Zoodiversity, **58**(1): 27–38, 2024
DOI 10.15407/zoo2024.01.27

UDC 599:575.858(1-021.21)

CRYPTIC SPECIATION AND CHARACTERISTICS OF THE TRANSITION BIAS FOLLOWING AN EXAMPLE OF THE CYTB GENE IN PALEARCTIC MAMMALS

S. V. Mezhzherin, S. Yu. Morozov-Leonov, O. V. Rostovska, V. O. Tereshchenko, E. I. Zhalay

Schmalhausen Institute of Zoology NAS of Ukraine,
vul. B. Khmelnytskogo, 15, Kyiv, 01054 Ukraine
E-mail: smezhzherin@gmail.com

S. V. Mezhzherin (http://orcid.org/0000-0003-2905-5235)
S. Yu. Morozov-Leonov (https://orcid.org/0000-0003-1784-7753)
O. V. Rostovska (https://orcid.org/0000-0001-6671-0265)
V. O. Tereshchenko (http://orcid.org /0000-0001-6671-0265)
E. I. Zhalay (https://orcid.org/0000-0003-3637-1156)

urn:lsid:zoobank.org:pub:EBC3FE43-DFD04092-AFF4-00A4DC042900

Cryptic Speciation and Characteristics of the Transition Bias Following an Example of the cytb Gene in Palearctic Mammals. Mezhzherin, S. V., Morozov-Leonov, S. Yu., Rostovska, O. V., Tereshchenko, V. O., Zhalay, E. I. — A study of taxonomic differentiation and transition bias of the cytb gene, focusing on the early stages of speciation, was carried out in 15 subfamilies/families of five orders of Palearctic mammals. It was confirmed that the genetic differentiation among taxa of small and large mammals is shifted by the same taxonomic level, in which connection the period of cryptic speciation of Micromammalia (semi- and allospecific levels) corresponds to the divergence of type species in their genus within Macromammalia. In all cases, both transition bias and its evolutionary compensation took place. The novelty of the study also lies in the proof that the alignment of the transitions and transversion rates in evolutionary lineages has the pattern of a phase transition, and the frequency of transversions in short-cycle species of the orders Insectivora and Rodentia is higher than in long-cycle orders Artiodactyla, Carnivora and Chiroptera while with relative equality of transitions. The latter circumstance might be associated with the traits of metabolism and determines the characteristics of the molecular evolution of mammalian groups with short and long life cycles.
Key words: genetic differentiation, transition bias, Micromammalia, Macromammalia, cytb, Palearctic, life cycles.

Early, as a rule, cryptic stages of speciation are a subject of particular interest in evolutionary biology and a problematic area of taxonomy. The latter circumstance is due to the fact that studies of the early stages of evolution are associated with genetic research. Whereas in taxonomic practice, generally, a standard typological approach is used applying morphological features that do not have high resolution in the early stages of evolution. Taxonomic levels correspond to the early stages of evolutionary divergence, which are interpreted as semispecies or allospecies (Anderson, 1977; Mezhzherin, 1994; Mallet, 2007). Semispecies are parapatric organisms that easily hybridize with each other in overlaping zones and with a level of genetic distances slightly exceeding the standard interpopulation differences. Trinominal names are used to designate them. A semispecies criterion is a fact of the gene flow restriction. Allospecies are defined as geographically isolated sibling species, their divergence is higher than that of semispecies and they are given binomial names. Semispecies and allospecies are characterized by the absence of reliable reproductive barriers, crypticity and vicariance. They form superspecies systems (Amadon, 1966) with regulated rules for their designation (Code of Zoollogical nomenclature).

The relevance of the studies of the cryptic stage of species formation is not only due to the difficulties of practical systematics, but also the need of obtaining an answer to one of the key evolutionary questions: if intraspecific and interspecific differentiations are, in essence, the same or something entirely different. The answer allows to present speciation either as a continuous linear process of accumulation of substituted genes, or as an abrupt phenomenon. In this regard, molecular and genetic studies are of particular interest.

The universal rule of molecular evolution is a general tendency of the prevalence of a certain type of nucleotide substitutions (Li & Graur, 1991). A special place is taken by a phenomenon of the transition bias — a predominance in the number of transitions over transversions (Fitch, 1967; Vogel & Kopun, 1977; Brown et al., 1982; Kumar, 1996; Belle et al., 2005; Stoltzfus & Norris, 2016), which should be considered as a special case of a general regularity. The phenomenon is observed in spontaneous mutations and is due to the relative ease of nucleotide substitutions within one class of chemical compounds of purines (A↔G) or pyrimidines (T↔C), compared to interclass substitutions of purines for pyrimidines or vice versa (A↔T, A↔C, G↔ T, G↔C). Of particular interest is the evolutionary compensation of transition bias (Rosenberg et al., 2003). Its essence is that with an increase in the level of divergence, there is a decrease in the scale of the bias, which persists at the intraspecific stages of evolution, but is leveled out at the order of interspecific divergence (Fitch, 1967; Brown et al., 1982; Collins & Jukes, 1994; Kumar, 1996; Ebersberger et al., 2002; Belle et al., 2005; Duchene et al., 2015). At the same time, the rates of evolutionary alignment of the transition bias are specific for different genes and parts of DNA (Keller et al., 2007; Mezhzherin & Tereshchenko, 2023). Likewise, there are reasons to believe that in different groups of family rank and above, transition bias and its evolutionary alignment have their own characteristics (Mezhzherin et al., 2023). The latter aspect remains insufficiently studied. Therefore, the aim of this study was to solve two problems: to determine the relationship between the stages of speciation and the nature of the compensation of the transition bias and how the nature of the bias is displayed in different groups of mammals in the taxonomic rank of family.

Materials and Methods

Subject of the study. The work uses complete sequences of the cytb gene submitted to GenBank and representing 15 families/subfamilies of five Palearctic orders of mammals. In total, the analyses included 2315 sequences of 577 species.

Micromammalia. Chiroptera Blumenbach, 1779 — Vespertilionidae Gray, 1821: *Barbastella* Gray, 1821 (6 species, 27 sequences), *Eptesicus* Rafinesque, 1820 (6/10), *Hypsugo* Kolenati, 1856 (3/7), *Murina* Gray, 1842 (6/9), *Myotis* Kaup, 1829 (38/93), *Pipistrellus* Schreber, 1774 (5/10), *Plecotus* Linnaeus, 1758 (7/14).

Insectivora Bowdich, 1821. Soricidae G. Fischer, 1817: *Sorex* Linnaeus, 1758 (33/69), *Neomys* Kaup, 1829 (3/11), *Crocidura* Wagler, 1832 (21/44), *Suncus* Ehrenberg, 1832 (2/8). Talpidae G. Fischer, 1817: *Talpa* Linnaeus, 1758 (13/64), *Mogera* Pomel, 1848 (8/36), *Uropsilus* Milne-Edwards, 1871 (8/47), *Desmana* Güldenstädt, 1777 (1/1), *Galemys* Kaup, 1829 (1/1); *Scaptonyx* Milne-Edwards, 1872 (2/2). Erinaceidae G. Fischer, 1817: *Atelerix* Pomel, 1948 (3/3) *Erinaceus* Linnaeus, 1758 (4/50), *Hemiechinus* Fitzinger, 1866 (1/10), *Mesechinus* Ognev, 1951 (2/14), *Paraechinus* Trouessart, 1879 (2/5).

Rodentia Bowdich, 1821. Myomorpha Brandt, 1855. Muroidea Fischer-Waldheim, 1817. Muridae Illiger, 1811. Murinae Illiger, 1811: *Apodemus* s. l. Kaup, 1829 (21/116), *Micromys* Dehne, 1841 (1/11), *Mus* Linnaeus, 1758 (19/127), *Rattus* Fischer von Waldheim, 1803 (3/20), *Tokudaia* Kuroda, 1943 (3/5). Gerbillinae Gray, 1825: *Dipodillus* Lataste, 1881 (4/31), *Gerbillus* Desmarest, 1804 (14/101), *Brachiones* Thomas, 1925 (1/1), *Meriones* Illiger, 1811 (10/39), *Psammomys* Cretzschmar, 1828 (1/4), *Rhombomys* Wagner, 1841 (1/7) *Sekeetamys* Ellerman, 1947 (1/1), *Desmodilus* Thomas & Schwann, 1904 (1/7), Gerbelliscus Thomas, 1897 (1/13), *Tatera* Lataste, 1882 (1/1). Cricetidae J. Fischer, 1817. Arvicolinae Gray, 1821: *Arvicola* Lacepede, 1799 (3/5), *Microtus* s. l. Schrank, 1798 (57/94), *Myodes* Pallas, 1811 (7/18), *Eothenomys* Miller, 1896 (11/16), *Alticola* Blanford, 1881

(11/18), *Chionomys* Miller, 1908 (3/14), *Dicrostonyx* Gloger, 1841 (1/1), *Dinaromys* Kretzoi, 1955 (1/2), *Prometheomys* Satunin, 1901 (1/2), *Myopus* Miller, 1910 (1/1), *Eolagurus* Argyropulo, 1946 (2/5), *Lagurus* Gloger, 1841 (1/5), *Lemmus* Link,1795 (1/1); *Ellobius* Fischer,1814 (7/14), *Ondatra* Link, 1795 (1/1). Cricetinae Fischer de Waldheim, 1817: *Cricetus* Leske, 1779 (1/10), *Mesocricetus* (Nehring, 1894) (3/22), *Cricetulus* Milne-Edwards, 1867 (5/40), *Phodopus* Miller, 1910 (3/15), *Allocricetulus* Argyropulo, 1933 (2/2), *Urocricetus* Satunin, 1903 (1/2), *Tscherskia* Ognev, 1914 (1/2). Dipodoidea Fischer de Waldheim, 1817. Dipodidae Fischer de Waldheim, 1817: *Cardiocranius* Satunin, 1903 (1/1), *Salpingotus* Vinogradov, 1922 (1/1), *Eremodipus* Vinogradov, 1930 (1/1), *Jaculus* Erxleben, 1777 (4/32), *Stylodipus* Allen, 1925 (3/7), *Chimaerodipus* Shenbrot, 2017 (1/5), *Euchoreutes* Sclater, 1891 (1/21), *Allactaga* F. Cuvier, 1836 (10/44), *Allactodipus* Kolesnikov, 1937 (1/3), *Pygeretmus* Gloger,1841 (3/11). Sciuromorpha Brandt, 1855. Sciuridae Fischer de Waldheim, 1817: *Spermophilus* F. Cuvier, 1825 (18/113), *Urocitellus* Obolenskij, 1927 (1/12), *Spermophilopsis* Blasius, 1884 (1/1), *Pteromys* G. Cuvier, 1800 (2/14), *Sciurus* Linnaeus, 1758 (2/31) *Tamias* Illiger, 1811 (1/11), *Atlantoxerus* Linnaeus, 1758 (1,1), *Marmota* Blumenbach, 1779 (9/17).

Macromammalia. Artiodactyla Owen, 1848. Cervidae Goldfuss, 1820: *Hydropotes* Swinhoe, 1870 (1/10), *Cervus* Linnaeus, 1758 (6/31), *Odocoileus* Rafinesque, 1832 (2/2), *Dama* Linnaeus, 1758 (1/5), *Rusa* C. H. Smith, 1827 (1/1) *Muntiacus* Rafinesque, 1815 (9/14), *Elaphodus* Milne-Edwards, 1872 (1/1), *Capreolus*, Gray, 1821 (2/32), *Alces* Gray, 1821 (1/7), *Axis* Smith,1827 (3/12). Moschidae Gray, 1821: *Moschus* Linnaeus, 1758 (4/4). Bovidae Gray, 1821: *Bos* Linnaeus, 1758 (4/27), *Bison* Smith, 1827 (2/15), *Capra* Linnaeus, 1758 (8/48), *Hemitragus* (Hodgson, 1841) (2/6), *Pseudois* Hodgson, 1846 (2/12), *Ammotragus* (Blyth, 1840) (1/6), *Ovis* Linnaeus, 1758 (7/59), *Rupicapra* Garsault, 1764 (1/8), *Naemorhedus* Smith, 1827 (3/21), *Capricornis* Gray, 1821(4/21), *Pantholops* Hodgson, 1834 (1/3), *Oryx* Blainville, 1816 (2/14), *Saiga* Gray, 1843 (1/9), *Antilope* Linnaeus, 1758 (1/12), *Gazella* Blainville, 1816 (6/51), *Nanger* Lataste, 1885 (1/3).

Carnivora Bowdich, 1821. Mustelidae Fischer von Waldheim: *Arctonyx* Cuvier, 1825 (1/2), *Meles* Brisson, 1762 (4/31), *Gulo* Pallas, 1780 (1/12), *Martes* Pinel, 1792 (5/29), *Vormela* Blasius, 1884 (1/8), *Enhydra* Fleming, 1828 (1/3), *Lutra* Brisson, 1762 (1/13), *Mustela* Brisson, 1762 (9/45). Felidae Fischer von Waldheim, 1817: *Acynonyx* Brookes, 1828 (1/1), *Catopuma* Severtzov, 1858 (1/1), *Caracal* Gray, 1843 (1/1); *Felis* Linnaeus, 1758 (2/13), *Lynx* Kerr, 1792 (4/6), *Pardofelis* Severtzov, 1858 (1/2), *Neofelis* Gray, 1867 (1/1), *Prionailurus* Severtzov, 1858 (3/7), *Otocolobus* Brandt, 1841 (1/1), *Panthera* Oken, 1816 (4/15). Canidae Fischer von Waldheim: *Canis* Linnaeus, 1758 (3/17), *Cuon* Hodgson, 1838 (1/1), *Nyctereutes* Temminck, 1838 (1/6), *Vulpes* Garsault, 1764 (7/24).

In this work, mostly complete sequences from 1140 to 1143 bp were taken. In some cases, sequences with a length of at least 1000 bp were used. In the latter case, alignment was carried out. For this, we used programs BioEdit (v7.2.5) and MEGA X (Kumar et al., 2018) using the ClustalW algorithm (Hall, 1999).

Characteristics of the analysis. Nucleotide substitutions are classified at the intraspecific, species, and genus levels of divergence in accordance with the taxonomy used when entering the cytb gene sequence in GenBank. The calculations did not take into account taxonomic incidents, which require revisions for their corrections. Moreover, to solve the problems and purposes of this work, taxonomic unambiguity is not a decisive factor.

Genetic differentiation was assessed by pairwise comparisons of nucleotide sequences. To do this, we used three indicators: the frequency of transitions (ts), the frequency of transversions (tv) and the substitution frequency in total (sub). In accordance with the scaling of the Murinae subfamily differentiation, which is one of the most studied molecular taxonomic groups of the Palearctic based on the cytb gene (Michaux et al., 2002; Ge et al., 2019), it will be correct to highlight the following ranges of evolutionary divergence (Mezhzherin & Tereshchenko, 2023). Individual and interpopulation variability within the species (range 0–0.02); a zone of mixed values of interpopulational differences and divergence of the initial stages of speciation (0.02–0.04); a zone of cryptic species formation (0.04–0.08) includes the subrange of 0.04–0.06 of the semispecies level, and the subrange of 0.06–0.08 to a greater extent of the allospecies level; a mixed zone (0.08–0.1) of allospecies and species divergence; a zone of species divergence (0.1–0.14); a range 0.14–0.16 — intergradation of species and genus levels; above 0.16 — genera differentiation.

The nucleotide sequence of the cytb gene is characterized by fairly stable ratios of nucleotides, which makes it possible to use simple, but at the same time, reliable methods for evaluating transition bias. The first method is the estimation of the average frequency of transitions and transversions (ts/tv-ratio) within each class of nucleotide substitutions.and the second is the ratio of transition and transversion sample deviation frequencies (F-index = SDts/SDtv) as well within each class.

Frequencies of the types of nucleotide substitutions were calculated directly by pairwise comparison of studied sequences.

The erroneous data was identified during phylogenetic analysis. Calculation of pairwise values of genetic distances was done using the MEGA program (v 11.0.11) (Tamura et al., 2021).

Results and Discussion

Scales of divergence. Frequency distributions of nucleotide substitutions, obtained by pairwise comparisons of cytb gene sequences at three taxonomic levels, represent continuous widely transgressive polygons (fig. 1).

The intraspecific level of nucleotide substitutions covers the range from 0 to 0.12 (fig. 1). Moreover, about 75 % of the cases fall into the class of minimal differentiation (0–0.02). About 16 % prove to be in the zone of 0.02-0.04. About 8 % of pairwise comparisons are in the range corresponding to the period of cryptic speciation (0.04–0.08), and only 1.5 % of comparisons are located within the range of species values of 0.08–0.12. In general, the average substitution frequency at this level reaches 0.018 with SD = 0.018. Mean values for individual families range from 0.005 (Soricidae) to 0.026 (Gerbelinae) (table 1).

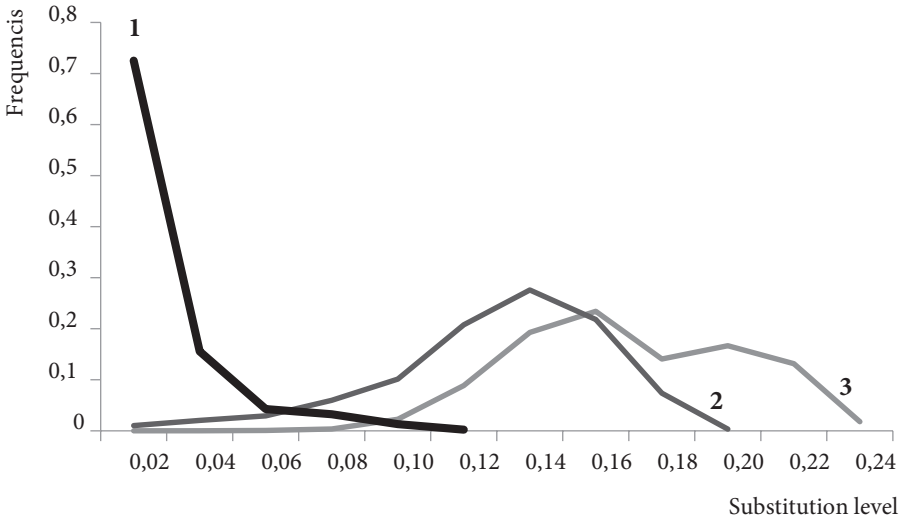

Fig. 1. Variation of summarized nucleotide substitution frequencies within Palearctic mammal subfamilies/families depending of substitution level. Taxonomic levels: 1 — intraspecies, 2 — species, 3 — genus.

Table 1. Average (M), sample deviations (SD) of nucleotide substitution, ts/tv and F indexes of different taxonomical levels within 15 Palearctic mammal families/subfamilies

Subfamily/ family	Intraspecies				Species				Genus			
	M	SD	ts/tv	F	M	SD	ts/tv	F	M	SD	ts/tv	F
Vespertilionidae	0.014	0.020	12.8	157.4	0.147	0.024	4.3	4.0	0.194	0.010	1.9	1.3
Soricidae	0.005	0.007	12.9	103.2	0.122	0.026	3.8	2.9	0.205	0.014	1.2	0.4
Talpidae	0.015	0.020	10.5	72.3	0.110	0.030	4.3	3.0	0.191	0.022	1.4	0.3
Erinaceidae	0.007	0.014	6.2	41.1	0.123	0.020	4.4	12.3	0.191	0.020	1.8	0.7
Murinae	0.022	0.017	7.0	24.5	0.142	0.024	2.1	0.5	0.169	0.010	1.4	1.0
Gerbilinae	0.026	0.028	6.3	26.5	0.124	0.030	3.5	3.2	0.168	0.024	1.8	0.4
Cricetinae	0.014	0.017	8.1	24.2	0.106	0.050	2.9	2.2	0.170	0.041	1.5	0.9
Arvicolinae	0.017	0.026	5.9	137.3	0.111	0.022	2.7	1.7	0.144	0.017	1.1	1.9
Dipodidae	0.018	0.017	5.9	29.7	0.136	0.032	2.0	1.1	0.182	0.026	1.4	0.6
Sciuridae	0.024	0.032	12.9	103.2	0.111	0.030	7.0	8.5	0.189	0.032	1.8	0.1
Mustelidae	0.012	0.020	9.2	80.3	0.074	0.030	8.2	22.0	0.151	0.022	3.2	1.1
Canidae	0.012	0.022	13.1	182.9	0.090	0.035	8.7	47.8	0.159	0.010	3.5	1.4
Felidae	0.006	0.007	8.2	45.1	0.067	0.035	27.1	237.3	0.121	0.010	9.0	8.5
Cervidae	0.009	0.010	25.9	135.2	0.068	0.022	17.4	98.4	0.131	0.020	4.9	1.0
Bovidae	0.009	0.014	9.1	77.5	0.049	0.020	15.6	40.3	0.127	0.022	3.3	0.4

The species level covers the substitutions range from 0 to 0.2. Cases of minimal substitutions, located in the range of 0–0.02, are a result of a false assignment of species status and account for 1 % only. Only 2 % of pairwise comparisons fall into the zone of 0.02–0.04, 8 % fall into the area of cryptic speciation (0.04–0.08), and 10.1 % fall into the intermediate range of 0.08–0.10. About 48.3 % falls directly into the species range of differentiation (0.1–0.14), and 21.8 % falls in the zone of intergradation of species and genus values (0.14–0.16). The cases of species differentiation at the genus level (over 0.16) account for 7.6 %. The average species differentiation calculated from the entire set of pairwise comparisons makes up M = 0.120 (SD = 0.033). The fluctuations in mean individual groups ranges from M = 0.047 (SD = 0.02) in Bovidae to M = 0.146 (SD = 0.024) in Vespertilionidae.

The genus level covers values from 0.02 to 0.24. The reassessment of the genus status is a range from 0.02 to 0.14 (30.9 %). The gray zone between species and genera division (0.14–0.16) is 23.4 %. The actual genus level covers the range from 0.16 to 0.24 (45.7 %). The mean genera differentiation calculated from the entire set of pairwise comparisons is M = 0.159 (SD = 0.034). Maximum differentiation takes place in Soricidae M = 0.205 (SD = 0.014), its high level is due to the divergence of subfamilies Soricinae – Crocidurinae. Minimal mean intergenera differences are observed in Felidae M = 0.121 (SD = 0.01) and Bovidae M = 0.127 (SD = 0.022).

Differences in the genetic differentiation of Micromammalia and Macromammalia are manifested at three levels of divergence. Moreover, the average values in the latter in all cases are significantly lower (fig. 2), and the species range of genetic differentiation of Macromammalia clearly corresponds to the zone of cryptic speciation of Micromammalia (fig. 3).

Transition bias. It takes place at all three levels of evolutionary divergence (table 1, fig. 4). The maximum values of the ts/tv-index fall at the intraspecific level M = 10.6 (SE = 1.43), the intermediate values at the species level M = 7.68 (SE = 1.89), and the minimum values at the genera level M = 2.65 (SE = 0.53). In general, the ts/tv-bias takes place at the early stages of speciation, while the frequencies of transitions and transversions are equalized at the later stages.

An analysis of the class average values of the ts/tv-index throughout the scale of nucleotide substitutions (fig. 5) shows the following regularities. First, the maximum values of the ts/tv-index fall into the zone of early and cryptic speciation (0.02–0.08), and only then

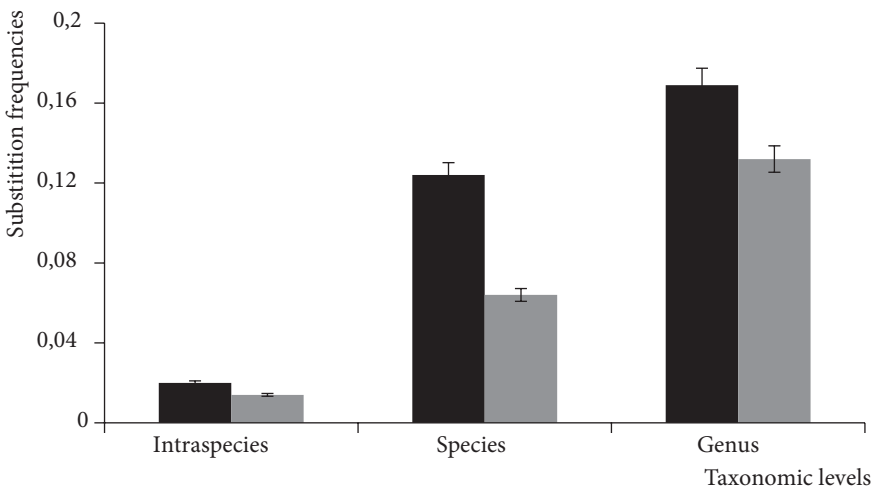

Fig. 2. Average frequencies of nucleotide substitutions frequencies (sub) and its standard errors of the three taxonomical levels in micromammals (black) and macromammals (gray).

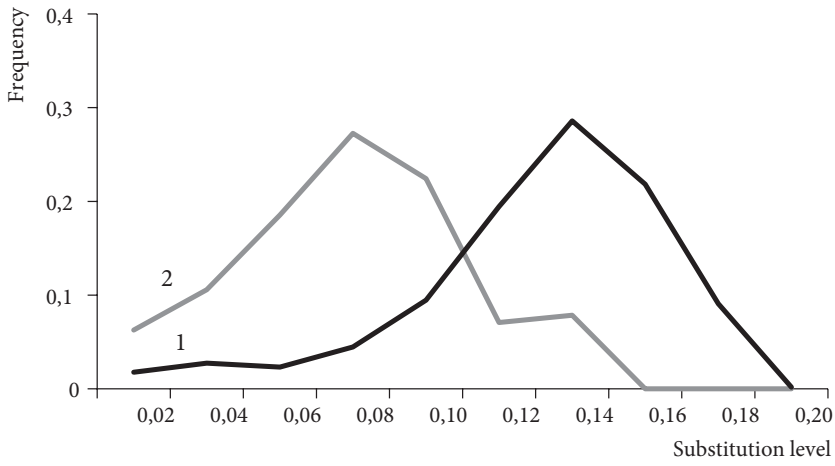

Fig. 3. Variation of summarized interspecies nucleotide substitution frequencies depending of substitution level in micromammals (1) and macromammals (2).

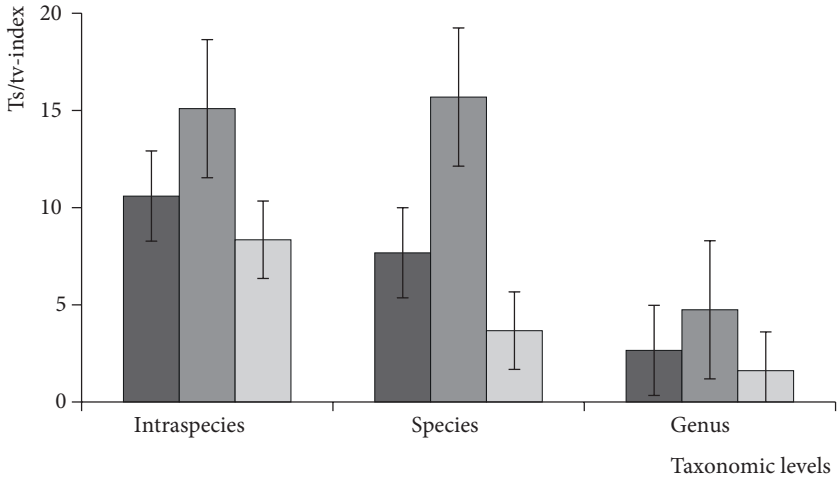

Fig. 4. Average frequencies of ts/tv-indexes and its standard errors on the three taxonomical levels: summarized data (dark gray), macromammals (gray), micromammals (light gray).

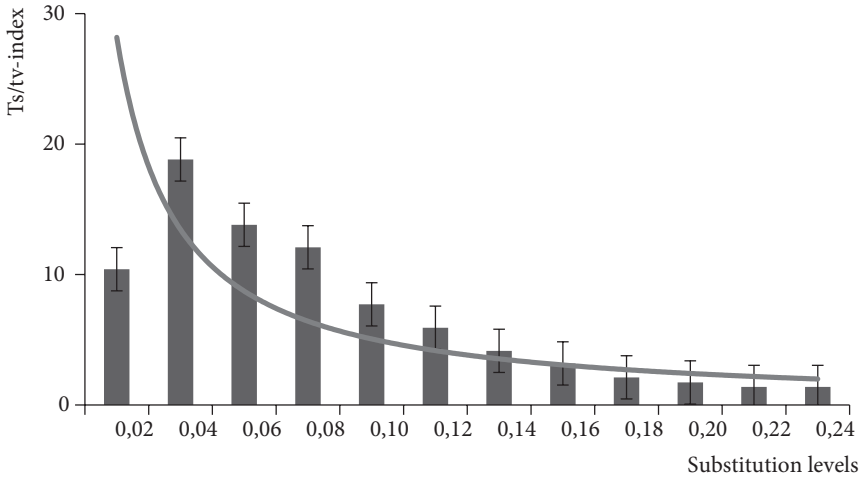

Fig. 5. Variation of summarized averages of ts/tv-index per substitution classes and its standard errors for subfamilies/families of Palearctic mammals on different nucleotide substitution levels. Line illustrates exponential approximation.

on the range of intraspecific differentiation (0–0.02). Second, the increase in the ts/tv-index values is caused by an abrupt increase in the frequency of transitions at the stages of minimal divergence, and their stabilization is due to a jump in the frequency of transversions and the relative constancy of transversions at the stages of species and genera divergence. Third, the change in the values of the ts/tv-index in the phyletic lineages is non-linear.

A similar situation, but with some specific characteristics, takes place when comparing the dispersions of transitions and transversions (table 1, fig. 6). The maximum values of the mean values of the F-index fall in the intraspecific level (M = 68.6; SE = 12.6); significantly lower in the species (M = 3.95, SE = 1.66) and minimal in the genus (M = 0.66, SE = 0.1).

When analyzing the F-index values along the entire scale of genetic differentiation, the range from 0 to 0.08 is clearly distinguished. In this interval, the differences between the two variances are significant at the highest level ($F = 11.3–28.5$; $\nu1 = 14$, $\nu2 = 15$, $p < 0.001$). Simultaneously, the maximum values are noted in the range of 0–0.02 (fig. 7). Within the frequency range of 0.08–0.1, the differences in dispersions are not so considerable, but significant ($F = 3.2$; $p < 0.05$). Above this level, the F-index values levels out and stabilizes ($F = 1.0–2.2$; $p > 0.05$). At the same time, the trend of changes in the F-index has an even more pronounced curvilinear nature than in the case with ts/tv-index.

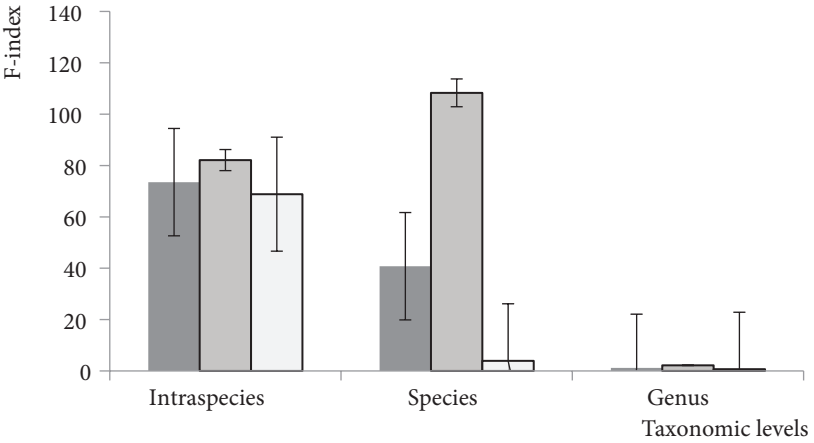

Fig. 6. Average values of F-indexes and its standard errors on the three taxonomical levels: summarized data (dark gray), macromammals (gray), micromammals (light gray).

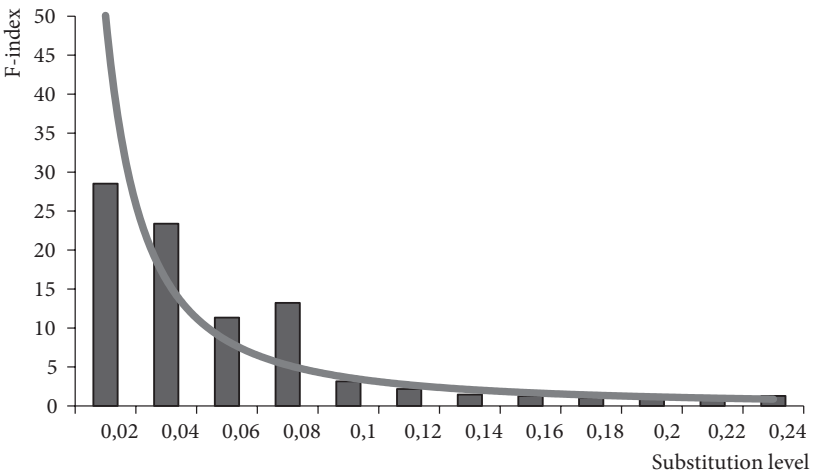

Fig. 7. Variation of summarized averages of F-index for subfamilies/families of Palearctic mammals of different nucleotide substitution level. Line illustrates exponential approximation.

Generalizations of families at the level of orders/suborders show that the pattern of the rates of changes in the frequencies of transitions and transversions in all cases is ambiguous. During the early stages of divergence in the range from 0.02 to 0.08, there is a sharp increase in the frequency of transitions, while maintaining the level of transversions (fig. 8). In the range of species values (from 0.12 and above), the situation is reversed: stabilization of the frequency of transitions and an abrupt increase in transversions. As a result of these non-linear processes, the ratio of transitions and transversions at the genera level of divergence levels off and the ts/tv-index tends to reach 1.

In Macromammalia, the values of the ts/tv-index, especially at the early stages of speciation, are higher than in Micromammalia (fig. 9), which is associated with a lag in the rate of accumulation of transitions compared to accumulation of transversions. This trend is especially pronounced in Macromammalia (Artidactyla, Carnivora) — animals with a long life cycle, compared to others with short life cycles (Insectivora, Rodentia). Small-sized bats (Chiroptera), as Macromammalia, are characterized by a long life cycle and stand apart from both. The mean values of the frequency of transitions at the species and genera levels of divergence of Micromammalia (0.097) and Macromammalia (0.103) are relatively equal

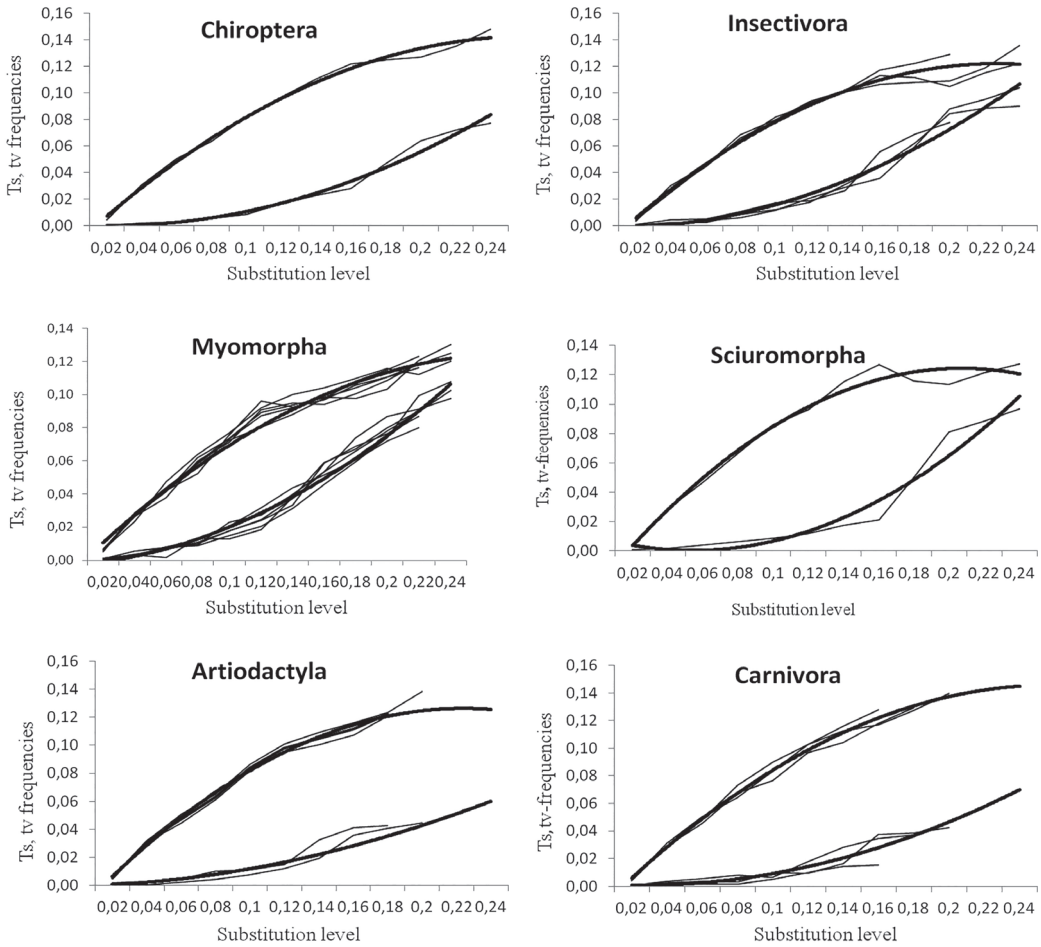

Fig. 8. Variation of transition (upper lines) and transversion (lower lines) frequencies accordingly to substitution frequencies level. Think lines are empirical data for each subfamily/family, thick ones — polynomial approximations of averaged data.

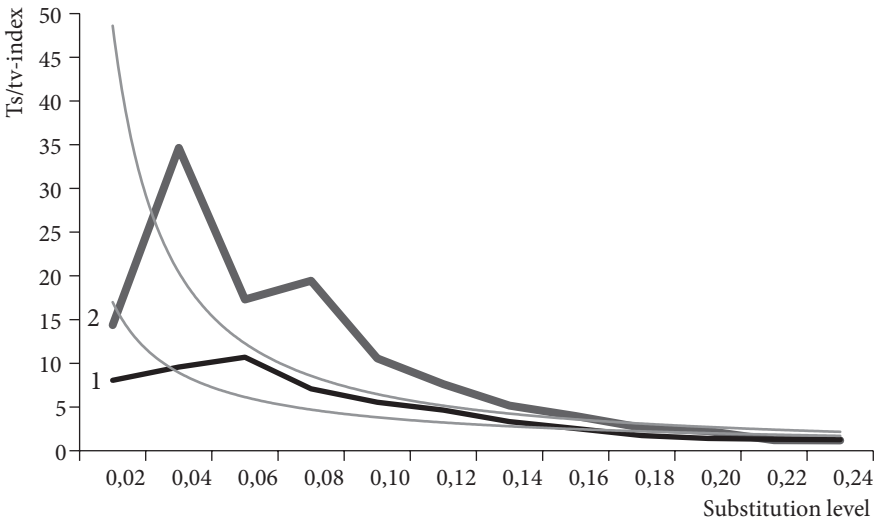

Fig. 9. Variation of summarized tv/ts-index in micro- (1) and macromammals (2) depending on nucleotide substitution level. Thick lines illustrate exponential approximation.

Table 2. Mean values (M), standard deviations (SD) and maximum class (Max) of values of transitions and transversions of the species and genus levels of divergence in the generalized samples

Orders	Transitions			Transversions			N
	M	SD	Max	M	SD	Max	
Chiroptera	0.124	0.014	0.14–0.16	0.054	0.019	0.08–0.10	2485
Artiodactyla. Carnivora	0.103	0.023	0.12–0.14	0.021	0.013	0.06–0.08	3792
Insectivora. Rodentia	0.097	0.017	0.14–0.16	0.053	0.036	0.12–0.14	11582

N — number of pairwise comparisons.

and clearly lower than in bats (0.124) (table 2). The situation is different for transversions: the average values for Micromammalia (0.053) and bats (0.054) are two times higher than for Macromammalia (0.021). As a result, in Insectivora and Rodentia orders the ratio between the mean values of transversions and transitions is 1.8, in Chiroptera it increases to 2.1, and in Macromammalia Artidactyla and Carnivora reaches 4.9.

Discussion

Comparative studies at the level of the cytb gene nucleotide sequences in the phyletic lineages of Palearctic mammals give basis to represent a general pattern of genetic differentiation and evolutionary bias. It can be presented as a certain triad of evolutionary ambiguity.

The first ambiguity. Homologous taxa of Micromammalia have a level of genetic differentiation on a clearly larger scale than that of Macromammalia. Moreover, the discrepancy between the mean values of the frequencies of nucleotide substitutions corresponds to a bias by one taxonomic level.

The fact that the taxa of small mammalian orders are genetically more distant from each other than large mammalian orders had been previously proven by multilocus allozyme analysis (Baccus et el., 1983; Mezhzherin & Morozov-Leonov, 1995). The reassessment of the taxonomic status of Macromammalia is caused by the presence of pronounced

traits in them, which are a result of prolonged ontogenesis (Mezhzherin & Morozov-Leonov, 1996). Small mammals are characterized by external uniformity (unimorphism), which leads to the cryptic nature of the early stages of speciation and their disregard during taxonomic constructions.

The second ambiguity. The nature of nucleotide substitutions at the early and late stages of cytb gene divergence is of an alternative nature. In the frequency range of substitutions from 0 to 0.1 there is a pronounced transition bias. At the late stages of divergence (range from 0.12 to 0.24), the frequency of transversions increases abruptly against the background of stabilization of the transitions levels. This causes the ts/tv-index values to approach 1. Such a ratio of transitions and transversions does not correspond to the model of random substitutions, in which the values of the ts/tv-index should be about 0.6. This means that, in contrast to the control region (Mezhzherin & Tereshchenko, 2023) representing the non-coding part of mtDNA, the deficiency of the cytb gene transversions persists even at high levels of divergence.

Directed evolutionary changes in the transition:transversion bias have the form of a phase transition caused by an abrupt increase in the frequencies of a certain type of nucleotide substitutions. At the early stages of speciation (0.02–0.08), there is a sharp increase in the frequency of transitions against the background of a slight increase in the frequency of transversions. At the later stages (0.14–0.24), on the contrary, there is a saltation in transversions and relative stabilization of transitions. There are two approaches to explain this consistent pattern. The first suggests it being a genetic saturation at the stages of divergence of species and genera (Philippe et al., 2011), which mechanism is associated with the accumulation of reverse mutations and their different probabilities in the case of transitions and transversions. The second is the ambiguity of the mutation process at different levels of evolutionary genetic divergence. It can be assumed that at the stages of latent speciation, covering periods of hundreds of thousands of years, a spontaneous mutation process prevails, accompanied by the dominance of transitions. With divergences, estimated in millions of years, a different pattern of mutation process took place, with the accumulation of transitions and transversions occurring more evenly. This means that speciation cannot be reduced to genetic processes occurring in modern populations, the idea of which has been repeatedly expressed (Altukhov, 1990; Stegnii, 1993).

Despite the similarity of interpretations of the mechanisms of evolutionary bias from the point of view of different concepts, the adoption of one of them has different theoretical consequences. In the first case, it is the recognition of the constancy of rates and the independent nature of mutations at early and late stages of evolution. In the second case, it is the unevenness of mutations at different stages of divergence and the related formation of an idea of the phylum as a holistically evolving system.

The third ambiguity. Long-cycle mammals of the orders Artiodactyla, Carnivora and Chiroptera have a lower frequency of transversions compared to short-cycle animals (Insectivora, Rodentia) with a relative equality of transitions. This leads to a more pronounced transition:transversion bias in them at all levels of divergence. It is logical to associate this trend with differences in the rate and nature of mutations in small and large mammals, short and long cycle mammals, caused by different rates of biological processes and, above all, metabolism. This point of view has not been previously confirmed (Belle et al., 2005). However, as this study shows, this hypothesis cannot be ruled out by making more large-scale comparisons. This explanation is also confirmed by the fact that the level of genetic variability in small groups of organisms within classes of vertebrates is significantly higher than in large organisms (Wooten & Smith, 1985; Mezhzherin, 2002).

Thereby, in the evolutionary lineages of Palearctic mammals, which differ in size and duration of life cycles, the ambiguity of genetic differentiation is clearly traced at the level of the cytb gene. It involves not only the different scale of taxa of the same level, which is a consequence of ignoring the stage of cryptic speciation by taxonomists, but also an objective circumstance — the nature of mutations, which determines the different rates of molecular evolution in short and long cycle mammals.

References

Altukhov, Yu. P. 1990. *Population genetics: diversity and stability*. Harwood Acad. Publ., London.

Amadon, D. 1966. The superspecies concept. *Systematic Zoology*, 15, 246–249.

Anderson, R. F. V. 1977. Ethological isolation and competition of allospecies in secondary contact. *The American Naturalist,* 111, 939–949.

Baccus, R., Ryman, N., Smith, M. H., Reuterwall, C. & Cameron, D. 1983. Genetic variability and differentiation of large grazing mammals. *Journal of Mammalogy,* 64, 109–120.

Belle, E., Piganeau, G., Gardner, M. & Eyre-Walker, A. 2005. An investigation of the variation in the transition bias among various animal mitochondrial DNA. *Gene, 355*, 58–66. doi:10.1016/j.gene.2005.05.019

Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. *Journal of Molecular Evolution,* 18*,* 225–239. doi:10.1007/BF01734101

Collins, D. W., & Jukes, T. H. 1994. Rates of transition and transversion in coding sequences since the human-rodent divergence. *Genomics,* **20** (3), 386–396. doi:10.1006/geno.1994.1192

Duchene, S., Ho, S. Y. & Holmes, E. C. 2015. Declining transition/transversion ratios through time reveal limitations to the accuracy of nucleotide substitution models. *BMC Evolutionary Biology*, *15*(36). doi:10.1186/s12862-015-0312-6

Ebersberger, I., Metzler, D., Schwarz, C. & Pääbo, S. 2002. Genomewide comparison of DNA sequences between humans and chimpanzees. *American Journal of Human Genetics,* 70, 1490–1497. doi:10.1086/340787

Fitch, W. M. 1967. Evidence suggesting a non-random character to nucleotide replacements in naturally occurring mutations. *Journal of Molecular Biology*, *26*, 499–507. doi:10.1016/0022-2836(67)90317-8

Ge, D., Feijo, A., Cheng, J., Lu, L., Liu, R., Abramov, A. V., Xia, L., Wen, Z., Zhang, W., Shi, L. & Yang, Q. 2019. Evolutionary history of field mice (Murinae: *Apodemus*), with emphasis on morphological variation among species in China and description of a new species. *Zoological Journal of Linnaen Society*, 187, 518–534. doi:10.1093/zoolinnean/zlz032

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence aligment editor and analysis program for Windows 95/98/NT. *Nucleic Acid Simposium Series*, 41, 95–98.

International code of zoological nomenclature. 1999. *Fourth Edition. International Commission on Zoological Nomenclature. The International Trust for Zoological Nomenclature*. Retrieved from https://www.iczn.org/the-code/the-code-online/

Keller, I., Bensasson, D. & Nichols, R. A. 2007. Transition-transversion bias is not universal: a counter example from grasshopper pseudogenes. *PLoS Genetics*, **3** (2), e22. doi:10.1371/journal.pgen.0030022

Kumar, S. 1996. Patterns of nucleotide substitution in mitochondrial protein coding genes of vertebrates. *Genetics*, 143, 537–548. doi:10.1093/genetics/143.1.537

Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution*, 35, 1547–1549.

Li, W.-H. & Graur, D. 1991. *Fundamentals of molecular evolution* (1st ed.). Sinauer Associates, Sunderland, MA.

Mallet, J. 2007. Subspecies, semispecies, superspecies. *Encyclopedia of Biodiversity*, 7, 45–48. doi:10.1016/B978-0-12-384719-5.00138-6

Mezhzherin, S. V. 1994. Taxonomy and modern views of the house mice system of Palearctic. *In*: Kotenkova, E. V. & Bulatovs, N. Sh., eds. *House mouse: origin, distribution, systematic, behavior*. Nauka, Moscow, 15–27 [In Russian].

Mezhzherin, S. V. 2002. Correlation between genetic variability and body size in vertebrates. *Russian Journal of Genetics*, 38, 1252–1258.

Mezhzherin, S. V. & Morozov-Leonov, S. Y. 1995. The genetic differentiation of mammalian taxa: their assessment by biochemical genetic markers. *Zhurnal obshchey biologii*, 56, 71–96.

Mezhzherin, S. V. & Morozov-Leonov, S. Y. 1996. The multifariousness of the gene differentiation of taxa at the species and genus levels and the biological heterogeneity of mammalian orders. *Zhurnal obshchey biologii*, 57, 79–94.

Mezhzherin, S. V., Morozov-Leonov, S. Y., Zhalay, O. I., Kokodiy, S. V., Tereshchenko, V. O., Rostovskaya, O. V. & Tsyba, A. O. 2023. Evolutionary transition/transversion bias by the example of the cytb gene of Palearctic Muridae (Rodentia) and Vespertilionidae (Chiroptera). *Repots of NAS of Ukraine*, 2, 93–98. /doi:10.15407/dopovidi2023.02.093

Mezhzherin, S. V. & Tereshchenko, V. O. 2023. Genetic divergence and evolutionary transition/transversion rate bias in control region of mitochondrial DNA of Palearctic mice (Murinae). *Cytology and Genetics*, 57, 213–223. doi:10.3103/s0095452723030076

Michaux, J. R., Chevret, P., Filippucci, M. G. & Macholan, M. 2002. Phylogeny of the genus *Apodemus* with a special emphasis on the subgenus *Sylvaemus* using the nuclear IRBP gene and two mitochondrial markers: Cytochrome b and 12S rRNA. *Molecular Phylogenetics and Evolution*, 23, 123–136. doi:10.1016/S1055-7903(02)00007-6

Philippe, H., Brinkmann, H., Lavrov, D. V., Timothy, D., Littlewood, J., Manuel, M., Wörheide, G. & Baurain, D. 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biology,* 9 (3). doi:10.1371/journal.pbio.1000602

Stegnii, V. N. 1993. *Genome architectonics: systemic mutations and evolution*. Novosibirsk State University, Novosibisk.

Stoltzfus, A. & Norris, R. W. 2016. On the causes of evolutionary transition:transversion bias. *Molecular Biology and Evolution, 33*, 595-602. doi:10.1101/027722

Tamura, K., Stecher, G. & Kumar, S. 2021. MEGA11: molecular evolutionary genetics analysis. Version 11. *Molecular Biology and Evolution*, *25*, 3022-3027. doi:10.1093/molbev/msab120

Vogel, F. & Kopun, M. 1977. Higher frequencies of transitions among point mutations. *Journal of Molecular Evolution*, 33, 595–602. doi:10.1007/BF01732746

Wooten, M. C. & Smith, M. H. 1985. Large mammals are genetically less variable? *Evolution*, 39, 210–212. doi:10.2307/2408532

Received 14 April 2023
Accepted 1 February 2024