genetic differentiation ( $F_{ST} = 0.038$ ) and similar levels of diversity among sites. Habitat connectivity revealed as an important element determining dispersal for *G. isabellae*, given the patchy distribution of the host plant (*Pinus sylvestris*) in the study area. Gene flow within and outside the National Park was proved, with a particular site of the buffer zone (Bujaruelo) acting as a source of migrants to other localities within and outside the National Park. This finding underlines the importance of considering buffer zones to preserve genetic diversity within protected areas, and that safeguarding the connectedness of pine patches is key to the conservation of this iconic moth.

**Keywords** *Pinus sylvestris* · Isolation by distance · Connectivity · Least cost path analysis · Socioeconomic influence zone · Genetic patch size

---

Inés González-Castellano, Neus Marí-Mena have contributed equally to this work.

---

✉ Marta Vila  
mvila@udc.es

<sup>1</sup> Departamento de Biología, Facultad de Ciencias, Universidade da Coruña, 15071 A Coruña, Spain

<sup>2</sup> Present Address: Centro de Investigación Mariña, Facultad de Biología, Universidade de Vigo, 36310 Vigo, Spain

<sup>3</sup> Present Address: AllGenetics & Biology SL, Rúa Cubelos 21 Oleiros, 15172 A Coruña, Spain

<sup>4</sup> Wildlife Ecology and Management, Albert-Ludwigs University, Tennenbacher Straße 4, 79106 Freiburg, Germany

<sup>5</sup> INRAE, UR633, Zoologie Forestière, F-45075 Orléans, France

<sup>6</sup> Institut de Recherche sur la Biologie de l’Insecte, UMR7261 CNRS-Université de Tours, Tours, France

<sup>7</sup> Centro de Investigacións Científicas Avanzadas (CICA), Biocost Research Group, Universidade da Coruña, 15071 A Coruña, Spain

## Introduction

Legal protection of territories and species is the mainstay of nature conservation. The International Union for Conservation of Nature (IUCN) established six protected area categories, from strict protection (I) to sustainable use of natural resources (VI) (Dudley 2008). In many countries, the areas with the highest levels of protection are frequently located at higher altitudes, as anthropogenic activities are concentrated in low-lying areas with more productive soils (Scott et al. 2001). For instance, 75% of the Spanish National Parks lie in mountain ranges. Some protected areas are therefore located in relatively extreme parts of the landscape in terms of climate, soil, altitude and/or water, which may lead to a lower habitat quality. These more extreme biophysical conditions may turn a protected area into a population sink for certain species, whose protected population/s actually depend on dispersal from source populations occurring in surrounding unprotected areas with higher quality habitats (Hansen 2011). Consequently, populations in protected areas may be vulnerable if land use changes or any other factor makes the external source to become a sink. Furthermore, protected areas are sometimes too small to maintain a full complement of species, excluding a part of the habitat that is necessary to maintain essential ecological processes, making the use and connectivity between core and buffer zones a major issue (Ghosh-Harihar et al. 2019). For all these reasons, knowledge of the connectivity of subpopulations within and outside protected areas is most desirable, especially in the light of the expanding and intensifying human use on the lands around them. The creation of the surrounding buffer zones, where land use is regulated to protect the core area, tries to minimise those putative negative edge effects (Götmark et al. 2000; Hansen and DeFries 2007). Given the difficulties that may arise when enlarging a protected area (e.g., reluctance of individual owners caused by fear to prohibition or control of profit-making activities in their property, see Cabalar Fuentes et al. 2011; Grodzinska-Jurczak and Cent 2011), any effort to alleviate boundary influences on a protected area should be scientifically based.

Landscape genetics facilitates the understanding of how environmental heterogeneity and intrinsic life-history traits (e.g., dispersal, lifespan) affect population connectivity and hence the geographic distribution of genetic diversity (Manel et al. 2003; Holderegger and Wagner 2008). Studying the genetic spatial patterns at fine-scale is particularly useful in the management of protected areas as they can identify the dispersal behaviour of a species, as well as reveal further temporal and spatial aspects preventing gene flow. Such baseline information can be used by conservation practitioners to better understand how to enhance connectivity among subpopulations and to prioritise particular

demes with respect to their potential to act as genetic sources (Howes et al. 2009; Bolliger et al. 2014; Keller et al. 2015).

The first Spanish National Park, Ordesa y Monte Perdido, was created in 1918. At present it is part of the Ordesa-Viñamala Biosphere Reserve (Central Pyrenees). As many as 172 butterfly species may occur in this Pyrenean National Park, i.e., 75% of the butterfly species of continental Spain (Murria-Beltrán 2022). Accordingly, the Park staff is working in several projects that concern lepidopterans, such as the Butterfly Monitoring Scheme and the monitoring of the nocturnal Spanish Moon Moth: *Graellsia isabellae* (Graells 1849), a protected species included in the Habitats Directive of the European Union (Annexes II and V).

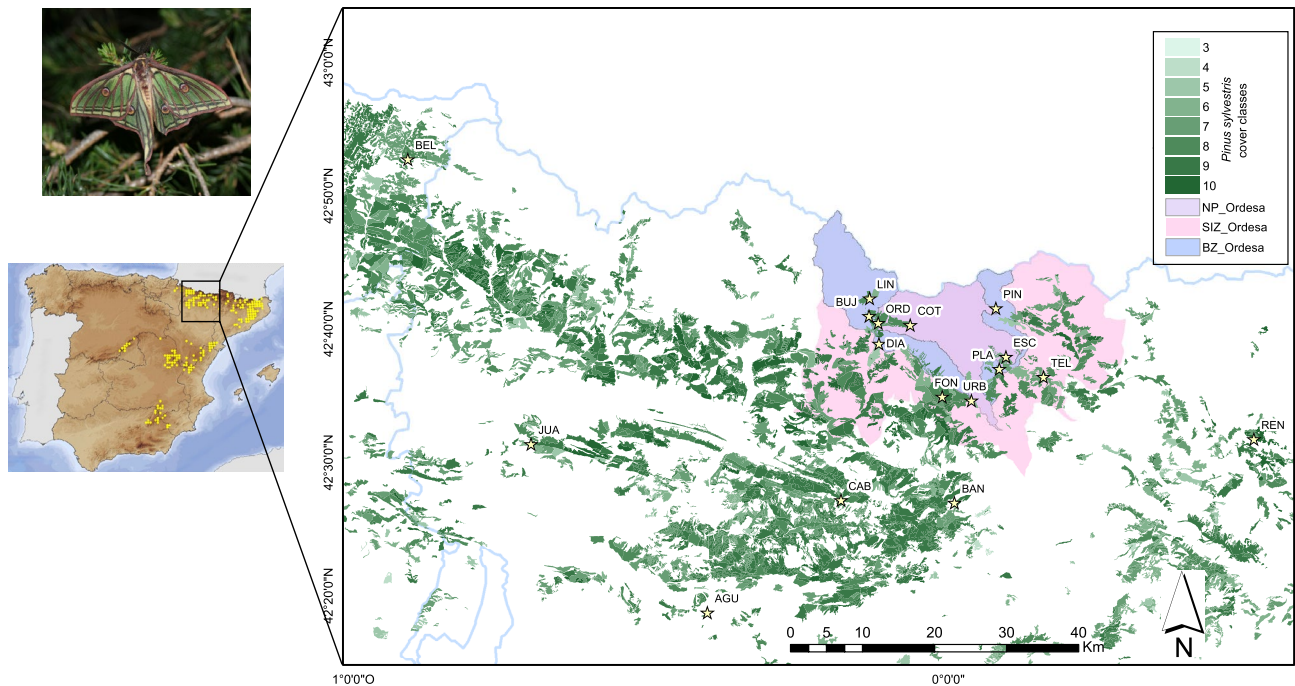
The Ordesa y Monte Perdido National Park (OMPNP) manifested its interest in knowing whether the genetic differentiation ( $F_{ST} = 0.087$ ) found by Marí-Mena et al. (2016) between two localities (west and east of the Monte Perdido massif) when surveying the phylogeographic pattern of the species was due to a real barrier to gene flow or just an artefact caused by insufficient sampling (Meirmans 2015) of the pine forest between those two sites (E. Villagrasa, chief executive officer of the OMPNP, pers. comm.). To address the degree of gene flow showed by *G. isabellae*, one must bear in mind that (i) several lines of evidence point to male-biased dispersal in this lepidopteran (Ylla i Ullastre 1997), (ii) males of *G. isabellae* can fly distances up to two km per night (Marí-Mena et al. 2019), and (iii) the only host plant used by *G. isabellae* in the Pyrenees is the Scots pine (*Pinus sylvestris*).

In particular, we aimed to: (1) check if genetic diversity is uniformly distributed all through the study area or peaks in protected areas, (2) describe the population structure of *G. isabellae* in the Ordesa y Monte Perdido National Park and surroundings, and (3) identify any potential source of migrants. Following Twyford et al. (2020) and to avoid overlooking critical aspects of population structure, we worked at two spatial scales (Western/Central Pyrenees—OMPNP and surroundings). Also, we used spatial analyses to prevent the putative overestimation of genetic differentiation caused by isolation by distance.

## Material and methods

### Study area

The Ordesa y Monte Perdido National Park (42°40'18''N 0°3'20''E, Spain) has a total surface area of 15,696.20 ha and its orography is dominated by the calcareous massif of Monte Perdido (maximum elevation 3355 m), located in the middle of the park and from which the valleys of Ordesa, Añisclo, Escuaín, and Pineta branch off. This national park



**Fig. 1** Distribution area of *Graellsia isabellae* in Spain (left, modified from Romo et al. 2012) and geographical location of samples along the study area (right). For site coordinates and details, see Table 1. Patches of *Pinus sylvestris* forest showing cover classes in 10% inter-

vals (value 10 represents 91–100% cover). Ordesa y Monte Perdido National Park (NP\_Ordesa) are shown with the buffer zone (BZ\_Ordesa) and the peripheral area (SIZ\_Ordesa). Inset: Adult male of *G. isabellae* after non-lethal sampling

has a protected buffer zone (“Zona periférica de protección”) consisting of 19,196.36 ha around it, which in turn is surrounded by a non-protected socioeconomic influence zone (89,290.44 ha) (Fig. 1) that benefits from the park’s recreational activities (600,000 tourists per year) (Ministerio para la Transición Ecológica y el Reto Demográfico 2022a).

As some results in landscape genetics may be scale dependant (Balkenhol et al. 2020), we decided to work both at a small (National Park and surrounding area, 30×20 km approx.) and large scales. The latter comprises the distribution area of the Western Pyrenean genetic group of *G. isabellae* (WP, sensu Marí-Mena et al. 2016, 125 km from west to east), where the National Park is included.

**Study species**

*Graellsia isabellae* is a univoltine lepidopteran that flies from dusk to midnight from mid-March to early July in montane forests of several mountain massifs of Spain (Central and Iberian Systems, Betic Mountains, and Pyrenees), as well as in the French Alps (Romo et al. 2012). The legal status of *G. isabellae* in Spain is Wild Species under Special Protection (“Especie Silvestre en Régimen de Protección Especial”). The yearly monitoring performed by the OMPNP staff confirmed the presence of the species in several sites within the park, the buffer, and peripheral areas.

**Sampling**

Some of the genetic data used in the present work were generated by Marí-Mena et al. (2016, 2019), but we now add data from eight new sites. Adult males were attracted by sexual synthetic pheromone (Millar et al. 2010) and caught with hand-held aerial nets during the flying season 2010. Overall, the present work uses samples from 17 sites (*Pinus sylvestris* woodland) (Fig. 1 and Table 1). The five core sites are located within the boundaries of the Ordesa y Monte Perdido National Park (OMPNP). The remaining samples were collected in sites outside the OMPNP, four of them within its buffer zone (protected), two sites within its socioeconomic influence zone (not protected and referred as “peripheral” hereafter), and six sites far off the OMPNP. We will use the following codes to refer to sampling sites: a one-letter prefix indicating its proximity to the OMPNP ([C]ore, [B]uffer, [P]eripheral, [F]ar) followed by a three-letter abbreviation of the name. Ordesa (C-ORD), Cotatuero (C-COT), San Úrbez (C-URB), Escuaín (C-ESC), Valle de Pineta (C-PIN), Línea de Bujaruelo (B-LIN), Bujaruelo (B-BUJ), Diazas (B-DIA), Plana Canal (B-PLA), Fondañóns (P-FON), Tella (P-TEL), Belabarce (F-BEL), San Juan de la Peña (F-JUA), Pico del Águila (F-AGU), Cabas (F-CAB), Barranco de la Bañera (F-BAN), and Renanué (F-REN). A mean of 23.6 moths

**Table 1** Genetic diversity for 17 sampling sites of *Graellsia isabellae* within and outside the Ordesa y Monte Perdido National Park (OMPNP, Spain)

Codes	Sites	Latitude	Longitude	<i>N</i>	<i>N<sub>A</sub></i>	<i>A<sub>R</sub></i>	<i>A<sub>PR</sub></i>	<i>H<sub>O</sub></i>	<i>H<sub>E</sub></i>	HWE ( <i>P</i> -value)	<i>F<sub>IS</sub></i>
F-BEL	Belabarce	42°52'38"N	0°53'13"W	31	3.444	2.61	0.08	0.436	0.502	0.997	0.149
F-JUA	San Juan	42°30'60"N	0°41'26"W	24	4.111	2.82	0.11	0.470	0.540	0.991	0.152
F-AGU	Pico del Águila	42°17'60"N	0°24'08"W	31	4.111	2.78	0.06	0.495	0.520	0.849	0.064
F-CAB	Cabas	42°26'06"N	0°10'17"W	32	4.222	2.77	0.05	0.504	0.519	0.982	0.045
B-LIN	Línea Bujaruelo	42°41'08"N	0°06'39"W	11	3.222	2.78	0.01	0.543	0.543	0.864	0.047
B-BUJ	Bujaruelo	42°39'48"N	0°06'47"W	27	4.111	2.83	0.02	0.462	0.537	0.999	0.159
B-DIA	Diazas	42°37'44"N	0°05'54"W	31	4.111	2.77	0.03	0.490	0.523	0.978	0.082
C-ORD	Ordesa	42°39'17"N	0°04'45"W	31	4.667	2.96	0.11	0.559	0.548	0.830	-0.004
C-COT	Cotatuero	42°39'03"N	0°02'39"W	6	3.333	2.98	0.12	0.519	0.521	0.835	0.104
P-FON	Fondañóns	42°33'36"N	0°00'20"E	31	4.667	2.83	0.10	0.504	0.534	0.998	0.072
F-BAN	B° Bañera	42°25'37"N	0°01'13"E	31	4.222	2.84	0.04	0.493	0.521	0.977	0.071
C-URB	San Úrbez	42°33'13"N	0°03'18"E	4	2.778	2.78	0.03	0.528	0.462	0.502	0.000
B-PLA	Plana Canal	42°35'31"N	0°06'14"E	14	3.889	2.96	0.06	0.545	0.527	0.601	0.004
C-ESC	Escuaín	42°36'24"N	0°06'58"E	31	4.444	2.90	0.05	0.505	0.540	0.943	0.081
P-TEL	Tella	42°34'48"N	0°10'39"E	31	4.778	3.06	0.08	0.521	0.576	1.000	0.113
C-PIN	Valle de Pineta	42°40'03"N	0°06'06"E	18	3.778	2.81	0.01	0.481	0.528	0.935	0.118
F-REN	Renanué	42°29'32"N	0°31'43"E	18	4.000	2.84	0.04	0.468	0.532	0.999	0.151

Locality codes have a prefix indicating whether the site is within the OMPNP (C, from “core”), the buffer zone (B), the peripheral area (P) or far off (F). Number of moths genotyped (*N*), mean number of alleles (*N<sub>A</sub>*), allelic richness (*A<sub>R</sub>*), allelic private richness (*A<sub>PR</sub>*), mean observed heterozygosity (*H<sub>O</sub>*), mean expected heterozygosity (*H<sub>E</sub>*), *P*-value of Hardy–Weinberg equilibrium test (HWE (*P*-value)), and inbreeding coefficient (*F<sub>IS</sub>*)

per site were collected (Table 1). Non-lethal tissue sampling was performed by clipping a ~ 130 mm<sup>2</sup> fragment of the right hind-wing tail following Vila et al. (2009). Each wing piece was dry-stored in an individual envelope and frozen at –20 °C upon arrival to the lab.

## Molecular procedures

Genomic DNA was extracted using a commercial kit (High Pure PCR Template Preparation Kit, Roche) following the manufacturer’s instructions. We screened the total 402 sampled individuals with a set of nine polymorphic microsatellite loci and the wet-lab protocol described by Vila et al. (2010). PCR products (1.2 µL) were mixed with 16 µL formamide containing GENESCAN-500 (ROX) Size Standard (Applied Biosystems, ABI) and the allele size of PCR products was determined on a 96-capillary 3730xl DNA Analyzer (ABI). Allele peaks were scored blind, i.e., independently called by three researchers, using Geneious Prime 2020.0.1 (Biomatters Limited, <https://www.geneious.com/>). Electropherograms previously generated by Mari-Mena et al. (2016; 2019) were thoroughly reviewed with the ones from the 191 individuals newly presented in this work.

Quality of the genotypic data was evaluated by checking for Hardy–Weinberg Equilibrium (HWE) across loci and sites using Fisher’s exact test implemented in GENEPOP 4.2 (Rousset 2008). Marker independence was confirmed

after testing for gametic disequilibrium (likelihood ratio test, GENEPOP). Significance was computed by a Markov chain method using 10,000 dememorizations, 5000 batches and 5000 iterations per batch. Significance levels (alpha = 0.05) were adjusted applying the sequential Bonferroni correction whenever multiple tests were performed. Null allele frequency was estimated using the EM algorithm as implemented in FreeNA (Chapuis and Estoup 2007) with a number of bootstrap replicates fixed to 10,000.

## Genetic diversity

Genetic variability was assessed with standard descriptive statistics. Number of alleles (*N<sub>A</sub>*), observed heterozygosity (*H<sub>O</sub>*), and expected heterozygosity (*H<sub>E</sub>*) for each site and/or locus were calculated using GenAIEx 6.5 (Peakall and Smouse 2006). Allelic richness (*A<sub>R</sub>*) and allelic private richness (*A<sub>PR</sub>*) were calculated using HP-RARE (Kalinowski 2005a), which takes into account uneven sample sizes by performing rarefaction. Inbreeding coefficient (*F<sub>IS</sub>*) values were calculated with FSTAT 2.9.4 (Goudet 2005).

## Genetic structure

First, genetic differentiation was estimated as pairwise *F<sub>ST</sub>* values (10,000 permutations) between localities using ARLEQUIN 3.5.1.2 (Excoffier and Lischer 2010).

Significance level was adjusted applying sequential Bonferroni correction for multiple comparisons.

Second, we performed three Mantel tests in GENODIVE 3.05 (Meirmans 2020) to compare genetic and geographic distances. Genetic differentiation among sites was linearized [ $F_{ST}/(1-F_{ST})$ ], changing negative values of  $F_{ST}$  to zero. Three spatial matrices were used: Euclidean distances and two cost matrices obtained by least-cost-path analyses. Euclidean distances between sites were calculated using Geographic Distance Matrix Generator 1.2.3 (Ersts 2009). For the least-cost path analyses, we used the Spanish fourth national forest inventory (IFN4; Ministerio para la Transición Ecológica y el Reto Demográfico 2022b) to produce a map of suitable habitat. Since *G. isabellae* exclusively uses *P. sylvestris* as its host plant in the study area, we filtered patches described as *Pinus sylvestris* stands, and included *P. sylvestris* cover values as a proxy of patch suitability. The database includes the dominant and additional tree species for each path, and cover values expressed in percentage of the total area. The vector layer was rasterized to a 0.5 km grid, and cell values calculated as cost for dispersal. We transformed cover values into discrete values of impedance (i.e., dispersal cost for the species to the nearest cell). Total cover of *P. sylvestris* (90–100%) were classified as Weight = 1. Weight increased one unit of impedance for each 10% interval for all cover values were higher than 40%, therefore weight ranged in the interval 1–6. All other cells were considered unsuitable, and given Weight = 10, and in a second analysis Weight = 100. This allowed us to consider a neutral weight for unsuitable habitat with the same proportion of tree cover for the 1–10 weights (hereafter impedance\_10), or a high-costly unsuitable barrier for the 1–100 weights (impedance\_100). We used the Least-Cost Path extension for QGIS 3.24.2 software ([www.qgis.org](http://www.qgis.org)) to calculate the least cost route for each pair of sampling locations and generated a symmetric matrix for each of the two analyses. Significance of each Mantel test was assessed based on 10,000 permutations. A higher correlation between the genetic distance and the least-cost matrices (than the one obtained between the genetic and Euclidean distances) would point to connectivity as an important element determining dispersal pathways (see Coulon et al. 2004) for the Spanish Moon Moth. In addition, we used the partial Mantel test implemented in GENODIVE 3.05 to explore the relationship between the genetic data and the two cost matrices (impedance\_10 and impedance\_100) after controlling for the effect of the geographic distance matrix. Significance was assessed based on 10,000 permutations and sequential Bonferroni.

As a significant Mantel test does not necessarily prove isolation-by-distance (IBD) (reviewed by Diniz-Filho et al. 2013), to properly address the “cline versus cluster dilemma” (Guillot et al. 2009), we followed Meirmans

(2012), Legendre et al. (2015), and Perez et al. (2018) and evaluated the shape of the genetic structure across the space calculating a Mantel correlogram. We used the matrix of Euclidean distances, calculated equifrequent geographic distance classes, and assessed significance after 10,000 permutations (and sequential Bonferroni) as implemented in GENODIVE 3.05. Following Klug et al. (2011) we calculated the correlogram using different number of distance classes (five, eight, and eleven).

Third, we used two different Bayesian approaches to assign individuals to clusters and to estimate the admixture proportions to each inferred cluster, taking the geographical position of the samples into account. (1) We ran STRUCTURE 2.3.4 (Pritchard et al. 2000) assuming admixture model and correlated allele frequencies. After preliminary runs and given the slight overall structuring ( $F_{ST} < 0.1$ ), each run was performed using the LOCPRIOR model implemented in STRUCTURE that uses sampling information as prior (Hubisz et al. 2009), with a burn-in of 100,000 iterations followed by 300,000 MCMC iterations for parameter estimation. Population structure was tested at  $K$  values ranging from 1 to  $N$ , where  $N$  was the number of sites included in the analysis, and with ten replicates per each  $K$ . Results were summarised using STRUCTURE HARVESTER (Earl and von Holdt 2012) and different partitions of the data were explored by examining both the log probability of the data ( $\ln Pr(X|K)$ ) and the  $\Delta K$  statistic following Evanno et al. (2005). CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007) was used to permute the admixture coefficients ( $Q$ -values) for the several independent runs resulting for the chosen  $K$ -value. Finally, DISTRUCT 1.1 (Rosenberg 2004) was employed to visualise the output from CLUMPP. (2) We used the Bayesian clustering algorithm implemented in TESS 2.3 (Durand et al. 2009) to model spatial dependencies in allele frequencies at a finer scale, i.e., including only the sites within the OMPNP boundaries and those from the buffer and peripheral areas (11 sampling sites, see Table 1). We used the BYM admixture model with a linear trend surface, updating the variance term (initially set to 1.0) during the course of the runs. We performed 100 independent runs for each  $K$  value, ranging from 2 to 11. The burn-in period was set to 10,000 sweeps, and MCMC sampling was run for 50,000 sweeps. We showed the partition of the data that produced the lowest value of the Deviance Information Criterion (DIC) for the 10% of the runs for each  $K$ . The estimated admixture coefficients of the 10% runs with the lowest DIC of the best  $K$  were averaged using CLUMPP 1.1.2. A graphical interpolation procedure was used to obtain a map showing the geographic variation of the  $Q$ -values. We displayed spatial patterns of admixture coefficients using kriging interpolation.

Four, we explored the spatial patterns of genetic variability of *G. isabellae* across the sampled area by performing

a spatial principal component analysis (sPCA) using the R package *adegenet* (Jombart 2008). sPCA is a multivariate analysis that ascertains the spatial autocorrelation (Moran's  $I$ ) among individual genotypes. The proximity of individuals was defined using a Gabriel graph (type = 2) as connection network, and two Monte Carlo tests were performed to assess the significance of global and local structures (9,999 permutations). Global scores can identify genetically distinguishable groups, clines in allele frequency, and intermediate individuals, whereas local scores may detect differentiation between neighbouring individuals.

## Migration

Contemporary estimates of gene flow ( $m$ : proportion of migrants per generation) among sampling sites were estimated using a Bayesian MCMC framework implemented in BAYESASS 3.0.4 (Wilson and Rannala 2003). The migration estimates produced by BAYESASS correspond to the last 1–3 generations and should be interpreted as upper limits of recent connectivity, bearing in mind that the actual level of connectivity may be lower (Samarasin et al. 2017). In order to increase accuracy and avoid overestimation of migration rates (Gottelli et al. 2012), sampling sites with less than twenty individuals were eliminated from the analysis. To estimate the posterior probability distributions of the parameters, MCMC were run for 50 million iterations, discarding the first 10 million iterations as burn-in, and sampling every 1000 generations. After preliminary runs, delta values were tuned individually to obtain acceptance rates within 40–60% of the total (Faubet et al. 2007) as follows:  $\Delta a = 0.35$ ,  $\Delta f = 0.95$ , and  $\Delta m = 0.95$ . Ten independent runs with different random starting seeds were performed to assess MCMC convergence and evaluate the consistency of the results. TRACER 1.7.1 (Rambaut et al. 2018) was used to visualise MCMC simulations and confirm convergence of chains among runs. We used the R script provided by Meirmans (2014) to calculate the Bayesian deviance for each run to determine the run that best fit the data (the one with the lowest deviancy).

## Results

### Dataset quality

We obtained multilocus genotypes for the whole set of 402 individuals. The number of individuals analysed from each sampling site varied from four at C-URB to 32 at F-CAB (Table 1). After the sequential Bonferroni correction, only two out of the 153 site-locus combinations (1.31%) showed significant deviation from Hardy–Weinberg equilibrium (HWE). Specifically, locus *GRAISA25* did not conform

HWE in F-BEL, and locus *GRAISA15* in B-DIA (Table S1). In both cases, departure from HWE as due to a heterozygote deficit, but no deviations from HWE were detected at locality level by the global tests (Table 1). Null allele frequencies were low, varying from 0.010 to 0.165 across loci (Table S2). In detail, only two loci exhibited a high null allele frequency (above 0.2 according to Chapuis and Estoup (2007)) in specific sites, *GRAISA17* in B-BUJ and C-URB, and *GRAISA25* in F-BEL (Table S1). Therefore, since none of the loci showed consistency in the presence of null alleles across localities, further analyses were performed on multi-locus data from all nine microsatellites. According to  $F_{IS}$  values, two out of 153 site-locus combinations showed a significant homozygote excess after sequential Bonferroni correction, both corresponding to site F-BEL and affecting loci *GRAISA17* and *GRAISA25* (Table S1). However, all the estimates of global  $F_{IS}$  resulted non-significant (Table 1).

No linkage disequilibrium was observed for any pair of loci after sequential Bonferroni correction. Hence, all microsatellites were considered as independent markers. Estimators of genetic variability per site, per locus, and per site and locus are available in Table 1, Table S1, and Table S2. Among the microsatellite loci, the number of alleles ranged from 3 (*GRAISA23*) to 14 (*GRAISA11*), with 11.29% of all scored alleles being private alleles (Table S1 and Table S2). Loci *GRAISA23* and *GRAISA11* were the least and most polymorphic of the microsatellite suite, averaging  $H_O$  0.028 and 0.829, and  $H_E$  0.038 and 0.778, respectively.

### Genetic diversity

The levels of allelic diversity and heterozygosity were relatively similar among all sites (Table 1). Over localities,  $H_O$  and  $H_E$  averaged 0.501 and 0.528, respectively. The proportion of unique alleles ranged from 1% in B-LIN and C-PIN to 12% in C-COT. The lowest values of allele richness and  $H_E$  were found at the westernmost site of Belabarce (F-BEL), whereas the highest ones were found at Tella (P-TEL) followed by Ordesa (C-ORD) and Escuaín (C-ESC). Genetic diversity seems to only slightly decrease toward the westernmost site (F-BEL) of the study area.

### Population structure

The overall genetic differentiation among sites was low ( $F_{ST} = 0.038$ ), with pairwise estimates between sites in the range (–0.013, 0.134) (Table 2).  $F_{ST}$  values revealed: (i) differentiation of F-BEL and F-REN (the westernmost and easternmost localities of the sampling, respectively) with most of the remaining sites; (ii) connectivity among F-JUA/F-AGU/F-CAB/B-LIN/B-BUJ/P-FON; and (iii) connectivity of F-BAN (south of the park, Fig. 1) with both western (F-JUA, F-CAB, B-LIN, B-BUJ, B-DIA, and

**Table 2** Genetic differentiation between sampled sites

	F-BEL	F-JUA	F-AGU	F-CAB	B-LIN	B-BUJ	B-DIA	C-ORD	C-COT	P-FON	F-BAN	C-URB	B-PLA	C-ESC	P-TEL	C-PIN	F-REN
F-BEL	–	<b>0.001</b>	<b>0.000</b>	<b>0.000</b>	0.002	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	0.435	<b>0.000</b>	<b>0.000</b>	0.038	0.050	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
F-JUA	0.040	–	0.006	<b>0.000</b>	0.054	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	0.042	<b>0.000</b>	0.009	0.178	0.027	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
F-AGU	0.041	0.026	–	<b>0.000</b>	0.001	0.013	<b>0.000</b>	<b>0.000</b>	0.081	0.001	<b>0.000</b>	0.057	0.253	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
F-CAB	0.064	0.039	0.036	–	0.005	<b>0.000</b>	<b>0.000</b>	0.065	0.001	0.034	0.380	0.242	0.001	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
B-LIN	0.056	0.027	0.047	0.031	–	0.021	0.005	0.009	0.242	<b>0.001</b>	0.017	0.481	0.022	0.002	0.005	<b>0.000</b>	<b>0.000</b>
B-BUJ	0.039	0.039	0.020	0.023	0.033	–	0.954	0.963	0.565	0.095	0.001	0.221	0.692	0.013	0.146	<b>0.000</b>	<b>0.000</b>
B-DIA	0.040	0.047	0.029	0.034	0.047	–0.008	–	0.034	0.459	0.072	0.001	0.137	0.175	0.015	0.041	<b>0.000</b>	<b>0.000</b>
C-ORD	0.044	0.064	0.036	0.028	0.031	–0.008	0.013	–	0.240	<b>0.000</b>	<b>0.000</b>	0.051	0.261	0.001	0.008	<b>0.000</b>	<b>0.000</b>
C-COT	0.006	0.048	0.030	0.031	0.017	0.001	0.003	0.010	–	0.239	0.052	0.451	0.670	0.400	0.760	0.076	0.002
P-FON	0.050	0.052	0.024	0.026	0.050	0.011	0.011	0.027	0.014	–	0.002	0.311	0.679	0.001	0.015	<b>0.000</b>	<b>0.000</b>
F-BAN	0.059	0.022	0.036	0.013	0.031	0.029	0.027	0.052	0.035	0.022	–	0.653	0.106	0.003	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
C-URB	0.067	0.035	0.048	0.008	0.001	0.027	0.033	0.037	0.002	0.014	–0.008	–	0.345	0.133	0.125	0.018	0.039
B-PLA	0.021	0.029	0.006	0.006	0.036	–0.004	0.009	0.004	–0.013	–0.004	0.013	0.006	–	0.462	0.879	0.137	<b>0.000</b>
C-ESC	0.051	0.058	0.047	0.027	0.043	0.020	0.017	0.026	0.006	0.025	0.022	0.033	0.000	–	0.685	0.001	<b>0.000</b>
P-TEL	0.042	0.054	0.034	0.036	0.038	0.009	0.014	0.017	–0.009	0.016	0.029	0.033	–0.009	–0.002	–	0.008	0.001
C-PIN	0.068	0.064	0.053	0.070	0.080	0.063	0.072	0.076	0.038	0.052	0.047	0.075	0.015	0.036	0.024	–	<b>0.000</b>
F-REN	0.134	0.089	0.107	0.104	0.080	0.076	0.084	0.103	0.077	0.070	0.059	0.060	0.075	0.060	0.039	0.062	–

Pairwise  $F_{ST}$  (below the diagonal), and the corresponding  $P$ -values (above the diagonal). Values in bold indicate significance after sequential Bonferroni correction. Site names are coded as in Table 1

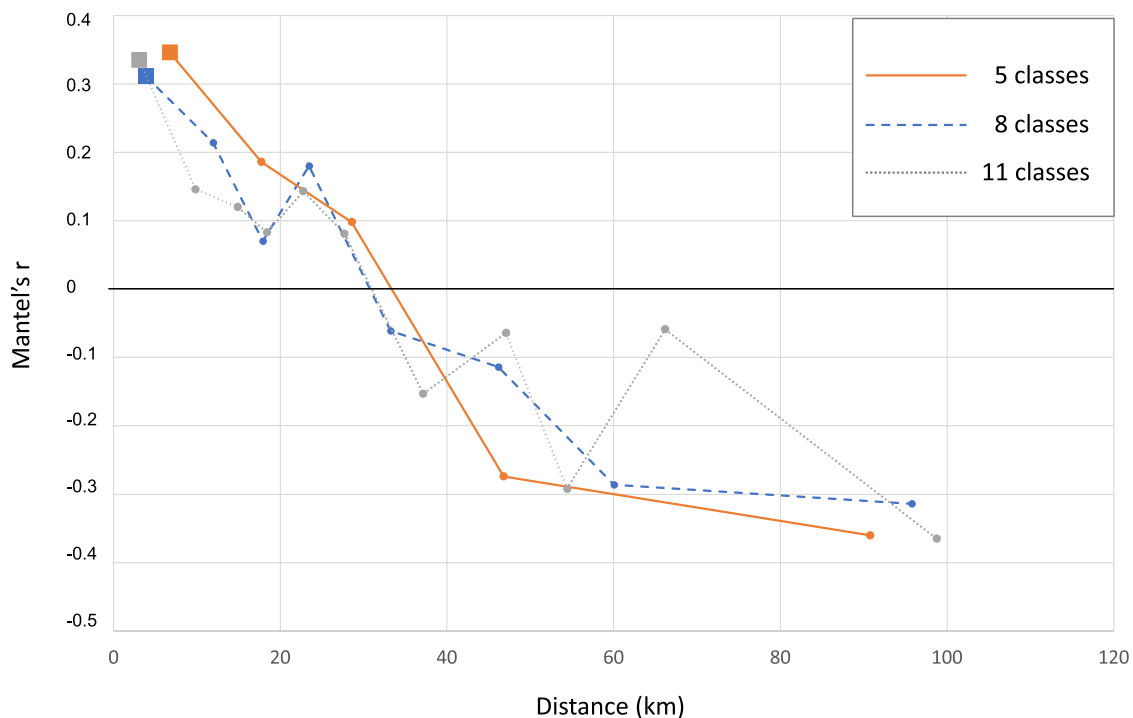
P-FON) and eastern sites (B-PLA and C-ESC). Focusing on the core sites of the OMPNP, buffer, and peripheral areas,  $F_{ST}$  values showed that: (i) the differentiation between C-PIN (the northeasternmost site) and the sites sampled to the west of the OMPNP (B-LIN, B-BUJ, B-DIA, C-ORD, and P-FON) was always significant; (ii) no significant differentiation was found among the samples located at the eastern side of the Añisclo Valley (B-PLA, C-ESC, and P-TEL), as well as between those and C-PIN; (iii) significant differentiation between sites in the northwest valleys (B-LIN, B-BUJ, C-ORD, B-DIA, and C-COT) and those in the southwest valleys (P-FON, C-URB, B-PLA, C-ESC, and P-TEL); and (iv) significant differentiation between P-FON and two localities in the western valleys of the National Park, B-LIN (Bujaruelo Valley) and C-ORD (Ordesa Valley).

Mantel tests always rejected the null hypothesis of absence of relationship between the genetic and geographic dissimilarity matrices. In addition, Mantel's  $r$  was always higher when using the geographic distance considering constraints from tree cover (whole dataset: Euclidean distance,  $r_M=0.585$ ,  $P=0.004$ , impedance\_10:  $r_M=0.764$ ,  $P=0.000$ ; impedance\_100:  $r_M=0.703$ ,  $P=0.000$ ). However, the association between genetic and geographic distance decreased at the small scale (11 sites dataset: Euclidean distance:  $r_M=0.242$ ,  $P=0.035$ ; impedance\_10:  $r_M=0.577$ ,  $P=0.001$ ; impedance\_100:  $r_M=0.664$ ,  $P=0.001$ ). The influence of

tree cover constrains on genetic distance is additionally supported by the partial Mantel tests controlling for geographic distance (whole dataset: impedance\_10,  $r_M=0.645$ ,  $P=0.002$ ; impedance\_100,  $r_M=0.544$ ,  $P=0.005$ ; 11 sites: impedance\_10  $r_M=0.622$ ,  $P=0.002$ ; impedance\_100,  $r_M=0.644$ ,  $P=0.008$ ).

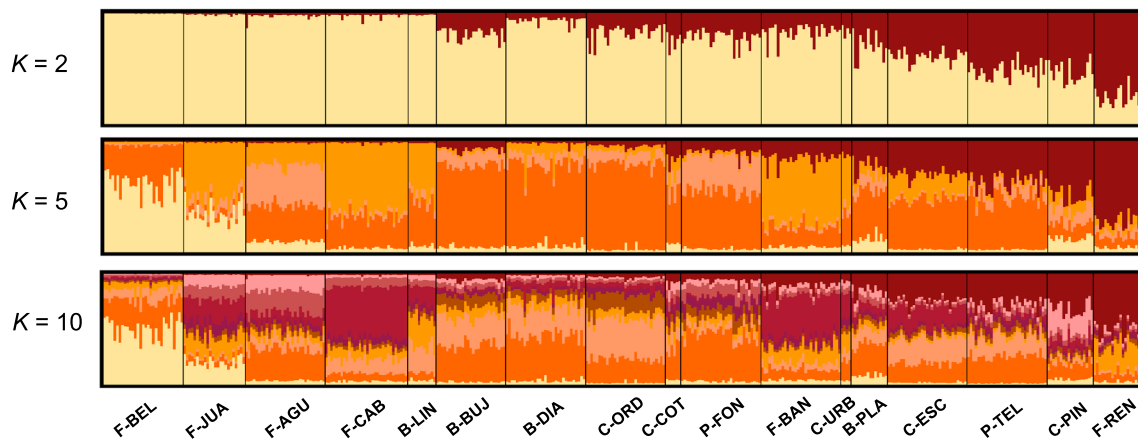
Dispersal was restricted within 6 km, as indicated by the significant and positive autocorrelation in the first distance class regardless of the number of distance classes used to calculate the correlogram (Fig. 2). When five distance classes were used, the  $r_M$  value was significantly positive at 12 km (midpoint of 0–25 km). When eight distance classes were used, the  $r_M$  value was significantly positive at 4 km (midpoint of 0–8 km). Lastly, when eleven distance classes were used, the  $r_M$  value was significantly positive at 3 km (midpoint of 0–6 km). In addition, the three correlograms firstly intercepted with the  $x$ -axis at 30–33 km (Fig. 2), which can be taken as an estimate of the genetic patch size (Diniz-Filho et al. 2002).

Both Bayesian clustering methods identified a west–east gradient in allele frequencies along the entire sampled area ( $K=2$ ) (Fig. 3). On the one hand, STRUCTURE pointed to  $K=5$  as a meaningful partition. However, two other groupings were also supported by Evanno's method:  $K=2$  and  $K=10$  (Fig. S1). None of these three partitions identified a sharp subdivision of the samples into distinct clusters. In order to get a more detailed picture of the manageable area



**Fig. 2** Spatial autocorrelogram for *G. isabellae* in the whole Pyrenean study area. Solid squares indicate significance after sequential Bonferroni correction





**Fig. 3** Bayesian assignment probabilities for  $K=2$ ,  $K=5$ , and  $K=10$  revealed by STRUCTURE. Analysis including 17 sampling sites from the studied area: Ordesa y Monte Perdido National Park (OMPNP), buffer zone, peripheral area, and Pyrenean surroundings. Each verti-

cal bar depicts an individual and each colour represents the probability of membership to an inferred cluster. Names above the plot represent site codes (Table 1)

in terms of conservation issues, another run of STRUCTURE was carried out including the 235 individuals sampled at the OMPNP, buffer and peripheral areas. Regardless of using  $\Delta K$  or  $\text{Pr}(X/K)$  as optimization criterion (Fig. S1), STRUCTURE indicated a main partition into two clusters ( $K=2$ ), and a secondary grouping into four clusters ( $K=4$ ) (Fig. S2). Once again, a clinal differentiation with a west–east pattern was evident throughout the OMPNP and its immediate vicinity. The membership coefficients of the sampled sites were used to generate pie charts separately for each location to illustrate the geographical pattern of assignment to each cluster at  $K=2$  (Fig. S3). On the other hand, the Bayesian spatial clustering method implemented in TESS was also used for seeking population structuring along the OMPNP dataset (OMPNP, buffer, and peripheral zones). According to the Deviation Information Criterion (DIC),  $K=2$  was the best clustering provided by TESS for the data (Fig. S4). TESS assigned individuals to each of the two inferred clusters similarly to STRUCTURE, unravelling again a clinal variation between westernmost and easternmost locations of the park, with the moths captured in the eastern localities showing similar membership coefficients for both groups (Fig. 4).

Lastly, the sPCA found a significant global spatial structure (global Monte Carlo test  $P=0.002$ ) across all sampling sites, while there was no evidence of local structures (local Monte Carlo test  $P=0.665$ ). Only the first positive eigenvalue was retained, so we assessed the first global scores, which suggested a clinal pattern (Fig. S5), supporting the previous results (Fig. 3). As for the OMPNP dataset, another sPCA was performed on this subset of samples, but no significant global ( $P=0.212$ ) or local ( $P=0.397$ ) structuring was retrieved (data not shown).

## Migration

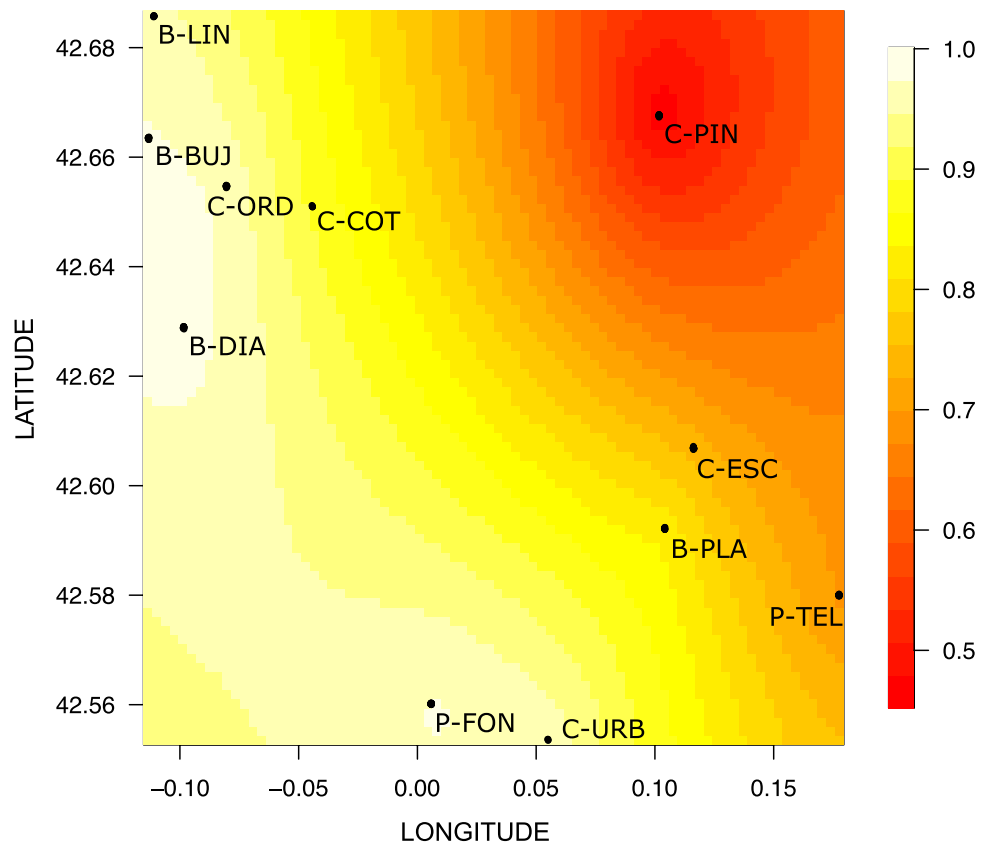
Almost all contemporary migration rates inferred by BAYESASS were low and not significant, i.e., 95% credible intervals included zero (Table 3). One particular site (B-BUJ, buffer zone of the OMPNP) showed higher rates of migration than the rest of localities used in this analysis ( $N > 20$ ), and seven out of them resulted significantly different from zero: from B-BUJ to C-ORD, C-ESC, B-DIA, P-FON, P-TEL, and F-BEL, F-AGU. Likewise, F-JUA was found to receive a significant proportion of migrants from B-BEL and F-BAN.

## Discussion

### A west–east cline

*Graellsia isabellae* showed a longitudinal cline of genetic differentiation not only in and around the Ordesa y Monte Perdido National Park (OMPNP), but also across the whole study area. Gene flow in this iconic moth thus proved dependent on both topography and the availability of suitable habitat: patches of *Pinus sylvestris*. The most plausible explanation for the cline is a discontinuous stepping stone model, where migration mainly occurs between adjacent subpopulations. However, geographic clines in genetic variation may be a consequence of local adaptation along an environmental gradient (Pruisscher et al. 2018). No such gradient was found by Chefaoui and Lobo (2007) and the microsatellite markers we applied proved to be selectively neutral (Marí-Mena 2013). Nevertheless, the genetic pattern resulting from a low number of markers might still be attributable to selection (Aguirre-Liguori et al. 2020).

**Fig. 4** Spatial Bayesian clustering revealed by TESS. Black circles indicate the relative positions of the sampled sites (coded as in Table 1). Darker and lighter shading are proportional to posterior probabilities of membership to clusters, with lighter yellow or darker red areas showing the highest posterior probabilities



The finding of restricted gene flow within at least 6 km fits the available information about dispersal ability of *G. isabellae*. Males, the dispersing sex, have a very short lifespan (ca. 6 days) (Ylla i Ullastre 1997) and fly only for a few hours per night (approximately from dusk to midnight). They can fly up to two km per night, which roughly makes a maximum dispersal of twelve km per male (Marí-Mena et al. 2019).

The correlograms also showed a “stabilizing profile” (sensu Diniz-Filho et al. 2002), which indicates a spatial range of genetic similarity that is congruent with the aforementioned cline. Using the intercept with the x-axis we estimated a genetic patch size of 30–33 km, so that moths occurring within that distance are expected to be genetically more similar than two random individuals taken from the study area, whereas specimens further apart might be considered as genetically “independent”. This 30–33 km result may be a first step to optimise sampling strategies to be used in conservation plans. However, the 30–33 km estimated patch size should be taken with caution, as our least-cost path analysis showed that pine tree forest cover influences genetic distances and patch sizes based on the intercept of the autocorrelogram are known to be sensitive to sampling design (Zeng et al. 2010). A patch size of 100–150 km was estimated for the Baltic Sea northern pike (*Esox lucius*), which showed a similar overall  $F_{ST}$  (0.037),

clinal pattern, and short dispersal (most individuals moving less than ten km, Laikre et al. (2005)). The difference between the patch sizes of *G. isabellae* and *E. lucius* may be related to their different habitats. A pelagic species may have similar dispersal costs all through the marine space, so that the patch size obtained from Euclidean distances can be accurate. However, the 30–33 km we estimated may require a higher dispersal cost for the moth in areas where pine tree forest cover decreases. Our work paves the way for future studies aiming to determine how the genetic patch size may vary depending on pine tree forest cover.

The genetic pattern of *G. isabellae* in the Western and Central Pyrenees is likely influenced by two Isolation by Distance (IBD) derived processes: Isolation-by-Barrier (IBB) and Isolation-by-Resistance (IBR) (Balkenhol et al. 2017). The former is best illustrated by the action of the Monte Perdido massif, which lies in the centre of the OMPNP and shapes its valleys, preventing straightforward gene flow between the eastern Pineta Valley and the westernmost valleys of Bujaruelo, Broto, Ordesa, and Vió. *P. sylvestris* is not present above 1800 m a.s.l. in the Monte Perdido massif, and *G. isabellae* males are rarely seen flying over those elevations (Breton et al. 2017). The genetic cline proves that the pine forests encircling the Monte Perdido massif (e.g., B-PLA, P-FON) allow the gene flow between

**Table 3** Migration rates between sites estimated by BAYESASS

into	from										
	F-BEL	F-JUA	F-AGU	F-CAB	B-BUJ	B-DIA	C-ORD	P-FON	F-BAN	C-ESC	P-TEL
<b>F-BEL</b>	0.827 (0.739-0.914)	0.009 (0-0.027)	0.011 (0-0.033)	0.015 (0-0.041)	0.083 ( <b>0.002-0.164</b> )	0.008 (0-0.025)	0.009 (0-0.027)	0.009 (0-0.027)	0.011 (0-0.031)	0.009 (0-0.027)	0.009 (0-0.026)
<b>F-JUA</b>	0.091 ( <b>0.022-0.160</b> )	0.687 (0.650-0.724)	0.024 (0-0.068)	0.038 (0-0.119)	0.015 (0-0.043)	0.011 (0-0.030)	0.012 (0-0.034)	0.010 (0-0.030)	0.092 ( <b>0.003-0.182</b> )	0.011 (0-0.031)	0.010 (0-0.029)
<b>F-AGU</b>	0.028 (0-0.084)	0.013 (0-0.039)	0.693 (0.644-0.742)	0.052 (0-0.135)	0.138 ( <b>0.025-0.250</b> )	0.009 (0-0.032)	0.011 (0-0.032)	0.011 (0-0.034)	0.024 (0-0.068)	0.011 (0-0.034)	0.011 (0-0.031)
<b>F-CAB</b>	0.020 (0-0.058)	0.009 (0-0.027)	0.015 (0-0.043)	0.762 (0.639-0.884)	0.056 (0-0.150)	0.008 (0-0.025)	0.013 (0-0.037)	0.011 (0-0.032)	0.082 (0-0.214)	0.015 (0-0.041)	0.010 (0-0.028)
<b>B-BUJ</b>	0.017 (0-0.048)	0.012 (0-0.036)	0.018 (0-0.050)	0.015 (0-0.044)	0.855 (0.781-0.928)	0.010 (0-0.032)	0.020 (0-0.056)	0.012 (0-0.036)	0.015 (0-0.042)	0.013 (0-0.037)	0.013 (0-0.039)
<b>B-DIA</b>	0.015 (0-0.042)	0.008 (0-0.024)	0.014 (0-0.040)	0.014 (0-0.040)	0.211 ( <b>0.154-0.268</b> )	0.675 (0.659-0.691)	0.013 (0-0.035)	0.009 (0-0.025)	0.024 (0-0.055)	0.009 (0-0.026)	0.009 (0-0.026)
<b>C-ORD</b>	0.022 (0-0.058)	0.009 (0-0.027)	0.015 (0-0.044)	0.028 (0-0.069)	0.181 ( <b>0.114-0.249</b> )	0.010 (0-0.029)	0.690 (0.657-0.724)	0.012 (0-0.033)	0.011 (0-0.032)	0.011 (0-0.031)	0.011 (0-0.032)
<b>P-FON</b>	0.015 (0-0.044)	0.009 (0-0.026)	0.018 (0-0.049)	0.019 (0-0.053)	0.207 ( <b>0.141-0.272</b> )	0.009 (0-0.027)	0.015 (0-0.039)	0.678 (0.655-0.700)	0.011 (0-0.031)	0.011 (0-0.031)	0.010 (0-0.028)
<b>F-BAN</b>	0.014 (0-0.040)	0.010 (0-0.030)	0.015 (0-0.044)	0.091 (0-0.204)	0.059 (0-0.123)	0.009 (0-0.025)	0.011 (0-0.032)	0.010 (0-0.030)	0.754 (0.657-0.850)	0.015 (0-0.047)	0.013 (0-0.037)
<b>C-ESC</b>	0.014 (0-0.039)	0.009 (0-0.027)	0.013 (0-0.036)	0.032 (0-0.086)	0.161 ( <b>0.079-0.242</b> )	0.009 (0-0.025)	0.011 (0-0.032)	0.010 (0-0.029)	0.037 (0-0.089)	0.690 (0.647-0.733)	0.016 (0-0.048)
<b>P-TEI</b>	0.011 (0-0.033)	0.009 (0-0.027)	0.014 (0-0.039)	0.017 (0-0.049)	0.183 ( <b>0.103-0.264</b> )	0.010 (0-0.029)	0.011 (0-0.032)	0.011 (0-0.033)	0.021 (0-0.054)	0.030 (0-0.092)	0.684 (0.649-0.718)

In parentheses, 95% credible intervals (CIs). Values in bold indicates significant migration values from the source into the recipient population (i.e., 95% CI does not include zero). The proportion of individuals that remain in their locality of origin are shown in grey. Sites with less than 20 individuals sampled were not included in this analysis. Site names are coded as in Table 1

the eastern cul-de-sac sites (such as C-PIN and C-ESC) and the western valleys of the OMPNP.

With regard to IBR, we showed that gene flow in *G. isabellae* is affected by patchiness of the host-plant distribution, as previously reported in other phytophagous insects (Crawford et al. 2011; Groot et al. 2011). Our least-cost path analysis using pine tree cover as a proxy of connectivity may be constrained by the following issues. Firstly, the calculated least-cost path was based on the assumption that dispersing individuals perfectly know the whole landscape and move through it in an optimal way (reviewed by Balkenhol et al. 2017). Secondly, the available vegetation map depicts the current vegetation, but not the historical changes in land cover. The increase in forest cover of the National Park that has occurred in recent decades as a consequence of the reduction of traditional agricultural activities in the province of Huesca (Martínez-Vega et al. 2017; García et al. 2019) obviously benefited the connectivity of *G. isabellae*. Combining the information from old vegetation maps and genetic data retrieved from museum specimens should help to estimate the distance among patches needed to preserve connectivity.

Our sampling design comprised most of the suitable habitat patches within the OMPNP/buffer/peripheral area and several others scattered throughout the Western and Central Pyrenees. This two-scale sampling resulted beneficial for different analyses, such as the correlations between genetic and geographic distances. Also, it was helpful to overcome the difficulty of Bayesian clustering algorithms to distinguish between a pattern of IBD and discrete population structure in continuously distributed organisms exhibiting IBD but not properly sampled (Anderson et al. 2010). The most popular clustering method, STRUCTURE, is sensitive to patterns of IBD and may overestimate a number of clusters where there is only a single large area with IBD. Indeed, several partitions ( $K=4$  for the protected area and  $K=5$  and  $K=10$  for the whole dataset, Fig. S1) supported by the most commonly used evaluation methods did not add any relevant information to  $K=2$ .

### Source-sink dynamics

We found that Bujaruelo (B-BUJ) acted as a source of eastward migrants into the following five sites (from west

to east): Diazas (B-DIA), Ordesa (C-ORD), Fondañóns (P-FON), Escuaín (C-ESC), and Tella (P-TEL). Moreover, Bujaruelo also contributed with a significant proportion of western migrants to Belabarce (F-BEL) and Pico del Águila (F-AGU). Given the dispersal ability of males of *Graellsia isabellae*, the migration obtained between distant locations must occur in a stepping-stone manner through intermediate patches of suitable habitat, as proved in other lepidopterans (Ugelvig et al. 2012 and references therein).

Our results revealed a tendency of gene flow to primarily occur from western to eastern sites in the National Park and adjoining areas. Source populations are thought to occur in good-quality habitat, produce a net surplus of offspring, and send out many emigrants. In contrast, sink populations are hypothetically located in habitat of lower quality, have a net deficit of offspring, and receive more immigrants than producing emigrants (Holderegger and Gugerli 2012). According to this, Bujaruelo (B-BUJ) likely occurs in a forest patch with high local breeding.

We can only speculate about the reason of asymmetrical migration in our results. Diazas (B-DIA) and Ordesa (C-ORD) are very close to Bujaruelo (B-BUJ) (~4 and 3 km, respectively), and occur in large stands of *P. sylvestris* virtually connected (Fig. S3). Wind might be facilitating the movement of males from B-BUJ into west and south direction as dominance of northern and eastern winds has been reported in the Ordesa Valley (Benito-Alonso 2006). Another cause of the asymmetric migration could be the transition between Atlantic and Mediterranean climatic/ecological conditions in the southern part of the OMPNP (Benito-Alonso 2006). The buffer zone is mainly occupied by Mediterranean scrubland in the south and alpine grassland and scrubland in the northwest (Benito-Alonso 2006), which could be affecting the population dynamics of *G. isabellae*, somehow enhancing the migration from the northwest of the park towards the southern valleys.

San Juan de la Peña (F-JUA) was identified as a sink of migrants from two sites located in opposite directions: Belabarce (F-BEL) to the west, and Barranco de la Bañera (F-BAN) to the east. San Juan is located within a protected area (Protected Landscape, Natura2000 network) where pine tree forest is well conserved. However, the surrounding pine tree forest is smaller than in F-BEL, F-Cabas or even F-BAN (Fig. 1). As landscape-scale intensive land use affects negatively to insects (Uhl et al. 2020), one might argue the larger unforested area around San Juan to be the reason of receiving migrants from west and east. However, a higher extent of habitat isolation does not explain unidirectional gene flow towards F-JUA.

The low overall genetic differentiation may have biased the migration rates estimated by BAYESASS (Faubet et al. 2007). Again, future studies using a higher number of markers will be necessary to test the population structure we are

reporting (Kalinowski 2005b; Aguirre-Liguori et al. 2020). In addition, not only capture-mark-recapture studies are needed to confirm the source-sink dynamics we are reporting at Bujaruelo and San Juan de la Peña, but also other factors need to be explored to explain asymmetrical gene flow, such as winds (e.g., Kling and Ackerly 2021) or predation pressure (e.g., Minnie et al. 2018).

## Conservation implications

Marí-Mena et al. (2019) estimated that *G. isabellae* had a low  $N_e$  (< 50) in Ordesa (C-ORD). They recommended to safeguard the connection of *P. sylvestris* patches to ensure gene flow as a cautionary measure for the conservation of this emblematic insect. Yet, C-ORD would not necessarily be at risk as long as gene flow with other populations is possible (Waples 2010). Forest patches of *Pinus sylvestris* currently enable population connectivity of *G. isabellae* in this Pyrenean region, and reforestation dynamics are expected to continue, at least, in the short-term (Vacquie et al. 2015). Therefore, conservation efforts should focus on maintaining the connectivity of Scots pine woodlands while avoiding the disappearance of sensitive open areas where several declining lepidopterans occur (Murria-Beltrán 2022). Maintaining the connectivity of Scot pine forests thus means to address the main threats shared by *P. sylvestris* and *G. isabellae*. Firstly, both species are negatively affected by summer drought (Chefaoui and Lobo 2007; Galiano et al. 2010). Severe droughts are predicted to increase with rapid climate change (Samaniego et al. 2018), which along with emerging diseases such as pine wilt (Vicente et al. 2012) could lead to strong changes in pine forests and cascade effects on its associated entomofauna. Secondly, heavy (> 50%) defoliation by the pine processionary moth (*Thaumetopoea pityocampa*) has been shown to negatively affect *G. isabellae* (Imbert et al. 2012). Lastly, large wildland fires are a huge problem in Spain. Luckily, the study area has not suffered any megafire yet (Cardil et al. 2016, Ministerio para la Transición Ecológica y el Reto Demográfico 2023).

Should Pyrenean Scots pine forests decline in the study area, conservation managers might be interested in designing reserves that preserve most of the current genetic diversity of *G. isabellae*. In that case, it will be useful to know that (i) most localities sampled for the present study harbour similar levels of genetic diversity, regardless of elevation, (ii) preserving populations in the extremes of the spatial range (Belabarce-Renanué) ensures that most of the among-population genetic divergence will be sampled, and (ii) the genetic patch size was estimated in 30–33 km (Euclidean distance).

Our results also highlighted the importance of buffer zones, which have a dual role: conservationist, providing additional protection of core conservation areas from

disturbance, and socioeconomic, benefiting local human communities with goods and services (Ahmad et al. 2013; Pringle 2017). On the one hand, the very low sampled obtained at C-URB might be a consequence of the lack of a buffer zone around the southern boundaries of the OMPNP. On the other hand, the finding of B-BUJ as a source sub-population reinforces the protector role of the buffer zone. Indeed, Bujaruelo is a Natura 2000 Site of Community Importance (code ES2410006) and, like the National Park, it belongs to the Ordesa-Viñamala Biosphere Reserve. Previous work has sought to raise the protection status of the adjoining areas of the National Park in order to prevent negative external pressures on sensitive landscape patches (Benito Alonso 2006, Martínez de Pisón, 2010), but the human activities permitted there make this attempt difficult.

## Conclusions

Given the strong interest of the Ordesa y Monte Perdido National Park in the conservation of the protected moth *G. isabellae*, we inform that the current forest patches of *P. sylvestris* maintain a west–east clinal distribution of genetic variation not only within the National Park, buffer, and peripheral areas, but also at a wider scale (provinces of Navarre and Huesca). We suggest to pay special attention to the southern site of San Úrbez (C-URB), given the narrow fringe of buffer zone in that area. In addition, a capture-mark-recapture study is needed to confirm whether Bujaruelo (B-BUJ) and San Juan (F-JUA) are acting as source and sink of individuals, respectively, as revealed by our genetic results. Lastly, we encourage the National Park to keep the monitoring of *G. isabellae*, an invaluable source of information in the face of rapid climate change.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10592-023-01536-z>.

**Acknowledgements** We thank N. Remón for her valuable help in the laboratory. J.C. Bellver, F. Carmena, O. Decker, M. Drago, G. Fernández, F. Goussard, G. Horcajada, M. Martínez, P. Lincango, J. Sevilla, S. Uceira, and M. Villanueva contributed to fieldwork with their advice, assistance, and equipment. We appreciate the support provided by the staff of the Ordesa y Monte Perdido National Park, especially E. Villagrana, V. Sanz, and F. Villaespesa. This work benefited from the discussions held with E. Murria-Beltrán, H. Naveira and L. Torrado-Blanco. We thank the two anonymous reviewers who carefully read our manuscript for their many insightful comments and suggestions. All samples were collected with appropriate licenses. Figure S3 was adapted from "Mapa de Hábitat de Aragón" by J.L. Benito with permission from "Dirección General de Conservación de la Naturaleza" (Gobierno de Aragón).

**Author contributions** MV, CLV, and GS conceived the study. MV and CLV obtained funding. NMM, MV, and CLV collected samples. NMM generated the raw data. IGC, NMM, GS, JGL, and JF analysed the data. IGS, JGL, and JF prepared the figures and tables. IGC, NMM, and

MV wrote the original manuscript. All authors reviewed drafts of the manuscript and approved the final version.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work was financially supported by grants from the Spanish Ministry of Education and Science (CGL2007-63549/BOS), the European Science Foundation [CONGEN short-visit (1635) and exchange (1683) Grants], and the University of A Coruña. IGC was funded by the Spanish Ministry of Universities (Margarita Salas postdoctoral fellowship, ref. RSU.UDC.MS17). NMM was funded by the Spanish Ministry of Science (FPI programme, ref. BES-2008–002571).

**Data availability** Accession codes: Sequences containing the microsatellite loci were deposited in GenBank under Accession Numbers FJ871975–FJ871983. A Genepop file with the multilocus genotypes used in the present study is available as Supplementary Material 3.

## Declarations

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aguirre-Liguori J, Luna-Sánchez J, Gasca-Pineda J, Eguiarte LE (2020) Evaluation of the minimum sampling design for population genomic and microsatellite studies: an analysis based on wild maize. *Front Genet* 11:870. <https://doi.org/10.3389/fgene.2020.00870>
- Ahmad CB, Abdullah J, Jaafar J (2013) Community perspectives on buffer zone for protected areas: a preliminary study. *Procedia Soc Behav Sci* 85:198–205. <https://doi.org/10.1016/j.sbspro.2013.08.351>
- Anderson CD, Epperson BK, Fortin M-J, Holderegger R, James PMA, Rosenberg MS et al (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Mol Ecol* 19:3565–3575. <https://doi.org/10.1111/j.1365-294X.2010.04757.x>
- Balkenhol N, Dudaniec RY, Krutovsky KV, Johnson JS, Cairns DM, Segelbacher G et al (2017) Landscape genomics: understanding relationships between environmental heterogeneity and genomic characteristics of populations. In: Rajora OM (ed) *Population genomics: concepts, approaches, and applications*. Springer, Cham, pp 261–322. [https://doi.org/10.1007/13836\\_2017\\_2](https://doi.org/10.1007/13836_2017_2)
- Balkenhol N, Schwartz MK, Inman RM, Copeland JP, Squires JS, Anderson NHJ et al (2020) Landscape genetics of wolverines (*Gulo gulo*): scale-dependent effects of bioclimatic, topographic, and anthropogenic variables. *J Mammal* 101:790–803. <https://doi.org/10.1093/jmammal/gyaa037>

- Benito-Alonso JL (2006) Vegetación del Parque Nacional de Ordesa y Monte Perdido (Sobrarbe, Pirineo central aragonés). Consejo de Protección de la Naturaleza de Aragón. [https://jolube.files.wordpress.com/2008/02/vegetacion\\_pnomp\\_benitoalonso\\_2006.pdf](https://jolube.files.wordpress.com/2008/02/vegetacion_pnomp_benitoalonso_2006.pdf). Accessed 10 June 2022
- Bolliger J, Lander T, Balkenhol N (2014) Landscape genetics since 2003: status, challenges and future directions. *Landsc Ecol* 29:361–366. <https://doi.org/10.1007/s10980-013-9982-x>
- Breton F, Maurel N, Braud Y, Lopez-Vaamonde C (2017) Étude des populations d'*Actias isabellae galliaegloria* Obth. à l'aide de la phéromone sexuelle synthétique dans le nord du Mercantour. *Alexandor* 28:143–148
- Cabalar Fuentes M, Pazos Otón M, Armas Quintá F, Macía Arce J (2011) The Natura 2000 network in Spain and its lack of protection. *Eur J Geogr* 2:50–60
- Cardil A, Merenciano D, Molina-Terrén DM (2016) Wildland fire typologies and extreme temperatures in NE Spain. *iForest* 10:9–14. <https://doi.org/10.3832/ifor1939-009>
- Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol* 24:621–631. <https://doi.org/10.1093/molbev/msl191>
- Chefaoui RM, Lobo JM (2007) Assessing the conservation status of an Iberian moth using pseudo-absences. *J Wildl Manage* 71:2507–2516. <https://doi.org/10.2193/2006-312>
- Coulon A, Cosson JF, Angibault JM, Cargnelutti B, Galan M, Morelet N et al (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Mol Ecol* 13:2841–2850. <https://doi.org/10.1111/j.1365-294X.2004.02253.x>
- Crawford LA, Desjardins S, Keyghobadi N (2011) Fine-scale genetic structure of an endangered population of the Mormon metalmark butterfly (*Apodemia mormo*) revealed using AFLPs. *Conserv Genet* 12:991–1001. <https://doi.org/10.1007/s10592-011-0202-4>
- Diniz-Filho JAF, De Campos Pires, Telles M (2002) Spatial autocorrelation analysis and the identification of operational units for conservation in continuous populations. *Conserv Biol* 16:924–935
- Diniz-Filho JA, Soares TN, Lima JS, Dobrovolski R, Landeiro VL, de Campos Telles MP et al (2013) Mantel test in population genetics. *Genet Mol Biol* 36:475–485. <https://doi.org/10.1590/S1415-47572013000400002>
- Dudley N (ed) (2008) Guidelines for applying protected area management categories. IUCN, Gland
- Durand E, Jay F, Gaggiotti OE, François O (2009) Spatial inference of admixture proportions and secondary contact zones. *Mol Biol Evol* 26:1963–1973. <https://doi.org/10.1093/molbev/msp106>
- Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Ersts PJ (2009) Geographic Distance Matrix Generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. [http://biodiversityinformatics.amnh.org/open\\_source/gdmg](http://biodiversityinformatics.amnh.org/open_source/gdmg). Accessed 10 June 2022
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10:564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Faubet P, Waples RS, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Mol Ecol* 16:1149–1166. <https://doi.org/10.1111/j.1365-294X.2007.03218.x>
- Galiano L, Martínez-Vilalta J, Lloret F (2010) Drought-induced multifactor decline of scots pine in the pyrenees and potential vegetation change by the expansion of co-occurring Oak species. *Ecosystems* 13:978–991. <https://doi.org/10.1007/s10021-010-9368-8>
- García MB, Errea P, García DG, Pizarro M (2019) Winners and losers of landscape changes over the last sixty years in one of the oldest and Southernmost National Parks of the European Alpine Region: Ordesa and Monte Perdido. *Cuadernos de investigación geográfica* 45:123–141. <https://doi.org/10.18172/cig.3711>
- Ghosh-Harihar M, An R, Athreya R, Borthakur U, Chanchani P, Chetry D et al (2019) Protected areas and biodiversity conservation in India. *Biol Conserv* 237:114–124. <https://doi.org/10.1016/j.biocon.2019.06.024>
- Götmark F, Söderlund H, Thorell M (2000) Buffer zones for forest reserves: opinions of land owners and conservation value of their forest around nature reserves in southern Sweden. *Biodivers Conserv* 9:1377–1390. <https://doi.org/10.1023/A:1008978831068>
- Gottelli D, Sillero-Zubiri C, Marino J, Funk JM, Wang J (2012) Genetic structure and patterns of gene flow among populations of the endangered Ethiopian wolf. *Anim Conserv* 16:234–247. <https://doi.org/10.1111/j.1469-1795.2012.00591.x>
- Goudet J (2005) FSTAT version 2.9.4: a program to estimate and test population genetics parameters. Updated from Goudet [1995]. <https://www2.unil.ch/popgen/softwares/fstat.htm>. Accessed 10 June 2022
- Grodzinska-Jurczak M, Cent J (2011) Expansion of nature conservation areas: problems with Natura 2000 implementation in Poland? *J Environ Manage* 47:11–27. <https://doi.org/10.1007/s00267-010-9583-2>
- Groot AT, Classen A, Inglis O, Blanco A, López J Jr, Téran Vargas A et al (2011) Genetic differentiation across North America in the generalist moth *Heliothis virescens* and the specialist *H. subflexa*. *Mol Ecol* 20:2676–2692. <https://doi.org/10.1111/j.1365-294X.2011.05129.x>
- Guillot G, Leblois R, Coulon A, Frantz AC (2009) Statistical methods in spatial genetics. *Mol Ecol* 18:4734–4756. <https://doi.org/10.1111/j.1365-294X.2009.04410.x>
- Hansen AJ (2011) Contribution of source-sink theory to protected area science. In: Liu J, Hull V, Morzillo AT, Wiens JA (eds) Sources, sinks, and sustainability across landscapes. Cambridge University Press, Cambridge, pp 339–360
- Hansen AJ, DeFries R (2007) Ecological mechanisms linking protected areas to surrounding lands. *Ecol Appl* 17:974–988. <https://doi.org/10.1890/05-1098>
- Holderegger R, Gugerli F (2012) Where do you come from, where do you go? Directional migration rates in landscape genetics. *Mol Ecol* 21:5640–5642. <https://doi.org/10.1111/mec.12032>
- Holderegger R, Wagner HH (2008) Landscape genetics. *Bioscience* 58:199–207. <https://doi.org/10.1641/B580306>
- Howes BJ, Brown JW, Gibbs HL, Herman TB, Mockford KAP, Weatherhead PJ (2009) Directional gene flow patterns in disjunct populations of the black ratsnake (*Pantheropsis obsoletus*) and the Blanding's turtle (*Emydoidea blandingii*). *Conserv Genet* 10:407–417. <https://doi.org/10.1007/s10592-008-9607-0>
- Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* 9:1322–1332. <https://doi.org/10.1111/j.1755-0998.2009.02591.x>
- Imbert C-E, Goussard F, Roques A (2012) Is the expansion of the pine processionary moth, due to global warming, impacting the endangered Spanish moon moth through an induced change in food quality? *Integr Zool* 7:147–157
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>

- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Kalinowski ST (2005a) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* 5:187–189. <https://doi.org/10.1111/j.1471-8286.2004.00845.x>
- Kalinowski ST (2005b) Do polymorphic loci require large sample sizes to estimate genetic distances? *Heredity* 94:33–36. <https://doi.org/10.1038/sj.hdy.6800548>
- Keller D, Holderegger R, van Strien MJ, Bolliger J (2015) How to make landscape genetics beneficial for conservation management? *Conserv Genet* 16:503–512. <https://doi.org/10.1007/s10592-014-0684-y>
- Kling MM, Ackerly DD (2021) Global wind patterns shape genetic differentiation, asymmetric gene flow, and genetic diversity in trees. *Proc Natl Acad Sci USA* 118:e2017317118. <https://doi.org/10.1073/pnas.2017317118>
- Klug PE, Wisely SM, With KA (2011) Population genetic structure and landscape connectivity of the Eastern Yellowbelly Racer (*Coluber constrictor flaviventris*) in the contiguous tallgrass prairie of northeastern Kansas, USA. *Landscape Ecol* 26:281–294. <https://doi.org/10.1007/s10980-010-9554-2>
- Laike L, Miller LM, Palmé A, Palm S, Kapuscinski AR, Thoreson G et al (2005) Spatial genetic structure of northern pike (*Esox lucius*) in the Baltic Sea. *Mol Ecol* 14:1955–1964. <https://doi.org/10.1111/j.1365-294X.2005.02570.x>
- Legendre P, Fortin M-H, Borcard D (2015) Should the Mantel test be used in spatial analysis? *Methods Ecol Evol* 6:1239–1247
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trend Ecol Evol* 18:189–197. [https://doi.org/10.1016/S0169-5347\(03\)00008-9](https://doi.org/10.1016/S0169-5347(03)00008-9)
- Marí-Mena N, Lopez-Vaamonde C, Naveira H, Auger-Rozenberg MA, Vila M (2016) Phylogeography of the Spanish Moon Moth *Graellsia isabellae* (Lepidoptera, Saturniidae). *BMC Evol Biol* 16:139. <https://doi.org/10.1186/s12862-016-0708-y>
- Marí-Mena N, Naveira H, Lopez-Vaamonde C, Vila M (2019) Census and contemporary effective population size of two populations of the protected Spanish Moon Moth (*Graellsia isabellae*). *Insect Conserv Divers* 12:147–160. <https://doi.org/10.1111/icad.12322>
- Marí-Mena N (2013) Conservation genetics of the protected moth *Graellsia isabellae* (Lepidoptera: Saturniidae). PhD Dissertation. University of A Coruna, Spain. <http://hdl.handle.net/2183/11510>
- Martínez de Pisón E (2010) Propuesta de ampliación del Parque Nacional de Ordesa y Monte Perdido. *Ería* 82:117–141
- Martínez-Vega J, Díaz A, Nava JM, Gallardo M, Echavarría P (2017) Assessing land use-cover changes and modelling change scenarios in two mountain Spanish National Parks. *Environments* 4:79. <https://doi.org/10.3390/environments4040079>
- Meirmans P (2012) The trouble with isolation by distance. *Mol Ecol* 21:2839–2846. <https://doi.org/10.1111/j.1365-294X.2012.05578.x>
- Meirmans PG (2014) Nonconvergence in Bayesian estimation of migration rates. *Mol Ecol Resour* 14:726–733. <https://doi.org/10.1111/1755-0998.12216>
- Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. *Mol Ecol* 24:3223–3231. <https://doi.org/10.1111/mec.13243>
- Meirmans PG (2020) GENODIVE version 3.0: easy-to-use software for the analysis of genetic data of diploids and polyploids. *Mol Ecol Resour* 20:1126–1131. <https://doi.org/10.1111/1755-0998.13145>
- Millar JG, McElfresh JS, Romero C, Vila M, Marí-Mena N, Lopez-Vaamonde C (2010) Identification of the sex pheromone of a protected species, the Spanish Moon Moth *Graellsia isabellae*. *J Chem Ecol* 36:923–932. <https://doi.org/10.1007/s10886-010-9831-1>
- Ministerio para la Transición Ecológica y el Reto Demográfico (2022a) Parque Nacional de Ordesa y Monte Perdido. <https://www.miteco.gob.es/es/red-parques-nacionales/nuestros-parques/ordesa/>. Accessed 10 June 2022
- Ministerio para la Transición Ecológica y el Reto Demográfico (2022b) Cuarto Inventario Forestal Nacional. [https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario-forestal-nacional/cuarto\\_inventario.aspx](https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario-forestal-nacional/cuarto_inventario.aspx). Accessed 10 June 2022
- Ministerio para la Transición Ecológica y el Reto Demográfico (2023) Estadística incendios forestales. <https://www.miteco.gob.es/es/biodiversidad/temas/incendios-forestales/estadisticas-incendios.aspx>. Accessed 19 Feb 2023
- Minnie L, Zalewski A, Zalewska H, Kerley GIH (2018) Spatial variation in anthropogenic mortality induces a source–sink system in a hunted mesopredator. *Oecologia* 186:939–951. <https://doi.org/10.1007/s00442-018-4072-z>
- Murria-Beltrán E (2022) Mariposas del Parque Nacional de Ordesa y Monte Perdido (Huesca). Atlas y guía de campo. Otro matiz, Zaragoza
- Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Perez MF, Franco FF, Bombonato JR, Bonatelli IAS, Khan G, Romeiro-Brito M et al (2018) Assessing population structure in the face of isolation by distance: are we neglecting the problem? *Divers Distrib* 24:1883–1889
- Pringle RM (2017) Upgrading protected areas to conserve wild biodiversity. *Nature* 546:91–99. <https://doi.org/10.1038/nature22902>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Pruisscher P, Nylín S, Gotthard K, Wheat CW (2018) Genetic variation underlying local adaptation of diapause induction along a cline in a butterfly. *Mol Ecol* 27:3613–3626. <https://doi.org/10.1111/mec.14829>
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* 67:901–904. <https://doi.org/10.1093/sysbio/syy032>
- Romo H, García-Barros E, Martín J, Ylla J, López M (2012) *Graellsia isabellae*. In: Galante E (ed) Bases ecológicas preliminares para la conservación de las especies de interés comunitario en España: Invertebrados. Ministerio de Agricultura, Alimentación y Medio Ambiente, Madrid, pp 1–53
- Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* 4:137–138. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>
- Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour* 8:103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Samaniego L, Thober S, Kumar R, Wanders N, Rakovec O, Pan M et al (2018) Anthropogenic warming exacerbates European soil moisture droughts. *Nature Clim Change* 8:421–426. <https://doi.org/10.1038/s41558-018-0138-5>
- Samarasin P, Shuter BJ, Wright SI, Rodd FH (2017) The problem of estimating recent genetic connectivity in a changing world. *Conserv Biol* 31:126–135. <https://doi.org/10.1111/cobi.12765>
- Scott JM, Davis FW, McGhie G, Wright RG, Groves C, Estes J (2001) Nature reserves: do they capture the full range of America's biological diversity? *Ecol Appl* 11:999–1007. [https://doi.org/10.1890/1051-0761\(2001\)011\[0999:NRDTC\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2001)011[0999:NRDTC]2.0.CO;2)
- Twyford AD, Wong ELY, Friedman J (2020) Multi-level patterns of genetic structure and isolation by distance in the widespread plant

- Mimulus guttatus*. Heredity 125:227–239. <https://doi.org/10.1038/s41437-020-0335-7>
- Ugelvig LV, Andersen A, Boomsma JJ, Nash DR (2012) Dispersal and gene flow in the rare, parasitic large blue butterfly *Maculinea arion*. Mol Ecol 21:3224–3236. <https://doi.org/10.1111/j.1365-294X.2012.05592.x>
- Uhl B, Wölfling M, Fiedler K (2020) Understanding small-scale insect diversity patterns inside two nature reserves: the role of local and landscape factors. Biodivers Conserv 29:2399–2418. <https://doi.org/10.1007/s10531-020-01981-z>
- Vacquie LA, Houet T, Shol TL, Roker R, Saylor KL (2015) Modeling regional land change scenarios to assess land abandonment and reforestation dynamics in the Pyrenees (France). J Mt Sci 12:905–920. <https://doi.org/10.1007/s11629-014-3405-6>
- Vicente C, Espada M, Vieira P, Mota M (2012) Pine wilt disease: a threat to European forestry. Eur J Plant Pathol 133:89–99. <https://doi.org/10.1007/s10658-011-9924-x>
- Vila M, Auger-Rozenberg MA, Goussard F, Lopez-Vaamonde C (2009) Effect of non-lethal sampling on life-history traits of the protected moth *Graellsia isabellae* (Lepidoptera: Saturniidae). Ecol Entomol 34:356–362. <https://doi.org/10.1111/j.1365-2311.2008.01084.x>
- Vila M, Marí-Mena N, Yen SH, Lopez-Vaamonde C (2010) Characterization of ten polymorphic microsatellite markers for the protected Spanish moon moth *Graellsia isabellae* (Lepidoptera: Saturniidae). Conserv Genet 11:1151–1154. <https://doi.org/10.1007/s10592-009-9905-1>
- Waples RS (2010) Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size. Mol Ecol Resour 10:785–796. <https://doi.org/10.1111/j.1755-0998.2010.02876.x>
- Wilson GAG, Rannala BB (2003) Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191. <https://doi.org/10.1093/genetics/163.3.1177>
- Ylla i Ullastre J, (1997) Història natural del lepidòpter *Graellsia isabellae* (Graells, 1849). Institut d'Estudis Catalans, Barcelona
- Zeng L-Y, Xu L-L, Tang S-Q, Tersing T, Geng Y-P, Zhong Y (2010) Effect of sampling strategy on estimation of fine-scale spatial genetic structure in *Androsace tapete* (Primulaceae), an alpine plant endemic to Qinghai-Tibetan Plateau. J Syst Evol 48:257–264. <https://doi.org/10.1111/j.1759-6831.2010.00084.x>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.