Phylogeography of the White-Toothed Shrews Crocidura suaveolens and Crocidura sibirica: Searching for the Geographical Homeland

V. A. Gritsyshin,*, I. V. Artyushin, V. O. Burskaya, B. I. Sheftel, V. S. Lebedev, and A. A. Bannikova

a Moscow State University, Moscow, 119991 Russia
b Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, 127051 Russia
c Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, 119071 Russia
d Zoological Museum, Moscow State University, Moscow, 125009 Russia

*e-mail: vladimir.sokol.gritsyshin@gmail.com
**e-mail: hylomys@mail.ru

Received November 29, 2020; revised June 23, 2021; accepted June 23, 2021

Abstract—We studied the polymorphism of the cytb gene in two forms of the Lesser White-toothed Shrew species complex: Crocidura suaveolens s. stricto and C. sibirica. The haplotypes of C. sibirica are found to be very similar to those of Crocidura suaveolens. They do not belong to a distinct haplogroup. The molecular diversity of the populations in the Asian part of the range is higher than in Eastern Europe. For the combined sample from Asia and Europe together, we revealed a significant signal of population expansion. Analysis of the expansion time showed that the Asian territory was colonized earlier (before the last glacial maximum) than the Eastern Europe (at the very end of the Late Pleistocene and in the early Holocene). The results of the ancestral area reconstruction are consistent with the hypothesis of a Middle Asian origin of the C. suaveolens/C. sibirica group, recent colonization of Inner Asia and later penetration into Eastern Europe.

Keywords: Crocidura suaveolens, Crocidura sibirica, mitochondrial phylogeography, ancestral area

DOI: 10.1134/S1062359022020091

INTRODUCTION

The Lesser White-toothed Shrew in a broad sense (Crocidura suaveolens sensu lato) is a large species complex, and its composition is still under discussion. Over a vast area stretching across all of Eurasia from the Iberian Peninsula to Tsushima Island through several natural zones, all forms of the Lesser White-toothed Shrew, except for C. sibirica, have identical karyotypes (Grafodatsky et al., 1988), and many authors believe that all populations with 2n = 40 on this territory should be considered a single species Crocidura suaveolens (Catzeflis et al., 1985; Vogel et al., 1986; Hutterer, 1993). However, there is also evidence of morphological (Zaitsev, 1991; Tembotova, 1999; Bannikova et al., 2001; Jiang and Hoffmann, 2001) and high molecular (Vogel et al., 2003; Ohdachi et al., 2004; Bannikova et al., 2006; Dubey et al., 2006, 2007a, 2007b) diversity of this group. At present, no fewer than seven forms are recognized within the “Lesser White-toothed Shrew” species group according to morphological data and mitochondrial DNA (mtDNA) sequences; however, the views on the species status of these form vary among different authors. Usually, the Lesser White-toothed Shrew C. suaveolens (Pallas, 1811), the Siberian Shrew C. sibirica (Dukelsky, 1930), the Caucasian Long-tailed Shrew C. gueldens-taeddii (Pallas 1811), the Caspian Shrew C. caspica (Thomas, 1907), the Manchurian Shrew C. shantungensis (Miller 1901), and Zarudnyi’s Rock Shrew C. zarudnyi (Ognev, 1928) (Zaitsev, 1991; Bannikova et al., 2006; Zaitsev et al., 2014; Burgin et al., 2018) are considered species. However, the species status is also assigned by some authors to C. mimula (Miller 1901) from Western Europe (Bannikova et al., 2006) or C. a. aleskandridis (Vesmanis, 1977) from Cyrenaica on the Mediterranean coast of North Africa (Burgin et al., 2018). According to mitochondrial data, there are some other forms in the “Lesser White-toothed Shrew” species group: the samples from the Iberian Peninsula, the south and west of France, and Southern Italy form a clade that includes the subspecies cantabra and ictus; the shrews of Cyprus (cypria), the samples from Lesbos and Vukarikisika islands and Western Turkey, the white-toothed shrews from Central and Southeastern Iran, and from Libya (Dubey et al., 2007a) all form distinct phylogroups. The taxonomic status of these forms remains unclear.

The Lesser White-toothed Shrew in a narrow sense (Crocidura suaveolens s.str.) (the type territory is the Crimea, Chersonesus) has the most extensive range
among all Palearctic species and has striking diversity of the habitats, occurring mainly in the intrazonal biotopes of the desert, steppe, and forest zones, as well as in anthropogenic habitats. In the deserts of Middle and Inner Asia, it inhabits dry stream beds (saires) and oases; in the forest zone of the European part of the range, it particularly tends to be synanthropic, occurring in various residential buildings (Zaitsev et al., 2014).

The Lesser White-toothed Shrew has a wide range of color and size variability, which explains the recognition of several forms by different authors: *C. lignicolor* (Miller 1900) from Xinjiang, *C. ilensis* (Miller 1901) from eastern Kazakhstan, *C. lar* (Allen 1928) from Mongolia, and *C. mordeni* (Goodwin 1934) from Central Kazakhstan. In some papers, the name *C. gmelini* was erroneously attributed to the form *suaveolens* s.str. (Hoffmann, 1996; Jiang and Hoffmann, 2001).

The phylogeographic structure and demographic history of *C. suaveolens* s.str. have been poorly studied, due to limited size of the previously analyzed samples. The vast range of the Lesser White-toothed Shrew demands a representative sample covering the entire geographical variation of the species. We should remark here that the Lesser White-toothed Shrew is the strict sense is genetically clearly distinct from other forms of the species complex, except for the Siberian Shrew. At the same time, the Siberian White-toothed Shrew has a wide range and both species occur in the Inner Asia, it inhabits dry stream beds (saires) and oases; in the forest zone of the European part of the range, it particularly tends to be synanthropic, occurring in various residential buildings (Zaitsev et al., 2014). Inter alia, difficulties (Zaitsev, 1991; Zaitsev et al., 2014). Inter alia, according to I.I. Stogov, only the Siberian White-toothed Shrew inhabits the right-bank part of the East Kazakhstan Region and both species occur in the Black Irtysh delta (Bekenov et al., 1985).

However, we have not found any morphologically identified specimens of *C. sibirica* in our sample from the latter locality (the Zaisan Basin). It is also worth noting that the shrews from the Middle and Inner Asia in some studies were erroneously assigned to *C. sibirica* instead of *C. suaveolens*, without examination of the data on *terra typica* of the latter form and without considering the morphological traits (Jiang and Hoffmann, 2001; Ohdachi et al., 2004; Dubey et al., 2006, 2007a). To address the problem of genetic mtDNA testing, we also included in our analysis the data on the samples with a geographical origin and morphological traits corresponding to the *C. sibirica* form.

In the present work, we studied for the first time the phylogeographic structure of the Lesser and Siberian White-toothed Shrews on the extensive geographical sample. The aim of our study was to explore intra-specific genetic variation, to identify potential phylogeographic breaks, and to find the center of genetic diversity and to determine the direction of dispersal.

### MATERIALS AND METHODS

**The volume of the material.** The sample analyzed included the mitochondrial cytochrome *b* (*cytb*) gene sequences of 160 individuals of the Lesser White-toothed Shrew and 14 individuals of the Siberian Shrew; among them, 109 samples of *C. suaveolens* and three samples of *C. sibirica* are the original material presented in this work for the first time. Other sequences (76 sequences), including the outgroup, were taken from the GenBank database (AB077076, AB077078, AB077080, AB077083–AB077089, AB077278, AB077279, AY843462–AY843486, AY994368, AY994370, AY994377, AY994386, AY994387, AY994389, DQ630062, DQ630063, EU742583, EU742593, EU742609–EU742614, HM586991–HM586996, KX354172–KX354175, KX354177, KX354178, KX354180, KX354181, MN690930–MN690935, MN690937–MN690939); we obtained 28 of them in the previous studies (Bannikova et al., 2006, 2009, 2011; Sheftel et al., 2018), and 48 were sequenced by other authors (Vogel et al., 2003; Ohdachi et al., 2004; Dubey et al., 2006, 2007a; Chen et al., 2020). The total sample includes specimens from 90 localities and covers the most part of the studied species range (Fig. 1).

**Isolation, amplification, and sequencing.** DNA was isolated by the standard phenol-chloroform method (Sambrook et al., 1989). External primers from the previous publications (Bannikova et al., 2006, 2011) were used for amplification; internal primers L15136 (Bannikova et al., 2006) and Cro_H481su (5'-ACA-GAGAATCTCCTCCTCAGATTCACTCTAC-3') specially designed for the “suaveolens” group were used for sequencing. Sequencing was performed commercially usingABI PRISM® BigDye™ Terminator v. 3.1, followed by analysis of the reaction products with an Applied Biosystems 3730 DNA Analyzer automatic sequencer at Evrogen JSC. The sequences are deposited in GenBank under accession numbers: MW297680–MW297791.

**Phylogenetic and phylogeographic analysis.** We conducted the phylogenetic analysis by the maximum likelihood estimation (MLE) method in the IQ-TREE (Nguyen et al., 2015; Hoang et al., 2017), by the neighbor joining (NJ) and maximum parsimony (MP) methods in the MEGA X program (Kumar et al., 2018), and by the Bayesian inference (BI) method in BEAST v. 1.10.4 (Drummond and Rambaut, 2007). In the latter case, the chain length was $3 \times 10^8$ generations; for constructing an ultrametric tree, we split the alignment into two partitions by codon positions 1 + 2 and 3 and set the molecular evolution models HKY and GTR + G4 for the first and second partitions, respectively. In the ML, NJ, and MP analyses, the reliability of topology was assessed by a bootstrap test.
with 1000 replications. The sequences of *C. shantu-
genesis* (AB077076, AB077078, AB077080, AB077278, EU742593), *C. caspica* (AY994368, AY994370), and *Crocidura aff. suaveolens* (KX354172–KX354175, KX354177, KX354178, KX354180) were taken from GenBank as the outgroup for trees.
group, the final alignment included 174 sequences for the ML, MP, and NJ reconstructions and 146 sequences for the BI analysis.

The median-joining haplotype network was constructed with NETWORK 5.0 (Bandelt et al., 1999) and edited using the standard Paint package in Microsoft Windows 10. The genetic distances between samples (p- and K2P distances) were calculated using MEGA X. The p-distances between large geographic samples were analyzed to determine their genetic similarity and to determine the number of groups in the subsequent analyses of the genetic diversity and demographic variables.

The genetic diversity estimators (nucleotide and haplotype diversity), the parameters of pairwise distances distribution (mismatch distribution), expansion time (τ), and the F_st values between samples and neutrality tests (Tajima, 1989; Fu, 1996) were calculated using Arlequin 3.5.2.2 (Excoffier et al., 2005) for the samples with n ≥ 10. The general pattern of demographic history was described using the Bayesian Skyline Plot (BSP) method (Drummond et al., 2005), which estimates the changes in the effective population size over time. We used BEAST v. 1.10.4 for demographic analysis. The strict clock model and the Bayesian Skyline model with ten groups for the tree were used. The suitability of a simpler demographic model, i.e., the model of exponential growth, was rejected using the Marginal Likelihood Estimator method.

**Estimation of the evolutionary rate of cytb.** There is a severe problem while estimating the time of recent divergence events (<1–2 million years). It’s the potential distortion of divergence times estimates due to the changes in the apparent rate of substitutions over time (rate decay) (Ho et al., 2005). As a consequence, the phylogenetic rates obtained by the examination of interspecific divergence events may be inapplicable to the dating of intraspecific (population) events.

The expansion time was determined by applying the rate of divergence estimated for all cytb positions, which is 16% per million years. It was obtained as follows: firstly, it was assumed that the rate of substitutions by transversions for the 3rd codon position (tv3) of cytb changes insignificantly over time and tv3 occurred at a rate close to the phylogenetic one. This substitution rate had been obtained for Crocidura earlier (Lavrenchenko et al., 2009): 1.87% per million years; secondly, based on the ratio of different types of cytb substitutions in each species of the C. suaveolens s.lato group, the tv3 rate was estimated to be 23.4% of the rate of all substitutions at all positions. Hence, the substitution rate of all substitutions in cytb of the C. suaveolens s.lato group species suitable for assessing recent times (<100000 years ago) is 8%, and, therefore, the haplotype divergence rate is 16%.

**Reconstruction of the center of origin.** The potential center of origin of the ancestral mitotype was reconstructed using the principle of maximum parsimony, as was proposed in a similar work on the Common Shrew (Raspopova et al., 2020). The purpose of this analysis was to use the connection between the geographical distribution of mitotypes and their phylogenetic relationships in order to determine the geographic placement of the ancestral mitotype, that makes the dispersal scenario of mitochondrial lineages over the area as parsimonious as possible. We approached this task using the parsimony algorithm adapted for landmark data and implemented in the TNT program version v. 1.5 Sept. 2020 (Catalano et al., 2010; Goloboff and Catalano 2016). Since the geographic coordinates cannot be used directly to describe the position of points in two-dimensional space, they were converted into arbitrary Cartesian coordinates using the linear multidimensional scaling of the matrix of distances between sampling points using the cmdscale package in R. The permitted distribution of ancestral mitotypes was confined to the set of sampling points and the nodes of a 12 × 12 coordinate grid covering the modern range. In addition, it was necessary to address the uncertainty of the phylogeny between haplotypes. For this purpose, the analysis was repeated 100 times using a sample of trees obtained as a result of BEAST (BSP) analysis. Then, we analyzed the spatial distribution of reconstructed placements of the ancestral node.

The resultant Cartesian coordinates were converted into true latitude/longitude using the original script provided by one of the authors (I. Artyushin). The distribution density was obtained by the kernel density estimation (KDE) using the von Mises–Fischer distribution with kappa = 200 as a kernel. The calculations were made in the R v3.6.1 environment using the nls {stats} and distGeo {geosphere} functions and the raster package for visualization.

**RESULTS**

**Phylogenetic and phylogeographic analysis.** None of the phylogenetic analyses (BI, ML, MP, NJ) yielded a resolved haplotype tree; however, several potential groups can be distinguished. Figure 2 shows the Bayesian tree constructed in the Beast software, which is also a chronogram of the major evolutionary events in the studied species group. The median network (Fig. 3) includes 79 haplotypes: 71 from the Lesser White-toothed Shrew, seven from the Siberian Shrew, one of them is common for both forms. The network has a star-shaped structure; nevertheless, the same groups of haplotypes as on the tree can be found here.

Group I is the most heterogeneous one; it lacks the central, most common haplotype, and its haplotypes differ from each other in the greater number of substitutions. This group consists of the haplotypes from the Balkhash Lake region, the Zaisan Basin, and Dzun-
garia, as well as three solitary haplotypes from Middle Asia. The haplotypes of the Siberian White-toothed Shrew within group II form a distinct and compact haplogroup. Group III includes two broadly dispersed haplotypes (h10, h69) and derived haplotypes and can be divided into two subgroups: IIIa and IIIb. Subgroup IIIa is formed mainly by the haplotypes of Eastern Europe and also includes some of the haplotypes from Mongolia, Dzungaria, the Zaisan Basin, and the Balkhash Lake region. Subgroup IIIb contains most of the haplotypes from Mongolia and Dzungaria, the Zaisan and Balkhash haplotypes, and one haplotype from Astrakhan, as well as some part of Siberian haplotypes; moreover, the samples of C. sibirica from Krasnoyarsk and one sample from Novosibirsk carry haplotype h69, which is common among the samples from other localities, including the Gobi.

In general, the samples from Middle Asia are most unlinked to each other and distant from the samples from other geographic regions. The haplotypes in common to the northernmost and southernmost regions of the range are found only in the Balkhash Lake region and the Zaisan Basin. In the sample from Eastern Europe, the haplotypes common with the Asian samples (primarily from the Balkhash Lake region) can be found to the east from the left bank of the Volga River and are absent on the right bank and further in the western part of the range.

As can be seen from the data presented, the haplotypes from the range of C. sibirica are very similar to those of C. suaveolens s.str.; they do not form a single monophyletic group, and the Lesser White-toothed Shrew is paraphyletic relative to the Siberian Shrew with respect to mtDNA. Given this we further analyze and discuss the sample of specimens of the two species (C. suaveolens/C. sibirica) as a single sample.

**Times of major divergence events.** Figure 2 shows the times of the main divergence events in the C. suaveolens/C. sibirica group. With the accepted divergence rate of ~16% per million years, the latest common ancestor of all C. suaveolens s.str./C. sibirica lines should have existed 42,000 years ago. The C. sibirica line, which is found within group II, was split 19,600 years ago, and its radiation occurred 15,000 years ago.

**Genetic diversity of geographic populations.** We split our sample into ten main samples (n ≥ 4) (Fig. 4) based on geographical and genetic distances (p-distances, 0.22–0.72%) between the samples from local populations. Seven of them (n ≥ 10) were included in
the analysis of genetic diversity and demography: *C. suaveolens* (1) Europe 1: to the west of the right bank of the Volga River; (2) Europe 2: from the left bank of the Volga River to the Ural; (3) the Balkhash Lake region; (4) the Zaisan Basin; (5) Southern and Central Mongolia (the Gobi); (6) Dzungaria; (7) *C. sibirica* Altai, Krasnoyarsk, Kemerovo, and Novosibirsk (Siberia). All available data were used to analyze the total sample for Asia \( (n = 97) \). The changes in the effective population size over time were modeled only for samples of more than 25 specimens.

The indices of the molecular diversity and demographic parameters for the samples of at least ten specimens are presented in Tables 1 and 2. Statistical parameters of the cytb diversity in the geographical
samples from all East European populations and several Asian populations reveal the low haplotype and nucleotide diversity of the population of Eastern Europe compared to individual Asian populations and to the joined Asian sample (Table 1). Among large samples, the maximum haplotype diversity is apparent in the population of the Zaisan Basin; the Balkhash Lake region has the maximum nucleotide diversity. The high $h$ and $\pi$ values in Siberia are due to existence of two sympatric haplogroups in $C. sibirica$ ($p$-distance, 0.76%); the diversity within these groups is low. Despite the small sizes of the samples from Middle Asia ($n = 7$) and Kashgar ($n = 4$), let us note that the haplotype and nucleotide diversity in them is very high ($\pi = 0.0046 \pm 0.0029$ in Middle Asia and $\pi = 0.0050 \pm 0.0036$ in Kashgar) with $h = 1$ in both cases.

**Demographic analysis.** Tajima’s and Fu’s neutrality tests are significantly negative only for the European population, indicating its recent expansion. There is also a significant increase in the case of the total Asian sample, but it is not significant for individual geographic populations (Table 1). The Rg and SSD indices do not reject the sudden expansion model for any of the samples (Table 2).

The modeling of the effective population size dynamic (BSP) for the joint $C. suaveolens$ s.str./$C. sibirica$ group indicates weak growth. It is difficult to determine the exact time when it happened given the pattern obtained (Fig. 5). The results of the BSP analysis for the Gobi and eastern Kazakhstan samples reveals relatively recent growth for the Gobi population and a noticeably slower and substantially earlier growth in eastern Kazakhstan. We failed to obtain the results of the BSP analysis for the sample from Europe due to extremely low nucleotide and haplotype variability.

However, the twofold difference in $\tau_a$ (the time in nucleotide substitutions, Table 2) in the samples from Middle Asia and Kazakhstan compared to Eastern Europe indirectly indicates their greater age. If we recalculate the $\tau_a$ index into absolute time at a $cytb$ divergence rate of 16% per million years for the two samples with a significant expansion (joint Asian and joint European, Table 2), we can obtain an approximate time of dispersal. On the Asian territory, the expansion occurred earlier: about 24000 years ago; the East European part of the range was colonized no more than 12000 years ago, which corresponds to the end of the Late Pleistocene and the beginning of Holocene. The Mismatch distribution analysis yielded quite similar distributions of pairwise distances in the total sample from Eastern Europe and in the large populations from Asia (Fig. 6). At the same time, the shape of the distribution is certainly unimodal in the former case (i.e., in Europe, it does not deviate from the shape expected from the sudden expansion model), while in the large Asian samples (South/Central Mongolia, eastern Kazakhstan, joint Asian sample) the observed distribution begins to deviate visually from the expected shape. The samples from the Balkhash Lake and Zaisan regions have the least smooth distribution, though their Rg values do not confirm this. Splitting the European sample into sub-
samples from the left and right banks of the Volga River to analyze populations of comparable size and compactness from Eastern Europe and Asia had not affected the result.

**Ancestral area.** Reconstruction of the center of origin (Fig. 7) showed that the latest common ancestor of modern mitochondrial lines of the Lesser White-toothed Shrew could have inhabited Middle Asia (except for the southwest), as well as the territory of Kazakhstan and further to the west, up to the Northern Caspian. Potential points of origin have the highest density of distribution in the area from Semirechye to Western Gissar-Alai mountain ranges. The area of possible distribution of the ancestral mitotype does not include Mongolia, North China, Altai, Iran, Turkmenistan, and the Black Sea region.

### Table 1. The indices of cytb diversity in geographic populations of the shrews and the results of neutrality tests: sample size \((n)\), haplotype diversity \((h)\), nucleotide diversity \((\pi)\), Tajima’s \(D\) values and Fu’s \(F_s\) and \(p\)-values. Valid neutrality indices are in bold.

| Sample                                    | \(N\) | \(H\)         | \(\pi\)          | Tajima’s \(D\) | \(F_u\) \(F_s\)  | \(p\)-value | \(R_g\) | \(p\)-value |
|-------------------------------------------|-------|---------------|-------------------|----------------|-------------------|-------------|---------|-------------|
| Siberia                                   | 11    | 0.9455 ± 0.0535 | 0.0045 ± 0.0026   | 0.6434         | 0.7730            | −0.3012     | 0.4130  |             |
| Dzungaria                                 | 10    | 0.9333 ± 0.0773 | 0.0041 ± 0.0025   | −0.2817        | 0.4190            | −2.1931     | 0.0840  |             |
| S. and C. Mongolia (Gobi)                 | 35    | 0.9647 ± 0.0201 | 0.0026 ± 0.0015   | −1.2325        | 0.1120            | −1.9485     | 0.1660  |             |
| Zaisan                                    | 14    | 0.9890 ± 0.0314 | 0.0038 ± 0.0022   | 0.1945         | 0.6280            | 1.0705      | 0.7300  |             |
| Balkhash                                  | 11    | 0.9455 ± 0.0659 | 0.0052 ± 0.0030   | −1.2674        | 0.1090            | −2.2619     | 0.1040  |             |
| E. Kazakhstan (Balkhash + Zaisan)         | 25    | 0.9867 ± 0.0167 | 0.0045 ± 0.0025   | −1.3985        | 0.0720            | −4.3467     | 0.0390  |             |
| All of Asia                               | 97    | 0.9888 ± 0.0052 | 0.0043 ± 0.0024   | −2.1728        | 0.0010            | −25.500     | 0.0000  |             |
| Europe 1: from the Volga River to the Urals | 15    | 0.9524 ± 0.0344 | 0.0024 ± 0.0015   | −1.06072       | 0.1440            | −3.242      | 0.0300  |             |
| Europe 2: to the west of the Volga River  | 31    | 0.9075 ± 0.0419 | 0.0017 ± 0.0011   | −1.4555        | 0.0610            | −5.4720     | 0.0030  |             |
| All of Europe                             | 46    | 0.9362 ± 0.0271 | 0.0020 ± 0.0012   | −1.9196        | 0.0070            | −12.233     | 0.0000  |             |

### Table 2. The estimation of demographic parameters of the Lesser White-toothed Shrew under the sudden expansion model: \(tau\) and relative \(tau\) are the time in substitutions per gene and per site, respectively; SSD is the sum of squared deviations from the model, \(R_g\) is the raggedness index of distribution of pairwise differences in a sample.

| Sample                                    | \(n\) | \(tau\) and 90% CI | Relative \(tau\) (substitutions per site) | SSD | SSD \(p\)-value | \(R_g\) | \(R_g\) \(p\)-value |
|-------------------------------------------|-------|---------------------|------------------------------------------|-----|-----------------|---------|---------------------|
| Siberia                                   | 11    | 10.736 (4.645–99.736) | 0.009 (0.004–0.088)                       | 0.0285 | 0.5940 | 0.0499 | 0.8060              |
| Dzungaria                                 | 10    | 4.141 (2.469–10.053)  | 0.004 (0.002–0.0090)                      | 0.2390 | 0.1370 | 0.1630 | 0.2390              |
| S. and C. Mongolia (Gobi)                 | 35    | 3.473 (1.889–5.563)   | 0.003 (0.002–0.005)                       | 0.0016 | 0.8520 | 0.0147 | 0.9250              |
| Zaisan                                    | 14    | 6.022 (2.986–9.793)   | 0.005 (0.003–0.009)                       | 0.0062 | 0.9150 | 0.0167 | 0.9720              |
| Balkhash                                  | 11    | 4.000 (2.910–11.811)  | 0.004 (0.003–0.010)                       | 0.0171 | 0.2700 | 0.0433 | 0.5180              |
| E. Kazakhstan (Balkhash + Zaisan)         | 25    | 6.107 (3.332–8.994)   | 0.005 (0.003–0.008)                       | 0.0021 | 0.9200 | 0.0095 | 0.9660              |
| All of Asia                               | 97    | 3.758 (2.521–10.352)  | 0.003 (0.002–0.009)                       | 0.0013 | 0.7330 | 0.0060 | 0.9230              |
| Europe 1: from the Volga River to the Urals | 15    | 2.041 (1.764–5.494)   | 0.002 (0.002–0.005)                       | 0.0052 | 0.6740 | 0.0373 | 0.7540              |
| Europe 2: to the west of the Volga River  | 31    | 2.027 (1.041–3.137)   | 0.002 (0.001–0.003)                       | 0.0031 | 0.5360 | 0.0453 | 0.5730              |
| All of Europe                             | 46    | 1.937 (1.320–3.586)   | 0.002 (0.001–0.003)                       | 0.0009 | 0.4470 | 0.0364 | 0.2760              |
DISCUSSION

The chronogram and the structure of group I in the phylogenetic network (probably, it is the most basal branch of the tree; it lacks a central haplotype common to most specimens and has a significant number of substitutions between haplotypes) suggest that this group, which contains the Middle Asian, Iranian, and Kashgar haplotypes predominantly is older than groups II and III; i.e., the dispersal of shrews in the southern part of the range is an earlier event than the colonization of the western and northeastern parts of the modern range. The European part of the range was colonized most recently. This is indicated by the lowest haplotype and nucleotide diversity in the Eastern European sample, the specimens of which belong only to group III which has a simple star-shaped structure with the central, most common haplotype. On the contrary, the haplotypes from eastern Kazakhstan occur in all groups of haplotypes and the nucleotide diversity of the eastern Kazakhstan populations is one of the highest, which indicates that it can be potentially one of if not the oldest one. The results of the ancestral locality reconstruction for the species are also consistent with the hypothesis of the Middle Asian and/or Kazakhstanian origin of the Lesser White-toothed Shrew and reveal the recent origin of the Inner Asian populations (Mongolia, China) and even later colonization of the European part of the range as a result of rapid expansion from the territory of Kazakhstan.

Thus, our data show that the Lesser White-toothed Shrew has colonized its range very quickly and relatively recently. Apparently, this is why phylogeographic breaks had not enough time to emerge over the vast and ecologically diverse territory and thus the variation in color and size reflects only the morphological adaptation to arid or humid landscapes and does not have an initial genetic basis.

The Lesser White-toothed Shrew resembles by its range type and its landscape and habitat preferences, certain other Palearctic species of small mammals. The Grey Dwarf Hamster, *Cricetulus migratorius* (Pallas 1773), inhabits steppes, forest steppes, deserts, and semi-deserts from Greece to southeastern Mongolia inclusive and from Central Russia to Balochistan (Lebedev, 2012). Thus, its range covers almost entirely the range of the Lesser White-toothed Shrew, though it doesn’t reach as far north and does not occur in the forest zone. The Grey Hamster is ecologically flexible and, similarly to the Lesser White-toothed Shrew, inhabits oases in the arid zone. It was shown that the species has three allopatric mitochondrial lineages
Fig. 6. The diagrams of distribution of pairwise genetic distances (Mismatch distribution) in the major samples of the Lesser White-toothed Shrew. The solid line is the actual distribution; the dashed line is the distribution according to the sudden expansion model. The X-axis is the distance in the number of substitutions; the Y-axis is the frequency of this distance.

(Lebedev et al., 2018) with a phylogeographic break along the Volga River between two broad-range forms (genetic distance of 2.5%). The Lesser White-toothed Shrew, unlike the hamster, has probably just recently colonized the territory to the west of the Volga River. The territories to the east of the Volga River and up to Mongolia are inhabited by the eastern lineage of the Grey Dwarf Hamster, and its phylogeographic structure resembles that of the Lesser White-toothed Shrew, with its star-shaped pattern and low nucleotide
diversity. Like the Lesser White-toothed Shrew, the Grey Dwarf Hamster has relatively recently populated Inner Asia (Mongolia) from Middle Asia and/or Kazakhstan; its expansion in the eastern part of current range actually coincides with that of the Lesser White-toothed Shrew and is estimated to be 25000 years ago (Lebedev et al., 2018).

We can find the common features in the shape of the recent range and the direction of dispersal between *Crocidura suaveolens* s.lato and the species complex of the House Mouse (*Mus musculus* s.lato). In particular, the area of the "musculus" form actually coincides with the area of *C. suaveolens* s.str., with the exception of the southernmost (Iran) and western (Central Europe) parts, and there are some similarities in the phylogeographic structure (Bonhomme and Searle, 2012). As in the case of the Lesser White-toothed Shrew, Middle Asia was colonized by the House Mouse before China, Mongolia, and Europe; published datings suggest an expansion close to the Pleistocene–Holocene boundary (Suzuki et al., 2013; Hardouin et al., 2015; Li et al., 2020).

The status of the Siberian White-toothed Shrew is still unclear. *C. sibirica* bears two mitochondrial lineages that are not connected with particular geographic localities and are not monophyletic, i.e., they are not species specific (it is similar to previous studies). One of these lineages (from the common group II) is distributed mainly in the western part of the range (Novosibirsk, Kemerovo) and the second one (the lineage from the common group IIIb) inhabits mainly the southeastern part (Krasnoyarsk, Teletskoe Lake).

The western Siberian genetic line, contains more haplotypes, compared to the southeastern line, none of them are found within the range of the Lesser White-toothed Shrew. This line branched ~20000 years ago. If we take this date as the time of divergence between the haplotypes of *C. sibirica* and *C. suaveolens* s.str., it is hardly reasonable to assume the species status of the Siberian form due to very recent time of its origin. The absence of the mitochondrial geographic structure also suggests recent genesis of the modern range.

In this case, the Siberian Shrew is only a morphotype of the Lesser White-toothed Shrew isolated during the last glacial maximum, in a steppe refugium bordering the forest e.g. in Altai mountains. Fixation of the morphological differences in such population, even those associated with adaptation to living in a forest environment, doesn’t necessarily lead to general divergence and reproductive isolation from the sister form. The existence of the second lineage (which is more recent and similar to the Mongolian haplotypes) can be explained by hybridization with the Lesser White-toothed Shrew during the secondary contact. Given the broad range of *C. suaveolens* s.str. and much smaller one of *C. sibirica*, which is also located at the northernmost limit of the Lesser White-toothed Shrew’s range, it’s reasonable to assume the directional introgression of mtDNA from the Lesser White-toothed Shrew to the Siberian Shrew.

Fig. 7. The distribution of tentative localities of the ancestral mitotype of the Lesser White-toothed Shrew. The background intensity corresponds to the density of the empirical distribution obtained by the KDE method.
Another hypothesis suggests that *C. sibirica* is actually a species, that lost it’s native mtDNA, since it was completely replaced by the mtDNA of the Lesser White-toothed Shrew due to hybridization during its active colonization of the current range. The coexistence of two mitochondrial lineages in this case can be explained by two hybridization events with the ancestral form (Lesser White-toothed Shrew) during non-simultaneous secondary contacts. We expect that a larger number of nuclear loci will help to unveil origin of discrepancy between the genetic and morphological variability of the Siberian Shrew.

ACKNOWLEDGMENTS

The authors are grateful to Yu.N. Litvinov, F.M. Gole
nishchev, V.G. Malikov, M.V. Kasatkin, Yu.M. Kovalskaya, A.V. Surov, and other colleagues involved in the collection of materials for genetic analysis.

FUNDING

This work was supported by the Russian Foundation for Basic Research, project no. 20-04-00081a.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of interest. The authors declare that they have no conflicts of interest.

Statement on the welfare of animals. All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

OPEN ACCESS

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

REFERENCES

Bandelt, H.J., Forster, P., and Röhl, A., Median-joining networks for inferring intraspecific phylogenies, *Mol. Biol. Evol.*, 1999, vol. 16, no. 1, pp. 37–48.

Bannikova, A.A., Oleinichenko, V.Yu., Lomov, A.A., and Dolgov, V.A., On the taxonomic relationship of the Lesser White-toothed Shrew and Gueldenstaedt’s Shrew, *Crocidura suaveolens*, and *C. gueldenstaedtii* (Insectivora, Soricidae), *Zool. Zh.*, 2001, vol. 80, no. 6, pp. 721–730.

Bannikova, A.A., Lebedev, V.S., Kramerov, D.A., and Zaitsev, M.V., Phylogeny and systematics of the *Crocidura suaveolens* species group: corroboration and controversy between nuclear and mitochondrial DNA markers, *Mammalia*, 2006, vol. 70, nos 1–2, pp. 106–119.

Bannikova, A.A., Sheftel’, B.I., Lebedev, V.S., Aleksandrov, D.Yu., and Myullenberg, M., *Crocidura shantungensis*, a new species for Mongolia and Buryatia, *Dokl. Biol. Sci.*, 2009, vol. 424, pp. 68–71.

Bannikova, A.A., Abramov, A.V., Borisenko, A.V., Lebedev, V.S., and Rozhnov, V.V., Mitochondrial diversity of the white-toothed shrews (Mammalia, Eulipotyphla, Crocidura) in Vietnam, *Zootaxa*, 2011, vol. 2812, no. 1, pp. 1–20.

Bekenov, A., Butovsksii, P.M., Kasabekov, B.V., Lankin, P.M., Strelkov, P.P., Stogov, I.I., Fedosenko, A.K., Shaimardanov, R.T., and Shubin, I.G., *Mlekopitayushchie Kazakhstana* (Mammals of Kazakhstan), vol. 4: *Nasekomoyadnye i rukokrylye* (Insectivores and Chiropterans), Gvozdev, E.B. and Stratuman, E.I., Eds., Kazakh SSR, Alma-Ata: Nauka, 1985.

Bonhomme, F. and Searle, J.B., House mouse phylogeography, *in Evolution of the House Mouse*, Cambridge Series in Morphology and Molecules, Machola’n, M., Baird, S.J.E., Munclinger, P., and Pi’take, J., Eds., Cambridge, UK: Cambridge University Press, 2012, pp. 278–296.

Burgin, C., He, K., Haslauer, R., Sheftel’, B.I., Jenkins, P.D., Ruedi, M., Hintscbe, S., Motokawa, M., Hinckley, A., and Hutterer, R., Species accounts of Soricidae, in *Handbook of the Mammals of the World*, vol. 8: Insectivores, Sloths and Colugos, Wilson, D.E. and Mittermeier, R.A., Eds., Spain, Barcelona: Lynx Edicions, 2018, pp. 332–551.

Catalano, S.A., Goloboff, P.A., and Giannini, N.P., Phylogenetic morphometrics (I): the use of landmark data in a phylogenetic framework, *Cladistics*, 2010, vol. 26, pp. 539–549.

Catzelis, F., Maddalena, T., Hellwing, S., and Vogel, P., Unexpected findings on the taxonomic status of east Mediterranean crocidura russula auct. (mammalia, insectivora), *Zeitschrift Sautergierkunde*, 1985, vol. 50, no. 4, pp. 185–201.

Chen, S., Qing, J., Liu, Z., Liu, Y., Tang, M., Murphy, R.W., Pu, Y., Wang, X., Tang, K., Guo, K., and Jiang, X., Multi-locus phylogeny and cryptic diversity of white-toothed shrews (Mammalia, Eulipotyphla, Crocidura) in China, *BMC Evol. Biol.*, 2020, vol. 20, no. 1, pp. 1–4.

Drummond, A.J., Rambaut, A., Shapiro, B., and Pybus, O.G., Bayesian coalescent inference of past population dynamics from molecular sequences, *Mol. Biol. Evol.*, 2005, vol. 5, pp. 1185–1192.

Drummond, A.J. and Rambaut, A., BEAST: Bayesian evolutionary analysis by sampling trees, *BMC Evol. Biol.*, 2007, vol. 7, p. 214.

Dubey, S., Zaitsev, M., Cosson, J.-F., Abdukadier, A., and Vogel, P., Pliocene and pleistocene diversification and multiple refugia in a Eurasian Shrew (*Crocidura suaveolens* group), *Mol. Phylogenet. Evol.*, 2006, vol. 38, no. 3, pp. 635–647.

Dubey, S., Cosson, J.-F., Magnanou, E., Vohralik, V., Bendal, P., Frynta, D., Hutterer, R., Vogel, V., and Vogel, P., Mediterranean populations of the Lesser White-toothed Shrew (*Crocidura suaveolens* group): an unexpected puzzle of Pleistocene survivors and prehistoric introductions, *Mol. Ecol.*, 2007a, vol. 16, no. 16, pp. 3438–3452.
Dubey, S., Nová, P., Vogel, P., and Vohralík, V., Cytogenetic and molecular relationships between Zarudny’s Rock Shrew (Crocidura zarudnyi; Mammalia: Soricomorpha) and Eurasian taxa, *J. Mammal.*, 2007b, vol. 88, no. 3, pp. 706–711.

Excoffier, L., Laval, G., and Schneider, S., Arlequin (version 3.0): an integrated software package for population genetics data analysis, *Evol. Bioinf.*, 2005, vol. 1, pp. 47–50.

Fu, Y., New statistical tests of neutrality for DNA samples from a population, *Genetics*, 1996, vol. 143, no. 1, pp. 557–570.

Goloboff, P.A. and Catalano, S.A., TNT version 1.5, including a full implementation of phylogenetic morphometrics, *Cladistics*, 2016, vol. 32, no. 3, pp. 221–238.

Grafodatskii, A.S., Radzhabli, S.I., Sharshov, A.V., and Zaitsev, M.V., Karyotypes of five species of white-toothed shrews of the fauna of the USSR, *Tsitologiya*, 1988, vol. 30, no. 10, pp. 1247–1250.

Hardouin, E.A., Orth, A., Teschke, M., Darvish, J., Tautz, D., and Bonhomme, F., Eurasian house mouse (*Mus musculus* L.) differentiation at microsatellite loci identifies the Iranian plateau as a phylogeographic hotspot, *BMC Evol. Biol.*, 2015, vol. 15, no. 26.

Ho, S.Y.W., Phillips, M.J., Cooper, A., and Drummond, A.J., Time dependency of molecular rate estimates and systematic overestimation of recent divergence times, *Mol. Biol. Evol.*, 2005, vol. 22, pp. 1561–1568.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S., UFBoot2: improving the ultrafast bootstrap approximation, *Mol. Biol. Evol.*, 2017, vol. 35, no. 2, pp. 518–522.

Hoffmann, R.S., Noteworthy shrews and voles from the Xizang-Qinghai Plateau, in *Contribution in Mammalogy: A Memorial Volume Honoring Dr. J. Knox Jones, Jr.*, 1996, pp. 155–168.

Hutterer, R., Order Insectivora, in *Mammal Species of the World: A Taxonomic and Geographic Reference*, Wilson, D.E., and Reeder, D.M., Eds., Washington, DC: Smithsonian Institution Press, 1993, pp. 69–130.

Jiang, X. and Hoffmann, R.S., A revision of the white-toothed shrews (Crocidura) of Southern China, *J. Mammal.*, 2001, vol. 82, no. 4, pp. 1059–1079.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K., MEGA X: molecular evolutionary genetics analysis across computing platforms, *Mol. Biol. Evol.*, 2018, vol. 35, no. 6, pp. 1547–1549.

Lavrenchenko, L.A., Bannikova, A.A., and Lebedev, V.S., Shrews (*Crocidura spp.*) endemic to Ethiopia: recent adaptive radiation of an ancient lineage, *Dokl. Biol. Sci.*, 2009, vol. 424, no. 5, pp. 705–708.

Lebedev, V.S., Family Cricetidae, in *Mlekopitayushchie Rossii: sistematika-geograficheskii spravochnik* (Mammals of Russia: A Taxonomic-Geographical Reference Book), Pavlinov, I.Ya. and Lisovskii, A.A., Eds., Sb. Tr. Zool. Muz. MGU, 2012, vol. 52, pp. 211–220.

Lebedev, V., Poplavskaya, N., Bannikova, A., Ryurikov, G., and Surov, A., Genetic differentiation in *Cricetulus migratorius* Pallas, 1773 (Rodentia, Cricetidae), *Mammal. Biol.*, 2018, vol. 92, no. 1, pp. 115–119.

Li, Y., Fujiwara, K., Osada, N., Kawai, Y., Takada, T., Kryukov, A.P., Abe, K., Yonekawa, H., Shiroishi, T., Moriwaka, K., and Saitou, N., House mouse *Mus musculus* dispersal in East Eurasia inferred from 98 newly determined complete mitochondrial genome sequences, *Heredity*, 2020, pp. 1–16.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q., IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies, *Mol. Biol. Evol.*, 2015, vol. 32, no. 1, pp. 268–274.

Ohdachi, S.D., Iwasa, M.A., Nesterenko, V.A., Abe, H., Masuda, R., and Haberl, W., Molecular phylogenetics of Crocidura shrews (Insectivora) in east and central Asia, *J. Mammal.*, 2004, vol. 85, no. 3, pp. 396–403.

Raspopova, A.A., Bannikova, A.A., Sheftel, B.I., Krystufek, B., Koupstov, A.V., Illarionova, N.A., Pavlova, S.V., and Lebedev, V.S., A never-ending story of the Common Shrew: searching for the origin, *Mammal. Res.*, 2020, vol. 65, pp. 729–742.

Sambrook, J., Fritsch, E.F., and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Laboratory Press, 1989.

Sheftel, B.I., Bannikova, A.A., Fang, Y., Demidova, T.B., Alexandrov, D.Y., Lebedev, V.S., and Sun, Y.H., Notes on the fauna, systematics, and ecology of small mammals in Southern Gansu, *China. Biol. Bull.* (Moscow), 2018, vol. 45, no. 8, pp. 898–912.

Suzuki, H., Nunome, M., Kinoshita, G., Aplin, K.P., Vogel, P., Kryukov, A.P., Jin, M.-L., Han, S.-H., Maryanto, I., Tsuchiya, K., Ikeda, H., Shiroishi, T., Yonekawa, H., and Moriwaki, K., Evolutionary and dispersal history of Eurasian house mice *Mus musculus* clarified by more extensive geographic sampling of mitochondrial DNA, *Heredity*, 2013, vol. 111, pp. 375–390.

Tajima, F., Statistical method for testing the neutral mutation hypothesis by DNA polymorphism, *Genetics*, 1989, vol. 123, no. 3, pp. 585–595.

Tembotova, F.A., Patterns of variability and evolution of insectivorous mammals of the Caucasus, *Doctoral (Biol.) Dissertation*, Na’chik: KBNTs RAN, 1999.

Vogel, P., Maddalena, T., and Catzeffis, F., A contribution to the taxonomy and ecology of shrews (*Crocidura zimmermanni* and *C. suaveolens*) from Crete and Turkey, *Acta Theriol.*, 1986, vol. 31, no. 39, pp. 537–545.

Vogel, P., Cosson, J.F., and Jurado, L.F.L., Taxonomic status and origin of the shrews (Soricidae) from the Canary Islands inferred from a mtDNA comparison with the European *Crocidura* species, *Mol. Phylogenet. Evol.*, 2003, vol. 27, no. 2, pp. 271–282.

Zaitsev, M.V., Species composition and issues of taxonomy of white-toothed shrews (Mammalia, Insectivora) of the fauna of the USSR, *Tr. Zool. Inst. Akad. Nauk SSSR*, 1991, vol. 243, pp. 3–46.

Zaitsev, M.V., Voitta, L.L., and Sheftel’, B.I., *Mlekopitayushchie fauny Rossii i sopredel’nykh territorii. Nasekomoyadnye* (Mammals of the Fauna of Russia and Adjacent Territories. Insectivores), Opredeliteli Faune Rossii, St. Petersburg: Zool. Inst. Ross. Akad. Nauk, 2014, no. 178.

Translated by E.V. Makeeva