
BIOCHEMISTRY, BIOPHYSICS
AND MOLECULAR BIOLOGY

Nuclear mtDNA Pseudogenes As a Source of New Variants of Mitochondrial Genes: A Case Study of Siberian Rubythroat *Luscinia calliope* (Muscicapidae, Aves)

L. N. Spiridonova^a, Ya. A. Red'kin^b, and O. P. Valchuk^a

Presented by Academician Yu.N. Zhuravlev June 17, 2015

Received October 22, 2015

Abstract—First evidence for the presence of copies of mitochondrial cytochrome *b* gene of the subspecies group *Luscinia calliope anadyrensis*—*L. c. camtschatkensis* in the nuclear genome of nominative *L. c. calliope* was obtained, which indirectly indicates the nuclear origin of the subspecies-specific mitochondrial haplotypes in Siberian rubythroat. This fact clarifies the appearance of mitochondrial haplotypes of eastern subspecies by exchange between the homologous regions of the nuclear and mitochondrial genomes followed by fixation by the founder effect. This is the first study to propose a mechanism of DNA fragment exchange between the nucleus and mitochondria (intergenomic recombination) and to show the role of nuclear copies of mtDNA as a source of new taxon-specific mitochondrial haplotypes, which implies their involvement in the microevolutionary processes and morphogenesis.

DOI: 10.1134/S1607672916010142

The exchange of genetic information between cellular organelles and the nucleus is considered to be a common phenomenon in all eukaryotes. The transfer of a large mtDNA fragment (7.9 kb) was first described for the domestic cat *Felis catus* [1]. The authors of this study proposed the term NUMT (“Nuclear copies of Mitochondrial genes”) for the nuclear copies of mtDNA genes and a hypothetical model of mtDNA fragment integration into the nuclear genome. However, the actual mechanisms of DNA movement and integration between the mitochondria and the nucleus have not yet been established. Currently, many cases of detecting nuclear copies of mtDNA genes in different groups of organisms (fungi, plants, arthropods, birds, and mammals) have been described. However, the facts of inverse symmetry (i.e., the insertion of nuclear DNA sequences in the mitochondrial genome) have not been known before.

In this work, in the course of phylogeographic studies of the population structure of Siberian rubythroat *Luscinia calliope* (Pallas, 1776), we for the first time established the fact of the transition of a nuclear copy sequence into the mitochondrial genome, which led to the occurrence of taxon-specific mitochondrial hap-

lotypes in several subspecies (geographical races) of this species. We have proposed and indirectly confirmed the existence of yet undescribed mechanism for the exchange of homologous regions between the nuclear and mitochondrial DNA via symmetric recombination. The results of our discovery broaden the understanding of the processes occurring in the cell, indicate the cases of interaction between the nuclear and mitochondrial genomes as a coupled genetic system of the cell, and allow rethinking the importance of the study of mitochondrial genes for assessing the rate of microevolutionary processes and morphogenesis.

In this paper, we present data on detecting copies of mitochondrial cytochrome *b* (*cyt b*) gene of the north-eastern subspecies group *Luscinia calliope anadyrensis*—*L. c. camtschatkensis* in the nuclear genome of the form *L. c. calliope*, which testified to the nuclear origin of the mitochondrial haplotypes in these subspecies of Siberian rubythroat. The presence of copies of mitochondrial genes of other subspecies in the nucleus of *L. c. calliope* indicates intergenomic recombination events that yielded new recombinant mitochondrial haplotypes of some geographical races of this species.

Siberian rubythroat *Luscinia calliope* is a widespread migratory Palearctic species of passerine birds. Molecular genetic studies of mtDNA *cyt b* gene of this species showed its division into three significantly different haplogroups [2, 3]: “western” (*calliope*) and two “eastern”—Sakhalin and Anadyr—Kamchatka, corresponding to forms *L. c. sachalinensis* (Portenko, 1937), *L. c. anadyrensis* (Portenko, 1939), and

^a Institute of Biology and Soil Sciences, Far East Branch, Russian Academy of Sciences, pr. Stoletiya Vladivostoka 159, Vladivostok, 690022 Russia

e-mail: spiridonova@biosoil.ru

^b Zoological Museum of Moscow State University, Moscow, Russia

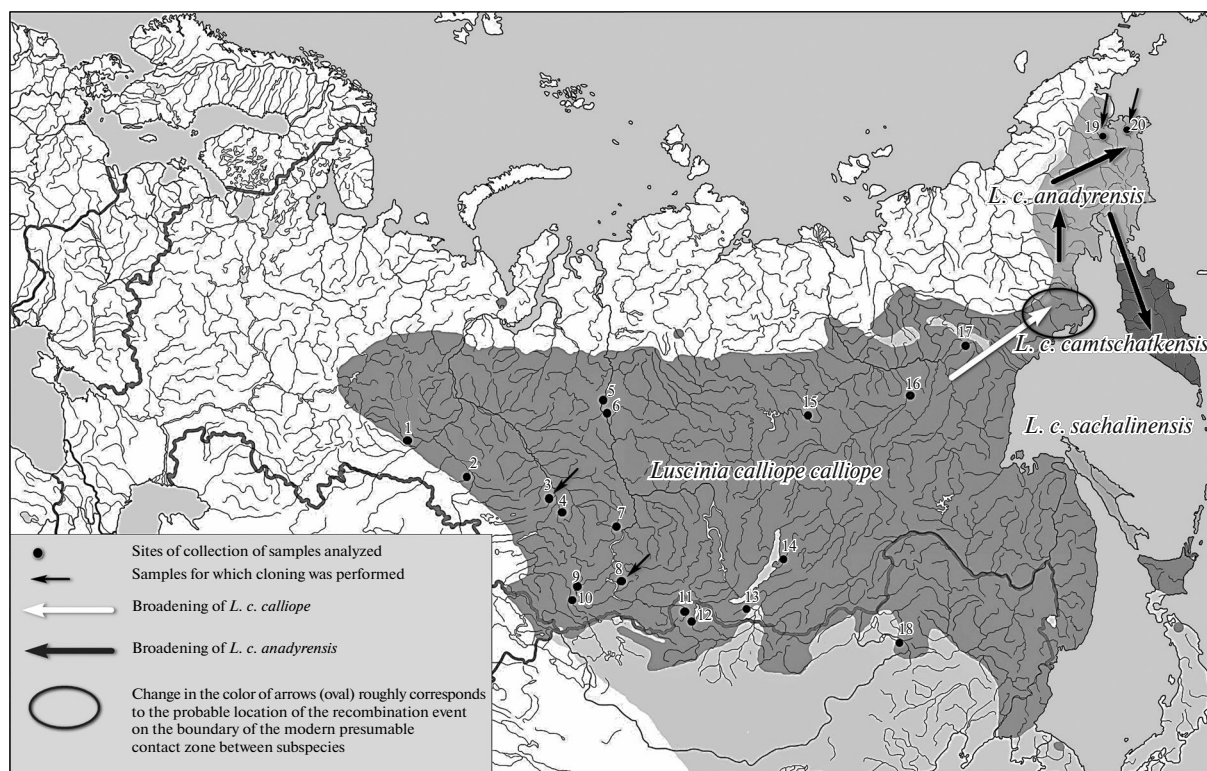


Fig. 1. Sites of collection of samples analyzed.

L. c. camtschatkensis (Gmelin, 1789). The strong genetic divergence among the haplotypes *calliope*, *anadyrensis*–*camtschatkensis*, and *sachalinensis* had no logical explanation, because there is no obvious geographical disjunctions between the ranges of subspecies of the “western” and “eastern” groups and the entire northeastern part of the range (Anadyr, Kamchatka, and Magadan region) is represented by populations with both mtDNA variants in different proportions. In addition, the issue on the origin of the haplotype *sachalinensis*, which is represented solely on Sakhalin and is absent in adjacent mainland areas, remained unresolved. To clear up these issues and a more complete phylogeographic study of haplotypes of the nominative subspecies *calliope*, we studied the birds from the previously unexplored parts of the range of Siberian rubythroat.

We analyzed 21 *L. c. calliope* samples from the collection of the Zoological Museum of Moscow State University, which were collected in 18 localities from 1906 to 2011 (Fig. 1). DNA was isolated from derivatives (foot pads and feathers) using the QIAgen DNeasy^R Tissue Kit (Qiagen, United States) according to protocol no. 9 [4]. The *cyt b* gene fragment and the nuclear copies of the same gene were amplified using different pairs of primers of our own design.

Visual testing for the quality of graphical chromatograms revealed the presence of equally repeating double peaks with varying signal intensity, against

the background of clear signals, in some sites in 19 samples. Since all sequences had a similar distribution of heteroplasmic sites, for preliminary analysis we randomly selected two samples from the western sites located at a distance of 650 km from each other (Tomsk and Ermak districts of Krasnoyarsk krai). To separate the sequences found in the heterogeneous fragment as a result of sequencing, the amplicons were cloned using the InsTAcloneTM PCR cloning kit (Fermentas, Lithuania) according to manufacturer's instruction. We selected, on average, ten clones for each sample. In total, 25 clones approximately 900 bp long were analyzed.

One variant of the clones of both samples corresponded to the mitochondrial *cyt b* gene of the nominative subspecies *L. c. calliope*, and the second variant was the nuclear *cyt b* pseudogene similar to the mitochondrial haplotypes of the eastern group, detected by us earlier and characteristic of the subspecies *L. c. anadyrensis* and *L. c. camtschatkensis* (98% similarity, Fig. 2). The nuclear pseudogenes contained determinant codons characteristic of NUMT. In addition, the high level of polymorphism, caused by randomly occurring mutations, also confirms the nuclear origin of these sequences.

We assumed that the mechanism of transfer of the nuclear copy identical to the subspecies *anadyrensis* into the mitochondrial DNA could be symmetric recombination between the homologous regions of the

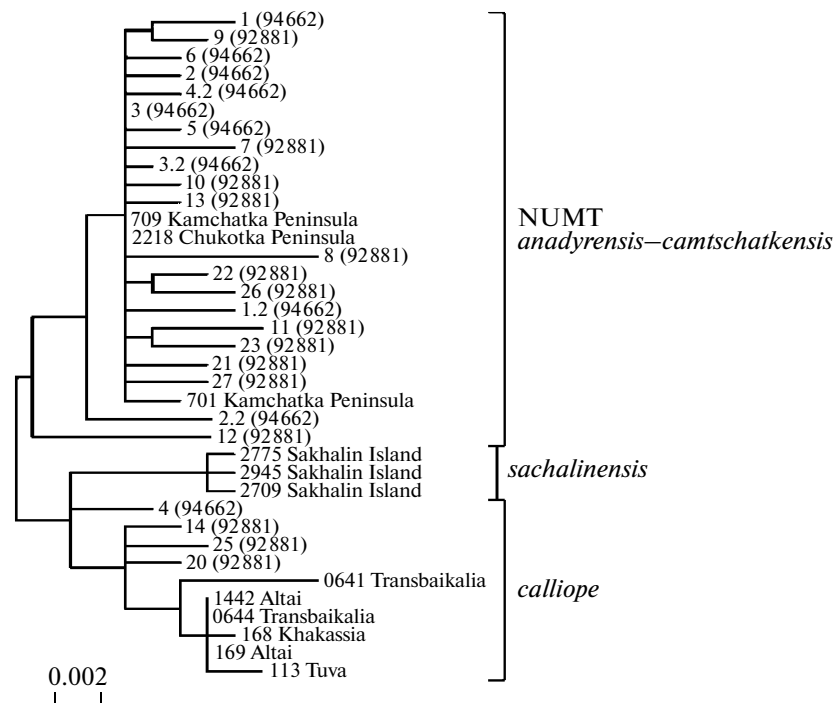


Fig. 2. Maximum Likelihood (ML) Tree for gene fragment clones *cyt b* (867 bp) from samples *L. c. calliope* Number 94662 and 92881. The numbers before the brackets denote the number of clones analyzed sample.

nuclear and mitochondrial genomes. Therefore, the nuclear genome of the Anadyr subspecies should contain a copy of the mitochondrial *cyt b* gene of the nominative subspecies *L. c. calliope*. To confirm this hypothesis, we first tested the samples of the Anadyr subspecies of Siberian rubythroat for the presence of heteroplasmic sites. Amplification of the *cyt b* gene with specific primers did yield a nuclear homologue. We have experimentally established that the primers for a longer fragment (3202 bp), comprising the *cyt b* gene and the control region (D-loop), covered both the mitochondrial and nuclear sequences. In view of this, we performed two PCR. In the first PCR, a long amplicon (3202 bp) comprising *cyt b* and *CR* genes was synthesized. In the second reaction, involving the first PCR product, the fraction of the nuclear homologue of the *cyt b* gene 1200 bp long was enriched. The products of the second PCR were sequenced, and the samples with the heteroplasmic sites were used for subsequent cloning. Nucleotide sequences of two *L. c. anadyrensis* samples with the previously determined mitochondrial haplotypes—no. 2053 (Acc. LK932608, haplotype *anadyrensis-camtschatkensis*, Chukotka Autonomous Area, Anadyr district) and no. 2251 (Acc. LK932617, haplotype *calliope*, Chukotka Autonomous Area, Bering district)—were cloned (19, 20, Fig. 1). The cloned sequences of bird no. 2053 were divided into two variants: *cyt b* gene of the group *anadyrensis-camtschatkensis* and a nuclear copy of the *cyt b* gene of the subspecies *calliope*, which

confirmed our hypothesis. Unexpected were the cloning results for bird no. 2251 with the *calliope* mitochondrial haplotype. One clone variant was the mitochondrial *cyt b* gene of the nominative subspecies *L. c. calliope*, and the second was a nuclear copy identical to the *cyt b* gene of the subspecies *L. c. sachalinensis*, which was previously found only on Sakhalin Island.

The discovered nuclear copies of mtDNA indirectly indicate that the mitochondrial haplotypes of the Anadyr, Kamchatka, and Sakhalin subspecies evolved as a result of symmetric recombination of homologous regions of the nuclear and mitochondrial genomes and the subsequent distribution of the recombinant haplotypes by the founder effect. Since in these cases a region of the mitochondrial DNA of *calliope* is replaced with the homologous nuclear copies of *anadyrensis-camtschatkensis* and *sachalinensis*, it can be assumed that the nuclear genome of the original nominative subspecies *L. c. calliope*, containing the copies of mitochondrial genes (in particular, *cyt b* gene) was the source of the new taxon-specific variants of mitochondrial haplotypes.

The probability of occurrence of the same complex molecular transfers simultaneously in several individuals is extremely low. We believe that the recombination event for each subspecies occurred once at the first stages of formation of germ cells, which in the embryogenesis of birds occurs very early, in the first third of embryogenesis, as was experimentally established in the studies of chicken embryos [5]. The

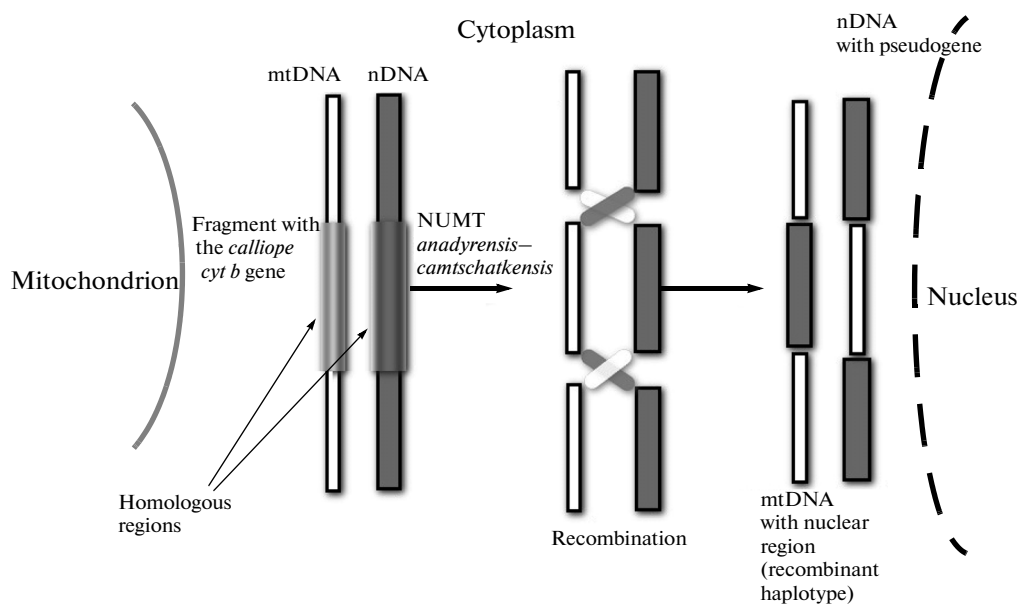


Fig. 3. Presumable scheme of recombination between mtDNA and nuclear DNA (nDNA), which yielded the recombinant haplotype of the subspecies *L. c. anadyrensis* and *L. c. camtschatkensis*.

zygote divisions after fertilization are not accompanied by the mitochondrial divisions, as a result of which the number of mtDNA copies per cell is reduced from 200 thousand in the egg to 5 thousand per cell in the blastocyst, which was shown in mammals [6]. After implantation, in the course of cell differentiation, primary germ cells (gonocytes), which contain approximately 10 mtDNA copies per cell, become distinguished. Mitochondria of germ cell progenitors account for a small proportion (0.01%) of the total initial mitochondrial pool of the zygote. Due to the drastic decrease in the number of mitochondria (approximately 20 thousand times), the mtDNA diversity in the cell is greatly reduced [7]. The intergenomic recombination, assumed in our work, apparently occurred at the stage of gonocytes, which contain the least number of mitochondria, which many times increases the possibility of maintaining and passing the new variants of mitochondrial DNA to the progeny.

The enzymology of general recombination has been well studied only in some prokaryotic organisms [8]. One of the specific enzymes required for successful homologous recombination is the *recA* protein. It catalyzes the single strand exchange using the energy of ATP hydrolysis to ADP and inorganic phosphate. The first stage of the recombination process in the Holliday scheme is *recA*-dependent insertion of single-stranded DNAs into the heteroduplex to form double-stranded breaks. The second enzyme, *recBCD* nuclease, which consists of three individual subunits (B, C, and D), exhibits endo- and exonuclease and helicase activities. Its mechanism of action is not understood completely; however, it is known that *recBCD* nuclease induces breaks in the DNA duplex

and, due to its intrinsic helicase activity, together with *recA* initiates the recombination process. General recombination involves several other enzymes that also catalyze DNA replication and repair [8].

The probability that intergenomic recombination occurred in the cytoplasm we explain by the known facts and respective assumptions. At the stage of meiosis, after the nuclear membrane destruction, the nuclear DNA gets into the cytoplasm (Fig. 3). The mtDNA can penetrate into the cytoplasm apparently under the influence of both genetic and environmental factors, such as exposure to mutagenic agents and other types of cell stress, which can damage mitochondria or their membranes [9].

The intergenomic exchange of nucleotide sequences in the case of *calliope*–*anadyrensis*–*camtschatkensis* presumably occurred in the northeastern peripheral part of the range of the subspecies *calliope* at a low abundance of birds, most likely in the region where the modern contact zone of this form with *L. c. anadyrensis* is located (Fig. 1). This allowed the new haplotype to fix in time and become taxon-specific upon the appearance of morphological characters of the new forms and their dispersal to new territories in the northeast Asia. Currently, in the Siberian rubythroat populations in the Anadyr Basin and Koryak Highlands, the mitochondrial haplotypes of *calliope* and the eastern recombinant haplotype are mixed in approximately equal proportions, with retaining the stability of the morphological traits characteristic of the Anadyr subspecies. Southwards, on the Kamchatka Peninsula, *L. c. camtschatkensis* populations, which are easily distinguishable from *anadyrensis*, are represented primarily by the birds with the same recombinant haplotype, whereas the *calliope*

haplotype is found here very rarely. Presumably, the recombination event for the case *calliope*–*sachalinensis* occurred on Sakhalin, because the mitochondrial haplotype of the Sakhalin subspecies is found only on this island. The phenomenon of recombination, in our opinion, is a relatively recent event, because the discovered nuclear copies of the *cyt b* gene have a high similarity (98–99%) with the mitochondrial genes of respective subspecies of the “eastern” group—*L. c. anadyrensis*, *L. c. camtschatkensis*, and *L. c. sachalinensis*.

In our opinion, the opposite assumption—the insertion of the eastern haplotype into the nuclear genome of *L. c. calliope*—it unlikely for several reasons. Firstly, the areas where the nuclear pseudogene was found (Western and Central Siberia) are remote from the breeding range of *L. c. anadyrensis*. Secondly, the Siberian rubythroat is characterized by a high nesting conservatism and has permanent migration routes. In addition, according to the observations of ornithologists, the range of *L. calliope* proceeded expanded from west to east [10, 11]. Furthermore, the eastern mitotype per se, according to our data, has not been found (except one case in Mongolia) westward of Blagoveshchensk (Amur region).

In the case of symmetric recombination between the homologous regions of the mitochondrial and nuclear genomes, the emergence of new mitochondrial haplotypes occurs on the basis of nuclear pseudogenes. Figure 3 shows the presumable scheme of the recombination event between the homologous fragments of the nuclear and mitochondrial DNA, which yielded new recombinant haplotypes, which are currently characteristic of the subspecies *L. c. anadyrensis*, *L. c. camtschatkensis*, and *L. c. sachalinensis*.

In the nuclear genome, mtDNA pseudogenes are more prone to recombination rearrangements, which leads to the emergence of their new variants. For example, the existence of several significantly diverged variants of haplotypes that are commonly divided by the geographical principle into the “western” and “eastern” ones, in the representatives of some corvid genera (genus *Corvus* [12] and genus *Pica* [13]), apparently illustrates the common phenomenon for different cases that is conventionally attributed to the consequences of glaciation, leading to insulation under conditions of long-term isolation of groups of populations of various species in remote refugia. Under certain circumstances described above, the random intergenomic recombination more rapidly lead to the emergence of new variants of mitochondrial gene haplotypes as compared to the spontaneous mutations in the mitochondrial genome itself and can be an alternative cause of the cases described above.

Thus, in this paper we for the first time described the cases of finding nuclear DNA in the mitochondrial genome and proposed the mechanism of DNA fragment exchange between the nucleus and mitochondria (intergenomic recombination). We also showed the

role of the nuclear copies of mtDNA as a source of new taxon-specific mitochondrial haplotypes, which indicates their importance in microevolutionary processes and morphogenesis. From a phylogenetic standpoint, NUMTs are often considered as a hindrance leading to erroneous interpretations of relationships between taxa [14]. Our findings provide a new look for the presence of NUMT sequences in the nuclear genome and open new opportunities for using similar cases in phylogenetic studies.

ACKNOWLEDGMENTS

We are grateful to Academician Yu.N. Zhuravlev and A.P. Kryukov for their attention to our research, assistance, and valuable comments, and to K.V. Rozhkov for methodological advice at the cloning stage. This study was supported by the Russian Science Foundation (project no. 14-50-00029 “Scientific Basis for Creating a National Depository Bank of Living Systems.”)

REFERENCES

1. Lopez, J.V., Yuhki, N., Masuda, R., Modi, W., and O'Brien, S.J., *J. Mol. Evol.*, 1994, vol. 39, no. 5, pp. 174–190.
2. Spiridonova, L.N., Val'chuk, O.P., Belov, P.S., and Maslovskii, K.S., *Russ. J. Genet.*, 2013, vol. 49, no. 6, pp. 638–644.
3. Spiridonova, L.N., Valchuk, O.P., Red'kin, Ya.A., and Saitoh, T., in *XXVI Intern. Ornithol. Congress*, Tokyo, 2014, p. 382.
4. Nishiguchi, M.K., Doukakis, P., Egan, M., et al., *Methods and Tools in Biosciences and Medicinem*, Washington, Basel: Birkhauser Verlag, 2002, pp. 249–287.
5. Strizhikova, S.V., Strizhikov, V.K., and Zhitenko, N.V., *Usp. Sovrem. Estestvozn.*, 2002, no. 4, pp. 77–78.
6. White, D.J., Wolff, J.N., Pierson, M., and Gemmell, N.J., Revealing the hidden complexities of mtDNA inheritance, *Mol. Ecol.*, 2008, vol. 17, no. 23, pp. 4925–4942.
7. Shoubridge, E.A. and Wai, T., *Curr. Top. Develop. Biol.*, 2007, vol. 77, pp. 87–111.
8. Singer, M. and Berg, P., *Geny i genomy* (Genes and Genomes), Moscow: Mir, 1998, vol. 2.
9. Blanchard, J.L. and Schmidt, G.W., *J. Mol. Evol.*, 1996, vol. 13, no. 6, pp. 537–548.
10. Nazarenko, A.A., *Zool. Zh.*, 1979, vol. 58, pp. 1680–1691.
11. Dorzhiev, Ts.Z., *Vestn. Buryat. Gos. Univ., Ser. Spets.*, 2006, no. 4, pp. 68–91.
12. Haring, E., Pinsker, W., Kryukov, A., and Gamauf, A., *J. Zool. Syst. Evol. Res.*, 2012, vol. 50, no. 3, pp. 230–246.
13. Kryukov, A., Iwasa, M.A., Kakizawa, R., Suzuki, H., Pinsker, W., and Haring, E., *J. Zool. Syst. Evol. Res.*, 2004, vol. 42, no. 4, pp. 342–351.
14. Bernt, M., Braband, A., Schierwater, B., and Stadler, P.F., *Mol. Phys. Evol.*, 2013, vol. 69, no. 2, pp. 328–338.

Translated by M. Batrukova