HAL
open science

Phylogeography of the Spanish Moon Moth Graellsia isabellae (Lepidoptera, Saturniidae)

Neus Mari-Mena, Carlos Lopez-Vaamonde, Horacio Naveira, Marie-Anne Auger-Rozenberg, Marta Vila

To cite this version:

Neus Mari-Mena, Carlos Lopez-Vaamonde, Horacio Naveira, Marie-Anne Auger-Rozenberg, Marta Vila. Phylogeography of the Spanish Moon Moth Graellsia isabellae (Lepidoptera, Saturniidae). BMC Evolutionary Biology, 2016, 16 (139), 20 p. 10.1186/s12862-016-0708-y . hal-02632543

HAL Id: hal-02632543

https://hal.inrae.fr/hal-02632543v1

Submitted on 27 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Marí-Mena et al. BMC Evolutionary Biology (2016) 16:139
DOI 10.1186/s12862-016-0708-y

BMC Evolutionary Biology

RESEARCH ARTICLE Open Access

Phylogeography of the Spanish Moon Moth *Graellsia isabellae* (Lepidoptera, Saturniidae)

Neus Marí-Mena[1,2], Carlos Lopez-Vaamonde[3,4], Horacio Naveira[1], Marie-Anne Auger-Rozenberg[3] and Marta Vila[1*]

Abstract

Background: Geographic and demographic factors as well as specialisation to a new host-plant may lead to host-associated differentiation in plant-feeding insects. We explored the phylogeography of a protected moth, *Graellsia isabellae*, and its two recognised host-plant species (*Pinus sylvestris* and *P. nigra*) in order to seek for any concordance useful to disentangle the evolutionary history of this iconic lepidopteran.

Results: DNA variation in one mitochondrial marker and nine nuclear microsatellite loci revealed a strong phylogeographic pattern across 28 populations of *G. isabellae* studied in Spain and France comprising six groups mostly distributed along different mountain ranges. Reanalysis of a previously published chloroplast microsatellite dataset revealed a three and two-group structure for Spanish *P. sylvestris* and *P. nigra*, respectively. Overall, the population groupings of this protected moth did not match the ones of *P. sylvestris* and *P. nigra*.

Conclusions: There was no evidence of host-associated differentiation between populations using *P. sylvestris* and the ones inhabiting *P. nigra*. The two major mitochondrial clades of *G. isabellae* likely diverged before the Last Glacial Maximum and geographically separated the species into a "southern" (Central and Southern Iberian clusters) and a "northern" lineage (Eastern Iberian, Pyrenean and French Alpine clusters). The Eastern Iberian System, where this insect uses both host-plants, harboured the highest level of genetic diversity. Such a group independently colonised the West and East parts of the Pyrenees. Our results point to a native origin for the French populations occurring in the Alps, genetically related to the Eastern Iberian and Pyrenean sites. The Central Iberian group derived from Southern Iberian ancestors. Secondary contacts were inferred between the Southern/Central Iberian populations and Eastern Iberian cluster as well as between the two Pyrenean ones. The mito-nuclear discordance observed with regard to the Eastern Iberian cluster is congruent with a secondary contact after the evolution of mito-nuclear incompatibilities in geographically isolated areas.

Keywords: *COI*, Microsatellites, cpDNA, *Pinus sylvestris*, *Pinus nigra*, Mito-nuclear discordance

Background

The diversity of plant-feeding insects is remarkable and host-associated differentiation has been advocated as a relevant trigger of such macroevolutionary diversification [1, 2]. However, not only specialisation to different hosts but genetic incompatibilities between geographically isolated populations may prevent successful hybridisation after a secondary contact [3]. At microevolutionary scale, knowledge about the relative roles played by specialisation to a given host and differentiation in separate areas is also valuable to account for the divergent lineages of oligophagous species of insects [4–6].

Here we set out to disentangle the microevolutionary history of the Spanish Moon Moth *Graellsia isabellae* (Graells 1849) (Lepidoptera, Saturniidae), a protected species in France and Spain after its inclusion in the Habitats Directive [7, 8]. This univoltine insect is mainly distributed in the mountains of the eastern half of the Iberian Peninsula and the French Alps (Fig. 1d) and it develops on two cold-adapted host-plants: the Scots (*Pinus sylvestris*) and Black (*P. nigra*) pines. *P. sylvestris* is the host species of *G. isabellae* in the Central Iberian System, Pyrenees and Alps, where the presence of *P. nigra* is scarcer. This spectacular silkmoth inhabits forests of both *P. sylvestris* and *P. nigra* in the Iberian System [9–11]. Although it is widely accepted

* Correspondence: mvila@udc.es
[1]Department of Molecular and Cell Biology, Evolutionary Biology Group (GIBE), Universidade da Coruña, A Fraga 10, E-15008 A Coruña, Spain
Full list of author information is available at the end of the article

© 2016 The Author(s). **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

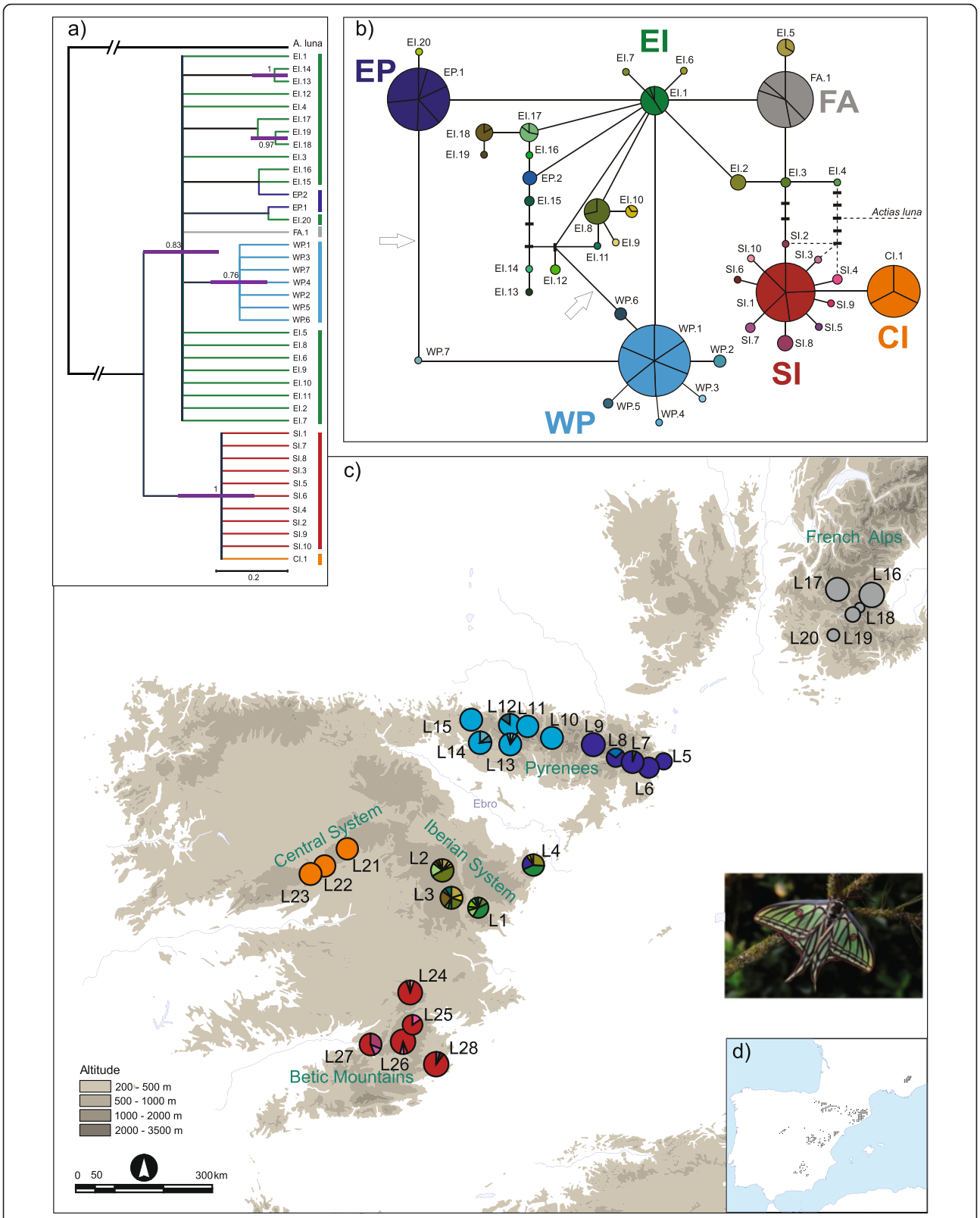

Fig. 1 (See legend on next page.)

(See figure on previous page.)
Fig. 1 Evolutionary relationships and geographical distribution of mtDNA *COI* haplotypes of *Graellsia isabellae*. The six clusters are displayed in different colours. Tones of dark and light blue refer to haplotypes from WP and EP, whereas tones of green and red correspond to EI and SI. **a** Bayesian phylogenetic tree. Violet bars at supported nodes indicate the temporal estimates (95 % HPD intervals); time scale indicates 0.2 Mya; numbers above branches represent posterior probabilities higher than 0.7. **b** 95 % Statistical Parsimony network, intraspecific connection limit = 12. *Actias luna* was equally connected to haplotypes EI.4, SI.2, SI, 3 and SI.4 by 66 mutational steps. Circle size is proportional to haplotype frequency. Names besides circles are haplotype codes. Solid lines connecting haplotypes represent a single mutational event, regardless of their length. Black rectangles represent missing or theoretical haplotypes. Arrows point to the most likely break of a loop according to coalescent predictions. **c** Geographical distribution of the 41 mitochondrial haplotypes in 28 populations across its entire distribution range. Population codes as labelled in Table 1. Inset: Adult male of *G. isabellae*. **d** Approximate geographic distribution of *G. isabellae* (redrawn from [67, 99])

that *G. isabellae* feeds exclusively on *P. nigra* in Southern Spain [12], the host-plant of the recently discovered population of Sierra-María (*L28*, Fig. 1), flying in a mixed forest of the thermophilous *P. halepensis/P. pinaster*, remains to be confirmed [13]. Our aims are to:

(1) Assess the levels of genetic variability and population structure of *Graellsia isabellae* across its entire distribution range.
(2) Evaluate whether host-associated differentiation and divergence of geographically isolated populations played a relevant role in the phylogeography of this iconic insect.

Pinus species are known for their variation in deterrent compounds (e.g. terpenoids) [14]. Thus, it is reasonable to expect populations of *G. isabellae* to have experienced host-associated differentiation. If so, congruence between the population structure of *G. isabellae* and its host plant use (*P. sylvestris*, *P. nigra* and *P. pinaster/P. halepensis*) should be found.

The evolutionary history of *G. isabellae* is expected to depend to some extent on the range shifts suffered by *P. sylvestris* and *P. nigra* during the Upper Quaternary [15]. Several Last Glacial Maximum (LGM) refugia have been inferred for *P. sylvestris* and *P. nigra* in the Iberian Peninsula [16, 17]. Both species expanded during the Holocene, especially *P. sylvestris* [18, 19]. However, the potentially suitable area for *G. isabellae* in Spain is about three times larger than the one it presently occupies [10]. Therefore, another plausible expectation is that the current degree of population structuring of *G. isabellae* was already present at the LGM, corresponding to the separate glacial refugia of *P. sylvestris*/*P. nigra*.

Methods

To achieve aim 1, we Sanger sequenced the second half of the cytochrome c oxidase subunit I gene (*COI*), a mitochondrial marker widely used in insect population genetics, and genotyped nine microsatellite markers specifically developed for *Graellsia isabellae* [20]. To attain our second aim, we used the genetic structure of *Pinus sylvestris* and *P. nigra* as a background. For this, we revised the available organelle and nuclear evidence for both pines at European level. We also surveyed the structure of their Iberian populations by reanalysing the chloroplast (cpDNA) microsatellite dataset published by Soto et al. [21].

Non-lethal sampling and DNA extraction

Between 2007 and 2009, 796 adult males of *G. isabellae* were collected from 28 locations, covering the whole distribution range (Table 1, Fig. 1c). Specimens were hand-netted after being attracted by either captive virgin females or by synthetic female pheromone placed near a light trap [22]. Tissue sampling was performed by clipping a ~ 130 mm^2 fragment of the right hind-wing tail [23]. Genomic DNA was extracted using a commercial kit (High Pure PCR Template Preparation Kit, Roche) following the manufacturer's instructions.

Mitochondrial DNA

We sequenced the 3′ end (832 bp) of the mitochondrial cytochrome c oxidase subunit I (*COI*) gene for 493 of the 796 sampled individuals. Amplifications were carried out in 30 μL volumes containing 1X PCR Buffer (5 PRIME), 1.5 mM $MgCl_2$, 1U *Taq*DNA Polymerase (5 PRIME), 0.2 mM of each dNTP, 0.2 μM of each primer C1-J-1751 and C2-N-3661 [24] and 30 ng of DNA. Reaction conditions consisted of 2 min 95 °C followed by 30 cycles of 1 min 94 °C; 1 min 57 °C; 1.5 min 68 °C and, lastly, a final extension of (7 min 68 °C). Amplified fragments were purified and bidirectionally sequenced at DNA Sequencing Service (Macrogen, Korea) using the following primers: ISAF (5′-GGTGACCCAATTCTTTACCAAC-3′, present work, positions 12,811–13,642 of the *Bombyx mandarina* mitochondrion genome), LepLEUr [25]. Inspection of electropherograms and alignmentents were performed in CODONCODES 3.7.1.1 (CodonCode, USA).

Microsatellite genotyping

The 796 samples were genotyped at ten microsatellite loci previously developed [20]. 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 3730*xl* DNA Analyzer (ABI). Allelic peaks were independently called by two researchers, using GENEMAPPER 4.0

Table 1 Genetic diversity within samples of *Graellsia isabellae* as from one mitochondrial and nine microsatellite markers

				mtDNA						Microsatellites									
Locality	Mountain range	Coordinates	Code	*N*	*S*	*Nh*	*h*	*π*	*k*	*N*	*NA*	A_E	*AR*	*APR*	H_o	F_{IS}	H_e	HWE	$H_{MD\text{-}SMM}$
Ademuz	Iberian	40°04′17.4″N 1°06′19.8″W	L1	17	10	8	0.824 ± 0.082	0.0028 ± 0.0005	2.324	32	84	6.0	5.2	0.36	0.655	**0.104**	0.731	***	NS
Bronchales	Iberian	40°31′19.5″N 1°38′56.1″W	L2	21	9	8	0.752 ± 0.086	0.0022 ± 0.0004	1.867	31	85	5.4	5.36	0.39	0.675	**0.106**	0.755	NS	NS
Huerta	Iberian	40°12′15.5″N 1°41′31.1″W	L3	20	8	8	0.874 ± 0.041	0.0032 ± 0.0003	2.637	32	93	5.8	5.47	0.36	0.672	**0.116**	0.760	***	NS
Els Ports	Iberian	40°47′29.9″N 0°18′45.8″E	L4	19	4	5	0.743 ± 0.064	0.0012 ± 0.0002	1.029	41	89	5.5	5.29	0.35	0.703	0.061	0.749	NS	NS
Albanyà	East Pyrenees	42°18′33.4″N 2°42′14.8″E	L5	11	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	11	28	2.3	2.78	0.03	0.444	0.075	0.481	NS	***
Castellfolit	East Pyrenees	42°13′22.9″N 2°32′40.1″E	L6	17	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	18	27	2.2	2.66	0	0.377	0.207	0.475	***	***
Montesquiu	East Pyrenees	42°14′14.6″N 2°21′14.4″E	L7	20	1	2	0.100 ± 0.088	0.0001 ± 0.0001	0.100	57	37	2.3	2.85	0.01	0.432	**0.123**	0.493	***	*
Montgrony	East Pyrenees	42°16′11.9″N 2°05′00.5″E	L8	14	2	2	0.440 ± 0.112	0.0011 ± 0.0003	0.879	45	36	2.3	2.98	0.02	0.449	**0.139**	0.521	***	***
Baiasca	East Pyrenees	42°30′07.2″N 1°08′59.0″E	L9	22	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	25	32	2.3	2.86	0.01	0.488	−0.061	0.460	NS	**
Renanué	West Pyrenees	42°29′31.7″N 0°31′42.6″E	L10	20	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	31	37	2.7	3.23	0.01	0.504	0.074	0.544	NS	**
La Sarra	West Pyrenees	42°38′23.7″N 0°09′13.1″E	L11	19	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	23	37	2.9	3.29	0.01	0.560	−0.000	0.560	NS	*
Ordesa	West Pyrenees	42°39′16.7″N 0°04′49.6″W	L12	20	1	2	0.268 ± 0.113	0.0003 ± 0.0001	0.268	27	34	2.7	3.05	0.05	0.482	0.112	0.542	NS	**
Cabas	West Pyrenees	42°26′06.1″N 0°10′16.7″W	L13	20	3	4	0.284 ± 0.128	0.0004 ± 0.0001	0.300	32	40	2.9	3.23	0.05	0.510	0.066	0.547	NS	***
San Juan	West Pyrenees	42°30′59.9″N 0°41′26.2″W	L14	21	2	3	0.410 ± 0.121	0.0005 ± 0.0002	0.438	24	40	3.1	3.24	0.06	0.472	0.149	0.551	***	***
Belabarce	West Pyrenees	42°52′38.4″N 0°53′12.9″W	L15	20	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	31	30	2.4	2.83	0.02	0.459	0.102	0.511	***	***
Ange Gardien	French Alps	44°44′14.6″N 6°46′8.31″E	L16	25	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	40	17	1.2	1.49	0.03	0.128	0.054	0.135	NS	***

Table 1 Genetic diversity within samples of *Graellsia isabellae* as from one mitochondrial and nine microsatellite markers *(Continued)*

Fournel	French Alps	44°45′55.3″N 6°31′51.4″E	L17	22	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	36	15	1.2	1.48	0	0.133	0.151	0.156	NS	***
Cristillan	French Alps	44°40′29.0″N 6°42′02.2″E	L18	4	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	12	13	1.2	1.38	0	0.120	0.047	0.126	NS	***
Guillestre	French Alps	44°42′19.8″N 6°43′24.2″E	L19	9	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	11	15	1.4	1.59	0	0.111	0.415	0.190	NS	***
Auzet	French Alps	44°16′55.6″N 6°17′52.6″E	L20	6	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	6	13	1.3	1.44	0	0.241	−0.413	0.170	NS	***
Rascafría	Central Iberian	40°51′24.7″N 3°53′31.6″W	L21	19	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	23	22	1.9	2.21	0.11	0.353	0.114	0.398	NS	**
Cercedilla	Central Iberian	40°45′28.5″N 4°04′17.7″W	L22	19	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	25	27	1.9	2.35	0	0.369	0.077	0.390	NS	*
Peguerinos	Central Iberian	40°40′57.1″N 4°10′40.9″W	L23	20	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	48	25	1.8	2.25	0	0.305	**0.209**	0.386	NS	**
Río Mundo	Betic Mountains	38°27′21.3″N 2°26′18.2″W	L24	19	2	3	0.205 ± 0.119	0.0003 ± 0.0002	0.211	22	49	3.8	3.81	0.09	0.485	0.166	0.581	***	NS
Guillimona	Betic Mountains	38°03′16.5″N 2°33′12.0″W	L25	13	0	1	0.000 ± 0.000	0.0000 ± 0.0000	0	13	45	4.1	3.95	0.04	0.556	0.031	0.573	NS	NS
Sagra	Betic Mountains	37°56′05.2″N 2°35′25.5″W	L26	20	1	3	0.195 ± 0.115	0.0002 ± 0.0002	0.200	35	62	4.0	4.12	0.1	0.545	0.080	0.592	NS	*
Cazorla	Betic Mountains	37°54′07.0″N 2°56′18.4″W	L27	16	2	3	0.608 ± 0.090	0.0008 ± 0.0002	0.692	34	58	4.0	4.05	0.16	0.542	0.086	0.592	NS	NS
Sierra María	Betic Mountains	37°44′01.9″N 2°01′18.7″W	L28	20	2	3	0.195 ± 0.115	0.0002 ± 0.0002	0.200	31	38	2.3	2.98	0.35	0.433	0.105	0.484	***	**
Global				493	33	41	0.868 ± 0.006	0.0044 ± 0.0001	3.650	796									

mtDNA: *N* sample size, *S* segregating sites, *Nh* number of haplotypes, *h* haplotype diversity, π nucleotide diversity, *k* average number of nucleotide differences

Microsatellites: *N* sample size, *NA* number of alleles per populations, *A* mean of alleles per population, *AR* allelic richness, *APR* private allelic richness, H_o observed heterozygosity, H_e expected heterozygosity, *HWE* deviations from Hardy-Weinberg equilibrium, *** significant after sequential Bonferroni correction ($p < 0.0017$), F_{IS} = inbreeding coefficient (significant values, $p < 0.0002$, are shown in bold), H_{MD} = expected heterozygosity at mutation-drift equilibrium under the SMM model: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

(ABI). Results from locus *GI03* were not included in the present study due to its complex allelic pattern.

Genetic diversity and population structure

Nucleotide and haplotype diversities were analysed for the mitochondrial *COI* dataset with DNASP 5.1 [26]. Nuclear gene diversity and F_{IS} values for each sampling location were calculated with FSTAT 2.9.3.2 [27]. Allelic richness and allelic private richness were obtained using HP-RARE [28]. Analyses of Hardy-Weinberg equilibrium (HWE) as well as tests for gametic phase disequilibrium were computed using the web version of GENEPOP 4.2 [29]. Genetic differentiation between populations was estimated using pairwise Φ_{ST} and F_{ST} as implemented in ARLEQUIN 3.5.1.2 [30].

We explored global genetic structure using both the mitochondrial and nuclear datasets. We applied the Spatial Analysis of Molecular Variance (SAMOVA 1.0) [31] to the former by using 100 simulated annealing processes for K values (groups of populations) from two to 10. Each run was repeated three times to check for consistency. In addition, we carried out population mixture analyses using the spatial clustering modules implemented for DNA sequence data in BAPS 6.0 [32, 33], both for individuals and groups of individuals (sampling localities). We ran the program ten times each, using $K = 10$, $K = 15$ and $K = 20$ as upper bounds to the number of populations. In each repetition the program explores the space of partitions and informs of the optimal one (highest log marginal likelihood value).

Two Bayesian clustering algorithms were used to infer population structure from the microsatellite dataset: STRUCTURE 2.3.4 [34] and BAPS 6.0. The former was run under the admixture model with correlated allele frequencies between populations [35]. We set a burn-in of 100,000 iterations followed by 500,000 iterations for parameter estimation. Each simulation was run 20 times, exploring values for K ranging from one to 29. We chose the best partition of the data by examining both the log probability of the data ($\ln \Pr(X|K)$) and the ΔK statistic, following Evanno et al. [36]. CLUMPP 1.1.2 [37] was used to permute the admixture coefficients for the several independent runs resulting for the chosen K-value. Finally, DISTRUCT 1.1 [38] was employed to visualize the output from CLUMPP. Multilocus genotypes were clustered using BAPS by grouping individuals (genetic mixture analysis) under a non-spatial model. We set 10 repetitions of the algorithm for each K ranging between 1 and 28. The best partition (optimal K) was identified as explained above.

In addition, Cavalli-Sforza and Edward's chord distances (D_C) were obtained with the aid of MICROSAT 1.5 [39] which were used to build Neighbor-joining (NJ) trees with NEIGHBOUR. Confidence in tree topology was assessed by bootstrapping over loci (10,000 samples), and resulting trees were summarized by CONSENSUS. Both programs are implemented in the PHYLIP 3.68 package [40].

Phylogenetic analyses of mitochondrial sequences

A 95 % Statistical Parsimony (SP) network was calculated for the *COI* haplotypes using TCS 1.2 [41]. Preliminary analyses using the homologous sequence from related species *Actias luna*, *A. selene* and/or *Argema mittrei* (KX302402, JX186589 and KX348010, respectively), revealed *Actias luna* (Linnaeus 1758) as an appropriate outgroup. In order to identify the earliest diverging haplotype/s, maximum connection limit was set to 70. Mitochondrial phylogenetic trees were reconstructed using Bayesian inference, with the aid of MRBAYES 3.2 [42], specifying a partitioned analysis by separating 1st and 3rd codon positions into different partitions, and eliminating the 2nd position of the *COI* alignment due to its lack of phylogenetic information (it only shows a single private change). We used a gamma model of rate variation across sites and sampled across the GTR model space in the Bayesian Markov chain Monte Carlo (MCMC) analysis itself [43], running two simultaneous, independent runs during 10^6 generations, with trees sampled every 10^3 generations. The analysis was carried out with default priors until the standard deviation of split frequencies dropped below 0.01, and the potential scale reduction factor for all parameters lay close to 1.0. Two completely independent analyses starting from different random trees were run. For the MCMC sampling of the target distribution, three heated chains and one cold chain were used. The first 25 % samples of the cold chain were discarded as burn-in. We used Bayes factor comparisons to test several topological hypotheses. Marginal model likelihoods were estimated by the stepping-stone method; strength of the evidence in favour of the better model was then assessed by the magnitude of the log-difference, following Kass & Raftery [44]. The strict clock model was tested against the non-clock model in this way, and once determined that our dataset evolved in a clock-like manner, we obtained a calibrated tree using the rate dating approach, based on the divergence rate of 3.54 % My^{-1} estimated for the *COI* gene in tenebrionid beetles [45], essentially due to changes in 1st or 3rd coding positions. The output cladogram summarizing the trees was visualized with FIGTREE 1.4.2 [46], retaining branch length information and expressing nodal support as posterior probabilities. Mean ages and 95 % highest posterior density (HPD) intervals of mtDNA phylogroups are used as estimates of divergence times.

Historical demography

We investigated the effects of fluctuations in population size on the genetic variation of the population clusters defined in this work. On the one hand, we used the *COI*

dataset to plot the observed frequency distribution of pairwise nucleotide differences with their expected distribution by performing a mismatch analysis. We estimated the demographic parameters θ_0, θ_1 and τ from the mismatch distribution using MISMATCHHETMUTRATE [47], based on a finite-sites model with heterogeneity of mutation rates; confidence intervals around the estimated parameters were obtained by bootstrapping (1000 samples). Goodness of fit to the expected mismatch distribution under different theoretical demographic scenarios was tested using the sum of squared deviations (SSD), by comparing observed values with those yielded by 1000 simulated populations. To obtain an approximate time for the putative expansion events, we followed [48], so that $t = \tau/2u$; $u = \mu Lg$, where μ is the estimated mutation rate (0.0177 substitutions/site/my), L is the length of the sequence (832 bp), and g is the generation time (1 year). In addition, we used Fu's Fs and R_2 tests, the more powerful ones for detecting departures from the null hypothesis of constant population size [49, 50]. Confidence intervals and P-values of these tests were obtained by Monte Carlo simulations (10,000 pseudo-replicates) carried out with DNASP, based on the neutral coalescent process (random-mating population of constant size, with all mutations selectively neutral and occurring at sites that have not previously mutated) and assuming no recombination. Coalescent simulations for the R_2 statistic were conditioned on the number of segregating sites.

As a separate approach, BOTTLENECK 1.2.02 [51] was used to detect reductions in population size using our microsatellite dataset as follows. Expected heterozygosities at mutation-drift equilibrium (h_{MD}), obtained from the observed number of alleles at each locus through simulating the coalescent process for n loci, are compared to heterozygosities at HWE (h_{HW}). During a bottleneck the number of alleles decreases faster than gene diversity, and therefore h_{HW} should be higher than h_{MD}. The analyses were performed across 10,000 iterations under both the infinite allele model (IAM) and the stepwise mutation model (SMM), which respectively correspond to the least and most stringent conditions to detect a bottleneck. We assessed the significance of the results using the Wilcoxon's signed ranks test, implemented in the program. We also investigated the distribution of allele frequencies, checking for the expected distortions (mode shifts) produced by recent bottlenecks [52].

Phylogeography of *Pinus sylvestris* and *P. nigra* as from published genetic data

With regard to *Pinus sylvestris*, we inspected the mitochondrial results published by Sinclair et al. ([53] haplotypes obtained by RFLP analysis of *COI*), Soranzo et al. [54] and Cheddadi et al. [16], both of whom analysed one intron (*nad1*, exon B/C). Then we surveyed the dataset from Naydenov et al. ([55] haplotypes revealed by the combination of the *nad7* intron 1 and the *nad1* intron B/C) and Pyhäjärvi et al. [56], whose mitochondrial haplotypes were combinations of the ones published by Soranzo et al. [54]. Then, we revised the interpretation of the results about population structure of Iberian Scots pine by Robledo-Arnuncio et al. [57] using chloroplast SSR and Prus-Glowacki et al. [58] based on isoenzyme data.

To the best of our knowledge, the only mitochondrial data from *Pinus nigra* related to forests of South Spain and Morocco [59]. In addition, we also revised the interpretation of the chloroplast SSR results by Afzal-Rafii & Dodd [17] and the genomic ISSR data by Rubio-Moraga et al. [60].

Genetic structure of *Pinus sylvestris* and *P. nigra* as from cpDNA

We reanalysed the cpDNA SSR dataset published by Soto et al. [21] for Iberian *P. sylvestris* and *P. nigra*. The former consisted of 706 individuals, sampled in 30 natural forests and genotyped for six loci. We excluded locus *PT1254* from our analyses as it contained 20.25 % of missing data. The percentage of missing data for the other five loci ranged between 0.56 (*PT30204*) and 2.12 (*PT1936*). The available cpDNA SSR dataset for *P. nigra* consisted of 326 individuals, sampled in 14 natural forests and genotyped for the same six markers. In this case, all loci were used, as missing data ranged from 0.92 % (*PT30204*) and 3.37 % (*PT1254*). Preliminary runs revealed that the use of missing data would produce slight differences in the clustering pattern of *P. nigra*. Therefore, 30 individuals of Black pine (1–6 individuals per sampling site, 11 localities affected) had to be discarded from the final analysis.

The overall genetic structure of both conifers was assessed using the Bayesian clustering implemented in BAPS 6.0. We applied a non-spatial genetic mixture analysis to groups (sampled localities) of haplotypes and the "linked loci" option. Again, we performed ten repetitions of the algorithm for each K ranging between 1 and the number of sampled sites (*P. sylvestris* = 30, *P. nigra* = 14). The best partition (optimal K) was identified as explained above.

The population structure of *Pinus nigra* in the South of Spain and Morocco inferred by Jaramillo-Correa et al. [59] was based on the same six markers we reanalysed from Soto et al. [21]. However, the population structure of *P. sylvestris* in the Northern Meseta (NW spain) surveyed by Robledo-Arnuncio et al. [57] relied on the five loci we reanalysed for *P. sylvestris* plus a sixth one, not included in Soto et al. [21]. So, concerning *P. sylvestris*, the only difference in terms of molecular markers is the exclusion of locus *PT1254* in our dataset and the information provided by locus *PT26081* in Robledo-Arnuncio et al. [57]. Lastly, it should be noted that 12 out of the

14 Spanish sites sampled by Prus-Glowacki et al. [58] and all (13) localities surveyed by Robledo-Arnuncio et al. [57] are included in the dataset by Soto et al. [21] we reanalysed.

Typing for three *Wolbachia* genes

Incompatibility between males from infected populations and females from uninfected ones may lead to asymmetrical gene flow, a phenomenon reported in moths [61]. The mitochondrial and nuclear discordant levels of differentiation between the Pyrenean sites *L9* and *L10* (see Results) might have been caused by reproductive parasites such as *Wolbachia*. We tested for presence of this bacterium in a set of 53 specimens of *Graellsia* following Kodandaramaiah et al. [62] (Additional file 1).

Results

Genetic diversity and population structure

Thirty-three polymorphic sites along the 832 bp alignment defined 41 mitochondrial haplotypes with a strong phylogeographic structure (Table 1). There were 15 parsimony informative sites (14 of them with two variants, whereas position 555 of the alignment showed three variants). Twelve out of the 34 detected mutations caused amino acid replacement (Additional file 2).

Haplotypes were not shared with localities from other mountain ranges, with the sole exception of variants EI.1 (mostly distributed in the Iberian System) and EP.1 (mainly occurring at the Eastern Pyrenees). The first one was shared by three localities of the Iberian System (*L1, L3* and *L4*) and Montesquiu (*L7*), whereas EP.1 was present at Els-Port (*L4*) and all the Eastern Pyrenean sites (*L5-L9*) (Additional file 3). The four sites sampled in the Iberian System (*L1-L4*) showed the highest haplotype and nucleotide diversity (Table 1).

Mitochondrial data revealed six groups of populations of *G. isabellae* (SAMOVA, Fig. 2a), which, with one exception (Pyrenees), showed a one to one correspondence with each of the main mountain ranges inhabited by the species: Iberian System (EI: Eastern Iberia, *L1, L2, L3, L4*, in green), Eastern Pyrenees (EP: *L5, L6, L7, L8, L9*, in dark blue), Western Pyrenees (WP: *L10, L11, L12, L13, L14, L15*, in light blue), French Alps (FA: *L16, L17, L18, L19, L20*, in grey), Central Iberian System (CI: Central Iberia, *L21, L22, L23*, in orange), and Betic Mountains (SI: Southern Iberia, *L24, L25, L26, L27, L28*, in red) (Fig. 1c). The spatial clustering of sampling localities calculated in BAPS produced exactly that same optimum partition in all the repetitions of the analysis. At the individual level, though, the most frequently favoured partition (25 out of 30 repetitions) consisted of just five mitochondrial clusters: CI-SI, WP, EP, FA and the EI core. This individual-based analysis assigned some samples from EI to other clusters: most of the specimens from Els-Ports (*L4*) to cluster EP and those specimens from Bronchales and Huerta (*L2* and *L3*) bearing haplotype EI.5 to cluster FA.

Overall, the number of alleles per nuclear marker ranged between six (locus *GI26*) and 37 (*GI11*). Loci *GI18* and *GI23* showed significant departures from HWE in more than two localities. This was likely due to segregating null alleles, as supported by the geographic distribution of these departures, i.e. locus *GI18* deviated from HWE at three sites (*L1-L3*) from the Iberian System and two (*L24, L26*) from the Betic Mountains, whereas HWE was rejected for *GI23* at three localities from the Eastern Pyrenees (*L5-L8*) (Additional file 4). Bayesian clustering and bottleneck analyses were unaffected by their exclusion, so we report nuclear results using all nine microsatellites. No linkage disequilibrium was observed for any pair of loci after Bonferroni correction.

The four sites sampled at the Iberian System (*L1-L4*) showed the highest levels of gene diversity ($H_e = 0.731–0.760$). When it comes to private allelic richness, the southern site of Sierra-María (*L28*) showed a similarly high level to the Iberian System localities (0.35–0.39) (Table 1). Seven of the 28 populations were significantly differentiated from all the others as from pairwise F_{ST} distances. Five of them are in the Pyrenees (*L9*, *L10*, *L11*, *L14* and *L15*), one in the Central Iberian System (*L21*) and the last one in the Betic Mountains (*L28*). The Western Pyrenean sites (*L10-L15*) revealed a particular substructure, as all pairwise comparisons but one (*L12-L13*) were significant (Additional file 5).

The most likely partition of the microsatellite dataset was six groups (EI, EP, WP, FA, SI and CI), as from STRUCTURE (Fig. 3a, upper panel, Additional file 6). A second partition ($K = 2$) was supported only by Evanno's method: the first cluster joined the French Alps, Eastern Pyrenees and Western Pyrenees (FA-EP-WP), whereas the second one grouped Eastern Iberia, Central Iberia, and Southern Iberia (EI-CI-SI) (Fig. 3a, lower panel). However, the overall population structure characterised by BAPS yielded $K = 19$ as the best partition (Fig. 3c). Three of these clusters were formed by a single individual (sampled at Ademuz *L1*, Bronchales *L2* and La Sarra *L11*, respectively). Cluster and sampling locality only matched at Rascafría *L21* and Sierra-María *L28*.

Phylogenetic analyses

The 33 polymorphic sites observed along the *COI* alignment produced a rather shallow haplotype phylogeny, with many polytomies. Both in the phylogenetic trees (Fig. 1a) and the statistical parsimony network (Fig. 1b) of the *COI* data, the EI cluster joined to FA, EP and WP. These two haplogroups (EI-FA-EP-WP and CI-SI) differ by three mutational steps. Although the first group was not as well supported in the Bayesian trees (posterior probability = 0.83) as the second one (posterior probability = 1), the

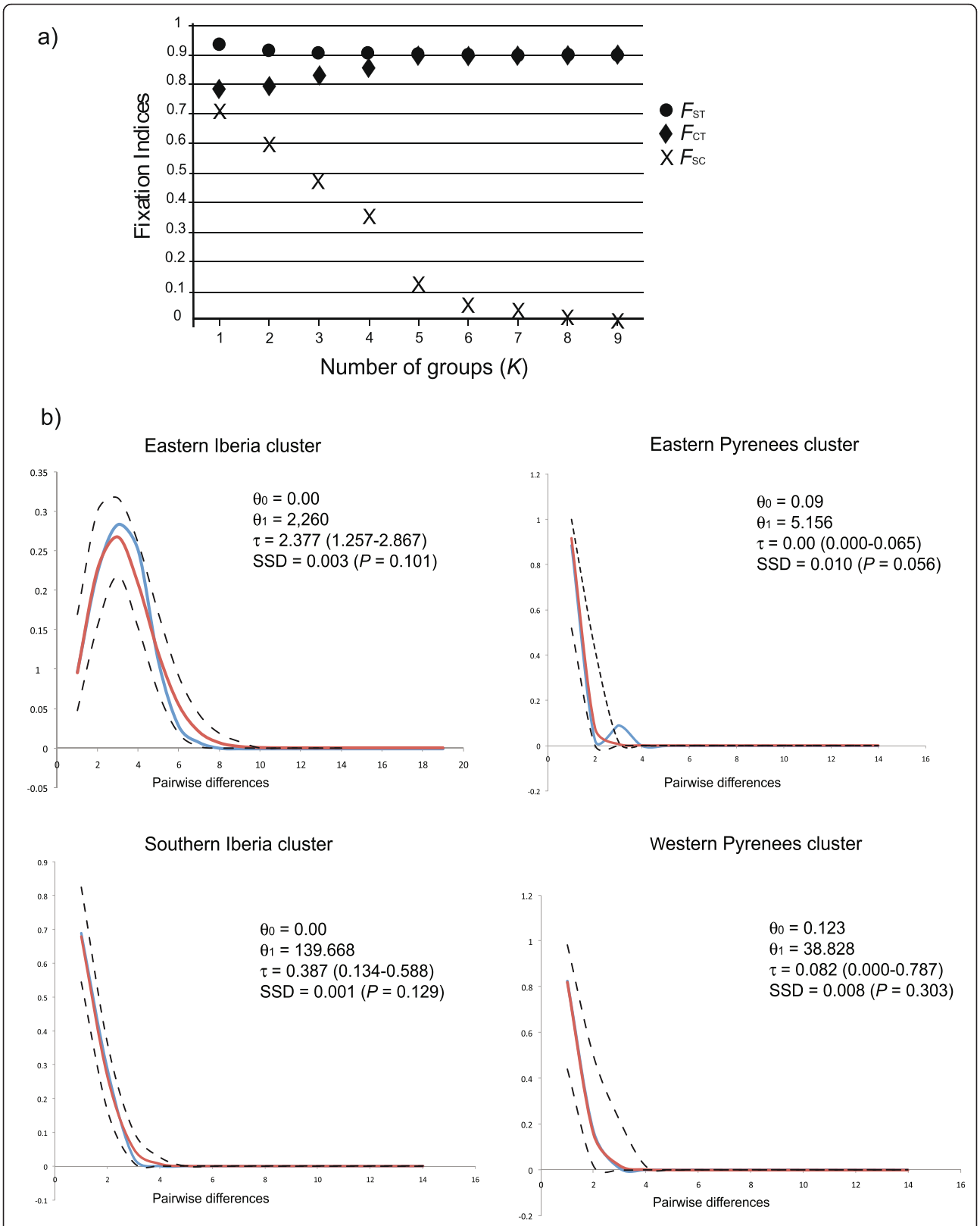

Fig. 2 (See legend on next page.)

(See figure on previous page.)
Fig. 2 Population structure and demography of *G. isabellae* as from mitochondrial data. **a** Values of fixation indices (F_{CT} = among groups differentiation; F_{SC} = among populations within groups differentiation; F_{ST} = total differentiation among populations) obtained by SAMOVA from a predefined number of groups (*K*) ranging from 1 to 9. **b** Mismatch distributions of polymorphic clusters resulting from STRUCTURE. Model of sudden expansion fit to data. Blue lines show the observed distribution of pairwise nucleotide differences. Red lines were obtained by fitting θ_0, θ_1 and τ by using the method of nonlinear least squares. Black dashed lines are the empirical 95 % confidence intervals for the mismatch distribution

monophyletic relationship between them was strongly preferred by the Bayes factor test (log difference = 21 log units). The clock-wise evolution of these sequences was also very well supported by the corresponding test (log difference of 35 log units in favour of the strict clock against the non-clock model). The split between these groups would have taken place ca. 400 kya (95 % HPD interval: 270-580 kya), with the most recent common ancestor (MRCA) of EI-FA-EP-WP dated at 300 kya (95 % HPD: 195–405) and that of CI-SI at 190 kya (95 % HPD: 95–310).

Clusters CI, SI, FA, EP and WP generally showed either a single haplotype or a highly predominant one with others arising from it by a single mutational step (Fig. 1b). The only exception to this rule was haplotype EP.2, found at the Eastern Pyrenean locality of Montgrony (*L8*), phylogenetically closer to the haplotypes of the Iberian System (EI cluster, Fig. 1b). The observations of either a single haplotype (CI, FA) or a star-like phylogeny (EP, WP, SI) stand in sharp contrast with the rich diversity and highly reticulated pattern displayed by the haplotypes of the EI cluster. Indeed, the only two terminal variants "geographically misplaced" were part of the EI cluster: EI.20 was found at the northernmost locality of the Eastern Iberian System (*L4*) but differed by a single mutational step from the dominant haplotype at the Eastern Pyrenees (EP.1), whereas EI.5 was sampled at *L2* and *L3* and also differed from the French haplotype (FA.1) by one substitution. Two more features of the mitochondrial network should be pointed out. Firstly, the only haplotype found at the Central Iberian System (CI.1) derives directly from the dominant haplotype at the Betic Mountains (SI.1), in agreement with their joining in a monophyletic group by the Bayesian tree reconstruction. Secondly, the also unique haplotype of the French Alps (FA.1) derives from haplotypes of the Iberian System (cluster EI), not from those of the geographically intervening Pyrenees (clusters EP and WP). Turning again to the phylogenetic tree for corroborative evidence, we used a Bayes factor test to contrast the hypothesis that FA and EI form a group with the hypothesis that EI forms a group with the populations from the Pyrenees (EP and WP) instead, obtaining a strong preference (5 log units) for the former topology.

As from the nuclear perspective, the NJ tree based on average microsatellite chord distances (D_C) between clusters showed the highest bootstrap support (93 %) for the branch separating EI-CI-SI from FA-EP-WP. The same partition was first unveiled by STRUCTURE (if $K = 2$) and recovered by BAPS (Fig. 3) and reveals a mito-nuclear discordance with regard to cluster EI. It also presented a noteworthy support (84 % bootstrap value) for the branch joining FA with EP, which makes considerable sense on geographical terms (Fig. 3b).

Historical demography

Clusters EI, SI and WP showed genetic signatures of population expansions, as revealed by the analysis of their mismatch distributions of mitochondrial haplotypes (Fig. 2a) and departures from constant population size models (Table 2). Based on τ estimates (Fig. 2b), the time since expansion for the EI, SI, and WP clusters would be 81 kya (43-97 kya), 13 kya (4.5-20 kya) and 2.8 kya (0-27 kya), respectively. The very low (EP) or downright absent (FA, CI) mitochondrial diversity observed in the three other clusters makes this kind of analysis inapplicable to them.

Sites from all clusters but EI likely suffered population size reductions in the near past, as revealed by the comparison of present gene diversity (H_{HW}) with its expected value in mutation-drift equilibrium (H_{MD}, conditioned on the number of observed alleles) (Table 1). We conservatively interpreted the results as indicative of bottleneck if the null hypothesis was rejected both under the IAM (all significant, data not shown) and SMM (Table 1) scenarios. The distributions of allele frequencies showed the characteristic L-shape of non-bottlenecked populations in the four localities sampled at EI, as well as the Eastern Pyrenean site of Montesquiu, the Central Iberian site of Peguerinos and the Western Pyrenean locality of Sarra.

Review on the phylogeography of *Pinus sylvestris* and *P. nigra*

The distribution area of *Graellsia isabellae* somehow resembles the one of *Pinus nigra salzmannii* [15, 63]. The few molecular data available show a strong population differentiation of the Black pine, not only at European level but also between the Northern and Southern Spanish populations [17]. Indeed, mitochondrial data revealed a particular phylogeographic pattern in South Spain, as the Betic Mountains harboured a genetic break, also shared by other conifers [59]. Lastly, nuclear ISSR markers confirmed the existence of population structure within the Betic Mountains as well as the affinity of the Moroccan and Southern Spanish populations found by Jaramillo-Correa et al. [59]. In addition, ISSR markers revealed a much deeper break

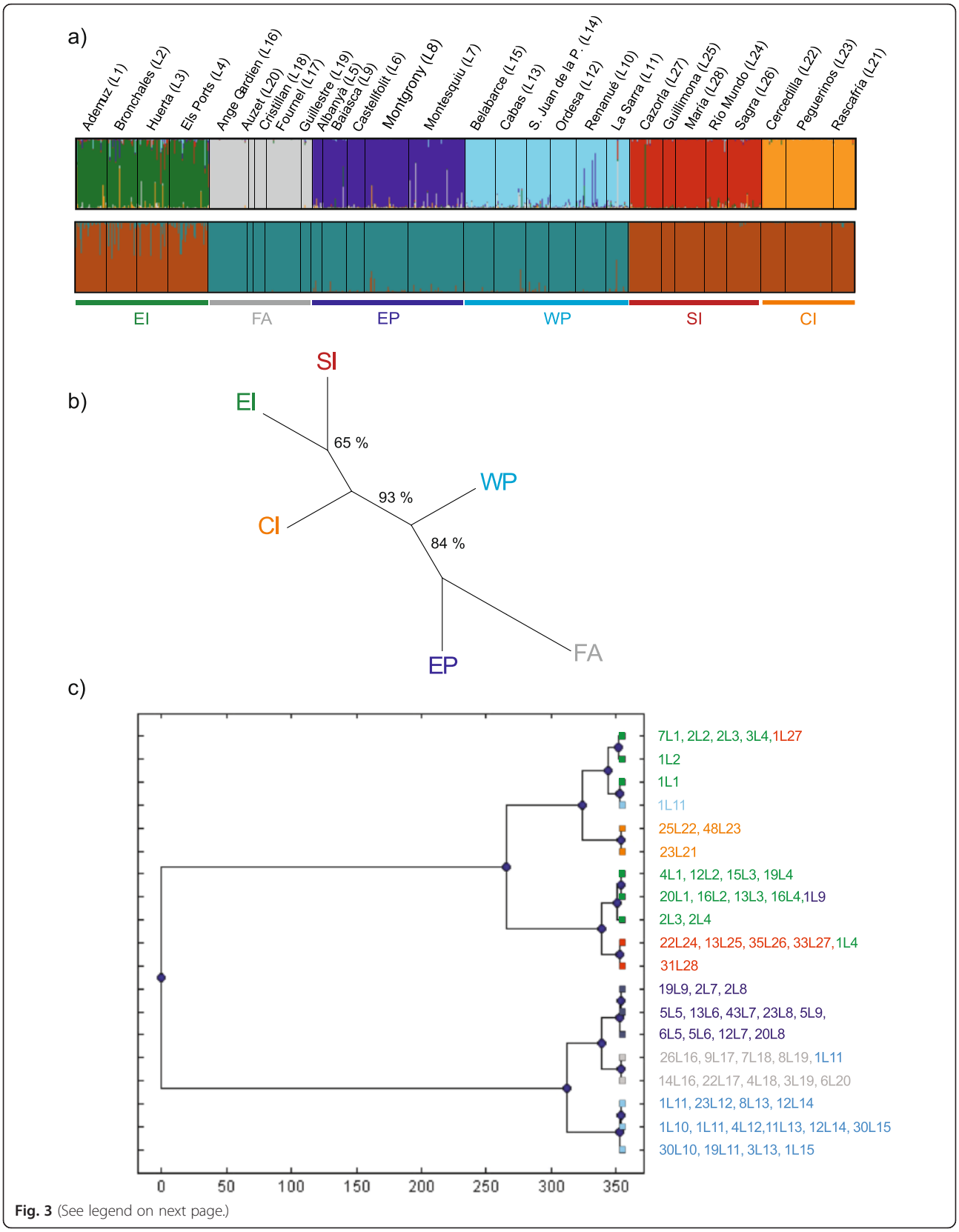

Fig. 3 (See legend on next page.)

(See figure on previous page.)
Fig. 3 Population structure and phylogeny of *G. isabellae* as from nuclear microsatellites. **a** Bayesian assignment probabilities for $K = 2$ and $K = 6$ revealed by STRUCTURE. Each vertical bar corresponds to one individual. Each background represents the proportion of membership of individuals to the inferred hypothetical groups of populations. Names above the plot represent population codes (Table 1). **b** Unrooted NJ tree of clusters based on D_C chord distance. Percentages correspond to the bootstrap support values. **c** UPGMA tree based on the estimated Kullback–Leibler divergence between the 19 homogeneous groups obtained by a non-spatial Bayesian cluster analysis implemented in BAPS. Information on the right (#LX) indicates number of individuals (#) from a given site (LX) assigned to that cluster

between these populations and the three sites from Iberian System [60]. More details available at Additional file 7.

Both chloroplast [16] and nuclear data [64–66] indicated the differentiation of Spanish *Pinus sylvestris* when compared to other European populations. Within the Iberian Peninsula, the mitochondrial evidence revealed a particular differentiation of Sierra Nevada, one of the two extant populations of the Scots pine in Southern Spain. However, the other southern locality of Baza (only 80 km apart) showed similar haplotype frequencies to the Central Iberian System [54–56]. The East–West Iberian differentiation as from haplotype frequencies was noticeable indeed [53–56]. The singularity of the Scots pine populations of the Iberian System gains additional interest because of sharing a haplotype with the Balkans [55]).

cpDNA structure of *Pinus sylvestris* and *P. nigra*

The overall population structure characterised by BAPS for Iberian *Pinus sylvestris* (five markers) yielded $K = 2$ as the best partition. The southernmost locality (Trevenque, #28, at Sierra Nevada) revealed as the first cluster, whereas the remaining 29 forests grouped together. Removal of this divergent population, led to a two-group partition with a clear West–east orientation (Fig. 4a). Therefore, we can safely infer a three-group population structure of the Spanish Scots pine surveyed by Soto et al. [21].

Concerning *Pinus nigra* (six markers) BAPS indicated again $K = 2$ as the most likely partition. The two southernmost Betic localities (Baza #II, Huelma #XII) grouped with the Central Spanish locality (Casavieja, #VI) and the Eastern Iberian forest coded as #XIV (Vistabella). All other localities clustered together. Such a two-group structure does not match the taxonomic division within *P. nigra salzmannii*, i.e. site #XIV (*pyrenaica*) grouped with the Central Iberian System and the southernmost Betic Mountains, whereas sites #IV and #X (*hispanica*) clustered with the Pyrenees and the northern Betic Mountains (Fig. 4b).

Typing for three *Wolbachia* genes

We found no evidence of infection by *Wolbachia* in any of the 53 specimens surveyed, 20 of them from *L9* and *L10* (Additional file 1).

Discussion

Genetic variability and population structure of *G. isabellae*

G. isabellae is structured into two main mitochondrial groups and six nuclear clusters. The mitochondrial "northern lineage" was formed by clusters EI, WP, EP and FA, classified as subspecies *G. isabellae isabellae, G. isabellae roncalensis, G. isabellae paradisea* and *G. isabellae galliaegloria*, respectively (revised by [67]). The EI cluster was the most diverse as measured from mitochondrial and nuclear data. A somehow similar result has been found in the Processionary Pine Moth (site 51 on *Pinus nigra*, [68]). Conversely, FA was the least variable. The mitochondrial "southern lineage" of *G. isabellae* was formed by clusters SI and CI, classified as subspecies *G. i. ceballosi* and *G. i. isabellae*, respectively (revised by [67]). The SI cluster was the second most diverse group in our dataset. Conversely, the CI cluster was mitochondrially monomorphic and its nuclear diversity was intermediate between the FA and EP groups.

The genetic pattern of the EI cluster points to the Iberian System as genetic sanctuary [69] for *G. isabellae*. This hypothesis is congruent with the basal position of haplotype EI.4 and the rich diversity and highly reticulated pattern displayed by the EI haplotypes (Fig. 1b). The ancestral character states of the EI cluster are also supported by the evolutionary history of *P. sylvestris* (Additional file 7).

The reticulated pattern of cluster EI was due to some haplotypes connected to variants from other clusters (WP, EP and FA) as well as to loops caused by ambiguous phylogenetic relationships. Firstly, we argue for incomplete lineage sorting to be the reason why variants EP.2 and EI.5 seem to be geographically "misplaced" if compared to their closest haplotypes (Fig. 1b, Additional file 3). Secondly, recombination and/or homoplasy could account for the twelve EI haplotypes that ambiguously connected to others. These loops were caused by eight third-codon synonymous substitutions (Additional file 2). Distinguishing between recombination after admixture of lineages, paternal leakage and homoplasies caused by reverse or parallel mutation will require larger sample sizes, additional markers as well as other analyses (e.g. [70, 71]. We acknowledge these mechanisms as plausible, as mitochondrial recombination has been reported in other orders of insects [72–74] and mtDNA was paternally inherited in *Antheraea* x *proylei*, an interspecific hybrid

Table 2 Signatures of population size changes on mitochondrial variation

mtDNA				
Cluster	*N*	*Nh*	Fu's *Fs*	*R2*
EI	77	21	−12.223***	0.055
EP	84	3	−1.095	0.051
WP	120	7	−8.346***	0.028**
FA	66	1	n.a.	n.a.
SI	88	10	−11.486***	0.030**
CI	58	1	n.a.	n.a.

N number of individuals, *Nh* number of haplotypes, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

from the same family and tribe, Saturniini, as *G. isabellae* [75]. Moreover, the production of new haplotypes by recombination requires heteroplasmy, a phenomenon we observed in two individuals (from France and Switzerland) not included in the present study.

We postulate that cluster WP has been isolated from EI for a longer period than the EP group. Once the coalescent predictions are applied to singletons WP.6 and WP.7 (Fig. 1b), cluster WP becomes the terminal star-like phylogeny typical of a population expansion, occurring at some point in the last 27 kya (Fig. 2b). Such a temporal framework has to be used with caution, both because of the uncertainty about the mutation rate applied and the singular substructure of *G. isabellae* in the Western Pyrenees revealed by the nuclear markers (Additional file 5). Indeed, the relatively wide curve of its mismatch distribution may have been produced not only by the time since expansion but also by the population substructure of the WP cluster [47], which was probably caused by the patchier distribution of *P. sylvestris* in that area (Additional file 8). Nevertheless, the weakly supported WP cluster (Fig. 1a) could have a pre-Holocene origin if genetic drift had shortened the coalescence time of the segregating alleles and produced the accumulation of fixed differences with other clusters. Populations of clusters WP and EP experienced population size fluctuations (Table 1), but our analyses cannot provide information about the timing or strength of those bottlenecks. Thus, we can only speculate about the independent and subsequent action of genetic drift on cluster EP as the main reason of its lower genetic diversity. We base this possibility on the severe fluctuations between arid and humid climatic stages in the Eastern Pyrenees associated with a decreasing genetic diversity from West to East in several Pyrenean plants [76].

We propose a native origin of the French populations (cluster FA) of the Spanish Moon Moth. Some authors considered them to be the result of a deliberate human introduction from Iberian individuals in the early 20th century, e.g. [77, 78]. However, all our evidence indicates that the French Alps were colonised by a small number of founders whose mitochondrial origin can be traced to the Iberian System (EI cluster), probably through an Eastern Pyrenean corridor. Indeed, the mitochondrial haplotype FA.1 was not found in the Iberian populations. FA.1 forms a loop with EI.1 (predominant in the Eastern Iberian System, but present in one individual from Montesquiu), EI.2 (found in five specimens from Els-Ports) and EI.3 (carried by two males from Ademuz). In addition, BAPS grouped the six individuals bearing the EI.5 terminal haplotype with the French specimens. Although we cannot completely rule out a deliberate human introduction from Iberian specimens, the natural foundation of the Alpine ancestral population following the expansions of *P. sylvestris* and/or *P. nigra* during the Holocene is supported by other lines of evidence. Firstly, the level of genetic diversity obtained for the FA cluster is very similar to that found in the Central Iberian System (CI), whose natural origin has never been in dispute. A similar result was obtained for the Northern Pine Processionary Moth (*Thaumetopoea pinivora*) [79]. Secondly, a Holocene spatial migration of some ancestors of the EP cluster northwards following pine expansions is plausible from a biogeographic perspective: (i) suitable habitat (*P. nigra salzmannii*) for *G. isabellae* existed in Southeastern France even before the LGM [80], (ii) the first postglacial expansion of the so-called "light-green haplotype" of *P. sylvestris* was traced (based on macrofossils) from the easternmost part of the Pyrenees northwards [16] and (iii) other montane species dependent on pine forest likely spread from the Eastern Pyrenees northward following Holocene pine expansions [81].

No host-associated differentiation

The population structure of *G. isabellae* did not support host-associated differentiation as a relevant factor preventing gene flow between populations using *Pinus sylvestris* or *P. nigra*. The indiscriminate use of the Scots and Black pine has been reported in other oligophagous plant-feeding insects [82]. Actually, local adaptations to *P. sylvestris* and *P. nigra* have been reported for the pine processionary moth *Thaumetopoea pityocampa* [83]. Females of this species may recognise the volatiles emitted by different host pine species and exhibit oviposition preference [84], but nevertheless host-associated mitochondrial differentiation was absent [68]. Absence of host-associated (*P. nigra* vs *P. sylvestris*) differentiation has also been reported in *T. pinivora* [85]. Actually, the recent discovery in South Spain of a population of this defoliator that uses *P. nigra* instead of its usual host *P. sylvestris* [85] strikingly resembles the case of *G. isabellae* at Sierra-María, likely using *P. pinaster/P. halepensis* [13]. The distance from Sierra-María to the closest patches of *P. nigra* (28 km North, 36 km West, [63]) makes the attraction and sampling of moths from the few stands of Black pine existing in that area unlikely. If *P. halepensis*

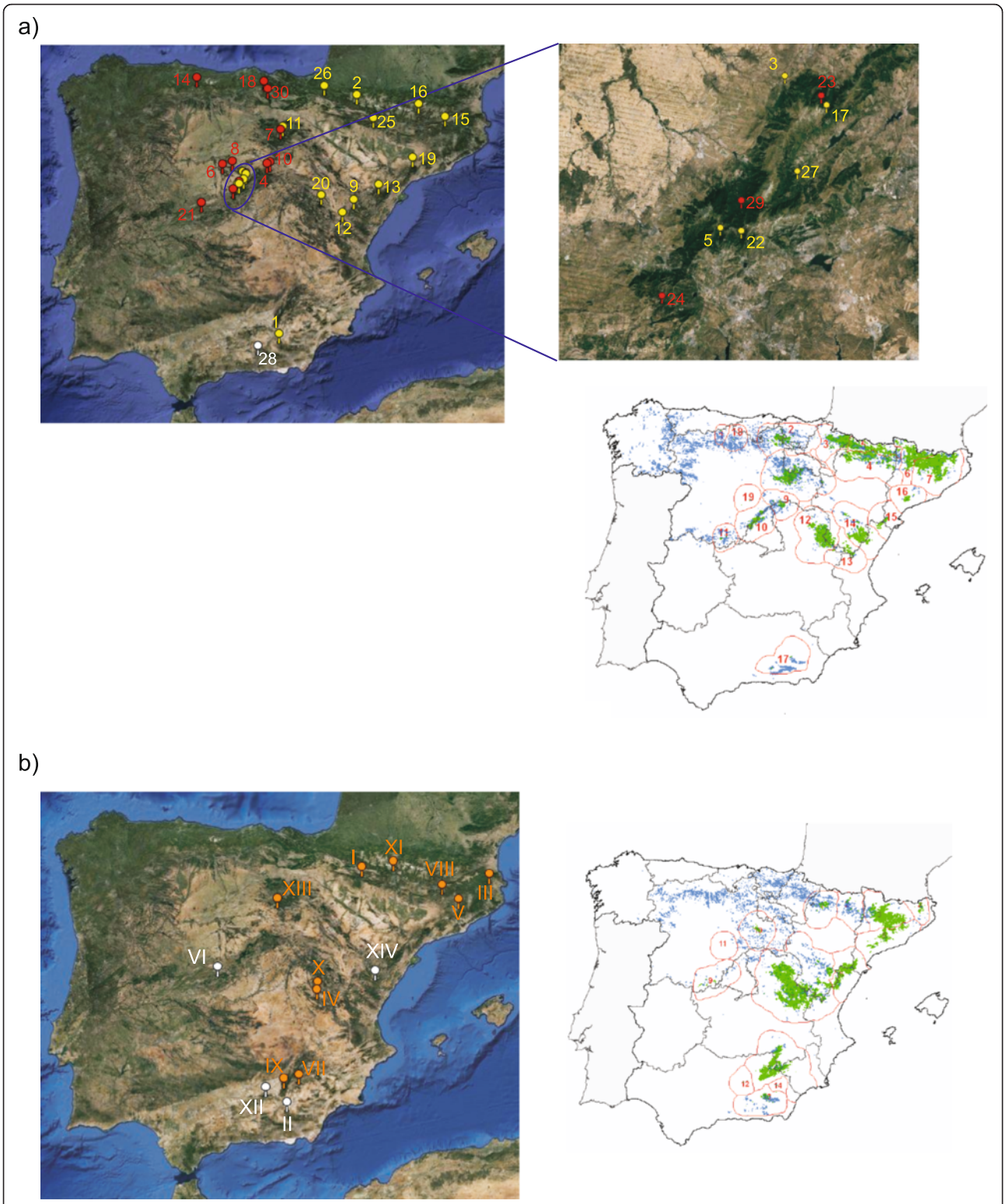

Fig 4 Population structure and distribution area of *Pinus sylvestris* and *P. nigra* in Spain. Results of the non-spatial Bayesian cluster analysis, BAPS, showing genetically homogenous groups of populations of (**a**) *P. sylvestris* and (**b**) *P. nigra* as from the the cpSSR dataset by Soto et al. [21]. Readers are referred to http://www.spatialepidemiology.net/user_maps/php/temp/06-23-15-74566.html and http://www.spatialepidemiology.net/user_maps/php/temp/06-25-15-94314.html for a closer inspection of these maps. Arabic numbers and roman numerals indicate population order as from [21]. Maps on the right [63] show the Spanish distribution range of each species, both natural (*green*) and reforested (*blue*) areas. Red lines and numbers refer to provenances

(or *P. pinaster*) is finally confirmed as the host-plant of *G. isabellae* at Sierra-María, future research will have to evaluate if its genetic differentiation of this southern locality is just the result of genetic drift (Fig. 3, Additional file 5) or whether adaptive divergence to a new, thermophilous, host is also involved (e.g. [6]).

G. isabellae and the LGM refugia of *Pinus* spp

The five differentiated groups of Spanish *Graellsia isabellae* reside in geographical areas identified as glacial refugia for *Pinus sylvestris* by Cheddadi et al. [16] (SI, EI, WP and EP) and Benito Garzón et al. [18] (the same ones plus CI). The origin of the two major mitochondrial clades observed in *G. isabellae*, the "southern" (CI and SI clusters) and "northern" lineages (EI, EP, WP and FA), fits a two glacial refugia scenario. These two groups would have initiated their divergence approximately 400 kya (Fig. 1a). Regardless of the actual date for their divergence, our data revealed an isolation scenario since, at least, the LGM. This hypothesis will be later revised in the light of the nuclear results (see section *Secondary contacts*), which grouped cluster EI with the so-called "southern lineage". The differentiated groups obtained for *P. sylvestris* and *P. nigra* (Fig. 4) probably remained isolated since the LGM as well. The deeper divergence of the southernmost population of the Scots pine (*P. sylvestris nevadensis*, revised by [15]) indicates its more ancient isolation, a common finding for other taxa occurring in Sierra Nevada (e.g. [86, 87]). The clear West–East differentiation of the other Spanish populations was congruent with the mitochondrial and nuclear data from prior literature (Additional file 7). Moreover, the Central Iberian System harbours a boundary between two gene zones (Fig. 4a) resembling the case of *Pinus pinaster* [88] and supporting the role of both sides of this mountain range as refugial areas [18, 19, 57].

Within the "northern" lineage of *G. isabellae*, we have just argued that genetic drift may account for a pre-Holocene origin of the WP cluster (Fig. 1a). This hypothesis would make plausible a Western Pyrenean refuge for this moth, inhabiting that refuge also postulated by Cheddadi et al. [16] and Benito Garzón et al. [18] for *Pinus sylvestris*. These last two works inferred the presence of *P. sylvestris* in the Western Pyrenees during the LGM, although the former presented a much larger potential habitat, covering the Eastern Pyrenees as well. Even assuming an isolated Western Pyrenean LGM refuge for the Scots pine as suggested at Fig. 3 (ECHAM3 scenario) by Benito Garzón et al. [18], we were unable to track the putative forest of *sylvestris/nigra* that connected at some point the Iberian System/Ebro Valley and the Western Pyrenees allowing *G. isabellae* to establish the WP cluster. Maybe the postglacial migration of the so-called "red-haplotype" of *P. sylvestris* from Eastern Iberia to the Western Pyrenees (Fig. 7 by [16]) was somehow singular and left no footprints in the Mid-Holocene expansion scenario displayed in Fig. 4 by [18]. Or maybe that connection occurred before the LGM, as in the case of the plant *Ramonda myconi* [76], and therefore was not reported in any of those works. The LGM-isolation hypothesis of cluster WP receives additional support from the phylogeography of *Carduelis citrinella*, a pine-associated bird, which shows remarkable coincidences with *G. isabellae* in terms of distribution area and genetic structure [81].

We advocate that cluster EP was founded by individuals belonging to EI, independently and after the establishment of cluster WP. The EI-EP corridor was noticeable in the reconstruction of potential habitat for *Pinus nigra* during the LGM and Mid-Holocene [18]. Therefore, we hypothesise that *P. nigra* was the host species that allowed *G. isabellae* to reach the Eastern Pyrenees at some point during the Holocene. It is worth noting that the expansion of *P. sylvestris* from both the Western and Eastern Pyrenean refugia had already covered most of the Spanish slopes of that mountain range at the Mid-Holocene (Fig. 4 by [18]), so the individuals arriving by means of the Black pine found a large area of suitable *P. sylvestris* habitat towards the West. We postulate the Holocene fragmentation of the EI-EP corridor of *P. nigra* as the most likely explanation for (i) an EP locality (Montesquiu) having one specimen carrying the most frequent haplotype (EI.1) of cluster EI, (ii) an EI site (Els-Ports) having four individuals showing the most frequent haplotype (EP.1) of cluster EP, and (iii) the seemingly "misplacement" of haplotype EI.20 being present in only one male from Els-Ports and just one substitution apart from EP.1 (also shared by four individuals from Els-Ports). Testing alternative hypotheses (unidirectional vs. bidirectional geneflow, founder effect vs. gradual fragmentation) for the colonisation of EP will require a more thorough sampling of the Catalonian areas where the distribution of both pine species overlap.

The terminal position of haplotype CI.1 (Fig. 1b) strongly suggests that the Central Iberian System derives from a small number of individuals with Betic ancestry (SI cluster). Indeed, the mismatch distributions of mitochondrial alleles in SI showed unequivocal footprints of a population expansion, likely after the LGM. The foundation of CI could have happened then by means of the northward expansion of *P. sylvestris* from its southernmost glacial refuge, as depicted at Fig. 7 by [16]. Unfortunately, the potential habitat reconstructions for the Scots pine at LGM and mid-Holocene failed to track such a SI-CI connection [18, 19]. The forest connecting SI and CI may have gone unnoticed by the potential habitat reconstructions for *P. sylvestris*, either because of a temporal mismatch or because it actually never existed and *G. isabellae* reached Central Spain by means of its other host-species, *P. nigra*. Both *P. sylvestris* and *P. nigra* were likely to be used by *G. isabellae* when the split between its northern and southern mitochondrial

lineages took place before the LGM. Then, the Central Iberian system (CI) was colonised by ancestors of the SI cluster. *P. nigra* was and currently still is most abundant than *P. sylvestris* in South Spain. By contrast, the occurrence of the Black pine in the Central Iberian System is currently minimal. Regardless of the pine species used to arrive to each of these mountain ranges, it is clear that the only possibility for the moth to survive was to use *P. nigra* in the South and *P. sylvestris* in Central Spain (Additional file 7).

Secondary contacts

We previously discussed that the two reciprocally monophyletic mitochondrial lineages of *G. isabellae* surely remained isolated since the LGM. This explanation is somehow jeopardised by the nuclear results, as the microsatellites revealed the EI cluster to be more closely related to the "southern-lineage": clusters SI and CI (Figs. 3b, c). According to microsatellites, the SI and EI clusters were connected when populations from EI and the Pyrenees became isolated. This mito-nuclear discordance might be related to male-biased dispersal, a life-history trait reported for *G. isabellae* [89] and invoked to account for the differences in nuclear and mitochondrial differentiation in another pine-dwelling lepidopteran [90]. However, given the geographical scale where the EI mito-nuclear discordance is found, a more plausible explanation is the existence of mito-nuclear incompatibilities, i.e. some kind of post-zygotic mechanism lowering the fitness of those Spanish Moon Moths with "southern lineage" mitochondrial haplotypes when carrying certain "northern" (EI) nuclear genotypes. Given the amount of non-synonymous substitutions in the "southern" lineage it is hardly surprising that selection favoured a specific "southern-mito + southern-nuclear" genetic combination during the period that those populations remained isolated from the rest of conspecifics. After the posterior mating between individuals from SI/CI and EI, the finding of a "southern-mito + northern-nuclear" combination would be expected in our dataset. That absence suggests that selection may have only allowed the "northern-mito + southern-nuclear" combination currently present at the Eastern Iberian System. The differential effect of specific mito-nuclear combinations on the fitness of interpopulation hybrids has been proved in other arthropods such as the copepod *Tigriopus californicus*, *Drosophila (Sophophora) melanogaster* and the seed beetle *Callosobruchus maculatus* (revised by [91, 92]).

A fine-scale survey in the Pyrenees is needed to confirm the presence of *G. isabellae* between Renanué and Baiasca (absent as from [67]) and unveil the factors preventing gene flow between clusters WP and EP. At present, two results from the present work are worth highlighting. The first is the lack of mitochondrial exchange between these two clusters (Fig. 1b). The second is a certain extent of asymmetric nuclear gene flow, as two individuals sampled at Renanué (WP) showed a high EP assignment probability, whereas individuals from Baiasca did not show any WP assignment probability (Fig. 3a). (1) The geographic coincidence of this phylogeographic break and the transition between the two genetically undifferentiated Pyrenean subspecies of *Pinus sylvestris* is remarkable. There are major climatic and ecological differences between the northern and southern slopes of the Pyrenees [93] as well as between the western and eastern parts of this mountain range (e.g. [76, 94–96]). The Spanish Moon Moth obviously needs a suitable host to complete its life cycle, but certain environmental variables, namely precipitation and temperature, have been inferred to be relevant for predicting its current distribution [10]. (2) Our preliminary results ruled out the effect of the maternally inherited bacterium *Wolbachia* as the cause for asymmetric gene flow between the two Pyrenean clusters. The westwards dispersal tracked by the high EP assignment probabilities of two individuals sampled at Renanué (WP) may reflect the onset of their hybridisation mediated by male gene flow. The increase of reforested pine woodland (Additional file 8) reinforces this hypothesis. The geographical pattern of mtDNA variability (higher in the central areas of WP and EP and monomorphic at both ends of each cluster, Fig. 1c) additionally agrees with a leading edge effect. However, further analyses are needed to confirm the aforementioned result as current gene flow.

Conclusions

Graellsia isabellae showed a strong phylogeographic pattern. The six differentiated groups revealed by the microsatellites and mtDNA showed a one to one correspondence with each of the main mountain ranges inhabited by the species. The mitochondrial data further clustered those groups into two major lineages, "southern" and "northern", which likely diverged before the LGM. There was no evidence of host-associated differentiation between populations using *P. sylvestris* and the ones utilising *P. nigra*. Eastern and Western Pyrenees were most likely colonised by individuals closely related to modern Eastern Iberian populations through independent asynchronous events. The past and present wider distribution of *P. nigra* when compared to *P. sylvestris* suggests that *G. isabellae* used the former to reach several mountain ranges where it currently lives on the latter. This seems to be the case for the populations of the Spanish Moon Moth from the Eastern Pyrenees and French Alps, which likely derived from the Holocene fragmentation of the continuous forest of *P. nigra* connecting different populations of *P. sylvestris* in Northeastern Spain and Southeastern France. The Central Iberian system descends from Betic ancestors, whose populations showed genetic footprints of both a population expansion and posterior bottlenecks. Subsequent gene flow between both the

Central/Betic mountains and the Eastern Iberian System was revealed by the nuclear dataset and was compatible with the overlapping potential distribution area of the Scots and Black pines during the Holocene. The mito-nuclear discordance involving the Eastern Iberian cluster is congruent with male-biased gene flow and/or a secondary contact after the evolution of mito-nuclear incompatibilities in geographically isolated areas.

Additional files

Additional file 1: Typing for three *Wolbachia* genes. Testing for the presence of *Wolbachia* in 53 specimens of *G. isabellae* following the protocol by Kodandaramaiah et al. [62]. (PDF 1008 kb)

Additional file 2: Details of the 95 % Statistical Parsimony network calculated for the 832 bp *COI* fragment of *G. isabellae* (Fig. 1b main text). a) Colours and numbers besides connections indicate the nucleotide where each mutation occurred. Stars point to ambiguous assignations as from software TCS 1.21; b) non-synonymous changes along the alignment. Details of the calculations used to test for no selection (H_O: $d_N/d_S = 1$) using MEGA 5; c) discussion. (PDF 180 kb)

Additional file 3: Geographic distribution of the mitochondrial haplotypes of *G. isabellae* and accession numbers. (PDF 149 kb)

Additional file 4: Genetic variation at microsatellite loci in populations of *G. isabellae*. The two values presented in the "Flobal F_{IS} column" correspond to the results of the entire dataset and results excluding loci *Gl23* and *Gl18*. Significant results at the 5 % after Bonferroni correction are shown in bold. (PDF 34 kb)

Additional file 5: Pairwise F_{ST} values for 28 localities of *G. isabellae*. Below diagonal: from the mitochondrial dataset (φ_{ST}). Above diagonal: from the microsatellite (nine) loci. Significant values are displayed in bold. Indicative adjusted nominal level (5 %) for multiple comparisons = 0.00013. (PDF 651 kb)

Additional file 6; Identification of the most likely number of *G. isabellae* populations by the analysis of microsatellite data with SRUCTURE 2.3.4. a) Estimated log probability of data for the different number of inferred clusters (*K*); bars correspond to standard deviation, after 20 independent runs; b) Rate of change in the log probability of data between successive *K* values (**Δ**k). Both figures were obtained with the aid of STRUCTURE HARVESTER 0.6.94 available at http://taylor0.biology.ucla.edu/struct_harvest/. (PDF 685 kb)

Additional file 7: Further details about the phylogeography of *Pinus sylvestris* and P. *nigra* and their relationship with *G. isabellae*. Population structure of *P. sylvestris* and *P. nigra* as revealed by different molecular markers. All figures are reproduced with kind permission of the copyright holder. (PDF 1273 kb)

Additional file 8: Detailed distribution of *Pinus sylvestris* in the Spanish Pyrenees, both natural (green) and reforested (blue) areas. a) Municipalities where samples of *G. isabellae* were obtained are highlighted; b) Enlarged view of the area between Baiasca (*L9*) and Renanué (*L10*). (PDF 3010 kb)

Abbreviations

Ca, circa; CI, Central Iberia; cpDNA, chloroplast DNA; EI, Eastern Iberia; EP, Eastern Pyrenees; FA, French Alps; HPD, High Posterior Density; ISSR, Inter Simple Sequence Repeat; Kya, kilo (10^3) years ago; MRCA, most recent common ancestor; mtDNA, mitocondrial DNA; mya, million (10^6) years ago; na, not applicable; SI, Southern Iberia; SSD, sum of squared deviations; SSR, single sequence repeats; vs, versus; WP, Western Pyrenees.

Acknowledgements

We thank N Remón and E Magnoux for help in the laboratory. J Baixeras, J Casaponsa, MA Gómez, S Ibáñez, DC Lees, R Maciá, MICROFAUNA, Y Monasterio, E Murria, JC Nevado, A Sáez, J Sevilla, T Latasa, J Ylla and, especially, F Goussard contributed to fieldwork with their advice, assistance and equipment. We appreciate the invaluable help provided by E Villagrasa (Parque Nacional de Ordesa y Monteperdido) and J Mestre (Parc Natural dels Ports). The male of *G. isabellae* included in Fig. 1c was photographed by T Decaens. U Kodandaramaiah and his team kindly provided controls for detection of *Wolbachia*. A Dolsa and J Baixeras generously sent us 13 of the samples tested for this intracellular parasite. This work benefited from the discussions held with D Buckley, M García-París, S González-Martínez, A Martínez-Abraín, W Nässig, JJ Pino, R Rougerie and J Rousselet. We thank the editor, R Vila, and two anonymous referees for their constructive comments. We are deeply indebted to DC Lees for his thorough revision of the final version of this work. Figures displayed in Additional file 7 are reproduced with permission of the copyright holder.

Funding

This work received funding from the following institutions: Ministerio de Educación y Ciencia (CGL2007-63549/BOS, BES-2008-002571, HF2007-0055), Xunta de Galicia (PGIDIT06PXIB103258PR, GRC2014/050), European Science Foundation (CONGEN 1635 & 1683), INRA (Projet-innovant-2007-EFPA) and Campus France (Partenariat-Hubert-Curien-Picasso-2008-7153UF).

Availability of data and materials

The mitochondrial haplotypes supporting the results of this article have been deposited on GenBank under accession numbers KX302361–KX302402 (see Additional file 3) and KX348010, whereas the microsatellite dataset is available in the figShare repository (entry DOI: 10.6084/m9.figshare.3423001 [97]). A fasta file containing the 41 mitochondrial haplotypes of *G. isabellae* can be found at figShare repository (entry DOI: 10.6084/m9.figshare.3423142 [98]).

Authors' contributions

MV and CLV conceived the study. MV, CLV, MAAR and HN obtained funding. NMM, MV and CLV collected samples. NMM and MV performed most lab and computer analyses. NMM, HN and MV drafted the manuscript. All authors contributed to the writing, read and approved the final manuscript.

Authors' information

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

Nonlethal sampling described in this paper has received prior and explicit approval from the competent authorities. Year 2007: fieldwork in France was carried out with collecting permit (Arrêté n° 2008.38.4) issued by Préfecture des Hautes Alpes. Fieldwork in Spain was specifically approved by the following institutions: Generalitat de Catalunya (Ref.: 0152S), Parc Natural dels Ports (Ref.: 0155S), Gobierno de Aragón (Ref.: LC/mp 24/2007/5682), Gobierno de Navarra (authorisation signed on May 7th, 2007, by Mr. Eraso Centelles (Director del Servicio de Conservación de la Biodiversidad), Mr. Martínez García (Técnico de la Sección de Hábitats) and Mr. Larumbe Arricibita (Jefe de la Sección de Hábitats)). Year 2008: Comunidad de Madrid (Ref.: ASP/mco), Generalitat de Catalunya (Ref.: SF/286, SF/287), Gobierno de Aragón (Ref.: LC/mp 24/20008/2015), Junta de Andalucía (Ref.: SGYB-AFR-CMM), Junta de Castilla y León (Ref.: EP/CYL/225/2008), Organismo Autónomo de Espacios Naturales de Castilla-La Mancha (Registro de salida 339796, 15 de abril de 2008). Year 2009: Generalitat de Catalunya (Ref.: SF/185, SF-186, SF/188), Generalitat Valenciana (Eixida 30297, 14 de abril de 2009), Gobierno de Aragón (Ref.: LC/mp 24/2009/1638), Organismo Autónomo de Espacios Naturales de Castilla-La Mancha (Registro de salida 425992, 8 de mayo de 2009). All permits were issued to N Marí Mena, C López Vaamonde and/or M Vila.

Consent for publication

Not applicable.

Author details

[1]Department of Molecular and Cell Biology, Evolutionary Biology Group (GIBE), Universidade da Coruña, A Fraga 10, E-15008 A Coruña, Spain. [2]AllGenetics & Biology, SL, Edificio de Servizos Centrais de Investigación, Campus de Elviña, E-15008 A Coruña, Spain. [3]INRA, UR633 Zoologie Forestière, F-45075 Orléans, France. [4]IRBI, UMR 7261, CNRS/Université François-Rabelais de Tours, 37200 Tours, France.

Received: 20 October 2015 Accepted: 12 June 2016
Published online: 24 June 2016

References

1. Matsubayashi KW, Kahono S, Katakura H. Divergent host plant specialization as the critical driving force in speciation between populations of a phytophagous ladybird beetle. J Evol Biol. 2011;24:1421–32.
2. Powell T, Forbes A, Hood G, Feder J. Ecological adaptation and reproductive isolation in sympatry: genetic and phenotypic evidence for native host races of *Rhagoletis pomonella*. Mol Ecol. 2014;23:688–704.
3. Nyman T, Vikberg V, Smith D, Boeve J. How common is ecological speciation in plant-feeding insects? A 'Higher' Nematinae perspective. BMC Evol Biol. 2010;10:266.
4. Brown J, LeebensMack J, Thompson J, Pellmyr O, Harrison R. Phylogeography and host association in a pollinating seed parasite *Greya politella* (Lepidoptera: Prodoxidae). Mol Ecol. 1997;6:215–24.
5. Mardulyn P, Othmezouri N, Mikhailov Y, Pasteels J. Conflicting mitochondrial and nuclear phylogeographic signals and evolution of host-plant shifts in the boreo-montane leaf beetle *Chrysomela lapponica*. Mol Phylogenet Evol. 2011;61:686–96.
6. Darwell C, Fox K, Althoff D. The roles of geography and founder effects in promoting host-associated differentiation in the generalist bogus yucca moth *Prodoxus decipiens*. J Evol Biol. 2014;27:2706–18.
7. Bensettiti F, Gaudillat V. *Graellsia isabellae* (Graëlls, 1849) L'Isabelle de France, le Papillon vitrail. Insectes, Lépidoptères, Saturniides. In: Naturelle MMMNdH, editor. Cahiers d'habitats Natura 2000 Connaissance et gestion des habitats et des espèces d'intérèt communautaire, Tome 7. Espèces animales. Paris: La documentation Française; 2002. p. 277–9.
8. Verdú JR, Numa C, Galante E. Atlas y libro rojo de los invertebrados amenazados de España (Especies Vulnerables), vol. I: Artrópodos. Madrid: Dirección General de Medio Natural y Política Forestal, Ministerio de Medio Ambiente, Medio Rural y Marino; 2011.
9. Montoya Moreno R, Hernández Alonso R. *Graellsia isabelae*. Vida Silvestre. 1975;12:207–20.
10. Chefaoui R, Lobo J. Assessing the conservation status of an Iberian moth using pseudo-absences. J Wildl Manage. 2007;71:2507–16.
11. de Arce Crespo JI, Jiménez Mendoza S, Sánchez Fernández P. Recopilación de la información biogeográfica, análisis de patrones ecológicos, conservación y mapa potencial de *Graellsia isabelae* (Graells, 1849) (Lepidoptera, Saturniidae) en la provincia de Cuenca. España Graellsia. 2010;66:9–20.
12. Gómez Bustillo MR, Fernández-Rubio F. Consideraciones sobre la planta nutricia de *Graellsia isabelae*. SHILAP Soc Hispano Luso Am Lepid. 1974;7:183–9.
13. Ibáñez Gázquez S, Nevado Ariza JC, Ylla Ullastre J. *Graellsia isabelae* (Graells, 1849), una nueva especie para la fauna lepidopterológica de Almería (España) (Lepidoptera: Saturniidae). SHILAP Soc Hispano Luso Am Lepid. 2008;36:427–30.
14. Fürstenberg-Hägg J, Zagrobelny M, Bak S. Plant defense against insect herbivores. Int J Mol Med Sci. 2013;14:10242–97.
15. Quézel P, Médail F. Écologie et biogéographie des forêts du bassin méditerranéen. Collection Environnement. Paris: Elsevier; 2003. p. 187–200.
16. Cheddadi R, Vendramin G, Litt T, Francois L, Kageyama M, Lorentz S, Laurent J, de Beaulieu J, Sadori L, Jost A, et al. Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. Glob Ecol Biogeogr. 2006;15:271–82.
17. Afzal-Rafii Z, Dodd R. Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (*Pinus nigra*) in Western Europe. Mol Ecol. 2007;16:723–36.
18. Benito Garzón M, Sánchez de Dios R, Sainz Ollero H. Predictive modelling of tree species distributions on the Iberian Peninsula during the Last Glacial Maximum and Mid-Holocene. Ecography. 2007;30:120–34.
19. Benito Garzón M, Sánchez de Dios R, Sainz Ollero H. The evolution of the *Pinus sylvestris* L. area in the Iberian Peninsula from the last glacial maximum to 2100 under climate change. Holocene. 2008;18:705–14.
20. Vila M, Marí-Mena N, Yen S, Lopez-Vaamonde C. Characterization of ten polymorphic microsatellite markers for the protected Spanish moon moth *Graellsia isabelae* (Lepidoptera: Saturniidae). Conserv Genet. 2010;11:1151–4.
21. Soto A, Robledo-Arnuncio J, González-Martínez S, Smouse P, Alía R. Climatic niche and neutral genetic diversity of the six Iberian pine species: a retrospective and prospective view. Mol Ecol. 2010;19:1396–409.
22. Millar J, McElfresh J, Romero C, Vila M, Marí-Mena N, Lopez-Vaamonde C. Identification of the Sex pheromone of a protected species, the Spanish moon moth *Graellsia isabellae*. J Chem Ecol. 2010;36:923–32.
23. Vila M, Auger-Rozenberg M, Goussard F, Lopez-Vaamonde C. Effect of non-lethal sampling on life-history traits of the protected moth *Graellsia isabelae* (Lepidoptera: Saturniidae). Ecol Entomol. 2009;34:356–62.
24. Simon C, Frati F, Beckenback A, Crespi B, Liu H, Flook P. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann Entomol Soc Am. 1994;87:651–701.
25. Vila M, Björklund M. The utility of the neglected mitochondrial control region for evolutionary studies in Lepidoptera (Insecta). J Mol Evol. 2004;58:280–90.
26. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25:1451–2.
27. Goudet J. FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered. 1995;86:485–6.
28. Kalinowski S. HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. Mol Ecol Notes. 2005;5:187–9.
29. Rousset F. GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Res. 2008;8:103–6.
30. Excoffier L, Lischer H. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res. 2010;10:564–7.
31. Dupanloup I, Schneider S, Excoffier L. A simulated annealing approach to define the genetic structure of populations. Mol Ecol. 2002;11:2571–81.
32. Corander J, Marttinen P, Sirén J, Tang J. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics. 2008;9.
33. Cheng L, Connor T, Sirén J, Aanensen D, Corander J. Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. Mol Biol Evol. 2013;30:1224–8.
34. Pritchard J, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945–59.
35. Falush D, Stephens M, Pritchard J. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2003;164:1567–87.
36. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005;14:2611–20.
37. Jakobsson M, Rosenberg N. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007;23:1801–6.
38. Rosenberg N. DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes. 2004;4:137–8.
39. Minch E, Ruiz-Linares A, Goldstein D, Feldman M, Cavalli-Sforza L. Microsat: a computer program for calculating various statistics on microsatellite allele data. Standford: Standford University Medical Center; 1996.
40. Felsenstein J. PHYLIP (Phylogenetic Inference Package) 3.68. Distributed by the author. Seattle: Department of Genome Sciences, University of Washington; 2005.
41. Clement M, Posada D, Crandall K. TCS: a computer program to estimate gene genealogies. Mol Ecol. 2000;9:1657–9.
42. Ronquist F, Teslenko M, van der Mark P, Ayres D, Darling A, Höhna S, Larget B, Liu L, Suchard M, Huelsenbeck J. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61:539–42.
43. Huelsenbeck J, Larget B, Alfaro M. Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. Mol Biol Evol. 2004;21:1123–33.
44. Kass R, Raftery A. Bayes factors. J Am Stat Assoc. 1995;90:773–95.
45. Papadopoulou A, Anastasiou I, Vogler A. Revisiting the insect mitochondrial molecular clock: the Mid-Aegean trench calibration. Mol Biol Evol. 2010;27:1659–72.
46. Rambaut A. FIGTREE 1.4.2. 2014; http://tree.bio.ed.ac.uk/software/figtree/ (Accessed 6 June 2016).
47. Schneider S, Excoffier L. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates very among sites: application to human mitochondrial DNA. Genetics. 1999;152:1079–89.
48. Rogers A, Harpending H. Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol. 1992;9:552–69.
49. Ramos-Onsins S, Rozas J. Statistical properties of new neutrality tests against population growth. Mol Biol Evol. 2002;19:2092–100.
50. Ramírez-Soriano A, Ramos-Onsins S, Rozas J, Calafell F, Navarro A. Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. Genetics. 2008;179:555–67.

51. Piry S, Luikart G, Cornuet J. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. J Hered. 1999;90:502–3.
52. Luikart G, Allendorf F, Cornuet J, Sherwin W. Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered. 1998;89:238–47.
53. Sinclair W, Morman J, Ennos R. The postglacial history of Scots pine (*Pinus sylvestris* L.) in Western Europe: evidence from mitochondrial DNA variation. Mol Ecol. 1999;8:83–8.
54. Soranzo N, Alia R, Provan J, Powell W. Patterns of variation at a mitochondrial sequence-tagged-site locus provides new insights into the postglacial history of European *Pinus sylvestris* populations. Mol Ecol. 2000;9:1205–11.
55. Naydenov K, Senneville S, Beaulieu J, Tremblay F, Bousquet J. Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. BMC Evol Biol. 2007;7:233.
56. Pyhäjärvi T, Salmela M, Savolainen O. Colonization routes of *Pinus sylvestris* inferred from distribution of mitochondrial DNA variation. Tree Genet Genomes. 2008;4:247–54.
57. Robledo-Arnuncio J, Collada C, Alía R, Gil L. Genetic structure of montane isolates of *Pinus sylvestris* L. in a Mediterranean refugial area. J Biogeogr. 2005;32:595–605.
58. Prus-Glowacki W, Stephan B, Bujas E, Alia R, Marciniak A. Genetic differentiation of autochthonous populations of *Pinus sylvestris* (Pinaceae) from the Iberian Peninsula. Plant Syst Evol. 2003;239:55–66.
59. Jaramillo-Correa J, Grivet D, Terrab A, Kurt Y, de-Lucas A, Wahid N, Vendramin G, González-Martínez S. The Strait of Gibraltar as a major biogeographic barrier in Mediterranean conifers: a comparative phylogeographic survey. Mol Ecol. 2010;19:5452–68.
60. Rubio-Moraga A, Candel-Perez D, Lucas-Borja M, Tiscar P, Viñegla B, Linares J, Gómez-Gómez L, Ahrazem O. Genetic diversity of *Pinus nigra* Arn. Populations in southern Spain and northern morocco revealed by inter-simple sequence repeat profiles. Int J Mol Sci. 2012;13:5645–58.
61. Sasaki T, Kubo T, Ishikawa H. Interspecific transfer of *Wolbachia* between two lepidopteran insects expressing cytoplasmic incompatibility: a *Wolbachia* variant naturally infecting *Cadra cautella* causes male killing in *Ephestia kuehniella*. Genetics. 2002;162:1313–9.
62. Kodandaramaiah U, Simonsen T, Bromilow S, Wahlberg N, Sperling F. Deceptive single-locus taxonomy and phylogeography: *Wolbachia*-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. Ecol Evol. 2013;3:5167–76.
63. García del Barrio J, Auñón F, Sánchez de Ron D. GIS of forest tree species of Spain. Instituto Nacional de Investigación y Tecnología Agraria y Alimentación. 2015. https://sites.google.com/site/sigtreeforestspeciesenglis/home/mapas-de-especies (Accessed 6 June 2016).
64. Dvornyk V, Sirvio A, Mikkonen M, Savolainen O. Low nucleotide diversity at the pal1 locus in the widely distributed *Pinus sylvestris*. Mol Biol Evol. 2002;19:179–88.
65. Pyhäjärvi T, García-Gil M, Knurr T, Mikkonen M, Wachowiak W, Savolainen O. Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. Genetics. 2007;177:1713–24.
66. Cipriano J, Carvalho A, Fernandes C, Gaspar M, Pires J, Bento J, Roxo L, Louzada J, Lima-Brito J. Evaluation of genetic diversity of Portuguese *Pinus sylvestris* L. populations based on molecular data and inferences about the future use of this germplasm. J Genet. 2013;92:e41–8.
67. Romo H, García-Barros E, Martín J, Ylla J, López M. *Graellsia isabelae*. In: Galante E, editor. Bases ecológicas preliminares para la conservación de las especies de interés comunitario en España: Invertebrados. Madrid: Ministerio de Agricultura, Alimentación y Medio Ambiente; 2012. p. 1–53.
68. Rousselet J, Zhao R, Argal D, Simonato M, Battisti A, Roques A, Kerdelhué C. The role of topography in structuring the demographic history of the pine processionary moth, *Thaumetopoea pityocampa* (Lepidoptera: Notodontidae). J Biogeogr. 2010;37:1478–90.
69. Recuero E, Garcia-París M. Evolutionary history of *Lissotriton helveticus*: Multilocus assessment of ancestral vs. recent colonization of the Iberian Peninsula. Mol Phylogenet Evol. 2011;60:170–82.
70. McCracken K, Sorenson M. Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear gene trees in the stiff-tailed ducks (*Nomonyx-Oxyura*)? Syst Biol. 2005;54:35–55.
71. Egger B, Koblmüller S, Sturmbauer C, Sefc KM. Nuclear and mitochondrial data reveal different evolutionary processes in the Lake Tanganyika cichlid genus *Tropheus*. BMC Evol Biol. 2007;7:137.
72. Piganeau G, Gardner M, Eyre-Walker A. A broad survey of recombination in animal mitochondria. Mol Biol Evol. 2004;21:2319–25.
73. Tsaousis A, Martin D, Ladoukakis E, Posada D, Zouros E. Widespread recombination in published animal mtDNA sequences. Mol Biol Evol. 2005;22:925–33.
74. Gotzek D, Clarke J, Shoemaker D. Mitochondrial genome evolution in fire ants (Hymenoptera: Formicidae). BMC Evol Biol. 2010;10:300.
75. Arunkumar K, Metta M, Nagaraju J. Molecular phylogeny of silkmoths reveals the origin of domesticated silkmoth, *Bombyx mori* from Chinese *Bombyx mandarina* and paternal inheritance of *Antheraea proylei* mitochondrial DNA. Mol Phylogenet Evol. 2006;40:419–27.
76. Dubreuil M, Riba M, Mayol M. Genetic structure and diversity in *Ramonda myconi* (Gesneriaceae): effects of historical climate change on a preglacial relict species. Am J Bot. 2008;95:577–87.
77. Templado J, Álvarez J, Ortiz E. Observaciones biológicas y citogenéticas sobre *Graellsia isabelae* (Graells 1849) (Lepidoptera: Saturniidae). Eos. 1975;49:285–92.
78. Fernández-Vidal EH. Comentarios acerca de la distribución geográfica francesa y notas taxonómicas sobre *Graellsia isabelae* (Graells, 1849) (Lepidoptera: Saturniidae). SHILAP Soc Hispano Luso Am Lepid. 1992;20:29–49.
79. Cassel-Lundhagen A, Ronnås C, Battisti A, Wallén J, Larsson S. Stepping-stone expansion and habitat loss explain a peculiar genetic structure and distribution of a forest insect. Mol Ecol. 2013;22:3362–75.
80. Roiron P, Chabal L, Figueiral I, Terral J, Ali AA. Palaeobiogeography of *Pinus nigra* Arn. subsp *salzmannii* (Dunal) Franco in the north-western Mediterranean Basin: a review based on macroremains. Rev Palaeobot Palynol. 2013;194:1–11.
81. Förschler M, Senar J, Borras A, Cabrera J, Björklund M. Gene flow and range expansion in a mountain-dwelling passerine with a fragmented distribution. Biol J Linn Soc Lond. 2011;103:707–21.
82. Masutti L, Battisti A, Faccoli M. Insect fauna of the *Pinus nigra* group in Italy. In: Lieutier F, Ghaioule D, editors. Entomological research in Mediterranean forest ecosystems. Paris: INRA editions; 2005. p. 79–87.
83. Zovi D, Stastny M, Battisti A, Larsson S. Ecological costs on local adaptation of an insect herbivore imposed by host plants and enemies. Ecology. 2008;89:1388–98.
84. Paiva M, Mateus E, Santos M, Branco M. Pine volatiles mediate host selection for oviposition by *Thaumetopoea pityocampa* (Lep., Notodontidae). J Appl Entomol. 2011;135:195–203.
85. Hódar JA, Cassel-Lundhagen A, Battisti A, Larsson S. A little further south: host range and genetics of the Northern pine processionary moth, *Thaumetopoea pinivora* (Lepidoptera: Notodontidae) at the southern edge of its distribution. Eur J Entomol. 2016;113:200–6.
86. Dinca V, Dapporto L, Vila R. A combined genetic-morphometric analysis unravels the complex biogeographical history of *Polyommatus icarus* and *Polyommatus celina* Common Blue butterflies. Mol Ecol. 2011;20:3921–35.
87. Miraldo A, Hewitt G, Paulo O, Emerson B. Phylogeography and demographic history of *Lacerta lepida* in the Iberian Peninsula: multiple refugia, range expansions and secondary contact zones. BMC Evol Biol. 2011;11:170.
88. Bucci G, González-Martínez S, Le Provost G, Plomion C, Ribeiro M, Sebastiani F, Alía R, Vendramin G. Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. Mol Ecol. 2007;16:2137–53.
89. Ylla i Ullastre J. Història natural del lepidòpter *Graellsia isabelae* (Graells, 1849). Barcelona: Institut d'Estudis Catalans; 1997.
90. Simonato M, Mendel Z, Kerdelhué C, Rousselet J, Magnoux E, Salvato P, Roques A, Battisti A, Zane L. Phylogeography of the pine processionary moth *Thaumetopoea wilkinsoni* in the Near East. Mol Ecol. 2007;16:2273–83.
91. Burton R, Barreto F. A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities? Mol Ecol. 2012;21:4942–57.
92. Wolff JN, Ladoukakis ED, Enríquez JA, Dowling DK. Mitonuclear interactions: evolutionary consequences over multiple biological scales. Phil Trans R Soc B. 2014;369:20130443.
93. Milá B, Surget-Groba Y, Heulin B, Gosá A, Fitze P. Multilocus phylogeography of the common lizard *Zootoca vivipara* at the Ibero-Pyrenean suture zone reveals lowland barriers and high-elevation introgression. BMC Evol Biol. 2013;13:192.
94. Jalut G, Marti J, Fontugne M, Delibrias G, Vilaplana J, Julia R. Glacial to interglacial vegetation changes in the northern and southern Pyrénées: Deglaciation, vegetation cover and chronology. Quat Sci Rev. 1992;11:449–80.
95. Van Dijk P, Bakx-Schotman T. Chloroplast DNA phylogeography and cytotype geography in autopolyploid *Plantago media*. Mol Ecol. 1997;6:345–52.

96. Charrier O, Dupont P, Pornon A, Escaravage N. Microsatellite marker analysis reveals the complex phylogeographic history of *Rhododendron ferrugineum* (ericaceae) in the Pyrenees. PLoS One. 2014;9:e92976.
97. Marí-Mena N, Lopez-Vaamonde C, Naveira H, Auger-Rozenberg MA, Vila M. Microsatellite datafile GENEPOP format.txt. Data from: Phylogeography of the Spanish Moon Moth *Graellsia isabellae* (Lepidoptera, Saturniidae). figshare. 2016;https://dx.doi.org/10.6084/m9.figshare.3423001.v2. (Accessed 15 June 2016).
98. Marí-Mena N, Lopez-Vaamonde C, Naveira H, Auger-Rozenberg MA, Vila M. Mitochondrial haplotype dataset. Data from: Phylogeography of the Spanish Moon Moth *Graellsia isabellae* (Lepidoptera, Saturniidae). figshare. 2016; https://dx.doi.org/10.6084/m9.figshare.3423142.v1. (Accessed 15 June 2016).
99. Collectif OPIE. Contribution à la connaissance de *Graellsia isabelae galliaegloria* Oberthur (Lepidoptera, Attacidae) connu uniquement en France. Rapport d'études de l'OPIE. 1998;3:1–36.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central