UNIL | Université de Lausanne

Faculté de biologie
et de médecine

Département d'Ecologie et d'Evolution

Chromosomal rearrangements and genetic structure in the shrews of the *Sorex araneus* group

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine

de l'Université de Lausanne

par

Patrick Basset

Diplômé en Biologie

Université de Lausanne

Jury

Prof. Kaethi Geering, Présidente
Prof. Jacques Hausser, Directeur de thèse
Prof. Jeremy Searle, Expert
Dr. Luca Fumagalli, Expert

Lausanne 2006

Unil

UNIL | Université de Lausanne

Faculté de biologie et de médecine

Ecole Doctorale

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président	*Madame Prof. Kaethi* ***Geering***
Directeur de thèse	*Monsieur Prof. Jacques* ***Hausser***
Rapporteur	
Experts	*Monsieur Prof. Jeremy B.* ***Searle***
	Monsieur Prof. Luca ***Fumagalli***

le Conseil de Faculté autorise l'impression de la thèse de

Monsieur Patrick Basset

Biologiste diplômé de l'Université de Lausanne

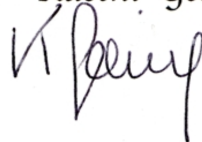
intitulée

Chromosomal rearrangements and gentetic structure in the shrews of the *Sorex araneus* group

Lausanne, le 10 février 2006

pour Le Doyen
de la Faculté de Biologie et de Médecine

Prof. Kaethi Geering



CONTENT

ACKNOWLEDGEMENTS **7**

ABSTRACT **9**

RESUME **11**

GENERAL INTRODUCTION **15**

CHROMOSOMAL SPECIATION 16
THE SHREWS OF THE SOREX ARANEUS GROUP 17
QUESTIONS ADDRESSED 21

CHAPTER 1 : GENETIC AND KARYOTYPIC STRUCTURE IN THE *SOREX ARANEUS* GROUP : ARE THEY INDEPENDENT ? **23**

ABSTRACT 24
INTRODUCTION 25
MATERIAL AND METHODS 27
RESULTS 32
DISCUSSION 38
ACKNOWLEDGMENTS 42

CHAPTER 2 : SPECIES IDENTIFICATION BY GENETIC ASSIGNMENT METHODS : WHAT ABOUT HYBRID ZONES ? **43**

ABSTRACT 44
INTRODUCTION 45
MATERIAL AND METHODS 46
RESULTS 48
DISCUSSION 52
ACKNOWLEDGEMENTS 55

CHAPTER 3 : CHROMOSOME LOCALIZATION OF MICROSATELLITE MARKERS IN THE SHREWS OF THE *SOREX ARANEUS* GROUP **57**

ABSTRACT 58
INTRODUCTION 59
MATERIAL AND METHODS 60
RESULTS 64
DISCUSSION 68
ACKNOWLEDGEMENTS 71

CHAPTER 4 : CHROMOSOME REARRANGEMENTS AND GENETIC DIFFERENTIATION AT DIFFERENT EVOLUTIONARY LEVELS OF THE *SOREX ARANEUS* GROUP **73**

ABSTRACT 74
INTRODUCTION 75
MATERIAL AND METHODS 77
RESULTS 81
DISCUSSION 86
ACKNOWLEDGEMENTS 89

CHAPTER 5 : RESTRICTED GENE FLOW AT SPECIFIC PARTS OF THE SHREW GENOME IN CHROMOSOMAL HYBRID ZONES **91**

ABSTRACT 92
INTRODUCTION 93
MATERIAL AND METHODS 97
RESULTS 99
DISCUSSION 108
ACKNOWLEDGMENTS 113

GENERAL CONCLUSIONS AND PERSPECTIVES **115**

REFERENCES **120**

APENDIXES **134**

ACKNOWLEDGEMENTS

First of all, I wish to thank my family for their constant love, patience and support during easy and less easy moments of this thesis. I would never have been able to complete this work without your help and a large part of the results obtained are yours!

My warmest thanks go to Prof. Jacques Hausser, my supervisor, who not only gave me the opportunity to write a thesis on a thrilling subject but also communicated to me his general enthusiasm and his particular interest for shrews.

I wish to warmly thank Glenn Yannic, who gave a large contribution to all the chapters of this thesis. It was a real pleasure to collaborate and imagine revolutionary projects about shrews and speciation.

I also thank Aline Dépraz, the “Tri” part of the “TriSo” group for vivacious discussions and especially for valuable comments and corrections of several chapters.

This thesis has benefited from the precious help or inspiration of many people: I wish especially to express my gratitude to Agata Banaszek, Christian Bernasconi, Laura Brändli, Thomas Broquet, Guillaume Evanno, Pierre Fontanillas, Lori Lawson Handley, Cynthia Hüsler, Julie Jaquiéry, Valérie Vogel and John Wang for their precious advices concerning lab work, data analyses and manuscripts writing.

I am grateful to Prof. Jeremy Searle, Luca Fumagalli and Kaethi Geering who kindly accepted to read and evaluate this study.

I also express my gratitude to Fabiana Bernasconi-Quadroni and Anne-Marie Mehmeti who helped me in the lab and karyotype preparations and who shared my hope to laser micro-dissect shrew chromosomes; to Nadège Duvoisin for a large piece of advice during lab work and animated discussions about “Calcio”; to Valérie Alonso, Sanije Jagodini, Emanuelle Müller who helped me with the lab work; to Patrick Moratal and Blaise Pavillard for technical assistance; to Fabien Fivaz, Nicolas Lugon-Moulin and Glenn Yannic who helped me during

my few excursions in the field to trap shrews; to Harald Brünner for letting me use his samples from the two Alpine hybrid zones of the *Sorex araneus* group.

I particularly thank all the people with whom I spent fundamental "Breaks" at 15h30 (or at other moments) and especially: Christian Bernasconi, Laura Brändli, Aline Déparz, Nadège Duvoisin, Véronique Helfer, Julie Jaquiéry, Jérôme Pellet, Glenn Yannic and Valérie Vogel for the friendly atmosphere.

I also would like to thank all the people of the DEE with whom I had the opportunity to share scientific and not so scientific discussions.

Finally, I wish to thank all the participants of the ISACC ("International *Sorex araneus* Cytogenetics Committee") meetings in Paris (2002) and St Petersbourg (2005) for stimulating discussions about shrews.

Financial support was provided by the following institutions: the Swiss National Funds for Scientific Research (grant n° 3100-05943 to Jacques Hausser), the Agassiz Foundation (Lausanne University), the Société Académique Vaudoise (SAV) and the Swiss Zoological Society.

ABSTRACT

The role of chromosomal rearrangements in the speciation process is much debated and many theoretical models have been developed. The shrews of the *Sorex araneus* group offer extraordinary opportunities to study the relationship between chromosomal variation and speciation. Indeed, this group of morphologically very similar species received a great deal of attention due to its karyotypic variability, which is mainly attributed to Robertsonian fusions.

To explore the impact of karyotypic changes on genetic differentiation, we first studied the relationship between genetic and karyotypic structure among Alpine species and among chromosome races of the *S. araneus* group using Bayesian admixture analyses. The results of these analyses confirmed the taxonomic status of the studied species even though introgression can still be detected between species. Moreover, the strong spatial sub-structure highlighted the role of historical factors (e.g. geographical isolation) on genetic structure.

Next, we studied gene flow at the chromosome level to address the question of the impact of chromosomal rearrangements on genetic differentiation. We used flow sorted chromosomes from three different karyotypic taxa of the *S. araneus* group to map microsatellite markers at the chromosome arm level. We have been able to map 24 markers and to show that the karyotypic organisation of these taxa is well conserved, which suggests that these markers can be used for further inter-taxa studies.

A general prediction of chromosomal speciation models is that genetic differentiation between two taxa should be larger across rearranged chromosomes than across chromosomes common to both taxa. We combined two approaches using mapped microsatellites to test this prediction. First, we studied the genetic differentiation among five shrew taxa placed at different evolutionary levels (i.e. within and among species). In this large scale study, we detected an overall significant difference in genetic structure between rearranged vs. common chromosomes. Moreover, this effect varied among pairwise comparisons, which allowed us to differentiate the role of the karyotypic complexity of hybrids and of the evolutionary divergence between taxa. Secondly, we compared the levels of gene flow measured across common vs. rearranged chromosomes in two karyotypically different hybrid zones (strong vs. low complexity of hybrids), which show similar levels of genetic structure. We detected a significantly stronger genetic structure across rearranged chromosomes in the hybrid zone showing the highest level of hybrid complexity. The large variance observed among loci suggested that other factors, such as the position of markers within the chromosome, also certainly affects genetic structure. In conclusion, our results strongly support the role of

chromosomal rearrangements in the reproductive barrier and suggest their importance in the speciation process of the *S. araneus* group.

RESUME

Le rôle des réarrangements chromosomiques dans les processus de spéciation est fortement débattu et de nombreux modèles théoriques ont été développés sur le sujet. Les musaraignes du groupe *Sorex araneus* présentent de nombreuses opportunités pour étudier les relations entre les variations chromosomiques et la spéciation. En effet, ce groupe d'espèces morphologiquement très proches a attiré l'attention des chercheurs en raison de sa variabilité caryotypique principalement attribuée à des fusions Robertsoniennes.

Pour explorer l'impact des changements caryotypiques sur la différenciation génétique, nous avons tout d'abord étudié les relations entre la structure génétique et caryotypique de races chromosomiques et d'espèces alpine du groupe *S. araneus* en utilisant des analyses Bayesiennes d' « admixture ». Les résultats de ces analyses ont confirmé le statut taxonomique des espèces étudiées bien que nous ayons détecté de l'introgression entre espèces. L'observation d'une sous structure spatiale relativement forte souligne l'importance des facteurs historiques (telle que l'isolation géographique) sur la structure génétique de ce groupe.

Ensuite, nous avons étudié le flux de gène au niveau des chromosomes pour aborder de manière directe la question de l'impact des réarrangements chromosomiques sur la différenciation génétique. En conséquence, nous avons utilisé des tris de chromosomes de trois taxons du groupe *S. araneus* pour localiser des marqueurs microsatellites au niveau du bras chromosomique. Au cours de cette étude, nous avons pu localiser 24 marqueurs et montrer une forte conservation dans l'organisation du caryotype de ces taxa. Ce résultat suggère que leur utilisation est appropriée pour des études entre taxa.

Une prédiction générale à tous les modèles de spéciation chromosomique correspond à la plus grande différenciation génétique des chromosomes réarrangés que des chromosomes communs. Nous avons combiné deux approches utilisant des microsatellites localisés au niveau du bras chromosomique pour tester cette prédiction. Premièrement, nous avons étudié la différenciation génétique entre cinq taxa du groupe *S. araneus* se trouvant à des niveaux évolutifs différents (i.e. à l'intérieur et entre espèce). Au cours de cette étude, nous avons détecté une différenciation globale significativement plus élevée sur les chromosomes réarrangés. Cet effet varie entre les comparaisons, ce qui nous a permis de souligner le rôle de la complexité caryotypique des hybrides et du niveau de divergence évolutive entre taxa. Deuxièmement, nous avons comparé le flux de gènes des chromosomes communs et réarrangés dans deux zones d'hybridation caryotypiquement différentes (forte vs. faible

complexité des hybrides) mais présentant un niveau de différenciation génétique similaire. Ceci nous a permis de détecter une structure génétique significativement plus élevée sur les chromosomes réarrangés au centre de la zone d'hybridation présentant la plus grande complexité caryotypic. La forte variance observée entre loci souligne en outre le fait que d'autres facteurs, tel que la position du marqueur sur le chromosome, affectent probablement aussi la structure génétique mesurée. En conclusion, nos résultats supportent fortement le rôle des réarrangements chromosomiques dans la barrière reproductive entre espèces ainsi que leur importance dans les processus de spéciation des musaraignes du groupe *S. araneus*.

GENERAL INTRODUCTION

How do new species emerge? This apparently simple question is still one of the most complex and controversial issue in evolutionary biology. One of the central topics in studies of speciation is the emergence of reproductive isolation. Numerous parameters have been proposed to play important roles in the evolution of the reproductive barrier between two populations. These parameters favoured the development of various models of speciation that can be classified according to the geographic mode of isolation (e.g. allopatry, parapatry, sympatry) or to the factors directly contributing to reproductive isolation (e.g. ecological, behavioural, genetic or chromosomal barriers). The literature concerning these models is extremely large (for reviews, see White 1978, King 1993, Howard & Berlocher 1998, Coyne & Orr 2004 or the recent special numbers of *Trends in Ecology and Evolution* 16(7) 2001 and *PNAS* 102(suppl.1) 2005). However, it is not my topic to discuss the particularities of each of these models. One of the factors commonly proposed to play an important role in the establishment of the reproductive barrier between populations is the variation of the karyotype. The models accounting for this variation are generally known as models of chromosomal speciation and will be of special interest throughout this study.

Chromosomal Speciation

It has often been noticed that closely related species can be distinguished by chromosomal changes whereas comparable rearrangements are only infrequently polymorphic within populations (e.g. King 1993). This observation led several authors to develop a large number of models proposing that chromosomal changes accelerate genetic differentiation between populations and therefore facilitate speciation (for reviews, see King 1993, Spirito 1998, Riesberg 2001, Coyne & Orr 2004, Ayala & Coluzzi 2005, Butlin 2005). These models can be separated into two main classes named the “hybrid dysfunction” and the “suppressed recombination” models of speciation (Ayala & Coluzzi 2005).

“Hybrid dysfunction models” claim that changes in the chromosome structure (i.e. chromosome number, chromosomal rearrangements) cause meiotic problems when heterozygous what will reduce the fertility and the reproductive fitness of heterozygous hybrids (White 1978, King 1993). Several authors (Wallace 1959, Lewis 1966, Grant 1981, White 1978, Baker & Bickham 1986) have offered a variety of models to account for evidences observed in plants or animals (for reviews, see Spirito 1998, Riesberg 2001). However, these models generally suffer from both empirical and theoretical difficulties (Rieseberg 2001, Navarro & Barton 2003a). For example, many chromosomal rearrangements

have little effects on fertility (Sites & Moritz 1987, Coyne *et al.* 1993, Coyne & Orr 2004). Moreover, a chromosomal rearrangement that has a major effect on fitness in heterozygotes is unlikely to be fixed whereas a rearrangement with little detrimental effects may spread more easily but will contribute little to reproductive isolation (Spirito 1998, but see White 1978 or Baker & Bickham 1986).

"Suppressed recombination models" account for a reduction or a suppression of recombination in heterokaryotypes. Chromosomal rearrangements influence recombination in a variety of ways: e.g. bringing pairs of loci that were unlinked into close linkage reduce or prevent recombination in heterozygotes and influence the distribution of crossing-over (Navarro & Barton 2003a). Therefore, since speciation can be viewed as the evolution of restrictions on the freedom of recombination, it is tempting to suggest that chromosomal rearrangements might play a role in speciation (Butlin 2005). Several of these models have recently been proposed by Rieseberg (2001), Noor *et al.* (2001) or Navarro & Barton (2003a) to account for speciation events between sunflowers, flies and human species respectively. Suppression of recombination in heterozygous individuals was reported in the case of mice (Davisson & Akeson 1993; Haigis & Dove 2003), but these characteristics are not restricted to Mammals (Marti & Bidau 1995; Rieseberg *et al.* 1999).

Most studies on the genetics of reproductive isolation and speciation have concentrated on model or laboratory species (e.g. the genus *Drosophila*). However, it is now possible and particularly interesting to extend these studies to other organisms. Besides to study the impact of chromosomal changes on the speciation process, it is important to work on groups showing large numbers of chromosomal rearrangements among taxa placed at different evolutionary levels (i.e. from chromosome variants to "full" species). Among mammals, the shrews of the *Sorex araneus* group are especially informative to study the relations between chromosomal variation and speciation.

The shrews of the Sorex araneus group

The shrews of the *Sorex araneus* group (Meylan & Hausser 1973) are small Insectivores with a large and mostly Palaearctic distribution. All the species included in this group (*Sorex araneus*, *S. antinorii*, *S. arcticus*, *S. asper*, *S. caucasicus*, *S. coronatus*, *S. daphaenodon*, *S. granarius*, *S. maritimensis* and *S. tundrensis*; Lugon-Moulin 2003) bear a karyotype with a low diploid number of chromosomes (2N = 20 – 42; Zima *et al.* 1998). Additionally, this group is characterised by the presence of a particular sexual chromosome system XY1Y2 in

males and XX in females (Sharman 1956). The X chromosome originated by the fusion of the ancestral X and one acrocentric autosome (Zima *et al.* 1998). The acrocentric homologue of the translocated autosome remained in its larger part unchanged and is designated Y2. The smaller odd acrocentric (Y1) represents the original Y chromosome (Pack *et al.* 1993). This composite sex chromosome system occurs invariably in all species and is commonly believed to be a character evidencing the monophyletic origin of this group.

Chromosomal polymorphism

The shrews of the *Sorex araneus* group have additionally received great deals of attention because of their highly variable karyotype. Therefore, the karyotype nomenclature (Searle *et al.* 1991) of *S. araneus* is noteworthy. Each of the 21 chromosome arm is labelled by a letter (*a* – *u*), with "*a*" indicating the largest arm and "*u*" the smallest. This nomenclature could easily be extended to inter-specific comparisons, at least in the western clade of the *S. araneus* group (Hausser 1994). Comparative analyses of banded karyotypes revealed high levels of chromosome arms homology (Volobouev 1989, Volobouev & Catzeflis 1989, Volobouev & Dutrillaux 1991) and karyotype differences between species can mainly be attributed to Robertsonian changes accompanied by telomere-centromere tandem translocations, centromeric shifts and pericentric inversions (Volobouev 1989).

In the type species of this group, *Sorex araneus*, Robertsonian polymorphisms are particularly prevalent. The ancestral karyotype of this species most likely consists of acrocentric chromosomes (for a review, see Wójcik *et al.* 2002) and most of the karyotypic variation is thought to have arisen through Robertsonian fusions, where two acrocentrics are combined to form a metacentric chromosome. Additional processes (i.e. Robertsonian fissions, Whole arm reciprocal translocations) have also been suggested to play important roles in the karyotypic evolution of this species (Searle & Wójcik 1998). However, the real importance of these processes is still discussed (e.g. Polyakov *et al.* 2001, Zdanova *et al.* 2005).

Twelve pairs of the primitive acrocentric chromosomes (*g* to *r*) are involved in the polymorphism, and thus the chromosome number for *S. araneus* varies between 2n = 20 and 2n = 33. All shrews in the *S. araneus* group have identical sex chromosomes and share three pairs of metacentric autosomes (*af*, *bc* and *tu*). Therefore, the number of chromosome arms is constant (fundamental number, NF = 40). Numerous karyotypic races (for definition, see Hausser *et al.* 1994) have been described all over the Palearctic range of *S. araneus*, each characterized by different sets of acrocentrics and metacentrics. In 2003, a list of 68 chromosome races was published by the "International *Sorex araneus* Cytogenetics Comittee

(ISACC)" (Wójcik *et al*. 2003). However, in spite of substantial chromosomal polymorphism, only very low concomitant morphological variation has been recorded (e.g. Hausser & Jammot 1984, Wójcik *et al.* 2000, Polly 2003).

Hybrid zones

The European species and chromosome races of the *S. araneus* group show a parapatric distribution and form contact or hybrid zones showing very variable levels of gene exchanges between adjacent populations (for reviews, see Searle & Wójcik 1998, Wójcik *et al.* 2002).

A significant topic about chromosomal hybrid zones concerns the fertility of Robertsonian heterozygotes (i.e. do they constitute "tension zones"?; Barton & Hewitt 1985). Therefore, Searle *et al.* (1990) made the distinction between "simple" and "complex" heterozygotes to account for the diversity of hybrids found in the *Sorex araneus* group. Simple heterozygotes produce trivalents during meiosis because at least one pair of homologous chromosomes is present both in the metacentric and acrocentric form (for example: *gi/g, i)*. During meiosis of a complex heterozygote, longer chain or ring elements are formed due to the presence of at least two metacentric chromosomes having only one arm in common (for example the chain: *m – mg – gi – ih – hj – jl – ol – on – nk – kr – r* or the ring: *kp – pq – oq – ko*). In house mice and other mammals, which display chromosomal polymorphism, individuals showing either multiple simple or complex heterozygotes almost always show substantial infertility and sometimes sterility (Searle 1993). However, data from the *Sorex araneus* group suggest that Robertsonian heterozygotes do not suffer from infertility as substantially as other taxa (Searle 1993, Narain & Fredga 1997, 1998). Nevertheless, complex heterozygotes for this group are assumed to be less fertile compared to simple heterozygotes (Hatfield *et al.* 1992, Banaszek *et al.* 2002). Furthermore complex heterozygotes forming chain configurations are less fertile than those forming ring configurations of equal length (Searle 1993, reviewed in Searle & Wójcik 1998).

As previously mentioned, hybrid zones are surprisingly varied in size and shape. This diversity can be illustrated by hybrid zones occurring in Great Britain or in the Swiss and French Alps. The area of polymorphism of the British hybrid zone involving two chromosome races of *S. araneus* (Oxford and Hermitage) is about 100 km. Interestingly, in the centre of this zone, there is a high frequency of acrocentric (acrocentric peak), reducing therefore the potential formation of complex heterozygotes (Hatfield *et al.* 1992, reviewed in Searle & Wójcik 1998). This mechanism has been suggested to favour gene flow between chromosome races ("de-speciation" process; Bengston & Frykman 1990). In contrast, the

alpine hybrid zones involve *S. antinorii* (the former Valais race of *S. araneus*, Brünner *et al.* 2002a) and two races of *S. araneus*, the Cordon race, which is one of the most "acrocentric" known races and the Vaud race. If hybrids were found in these two zones, the general situation strongly differs from the cases described above: *S. antinorii* was found to be clearly genetically differentiated from *S. araneus*, as shown by protein electrophoresis (Brünner and Hausser 1996), mtDNA (Taberlet *et al.* 1994) and microsatellites analyses (Lugon-Moulin *et al.* 1996, 1999a). These hybrid zones are extremely narrow (less than one km) and the clines of genetic markers used are very steep (Brünner and Hausser 1996). In the case of the *S. antinorii – S. araneus* Cordon hybrid zone, a Y-linked microsatellite showed a complete absence of male-mediated gene flow (Balloux *et al.* 2000), which actually suggests a genic differentiation acquired in allopatry. Autosomic microsatellites clearly showed that specific status is the main cause of genetic divergence between populations, the effect of distance or geographical barriers being weak (Lugon-Moulin *et al.* 1999b, but see discussion in Brünner *et al.* 2002b). A direct role of chromosomal differentiation in gene flow restriction could be deduced from the nature of hybrids: in the S. *antinorii – S. araneus* Cordon case, where few or no complex Rb heterozygotes are expected, several hybrids were found (17 % of the individuals studied within the "central" kilometre of the hybrid zone were of hybrid origin) but no single F1 hybrid was detected. Genetically speaking, these backcross hybrids were indiscernible from the "pure" race individuals of the locality in which they were caught. In contrast, in the *S. antinorii – S. araneus* Vaud case, practically every hybrid had an F1 karyotype, usually with a CXI (chain-eleven) multivalent, and they were, as expected, genetically intermediate between the two species (see Brünner *et al.* 2002a). Thanks to the polymorphism of the *lo* chromosome in *S. antinorii*, it was however possible to detect two backcross hybrids indicating that gene flow is not absolutely suppressed. However, this comparison suggests that even under a rather low rate of hybridisation, introgression (female-mediated) is higher in the S. *antinorii – S. araneus* Cordon case.

Additional contact or hybrid zones have been studied among the taxa of the *Sorex araneus* group (Searle & Wójcik 1998, Wójcik *et al.* 2002). As previously mentioned they differ by their genetic and karyotypic characteristics but several of these zones show particular geographic (e.g. contact at a river, railway embankment), karyotypic (e.g. acrocentric or recombinant peaks, formation of hybrid races; Fedyk *et al.* 1991), or genetic (e.g. contact between races within *S. araneus*, species) features. This diversity makes therefore this group extraordinarily informative to study the impacts of chromosome rearrangements on the genetic structure among populations.

Questions addressed

Altogether, the shrews of the *Sorex araneus* group offer a complete array of every possible level of chromosomal and genetic differentiation. In South-Western Europe, four species are recognised: *S. antinorii*, *S. araneus*, *S. coronatus* and *S. granarius*, which differ essentially by the amount and the composition of Robertsonian metacentrics. Additionally, several chromosome races of *S. araneus* are also present in the same region (i.e. Bretolet, Carlit, Cordon, Jura, Vaud and Mooswald).

During the last Pleistocene glaciations, the Alpine barrier played a major role in separating the Italian peninsula from the rest of the continent. For numerous taxa (Taberlet *et al.* 1998), this geographical isolation led to genetic divergence and thereafter influenced postglacial re-colonisation of Europe. The shrews of the *Sorex araneus* group did not escape this pattern and at least five taxa meet in this region (i.e. *S. coronatus*, *S. antinorii* and *S. araneus* Cordon, Bretolet and Vaud). Several contact or hybrid zones between these taxa have moreover been much studied (e.g. Neet & Hausser 1990, Lugon-Moulin *et al.* 1996, Brünner *et al.* 2002b). These taxa are morphologically very similar but show different genetic and karyotypic characteristics. However, almost no trial has been done to link genetic and karyotypic structure at a larger scale than the hybrid zone. Therefore the respective role of genetic and karyotypic differences in structuring the populations of these taxa still needed to be assessed. Consequently, we used Bayesian admixture analyses (Pritchard *et al.* 2000) to check the concordance between genetic and karyotypic structure and tried to identify cryptic substructure among these taxa (Chapter 1). Additionally, we estimated the utility of genetic markers in the identification of several Alpine *S. araneus* group taxa (Chapter 1). Then, we explored the potential and efficiency of the same Bayesian assignment method in combination with the genetic dataset developed in Chapter 1 to study admixture and individual assignment in the difficult context of hybrid zones (Chapter 2).

It is only by studying gene flow at the chromosome level that the question of an impact of chromosomal rearrangements on the genetic structure among the Alpine shrews of the *Sorex araneus* group can be addressed. Microsatellite markers seem especially effective to study genetic structure among these closely related taxa (e.g. Lugon-Moulin *et al.* 1999a, Wyttenbach *et al.* 1999, Brünner *et al.* 2002b, Andersson *et al.* 2004). Consequently, we used flow sorted chromosomes from three different karyotypic taxa (*S. granarius*, *S. araneus*

Cordon and Novosibirsk) to map microsatellite markers at the chromosome arm level (Chapter 3). The comparison of the results among the three taxa allowed identifying markers appropriate for further inter-taxa population genetics studies.

A common prediction to chromosomal speciation models (e.g. Rieseberg 2001) is that when studying a pair of species differing by chromosomal rearrangements, genetic structure should be larger over rearranged chromosomes than over chromosomes common to both species. To test this prediction, we combined two approaches.

First, we compared the genetic differentiation measured over "common" and "rearranged" chromosomes among five karyotypic taxa of the *Sorex araneus* group placed at different evolutionary levels (i.e. chromosome races, partial reproductive isolation, complete reproductive isolation) (Chapter 4). Our prediction was that if karyotypic differences influence the genetic differentiation of this group, rearranged chromosomes would in general be more structured than common chromosomes.

Second, we compared the levels of genetic structure measured over "common" and "rearranged" chromosomes in two hybrid zones between *S. antinorii* and two genetically very similar but karyotypically different chromosome races of *S. araneus* (i.e. Cordon and Vaud) (Chapter 5). Our primary prediction was that if karyotypic differences act as a reproductive barrier, genetic structure would be higher for rearranged chromosomes than for common chromosomes. As the complexity of the hybrids produced was not the same in both zones, our second prediction was that the difference between the two categories of chromosomes would be larger in the most complex hybrid zone.

CHAPTER 1

Genetic and karyotypic structure in the shrews of the Sorex araneus group: Are they independent?

Patrick Basset, Glenn Yannic and Jacques Hausser

(*Molecular Ecology* 2006, in press)

ABSTRACT

The species of the common shrew (*Sorex araneus*) group are morphologically very similar but exhibit high levels of karyotypic variation. Here we used genetic variation at 10 microsatellite markers in a dataset of 212 individuals mostly sampled in the western Alps and composed of five karyotypic taxa (*S. coronatus*, *S. antinorii* and the *S. araneus* chromosome races Cordon, Bretolet and Vaud) to investigate the concordance between genetic and karyotypic structure. Bayesian analysis confirmed the taxonomic status of the three sampled species since individuals consistently grouped according to their taxonomical status. However, introgression can still be detected between *S. antinorii* and the race Cordon of *S. araneus*. This observation is consistent with the expected low karyotypic complexity of hybrids between these two taxa. Geographically based cryptic substructure was discovered within *S. antinorii*, a pattern consistent with the different post-glaciation recolonization routes of this species. Additionally, we detected two genetic groups within *S. araneus* notwithstanding the presence of three chromosome races. This pattern can be explained by the probable hybrid status of the Bretolet race but also suggests a relatively low impact of chromosomal differences on genetic structure compared to historical factors. Finally, we propose that the current dataset (available at http://www.unil.ch/dee/page7010_en.html#1) could be used as a reference by those wanting to identify *Sorex* individuals sampled in the western Alps.

INTRODUCTION

Closely related species and even populations of the same species can exhibit a large amount of karyotype variation. In such situations, the real impact of karyotypic differences on the genetic relationships among taxa is often particularly difficult to evaluate (King 1993, Rieseberg 2001). The closely related shrews of the *Sorex araneus* group constitute an illustrative example. This group of morphologically very similar species is characterised by a XY1Y2 sex chromosome complex in males (Sharman 1956) and is well known for its spectacular chromosomal evolution. Considerable autosomal variation (mainly Robertsonian changes) can be observed not only among the species of this group but also within its type species, *Sorex araneus*. At least 60 chromosomal races distributed all over Europe and Siberia make this species one of the most chromosomally polymorphic among mammals (Wójcik *et al.* 2002). Different models of chromosomal evolution have been developed to account for the large chromosomal variation found in these species (for recent reviews see Searle & Wójcik 1998) but the comparative analyses of karyotypic, biochemical or mitochondrial DNA data often show contradictory results (Taberlet *et al.* 1994, Fumagalli *et al.* 1996, Ratkiewicz *et al.* 2002).

During the last Pleistocene glaciations, the Alpine barrier played a major role in separating the Italian peninsula from the rest of the continent. For numerous taxa (Taberlet *et al.* 1998), this geographical isolation led to genetic divergence and thereafter influenced postglacial re-colonisation of Europe. The shrews of the *Sorex araneus* group did not escape this common pattern and at least five chromosomal races and/or species of this group meet in this region. Actually, *S. coronatus* would have diverged in glacial refugia situated in south-western France or Spain (Hausser 1978) and then would have colonised the pre-Alpine lowlands and large Alpine valleys from the west. The refugia of *S. antinorii* were certainly situated in the Italian peninsula (Brünner *et al.* 2002b). This species (formerly considered as a chromosome race of *S. araneus*, Brünner *et al.* 2002a) crossed several lower Alpine passes in the Swiss and French Alps but mostly remained restricted to Italy (Lugon-Moulin & Hausser 2002). Finally, *S. araneus* certainly presents the most complex evolutionary history. This species may have been restricted to several refugia during the past glacial periods (Taberlet *et al.* 1994) but probably colonised the Alps mostly from south-eastern refugia. Three genetically closely related chromosome races of this species (Cordon, Bretolet and Vaud) are presently distributed in the western Alps.

Species and chromosome races of the *Sorex araneus* group are morphologically very similar and impossible to tell apart in the field. Identification of individuals can be performed from karyotypes and chromosome counts, allozymes (Hausser & Zuber 1983, Neet & Hausser 1991), morphometric measures (Hausser *et al.* 1991) or analysis of diagnostic markers (Basset & Hausser 2003) but all these methods require destructive sampling or have drawbacks in field studies (reviewed in Basset & Hausser 2003). Recently, Pfunder *et al.* (2004) proposed an attractive microarray-based diagnostic test for shrew species; however it did not allow discrimination between all *Sorex* species present in the Alps (e.g. between *S. antinorii* and *S. araneus*). Additionally, this method could be somewhat expensive and time consuming when used at small scale.

Recently, numerous methods based on genetic assignation of individuals to a given group have been developed and seem especially effective (Paetkau *et al.* 1995, Rannala & Mountain 1997, Cornuet *et al.* 1999, Pritchard *et al.* 2000, Vázquez-Dominguez *et al.* 2001; Wilson & Rannala 2003). However, before using these techniques to allocate unknown samples to taxa, it is necessary to check the agreement between the genetic and the taxonomic structure. This last point should not be underestimated in the *Sorex araneus* group as the genetic relationships among taxa are not straightforward. European species and chromosome races of this group show parapatric distributions and typically form various contact or hybrid zones. Several of these zones have been studied all over Europe (Searle & Wójcik 1998, Fredga & Narain 2000, Ratkiewicz *et al.* 2003, Andersson *et al.* 2004) including the Alps (Neet & Hausser 1990, Neet 1992, Lugon-Moulin *et al.* 1996, Brünner *et al.* 2002b). These zones generally showed a large variety of levels of gene exchange between adjacent populations, from total isolation to almost free gene flow. For example, species identity was clearly the greatest factor structuring the hybrid zone between *S. antinorii* and *S. araneus* (Brünner *et al.* 2002b) whereas Andersson *et al.* (2004) found a similar level of genetic structure within and between two chromosome races belonging to two different karyotypic groups (group of chromosome races characterized by some shared metacentrics (Searle & Wójcik 1998)). Actually, the effect of chromosomal differences on the gene flow between two populations is much debated (Rieseberg 2001, Navarro & Barton 2003a, Panithanarak *et al.* 2004) and still needs to be assessed in the case of the *Sorex araneus* group. Almost no real trial has been done to link genetic and karyotypic structure at a larger scale than the hybrid zone. Generally, to estimate population structure and/or assign individuals to a population it is necessary to *a priori* define discrete populations following subjective criteria. However, it seems important to check whether these *a priori* assumptions match genetic data in natural populations,

particularly when population limits are not totally clear (for example in hybridizing taxa). Some of the recent assignment methods (Pritchard *et al.* 2000, Vázquez-Dominguez *et al.* 2001, Wilson & Rannala 2003) allow description of population structure without requiring predefined groups, providing new opportunities for checking the relationships between expected and real population structure. The methods of Pritchard *et al.* (2000) and Wilson & Rannala (2003) are particularly informative when studying a possibly hybridizing group of species as they consider that an individual could originate from more than one population.

In the present study we genotyped 212 individuals at 10 microsatellite loci in order to (1) check the concordance between genetic structure and karyotypic structure in the *Sorex araneus* group; (2) identify potential cryptic substructure; (3) estimate the utility of genetic markers in the identification of different species or chromosomal races of the *Sorex araneus* group in the Alps; and (4) develop a genetic reference to allocate individuals of unknown origin to species and/or population with Bayesian assignment techniques.

MATERIAL AND METHODS

Sample Collection

A total of 33 Jersey shrews (*Sorex coronatus*), 83 Valais shrews (*S. antinorii)* and 96 Common shrews (*S. araneus*), subdivided into three chromosome races: 30 *S. araneus* Cordon, 25 *S. a.* Bretolet and 41 *S. a.* Vaud were analyzed during this study. All these individuals were sampled by various collectors (DEE collection, Lausanne University) from 1985 to 2003. Distribution of sampling localities is shown in Fig. 1 and covers a large part of the European distribution of the studied taxa. Species and chromosome race identification of most individuals followed karyotype analysis but in a few unambiguous cases, it was deduced from sampling localities and morphological analysis.

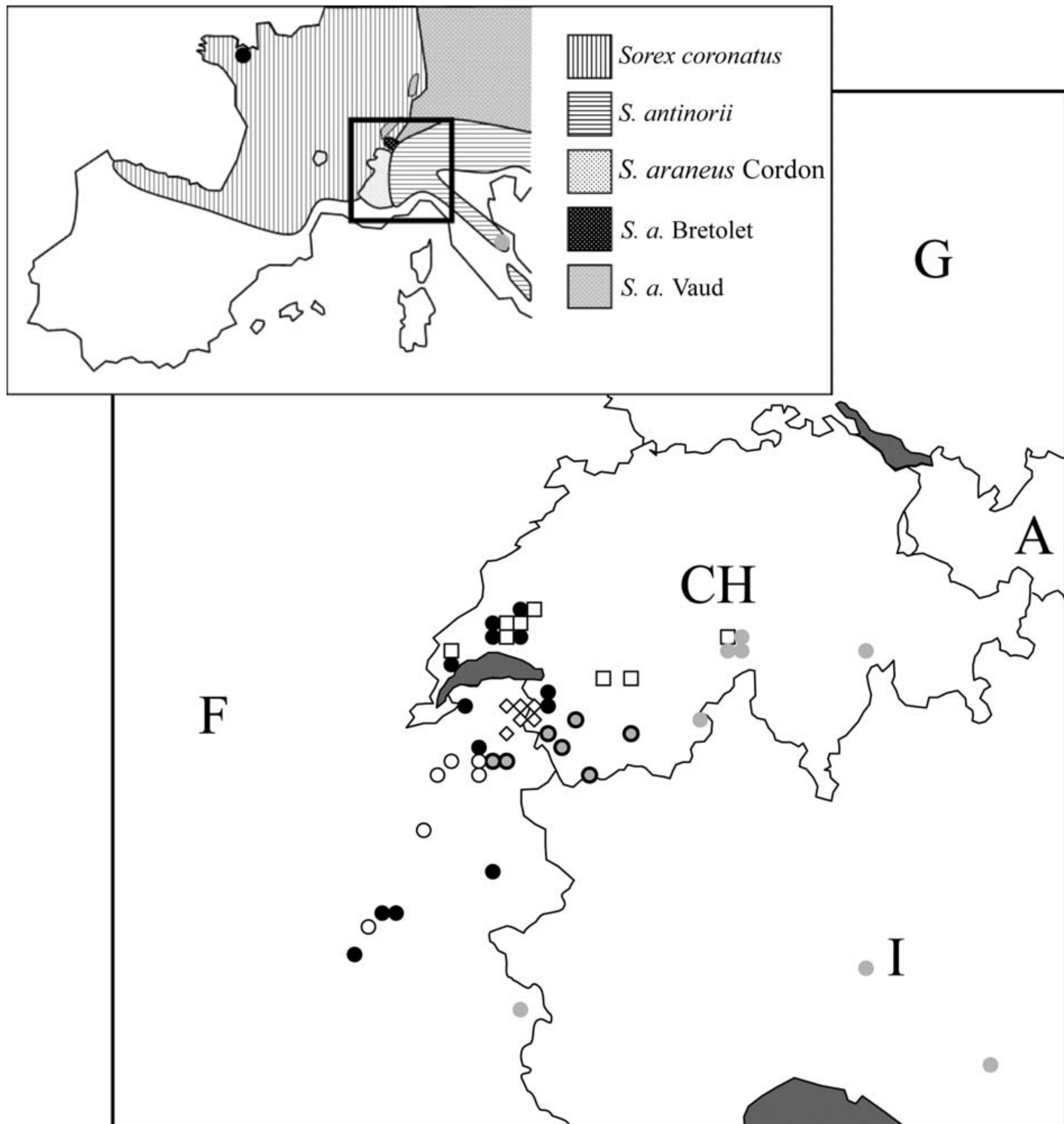


Figure 1 Insert map: approximate distribution of the studied taxa in south-western Europe, including remote sampling localities. Below, sampling localities of the five taxa studied; black circles: *Sorex coronatus*, grey circles: *S. antinorii* (framed grey circles: St-Bernard group as defined by our analysis), open circles: *S. araneus* Cordon, open squares: *S. a.* Vaud and open diamonds: *S. a.* Bretolet.

DNA extraction and microsatellite analysis

Tissue samples (liver, heart or spleen) were stored at -70°C and total genomic DNA was extracted using the DNeasy Tissue Kit (Qiagen).

Ten microsatellite loci were used in this study and included loci L9, L13, L67, L99 (PCR conditions given in Balloux *et al.* 1998 and Lugon-Moulin *et al.* 2000) and B3, B5, B10, B15,

C5, C19 (Table 1). These last six loci were extracted from two enriched *Sorex araneus* microsatellite libraries developed by a commercial company (Genetic Identification Services, Inc. Chatsworth, CA). PCR conditions for these loci were as follows: 0.2 mM dNTPs, 0.325 µM of each primer, 10x PCR buffer (Qiagen) and 0.4U Taq polymerase (Qiagen). $MgCl_2$ concentration as well as annealing temperature varied (Table 1). For all primers, PCR amplifications were performed in a 20 µl total volume and cycling was carried out in a PE 9700 (Applied Biosytems) using the following profile: 95°C for 5 min, 35 cycles of 30 s at 94°C, 30 s at the annealing temperature (Table 1), 30 s at 72°C; and a final extension at 72°C for 4 min. One primer of each pair was labeled with a fluorescent dye on the 5'end, which allowed analyses on an ABI 377XL sequencer (Applied Biosystems). Data collection, sizing of the bands and analyses were done using GENESCAN software (Applied Biosystems).

Table 1 Primer sequence and PCR conditions of six microsatellite loci isolated from two *Sorex araneus* microsatellite libraries developed by Genetic Identification Services.

Locus	Primer Sequence (5'-3')	Annealing T [°C]	MgCl2 [mM]	Size [bp]	Nb repeats	Accession numbers
B3	F: CTT GCC ACA TTC CCA CAT C R: AGC CCC ACA GCT TTC TCC	57	1.0	208	30	DQ074646
B5	F: ATG TCT TGC TGG CTG AAG G R: CTG CTG TTC ACA AAC TCC AAG	55	1.5	196	19	DQ074647
B10	F: CTC CAA ACC CTA ACA CTC TGT C R: TTC ACG TGT TCT TTG CTT CC	55	1.5	434	18	DQ074648
B15	F: GTA GAG TTG CTG GCT CAA AGG R: ATG GGA AGA CAT TGG ATT GG	55	1.5	299	18	DQ074650
C5	F: TAG ATG ACT CTG TGT TCA GGC R: GTT GGG AAG GTA AGA TCA GG	55	1.5	236	16	DQ074649
C19	F: TGC CAT AAA CAC CAC TTA CC R: GTG ATC AAT ACC CTG TGG AG	60	1.5	211	12	DQ074651

Hardy-Weinberg and linkage equilibrium

The software package Genetix 4.02 (Belkir *et al.* 2001; http://www.univ-montp2.fr/~genetix/genetix.htm) was used to calculate the allele frequencies, allele number, observed (H_O) and expected (H_E) heterozygosities for each species or chromosome races. Genotypic linkage disequilibria were tested using GENEPOP version 3.4 (updated from Raymond & Rousset 1995; http://wbiomed.curtin.edu.au/genepop).

Deviations from Hardy-Weinberg Equilibrium (HWE) and the significance of Weir & Cockerham (1984) F-statistics were evaluated using FSTAT 2.9.4 (Goudet 2001; http://www.2.unil.ch/popgen/sofwares/fstat.html).

Bayesian analyses

To check the concordance between karyotypic and genetic structure, all the genotypes were screened using a Bayesian admixture procedure implemented in STRUCTURE 2.1 (Pritchard *et al.* 2000; http://pritch.bsd.uchicago.edu). This model was designed to identify the K (unknown) genetic clusters (or populations) of origin of individuals, and simultaneously to probabilistically assign individuals to one cluster or more than one cluster if they are genetically admixed as a result of hybridization. STRUCTURE was run with the "admixture model", and five repetitions of 100,000 iterations following a burn-in period of 20,000 iterations.

We first assessed population structure using the total dataset (n = 212), assuming that sampled individuals belong to an unknown number of K genetically distinct clusters. Posterior probability values for K ("Log probability of data"; L(K)) were estimated assigning a prior from one to ten. Using only this parameter as described by Pritchard *et al.* (2000) it was not obvious which number of clusters (K) best fits our dataset (Fig. 2A). Therefore, we followed the recommendation of Evanno *et al.* (2005) and calculated the ΔK statistic, which is based on the rate of change in the "Log probability of data" between successive K values. We chose the value of K = 3, which showed the highest ΔK and then evaluated the individual membership coefficient (q_{ind}) to the three inferred clusters. Individuals with a proportion of membership to each cluster $q_{ind} < 0.90$ (admixed individual) were assigned to more than one cluster whereas individuals with $q_{ind} \geq 0.90$ were assigned to only one cluster. The threshold value of 0.90 was arbitrarily defined to be sure that at least 90% of an individual's genome is assigned to one cluster (Manel *et al.* 2002, Cegelski *et al.* 2003). Then, we assessed the average membership coefficient (q_{group}) of each taxon (species or chromosomal race) to each cluster. Similarly, each sampled taxa (species or chromosome race) was assigned to one cluster if its q_{group} was ≥ 0.90, or jointly to more than one cluster, if its q_{group} to each cluster was < 0.90 (admixed taxon).

In cases of hierarchical population structure, STRUCTURE is known to preferentially detect the uppermost structure level (Evanno *et al.* 2005). First analysis of our dataset and the presence of three chromosome races within *Sorex araneus* suggested that such a situation was present in our case. To explore whether substructure could be detected within each species, the dataset of each species was then analyzed independently.

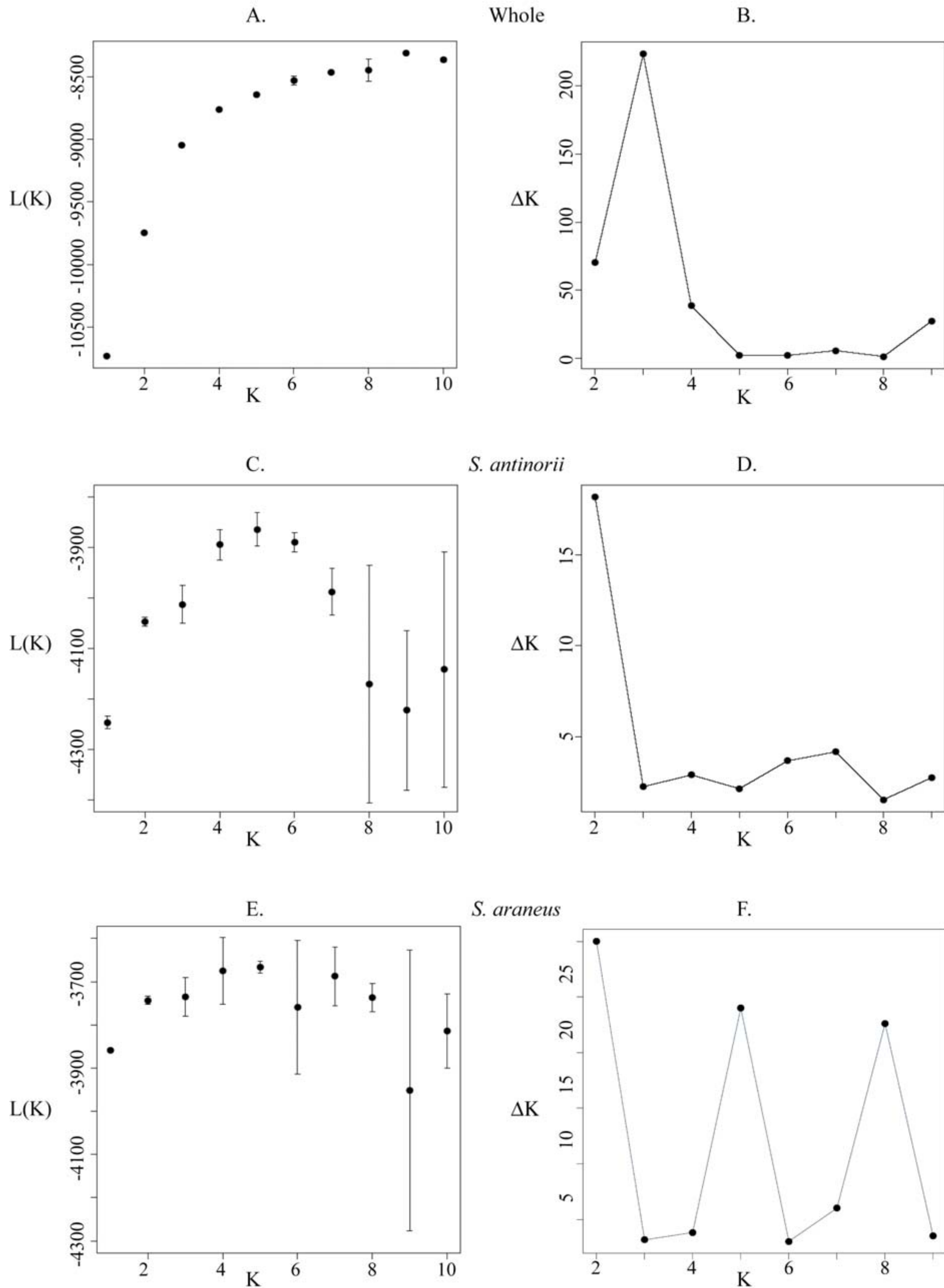


Figure 2 Detection of the number of groups in the whole dataset (A and B), *Sorex antinorii* dataset (C and D) and *S. araneus* dataset (E and F). (A), (C) and (E): Mean L(K) (± s.d.) over five runs as a function of K. (B), (D) and (F): ΔK following Evanno *et al.* (2005) as a function of K.

Additionally, we investigated the power of our dataset to act as a reference to identify individuals of unknown origin. Therefore, we used STRUCTURE with K = 3, using the available prior population information (species classification), and options USEPOPINFO = 1. In this way, each shrew of the dataset was forced to have its genotype assigned to one of the three species, or, if admixed to more than one species. The same analysis was then performed on the *Sorex araneus* dataset only, to discriminate among the three chromosome races. Finally, we tested the real efficiency of our dataset in the identification of unknown individuals using a "leave one out" procedure. We chose a random subset of individuals (representing about 10% of each taxa) as test individuals. We ran STRUCTURE using the available prior species information (K = 3) for all individuals (USEPOPINFO = 1) except for the test individuals (4 *Sorex coronatus*, 9 *S. antinorii*, 3 *S. araneus* Cordon, 3 *S. a.* Bretolet and 4 *S. a.* Vaud) treated as having unknown origin (USEPOPINFO = 0). The same analyses were then repeated ten times, each time randomly selecting the same number of test individuals. Assignation results of the test individuals were then used to estimate the percentage of correct assignations (individuals correctly assigned with $q_{ind} \geq 0.90$), unassigned individuals (showing admixture $0.10 < q_{ind} < 0.90$) and assignment mistakes (shrews assigned to an incorrect cluster with high probability $q_{ind} \geq 0.90$). The same procedure was then repeated on the *S. araneus* dataset only.

RESULTS

Population genetic diversity

We determined the individual genotypes at 10 microsatellite loci in 212 shrews. All microsatellite markers were polymorphic and the number of alleles per locus ranged from five to 43 (average 28.5 ± 13.4). In the following analyses, the dataset was first subdivided into the three species included in this study (*Sorex coronatus, S. antinorii* and *S. araneus*). Then *S. araneus* was subdivided into the three chromosome races, *S. a.* Cordon, *S. a.* Bretolet and *S. a.* Vaud. Number of alleles per taxon ranged from 70 (*S. coronatus*) to 221 (*S. antinorii*). The number of private alleles ranged from 3 (*S. a.* Bretolet) to 75 (*S. antinorii*) with an average of 43.7 ± 28.0 when pooling the three chromosome races of *S. araneus* or 24.2 ± 29.2 when subdividing *S. araneus* in the three races (Table 2).

Tests of fit to HWE, linkage equilibrium and divergence among shrew taxa

All five species and chromosome races showed H_O values lower than expected, with average F_{IS} values that were positive (from 0.232 to 0.454), highly significantly different from 0 ($P \leq 0.002$) indicating deviation from HWE (Table 2). This suggests a strong Wahlund effect, probably resulting from the pooling of individuals of geographically and genetically different origins. At least a part of this homozygote excess could be explained by the presence of null alleles (Pemberton *et al.* 1995). To test for this effect, the correlation between the number of non-amplifying samples for each locus and F_{IS} was examined. In each taxa, individuals were grouped into geographic sampling localities. Localities with less than three individuals were left out of the analysis. A significant correlation was found only for locus L99 ($R^2 = 0.942$, $P \leq 0.001$). This result can easily be explained by the low allele number of this locus, which is confirmed by the fact that this locus did not show homozygote excess in any of the populations tested. Thus, it seems more likely that the general homozygote excess is due to the pooling of genetically differentiated individuals. After Bonferroni correction for multiple comparisons, the loci B10 and L13 were not in linkage equilibrium in *Sorex antinorii*. However these loci map to different chromosomes (P. Basset, unpublished data).

Table 2 Genetic diversity in the three species of shrews and in the three chromosome races of *S. araneus* over the 10 microsatellite loci.

	No. of alleles	**No. of private alleles**	**H_E**	**H_O**	**Hardy-Weinberg Equilibrium (F_{IS})**
1 *S. coronatus* (n = 33)	70	21	0.558 (0.273)	0.315 (0.239)	0.454***
2 *S. antinorii* (n = 83)	221	75	0.779 (0.256)	0.561 (0.284)	0.286***
3 *S. araneus* all (n = 96)	168	35	0.738 (0.273)	0.527 (0.295)	0.291***
3.1 Cordon (n = 30)	126	14	0.727 (0.307)	0.573 (0.300)	0.232***
3.2 Bretolet (n = 25)	104	3	0.717 (0.241)	0.542 (0.289)	0.263***
3.3 Vaud (n = 41)	111	8	0.688 (0.283)	0.486 (0.310)	0.305***
Overall	285		0.829 (0.192)	0.501 (0.271)	0.296***

H_E =expected heterozygosity without bias (Nei 1978), H_O = observed heterozygosity, Standard Deviation in brackets. F_{IS} = Deviation from Hardy weinberg equilibrium following Weir & Cockerham 1984, *** = P-value ≤ 0.002.

Genetic diversity was significantly partitioned among the three species ($F_{ST} = 0.199$, $P \leq 0.001$) and chromosome races ($F_{ST} = 0.047$, $P \leq 0.001$). All pairwise F_{ST} values (Table 3) were significantly different from 0 ($P \leq 0.01$).

Table 3 Estimate of pairwise genetic differentiation (F_{ST}) among shrew taxa. *** = P-value ≤ 0.01

Taxa		*S. coronatus*	*S. antinorii*	*S. araneus*		
				Cordon	Bretolet	Vaud
S. coronatus		-	0.172***	0.300***	0.312***	0.326***
S. antinorii			-	0.127***	0.153***	0.174***
S. araneus	Cordon			-	0.046***	0.064***
	Bretolet				-	0.028***
	Vaud					-

Genetic admixture and assignment analysis

We used Bayesian analyses (STRUCTURE, POPINFO = 0) to detect admixture and possible cryptic substructure in our dataset (n = 212). Using only the "Log probability of data" (as described in Pritchard *et al.* 2000) it was not clear which number of clusters best fits our data (Fig. 2A). However, the statistic ΔK described by Evanno *et al.* (2005) clearly indicates that the sample included at least three distinct groups (the highest ΔK was obtained with K = 3, Fig. 2B).

The average proportions of membership (q_{group}) of each sampled taxa in the three clusters (Table 4) showed that all the *Sorex coronatus* individuals grouped in cluster I ($q_{coronatus}$ I = 0.99) while *S. antinorii* was significantly assigned to cluster II ($q_{antinorii}$ II = 0.97). However, five individuals (out of 83) of this species showed signs of admixture (q_{ind} II < 0.90) with *S. araneus*. Cluster III represented the "*S. araneus* cluster" since this species grouped in this cluster with $q_{araneus}$ III = 0.95. Cordon was the only chromosome race to show admixture with *S. antinorii* (q_{Cordon} III = 0.89 and 8/30 individuals had q_{ind} III between 0.33 and 0.83). In contrast, every Vaud and all but three Bretolet individuals grouped in cluster III (q_{Vaud} III = 0.99 and $q_{Bretolet}$ III = 0.96 respectively). To summarise, of the 212 individuals tested, no individual grouped with a cluster different than its putative origin and only 16 (7.5%) showed signs of admixture. Half of these individuals were karyotypically identified as *S. araneus* Cordon and showed signs of admixture with *S. antinorii*.

Table 4 Bayesian clustering analyses for the *Sorex* reference dataset (212 individuals; 10 loci) performed using STRUCTURE (Pritchard *et al.* 2000).

	Cluster		
Taxa	**I**	**II**	**III**
***S. coronatus* (n = 33)**	**0.991**	**0.004**	**0.005**
***S. antinorii* (n = 83)**	**0.007**	**0.969**	**0.024**
Ind n° 2883	0.003	0.899	0.098
Ind n° 3070	0.003	0.845	0.152
Ind n° 3071	0.002	0.885	0.113
Ind n° 3134	0.004	0.560	0.436
Ind n° 5319	0.003	0.893	0.104
***S. araneus* Cordon (n = 30)**	**0.006**	**0.105**	**0.890**
Ind n° 3336	0.002	0.249	0.749
Ind n° 3337	0.002	0.322	0.676
Ind n° 3341	0.002	0.423	0.575
Ind n° 3342	0.003	0.672	0.325
Ind n° 3348	0.009	0.444	0.547
Ind n° 3373	0.003	0.167	0.831
Ind n° 3379	0.009	0.313	0.677
Ind n° LC2	0.044	0.185	0.771
***S. araneus* Bretolet (n = 25)**	**0.019**	**0.024**	**0.956**
Ind n° 2345	0.209	0.003	0.787
Ind n° 3274	0.185	0.005	0.810
Ind n° 3275	0.006	0.454	0.540
***S. araneus* Vaud (n = 41)**	**0.005**	**0.007**	**0.988**

N.B. In bold, average proportion of membership (q_{group}) of each predefined population in each of K = 3 inferred clusters. Admixed individuals ($q_{ind} < 0.90$) are indicated under each population lines.

Preliminary analysis and the presence of three different chromosome races within *S. araneus* prompted us to explore whether substructure could be detected within each of the three species. No substructure was detected within *S. coronatus* as the most likely K for this species is one. For *S. antinorii* and *S. araneus* two distinct groups were detected within each of these species. Again it was necessary to estimate the ΔK statistic to decide which K best fits the data, Fig. 2C to F. For *S. araneus*, a careful comparison of this statistic with the L(K) was necessary as more than one ΔK peak were detected.

The *Sorex antinorii* dataset was split into two well geographically differentiated clusters (Fig. 3A): individuals sampled in Italy, eastern Switzerland and the southern French Alps (Hautes-Alpes) grouped in one cluster while individuals sampled in western Switzerland and the northern part of the French Alps grouped in a second cluster (framed grey circles in Fig. 1).

Two clusters were also detected within the *S. araneus* dataset ($q_{araneus}$ I = 0.536, Fig. 3B). Each of the three chromosome races showed signs of admixture between the two clusters with q_{Cordon} I = 0.865, $q_{Bretolet}$ I = 0.743 and q_{Vaud} I = 0.168.

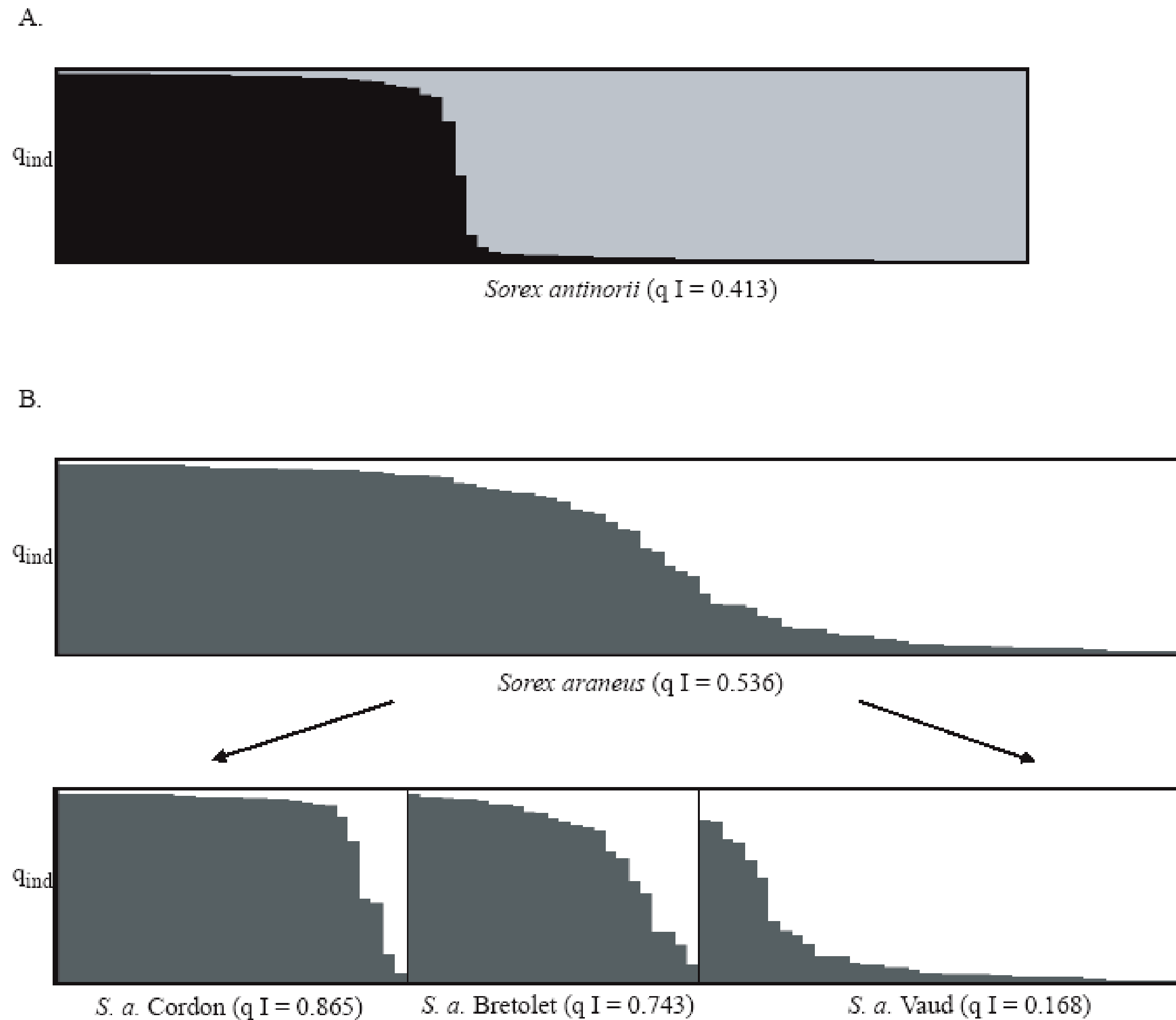


Figure 3 Distruct plots (Rosenberg 2004) for *Sorex antinorii* (A) and *S. araneus* (B). Each individual is represented by a line partitioned into two colours (K = 2) representing its genotypic assignment to one cluster or the other (q_{ind}). In both species, individuals are sorted according to decreasing q_{ind} values. The *S. araneus* dataset is then sorted according to the three chromosome races Cordon, Bretolet and Vaud. Mean q I values are given in parentheses.

Shrew ancestry was then estimated using prior information (POPINFO = 1) about species (K = 3). Each species grouped in their respective cluster with high probability ($q_{species} \geq 0.99$). Of the 212 individuals tested, none was assigned to a species different than its putative origin but five (2.3%) showed signs of admixture (ind. n° 3071, 3134, 3341, 3379 and 2345). Finally, ancestry was estimated on the *S. araneus* dataset using prior information (POPINFO = 1)

about chromosomal race (K = 3). Individuals clustered into three groups corresponding to the three chromosome races studied (Cordon, Bretolet and Vaud) with high probability ($q_{race} \geq 0.96$). Again, of the 96 individuals tested, only five (5.2%) showed signs of admixture (ind. n° 3349, LC6, 2596, 3021 and 3261).

We tested the efficiency of our dataset to act as reference to identify individuals with unknown species identity and to identify *Sorex araneus* individuals with unknown chromosome race identity. All individuals showing admixture signs in the previous analyses were left out from this reference dataset. Then assignment tests were performed on our global sample (including admixed individuals) using this reduced dataset (n = 207 for the species dataset and n = 91 for the *S. araneus* dataset). Species identification was correct in 100% of *S. coronatus*, 92.2% of *S. antinorii* and 86.4% of *S. araneus* using the criterion $q_{ind} \geq 0.90$ (Table 5A). For these last two species, 8.8% and 13.4% respectively of the individuals were identified as admixed ($q_{ind} < 0.90$ level). None of the individuals with q_{ind} values ≥ 0.90 were assigned to a cluster different than its correct origin (assignment mistakes).

Chromosome race identification within *Sorex araneus* was much lower (Table 5B). Only 48.3%, 22.5% and 41.5% of individuals belonging to the Cordon, Bretolet and Vaud race respectively were correctly assigned using the criterion $q_{ind} \geq 0.90$. A large percentage of the individuals (respectively 48.3%, 75.0% and 56.1%) could not be assigned to any cluster and showed clear admixture. Finally, in 3.4%, 2.5% and 2.4% of the cases respectively, individuals were assigned to a wrong cluster with a high q_{ind} value ($q_{ind} \geq 0.90$).

Table 5 Percentage of correct assignation ($q_{ind} \geq 0.90$), assignation with admixture ($q_{ind} < 0.90$) and wrong assignation ($q_{ind} \geq 0.90$ for wrong taxa) of species (A) and *Sorex araneus* chromosome race (B).

A

	S. coronatus	*S. antinorii*	*S. araneus*
Correct ($q_{ind} \geq 0.90$)	100.0%	92.3%	86.4%
Admix ($q_{ind} < 0.90$)	0.0%	7.7%	13.6%
Wrong ($q_{ind\ wrong} \geq 0.90$)	0.0%	0.0%	0.0%

B

	S. araneus		
	Cordon	Bretolet	Vaud
Correct ($q_{ind} \geq 0.90$)	48.3%	22.5%	41.5%
Admix ($q_{ind} < 0.90$)	48.3%	75.0%	56.1%
Wrong ($q_{ind\ wrong} \geq 0.90$)	3.4%	2.5%	2.4%

DISCUSSION

Species introgression

Species and chromosome races of the *Sorex araneus* group form numerous contact or hybrid zones (Searle & Wójcik 1998). This observation is particularly true in the Alpine region, which is known to be a suture zone for numerous taxa (Taberlet *et al.* 1998, Hewitt 2001). Recent and/or past introgression among these taxa explain why relationships among taxa of the *Sorex araneus* group in the Alps are challenging issues that are much debated (Taberlet *et al.* 1994, Brünner *et al.* 2002b, Lugon-Moulin & Hausser 2002). However, no trial has been done to study genetic structure without using prior population information and to detect hidden or unexpected structure. The first goal of our work was to check if the genetic structure matched the structure defined by the different karyotypes involved in this group (defined as the taxonomical structure) without using prior information about population. Results of this study confirmed the taxonomic status of the three species *Sorex coronatus*, *S. antinorii* and *S. araneus* since three distinct clusters were detected by STRUCTURE. Each cluster consistently grouped individuals according to their taxonomic status but it is interesting to note that weak introgression can still be detected between *S. antinorii* and *S. araneus* (Table 4). It should be noticed that a specific status was given to *S. antinorii* (Brünner *et al.* 2002a) notwithstanding a very limited gene flow with adjacent populations of *S. araneus* (see discussion Brünner *et al.* 2002a).

Five individuals (out of 83) of *S. antinorii* showed signs of admixture with *S. araneus*. About the same proportion of individuals of *S. araneus* (11 out of 96) showed admixture with *S. antinorii*. However, eight of these individuals belonged to the chromosome race Cordon and admixed individuals of this race showed a much higher level of introgression ($0.33 \leq q_{ind}$ III ≤ 0.83) than the other taxa. Such introgression could be explained by the presence of at least one hybrid zone with *S. antinorii* (Brünner & Hausser 1996). This hypothesis is confirmed by the sampling localities of five admixed individuals situated close to the contact zone. However, we did not detect any introgression of *S. antinorii* into the *S. araneus* Vaud chromosome race notwithstanding the occurrence of hybridization between these two taxa (Brünner *et al.* 2002b). Differences in the karyotype complexity of these two chromosome races could explain the disparity in introgression level. The Cordon race is known for its acrocentric karyotype; hybrids with *S. antinorii* should form mostly trivalents and encounter only mild problems at meiosis (Brünner *et al.* 2002b). In contrast, the Vaud race is defined by a much more metacentric karyotype showing up to eight monobrachial homologies with *S.*

antinorii. Hybrids should then meet more chromosomal incompatibilities (Brünner *et al.* 2002b) and gene flow between these two taxa should be greatly reduced.
Only a few (five out of 83) individuals of *S. antinorii* showed signs of admixture with *S. araneus*, therefore introgression seems to be unidirectional (*S. antinorii* into *S. araneus*). This pattern could be explained by differences in the sampling distribution of the two taxa, i.e. most Cordon individuals were sampled close to some *S. antinorii* localities, whereas sample distribution of *S. antinorii* is much larger and only a few of these individuals were sampled close to Cordon localities. However, more comprehensive studies of contact zones between these two taxa are necessary to clarify this pattern.

Structure within species

Evanno *et al.* (2005) showed that when confronted with complex migration schemes, STRUCTURE detects the uppermost hierarchical level of population structure. Our study illustrates this phenomenon as we detected substructure into two out of three species, a result not evidenced by the first analyses of the dataset.
Bayesian clustering of the *Sorex antinorii* dataset revealed two different, geographically based clusters. One group contained all individuals sampled in the northern part of the French Alps and western Switzerland (St-Bernard region, Fig. 1). The second group contained the remaining individuals sampled in Italy, eastern Switzerland and the southern French Alps. This geographical subdivision confirmed the possible presence of at least two different recolonization routes (i.e. Simplon and St-Bernard pass) of *S. antinorii* from the Italian peninsula, as postulated by Lugon-Moulin & Hausser (2002) and Fivaz *et al.* (2003).
In addition, the present work clarified some of the hypotheses developed by the same authors to account for the near absence of gene flow between the two regions. These authors postulated that two distinct male lineages could have already been differentiated in Italy. In our study, all individuals sampled in Italy (Apennine region) clustered with the eastern Switzerland and southern French Alps group. Further Italian samples are needed to properly address the recolonization of Switzerland by *S. antinorii*, and in particular to verify if the secluded Aosta valley, leading to the St-Bernard pass, actually hosts populations differentiated from the other Italian ones.
Despite the presence of three chromosome races, analyses of the *Sorex araneus* dataset revealed only two different clusters. However, partition of individuals in these clusters generally followed karyotypic identification, i.e. most individuals of the race Cordon grouped in one cluster (q_{Cordon} I = 0.865) while most individuals of the Vaud race grouped in the

second cluster (mean q_{Vaud} I = 0.168). The chromosome race Bretolet showed an admixed status between these two clusters although it shows closer links to the Cordon cluster ($q_{Bretolet}$ I = 0.743), suggesting the hybrid status of this chromosome race. Actually, this karyotypic group was postulated to be formed by introgression of "Vaud" metacentric chromosomes into Cordon populations (Hausser *et al.* 1991).

We did not find any population structure within *Sorex coronatus* but it should be noted that the sampling effort was weaker for this species (n = 33). This species has a larger allelic diversity compared to *S. araneus* suggesting that substructure could still be hidden.

Effect of karyotype on genetic structure

Differences in karyotypes are thought to reduce gene flow among different karyotypic groups (Rieseberg 2001, Navarro & Barton 2003a, Panithanarak *et al.* 2004). Data concerning the *Sorex araneus* group are scarce but in our case, this should be translated into reduced gene flow among the three chromosome races. We were therefore expecting three distinct clusters within *S. araneus* (corresponding to the three chromosome races sampled) and only one cluster within *S. antinorii* (this species has a homogenous karyotype). Surprisingly, our study did not confirm this prediction since two clusters were detected in both species. In addition, levels of introgression between clusters seemed larger within *S. araneus* (Fig. 2B) compared to *S. antinorii* (Fig 2A). This is also suggested by the rather low F_{ST} values calculated among chromosome races within *S. araneus* ($F_{ST\ araneus}$ from 0.028 to 0.064) compared to the larger F_{ST} between the two geographical clusters within *S. antinorii* ($F_{ST\ antinorii}$ = 0.082). These results suggest that in the taxa studied, karyotypic differences played only a minor role in structuring the populations relative to historical and/or geographical factors. Andersson *et al.* (2004) obtained similar results in a hybrid zone between the North (Abisko race) and West (Sidensjö race) European karyotypic group of *S. araneus*. However, the real impact of karyotypic changes on the genetic structure among taxa of the *S. araneus* group can only be addressed through detailed individual analyses of the chromosomes involved in these changes.

Species identification

Sibling species can be defined as those sister species that are very difficult or virtually impossible to distinguish by morphological characters (Futuyma 1986). This applies to species and chromosome races of the *Sorex araneus* group, and identification methods have already been the topic of several papers (Hausser *et al.* 1991, Wójcik *et al.* 1996, Basset &

Hausser 2003 and Pfunder *et al.* 2004). These authors used morphometric measures, diagnostic microsatellite markers or microarray-based analysis to distinguish the different species of this group. However, drawbacks were found with each method (reviewed in Basset & Hausser 2003) and only morphometric or microsatellite analyses gave information about the possible hybrid status of an individual; a crucial point when studying potentially hybridizing groups of species. Our method seemed to perform well at the species level since we identified with a $q_{ind} \geq 0.90$ criterion, 100% of the S*orex coronatus* individuals, 92.2% of *S. antinorii* and 86.4% of *S. araneus* (Table 5A). These values are close to those obtained by Hausser & Jammot (1984) and Hausser *et al.* (1991) using morphological measures combined with discriminant analyses. These authors found correct classification in more than 95% of the cases but this technique is particularly time-consuming, and requires that the same person analyses the reference and the individuals to assign.

Identification efficiency strongly decreased when we tried to discriminate among the different chromosomal races within *S. araneus*. Using the $q_{ind} \geq 0.90$ criterion, identification within this species did not exceed 48.3% (Cordon race). Such low assignment power could be explained by the low number of individuals sampled for each of these chromosome races (Evanno *et al.* 2005). But F_{ST} values among these taxa were also the lowest observed in this study (Table 3). Assignment power of STRUCTURE is known to be well correlated with measures such as F_{ST} (Berry *et al.* 2004). *Sorex coronatus* also illustrates this relationship, as this species showed the highest F_{ST} values with all other taxa ($F_{ST} \geq 0.17$) (Table 3) and was always correctly identified or separated from the other taxa. Berry *et al.* (2004) have shown that increasing the number of genetic markers could increase the assignment power particularly when F_{ST} values are rather low ($F_{ST} \leq 0.08$). Genetic discrimination between other chromosome races of *S. araneus* has already proven to be difficult (Andersson *et al.* 2004).

In conclusion, Bayesian assignment analyses revealed large scale introgression of *Sorex antinorii* into one chromosome race (Cordon) of *Sorex araneus*. In addition, geographically based cryptic substructure was discovered within *S. antinorii*, a pattern consistent with the different putative post-glacial recolonization routes of this species. This study furthermore illustrates the relatively low impact of chromosomal differences on the genetic structure compared to historical factors. Finally, the dataset described in this article is available to researchers at http://www.unil.ch/dee/page7010_en.html#1, so that it can be used as a reference dataset to help identify unknown *Sorex* individuals sampled in the western Alps or clarify the relationships among individuals sampled in hybrid zones between species of the

Sorex araneus group. This last point is indeed particularly important as it is generally very difficult to obtain reference datasets when working on hybrid zones.

ACKNOWLEDGMENTS

We are particularly grateful to Laura Brändli, Guillaume Evanno, Julie Jaquiéry, Lori Handley and Valérie Vogel for helpful comments on the manuscript. This work was supported by the Swiss NSF (grants n°: 3100-05943), a Société Académique Vaudoise (Switzerland) grant to PB and a Fondation Agassiz (Switzerland) grant to PB.

CHAPTER 2

Species identification by genetic assignment methods: what about hybrid zones?

Patrick Basset, Glenn Yannic, Harald Brünner and Jacques Hausser

(In preparation)

ABSTRACT

Traditionally, individuals are analysed according to their sampling locality, morphology, behaviour or karyotype. But the increasing availability of genetic information, more and more favours its use for individual sorting purposes and numerous assignment methods based on the genetic composition of individuals have been developed. The shrews of the *Sorex araneus* group offer good opportunities to test the application of classical and genetic assignment methods. Here we explore the potential and efficiency of a Bayesian assignment method combined or not with a dataset reference to study admixture and individual assignment in the difficult context of two hybrid zones between karyotypic species of the *Sorex araneus* group. As a whole, we assigned more than 80% of the individuals to their respective karyotypic categories (i.e. "pure" species or hybrids). Additionally, we showed that assignment efficiency depends on the characteristics of the hybrid zones and on how it is analysed (i.e. independently or in comparison to a dataset reference). These results are then discussed in the context of the karyotypic complexity of the hybrids and the importance of using a reference population when analysing hybrid zones is then assessed.

INTRODUCTION

In population biology, individuals are commonly analysed according to their sampling localities. But in numerous situations, it is important to classify individuals into groups using other criteria such as morphology (e.g. Hausser *et al.* 1991, Polly 2003, Motokawa 2004, Yamaguchi *et al.* 2004, Stanley *et al.* 2005), behavioural patterns (e.g. Chapuisat 1998, Hoelzel *et al.* 1998, Jaquiéry *et al.* 2005), or karyotype (e.g. Searle & Wójcik 1998, Dobigny *et al.* 2003, Morgan-Richards & Wallis 2003, Panithanarak *et al.* 2004). Such classification processes are essential when individuals belonging to different groups occur in syntopy (i.e. could be sampled in the same localities). In several circumstances, access to the sorting criteria (e.g. karyotype, behaviour) is difficult and/or requires particular sampling strategies. The increasing availability of information from neutral genetic markers, such as microsatellites, is an alternative approach for individual sorting. Numerous assignment methods based on the genetic composition of individuals have been developed and seem effective in a variety of situations (reviewed in Manel *et al*. 2005).

The shrews of the *Sorex araneus* group offer good opportunities to test the application of classical and genetic assignment methods. This group of Palaearctic species displays a remarkable chromosomal variation and its type species, *S. araneus*, is one of the most chromosomally polymorphic of mammals (Wójcik *et al.* 2003). Species and chromosome races of this group are usually morphologically very similar and almost impossible to tell apart in the field. Traditionally, identification of individuals required karyotype analysis, which was moreover useful for sorting and analysing population occurring in sympatry (for recent examples, see Brünner *et al.* 2002b, Andersson *et al.* 2004). Unfortunately, this procedure generally requires destructive sampling. This can be avoided (Brünner & Hausser 1996) but only using a demanding and expensive cell culture technique. Several other identification methods have been proposed but they all have more or less restrictive drawbacks (reviewed in Basset & Hausser 2003). Recently, Basset *et al.* (in press, Chapter 1) used genetic variation at 10 microsatellite markers in five karyotypic taxa of the *S. araneus* group sampled in the western Alps to investigate the concordance between genetic and karyotypic structure. Bayesian clustering analysis (Pritchard *et al.* 2000) of their dataset produced good assignment results at the species level (although this method performed poorly at the chromosome race level). Therefore, these authors proposed that their current dataset could be used as a reference by those wanting to identify *Sorex* individuals at the species level in the western Alps.

Hybrid zones are of paramount interest in evolutionary biology and are therefore the topic of numerous studies (for reviews, see Harrison 1990, Arnold 1997, Hewitt 2001). Species and chromosome races of the *Sorex araneus* group form various contact or hybrid zones showing an extraordinary variety of gene exchange levels among populations (Searle & Wójcik 1998). Three contact zones between different karyotypic taxa of this group have been studied in the western Alps (Neet & Hausser 1990, Lugon-Moulin *et al.* 1999b, Brünner *et al.* 2002b). By definition, individuals sampled into hybrid zones could include parts of the different hybridizing genomes. In such a situation, utilisation of genetic assignment methods to sort individuals in different categories is not straightforward. However, several assignment methods address the issue of hybridization. In such a context, the method of Pritchard *et al.* (2000) is particularly relevant as it considers that an individual could originate from more than one population.

The goal of our study was first to explore the potential and efficiency of the Bayesian assignment method developed by Pritchard *et al.* (2000) to study admixture and individual assignment in the context of hybrid zones. Additionally, we checked for the same purposes the usefulness of the reference dataset developed by Basset *et al.* (in press, Chapter 1) for studying the Alpine hybrid zones of the *Sorex araneus* group.

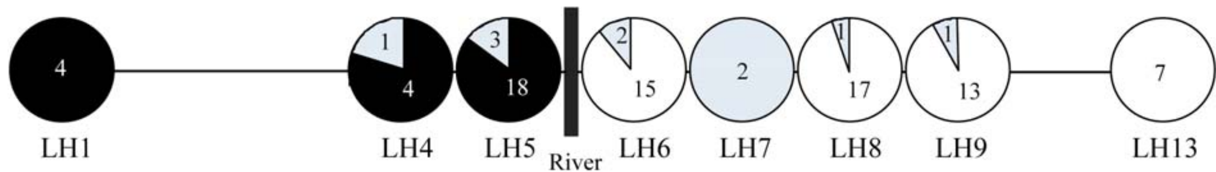
MATERIAL AND METHODS

Sample Collection

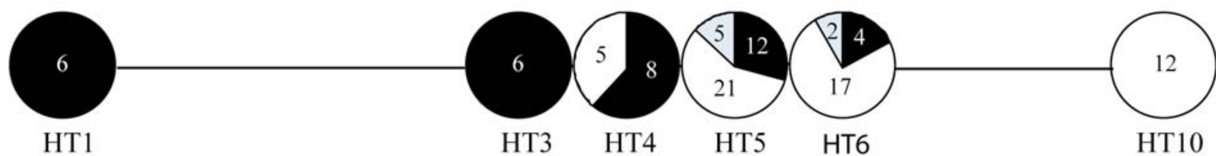
A total of 184 shrews sampled from two hybrid zones were analysed during this study (Fig. 1). Three different karyotypic taxa were found in these zones. The first one, *S. antinorii* (karyotype: XX/XY1Y2, *af*, *bc*, *gi*, *hj*, *kn*, *l/o*, *m*, *p*, q, *r*, *tu*; Brünner *et al.* 2002a) meets *S. araneus* Cordon (XX/XY1Y2, *af*, *bc*, *g*, *h*, *i*, *jl*, *k*, *m*, *n*, *o*, *p*, *q*, *r*, *tu*; Wójcik *et al.* 2003) in the Les Houches hybrid zone (hereafter LH) and *S. araneus* Vaud (XX/XY1Y2, *af*, *bc*, *gm*, *hi*, *jl*, *kr*, *no*, *p*, *q*, *tu*; Wójcik *et al.* 2003) in the Haslital hybrid zone (hereafter HT). Each individual in both hybrid zones was classified into “pure” species or hybrid categories following karyotype analysis according to earlier studies (Brünner & Hausser 1996, Brünner *et al.* 2002b). Eighty-six of these individuals were sampled from 8 localities within the LH hybrid zone (26 *S. araneus* Cordon, 50 as *S. antinorii* and 10 hybrids; Fig. 1A) and the remaining 98 individuals were sampled from 6 localities of the HT hybrid zone (36 *S. araneus* Vaud, 55 *S. antinorii* and 7 hybrids; Fig. 1B). Sampling procedure, characteristics and summaries of the

main results obtained in these two zones were described in Lugon-Moulin *et al.* (1999) and Brünner *et al.* (2002b).

(A) Les Houches



(B) Haslital



1km

Figure 1 Sampling localities in the Les Houches (A) and Haslital (B) hybrid zones. The number of individuals sampled is indicated in each part of the circles. Black: *S. araneus*, white: *S. antinorii*, grey: hybrids.

DNA extraction and microsatellite analysis

Tissue samples were preserved in 100% ethanol and total genomic DNA was extracted using the DNeasy Tissue Kit (Qiagen). The same ten microsatellite loci used in Basset *et al.* (in press, Chapter 1): L9, L67 (Balloux *et al.* 1998), L13, L99 (Lugon-Moulin *et al.* 2000), B3, B5, B10, B15, C5, C19 (Basset *et al.* in press, Chapter 1) were amplified in each individual by polymerase chain reaction and analysed using protocols described in Basset *et al.* (in press, Chapter 1).

Bayesian analyses

Recently, Basset *et al.* (in press, Chapter 1) investigated the genetic structure of a dataset composed of 212 shrews of the *Sorex araneus* group using a Bayesian admixture procedure implemented in STRUCTURE (Pritchard *et al.* 2000; http://pritch.bsd.uchicago.edu). In addition, these authors proposed their dataset (accessible at http://www.unil.ch/dee/page7010.html#5) as a reference to identify individuals sampled in

hybrid zones. Here we tested this assumption on karyotyped individuals sampled in two hybrid zones. In a first step, we studied admixture in both hybrid zones independently. For these analyses, we used the admixture model implemented in STRUCTURE, assuming that sampled individuals belonged to K = 2 genetically distinct clusters without using any prior population information. In a second step, we studied individual assignment in both hybrid zones using the dataset developed by Basset *et al.* (in press, Chapter 1) as a reference. Therefore, we ran STRUCTURE using species information for all individuals of the proposed reference dataset (USEPOPINFO = 1) and we treated the 184 individuals sampled in hybrid zones as having unknown origin (USEPOPINFO = 0). The reference dataset contains three distinct species (*Sorex araneus*, *S. antinorii* and *S. coronatus*) so that we fixed K = 3 for this analysis. In both analyses (with or without comparison to the dataset reference), we performed five repetitions of 100,000 iterations after a burn-in period of 20,000 iterations. Individuals with a proportion of membership to each cluster $0.10 < q_{ind} < 0.90$ (admixed individuals) were assigned to more than one cluster whereas individuals with $q_{ind} \geq 0.90$ were assigned to only one cluster. The threshold value of 0.90 was arbitrarily defined. This means that at least 90% of an individual's genome is assigned to one cluster (Manel *et al.* 2002, Cegelski *et al.* 2003) and seems efficient to discriminate among *Sorex* species (Basset *et al.* in press, Chapter 1).

RESULTS

Les Houches hybrid zone

Independent Bayesian admixture analyses of the LH sample (n = 84) grouped 25/26 *S. araneus* Cordon individuals in one cluster and 42/50 *S. antinorii* in a second cluster (Table 1A and Fig. 2A). Nine of these karyotypically pure individuals showed admixture ($0.10 < q_i < 0.90$) between the two clusters. Four karyotypic hybrids grouped with the *S. araneus* Cordon cluster, five grouped with the *S. antinorii* cluster and one showed admixture between the two clusters (Table 2A).

Analysing the same individuals with reference to the dataset developed in Basset *et al.* (in press, Chapter 1), we assigned 17 of the *S. araneus* to their species (Table 1A and Fig. 2B). Five individuals of this species were assigned to *S. antinorii* and the remaining four showed admixture between species. Forty-four of the *S. antinorii* individuals were assigned to their

species and six showed admixture between *S. araneus* and *S. antinorii*. Three karyotypic hybrids were assigned to *S. araneus* and six to *S. antinorii* (Table 2A).

Table 1 Number of individuals of each karyotypic category assigned to each genetic categories in Les Houches (LH) and Haslital (HT) hybrid zones.

			Genetic assignation		
			S. araneus	***S. antinorii***	**Admixed**
Karyotype	**N**		**[$q_{araneus} \geq 0.90$]**	**[$q_{antinorii} \geq 0.90$]**	**[$0.10 < qi < 0.90$]**
(LH)					
S. araneus	26	Independent	25	0	1
		With ref.	17	5	4
S. antinorii	50	Independent	0	42	8
		With ref.	0	44	6
Hybrids	10	Independent	4	5	1
		With ref.	3	7	0
(HT)					
S. araneus	36	Independent	34	1	1
		With ref.	33	1	2
S. antinorii	55	Independent	0	52	3
		With ref.	0	52	3
Hybrids	7	Independent	1	0	6
		With ref.	1	1	5

NB. For each karyotypic category, the results are given after independent analyses of the hybrid zone dataset (Independent) and with comparison to the reference dataset developed in Basset et al. (in press, Chapter 1), (With ref.).

Thus, 19 individuals gave contradictory results depending on how they were analysed (independent/reference; rectangles in Fig. 2). Six individuals were assigned to one species with the first analysis but to the other with the second, and 13 showed admixture in one analysis but not in the second. To summarise, 9/26 *S. araneus* Cordon, 10/50 *S. antinorii* and 2/10 hybrids showed admixture or incorrect assignment in one or both analyses. These ambiguous individuals had been sampled in localities: LH4 (3 individuals), LH5 (6), LH6 (4), LH8 (4), and LH13 (3).

Table 2 Genetic assignation (independent / with comparison to the dataset reference) of hybrid individuals sampled in Les Houches (LH) and Haslital (HT) hybrid zones according to the most frequent species of their sampling localities.

Sampling Locality	Most frequent species	Nb of hybrids	Genetic assignation *S. araneus* [$q_{araneus} \geq 0.90$]	*S. antinorii* [$q_{antinorii} \geq 0.90$]	Admixed [$0.10 < qi < 0.90$]
(LH)					
4	*S. araneus*	1	1/ 0	0/ 1	0/ 0
5	*S. araneus*	3	3/ 3	0/ 0	0/ 0
6	*S. antinorii*	2	0/ 0	2/ 2	0/ 0
7	*S. antinorii*	2	0/ 0	1/ 2	1/ 0
8	*S. antinorii*	1	0/ 0	1/ 1	0/ 0
9	*S. antinorii*	1	0/ 0	1/ 1	0/ 0
(HT)					
5	Sympatry	5	1/ 1	0/ 1	4/ 3
6	Sympatry	2	0/ 0	0/ 0	2/ 2

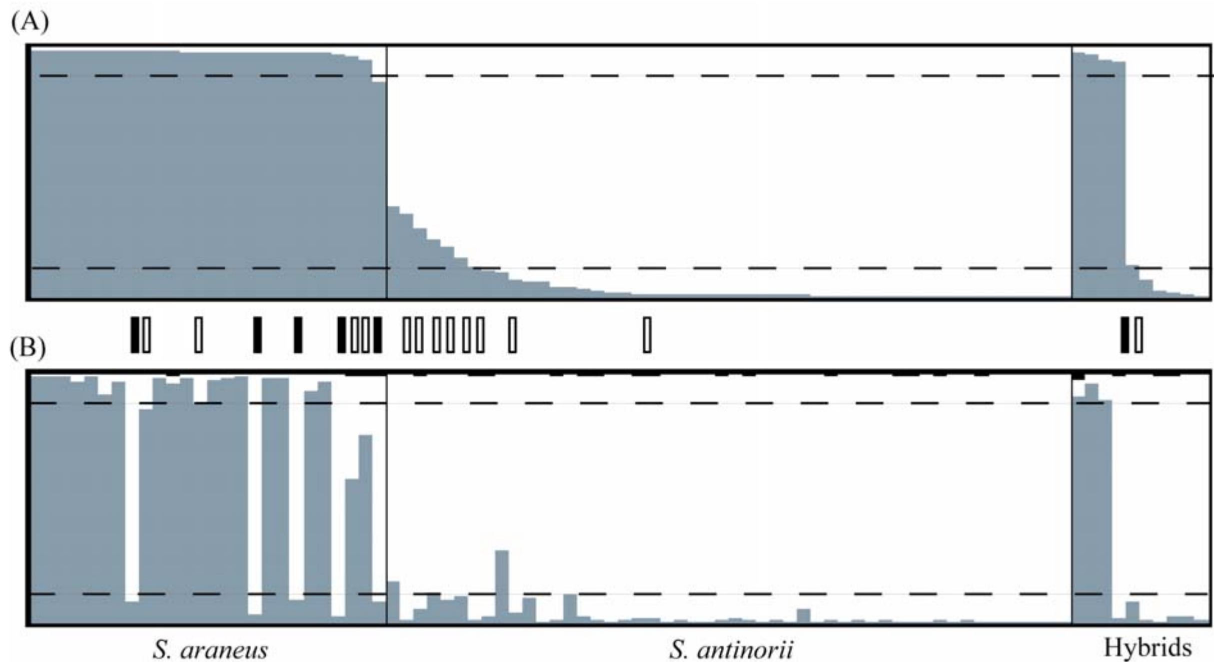


Figure 2 Distruct plots (Rosenberg 2004) for the LH hybrid zone analyses with (A) independent analysis of the 86 individuals assuming two species (K = 2) and (B) comparison to the dataset reference developed in Basset *et al.* (in press, Chapter 1) assuming three species (K = 3). Each individual is represented by a line partitioned into two (A) or three (B) colours representing genotypic assignment to one cluster or the other (q_i). In both analyses, individuals are sorted according to their karyotypic category (*S. araneus*, *S. antinorii* and hybrids), and then individuals are sorted according to decreasing q_i value obtained in the first analysis. Assignment limits ($q_i = 0.10$ and 0.90) are indicated by dashed lines and individuals showing different results according to analyses are indicated by open (change from admixed to one species) or black (change from one species to the other) rectangles.

Haslital hybrid zone

Independent Bayesian admixture analyses of the HT sample (n = 98) grouped 34/36 *S. araneus* Vaud in one cluster and 52/55 *S. antinorii* in a second cluster (Table 1B and Fig. 3A). Four of these karyotypically pure individuals showed admixture ($0.10 < q_i < 0.90$) between these clusters and one *S. araneus* grouped with the *S. antinorii* cluster. Six of the karyotypic hybrids showed admixture between the two clusters and one grouped with the *S. araneus* cluster (Table 2B).

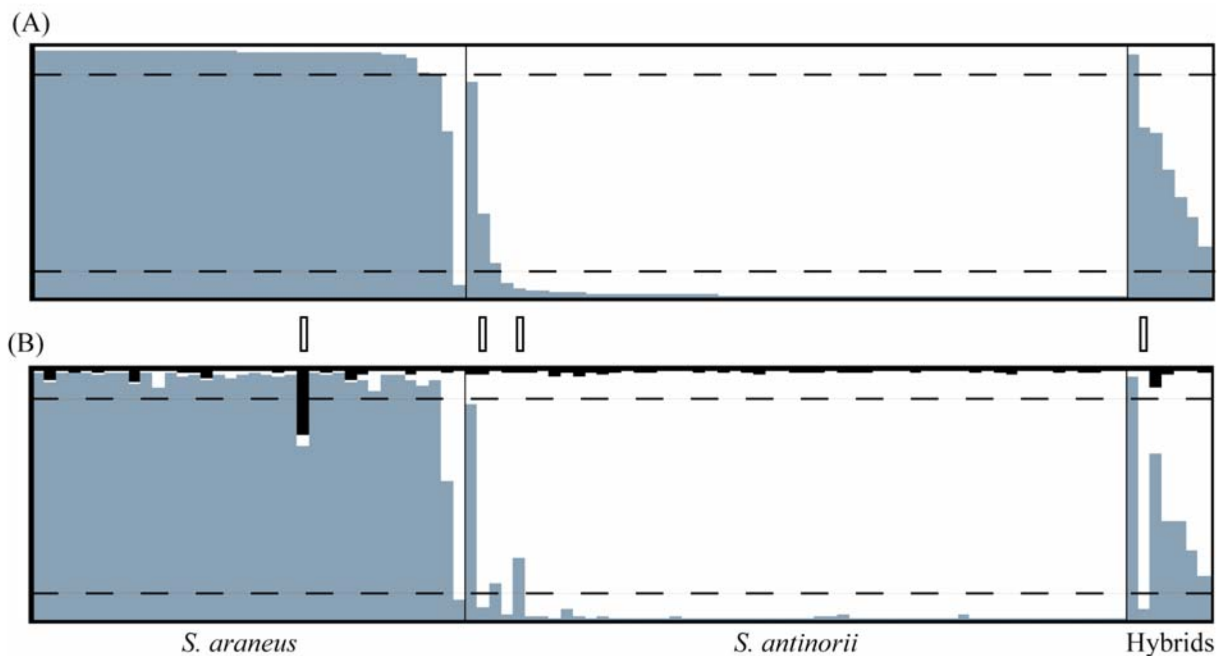


Figure 3 Distruct plots (Rosenberg 2004) for the HT hybrid zone analyses with (A) independent analysis of the 98 individuals assuming two species (K = 2) and (B) comparison to the dataset reference developed in Basset *et al.* (in press, Chapter 1) assuming three species (K = 3). Each individual is represented by a line partitioned into two (A) or three (B) colours representing genotypic assignment to one cluster or the other (q_i). In both analyses, individuals are sorted according to their karyotypic category (*S. araneus*, *S. antinorii* and hybrids), and then individuals are sorted according to decreasing q_i value obtained in the first analysis. Assignment limits (q_i = 0.10 and 0.90) are indicated by dashed lines and individuals showing different results according to analysis are indicated by open (change from admixed to one species) rectangles.

Analysing the same individuals with reference to the dataset developed in Basset *et al.* (in press, Chapter 1), we assigned 33/36 *S. araneus* to their species (Table 1B and Fig. 3B). One individual of this species was assigned to *S. antinorii*, one showed admixture with *S. antinorii* and one showed admixture with *S. coronatus*. Fifty-two *S. antinorii* were assigned to their species and the remaining three individuals showed admixture with *S. araneus*. Five of the

seven karyotypic hybrids showed admixture between *S. araneus* and *S. antinorii*, one was assigned to *S. araneus* and one was assigned to S. *antinorii* (Table 2B).

Thus, only four individuals gave contradictory results according to the type of analysis (rectangles in Fig. 3) and they all showed admixture in one analysis but not in the other. To summarise, 3/36 *S. araneus* Vaud, 4/55 *S. antinorii* and 6/7 hybrids showed admixture or incorrect assignment in one or both analyses. These ambiguous individuals were sampled in locality HT3 (1 individual), HT4 (1), HT5 (4) and HT10 (1).

DISCUSSION

Our study illustrates that assignment power of genetic Bayesian clustering methods (STRUCTURE, Pritchard *et al.* 2000) in two hybrid zones between karyotypic species of the *Sorex araneus* group is generally good. A total of 149 of the 184 individuals (81%) tested were correctly assigned to one of the three categories sampled (karyotypically pure *S. araneus* or *S. antinorii* and hybrids) whatever the type of analysis performed (independent or comparison to a reference dataset). Although individuals were sampled in the difficult context of hybrid zones, the frequency of correct classification is only slightly lower to that Basset *et al.* (in press, Chapter 1) noticed in their dataset of individuals sampled away from hybrid zones. These authors found correct classifications in 92% of *S. antinorii* and 86% of *S. araneus* and their values were similar to those obtained using morphometric assignment techniques (Hausser 1984, Hausser *et al*. 1991, Brünner *et al*. 2002a). This concordance illustrates the general strong genetic differentiation already observed between these two species (Taberlet *et al.* 1994, Brünner *et al.* 2002b) compared to chromosome races within *S. araneus*. The situation is for example quite different to that noticed in a hybrid zone between chromosome races of *S. araneus* in Sweden (Andersson *et al.* 2004). Using the same Bayesian admixture protocol (STRUCTURE), these authors did not manage to distinguish any population structure among two karyotypic groups of *S. araneus*. It is important to note the weak genetic structure among chromosome races observed in their dataset (F_{ST} = 0.015) compared to the genetic structure between species observed in our hybrid zones (F_{ST} LH = 0.103, Lugon-Moulin *et al.* 1999; F_{ST} HT = 0.107, Brünner *et al.* 2002b) since assignment power of Bayesian clustering methods is generally well correlated with genetic structure (Berry *et al.* 2004). Additionally, success of assignment of alpine *S. araneus* chromosome races has already proven to be particularly low (Basset *et al.* in press, Chapter 1). Finally, several characteristics of the genetic markers (e.g. homoplasy, Zhang & Hewitt 2003) or

methodological factors (e.g. number of markers, stringency level of assignation; Berry *et al.* 2004) might explain some of the assignment uncertainties observed. However, it is essential to note that the number of ambiguously assigned individuals also probably reflect the real genetic introgression between the studied species (Basset *et al.* in press, Chapter 1). For example, decreasing the stringency of the assignation threshold to $q_i = 0.80$ increases correct assignation to 85% (instead of 81%).

Difference between LH and HT hybrid zones

Interestingly, the two hybrid zones did not show similar levels of admixture. The number of admixed or ambiguous individuals detected in the LH hybrid zone is indeed much larger than in the HT hybrid zone (31% vs. 9%; Table 1). Differences between the two hybrid zones were expected since the chromosome composition of the *S. araneus* races in contact with *S. antinorii* are not the same. The Cordon chromosome race present in the LH hybrid zone has one of the most acrocentric karyotypes known in *S. araneus*. Most hybrids with *S. antinorii* form well tolerated trivalents at meiosis (Brünner *et al.* 2002b). In the HT hybrid zone, both hybridizing taxa have distinct metacentric chromosomes. Therefore hybrids between these taxa form complex chains of elements and should meet much more serious chromosomal incompatibilities (Narain & Fredga 1997, 1998, Banaszek *et al.* 2002). These differences are well illustrated by the situation of hybrids in both zones. In the LH hybrid zone, all karyotypic hybrids detected were F_x backcrosses (with $x > 1$; Brünner *et al.* 2002b) and in our analysis they show very low signs of admixture. Most of them (8/10) strongly clustered with the most common species present in their sampling localities (Table 2A). In contrast, in the HT, most of the analysed hybrids were F1 (Brünner *et al.* 2002b) and they showed in our analysis clear admixture signs between the two parent species ($0.10 < q_i < 0.90$; Table 2B). Interestingly, the average q_i value for these F1 hybrids was not significantly different from the expected $q_{antinorii} \approx q_{araneus} \approx 0.5$ value ($P < 0.001$, t-test), which validates the F1 status of these individuals. It is worth noting that we analysed a backcross hybrid with *S. araneus* and this individual strongly clustered with this species. Similar results were obtained by Lugon-Moulin *et al.* (1999b) and Brünner *et al.* (2002b) using principal component analysis (PCA) on microsatellites.

Independent analysis or comparison to a reference?

Surprisingly, depending on how they were analysed several individuals showed very contrasting results in the LH hybrid zone. Five karyotypic *Sorex araneus* were assigned to *S. antinorii* (Fig. 2B) when analysed in comparison to the dataset reference developed in Basset

et al. (in press, Chapter 1), but the same individuals clustered with *S. araneus* when this hybrid zone was analysed independently (Fig. 2A). The large difference observed between the two analyses is explained by the particular genetic composition of ambiguous individuals and by differences in analysis processes. When a dataset is analysed with a reference, each individual is analysed independently and compared to the genetic composition of the reference only. In contrast, when a dataset is analysed independently, the genetic composition of an individual is compared to the rest of the dataset. In our situation, ambiguous *S. araneus* bear alleles present in the *S. araneus* sampled in LH and in the *S. antinorii* reference. Moreover, some of these alleles are absent from the *S. antinorii* sampled in LH. Consequently, these individuals cluster to *S. antinorii* when compared to the reference but to *S. araneus* when compared to the rest of the dataset. The presence of numerous alleles apparently specific to *S. antinorii* in individuals with a *S. araneus* karyotype remains difficult to explain. Introgression of *S. antinorii* alleles into *S. araneus* karyotypes is expected if hybrids are fertile and backcrosses frequent. However, as previously mentioned, some of the *S. antinorii* discriminative alleles have not been sampled in LH. This paradox could be explained by the detection of two genetically distinct groups in *S. antinorii* by Lugon-Moulin *et al.* (2002) and Basset *et al.* (in press, Chapter 1). These groups were consistent with the post-glaciations recolonization routes of this species (i.e. St Bernard pass region vs. Simplon pass region) but the exact geographical distribution of both groups was unknown. Interestingly, if we take into account these two groups in our analyses, the *S. antinorii* sampled in LH cluster with the "St Bernard group" but the ambiguous *S. araneus* cluster with the "Simplon group". *Sorex antinorii* sampled in the southern French Alps have also been found to cluster with the "Simplon group" (Basset *et al.* in press, Chapter 1). It is therefore likely that *S. araneus* Cordon has – or had – contacts and genetic exchanges with *S. antinorii* from the "Simplon group". Further studies about the distribution, recolonization routes and genetic composition of *S. antinorii* populations in the LH region are necessary to clarify this topic.

These observations illustrate the importance of sampling in hybrid zone analyses. Moreover, the fact that correct assignment of these ambiguous individuals was obtained when analysing each hybrid zone independently and that introgression was suggested by analyses using the reference dataset, illustrates that these two analyses should not be used for the same purposes. Thus, if the main goal of a study is to locally distribute individuals into two or more categories, it is probably better to analyse the dataset independently. In contrast, if the aim of

a study is to detect introgression between taxa, utilisation of a reference is then probably recommended.

Conclusions

Using Bayesian clustering analysis, we have been able to assign more than 80% of individuals sampled into two hybrid zones between species of the *Sorex araneus* group to their respective karyotypic categories (i.e. “pure” species or hybrids). Additionally, we have shown that assignment efficiency depends on the characteristics of the hybrid zones (i.e. assignment power was larger in HT than in LH) and on how it is analysed (i.e. independently or in comparison to a dataset reference).

Although for fine scale studies, karyotyping individuals is always recommended, genetic methods represent good alternatives in numerous situations such as localization and characterization of new hybrid zones, analyses of numerous individuals or when it is necessary to decrease levels of invasiveness (e.g. for conservation purposes). This topic is particularly relevant when the genetic differentiation between taxa is large. Additionally, it offers complementary and interesting insights into the processes actually acting in and across hybrid zones.

ACKNOWLEDGEMENTS

We are particularly grateful to Lori Handley for checking our English. This work was supported by the Swiss NSF (grants n°: 3100-05943), a Société Académique Vaudoise (Switzerland) grant to PB and a Fondation Agassiz (Switzerland) grant to PB.

CHAPTER 3

Chromosome localization of microsatellite markers in the shrews of the *Sorex araneus* group

Patrick Basset, Glenn Yannic, Fengtang Yang, Patricia C.M.O'Brien, Alexander S. Graphodatsky, Malcolm A. Ferguson-Smith, Gabriel Balmus, Vitaly T. Volobouev and Jacques Hausser

(A slightly modified version of this chapter has been accepted in *Chromosome research*)

ABSTRACT

The shrews of the *Sorex araneus* group are characterized by spectacular karyotypic evolution. This makes this group an exceptionally interesting model for population genetics and evolutionary studies. Here, we mapped 46 microsatellite markers at the chromosome arm level using flow sorted chromosomes of three karyotypically different taxa of the *Sorex araneus* group (*S. granarius*, *S. araneus* chromosome races Cordon and Novosibirsk). Twenty-five loci were unambiguously mapped to only one chromosome arm in the three taxa, whereas 21 loci were assigned to multiple chromosomes. Unambiguously mapped loci marked the three sexual chromosomes (XY1Y2) and 9 of the 18 autosomal arms of the *S. araneus* group. Only one locus showed discordance among the taxa studied, suggesting that despite the presence of numerous Robertsonian rearrangements, the organisation of the genome in the *S. araneus* group is well conserved. Consequently, we propose that these markers could be used to compare genetic structure among taxa of the *S. araneus* group at the chromosome level. This would constitute a valuable tool for identifying the role of chromosomal rearrangements in the genetic diversification and speciation process of this group.

INTRODUCTION

Comparative gene mapping is of primary interest to understand the evolution of the mammalian karyotype (O'Brien *et al.* 1999). Among mammals, efforts have mostly been concentrated on humans or laboratory and farm animals and have revealed a surprising conservation in mammalian genome organizations. However, to get a satisfactory picture of the genome changes that have occurred during mammalian radiation, it is important to expand the range of Orders examined (O'Brien *et al.* 2001). In this context, Insectivores (shrews, hedgehogs or moles) are of special interest. Morphologically, they appear to be the closest to the ancestral eutherian condition (Nowak 1991) and DNA markers suggest their paraphyletic origin (Arnason & Janke 2002, Murphy *et al.* 2004).

Among Insectivores, the shrews of the *Sorex araneus* group have been subjected to detailed phylogenetic and population genetic analyses involving chromosome and genetic markers (for reviews Searle & Wójcik 1998). This monophyletic group of 10 species is characterized by the particular sexual chromosome complex XY1Y2 in males (Sharman 1956) and by a spectacular karyotypic diversification, in spite of their high-degree of morphological similarity (Zima *et al.* 1998). In the type species of this group, *Sorex araneus*, Robertsonian polymorphisms are particularly prevalent. The ancestral karyotype of this species is thought to consist of acrocentric chromosomes only. Then, repeated Robertsonian fusions led to the formation of various metacentric chromosomal complements seen in more than 60 different chromosomal races, each characterized by a particular set of metacentrics and acrocentrics (Wójcik *et al.* 2003). This outstanding karyotypic variation makes *S. araneus* an exceptionally interesting model species for population genetics and evolutionary studies.

Given its large distribution, its abundance, its ancestral mammalian morphology and its peculiar genome evolution, *S. araneus* represents an obvious "type" species for the Insectivores. Consequently, a gene mapping project of the common shrew chromosomes was undertaken (Pack *et al.* 1995) and its genome selected as candidate for complete sequencing (O'Brien *et al.* 2001, Pennisi 2004). The *S. araneus* genetic map currently contains 53 markers (Zhdanova *et al.* 2003) which mark the X (*de*) chromosome and 7 out of the 9 metacentric chromosomes of the Novosibirsk race. There are, however, some gaps to be filled in order to take full advantage of this map to address many important evolutionary problems. For example, a genetic map of microsatellite loci would provide a powerful tool to unravel the role of chromosomal rearrangements in the speciation process. Indeed, over the entire range of the *S. araneus* group, several populations characterized by different karyotypes come into

contact and hybridize. Recently, new models of speciation have been developed to account for the proposed role of chromosomal rearrangements in speciation (Rieseberg 2001, Noor *et al.* 2001, Navarro & Barton 2003a). But, in the case of the *S. araneus* group, it is currently not possible to distinguish between a reproductive barrier caused by genetical effects or by karyotypic differences (Lugon-Moulin *et al.* 1999b, Brünner *et al.* 2002b). This question can be addressed with a battery of microsatellite markers mapped at the chromosome level. However, only four of these markers have been located so far (Zhdanova *et al.* 2003).

Traditionally, gene mapping in the *S. araneus* group dealt with somatic cell hybrid panels (Pack *et al.* 1995). But, due to numerous hidden chromosome rearrangements in panels this method proved laborious (Zhdanova *et al.* 2003). Mapping accuracy would be further improved with the use of mostly acrocentric taxa, but only the all metacentric *S. araneus* Novosibirsk chromosome race has been used so far.

In this study, we aimed to map 46 microsatellite loci at the chromosome arm level using flow sorted chromosomes of three karyotypically different taxa of the *S. araneus* group. The first taxon is *Sorex granarius*, a species genetically very similar to *S. araneus* (Fumagalli *et al.* 1999), which has retained the ancestral karyotype of the *S. araneus* group (Wójcik & Searle 1988, Volobouev 1989). Indeed, all its autosomes (except the smallest one *tu*) are in an acrocentric state. The second taxon, *S. araneus* Cordon, is one of the most acrocentric *S. araneus* chromosome races since most of its autosomes (except the two large *af* and *bc*, the smallest one *tu* and the polymorphic *j*/*l*) are acrocentric. Finally, we added *S. araneus* Novosibirsk, the traditional "gene mapping race" for this group of taxa, although all its autosomes are metacentric. The comparison of the results among three taxa will identify markers appropriate for further inter-taxa population genetics studies.

MATERIAL AND METHODS

Chromosome isolation

Chromosomes from fibroblast cell lines of one female of *Sorex granarius* and *S. araneus* Cordon and one male of *S. araneus* Novosibirsk were sorted on a bivariate fluorescence activated flow sorter as described previously (Yang *et al.* 1995). Their respective karyotypes consisted of 2n = 36 (*a*, *b*, *c*, *f*, *g*, *h*, *i*, *j*, *k*, *l*, *m*, *n*, *o*, *p*, *q*, *r*, *tu*, X(*de*)), 2n = 30 (*af*, *bc*, *g*, *h*, *i*, *jl*, *k*, *m*, *n*, *o*, *p, q*, *r*, *tu*, X (*de*); Fig. 1A) and 2n = 21 (*af*, *bc*, *go*, *hn*, *ik*, *jl*, *mp*, *qr*, *tu*, X(*de*), Y1(*s*), Y2(*d*)) chromosomes. Note that each of the 21 chromosome arms of the *S. araneus*

group is labelled by a letter (*a* – *u*), with "*a*" indicating the largest arm and "*u*" the smallest (Searle *et al.* 1991). Chromosome-specific DNA from these individuals was obtained by degenerate oligonucleotide primed PCR (DOP-PCR) amplification of flow-sorted chromosomes following standard procedures (Telenius *et al.* 1992, Yang *et al.* 1995). The characterization and purity of each chromosome sort was checked by fluorescence *in situ* hybridization (FISH). Painting probes of each sort were hybridized to standard metaphase preparations of their respective species and examined by digital fluorescence microscopy according to standard protocols (Yang *et al.* 1995, Fergusson-Smith 1997). Images were captured and processed as described in Yang *et al.* (1999).

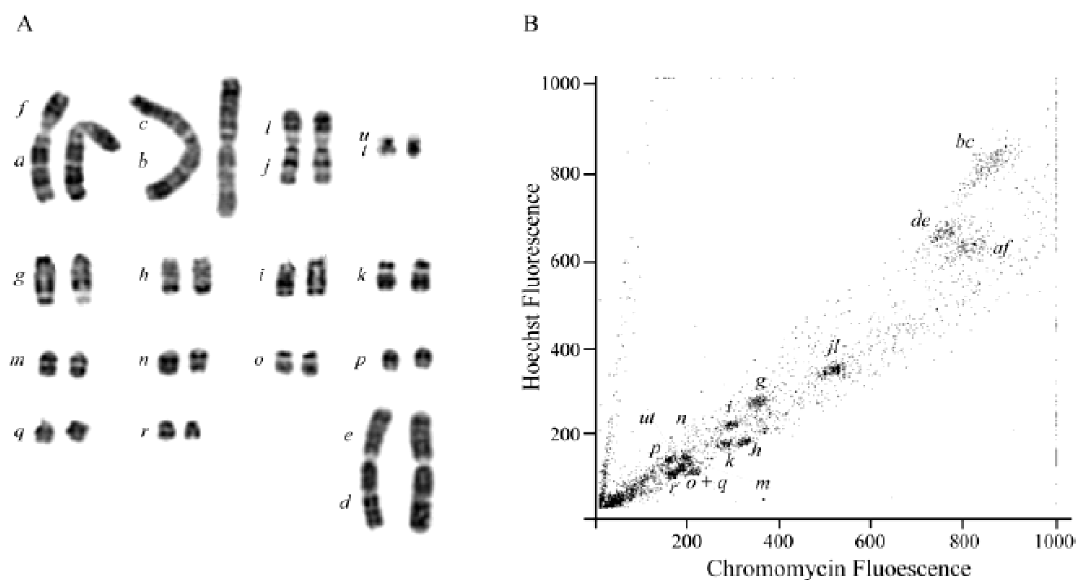


Figure 1. G-banded karyotype (A) and flow karyotype (B) of the female *Sorex araneus* Cordon.

Sorex araneus group microsatellite loci

Twenty-two markers have already been described in the *Sorex araneu*s group (Wyttenbach *et al.* 1997, Balloux *et al.* 1998, Balloux *et al.* 2000, Lugon-Moulin *et al.* 2000 and Basset *et al.* in press, Chapter 1). Mapping efficiency is sometimes low (Zhdanova *et al.* 2003), therefore this number was not sufficient to ensure that our microsatellite genetic map would mark most of the chromosomes. For this reason, we described another 24 loci extracted from four microsatellite enriched libraries developed by a commercial company (Genetic Identification Services, Inc. Chatsworth, CA). These loci were found to be highly polymorphic and add to the 22 previously characterized loci to give a total of 46 microsatellite markers described for the *Sorex araneus* group (Table 1).

Table 1. PCR primers and amplification conditions used for mapping the *Sorex araneus* group microsatellite loci on *Sorex* chromosomes.

MS	Accession Number	Primers (5' - 3' ; For/Rev)	Motifs [a]	Annealing [C°]	MgCl2 [mM]	Refs
L2	U82710	CAAAAACAAAAAAGAAGAAAGAAG TCTTTATCCTCCATTCCCTC	(GGA)12	55	1.5	1
L9	U82711	TCATGGACTTTTCTGTGCTG CTTTGGCATGAATTTGCC	(AC)29	55	1.0	1
L13	AF175741	TATAACTGTTATCTCACAGCGATTCA ATCCATCTTTATCTTTTCCATTGC	(AATT)6	55	1.5	3
L14	AF032911	AGGGAGGGAAACTTGTTAAAGG GTGAGGTCCTGGAATAGTGTCC	(AC)14	55	1.0	2
L16	U82712	TCAGAGTCAGAATTTCTAATTTGGC TTAGTGTATTATGACAGATGCGGG	(AC)16	55	1.5	1
L33	AF032912	TGATTGTATGGAAAAAAGCCG ATACTAAGACCACCCTACCAATGC	(AC)19	55	1.5	2
L45	U82713	CTTAAACGTTCTTATCTATTTGGTTG GACATATGTGCACTATGAAATTATTG	(AC)10	55	1.5	1
L57	U82714	CTGTTTTTCTGTCCCTCATAGC TGTCCTAGTGACATTATCCTATTGG	(AC)10	55	2.0	1
L62	U82715	CAGTCTCTCACTGTGGCACTATG GTCATTCTGGATAAGAACCATATGC	(AC)16	55	1.5	1
L67	U82716	GAAGTGATACATGAGTGCATGAG GTTGTTAACAAGAGAGGTATTACACC	(AC)17	55	1.5	1
L68	AF032913	TCATGGTCATTTCATCACATACC GTAGATGTTGCCACTGGTGG	(AC)14	55	1.5	2
L69	U82717	CTTTATGGTAGAAAATGGTG GACCATATACTAAGTTGTTTTG	(AC)17	57	1.5	1
L92	AF032914	ACTGGTGCCCAATCGATAAG GAGAATTGTTGGATGTGCCC	(AC)7	55	1.5	2
L97	AF032915	ATTCTCGTGGGTAGACCGTG ATAAATGTGGGAAATGGACAGG	(AC)56	55	1.5	2
L99	AF175744	AAATAATTTCTTCCTGGCAAG ATAAATGCAGCAAAGTTATAAACTT	(AC)6	55	1.5	3
L8Y	AF175743	CCTCTTTTGTTTCTCATCATTTTC AATGATGAAGTGGATGAGTTAGATACC	(GAA)20	55	1.5	4
A8	DQ247975	CTACGCGCCTTCTTTCAGTC GAAGCTGTCCACTGTGTAAACG	(CA)26	57	1.5	6
A25	DQ247976	GGCAGTGCTCAGGGATAAC AGTGAGGACAGAATTTCAGGTG	(CA)25	57	1.0	6
B3	DQ074646	CTTGCCACATTCCCACATC AGCCCCACAGCTTTCTCC	(GA)30	57	1.0	5
B5	DQ074647	ATGTCTTGCTGGCTGAAGG CTGCTGTTCACAAACTCCAAG	(CA)12(GA)19	55	1.5	5
B7	DQ247977	AGACGCCCTTGTTCTCTCC CCCAGGACTTTCGGTTCTTAC	(GA)29	60	1.5	6
B10	DQ074648	CTCCAAACCCTAACACTCTGTC TTCACGTGTTCTTTGCTTCC	(GA)30	55	1.5	5
B12	DQ247978	GGGGCTTTCTCCACTCTTG TGCTCAGACCTTGATTAGACACTC	(GA)34	60	1.5	6

N.B. References: (1) Wyttenbach *et al.* 1997, (2) Balloux *et al.* 1998, (3) Lugon-Moulin *et al.* 2000, (4) Balloux *et al.* 2000, (5) Basset *et al.*, in press (Chapter 1), (6) This study; [a] Represents the size of the original clone.

Table 1. continued.

MS	Accession Number	Primers (5' - 3' ; For/Rev)	Motifs	Annealing [C°]	MgCl2 [mM]	Refs
B15	DQ074650	GTAGAGTTGCTGGCTCAAAGG ATGGGAAGACATTGGATTGG	(GA)31	55	1.5	5
B30	DQ247979	TCTCCCTTATCCCGCTGTC ACGAAAGGCTGCAACTCAAC	(GA)26	55	1.5	6
C5	DQ074649	TAGATGACTCTGTGTTCAGGC GTTGGGAAGGTAAGATCAGG	(GA)35(CA)12	55	1.5	5
C19	DQ074651	TGCCATAAACACCACTTACC GTGATCAATACCCTGTGGAG	(GA)22	60	1.5	5
C25	DQ247980	CCCAGGCATAAGTTTCAGG TGTGAACTGTGGTGGATAGATG	(GA)29	57	1.5	6
C100	DQ247981	CTCGGTGTTTCTACGAT CAGAGATAGAAGAGGCCAAG	(GA)21	55	2.0	6
C117	DQ247982	TAGATGACCAGGATGGAG ACAGAGCTGGGAATCAGT	(GAT)24	55	1.5	6
C119	DQ247983	CCAGCCTTTACTTCTGCTAC TGGGTCTCATTCCTCTGAC	(CAT)29	50	1.5	6
C122	DQ247984	AGTTTTCTTCTCGCCCGTCT CCACTGTGCCAAGGATAGTT	(CT)17	57	1.5	6
C151	DQ247985	CAACGGAGACATTACTGGTG CCAAACTCAAAGGCAGGA	(TGA)30	55	1.5	6
C171	DQ247986	GTGACTGTTCCCATGATGAC ACCAATGTCCCCAGTTTC	(GA)25	55	1.5	6
C240	DQ247987	GGGTTCAATCTCCAACATCC ATCCTGCCCTTCTTTTCCTC	(GA)22	55	1.5	6
D11	DQ247988	GTGTCGAGAGTCGGAAAACC AGCCAGGAACAAGCCCTAC	(TAGA)15	57	1.5	6
D23	DQ247989	ATGGTGGAAAGGCTCAAG AAAAGCAGTATTGGGTCTGG	(TAGA)23	57	1.5	6
D24	DQ247990	CCCAGAGTTACCTTTGAGATATGC TCAATTTTCCCTGGAGGATG	(GA)38	55	1.5	6
D29	DQ247991	GGAAGCAGCGTGAGACTACC AATGGAGACGTGACTGAGACC	(ATC)9	65	1.5	6
D103	DQ247992	TTATGCCACTGATACACCAA ATCCAAAAGGGTTTCCTTAC	(CTAT)12	55	2.0	6
D106	DQ247993	ATTTCTCCCTTCAATCTGGT AGGAGTACCTCTGGGTGTG	(CTAT)10	55	2.0	6
D107	DQ247994	AGGAAGACTGGGGGTATGTT TAGGTCTGCTGCCTGCAT	(CTAT)17	55	2.0	6
D109	DQ247995	TGAACTTGGGAGATGCAAT ATAGGAGAGGGCAAGCAG	(CTAT)15CG (CT)15	55	2.0	6
D110	DQ247996	TGTTTTGGTTGAGGTTGG TCACACGCCATCAGTAAGT	(CTAT)36	55	2.0	6
D112	DQ247997	GCAAACTACCTGTGGCGTATT CCAGCCCTCTTATGAAACTCTT	(CTAT)20	60	2.0	6
D138	DQ247998	ACCTGGAGTGACAGTGAGC GGGTGCTGGAGTGACAGTAT	(CTAT)21	55	1.5	6

Mapping microsatellite loci to chromosome

Flow-sorted chromosomes were amplified a second time by DOP-PCR (Telenius *et al.* 1992). This second DOP-PCR amplification step was necessary to have enough material and sensitivity to test all loci. Each chromosome-specific DNA was then screened for the presence of each *Sorex araneus* group microsatellite locus by standard PCR using conditions as described in Table 1.

PCR conditions of formerly described markers are given in Wyttenbach *et al.* (1997), Balloux *et al.* (1998), Balloux *et al.* (2000), Lugon-Moulin *et al.* (2000) and Basset *et al.* (in press, Chapter 1). PCR conditions of the markers described in this study were as follow: 0.2 mM dNTPs, 0.325 μM of each primer, 1x PCR buffer (Qiagen) and 0.5U Taq polymerase (Qiagen). $MgCl_2$ concentration as well as annealing temperature varied (Table 1). For all primers, PCR amplifications were performed on 20-50 ng of DOP-PCR product in a final volume of 20μl and cycling was carried out in a PE 9700 thermal cycler (Applied Biosystems) using the cycling profile: 95°C for 5 min; 35 cycles of 30s at 94°C, 30s at the annealing temperature, 30s at 72°C; and a final extension at 72°C for 4min.

To guarantee correct assignments, each locus was tested against each chromosome sort of the three taxa studied, two to four times. Preliminary analyses showed that unspecific amplifications were sometimes observed in several chromosome sorts. Therefore, we used whole DNA from sorted individuals as positive controls whose amplification product sizes were used as references. To better estimate the size of the amplification product, one primer of each pair was labelled with a fluorescent dye on the 5'end, and run on an ABI 377XL automated sequencer (Applied Biosystems). Data collection, sizing and analyses of PCR products were done using GENESCAN software (Applied Biosystems).

RESULTS

Chromosome flow-sorting

Bivariate chromosome flow sorting allowed obtaining chromosome specific sorts for the three taxa (Table 2). Sixteen distinct peaks were identified in the flow karyotype of *Sorex granarius*, 14 in *S. araneus* Cordon (Fig. 1B) and 12 in *S. araneus* Novosibirk. To assign the content of each peak to particular chromosomes, painting probes from each peak were hybridized to metaphase preparations of each taxon (see Fig. 2 for examples in a male of *S. araneus* Cordon). In *S. granarius*, a single chromosome was found in 12 peaks, whereas the

other four peaks contained multiple chromosomes (*j* + *k* + *l*, *r* + *tu*, *q* + *r*, *m* + *o*; Table 2). In addition, chromosomes *m* and *o* were each represented in a second individual peak. In *S. araneus* Cordon, a single chromosome was found in 13 peaks and one peak contained two chromosomes (*o* + *q*; Table 2). Finally, all 12 peaks of *S. araneus* Novosibirsk each contained a single chromosome (Table 2). Most chromosomes could reliably be assigned to a specific sort; therefore, these sorts may be used for genetic mapping purposes.

Table 2. Chromosome-specific sort compositions of *Sorex granarius*, *S. araneus* Cordon and *S. araneus* Novosibirsk.

S. granarius	*S. araneus* Cordon	*S. araneus* Novosibirsk
a	*af*	*af*
f		
b	*bc*	*bc*
c		
g	*g*	*go*
o	*o* + *q*	
h	*h*	*hn*
n	*n*	
i	*i*	*ik*
j + *k* + *l*	*k*	
	jl	*jl*
m	*m*	*mp*
m + *o*		
p	*p*	
q + *r*	r	*qr*
r + *tu*	*tu*	*tu*
X(*de*)	X(*de*)	X(*de*)
		Y2(*d*)
		Y1(*s*)

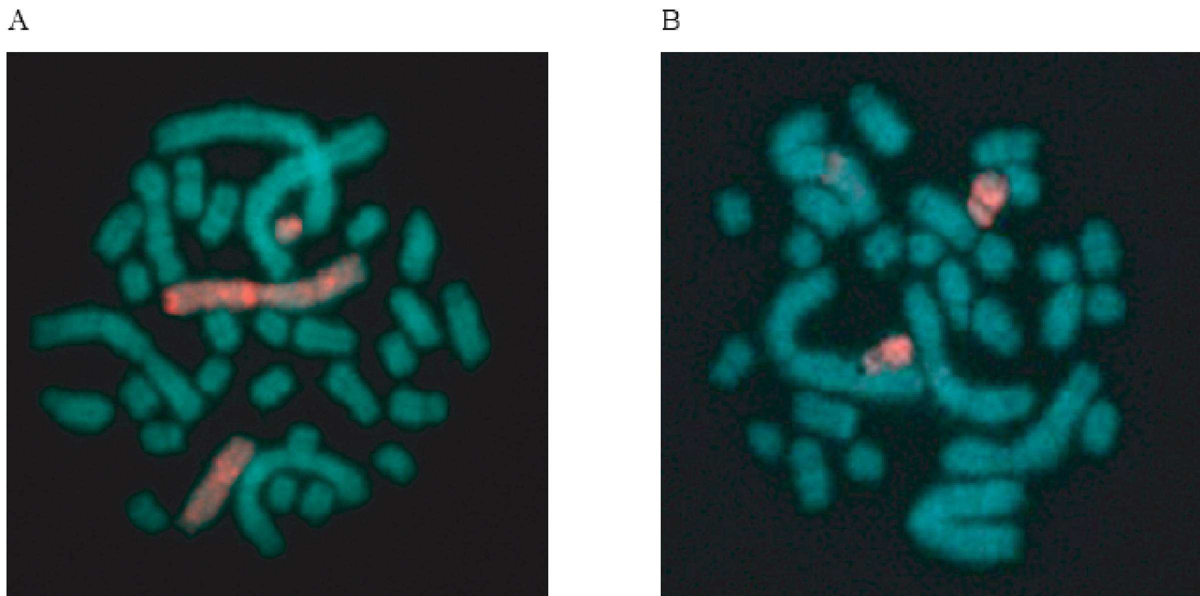


Figure 2. Examples of *Sorex araneus* Cordon male metaphase spreads showing fluorescence of individual chromosome painted by chromosome-specific probes generated from flow sorted chromosomes: (A) Chr X (*de*); (B) Chr *k*.

Microsatellite mapping

The mapping results of the 46 microsatellite loci for the three *Sorex* taxa are summarized in Table 3 and 4.

Sorex granarius. Forty-one primer pairs amplified the expected size fragment in at least one chromosome sort. Three pairs amplified the expected size fragment in the positive control only and 2 failed to amplify in both positive and chromosome sorts in spite of several attempts to optimize PCR conditions. Among the 41 loci showing positive amplification in *S. granarius*, 26 were each assigned to only one chromosome sort and 15 were assigned to multiple sorts.

S. araneus Cordon. Thirty-eight primer pairs amplified the expected size fragment in at least one chromosome sort. Five pairs amplified the expected size fragment in the positive control only and 3 failed to amplify in both positive and chromosome sorts in spite of several attempts to optimize PCR conditions. Among the 38 loci showing positive amplification in *S. araneus* Cordon, 23 were each assigned to only one chromosome sort and 15 were assigned to multiple sorts.

S. araneus Novosibirsk. Forty primer pairs amplified the expected size fragment in at least one chromosome sort. One pair amplified the expected size fragment in the positive control only and 5 failed to amplify in both positive and chromosome sorts in spite of several

attempts to optimize PCR conditions. Among the 40 loci showing positive amplification in *S. araneus* Novosibirsk, 21 were each assigned to only one chromosome sort and 19 were assigned to multiple sorts.

Table 3. Unambiguous mapping of 25 microsatellite markers in the three karyotypic taxa *Sorex granarius*, *S. araneus* Cordon and *S. araneus* Novosibirsk.

Locus	*Sorex granarius*	*S. araneus* Cordon	*S. araneus* Novosibirsk
L2	*b*	n.a.	*bc*
L9	*c*	*bc*	-
L13	*de*	*de*	*de*
L16	*a*	*af*	*af*
L57	*de*	*de*	*de /y2*
L62	*g*	*g*	*go*
L68	*b*	*bc*	*bc*
L69	*f*	-	*af*
L99	*n*	*n*	*hn*
L8Y	n.a.	n.a.	*Y1*
A8	-	*jl*	-
B3	*f*	*af*	*af*
B10	*b*	*bc*	*bc*
B15	*f*	*af*	-
B30	*o*	*o, q*	*go*
C19	*de*	*de*	*de / y2*
C100	*b*	*bc*	*bc*
C117	*b*	*bc*	*bc*
C171	*de*	n.a.	*de / y2*
D24	*j, k, l*	*jl*	*jl*
D106	*h*	*h*	-
D107	*a*	*af*	*af*
D109	*o*	*o, q*	*go*
D112	*a*	*af*	*af*
D138	*de*	*de*	*de / y2*

N.B. The chromosome names followed the traditional chromosome nomenclature in the *S. araneus* group (Searle *et al.* 1991). No amplification of correct size on sorted chromosome (-) nor- on both sorted chromosome and positive control (n.a.).

Table 4. Ambiguous assignation of 21 microsatellite markers in the three karyotypic taxa *Sorex granarius*, *S. araneus* Cordon and *S. araneus* Novosibirsk.

Locus	*Sorex granarius*	*S. araneus* Cordon	*S. araneus* Novosibirsk
L14	*f / m /* ***o*** */ de*	*g / i /* ***o****, q*	*af /* ***go***
L33	*a /* ***m****, o*	*i/ o, q/* ***m***	***mp***
L45	***b*** */ i /de*	***b****c/ af / jl / de / ut*	***b****c / af*
L67	*b / i / q, r*	*af / g*	*bc / qr*
L92	***b*** */de*	***b****c / de*	*af /* ***b****c*
L97	-	***bc*** */ i*	***bc***
A25	*b/****f***	***af*** */ r / mp*	***af*** */ mp*
B5	*m/ i*	*af / jl / g / p*	*af / mp/ y2*
B7	n.a.	***g*** */ k*	*af /* ***g****o*
B12	n.a.	*r / m*	n.a.
C5	*a / i / de*	-	-
C25	***f*** */* ***i***	*af /* ***i***	***af*** */ mp /* ***i****k / go / y1*
C119	*b/* ***f***	n.a.	***af*** */ ik/ go*
C122	***a****/* ***g***	***af*** */* ***g***	***af*** */* ***g****o*
C151	***a****/****f***	n.a.	***af*** */ mp*
C240	*i*	*jl / p*	*af*
D11	*b*	*af*	*af*
D23	***f*** */ j,k,l*	-	***af*** */ go*
D29	*b / de / m, o*	*ut*	*af / mp*
D103	*i*	*m / ut*	*mp / af*
D110	***b*** */ de/* ***c***	***bc*** */ jl / ut*	***bc*** */ af / qr / de / y2*

N.B. The chromosome names followed the traditional chromosome nomenclature in the *S. araneus* group (Searle *et al.* 1991). No amplification of correct size on sorted chromosome (-) nor- on both sorted chromosome and positive control (n.a.). In bold are the most probable localizations of the microsatellite markers at the chromosome or chromosome arm level (chromosome sorts giving positive amplification in each taxa).

DISCUSSION

Twenty-six microsatellite loci have been assigned to a specific chromosome in *Sorex granarius*, *S. araneus* Cordon and *S. araneus* Novosibirsk. In addition, 25 of these were localised on the same chromosome arm in the three taxa: they are therefore considered as unambiguously mapped (Table 3). At least nine of the 18 autosomal chromosome arms found in the *S. araneus* group were hybridized by these markers: *a* (3 loci), *b* (5), *c* (1), *f* (3), *g* (1), h (1), *j* or *l* (2), *n* (1) and *o* (2). We were not able to discriminate between the chromosome arms of the metacentric chromosome *jl* since these arms were part of the same chromosome sort in *S. granarius* (Table 2). The remaining six loci mapped to the sex chromosomes. The chromosome X (*de*) is the product of a Robertsonian fusion between the "original" mammalian X (most of arm *e*) and an autosome (all of arm *d* and part of arm *e*; Pack *et al.* 1993). Chromosome Y2 in males corresponds to this autosome and is therefore homologous

to chromosome arm *d*. Four loci (L57, C19, C171 and D138) were assigned to chromosomes X and Y2 in *S. araneus* Novosibirsk and probably map to the chromosome arm *d*. One locus (L13) was assigned only to the X chromosome in the same species suggesting localization on chromosome arm *e*, the "true" X chromosome. Finally, as expected the male specific locus L8Y (Balloux *et al.* 2000) mapped to chromosome Y1, the true Y.

The distribution of our mapped loci was not significantly different from the expected null distribution according to chromosome arm size (χ^2 = 0.584, d.f. = 17; P = NS). Microsatellite loci thus did not appear to cluster on any particular chromosome.

Only one locus (Table 4) assigned to a single chromosome showed discordance among the taxa (locus D11 mapped on chromosome *b* in *S. granarius* and *af* in *S. araneus* Cordon and Novosibirsk). Data about genetic exchange between chromosome arms in the *S. araneus* group are scarce. Using high resolution chromosome analysis Volobouev & Catzeflis (1989) detected only several centromeric shifts between *S. granarius* and *S. araneus*. Our study corroborates these findings and suggests that despite the presence of numerous Robertsonian rearrangements, the organization of the genome in the *S. araneus* group is well conserved. Consequently, this high level of conservation suggests that our markers can be used to study and compare the genetic structure within and among the different species and chromosome races of the *S. araneus* group at a chromosome specific level.

Multiple assignations

Due to PCR amplification in more than one chromosome specific sort, 20 loci could not be unambiguously assigned to a specific chromosome (Table 4). Interestingly, these loci were in general ambiguous in all three taxa. Several non exclusive reasons may explain this pattern: i.e. lack of specificity of markers, contamination between sorts or genomic rearrangements.

First, the competition for binding sites on a single chromosome is probably weaker than over the whole genome. Therefore, the specificity of a primer pair may decrease when tested on isolated chromosomes and allow unspecific amplification. Nevertheless, most of our attempts to increase PCR specificity and design new primer pairs did not improve our capacity to assign loci to a single chromosome sort.

Second, the purity by which individual chromosomes can be sorted can approach but never reach 100% (Doležel *et al.* 2004, Ibrahim & van den Engh 2004). Therefore, chromosomes of similar sizes and GC/AT ratios could contaminate the desired chromosome sort. However, some markers were mapped to chromosomes that are quite distant on the flow karyotype. In such cases, flow rate might have been too fast, allowing more than one chromosome to be

sorted together. Alternatively, the fragmentation and/or clumping of chromosomes might have altered their size and caused them to be sorted in the wrong sort. No apparent contaminations (except sorting of multiple chromosomes in several sorts) were detected by our painting experiments. But, these experiments may fail to detect low level DNA contamination, which is not the case in PCR.

Finally, sequence duplication is common in the mammalian genome (Samonte & Eichler 2002, Thomas *et al.* 2004). Although the evolution of duplicated non-coding regions (including microsatellites) is still poorly investigated, it is likely that some of our loci belong to such regions. When a locus with high mutation rate (such as microsatellites) is duplicated, it is expected that some individuals bear more than two alleles. The locus C122 probably illustrates this situation since several tri- or tetraploid individuals were detected for this locus. Additionally, this locus was assigned (after numerous attempts to increase PCR specificity) to the same two chromosome arms (*a* and *g*) in the three taxa. Contamination between these chromosomes in the three taxa is highly unlikely since these chromosome arms are combined in acrocentric or metacentric chromosomes of different size (i.e. *a* / *g* in *S. granarius*, *af* / *g* in *S. araneus* Cordon and *af* / *go* in *S. araneus* Novosibirsk).

To conclude, at least one chromosome showing positive amplification was common to the three taxa (in bold in Table 4) in 13/20 of the multiply assigned loci. These chromosomes therefore represent the most probable localization for these loci. Further studies are necessary to confirm these possible localizations.

Comparison with previous studies

As previously mentioned, four microsatellite loci have already been mapped by Zhdanova *et al.* (2003) using somatic cell hybrid panels of *Sorex araneus* Novosibirsk. For two loci, the situation was consistent between the two studies: locus L16 was unambiguously mapped to chromosome *af* and the most likely localization for locus L14 is on chromosome arm *o*. The two other loci (L92 and L67) however give conflicting results. These loci were respectively mapped to chromosome *af* and *jl* by Zhdanova *et al.* (2003) whilst both loci amplified in more than one chromosome sort in our study (Table 4). Our results for locus L92 point toward chromosome arm *b* although an amplification product was also detected on chromosome *af* in *S. araneus* Novosibirsk. As for locus L67, the situation is even more complex: our results do not allow for any conclusions but no amplification on chromosome *jl* in any of the three taxa was detected. A possible explanation for these discrepancies could be the high rate of hidden chromosomal rearrangements noticed in the shrew somatic cell hybrid panels (Zhdanova *et al.*

2003). However, as already mentioned, our method suffers from several uncertainties that cannot be ruled out for these loci.

Conclusions

PCR-based localization of genetic markers on flow sorted chromosomes has been used in a variety of groups (e.g. Sargan *et al.* 2000, Doležel *et al.* 2004). Although this method suffers from several limitations (e.g. genetic markers are assigned to chromosomes but not positioned on them), we have been able to map 25 microsatellite markers on three karyotypically different taxa of the *Sorex araneus* group. Also, the use of two particularly acrocentric taxa allowed for most localizations at the chromosome arm level. This last point should not be underestimated since this group shows an extraordinarily large variety of Robertsonian rearrangements. Therefore we provide a microsatellite markers map that includes the Y chromosome (Y1), the two arms *d* and *e* of the X chromosome and at least 9 of the 18 autosomal chromosome arms observed in this group.

The role of chromosomal rearrangements in the speciation process is still subject to controversy (for recent review, see Coyne & Orr 2004). Recently, several studies used mapped genetic markers to show that some genomic regions experience stronger barriers to gene flow than others and that these regions are linked to rearranged chromosomes (e.g. Rieseberg *et al.* 1999, Panithanarak *et al.* 2004). As previously mentioned, the outstanding karyotypic variation of the *S. araneus* group makes it an interesting model for studying the role of chromosomal rearrangement in the speciation process. The high level of conservation in the localization of markers observed among the studied taxa suggests their potential utility to compare genetic structure among taxa of the *S. araneus* group. Therefore, we propose these markers could be used to identify the role of chromosomal rearrangements in the genetic diversification and speciation process of this group.

ACKNOWLEDGEMENTS

We are particularly grateful to Aline Dépraz for helpful comments on the manuscript. This work was supported by the Swiss NSF (grants no: 3100-05943), a Société Académique Vaudoise (Switzerland) grant to P.B. and a Fondation Agassiz (Switzerland) grant to P.B.

CHAPTER 4

Chromosome rearrangements and genetic differentiation at different evolutionary levels of the Sorex araneus group

Patrick Basset, Glenn Yannic and Jacques Hausser

(In preparation)

ABSTRACT

Robertsonian (Rb) fusions received large theoretical support for their role in animal speciation. But empirical evidences are often lacking because of the difficulty to discriminate between incompatibilities generated by genes only or also by chromosomes. Here, we address the role of Rb rearrangements on the genetic diversification of the karyotypically and genetically diversified shrews of the *Sorex araneus* group, using microsatellite markers mapped at the chromosome arm level. We compared genetic structure and genetic distance between rearranged and common chromosomes in pairwise comparisons of five karyotypic taxa of the *S. araneus* group with the prediction that rearranged chromosomes show larger levels of genetic differentiation. Inter-specific structure and distance were larger across rearranged chromosomes for most of the comparisons although these differences were in general not significant. This last result could be explained by the large variance observed among microsatellite estimates. Considering all possible comparisons, we found a significantly larger differentiation of rearranged chromosomes supporting the role of chromosomal rearrangements in the general genetic diversification of this group. In addition, the large variance observed among the pairs of taxa analysed supports the role of both the hybrid karyotypic complexity and the level of evolutionary divergence.

INTRODUCTION

Recent theoretical studies have confirmed the plausibility of speciation driven by chromosomal rearrangements (Rieseberg 2001, Noor *et al.* 2001a, Navarro & Barton 2003a). However, since such theories are often difficult to test empirically, the real impact of rearrangements on speciation remains much debated (e.g. Coyne & Orr 2004). The main problem lies in the almost systematic correlation between chromosomal and genetic differentiation. It is thus difficult to assess the exact sequence of events leading to reproductive isolation and particularly to discriminate between incompatibilities generated by genes only, or also by chromosomes (for example, see the highly discussed case of human and chimpanzee in Navarro & Barton 2003b, Lu *et al.* 2003, Hey 2003 or Zhang *et al.* 2004). To help teasing apart the respective roles of genes and chromosomes, karyotypic variable taxa at different stages of evolutionary divergence are needed. In such a context, one can assume that if chromosomal rearrangements affect reproductive isolation, the genetic divergence between two taxa should be greater in the regions of their genome located on chromosomes differently rearranged (Hey 2003, Ayala & Coluzzi 2005).

Robertsonian (Rb) fusions (i.e. the fusion of primitive acrocentric chromosomes into metacentrics) probably received the largest support for their role in animal, and especially in mammal, speciation (Baker & Bickham 1986, King 1993, Searle 1993, Coyne & Orr 2004) and this for several reasons. First of all, Rb fusions are repeatedly polymorphic within mammalian populations (Nachman & Searle 1995) and appear to be the most common type of rearrangements fixed between mammalian species (Baker & Bickham 1986). In addition, hybrids that are heterozygotes for many Rb fusions might suffer from low fitness (Searle 1993, Hauffe & Searle 1998, Castiglia & Capanna 2000, Piálek *et al.* 2001). Finally, recombination is expected to be suppressed or reduced in heterozygotes Rb chromosomes (Davisson & Akeson 1993, Haigis & Dove 2003), which is a prerequisite for most recent chromosomal speciation models (for a review, see Butlin 2005).

In such a context, the shrew species within the *Sorex araneus* group (e.g. *S. antinorii*, *S. araneus*, *S. coronatus* or *S. granarius*) offer an extraordinary opportunity to study the role of Rb rearrangements on reproductive isolation and genetic structure. Indeed, this group of morphologically very similar species received great deals of attention due to its karyotypic variability mainly attributed to Rb fusions (Volobouev 1989). In *S. araneus*, the type species of this group, Rb polymorphism is so prevalent that more than 60 chromosome races have been described (Wójcik *et al.* 2003). With the exception of the sexual chromosomes and three

pairs of metacentric autosomes (*af*, *bc* and *tu*; nomenclature of chromosome arms according to Searle *et al.* 1991) that are invariants, all other autosomal arms (*g* – *r*) may occur as acrocentrics and/or combined as different metacentrics. As a result, this species covers, together with the remaining species of the *S. araneus* group, wide ranges of chromosomal rearrangements, from similar karyotypes to extremely divergent ones. Therefore, when comparing two taxa of this group (i.e. chromosome races or species), it is possible to identify some chromosome arms that are identically arranged into acrocentric or metacentric chromosomes and other that are rearranged in different acrocentrics and/or metacentrics. Throughout this study, we will consider these two classes of chromosomes as the "common" and "rearranged" chromosomes.

In addition to providing a remarkable karyotypic diversity, the *S. araneus* group provides all the main levels of evolutionary divergence expected during chromosomal speciation, that is: (1) chromosome races within a species, (2) restricted gene flow between karyotypic species and (3) complete reproductive isolation. The first situation is exemplified by the chromosome races of *S. araneus*: only low levels of genetic divergence have generally been detected among these races (Ratkiewicz *et al.* 2002, Andersson *et al.* 2004, Basset *et al.* in press (Chapter 1), and see Wójcik *et al.* 2002 for recent review). The second situation is illustrated by *S. araneus* and *S. antinorii*. The latter was until recently considered as a chromosome race of *S. araneus*, but given its karyotypical, morphological, biochemical and genetic distinctness, it was promoted to the species rank (Brünner *et al.* 2002a). Actually, *S. antinorii* and *S. araneus* meet in at least two hybrid zones in the Western Alps (Brünner & Hausser 1996, Brünner *et al.* 2002b) and, in spite of an apparently strongly reduced gene flow, introgression between these species is still detected (Lugon-Moulin *et al.* 1999a, Brünner *et al.* 2002b, Basset *et al.* in press (Chapter 1)). Finally, the third and last level is represented by the remaining eight species of the *Sorex araneus* group (*S. arcticus*, *S. asper*, *S. caucasicus*, *S. coronatus*, *S. daphaenodon*, *S. granarius*, *S. maritimensis* and *S. tundrensis*), which appear to be reproductively isolated from *S. araneus*. For example, several contact zones between *S. coronatus* and *S. araneus* have been detected in the Western Alps, but extensive karyological studies over several decades did not provide any evidence for ongoing hybridization between these two species (Neet 1989, Neet & Hausser 1990).

The goal of the present study is to use microsatellite markers mapped at the chromosome arm level (Basset *et al.* in prep, Chapter 3) to address the role of chromosomal rearrangements on the genetic differentiation of the karyotypically and genetically diversified *S. araneus* group.

Our prediction is that if karyotypic differences influence the genetic diversification of this group, genetic differentiation will be higher for rearranged than for common chromosomes.

MATERIAL AND METHODS

Taxa analysed

Five karyotypic taxa of the *Sorex araneus* group (*S. coronatus*, *S. antinorii* and *S. araneus* chromosome race Vaud, Cordon and Białowieza) were analysed during this study. *Sorex coronatus* is distributed from northern Spain to eastern Germany (generally below 1,000m above sea level) and would have diverged during the last Pleistocene glaciations in refugia in south-western France or Spain (Hausser *et al.* 1978). Karyotypically, this species (karyotype: XX/XY1Y2, *af*, *b**, *ci*, *gr*, *h**, *jn*, *kq*, *lo*, *mp*, *tu*; Hausser *et al.* 1991) is believed to be the sister group of both *S. araneus* and *S. antinorii* (Taberlet *et al.* 1994). Note that Rb fusions are not the only kind of rearrangements differentiating this species from *S. araneus* and *S. antinorii* (e.g. * centromeric shifts; Volobouev & Catzeflis 1989). Although these three species are parapatric, almost no admixture has been detected (Basset *et al.* in press, Chapter 1). The second species, *S. antinorii,* survived in the Apennine Peninsula during the last glaciations and currently occurs in Italy, south-eastern France, and southern Switzerland (Brünner *et al.* 2002a). Karyotypically, it is the sister group of *S. araneus* (Searle & Wójcik 1998) and it is characterized as: XX/XY1Y2, *af*, *bc*, *gi*, *hj*, *kn*, *l/o*, *m*, *p*, *r*, *tu* (Brünner *et al.* 2002a). As previously mentioned, this species is known to naturally hybridize with *S. araneus* (Brünner & Hausser 1996, Brünner *et al.* 2002b) with detectable levels of introgression at large geographical scales (Basset *et al.* in press, Chapter 1). Finally, the remaining three taxa analysed consist of chromosome races belonging to *S. araneus*. The Vaud race (XX/XY1Y2, *af*, *bc*, *gm*, *hi*, *j/l*, *kr*, *n/o*, *p*, *q*, *tu*; Wójcik *et al.* 2003) occurs in south-western Switzerland and belongs to the western karyotypic group characterized by the metacentrics *gm* and *hi*. The Cordon race, in the French Alps, is the most acrocentric race of *S. araneus* (XX/XY1Y2, *af*, *bc*, *g*, *h*, *i*, *j/l*, *k*, *m*, *n*, *o*, *p*, *q*, *r*, *tu*; Wójcik *et al.* 2003) since all of its variable chromosomes are in an acrocentric state except the polymorphic *j/l*. Although strictly speaking it does not belong to the western karyotypic group (it lacks the metacentrics *gm* and *hi*), this race is genetically very close to the Vaud race (Taberlet *et al.* 1994). Finally, the race Białowieza from north-eastern Poland is geographically far apart from the other taxa analysed in this

study. This race (XX/XY1Y2, *af*, *bc*, *g/r*, *h/n*, *ik*, *j/l*, *m/p*, *q*, *tu*; Wójcik *et al.* 2003) belongs to the eastern karyotypic group characterized by the metacentrics *gr*.

For each of the 10 comparisons, chromosomes common to both taxa and chromosomes differentially rearranged in each taxon are listed in Figure 1. In the same figure, each comparison is ordered according to its level of genetic divergence.

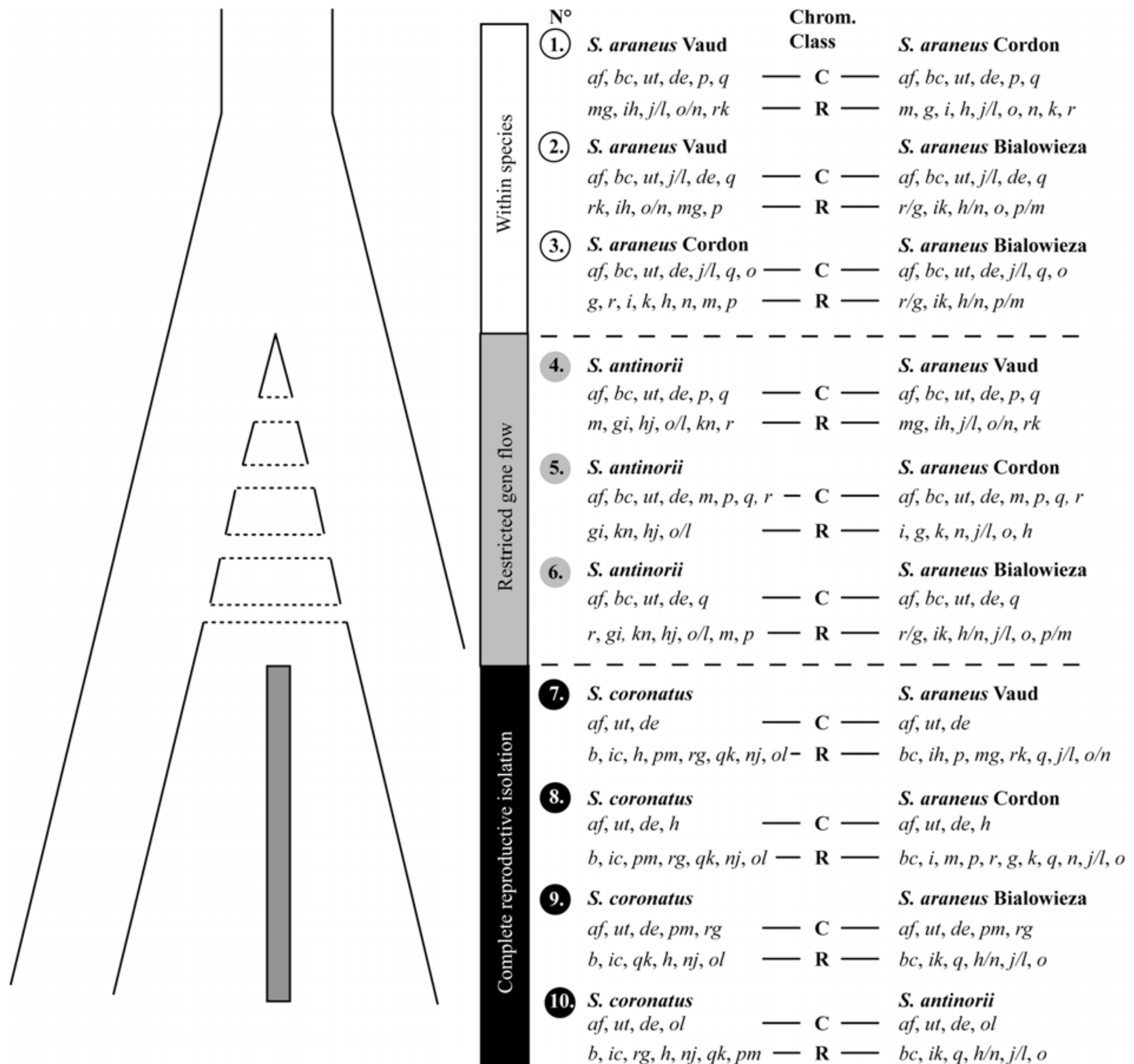


Figure 1 Details of the 10 pairwise comparisons among the five karyotypic taxa of the *S. araneus* group. For each comparison, chromosomes of both taxa are sorted in common (C) or rearranged (R) groups. Each pair of taxa is then classified into three evolutionary levels (left part of the figure) according to expected reproductive isolation: i.e. within species, restricted gene flow and complete reproductive isolation.

Sampling

A total of 19 *Sorex coronatus*, 55 *S. antinorii* and 110 *S. araneus*, subdivided into three chromosome races: 38 *S. araneus* Vaud, 35 *S. a.* Cordon and 37 *S. a.* Białowieza were analyzed during this study (Table 1). Since no more than 3 individuals were sampled in each locality for *S. coronatus*, all were grouped into one large heterogeneous population (represented by a dashed symbol in Fig. 2).

Table 1. Number of individuals analysed in each taxon. For each taxon, individuals have been grouped into one to four populations according sampling localities. No population name was assigned to *Sorex coronatus* since the number of individuals sampled in each locality of this species is always low (≤ 3), therefore only one large population of this species is considered. Figure references correspond to the localities in Fig. 2.

Fig. ref.	Species	Chromosome race	Population	N° of ind.
1	*S. coronatus*		-	19
2	*S. antinorii*		Trient (CH)	22
3	*S. antinorii*		Herrens (CH)	8
4	*S. antinorii*		Chastlerra (CH)	12
5	*S. antinorii*		Tännerweide (CH)	13
6	*S. araneus*	Vaud	Bassins (CH)	10
7	*S. araneus*	Vaud	Jorat (CH)	21
8	*S. araneus*	Vaud	Champittet (CH)	7
9	*S. araneus*	Cordon	La Clusaz (F)	6
10	*S. araneus*	Cordon	Cordon (F)	29
11	*S. araneus*	Białowieza	Jurowce (PL)	20
12	*S. araneus*	Białowieza	Gugny (PL)	17
	Total			184

DNA extraction and microsatellite analysis

Tissue samples (liver, heart, spleen or phalanxes) were stored at -70°C or in alcohol (100%) before total genomic DNA was extracted using the DNeasy Tissue Kit (Qiagen).

Sixteen microsatellite loci were chosen among those unambiguously mapped at the chromosome level in Basset *et al.* (in prep, Chapter 3), the letter in parentheses indicating their chromosome arm localization: L16 (*a*), D107 (*a*), L69 (*f*), B3 (*f*), L68 (*b*), C117 (*b*), L9 (*c*), C171 (*de*), L13 (*de*), L57 (*de*), L62 (*g*), D106 (*h*), D24 (*jl*), L99 (*n*), B30 (*o*), D109 (*o*). Each locus is thus part of either the common or rearranged group depending on its chromosome localization and the pair of taxa under study (Fig. 1).

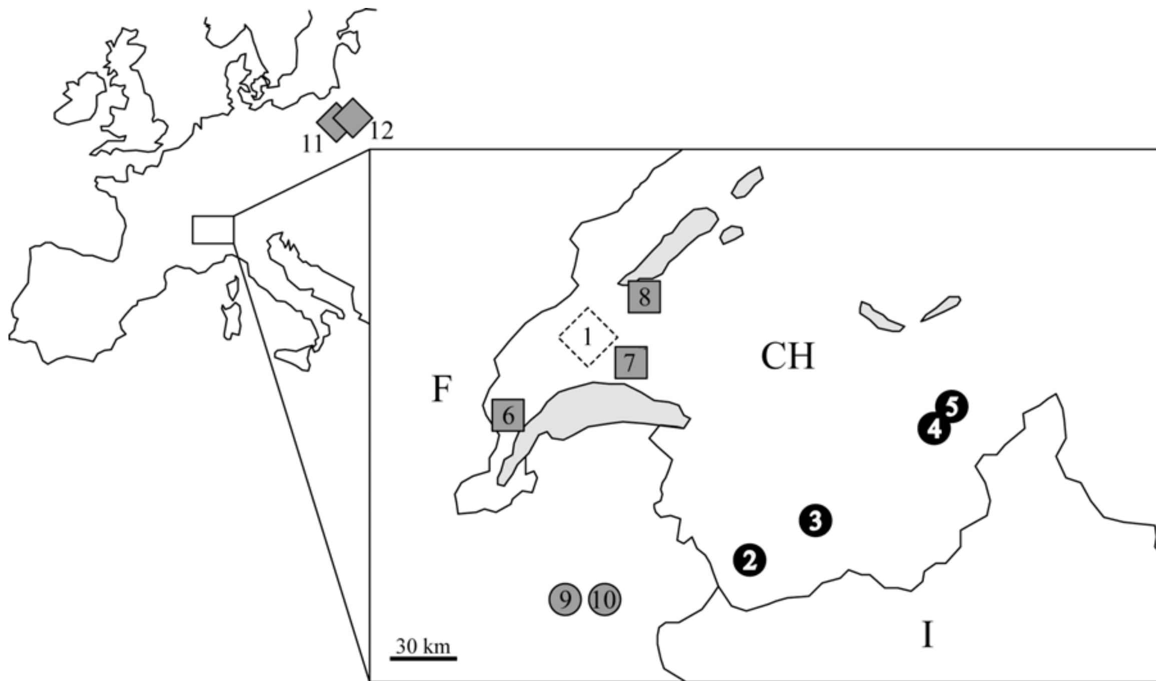


Figure 2 Sampling localities of the five karyotypic taxa analysed during this study. Numbers correspond to the localities described in Table 1.

PCR conditions are given elsewhere (Wyttenbach *et al.* 1997, Balloux *et al.* 1998, Lugon-Moulin *et al.* 2000, Basset *et al.* in press (Chapter 1), and Basset *et al.* in prep (Chapter 3)) but for all primers, PCR amplifications were performed in a 20 µl total volume. Cycling was carried out in a PE9700 (Applied Biosystems) using the following profile: 95°C for 5 min, 35 cycles of 30 s at 94°C, 30 s at annealing temperature (Basset *et al.* in prep, Chapter 3), 30 s at 72°C; and a final extension at 72°C for 4 min. One primer of each pair was labelled with a fluorescent dye (HEX, FAM or NED) on the 5'end, which allowed analyses on an ABI 377XL sequencer (Applied Biosystems). Data collection, sizing of the bands and analyses were done using the GENESCAN software (Applied Biosystems).

Divergence measures

Since chromosome changes are expected to influence genetic structure between but not within karyotypically identical taxa, we estimated the structure first within the "common" and then within the "rearranged" groups of chromosomes with hierarchical F-statistics (Weir 1996) using the software package ARLEQUIN version 2.00 (Schneider *et al.* 2000; http://anthropologie.unige.ch/arlequin). For each pair of taxa analysed, two levels of structure were considered: the intra-taxon structure F_{SR} (between populations within taxon) and the inter-taxa structure F_{RT} (between populations of different taxa).

Differentiation measures based on F-statistics (for a review, see Excoffier 2001) are closely tied to the infinite allele model of mutation (IAM), where each mutation can result in an allele of any size (Kimura & Ohta 1978). In addition, F-statistics tend to reach a plateau, not reflecting any more the increasing divergence with increasing time (Kalinowski 2002). Therefore, we additionally estimated the genetic distance $(\delta\mu)^2$ of Goldstein *et al.* (1995) that has been designed to avoid such plateau and to increase linearly with time under the stepwise mutation model (SMM) (Goldstein & Pollock 1997). It should be noted that this distance account for differences in allele sizes. Again, for each comparison, we estimated genetic distance across common and rearranged chromosomes at the intra- and inter-taxa levels.

Differentiation between common and rearranged chromosomes

For each pair of taxa and for the two divergence measures (i.e. F-statistics and $(\delta\mu)^2$), the difference between the two classes of chromosomes (common vs. rearranged) were tested by comparing the observed value to a null distribution of no difference between groups based on 10'000 permutations of microsatellites between groups. In addition, the overall difference between common and rearranged chromosomes (across the 10 comparisons) was tested against the expected "rearranged/common" ratio of 1 if chromosomes have no effect with a Wilcoxon's signed rank test (Crawley 2002).

RESULTS

Genetic structure

Results of genetic structure estimated by hierarchical F-statistics (F_{SR} and F_{RT}) across common and rearranged chromosomes for the 10 pairs of taxa analysed in our study are indicated in Figure 3 and individual microsatellite estimates in each comparison are given in annexe 1.

Intra-taxon (F_{SR}). None of the pairs of taxa tested show significant differences between the two groups of chromosomes ($P > 0.100$, permutation tests; Fig. 5A). The results of the three evolutionary levels do not differ significantly from each other ($P > 0.10$, Wilcoxon's test). Pooling all pairs of taxa, the ratio between the genetic structure (F_{SR}) across rearranged and common chromosomes is not significantly different than one ($P = 0.323$, Wilcoxon's signed rank test; Fig. 5A).

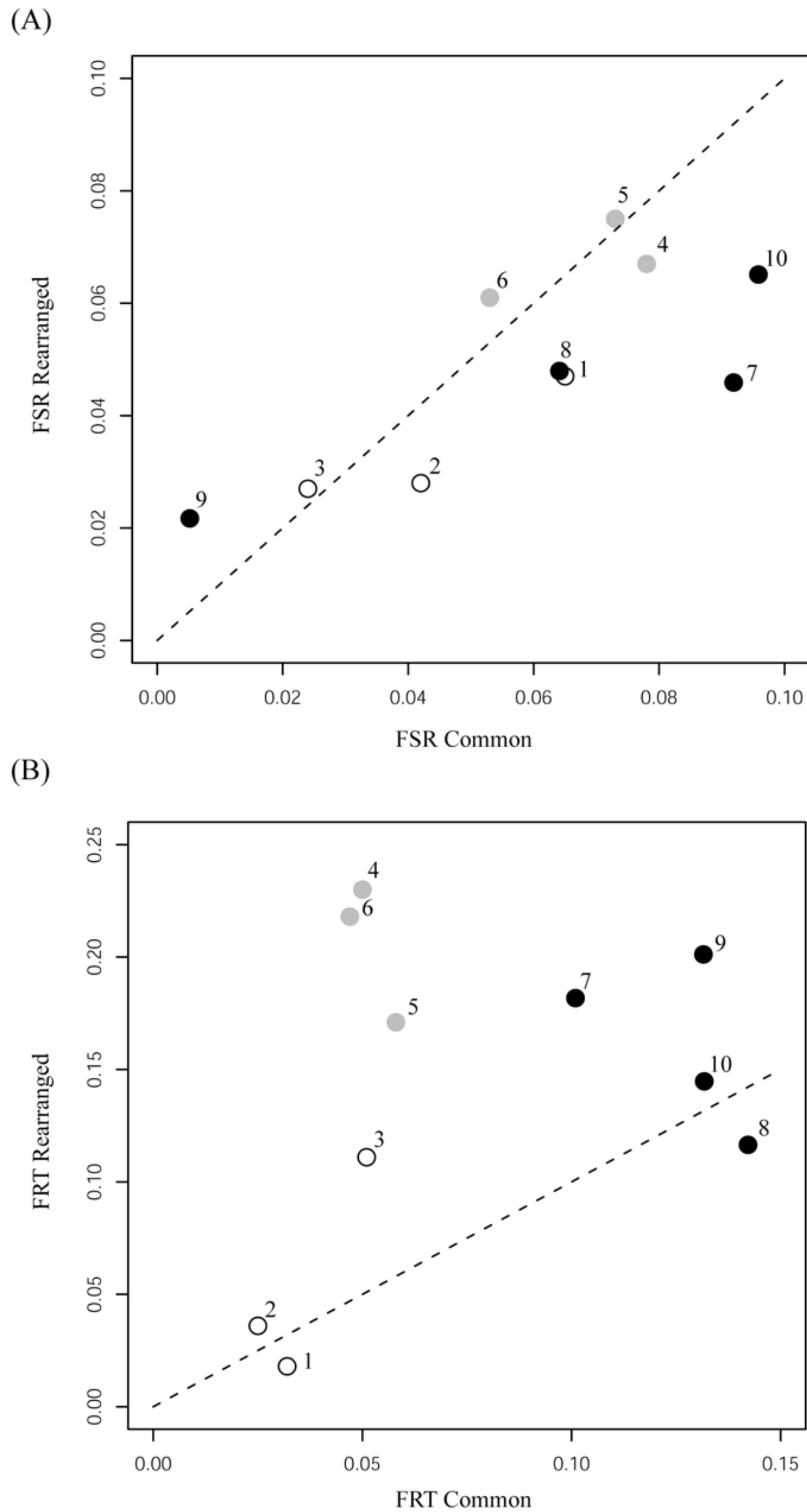


Figure 3 Values of hierarchical F-statistics across common and rearranged loci for each pair of taxa analysed, (A) intra-taxon and (B) inter-taxa comparisons. Open circles: comparisons within *S. araneus*, grey: comparisons between *S. antinorii* and *S. araneus* and black: comparisons involving *S. coronatus*. Numbers correspond to the comparisons described in Figure 1 and the dashed line corresponds to the line of slope 1.

Inter-taxa (F_{RT}). Although most of the comparisons show larger estimates across rearranged than common chromosomes, only two pairs of taxa show a marginally significantly larger genetic structure on the rearranged chromosomes (*S. antinorii* – *S. a.* Vaud: P = 0.060; *S. antinorii* – *S. a.* Białowieza: P = 0.074; permutation tests) and no significant difference is detected among the remaining eight pairs analysed (Fig. 5C). Interestingly, the comparisons between *S. antinorii* and *S. araneus* show significantly larger ratio between rearranged and common chromosomes than the two other evolutionary levels (P = 0.016; Wilcoxon's test; Fig. 5C). Finally, pooling all pairs of taxa, the ratio between the genetic structure (F_{RT}) measured across rearranged and common chromosomes is significantly larger than 1 (P = 0.018, Wilcoxon's Signed rank test; Fig. 5C).

Genetic distance

Results of the genetic distance $(\delta\mu)^2$ estimated across common and rearranged chromosomes at the intra-taxon and the inter-taxa levels for the 10 comparisons analysed in our study are indicated in Figure 4 and individual microsatellite estimates in each comparisons are indicated in annexe 1.

Intra-taxon ($(\delta\mu)^2$). None of comparison show a significant difference between genetic distance across common and rearranged chromosomes, but for one pair rearranged estimates are marginally significantly higher than common estimates (*S. a.* Vaud – *S. a.* Cordon; P = 0.056, permutation test). The results of the three evolutionary levels do not differ significantly from each other (P > 0.10, Wilcoxon's test) and pooling all pairs of taxa, we did not detect a ratio between rearranged and common chromosomes that is significantly different than one (P = 0.625, Wilcoxon's signed rank test; Fig. 5B).

Inter-taxa ($(\delta\mu)^2$). Although, most of the comparisons show larger estimates across rearranged than common chromosomes, only one pair show a marginally significantly larger genetic distance across the rearranged chromosomes (*S. antinorii* – *S. a.* Białowieza: P = 0.072; permutation test) and no significant difference is detected among the remaining nine pairs analysed (Fig. 5D). The comparisons between *S. antinorii* and *S. araneus* show marginally significantly larger ratio between rearranged and common chromosomes than the two other evolutionary levels (P = 0.056; Wilcoxon's test; Fig. 5D). Again, pooling all pairs of taxa, the ratio between genetic distance measured across rearranged and common chromosomes is significantly larger than one (P = 0.048, Wilcoxon's signed rank test; Fig. 5D).

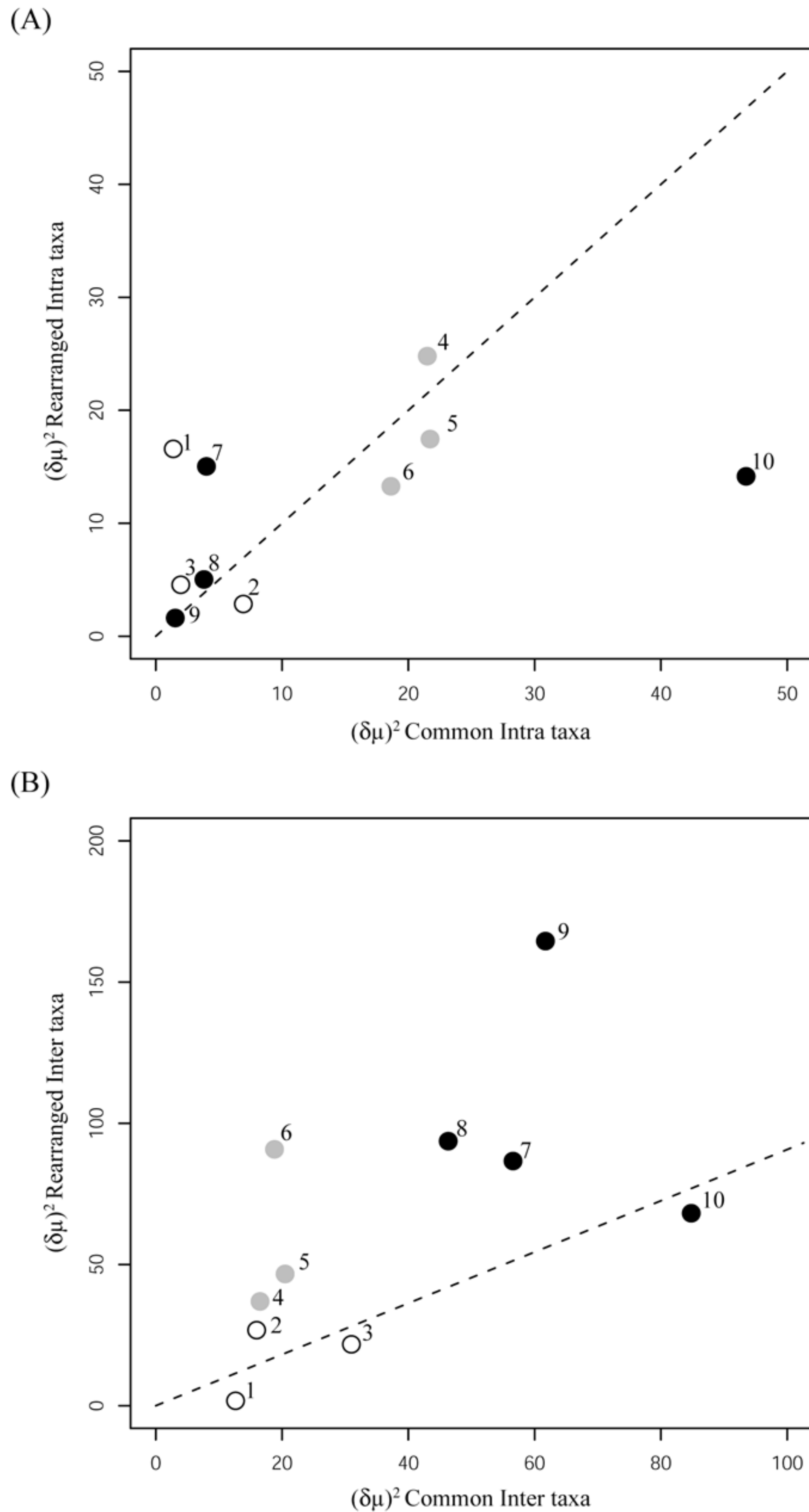


Figure 4 Values of genetic distance $(\delta\mu)^2$ across common and rearranged loci for each pair of taxa analysed, (A) intra-taxon and (B) inter-taxa comparisons. Open circles: comparisons within *S. araneus*, grey: comparisons between *S. antinorii* and *S. araneus* and black: comparisons involving *S. coronatus*. Numbers correspond to the comparisons described in Figure 1 and the dashed line corresponds to the line of slope 1.

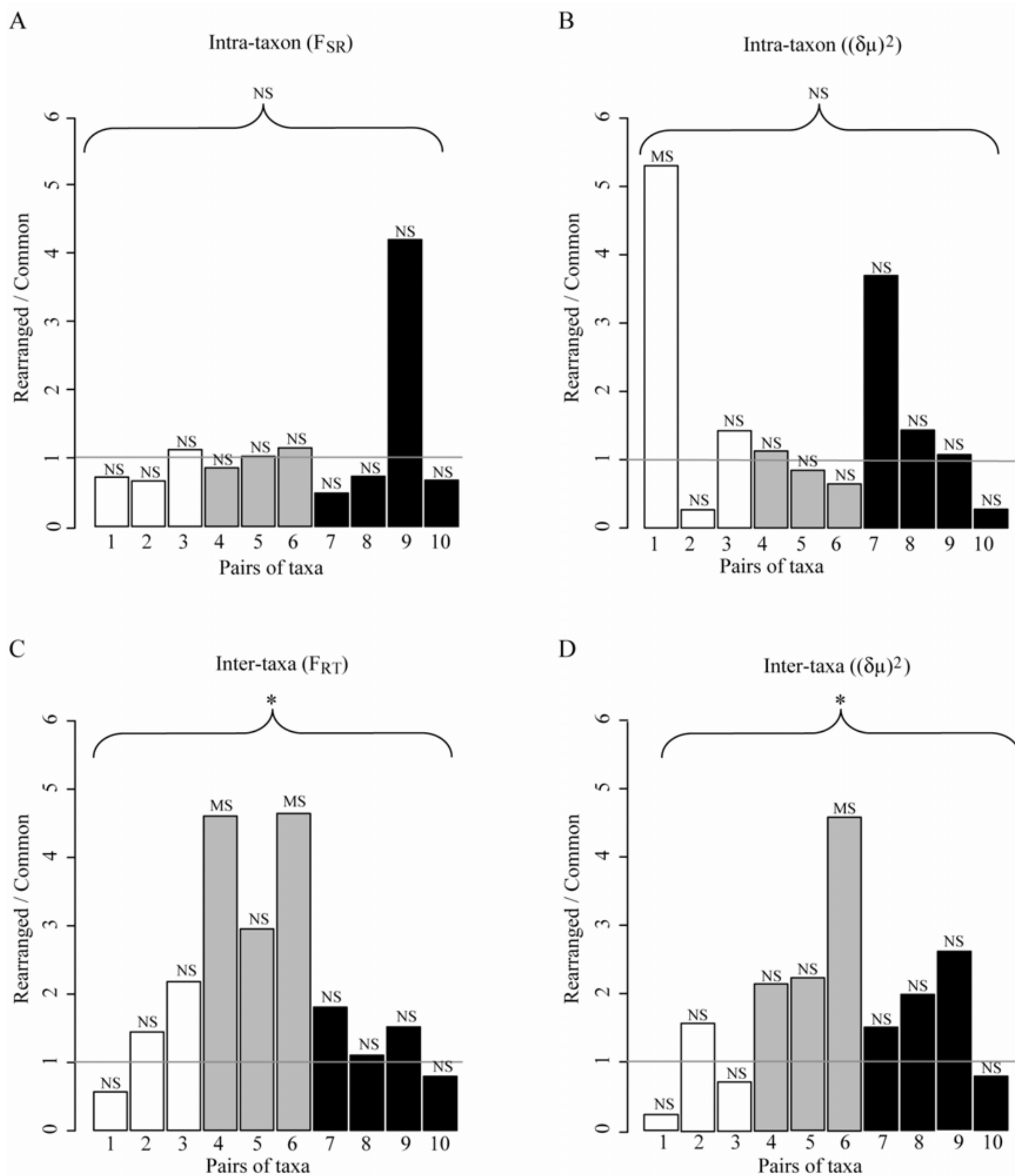


Figure 5 Ratio between genetic differentiation estimates across rearranged and common chromosomes (Rearranged/ Common) based on intra- (A and B) or inter- (C and D) taxa comparisons. Genetic differentiation is measured according to hierarchical F-statistics (A and C) or $(\delta\mu)^2$ genetic distance (B and D). White: comparisons within *S. araneus*, grey: comparisons between *S. antinorii* and *S. araneus* and black: comparisons involving *S. coronatus*. The grey lines indicate the expected ratio (i.e. 1.0) if chromosomal rearrangements do not influence the genetic structure of our samples. * = $P < 0.05$, MS = marginally significant ($0.05 \leq P \leq 0.10$), NS = not significant ($P > 0.10$). Numbers of comparisons (1 – 10) correspond to: 1 *S. a.* Vaud – *S. a.* Cordon, 2 *S. a.* Vaud – *S. a.* Białowieza, 3 *S. a.* Cordon – *S. a.* Białowieza, 4 *S. antinorii* – *S. a.* Vaud, 5 *S. antinorii* – *S. a.* Cordon, 6 *S. antinorii* – *S. a.* Białowieza, 7 *S. coronatus* – *S. a.* Vaud, 8 *S. coronatus* – *S. a.* Cordon, 9 *S. coronatus* – *S. a.* Białowieza, 10 *S. coronatus* – *S. antinorii.*

DISCUSSION

We used microsatellite loci mapped at the chromosome arm level to estimate the importance of chromosomal rearrangements in the genetic diversification of the *Sorex araneus* group. Although rearranged chromosomes are more structured and separated by larger distances than common chromosomes in most of the pairwise inter-taxa comparisons, these differences are never statistically significant (P > 0.05, permutation tests; Fig. 5). Such a result might stem from the large variance observed among microsatellite loci in both groups of chromosomes (Annexe 1), which reduces the power of permutation tests. Using a simple strategy, we tested the role of the number of microsatellite analysed in the power of our permutation tests. We doubled the number of microsatellites analysed (estimating that added microsatellites are similarly informative) and re-tested each comparison with this new dataset. Interestingly, each of the comparison that was marginally significant before this test became highly significant. Moreover, every comparison between *S. araneus* and *S. antinorii* showed significantly larger genetic structure across rearranged chromosomes. Although these results must be interpreted with caution, they suggest that differences between common and rearranged chromosomes actually exist and that a global study of these comparisons is of interest.

Variance among pairs of taxa?

The 10 comparisons analysed in this study are not similarly influenced by chromosomal rearrangements (Fig. 5). For example, the ratio Rearranged/Common is lower than one in the comparison between *S. a.* Vaud and *S. a.* Cordon but close to five in the comparison between *S. antinorii* and *S. a.* Białowieza. At least two non exclusive reasons can be put forward: differences in karyotypes and differences in evolutionary divergence.

To start with, the karyotypic complexity of the hybrids that would be produced by each pair of taxa analysed is not always the same. Searle *et al.* (1990) made an important distinction between simple (i.e. which produce trivalents at meiosis I) and complex (i.e. which produce longer configurations) heterozygotes. Although Rb heterozygotes in the *S. araneus* group do not seem to suffer from infertility as substantially as other mammals (e.g. Searle 1993, Narain Fredga 1997, 1998, Banaszek *et al.* 2000), complex heterozygotes are assumed to be less fertile (e.g. Banaszek *et al.* 2002) and form larger linkage block (e.g. Brünner *et al.* 2002b). The Cordon race, with its almost all acrocentric karyotype, is expected to mostly form simple heterozygote hybrids with any other taxa. As expected, when this chromosome race is compared with *S. antinorii*, the effect of chromosomal rearrangements is the lowest observed among the comparisons between *S. araneus* and *S. antinorii* (Fig. 5C). In contrast, *S. a.*

Białowieza, which has a much more metacentric karyotype, shows differences between common and rearranged chromosomes that are much larger although it is geographically well separated from *S. antinorii* and belongs to the eastern karyotypic group. This surprising observation suggests that the differences between common and rearranged chromosomes have initiated before the separation of the eastern and western karyotypic groups and are therefore important for the evolutionary history of the *S. araneus* group. Further comparisons between representatives of the western and eastern karyotypic groups would be of primary importance to confirm this result.

Second, mutation rates of microsatellite markers are known to be high (Ellegren 2004) and the range of allele size found at one locus is limited (e.g. Garza *et al.* 1995). Microsatellites are thus thought to be subjected to homoplasy (i.e. identity in state although not by descent) and the strength of this factor tends to increase with divergence time (Estoup *et al.* 2002). Using a linear reference calibrated on mtDNA, it is possible to address the issue of possible homoplasy in our microsatellite dataset. Plotting the genetic differentiation measures estimated from our complete microsatellite dataset on a mtDNA distance estimated using the P-distance model (Glenn Yannic, unpublished data), we detect a stronger departure from linearity of inter-taxa genetic structure (F_{RT}; Fig 6A) than of inter-taxa $(\delta\mu)^2$ genetic distance (Fig 6B). Moreover the ratio between rearranged and common chromosomes is significantly larger in the three comparisons involving *S. antinorii* and *S. araneus* than in the remaining seven comparisons if we use F_{RT} ($P = 0.016$; Wilcoxon's test) but only marginally significant if we use $(\delta\mu)^2$ ($P = 0.067$; Wilcoxon's test). These two species are only partially reproductively isolated since they hybridize in nature (e.g. Lugon-Moulin *et al.* 1999a, Brünner *et al.* 2002b, Chapter 2 and 5). Furthermore, large scale introgression can still be detected between these species (Basset *et al.* in press, Chapter 1). As a result, it seems that these two species, only partially reproductively isolated with detectable introgression are placed at an ideal evolutionary level to detect differences between common and rearranged chromosomes with microsatellites. But the situation of the two other evolutionary levels addressed in this study is different. Indeed, the divergence time between *S. coronatus* and *S. araneus* or *S. antinorii* is probably long enough for homoplasy to mask, at least in part, differences between rearranged and common chromosomes. Measures of differentiation linear with time, such as $(\delta\mu)^2$, can in part allow to address this issue. In contrast, the low differences observed between common and rearranged chromosomes among chromosome races of *S. araneus* cannot be explained by homoplasy. However, at this evolutionary scale,

differences between common and rearranged chromosomes may be difficult to detect, and may require larger microsatellite resolution.

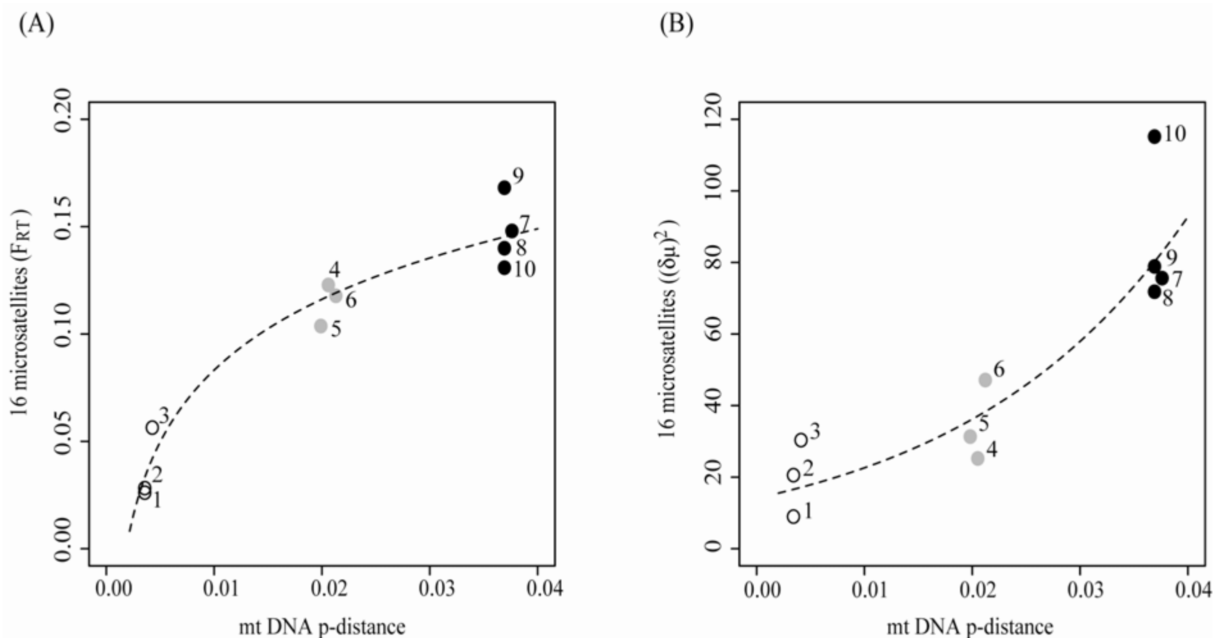


Figure 6 Plot of the inter-taxa genetic structure F_{RT} (A) or genetic distance $(\delta\mu)^2$ (B) based on the global microsatellite dataset on a mtDNA distance estimated using the P-distance model (Glenn Yannic, unpublished data). White circles: comparison within *S. araneus*, grey: comparisons between *S. antinorii* and *S. araneus* and black: comparisons involving *S. coronatus*. Numbers of comparisons (1 – 10) correspond to the pairs described in Figure 1.

Finally, it is worth noting that the variance among pairs of taxa tends to be larger across the rearranged than the common chromosomes (Fig. 3 and 4). This difference may probably be explained by the observation that genetic differentiation of common chromosomes is mostly influenced by the genetic incompatibilities encountered between pairs of species, whereas the genetic differentiation of rearranged chromosomes is expected to be influenced by both genetic and chromosomal incompatibilities.

Overall difference between common and rearranged chromosomes

Considering all the comparisons, we show that rearranged chromosomes are significantly more structured and separated by larger genetic distance than common chromosomes in inter-taxa comparisons (Fig. 5C and D). In contrast, no difference between the same two groups of chromosomes could be detected at the intra-taxa level (Fig. 5A and B). These results provide empirical supports for the role of Robertsonian rearrangements in the general genetic structure of the *S. araneus* group and highlight the potential impact of these rearrangements in the

speciation process of this group. To our knowledge, it constitutes the first evidence of an impact of chromosomal rearrangements on the global genetic structure of an extended group of karyotypic races and species. Lower gene exchanges or higher genetic divergences across genomic regions differing by chromosomal rearrangements have also been detected between species of flies (e.g. *Drosophila*; Noor et al. 2001a,b, Machado *et al.* 2002, Oritz-Barientos *et al.* 2002) or sunflowers (e.g. *Helianthus*; Rieseberg *et al.* 1999), between chromosomal races of house mouse (Panithanarak *et al.* 2004) or between human and chimpanzee (Navarro & Barton 2003b). However, these studies generally deal with only one single comparison of two taxa.

Conclusion

Although none of the individual comparisons were significant, we have shown that chromosomal rearrangements influence the overall genetic differentiation of the *Sorex araneus* group. Moreover, our results highlight that at least two reasons explain the differences between the comparisons under study: i.e. the karyotypic complexity of the hybrids produced and the level of evolutionary divergence. The first point could be addressed by increasing the resolution of microsatellite loci (e.g. increasing the number of mapped loci) in systematic studies of pairs of taxa. In contrast, microsatellites should be combined with other categories of markers (e.g. with lower mutation rates) to address the exact role of evolutionary divergence. Finally, the use of a geographically and karyotypically distant chromosome race suggests that differentiation between common and rearranged chromosome has had a strong impact on the karyotypic history of this group.

ACKNOWLEDGEMENTS

We are grateful to Agata Banaszek and Mirek Ratkiewicz for providing us the samples of *S. araneus* Białowieza from Poland. We would like to thank Rebecca Savioz and Maxence Antonin for their help during lab work. Finally, we appreciate the valuable comments made on the manuscripts by Aline Dépraz. This work was supported by the Swiss NSF (grants n°: 3100-05943), a Société Académique Vaudoise (Switzerland) grant to PB and a Fondation Agassiz (Switzerland) grant to PB.

CHAPTER 5

Restricted gene flow at specific parts of the shrew genome in chromosomal hybrid zones

Patrick Basset, Glenn Yannic, Harald Brünner and Jacques Hausser

(In preparation)

ABSTRACT

The species and races of the shrews of the *Sorex araneus* group exhibit a huge range of chromosomal polymorphisms. European taxa of this group are parapatric and form contact or hybrid zones which span an extraordinary variety of situations ranging from absolute genetic isolation to almost free gene flow. This variety seems to depend for a large part on the chromosome composition of populations, which are primarily differentiated by various Robertsonian fusions of a subset of acrocentric chromosomes. Various data suggest that chromosomal rearrangements play a causative role in the speciation process. In such models, gene flow should be more restricted for markers on chromosomes involved in rearrangements than on chromosomes common in both parent species. In the present study we address the possibility of such differential gene flow in the context of two genetically very similar but karyotypically different hybrid zones between species of the *Sorex araneus* group using microsatellite loci mapped to the chromosome arm level. Inter-specific genetic structure across rearranged chromosomes was in general larger than across common chromosomes. However, the difference between the two classes of chromosomes was only significant in the hybrid zone where the complexity of hybrids is expected to be larger. These differences were not found to distinguish populations within species. Therefore, the rearranged chromosomes appear to affect the reproductive barrier between karyotypic species, although the strength of this effect depends on the complexity of the hybrids produced.

INTRODUCTION

Closely related species or even populations within one species are often characterized by differences in karyotype. This observation has prompted several authors to argue that chromosomal rearrangements, such as Robertsonian fusions and fissions, translocations, and inversions, may play a causative role in speciation (e.g. King 1993, Noor *et al.* 2001a, Rieseberg 2001, Navarro & Barton 2003a). Many models suggest that chromosomal rearrangements facilitate speciation by accelerating genetic differentiation between populations. Traditional models claimed that rearrangements cause meiotic problems for heterozygous individuals and therefore reduce their fertility and reproductive fitness (White 1978, King 1993). In contrast, recent models emphasize a reduction or a suppression of recombination in heterokaryotypes (Rieseberg 2001, Noor *et al.* 2001a, Navarro & Barton 2003a). While there are particularities to each model (for recent reviews, see Spirito 2000, Rieseberg 2001, Coyne & Orr 2004, Ayala & Coluzzi 2005, Butlin 2005) all suggest an interesting possibility: gene exchange frequencies should be differentiated according to the chromosome on which they are located. Therefore, chromosomal rearrangements should induce barriers or filters to gene flow that would be specific to some parts of the genome.

The shrews of the *Sorex araneus* group offer an exceptional opportunity to study the impact of chromosomal rearrangements on gene flow. They display one of the most outstanding chromosomal polymorphism rates found among mammals and offer a complete array of every possible levels of chromosomal and genetic differentiation. The species of the western clade of this group (*S. araneus*, *S. antinorii*, *S. coronatus* and *S. granarius*) present the same chromosome arms, which are labelled from *a* to *u* according to their size (Searle *et al.* 1991). In the type species, *Sorex araneus*, Robertsonian polymorphisms are particularly prevalent. The three pairs of metacentric autosomes *af*, *bc* and *tu* as well as the sexual chromosomes are invariant whereas the primitive acrocentric autosomes *g* to *r* may be distributed into various acrocentric and metacentric combinations. All in all, these polymorphisms allowed describing more than 60 chromosome races (Wójcik *et al*. 2003).

Hybrid zones are often cited as “natural laboratories for evolutionary studies” (Hewitt 1988) and constitute unique opportunities to understand the early processes involved in the establishment of barriers to gene flow and speciation (Barton & Hewitt 1985, Harrisson 1990). These systems are therefore among the most interesting for studying the role of chromosomal rearrangements in speciation. European species and chromosome races of the *Sorex araneus* group are often parapatric and form hybrid zones of various sizes and shapes

(for reviews, see Searle & Wójcik 1998 and Wójcik *et al.* 2002). In such a context, the hybrid zones involving *S. araneus* and *S. antinorii* are of special interest. These species most likely diverged genetically in allopatry during the last Pleistocene glaciations and are likely to have had a long period of independent evolution. The glacial refugia of *S. araneus* were probably situated in south-eastern Europe (Taberlet *et al.* 1994) whereas *S. antinorii* was certainly restricted to refugia situated in the Apennine peninsula (Brünner *et al.* 2002a, b, Lugon-Moulin & Hausser 2002). This latter species recently crossed several lower alpine pass in the Swiss and French Alps (Lugon-Moulin & Hausser 2002) and came into contact with *S. araneus*. In this study, we propose examining the role of karyotypic differences on the genetic structure of two hybrid zones between these species. Both hybrid zones are extremely narrow (less than a kilometre wide) and the observed genetic clines are very steep (Brünner & Hausser 1996, Brünner *et al.* 2002b). Interestingly, the *S. araneus* chromosome races involved in each hybrid zone are not the same. In the first one, *S. antinorii* (characterized by the metacentrics *gi*, *hj*, *kn*, and *lo*; Fig. 1) meets the Cordon race at Les Houches (hereafter LH) in the French Alps (Fig. 2). This chromosome race is one of the most acrocentric and F1 hybrids with *S. antinorii* (although never found, Brünner & Hausser 1996) should encounter only relatively mild problems at meiosis. Most hybrids are expected to carry four trivalents ("simple" heterozygous) which are well tolerated by these shrews (Narain & Fredga 1997, 1998, Searle & Wójcik 1998, Banaszek *et al.* 2002).

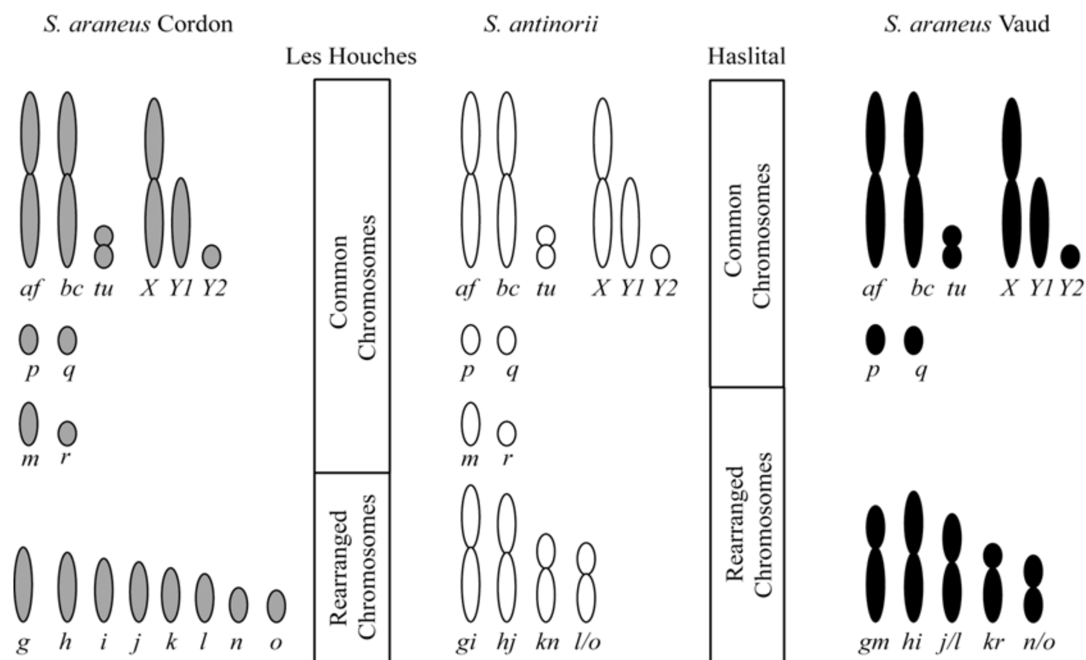


Figure 1 Schematised karyotypes of the three taxa involved in the two hybrid zones (*Sorex antinorii* and *S. araneus* Cordon in LH, and *S. a.* Vaud in HT).

In the second hybrid zone, *S. antinorii* meets the Vaud race of *S. araneus* in the Haslital (hereafter HT) in the Swiss Alps (Fig. 2). This chromosome race is genetically very similar to the Cordon race (Taberlet *et al.* 1994, Basset *et al.* in press (Chapter 1)) but it has a much more metacentric karyotype characterized by *mg*, *hi*, *jl*, *kr* and *no*. Most of the F1 hybrids with *S. antinorii* would present a long chain of eleven elements ("complex" heterozygous, CXI) which allows producing viable gametes only if equilibrated for the parental types and should severely impede recombination and fertility of hybrids (Narain & Fredga 1997, 1998, Searle & Wójcik 1998, Banaszek *et al.* 2002). When comparing the karyotype of the taxa involved in the two hybrid zones, it is possible to define: (i) one group of chromosomes similarly arranged as common acrocentrics or metacentrics, and (ii) one group of chromosomes rearranged in different acrocentrics or metacentrics (Fig. 1). We will use the nomenclature "common" or "rearranged" throughout this paper to identify these two groups.

Genetic analyses of both hybrid zones showed that the specific status is the main cause of genetic divergence among populations with the effect of distance or geographic barriers being weak (Lugon-Moulin *et al.* 1999, Brünner *et al.* 2002b). However, comparing these genetically very similar but karyotypically very different hybrid zones, Brünner *et al.* (2002b) did not succeed discriminating between reproductive barriers caused only by genetic factors or in combination with chromosomal differences.

Only by studying gene flow at the chromosome level can the role of their rearrangements as a reproductive barrier be addressed. Few studies have tackled the question of differential gene flow in natural hybrid zones (e.g. Rieseberg *et al.* 1999, Panithanarak *et al.* 2004) and they usually examined model species for which genetic maps of high densities are available (e.g. sunflowers or house mice). Recently Basset *et al.* (Chapter 3) mapped more than 20 microsatellite markers to the chromosome arm level. Several of these markers are located on chromosome arms belonging to the "common" group, while others are located on chromosome arms belonging to the "rearranged" group. Thus, the goals of the present study are first to compare the genetic structure measured over the "common" and "rearranged" groups of chromosomes, and second to compare the levels of genetic structure observed in the two *Sorex* hybrid zones. If karyotypic differences act as a reproductive barrier, our primary prediction is that genetic structure is higher for rearranged chromosomes than for common chromosomes. Additionally, as the complexity of the hybrids produced in the HT hybrid zone (complex heterozygotes) is larger than in the LH hybrid zone (simple heterozygotes), our second prediction is that the difference between the two groups of chromosomes is larger in the HT than in the LH hybrid zone.

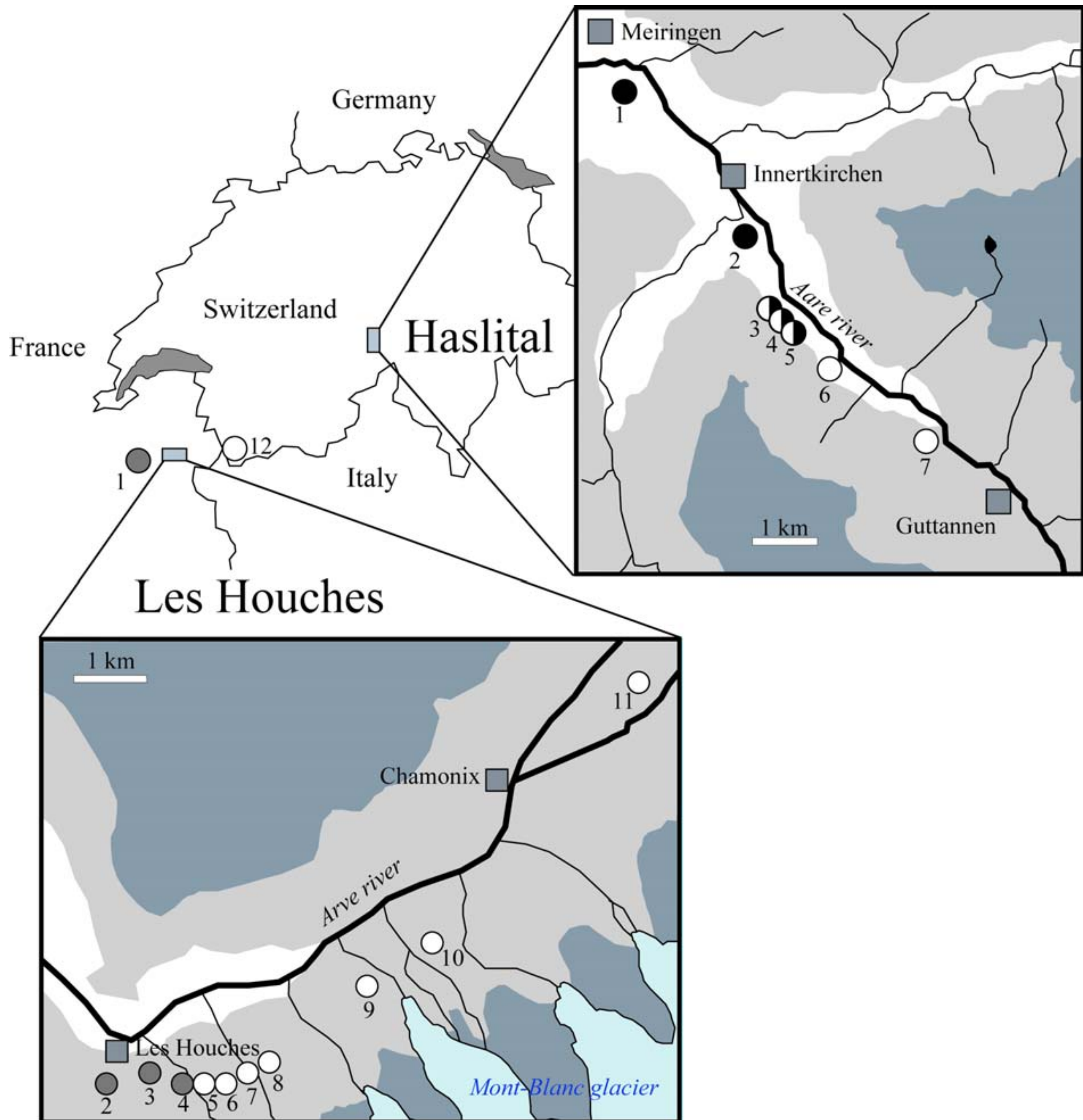


Figure 2 Study area and sampling localities in the Les Houches and Haslital hybrid zones. Open circles: *Sorex antinorii*, black circles: *S. araneus* Vaud, grey circles: *S. a.* Cordon. Both species are present in localities 3, 4 and 5 of the Haslital hybrid zone. In hybrid zones insert maps, light grey: area above 1,000m above sea level; dark grey: area above 2,000m above sea level.

MATERIAL AND METHODS

Specimens from the hybrid zones

Most of the specimens analysed in the LH hybrid zone were collected in 1992 – 1995 (except localities 1 and 12 sampled in 2003) as part of earlier studies (Brünner & Hausser 1996, Lugon-Moulin *et al.* 1996). These shrews were trapped in 12 localities and the mean sample size for each locality was 15 (range 4 – 29; Table 1). In the HT hybrid zone, all specimens analysed were collected in 1992 – 1995 as part of an earlier study (Brünner *et al.* 2002b). These shrews were trapped in seven localities and the mean size for each locality was 16 (range 6 – 33; Table 1).

According to karyotype analysis (Brünner & Hausser 1996, Brünner *et al.* 2002b), all localities could be classified as containing one or both species (Table 1). Note that the two species are only found in sympatry in the three central localities of the HT hybrid zone. As the goal of this study was to compare levels of genetic structure between species, these three localities (n° 3, 4 and 5) were split according to karyotype into monospecific sub-samples. It should also be noted that no hybrids were used in this study. Trapping and karyotype preparation conditions were described in Brünner & Hausser (1996) and Brünner *et al.* (2002b).

Table 1 Number of individuals analysed of each taxa per sample sites (see maps in Fig. 1 for trapping localities) in both hybrid zones. Number of karyotyped individuals are given in parentheses.

Les Houches			**Haslital**		
Locality	*S. a.* Cordon	*S. antinorii*	Locality	*S. a.* Vaud	*S. antinorii*
1	29(4)	0	1	6(6)	0
2	4(1)	0	2	7(6)	0
3	5(4)	0	3	8(8)	5(5)
4	21(21)	0	4	12(12)	21(21)
5	0	17(12)	5	7(7)	25(25)
6	0	20(14)	6	0	11(5)
7	0	18(13)	7	0	13(11)
8	0	14(13)	Total	40(39)	75(67)
9	0	8(6)			
10	0	8(7)			
11	0	9(8)			
12	0	22 (4)			
Total	59(30)	116(77)			

DNA extraction and microsatellite typing

Tissue samples (phalanges) were stored in alcohol (70%) at 4°C and total genomic DNA was extracted using the DNeasy Tissue Kit (Qiagen).

Seventeen microsatellite loci were chosen from the loci unambiguously mapped to the chromosome arm level in Basset *et al.* (Chapter 3), and the letters in parentheses indicate their chromosome localization: L16 (*a*), L69 (*f*), B3 (*f*), D107 (*a*), D112 (*a*), L9 (*c*), L68 (*b*), C117 (*b*), L13 (*de*), C171 (*de*), L57 (*de*), L62 (*g*), D24 (*jl*), D106 (*h*), L99 (*n*), B30 (*o*), D109 (*o*). Thus, the first 11 loci belong to the "common group" and the last six to the "rearranged group".

PCR conditions are described elsewhere (Wyttenbach *et al.* 1997, Balloux *et al.* 1998, Lugon-Moulin *et al.* 2000, Basset *et al.* in press (Chapter 1), and Basset *et al.* (Chapter 3)) except that all PCR amplifications were performed in a 20 µl total volume. Cycling was carried out in a PE9700 (Applied Biosystems) using the following profile: 95°C for 5 min, 35 cycles of 30 s at 94°C, 30 s at annealing temperature (Basset *et al.* (Chapter 3)), 30 s at 72°C; and a final extension at 72°C for 4 min. One primer of each pair was labelled with a fluorescent dye (HEX, FAM or NED) on the 5'end, which allowed analyses on an ABI 377XL sequencer (Applied Biosystems). Data collection, sizing of the bands and analysis were done using the GENESCAN software (Applied Biosystems). The individuals already genotyped at some loci (L9, L16, L57, L62 and L69) in other studies (Lugon-Moulin *et al.* 1999 and Brünner *et al.* 2002b) were not re-analysed for this study.

Statistical analyses

The software package FSTAT version 2.9.3 (updated from Goudet 1995; http://www2.unil.ch/popgen/softwares/fstat.htm) was used to calculate allele frequencies, allele numbers, observed heterozygosities (H_O), and expected heterozygosities within (H_S) and between (H_T) samples, following Nei (1987).

Both F-statistics (Wright 1965) and R-statistics (Slatkin 1995) have their strength and drawbacks in inferring genetic structure from microsatellite data (for a review, see Balloux & Lugon-Moulin 2002). However, it is not our purpose in this study to compare the relative efficiency of these statistics. Therefore, we decided to analyse genetic structure using estimates derived from F-statistics according to Weir & Cockerham (1984), using FSTAT version 2.9.3. Allele frequencies were weighted according to sample size. Heterozygote deficit within populations ($F_{IS} > 0$) was tested using a permutation procedure (10,000 randomizations) to infer random mating. The exact G-test (Goudet *et al.* 1996), as

implemented in FSTAT 2.9.3, was used to assess the significance of genetic differentiation. To assess population structure within and between the different taxa, only mono-specific populations were analysed (i.e. central populations in HT were split into mono-specific subpopulations). Hierarchical estimates of F-statistics (Weir 1996) were obtained using the software package ARLEQUIN version 2.000 (Schneider *et al.* 2000; http://anthropologie.unige.ch/arlequin). The genetic variance at different hierarchical levels (within and between species) was estimated using AMOVA (Michalakis & Excoffier 1996) in Arlequin 2.000.

When heterozygote deficit was found for a population, the software MICRO-CHECKER version 2.2.3 was used to look for genotyping errors (e.g. null alleles; Van Oosterhout *et al.* 2004, http://www.microchecker.hull.ac.uk). In cases where null alleles were found, adjusted genotypic frequencies (Chakraborty *et al.* 1992, Brookfield 1996, Van Oosterhout *et al.* 2004) were used to insure that this did not influence the genetic structure results.

Differences between the two groups of chromosomes (common and rearranged) were tested by permutation tests. A distribution of differences between groups was generated by doing 10,000 permutations of microsatellite loci between the two groups and the observed difference was compared to this distribution.

Preliminary results and the large genetic variance observed among microsatellite loci estimations seemed to indicate that historical factors related to markers (such as homoplasy) could mask a part of the difference observed between the two classes of chromosomes. Therefore, we re-analysed the centre of each hybrid zone independently since the impact of chromosomal rearrangements on gene exchange should be the stronger in these regions. The centre of both zones was defined by the localities where karyotypic hybrids had been found (Brünner *et al.* 1996; Brünner *et al.* 2002b), which correspond to populations 3, 4, 5, 6, 7 and 8 in LH, and 3, 4 and 5 in HT.

RESULTS

Polymorphism, genetic variability of loci and heterozygote deficit within population

The number of total alleles and species-specific alleles as well as the observed and expected heterozygosities from the LH and HT hybrid zones are detailed in Table 2 and 3 respectively.

LH hybrid zone. The number of total alleles and LH-specific alleles did not significantly differ between loci located on common or rearranged chromosomes (t-test, P = 0.886 and

0.645 respectively). Expected heterozygosities within samples (H_S) were generally high (with the exception of locus L13), ranging from 0.13 to 0.91, with an average of 0.72, and likewise expected heterozygosities between samples (H_T) averaged 0.78 (0.23 – 0.95). Observed heterozygosities (H_O) were equal or lower in magnitude (0.06 – 0.91; average, 0.65) than expected values.

Table 2 Number of alleles (Na) and specific alleles (Na_S) found in *Sorex araneus* Cordon, *S. antinorii* and in the whole Les Houches hybrid zone; observed heterozygozitiy (H_O), and expected heterozygosity within (H_S) and between (H_T) samples, per locus, per chromosome class (Common and Rearranged) and across all loci.

		S. a. Cordon		*S. antinorii*		Whole Les Houches hybrid zone				
Chrom. Class	Locus	Na	Na_S	Na	Na_S	Na	Na_S	H_O	H_S	H_T
C	L16	5	2	3	0	5	2	0.29	0.34	0.41
C	L69	20	6	18	4	24	10	0.73	0.86	0.90
C	B3	24	9	24	9	33	18	0.81	0.89	0.94
C	D107	18	6	13	1	19	7	0.89	0.86	0.91
C	D112	32	8	52	28	60	36	0.82	0.91	0.95
C	L9	21	9	20	6	29	15	0.87	0.88	0.92
C	L68	10	1	12	3	13	4	0.82	0.81	0.87
C	C117	15	6	11	2	17	8	0.65	0.74	0.79
C	L13	4	2	2	0	4	2	0.06	0.13	0.23
C	C171	11	5	17	11	22	16	0.56	0.62	0.63
C	L57	18	4	20	6	25	10	0.79	0.86	0.92
R	L62	16	7	13	4	20	11	0.84	0.80	0.90
R	D24	33	9	35	11	44	20	0.49	0.91	0.94
R	D106	9	2	8	1	10	3	0.48	0.51	0.60
R	L99	4	0	4	0	4	0	0.41	0.43	0.58
R	B30	10	4	12	6	16	10	0.55	0.73	0.82
R	D109	28	6	30	8	36	14	0.91	0.91	0.95
Mean Common		16.2	5.3	17.5	6.4	22.8	11.6	0.66	0.72	0.77
Mean Rearranged		16.7	4.7	17.0	5.0	21.7	9.7	0.60	0.71	0.80
Mean all loci		16.4	5.1	17.3	5.9	22.4	10.9	0.64	0.72	0.78

HT hybrid zone. Again, the number of total alleles and HT-specific alleles did not differ significantly between loci located on common or rearranged chromosomes (t-test, P = 0.782 and 0.997 respectively). Expected heterozygosities within samples (H_S) were generally high, ranging from 0.26 to 0.90, with an average of 0.72, and likewise expected heterozygosities

between samples (H_T) averaged 0.80 (0.41 – 0.95). Observed heterozygosities (H_O) were equal or lower in magnitude (0.20 – 0.91; average, 0.63) than expected values.

Table 3 Number of alleles (Na) and specific alleles (Na_S) found in *Sorex araneus* Cordon, *S. antinorii*, and in the whole Haslital hybrid zone; observed heterozygozitiy (H_O), and expected heterozygosity within (H_S) and between (H_T) samples, per locus, per chromosome class (Common and Rearranged) and across all loci.

		S. a. Vaud		*S. antinorii*		Whole Haslital hybrid zone				
Chrom. Class	Locus	Na	Na_S	Na	Na_S	Na	Na_S	H_O	H_S	H_T
C	L16	3	1	3	1	4	2	0.29	0.33	0.41
C	L69	18	5	17	4	22	9	0.87	0.89	0.92
C	B3	14	3	25	14	28	17	0.67	0.85	0.88
C	D107	9	2	11	4	13	6	0.79	0.83	0.83
C	D112	20	4	30	14	34	18	0.88	0.90	0.94
C	L9	11	3	22	14	25	17	0.91	0.89	0.94
C	L68	11	1	11	1	12	2	0.72	0.82	0.87
C	C117	15	4	14	3	18	7	0.79	0.84	0.90
C	L13	2	0	3	1	3	1	0.20	0.44	0.55
C	C171	20	5	25	10	30	15	0.68	0.78	0.84
C	L57	16	8	18	10	26	18	0.78	0.88	0.91
R	L62	11	4	15	8	19	12	0.79	0.84	0.87
R	D24	22	9	22	9	31	18	0.55	0.88	0.94
R	D106	11	5	7	1	12	6	0.31	0.41	0.63
R	L99	2	0	3	1	3	1	0.19	0.26	0.59
R	B30	6	1	7	2	8	3	0.46	0.51	0.68
R	D109	17	4	30	17	34	21	0.85	0.89	0.95
Mean Common		12.6	3.3	16.3	6.9	19.5	10.2	0.69	0.77	0.82
Mean Rearranged		11.5	3.8	14.0	6.3	17.8	10.2	0.52	0.63	0.78
Mean all loci		12.2	3.5	15.5	6.7	18.9	10.2	0.63	0.72	0.80

In both hybrid zones and across all loci, within population heterozygote deficit was highly significantly different from 0 ($F_{IS\ LH}$ = 0.100, $F_{IS\ HT}$ = 0.133; Table 4). Although, heterozygote deficit was slightly higher in the rearranged than in the common group, no significant difference was observed between these two groups (permutation test: P = 0.301 for LH and 0.289 for HT; Table 4). At least a part of this deficit could be explained by the presence of genotyping errors (Van Oosterhout *et al.* 2004). In both hybrid zones, null alleles were detected in 12 of the 17 loci in at least one population. Using adjusted frequencies (Van Oosterhout *et al.* 2004) for these loci, among populations genetic structures across common, rearranged and all loci were in the same order of magnitude as unadjusted frequencies.

Therefore, only the genetic structure results based on observed frequencies are presented in the following sections.

Table 4 Values of F_{IS} estimated from the loci located on common or rearranged chromosomes and acorss all loci for the *Sorex araneus* side, the *S. antinorii* side and the whole Les Houches and Haslital hybrid zones.

	Les Houches					
	S. araneus side		*S. antinorii* side		Overall	
	$F_{IS(Cordon)}$	Test diff.	$F_{IS(ant)}$	Test diff.	F_{IS}	Test diff.
Common loci	0.089 ***	NS	0.062 ***	NS	0.071 ***	NS
Rearranged loci	0.161 ***		0.155 ***		0.152 ***	
All loci	0.115 ***		0.090 ***		0.100 ***	

	Haslital					
	S. araneus side		*S. antinorii* side		Overall	
	$F_{IS(Vaud)}$	Test diff.	$F_{IS(ant)}$	Test diff.	F_{IS}	Test diff.
Common loci	0.139 ***	NS	0.094 ***	NS	0.109 ***	NS
Rearranged loci	0.183 ***		0.183 ***		0.183 ***	
All loci	0.152 ***		0.123 ***		0.133 ***	

N.B. Test diff. indicates if the values given by the common and rearranged loci are different. Asterisks indicate significant values for the estimators: *** P < 0.001; NS = not significant.

Genetic structuring

The genetic structure parameters according to Weir & Cockerham (1984) of both hybrid zones are summarized in Figure 3.

LH hybrid zone. Across all loci, the highly significant F_{ST} (0.082, P < 0.001) suggests a moderate genetic structuring. The genetic structure estimated across loci located on rearranged chromosomes is slightly higher than across loci located on common chromosomes ($F_{ST\text{-rearranged}}$ = 0.109 vs. $F_{ST\text{-common}}$ = 0.067) but this difference is not significant (permutation test: P = 0.159). As two species are involved, F-statistics were also estimated for *S. araneus* Cordon and *S. antinorii* samples independently (Fig. 3A). In both species, we found highly significant (P < 0.001) over all loci F_{SR} values (0.065 for *S. araneus* Cordon and 0.032 for *S. antinorii*) but we did not find any significant difference between the two groups of loci (Fig. 3A; $F_{SR\text{-Cordon}}$ = 0.060 and 0.073, $F_{SR\text{-antinorii}}$ = 0.029 and 0.037 for common and rearranged loci respectively; permutation tests: P = 0.736 and 0.403).

HT hybrid zone. Here also, the highly significant across all loci F_{ST} (0.104, P < 0.001) suggests a moderate genetic structuring of this zone. Interestingly, the genetic structure estimated across the rearranged loci was significantly higher than across the common loci

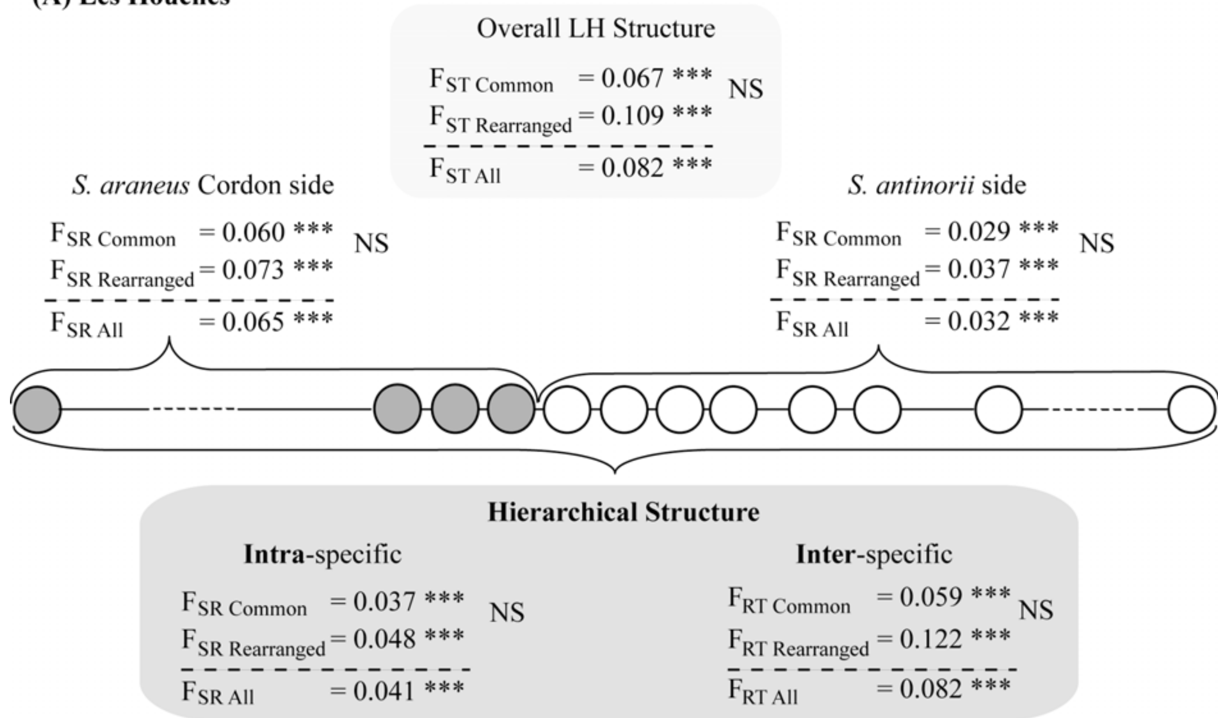
($F_{ST\text{-rearranged}}$ = 0.181 vs. $F_{ST\text{-common}}$ = 0.063; permutation test: P = 0.023). This suggests that chromosomal rearrangements have an effect on gene exchange in this hybrid zone. Although, highly significant genetic structures were observed across all loci in both species (F_{SR} = 0.029 for *S. araneus* Vaud and F_{SR} = 0.025 for *S. antinorii*), we did not find any significant difference between the common and rearranged loci within each species (Fig. 3B; F_{SRVaud} = 0.024 and 0.040, $F_{SR\ antinorii}$ = 0.016 and 0.043 for common and rearranged loci respectively; permutation tests: P = 0.398 and 0.191).

Hierarchical F-statistics

In a second step, hierarchical F-statistics (deviations from Hardy-Weinberg equilibrium within population (F_{IS}) and within the entire zone (F_{IT}), differentiation of populations within species (intra-specific, F_{SR}), and differentiation between populations of the two species (inter-specific, F_{RT})) were estimated for both hybrid zones. Per locus, across common, across rearranged and across all loci values of both hybrid zones are presented in Tables 5 and 6 and a summary of the genetic structure is given in Figure 3.

LH hybrid zone. Differentiation of populations within each species (F_{SR}) was highly significant across common (0.037; P < 0.001), rearranged (0.048; P < 0.001) and all loci (0.041; P < 0.001). The values across common and rearranged loci were not significantly different (permutation test: P = 0.487). Interestingly, differentiations of populations between species (F_{RT}) strongly varied across loci (Table 5, Fig. 4). Across all loci, this last parameter was highly significant (0.082; P < 0.001) and higher than within species. Moreover, this genetic differentiation was higher across rearranged ($F_{RT\ LH\ specific}$ = 0.122; P < 0.001) than across common loci ($F_{RT\ LH\ common}$ = 0.059; P < 0.001) but the difference between these two groups was not significant (permutation test: P = 0.182).

HT hybrid zone. Differentiation of populations within each species (F_{SR}) was highly significant across common (0.019; P < 0.001), rearranged (0.041; P < 0.001) and all loci (0.026; P < 0.001). The values of common and rearranged loci were not significantly different (permutation test: P = 0.222). Again, differentiation of populations between species (F_{RT}) strongly varied across loci (Table 6, Fig. 4), was highly significant across all loci (0.143; P < 0.001) and was much higher than within species. This genetic differentiation was marginally significantly higher across the rearranged chromosomes ($F_{RT\ HT\ rearranged}$ = 0.248; P < 0.001) than across the common chromosomes ($F_{RT\ HT\ common}$ = 0.083; P < 0.001) (permutation test: P = 0.063).



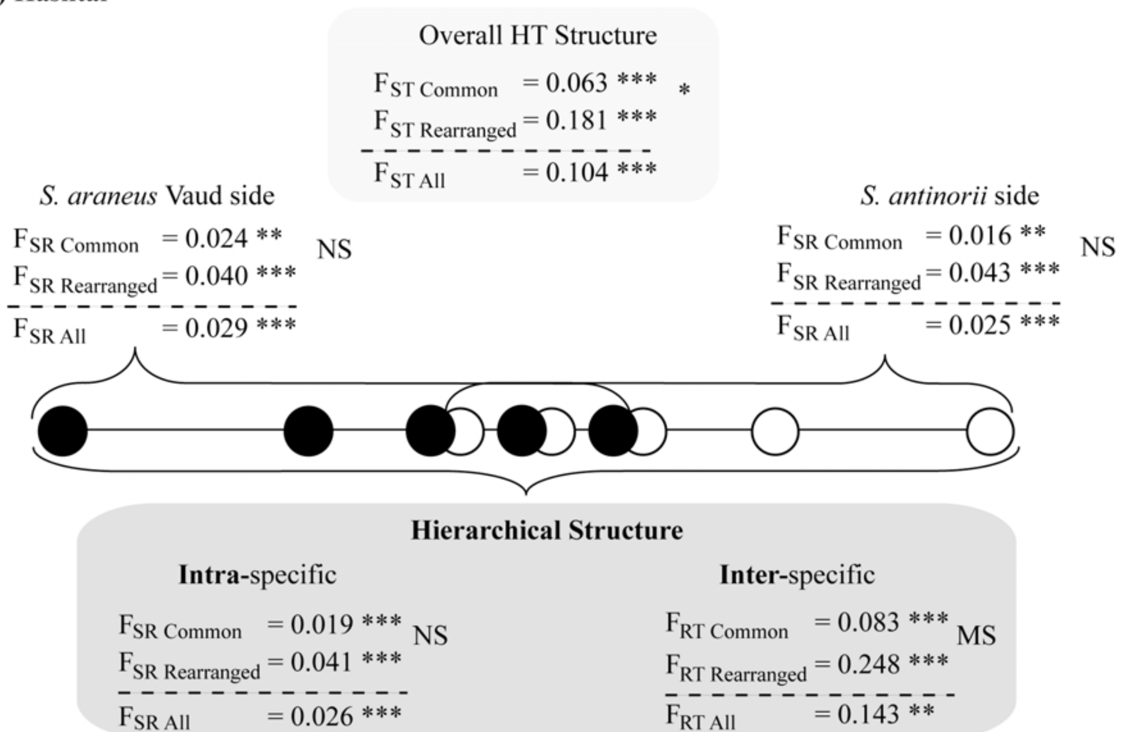


Figure 3 Summary of the genetic structure observed in the Les Houches and Haslital hybrid zones over "common", "rearranged" and all loci. Open, grey and black circles represent the *Sorex antinorii*, *S. araneus* Cordon and *S. araneus* Vaud populations, respectively.

Table 5 Hierarchical F-statistics[a] per locus, across Common, Rearranged and across all loci in the Les Houches hybrid zone.

Les Houches Hybrid Zone									
		Hierarchical F-statistics							
Chrom. Class	Locus	F_{IS}		F_{SR}		F_{RT}		F_{IT}	
C	L16	0.064 NS		0.149 ***		0.148 *		0.322 ***	
C	L69	0.114 **		0.028 ***		0.025 **		0.160 ***	
C	B3	0.123 ***		0.035 ***		0.012 **		0.164 ***	
C	D107	-0.023 NS		0.032 ***		0.029 **		0.038 NS	
C	D112	0.069**		0.046 ***		0.011 NS		0.122 ***	
C	L9	0.019 NS		0.009 NS		0.053 **		0.078 *	
C	L68	0.012 NS		0.026 **		0.104 **		0.138 **	
C	C117	0.135 **		0.017 NS		0.081 **		0.218 ***	
C	L13	0.470 ***		0.328 ***		0.021 NS		0.651 ***	
C	C171	0.085 *		0.036 **		0.019 NS		0.134 **	
C	L57	0.087 **		0.003 NS		0.139 **		0.215 ***	
R	L62	-0.033 NS		0.062 ***		0.089 **		0.117 **	
R	D24	0.475 ***		0.032 **		0.012 NS		0.497 ***	
R	D106	0.044 NS		0.071 ***		0.127 **		0.225 ***	
R	L99	0.065 NS		0.072 **		0.375 **		0.458 ***	
R	B30	0.251 ***		0.037 **		0.181 **		0.409 ***	
R	D109	0.023 NS		0.033 ***		0.008 NS		0.062 **	
			Test diff.		Test diff.		Test diff.		Test diff.
Common		0.072 ***	NS	0.037 ***	NS	0.059 ***	NS	0.159 ***	NS
Rearranged		0.152 ***		0.048 ***		0.122 ***		0.291 ***	
All loci		0.100 ***		0.041 ***		0.082 ***		0.208 ***	

[a] Subscripts I, S, R, T stand for individuals, samples, species, total respectively. Test diff. indicates if the values given by the common and rearranged loci are different. Asterisks indicate significant values for the estimators: * P < 0.05, ** P < 0.01, *** P < 0.001; NS = not significant.

Table 6 Hierarchical F-statistics[a] per locus, across Common, Rearranged and across all loci in the Haslital hybrid zone.

Haslital Hybrid Zone									
		Hierarchical F-statistics							
Chrom. Class	Locus	F_{IS}		F_{SR}		F_{RT}		F_{IT}	
C	L16	0.098 NS		0.051 NS		0.328 *		0.424 **	
C	L69	0.032 NS		-0.006 NS		0.073 **		0.097 *	
C	B3	0.204 ***		0.010 NS		0.065 **		0.263 ***	
C	D107	0.045 NS		0.003 NS		0.002 NS		0.050 NS	
C	D112	0.045 NS		0.020 *		0.043 **		0.104 **	
C	L9	-0.006 NS		0.000 NS		0.089 *		0.084 NS	
C	L68	0.076 *		0.024 *		0.043 **		0.138 **	
C	C117	0.068 *		0.038 **		0.079 *		0.174 ***	
C	L13	0.542 ***		-0.019 NS		0.319 **		0.683 ***	
C	C171	0.197 ***		0.062 ***		0.022 NS		0.263 ***	
C	L57	0.110 **		0.033 **		0.026 NS		0.161 ***	
R	L62	0.076 *		0.022 *		0.036 *		0.128 ***	
R	D24	0.382 ***		0.037 **		0.049 *		0.434 ***	
R	D106	0.308 ***		0.000 NS		0.530 **		0.675 ***	
R	L99	0.322 ***		0.009 NS		0.638 *		0.757 ***	
R	B30	0.060 NS		0.147 ***		0.203 *		0.361 ***	
R	D109	0.065 *		0.017 *		0.066 **		0.142 ***	
			Test diff.		Test diff.		Test diff.		Test diff.
Common		0.109 ***	NS	0.019 ***	NS	0.083 ***	MS	0.198 ***	MS
Rearranged		0.183 ***		0.041 ***		0.248 ***		0.411 ***	
All loci		0.133 ***		0.026 ***		0.143 **		0.276 ***	

[a] Subscripts I, S, R, T stand for individuals, samples, species and total respectively. Test diff. indicates if the values given by the common and rearranged loci are different. Asterisks indicate significant values for the estimators: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; MS = marginally significant ($0.05 \leq P \leq 0.10$), NS = not significant.

Centre of hybrid zones

Because the centre of hybrid zones might provide higher resolution, we estimated hierarchical F-statistics in the centre of both zones (Table 7). In both zones, heterozygote deficit was highly significant ($F_{IS\ centre\ LH} = 0.103$, $F_{IS\ centre\ HT} = 0.148$). However, no significant differences were observed between the values estimated across common or rearranged loci (permutation tests: $P = 0.220$ for LH and 0.381 for HT). Population differentiation within species across all loci was highly significant in the centre of the LH hybrid zone ($F_{SR\ LH}$ =

0.024) but not significant in the centre of the HT hybrid zone ($F_{SR\ HT}$ = 0.006). Again, in both zones, no significant differences were observed between the values estimated across common or rearranged loci (permutation tests: P = 0.677 for LH and 0.694 for HT). Population differentiation between species in the centre of the LH hybrid zone across all loci gave a slightly lower value ($F_{RT\ centre\ LH}$ = 0.071; P < 0.001) than across the whole hybrid zone and no difference was observed between the common and rearranged loci ($F_{RT\ centre\ LH\ common}$ = 0.062, P < 0.001; $F_{RT\ centre\ LH\ rearranged}$ = 0.085, P < 0.001; permutation test: P = 0.398). In the centre of the HT hybrid zone, population differentiation between species gave a slightly higher value ($F_{RT\ centre\ HT}$ = 0.163, P < 0.001) than over the whole hybrid zones. Interestingly, the loci located on rearranged chromosomes ($F_{RT\ centre\ HT\ rearranged}$ = 0.276, P < 0.001) were significantly more structured than the loci located on common chromosomes ($F_{RT\ centre\ HT\ common}$ = 0.098, P < 0.001) (permutation test: P = 0.036). This last result clearly support the hypothesis that in the centre of this hybrid zone, chromosomal rearrangements significantly act as a barrier to gene flow for only some parts of the genome.

Figure 4 Individual intra- (F_{SR}) and inter- (F_{RT}) specific values for each loci located across common (black symbols) and rearranged (open symbols) chromosomes in Les Houches and Haslital hybrid zones. Dotted lines indicate the values across each group of chromosomes.

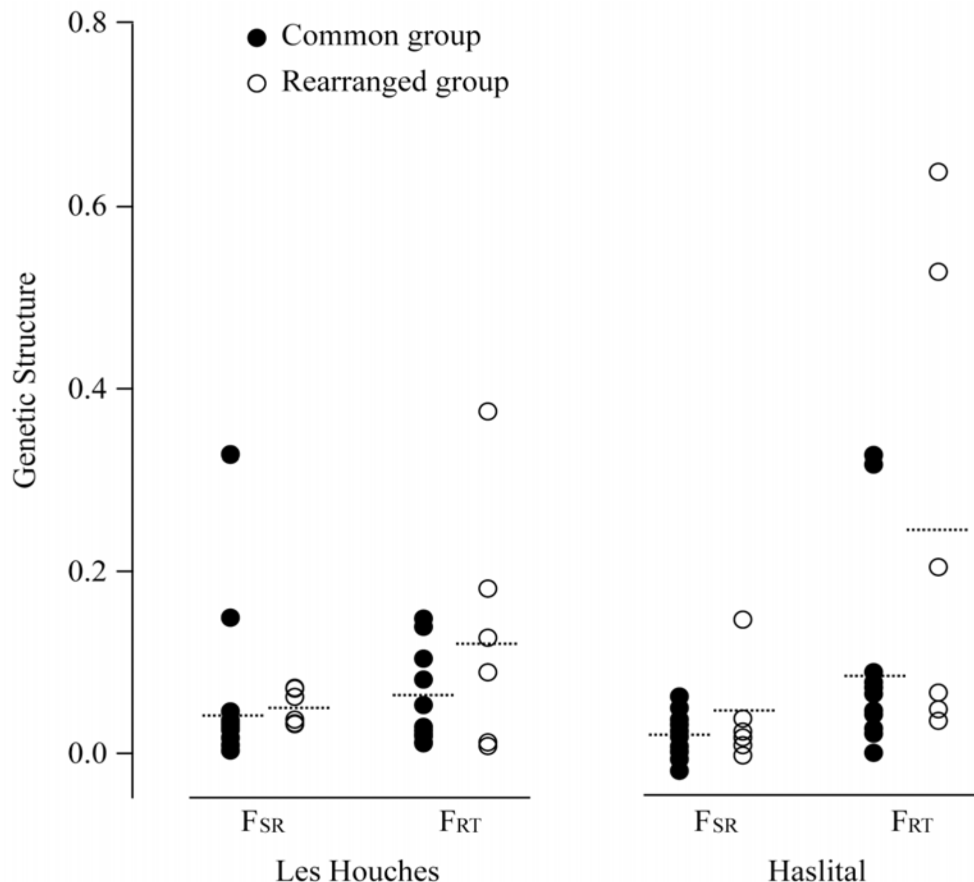


Table 7 Hierarchical F-statistics[a] over common, rearranged and over all loci in the centre of the Les Houches and Haslital hybrid zones.

Centre Les Houches hybrid zone								
	Hierarchical F-statistics							
	F_{IS}	Test diff.	F_{SR}	Test diff.	F_{RT}	Test diff.	FIT	Test diff.
Common	0.066 ***	NS	0.022 ***	NS	0.062 ***	NS	0.143 ***	NS
Rearranged	0.169 ***		0.029 ***		0.085 ***		0.262 ***	
All loci	0.103 ***		0.024 ***		0.071 ***		0.186 ***	

Centre Haslital hybrid zone								
	Hierarchical F-statistics							
		Test diff.		Test diff.		Test diff.		Test diff.
Common	0.127 ***	NS	0.005 NS	NS	0.098 ***	*	0.216 ***	MS
Rearranged	0.193 ***		0.008 NS		0.276 ***		0.420 ***	
All loci	0.148 ***		0.006 NS		0.163 ***		0.291 ***	

[a] Subscripts I, S, R, T stand for individuals, samples, species and total respectively. Test diff. indicates if the values given by the common and rearranged loci are different. Asterisks indicate significant values for the estimators: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; MS = marginally significant ($0.05 \leq P \leq 0.10$), NS = not significant.

DISCUSSION

Overall genetic structure in the LH and HT hybrid zones

Our results using 17 microsatellite loci indicate a relatively high level of genetic differentiation in both hybrid zones ($F_{ST} \approx 0.10$; Table 4 and 5) compared to other hybrid zones between karyotypic taxa of the *Sorex araneus* group (e.g. $F_{ST} \approx 0.02$; Wyttenbach *et al.* 1999, Andersson *et al.* 2004). These differences are probably explained by the occurrence in the two alpine hybrid zones of the genetically differentiated *S. antinorii* (Brünner *et al.* 2002a, Basset *et al.* in press (Chapter 1)) as suggested by the larger inter- than intra-specific structure observed in both hybrid zones (Fig. 3). Interestingly, the genetic structure values measured in this study are globally similar to what Lugon-Moulin *et al.* (1999) and Brünner *et al.* (2002b) observed in the same hybrid zones using only 7 microsatellite loci. This implies that the set of 10 loci added in this study carry similar information to the first seven loci.

Genetic structure over all populations is slightly larger in HT than LH ($F_{ST\ All\ HT} = 0.104$ vs. $F_{ST\ All\ LH}$ 0.082). The only probable explanation for this difference relates to differences

between the taxa in contact (*S. antinorii* vs. *S. araneus* Cordon in LH and *S. araneus* Vaud in HT). This is further supported by the much larger inter-specific structure detected in HT compared to LH ($F_{RT\ HT}$ = 0.143 and $F_{RT\ LH}$ = 0.082). Moreover, larger levels of introgression between *S. araneus* Cordon and *S. antinorii* than between *S. araneus* Vaud and *S. antinorii* were already noticed by Basset *et al.* (in press (Chapter 1)) at a larger geographical scale. Genetic incompatibilities are predicted to be similar in both hybrid zones since genetic differentiation between the Cordon and Vaud chromosome races is extremely low (Taberlet *et al.* 1994, Basset *et al.* in press (Chapter 1)) but as previously mentioned chromosomal incompatibilities are expected to be stronger in HT. The presence of numerous rivers in the LH hybrid zone could also influence the overall genetic structure observed in this zone (Brünner *et al.* 2002b). However, rivers have been shown not to constitute a strong barrier to gene flow for these shrews (Lugon-Moulin *et al.* 1999b). Furthermore, these geographical factors should increase the genetic structure in LH instead of decreasing it. These observations therefore support our hypothesis of an impact of chromosomal rearrangements on the genetic structure of these hybrid zones. Nevertheless, only by individually studying each category of chromosomes (common and rearranged) can other historical factors be ruled out.

Differences between common and rearranged chromosomes

As expected from our primary prediction, in the HT hybrid zone we detected significantly higher levels of genetic structure and inter-specific structure (although only marginally significant for the latter) across loci located on the rearranged chromosomes than across loci located on common chromosomes (Table 6, Fig. 3B). Intra-specific structure could be used as a control for the real significance of observed differences since there are no karyotypic differences within species. Not surprisingly, no significant difference was observed between the two classes of chromosomes within species. Therefore, our hypothesis of chromosomal rearrangements playing a role on the reproductive barrier of this hybrid zone is strongly supported.

As previously mentioned, *Sorex antinorii* probably diverged genetically in allopatry during the last glaciations period (Brünner *et al.* 2002a). After re-colonization, this species made contact with *S. araneus* in several valleys. Through this contact, gene flow between *S. antinorii* and *S. araneus* reduced interspecific differences for most regions of the genome. However, in the HT hybrid zone, hybrids produce a chain of rearranged chromosomes (i.e. “complex” heterozygotes), which compose a large linkage block. Introgression of alleles into

the heterospecific background should be strongly impeded by this block. Therefore, loci located on this block would have been protected from gene flow and have remained strongly differentiated between *S. araneus* Vaud and *S. antinorii*.

In the LH hybrid zone, hybrids mostly produce trivalents (i.e. "simple" heterozygotes). These hybrids have been shown to be well supported by these shrews (Narain & Fredga 1997, 1998, Searle & Wójcik 1998). Therefore, rearranged chromosomes in the LH hybrid zone should introgress much more easily than in the HT. Our results support this second prediction since the difference in the genetic structure of common and rearranged chromosomes in this zone is not significant and is of lesser magnitude than in HT hybrid zone (Table 5, Fig. 3A). However, it should be noted that even in this hybrid zone, rearranged chromosomes are more structured than common chromosomes, suggesting that karyotypic differences influence the genetic structure of this zone as well.

Interestingly, most of the discrepancies observed between LH and HT are carried by the rearranged chromosomes ($F_{RT\ rearranged\ LH}$ = 0.122, $F_{RT\ rearranged\ HT}$ = 0.248). This difference probably reflects the larger impact of chromosomal rearrangements on the genetic structure in the HT hybrid zone. Finally, the genetic differentiation observed across common chromosomes in both zones ($F_{RT\ common\ LH}$ = 0.059, $F_{RT\ common\ HT}$ = 0.083), certainly reflects the genetic differences accumulated between the two species involved in these zones.

Variation across loci

The loci within each chromosome class are not equally informative and the variance observed across loci is large (Fig. 4). For example, inter-specific structure (F_{RT}) across rearranged chromosomes vary from 0.008 (locus D109) to 0.375 (L99) in LH and 0.036 (L62) to 0.638 (L99) in HT. Several non exclusive factors may explain this strong variance: i.e. intrinsic characteristics of the loci (e.g. number of alleles, evolution patterns), which chromosome a locus is located on or the position within the chromosome.

First, we observed large variations in the allele number per loci and this may bias the estimated population differentiation. However, this should not influence our conclusions as the number of alleles observed in both hybrid zones is similar for the common and rearranged categories (Table 2 and 3).

Secondly, in our analyses we pooled loci located on different chromosomes (e.g. the six loci that compose the rearranged group are located on five different chromosomes). For example, loci contributing to reproductive isolation in animals are disproportionately found on particular chromosomes (e.g. chromosome 17 in mice, Yeom *et al.* 1992) and in this context,

the X chromosome plays an important role (Coyne & Orr 1989). Moreover, X-linked markers show reduced introgression across a number of mice hybrid zones (Tucker *et al.* 1992, Dod *et al.* 1993, Payseur *et al.* 2004, Payseur & Nachman 2005). In the *Sorex araneus* group, the sex chromosome system in males is unusual (i.e. XY1Y2) and only the smallest arm *e* can be considered as the "real" X chromosome (Zima *et al.* 1998). In our study, this arm has been treated identically to the other chromosomes of the common group but more extensive studies of its role as a reproductive barrier between *Sorex* species are necessary to confirm the appropriateness of this decision.

Third, genetic differentiation estimates strongly differ among loci situated over the same chromosome arm. For example, loci L16, D107 and D112 all map to chromosome arm *a,* but show F_{RT} values in HT ranging from 0.002 to 0.328. Several genetic factors are suspected to play important roles on gene flow among populations. One is the rate of recombination along the chromosome (e.g. Ortíz-Barrientos *et al.* 2002, Butlin 2005, Stump *et al.* 2005). Studies in numerous organisms demonstrate that recombination is not uniformly distributed along the genome and that most recombination events occur at highly localized "hot spots" (e.g. Kauppi *et al.* 2004). In general, lower recombination rates are observed near the centromere of metacentric chromosomes (Nachman 2001). The potential impact of the variation of recombination rate along chromosomes is exemplified by Panithanarak *et al.* (2004). These authors showed in a mouse hybrid zone that loci near the centromere (i.e. experiencing low levels of recombination) of rearranged chromosomes were protected from gene flow which was not the case for loci located close to the telomeres. Unfortunately, data about the localization of loci within chromosomes in the *Sorex araneus* group are lacking and fine scale localizations are necessary to test a possible impact of recombination.

Finally, low genetic differentiation for several loci located on rearranged chromosomes could reflect the retention of same ancestral polymorphism or homoplasy. The study of the centre of these hybrid zones brings an important perspective to this question. It is indeed in localities where hybrids were identified that "effective" gene flow occurs and that differences in chromosomal rearrangements will have the strongest impact. Interestingly, inter-specific structure (F_{RT}) increases in HT in the localities where the two species occur in sympatry (Table 7). Moreover, the difference between common and rearranged inter-specific structure increases over the same localities and is significant ($P = 0.038$). This highlights again the role of chromosomal rearrangements in the reproductive barrier between *S. antinorii* and *S. araneus* Vaud and suggests an impact stronger and/or less masked in the centre than over the whole hybrid zone. Therefore, when studying similar hybrid zones, examining loci with

comparable intrinsic characteristics or to increasing the number of markers used in each chromosomal category is recommended.

Hybrid dysfunction vs. suppressed recombination?

Models proposing that chromosomal rearrangements facilitate speciation fall into two main categories: the “hybrid dysfunction” and the “suppressed recombination” models (Ayala & Coluzzi 2005).

Hybrid dysfunction models claim that rearrangements reduce the fertility and the reproductive fitness of heterozygous hybrids (White 1978, King 1993), but suffer from several empirical and theoretical difficulties (e.g. Rieseberg 2001, Navarro & Barton 2003b, Coyne & Orr 2004). Data from the *Sorex araneus* group suggest that Robertsonian heterozygotes do not suffer from infertility as substantially as other taxa (Searle 1993, Narain & Fredga 1997, 1998, Banaszek 2000). Nevertheless, Banaszek *et al.* (2002) observed levels of nondisjunction high enough to affect fertility of complex heterozygotes. These authors furthermore highlighted the difference between simple and complex heterozygotes since they did not find any evidence that simple heterozygotes are less fit than homozygotes. Furthermore, the absence of male mediated gene flow detected in the LH hybrid zone (Balloux *et al.* 2000) suggests that male F1 hybrids could be sterile in this zone.

Suppressed recombination models claim that suppression of recombination by chromosomal rearrangements could be more important than their effect on fitness (Rieseberg 2001). Data concerning reduction of recombination in the *Sorex araneus* group are scarce but suppression of recombination in heterozygous Robertsonian individuals was reported in the case of mice (Davisson & Akeson 1993; Haigis & Dove 2003, Merico *et al.* 2003).

With our data teasing apart the two categories of models is not possible. As previously mentioned, the high variance observed across the loci of the rearranged group is concordant with variation of recombination along chromosomes and therefore concordant with the suppressed recombination model. However, if recombination only affects the differential genetic structure of these hybrid zones, we expect structure to be similar for common chromosomes in both hybrid zones. Inter-specific structure of common chromosomes is larger in HT than in LH ($F_{RT\ common}$ = 0.083 in HT and 0.059 in LH) suggesting that hybrid dysfunction also acts in these hybrid zones. Thus, both hybrid dysfunction and reduced recombination likely contribute to the genetic structure of these hybrid zones.

Conclusions

Using mapped genetic markers we have been able to show differences in the porosity of gene flow of some regions of the genome. Furthermore, our data add to recent studies (e.g. Rieseberg *et al.* 1999, Noor *et al.* 2001, Machado *et al.* 2002, Navarro & Barton 2003b, Panithanarak *et al.* 2004) supporting the role of chromosomal rearrangements in the reproductive barrier between species. Even if it is likely that other factors, such as genetic incompatibilities accumulated in allopatry, also affect the gene flow between taxa and even if the reduction of gene flow for some parts of the genome does not necessarily imply speciation, our results strongly suggest that chromosomal rearrangements if linked to "isolation" genes (e.g. Rieseberg 2001) could facilitate the genetic diversification of the *S. araneus* group, finally promoting speciation events.

Comparing two hybrid zones with different characteristics allowed us to highlight the importance of the chromosome composition of hybrids. Although the genetic incompatibilities were essentially the same for both zones, an effect of chromosomal rearrangements was only detected in the zone with the larger chromosomal incompatibilities. More studies of *Sorex* hybrid zones with diverse chromosomal (e.g. complex heterozygotes forming rings or shorter chains of chromosome) and genetic (e.g. within *S. araneus*) characteristics would thus allow to address further important issues about the relative roles of genes and chromosomes in the evolution of reproductive barriers between chromosomal variants.

ACKNOWLEDGMENTS

This work benefited from the invaluable advice of Agata Banaszek, Laura Brändli, Aline Depraz, Valérie Vogel and John Wang. This work was supported by the Swiss NSF (grants n°: 3100-05943), a Société Académique Vaudoise (Switzerland) grant to PB and a Fondation Agassiz (Switzerland) grant to PB.

General Conclusions and Perspectives

In the introduction of this thesis, we addressed the question of a causative role of chromosomal rearrangements in the speciation process in general and more specifically in the genetic diversification of the shrews of the *Sorex araneus* group. All along this work, we have tried to tackle these issues using microsatellite markers in the framework of two different approaches.

Bayesian admixture analyses

We used Bayesian admixture protocols (Pritchard *et al.* 2000) to check the relationships between genotypic and karyotypic structure in five taxa of the *S. araneus* group with the prediction that genetic structure should be closely associated to karyotypic structure if this last parameter influence the first one (Chapter 1 and 2). Although, we have clearly established an association between the genetic and karyotypic structure at the species level (*S. araneus*, *S. antinorii* and *S. coronatus*) of our dataset, this association was much weaker among the three chromosome races studied (*S. araneus* Cordon, Bretolet and Vaud). Moreover, the detection within *S. antinorii* (although its constant karyotype) of geographically based cryptic substructure as strong as among chromosome races suggested that chromosomal differences have a relatively low impact on genetic structure compared to historical factors.

These differences between the species and chromosome races levels may be accounted for by various factors, therefore some of the conclusions drawn in these chapters need to be moderated. First of all, the karyotype of *S. araneus* Bretolet is intermediate between *S. a.* Vaud and Cordon and it was postulated to be formed by the hybridization of these two taxa (Hausser *et al.* 1991). Therefore it is likely that this race may have facilitated the genetic exchanges between the two karyotypically more distinct taxa and subsequently reduced the impact of karyotypic differences among these three taxa. Moreover, it should be noted that only well tolerated simple heterozygotes could be formed by these three chromosome races. Second, two hierarchical levels of evolutionary divergence have been considered in this work: the within and the between species levels. As expected the level of genetic structure measured among species was higher than among the chromosome races. The performances of Bayesian admixture protocols are sensitive to both the level of genetic differentiation and the number of marker analysed (Berry *et al.* 2004, Vähä & Primmer 2005). Accuracy of results improves with levels of divergence and in cases of low divergence more markers are necessary to reach similar accuracy. For this study, we analysed 10 markers, which is enough when considering the differentiation among our species (i.e. $F_{ST} \geq 0.12$) but which is probably low when considering the differentiation among the chromosome races within *S. araneus* (i.e. $F_{ST} \leq$

0.06). Consequently, some of the discordance noticed between the two hierarchical levels could be explained by differences in the efficiency of the performed analyses. One way to address these issues would be to analyse a larger number of chromosome races and to adapt the number of markers according to the smallest level of differentiation detected. However it is necessary to keep in mind that such approaches will only allow correlative but no causative conclusions.

Differential genetic diversification

We analysed independently the two classes of chromosomes (i.e. common and rearranged) in order to focus on the role of rearranged chromosomes on the genetic differentiation among taxa. In a first step, we used flow sorted chromosomes to map microsatellite markers at the chromosome arm level in three karyotypic taxa (*S. granarius*, *S. araneus* Cordon and Novosibirsk). Although it has some limitations, (e.g. relatively low efficiency and lack of resolution within chromosome), this technique allowed us to map 25 markers and to show their potential efficiency for inter-taxa comparisons (Chapter 3). These markers have then been used to test the expectation that if chromosomal rearrangements affect genetic diversification, the genetic divergence and the genetic structure between two taxa should be greater in the regions of their genome located on chromosomes differently rearranged. This prediction was first tested in pairwise comparisons of five karyotypic taxa (*S. coronatus*, *S. antinorii*, *S. araneus* Vaud, Cordon and Białowieza) placed at different evolutionary levels of the *S. araneus* group (Chapter 4) and then in two hybrid zones between *S. antinorii* and *S. araneus* (Chapter 5). As expected these studies indicated a generally higher genetic differentiation and genetic structure of rearranged chromosomes. Such conclusions strongly support the role of the rearrangements in the genetic differentiation of the *S. araneus* group. As highlighted by several authors (e.g. King 1993, Noor *et al.* 2001a, Rieseberg 2001, Navarro & Barton 2003a), the restriction of gene exchanges across large blocks of rearranged chromosomes may allow the accumulation of isolation genes and extend their effects over a larger fraction of the genome, hence favouring the establishment of complete reproductive barriers between taxa, finally meaning speciation.

We have nevertheless come across several limitations during these studies and taking into account these factors in future studies of the *S. araneus* group would address into more details several important issues of chromosomal speciation. A first limitation is that we have only been able to analyse each class of chromosomes as a whole. However, as highlighted by our results, the strength of the impact of rearrangements depends on the karyotypic configurations

of hybrids (e.g. trivalents vs. monobrachial homologies) and this complexity may vary according to the type of chromosomes (i.e. acrocentric vs. metacentrics) involved in the rearrangements. Therefore, an individual analysis of each chromosome would be of primary interest to address the relative importance of each chromosomal configuration or to tackle broader issues such as the higher prevalence of some metacentric chromosomes.

A second limitation was that the markers used in this study have only been mapped at the chromosome arm level. However, as exemplified by Panithanarak *et al.* (2004) the impact of a rearrangement may affect differentially the regions of a single chromosome. Moreover, a precise localization on the chromosome arm is essential to differentiate between the two main classes of chromosomal speciation models (reviewed in Ayala & Coluzzi 2005, Butlin 2005). Indeed, the recombination rate is of primary importance for one category of these models (i.e. "suppressed recombination" models). But the outcome of a rearrangement on recombination may vary along the chromosome. Hence, the position of the markers within the chromosomes may help teasing apart the importance of the two classes of models.

Finally, only one category of markers has been used all along these studies: i.e. microsatellites. Although their properties (i.e. high polymorphism, abundance and codominance) make these markers some of the most popular in population genetics (Goldstein & Schlötterer 1999, Schlötterer 2004), they are maybe not the best suited to deal with any evolutionary levels and all situations. First of all, this category of marker is subjected to homoplasy (Estoup *et al.* 2002), hence it is expected that the genetic divergence estimated by these markers reaches a plateau if divergence time between taxa is long enough. In such situation it is thus likely that the difference between common and rearranged chromosomes will no more be detectable. Furthermore, the expected neutral properties of microsatellites make them inappropriate to detect differential levels of selective divergence between the two classes of chromosomes (e.g. Navarro & Barton 2003b). Although differences in divergence levels should apply to all sequences (neutral or not), neutral differences may be less marked and more difficult to detect. Therefore to account for as many situations as possible and to detect slight restrictions of gene exchanges, future speciation studies should combine the information of high-resolution markers such as microsatellites with other markers categories such as DNA-sequence polymorphisms (Schlötterer 2004).

To conclude, we have shown that chromosomal rearrangements undoubtedly favoured the genetic divergence and the establishment of reproductive barriers between karyotypic taxa of the *Sorex araneus* group. Although it is important to stress that genetic incompatibilities (e.g.

accumulated in allopatry during periods of geographic isolation) may also affect the differentiation of this group, this add to the recent data collected in a variety of plant or animal species to demonstrate the role of chromosome in speciation. However, many details of the process still need to be assessed. In such a context, the recent complete sequencing project of *Sorex araneus* (Chang *et al.* 2005), the type species of the *S. araneus* group, offers extraordinary perspectives and will further promote this group as an ideal model to study the detailed mechanisms of chromosomal speciation.

REFERENCES

Andersson AC, Narain Y, Tegelström H, Fredga K (2004) No apparent reduction of gene flow in a hybrid zone between the West and North European karyotypic groups of the common shrew, *Sorex araneus*. *Molecular Ecology*, **13**, 1205 – 1215.

Arnason U, Janke A (2002) Mitogenomics analyses of eutherian relationships. *Cytogenetic and Genome Research*, **96**, 20 – 32.

Arnold ML (1997) Natural hybridization and evolution. Oxford University Press, New York.

Ayala FJ, Coluzzi M (2005) Chromosome speciation: Humans, *Drosophila* and mosquitoes. *Proc. Natl. Acad. Sci. USA*, **102**, 6535 – 6542.

Baker RJ, Bickham JW (1986) Speciation by monobrachial centric fusions. *Proc. Natl. Acad. Sci. USA*, **83**, 8245 – 8248.

Balloux F, Ecoffey E, Fumagalli L, Goudet J, Wyttenbach A, 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. Brünner H, Lugon-Moulin N, Hausser J, Goudet J (2000) Microcatellite can be misleading: an empirical and simulation study. *Evolution*, **54**, 1414 – 1422.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, **11**, 155 – 165.

Banaszek A, Fedyk S, Szałaj KA, Chętnicki W (2000) A comparison of spermatogenesis in homozygotes, simple Robertsonian heterozygotes and complex heterozygotes of the common shrew (*Sorex araneus* L.). *Heredity*, **84**, 570 – 577.

Banaszek A, Fedyk S, Fiedorczuk U, Szałaj KA, Chętnicki W (2002) Meiotic studies of males common shrews (*Sorex araneus* L.) from a hybrid zone between chromosome races. *Cytogenetic and genome research*, **96**, 40 – 44.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual review of Ecology and Systematics*, **16**, 113 – 148.

Basset P, Hausser J (2003) Identification of three *Sorex* species with microsatellite markers. *Mammalia*, **68**, 245 – 252.

Basset P, Yannic G, Hausser J (in press) Genetic and karyotypic structure in the shrews of the *Sorex araneus* group: Are they independent? *Molecular Ecology*.

Belkir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2001) *Genetix 4.03, logiciel sousWindows TM pour la génétique des populations.* Laboratoire Génome,

Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Bengston BO, Frykman I (1990) Karyotype evolution: Evidence from the common shrew (*Sorex araneus* L.). *Journal of Evolutionary Biology*, **3**, 85 – 101.

Berry O, Tocher MD, Sarre SD (2004) Can assignment tests measure dispersal? *Molecular Ecology*, **13**, 551 – 561.

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, **5**, 453 – 455.

Brünner H, Fumagalli L, Hausser J (1994) A comparison of two contact zones between chromosomal races of *Sorex araneus* in the western Alps: Karyology and genetics. – Abstract in: *Folia Zoologica*, **43**, 113.

Brünner H, 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.

Brünner H, Lugon-Moulin N, Balloux F, Fumagalli L, Hausser J (2002a) A taxonomical re-evaluation of the Valais chromosome race of the common shrew *Sorex araneus* (Insectivora: Soricidae). *Acta Theriologica*, **47**, 245 – 275.

Brünner H, Lugon-Moulin N, Hausser J (2002b) Alps, genes, and chromosomes: their role in the formation of species in the *Sorex araneus* group (Mammalia, Insectivora), as inferred from two hybrid zones. *Cytogentic and Genome Research*, **96**, 85 – 96.

Butlin RK (2005) Recombination and speciation. *Molecular Ecology*, **14**, 2621 – 2635.

Castiglia R, Capanna E (2000) Contact zone between chromosomal races of *Mus musculus domesticus*. 2. Fertility and segregation in laboratory-reared and wild mice heterozygous for multiple Robertsonian rearrangements. *Heredity*, **85**, 147 – 156.

Cegelski CC, Waits LP, Anderson NJ (2003) Assessing population structure and gene flow in Montana wolverines (*Gulo gulo*) using assignment-based approaches. *Molecular Ecology*, **12**, 2907 – 2918.

Chakraborty R, De Andrade M, Daiger SP, Budowle B (1992) Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Annals of Human Genetics*, **56**, 45 – 47.

Chang J, Gnerre S, Wade C, Clamp M, Lindblad-Toh K (2005) Sequencing *Sorex araneus*: implication of a low coverage (2x) genome. Abstract in: Seventh meeting of the international *Sorex araneus* Cytogenetics Committee.

Chapuisat M (1998) Mating frequency of ant queens with alternative dispersal strategies, as revealed by microsatellite analysis of sperm. *Molecular Ecology*, **7**, 1097 – 1105.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, **153**, 1989 – 2000.

Coyne JA, Orr HA (1989) Patterns of speciation in *Drosophila*. *Evolution*, **43**, 362 – 381.

Coyne JA, Meyers W, Crittenden AP, Sniegowski P (1993) The effects of pericentric inversions in *Drosophila melanogaster*. *Genetics*, **134**, 487 – 496.

Coyne JA, Orr AH (2004) *Speciation*. Sinauer Associates, Sunderland, Mass. USA.

Crawley MJ (2002) *Statistical Computing: An introduction to data analysis using S-Plus.* John Wiley & Sons, Chichester, UK.

Davisson MT, Akeson EC (1993) Recombination suppression by heterozygous Robertsonian chromosomes in the Mouse. *Genetics*, **133**, 649 – 667.

Dobigny G, Granjon L, Aniskin V, Ba K, Volobouev V (2003) A new sibling species of *Taterillus* (Muridae, Gerbillinae) from West Africa. *Mammalian Biology*, **68**, 299 – 316.

Dod B, Jermiin LS, Boursot P, Chapman VH, Nielsen JT, Bonhomme F (1993) Counterselection on sex chromosomes in the *Mus musculus* European hybrid zone. *Journal of evolutionary Biology*, **6**, 529 – 546.

Doležel J, Kubalákova M, Bartoš J, Macas J (2004) Flow cytogenetics and plant genome mapping. *Chromosome research*, **12**, 77 – 91.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature reviews gentics*, **5**, 435 – 445.

Estoup A, Jarne PJ, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, **11**, 1591 – 1604.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611 – 2620.

Excoffier L (2001) Analysis of population subdivision. In: *Handbook of statistical Genetics* (eds Balding DJ, Bishop M, Cannings C), pp 271 – 307. John Wiley & Sons, Chichester, UK.

Fedyk S, Chętnicki W, Banaszek A (1991) Genetic differentiation of Polish populations of *Sorex araneus* L. III. Interchromosomal recombination in a hybrid zone. *Evolution*, **45**, 1384 – 1392.

Ferguson-Smith MA (1997) Genetic analysis by chromosome sorting and painting: Phylogenetic and diagnostic applications. *European Journal of Human Genetic*, **5**, 253 – 265.

Fivaz F, Basset P, Lugon-Moulin N, Hausser J (2003) Postglacial recolonization of the Valais (Switzerland) by the shrew *Sorex antinorii*: is dispersal sex-biased? A preliminary study. *Mammalia*, **68**, 253 – 262.

Fredga K, Narain Y (2000) The complex hybrid zone between the Abisko and Sidenjö chromosome races of the *Sorex araneus* in Sweden. *Biological Journal of the Linnean Society*, **70**, 285 – 307.

Fumagalli L, Hausser J, Taberlet P, Gielly L, Stewart DT (1996) Phylogenetic structure of the holarctic *Sorex araneus* group and its relationships with *S. samniticus*, as inferred from mtDNA sequences. *Hereditas*, **125**, 191 – 199.

Fumagalli L, Taberlet P, Stewart D, Gielly L, Hausser J, Vogel P (1999) Molecular phylogeny and evolution of *Sorex* shrews (Soricidae: Insectivora) inferred from mitochondrial DNA sequence data. *Molecular phylogenetics and Evolution*, **11**, 222 – 235.

Futuyma DJ (1986) *Evolutionary Biology, second edition*. Sinauer Associates, Sunderland, Mass. USA.

Garza JC, Slatkin M, Freimer NB (1995) Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. *Mol. Biol. Evol.*, **12**, 594 – 603.

Goldstein DB, Linares AR, Cavali-Sforza LL, Feldman MW (1995) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl, Acad. Sci. USA,* **92**, 6723 – 6727.

Goldstein DB, Pollock DD (1997) Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of Heredity*, **88**, 335 – 342.

Goldstein DB, Schlötterer C (1999) *Microsatellites: Evolution and applications*. Oxford Univ. Press, Oxford, UK.

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.

Grant V (1981) *Plant speciation* (second edition). Columbia University Press, New York, USA.

Haigis KM, Dove WF (2003) A Robertsonian translocation suppresses a somatic recombination pathway to loss of heterozygosity. *Nature Genetics*, **33**, 33 – 39.

Harrison RG (1990) Hybrid zones: windows on eloutionary process. In: *Oxford surveys in Evolutionary Biology* (eds. Futuyma D and Antonovics J), pp 69 – 128, Oxford University Press, Oxford, UK.

Hatfield T, Barton NH, Searle JB (1992) A model of a hybrid zone between two chromosomal races of the common shrew (*Sorex araneus*). *Evolution*, **46**, 1129 – 145.

Hausser J (1978) Répartition en Suisse et en France de *Sorex araneus* L., 1758 et de *Sorex coronatus* Millet, 1828 (Mammalia, Insectivora). *Mammalia*, **42**, 329 – 341.

Hausser J (1994) The *Sorex* of the *araneus-arcticus* group (Mammalia: Soricidae): do they actually speciate? *Carnegie Mus. Nat. Hist. Spec. Publ.*, **18**, 295 – 306.

Hausser J, Zuber N (1983) Détérmination spécifique d'individus vivants des deux espèces jumelles *Sorex araneus* et *S. coronatus*, par deux techniques biochimiques (Insectivora, Soricidae). *Revue Suisse de Zoologie*, **90**, 857 – 862.

Hausser J, Jammot D (1984) Etude biométrique des mâchoires chez les *Sorex* du groupe *araneus* en Europe continentale (Mammalia, Insectivora). *Mammalia*, **38**, 324 – 343.

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

Hewitt GM (1988) Hybrid zones – Natural laboratories for evolutionary studies. *Trends in Ecology and Evolution*, **3**, 158 – 167.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology*, **10**, 537 – 549.

Hey J (2003) Speciation and inversions: chimps and humans. *BioEssays*, **25**, 825 – 828.

Hoelzel AR, Dahlheim M, Stern SJ (1998) Low genetic variation among killer whales (*Orcinus orca*) in the eastern north Pacific and genetic differentiation between foraging specialists. *Journal of Heredity*, **89**, 121 – 128.

Howard DJ, Berlocher SH (eds) (1998) *Endless forms: Species and speciation*. Oxford University Press,New York, USA.

Ibrahim SF, van den Engh G (2004) High-speed chromosome sorting. *Chromosome research*, **12**, 5 – 14.

Jaquiéry J, Vogel V, Keller K (2005) Multilevel genetic analyses of two European supercolonies of the Argentine ant *Linepithema humile*. *Molecular Ecology*, **14**, 589 – 598.

Kalinowski ST (2002) Evolutionary and statistical properties of three genetic distances. *Molecular Ecology*, **11**, 1263 – 1273.

Kauppi L, Jeffreys AJ, Keeney S (2004) Where the crossovers are: recombination distributions in mammals. *Nature reviews genetics*, **5**, 413 – 424.

Kimura M, Ohta T (1978) Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proc. Nat. Acad. Sci. USA*, **75**, 2868 – 2872.

King M (1993) Species evolution, the role of chromosomal changes. Cambridge University Press, UK.

Lewiss H (1966) Speciation in flowering plants. *Science*, **152**, 167 – 172.

Lu J, Li WH, Wu CI (2003) Comments on "Chromosomal speciation and molecular divergence – Accelerated Evolution in rearranged chromosomes". *Science*, **302**, 988b.

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

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

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

Lugon-Moulin N, Balloux F, Hausser J (2000) Genetic differentiation of common shrew (*Sorex araneus*) populations among different alpine valleys revealed by microsatellites. *Acta Theriologica*, **45**, 103 – 117.

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

Lugon-Moulin N (2003) *Les musaraignes: Biologie, écologie, répartition en Suisse.* Editions Porte-Plumes, VS, Suisse.

Machado CA, Kliman RM, Markert JA, Hey J (2002) Inferring the history of speciation from multilocus DNA sequence Data: the case of *Drosophila pseudoobscura* and close relatives. *Molecular Biology and Evolution*, **19**, 472 – 488.

Manel S, Berthier P, Luikart G (2002) Detecting wildlife poaching: Identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. *Conservation Biology*, **16**, 650 – 659.

Manel S, Gaggiotti OE, Waples RS (2005) Assignement methods: matching biological questions with appropriate techniques. *Trends in Ecology and Evolution*, **20**, 136 – 142.

Marti DA, Bidau CJ (1995) Male and female meiosis in a natural population of *Dichroplus pratensis* (Arcrididae) polymorphic for Robertsonian translocations: A study of chiasma frequency and distribution. *Hereditas*, **123**, 227 – 235.

Merico V, Pigozzi MI, Esposito A, Merani MS, Garagna S (2003) Meiotic recombination and spermatogenesis impairment in *Mus musculus domesticus* carrying multiple simple Robertsonian translocations. *Cytogentic and genome research*, **103**, 321 – 329.

Meylan A, Hausser J (1973) Les chromosomes des *Sorex* du groupe *araneus-arcticus* (Mammalia, Insectivora). *Z. Säugetierkunde*, **38**, 143 – 158.

Michalakis Y, Excoffier L (1996) A genetic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*, **142**, 1061 – 1064.

Morgan-Richards M, Wallis GP (2003) A comparison of five hybrid zones of the weta *Hemideina thoracica* (Orthoptera: Anostostomatidae): Degree of cytogenetic differentiation fails to predict zone width. *Evolution*, **57**, 849 – 861.

Motokawa M (2004) Phylogenetic relationships within the family Talpidae (Mammalia: Insectivora). *Journal of Zoology*, **263**, 147 – 157.

Murphy WJ, Pevzner PA, O'Brien SJ (2004) Mammalian phylogenomics comes of age. *Trends in genetics*, **20**, 631 – 639.

Nachman MW (2001) Single nucleotide polymorphisms and recombination rate in humans. *Trends in Genetics*, **17**, 481 – 485.

Narain Y, Fredga K (1997) Meiosis and fertility in common shrews, *Sorex araneus*, from chromosomal hybrid zone in central Sweden. *Cytogenetics and cell genetics*, **78**, 253 – 259.

Narain Y, Fredga K (1998) Spermatogenesis in common shrews, *Sorex araneus*, from a hybrid zone with extensive Robertsonian polymorphism. *Cytogenetics and cell genetics*, **80**, 158 – 164.

Navarro A, Barton NH (2003a) Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution*, **57**, 447 – 459.

Navarro A, Barton NH (2003b) Chromosomal speciation and molecular divergence – Accelerated evolution in rearranged chromosomes. *Science*, **300**, 321 – 324.

Neet CR (1989) Ecologie comparée et biogéographie évolutive de deux espèces parapatriques : *Sorex araneus* et *Sorex coronatus* (Mammalia, Insectivora, Soricidae). Ph. D. thesis, University of Lausanne, Switzerland.

Neet CR, Hausser J (1990) Habitat selection in zones of parapatric contact between the common shrew *Sorex araneus* and Millet's shrew *S. coronatus*. *Journal of Animal Ecology*, **59**, 235 – 250.

Neet CR, Hausser J (1991) Biochemical analysis and determination of living individuals of the alpine karyotypic races and species of the *Sorex araneus* group. *Mémoire de la Société Vaudoise Sciences Naturelles*, **19**, 97 – 106.

Neet CR (1992) The use of fur colour characters to distinguish the sibling species *Sorex araneus* and *Sorex coronatus* (Insectivora, Soricidae) – a field test in a zone of parapatric contact. *International Journal of Mammalian Biology*, **56**, 176 – 178.

Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press, New-York, USA.

Noor MAF, Grams KL, Bertucci LA, Reiland J (2001a) Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. USA*, **98**, 12084 – 12088.

Noor MAF, Grams KL, Bertucci LA, Almendarez Y, Reiland J, Smith KR (2001b) The genetic of reproductive isolation and the potential for gene exchange between *Drosophila pseudoobscura* and *D. persimilis* via backcross hybrid males. *Evolution*, **55**, 512 – 521.

Nowak RM (1991) *Walker's mammals of the world, 5th ed*. John Hopkins University Press, Baltimore, MD, USA.

O'Brien SJ, Menotti-Raymond M, Murphy WJ, Nash WG, Wienberg J, Stanyon R, et al. (1999) The promise of comparative genomics in mammals. *Science*, **286**, 458 – 481.

O'Brien SJ, Eizirik E, Murphy WJ (2001) On choosing Mammalian genomes for sequencing. *Science*, **292**, 2264 – 2266.

Ortíz-Barrientos D, Reiland J, Hey J, Noor MAF (2002) Recombination and the divergence of hybridizing species. *Genetica*, **116**, 167 – 178.

Pack SD, Borodin PM, Serov OL, Searle JB (1993) The X-autosome translocation in the common shrew (*Sorex araneus* L.): late replication in female somatic cells and pairing in male meiosis. *Chromosoma*, **102**, 355 – 360.

Pack SD, Kolonin MG, Borodin PM, Searle JB, Serov OL (1995) Gene mapping in the common shrew (*Sorex araneus*; Insectivora) by shrew – rodent cell hybrids:

chromosome localization of the loci for HPRT, TK, LDHA, MDH1, G6PD, PGD, and ADA. *Mammalian genome*, **6**, 784 – 787.

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, **4**, 347 – 354.

Panithanarak T, Hauffe HC, Dallas JF, Glover A, Ward RG, Searle JB (2004) Linkage-dependent gene flow in a house mouse chromosomal hybrid zone. *Evolution*, **58**, 184 – 192.

Payseur BA, Krenz JG, Nachman MW (2004) Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution*, **58**, 2064 – 2078.

Payseur B, Nachman MW (2005) The genomics of speciation: investigating the molecular correlates of X chromosome introgression across the hybrid zone between *Mus domesticus* and *Mus musculus*. *Biological Journal of the Linnean Society*, **84**, 523 – 534.

Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Nonamplifying alleles at microsatellite loci-a caution for parentage and population studies. *Molecular Ecology*, **4**, 249 – 252.

Pennisi E (2004) More genomes, but shallower coverage. *Science*, **304**, 1227.

Pfunder M, Holzgang O, Frey E (2004) Development of microarray-based diagnostics of voles and shrews for use in biodiversity monitoring studies, and evaluation of mitochondrial cytochrome oxidase I vs. cytochrome b as genetic markers. *Molecular Ecology*, **13**, 1277 – 1286.

Piálek J, Hauffe HC, Rodríguez-Clark KM, Searle JB (2001) Raciation and speciation in house mice from the Alps: the role of chromosomes. *Molecular Ecology*, **10**, 613 – 625.

Polly PD (2003) Paleophylogeography of *Sorex araneus* (Insectivora, Soricidae): molar shape as a morphological marker for fossil shrews. *Mammalia*, **68**, 233 – 243.

Polyakov AV, Panov VV, Ladygina T.Y, Bochkarev MN, Rodionova MI, Borodin PM (2001) Chromosomal evolution of the Common shrew *Sorex araneus* L. from the Southern Urals and Siberia in the postglacial Period. *Russian Journal of Genetic*s, **37**, 351 – 357.

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

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proc. Natl. Acad. Sci. USA*, **4**, 9197 – 9201.

Ratkiewicz M, Fedyk S, Banaszek A, *et al.* (2002) The evolutionary history of the two karyotypic groups of the common shrew, *Sorex araneus*, in Poland. *Heredity*, **88**, 235 – 242.

Ratkiewicz M, Banaszek A, Jadwiszczak K, Chetnicki W, Fedyk S (2003) Genetic diversity, stability of population structure and barriers to gene flow in a hybrid zone between two *Sorex araneus* chromosome races. *Mammalia*, **67**, 275 – 283.

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

Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflowers species. *Genetics*, **152**, 713 – 727.

Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution*, **16**, 351 – 358.

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

Samonte RV, Eichler EE (2002) Segmental duplication and the evolution of the primate genome. *Nature Reviews Genetics*, **3**, 65 – 72.

Sargan DR, Yang F, Squire M, Milne BS, O'Brien PCM, Ferguson-Smith MA (2000) Use of flow-sorted canine chromosomes in the assignment of canine linkage, radiation hybrid, and syntenic groups to chromosomes: Refinement and verification of the comparative chromosome Map for dog and human. *Genomics*, **69**, 182 – 195.

Schlötterer C (2004) The evolution of molecular markers – just a matter of fashion? *Nature reviews genetics*, **5**, 63 – 69.

Schneider S, Roessli D, Excoffier L (2000) Arlequin: a software for population genetics data analysis. Genetics and Biometry Lab, Department of Anthropology, University of Geneva.

Searle JB, Hübner R, Wallace BMN, Garagna S (1990) Robertsonian variation in wild mice and shrews. *Chromosomes Today*, **10**, 253 – 263.

Searle JB, Fedyk S, Fredga K, Hausser J, Volobouev VT (1991) Nomenclature for the chromosomes of the common shrew *Sorex araneus*. *Mémoire de la Société Vaudoise en Sciences Naturelles*, **19**, 13 – 22.

Searle JB (1993) Chromosomal hybrid zones in eutherian mammals. In: *Hybrids Zones and Evolutionary Process* (ed Harrison RG), pp. 309 – 353. Oxford University Press, New York, USA.

Searle JB, Wójcik JM (1998) Chromosomal evolution: the case of *Sorex araneus*. In Evolution of shrews (eds Wójcik JM and Wolsan M), pp 219 – 262. Mammal Research Institute, Polish Academy of Sciences, Białowieza, Poland.

Sharman GB (1956) Chromosome of the common shrew. *Nature*, **177**, 941 – 942.

Sites JW Jr, Moritz C (1987) Chromosomal evolution and speciation revisited. *Syst. Zool.*, **36**, 153 – 174.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457- 462.

Spirito F (2000) The role of chromosomal rearrangements in speciation. In *Endless Forms* (eds Howard DJ and Berlocher SH), pp 320 – 329. Oxford University Press, New York, USA.

Stanley WT, Rogers MA, Hutterer R (2005) A new species of *Congosorex* from the Eastern arc mountains, Tanzania, with significant biogeographical implications. *Journal of Zoology*, **265**, 269 – 280.

Stump AD, Fitzpatrick MC, Lobo NF, Traoré S, Sagnon NF, Costantini C *et al.* (2005) Centromere-proximal differentiation and speciation in *Anopheles gambiae*. *Proc. Natl. Acad. Sci. USA*, **102**, 15930 – 15935.

Taberlet P, Fumagalli L, 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.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453 – 464.

Telenius H, Pelmear AH, Tunnacliffe A, Carter NP, Behmel A, Ferguson-Smith MA, *et al.* (1992) Cytogenetic analysis by chromosome painting using DOP PCR amplified flow sorted chromosomes. *Genes Chromosomes Cancer*, **4**, 257 – 263.

Thomas EE, Srebro N, Sebat J, Navin N, Healy J, Mishra B, *et al.* (2004) Distribution of short paired duplications in mammalian genomes. *Proc. Natl. Acad. Sci. USA*, **101**, 10349 – 10354.

Tucker PK, Lee BK, Lundrigan BL, and Eicher EM (1992) Geographic origin of the Y chromosomes in “old” inbred strains of mice. *Mammalian Genome* **3**: 254 – 261.

Vähä JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, **15**, 63 – 72.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKERS: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535 – 538.

Vázquez-Dominguez E, Paetkau D, Tucker N, Hintens G, Moritz C (2001) Resolution of natural groups using iterative assignment tests: an example from two species of Australian native rats (*Rattus*). *Molecular Ecology*, **10**, 2069 – 2078.

Volobouev VT (1989) Phylogenetic-relationships of the *Sorex araneus-arcticus* species complex (Insectivora, Soricidae) based on high-resolution chromosome analysis. *Journal of Heredity*, **80**, 284 – 290.

Volobouev VT, Catzeflis F (1989) Mechanisms of chromosomal evolution in three European species of the *Sorex araneus-arcticus* group (Insectivora: Soricidae). *Z. zool. Syst. Evolut.-forsch.*, **27**, 252 – 262.

Volobouev VT, Dutrillaux B (1991) Chromosomal evolution and phylogenetic relationships of the *Sorex araneus-arcticus* group. *Mémoire de la Société Vaudoise Sciences Naturelles*, **19**, 131 – 139.

Wallace B (1959) The influence of genetic systems on geographical distribution. *Cold Spring Harb. Symp. Quant. Biol.*, **20**, 16 – 24.

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

Weir BS (1996) *Genetic data analysis II: Methods for Discrete Population Genetic Data*. Sinauer Associates, Inc., Sunderland, MA, USA.

White MJD (1978) *Modes of Speciation*. WH Freeman & Co, San Francisco, USA.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, **163**, 1177 – 1191.

Wójcik JM, Searle JB (1988) The chromosome complement of *Sorex granarius* – the ancestral karyotype of the common shrew (*Sorex araneus*)? *Heredity*, **61**, 225 – 229.

Wójcik JM, Wójcik AM, Zalewska H, Rychlik L (1996) Allozyme differentiation in four sympatric species of European shrews (Soricidae: Mammalia). *Biochemical Systematics and Ecology*, **24**, 291 – 298.

Wójcik JM, Bogdanowicz W, Pucek Z, Wójcik AM, Zalewska H (2000) Morphometric variation of the common shrew *Sorex araneus* in Poland, in relation to karyotype. *Acta theriologica*, **45**, 161 – 172.

Wójcik JM, Ratkiewicz M, Searle JB (2002) Evolution of the common shrew *Sorex araneus*: chromosomal and molecular aspects. *Acta Theriologica*, **47**, 139 – 167.

Wójcik JM, Borodin PM, Fedyk S, Fredga K, Hausser J, Mishta A, *et al.* (2003) The list of the chromosome races of the common shrew *Sorex araneus* (updated 2002). *Mammalia*, **67**, 169 – 178.

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, Hausser J (1997) Isolation and characterization of simple sequence repeats in the genome of the common shrew. *Molecular Ecology*, **6**, 797 – 800.

Wyttenbach A, Narain Y, Fredga K (1999) 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.

Yamaguchi N, Driscoll CA, Kitchener AC, Ward JM, Macdonald DW (2004) Craniological differentiation between European wildcats (*Felis silvestris silvestris*), African wildcats (*F. s. lybica*) and Asian wildcats (*F. s. ornata*): implication for their evolution and conservation. *Biological journal of the Linnean Society*, **83**, 47 – 63.

Yang F, Carter NP, Shi L, Ferguson-Smith MA (1995) A comparative study of karyotypes of muntjacs by chromosome painting. *Chromosoma*, **103**, 642 – 652.

Yang F, O'Brien PCM, Milne BS, Graphodatsky AS, Solanky N, Trifonov *et al.* (1999) A Complete Comparative chromosome map for the dog, red fox, and human and its integration with canine genetic maps. *Genomics*, **62**, 189 – 202.

Yeom YI, Abe K, Bennett D, Artzt K (1992) Testis-/embryo-expressed genes are clustered in the mouse H-2K egion. *Proc. Natl. Acad. Sci. USA*, **89**, 773 -777.

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **12**, 563 – 584.

Zhang J, Wang X, Podlaha O (2004) Testing the chromosomal speciation hypothesis for humans and chimpanzees. *Genome research*, **14**, 845 – 851.

Zhdanova NS, Fokina VM, Balloux F, Hausser J, Borodin PM, Volobouev VT, *et al.* (2003) Current cytogenetic map of the common shrew, *Sorex araneus* L.: localization of 7 genes and 4 microsatellites. *Mammalia*, **68**, 285 – 293.

Zhdanova NS, Karamisheva TV, Minina J, Astakhova NM, Lansdorp P, Kammori M, *et al.* (2005) Unusual distribution pattern of telomeric repeats in the shrews *Sorex araneus* and *Sorex granarius*. *Chromosome research*, **13**, 617 – 625.

Zima J, Lukáčová L, Macholán M (1998) Chromosomal evolution in shrews (eds Wójcik JM and Wolsan M), pp 175 – 218. Mammal Research Institute, Polish Academy of Sciences, Białowieza, Poland.

APENDIXES

Appendix 1: Hierarchical genetic structure and genetic distance ($(\delta\mu)^2$) per locus, across Common (C), Rearranged (R) and across all loci at the intra-taxon and inter-taxa levels for each of the 10 comparisons analysed in chapter 4.

			1. *S. a.* Vaud - *S. a.* Cordon			
			Intra-taxon		**Inter-taxa**	
Group	**Chrom.**	**Locus**	F_{SR}	$(\delta\mu)^2$	F_{RT}	$(\delta\mu)^2$
C	***a***	**L16**	0.038 NS	0.759	0.023 ***	0.213
C	***a***	**D107**	0.032 *	8.045	-0.002 NS	3.351
C	***f***	**L69**	0.011 NS	0.856	0.031 ***	1.404
C	***f***	**B3**	0.094 ***	10.218	-0.038 NS	1.130
C	***b***	**L68**	0.033 NS	0.124	0.149 ***	0.053
C	***b***	**C117**	0.062 *	6.440	0.017 ***	46.814
C	***c***	**L9**	0.018 NS	0.827	0.045 ***	7.448
C	***de***	**C171**	-0.007 NS	2.500	0.039 ***	9.238
C	***de***	**L13**	0.468 **	0.260	-0.074 NS	0.474
C	***de***	**L57**	0.120 ***	2.888	0.075 ***	63.654
R	***g***	**L62**	0.091 ***	4.753	-0.001 NS	1.397
R	***h***	**D106**	0.043 *	9.374	0.066 ***	1.185
R	***jl***	**D24**	0.050 **	68.641	0.009 ***	2.195
R	***n***	**L99**	-0.033 NS	0.002	0.154 ***	0.015
R	***o***	**B30**	0.029 NS	3.791	-0.019 NS	0.324
R	***o***	**D109**	0.034 **	18.143	-0.011 NS	4.263
		Common	0.065 ***	3.291	0.032 ***	13.378
		Rearranged	0.047 ***	17.451	0.018 ***	1.563
		Test permut.	P = 0.658	P = 0.056	P = 0.615	P = 0.232
		Overall	0.059 ***	8.601	0.027 ***	8.947

			2. *S. a.* Vaud - *S. a.* Białowieza			
			Intra-taxon		**Inter-taxa**	
Group	**Chrom.**	**Locus**	F_{SR}	$(\delta\mu)^2$	F_{RT}	$(\delta\mu)^2$
C	***a***	**L16**	0.004 NS	0.331	0.095 ***	0.279
C	***a***	**D107**	0.026 *	2.637	0.011 ***	117.906
C	***f***	**L69**	0.005 NS	0.590	0.016 ***	16.108
C	***f***	**B3**	0.068 ***	9.826	0.003 ***	12.262
C	***b***	**L68**	0.030 *	0.120	0.057 ***	7.726
C	***b***	**C117**	0.033 NS	3.645	0.078 ***	13.116
C	***c***	**L9**	0.021 *	1.980	0.048 ***	3.139
C	***de***	**C171**	0.004 NS	4.282	0.018 ***	0.223
C	***de***	**L13**	0.259 **	0.106	-0.138 *	0.002
C	***de***	**L57**	0.042 *	2.444	0.040 ***	2.870
R	***g***	**L62**	0.040 **	2.801	0.033 ***	1.754
R	***h***	**D106**	0.008 NS	0.843	0.038 ***	41.422
C	***jl***	**D24**	0.045 ***	66.141	-0.006 NS	17.073
R	***n***	**L99**	0.003 NS	0.001	0.012 ***	0.001
R	***o***	**B30**	0.042 NS	2.360	0.079 ***	12.345
R	***o***	**D109**	0.024 **	5.127	0.010 ***	81.693
		Common	0.042 ***	8.373	0.025 ***	17.337
		Rearranged	0.028 ***	2.226	0.035 ***	27.443
		Test permut.	P = 0.643	P = 0.625	P = 0.633	P = 0.640
		Overall	0.038 ***	6.452	0.028 ***	20.495

			3. *S. a.* Cordon - *S. a.* Białowieza			
			Intra-taxon		**Inter-taxa**	
Group	**Chrom.**	**Locus**	F_{SR}	$(\delta\mu)^2$	F_{RT}	$(\delta\mu)^2$
C	***a***	**L16**	0.035 NS	0.988	0.167 ***	0.005
C	***a***	**D107**	-0.006 NS	7.094	0.032 ***	81.503
C	***f***	**L69**	-0.010 NS	0.538	0.030 ***	8.002
C	***f***	**B3**	0.023 NS	2.348	-0.003 NS	5.949
C	***b***	**L68**	0.030 *	0.004	0.153 ***	6.504
C	***b***	**C117**	0.027 NS	3.500	0.072 ***	109.489
C	***c***	**L9**	-0.002 NS	1.510	0.042 ***	20.257
C	***de***	**C171**	0.015 NS	1.782	0.007 ***	6.592
C	***de***	**L13**	-0.008 NS	0.167	0.167 ***	0.543
C	***de***	**L57**	0.077 **	0.691	0.110 ***	93.557
R	***g***	**L62**	0.058 **	4.110	0.005 ***	6.2808
R	***h***	**D106**	0.039 *	9.751	0.086 ***	56.621
C	***jl***	**D24**	0.026 NS	7.347	-0.001 ***	31.512
R	***n***	**L99**	-0.026 NS	0.000	0.207 ***	0.022
C	***o***	**B30**	0.080 **	2.090	0.013 ***	8.670
C	***o***	**D109**	0.000NS	14.1558	0.014 ***	48.633
		Common	0.024 NS	3.247	0.051 ***	32.401
		Rearranged	0.027 NS	4.622	0.111 ***	20.975
		Test permut.	P = 0.315	P = 0.639	P = 0.864	P = 0.692
		Overall	0.025 NS	3.505	0.056 ***	30.259

			4. *S. antinorii* - *S. a.* Vaud			
			Intra-taxon		**Inter-taxa**	
Group	**Chrom.**	**Locus**	F_{SR}	$(\delta\mu)^2$	F_{RT}	$(\delta\mu)^2$
C	***a***	**L16**	0.120 *	0.288	0.188 ***	0.100
C	***a***	**D107**	0.120 ***	42.805	0.023 ***	77.448
C	***f***	**L69**	0.046 ***	0.919	0.012 ***	0.251
C	***f***	**B3**	0.071 ***	10.299	-0.017 NS	9.317
C	***b***	**L68**	0.028 *	1.205	0.077 ***	0.135
C	***b***	**C117**	0.058 **	4.300	0.077***	46.879
C	***c***	**L9**	0.038 **	10.251	0.068 ***	6.469
C	***de***	**C171**	0.073 ***	127.011	0.039 ***	34.191
C	***de***	**L13**	0.336 ***	0.101	0.119 ***	0.092
C	***de***	**L57**	0.068 ***	15.934	0.018 ***	0.780
R	***g***	**L62**	0.094 ***	3.305	0.023 ***	0.405
R	***h***	**D106**	0.067 **	6.211	0.203 ***	53.759
R	***jl***	**D24**	0.066 ***	90.955	0.035 ***	18.787
R	***n***	**L99**	0.025 NS	0.015	0.796 ***	0.858
R	***o***	**B30**	0.066 *	19.909	0.335 ***	61.914
R	***o***	**D109**	0.052 ***	23.714	0.020 ***	90.997
		Common	0.076 ***	21.311	0.050 ***	17.566
		Rearranged	0.067 ***	24.018	0.230 ***	37.787
		Test permut.	p = 0.6774	P = 0.909	p = 0.064	P = 0.224
		Overall	0.0729 ***	22.326	0.122 ***	25.149

			5. ***S. antinorii*** **-** ***S. a.*** **Cordon**			
			Intra-taxon		**Inter-taxa**	
Group	**Chrom.**	**Locus**	F_{SR}	$(\delta\mu)^2$	F_{RT}	$(\delta\mu)^2$
C	***a***	**L16**	0.146 ***	0.944	0.238 ***	0.605
C	***a***	**D107**	0.132 ***	47.261	-0.009 NS	48.579
C	***f***	**L69**	0.047 **	0.868	0.048 ***	2.843
C	***f***	**B3**	0.046 **	2.821	0.008 ***	3.959
C	***b***	**L68**	0.029 NS	1.089	0.181 ***	0.019
C	***b***	**C117**	0.059 **	4.154	-0.009 NS	0.000
C	***c***	**L9**	0.026 **	9.778	0.018 ***	0.035
C	***de***	**C171**	0.102 ***	124.511	-0.002 NS	78.975
C	***de***	**L13**	0.038 NS	0.162	-0.025 NS	0.149
C	***de***	**L57**	0.098 ***	14.182	0.085 ***	78.525
R	***g***	**L62**	0.115 ***	4.618	0.011 ***	0.298
R	***h***	**D106**	0.098 ***	15.120	0.112 ***	70.910
R	***jl***	**D24**	0.059 **	32.160	0.033 ***	8.139
R	***n***	**L99**	0.007 NS	0.014	0.631 ***	1.096
R	***o***	**B30**	0.089 **	19.640	0.272 ***	71.194
R	***o***	**D109**	0.045 ***	32.743	0.009 ***	134.650
		Common	0.073 ***	20.577	0.058 ***	21.369
		Rearranged	0.075 ***	17.382	0.171***	47.714
		Test permut.	P = 0.944	P = 0.923	P = 0.2351	P = 0.289
		Overall	0.074 ***	19.379	0.103 ***	31.248

			6. ***S. antinorii*** **-** ***S. a.*** **Białowieza**			
			Intra-taxon		**Inter-taxa**	
Group	**Chrom.**	**Locus**	F_{SR}	$(\delta\mu)^2$	F_{RT}	$(\delta\mu)^2$
C	***a***	**L16**	0.124 **	0.517	0.026 ***	0.714
C	***a***	**D107**	0.108 ***	41.854	0.002 ***	4.235
C	***f***	**L69**	0.037 **	0.602	0.038 ***	20.383
C	***f***	**B3**	0.033 *	2.429	0.042 ***	0.202
C	***b***	**L68**	0.027 *	1.086	0.027 ***	5.820
C	***b***	**C117**	0.037 *	1.358	0.155 ***	109.589
C	***c***	**L9**	0.027 **	10.931	0.027 ***	18.620
C	***de***	**C171**	0.091 ***	126.294	0.003 ***	39.933
C	***de***	**L13**	0.002 NS	0.008	0.175 ***	0.123
C	***de***	**L57**	0.044 ***	13.738	0.050 ***	0.658
R	***g***	**L62**	0.076 ***	2.666	0.029 ***	3.844
R	***h***	**D106**	0.062 ***	6.589	0.176 ***	0.803
R	***jl***	**D24**	0.054 **	29.660	0.030 ***	71.679
R	***n***	**L99**	0.039 NS	0.014	0.817 ***	0.810
R	***o***	**B30**	0.088 ***	18.209	0.281 ***	129.551
R	***o***	**D109**	0.036 ***	19.727	0.018 ***	345.129
		Common	0.053 ***	19.882	0.047 ***	20.028
		Rearranged	0.061 ***	12.811	0.218 ***	91.969
		Test permut.	P = 0.661	P = 0.842	P = 0.074	P = 0.072
		Overall	0.056 ***	17.230	0.117 ***	47.006

Group	Chrom.	Locus	7. *S. coronatus* - *S. a.* Vaud			
			Intra-taxon		Inter-taxa	
			F_{SR}	$(\delta\mu)^2$	F_{RT}	$(\delta\mu)^2$
C	*a*	L16	0.038NS	0.102	0.325 ***	1.108
C	*a*	D107	0.047 **	3.588	-0.014 NS	13.380
C	*f*	L69	0.021 NS	0.907	0.107 ***	102.022
C	*f*	B3	0.130 ***	17.696	0.131 ***	100.729
R	*b*	L68	0.037 NS	0.240	0.087 ***	0.754
R	*b*	C117	0.062 NS	6.587	0.113 ***	40.729
R	*c*	L9	0.038 NS	1.300	0.131 ***	29.674
C	*de*	C171	-0.005 NS	4.999	0.285 ***	30.073
C	*de*	L13	0.383 ***	0.199	-0.065 NS	155.265
C	*de*	L57	0.074 *	4.641	-0.047 NS	7.814
R	*g*	L62	0.064 **	3.440	-0.003 NS	19.546
R	*h*	D106	0.012 NS	0.465	0.054 ***	0.709
R	*jl*	D24	0.065 ***	127.436	0.053 ***	541.515
R	*n*	L99	-0.019 NS	0.003	0.928 ***	0.704
R	*o*	B30	0.020 NS	4.060	0.236 ***	28.690
R	*o*	D109	0.047 ***	9.114	0.012 ***	134.072
		Common	0.092 ***	4.590	0.101 ***	58.627
		Rearranged	0.045 ***	16.961	0.182 ***	88.488
		Test permut.	p = 0.252	P = 0.817	P = 0.749	P = 0.878
		Overall	0.066 ***	11.549	0.147 ***	75.424

Group	Chrom.	Locus	8. *S. coronatus* - *S. a.* Cordon			
			Intra-taxon		Inter-taxa	
			F_{SR}	$(\delta\mu)^2$	F_{RT}	$(\delta\mu)^2$
C	*a*	L16	0.152 NS	1.415	0.348 ***	2.293
C	*a*	D107	-0.023 NS	12.501	0.026 ***	3.339
C	*f*	L69	-0.004 NS	0.804	0.109 ***	79.491
C	*f*	B3	0.064 *	2.740	0.168 ***	123.191
R	*b*	L68	0.023 NS	0.007	0.025 ***	0.408
R	*b*	C117	0.074	6.293	0.012 ***	0.212
R	*c*	L9	-0.017 NS	0.354	0.163 ***	66.855
C	*de*	C171	0.001 NS	0.000	0.245 ***	5.975
C	*de*	L13	-0.038 NS	0.321	0.425 ***	138.584
C	*de*	L57	0.195 ***	1.135	-0.017 NS	26.864
R	*g*	L62	0.138 **	6.066	-0.086 NS	10.493
C	*h*	D106	0.097 *	18.283	-0.053 NS	3.728
R	*jl*	D24	0.028 NS	9.847	0.091 ***	474.756
R	*n*	L99	-0.034 NS	0.001	0.710 ***	0.921
R	*o*	B30	0.115 NS	3.521	0.116 ***	22.918
R	*o*	D109	0.007 NS	27.172	0.032 ***	186.148
		Common	0.064 *	4.650	0.132 ***	47.933
		Rearranged	0.047 *	6.658	0.145 ***	95.339
		Test permut.	P =0.677	P = 0.641	P = 0.939	P = 0.574
		Overall	0.055 *	5.654	0.139 ***	71.636

			9. ***S. coronatus*** **-** ***S. a.*** **Białowieza**			
			Intra-taxon		**Inter-taxa**	
Group	**Chrom.**	**Locus**	F_{SR}	$(\delta\mu)^2$	F_{RT}	$(\delta\mu)^2$
C	***a***	**L16**	0.002 NS	0.560	0.156 ***	2.500
C	***a***	**D107**	0.000 NS	1.686	0.021 ***	51.848
C	***f***	**L69**	-0.011 NS	0.273	0.107 ***	37.053
C	***f***	**B3**	0.012 NS	1.956	0.226 ***	183.282
R	***b***	**L68**	0.036 *	0.000	0.095 ***	3.652
R	***b***	**C117**	0.004 NS	0.703	0.113 ***	100.071
R	***c***	**L9**	0.009 NS	2.660	0.109 ***	13.511
C	***de***	**C171**	0.030 NS	3.565	0.294 ***	25.120
C	***de***	**L13**	-0.007 NS	0.013	0.194 ***	156.481
C	***de***	**L57**	0.001 NS	0.247	0.054 ***	20.156
C	***g***	**L62**	0.010 NS	2.163	0.032 ***	33.009
R	***h***	**D106**	0.001 NS	1.220	0.076 ***	31.292
R	***jl***	**D24**	0.029 *	4.847	0.079 ***	750.892
R	***n***	**L99**	-0.028 NS	0.000	0.964 ***	0.660
R	***o***	**B30**	0.113 **	0.659	0.001 ***	3.396
R	***o***	**D109**	-0.003 NS	1.140	0.044 ***	425.075
		Common	0.005 NS	1.308	0.132 ***	63.681
		Rearranged	0.021 *	1.403	0.200 ***	166.069
		Test permut.	P = 0.180	P = 0.613	P = 0.882	P = 0.404
		Overall	0.013 *	1.356	0.167 ***	114.875

			10. ***S. coronatus*** **-** ***S. antinorii***			
			Intra-taxon		**Inter-taxa**	
Group	**Chrom.**	**Locus**	F_{SR}	$(\delta\mu)^2$	F_{RT}	$(\delta\mu)^2$
C	***a***	**L16**	0.267 ***	0.474	-0.059 NS	0.542
C	***a***	**D107**	0.154 ***	82.021	-0.061 NS	26.446
C	***f***	**L69**	0.064 ***	0.932	0.121 ***	112.399
C	***f***	**B3**	0.052 ***	2.901	0.190 ***	171.316
R	***b***	**L68**	0.030 NS	2.171	0.178 ***	0.251
R	***b***	**C117**	0.057 *	2.014	0.125 ***	0.216
R	***c***	**L9**	0.041 **	19.202	0.094 ***	63.852
C	***de***	**C171**	0.140 ***	249.023	0.201 ***	128.397
C	***de***	**L13**	0.001 NS	0.003	0.440 ***	147.814
C	***de***	**L57**	0.064 ***	27.228	0.013 ***	13.531
R	***g***	**L62**	0.106 ***	3.169	-0.008 NS	14.324
R	***h***	**D106**	0.090 **	11.957	0.238 ***	42.120
R	***jl***	**D24**	0.071 *	54.474	0.087 ***	358.574
R	***n***	**L99**	0.036 NS	0.028	0.018 ***	0.008
C	***o***	**B30**	0.115 ***	35.758	0.429 ***	174.898
C	***o***	**D109**	0.055 ***	38.314	-0.004 NS	4.160
		Common	0.096 ***	48.517	0.143 ***	86.612
		Rearranged	0.065 ***	13.288	0.113 ***	68.478
		Test permut.	P = 0.184	P = 0.326	P =0.713	P = 0.756
		Overall	0.082 ***	33.104	0.130 ***	78.678