The evolutionary history of the two karyotypic groups of the common shrew, *Sorex araneus*, in Poland

M Ratkiewicz¹, S Fedyk¹, A Banaszek¹, L Gielly², W Chętnicki¹, K Jadwiszczak¹ and P Taberlet²

¹University of Białystok, Institute of Biology, Świerkowa 20 B, 15–950 Białystok, Poland; ²Laboratoire de Biologie des Populations d'Altitude, Université Joseph Fourier, CNRS, UMR 5553, BP 53 38 041 Grenoble, Cedex 9 France

Genetic variability within and among two karyotypic groups and five chromosome races of the common shrew (*Sorex araneus*) in Poland was assayed by sequencing a 1023 bp part of the cytochrome *b* gene (mtDNA) from 28 individuals. Thirty-four variable positions defined 21 distinct haplotypes with a maximum sequence divergence of 0.88%. No significant differentiation in the cytochrome *b* gene between Western and Eastern Karyotypic groups was found. Haplotype diversity estimates within the races and groups sampled were high ($h = 0.800–0.928$), while nucleotide diversity estimates were low ($\pi = 0.0034–0.0053$). The distribution of pairwise nucleotide differences fits well with expectations of

a 'sudden expansion' model. High haplotype diversity was accompanied by relatively high expected heterozygosity (H_E) values in nuclear genes (calculated over 47 enzyme loci: $H_E = 0.031–0.049$), giving no evidence for a recent bottleneck after the process of post-Pleistocene recolonization of Poland by the shrews. Thus, for *S. araneus* chromosome races in Poland, the data on the cytochrome *b* gene variability support the hypothesis assuming the Robertsonian fusions having spread into an ancestral acrocentric distribution.

Heredity (2002) 88, 235–242. DOI: 10.1038/sj/hdy/6800032

Keywords: common shrew; *Sorex araneus*; cytochrome *b*; chromosome races; evolutionary history; chromosomal evolution

Introduction

The enormous karyotype variation in the common shrew, *Sorex araneus* has been extensively studied for several decades. So far, more than 50 chromosome races have been described (Zima *et al*, 1996; Bulatova *et al*, 2000; Mishta *et al*, 2000; Polyakov *et al*, 2000) and the list of the *S. araneus* races is probably still not completed. The chromosome races of the *S. araneus* species are characterized by different sets of acrocentrics and metacentrics. As the ancestral karyotype for *S. araneus* was most probably acrocentric (Wójcik and Searle, 1988), Robertsonian (centric) fusions leading to metacentric state must have occurred and become fixed. The conditions under which a fixation of new Rb fusion could have occurred, can be summarized in two most probable models. First is the chain processes variant of White's (1978) stasipatric model of chromosomal evolution. The distribution of metacentrics and chromosome races in Poland fits White's stasipatric model remarkably well (Wójcik, 1993). Different metacentrics (arising from Rb fusions) apparently have spread different distances into an ancestral acrocentric distribution (Wójcik, 1993). Two factors give evidence for this model of local fixation of Rb metacentrics in the common shrew. First, simple Rb heterozygotes in the common shrew show little disadvantage in comparison to homozygotes (Searle, 1993). Second, Wytenbach

et al (1998) showed meiotic drive in favour of some metacentrics in the male common shrews.

On the other hand, the simplest model for local fixation is by genetic drift only. At present, no evidence for genetic drift comes from allozyme (Bengtsson and Frykman, 1990) or microsatellites (Wytenbach and Hausser, 1996) in studies of *S. araneus*. However, the conditions for local fixation of metacentrics could have been suitable in the past. At the end of glaciation, small populations were probably formed at the leading edge of expansion by constant range changes. Such bottlenecked populations could have been the sites of fixation for new chromosomal variants.

The possibility of determining past population sizes comes from molecular data collected from present-day populations. In our study we used cytochrome *b* gene sequence data of shrews from five chromosomal races in Poland that belong to two karyotypically distant groups: the West European Karyological Group (WEKG) and the East European Karyological Group (EEKG), see Searle (1984), Wójcik (1993), Searle and Wójcik (1998). The WEKG races in Poland are characterized by (at maximum) five different Rb metacentrics. While within the EEKG races as many as nine different fusions must have occurred (Wójcik, 1993; Fedyk, 1995). Thus, it cannot be ruled out that at least some metacentrics within EEKG could have arisen by whole arm reciprocal translocations (WARTs, see Hauffe and Pialek, 1997). It seems that WARTs may be fixed in a bottlenecked populations only, as the fertility of heterozygotes for a WART (forming at least four element meiotic complexes) is almost certainly lower than Rb heterozygotes (forming

Correspondence: M Ratkiewicz, University of Białystok, Institute of Biology, Świerkowa 20 B, 15–950 Białystok, Poland.

E-mail: ermi@cksr.ac.bialystok.pl

Received 19 March 2001; accepted 24 October 2001

meiotic trivalents) (Searle and Wójcik, 1998). Thus, the lack of a severe bottleneck in the common shrew populations would be an indirect argument for Rb fusions, as a main type of chromosomal mutation type in *S. araneus* from Poland and give support for Wójcik's (1993) model of fixation of new Rb metacentrics in large continuous populations. In this paper, we determine past population sizes for *S. araneus*, making use of molecular data collected from present-day populations. We establish the long-term effective female population size and investigate whether chromosome races represented by Polish populations of *S. araneus* have experienced population bottlenecks at the time of fixation of the chromosomal variants that define them. We also include results of an allozyme study based on 47 protein loci in the common shrew populations from WEKG and EEKG in Poland, to give a wider picture of genetic variation in this area. Finally, we compare results from allozymes with the cytochrome *b* gene analyses.

Materials and methods

Tissue samples were collected in 1993–1994 years from 10 different *S. araneus* populations (Figure 1). The sampling localities were (abbreviations, coordinates, sample size for cytochrome *b* gene analyses): Blizocin (BLI, 51°36'N 22°10'E, *n* = 5), Jurowce (JUR, 53°11'N 23°07'E, *n* = 6), Ulaski (ULA, 52°52'N 21°06'E, *n* = 1), Zubrówka (ZUB, 54°05'N 23°10'E, *n* = 2), Ramoty (RAM, 53°57'N 19°14'E, *n* = 2), Kępki (KEP, 54°12'N 19°18'E, *n* = 1), Lubstowo (LUB, 54°08'N 19°12'E, *n* = 1), Jabłonowiec (JAB, 51°41'N 21°44'E, *n* = 6), Rokitnia (ROK, 51°35'N 21°46'E, *n* = 2) and Łutynowo (LUT, 53°34'N 20°19'E, *n* = 2). Sample sizes for the electrophoretic study ranged from 12 to 42 individuals (219 individuals all together). Allozymes were not studied in the JAB population. Sampling

localities BLI, JUR and ULA were located within Białowieża chromosome race (EEKG), ZUB within Goldap race (EEKG), RAM within Łęgucki Młyn race (EEKG), KEP and LUB within Nogat race (WEKG), while JAB, ROK and LUT within Drnholec race (WEKG).

Chromosome preparations were done from spleen cells and stained for G-bands with Giemsa reagent after treatment with trypsin (Seabright, 1971). Chromosome arms were labelled according to the nomenclature proposed by Searle *et al* (1991).

DNA isolation and amplification

Total DNA was extracted from dried toes using Qiagen DNAeasy Tissue Kit following the manufacturer's instructions. Amplification took place in a 25 µl volume containing 50 µM of dNTP each, 2 mM MgCl₂, PCR buffer (Qiagen), 1 µM primers (both), 1 U Taq DNA polymerase and 2 µl DNA template per tube in a Perkin Elmer thermal cycler using the following profile: initial denaturation for 3 min at 93°C, denaturation for 30 s at 93°C, annealing for 30 s at 50°C and elongation for 3 min at 72°C. The primers used to amplify the cytochrome *b* gene fragment were: 'cytb L14841' and 'cytb H15915' (Irwin *et al*, 1991). Amplified DNA was cleaned using the QIAquick PCR Purification Kit (Qiagen) following the manufacturer's instructions. PCR products were sequenced on an ABI 377 automated sequencer (using PRISM™ Ready Reaction DyeDeoxy Terminator Cycle Sequencing chemistry: ABI) according to the manufacturer instructions.

Descriptive statistics

Estimates of haplotypic diversity (*h*) and nucleotide diversity (π) were calculated according to Nei (1987). Relationships among different haplotypes were estimated using minimum spanning network performed with the help of the ARLEQUIN program (Version 2.0, Schneider *et al*, 2000).

Neutrality test for mtDNA

There is some evidence that mtDNA may be not a strictly neutral marker (William *et al*, 1995). As a test of neutrality of different mtDNA haplotypes, we tested whether sequence variation in different populations conforms to a neutral infinite allele distribution (Tajima, 1989). The departures from neutrality could be detected by measuring the magnitude of differences, as measured by Tajima's *D* statistics (1989). This test assumes no severe bottleneck and recent expansion, which would lower the values of *D* statistics, leading to increased probability of erroneously rejecting neutrality.

Historical demography

To study whether a reduction in population size has occurred in the past we adopted two different approaches. The estimation of the long-term effective female population size was done from the equation: $Ne = [10^6\pi]/[2sg]$ (Avise *et al*, 1988), where π is the nucleotide diversity (mean number of base substitutions per nucleotide between individuals within the given population), *s* is the evolutionary rate (*s* = % substitutions per genome per million years), *g* is the generation time. We adopted the estimate of *s* = 2.5% divergence rate for the cytochrome *b* gene in mammals (Meyer *et al*, 1990) and *g* = 1. Another way to gain insight into the historical

Figure 1 Map of the locations of the populations sampled and the distribution of the chromosome races of the common shrew. Race abbreviations the after Zima *et al* (1996). Race ranges after Fedyk *et al* (2000). Bold line = the position of contact zone between WEKG and EEKG karyotypic groups.

demography of the populations studied is the examination of the distribution of pairwise sequence differences within populations (Slatkin and Hudson, 1991; Rogers and Harpending, 1992; Rogers, 1995). This information from pairwise sequence differences ('mismatch distribution') was used to test whether current haplotypic variation in the common shrew might fit better with expectations of the 'equilibrium' model (constant long term N_e) or with the 'sudden expansion' model which postulates a recent expansion in population size (Rogers, 1995). Moreover, parameters derived from mismatch distributions can be used to get rough estimates of three demographic parameters: effective female population size before (N_0) and after (N_1 = current size) a hypothesized expansion, and the time (t) to the expansion in generations (Rogers and Harpending, 1992). The initial population size (N_0) was estimated by equating $\theta_0 = 2N_0u$, where $u = 2sk$, where s is the mutation rate (2.5% per Myr, see Meyer *et al*, 1990) and k is the length of the sequence (1023 bp). θ_0 (the expected distribution of pairwise differences before expansion) was estimated as $\sqrt{v - m}$, where m and v are the observed mean and variance of pairwise sequence differences. The time (τ) in units of $1/2u$ generations (where u is the sum of per-nucleotide mutation rate in the mtDNA region under study) was estimated as $m - \theta_0$ (Rogers, 1995). This estimation, however, cannot be done if the variance of the mismatch distribution is smaller than the mean. To overcome this, θ_0 was set to zero if $v < m$ following Rogers (1995). It should be noted that if the expected distribution of pairwise differences before expansion (θ_0) was equal zero, the initial population size (N_0) was also zero. This does not however, mean that the ancestral population was of zero size, it rather means that before expansion all haplotypes in the population were identical. An estimate for current population size (N_1) was derived from the equation: $\theta_1 = 2N_1u$, where θ_1 is the expected distribution of pairwise differences after expansion (Rogers and Harpending, 1992). The estimates of time since expansion were obtained from the equation: $\tau = 2ut$ where t is the time elapsed between N_0 and N_1 (Rogers and Harpending, 1992). The validity of the estimated expansion model was tested using the parametric bootstrap approach (Schneider and Excoffier, 1999). The sum of square deviations (SSD) between the observed and the expected mismatch was used as a test statistics (Schneider and Excoffier, 1999). All the calculations mentioned above were done when sample size was not less than five individuals (samples BLI, JUR and JAB). Analysis of 'full sample' and separate calculations for combined samples of different karyotypic groups: WEKG and EEKG were also performed. All the analyses were done with the help of the ARLEQUIN program (Version 2.0, Schneider *et al*, 2000).

Population structure

To test for differentiation among populations, an exact test for population differentiation (the hypothesis of a random distribution of k different haplotypes among r populations) was performed using Markov Chain Algorithm (Raymond and Rousset, 1995). Furthermore, using the cytochrome *b* gene data, pairwise F_{ST} values among BLI, JAB and JUR populations, races and karyotypic groups under study, were calculated. To provide a comparison of divergence between WEKG and EEKG using

two types of markers (mtDNA and allozymes) we calculated Nei's (1978) genetic distance between populations of different races and karyotypic groups studied. Observed and expected heterozygosity values for the common shrew populations were also calculated to see possible loss of genetic variation due to a population bottleneck. The program BOTTLENECK (Cornuet and Luikart, 1996) was used to test for a recent reduction of effective population size. The data for 47 enzyme loci were obtained using conventional starch gel and cellulose acetate electrophoresis of allozymes (Richardson *et al*, 1986; Murphy *et al*, 1996).

Results

Characterization of cytochrome *b* gene sequence

The sequenced fragment of the cytochrome *b* gene was 1023 nucleotides long (GenBank accession numbers: AJ409867–AJ409894). Thirty-four variable positions, defining 21 distinct haplotypes were found among 28 individuals studied (Table 1). Most of substitutions were transitions (30 out of 35, TS: TV ratio was 6), like in other mammalian species, including *S. araneus* (Irwin *et al*, 1991; Taberlet *et al*, 1994). As in the other study of cytochrome *b* gene of six *Sorex* species (Fumagalli *et al*, 1996), we found for the light strand of cyt *b* gene a deficiency of guanine (13.70%); the other nucleotides were more balanced (29.49% thymine, 27.96% cytosine and 28.85% adenosine). The mean pairwise nucleotide divergence (π) among the haplotypes was 0.0043, SD $\pm$ 0.002, n = 21) and π = 0.0036, SD $\pm$ 0.002 among all individuals in the entire sample (k = 378 comparisons). Assuming a rate of 2.5% divergence for the cytochrome *b* gene in mammals (Meyer *et al*, 1990), a divergence rate per Myr, then the divergence equal to 0.88% between the two most divergent haplotypes indicates that the common ancestor of all current *S. araneus* haplotypes existed about 350 000 years ago.

Interpopulation variation of the haplotypes

The most common haplotype was found in only four out of 10 populations studied and the majority of haplotypes (18 out of 21) were represented only once (Table 1). The most common haplotype was present both, in EEKG and WEKG, however in WEKG it was found only once. Haplotype diversity estimates within the populations sampled were high (h = 0.800–0.928), while nucleotide diversity estimates were low (π = 0.0034–0.0053 $\pm$ 0.002, Table 2). The question arises whether the haplotype and nucleotide diversity estimates can be explained by selection or historical demography of populations.

The Tajima's (1989) test of neutrality was performed. The distribution of haplotypic variability within populations, races and karyotypic groups was compared to that expected under infinite-site model without recombination. The hypothesis of neutrality was not rejected in the BLI, JUR and the JAB samples. We rejected the null hypothesis of neutrality in the combined samples: of the Białowieża race, EEKG and WEKG karyotypic groups and in a pooled sample of all specimens (Table 2). The rejection of neutrality, however may be due to factors other than selection acting on haplotypes, for example population expansion, which in fact is supported by significantly negative Tajima's *D* values (Table 2). Thus, we

Table 1 Cytochrome *b* haplotypes found in 10 *Sorex araneus* populations studied. For population abbreviations and their chromosome constitution see Materials and methods

No	Haplotype ^a	JUR	BLI	ULA	RAM	ZUB	JAB	ROK	LUT	KEP	LUB	E	W	Σ
1	AACCTACTACTCAGTGACCAATCCATCTCCCTTT	1	0	0	0	0	0	0	0	0	0	1	0	1
2	C.....AT..TT...	2	0	0	0	0	1	0	0	0	0	2	1	3
3	T.....AA.....C.....T..TT...	1	0	0	0	0	0	0	0	0	0	1	0	1
4	CT...T..TTC...	1	0	0	0	0	0	0	0	0	0	1	0	1
5	C.....T..TT...	1	0	1	0	2	0	0	1	0	0	4	1	5
6	.G.T.....T.....C.....T..TT...	0	1	0	0	0	0	0	0	0	0	1	0	1
7	.T....C.....C..T...C...AT..TT...	0	1	0	0	0	0	0	0	0	0	1	0	1
8	C.....C.....T..TT...	0	1	0	0	0	0	0	0	0	0	1	0	1
9	A.....GC...AT..TT...	0	1	0	0	0	0	0	0	0	0	1	0	1
10	G.....A.....GC...AT..TT...	0	1	0	0	0	0	0	0	0	0	1	0	1
11	T.....C.....T..TT...	0	0	0	1	0	0	0	0	0	0	1	0	1
12	.T.....A.....C.....T..TT...	0	0	0	1	0	0	0	0	0	0	1	0	1
13	G.....C..G.T...T...	0	0	0	0	0	2	0	0	0	0	0	2	2
14	A.....C.....T..TT...	0	0	0	0	0	1	0	0	0	0	0	1	1
15	C.....T.TTA...	0	0	0	0	0	1	0	0	0	0	0	1	1
16	G.C.....T..TT.C.	0	0	0	0	0	1	0	0	0	0	0	1	1
17	C.....T..TT..C	0	0	0	0	0	0	1	0	0	0	0	1	1
18	G.....T...C...AT..TT...	0	0	0	0	0	0	1	0	0	0	0	1	1
19	C.T...TC.TT...	0	0	0	0	0	0	0	1	0	0	0	1	1
20	A..G.T..C.....T..TT...	0	0	0	0	0	0	0	0	1	0	0	1	1
21	A.....C.....TT...	0	0	0	0	0	0	0	0	0	1	0	1	1
Σ		6	5	1	2	2	6	2	2	1	1	16	12	28

^aVariable nucleotide sites (34) in the sequential order. E = EEKG, W = WEKG.

Table 2 Cytochrome *b* haplotype (*h*) and nucleotide (π) diversities, Tajima (1989) *D* statistics and estimated long-term effective female population size (*N_e*) for *Sorex araneus* samples

Sample	No.	<i>n_H</i>	<i>h</i>	$\pi \pm SE$	<i>N_e</i>	<i>D</i>
JUR	6	5	0.861	0.0035 ± 0.002	70,000	-1.16
BLI	5	5	0.800	0.0053 ± 0.004	106,000	-0.45
JAB	6	4	0.861	0.0037 ± 0.003	74,000	-0.21
Białowieża race	12	10	0.875	0.0040 ± 0.002	80,000	-1.79*
Drnholec race	10	9	0.880	0.0035 ± 0.002	70,000	-1.31
EEKG	16	12	0.918	0.0034 ± 0.002	68,000	-1.91**
WEKG	12	11	0.903	0.0037 ± 0.002	74,000	-1.59*
Pooled	28	21	0.928	0.0035 ± 0.002	70,000	-2.17**

No. = number of individuals; *n_H* = number of distinct haplotypes; **P* > 0.05, ***P* > 0.01 for statistically significant *D* values; SE = standard error.

also performed analyses to evidence possible historical explanations.

The long-term effective population size (*N_e*) calculated using mean pairwise divergence and evolutionary rate (2.5% per Myr) for the pooled sample was 70 000. *N_e* values for EEKG and WEKG were 68 000 and 74 000, respectively (Table 2). We also estimated the effective female population size after expansion (*N₁*, Table 3), using the method of Rogers and Harpending (1992). The *N₁* values from Table 3 seem to be overestimated, however it is evident, they are remarkably high, indicating considerable increase of population size during population expansion. Thus, no recent population bottleneck was probably present after expansion of *S. araneus* populations, after the last glaciation in Poland. For the groups studied, the observed distribution of pairwise sequence differences (the mismatch distribution) was unimodal and followed the expected distribution of a growing population (see Figure 3 for EEKG as an example). The time since the beginning of expansion (*t* = 37 400–62 400

Table 3 Results of pairwise sequence difference showing approximate time to the beginning of the expansion (*t*), effective female population size at the beginning (*N₀*) and at the end (= present day) of expansion (*N₁*), and the probability (*P*) that the simulated sum of squared deviations is greater or equal than the observed sum of squared deviations (the fit to the predicted expansion scenario)

Sample	No.	<i>t</i>	<i>N₀</i>	<i>N₁</i>	<i>P</i>
JUR	6	42,500	19	409,800	0.85
BLI	5	62,400	0	91,544,000	0.46
JAB	6	47,900	0	91,544,400	0.12
Białowieża race	12	40,500	9012	582,000	0.88
Drnholec race	10	38,900	0	93,157,300	0.35
EEKG	16	37,700	2463	130,107,000	0.95
WEKG	12	41,800	0	1,270,000	0.12
Pooled	28	37,400	39	3,091,000	0.72

years ago, Table 3), estimated according to Rogers and Harpending (1992), coincides with the last glacial period.

Population differentiation

We found no statistically significant differentiation in the cytochrome *b* gene between karyotypic groups (group pairwise differentiation test for non-differentiation, NS); races (NS) and populations (NS). All the F_{ST} values were equal to zero. Hence, due to low degree of absolute divergence and fairly short time to common ancestry no population subdivision was found in *S. araneus* studied. The lack of phylogeographical structure is also evident from the minimum spanning network (Figure 2), which seems to represent a 'sudden expansion' model.

Allozyme analysis

Eight loci were found to be polymorphic with two or more alleles in at least one population: *aminoacylase*, *esterase-1*, *esterase-D*, *isocitric dehydrogenase-2*, *lactate dehydrogenase-2*, *mannosephosphate isomerase*, *phosphoglucumutase-1* and *phosphoglucumutase-3*. The other loci were monomorphic. None of the alleles was diagnostic for any race or karyotypic group. The full list of loci and allele frequencies are available from the authors upon request. Heterozygosity values (both observed and expected) were similar in all the populations studied and did not

differ substantially (Table 4). Tests for a population bottleneck revealed no evidence for a recent reduction in effective population size in any sampling area (Table 4). The values of Nei's (1978) genetic distances (*D*) among populations studied were low. Values of *D* among populations within EEKG varied from 0.000 to 0.006, within WEKG *D* values were from 0.000 to 0.009, with average *D* value between WEKG and EEKG 0.003 (0.000 – 0.006), showing no divergence between WEKG and EEKG on allozyme loci.

Discussion

Selection on mtDNA and demographic explanations

We did not notice an increased frequency of any haplotype in any chromosome race or karyotypic group, what could be indicative for the process of positive selection favouring the common haplotype. Surprisingly, the test of neutrality was rejected in the case of Białowieża race, WEKG, EEKG and in a pooled sample (Table 2). This, however, could be expected if the populations under study have not been of constant size for considerable periods of time. In such a case, the approach cannot strictly separate selective and demographic explanations (Tajima, 1989; William *et al*, 1995). An excess of singletons may be an evidence for expanding populations (William *et al*, 1995). Indeed, the expansion scenario is supported by significantly negative Tajima's *D* values (Table 2) and the star-like haplotype network (Figure 2).

The observed distribution of pairwise sequence differences (the mismatch distribution) was unimodal (Figure 3), and followed the expected distribution of a growing population (Rogers and Harpending, 1992). The long-term effective female population sizes (*N_e*) estimated from nucleotide diversity (π) seem to be relatively high (Table 2), indicating that the population size probably has not dropped to a few individuals since the last glaciation. Furthermore, the estimates of present-day female population sizes (*N_i*), based on mismatch distribution of pairwise sequence differences (Rogers and Harpending, 1992) were around tens of millions of individuals (Table 3). The *N_i* values from Table 3 are probably overestimated and should be regarded as rough approximations. Hence, although there is a large uncertainty on true values of estimated parameters, calculated *N_i* values (Table 3) indi-

Figure 2 The minimum spanning network obtained for nine populations of the common shrew in Poland. For abbreviations see Materials and methods. The small black nodes refer to one substitution. The alternative links between haplotypes indicated with an asterisk (*).

Table 4 Observed (*H_o*) and expected (*H_e*) heterozygosity values calculated over 47 allozyme loci for nine *Sorex araneus* populations studied (standard error in parentheses). The probability (*P*) that the populations studied fit to mutation-drift equilibrium is also given as computed using BOTTLENECK (Wilcoxon test; Cornuet and Luikart, 1996)

Sample	<i>H_o</i>	<i>H_e</i>	<i>P</i>
JUR	0.039 (0.017)	0.042 (0.018)	0.44
BLI	0.042 (0.017)	0.045 (0.02)	0.69
ULA	0.045 (0.021)	0.042 (0.019)	0.63
ZUB	0.038 (0.017)	0.042 (0.019)	0.56
RAM	0.048 (0.02)	0.049 (0.021)	0.21
JAB	not studied	not studied	–
LUT	0.035 (0.014)	0.039 (0.016)	0.64
ROK	0.026 (0.013)	0.031 (0.015)	0.06
KEP	0.047 (0.02)	0.047 (0.021)	0.69
LUB	0.052 (0.023)	0.049 (0.021)	0.09

Figure 3 The observed and expected distribution of pairwise differences under the model of population expansion (Rogers and Harpending, 1992). The parameters used were for the EEKG group of the common shrew: $\tau = 3.856$, $\theta_0 = 0.126$ and $\theta_1 = 16.977$.

cate very large present-day female population sizes. It can be assumed, for the common shrew, that when the glacial period ended, the species rapidly occupied new habitats and the population size increased. This is in agreement with ecological data: in favourable habitats *S. araneus* presumably exists in large, continuous populations (Croin-Michielsen, 1966). Since mismatch distributions were not strongly L-shaped (left truncated), which would be rather typical for strongly bottlenecked populations (Marjoram and Donnley, 1994), we assume there has been no population bottleneck for *S. araneus* populations since last glaciation.

Concordance between molecular data and models of chromosomal evolution

The 'sudden expansion' model suggested in our study of the *cyt b* gene of the common shrew is in agreement with the chain variant of the stasipatric model of chromosome evolution (White, 1978) proposed for chromosome races in Poland (Wójcik, 1993). The race-diagnostic metacentrics may have arisen within the species range by the process of Rb fusions during expansion through Europe. According to Wójcik (1993) the area of Poland was populated from the southwest by shrews possessing *jl* and *hi* metacentrics, while shrews with *jl* and *gr* were colonizing this area from the east during early post-glacial period. After that different metacentrics subsequently spread over populations (Wójcik, 1993). Our data from the cytochrome *b* gene and allozyme variation seem to favour the model assuming that Rb fusions rather than WARTs played main role in the chromosome evolution of the common shrew in Poland. The lack of a population bottleneck probably rules out the possibility of WARTs in Polish chromosome races of the common shrew. This is because WARTs may be fixed in bottlenecked populations only (Searle and Wójcik, 1998). This does not, however, mean that no shrew metacentrics have been formed by WARTs. As the distribution and chromosomal relationships of some races in Finland, Sweden and in Siberia could be most easily explained using WARTs (Halkka *et al*, 1987; Fredga, 1996; Polyakov *et al*, 2000), it would be interesting to use molecular markers to study

whether populations belonging to those races were bottlenecked or not in the past.

Divergence between karyotypic groups on molecular ground

In the study of cytochrome *b* gene of *S. araneus* we showed somewhat intriguing picture of population structure. The two aspects of our study seem to be most striking: the lack of ancient divergence between two karyotypic groups: EEKG and WEKG of *S. araneus* on molecular grounds and the star-like phylogeny of different haplotypes (Figure 3). We found only slight differences in haplotype frequencies among different populations, despite the fact that they represent different karyotypic groups and chromosome races. The biggest pairwise nucleotide divergence among the shrews studied was 0.88%, despite the fact that they represent different karyotypic groups and chromosome races. This does not, however, rule out the possibility of the common shrews from WEKG and EEKG having survived in different refuges. Unfortunately, the cytochrome *b* gene does not seem to be an appropriate marker to distinguish such alternatives. Bilton *et al* (1998) showed that the same branch of NJ tree of *cyt b* haplotypes links together individuals from a huge geographical area extending from Western Europe to Eastern Siberia, although most certainly they did not survive the last glaciation in a single refugium (see Searle and Wójcik, 1998). The estimated time to the common ancestor of all current *S. araneus* haplotypes (350 000 years ago) is certainly too long for the last recolonization of Poland. To explain this, two hypotheses may be considered: (1) the molecular clock is faster in *Sorex*, or (2) *S. araneus* did not suffer from a bottleneck before the last recolonization. At the moment, it is not possible to say, which hypothesis is more probable. It should be noted, however, that evolutionary rate of mtDNA in Insectivora may show 1.6-fold increase when compared to the slowest evolving orders of mammals (Grissi *et al*, 2000). If this holds true for *Sorex*, all the haplotypes studied coalesce to the common ancestor that lived around 220 000 years ago.

The lack of population structuring in present-day populations most probably reflects the retention of shared ancestral polymorphism spread over large geographical areas by recent population expansion. The expansion probably started about 37 400 to 62 400 years ago (Table 3) and it coincides with the last glacial period. Thus, during expansion process through Europe different chromosome races may have arisen within the species range.

Although the chromosomal differences between distinct races of the *S. araneus* in Poland are evident, the pattern of genetic variation between karyotypic groups revealed for the *cyt b* gene does not represent secondary contact between previously isolated populations. Thus, the results from the cytochrome *b* gene studies are in an agreement with the hypothesis of the recentness of the chromosomal variation in the common shrew (Taberlet *et al*, 1994).

Acknowledgements

This work was supported by the Polish Scientific Committee (grant KBN no 6P04C 063 09 to M Ratkiewicz and 6P204 03405 to S Fedyk) and Université Joseph Fourier

and Centre National de la Recherche Scientifique (France). We thank Agnieszka Radulska, Monika Zaleska and Konrad Fedyk for their invaluable help in the field.

References

- Avice JC, Ball RM, Arnold J (1988). Current versus historical populations sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Mol Biol Evol* 5: 331–344.
- Bengtsson BO, Frykman I (1990). Karyotype evolution: evidence from the common shrew (*Sorex araneus* L.). *J Evol Biol* 3: 85–101.
- Bilton DT, Mirol PM, Mascheretti S, Fredga K, Zima J, Searle JB (1998). Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards post-glacial colonization. *Proc R Soc Lond B* 265: 1219–1226.
- Bulatova N, Searle JB, Bystrakova N, Nadjafova R, Shchipanov N, Orlov V (2000). The diversity of chromosome races in *Sorex araneus* from European Russia. In: Searle JB, Wójcik JM (eds) *Evolution in the Sorex araneus group: cytogenetic and molecular aspects*. *Acta Theriol* 54 (Suppl 1): 33–46.
- Cornuet JM, Luikart G (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144: 2001–2014.
- Croin-Michielsen N (1996). Intraspecific and interspecific competition in the shrews *Sorex araneus* L. and *S. minutus* L. *Arch Neerlandaises de Zoologie* 17: 73–174.
- Fedyk S (1995). *Geographic chromosomal differentiation and hybrid zones between chromosome races of Sorex araneus in north-eastern Poland*. Dissertationes Universitatis Varsoviensis 439, Białystok [in Polish].
- Fedyk S, Banaszek A, Chętnicki W, Cichomska A, Szalaj KA (2000). Reassessment of the range of the Drndolec race: studies on meiosis in *Sorex araneus* hybrids. In: Searle JB, Wójcik JM (eds) *Evolution in the Sorex araneus group: cytogenetic and molecular aspects*. *Acta Theriol* 54 (Suppl 1): 59–67.
- Fredga K (1996). The chromosome races of *Sorex araneus* in Scandinavia. In: Fredga K, Searle JB (eds) *Evolution in the Sorex araneus group. Cytogenetic and molecular aspects: Proceedings of the ISACC's Fourth International Meeting*. *Hereditas* 125: 123–135.
- Fumagalli L, Taberlet P, Favre L, Hausser J (1996). Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. *Mol Biol Evol* 13: 31–46.
- Grissi C, Reyes A, Pesole G, Saccone C (2000). Lineage-specific evolutionary rate in mammalian mtDNA. *Mol Biol Evol* 17: 1022–1031.
- Halkka L, Soderlund V, Skaren U, Heikkilä J (1987). Chromosomal polymorphism and racial evolution of *Sorex araneus* L. in Finland. *Hereditas* 106: 257–275.
- Hauffe HC, Pialek J (1997). Evolution of chromosomal races of *Mus musculus domesticus* in the Rhaetian Alps: the role of whole-arm reciprocal translocation and zonal raiation. *Biol J Linn Soc* 62: 255–278.
- Irwin DM, Kocher TD, Wilson AC (1991). Evolution of cytochrome *b* gene of mammals. *J Mol Evol* 32: 128–144.
- Marjoram P, Donnelly P (1994). Pairwise comparisons of mitochondrial DNA sequences in subdivided populations and implications for early human evolution. *Genetics* 136: 673–683.
- Meyer A, Kocher TD, Basasibwaki P, Wilson AC (1990). Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* 347: 550–553.
- Mishta AV, Searle JB, Wójcik JM (2000). Karyotypic variation of the common shrew *Sorex araneus* in Belarus, Estonia, Latvia, Lithuania and Ukraine. In: Searle JB, Wójcik JM (eds) *Evolution in the Sorex araneus group: cytogenetic and molecular aspects*. *Acta Theriol* 54 (Suppl 1): 47–58.
- Murphy RW, Sites JW, Buth DG, Haufler CH (1996). Proteins: isozyme electrophoresis. In: Hillis DM, Moritz C, Mable BK (eds) *Molecular Systematics*, Sinauer Associates: Sunderland, MA. pp 51–120.
- Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.
- Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University Press: New York.
- Polyakov AV, Zima J, Banaszek A, Searle JB, Borodin PM (2000). New chromosome races of the common shrew *Sorex araneus* from Eastern Siberia. In: Searle JB and Wójcik JM (eds) *Evolution in the Sorex araneus group: cytogenetic and molecular aspects*. *Acta Theriol* 54 (Suppl 1): 11–18.
- Raymond ML, Rousset F (1995). An exact test for population differentiation. *Evolution* 49: 1280–1283.
- Richardson BJ, Baverstock PR, Adams M (1986). *Isoenzyme Electrophoresis. A Handbook for Animal Systematics and Population Studies*. Academic Press: San Diego, California.
- Rogers A (1995). Genetic evidence for a Pleistocene population explosion. *Evolution* 49: 608–615.
- Rogers AR, Harpending H (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol* 9: 552–569.
- Schneider S, Excoffier L (1999). Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 152: 1079–1089.
- Schneider S, Roessli D, Excoffier L (2000). Arlequin, ver. 2.000. A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Seabright M (1971). A rapid banding technique for human chromosomes. *Lancet* 2: 971–972.
- Searle JB (1984). Three new karyotypic races of the common shrew *Sorex araneus* (Mammalia: Insectivora) and a phylogeny. *Syst Zool* 33: 184–194.
- Searle JB (1993). Chromosomal hybrid zones in eutherian mammals. In: Harrison RG (ed) *Hybrid zones and the Evolutionary Process*, Oxford University Press: New York. pp 309–353.
- Searle JB, Fedyk S, Fredga K, Hausser J, Volobouev V T (1991). Nomenclature for the chromosomes of the common shrew (*Sorex araneus*). In: Hausser J (ed) *The Cytogenetics of the Sorex araneus group and related topics: Proceedings of the ISACC's Second International Meeting*. *Mem Soc Vaud Sci Nat* 19: 13–22.
- Searle JB, Wójcik JM (1998). Chromosomal evolution: the case of *Sorex araneus*. In: Wójcik JM and Wolsan M (eds) *Evolution of Shrews* Mammal Research Institute: Polish Academy of Sciences, Białowieża, pp 219–268.
- Slatkin M, Hudson RR (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129: 555–562.
- 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.
- Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
- White MJD (1978). *Models of Speciation*. Freeman: San Francisco.
- William J, Ballard O, Kreitman M (1995). Is mitochondrial DNA a strictly neutral marker? *Trends Ecol Evol* 10: 485–488.
- Wójcik JM (1993). Chromosome races of the common shrew *Sorex araneus* in Poland: a model of karyotype evolution. *Acta Theriol* 38: 315–338.
- Wójcik JM, Searle JB (1988). The chromosome complement of *Sorex granarius*—the ancestral karyotype of the common shrew (*Sorex araneus*)? *Heredity* 61: 225–229.
- Wytenbach A, Borodin P, Hausser J (1998). Meiotic drive favors Robertsonian metacentric chromosomes in the common shrew (*Sorex araneus*, Insectivora, Mammalia). *Cytogenet Cell Genet* 83: 199–206.
- Wytenbach A, Hausser J (1996). The fixation of metacentric

chromosomes during the chromosomal evolution in the common shrew (*Sorex araneus*, Insectivora). *Hereditas* **125**: 209–217.

Zima J, Fedyk S, Fredga K, Hausser J, Mishta A, Searle JB, Volobouev V, Wójcik JM (1996). The list of the chromosome

racies of the common shrew (*Sorex araneus*). In: Fredga K, Searle JB (eds) *Evolution in the Sorex araneus group. Cytogenetic and Molecular Aspects: Proceedings of the ISACC's Fourth International Meeting*. *Hereditas* **125**: 97–107.