

ORIGINAL ARTICLE

Complex postglacial history of the temperate bark beetle *Tomicus piniperda* L. (Coleoptera, Scolytinae)A Horn^{1,4}, C Stauffer², F Lieutier¹ and C Kerdelhué³¹Université d'Orléans, Laboratoire de Biologie des Ligneux et des Grandes Cultures UPRES EA 1207, Orléans, France; ²Department of Forest and Soil Sciences, Institute of Forest Entomology, Forest Pathology and Forest Protection, Boku, University of Natural Resources and Applied Life Sciences, Vienna, Austria and ³INRA, UMR1202 BIOGECO, Cestas, France

Tomicus piniperda is an economically important pine bark beetle infesting European *Pinus* spp. stands. We sequenced and analyzed 797 bp of the mitochondrial genome from individuals obtained from 34 populations sampled throughout the European range. We obtained 36 haplotypes, from which a haplotype network was constructed. In the Iberian Peninsula, high-genetic variability was detected with numerous endemic haplotypes. In contrast, the other European populations were less diverse with a single haplotype predominating from the Pyrenees to Scandinavia. Nevertheless, even within Europe, a few populations showed significant amounts of diversity. Four groups were obtained by Spatial Analysis of Molecular Variance, illustrating the regional characteristics of the species. *T. piniperda* had multiple fragmented refugia in the Iberian

Peninsula. These currently isolated populations only partly contributed to postglacial re-colonizations of Northern Europe during interglacials. Nevertheless, few long-range migration events up to Northern Europe were detected, mostly originating from the Pyrenees. In the rest of Europe, the phylogeographical patterns were unclear, because of repeated cycles of contraction and expansion. The genetic analysis showed one glacial refugium in North-Central Europe, whereas other refugia most likely occurred in the Southern Alps, Apennine and the Balkans. The phylogeographical pattern depicted here reflects partly the postglacial history of the beetles' main host tree *P. sylvestris*.

Heredity (2009) **103**, 238–247; doi:10.1038/hdy.2009.48; published online 29 April 2009

Keywords: phylogeography; mitochondrial DNA; Europe; glacial refugia; *Tomicus piniperda*; *Pinus* species

Introduction

Quaternary climatic oscillations have had dramatic effects on the evolution of species. The contemporary distribution of genetic diversity cannot be understood without studying how organisms responded to climatic history in geological times (Hewitt, 2000). The distribution ranges of temperate species were restricted during glacial maxima to a few glacial refugia and the organisms re-colonized northwards during interglacial periods of temperature amelioration. These re-colonization routes were often blocked by geographical barriers or by expansion routes from other lineages (Hewitt, 1996). The suture zones formed by lineages originating from different refugia and coming into secondary contact are considered to show higher genetic diversity than other geographic regions (Petit *et al.*, 2003). In most cases, refugial areas were localized in the southernmost regions (for example, Schmitt, 2007). Yet, cold-tolerant species may also have survived in northern refugia (Stewart and

Lister, 2001; Hewitt, 2004), such as the Alps, or central, eastern and northern Europe, leading to complex postglacial patterns (for example, Ursenbacher *et al.*, 2006). Moreover, some species strictly depend on other organisms (hosts, mutualists, symbionts, and so on) for their development or dispersal. In that case, the phylogeographic pattern of the dependent species can be influenced by that of its partner, and a certain level of similarity may be highlighted (see for instance Burban *et al.*, 1999; Burban and Petit, 2003). However, this is not always the case and the genetic structure of insects does not necessarily reflect that of its host (Stauffer *et al.*, 1999; Sallé *et al.*, 2007).

The pine shoot beetle *Tomicus piniperda* (L.) (Coleoptera: Scolytinae) has a Palearctic distribution (Balachowsky, 1949). The exact eastern limits of its geographical distribution are still unclear. It develops on various pine species but is mainly found on *Pinus sylvestris* in Europe (Postner, 1974; Pfeffer, 1994). Moreover, its present distribution and host association suggest that *T. piniperda* survives poorly on Mediterranean pine species such as *P. halepensis* and *P. pinea*, and that it cannot develop in warm and dry environments (Gallego *et al.*, 2004). The life cycle of this species is characterized by two phases of dispersal and one over-wintering period (Långström, 1983). The first dispersal phase occurs during the reproductive process, when adults fly to an attractive host and oviposit in the inner bark. The second dispersal

Correspondence: Dr A Horn, Department of Ecology and Evolution, University of Lausanne, Biophore Building, Lausanne CH-1015, Switzerland.

E-mail: agneshorn@gmail.com

⁴Current address: Department of Ecology and Evolution, Biophore Building, University of Lausanne, CH-1015 Lausanne, Switzerland.

Received 13 October 2008; revised 16 March 2009; accepted 23 March 2009; published online 29 April 2009

phase takes place after nymphosis, when young adults fly to the shoots of the host plant for maturation feeding.

As *T. piniperda* has an obligate relationship with a pine host for both reproduction and maturation feeding, we hypothesized that its phylogeographic pattern is closely linked to the past history of its main hosts. There is evidence that *P. sylvestris* had glacial refugia in Central Europe and in the Alps, whereas different pine species were also present in the Iberian Peninsula and Italy during the coldest period (Sinclair *et al.*, 1999; Soranzo *et al.*, 2000; Burbán and Petit, 2003; Bucci *et al.*, 2007). We thus expect that *T. piniperda* exhibits a complex contemporary genetic structure. A phylogeographic study with a restricted number of populations suggested that the evolutionary history of *T. piniperda* reflects that of *P. sylvestris* with both southern and northern glacial refugia (Ritzlerow *et al.*, 2004). However, samples were lacking from the southern range and only one population was analyzed from the Iberian Peninsula. Thus, the recent past history of *T. piniperda* could not be solved.

Here we present a phylogeographic study of *T. piniperda* in its known distribution range in Europe using a mitochondrial marker. The objectives of our work were: (i) to reconstruct the phylogeography of *T. piniperda* with representative sampling of the species; (ii) to confirm the presence of northern refugia during the glaciations; and (iii) to infer if the past history of this cold-tolerant species was related to that of its pine hosts.

Materials and methods

Beetle sampling

Beetles were collected throughout Europe from 1999 to 2004. Thirty-four populations were found in sixteen countries and on six *Pinus* species, to complement the existing samples from France (Kerdelhué *et al.*, 2002) and to obtain a representative sampling of the species across its distribution range in Europe. No insects were trapped in the southern part of the Iberian Peninsula, of Italy and of the Balkans, where only the congeneric *T. destruens* was found (Figure 1). Samples were stored in absolute ethanol. All relevant information is gathered in Table 1 and the localities are shown in Figure 1.

DNA extraction

DNA from the imago stage was extracted from head, thorax and legs, whereas DNA was extracted from the entire body for larval or pupal samples. The abdomen, elytra and antennae of adults were kept as vouchers in ethanol. DNA was extracted and isolated with the GenElute Mammalian Genomic DNA miniprep kit (Sigma, St Louis, MO, USA).

Mitochondrial DNA amplification

T. piniperda specific primers were used to amplify a partial region of the cytochrome oxidase I and II genes (Kerdelhué *et al.*, 2002). The annealing temperature was 50 °C, and a total of 30 cycles of amplification was carried out in 50 µl reaction volume. PCR products were purified using the GenElute PCR Clean-Up kit (Sigma) and direct sequencing was carried out systematically with both PCR primers using the BigDye Terminator sequencing kit (PE Applied Biosystems, Foster City, CA, USA) and carried out with an ABI 3100 automatic sequencer.

All sequences were carefully checked by hand before analysis.

Data analysis of mitochondrial sequences

All the obtained sequences, and sequences from France (GenBank accession numbers AF457799-AF457803, AF457804-AF457808, AF457819-AF457820, AF457822-AF457826; Kerdelhué *et al.*, 2002), were aligned using ClustalW v1.4 (Thompson *et al.*, 1994) as implemented in BIOEDIT v7.0.5.

A statistical parsimony network was computed using TCS v1.21 (Clement *et al.*, 2000), which estimates genes genealogies from DNA sequences following the method described in Templeton *et al.* (1992). We used topological and frequency criteria (Crandall and Templeton, 1993; see also Pfenninger and Posada, 2002) to solve the few cladogram ambiguities that occurred.

Allelic richness R was computed after rarefaction to three individuals (Petit *et al.*, 1998) using CONTRIB, for all populations including at least three individuals. Gene diversity H and within population mean number of pairwise differences π were calculated using ARLEQUIN v3.1 (Excoffier *et al.*, 2005).

We used a spatial analysis of variance (SAMOVA 1.0, Dupanloup *et al.*, 2002) to identify groups of population that are geographically homogeneous and maximally differentiated. By using a simulated annealing procedure, the program maximizes the proportion of total genetic variance because of differences among groups of populations (F_{CT}). The program was run for 10000 permutations from 100 random initial conditions for two to 15 differentiated groups ($K = 2$ to $K = 15$).

To test for a host effect on the distribution of genetic diversity, we carried out an Analysis of Molecular Variance (AMOVA, Excoffier *et al.*, 1992) on the whole data set using ARLEQUIN v3.1. Populations were grouped according to the pine species from which they were sampled (see Table 1).

Occurrence of a significant phylogeographic structure was assessed by testing if G_{st} (coefficient of genetic variation over all populations) was significantly smaller than N_{st} (equivalent coefficient taking into account the similarities between haplotypes) by the use of 1000 permutations (see Pons and Petit, 1996) in the program PERMUT. Pairwise G_{st} and N_{st} were calculated using DISTON, and Nei's average number of differences between populations (corrected Nei's D ; Nei, 1975) was calculated using ARLEQUIN. The geographic distances were computed between all sampling locations using the geographic coordinates (see <http://jan.ucc.nau.edu/~cvm/latlongdist.html>). To detect isolation by distance, matrices of genetic distances (pairwise G_{st} and N_{st} , and corrected Nei's D) were compared with the matrix of geographic distances by means of a simple Mantel test (Legendre and Legendre, 1998) using the R software (R Development Core Team, 2005). We used 999 random permutations to test the significance of the Mantel test statistic. Correlations between intra-population parameters (H and π) and latitude were assessed by means of linear regressions (using R). These tests were carried out to assess if genetic diversity significantly decreased with latitude. CONTRIB, PERMUT and DISTON are available at <http://www.pierroton.inra.fr/genetics/labo/Software/>. Both Mantel tests and diversity analyses were carried out

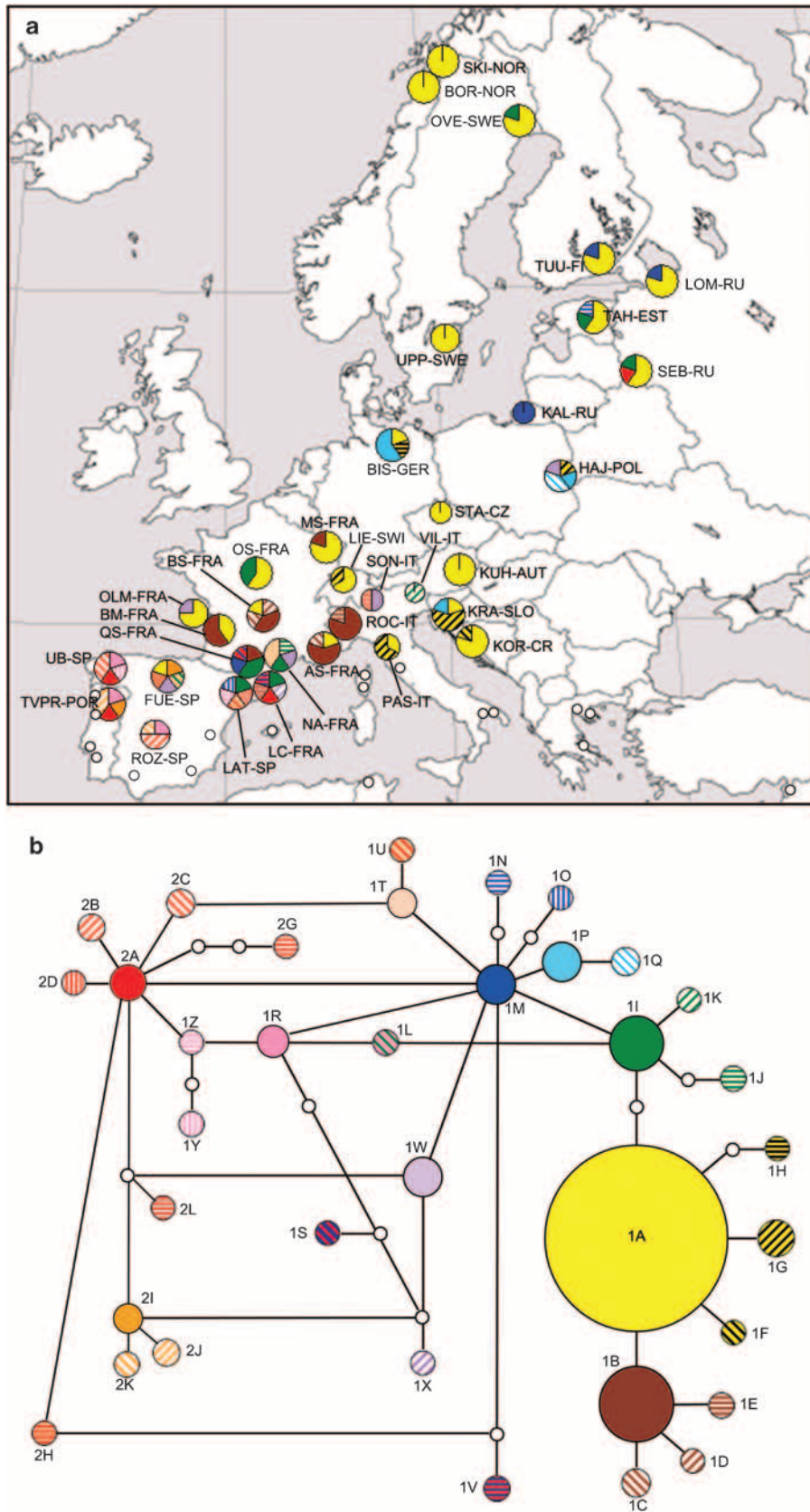


Figure 1 Geographical distribution of *T. piniperda* populations and their associated haplotypes of the cytochrome oxidase I and II sequences. (a) Geographic distribution of the 34 populations represented with the proportion of different haplotypes for each population. Small open circles show unsuccessful trappings. Codes for the localities are given in Table 1. Color codes refer to the color used in the haplotype network. (b) Haplotype network of the 36 haplotypes. Each line corresponds to a mutational step and each empty circle to a missing intermediate. Haplotype frequencies are represented by the size of the circle.

Table 1 Sampling sites, date of capture, host tree, geographic coordinates and abbreviations of *T. piniperda*'s populations

Code	Country	Location	Host species	Date	Collected by	Latitude	Longitude
KUH-AU	Austria	Kühnsdorf	<i>P. sylvestris</i>	06/2004	C Stauffer	46°37' N	14°36' E
KOR-CR	Croatia	Korenica	<i>P. sylvestris</i>	04/2002	B Hrasovec	44°46' N	13°41' E
STA-CZ	Czech Republic	Stara Boleslav	<i>P. sylvestris</i>	—	M Knizek	50°12' N	14°44' E
TAH-EST	Estonia	Tahkurana	<i>P. sylvestris</i>	07/2004	M Robba	58°24' N	24°32' E
TUU-FI	Finland	Tuusala	<i>P. sylvestris</i>	07/2002	C Stauffer	60°24' N	25°01' E
NA-FRA	France	Narbonne	<i>P. halepensis</i>	03/2000	J Garcia	43°11' N	2°52' E
OS-FRA	France	Orléans	<i>P. sylvestris</i>	03/2000	Kerdelhué et al. 2002	47°51' N	1°56' E
MS-FRA	France	Mulhouse	<i>P. sylvestris</i>	03/2000	Kerdelhué et al. 2002	47°44' N	7°21' E
AS-FRA	France	St André les Alpes	<i>P. sylvestris</i>	03/2000	J Garcia	43°96' N	6°30' E
QS-FRA	France	Quillan	<i>P. sylvestris</i>	03/2000	Kerdelhué et al. 2002	43°20' N	2°15' E
LC-FRA	France	Mont Louis	<i>P. uncinata</i>	03/2000	Kerdelhué et al. 2002	42°30' N	2°07' E
BS-FRA	France	Brioude	<i>P. sylvestris</i>	03/2000	J Garcia	45°17' N	3°22' E
BM-FRA	France	Bordeaux	<i>P. pinaster</i>	03/2000	J Garcia	44°50' N	0°34' W
OLM-FRA	France	Oléron	<i>P. pinaster</i>	09/2002	J Garcia	45°55' N	1°18' W
BIS-GER	Germany	Bispingen	<i>P. sylvestris</i>	06/2004	C Stauffer	53°05' N	10°00' E
ROC-IT	Italy	Roccia Melone	<i>P. sylvestris</i>	05/2002	M Faccoli	45°06' N	8°08' E
PAS-IT	Italy	Passo del Bocco	<i>P. nigra</i>	04/2002	M Faccoli	44°30' N	9°04' E
SON-IT	Italy	Sonico	<i>P. sylvestris</i>	11/2002	M Faccoli	46°09' N	10°20' E
VIL-IT	Italy	Villasantina	<i>P. sylvestris</i>	03/2001	M Faccoli	46°25' N	12°55' E
SKI-NOR	Norway	Skril	<i>P. sylvestris</i>	07/2004	C Stauffer	68°08' N	16°00' E
BOR-NOR	Norway	Borkam	<i>P. sylvestris</i>	07/2004	C Stauffer	67°15' N	15°24' E
HAJ-POL	Poland	Hajno' Wka	<i>P. sylvestris</i>	07/2004	C Stauffer	52°45' N	23°36' E
TVPR-POR	Portugal	Vila Real	<i>P. pinaster</i>	09/2002	T Vasconcelos	41°30' N	7°33' W
PET-RU	Russia	Lomonosor	<i>P. sylvestris</i>	—	M Mandelshtam	59°54' N	29°44' E
KAL-RU	Russia	Zelenogradsk	<i>P. sylvestris</i>	—	M Mandelshtam	54°57' N	20°30' E
SEB-RU	Russia	Sebez	<i>P. sylvestris</i>	—	M Mandelshtam	56°08' N	28°39' E
KRA-SLO	Slovenia	Kranj-Mlaka	<i>P. sylvestris</i>	05/2002	R Pavlin	46°15' N	14°21' E
UB-SP	Spain	Uriz-Begonte	<i>P. radiata</i>	01/2003	M Lombardero	43°08' N	7°39' W
FUE-SP	Spain	Vitoria	<i>P. sylvestris</i>	03/2002	D Gallego	42°47' N	02°56' W
ROZ-SP	Spain	Segovia	<i>P. sylvestris</i>	04/2002	D Gallego	40°53' N	4°00' W
LAT-SP	Spain	La Torre de Cadi	<i>P. sylvestris</i>	05/2002	D Gallego	42°32' N	1°51' E
UPP-SWE	Sweden	Uppsala	<i>P. sylvestris</i>	06/1998	C Stauffer	59°52' N	17°38' E
OVE-SWE	Sweden	Overkalix	<i>P. sylvestris</i>	07/2004	C Stauffer	66°21' N	22°56' E
LIE-SWI	Switzerland	Liesberg	<i>P. sylvestris</i>	05/2002	M Kenis	47°38' N	7°42' E

Collectors' names are also listed.

on the whole data set and among the two major groups defined by SAMOVA (see below).

The demographic history of populations belonging to the different geographical groups defined by SAMOVA was inferred using different methods. Owing to the insufficient sample sizes, only the two major groups were analyzed (see below). First, mismatch distributions of the pairwise genetic differences (Rogers and Harpending, 1992) within each of the two geographical groups were conducted using ARLEQUIN and their goodness-of-fit to a sudden expansion model was tested using parametric bootstrap approaches (1000 replicates). The sum of squared deviations (SSD) between the observed and expected mismatch distributions was used to assess the significance of the test. Pairwise differences typically form two main patterns, that is, multimodal distributions which are consistent with demographic stability, and unimodal distributions which reflect recent population expansion (Slatkin and Hudson, 1991). In addition, the significance of population expansion was tested using Tajima's *D* (Tajima, 1989) and Fu's *F*_s neutrality tests (Fu, 1997) implemented in ARLEQUIN, and the Ramos-Onsins and Rozas's *R*₂ statistics (Ramos-Onsins and Rozas, 2002) carried out using DnaSP version 4.20 (Rozas et al., 2003). Mismatch analyses were also used to estimate the approximate timing of expansion of *T. piniperda*'s populations within the geographical groups defined by the SAMOVA. We used the relationship $\tau = 2ut$ (Rogers and Harpending, 1992), τ being the age of expansion

measured in units of mutational time, *t* the expansion time in number of generations, and *u* the mutation rate per sequence and per generation. This last value was calculated using the relationship $u = 2\mu k$, with μ the mutation rate per nucleotide and *k* the length of the sequence in nucleotides. The 2.3% pairwise sequence divergence for arthropods mitochondrial genes defined by Brower (1994) was used to approximate μ as 1.15 per million years per nucleotide.

Results

Mitochondrial variability

We obtained a final alignment of 150 sequences that were 797 bp long, including 463 bp in COI, 69 bp tRNA Leu and 265 bp in COII.

A total of 36 haplotypes (HT) were identified with 35 polymorphic sites and were named 1A, 1B, 1C, ... 1Z, 2A, 2B, 2C, 2D, 2G, ... - 2L (Table 2). Haplotype sequences are available in GenBank under accession numbers FJ619352-FJ619381, FJ619384-FJ619389. Haplotype 1A was shared by 63 individuals mostly outside the Iberian Peninsula and Italy (Figure 1). Haplotype 1B was found in 14 individuals mainly from France and Italy, haplotype 1I was shared by 10 individuals distributed from Pyrenees to Sweden, and haplotype 1G was found for seven individuals in the Eastern part of the distribution range. All other haplotypes were shared by a

Table 2 Haplotypes (HT) found in each population and population parameters

Code	N	# HT	Haplotypes	H	R	π
KUH-AU	5	1	1A(5)	0	0	0
KOR-CR	5	2	1A(4), 1F	0.4	0.6	0.4
STA-CZ	2	1	1A(2)	—	—	—
TAH-EST	5	3	1A(3), 1I, 1N	0.7	1.2	2.4
TUU-FI	5	2	1A(4), 1M	0.4	0.6	2
NA-FRA	5	4	1J, 1W, 1I, 1T(2)	0.9	1.7	3
OS-FRA	5	2	1A(3), 1I(2)	0.6	0.9	1.2
MS-FRA	5	2	1A(4), 1B	0.4	0.6	0.4
AS-FRA	5	3	1A, 1B(3), 1C	0.7	1.2	0.8
QS-FRA	5	4	1B, 1I(2), 1M, 1S	0.9	1.7	4.2
LC-FRA	5	5	1I, 1V, 1X, 2A, 2L	1	2	3.4
BS-FRA	5	4	1A, 1B(2), 1C, 1D	0.9	1.7	1.2
BM-FRA	5	2	1A(2), 1B(3)	0.6	0.9	0.6
OLM-FRA	4	2	1A(3), 1W	0.5	0.75	2
BIS-GER	5	3	1A, 1H, 1P(3)	0.7	1.2	2.8
ROC-IT	5	2	1B(4), 1E	0.4	0.6	0.4
PAS-IT	3	2	1A, 1G(2)	0.67	1	1.33
SON-IT	2	2	1W, 2G	—	—	—
VIL-IT	1	1	1K	—	—	—
SKI-NOR	5	1	1A(5)	0	0	0
BOR-NOR	5	1	1A(5)	0	0	0
HAJ-POL	5	4	1G, 1P, 1Q(2), 1W	0.9	1.7	2.8
TVPR-POR	5	4	1R, 2A, 2I, 2J(2)	0.9	1.7	2.8
LOM-RU	5	2	1A(4), 1M	0.4	0.6	2
KAL-RU	2	1	1M(2)	—	—	—
SEB-RU	5	3	1A(3), 1I, 2A	0.7	1.2	2
KRA-SLO	5	3	1A, 1G(3), 1P	0.7	1.2	2.2
UB-SP	5	4	1R, 1Y, 2A, 2C(2)	0.9	1.7	2.4
FUE-SP	5	5	1A, 1L, 1W, 2H, 2I	1	2	3.4
ROZ-SP	4	3	1R, 2B(2), 2K	0.83	1.5	3.16
LAT-SP	5	5	1I, 1O, 1U, 1Z, 2D	1	2	2.6
UPP-SWE	4	1	1A(4)	0	0	0
OVE-SWE	5	2	1A(4), 1I	0.4	0.6	0.8
LIE-SWI	3	2	1A(2), 1G	0.67	1	1.33

N, number of sequenced individuals; H, gene diversity; R, allelic richness after rarefaction to three; π , nucleotide diversity. Codes for populations are given in Table 1. Number in bracket after haplotype name is the number of individuals with that haplotype. H, R and π were not computed for populations SON-IT, VIL-IT, STA-CZ and KAL-RU because the number of individuals was too low.

maximum of five individuals. Further, 25 private haplotypes (that is, haplotypes found in one population only) were identified, 17 of which were located in the Iberian Peninsula or the Pyrenees, four in Italy or Croatia, one in France and three in Northern latitudes (Germany, Poland and Estonia). Moreover, thirteen haplotypes (1L, 1O, 1R, 1U, 1Y, 1Z, 2B to 2D, and 2H to 2K) were found exclusively in the Iberian Peninsula, although in several populations.

The 36 haplotypes were joined in a single haplotype network with 95% probability (Figure 1b). The three most frequent haplotypes, namely 1A, 1B and 1I, were closely related. Each of them had 2–3 rare satellite haplotypes that diverged from the major haplotype (either 1A, 1B or 1I) by one or two mutations. The rest of the network grouped all other haplotypes without any clear structure, as many loops occurred. There were very few missing haplotypes in the entire network. It is interesting that the three major haplotypes and their satellites were mostly found outside the Iberian Peninsula, whereas the rarer haplotypes present in the rest of the network were mostly found in Spain and Portugal (Figure 1a).

Mitochondrial population genetic parameters, spatial structure and host effect

Total gene diversity H_T was 0.82, whereas the average within-population diversity H_S was 0.59. The indices of population structure G_{ST} and N_{ST} were 0.276 and 0.439, respectively, and did not differ significantly from each other, indicating a weak phylogeographical structure. For each population, gene diversity H , allelic richness R and mean number of pairwise differences π are given in Table 2.

The SAMOVA analysis did not allow to unambiguously identify the optimal K , that is, the number of groups that shows the highest F_{CT} value. F_{CT} first increased from 0.510 to 0.520, reached a local maximum for $K=4$ and slightly decreased. However, it gradually increased again to reach 0.537 for $K=15$. However, as soon as K was higher or equal to five, the new groups consisted of single populations with no strong changes to the structure identified for $K=4$. This configuration was thus retained, with $F_{CT}=0.52$ ($P<0.001$), $F_{ST}=0.56$ ($P<0.001$) and $F_{SC}=0.09$ ($P<0.001$). The four groups obtained with the SAMOVA were (see Figure 2) (i) an 'Iberian' group clustering the six populations from Spain and Portugal and the Pyrenean population of highest altitude (LC-FRA); (ii) a 'Pyrenean' group comprising QS-FRA and NA-FRA; (iii) a 'Central-European' group clustering 4 populations from Italy, Germany, Poland and Russia (SON-IT, BIS-GER, HAJ-POL and KAL-RU); (iv) a 'Main-European' group gathering all the other populations. Owing to the insufficient sample size, only Iberian and Main-European groups were studied for within-group analyses (see below).

AMOVA showed that the proportion of molecular variance explained by hosts was not significant (14.36%, $0.05<P<0.10$), whereas a significant amount of genetic diversity was found among populations within hosts (33.79%, $P<0.001$).

The Mantel test showed a significant effect of isolation by distance as the matrix of geographic distances was significantly correlated to either G_{ST} (standardized Mantel statistics $r_M=0.2798$, $P=0.002$) or N_{ST} ($r_M=0.2395$, $P=0.003$) in the whole data set, and to G_{ST} (Mantel statistics $r_M=0.1527$, $P=0.037$) and N_{ST} (Mantel statistics $r_M=0.6033$, $P=0.045$) in the Main-European and in the Iberian groups, respectively.

H and π were found to be negatively and significantly correlated with latitude in the whole data set (H : $R^2=0.430$, $P<0.001$ and π : $R^2=0.169$, $P=0.024$), whereas only H was negatively and significantly correlated with latitude in the Main-European group ($R^2=0.290$, $P<0.014$). None of the two parameters were correlated with latitude within the Iberian group.

Demographic history

Demographic inferences were carried out within the Iberian and Main-European groups found by the SAMOVA analysis. Mismatch analyses for both groups were consistent with the sudden expansion model (Main-European: $SDD=0.0007$, $P>0.10$; Iberian: $SDD=0.0079$, $P>0.10$) and showed unimodal distributions that closely fit the expected distributions (Figure 3). Both Fu's F_s statistics were negative and significant (Main-European: $F_s=-6.72$, $P<0.01$; Iberian: $F_s=-15.46$, $P<0.001$). Tajima's D values were also negative but significant in

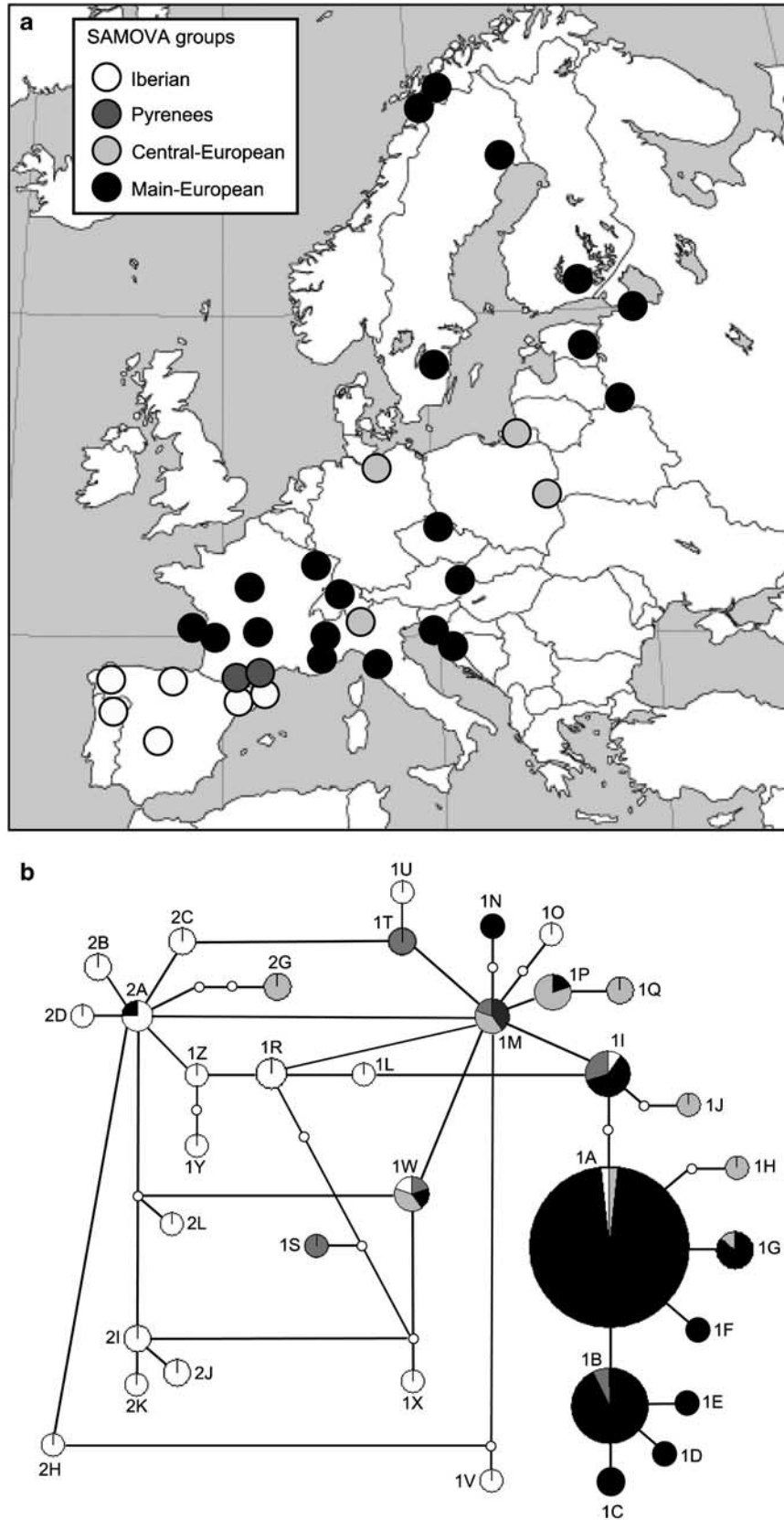


Figure 2 Results of the Spatial Analysis of Molecular Variance (SAMOVA). (a) Geographical distribution of the four identified groups (see text for details). (b) Haplotype network of the 36 haplotypes showing the proportion of individuals belonging to each of the four groups. Color codes are the same as for the map above.

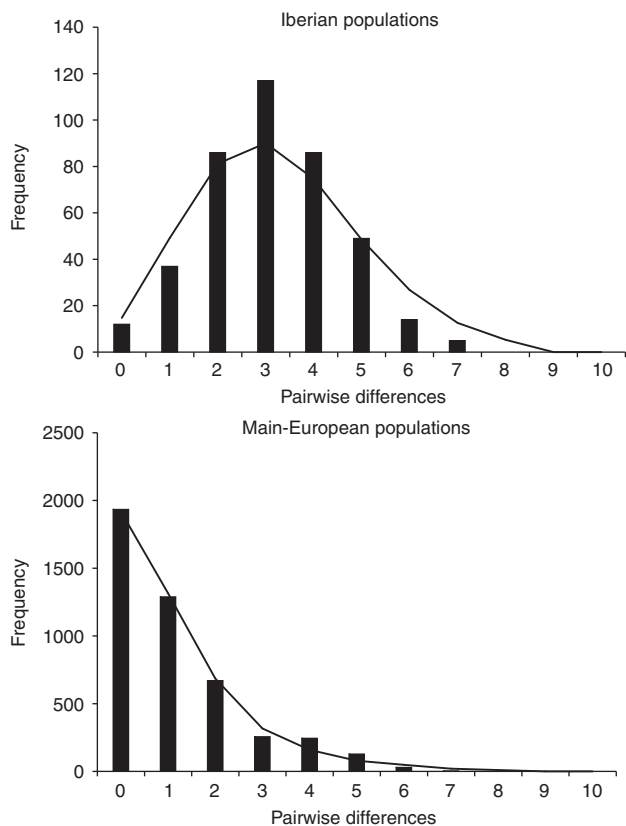


Figure 3 Mismatch distributions for haplotypes of *T. piniperda* among the Iberian and Main-European groups. The bars represent the observed distributions. Solid lines describe the distributions expected under a sudden expansion model.

neither of the two groups (Main-European: Tajima's $D = -1.36$, $P > 0.05$; Iberian: Tajima's $D = -1.01$, $P > 0.1$). Low values were obtained for R_2 in both groups but they were not significant (Main-European: $R_2 = 0.06$, $P > 0.1$; Iberian: $R_2 = 0.08$, $P > 0.05$). The approximate timing of demographic expansion t was estimated for both groups using τ values (Main-European: $\tau = 0.285$; Iberian: $\tau = 3.281$), mutation rate of 1.15% per million year and per nucleotide (Brower, 1994), and a generation time of one year. For the Main-European group, the expansion was estimated to date back 7800 years, whereas the estimated age of the expansion for the Iberian group was 90 000 years.

Discussion

Our phylogeographical study of the bark beetle *T. piniperda* in its known European range showed that this species is genetically diverse, with 36 haplotypes identified from a sample of 150 individuals. A majority of these haplotypes were found in single populations, especially in the Iberian Peninsula. Even if the corresponding genetic network was not clearly structured and if no evidence of a strong phylogeographical signal was found, our results permitted us to decipher the recent past history of the species, and to show that its Quaternary history varied according to the region considered. In particular, the spatial analysis of molecular variance allowed us to identify four main groups of

populations that illustrate the regional characteristics of this cold-tolerant species.

T. piniperda's Iberian populations: a past history

consistent with the 'refugia-within-refugia' hypothesis

The Iberian Peninsula has long been identified as one of the most important Pleistocene glacial refugia for European species (Hewitt, 2004). Yet, it has recently been proposed that this region should rather be seen as a complex of several separate refugia, in which species or populations could independently survive during the glaciations (Gomez and Lunt, 2006). Many species actually show a strong population sub-structure within Iberia. This 'refugia-within-refugia' scenario can hold true for *T. piniperda*, as we found high-allelic diversity and high levels of endemism in this region. The Iberian populations (including populations found near the Pyrenees) are characterized by a high number of rare haplotypes. Moreover, genetic diversity in this region did not decrease with latitude, and the genetic divergence between populations was not significantly related to geographic distances. The data suggest that *T. piniperda* survived the recurrent Quaternary climatic oscillations in several refugia within the Iberian Peninsula without major movements or severe bottlenecks. Such an evolution is expected at the rear edge of distributions, in places where short altitudinal or latitudinal shifts were sufficient to track the most suitable habitats and where allelic diversity was thus maintained (Hewitt, 2001). We could estimate the date of population expansion in the Iberian group to be as old as 90 000 years, which is consistent with the hypothesis of local survival of the species during the repeated Quaternary climatic oscillations. *T. piniperda* is a cold-tolerant species that cannot develop under warm and dry climates (Gallego et al., 2004; Horn, unpublished data). It can thus be hypothesized that it was present in the southern part of the Iberian Peninsula and in low altitudes during glacial maxima, and that it survived the interglacials at higher altitudes and latitudes. Its present-day distribution is probably similar to its interglacial distributions, as was suggested for its main host *P. sylvestris* (Cheddadi et al., 2006).

It is interesting that most of the haplotypes found in Iberia were restricted to this region, suggesting a very limited number of long-distance migrations to the rest of Europe during the warming periods. This can be because of the Pyrenees, a physical barrier that efficiently hindered migration events. Dynesius and Jansson (2000) also suggested that populations that evolved in regions where climatic oscillations were limited were not selected for high-dispersal ability, in contrast to populations that underwent high intensities of climatic oscillations. However, we hypothesized that long-distance movements of individuals originating from the Iberian Peninsula could also have occurred, as can be seen from the presence of haplotypes 2A and 2G up to Russia and to the Italian Alps, respectively. Moreover, the spatial analysis of molecular variance suggested that the two populations located on the French side of the Pyrenees, namely QS-FRA and NA-FRA, should be seen as a separate group. Some of the haplotypes found in these populations are endemic to the Pyrenees and closely related to the Iberian ones, but others have a larger geographical range, including either France, Italy, or

Northern Europe. This suggests that these two north-Pyrenean populations contain both individuals that belong to the Iberian group and individuals that migrated from other refugia or to northern regions.

Outside Iberia: a complex history made of multiple refugia and repeated long-distance dispersal events

Our results show that all populations sampled outside the Iberian Peninsula and the Pyrenees can be separated in two groups, (i) the Central-European populations from Germany, Poland, and Kaliningrad (on the Baltic Sea shore) together with one peculiar Italian population (SON-IT) that will be discussed below; and (ii) all other sampling sites. Some haplotypes are endemic from North or Central Europe (1N, 1H, 1Q), and others (1M, 1P) are also predominantly found in these regions which can explain why the Central-European populations form a separate group. Taken together, these data suggest that *T. piniperda* probably survived in a northern refugium, even if its location cannot be identified with confidence. On the basis of the macrofossils and pollen records, Cheddadi *et al.* (2006) showed that the host plant *P. sylvestris* survived the last ice age in the Mediterranean Peninsulas but also in the Southern Alps, the Danube region and the Hungarian plain. They argue that the populations are today present in their interglacial refugia, located north to their last glacial refugia in which they cannot survive today due to high temperatures. Haplotypes that survived in southern refugia are thus now located at higher latitudes or altitudes. Owing to the obligate relationship between *T. piniperda* and its pine hosts, we can hypothesize that the species had to survive the glacial maxima in regions where pines were present. After the arguments presented by Cheddadi *et al.* (2006), we suggest that North-Central Europe is the current interglacial refugium for haplotypes that probably survived the glaciations in Central or Eastern Europe, where pines were also present during the Last Glacial Maximum (LGM). As very little is known about the distribution of *T. piniperda* outside Europe, we cannot rule out the presence of refugia elsewhere. Genetic data for beetles sampled from eastern populations of *P. sylvestris* or from Asia Minor could help to distinguish between different hypotheses. It is interesting that the occurrence of northern glacial refugia from either Central or Eastern Europe was suggested for the spruce *Picea abies* and two of its associated bark beetles *Ips typographus* and *Pityogenes chalcographus* (Lagercrantz and Ryman, 1990; Stauffer *et al.*, 1999; Avtzis *et al.*, 2008) that are ecologically similar to *T. piniperda*.

Most of the other sampling sites are characterized by the occurrence of one of the two main haplotypes, 1A and 1B, and the whole group of populations bears signs of fairly recent population expansion, which is expected during the interglacial northern colonization of suitable habitats, after the range contraction imposed by the glacial maximum (Slatkin and Hudson, 1991; Avise, 2000). It is interesting that the date of the expansion was estimated to ca. 7500 years, that is, after the LGM. Contrarily to the situation described for the Iberian Peninsula, the geographic distributions of several haplotypes exhibit repeated long-distance migration events, which are because of the contraction-expansion cycles after the Quaternary climatic oscillations. Glaciations

were more severe in this part of Europe than they were in Iberia, which can explain the contrast between regions (Hewitt, 1996; Dynesius and Jansson, 2000). Nevertheless, the location of some of the rare haplotypes allows us to formulate hypotheses about the most plausible refugia and the directions of post-glacial re-colonization. For instance, the haplotypes 1F and 1G, which are closely related to the main haplotype 1A, are found mostly in Italy, Slovenia and Croatia, which is consistent with a glacial refugium located south of the Alps. As they are found also in more northern locations like Germany or Poland, and as the main 1A occurs as north as Northern Sweden and Norway, we suggest that this refugium was one of the main sources for the post-glacial northward colonizations that took place during the last interglacial. Concerning the second most frequent haplotype 1B, the Southern Alps and the Massif Central harbor its closest rare haplotypes 1C, 1D and 1E, and the major haplotype 1B is restricted to France and the Italian Alps. This distribution suggests that this group of related haplotypes originated from an Italian refugium and that they are found today in their interglacial range, north of the places in which they survived the LGM. It is worth noting that the same scenario was suggested for *P. sylvestris* (Cheddadi *et al.*, 2006; Naydenov *et al.*, 2007).

Finally, some closely related rare haplotypes are found both in the Pyrenees and in Italy (1K and 1J, closely related to the more frequent 1I that is mostly found in the Pyrenees). This pattern can be explained if this group of haplotypes survived in a glacial refugium located in the north of the Iberian Peninsula and reached Northern Europe during an interglacial, as is suggested by the occurrence of haplotype 1I up to Estonia and Russia, in long distance movements similar to that of Iberian haplotypes 2A and 2G (Nichols and Hewitt, 1994). During the following glacial episode, some individuals may have been trapped in Italy and the Balkans, whereas some were still present in the Pyrenees. In the present time, individuals bearing these haplotypes are found both in the North of the Iberian refugia and in the North of the Italian refugia. Similarly, the haplotype 1W has a peculiar contemporary distribution, with individuals in the Iberian Peninsula, France, Italy and up to Poland. Once again, we can hypothesize that such a pattern is due to the fact that a given haplotype survived in different refugia after recurrent glacial episodes. The complex distribution pattern of haplotype 1W probably explains why the SAMOVA analysis grouped the Italian population SON-IT with the North-Central European populations. It is interesting that one haplotype of *P. sylvestris* found mostly in the Iberian Peninsula was also present in low proportions in the Balkans, which suggest a similar history (Naydenov *et al.*, 2007), with different refugia sharing similar rare haplotypes.

T. piniperda is known to have good dispersal abilities (Kerdelhué *et al.*, 2006), which may explain the complex phylogeographic pattern highlighted here. It is interesting that its sibling species *T. destruens* was recently proved to show a similar complex history because of high migration events in Europe (Horn *et al.*, 2006). Moreover, our results show that the recent history of *T. piniperda* is highly similar to that of its main host *P. sylvestris*, even if it is not strictly associated to that particular host species. Their ecological requirements are very close, and the dispersal ability of the beetle allowed

it to track its host both during glacial periods and during interglacials. Both species apparently shared the same refugia, in the Mediterranean Peninsulas, in the Southern Alps but also in more cryptic northern refugia.

Acknowledgements

We are grateful to M Faccoli, D Gallego, J Garcia, B Hrasovec, M Kenis, M Knizek, M Lombardero, M Mandelshtam, R Pavlin, M Robba for providing samples, and to T Vasconcelos for her help in lab work with the Portuguese samples. We thank E Magnoux for her technical assistance with the automatic sequencer ABI 3100, and N Bruyndonckx, C Burban, RJ Petit, G Roux-Morabito, M Stöck and the three anonymous referees for their helpful comments on the manuscript. We also thank especially Patrick Nosil and Dorothea Pio who corrected the English.

References

- Avice JC (2000). *Phylogeography: the History and Formation of Species*. Harvard University Press: Cambridge, MA, USA.
- Avtzis DN, Arthofer W, Stauffer C (2008). Sympatric occurrence of diverged mtDNA lineages of *Pityogenes chalcographus* (Coleoptera, Scolytinae) in Europe. *Biol J Linn Soc* **94**: 331–340.
- Balachowsky A (1949). *Faune de France: 50 Coléoptères Scolytides*. Librairie de la Faculté des Sciences: Paris.
- Brower AVZ (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc Natl Acad Sci USA* **91**: 6491–6495.
- Bucci G, Gonzalez-Martinez SC, Le Provost G, Plomion C, Ribeiro MM, Sebastiani F et al. (2007). Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. *Mol Ecol* **16**: 2137–2153.
- Burban C, Petit RJ (2003). Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Mol Ecol* **12**: 1487–1495.
- Burban C, Petit RJ, Carcreff E, Jactel H (1999). Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. *Mol Ecol* **8**: 1593–1602.
- Cheddadi R, Vendramin GG, Litt T, Francois L, Kageyama M, Lorentz S et al. (2006). Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. *Global Ecol Biogeogr* **15**: 271–282.
- Clement M, Posada D, Crandall KA (2000). TCS: a computer program to estimate gene genealogies. *Mol Ecol* **9**: 1657–1659.
- Crandall KA, Templeton AR (1993). Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* **134**: 959–969.
- Dupanloup I, Schneider S, Excoffier L (2002). A simulated annealing approach to define the genetic structure of populations. *Mol Ecol* **11**: 2571–2581.
- Dynesius M, Jansson R (2000). Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proc Natl Acad Sci USA* **97**: 9115–9120.
- Excoffier L, Laval G, Schneider S (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol Bioinform Online* **1**: 47–50.
- Excoffier L, Smouse PE, Quattro JM (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.
- Fu YX (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915–925.
- Gallego D, Canovas F, Esteve M, Galian J (2004). Descriptive biogeography of *Tomicus* (Coleoptera: Scolytidae) species in Spain. *J Biogeogr* **31**: 2011–2024.
- Gomez A, Lunt DH (2006). Refugia within refugia: Patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S, Ferrand N (eds). *Phylogeography in Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of European Biodiversity*. Springer: Dordrecht, the Netherlands, pp 1–42.
- Hewitt GM (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc* **58**: 247–276.
- Hewitt GM (2000). The genetic legacy of the quaternary ice ages. *Nature* **405**: 907–913.
- Hewitt GM (2001). Speciation, hybrid zones and phylogeography—or seeing genes in space and time. *Mol Ecol* **10**: 537–549.
- Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. *Phil Trans R Soc Lond B Biol Sci* **359**: 183–195.
- Horn A, Roux-Morabito G, Lieutier F, Kerdelhué C (2006). Phylogeographic structure and past history of the circum-Mediterranean species *Tomicus destruens* Woll. (Coleoptera: Scolytinae). *Mol Ecol* **15**: 1603–1615.
- Kerdelhué C, Magnoux E, Lieutier F, Roques A, Rousselet J (2006). Comparative population genetic study of two oligophagous insects associated with the same hosts. *Heredity* **97**: 38–45.
- Kerdelhué C, Roux-Morabito G, Forichon J, Chambon J-M, Robert A, Lieutier F (2002). Population genetic structure of *Tomicus piniperda* L. (Curculionidae: Scolytinae) on different pine species and validation of *T. destruens* (Woll.). *Mol Ecol* **11**: 483–494.
- Lagercrantz U, Ryman N (1990). Genetic structure of Norway spruce (*Picea abies*)—Concordance of morphological and allozymic variation. *Evolution* **44**: 38–53.
- Långström B (1983). Life cycle and shoot-feeding of the pine shoot beetles. *Stud For Suec* **163**: 1–29.
- Legendre P, Legendre L (1998). *Numerical Ecology*. Elsevier: Amsterdam.
- Naydenov K, Senneville S, Beaulieu J, Tremblay F, Bousquet J (2007). 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* **7**: 233.
- Nei M (1975). *Molecular Population Genetics and Evolution*. North-Holland Research Monographs, Frontiers in Biology, vol. 40. North Holland Publishing Company: Amsterdam, The Netherlands.
- Nichols RA, Hewitt GM (1994). The genetic consequences of long distance dispersal during colonization. *Heredity* **72**: 312–317.
- Petit RJ, Aguinalde I, Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R et al. (2003). Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science* **300**: 1563–1565.
- Petit RJ, El Mousadik A, Pons O (1998). Identifying populations for conservation on the basis of genetic markers. *Conserv Biol* **12**: 844–855.
- Pfeffer A (1994). Zentral—und Westpaläarktische Borken—und Kernkäfer (Coleoptera, Scolytidae, Platypodidae). *Entomol Basil* **17**: 5–310.
- Pfenninger M, Posada D (2002). Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution* **56**: 1776–1788.
- Pons O, Petit RJ (1996). Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **144**: 1237–1245.
- Postner M (1974). Scolytidae (=Ipidae), Borkenkäfer. In: Schwenke W (ed.). *Die Forstschädlinge Europas*. Parey: Berlin, pp. 334–482.

- R Development Core Team (2005). R: a language and environment for statistical computing. In: R Foundation for Statistical Computing: Vienna, Austria <http://www.R-project.org>.
- Ramos-Onsins SE, Rozas J (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**: 2092–2100.
- Ritzerow S, Konrad H, Stauffer C (2004). Phylogeography of the Eurasian pine shoot beetle *Tomicus piniperda* L. (Coleoptera, Scolytidae). *Eur J Entomol* **101**: 13–19.
- Rogers AR, Harpending H (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* **9**: 552–569.
- Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.
- Sallé A, Artofer W, Lieutier F, Stauffer C, Kerdelhué C (2007). Phylogeography of a host-specific insect: genetic structure of *Ips typographus* in Europe does not reflect the past fragmentation of its host. *Biol J Linn Soc* **90**: 239–246.
- Schmitt T (2007). Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Front Zool* **4**: 11.
- Sinclair WT, Morman JD, Ennos RA (1999). The postglacial history of Scots pine (*Pinus sylvestris* L.) in western Europe: evidence from mitochondrial DNA variation. *Mol Ecol* **8**: 83–88.
- Slatkin M, Hudson RR (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**: 555–562.
- Soranzo N, Alia R, Provan J, Powell W (2000). Patterns of variation at a mitochondrial sequence-tagged-site locus provides new insights into the postglacial history of European *Pinus sylvestris* populations. *Mol Ecol* **9**: 1205–1211.
- Stauffer C, Lakatos F, Hewitt GM (1999). Phylogeography and postglacial colonization routes of *Ips typographus* L. (Coleoptera, Scolytidae). *Mol Ecol* **8**: 763–773.
- Stewart JR, Lister AM (2001). Cryptic northern refugia and the origins of the modern biota. *Trends Ecol Evol* **16**: 608–613.
- Tajima F (1989). The effect of change in population size on DNA polymorphism. *Genetics* **123**: 597–601.
- Templeton AR, Crandall KA, Sing C (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619–633.
- Thompson JD, Higgins DG, Gibson TJ (1994). Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acid Res* **22**: 4673–4680.
- Ursenbacher S, Carlsson M, Helfer V, Tegelstrom H, Fumagalli L (2006). Phylogeography and Pleistocene refugia of the adder (*Vipera berus*) as inferred from mitochondrial DNA sequence data. *Mol Ecol* **15**: 3425–3437.