Received: 14 September 2017 | Revised: 24 July 2018 | Accepted: 28 July 2018

DOI: 10.1111/jbi.13429

RESEARCH PAPER

WILEY Journal of Biogeography

Phylogeography of the endangered saproxylic beetle Rosalia longicorn, *Rosalia alpina* (Coleoptera, Cerambycidae), corresponds with its main host, the European beech (*Fagus sylvatica*, Fagaceae)

Lukas Drag[1,2] | David Hauck[2] | Oldrich Rican[1] | Thomas Schmitt[3,4] | Dmitri F. Shovkoon[5,6] | Roman J. Godunko[2] | Gianfranco Curletti[7] | Lukas Cizek[1,2]

[1]Faculty of Science, University of South Bohemia, Ceske Budejovice, Czech Republic

[2]Institute of Entomology, Biology Centre CAS, Ceske Budejovice, Czech Republic

[3]Department of Taxonomy and Biogeography, Senckenberg German Entomological Institute, Müncheberg, Germany

[4]Faculty of Natural Sciences I, Zoology, Institute of Biology, Martin Luther University Halle-Wittenberg, Halle, Germany

[5]Institute of Steppe of the Ural branch of the Russian Academy of Sciences, Orenburg, Russian Federation

[6]Laboratory of Animal Systematic and Faunistic, Samara State University, Samara, Russian Federation

[7]Department of Entomology, Museo Civico di Storia Naturale, Carmagnola, Italy

Correspondence
Lukas Drag, Institute of Entomology, Biology Centre CAS, Branišovská 1160/31, 370 05 České Budějovice, Czech Republic
Email: lukasdrag@gmail.com

Funding information
Czech Science Foundation, Grant/Award Number: 17-21082S; Grant Agency of the University of South Bohemia, Grant/Award Number: 04-168/2013/P; Project Biodiversity of Forest Ecosystems, Grant/Award Number: CZ.1.07/2.3.00/20.0064; European Social Fund; state budget of the Czech Republic

Editor: Hans-Peter Comes

Abstract

Aim: The Rosalia longicorn (*Rosalia alpina*) is an internationally protected icon of biodiversity associated with old trees and dead wood. Although the beetle regularly exploits several *marginal* hosts, its preferred *main* host is European beech (*Fagus sylvatica* s.l.). Moreover, the geographical ranges of *R. alpina* and beech closely overlap. To assess whether their spatial association is mirrored in the genetic patterns of both species, we investigated the phylogeography of *Rosalia alpina* over its entire geographical range and compared it with the known genetic patterns of its hosts.

Location: Europe and western Asia.

Methods: Using both mitochondrial (COI) and nuclear (14 microsatellite loci) markers, we analysed 148 (444, respectively) individuals from 31 (30, respectively) sites. We constructed a Bayesian Inference tree and a haplotype network, calculated the spatial analysis of molecular variance and assessed the population structure of our dataset using two Bayesian clustering methods (STRUCTURE and BAPS).

Results: Mitochondrial markers suggested existence of five clades in *R. alpina* populations. Two of them were endemic to the Italian mainland, one to Sicily, and another to southern Turkey. The remaining clade probably originated in the Balkans and colonized the rest of the species' range. Nuclear markers supported this division. They also suggested two main recolonization routes from the Balkans; one heading north and then both west and east, the second expanding eastwards as far as the Caucasus. The observed genetic patterns were largely congruent with those of European beech.

Main conclusions: The results of both markers were mostly congruent, suggesting at least four potential refugia for *R. alpina* located in the southernmost parts of its geographical range. Its populations from a large part of Europe and western Asia, however, were genetically poor, dominated by a single haplotype. Phylogeographies of the beetle and its *main* host seem to be tightly matched, reflecting their common history.

KEYWORDS
biogeography, conservation, insect-plant interactions, postglacial recolonization, refugium, spatial genetic structure

1 | INTRODUCTION

Host-specific organisms, such as many parasites, symbionts, or phytophages, are forced to react to environmental changes within the frame set by their hosts (Ahern, Hawthorne, & Raupp, 2009; Nieberding, Morand, Libois, & Michaux, 2004). If the relationship between a specialist and its host is long-term, the population genetic structure of the specialist should be strongly influenced by its host, likely mirroring its phylogeographical structure. Numerous studies have already documented this relationship for host-parasite systems (Blouin, Yowell, Courtney, & Dame, 1995; Nadler, 1995; Nieberding & Morand, 2006), but comparative studies investigating the role of the host in shaping phytophage phylogeographical structures are still rather rare (but see Avtzis, Bertheau, & Stauffer, 2012).

This is particularly regrettable as phytophages are among the most diverse and ecologically important organisms in terrestrial systems (Hamilton et al., 2013). In this group, insects are by far the most numerous herbivores, and are often highly specialized (Novotny et al., 2010). Evolution of insect herbivore diet breadth has therefore received substantial attention (reviewed in Jaenike, 1990). Ecological monophagy is a specific type of specialization where local populations are highly host-specific, but the species as a whole exploits different hosts over its range (Fox & Morrow, 1981; Scriber, 1986). This phenomenon is surprisingly common among insects (e.g., Jermy & Szentesi, 2003; Scriber, Allen, & Walker, 2006). Some suggestions to explain the occurrence of locally narrowed host use include strong ecological selection, natural enemies (Murphy, 2004), and the absence of host–plant species used elsewhere (Scriber, 1986).

The Rosalia longicorn (*Rosalia alpina*; Linnaeus, 1758) is a xylophagous beetle that represents a specific example of ecological monophagy. It is able to exploit a wide range of distantly related hosts (e.g., Cizek, Schlaghamersky, Borucky, Hauck, & Helesic, 2009), but it is mainly associated with just one, European beech (*Fagus* spp.). This host is preferred whenever present, although many populations also exploit *marginal* hosts including maple (*Acer* spp.), ash (*Fraxinus* spp.), lime (*Tilia* spp.), and elm (*Ulmus* spp.). In the absence of the *main* host, *R. alpina* populations are able to survive solely on the *marginal* hosts (Cizek et al., 2009; Picard, 1929; Shapovalov, 2012). Development of the beetle has also been recorded in a wide range of *occasional* hosts, including e.g., hornbeam (*Carpinus* spp.), horse chestnut (*Aesculus hippocastanum*), and walnut (*Junglans* spp.) (Hovorka, 2011; Merkl, Hegyessy, & Kovács, 1996). The *main* and *marginal* hosts include members of Fagales, Sapindales, Lamiales, Rosales, and Malvales (Apg, 2009). *Rosalia alpina*'s host range thus encompasses trees as phylogenetically distant as possible for a beetle developing in the wood of broad-leaved trees in Europe and western Asia.

Despite the wide use of different host species, the distribution of *R. alpina* closely overlaps with the *main* host, *Fagus sylvatica* s.l. (i.e., the complex of the European beech *F. sylvatica* s.str. L. and the Oriental beech *F. orientalis* Lipsky; Greuter & Burdet, 1981). Within the geographical range of beech, the beetle is absent from Great Britain and most of NW Europe; it has recently disappeared from Scandinavia, most of Germany, Poland, and the Czech Republic (Lindhe, Jeppsson, & Ehnström, 2011; Michalcewicz & Ciach, 2015). Several lowland populations inhabit forests where beech is locally absent (Drag et al., 2015). Outside the beech range, the beetle is known from the Urals and several sites along major rivers of E Europe (Shapovalov, 2012).

In this study, we examined whether the close association between *R. alpina* and beech is reflected in similar genetic patterns of both species. To do this, we studied the geographical pattern of genetic variation in *R. alpina* over its entire geographical range. Using both mitochondrial (cytochrome c oxidase subunit I) and nuclear (14 microsatellite loci) markers, we analysed 148 (444, respectively) individuals from 31 (30, respectively) localities across Europe and western Asia. Furthermore, we compared the phylogeographic structure and demographical history of *R. alpina* to patterns of known genetic variation for its *main* and *marginal* host species.

2 | MATERIALS AND METHODS

2.1 | Sampling

Previously published data from Central and South-eastern Europe (Drag et al., 2015) were supplemented by data from new populations originating from various parts of the species' range in Europe (Alps, Pyrenees, Italy) and western Asia (Urals, Caucasus). We also included one population that is considered a separate subspecies, *Rosalia alpina syriaca* (Pic, 1894), from the Hatay province in southern Turkey (for more details see Table 1). In total, beetles were sampled from 31 localities across most of the geographical range of *R. alpina*. Part of a middle leg from all sampled specimens was taken in the field and stored in vials containing 96% ethanol for molecular analyses. Genomic DNA was extracted from the sampled tissue using the Genomic DNA Mini Kit Tissue (Geneaid Biotech Ltd., Taiwan) following the manufacturer's instructions.

2.2 | COI

2.2.1 | Amplification, sequencing, and alignment

A partial (766 bp) fragment of the mitochondrial gene cytochrome c oxidase subunit I (COI) was amplified and sequenced for 68

TABLE 1 Sampled populations of the Rosalia longicorn (*Rosalia alpina*) including information about their host tree, GPS coordinates, and the number of individuals analysed for cytochrome c oxidase I (COI) and 14 nuclear microsatellite loci. In total, 148 individuals from 31 sites were analysed for COI. For microsatellites, 444 individuals from 30 sites were grouped to 25 populations

						Number of individuals analysed	
Country	Locality	Code	Host tree	Latitude	Longitude	COI	Microsatellites
Albania	Deje Mts	AL1	FS	41.7094	20.158686	1	0
Austria	Kalkalpen	AT1	FS	47.805571	13.950015	5	20
Bulgaria	Stara planina II	BG1	FS	42.784421	23.790154	5	20
Bulgaria	Strandja	BG2	FO	42.08869	27.750227	5	20
Croatia	Lonsko Polje	HR1	FS	45.196004	17.128523	5	18
Czech R	Bezdez	CZ1	FA	50.539185	14.720318	5	20
France	Marais-Poitevin	FR1	FE	46.301939	–0.604017	4	20
France	Rhone Alpes	FR2	FS, APs	44.953561	5.268031	4	20
France	Ferrette	FR3	FS	47.449193	7.325025	5	15
Georgia	Lagodekhi	GE1	FO	41.848542	46.301536	8	20[A]
Georgia	Kvetera	GE2	FO	42.056733	45.098442	1	1[A]
Georgia	Omalo	GE3	FO	42.242475	45.495175	1	1[A]
Greece	Olymp	GR1	FS	40.108456	22.460764	5	20
Greece	Pindos	GR2	FS	39.959741	20.906086	6	16
Greece	Vermio	GR3	FS	40.589685	22.042605	5	7
Greece	Rhodope	GR4	FS	40.921661	24.189622	5	20
Greece	Evros	GR5	FS	41.109798	25.962128	5	9
Hungary	Borzsony	HU1	FS	47.917322	18.977368	5	20
Italy	Livata	IT1	FS	41.922908	13.174433	5	7[B]
Italy	Abruzzo	IT2	FS	41.798529	13.770105	4	5[B]
Italy	Alumiere	IT3	FS	42.151168	11.90679	3	3[B]
Italy	Sicily	IT4	FS	37.957662	15.044074	7	27
Romania	Comana forest	RO1	AC, T, F	44.15751	26.100216	5	18
Romania	Apuseni	RO2	FS	46.461244	23.374803	5	20
Russia	Urals	RU1	APl, UG	52.555806	56.165222	8	21
Serbia	Stara planina I	SR1	FS	44.173172	22.123664	5	9
Slovakia	Vihorlat	SK1	FS	48.886894	22.242261	5	20
Spain	Aralar	ES1	FS	42.984417	–2.12885	7	17[C]
Spain	San Sebastian	ES2	FS	43.24897	–1.821457	1	1[C]
Turkey	Hatay	TR1	FO	36.840639	36.36629	6	21
Crimea	Crimea	CR1	FO	44.74788	34.333431	7	8

[A,B,C] in microsatellites populations from Georgia (A), Italy (B), and Spain (C) were merged together as GE1-3, IT1-3, and ES1-2; FS: *Fagus sylvatica*; FO: *Fagus orientalis*; AC: *Acer campestre*; APs: *Acer pseudoplatanus*; APl: *Acer platanoides*; UG: *Ulmus glabra*; FA: *Fraxinus angustifolia*; FE: *Fraxinus excelsior*; F: *Fraxinus* sp.; T: *Tilia* sp.

individuals of *R. alpina*. Together with 80 COI sequences previously published in Drag et al. (2015), we analysed 148 sequences from 31 localities across Europe and western Asia. As outgroups, we used one individual of *Rosalia coelestis* (Semenov, 1911) and one individual of *Rosalia lameerei* (Brogn, 1890). We used universal forward and reverse primers (Appendix S1 in Supporting Information). Sequencing was performed by Macrogen Inc. The sequences from each individual were edited and aligned (MUSCLE, default settings) in GENEIOUS, 6.1.6 (Biomatters).

2.2.2 | Phylogenetic analysis

Phylogenetic relationships among individual haplotypes were reconstructed using Bayesian Inference analysis (BI) in BEAST 1.8.4 (Drummond, Suchard, Xie, & Rambaut, 2012) and Maximum Likelihood analysis (ML) in PHYML 3.1 (Guindon & Gascuel, 2003). The GTR+G model was identified as the best fitting model by the JMODELTEST 2.1.7 (Darriba, Taboada, Doallo, & Posada, 2012) using the Akaike information criterion (AIC). For the BI analysis we used the strict

molecular clock model and for tree prior the coalescent model with constant size. The analyses were performed with several independent runs for 10 million generations with trees sampled every 3,000 generations. The analyses were run at the freely available The CIPRES Science Gateway 3.3 (www.phylo.org). Runs were checked for convergence diagnostics with Tracer 1.8 (Rambaut, Drummond, Xie, Baele, & Suchard, 2018). Four well-converged runs were combined in LogCombiner 1.8 with a burn-in of 10% for each of the data partition schemes. The final tree was produced from these data with TreeAnotator 1.8.

In the ML analysis, the tree searches were performed with the NNIs search option and a parsimony initial tree. Branch support for the resulting tree was evaluated by the nonparametric bootstrap method with 1,000 replicates. The final phylogenetic tree was visualized in FigTree 1.4.2 (Rambaut, 2014).

A haplotype network was produced using the statistical parsimony method (95% connection limit) implemented in tcs, 1.21 (Clement, Posada, & Crandall, 2000).

2.2.3 | Genetic diversity and differentiation

Standard genetic indices such as the number of haplotypes (*H*), the haplotype (*h*) and nucleotide (π) diversities, and the number of polymorphic sites (*P*) were computed for each clade, as well as for both subclades within the clade C5 (see below) using DnaSP, 5.10 (Librado & Rozas, 2009). Divergence, expressed as the uncorrected p-distance, within and among clades and two outgroups was measured in mega, 6 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013). A spatial analysis of molecular variance (SAMOVA; Dupanloup, Schneider, & Excoffier, 2002) was used to investigate the geographical structure of studied populations. samova 1.0 (Dupanloup et al., 2002) was run for K = 2–15 with 1,000 permutations from 100 initial conditions. To examine signs of past population size changes, we calculated two commonly used indices for each clade as well as for both subclades within the clade C5 in Arlequin 3.5.1.3 (Excoffier & Lischer, 2010): Tajima's *D* (Tajima, 1989) and Fu's F_S (Fu, 1997). The significance of both values was tested by 10,000 permutations. According to Tajima (1989) and Fu (1997), *D* and F_S indices are sensitive to population demographical expansion, which normally leads to significantly negative values.

2.3 | Microsatellites

2.3.1 | Amplification

Microsatellite analyses comprised a total of 444 individuals, genotyped for 14 polymorphic loci. Eight loci were previously used in Drag et al. (2015). The additional six loci were newly developed for this species using the same method and PCR protocol as described in Drag, Zima, and Cizek (2013) (Appendix S2). All 444 individuals were collected from the same populations as the individuals used for COI analyses, except for one individual from the AL1 population. Unlike COI, some geographically close sampling localities had to be merged to one population for the microsatellite analyses in order to maintain sufficient sample sizes resulting in 25 populations (Table 1). PCR products were analysed with the automated sequencer ABI 3730XL (Applied Biosystems, USA) by Macrogen Inc., Korea. Allelic patterns were scored using genemapper, 3.7 (Applied Biosystems).

2.3.2 | Loci characteristics and genetic diversity

The occurrence of null alleles was investigated by FreeNA (Chapuis & Estoup, 2007). For each population, we calculated the number of alleles (N_A), the number of effective alleles (N_E), the number of polymorphic loci (*P*), observed (H_O) and expected (H_E) heterozygosity, and the number of private alleles (N_{PA}) using genAlEx, 6.5 (Peakall & Smouse, 2012). Linkage disequilibrium (LD) between all pairs of loci, as well as the Hardy–Weinberg equilibrium (HWE) across loci and populations were tested with genepop, 4.1.3 (Rousset, 2008) using default settings.

2.3.3 | Population structure

The population structure of our dataset was assessed using two Bayesian clustering methods. structure 2.3.4 (Pritchard, Stephens, & Donnelly, 2000) attempts to minimize departures from HWE and LD, while baps 6.0 (Corander & Marttinen, 2006) seeks genetically homogeneous clusters employing information on geographic coordinates of individuals. In both programs, the single outlying population from Hatay was excluded from the analysis because of its remote genetic and geographic position (for more information, see COI results). First, STRUCTURE (assuming an admixture model and correlated allele frequencies) was run for values of *K* ranging from 1 to 15 with 100,000 burn-in and 1,000,000 Markov chain Monte Carlo (MCMC) steps with 10 replicates for each *K*. Each resulting value of ln $P(D|K)$ was plotted against *K* and the best value of *K* was chosen to correspond to the point at which the curve plateaus. The best *K* value was also chosen according to Evanno, Regnaut, and Goudet (2005), using structure harvester Web 0.6.94 (Earl & vonHoldt, 2012). The results obtained for a given *K* were then processed with the Greedy algorithm in clumpp 1.1.2 (Jakobsson & Rosenberg, 2007) and visualized as barplots in distruct 1.1 (Rosenberg, 2004). Results for K = 5 that seemed to be the most informative were used to generate pie charts, illustrating the geographical structure of each population. Since it is recommended to use different Bayesian clustering approaches (e.g., Pearse & Crandall, 2004), we also used BAPS to infer the population structure with the inclusion of spatial data. We ran BAPS for 100 runs, assuming the spatial clustering models with fixed *K* (fixed-*K* clustering) ranging from 2 to 15.

2.3.4 | Genetic differentiation

To construct a phylogenetic tree of populations, we performed an evolutionary analysis of allele frequencies using the neighbour-joining (NJ) method as implemented in poptree2 (Takezaki, Nei, & Tamura, 2010). As a genetic measure, we used the D_A distance

values (Nei & Chesser, 1983) that appear to be more suitable for microsatellite data than other distance measures (Takezaki & Nei, 2008). As three loci (RA_08, RA_40 and RA_37) had missing data for all individuals within one population (see results), they were excluded from this analysis, thus leaving eleven loci. Finally, we edited the constructed NJ tree in MEGA, 6 (Tamura et al., 2013).

Gene flow among populations was estimated as pairwise F_{ST} values calculated in ARLEQUIN, 3.5.1.3 (Excoffier & Lischer, 2010). The significance of the derived genetic distances was tested by 10,000 permutations.

3 | RESULTS

3.1 | COI

3.1.1 | Phylogenetic analysis

We identified 37 different haplotypes based on 148 sequences of the 766 bp fragment of the mitochondrial gene COI (Appendix S3). Both phylogenetic analyses, Bayesian Inference (Figures 1a and 2) and Maximum Likelihood (Appendix S4) revealed five clades (C1, C2, C3, C4, C5), which corresponds with the results gained from the statistical parsimony analysis (Figure 1b). The strongly differentiated clade C1 was formed by the six individuals from the Hatay province in southern Turkey (TR1). Clade C2, also markedly different from all other clades, consisted of a single individual sampled in the Abruzzo region in Italy (IT2). Clade C3 included all individuals from Sicily (IT4) while clade C4 grouped all individuals from mainland Italy (IT1, IT2, IT3). Clade C5 included all remaining individuals. This geographically widely distributed and haplotypically heterogeneous clade C5 was subdivided in two subclades. The heterogeneous subclade C5a contained mostly individuals from the three western and central Greek populations (GR1, GR2, GR3) and from Albania (AL1), whose haplotypes were separated by three or more mutations from the central haplotype HT_1; the genetically poor but geographically differentiated subclade C5b embraced remaining individuals carrying the haplotypes differentiated by one or two mutations from HT_1 (Figure 1b). The division of the C5 clade was thus admittedly arbitrary in order to allow for distinguishing the genetically rich populations of the SW Balkans from the rest of the clade.

3.1.2 | Genetic diversity and differentiation

Overall haplotype diversity (h) was 0.687 and nucleotide diversity (π) was 0.0061 (Table 2). Subclade C5b, represented by 109 individuals distributed over a substantial part of the species' range, had a comparatively low number of haplotypes (16). Also its haplotype (h = 0.429) and nucleotide (π = 0.00071) diversity were the lowest of all clades represented by more than one individual. By contrast, C5a was the most diverse clade, despite being restricted to western and central Greece and Albania. The 14 individuals sampled represented eleven haplotypes and had the highest diversity indices of all clades (h = 0.967, π = 0.00374).

The proportion of sequence differences (p-distance) among the clades ranged from 0.009 (between C4 and C5) to 0.03 (between C1 and C5) (Table 3). The percentage of differences between the basal Hatay population (C1) and all other clades were always about 3%. The genetic divergence between *R. alpina* and other species of the same subgenus (*R. lameerei*, *R. coelestis*) was 10% and 14%.

Congruent results were obtained from the spatial analysis of molecular variance (SAMOVA; Appendix S5). The highest F_{CT} value (0.8614, $p < 0.05$) was recorded for the division into two clusters (K = 2), which separated clade C1 from all other populations. Nevertheless, a similar F_{CT} value was recorded for K = 6 (0.8613, $p < 0.01$) closely followed by K = 8 (0.8603, $p < 0.01$). Calculated for each clade, Tajima's D and Fu's F_S had significantly negative values only for clade C5. Concerning the subclades within C5, C5b had a significant negative value for both parameters, while C5a had significant value only for Fu's F_S, but not for Tajima's D (Table 2).

3.2 | Microsatellites

3.2.1 | Loci characteristics and genetic diversity

In total, 444 individuals from 25 populations were genotyped for 14 microsatellite loci. All analysed loci were polymorphic with the total number of alleles per locus ranging from 3 to 16 (mean 7.1 ± 4.2). Some loci could not be genotyped for some populations (locus RA_08 and RA_40 for Hatay; RA_37 for the Urals), probably due to long-term isolation resulting in the changes of the primer sequences. The overall occurrence of null alleles was low, and the mean estimated frequency per locus across all populations never exceeded 10% (except for RA_15, with 16.5%). No linkage disequilibrium was found between any pairs of loci for each population after Bonferroni correction for multiple tests. Five loci (RA_08, RA_15, RA_40, RA_37, and RA_28) deviated from HWE in at least one population (after Bonferroni correction),) probably as a consequence of the violation of some assumptions (Frankham, Ballou, & Briscoe, 2010).

The mean number of alleles per locus and population ranged from 1.4 in CR1 to 4.1 in GR1 and GR2 (Appendix S6). These values corresponded with the ones for the expected heterozygosity (H_E), with the lowest in CR1 (0.095) and the highest in GR1 (0.455) and GR2 (0.523). The H_E values also significantly correlated with both mitochondrial diversity indices (haplotype diversity: R^2 = 0.243, $p < 0.05$; nucleotide diversity: R^2 = 0.330, $p < 0.005$). Private alleles were recorded for thirteen populations, with the highest number in GR2 and TR1 (four private alleles each).

3.2.2 | Population structure

According to ΔK, the best number of clusters in STRUCTURE (excluding the Hatay population) was two (K = 2, but see Janes et al., 2017) (Appendix S7A). Such division separated Italy (IT1-3), including Sicily (IT4), from all other populations, and the three Greek populations (GR1, GR2 and GR3) had an intermediate assignment (Figure 3a). For K = 3, these Greek populations were assigned as an independent

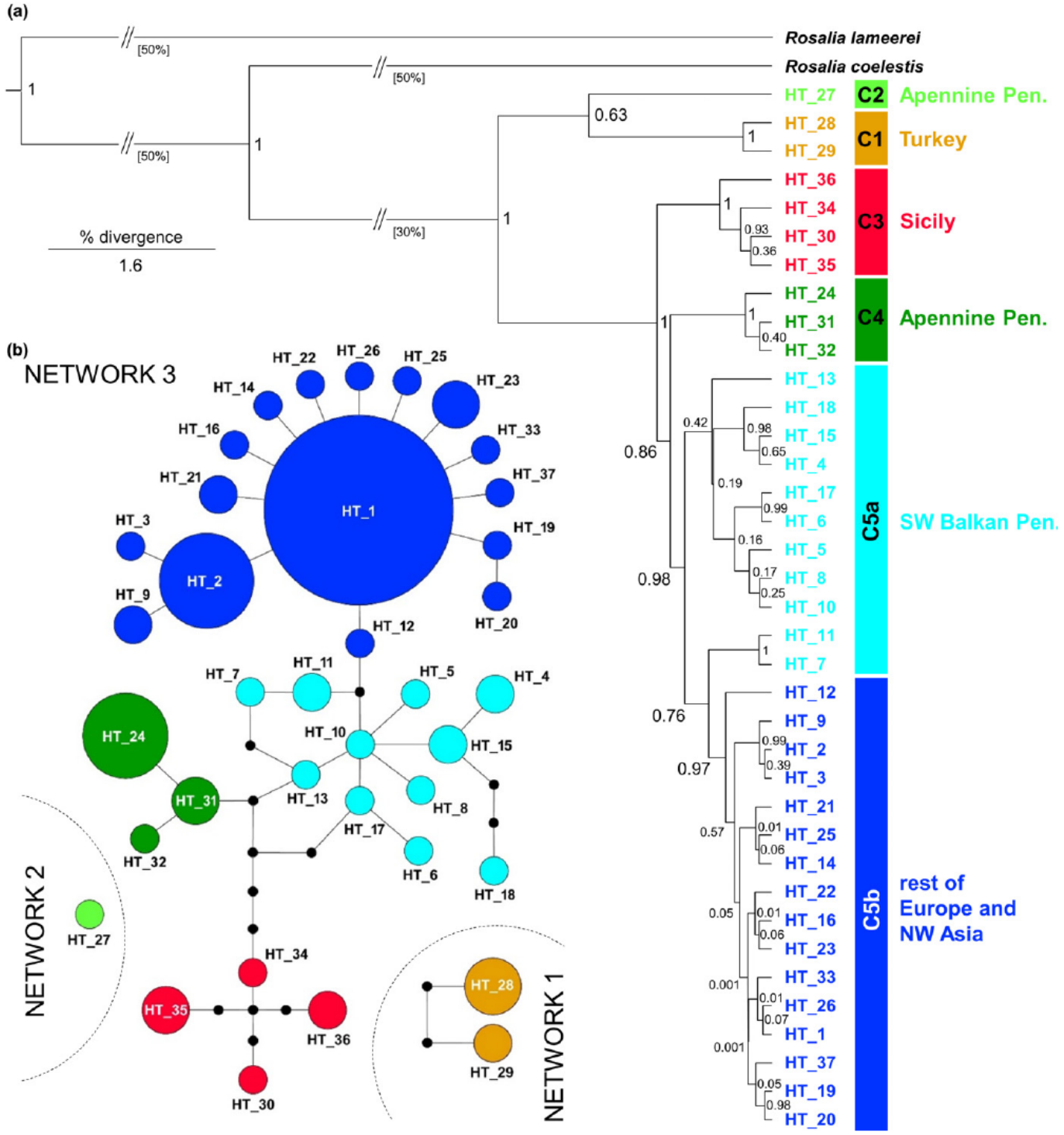

FIGURE 1 (a) Bayesian Inference tree (BEAST; Drummond et al., 2012) of Rosalia longicorn (*Rosalia alpina*) based on mitochondrial cytochrome c oxidase I (COI) sequence data. *Rosalia coelestis* and *R. lameerei* were used as outgroups. The scale bar represents 1.6% of pairwise divergence. The five identified clades are marked with different colours. Clade C5 was further divided into two subclades (C5a and C5b). The posterior probabilities are shown for each branch. Branch lengths of some groups were shortened (values in the brackets). (b) Three haplotype networks constructed using the statistical parsimony method (TCS; Clement et al., 2000) with 95% connection limit. Each haplotype is represented by a circle whose size is proportional to the haplotype frequency. Small black circles indicate missing haplotypes necessary to link all observed haplotypes to the network [Colour figure can be viewed at wileyonlinelibrary.com]

cluster as well as the Italian ones. This division was also obtained for all subsequent *K*s. The third group, represented by the rest of the analysed populations, largely remained homogeneous only until *K* = 4. For *K* = 5, it was divided into a more eastern and a more western lineage meeting in the Carpathians. These two lineages, however, were not well separated geographically and the Carpathian

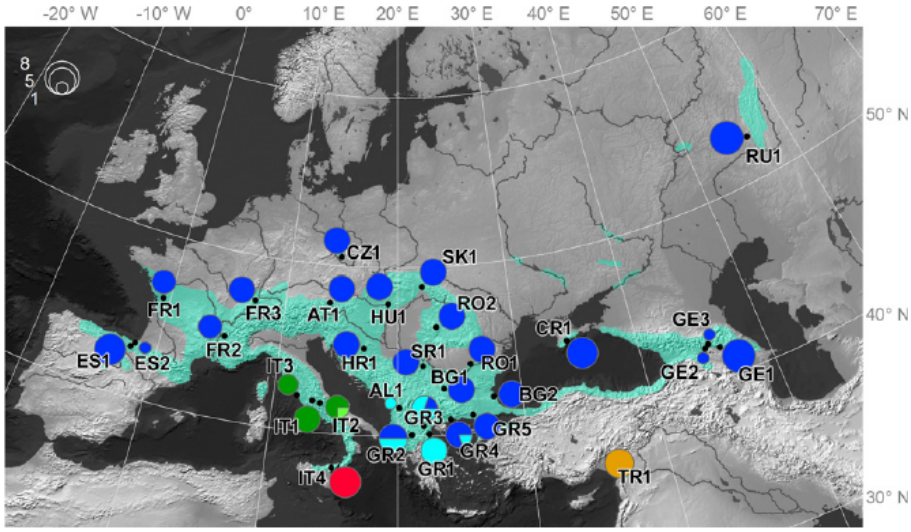

FIGURE 2 Geographical distribution of five clades (with two subclades of clade C5) across 31 sampled populations (full names are listed in Table 1) of Rosalia longicorn (*Rosalia alpina*) based on COI sequences. Colours correspond with Figure 1a,b. The circle sizes are proportional to the number of individuals from each population. The green colour in the background represents the distribution range of *R. alpina* [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Genetic diversity indices of Rosalia longicorn (*Rosalia alpina*) calculated for each clade and both subclades within C5. Results derived from cytochrome c oxidase I (COI)

Clade	*N*	*H*	*h*	π	*P*	Tajima's *D*	Fu's F_S
C1	6	2	0.533	0.00209	3	1.124	2.506
C2	1	1	—	—	—	—	—
C3	7	4	0.81	0.00385	7	0.173	0.627
C4	11	3	0.564	0.0009	2	0.036	−0.113
C5	123	27	0.552	0.00179	26	−2.094*	−26.755*
C5a	14	11	0.967	0.00374	12	−0.956	−6.598*
C5b	109	16	0.429	0.00071	15	−2.214*	−19.179*

Note. *N*: number of individuals within each clade; *H*: number of haplotypes; *h*: haplotype diversity; π: nucleotide diversity; *P*: number of polymorphic sites.
*$p < 0.01$.

populations represented a mixture of both. Further phylogeographical structures were unravelled with increasing *K*. For example, the Urals (RU1), as the most distant population, represented a separate cluster for $K \geq 5$. Unexpectedly, the Crimea (CR1) population clustered together with populations in the Pyrenees (ES1-2) and the western Alps (FR2) for $K \geq 6$. However, this grouping might be an artefact, due to the limited number of individuals from Crimea. At $K = 7$, the population from the Caucasus (GE1-3) and a substantial number of individuals from eastern Bulgaria (BG2) also indicated a separate group.

BAPS detected $K = 10$ as the best supported. Nevertheless, we also performed analyses for smaller *K*s to detect gradual clustering and to compare these results with those from STRUCTURE (Appendix S8). Generally, the results obtained from BAPS were mostly in accordance with the STRUCTURE assignments. Most clusters suggested for any given *K* included the same populations in both analyses, with the exception that BAPS analysis split the Italian and Sicilian populations into two separate clusters for $K \geq 5$.

3.2.3 | Genetic differentiation

A neighbour-joining (NJ) tree, based on D_A distances between all populations (Figure 4), was similar to the Bayesian Inference tree based on COI sequences. Correspondingly, the NJ tree showed the same main clades (i.e. C1, C3, C4, C5; the C2 mtDNA clade was represented by a single individual and, therefore, was combined with C4 in this analysis), although relationships between clades were

TABLE 3 Estimates of genetic divergences within and between five clades of Rosalia longicorn (*Rosalia alpina*) and its two outgroups (*R. coelestis* and *R. lameerei*). Results derived from mitochondrial cytochrome c oxidase I (COI)

		Between clades					
Clade	Within clade	C1	C2	C3	C4	C5	*R. coelestis*
C1	0.0021						
C2	n.c.	0.023					
C3	0.0039	0.029	0.024				
C4	0.0009	0.028	0.023	0.010			
C5	0.0018	0.030	0.025	0.012	0.009		
R. coelestis	n.c.	0.108	0.102	0.102	0.100	0.102	
R. lameerei	n.c.	0.144	0.142	0.141	0.140	0.145	0.146

Note. n.c.: the clade C2 and two outgroups were represented by a single individual, hence genetic divergence within the clade could not be calculated.

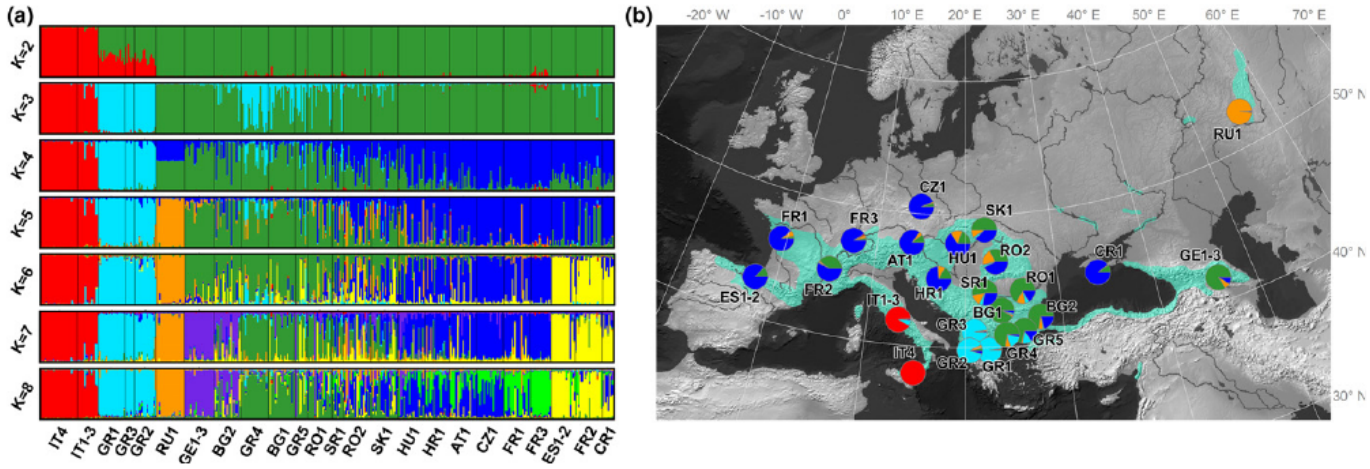

FIGURE 3 Bayesian clustering analysis (STRUCTURE; Pritchard et al., 2000) of 24 populations (full names are listed in Table 1) of Rosalia longicorn (*Rosalia alpina*) based on 14 nuclear microsatellite loci. (a) Barplots illustrate the estimated assignment of each individual into *K* clusters (*K* = 2–8). According to the method of Evanno et al. (2005), the best number of clusters was two (*K* = 2; Appendix S7). Solid black lines define the boundaries between the same populations as in the map. (b) Geographical distribution and the genetic structure of 24 studied populations. Each pie chart represents the proportion of membership of individuals from a given population in each of five clusters (*K* = 5) as revealed by STRUCTURE. The green colour in the background represents the distribution range of *R. alpina* [Colour figure can be viewed at wileyonlinelibrary.com]

slightly distinct. In agreement with the mtDNA BI tree, the nuclear tree also showed that populations from western and central Greece (i.e. C5a) were not monophyletic and that the populations from north-eastern Greece (GR4, GR5) were closely related to the populations occupying large parts of the species' distribution (C5b).

The pairwise F_{ST} values indicated a large range of genetic differentiation, from values close to zero to 0.79 (between the populations in Crimea and Sicily) (Appendix S9). In general, the lowest F_{ST} values were found between many populations within the clade C5b, and the highest values were present between Hatay (TR1), Sicily (IT4) and mainland Italy (IT1-3) [on one side], and any of the other populations [on the other side].

4 | DISCUSSION

We present a new phylogeographical reconstruction for the Rosalia longicorn (*Rosalia alpina*) across its geographical range. Information about the genetic structure of both the beetle and its *main* host, the European beech (*Fagus sylvatica* s.l.), allowed us to make a direct comparison of their phylogeographical patterns and infer conclusions regarding their common history. To our knowledge, this is the first such study of an invertebrate species restricted to European broad-leaved forests.

4.1 | *Rosalia alpina* phylogeography

The mtDNA-based clade C1 formed by the Hatay population from southern Turkey (TR1) is well-supported, and together with C2 is the most basal. There is no indication of recent gene flow between the Hatay population and other populations. However, the genetic distance of the Hatay population to other clades (p-distance ~3%) was lower than is usually considered for a species limit among taxonomically or ecologically related taxa (Audisio et al., 2009; Nakamine & Takeda, 2008; Solano et al., 2013). The average genetic distance between *R. alpina* and other species from the same subgenus (*R. lameerei*, *R. coelestis*) used as outgroups ranged from 10% to 14%. The small enclave in southern Turkey therefore represents a highly differentiated, most likely endemic, group of *R. alpina* associated with *Fagus orientalis*. The high degree of differentiation has probably accumulated due to long-term isolation over several glacial-interglacial cycles and well justifies its subspecific status as *R. alpina syriaca*.

The mtDNA and nuclear microsatellite analyses clearly distinguished the Italian populations (including Sicily) from the rest of Europe, implying that neither of the studied populations was involved in postglacial colonization of Europe. However, the high genetic differentiation between Sicily and central Italy in both marker systems support two separate gene pools, calling for at least two different centres of differentiation: one in mainland Italy and one in Sicily, both most likely having represented glacial refugia following the refugia-within-refugia theory (Gómez & Lunt, 2007). Other studies also revealed endemic lineages for Sicily, thus supporting its importance as a differentiation centre and refugium (e.g., *Erinaceus europaeus:* Santucci, Emerson, & Hewitt, 1998; *Osmoderma eremita* s.l.: Audisio et al., 2009; *Melanargia galathea*: Habel, Lens, Rödder, & Schmitt, 2011).

The provenances of mainland Italy embraced two mtDNA clades C2 and C4. The clade C2, represented by a single individual, was strongly differentiated from all other clades, being similar to the genetic distance between *R. alpina syriaca* (C1) and other clades. To rule out any possibility of a mistake, the single C2 individual was sequenced again. However, its nuclear DNA was not exceptional in the context of the mainland Italian populations. Therefore, the existence of the clade C2 most likely points to the survival of an ancient mtDNA lineage that is still detectable owing to the nonrecombinant nature of mtDNA.

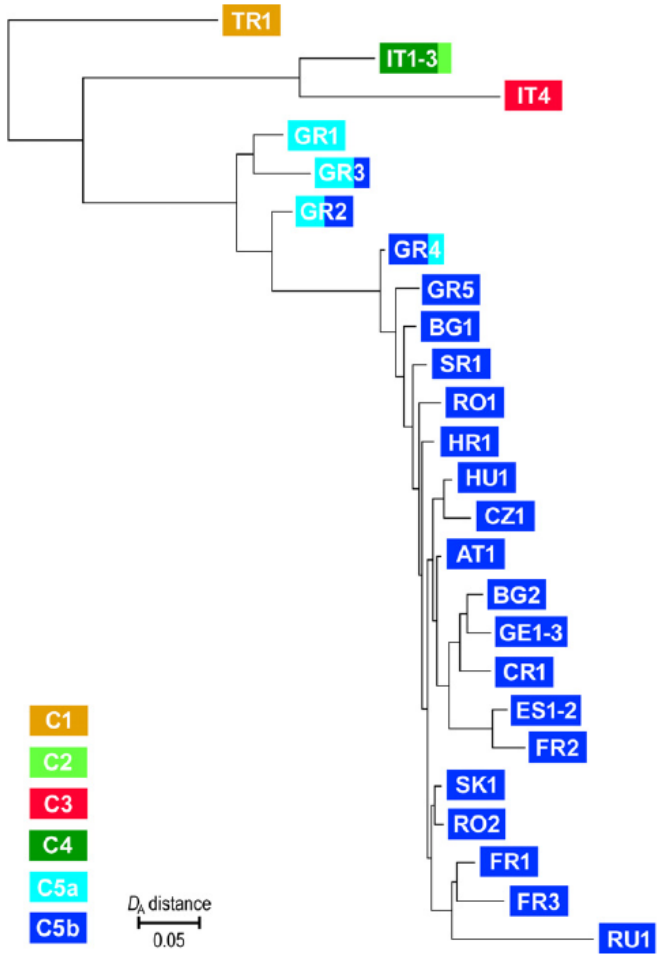

FIGURE 4 Unrooted neighbour-joining tree of 25 populations of Rosalia longicorn (*Rosalia alpina*) based on D_A distance values gained from 11 nuclear microsatellite loci in POPTREE2 (Takezaki et al., 2010). The scale bar represents 0.05 of D_A values. Full names of the locations are listed in Table 1. Colours indicate five clades and two subclades gained from the Bayesian Inference tree in Figure 1 [Colour figure can be viewed at wileyonlinelibrary.com]

One monophyletic clade (C5) inhabited the remainder of the geographical range of *R. alpina*. The paraphyletic subclade C5a restricted to the SW Balkans had the highest genetic diversity, with numerous unique haplotypes. The genetically poor subclade C5b occupied the eastern Balkans and the rest of the species' geographical range from Spain to the Urals. A single haplotype (HT1) dominated most populations throughout this range. Our data thus corroborate the hypothesis that the mountains of the Balkans served as a glacial refugium for *R. alpina* (Drag et al., 2015). The high amount of unique haplotypes in the southern Balkans supports the assumption of a structured refugium and long-term continuous persistence in this region. Similar findings exist for many species, indicating a refugia-within-refugia structure in the Balkans during glacial periods (e.g., *Arion fuscus:* Pinceel, Jordaens, Pfenninger, & Backeljau, 2005; *Sus scrofa*: Alexandri et al., 2012; *Lissotriton vulgaris:* Pabijan et al., 2015).

Since we found no indication that *R. alpina* survived in an extra-Mediterranean glacial refugia (e.g., the Alps, Pyrenees, or Caucasus; Schmitt, 2007; Schmitt & Varga, 2012), it is likely that the Balkans acted as the only source for postglacial recolonization of most of Europe and western Asia. Because the GR4 and GR5 populations in the eastern Balkans are very similar to the populations from the postglacially colonized region, a colonization mostly originating from that region is the most parsimonious scenario. If so, populations in north-eastern Greece or southern Bulgaria would represent the leading edge, while the central and western Greek populations GR1, GR2 and GR3 would be rear edge populations (Hampe & Petit, 2005) that did not contribute to postglacial recolonization.

Based on the microsatellite pattern and previous findings (Drag et al., 2015), two main recolonization routes can be proposed. One probably went north and then split west and east as far as Spain and the Urals. The second expanded eastwards to the Caucasus, possibly following the southern coast of the Black Sea. The Carpathians represents a mixture of both. However, it still remains unclear if both routes originated from a single refugium in north-eastern Greece, or if another refugium, probably in the mountains along the east coast of the Adriatic Sea, was involved.

4.2 | Comparative phylogeography

To assess whether *R. alpina* was historically closely associated with its hosts, we compared the phylogeographical patterns of *R. alpina* to those of its *marginal* hosts (*Fraxinus*, *Ulmus*, *Acer*, and *Tilia*) and the *main* host (*Fagus sylvatica* s.l.). Even though information about some of the *marginal* hosts is rather fragmentary, some basic patterns can be inferred (Grimm & Denk, 2014; Heuertz et al., 2006; Myking & Yakovlev, 2006; Phuekvilai, 2014; Whiteley, 2004). Unlike *R. alpina*, the *marginal* hosts consist of several genetically diversified lineages surviving in multiple refugia throughout Europe. Since these refugia were often involved in subsequent recolonizations of Europe, the current genetic structure of any of the *marginal* hosts does not seem to be congruent with the patterns found for the beetle.

The best correspondence was thus found between the beetle and its *main* host (*Fagus sylvatica* s.l.; Figure 5). Oriental beech (*F. orientalis*) is ancestral to European beech (*Fagus sylvatica* s.str.) (Gömöry & Paule, 2010), and the only *R. alpina* population ancestrally exploiting Oriental beech (the Hatay population) is a sister clade to all remaining populations. Virtually the whole range of *Fagus orientalis* currently inhabited by *R. alpina* was most likely colonized by the lineage originally exploiting *F. sylvatica* s.str. (Figure 5). In *F. sylvatica* s.str., three main haplotype groups exist for chloroplast markers (Demesure, Comps, & Petit, 1996; Magri et al., 2006). One group was observed in the Apennine Peninsula, the second in the Balkans, and the third over the rest of Europe. These results are complementary to those obtained for *R. alpina* from mitochondrial (COI) sequences. The recorded haplotype diversity observed for cpDNA in beech also corresponds with the beetle's distribution pattern, showing that the southernmost parts of the natural range of beech had higher diversity than the central and northern populations (Magri et al., 2006). A comparison between nuclear markers of the beetle and

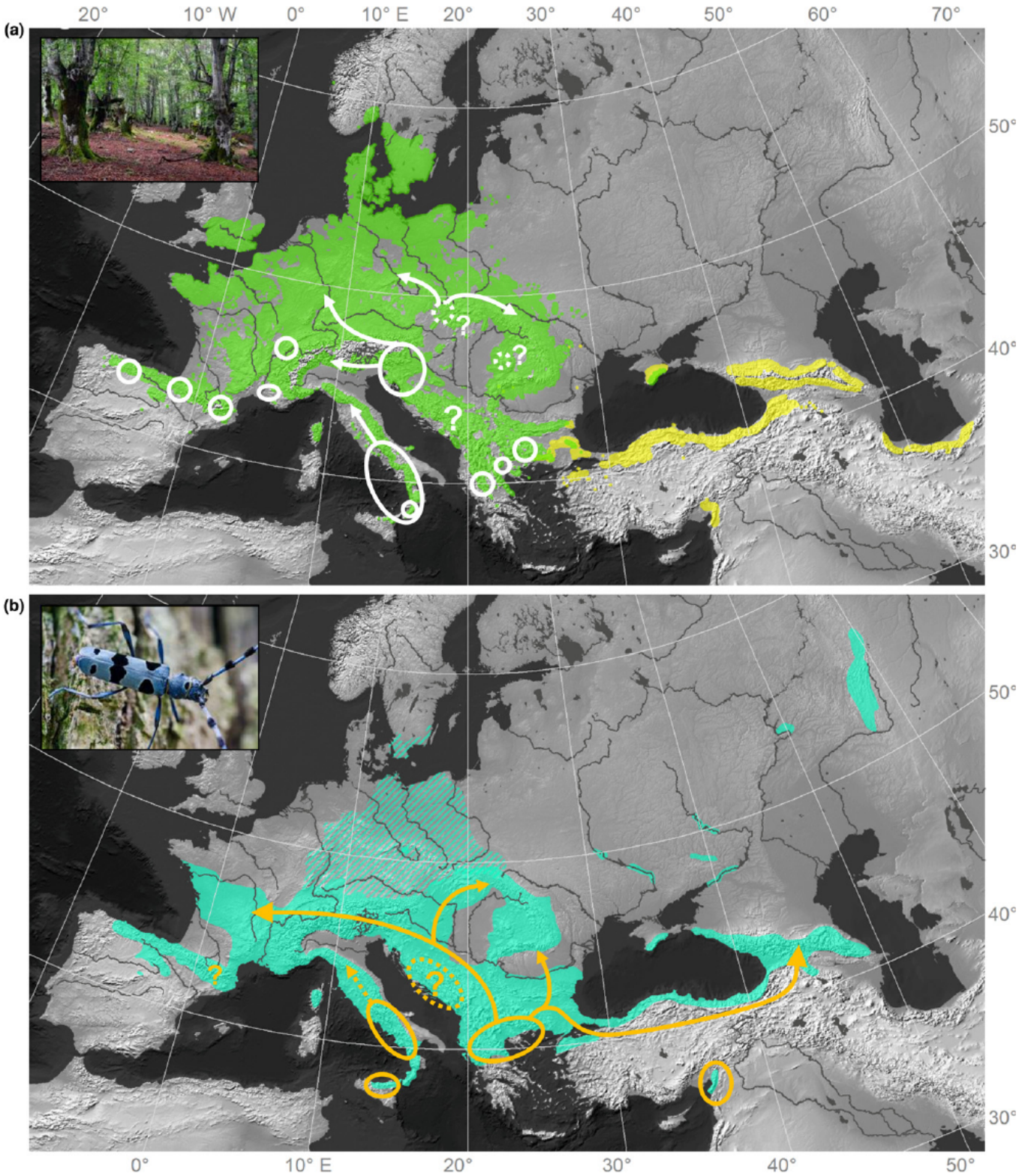

FIGURE 5 (a) Distribution range of *Fagus sylvatica* s.str. (green) and *Fagus orientalis* (yellow) and the tentative location of refugia and the main postglacial recolonization routes for *F. sylvatica* s.str. based on Magri et al., 2006. (b) Approximate distribution range of Rosalia longicorn (*Rosalia alpina*) and its tentative location of refugia and the postglacial recolonization routes. The dashed area represents the approximate historical distribution [Colour figure can be viewed at wileyonlinelibrary.com]

its *main* host (microsatellites and isozymes) was less clear, yet the main patterns were also congruent (Comps, Gömöry, Letouzey, Thiébaut, & Petit, 2001; Magri et al., 2006). The highest discrepancy exists in south-western Europe, where the beech populations appear to be more differentiated than those of the beetle (De Lafontaine, Ducousso, Lefevre, Magnanou, & Petit, 2013).

To conclude, the main congruent phylogeographical patterns between *R. alpina* and *F. sylvatica* s.str. are that populations of both species survived the last glaciation in the Apennine Peninsula, but did not cross the Alps; the south-western Balkans served as an important refugium displaying the highest genetic diversity; and that populations over a large part of Europe are dominated by a single haplotype. The phylogeographies of *R. alpina* and its *main* host thus seem to be tightly matched in many aspects.

Unlike the beetle, the biogeography of its host can be studied through additional methods. Paleobotanical analyses, including pollen and macrofossils records, indicated the existence of further refugia located north of the typical Mediterranean refugia, e.g., southern France, the eastern Alps–Slovenia–Istria, the Carpathian Basin and possibly even further north (De Lafontaine, Amasifuen Guerra, Ducousso, & Petit, 2014; Magri et al., 2006; Magyari et al., 2014). The available genetic data do not support the existence of such refugia for *R. alpina*. The potential beech refugia at more northern latitudes were probably small and scattered in favourable locations (Allen et al., 2010; Tzedakis, Emerson, & Hewitt, 2013), as is typical for many extra-Mediterranean refugia (Schmitt & Varga, 2012). They likely did not offer favourable conditions allowing for the long-term survival of *R. alpina*, which seems to be unable to persist in small, isolated populations (Drag et al., 2015). Hence, the absence of evidence for such extra-Mediterranean refugia for *R. alpina* does not mean that such refugia did not exist for beech. Similarly, the proposed beech refugia in the Pyrenees and their vicinity (Magri et al., 2006) were probably too small to sustain the beetle through an ice age, although samples of *R. alpina* from other parts of the Iberian Peninsula would be required to substantiate any possible refugium in this part of their current range.

4.3 | Ecological monophagy

Rosalia alpina is apparently able to maintain viable populations on a phylogenetically wide range of trees. While potentially a polyphagous species, it nevertheless does exhibit a close association with its *main* host (European beech complex). Today, *Fagus* serves as the *main* host in the regions geographically matching the assumed glacial refugia of *R. alpina* (Hatay, SW Balkans, Sicily, mainland Italy). Consequently, it is highly unlikely that the *main* host has been incorporated into the diet of a previously polyphagous beetle only recently.

Our data do not allow a conclusion on whether the *marginal* hosts were always part of the species' diet or whether their host list has broadened relatively recently. However, related *Rosalia* species inhabiting eastern parts of Asia exploit *Acer* spp. as their main host (Cherepanov, 1981). Further, many of the *marginal* hosts probably shared glacial refugia with beech. *Rosalia alpina* thus probably exploited both the *main* and the *marginal* hosts during the last glaciation. Nevertheless, beech was one of the last tree species to spread across Europe in the Holocene, while many *marginal* hosts such as elm, lime, and ash spread earlier, together with trees such as birch and hazel (Giesecke, Brewer, Finsinger, Leydet, & Bradshaw, 2017; Tinner & Lotter, 2006).

The question remains why *R. alpina* was not able to follow the spread of its *marginal* hosts for several thousand years, but instead followed the much slower spread of beech. Possible explanations include the nutritional qualities of beech wood for the beetle larvae, or shared ecological requirements of both species. Since individuals from *marginal* hosts were not smaller than those from the *main* host (Michalcewicz & Ciach, 2012), the nutrition hypothesis seems unlikely, although further data would be required. The most likely explanation of the close relationship between *R. alpina* and beech thus seems to be their shared ecological requirements for climatic or habitat conditions.

To conclude, our analyses suggest that *R. alpina* has a close and long-lasting, but not currently exclusive, relationship with its *main* host, the European beech (*Fagus sylvatica* s.l.). Despite the fact that beech and *R. alpina* are widespread in Europe and western Asia, most of the genetic diversity of *R. alpina* is found in rather small enclaves of beech forest in the extreme south of this tree's range, including the mountains of Sicily, Greece, and the Hatay province in southern Turkey. These forests are, however, vulnerable to climate change and human pressure due to their limited size, isolation, and high humidity requirements in otherwise relatively dry regions. The patterns of genetic diversity observed in *R. alpina* are likely to be shared by many species in European broad-leaved forests. Conservation of beech forest in the extreme south of its range should thus be given a high priority.

ACKNOWLEDGEMENTS

We thank Andrej Shapovalov (Saint Petersburg), Danilo Russo (Napoli), Xanti Pagola (Spain), Sandor Berces (Budapest), Christian Komposch and Sandra Aurenhammer (Graz), Evgeni Chehlarov (Sofia), Dan Lestina and Martin Slachta (Ceske Budejovice), Robert Tropek (Prague), and Jiri Prochazka (Brno) for assistance with obtaining samples; Jan Miklín (Ostrava) for creating the maps; Martin Volf (Ceske Budejovice) for his valuable comments on the manuscript and Amelia Hood, Matthew Sweney and Carl Wardhaugh for language correction. We also thank two anonymous reviewers for helpful comments on the manuscript. The study and its authors were supported by the Czech Science Foundation (17-21082S), the Grant Agency of the University of South Bohemia (04-168/2013/P), and the Project Biodiversity of Forest Ecosystems (CZ.1.07/2.3.00/20.0064), cofinanced by the European Social Fund and the state budget of the Czech Republic.

ORCID

Lukas Drag http://orcid.org/0000-0002-6002-9214

REFERENCES

Ahern, R. G., Hawthorne, D. J., & Raupp, M. J. (2009). Phylogeography of a specialist insect, Adelges cooleyi: Historical and contemporary processes shape the distribution of population genetic variation. *Molecular Ecology*, *18*, 343–356. https://doi.org/10.1111/j.1365-294X.2008.04034.x

Alexandri, P., Triantafyllidis, A., Papakostas, S., Chatzinikos, E., Platis, P., Papageorgiou, N., … Triantaphyllidis, C. (2012). The Balkans and the colonization of Europe: The post-glacial range expansion of the wild boar, *Sus scrofa*. *Journal of Biogeography*, *39*, 713–723. https://doi.org/10.1111/j.1365-2699.2011.02636.x

Allen, J. R. M., Hickler, T., Singarayer, J. S., Sykes, M. T., Valdes, P. J., & Huntley, B. (2010). Last glacial vegetation of northern Eurasia. *Quaternary Science Reviews*, *29*, 2604–2618. https://doi.org/10.1016/j.quascirev.2010.05.031

Apg, I. I. I. (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, *161*, 105–121.

Audisio, P. H., Brustel, H., Carpaneto, G. M., Coletti, G., Mancini, E., Trizzino, M., … De Biase, A. (2009). Data on molecular taxonomy and genetic diversification of the European Hermit beetles, a species complex of endangered insects (Coleoptera: Scarabaeidae, Cetoniinae, Osmoderma). *Journal of Zoological Systematics and Evolutionary Research*, *47*, 88–95. https://doi.org/10.1111/j.1439-0469.2008.00475.x

Avtzis, D. N., Bertheau, C., & Stauffer, C. (2012). What is next in bark beetle phylogeography? *Insects*, *3*, 453–472. https://doi.org/10.3390/insects3020453

Blouin, M. S., Yowell, C. A., Courtney, C. H., & Dame, J. B. (1995). Host movement and the genetic structure of populations of parasitic nematodes. *Genetics*, *141*, 1007–1014.

Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, *24*, 621–631. https://doi.org/10.1093/molbev/msl191

Cherepanov, A. I. (1981). *Cerambycidae of Northern Asia (Cerambycinae). [Usachi Severnoi Azii (Cerambycinae)]*. Novosibirsk: Nauka.

Cizek, L., Schlaghamersky, J., Borucky, J., Hauck, D., & Helesic, J. (2009). Range expansion of an endangered beetle: Alpine longhorn *Rosalia alpina* (Coleoptera: Cerambycidae) spreads to the lowlands of Central Europe. *Entomologica Fennica*, *20*, 200–206.

Clement, M., Posada, D., & Crandall, K. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, *9*, 1657–1660. https://doi.org/10.1046/j.1365-294x.2000.01020.x

Comps, B., Gömöry, D., Letouzey, J., Thiébaut, B., & Petit, R. J. (2001). Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics*, *157*, 389–397.

Corander, J., & Marttinen, P. (2006). Bayesian identification of admixture events using multi-locus molecular markers. *Molecular Ecology*, *15*, 2833–2843. https://doi.org/10.1111/j.1365-294X.2006.02994.x

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, *9*, 772. https://doi.org/10.1038/nmeth.2109

De Lafontaine, G., Amasifuen Guerra, C. A., Ducousso, A., & Petit, R. J. (2014). Cryptic no more: Soil macrofossils uncover Pleistocene forest microrefugia within a periglacial desert. *New Phytologist*, *204*, 715–729. https://doi.org/10.1111/nph.12833

De Lafontaine, G., Ducousso, A., Lefevre, S., Magnanou, E., & Petit, R. J. (2013). Stronger spatial genetic structure in recolonized areas than in refugia in the European beech. *Molecular Ecology*, *22*, 4397–4412. https://doi.org/10.1111/mec.12403

Demesure, B., Comps, B., & Petit, J. (1996). Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution*, *50*, 2515–2520. https://doi.org/10.1111/j.1558-5646.1996.tb03638.x

Drag, L., Hauck, D., Berces, S., Michalcewicz, J., Jelaska-Šerić, L., Aurenhammer, S., & Cizek, L. (2015). Genetic differentiation of populations of the threatened saproxylic beetle Rosalia longicorn, *Rosalia alpina* (Coleoptera, Cerambycidae) in Central and Southeast Europe. *Biological Journal of the Linnean Society*, *116*, 911–925. https://doi.org/10.1111/bij.12624

Drag, L., Zima, J. Jr., & Cizek, L. (2013). Characterization of nine polymorphic microsatellite loci for a threatened saproxylic beetle *Rosalia alpina* (Coleoptera: Cerambycidae). *Conservation Genetics Resources*, *5*, 903–905. https://doi.org/10.1007/s12686-013-9929-1

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*, 1969–1973. https://doi.org/10.1093/molbev/mss075

Dupanloup, I., Schneider, S., & Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, *11*, 2571–2581. https://doi.org/10.1046/j.1365-294X.2002.01650.x

Earl, D. A., & vonHoldt, B. M. (2012). Structure Harvester: A website and program for visualizing Structure output and implementing the Evanno method. *Conservation Genetetics Resources*, *4*, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software Structure: A simulation study. *Molecular Ecology*, *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fox, L. R., & Morrow, P. A. (1981). Specialization: Species property or local phenomenon? *Science*, *211*, 887–893. https://doi.org/10.1126/science.211.4485.887

Frankham, R., Ballou, J. D., & Briscoe, D. A. (2010). *Introduction to conservation genetics* (2nd ed.). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511809002

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, *147*, 915–925.

Giesecke, T., Brewer, S., Finsinger, W., Leydet, M., & Bradshaw, R. H. W. (2017). Patterns and dynamics of European vegetation change over the last 15,000 years. *Journal of Biogeography*, *44*, 1441–1456. https://doi.org/10.1111/jbi.12974

Gómez, A., & Lunt, D. H. (2007). Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. In S. Weiss & N. Ferrand (Eds.), *Phylogeography of southern European refugia* (pp. 155–188). Dordrecht, the Netherlands: Springer. https://doi.org/10.1007/1-4020-4904-8

Gömöry, D., & Paule, L. (2010). Reticulate evolution patterns in western-Eurasian beeches. *Botanica Helvetica*, *120*, 63–74. https://doi.org/10.1007/s00035-010-0068-y

Greuter, W., & Burdet, H. M. (1981). Fagus sylvatica subsp. orientalis. In W. Greuter & T. Raus (Eds.), *Med-checklist notulae*, 4. *Willdenowia 11*, 271–280.

Grimm, G. W., & Denk, T. (2014). The Colchic region as refuge for relict tree lineages: Cryptic speciation in field maples. *Turkish Journal of Botany*, *38*, 1050–1066. https://doi.org/10.3906/bot-1403-87

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, *52*, 696–704.

Habel, J. C., Lens, L., Rödder, D., & Schmitt, T. (2011). From Africa to Europe and back: Refugia and range shifts cause high genetic differentiation in the Marbled White butterfly *Melanargia galathea*. *BMC Evolutionary Biology*, *11*, 215. https://doi.org/10.1186/1471-2148-11-215

Hamilton, A. J., Novotny, V., Waters, E. K., Basset, Y., Benke, K. K., Grimbacher, P. S., … Stork, N. E. (2013). Estimating global arthropod species richness: Refining probabilistic models using probability bounds analysis. *Oecologia*, *171*, 357–365. https://doi.org/10.1007/s00442-012-2434-5

Hampe, A., & Petit, R. J. (2005). Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, *8*, 461–467. https://doi.org/10.1111/j.1461-0248.2005.00739.x

Heuertz, M., Carnevale, S., Fineschi, S., Sebastiani, F., Hausman, J. F., Paule, L., & Vendramin, G. G. (2006). Chloroplast DNA phylogeography of European ashes, *Fraxinus* sp. (Oleaceae): Roles of hybridization and life history traits. *Molecular Ecology*, *15*, 2131–2140. https://doi.org/10.1111/j.1365-294X.2006.02897.x

Hovorka, W. (2011). *Nachweise des Alpenbocks (Rosalia alpina*, L.) bei Rabensburg im Weinviertel (Coleoptera: Cerambycidae). Beiträge zur Entomofaunistik 12 Kurzmitteilungen. Wien: Österreichische Gesellschaft fur Entomofaunistik.

Jaenike, J. (1990). Host specialization in phytophagous insects. *Annual Review of Ecology, Evolution, and Systematics*, *21*, 243–73. https://doi.org/10.1146/annurev.es.21.110190.001331

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The K = 2 conundrum. *Molecular Ecology*, *26*, 3594–3602. https://doi.org/10.1111/mec.14187

Jermy, T., & Szentesi, A. (2003). Evolutionary aspects of host plant specialisation—A study on bruchids (Coleoptera: Bruchidae). *Oikos*, *101*, 196–204. https://doi.org/10.1034/j.1600-0706.2003.11918.x

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*, 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Lindhe, A., Jeppsson, T., & Ehnström, B. (2011). *Longhorn beetles in Sweden—Changes in distribution and abundance over the last two hundred years*. Färjestaden: Sveriges Entomologiska Förening.

Magri, D., Vendramin, G. G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., … de Beaulieu, J.-L. (2006). A new scenario for the Quaternary history of European beech populations: Paleobotanical evidence and genetic consequences. *New Phytologist*, *171*, 199–221. https://doi.org/10.1111/j.1469-8137.2006.01740.x

Magyari, E. K., Kuneš, P., Jakab, G., Sümegi, P., Pelánková, B., Schäbitz, F., … Chytrý, M. (2014). Late Pleniglacial vegetation in eastern-central Europe: Are there modern analogues in Siberia? *Quaternary Science Reviews*, *95*, 60–79. https://doi.org/10.1016/j.quascirev.2014.04.020

Merkl, O., Hegyessy, G., & Kovács, T. (1996). Cerambycidae (Coleoptera) from the Bükk National Park. In S. Mahunka (Eds.), *The Fauna of the Bükk National Park*, *II* (p. 655). Budapest, Hungary: Hungarian Natural History Museum.

Michalcewicz, J., & Ciach, M. (2012). Biometry of adult rosalia longicorn *Rosalia alpina* (L.)(Coleoptera: Cerambycidae) from the Polish Carpathians: A preliminary study. *Polish Journal of Entomology*, *81*, 311–320. https://doi.org/10.2478/v10200-012-0011-1

Michalcewicz, J., & Ciach, M. (2015). Current distribution of the Rosalia longicorn *Rosalia alpina* (LINNAEUS, 1758) (Coleoptera: Cerambycidae) in Poland. *Polish Journal of Entomology*, *84*, 9–20. https://doi.org/10.1515/pjen-2015-0002

Murphy, S. M. (2004). Enemy-free space maintains swallowtail butterfly host shift. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 18048–18052. https://doi.org/10.1073/pnas.0406490102

Myking, T., & Yakovlev, I. (2006). Variation in leaf morphology and chloroplast DNA in *Ulmus glabra* in the northern suture zone: Effects of distinct glacial refugia. *Scandinavian Journal of Forest Research*, *21*, 99–107. https://doi.org/10.1080/02827580500539265

Nadler, S. (1995). Microevolution and the genetic structure of parasite populations. *Journal of Parasitology*, *81*, 395–403. https://doi.org/10.2307/3283821

Nakamine, H., & Takeda, M. (2008). Molecular phylogenetic relationships of flightless beetles belonging to the genus Mesechthistatus Breuning, (Coleoptera: Cerambycidae) inferred from mitochondrial COI gene sequences. *Journal of Insect Science*, *8*, 1–11. https://doi.org/10.1673/031.008.7001

Nei, M., & Chesser, R. K. (1983). Estimation of fixation indices and gene diversities. *Annals of Human Genetics*, *47*, 253–259. https://doi.org/10.1111/j.1469-1809.1983.tb00993.x

Nieberding, C. M., & Morand, S. (2006). Comparative phylogeography: The use of parasites for insights into host history. In S. Morand, B. R. Krasnov, & R. Poulin (Eds.), *Micromammals and Macroparasites* (pp. 277–293). New York: Springer. https://doi.org/10.1007/978-4-431-36025-4

Nieberding, C., Morand, S., Libois, R., & Michaux, J. R. (2004). A parasite reveals cryptic phylogeographic history of its host. *Proceedings of the Royal Society of London*, *271*, 2559–2568. https://doi.org/10.1098/rspb.2004.2930

Novotny, V., Miller, S. E., Baje, L., Balagawi, S., Basset, Y., Cizek, L., … Weiblen, G. D. (2010). Guild-specific patterns of species richness and host specialization in plant-herbivore food webs from a tropical forest. *Journal of Animal Ecology*, *79*, 1193–1203. https://doi.org/10.1111/j.1365-2656.2010.01728.x

Pabijan, M., Zieliński, P., Dudek, K., Chloupek, M., Sotiropoulos, K., Liana, M., & Babik, W. (2015). The dissection of a Pleistocene refugium: Phylogeography of the smooth newt, *Lissotriton vulgaris*, in the Balkans. *Journal of Biogeography*, *42*, 671–683. https://doi.org/10.1111/jbi.12449

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*, 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pearse, D. E., & Crandall, K. A. (2004). Beyond FST: Analysis of population genetic data for conservation. *Conservation Genetics*, *5*, 585–602. https://doi.org/10.1007/s10592-003-1863-4

Phuekvilai, P. (2014). Relicts, refugia and reticulation: A study of population history, hybrids and phylogeny in the long-lived flowering tree Genus Tilia. Ph.D. thesis. School of Biology, Newcastle University, Newcastle upon Tyne, UK.

Picard, F. (1929). *Faune de France*. Paris: Coleopteres. Cerambycidae.

Pinceel, J., Jordaens, K., Pfenninger, M., & Backeljau, T. (2005). Range-wide phylogeography of a terrestrial slug in Europe: Evidence for Alpine refugia and rapid colonization after the Pleistocene glaciations. *Molecular Ecology*, *14*, 1133–1150. https://doi.org/10.1111/j.1365-294X.2005.02479.x

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

Rambaut, A. (2014). *FigTree v1.4.2: Tree figure drawing tool*. Retrieved from http://tree.bio.ed.ac.uk/software/figtree/.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018) Tracer v1.7. Retrieved from http://beast.community/tracer.

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, *4*, 137–138.

Rousset, F. (2008). Genepop'007: A complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, *8*, 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Santucci, F., Emerson, B. C., & Hewitt, G. M. (1998). Mitochondrial DNA phylogeography of European hedgehogs. *Molecular Ecology*, *7*, 1163–1172. https://doi.org/10.1046/j.1365-294x.1998.00436.x

Schmitt, T. (2007). Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, *4*, 11. https://doi.org/10.1186/1742-9994-4-11

Schmitt, T., & Varga, Z. (2012). Extra-Mediterranean refugia: The rule and not the exception? *Frontiers in Zoology*, *9*, 22. https://doi.org/10.1186/1742-9994-9-22

Scriber, J. M. (1986). Origins of the regional feeding abilities in the tiger swallowtail butterfly: Ecological monophagy and the Papilio glaucus

australis subspecies in Florida. *Oecologia*, *71*, 94–103. https://doi.org/10.1007/BF00377326

Scriber, J. M., Allen, G. R., & Walker, P. W. (2006). Ecological monophagy in Tasmanian *Graphium macleayanum moggana* with evolutionary reflections of ancient Angiosperm hosts. *Insect Science*, *13*, 325–334.

Shapovalov, A. M. (2012). The longicorn beetles of Orenburg Province: Fauna, distribution, and bionomy. *Archives of Orenburg Branch of Russian Entomological Society*, *3*, 1–224.

Solano, E., Mancini, E., Ciucci, P., Mason, F., Audisio, P., & Antonini, G. (2013). The EU protected taxon *Morimus funereus* Mulsant, 1862 (Coleoptera: Cerambycidae) and its western Palaearctic allies: Systematics and conservation outcomes. *Conservation Genetics*, *14*, 683–694. https://doi.org/10.1007/s10592-013-0461-3

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*, 585–595.

Takezaki, N., & Nei, M. (2008). Empirical tests of the reliability of phylogenetic trees constructed with microsatellite DNA. *Genetics*, *178*, 385–392. https://doi.org/10.1534/genetics.107.081505

Takezaki, N., Nei, M., & Tamura, K. (2010). POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows-interface. *Molecular Biology and Evolution*, *27*, 747–752. https://doi.org/10.1093/molbev/msp312

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, *30*, 2725–2729. https://doi.org/10.1093/molbev/mst197

Tinner, W., & Lotter, A. F. (2006). Holocene expansions of *Fagus silvatica* and *Abies alba* in Central Europe: Where are we after eight decades of debate? *Quaternary Science Reviews*, *25*, 526–549. https://doi.org/10.1016/j.quascirev.2005.03.017

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

Whiteley, R. E. (2004). Quantitative and molecular genetic variation in *Ulmus laevis* Pall. Doctor's dissertation. Swedish University of Agricultural Sciences, Uppsala, Sweden.

DATA ACCESSIBILITY

All unique sequences were submitted to GenBank under accession numbers: MH795309-MH795323. Microsatellite loci newly developed for this study can be found under GenBank accession numbers: MH795324-MH795329.

BIOSKETCH

Lukas Drag is broadly interested in the conservation of saproxylic beetles. This paper represents one chapter in his PhD work at the University of South Bohemia on the phylogeography and conservation genetics of endangered saproxylic beetles in Europe. LD and the other authors are members of a research team focused on various aspects of the ecology, biology and diversity of insects associated with woodlands (see www.oldtree.cz).

Author contributions: L.C. and L.D. designed the study, L.D. performed the research, D.H., D.F.S., R.G., G.C. contributed samples, T.S. and O.R. provided useful comments, L.C. and L.D. wrote the paper.

SUPPORTING INFORMATION

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

How to cite this article: Drag L, Hauck D, Rican O, et al. Phylogeography of the endangered saproxylic beetle Rosalia longicorn, *Rosalia alpina* (Coleoptera, Cerambycidae), corresponds with its main host, the European beech (*Fagus sylvatica*, Fagaceae). *J Biogeogr*. 2018;45:2631–2644. https://doi.org/10.1111/jbi.13429