Acta Theriologica 45, Suppl. 1: 103–117, 2000.
PL ISSN 0001-7051

Genetic differentiation of common shrew *Sorex araneus* populations among different alpine valleys revealed by microsatellites

Nicolas LUGON-MOULIN, François BALLOUX and Jacques HAUSSER

Lugon-Moulin N., Balloux F. and Hausser J. 2000. Genetic differentiation of common shrew *Sorex araneus* populations among different alpine valleys revealed by microsatellites. [In: Evolution in the *Sorex araneus* group: Cytogenetic and molecular aspects. J. B. Searle and J. M. Wójcik, eds]. Acta Theriologica 45, Suppl. 1: 103–117.

Geographical barriers may affect the genetic structure of populations by reducing gene exchanges among them. In Switzerland, the common shrew *Sorex araneus* Linnaeus, 1758 is mostly confined to mountainous areas because of a competing sister species, Millet's shrew *S. coronatus* Millet, 1828, which occupies most of the Swiss lowlands. The structure of common shrew populations found in different alpine valleys may therefore be affected by the topography. Using microsatellites, genetic structuring of seven shrew populations is investigated among four different valleys of the Swiss Alps. Using the exact *G*-test, significant genetic structuring is detected between several valleys. Isolation by distance does not fully explain our results. It appears that high mountain ridges (> 2400 m) can significantly reduce gene flow. *F*- and *R*-statistics are estimated and compared to the exact *G*-tests results. Mantel tests show that F_{ST}, unlike R_{ST}, is significantly correlated with differentiation. F_{ST} remains however low even at high differentiation levels, while R_{ST} has a high variance. We discuss how these results may have wider implications with regards the interpretation of microsatellite data. Finally, a new microsatellite locus, L99, appears to discriminate *S. araneus* of the Vaud and Cordon races from both *S. araneus* Valais and *S. coronatus*.

Institut d'Ecologie, Laboratoire de Zoologie et Ecologie Animale, Université de Lausanne, 1015 Lausanne-Dorigny, Switzerland, e-mail: Nicolas.Lugon-Moulin@ie-zea.unil.ch

Key words: *Sorex araneus*, microsatellites, exact *G*-tests, *F*-statistics, *R*-statistics, genetic differentiation

Introduction

Geographic barriers such as mountain ridges may affect the genetic architecture of populations by reducing gene flow among them. Indeed, geographical barriers often coincide with genetic discontinuities. As such, the Alps in Central Europe represent a potentially important physical obstacle for the dispersal of many organisms. Their role in shaping the genetic structure of animal populations by isolating them from each other may be important. The common shrew *Sorex araneus* Linnaeus, 1758, a small insectivore subdivided into numerous chromosome races (Zima *et al.* 1996), is known to occur from the sea level up to the timber line. Over most of its wide range extending from western Europe as far as Lake Baikal,

S. araneus is usually common at low altitudes. In Switzerland however, this species is mostly confined to mountains because of a competing sister species, Millet's shrew *S. coronatus* Millet, 1828, occupying most of the Swiss lowlands. The situation found in the Swiss Alps implies that the population structure of *S. araneus* may be affected by the alpine topography.

In the Alps, records at 2500 m exist for this species (Winding *et al.* 1995, Reiter and Winding 1997), indicating that occasional dispersal may occur at such altitudes. Genetic structuring of alpine populations in this species is however, poorly known. The recent development of microsatellites in the common shrew opened new possibilities in the study of fine-scale population structuring in this species (Wyttenbach *et al.* 1997, Balloux *et al.* 1998), since intraracial allozyme polymorphism is weak in this organism (eg Wójcik and Wójcik 1994). Microsatellites are highly polymorphic, codominant genetic markers now of current use in population genetics (Jarne and Lagoda 1996). They appear therefore of particular value to study how the alpine topography affects the genetic structuring and relationships among common shrew populations. Using these genetic markers, significant genetic structuring was detected among shrew populations of the Bretolet chromosome race found in four different alpine valleys in the French and Swiss Alps, but differentiation between valleys was not significant (Wyttenbach *et al.* 1999a).

The aim of this paper is to study the genetic structuring of common shrew populations of the Vaud chromosome race found in different valleys of the Swiss Alps. This chromosome race is mainly confined to alpine areas of the Berner Oberland (Switzerland). Using 12 microsatellite loci, we try to assess whether mountain ridges promote differentiation of populations from adjacent alpine valleys. The exact G-test, shown to be the most powerful test of differentiation for diploid populations (Goudet *et al.* 1996), is used to test for genetic differentiation. We estimate population structuring among samples from four different valleys using the traditional measure of differentiation, F_{ST} (Wright 1951, 1965). Since most microsatellite mutations appear to follow a stepwise mutation model (SMM; Kimura and Ohta 1978), Slatkin (1995) recently proposed a statistic for microsatellites, R_{ST}, based on the variance of allele size. In this paper, both F- and R-statistics are estimated and compared to the exact G-test results to investigate their respective efficiencies in estimating differentiation.

Material and methods

Study area

The study area is located in the Swiss Prealpine range, in the Berner Oberland, canton of Bern, and consists of eight sampling sites located in five different valleys (Fig. 1): the Frutigtal, Engstliegental, Kandertal, Ueschinental and Gasteretal valleys (Table 1; Fig. 1). Altitudes of the sampling sites ranged from 710 m (site 8; Frutigtal) to 1960 m a.s.l. (site 1; Ueschinental). Samples were separated from each other by 3 to 14 km (Fig. 1). These valleys are isolated from each other to different extents. First, the Kandertal and Engstliegental are adjacent valleys originating from the Frutigtal valley. That is, they are first separated by very low, and progressively higher mountain ridges. Therefore, only weak

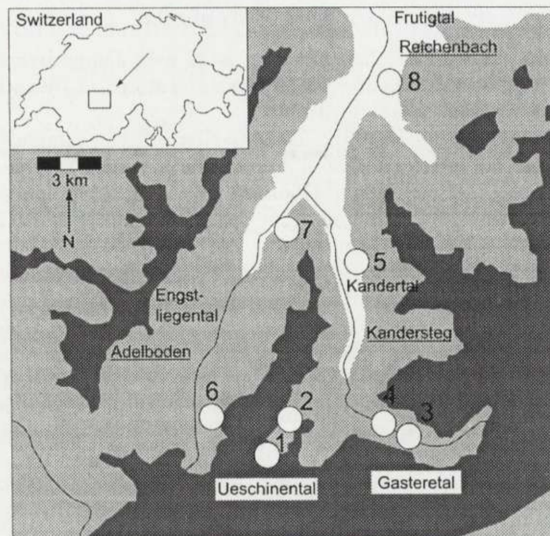

Fig. 1. Map of the study area showing the seven sites where *S. araneus* of the Vaud race were collected (additionally, *S. coronatus* was collected at site 8). Dark areas > 2000 m, grey areas > 1000 m a.s.l.; fine lines indicate rivers. Villages (underlined) and valleys names are indicated.

genetic differentiation is expected among these two valleys. The Ueschinental and Gasteretal valleys both originate from the Kandertal and can be viewed as geographically rather isolated. They are connected by lower passages in the north, which should allow dispersal. Ridges > 2100 m separate sample 1 from a direct connection to the Gasteretal (Fig. 1). The latter valley is surrounded by high mountain ridges (> 2100 m) except for its access from the Kandertal (Fig. 1). The southern Ueschinental and southern Engstliegental are geographically proximate, but separated by ridges at least 2400 m high and are expected to be well differentiated from each other.

Table 1. Details of each trapping site including sample size (n). *S. araneus* were collected as sites 1–7 and *S. coronatus* at site 8.

No.	Trapping site	Coordinates	Valley	Altitude (m)	n
1	Sulzweng	613.2/143.7	Ueschinental	1960	6
2	Balme	615/146.2	Ueschinental	1650	6
3	Staldi	620.7/144.7	Gasteretal	1430	6
4	Near Kander river	619.2–6/145.2	Gasteretal	1400	5
5	Ronewald	617.6/154	Kandertal	890	5
6	Huserweid	609.5/146.4	Engstliegental	1340	13
7	Meise	612.9–613.4/156.0–3	Engstliegental	930	5
8	Kien (Chiene river)	618.9/162.9	Frutigtal	710	3
Total					49

Sampling of individuals

Shrews (*S. araneus* and/or *S. coronatus*) were collected with Longworth traps from October to December 1997. Tissue samples were taken by toe-clipping and stored in 70% ethanol. All shrews were then immediately released at their collecting sites.

To verify the specific status of individuals collected in Kien (710 m, sample 8; Fig. 1), we preserved the toe-clippings in culture medium at 4°C and fibroblast cell cultures were set up. Chromosome preparations were made following Brünner and Hausser (1996) and G-banding performed according to a modified method from Seabright (1971). In addition, toe-clippings from a few other individuals were preserved in culture medium to confirm their racial status (Vaud race).

The Vaud chromosome race of *S. araneus* can be viewed as a genetic isolate mainly located in the Berner Oberland (Switzerland). To the north, its distribution is limited by the presence of *S. coronatus* in the lowlands. To the south, high mountain ridges separate the Berner Oberland from the Valais region of Switzerland, where another race, the Valais race, occurs. These two chromosome races are in contact in the Haslital valley, in the eastern Berner Oberland (Hausser *et al.* 1991). These two races are well differentiated from each other in terms of karyotypes, morphometrics, allozymes and mitochondrial DNA (Hausser *et al.* 1991, Taberlet *et al.* 1994). Genetic exchanges between these races are expected to be strongly reduced (Hausser *et al.* 1991, Brünner and Hausser 1996). *S. araneus* populations occurring in western Switzerland were formerly considered to belong to the Vaud race. Recently however, following the nomenclature of Hausser *et al.* (1994), they were recognised as a separate race, namely the Jura race (Hausser *et al.* 1994, Zima *et al.* 1996). However, genetic exchange is not expected between the Jura and Vaud races, since they are separated from each other by *S. coronatus*.

Genetic analysis of microsatellites

A total of 12 microsatellite loci were used for analysis: Loci L9, L16, L45, L57, L67 (Wyttenbach *et al.* 1997), L14, L33, L68, L92, L97 (Balloux *et al.* 1998) and two new loci, L13 and L99 (Genbank accession numbers: AF175741 and AF175744, respectively). PCR conditions are described in Balloux *et al.* (1998) and Lugon-Moulin *et al.* (1999b). Conditions for L13 and L99 are the same as for the loci described in Balloux *et al.* (1998). All samples were electrophoresed for 2.5–7.5 hours on a denaturing polyacrylamide gel (6%, 8M urea) along with a sequencing reaction and a clone or an individual of known size as size markers. Fixation, drying and autoradiography followed standard procedures (Sambrook *et al.* 1989).

Genetic polymorphism and overall structuring

In a first step, allele frequencies, observed (*Ho*) and expected heterozygosities within (*Hs*) and between (*Ht*) samples were calculated with the software package FSTAT 2.8, updated from Goudet (1995). Wright's (1951) *F*-statistics were estimated according to Weir and Cockerham (1984) using FSTAT 2.8. Within population heterozygote deficiency due to non random mating was estimated by F_{IS} and tested using 5000 permutations of alleles within samples. R_{ST} (Slatkin 1995) was estimated according to Rousset (1996). However, neither F_{ST} nor R_{ST} were tested as such. Population differentiation was tested using the exact *G*-test (Goudet *et al.* 1996). A total of 10 000 permutations of genotypes among samples were performed. The test probability was given as the proportion of *G*-values obtained from the permuted datasets which were as high or higher than the observed one. All calculations were performed using FSTAT 2.8.

Pairwise sample and valley comparisons

In a second step, pairwise sample comparisons were performed. Both F_{ST} and R_{ST} were estimated. Differentation was assessed using the exact *G*-test and significance tested using 15 000 permutations of genotypes among samples. To compare F_{ST} and R_{ST} in their efficiencies to detect structuring, a reference for assessing the degree of population differentiation is needed. In particular, neither F_{ST} nor R_{ST} can be used as reference since theoretical assumptions underly both estimators. We use the exact

G-test p-values as references. Indeed, the exact G-test will assess genetic differentiation without an underlying model and was shown to be the most powerful method to detect genetic differentiation in diploid populations (Goudet *et al.* 1996). Therefore, a comparison of these two statistics can be carried out. The lowest possible p-value is < 0.00007 since the highest possible number of permutations which can be performed is 15000 with FSTAT 2.8. Three matrices were constructed: one for F_{ST}, one for R_{ST} and the last for the exact G-test p-values. To test whether these two statistics (F_{ST} and R_{ST}) are correlated to the differentiation level (ie exact G-test), Mantel tests (Manly 1991) were performed. A total of 10000 randomizations of rows and columns of the matrix to be explained were carried out to test whether the amount of variance explained by the explanatory matrix (F_{ST} or R_{ST} matrix) significantly differs from zero. Computations were carried out using a program written by J. Goudet (available upon request from its creator at the same institutional address as the authors). We also tested whether p-values, F_{ST} and R_{ST} are correlated with geographical distances, using Mantel tests as described above.

In a third step, all *S. araneus* individuals collected in the same valley were pooled together to perform pairwise valley comparisons. This pooling procedure enables us to estimate the between-valley differentiation.

Evolutionary relationships among samples

Cavalli-Sforza and Edwards chord distances (Cavalli-Sforza and Edwards 1967) were calculated from allele frequency data using PHYLIP 3.57 (Felsenstein 1995). This distance was shown to perform well with microsatellites, outperforming distances designed for the stepwise mutation model (Takezaki and Nei 1996). A neighbor-joining tree was constructed (Saitou and Nei 1987). The tree was rooted using the sister species *S. coronatus*. To assess tree robustness, gene frequency data were bootstrapped 1000 times (SEQBOOT subroutine in PHYLIP). A consensus tree was computed using the CONSENSE subroutine in PHYLIP.

Results

Sampling and karyotypes

A total of 46 common shrews of the Vaud chromosome race were collected at seven localities (Table 1). The three individuals collected at the eighth locality (Kien; Table 1) were Millet's shrew (*S. coronatus*), as revealed by karyotype and microsatellite analysis (see below). Sample sizes per collecting site are given in Table 1. Three additional, but isolated *S. araneus* individuals were collected. They will be used in the pairwise valley comparisons (see below). The first of these shrews was collected in the upper part of the Engstliegental, about 1 km south-east of sample 6 (coordinates: 610/145.2; 1450 m); the two others were trapped between sample 1 and 2 in the Ueschinental (coordinates: 614.7/145.6; 1740 m).

Microsatellite polymorphism: differences between *S. araneus* and *S. coronatus*

Although the small *S. coronatus* sample does not enable any statistical interspecific inferences, a few interesting features neverthless appear. At locus L14, *S. coronatus* is monomorphic for an allele not found in *S. araneus*. At locus L97, the two alleles displayed by *S. coronatus* differed by a single base pair instead of two with all alleles found in *S. araneus*. At locus L57, a dinucleotide repeat, such discrepancies (ie single base pair differences) were also found among *S. araneus*

samples. Finally, no amplification product was observed at locus L13 for *S. coronatus*, suggesting a mutation at the primer annealing site.

However, the most interesting feature was clearly at locus L99: two alleles were found, one in *S. araneus* samples, the other allele being exclusively present in the three *S. coronatus*. These two species could therefore be characterized by alternative fixed allele at this locus. This result prompted us to test this locus on further individuals of *S. coronatus*. Moreover, individuals of three other alpine chromosome races of *S. araneus* (races Vaud, Valais and Cordon), *S. granarius* and *S. raddei*, were also tested. Nine individuals per taxon were analysed ($n = 54$). Locus L99 was only polymorphic in *S. raddei*. In all other taxa, it was fixed. *S. coronatus* and *S. araneus* Valais shared the same allele, whereas *S. araneus* Vaud and Cordon, as well as *S. granarius*, shared the alternative allele.

Population genetic structuring of the common shrew

Genetic variability of microsatellites in *S. araneus*

Locus L99 being monomorphic, thus uninformative for *S. araneus*, will no longer be considered in this study. The number of alleles ranged from 2 for locus L16 to 21 for locus L33 (Table 2), with an allele mean number of 11.1. The overall level of observed heterozygosity (*Ho*) was 0.64, with a range from 0.24 (L16) to 0.84 (L57). Expected heterozygosities within samples (*Hs*) ranged from 0.36 to 0.92, with an average of 0.70, whereas expected heterozygosities between samples (*Ht*) averaged 0.73 (range: 0.35–0.92; Table 2).

Table 2. Number of alleles (*Na*), observed heterozygosity (*Ho*), expected heterozygosities within (*Hs*) and between (*Ht*) *S. araneus* samples, per locus and over all loci.

Locus	*Na*	*Ho*	*Hs*	*Ht*
L33	21	0.779	0.851	0.895
L97	18	0.773	0.924	0.924
L68	8	0.761	0.728	0.801
L13	3	0.317	0.506	0.547
L14	9	0.694	0.795	0.829
L92	6	0.643	0.728	0.748
L45	9	0.473	0.458	0.487
L57	19	0.838	0.817	0.858
L9	17	0.729	0.799	0.818
L16	2	0.236	0.361	0.352
L67	10	0.802	0.755	0.803
All loci	122	0.640	0.702	0.733

Random mating within samples

The overall F_{IS} value is highly significant (over all loci, over all samples: F_{IS} = 0.082, $p < 0.001$). When looking at F_{IS} values for each locus (over all samples), only two loci out of 11 significantly differ from zero (L13 and L97: F_{IS} = 0.38 and 0.16 with $p = 0.006$ and 0.003, respectively). Locus L16 displays quite a high F_{IS} value (F_{IS} = 0.35). Wyttenbach *et al.* (1999b) suggested that locus L16 may show null alleles or be under selective pressure. Lugon-Moulin *et al.* (1999b) also noticed the peculiar behaviour of this locus. There may be similar problems with locus L13 and L97. Consequently, these three loci (L13, L16 and L97) will no longer be considered in this study. When these three loci are removed from the analysis, the overall F_{IS} is no longer significant (F_{IS} = 0.029, $p = 0.13$, Table 3). Over all loci F_{IS} values now range from −0.009 (site 6) to 0.127 for site 2 and no sampling site shows significant deviation from Hardy-Weinberg expectations.

Overall population structuring

F_{ST} values are rather concordant across loci, ranging from 0.026 (L9 and L92) to 0.088 (L68), with an averaged value of 0.054 (Table 3). R_{ST} on the other hand display a larger range of values than F_{ST}, from −0.002 (L67) to 0.196 (L9), with an averaged value of 0.114. Population differentiation assessed by the exact *G*-test is significant for all but one locus (Table 3). Population differentiation is highly significant over all loci (exact *G*-test: $p < 0.0001$).

Pairwise sample comparisons

F_{ST}, R_{ST}, the exact *G*-test *p*-values and geographical distances among all pairs of *S. araneus* samples are presented in Table 4. F_{ST} range from 0.007 to 0.120, and R_{ST} from −0.049 to as high as 0.501. Significant within-valley genetic structuring was evident between Engstliegental samples 6 and 7, and Ueschinental samples 1

Table 3. F_{IT}, F_{ST} and F_{IS}, as well as R_{ST} values, and exact *G*-test results, per locus and over all loci. Asterisks for the exact *G*-test indicate significant genetic structuring among samples: * – $p < 0.05$, ** – $p < 0.01$, *** – $p < 0.001$, ns – not significant).

Locus	F_{IT}	F_{ST}	F_{IS}	R_{ST}	Exact *G*-test
L33	0.121	0.060	0.065 ns	0.110	***
L68	0.046	0.088	−0.046 ns	0.018	***
L14	0.170	0.050	0.127 ns	0.033	**
L92	0.072	0.026	0.047 ns	0.090	*
L45	0.053	0.070	−0.018 ns	0.065	*
L57	0.063	0.048	0.016 ns	0.050	**
L9	0.086	0.026	0.062 ns	0.196	ns
L67	0.024	0.073	−0.053 ns	−0.002	**
All loci	0.082	0.054	0.029 ns	0.114	***

Table 4. Geographical distances (in kilometers), F_{ST}, R_{ST}, and p-values of the exact G-test among all *S. araneus* sample pairs. Asterisks indicate significant p-values (ie significant genetic differentiation among samples): * – $p < 0.05$, ** – $p < 0.01$, *** – $p < 0.001$, ns – not significant.

Sample pairs	Distances (km)	F_{ST}	R_{ST}	p
1–2	3.0	0.071	0.144	0.0343 *
1–3	7.6	0.066	0.140	0.0167 *
1–4	6.2	0.007	0.072	0.8343 ns
1–5	11.1	0.096	0.199	0.0039 **
1–6	4.5	0.054	0.258	< 0.00007 ***
1–7	12.4	0.120	0.501	0.0022 **
2–3	6.0	0.030	−0.005	0.0821 ns
2–4	4.8	0.023	0.007	0.2044 ns
2–5	8.3	0.038	−0.049	0.0347 *
2–6	5.5	0.026	0.011	0.0131 *
2–7	10.1	0.029	0.163	0.0257 *
3–4	1.5	0.023	−0.029	0.2591 ns
3–5	9.9	0.094	0.082	< 0.00007 ***
3–6	11.3	0.062	−0.019	< 0.00007 ***
3–7	13.8	0.085	0.154	0.0029 **
4–5	9.0	0.088	0.091	< 0.00007 ***
4–6	9.9	0.058	0.065	0.0024 **
4–7	12.7	0.085	0.320	< 0.00007 ***
5–6	11.2	0.046	0.095	0.0065 **
5–7	5.0	0.029	0.194	0.0648 ns
6–7	10.4	0.049	0.034	0.0168 *

and 2. In contrast, no significant genetic differentiation could be found among Gasteretal samples 3 and 4.

Several samples not collected in the same valley are not significantly differentiated from each other (Table 4). Samples from the Gasteretal are not differentiated from Ueschinental samples, except among samples 1 and 3. As expected, the proximate samples 5 (Kandertal) and 7 (Engstliegental) are not differentiated from each other. All other pairwise comparisons yielded significant results (Table 4). Samples located in the Gasteretal are highly significantly differentiated both from the Kandertal and Engstliegental samples (all $p < 0.003$). Also, the Ueschinental sample 1 is highly differentiated from both the Engstliegental and Kandertal samples.

Between-valley analysis

For this analysis, individuals obtained from the same valley were pooled together ($n = 49$; including the three isolated *S. araneus* individuals). No significant differentiation is found between the Gasteretal and Ueschinental valleys (Table 5), as suggested by the above pairwise sample comparisons. However, very highly significant ($p < 0.001$) differentiation is found in all pairwise comparisons including either the Ueschinental or the Gasteretal (Table 5). Weaker, but significant

Table 5. F_{ST}, R_{ST}, and p-values of the exact G-tests for all pairs of valleys. Asterisks indicate significant p-values (ie significant genetic differentiation among valleys): * – $p < 0.05$, ** – $p < 0.01$, *** – $p < 0.001$, ns – not significant.

Pair of valleys	F_{ST}	R_{ST}	p
Ueschinental/Engstliegental	0.020	0.148	< 0.0001 ***
Ueschinental/Kandertal	0.045	0.039	0.0004 ***
Ueschinental/Gasteretal	0.013	0.018	0.0723 ns
Gasteretal/Engstliegental	0.049	0.109	< 0.0001 ***
Gasteretal/Kandertal	0.083	0.109	0.0004 ***
Kandertal/Engstliegental	0.026	0.117	0.0158 *

differentiation is found among the Engstliegental and Kandertal valleys (p = 0.016). This result likely accounts for the differentiation among sample 5 and 6 (p = 0.006), since samples 5 and 7 are not differentiated (p = 0.06; Table 4). For the between-valley comparisons, F_{ST} values ranged from 0.013 to 0.083 whereas R_{ST} values ranged from 0.018 to 0.148 (over all loci; Table 5).

Since multiple pairwise tests were performed (27 tests = 21 pairwise sample and 6 pairwise valley tests), the risk of type I error is enhanced. When a sequential Bonferroni correction (Rice 1989) is applied to these multiple tests, all very highly significant p-values (ie $p < 0.001$) remain significant. Moreover, p-values < 0.01 also remain significant except for sample pairs 1–5 and 5–6 in Table 4.

Mantel tests

Neither the exact G-test p-values nor R_{ST} are correlated with geographical distances (Mantel tests: $R^2 = 0.13$, $p = 0.12$; $R^2 = 0.13$, $p = 0.10$, respectively). However, F_{ST} is significantly correlated with the geographical distances ($R^2 = 0.38$, $p = 0.003$). Interestingly, F_{ST} is significantly correlated with the exact G-test ($R^2 = 0.29$, $p = 0.006$) while, in sharp contrast, R_{ST} explains less than 5% of the variance contained in the p-values dataset ($p = 0.37$). Significant Mantel test results remain so after a sequential Bonferroni correction is applied.

Phylogenetic analysis

The consensus tree, reconstructed from Cavalli-Sforza and Edwards (1967) distances calculated from allele frequencies at eight microsatellite loci, reveals two major groupings (Fig. 2). These are in general agreement with above analyses. The first is formed by samples from the Ueschinental and Gasteretal (branch support: 73%). Gasteretal sample 4 and Ueschinental sample 1 are grouped in 80% of the resampled trees, and Gasteretal sample 3 is grouped to them in 88% of the resampled trees (Fig. 2). The second grouping is formed by Kandertal sample 5 and Engstliegental sample 7 (branch support: 75%). There is no clear relationship between Engstliegental sample 6 and the two major groupings.

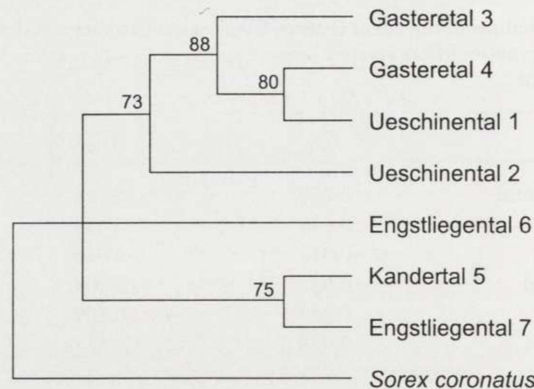

Fig. 2. 50% majority-rule consensus tree based on Cavalli-Sforza and Edwards (1967) chord distances derived from allele frequencies at eight microsatellites (neighbor-joining method of tree reconstruction). The tree was rooted with the sister species *Sorex coronatus*. Numbers at the nodes indicate the percentage of 1000 bootstrap replicates (> 50%) that supported the node. Sample numbers correspond to data given in Table 1. The valley in which samples were collected is also indicated.

Discussion

Utility of microsatellite locus L99 as a diagnostic marker

Locus 99 appears to be a new diagnostic marker for discrimating between *S. araneus* of the Vaud and Cordon chromosome races, from *S. araneus* of the Valais race and *S. coronatus*. This locus is therefore of particular interest in studying hybrid zones between the Vaud and Valais races and between the Cordon and Valais races (Hausser *et al.* 1991, Brünner and Hausser 1996, Lugon-Moulin *et al.* 1996, 1999a, b). Another method, based on albumin electrophoresis, displays the same biallelic pattern as locus L99 (Neet and Hausser 1991). *S. coronatus* and the Valais race share a slow serum albumin band, while the Vaud and Cordon races possess a rapid albumin. It is interesting to further note that *S. coronatus* and the Valais race share the metacentric chromosome *lo* (arm designation following Searle *et al.* 1991), although it is fixed in *S. coronatus* and polymorphic in the Valais race (ie arms *l* and *o* are either found as acrocentric chromosomes or fused together as a metacentric). In contrast, the Vaud and Cordon races, as well as all other known races of *S. araneus*, never possess this metacentric *lo* (eg Hausser *et al.* 1991, Zima *et al.* 1996). These concordant results across several classes of markers (microsatellites, albumin proteins, chromosomes) suggest possible ancient hybridization events between *S. coronatus* and the Valais race, possibly when arms *l* and *o* were not yet fixed. However, retention of ancestral polymorphism as opposed to ancestral hybridization cannot be excluded.

The utility of locus L99 in identifying individuals of *S. coronatus* from alpine races of *S. araneus* (except the Valais race) in the field is noteworthy. Although

chromosome analysis and urinary pepsinogens (Neet and Hausser 1991) enable discrimination between *S. coronatus* and all alpine races of *S. araneus* (ie including the Valais race), these two techniques also have their drawbacks. Direct preparation of karyotypes implies the sacrifice of the animals. The alternative, indirect method using cell cultures is both time-consuming and expensive. The biochemical technique (urinary pepsin electrophoresis) is somewhat limited since it appears that samples must be analysed within two weeks or less, and field sampling is not practical (Neet and Hausser 1991).

The microsatellite-based method is practical in the field. Sampling by toe--clipping is straightforward. Only minor quantities of DNA enable successful PCR amplification. Although the validity of locus L99 as diagnostic marker has to be confirmed on a larger dataset, its use for field work in the Swiss Alps appears promising. Moreover, when combined with other loci (L14, L13), it should also enable the discrimination of *S. coronatus* and *S. araneus* Valais.

Population structuring of the common shrew

The average F_{ST} value reported in the present study (F_{ST} = 0.054) is higher than previously reported values for this species using microsatellites. Using these genetic markers, Wyttenbach and Hausser (1996) reported a value of F_{ST} = 0.015–0.026 in a study of populations of the Bretolet race found on a 6 km stretch in a valley in the Morzine region (Western Alps, France). Using samples found in four different valleys, Wyttenbach *et al.* (1999a) detected a significant genetic structuring (F_{ST} = 0.032), but genetic differentiation between valleys was insignificant. Lugon-Moulin *et al.* (1999b) estimated a similar F_{ST} = 0.033 among samples of the Valais race found in or at proximity of a hybrid zone with the Cordon race. These samples formed a transect of about 6 km in the Arve valley (Western Alps, France). This value reflected however mainly introgression from the Cordon race (Lugon-Moulin *et al.* 1999a). On a 7 km transect, Lugon-Moulin *et al.* (1996) did not detect isolation by distance using partial Mantel tests. The geographical scale of the present study is slightly larger than those described above. Roughly 14 km separate the two most distant samples (see Fig. 1; Table 4). Mantel test show that F_{ST} is significantly correlated with geographical distances. However, because sampling was purposely performed in different valleys separated by ridges of various heights, other factors such as the effect of mountain ridges may partly account for this correlation. Such spurious correlations were found by Lugon--Moulin *et al.* (1996, 1999a) who studied the genetic structuring of *S. araneus* populations sampled along a transect. Using multiple matrices Mantel tests, these authors found that the effect of geographical distance was no longer significant when other possible effects were accounted for. On the other hand, the exact *G*-test is surprisingly not significantly correlated with geographical distances.

Significant differentiation occurs within as well as among valleys. Significant genetic structuring occurs between samples from the Ueschinental and Gasteretal relative to the other two valleys, as further suggested by the phylogenetic analysis

clustering together the samples collected in these two valleys. However, no significant differentiation occurs between the Ueschinental and Gasteretal. Dispersal probably occurs via the northern parts of these valleys (see Fig. 1), although it cannot be ruled out that direct dispersal occurs between sample 1 and the Gasteretal. In contrast, Ueschinental sample 1 and Engstliegental sample 6, only 4.5 km apart, are significantly differentiated from each other (exact G-test: $p < 0.00007$). The most likely explanation for the significant structuring among these two rather proximate samples is the presence of the high mountain ridges (> 2400 m) separating them. The ridges could act as a barrier to dispersal because they represent rocky habitats with a lack of vegetation cover. The allelic distributions among these two samples are quite different, as suggested by the high R_{ST} value ($R_{ST} = 0.258$), and direct gene exchanges (ie over the ridges) should be a rare event among these two samples. Both Gasteretal samples are differentiated from the Kandertal valley, despite the fact that rather continuous forest habitat join these two valleys. It is possible that other unsuitable areas for dispersal exist. Other processes which would result in genetic drift (extinction/re-colonization, fluctuation in population size) could also occur. More and larger samples would be required to test some of these hypotheses.

Comparison of *F*- and *R*-statistics

The use of microsatellite markers has dramatically increased in the last few years. It is becoming evident that mutation models and patterns of microsatellites are complex (eg Estoup and Cornuet 1999). High mutation rates are predicted to decrease F_{ST} values (eg Slatkin 1995, Hedrick 1999). Recently, Balloux *et al.* (2000a) used simulations to show that F_{ST} is quite sensitive to polymorphism when the mutation rate is higher than the migration rate. Therefore, this estimator will underestimate true levels of differentiation when migration is low and hence, overestimate migration (eg Hedrick 1999, Balloux *et al.* 2000a, b). On the other hand, R-statistics, devised for the stepwise mutation model, are not sensitive to polymorphism. However, R_{ST} estimation is seriously affected if mutation only slightly deviates from a strict stepwise pattern (Balloux *et al.* 2000a). This is the case for several studies (eg Angers and Bernatchez 1998) and for several loci used in the present study (Wyttenbach *et al.* 1999a). Because microsatellites are characterized by a high polymorphism and their mutation process likely deviates from a strict stepwise mutation model, both estimators will underestimate the true level of differentiation.

However, the difference between these two estimators do not only depend on the mutation rate and model (although their effect can be quite important; Balloux *et al.* 2000a), but also on differences in average coalescence time between populations (Slatkin 1995). In a cursory comparison of different recent studies in which both F- and R-statistics were used, Lugon-Moulin *et al.* (1999b) noticed that when R_{ST} is higher than or equal to 0.06, the corresponding F_{ST} is usually lower. It appears that when levels of differentiation are high, R_{ST} is better at unravelling structuring

than F_{ST}. Our mean R_{ST} is more than twice as large as the mean F_{ST} (0.114 vs 0.054), fitting the above observation.

Mantel tests indicate that F_{ST}, unlike R_{ST}, is significantly correlated to the degree of differentiation. As expected when using markers with high mutation rates (Slatkin 1995), F_{ST} values are rather low, even at high differentiation levels (highest $F_{ST} = 0.120$; exact G-test: $p < 0.0022$; Table 4). But interestingly, the F-statistics, even if deflated, remains correlated with the differentiation levels. Comparisons of F_{ST} values across different studies can be carried out, but the polymorphism of the loci used should be comparable and mutation rate should not exceed migration levels (see Balloux *et al.* 2000b). A different picture appears with R_{ST}. This statistics is known to have a large variance (Slatkin 1995). In the present study, at the highest levels of structuring (corresponding to $p < 0.00007$), R_{ST} values range from as low as -0.019 to as high as 0.320 (Table 4), although our observations above suggest that R_{ST} should be less biased when levels of differentiation are higher. R_{ST} also displays a high variance at lower differentiation levels (Table 3 and 4). These results cast some doubt about the utility of R-statistics with microsatellites. Larger datasets are needed to validate our results.

Conclusion

This study brings a better understanding of the potential effect of mountain ridges on isolating populations from adjacent valleys. Our results suggest that dispersal should be reduced at altitudes > 2400 m, affecting the population structure of populations. Our mean F_{ST} of 0.054 is higher than previously reported values for this species using microsatellites. It should at least partly be accounted for by topography, although isolation by distance is correlated to F_{ST}. Being free of any assumption of a mutation model, the exact G-test should be a valuable tool in (a) assessing the genetic differentiation of diploid population, and (b) the reliability and efficiency of F_{ST} and R_{ST} in estimating the degree of this differentiation.

Aknowledgements: We wish to thank J. Aeschimann, H. Brünner and E. Lugon-Moulin for invaluable help in the field; A.-M. Mehmeti for assistance in cell cultures and karyotype preparations; N. diMarco for extractions, P. Moratal for help in elaborating the map, J. Goudet for kindly providing the program used to perform Mantel tests, and two anonymous referees for helpful comments on the manuscript.

References

Angers B. and Bernatchez L. 1998. Combined use of SMM and non-SMM methods to infer fine structure and evolutionary history of closely related brook charr (*Salvelinus fontinalis*, Salmonidae) populations from microsatellites. Molecular Biology and Evolution 15: 143–159.

Balloux F., Brünner H., Lugon-Moulin N., Hausser J. and Goudet J. 2000a. Microsatellites can be misleading: an empirical and simulation study. Evolution, in press.

Balloux F., Ecoffey E., Fumagalli L., Goudet J., Wyttenbach A. and Hausser J. 1998. Microsatellite conservation, polymorphism, and GC content in shrews of the genus *Sorex* (Insectivora, Mammalia). Molecular Biology and Evolution 15: 473–475.

Balloux F., Lugon-Moulin N. and Hausser J. 2000b. Estimating gene flow across hybrid zones: how reliable are microsatellites? [In: Evolution in the *Sorex araneus* group: Cytogenetic and molecular aspects. J. B. Searle and J. M. Wójcik, eds]. Acta Theriologica 45, Suppl. 1: 93–101.

Brünner H. and Hausser J. 1996. Genetic and karyotypic structure of a hybrid zone between the chromosomal races Cordon and Valais in the common shrew, *Sorex araneus*. Hereditas 125: 147–158.

Cavalli-Sforza L. L. and Edwards A. W. F. 1967. Phylogenetic analysis: models and estimation procedures. American Journal of Human Genetics 19: 233–257.

Estoup A. and Cornuet J.-M. 1999. Microsatellite evolution: inferences from population data. [In: Microsatellites, evolution and applications. D. B. Goldstein and C. Schlötterer, eds]. Oxford University Press, Oxford: 49–65.

Felsenstein J. 1995. PHYLIP. Phylogeny Inference Package. Version 3.57c. University of Washington, Seattle.

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. and Rousset F. 1996. Testing differentiation in diploid populations. Genetics 144: 1933–1940.

Hausser J., Bosshard E., Taberlet P. and Wójcik J. 1991. Relationship between chromosome races and species of the *Sorex* of the *araneus* group in the western Alps. Mémoires de la Société Vaudoise de Sciences Naturelles 19: 79–95.

Hausser J., Fedyk S., Fredga K., Searle J. B., Volobouev V., Wójcik J. M. and Zima J. 1994. Definition and nomenclature of the chromosome races of *Sorex araneus*. Folia Zoologica 43, Suppl. 1: 1–9.

Hedrick P. W. 1999. Perspective: Highly variable loci and their interpretation in evolution and conservation. Evolution 53: 313–318.

Jarne P. and Lagoda P. J. L. 1996. Microsatellites, from molecules to populations and back. Trends in Ecology and Evolution 11: 424–429.

Kimura M. and Ohta T. 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. Proceedings of the National Academy of Sciences of the USA 75: 2868–2872.

Lugon-Moulin N., Brünner H., Balloux F., Hausser J. and Goudet J. 1999a. Do riverine barriers, history or introgression shape the genetic structuring of a common shrew (*Sorex araneus*) population? Heredity 83: 155–161.

Lugon-Moulin N., Brünner H., Wyttenbach A., Hausser J. and Goudet J. 1999b. Hierarchical analyses of genetic differentiation in a hybrid zone of *Sorex araneus* (Insectivora, Soricidae). Molecular Ecology 8: 419–431.

Lugon-Moulin N., Wyttenbach A., Brünner H., Goudet J. and Hausser J. 1996. Study of gene flow through a hybrid zone in the common shrew (*Sorex araneus*) using microsatellites. Hereditas 125: 159–168.

Manly B. J. F. 1991. Randomization and Monte Carlo methods in biology. Chapman and Hall, London: 1–281.

Neet C. R. and Hausser J. 1991. Biochemical analysis and determination of living individuals of the Alpine karyotypic races and species of the *Sorex araenus* group. Mémoires de la Société Vaudoise de Sciences Naturelles 19: 97–106.

Reiter G. and Winding N. 1997. Verbreitung und Ökologie alpiner Kleinsäuger (Insectivora, Rodentia) an der Südseite der Hohen Tauern, Österreich. Wissenschaftliche Mitteilungen aus dem Nationalpark Hohe Tauern 3: 97–135.

Rice W. R. 1989. Analyzing tables of statistical tests. Evolution 43: 223–225.

Rousset F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation process. Genetics 142: 1357–1362.

Saitou N. and Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4: 406–425.

Sambrook J., Fritsch E. F. and Maniatis T. 1989. Molecular cloning. Cold Spring Harbor Lab Press. Cold Spring Harbor, NY.

Seabright M. 1971. A rapid banding technique for human chromosomes. Lancet 2: 971–972.

Searle J. B., Fedyk S., Fredga K., Hausser J. and Volobouev V. T. 1991. Nomenclature for the chromosomes of the common shrew *Sorex araneus*. Mémoires de la Société Vaudoise de Sciences Naturelles 19: 13–22.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139: 457–62.

Taberlet P., Fumagalli L. and Hausser J. 1994. Chromosomal versus mitochondrial DNA evolution: tracking the evolutionary history of the southwestern European populations of the *Sorex araneus* group (Mammalia, Insectivora). Evolution 48: 623–636.

Takezaki N. and Nei M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics 144: 389–399.

Weir B. S. and Cockerham C. C. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38: 1358–1370.

Winding N., Werner S., Slotta-Bachmayer L., Lindner R. and Illich I. 1995. Zoologische Dauerbeobachtung im Sonderschutzgebiet Piffkar. Unpublished manuscript. Nationalparkinstitut des Hauses der Natur, Salzburg.

Wójcik J. M. and Wójcik A. M. 1994. Protein variation in the common shrew (*Sorex araneus* L.) in Poland, in relation to karyotype. Folia Zoologica 43, Suppl. 1: 53–61.

Wright S. 1951. The genetical structure of populations. Annals of Eugenics 15: 323–354.

Wright S. 1965. The interpretation of population structure by *F*-statistics with special regard to systems of mating. Evolution 19: 395–420.

Wyttenbach A., Favre L. and Hausser J. 1997. Isolation and characterization of simple sequence repeats in the genome of the common shrew. Molecular Ecology 6: 797–800.

Wyttenbach A., Goudet J., Cornuet J.-M. and Hausser J. 1999a. Microsatellite variation reveals low genetic subdivision in a chromosome race of *Sorex araneus* (Mammalia, Insectivora). Journal of Heredity 90: 323–327.

Wyttenbach A. and Hausser J. 1996. The fixation of metacentric chromosomes during the chromosomal evolution in the common shrew (*Sorex araneus*, Insectivora). Hereditas 125: 209–217.

Wyttenbach A., Narain Y. and Fredga K. 1999b. Genetic structuring and gene flow in a hybrid zone between two chromosome races of the common shrew (*Sorex araneus*, Insectivora) revealed by microsatellites. Heredity 82: 79–88.

Zima J., Fedyk S., Fredga K., Hausser J., Mishta A., Searle J. B., Volobouev V. T. and Wójcik J. M. 1996. The list of the chromosome races of the common shrew (*Sorex araneus*). Hereditas 125: 97–107.

Received 1 September 1999, accepted 28 January 2000.