Molecular Ecology (2004) **13**, 1883–1897 doi: 10.1111/j.1365-294X.2004.02212.x

Isolation by distance and sharp discontinuities in gene frequencies: implications for the phylogeography of an alpine insect species, *Carabus solieri*

S. GARNIER,*† P. ALIBERT,† P. AUDIOT,* B. PRIEUR* and J.-Y. RASPLUS*
**INRA, Centre de Biologie et de Gestion des Populations, Campus International de Baillarguet, CS 30016, 34988 Montferrier sur Lez, France, †UMR-CNRS 5561 Biogéosciences, Université de Bourgogne, 6 blvd Gabriel, 21000 Dijon, France*

Abstract

Analysis of genetic isolation by distance (IBD) is of prime importance for the study of processes responsible for spatial population genetic structure and is thus frequently used in case studies. However, the identification of a significant IBD pattern does not necessarily imply the absence of sharp discontinuities in gene frequencies. Therefore, identifying barriers to gene flow and/or secondary contact between differentiated entities remains a major challenge in population biology. Geographical genetic structure of 41 populations (1080 individuals) of an alpine insect species, *Carabus solieri*, was studied using 10 microsatellite loci. All populations were significantly differentiated and spatially structured according to IBD over the entire range. However, clustering analyses clearly identified three main clusters of populations, which correspond to geographical entities. Whereas IBD also occurs within each cluster, population structure was different according to which group of populations was considered. The southernmost cluster corresponds to the most fragmented part of the range. Consistently, it was characterized by relatively high levels of differentiation associated with low genetic diversity, and the slope of the regression of genetic differentiation against geographical distances was threefold those of the two other clusters. Comparisons of within-cluster and between-cluster IBD patterns revealed barriers to gene flow. A comparison of the two approaches, IBD and clustering analyses, provided us with valuable information with which to infer the phylogeography of the species, and in particular to propose postglacial colonization routes from two potential refugia located in Italy and in southeastern France. Our study highlights strongly the possible confounding contribution of barriers to gene flow to IBD pattern and emphasizes the utility of the model-based clustering analysis to identify such barriers.

Keywords: barrier to gene flow, clustering analysis, colonization, ground beetle, isolation by distance, microsatellites

Received 4 December 2003; revision received 19 March 2004; accepted 19 March 2004

Introduction

The evolutionary potential of a species depends mainly on population structure, which results from a balance of evolutionary forces producing either local differentiation or homogeneity (Slatkin 1987). The relative importance of each factor may change in space and time, and some events can leave imprints for a long period of time (Hewitt 2000). In such a context the distinction between historical and present processes is considered a key point in studies of population differentiation. For species with low dispersal ability, a higher genetic similarity is expected between neighbouring individuals or populations than between distant ones. This pattern of genetic structure is called isolation by distance (IBD, Wright 1943). At equilibrium, under dispersal and genetic drift, IBD pattern is revealed by a positive and significant correlation between genetic

Correspondence and present address: S. Garnier. UMR-CNRS 5561 Biogéosciences, Université de Bourgogne, 6 blvd Gabriel, 21000 Dijon, France. Fax: + 33 (0) 3 80 39 62 31;
E-mail: garniers@ensam.inra.fr

differentiation and geographical distances (Slatkin 1993; Rousset 1997). Identification of IBD can help to show equilibrium between migration and genetic drift (contemporary processes), or to link limited dispersal ability and genetic differentiation. It can also allow the determination of the neighbouring size, the estimation of demographic parameters or the comparison of the relative influences of gene flow and drift on population structure among different regions (Slatkin 1993; Rousset 1997; Hutchison & Templeton 1999; Rousset 2000; Pogson *et al*. 2001). Although empirical studies provide numerous examples of IBD occurrence in natural populations (Slatkin 1993; Peterson 1996; Pogson *et al*. 2001 and references therein), there is growing evidence that many species have not yet reached migration-drift equilibrium and that observed patterns of genetic differentiation reflect population history rather than current levels of gene flow (e.g. Latta & Mitton 1999; Pogson *et al*. 2001; Turgeon & Bernatchez 2001). In such nonequilibrium systems, the study of the relationship between genetic and geographical distances can still be informative revealing, for instance, recent range expansion or a quasi absence of gene flow (Slatkin 1993; Hutchison & Templeton 1999). Hence, analyses of IBD patterns are of prime importance in identifying processes responsible for spatial population genetic structure and are thus used widely in case studies.

The occurrence and pattern of IBD depends on the spatial scale considered and can also change according to region and time, due to spatial and/or temporal variation in relative influences of forces moulding population structure (Slatkin 1993; Rousset 1997; Castric & Bernatchez 2003). Moreover, the assumptions of spatial and temporal stability in some models contrast with heterogeneity in terms of demography, environmental conditions or history characterizing natural populations. As a consequence, a significant IBD pattern does not necessarily mean a spatially homogeneous gene flow. Indeed, some authors have pointed out that signficiant IBD pattern could exist despite the presence of barriers to gene flow (Bossart & Prowell 1998; Lugon-Moulin & Hausser 2002). In such a context, it is of prime importance to be able to identify heterogeneity of gene flow due to such barriers. A model-based clustering method (Pritchard *et al*. 2000) can help to approach this problem. This method was developed to infer population structure and to assign individuals to populations using multilocus genotype data. Unlike previous assignment methods, the genetic composition of source populations is unknown (populations are not defined a priori). We will show that the combined analysis of IBD patterns and clustering analysis is a powerful approach to detect sharp discontinuities in gene frequencies due to physical barriers to gene flow and/or secondary contact.

In the present study we used 10 microsatellite loci to describe the genetic population structure of *Carabus solieri* Dejean (Coleoptera, Carabidae), a ground beetle distributed in the Southern Alps of France and Italy. This species exhibits high levels of diversity and differentiation for several characters including neutral genetic markers, morphology and colour (Bonadona 1973; Darnaud *et al*. 1978; Rasplus *et al*. 2001). As with many other ground beetles species, this brachypterous insect has limited dispersal abilities and is therefore susceptible to exhibit an IBD pattern if it has reached migration-drift equilibrium. However, it is suspected that current population structure still contains historical imprints and that the pattern of population structure is due partly to historical events, i.e. hybridization after a secondary contact between two differentiated entities considered as subspecies (Rasplus *et al*. 2001). In addition, the habitat structure within its range appears rather heterogeneous (due, in particular, to the mountainous relief) and provides numerous potential barriers to gene flow. Hence, our objectives were (i) to test for migration-drift equilibrium by studying the IBD pattern, (ii) to detect and identify potential barriers to gene flow, (iii) to determine the nature of such barriers (secondary contact vs. physical barriers) and (iv) to discuss plausible phylogeographical scenarios and in particular to infer recolonization routes and the origin of the different entities recognized in *C. solieri*.

Materials and methods

Species studied and sampling scheme

C. solieri occurs in a relatively restricted area in the southern and Ligurian Alps (Fig. 1). This species is associated mainly with humid forests, either deciduous or coniferous, but the species also occurs in dry Mediterranean forest and alpine grasslands. It is a spring breeder, laying eggs in spring and summer, depending on environmental conditions. Larval development occurs in summer and tenerals emerge in late summer or autumn, and overwinter in the soil. Mating will occur during the following spring. This ground beetle is endangered mainly by habitat destruction and fragmentation, especially in highly man-modified habitats in Estérel (southern France), and to a lesser extent by exhaustive collection by entomologists in specific areas. Consequently, *C. solieri* is protected in France.

Despite its limited range, *C. solieri* exhibits important variation in colour, morphology and genes (see Rasplus *et al*. 2001). As a result, intraspecific classification of *C. solieri* is unclear. Depending on authors, three to six subspecies are recognized, but some authors describe more. Rasplus *et al*. (2001) have proposed that *C. solieri* could be subdivided into two distinctive entities considered as subspecies, which differentiated during last glaciations within two different refugia, one in the south of France and one in Italy. After postglacial recolonization, these entities

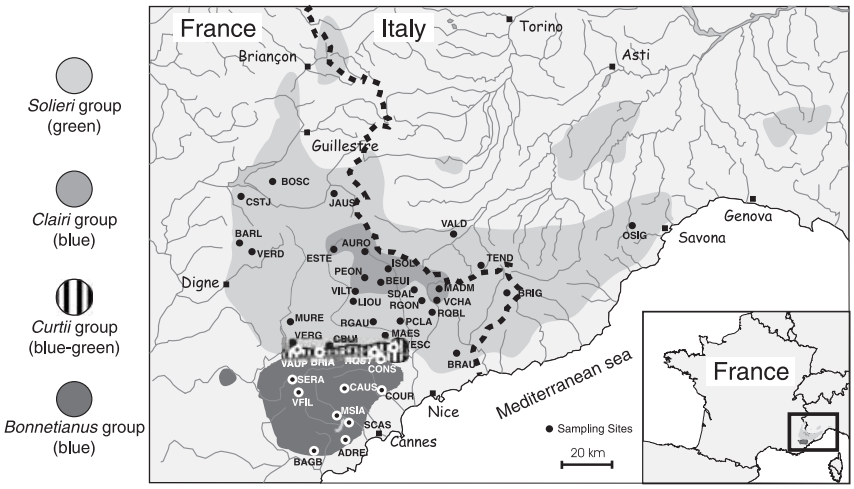

Fig. 1 Range of the different groups of *C. solieri* (see text for details) and location of sampling sites. Colour of individuals is given below the group label.

would have encountered each other, hybridized and introgressed with each other. Here, we will simply consider four groups of *C. solieri* without considering them as real taxonomic entities. Such groups are convenient because they have a geographical coherence and are consistent with the colour of individuals (Fig. 1). The *Bonnetianus* group occurs in the most southern part of the range, from the Estérel to the Montagne du Cheiron. It represents one of the two subspecies defined by Rasplus *et al*. (2001). The *Clairi* group inhabits mountain forests in the Mercantour massif. Individuals from both groups are deep metallic blue. Elsewhere a metallic green form occurs which we refer to as the *Solieri* group. Individuals with intermediate colour (blue-green) are found in the contact zone between *Bonnetianus* and *Solieri* groups. They are sometimes considered as a subspecies but are suspected to be hybrids between these two groups. We refer to them as the *Curtii* group.

Adults were collected with permission using rows of 20–60 unbaited pitfall traps during spring and summer 1997, 1998, 2000 and 2001. Pitfall traps were checked weekly or every 2 weeks all along the adult activity period (from April to August). Specimens were killed and stored in 100% ethanol at –22 °C until analysis. A total of 1080 individuals was collected from 41 localities (see details and localization of sampling sites in Table 1 and Fig. 1). We pooled samples from different years for four populations (AURO, CAUS, PEON and RQST, Table 1) because pairwise exact tests, performed with GENEPOP 3.3 (Raymond & Rousset 1995b) showed no significant genotypic differentiation between years.

Microsatellite analysis

Total DNA was extracted from muscles of a leg using standard phenol–chloroform extraction (Doyle & Doyle 1987). Individuals were genotyped at 10 microsatellite DNA loci: eight were cloned and characterized in *C. solieri*, one in *C. punctatoauratus*, and one in *C. nemoralis* (see Table 2). Primer sequences and other information are given in Garnier *et al*. (2002), except for locus Cn2B (isolated in *C. nemoralis*), whose unpublished primer sequences are ACA-GAA-CGA-CTA-AAA-TGA-ACA-C and ATA-ATC-TAC-CCA-CAA-CCG-AC (GenBank Accession no. AY464122). Genotyping was performed following two methods, depending on the locus (Table 2). In the first method, polymerase chain reaction (PCR) amplifications were carried out with radiolabelled dATP with αP^{33} (see Garnier *et al*. 2002 for details). PCR products were run on 6% denaturing polyacrylamide gels, and the fragments were visualized by autoradiography. Allele size was determined by reference to the original cloned allele. In the second method, one of the two primers was labelled at the 5′ end with one of the fluorescent dyes FAM, HEX or NED (Table 2). Loci giving sufficiently different size ranges were labelled with the same dye, in order to maximize the number of loci multiloaded. PCR amplification were performed as described in Garnier *et al*. (2002), but the reaction mix was slightly different as there was no αP^{33} and the mix contained 4 μM of the nonfluorescent primer and × μM of the fluorescent primer (× depending on the locus, Table 2). PCR products were analysed in the ABI 310 automated sequencer following the manufacturer's protocol using GENESCAN analysis 3.1 and GENOTYPER 2.5 software.

Intrapopulation genetic diversity

Intrapopulation genetic variation was estimated by the number of polymorphic loci (N_{pol}), observed (H_O) and expected (H_E) heterozygosities (unbiased estimate, Nei 1978), using GENETIX 4.04 (Belkhir *et al*. 2001). The observed number of alleles in a sample is highly dependent on sample size, consequently allelic richness was calculated using a rarefaction index, as suggested by El Mousadik & Petit (1996). Given that 2*N* genes have been sampled, the

Locality		Year of sampling	N	N_{pol}	A	H_O	H_E
Les Adrets de l'Estérel	ADRE	1997	20	7	3.0	0.43	0.41
Auron	AURO	1998, 2001	20	9	4.0	0.46	0.46
Bagnols en Forêt	BAGB	2000	24	5	1.5	0.15	0.14
Barles	BARL	2001	25	8	2.9	0.37	0.35
Beuil	BEUI	1998	27	9	4.5	0.52	0.53
Boscodon	BOSC	2001	29	8	2.6	0.28	0.30
Col de Braus	BRAU	1998	24	10	3.5	0.47	0.50
Briançonnet	BRIA	2000	30	8	2.9	0.45	0.42
La Brigue	BRIG	1998	21	10	5.3	0.60	0.67
Caussols	CAUS	1998, 2000	27	9	2.8	0.48	0.49
Col du Buis	CBUI	2000	30	8	3.3	0.43	0.45
Conségudes	CONS	2000	30	10	4.6	0.58	0.56
Courmette	COUR	1998	20	8	2.5	0.34	0.34
Col Saint-Jean	CSTJ	2001	15	8	2.8	0.30	0.31
Esteng	ESTE	1998	11	8	4.6	0.45	0.50
Isola	ISOL	1998	21	10	4.5	0.55	0.56
Jausiers	JAUS	2001	25	5	1.7	0.23	0.21
Le Liouc	LIOU	2001	41	8	3.1	0.38	0.37
La Madone de Fenestre	MADM	1998	22	10	4.2	0.50	0.51
Malaussène	MAES	2000	26	10	4.0	0.48	0.49
Montauroux	MSIA	2001	31	5	2.2	0.25	0.26
Mure	MURE	1998	18	6	3.1	0.33	0.34
Osiglia	OSIG	1997	44	10	5.2	0.58	0.64
Pont de Clans	PCLA	2000	36	8	3.4	0.42	0.44
Péone	PEON	1998, 2000	33	9	4.8	0.53	0.54
Rigaud	RGAU	1998	21	8	3.3	0.40	0.41
Rigons	RGON	2000	30	10	4.9	0.60	0.60
Roquebillière	RQBL	2000	30	9	3.9	0.54	0.57
Roquestéron	RQST	1997, 1998, 2000	26	10	3.7	0.47	0.47
Saint-Cassien	SCAS	1997	28	7	2.4	0.32	0.34
Saint-Dalmas	SDAL	2000	17	9	4.7	0.56	0.57
Séranon	SERA	2000	30	9	3.9	0.48	0.47
Tende	TEND	1998	25	10	5.1	0.58	0.63
Valdieri	VALD	1998	25	9	2.3	0.36	0.32
Vauplane	VAUP	1998	24	9	3.7	0.50	0.48
Vallon des Châtaigniers	VCHA	2000	30	10	3.8	0.43	0.45
Verdaches	VERD	2001	35	7	2.9	0.33	0.34
Vergons	VERG	1998	20	9	4.1	0.52	0.52
Vescous	VESC	2000	30	8	2.7	0.35	0.37
Vallon du Fil	VFIL	2000	30	9	2.9	0.32	0.35
Villetalle	VILT	2000	29	9	4.1	0.49	0.52

Table 1 Sampling sites and genetic polymorphism of the 41 populations studied. N, sample size; N_{pol}, number of polymorphic loci; A, allelic richness (estimated for a sample of 11 individuals); H_O, observed heterozygosity; H_E, gene diversity; A, H_O, H_E are averaged over 10 loci

expected number of alleles in a subsample of $2n$ genes ($n \leq N$) was estimated using FSTAT 2.9.3 (Goudet 1995) [n is fixed as the smallest number of individuals genotyped for a locus in a sample, i.e. $n = 11$ (ESTE sample)]. Different statistics were computed because the behaviour of these statistics can vary according to population history (Cornuet & Luikart 1996). Intrapopulation genetic diversity was assessed because it can reflect past range expansion (Le Corre & Kremer 1998; Hewitt 2000) or current habitat fragmentation.

Linkage disequilibrium between all pairs of loci and departure from Hardy–Weinberg equilibrium for each locus were tested within each population using exact tests. A global test for Hardy–Weinberg equilibrium across all loci was constructed using Fisher's method (Sokal & Rohlf 1995), providing that statistical independence of loci was established previously. Calculations were performed using GENEPOP 3.3 (Raymond & Rousset 1995b). Deviations from Hardy–Weinberg proportions were quantified by the unbiased estimator of Wright's inbreeding coefficient F_{IS} calculated according to Weir & Cockerham (1984).

Population differentiation and IBD

In order to assess the among-population variability, we first considered each sampling site as a distinct population. Differentiation for both all populations and all

Table 2 Polymorphisms at 10 microsatellite loci over 41 populations of *C. solieri*. Loci were isolated in *Carabus solieri* (Csol), in *C. nemoralis* (Cn) or in *C. punctatoauratus* (Cp). Labelling corresponds to radiolabelled dATP (= radio) or fluorescent primer [= fluo (dye group, concentration in mix reaction in μM)]. Tm is the annealing temperature in °C (TD means 'touchdown' procedure, see Garnier *et al*. 2002); range is the size in base pairs of the smallest and the largest allele; N_A is the total number of alleles found; n_A is the mean number of alleles found in one population (averaged over localities) with its standard deviation (SD); n_{min} and n_{max} are the minimum and the maximum number of alleles found in one locality, respectively

Locus	Labelling	Tm	Range	N_A	n_A (SD)	n_{min}	n_{max}
Csol 10129B	radio	TD 60	156–208	22	5.00 (2.07)	2	11
Csol 1122	fluo (FAM; 0.15)	TD 62	113–166	22	5.90 (2.60)	1	13
Csol 1259	fluo (HEX; 0.10)	TD 60	161–213	21	6.10 (2.66)	2	13
Csol 13F	fluo (NED; 0.40)	TD 60	163–171	6	1.88 (0.87)	1	5
Csol 6103	fluo (NED; 0.15)	TD 60	230–240	6	2.46 (0.92)	1	5
Csol 8155	radio	TD 55	126–192	30	5.80 (2.93)	1	14
Csol 828	fluo (FAM; 0.10)	TD 62	176–192	6	1.93 (0.96)	1	6
Csol 9170	fluo (FAM; 0.10)	54	208–250	11	2.68 (1.21)	1	5
Cn 2B	fluo (HEX; 0.15)	TD 58	272–444	62	7.93 (4.81)	1	18
Cp 1/24	radio	TD 55	133–222	30	2.95 (2.30)	1	13

population pairs was tested using a log-likelihood (G)-based exact test (Goudet *et al*. 1996). These tests were performed for each locus and then combined in a global test with Fisher's method (Manly 1985). In addition, both global and pairwise estimates of F_{ST} were computed following Weir & Cockerham (1984) to quantify levels of differentiation.

Finally, IBD over the distribution area was assessed by testing the correlation between genetic and geographical distance considering all population pairs in using the regression of $F_{ST}/(1–F_{ST})$ estimates on logarithm of distance for populations, as suggested by Rousset (1997). This model was tested using Mantel's tests. For geographical distances, we considered straight-line distances between all pairs of sampling sites. IBD was also tested in each cluster of populations identified according to the results of the clustering analysis (see Results section) with the same procedure. All these tests and calculations were performed with GENEPOP 3.3 (Raymond & Rousset 1995b).

Genetic model-based clustering

We used the model-based clustering method described by Pritchard *et al*. (2000) to infer population structure and assign individuals to populations using multilocus genotype data, as implemented in the program STRUCTURE. The model assumes K populations (or clusters; K may be unknown, as in the present case) modelled each by its own set of allele frequencies at each locus. The genetic composition of these populations and the assignation of individuals are both unknown. Assuming Hardy–Weinberg equilibrium and complete linkage equilibrium between loci within populations, population allele frequencies and assignation of individuals to populations were inferred simultaneously using a Bayesian approach.

Generally, the number of clusters (K) in the data is inferred from the posterior probability distribution $\Pr(K \mid X)$ calculated from the posterior probability of the data $\Pr(X \mid K)$ (X being the genotypes of the sampled individuals). Choosing K that maximizes the posterior probability of the data (PPD) can be difficult to apply for complex data sets including many groups (Rosenberg *et al*. 2002). In this case of highly structured data, as K is increased the most divergent groups separate into distinct clusters first, in some cases analogously to the hierarchical branching of tree diagrams (Pritchard *et al*. 2000; Rosenberg *et al*. 2002). As we should aim for the smallest value of K that captures the major structure in the data, a second way to choose K is to consider the successive increase of the PPD for increasing values of K, which can be regarded as the gain of information at each addition of a set of allele frequencies. However, it should be emphasized that the PPD is not an accurate estimate and should be regarded as a heuristic guide to which models are most consistent with the data (see Pritchard *et al*. 2000). In the simple case of individuals sampled from two differentiated locations, the more differentiated these two populations, the more important would be the increase of PPD between $K = 1$ and $K = 2$. Suppose now a set of differentiated populations exhibiting an IBD pattern. As runs are performed with increasing K, the study area would be partitioned into smaller and smaller subareas, until each sampling location, and populations would be clustered according to their geographical proximities. Differences in allelic frequencies between clusters would be less important as K is incremented. Thus, we would expect the PPD increase from $K = 1$ to K to be equal

to the number of sampling sites but to a lesser and lesser extent. In other words, the information brought by each successive additional set of allele frequencies would gradually decrease. By adopting this approach, it is then theoretically possible to identify clusters of populations separated by barriers to gene flow (or vicariant events), as strong allele frequency changes are usually evident on either side of these barriers. In such a context, the gain of PPD would be high until the value of K is equal to the number N of such clusters; then, this gain would drop for a value of $K = N + 1$ and eventually gradually decrease for greater values of K if IBD is occurring in some clusters. We then paid attention both to values of the PPD for each run and to the importance of the increase of the PPD for successive values of K.

Independent runs of the program were carried out for the total data set for values of K comprised between one and 40 (almost the number of sampling sites). As three main clusters were identified (see Results section), STRUCTURE runs were then performed within each cluster. All runs were based on 100 000 iterations after a burn-in period of 20 000 iterations. A minimum of five independent runs were conducted for each situation (number of cluster — data set combination) in order to assess the consistency of the results across runs, using admixture model and uncorrelated allele frequencies model without incorporating population information (see Pritchard *et al*. 2000 for details). We used the program DISTRUCT (Rosenberg 2002) to display individuals' membership coefficients for each cluster. Finally, we also performed an analysis of molecular variance (AMOVA, Excoffier *et al*. 1992) as implemented in ARLEQUIN 2.0 (Schneider *et al*. 2000) to quantify the different genetic variance components (among groups, among populations within groups, within populations).

Results

Intrapopulation genetic diversity

Data relative to polymorphism of each locus are presented in Table 2. Considering all the loci, a total number of 216 alleles were found. The amount of polymorphism varied greatly among loci, ranging from six to 62 alleles (Table 2). The mean number of alleles per locus found in one locality varied from 1.88 to 7.93. Some loci presented high intrapopulation polymorphisms, e.g. 18 alleles found in the locality OSIG (44 individuals) for the locus Cn 2B (Table 2).

Thirty of 41 localities had a fixed allele for at least one locus, some of which BAGB, JAUS and MSIA showed fixed alleles for five of the 10 loci (Table 1). Allelic richness averaged over loci ranged from 1.5 to 5.3 (Table 1). Gene diversity also varied among localities, ranging from 0.14 for BAGB to 0.67 for BRIG (Table 1). Both allelic richness and gene diversity exhibited the same clear spatial pattern,

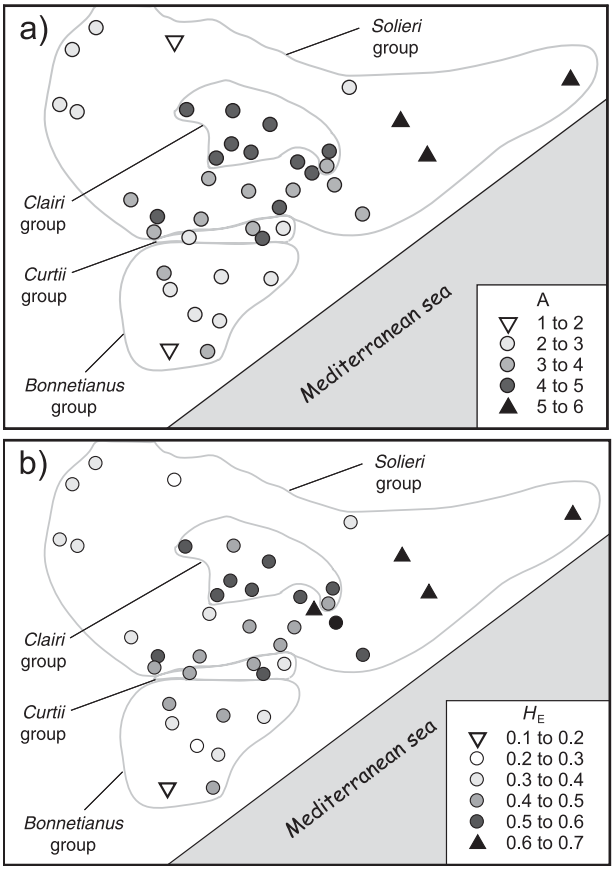

Fig. 2 Geographical distribution of (a) mean allelic richness estimated for a standardized sample size of 11 individuals [A] and (b) gene diversity [H_E]. Limits of the four groups of *C. solieri* are enclosed by grey lines.

i.e. an increase from south to northeast and from northwest to east (Fig. 2).

Out of the 1335 exact tests performed (population–loci pair combinations) for genotypic disequilibrium, only 57 (4.27%) were significant at the 0.05 level. This was less than the 5% expected to be significant by chance alone. Moreover, a single test remained significant after Bonferroni correction. There was therefore no evidence of linkage between loci, which were then considered statistically independent.

Hardy–Weinberg equilibrium was tested for each locus in all populations. A total of 335 exact tests was performed and 23 (6.9%) were significant at the 0.05 level, which is slightly over the proportion expected by chance alone. These significant tests concerned eight of the 10 loci and 16 of the 41 localities (results not shown). Four tests remained significant after Bonferroni correction: loci Csol1259 for BRIG, Csol13F and Cn2B for OSIG, and Csol13F for RQBL. Global test was significant for nine populations (two heterozygote excess and seven heterozygote deficiency) at the 0.05 level and only two remained significant (BRIG and OSIG) after Bonferroni correction. Amplification of locus

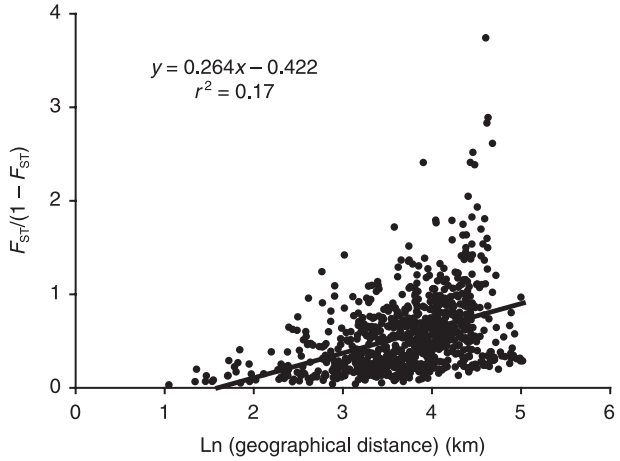

Fig. 3 Isolation by distance pattern in the distribution area. Regression of genetic differentiation [estimated by $F_{ST}/(1 - F_{ST})$] against logarithm of geographical distances (km) for all pairs of sampled populations.

Csol13F could not be obtained for five individuals from the OSIG population. Microsatellite isolation used an individual from SCAS population, one of the furthest populations from OSIG and belonging to a different subspecies (Rasplus *et al*. 2001). Therefore, we cannot exclude the presence of null alleles, e.g. due to mutation in the flanking region as a cause of the heterozygote deficiency observed. However, as there was no tendency for a particular locus to present systematically a heterozygote excess or deficiency, all loci were included in the following analyses.

Population differentiation and IBD

Genetic differentiation across all populations was highly significant for each locus and over all loci ($P < 0.0001$). Values of F_{ST} ranged from 0.256 for locus Csol1259 to 0.550 for locus Csol6103, and F_{ST} was 0.335 over all loci.

Of 8200 exact tests for single-locus genotypic differentiation between population pairs, 7090 (86.5%) were significant at the 0.05 level. The proportion of significant tests varied from 61.34% for locus Csol 13F to 98.90% for locus Csol 1122, which is much more than expected under the null hypothesis of identical genotypic distribution across populations. All pairwise tests (820 population pairs) over all loci were highly significant, even after Bonferroni correction. Multilocus F_{ST} ranged from 0.022 (SDAL vs. RGON) to 0.789 (BAGB vs. JAUS). The overall level of differentiation is high as nearly 50% of the pairwise multilocus F_{ST} were over 0.3 and 11% were greater than 0.5, whereas only 5% were lower than 0.1. As seen previously, this level of differentiation was not due to one or two particular loci. Finally, genetic differentiation between population pairs increased significantly with geographical distance ($r = 0.41$, $P < 10^{-5}$; Fig. 3).

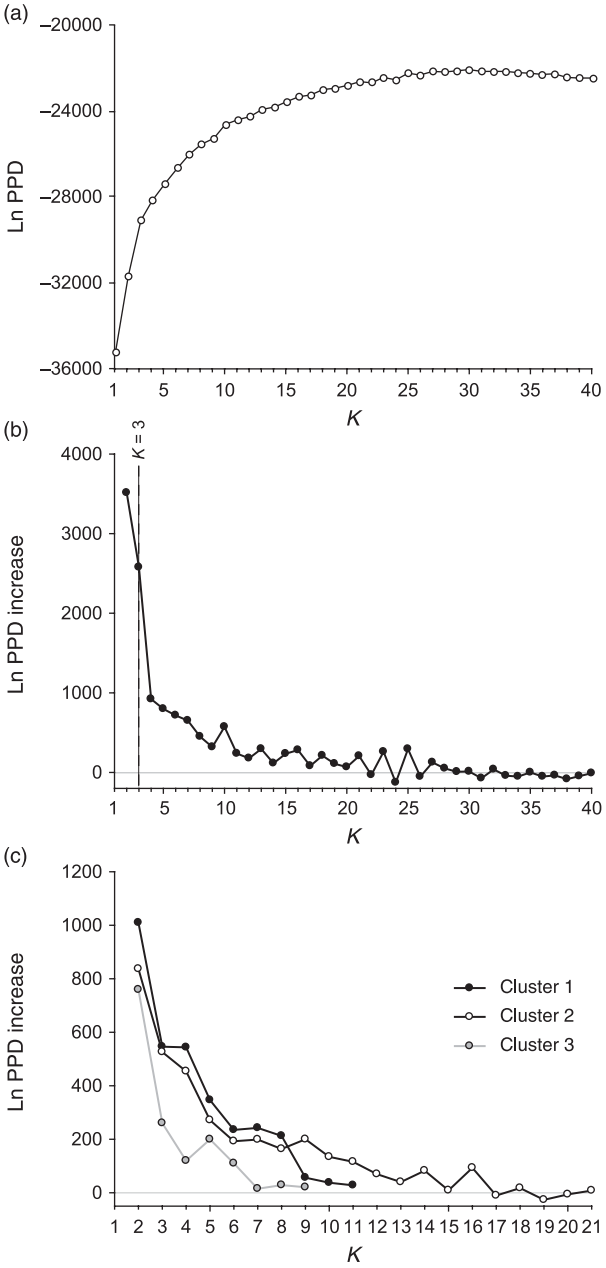

Fig. 4 Posterior probability of the data (PPD) against the maximum number of clusters (K) considered (a) and increase of PPD given K (b and c). For K clusters, this increase is calculated as (Ln PPD_K – Ln PPD_{K-1}). (a and b) Analysis for all data. (c) Analysis for the three main clusters (clusters 1, 2 and 3).

Genetic model-based clustering

Genetic structure over the whole distribution range. When considering all individuals, the PPD increased from $K = 1$ to $K = 30$, where it reached its maximum value and exhibited a plateau (Fig. 4a). This result indicates that differentiation occurs between most of the sampling sites. Concurrently, almost all individuals from the same population had

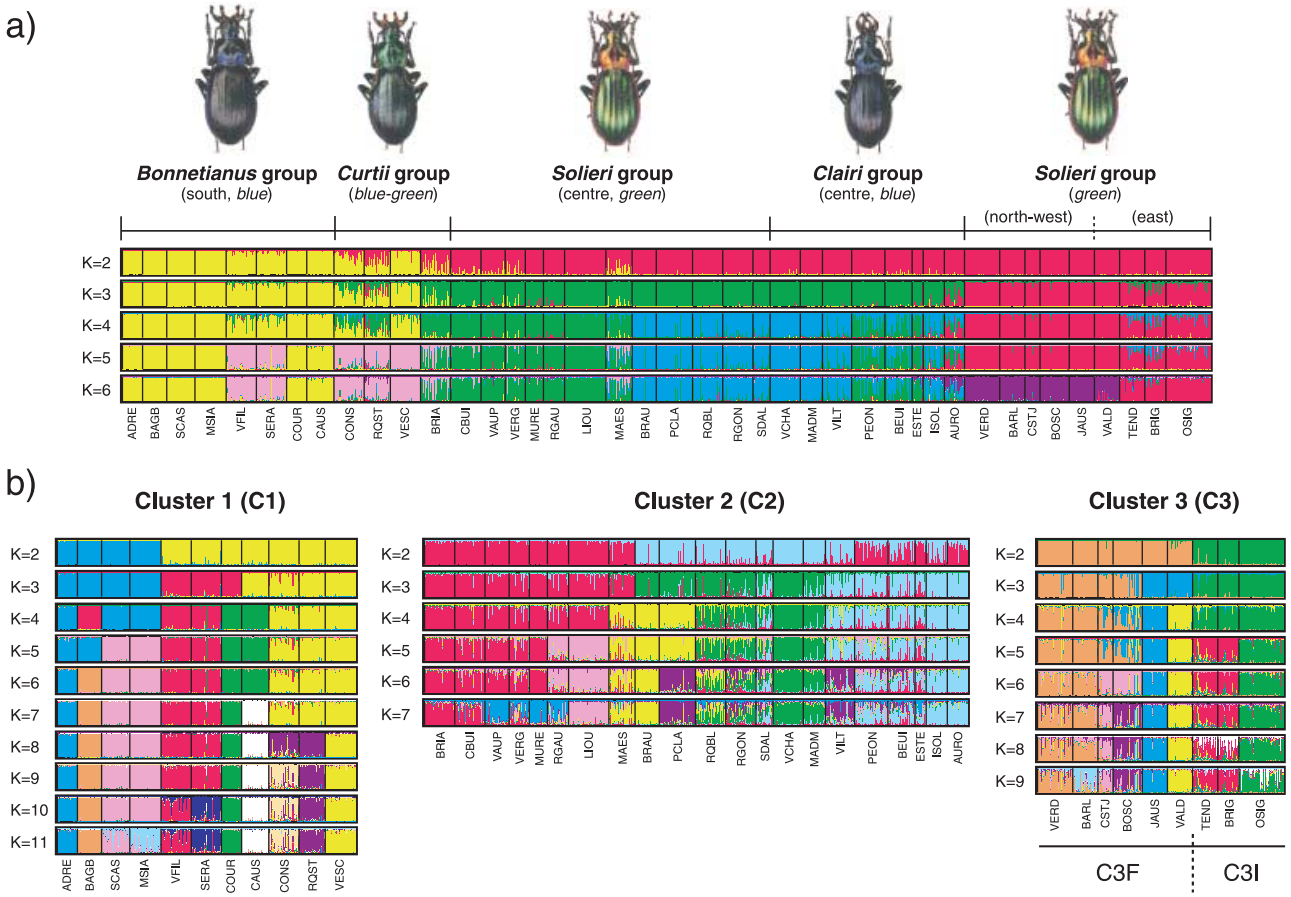

Fig. 5 Clustering results (a) for all populations sampled (for $K = 2$–6) and (b) for the three main clusters (clusters 1, 2 and 3 for $K = 2$–11, 7 and 9, respectively). Each individual is represented as a vertical line partitioned into K coloured segments, whose length is proportional to the individual's estimated membership coefficients in the K clusters. Individuals of different populations are separated by a black line. Populations are labelled below the figure and are approximately sorted from south to north (from the left to the right of the figure). Group affiliation (see text for details), localization in the distribution range and colour of individuals are given above the figure. For each data set and each K, figures are based on the run with the highest posterior probability of the data. There is no colour correspondence across figures based on different data sets.

similar membership coefficients (Fig. 5a). Generally, one of the clusters of individuals obtained for a given value of K is split into two clusters for $K + 1$. Furthermore, clusters always corresponded to geographical groups of populations in a smaller and smaller spatial scale. Theoretically, the PPD would always increase for growing values of K. However, as specified by Pritchard *et al*. (2000), the PPD is not really estimated but rather approximated in an *ad hoc* way, which explains the presence of a more or less plateau for values of K equal or superior to the actual number of populations in the sample.

Whereas these results are consistent with an IBD pattern, we could nevertheless identify three main clusters of populations. First, the increase of the PPD is high for $K = 2$ and $K = 3$, then for $K > 3$ the gain of information was definitely less and exhibited gradually decreasing values (Fig. 4b). This result means that the information brought by the fourth cluster (and the following) is much less important than the information brought by the former three. Second, once two populations have been assigned to different clusters for $K = 3$, they never belonged to the same cluster for greater values of K (partially shown in Fig. 5a). Third, whereas several clustering solutions appeared (with similar PPD) for $K = 2$, 10 runs gave exactly the same clustering solution for $K = 3$. When trying to characterize more gene pools than there are sets of allele frequencies in the model, several solutions can be found.

The three main clusters, hereafter called C1, C2 and C3, corresponded to the *Bonnetianus* and *Curtii* groups (except BRIA) for C1, the *Solieri* group from the centre of the range and the *Clairi* group and population BRIA for C2, the *Solieri* group from the northwest and the east of the range for C3 (Figs 1 and 5a). When considering the total sample, results of the AMOVA (Table 3) show that the among-cluster

Table 3 Results of the analysis of molecular variance (AMOVA). C1, C2 and C3 are the three main clusters defined according to the clustering analysis

Data set	No. of populations	Variance components		
		Among clusters	Among populations within cluster	Within populations
All populations	41	19.5	18.7	61.8
C1 + C2	32	16.5	19.9	63.6
C1 + C3	20	26.1	21.3	52.6
C2 + C3	30	19.0	15.3	65.7
C1	11		34.7	65.3
C2	21		18.0	82.0
C3	9		20.8	79.2

variance component (19.5%) is nearly equal to the among-population within-cluster component (18.7%). Thus, the within-population variance component accounted for around 60% of the genetic diversity of *C. solieri*.

Genetic structure at regional scale. The clustering analysis was conducted independently for each of the three main clusters. Individuals from RQST and CONS (*Curtii* group) had partial membership in two or three clusters (Fig. 5a). Two reasons enabled us to include these populations in C1: (i) individuals were assigned mainly to C1 (average membership coefficients to C1, C2 and C3 were 0.55, 0.26 and 0.19 for RQST and 0.68, 0.29 and 0.03 for CONS); and (ii) for $K > 4$, individuals were assigned with other populations of C1. Similarly, AURO was included in C2.

For each cluster, we found similar patterns: (i) the PPD increased with increasing value of K; (ii) similar membership coefficients were observed for individuals sampled in the same locality; and (iii) nearby populations were always clustered. However, several elements show that population structure is different according to the cluster considered. Individuals from C1 were more strongly assigned to populations than individuals from C2 and C3 clusters (Fig. 5b). Furthermore, the among-population variance component accounted for a third of the genetic diversity of C1, while it accounted for only a fifth for C2 and C3 (Table 3). Finally, the increase of PPD in C3 was important between $K = 1$ and 2, but sharply diminished from $K = 3$ (Fig. 4c). Populations assigned to different clusters for $K = 2$ never clustered together for greater values of K. As a result, C3 could be subdivided into two subgroups: C3F (for France) and C3I (for Italy) (Fig. 5b). On the other hand, such subdivision was not feasible in C1 and C2.

Within- and between-cluster IBD patterns. The relationship between genetic differentiation and geographical distance was assessed in each cluster independently, and between adjacent clusters. Within each cluster, genetic differentiation was correlated positively and significantly to geographical distance ($P = 0.035$, $P < 0.001$ and $P = 0.011$ for C1, C2 and C3, respectively). However, the slope of the regression was three times higher in C1 than in both C2 and C3 (Fig. 6). We plotted population pairs belonging to different clusters on the same figures. In the case of C1 and C2, for the same geographical distances, between-cluster differentiation (C1C2) was higher than within-C2 differentiation (C2C2) but was equivalent or slightly lower than within-C1 differentiation (C1C1, Fig. 6a). This indicates that the strong allele frequencies changes between C1 and C2 separated a homogeneous cluster (C2, among-population variance component = 18.0%, Table 3) and a heterogeneous cluster (C1, among-population variance component = 34.7%, Table 3). C1 seems more prone to genetic drift compared to gene flow than C2.

In order to assess the spatial variation of the relative influence of drift and gene flow in the contact zone between C1 and C2, we defined six 'latitudinal groups of populations' (LG, Fig. 7). We assessed trends of the regression slopes from IBD analysis within each of these groups (Fig. 6c). Slopes of the regressions were gentle for LG at the north of the *Curtii* group (LG4, LG5, LG6). The slope was steep for LG3, and even steeper when considering southern groups (LG1 and LG2), both belonging to the *Bonnetianus* group.

In the case of clusters 2 and 3, we distinguished pairs implicating populations from C3F and C3I subgroups in the intercluster population pairs because they clearly exhibit a different pattern. When considering C3I and C2, between-cluster differentiation (C2C3I) is the same as within-cluster differentiation (C2C2 and C3C3, Fig. 6b) for comparable geographical distances. Slopes of the regressions are also similar. On the other hand, when considering C3F, slopes of regressions are similar but between-cluster differentiation (C2C3F) is much higher than within-cluster differentiation (C2C2 and C3C3, Fig. 6b) for similar geographical distances, thus clearly indicating the presence of a barrier to gene flow between C2 and C3F. Neither comparison of regression slope in IBD analysis in the two subgroups C3F and

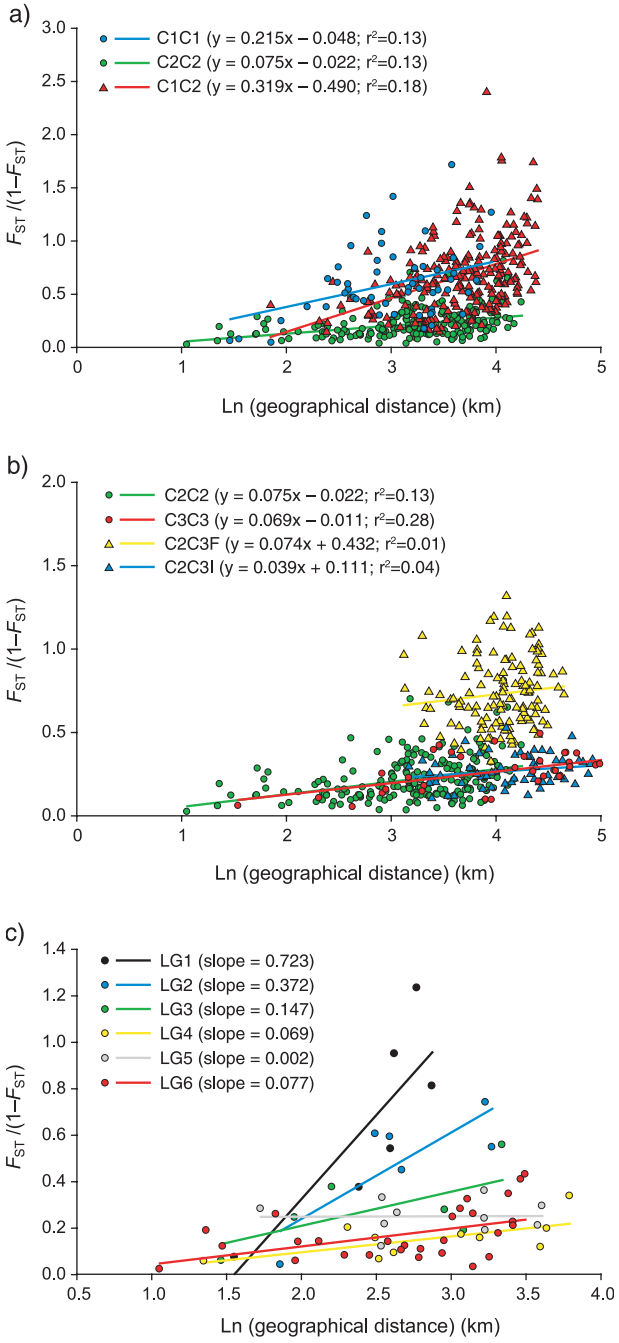

Fig. 6 Relationship between genetic differentiation [estimated by $F_{ST}/(1\text{-}F_{ST})$] and logarithms of geographical distances (km) in parts of the distribution area. Scatterplots and regression for (a) clusters 1 and 2: within-cluster 1 population pairs (G1G1), within-cluster 2 population pairs (G2G2) and between-cluster population pairs (G1G2); (b) clusters 2 and 3: within-cluster 2 population pairs (G2G2), within-cluster 3 population pairs (G3G3), and between-cluster 2 and subclusters 3F (G2G3F) and 3I (G2G3I), respectively. (c) Latitudinal groups of populations (LG): populations were pooled according to their latitude, from the south (LG1, within-LG1 population pairs) to the north (LG6, within-LG6 population pairs) (see Fig. 7 for the definition of these groups).

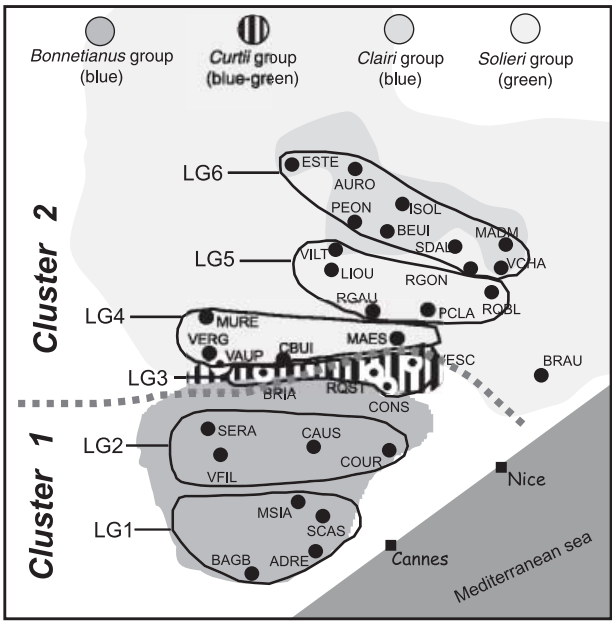

Fig. 7 Definition of six latitudinal groups of populations called LG1 to LG6 from the south to the north. BRAU population was included in any group because of its outlying position.

C3I nor comparison of intra- and intersubgroup regressions were performed because C3I subgroup is constituted of only three populations and intra- and intersubgroup spatial scales were different.

Discussion

Population differentiation and IBD

An important first result of our study is that there is significant genetic differentiation among *C. solieri* populations and that these populations are geographically structured according to an IBD pattern. All populations pairs were significantly differentiated, even at a geographical distance as small as 3 km (e.g. between RGON and SDAL). As a whole, levels of differentiation were high, as 40% of the genetic variance was due to differences among populations (Table 3), and pairwise values of F_{ST} reached a maximum value of 0.789 (between BAGB and JAUS). Previous studies on ground beetles have reported population differentiation at local scale (less than 15 km) even in relatively homogeneous areas (Assmann & Weber 1997; Brouat *et al*. 2003; Keller & Largiadèr 2003) and identified roads and nonforest areas as effective barriers to gene flow. Here, however, differentiation levels are much higher. For instance, for similar geographical distances (less than 15 km), pairwise F_{ST} range from 0.07 to 0.28 for *C. solieri*, whereas they range from 0.01 to 0.06 for *Carabus punctatoauratus*, another forest brachypterous ground beetle (Brouat *et al*.

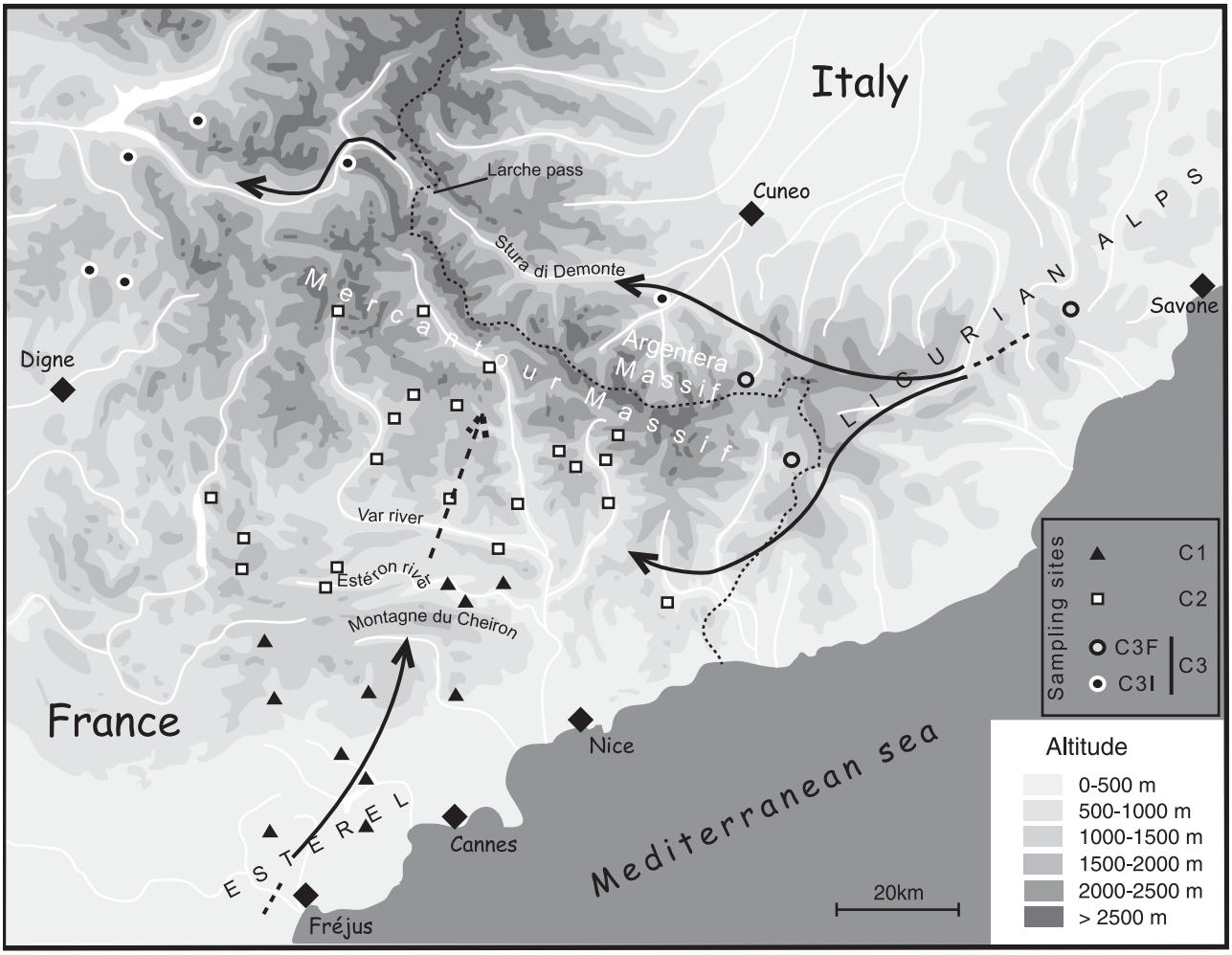

Fig. 8 Hypothetical postglacial colonization routes of *C. solieri*. Black arrows indicate the two colonization routes from the Italian refuge and the northward expansion from the French refuge (the possible colonization of the Mercantour from this latter refuge is shown by the dashed arrow). Symbols of sampling sites indicate the cluster to which they were assigned (see text for details).

2003). However, this result is not surprising because the habitat of *C. solieri* is heterogeneous and potential barriers to gene flow are numerous (e.g. relief, nonforest areas, roads, rivers).

IBD in the whole range of this beetle was revealed through a significant positive and monotonic relationship between genetic differentiation and geographical distance (Fig. 3). Similar conclusions can be drawn from the clustering analysis. A decline of gene flow among populations in relation to geographical distance has been shown for numerous insect species (Peterson 1996 and references therein), although not always at all geographical scales. For instance, Peterson (1995) reported the absence of correlation between gene flow and geographical distances in the butterfly *Euphilotes enoptes* over distances up to 30 km, and in contrast a significant IBD over a larger spatial scale (Peterson 1996). No data related to dispersal exist for *C. solieri*. However, this species is brachypterous and has clearly limited dispersal abilities, as with many other large ground beetles. Data published for other *Carabus* species range from less than 1 m to several tens of metres in 24 h and a few hundred metres in longer time-spans (den Boer 1970; Thiele 1977; Niehues *et al*. 1996). Because of this limited dispersal power and the heterogeneity of its habitat, *C. solieri* probably approaches a stepping-stone model of regional population structure.

Genetic model-based clustering and hierarchical population structure

A second important result is that although being consistent with an IBD pattern, clustering analysis clearly identified three main clusters of populations isolated by barriers to gene flow and exhibiting different patterns of differentiation.

Within-clusters population structure. A significant IBD pattern was revealed within each of the three clusters (C1, C2 and C3). However, population structure is different according to the cluster considered. First, population subdivision is stronger in C1 than in C2 and C3, as shown by a greater among-population variance component, a stronger assignment of individuals from C1 in clustering analysis and a steeper slope of the regression from the IBD analysis in C1. Moreover, intrapopulation genetic diversity is low in C1 relatively to C2 and C3 to the exception of the northwestern populations. Hedrick (1999) pointed out that levels of within-group variation could heavily influence measures of differentiation between groups for highly variable loci such as microsatellites, the maximum value of F_{ST} being at maximum less than the average within-population homozygosity. It could thus be argued that the difference of levels of differentiation observed between C1 and C2 would be due only to their difference of polymorphism (higher for C2, Table 1 and Fig. 2). However, levels of differentiation similar to those found in C2 were observed in samples originating from the northwest of the range and were associated with low levels of polymorphism (as those observed in C1). The high level of differentiation and the low genetic diversity in the southern part of the distribution of *C. solieri* are thus certainly related to the strong habitat fragmentation, which probably reduces both favourable habitat size and gene flow between habitats.

Among-cluster population structure. When considering C1 plus C2, we found between-cluster levels of differentiation higher than within C2 alone and about the same order as within C1 (Fig. 6a). Gene flow seemed to be more important in the north of the *Curtii* group range — which corresponds to the limit between C1 and C2 — and decreases strongly towards the south (Fig. 6c). This pattern could be due partly to the relief in this region. The *Curtii* group is located in the valley of the Estéron river (see Fig. 8), lined in the north by a crest, orientated east–west, and ranging from 1100 to 1550 m above sea level. This mountain range exhibits dry and open habitats unfavourable to *C. solieri* on its south slope and is also the southern limit of the Var river valley, another potential barrier to dispersal for the beetle. In the south of the Estéron valley a succession of mountain ranges, with an east–west orientation and a maximal elevation of 1777 m, probably restrict north–south gene flow. These ranges are covered by forests on their northern slopes and by dry Mediterranean vegetation on their southern slopes, with numerous cliffs. However, it remains to be determined whether the barrier between C1 and C2 is due to the present habitat structure and/or to a secondary contact between two genetically differentiated entities (see below).

When considering pairwise population differentiation between C2 and C3, results are different depending on the C3 subgroup considered. Obviously, a strong barrier to gene flow occurs between C2 and C3F but not between C2 and C3I (Fig. 6b). Even though the difference of genetic diversity between C3F and C3I (low and high, respectively) could cause a bias in the observed differentiation with C2 — the theoretical maximum value of F_{ST} being higher for C3F than for C3I (Hedrick 1999) — this is not sufficient to explain this result. If one considers two populations of C2 and C3F with similar genetic diversity level, for instance BRIA and BARL, respectively, pairwise differentiation with any population from C2 is always higher when considering BARL (pairwise F_{ST} from 0.30 to 0.46) than when considering BRIA (pairwise F_{ST} from 0.13 to 0.30) for similar geographical distances. The barrier to gene flow evidenced between C2 and C3F fit to a crest ranging from 2350 to 3300 m elevation, from the Argentera massif to the north of the Mercantour Massif (see Fig. 8). Also, the absence of barrier to gene flow between C2 and C3I is consistent with the clustering analysis. Indeed, individuals from C3I sites have partial mean memberships also in cluster C2 (probabilities of 0.07, 0.11 and 0.25 for OSIG, TEND and BRIG, respectively; $K = 3$, Fig. 5a). This result should be related to the absence of a strong barrier between the Italian part and the centre of the range of *C. solieri* and provides interesting insights to the phylogeography of the species.

Phylogeography of C. solieri

We have shown that if IBD is the basic process of population structuring in *C. solieri*, the current population structure is likely to also show historical imprints. According to morphological and molecular markers, Rasplus *et al*. (2001) recently proposed that *C. solieri* has been isolated during the last Quaternary ice age into two refugia, one probably located in Italy and the other in south of France (Estérel and/or Maures massifs). The Iberic peninsula, Italy and the Balkans are often identified as important refugia for numerous species in Europe (Taberlet *et al*. 1998; Hewitt 1999), but other places have also been proposed, in particular in the south of France (Pons 1981; Konnert & Bergmann 1995; Blondel & Aronson 1999; Vogel *et al*. 1999; Kropf *et al*. 2002). In the case of *C. solieri*, location of the refugia suited the fact that the most southern extension of the Alpine ice sheet during the last glacial period (Würm) was in the Alpes Maritimes and Ligurian region (see Kropf *et al*. 2002).

Postglacial recolonization would have then occurred northward as the blue form differentiated in the south of France and westward as the green form differentiated in Italy, probably following two ways as suggested by the present results (see Fig. 8). A first postglacial colonization route from Italy could have occurred in the north of Argentera and Mercantour massifs following the Stura di

Demonte valley, and crossed the Larche pass (1991 m elevation) to reach the present northwest part of the range of *C. solieri*. This colonization area corresponds to the range of the C3 cluster. Genetic variability estimated either by allelic richness or genetic diversity declines from east to west (Fig. 2). This is consistent with theoretical and empirical works reporting a general pattern of gradual loss of genetic diversity produced by colonization (Austerlitz *et al*. 1997; Le Corre & Kremer 1998; Hewitt 1999; but see Comps *et al*. 2001). The most frequent alleles in populations from northwest of the range (C3F) were also among the most frequent in populations from Italy (C3I), whereas most alleles with low frequencies observed in populations from C3I were absent in populations from C3F. The loss of diversity in C3F relative to C3I could be explained by the cross of the Larche pass and/or the colonization process through the narrow Stura di Demonte valley. A second postglacial colonization route from Italy probably progressed between the Mercantour massif and the Mediterranean Sea. The range from the Argentera massif to the north of the Mercantour massif (2350–3300 m high) could have prevented the mix of populations originating from the two colonization routes. This is illustrated by the barrier to gene flow identified between C3F and C2.

Hybridization and introgression between the two differentiated entities are supposed to have occurred following their secondary contact (Rasplus *et al*. 2001). The Alpine barrier is indeed one European region concentrating a high number of hybrid zones resulting from postglacial secondary contact (the so-called suture zones, Taberlet *et al*. 1998; Hewitt 2000). However, there is still uncertainty about the exact location of the first contact zone between the two original subspecies of *C. solieri*. It could correspond to the present *Curtii* group range. This would mean that the blue colour of the *Clairi* group results from an independent acquisition. An alternative hypothesis could be that the initial contact zone corresponds to the current limit between C2 and C3I. In this case the *Clairi* group would illustrate the northernmost expansion of the blue entity differentiated in south of France. Even though the present study does not show more genetic proximity between the *Clairi* and *Bonnetianus* groups than between the former and the *Solieri* group in the centre of the range, as expected under the alternative hypothesis, this does not allow us to firmly reject this second scenario. Additional studies are required.

Conclusions

In the present study, we showed empirically that a significant IBD pattern suggesting gradual and steady change in gene frequencies can nevertheless hide sharp discontinuities in gene frequencies. Indeed, we identified two major barriers to gene flow in the distribution area of *C. solieri*. If the physical nature of the first (i.e. a crest range) is evident, it remains to determine the origin of the second (secondary contact and/or habitat structure). IBD patterns are quite different when considering clusters of populations defined by such barriers independently or together. The slope of the regression of genetic differentiation against geographical distances is much higher when considering populations altogether (for instance 0.173 for C2 and C3) than when conducting independent analyses for each cluster (0.075 and 0.069 for C2 and C3, respectively). This result highlights strongly the possible confounding contribution of barriers to gene flow to IBD pattern and emphasizes the utility of the model-based clustering analysis used in this study to investigate genetic population structure more effectively, and in particular to detect barriers to gene flow or secondary contact. The three main clusters identified were characterized by different population structures, suggesting that relative influences of evolutionary forces are quite different between the corresponding regions. Here, the combined analyses of IBD patterns and clustering provided new and valuable insights to the study of *C. solieri* phylogeography, allowing in particular the identification of colonization routes from the Italian refuge. Whereas the main clusters and subgroups identified are probably related to its colonization history, the significant IBD pattern determined within each cluster suggests that the current equilibrium between migration and genetic drift is mainly responsible for the observed population structures.

Acknowledgements

We thank Serge Meusnier, Olivier Duron, Elisabeth Pêcheur and Ariane Bernard-Laurent for their significant help during field sampling. We are also grateful to Thomas Lenormand and Raphaël Leblois for valuable discussions during the preparation of the manuscript. Comments of T. Lenormand improved an earlier version of the manuscript. Thanks also to Louis Bernatchez and to two anonymous reviewers for their constructive criticisms and to Jonathan Pritchard for a discussion about PPD. Eleanor Haine helped with the English language revisions. This work was supported partially by research funds from the Bureau des Ressources Génétiques and by the Institut Français de la Biodiversité. S.G. was supported by a grant from the French Ministère de l'Education Nationale, de la Recherche et de la Technologie. This work is a contribution of the 'Equipe Systématique, Phylogénie et Phylogéographie' from the CBGP and the 'Equipe Différenciation et Espèces' from the UMR-CNRS 5561.

References

Assmann T, Weber F (1997) On the allozyme differentiation of *Carabus punctatoauratus* Germar (Coleoptera, Carabidae). *Journal of Zoological Systematics and Evolutionary Research*, **35**, 33–43.

Austerlitz F, Jung-Muller B, Godelle B, Gouyon P-H (1997) Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology*, **51**, 148–164.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2001) *Genetix 4.02, Logiciel Sous Windows TM Pour La Génétique Des Populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR. 5000, Université de Montpellier II, Montpellier.

Blondel J, Aronson J (1999) *Biology and Wildlife of the Mediterranean Region*. Oxford University Press, Oxford.

den Boer PJ (1970) On the significance of dispersal power for populations of Carabid-beetles (Coleoptera, Carabidae). *Oecologia*, **4**, 1–28.

Bonadona P (1973) Nouvelle contribution à la connaissance des races françaises de *Chrysocarabus solieri* (Col, Carabidae). *Annales de la Société Entomologique de France*, **9**, 759–812.

Bossart JL, Prowell DP (1998) Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends in Ecology and Evolution*, **13**, 202–206.

Brouat C, Sennedot F, Audiot P, Leblois R, Rasplus J-Y (2003) Fine-scale genetic structure of two carabid species with contrasted levels of habitat specialization. *Molecular Ecology*, **12**, 1731–1745.

Castric V, Bernatchez L (2003) The rise and fall of isolation by distance in the anadromous brook charr (*Salvelinus fontinalis* Mitchill). *Genetics*, **163**, 983–996.

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

Cornuet J-M, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Darnaud J, Lecumberry M, Blanc R (1978) Coléoptères Carabidae *Chrysocarabus solieri* Dejean 1826. *Iconographie entomologique. Coléoptères. Planche*, **4**, 1–6.

Doyle JJ, Doyle JL (1987) A rapid isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11–15.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L.) Skeels) endemic to Morocco. *Theoretical Applied Genetics*, **92**, 832–839.

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.

Garnier S, Brouat C, Mondor-Genson G, Prieur B, Sennedot F, Rasplus J-Y (2002) Microsatellite DNA markers for two endemic ground beetles: *Carabus punctatoauratus* and *C. solieri*. *Molecular Ecology Notes*, **2**, 572–574.

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate *F*-statistics. *Journal of Heredity*, **86**, 485–486.

Goudet J, Raymond M, de Meeüs T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics*, **144**, 1933–1940.

Hedrick PW (1999) Highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.

Hewitt G (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898–1914.

Keller I, Largiadèr CR (2003) Recent habitat fragmentation caused by major roads leads to reduction of gene flow and loss of genetic variability in ground beetles. *Proceedings of the Royal Society of London B*, **270**, 417–423.

Konnert M, Bergmann F (1995) The geographical distribution of genetic variation of silver fir (*Abies alba, Pinaceae*) in relation to its migration history. *Plant Systematics and Evolution*, **196**, 19–30.

Kropf M, Kadereit JW, Comes HP (2002) Late Quaternary distributional stasis in the submediterranean mountain plant *Anthyllis montana* L. (Fabaceae) inferred from ITS sequences and amplified fragment length polymorphism markers. *Molecular Ecology*, **11**, 447–463.

Latta RG, Mitton JB (1999) Historical separation and present gene flow through a zone of secondary contact in ponderosa pine. *Evolution*, **53**, 769–776.

Le Corre V, Kremer A (1998) Cumulative effects of founding events during colonization on genetic diversity and differentiation in an island and stepping-stone model. *Journal of Evolutionary Biology*, **11**, 495–512.

Lugon-Moulin N, Hausser J (2002) Phylogeographical structure, postglacial recolonization and barriers to gene flow in the distinctive Valais chromosome race of common shrew (*Sorex araneus*). *Molecular Ecology*, **11**, 785–794.

Manly BJF (1985) *The Statistics of Natural Selection*. Chapman & Hall, London.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Niehues FJ, Hockmann P, Weber F (1996) Genetics and dynamics of a *Carabus auronitens* metapopulation in the Westphalian Lowlands (Coleoptera, Carabidae). *Annales Entomologici Fennici*, **33**, 85–96.

Peterson MA (1995) Phenological isolation, gene flow and developmental differences among low- and high-elevation populations of *Euphilotes enoptes* (Lepidoptera: Lycaenidae). *Evolution*, **49**, 446–455.

Peterson MA (1996) Long-distance gene flow in the sedentary butterfly, *Euphilotes enoptes* (Lepidoptera: Lycaenidae). *Evolution*, **50**, 1990–1999.

Pogson GH, Taggart CT, Mesa KA, Boutilier RG (2001) Isolation by distance in the atlantic cod, *Gadus morhua*, at large and small geographic scales. *Evolution*, **55**, 131–146.

Pons A (1981) The history of the Mediterranean shrublands. In: *Maquis and Chaparrals. Ecosystems of the World* (eds diCastri F, Goodall DW, Specht RL), pp. 131–138. Elsevier, Amsterdam.

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

Rasplus J-Y, Garnier S, Meusnier S, Piry S, Mondor G, Audiot P, Cornuet J-M (2001) Setting conservation priorities: the case study of *Carabus solieri* (Col. Carabidae). *Genetics, Selection and Evolution*, **33**, S141–S175.

Raymond M, Rousset F (1995a) An exact test for population differentiation. *Evolution*, **49**, 1280–1283.

Raymond M, Rousset F (1995b) GENEPOP version 1.2: population genetics software for exact tests and ecumenism. *Journal of Heredity*, **86**, 248–249.

Rosenberg NA (2002) *DISTRUCT: a Program for the Graphical Display of Structure Results*. http://www-hto.usc.edu/~noahr/distruct.html.

Rosenberg NA, Pritchard JK, Weber JL *et al.* (2002) Genetic structure of human populations. *Science*, **298**, 2381–2385.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology*, **13**, 58–62.

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN 2.000. a Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Geneva.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Sokal RR, Rohlf FJ (1995) *Biometry: the Principles and Practices of Statistics in Biology Research*, 3rd edn. W.H. Freeman, New York.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cossons J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Thiele H (1977) *Carabid Beetles in Their Environments*. Springer-Verlag, Berlin.

Turgeon J, Bernatchez L (2001) Clinal variation at microsatellite loci reveals historical secondary intergradation between glacial races of *Coregonus artedi* (Teleostei: Coregoninae). *Evolution*, **55**, 2274–2286.

Vogel JC, Rumsey FJ, Schneller JJ, Barrett JA, Gibby M (1999) Where are the glacial refugia in Europe? Evidence from pteridophytes. *Biological Journal of the Linnean Society*, **66**, 23–37.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

This work constitutes part of S. Garnier's PhD thesis. He studied differentiation and hybridization in *Carabus solieri* using population genetics and geometric morphometrics approaches. P. Alibert is a senior scientist working on population differentiation and speciation. J.-Y. Rasplus is a senior scientist working on systematic, evolution and conservation genetics of insects.
