Published by Associazione Teriologica Italiana
Hystrix, the Italian Journal of Mammalogy
Volume 34 (1): 33–38, 2023
OPEN ACCESS
Available online at:
http://www.italian-journal-of-mammalogy.it
doi:10.4404/hystrix-00590-2022

Research Article

# The first phylogenetic data on the elusive shrews of the *Crocidura pergrisea* species complex

Anna Andreevna Bannikova[1,*], Alexandra Andreevna Lisenkova[1], Evgeniya Nikolaevna Solovyeva[2], Alexei Vladimirovich Abramov[3], Boris Ilyich Sheftel[4], Boris Kryštufek[5], Vladimir Svyatoslavovich Lebedev[2]

[1]*Department of Vertebrate Zoology, Lomonosov Moscow State University, Vorobievy Gory, 119991 Moscow, Russia*
[2]*Zoological Museum of Moscow State University, 125009 Moscow, Russia*
[3]*Zoological Institute, Russian Academy of Sciences, Universitetskaya nab. 1, Saint Petersburg 199034, Russia*
[4]*A.N. Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Leninskii pr., 33, Moscow, Russia*
[5]*Slovenian Museum of Natural History, Prešernova 20, SI-1000 Ljubljana, Slovenia*

*Keywords:*
*Crocidura*
molecular phylogeny
South West Asia

*Article history:*
Received: 20 November 2022
Accepted: 21 April 2023

*Acknowledgements*
The sample of *Crocidura arispa* was kindly provided by Dr. Frank E. Zachos, Natural History Museum of Vienna. The work was funded by the Russian Scientific Foundation, project 21-14-00007. Participation of BK was supported by program P1-0255. We are grateful to anonymous reviewers for their detailed and useful remarks that helped to make our work much better.

**Abstract**

Within the most speciose genus of extant mammals – *Crocidura*, the *pergrisea* species complex distributed in South West Asia remains the least studied, largely due to the rarity of its representatives. We examined the phylogenetic position of two putative species of the pergrisea species complex (*C. serezkyensis* and *C. arispa*) using historical DNA isolated from museum specimens. On the basis of sequence data for two nuclear and one mitochondrial genes we have come to the following conclusions: *C. serezkyensis* and *C. arispa* are rather close to each other and belong to a separate lineage of white-toothed shrews for which *C. ramona* from Israel is the relatively close sister branch. The *pergrisea* species complex does not include *C. zarudnyi*, which was previously shown to be close to *C. suaveolens*. The clade including *C. pergrisea* species complex and *C. ramona* likely belongs to a large Afro-Mediterranean clade, which includes also the Afromontane clade, the Mediterranean clade and *C. leucodon*. The problems of systematics within the *pergrisea* species complex are discussed.

## Introduction

Over the past two decades, the field of molecular phylogenetics of mammals has been so successful, that there are currently almost no unresolved phylogenetic nodes at the level of high-ranking taxa. However, there are many large genus level taxa in which complex relationships between species and species groups remain unclear. Among such problematic groups is the genus *Crocidura*, which is the most speciose genus within the family Soricidae as well as the most speciose genus of extant mammals, represented by more than 200 currently recognized species distributed throughout Eurasia and Africa (Burgin et al., 2018; Esselstyn et al., 2021).

The molecular phylogenetic hypothesis of the genus *Crocidura* was proposed on the basis of two mitochondrial and two nuclear genes (Dubey et al., 2008). Since then, the study of shrew phylogenetics and systematics is directing taxonomic diversity and evolutionary history of individual species and / or regional groups in Southeast Asia or Africa. Less studies focus on shrews of China, and there are only few works on the Middle Eastern part of the distribution range of the genus (Dubey et al., 2007a,b; Shpirer et al., 2021; İbiş et al., 2022). Besides, many species groups are still ignored. For example, genetic diversity within the complexes of Palearctic species: *C. leucodon*, *C. russula*, and *C. suaveolens*, have so far been studied mainly only on the basis of the mitochondrial DNA and the *C. pergrisea* species complex has never been studied by any molecular methods.

Shrews of the *C. pergrisea* species complex were named as “flat-headed rock-shrews" by Burgin et al. (2018). These extremely rare petrophilic shrews, which occur in the mountains of the Caucasus, the Middle East and Central Asia, have a mode of life that is not typical for other *Crocidura* species, as most of the specimens were collected in caves, gorges and stone screes (Spitzenberger, 1971; Bekenov et al., 1985). These shrews are scarce in museum collections since their capture rate is several orders lower than those of the *C. suaveolens* species complex, which are sympatric with *pergrisea* in most of its distribution range. According to Zaitsev (1993), the total number of vouchers of the flat-headed rock-shrews in the world museums was below 30. Descriptions of various species were based on few specimens and frequently on a single specimen. It therefore poses no surprise that their taxonomic status is still uncertain. Phylogenetic position of the *pergrisea* species complex is further puzzled by a lack of cytogenetic and molecular studies for the majority of species in this group.

According to the morphological revision by Zaitsev (1993) the *pergrisea* species complex includes the following taxa: pale gray shrew *C. pergrisea* Miller, 1913; lesser rock shrew *C. serezkyensis* Laptev, 1929; Armenian shrew *C. armenica* Gureev, 1963; and Zarudny’s rock shrew *C. zarudnyi* Ognev, 1928. The Jackass shrew was included in the nominal species as a subspecies *C. arispa* (Spitzenberger, 1971). Later, it was given full species rank by Kryštufek and Vohralík (2001). As for Zarudny‘s rock shrew, previously it was also considered a subspecies of *C. pergrisea* by Ellerman and Morrison-Scott (1951) and Lay (1967). However, mitochondrial data from a single specimen from Baluchestan identified as Zarudny‘s rock shrew showed that this species is likely the closest sister group of the lesser white-toothed shrew *C. suaveolens* species complex, which is also confirmed by the similarity of their karyotypes (Dubey et al., 2007a). Based on these results one may expect that if *C. zarudnyi* is really a member of the “pergrisea" species complex the latter should be related to the *C. suaveolens* species complex.

In this study, we elucidate the phylogenetic position of the *pergrisea* species complex within Eurasian crocidurine shrews based on the his-

*Corresponding author
*Email address:* hylomys@mail.ru (Anna Andreevna Bannikova)

torical museum DNA of several potential species of the group using the mitochondrial *cytb* gene and two nuclear genes. In comparison with Dubey et al. (2008) we have increased the number of sequences of African species including the Eastern Afromontane group (endemic to Ethiopia and Tanzania); added the endemics of East Indochina (Vietnam) to the species of SE Asia; and studied the phylogenetic relationships of the Palearctic *C. pergrisea* species complex for the first time.

## Materials and methods

### Taxon sampling and tissue collection

The materials used in our study are presented in the Table 1, Table S1 and Fig. S4. In total, we obtained sequences of one mitochondrial and two nuclear genes from 24 recent and four historical specimens of 20 species of *Crocidura* and *Diplomesodon pulchellum.* For the phylogenetic analysis of the combined datasets, 327 sequences of mitochondrial and two nuclear genes of different species of *Crocidura* and *Suncus* from GenBank were used (Table S1). For the separate analysis of the *cytb* gene, 1397 sequences were retrieved from GenBank (Fig. S4).

The sample of *Crocidura* species complex includes two specimens of *C. serezkyensis* and a single specimen (paratype) of *C. arispa*. The *C. pergrisea* species proper was not included in our analysis. However, based on its undoubted morphological similarity with *C. serezkyensis* and *C. arispa* (Spitzenberger, 1971; Zaitsev, 1993; Kryštufek and Vohralík, 2001) we apply the name *pergrisea* species complex for the studied group.

Most of newly acquired recent specimens included in this study were collected as part of the biodiversity surveys conducted by the Joint Vietnam–Russian Tropical Research and Technological Centre, the 2011 Joint Russian-Mongolian Biological expedition and over the course of fieldwork in Kazakhstan, Central Russia, Siberia, Russian Far East and North Caucasus. The vast majority of the samples belong to the collection of alcohol tissues of ZMMU. Among them, tissues of *C. mimula* from Italy and Germany and *C. russula* from Spain were kindly provided by Dr. P. Vogel during our personal scientific collaboration in 2004–2006. A large part of the sample set was obtained by small tissue biopsies of live-trapped animals. Voucher specimens were deposited in the Zoological Museum of Lomonosov Moscow State University (ZMMU) and the Zoological Institute of the Russian Academy of Sciences (ZIN, Saint Petersburg).

### DNA extraction, PCR amplification, and sequencing of the recent specimens

Genomic DNA from ethanol-preserved tissues of the recent specimens was extracted using a standard protocol of proteinase K digestion, phenol-chloroform deproteinization, and propanol precipitation (Sambrook et al., 1989). We sequenced the complete mitochondrial cytochrome *b* (*cytb*) gene and fragments of two nuclear loci: exon 26 of the apolipoprotein B (*ApoB*) and exon 11 of the breast cancer type 1 susceptibility protein (*BRCA1*). Primers and polymerase chain reaction protocols for the amplification and sequencing of *cytb* and *BRCA1* from recent samples are described in Bannikova et al. (2011, 2021). The primers for amplification and sequencing of *ApoB* gene are given in Esselstyn et al. (2009). PCR products were sequenced on an ABI 3100-Avant autosequencing using ABI PRISM®BigDyeTM Terminator v. 3.1 (Applied Biosystems, Foster City, CA, United States).

### DNA extraction, PCR amplification, and sequencing of the historical specimens

Total DNA of *C. serezkyensis* (S111841 and S111842) and *C. ramona* (S165754) was extracted from dried skin of museum vouchers deposited in the collection of ZMMU. DNA of *C. arispa* was isolated from a fragment of dried skin of a paratype (NMV 13284) deposited in the Natural History Museum, Vienna.

Extraction of the *C. ramona* DNA was performed using the protocol for tissues of QIAamp DNA Investigator kit (Qiagen). DNA from the *C. serezkyensis* and *C. arispa* samples was extracted using QIAamp DNA Mini Kit (Qiagen) for the former and QIAquick PCR Purification Kit (Qiagen) with a modified lysis buffer containing 0.5M EDTA, 10 % SDS and proteinase K for the latter. For all samples lysis continued overnight in a thermoshaker at 56 °C and 700 rpm. The DNA was then eluted in 100 $\mu$l of elution buffer provided in the corresponding kit. A blank was also used as a negative control in each extraction procedure.

The short fragments of *cytb*, *BRCA1*, and *ApoB* genes were amplified using the combinations of primers designed for this study (Tab. S2) or published previously (Bannikova et al., 2011, 2021; Esselstyn et al., 2009). The fragments of *cytb* gene were obtained with the primers L14723_Cr and L14728_Cr (Bannikova et al., 2011) and six new primers listed in Table S2. Two *ApoB* gene fragments were amplified at annealing temperature of 50º C using primer combinations ApoB-F1 Esselstyn et al. (2009)) / R327a (Tab. S2) and ApoB-R1 (Esselstyn et al., 2009) / F238b (Tab. S2). Four *BRCA1* gene fragments were amplified at annealing temperature of 52 °C with the following combinations of original and previously designed primers: F60 (Bannikova et al., 2021) / R262b (Tab. S2); F164a / R400a; F606a / R843a (Bannikova et al., 2021); F330a / R636a.

Amplification of mitochondrial and nuclear fragments was performed in 20 $\mu$l reaction volume containing 2–3 $\mu$l DNA, 4 $\mu$l ScreenMix-HS PCR master mix (Evrogen, Russia) and 1 $\mu$l of each primer (10 pmol/$\mu$l). Double-stranded polymerase chain reaction PCR was performed under the following conditions: 95 °C for 3 min; ( 95 °C for 30 sec, 50–54 °C for 30 sec, and 72 °C for 30 sec) x45 cycles; 72 °C for 7 min; hold at 12 °C. Negative controls were used for both extraction and PCR. The PCR products were visualized and verified in 1.5 % agarose gel. Amplicons were sequenced directly by Sanger sequencing on Applied Biosystems®3130xl Genetic Analyzer. Each fragment was sequenced several times from both forward and reverse primers to ensure the authenticity of the sequence.

It is known that DNA undergoes degradation over time (Hofreiter et al., 2001) and accumulates postmortal mutations due to the conditions in which the sample was preserved. Thus, depending on the conservation method, age of the sample, its history and many other factors, mtDNA from even relatively recent museum specimens can produce PCR and sequencing artifacts (Sefc et al., 2007). One of the most common postmortal mutations, considered a hallmark and a defining feature of ancient DNA, involves C→T substitutions, presumably due to the deamination of cytosine bases in the template (Hofreiter et al., 2001). Here, neither clear double C/T peaks nor an excess of C→T transitions were observed in the *cytb* sequences obtained from the historical *Crocidura* samples relative to modern samples, suggesting that the substitutions observed in them are authentic.

As a result of the sequencing of historical museum specimens, we obtained the following fragments of three genes: *cytb* – 1140 bp for *C. serezkyensis* and two fragments of 469 bp and 629 bp for *C. arispa*; *ApoB* – 559 bp and *BRCA1* – 744 bp for *C. serezkyensis*, *C. arispa* and *C. ramona*. The sequences obtained in this study can be accessed via GenBank under accession numbers: OP599553–OP599614 (Table 1).

### Alignment, partitioning and datasets

All sequences were aligned by eye using Bioedit version 7.0.9.0 (Hall 1999). In all analyses, sequences were used as unphased genotypes. Heterozygous positions (at which two peaks of approximately equal intensity are observed) were coded using the IUB ambiguity codes. Phylogenetic analyses were performed on the dataset consisting of 116 operational taxonomic units (OTUs) of 90 taxa with complete sequence data (i.e., concatenated two nuclear and one mitochondrial concatenated genes, – dataset I). In this dataset, only transversions of the 3$^{rd}$ codon positions of *cytb* were used; the 1$^{st}$ and 2$^{nd}$ codon positions were removed from the analyses to balance the number of informative sites in the mitochondrial and nuclear genes and avoid the selection signal. Additionally, we performed the analyses on a larger taxon sampling (143 OTUs of 106 taxa) but two nuclear genes only were performed (*ApoB* and *BRCA1* – dataset II). Finally, the largest dataset contained 1400 sequences of *cytb* (all codon positions – dataset III) of 127 species of *Crocidura*, *Diplomesodon pulchellum* and *Paracrocidura schoutedeni*.

**Table 1** – List of original material used in the study. The asterisks mark our sequences obtained in the previous studies.

| species/specimen and name in the concatenation | specimen code | specimen vaucher | locality | GB Acc No | | |
|---|---|---|---|---|---|---|
| | | | | *cytb* | *ApoB* | *BRCA1* |
| *C. arispa* | NHMV13284 | NHMV13284 | Turkey, Ulukişla, Niğde | OP599553 | OP599590 | OP599565 |
| *C. serezkyensis* 1 | 111841 | ZMMU:S111841 | Tajikistan, Pamir, Peter I Range, Pashimgar | OP599554 | OP599591 | OP599566 |
| *C. serezkyensis* 2 | 111842 | ZMMUS111842 | Tajikistan, Pamir, Peter I Range, Pashimgar | OP599555 | OP599592 | OP599567 |
| *C. ramona* | S165754 | ZMMU:S165754 | Israel:Makhtesh Ramon | – | OP599593 | OP599568 |
| *C. leucodon* 1 | 907A_Dag | – | Russia: Dagestan, Majalis | OP599556 | MW381917∗ | MW410132∗ |
| *C. russula* 1 | Spain1 | – | Spain: Andalusia, Sierra Nevada | – | MW381918∗ | MW410133∗ |
| *C. shantungensis* 1 | AVA16-175 | ZIN:104244 | Russia: Southern Primorye, Khasan region | OP599559 | OP599603 | OP599578 |
| *C. shantungensis* 2 | Khingan16 | – | Mongolia:Khingan Range, Nutryk River | OP599558 | OP599602 | OP599577 |
| *C. gueldenstaedtii* | Cgu_KA2 | – | Azerbaijan: Kyzyl-Agach | – | OP599606 | OP599581 |
| *C. mimula* 1 | It_A-0V-21-A | – | Italy: Venice, Latisana | AY994388∗ | OP599604 | OP599579 |
| *C. mimula* 2 | C_mimula_DD | – | Germany | OP599560 | OP599605 | OP599580 |
| *C. caspica* 1 | Ccasp2003-1 | – | Azerbaijan: Talysh | OP599561 | OP599607 | OP599582 |
| *C. caspica* 2 | Ccasp2003-31 | – | Azerbaijan: Lenkoran | AY994370∗ | OP599608 | OP599583 |
| *C. suaveolens* 1 | Csu_M11-80 | ZMMU:S189056 | Mongolia: Gun-Tamga Bulag | – | OP599609 | OP599584 |
| *C. suaveolens* 2 | Csu_Crimea1 | – | Crimea | – | OP599610 | OP599585 |
| *C. sibirica* 1 | Csib_Tel07-15 | – | Russia:Altai Republic, Teletskoye Lake | HM586994∗ | OP599611 | OP599586 |
| *C. sibirica* 2 | Csib_1651 | – | Russia:Kemerovo | AY994389∗ | OP599612 | OP599587 |
| *C. dracula* | B51 | ZIN:101406 | Viet Nam: Yen Bai Province | – | OP599594 | OP599569 |
| *C. attenuate* 2 | CatBa12-6 | ZIN:100769 | Viet Nam: Cat Ba Isl | – | OP599596 | OP599571 |
| *C. wuchihensis* 1 | AVA15-305 | ZIN:103280 | Viet Nam: Tuyen Quang Province, Na Hang | OP599557 | OP599595 | OP599570 |
| *C. indochinensis* | CVN266 | ZIN:97670 | Viet Nam: Lam Dong Province, Bi Doup | HM587024∗ | OP599597 | OP599572 |
| *C. sapaensis* | CVN94 | ZIN:96264 | Viet Nam: Lao Cai Province, Sa Pa | HM587006∗ | OP599598 | OP599573 |
| *Crocidura* sp.*AB2* | CVN201 | ZIN:97089 | Viet Nam: Ba Ria-Vung Tau, Binh Chau | HM587018∗ | OP599599 | OP599574 |
| *C. phanluongi* | MaDa4 | ZIN:100306 | Viet Nam: Dong Nai Province, Ma Da | JX181939∗ | OP599600 | OP599575 |
| *C. sokolovi* | CVN2 | ZIN: 96394 | Viet Nam: Kon Tum Province, Ngoc Linh Mt. | HM586999∗ | OP599601 | OP599576 |
| *Diplomesodon pulchellum* 1 | Dipl07-1 | – | Russia: Astrakhan region, Dosang | OP599562 | OP599613 | OP599588 |
| *Diplomesodon pulchellum* 2 | Dipl08-1 | – | Russia: Astrakhan region, Dosang | OP599563 | OP599614 | OP599589 |
| *Diplomesodon pulchellum* | Kb16-105 | ZMMU:S197309 | Kazakhstan: Almaty region, Karatal district | OP599564 | – | – |

## Phylogenetic tree reconstruction

Trees were reconstructed under Bayesian (BI) and maximum likelihood (ML) criteria and rooted with the outgroups, including *Suncus murinus* and *Suncus montanus.* ML reconstructions were performed in IQTree v.1.6 (Nguyen et al., 2015). The ModelFinder routine (Kalyaanamoorthy et al., 2017) was used to determine the optimal partitioning scheme and best-fit substitution models for each subset under the Bayesian information criterion (BIC). Clade stability was tested using Ultrafast Bootstrap (Minh et al., 2013), with 10,000 replicates.

For the datasets I and II the Bayesian tree reconstruction was performed in MrBayes 3.2 Ronquist et al. (2012), assuming separate models for each of the subsets identified by ModelFinder. The analysis included two independent runs with the chain length of two million generations and sampling rate of 2,000 generations. Using these settings, the effective sample size exceeded 200 for all estimated parameters. Tracer 1.7 software (Rambaut and Drummond 2005) was used to check for convergence, and the burn-in fraction was 10 % of the chain length.

In case of the *cytb* dataset III, the ultrametric Bayesian tree was reconstructed in BEAST 1.10.4 (Suchard et al., 2018) assuming a lognormal relaxed clock model and employing a birth-death tree prior. Separate models were used for each of the codon positions (Tab. S3). Chain length was set to 200 million generations. Tracer 1.7 and TreeAnnotator 1.10.4 were used to analyze the program output and to generate the Maximum Clade Credibility tree, correspondingly.

## Results

### Combined nuclear-mitochondrial tree

In a combined analyses of two nuclear and one mitochondrial genes, the final alignment consisted of 1669 nucleotide positions, including 561 bp of *ApoB*, 728 bp of *BRCA1*, and 380 bp of *cytb*. The optimal partitioning scheme for each nuclear gene, as identified by ModelFinder under the BIC criterion corresponded to the partitioning by codon position, but with the data from the two genes combined. The best-fit substitution models employed for each of the subsets are given in Table S3.

All methods of phylogenetic reconstruction (ML and BI), based on the combined mito-nuclear dataset, were congruent in recovering the same main clades with high and/or moderate support. Thus, only the Bayesian tree is shown here (Fig. 1). *C. serezkyensis* and *C. arispa* are undoubtedly closely related forms, which is consistent with their position within the same species complex. *C. ramona*, – the endemic species of Israel, appears to be their closest sister group. The clade (*C. serezkyensis*+*C. arispa*) / *C. ramona* is placed as the sister group to *C. leucodon* (0.63/95), but with moderate support. This group of Palearctic species (*C. serezkyensis* + *C. arispa*) / *C. leucodon* is re-

covered as the closest relative of the Mediterranean 36-chromosomal species (*C. sicula*, *C. canariensis*, *C tarfayensis*, and *C. zimmermanni*), albeit without strong support (0.51/65). In turn, all of the above species are part of a larger Afro-Mediterranean clade (0.94/73), which also includes Eastern Afromontane (mainly species of Ethiopia and Tanzania) and West African (*C. obscurior* / *C. eburnea* species group) clades.

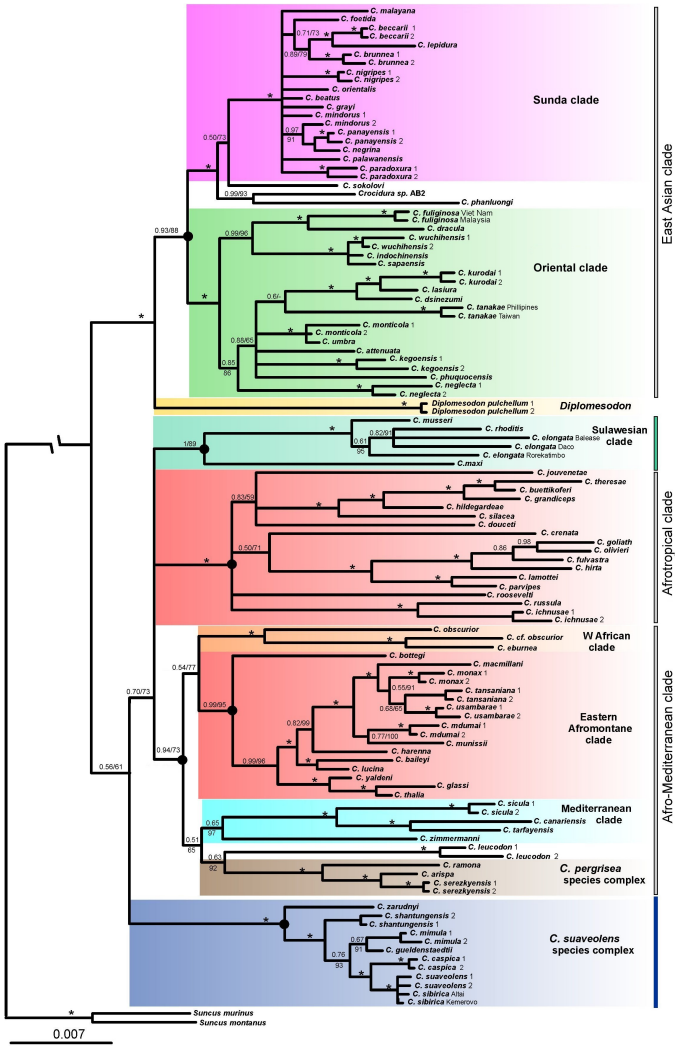


**Figure 1** – The Bayesian phylogeny of *Crocidura* as deduced from concatenated alignment of the fragments of *ApoB* and *BRCA1* exons and transversions of the 3rd codon positions of *cytb* gene. Values near the nodes correspond to posterior probabilities in the MrBayes and bootstrap support (1000 pseudoreplicates) in the ML analyses respectively (BPP/ML). Asterisks denote support values ≥0.98 BPP/98 % ML. *Suncus murinus* and *Suncus montanus* serve as outgroups .

There is no distinct Asian clade. On the contrary, several monophyletic groupings of Asian species can be identified, of which only two sister clades (Sunda and Oriental) show obvious close relationships. Their association, which includes the majority of species of South and East Asia, is inferred to be a sister branch to *Diplomesodon*. However, Sulawesian species are not directly related to species of the Malay Archipelago and Indochina and form a separate clade, which also includes *C. maxi* from Java. The trans-Palearctic *C. suaveolens* species complex has no obvious close relatives. Finally, as mentioned above, *C. leucodon* as well as the *C. pergrisea* species complex with *C. ramona* belong to the Afro-Mediterranean clade.

There is also no unified African clade. Instead, at least two unrelated African clades with sub-Saharan distribution are identified. One of them is the Afrotropical clade described earlier by Dubey et al. (2008) and the other is the Eastern Afromontane clade recently described by Bannikova et al. (2021). The third potential group corresponds to the West African *C. obscurior* / *C. eburnea* species group, which belongs to the Afro-Mediterranean clade, however its precise position therein is unclear.

### Concatenation of *ApoB* and *BRCA1* genes

The tree of concatenated *ApoB+BRCA1* sequences (Fig. 2a) is similar to the combined mito-nuclear tree and also demonstrates a division into several Asian and two African clades. The support of Afro-Mediterranean clade is retained. Within this clade, the support of the association of west Palearctic species (*C. serezkyensis*+*C. arispa*) *C. ramona*)) *C. leucodon*))) with Mediterranean subclade is increased. The position of the Sulawesian clade is better resolved than in the mito-nuclear tree, as it occupies a sister position to the Afro-Mediterranean clade with moderate support (0.75/88). Finally, the isolated position of the *C. suaveolens* species complex placed as the sister branch to the association of Afro-Mediterranean, Afrotropial and Sulawesian clades is preserved but with higher support in both BI and ML analyses. Thus, the reduced support of several key nodes in the mito-nuclear tree in comparison with the nuclear tree can be attributed to the impact of the mitochondrial *cytb* data.

### *Cytb* tree

The Bayesian ultrametric *cytb* tree (BEAST 1.10.4) on the basis of all codon positions of 1400 sequences is presented in Fig. 2b and the Supporting Information 2. Same as in the nuclear and mito-nuclear trees, the basal radiation is not resolved and the relationships between numerous clades are less clear. In total, 14 clades and seven branches formed by single species can be distinguished. Not all large clades identified in the nuclear tree are supported by the mitochondrial data. The Afrotropical clade is split into several non-sister lineages, and some species found within Afrotropical clade form separated branches (for example, *C. russula* species complex). The Afro-Mediterranean clade is not recovered in the mtDNA analysis.

At the same time, the *C. serezkyensis* + *C. arispa* clade as well as its grouping with *C. ramona* (p-distance 10 %) have robust support. The genetic distance between *C. serezkyensis* and *C. arispa* is only 3.6 %. Interestingly, *C. leucodon* and the Mediterranean clade appear to be distant from the *pergrisea* species complex and do not cluster with it even under low support.

## Discussion

### Phylogenetic position of the *C. pergrisea* species complex

Both nuclear and mitochondrial data clearly demonstrate that the *C. pergrisea* species complex forms a distinct genetic lineage, which is neither part of nor closely related to any of the known species groups within *Crocidura* (e.g. *C. suaveolens* species complex). Moreover, *C. serezkyensis* and *C. arispa* are without a doubt not directly related to *C. zarudnyi*, which was traditionally included in the *C. pergrisea* species complex (Hutterer, 1993; Zaitsev, 1993) but is genealogically close to *C. suaveolens* based on genetic data (Dubey et al., 2007a). This result is consistent with chromosomal data on *C. zarudnyi* (2n=40, FN=50) which is chromosomally highly similar to the *C. suaveolens* species complex (Dubey et al., 2007a). Contrary to this, shrews of the *C. pergrisea* species complex from Dzulfa (Azerbaijan) displayed a very unique karyotype (2n=22, FN=34) (Grafodatskii et al. 1988). The latter karyotype was originally reported under the name *C. pergrisea* but interpreted as *C. serezkyensis* by (Hutterer, 1993, 2005) and as *C. arispa* by Burgin et al. (2018). However, taking into account difficulties in distinguishing between crocidurine species in general and the *pergrisea* species complex in particular, the holotype of *C. zarudnyi* also needs to be screened for its molecular makeup. In the lack of such data, we have to keep in mind as an option that Dubey et al. (2007a) and Mohammadi et al. (2013) might have sequenced not *C. zarudnyi* proper but some other yet unrecognized species of the *C. suaveolens* species complex.

Based on both mitochondrial and nuclear data, the *C. serezkyensis* + *C. arispa* clade is supported as the sister group to the Negev shrew *C. ramona* Ivanitskaya et al. (1996), which is endemic to Israel. On the basis of the values of *cytb* distances, *C. ramona* is closer to *C. serezkyensis* + *C. arispa* (*cytb* distance=10.5 %) than any other *Crocidura* species (12–18 %); at the same time *ramona* is substantially more distant

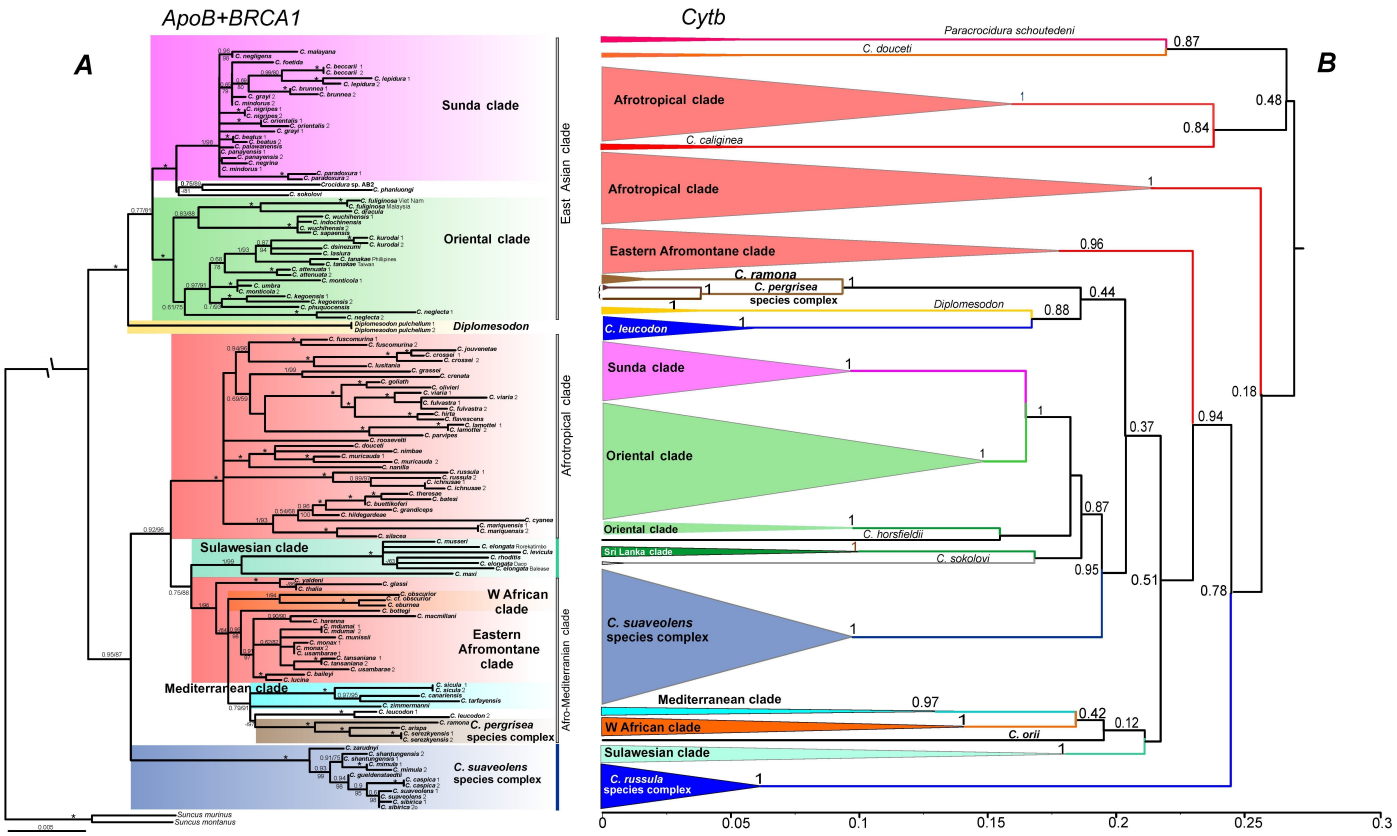


**Figure 2** – (A) Maximum likelihood tree as deduced from concatenated alignment of the fragments of *ApoB* and *BRCA1* exons. **(B)** Bayesian ultrametric tree reconstructed in BEAST for all codon positions of *cytb* gene for 1400 sequences of *Crocidura* spp., *Diplomesodon pulchellum* and *Paracrocidura schoutedeni.* Values near the nodes correspond to posterior probabilities and bootstrap support (1000 pseudoreplicates) in the MrBayes and ML analyses respectively (BPP/ML) in the *ApoB* + *BRCA1* tree, and posterior probabilities in BEAST in the *cytb* tree. Asterisks denote support values ≥0.98 BPP/98 % ML .

from *serezkyensis* and *arispa* than they are from each other (3.6 %). Our molecular data clearly support Ivanitskaya et al. (1996) who suggested that *C. pergrisea* is the nearest albeit not necessarily close relative of *C. ramona.* This conclusion was derived mainly from a comparison between karyotypes of *C.* cf. *pergrisea* from Dzulfa (see above) and *C. ramona* (2n=28, FN=44). In particular, it was suggested that a potential karyological synapomorphy for the grouping of these two species is the shared morphology of the X-chromosome, which is acrocentric and demonstrates the inverted G-banding pattern in comparison to other studied *Crocidura* species.

From a morphological perspective, shrews of the *C. pergrisea* species complex are characterized by a flattened skull (Zaitsev, 1993). According to Ivanitskaya et al. (1996) the relative skull height in *C. ramona* is higher than in *C. pergrisea* but, at the same time, much lower than in other *Crocidura*, which was viewed as another indication of closer relationships between *C. ramona* and the *pergrisea* species complex.

## Systematics of the "*pergrisea*" species complex

Taxonomy of the *pergrisea* species complex has always been controversial, which is a direct consequence of the small number of preserved vouchers. Available data are usually insufficient to provide a reliable diagnoses for most of putative species. In most cases, the revisers suggested the existence of several species within the *pergrisea* species complex and focused on diagnostics between the latter and other similarly sized sympatric species of the genus *Crocidura* (e.g., Jenkins, 1976).

Spitzenberger (1971) separated *C. pergrisea* from *C. zarudnyi* on the basis of the skull shape (not flattened in the latter) and described *arispa* as subspecies of *C. pergrisea*, naming the small size, relatively short tail, and some dental features among the differences between *arispa* and the nominal form. This author included *C. serezkyensis* in *C. pergrisea* and questioned whether or not *C. armenica* belongs to the *pergrisea* species complex.

Zaitsev (1993) clearly separated *C. serezkyensis* as a species distinct from *C. pergrisea* and *C. armenica* based on the unique structure of the third upper unicuspid tooth and the characteristic pelage color which is ashy–grey to silver–grey in *serezkyensis*, in contrast to brownish–grey or pale–yellow in other species of the group. According to Zaitsev (1993), *C. armenica* differs from a similarly small-sized *C. serezkyensis* also by the mandible shape. He considered the taxonomic status of *C. armenica* as unclear suggesting that it may be a subspecies of *C. pergrisea* and at the same time close to *C. arispa* which has a similar skull size and pelage coloration.

Hutterer (1993) supported the identification of *C. serezkyensis* as distinct from *C. pergrisea*, however, he included *C. arispa* into *C. serezkyensis* based on the similar size and also retained *C. armenica* as a separate species. Having revised the taxonomy of the group, Kryštufek and Vohralík (2001) suggested that *C. arispa* deserves species rank and identified the diagnostic features separating it from *C. pergrisea*, *C. serezkyensis*, *C. armenica* and *C. zarudnyi*. This arrangement was accepted by Hutterer (2005) in the 3rd edition of The Mammal Species of the World, however, in the Handbook of the Mammals of the World (Burgin et al., 2018), the authors retained all species with the exception of *C. armenica* which was not considered a valid species (without proper justification).

In summary, with the exception of *C. zarudnyi*, the *pergrisea* complex is now believed to contain four presumably allopatric species, of which we have examined just two. Although these two species originate from the opposite parts of the group distribution range and have discriminating traits, the genetic divergence between them appears to be relatively low. For example, the *cytb p*-distance between *C. arispa* and *C. serezkyensis* (3.6 %) falls into the range potentially corresponding to both inter- and intraspecific differentiation, albeit closer to its lower limit (Bradley and Baker, 2001). This value is similar to that between Western and Eastern clades of *C. leucodon* (Dubey et al., 2007b) but is lower than the distances within pairs of closely related species *C. suaveolens / C. caspica* (5.3 %) and *C. gueldenstaedtii / C. mimula* (4.8 %)

(Bannikova et al., 2006). This result is consistent with the hypothesis that the *pergrisea* species complex is, in fact, a single superspecies consisting of a number of recently diverged and geographically isolated allospecies. However, this thesis should be verified using additional sample sets including additional data from museum vouchers.

### Comments on the Afro-Mediterranean Clade

Based on the nuclear data, the species of the *pergrisea* complex as well as *C. ramona* undoubtedly belong to the large Afro-Mediterranean clade, which corresponds to the Old World clade sensu Dubey et al. (2008) but with the exclusion of the Sulawesian clade. Compared to Dubey et al. (2008), our sample set of *Crocidura* species was substantially expanded, which allowed us to better assess the taxonomic diversity in the Afro-Mediterranean clade. According to our results the Afro-Mediterranean clade includes several subclades and species groups, such as the species-rich Eastern Afromontane clade (containing many Ethiopian and Tanzanian-Kenyan endemics), Mediterranean clade (*C. sicula, C. canariensis*, *C. tarfayensis*, and *C. zimmermanni*), West African *C. obscurior/C. eburnea* species group, *C. leucodon* lineage and the *C. pergrisea* species complex+*C. ramona* clade. Our data confirm the suggestion of Lavrenchenko et al. (2009) that the Eastern Afromontane clade is phylogenetically related to the Mediterranean species of *Crocidura*. Moreover, the Mediterranean clade is a possible candidate for the closest relative of the *C. ramona / C. pergrisea* species complex. To achieve a robust support for all members of the Mediterranian clade, further studied should focus on multi-locus data.

## References


Bannikova A.A., Abramov A.V., Borisenko A.V., Lebedev V.S., Rozhnov V.V., 2011. Mitochondrial diversity of the white-toothed shrews (Mammalia, Eulipotyphla, *Crocidura*) in Vietnam. Zootaxa 2812: 1–20.

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

Bannikova A.A., Zemlemerova E.D., Lebedev V.S., Lavrenchenko L.A., 2021. The phylogenetic relationships within the Eastern Afromontane clade of *Crocidura* based on mitochondrial and nuclear data. Mamm. Biol. 101(6): 1005–18. doi:10.1007/s42991-021-00120-7

Bekenov A., Butovsky P.M., Kasabekov B.B., Lankin P.M., Strelkov P.P., Stogov I.I., Fedosenko A.K., Shaimardanov R.T., Shubin I.G, 1985. Insectivora and Chiroptera. In: Gvozdev E. I., Strautman E. I. (Eds.) Mammals of Kazakhstan, 4. Nauka, Alma-Ata. 1–280 [book in Russian].

Bradley R.D., Baker R.J., 2001. A test of the genetic species concept: cytochrome-b sequences and mammals. J. Mammalogy 82(4): 960–973. doi:10.1644/1545-1542(2001)

Burgin C.J., He K., Haslauer R., Sheftel B.I., Jenkins P.D., Ruedi M., Hintsche S., Motokawa M., Hinckley A., Hutterer R., 2018. Family Soricidae (shrews). In: Wilson, D.E. and Mittermeier, R.A. (Eds.), Handbook of the Mammals of the World. Vol. 8. Insectivores, Sloths and Colugos. Lynx Edicions, Barcelona. 332–551.

Dubey S., Cosson J.F., Vohralík V., Kryštufek B., Diker E., Vogel P., 2007b. Molecular evidence of Pleistocene bidirectional faunal exchange between Europe and the Near East: the case of the bicoloured shrew (*Crocidura leucodon*, Soricidae). J. Evol. Biol. 20(5): 1799–1808. doi:10.1111/j.1420-9101.2007.01382.x

Dubey S., Nová P., Vogel P., Vohralík V., 2007a. Cytogenetic and Molecular Relationships between Zarudny'S Rock Shrew (*Crocidura zarudnyi*; Mammalia: Soricomorpha) and Eurasian Taxa. J. Mammalogy 88(3): 706–711. doi:10.1644/05-MAMM-A-409R.1.0R.19R1.1

Dubey S., Salamin N., Ruedi M., Barriere P., Colyn M., Vogel P., 2008. Biogeographic origin and radiation of the Old World crocidurine shrews (Mammalia: Soricidae) inferred from mitochondrial and nuclear genes. Mol. Phylogenet. Evol. 48: 953–963. doi:10.1016/j.ympev.2008.07.002

Ellerman J.R., Morrison-Scott T.C., 1951. Checklist of Palearctic and Indian mammals 1758 to 1946. Brit. Mus. Nat. Hist, London.

Esselstyn J.A., Timm R.M., Brown R.M., 2009. Do geological or climatic processes drive speciation in dynamic archipelagos? The tempo and mode of diversification in Southeast Asian shrews. Evolution 63: 2595–2610. doi:10.1111/j.1558--5646.2009.00743.x

Esselstyn J.A., Achmadi A.S., Handika H., Swanson M.T., Giarla T.C., Rowe K.C., 2021. Fourteen new, endemic species of shrew (genus *Crocidura*) from Sulawesi reveal a spectacular island radiation. Bull. of the Am. Mus. Nat. Hist. 454(1): 1–108. doi: 10.1206/0003-0090.454.1.1

Grafodatsky A.S., Radzhabli S.I., Sharshov A.V., Zaitsev M.V., 1988. Karyotypes of five *Crocidura* species of the USSR fauna. Tsitologiya 30: 1247–1250 (In Russian).

Hall T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Proceedings of the Conference held in Nucleic acids symposium series 41: 95–98. http://jwbrown.mbio.ncsu.edu/JWB/papers/1999Hall1.pdf

Hofreiter M., Jaenicke V., Serre D., von Haeseler A., Pääbo S., 2001. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucl. Acids Res. 29: 4793–4799. doi:10.1093/nar/29.23.4793

Hutterer R., 1993. Order Insectivora. In: Wilson D.E. & Reeder D.M. (Eds.) Mammal Species of the World. A Taxonomic and Geographic Reference (2nd edn). Smithsonian Institution Press, Washington and London, 69–130.

Hutterer R., 2005. Order Erinaceomorpha, Order Soricomorpha. In: Wilson D.E., Reeder D.M. (Eds.) Mammal Species of the World. A Taxonomic and Geographic Reference (3rd edn). John Hopkins Press, Baltimore, 212–311.

İbiş O, Koepfli KP, Özcan S, Tez C 2022. Whole mitogenomes of Turkish white-toothed shrews, genus *Crocidura* (Eulipotyphla: Soricidae), with new insights into the phylogenetic positions of *Crocidura leucodon* and the *Crocidura suaveolens* group. Org. Divers. Evol. doi:10.1007/s13127-022-00579-3

Ivanitskaya E., Shenbrot G., Nevo E., 1996. *Crocidura ramona* sp. nov. (Insectivora, Soricidae): A new species of shrew from the central Negev desert, Israel. Zeitschrift fur Säugetierkunde 61(2): 93–103.

Jenkins P.D., 1976. Variation in Eurasian shrews of the genus *Crocidura* (Insectivora: Soricidae). Bull. Brit. Mus., Nat. Hist., Zool.; G.B.; DA. 30(7): 271–309. http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=PASCAL7750180075

Kalyaanamoorthy S., Minh B.Q., Wong T.K.F., von Haeseler A., Jermiin L.S., 2017. ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates. Nature Methods 14: 587–589. doi:10.1038/nmeth.4285

Kryštufek B., Vohralík V., 2001. Mammals of Turkey and Cyprus. Introduction. Checklist. Insectivora. Knjižnica Annales majora, Koper, p. 88–91.

Lavrenchenko L.A., Bannikova AA., Lebedev V.S., 2009. Shrews (*Crocidura* spp.) endemic to Ethiopia: recent adaptive radiation of an ancient lineage. Dokl. Biol. Sci. 424:57–60. doi:10.1134/S0012496609010177 (Original Russian published in Doklady Akademii Nauk, 2009, 424: 705–708)

Lay D.A., 1967. A study of the mammals of Iran, resulting from the Street Expedition of 1962-1963. Field. Zool. 54: 1– 282.

Minh B.Q., Nguyen M.A.T., von Haeseler A., 2013. Ultrafast approximation for phylogenetic bootstrap. Molecular Biology and Evolution 30: 1188–1195. doi:10.1093/molbev/mst024

Mohammadi S., Dubey S., Sabbaghzadeh A., 2013. New record of *Crocidura zarudnyi* from Zabol, Iran, Zool. Ecol. 23: 2, 162–164. doi:10.1080/21658005.2013.802105

Nguyen L.-T., Schmidt H.A., von Haeseler A., Minh B.Q., 2015, IQTREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Mol. Biol. Evol. 32: 268–274. doi:10.1093/molbev/msu300

Rambaut A., Drummond A.J., Xie D., Baele G., Suchard M.A., 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. syy032. doi:10.1093/sysbio/syy032

Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Hohna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61: 539–542. doi: 10.1093/sysbio/sys029

Sambrook J., Fritsch E.F, Maniatis T., 1989. Molecular cloning:a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sefc K.M., Payne R.B., Sorenson M.D., 2007. Sorenson Single base errors in PCR products from avian museum specimens and their effect on estimates of historical genetic diversity. Conservation Genetics 8: 879–884. doi:10.1007/s10592-006-9240-8

Shpirer E., Haddas-Sasson M., Spivak-Glater M., Feldstein T., Meiri S., Huchon D., 2021. Molecular relationships of the Israeli shrews (Eulipotyphla: Soricidae) based on cytochrome b sequences. Mammalia 85(1): 79–89. doi:10.1515/mammalia-2019-0143

Spitzenberger F., 1971. Eine neue, tiergeographish bemerkenswerte *Crocidura* (Insectivora, Mammalia) aus der Türkei. Ann. Naturhistor. Mus. Wien B 75: 539–552.

Suchard M.A., Lemey P., Baele G., Ayres D.L., Drummond A.J., Rambaut A., 2018. Rambaut Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 4(1), vey016.

Zaitsev, M.V. 1993. Species composition and questions of systematics of white-toothed shrews (Mammalia, Insectivora) of the fauna of USSR. In: Zaitsev M.V., (Ed.) Questions of systematics, faunistics and palaeontology of small mammals. Proc. Zool. Inst. USSR Acad. Sci. 243: 3–46. [In Russian with English abstract]


*Associate Editor:* P. Colangelo

## Supplemental information

Additional Supplemental Information may be found in the online version of this article:

**Table S1** Data on nuclear genes and *cytb* used in the phylogenetic reconstructions based on the concatenated sequences. The list of sequences retrieved from GenBank and used in the analysis of dataset 1 (concatenated alignment of *cytb+ApoB+BRCA1*) and dataset 2 (*ApoB+BRCA1*).

**Table S2** Data on nuclear genes and *cytb* used in the phylogenetic reconstructions based on the concatenated sequences. Original primers designed for amplification and sequencing of short fragments of *cytb*, *ApoB* and *BRCA1* in *Crocidura serezkyensis*, *C. arispa* and *C. ramona*.

**Table S3** Data on nuclear genes and *cytb* used in the phylogenetic reconstructions based on the concatenated sequences. Models for the fragments of nuclear *ApoB*, *BRCA1* and mitochondrial *cytb* genes employed in the maximum likelihood analysis.

**Supplemental material S4** Data on mitochondrial *cytb* gene used in this study. List of *cytb* sequences retrieved from GenBank and used in the reconstruction of the Bayesian ultrametric tree reconstructed performed in BEAST for all codon positions of *cytb* gene. The Bayesian ultrametric tree reconstructed in BEAST for all codon positions of *cytb* gene for 1400 sequences of *Crocidura* spp, *Diplomesodon pulchellum* and *Paracrocidura schoutedeni* in newick format.