

Different rates of molecular evolution of mitochondrial genes in Baikalian and non-Baikalian amphipods

Romanova E.V.¹, Sherbakov D.Y.^{1,2*}¹ Limnological Institute, Siberian Branch of the Russian Academy of Sciences, Ulan-Batorskaya Str., 3, Irkutsk, 664033, Russia² Irkutsk State University, Karl Marx Str., 1, Irkutsk, 664003, Russia

ABSTRACT. Rates of molecular evolution in eight protein-coding mitochondrial genes from two parallel lineages of Baikalian amphipods were compared to those in the representatives of gen. *Gammarus*. In six genes (Atp6, Cox2, Cox3, Nad1, Nad5, Nad6), there was a significant acceleration in the evolutionary rate of Baikalian species over those from *Gammarus* group. Correlation of the absolute rate of base substitution accumulation and the acceleration was insignificant, that allowed us to propose mostly adaptive reasons for the effect found.

Keywords: Rate of evolution, mitogenome, genetic distances, Baikal

1. Introduction

The accumulation rate of base substitutions (rate of molecular evolution) depends on various factors, including physiological characteristics, environmental conditions and life history of species (Fazalova et al., 2010; Ellegren and Galtier, 2016). Therefore, significant changes in the rate of molecular evolution in sister groups may provide valuable insight into driving forces of adaptive radiation and speciation. In particular, some studies showed a positive correlation between the rate of molecular evolution and the rate of species diversification (Lanfear et al., 2010; Bromham et al., 2013).

One of the most attractive and mysterious cases of a wide variation of diversification rates is the adaptive radiation of amphipods in Lake Baikal where they evolved into two very diverse species flocks. These flocks result from two independent introductions of the gen. *Gammarus* representatives into the lake (Macdonald et al., 2005; Hou and Sket, 2016; Naumenko et al., 2017). Since the species of gen. *Gammarus* are very uniform morphologically (Pinkster, 1983), the adaptive radiations of the *Gammarus* descendants in Lake Baikal led to the appearance of two very diverse morphologically and ecologically amphipod lineages (Kamaltynov, 2009).

This peculiar evolutionary background makes it possible to test the hypothesis, according to which an acceleration of speciation and morphological diversification is accompanied by an acceleration of the rate at which base substitutions are fixed in genomes (Rabosky, 2019; Wiens and Scholl, 2019).

Several mitochondrial genomes (mitogenomes) have been sequenced recently from the representatives of both Baikalian species and several species of gen. *Gammarus* (Krebs and Bastrop, 2012; Romanova et al., 2016; Macher et al., 2017; Cormier et al., 2018; Sun et al., 2020). Since the proteins coded in mitochondrial DNA are not directly involved in morphogenesis, the deviations in the rate of their molecular evolution may serve as an indication of changes in the average rate of molecular evolution in the respective species.

In this paper, we use nucleotide sequences of protein-coding genes of mitogenomes to address the question if there were parallel changes in the rates of molecular evolution during adaptive radiation in the two species flocks of Baikalian amphipods.

2. Materials and methods

2.1. Mitochondrial genome sequences

Complete and nearly complete mitogenomes of amphipods available in GenBank by June 2019 were used in this paper: KX341962, KP161875 GU130250 AB539699 KX341963 MF361127 MF361126 KX341964 KF690638 KM287572 MF766257 JN704067 KY197961 MG779536 KX341965, KX341966 JN827386 LT594768 MH542433 LT594767 KX341967, FJ555185 KX341968 AY639937 MH542432 MG010370, MG010371 MH542430 MH229998, FR872382 - FR872383, HE860495 - HE860513, HE861923, FR872383, LN871175, LN871176, MH592123, MH592125 - MH592152, KU869711 - KU869713, MH542431,

*Corresponding author.

E-mail address: sherb@lin.irk.ru (D.Y. Sherbakov)

MK644228, MK354235, MK354236, and MH294484. Protein-coding genes were obtained according to the original annotations. The codon tables were also not verified, although the deduced amino acid sequences were checked for the absence of stop codons. All mitogenomes were downloaded separately and split into separate protein-coding genes that were translated and aligned using SeaView (Gouy et al., 2010) coupled with Mafft (Katoh and Standley, 2013) and then operated with custom Biopython scripts (Cock et al., 2009) available from the authors.

2.2. Phylogenetic inferences

The Maximum Likelihood (ML) tree topology based on mitochondrial protein-coding genes of all amphipods were inferred from their combined amino acid sequences added by three crustacean species: *Eophreatoicus* sp. 14FK-2009 (FJ790313), *Ligia oceanica* (DQ442914), and *Neomysis japonica* (KR006340) as an outgroup, using IQ_TREE software (Nguyen et al., 2014). It was applied to the single partition containing

all combined amino acid sequences. The mtMet + F + R9 (nine categories of substitution rates) model was chosen using ModelFinder (Kalyaanamoorthy et al., 2017) built-in IQ_TREE. Both ST-*alrt* and ultra-fast bootstrap measures of statistical confidence were used. The tree obtained was further used as the topological constraint for the calculations of the base substitution rates within the separate protein-coding genes of the amphipods.

2.3. Rates of molecular evolution

Differences in rates of molecular evolution were estimated based on the nucleotide sequences of mitogenomes of the two independent lineages of Amphipods endemic to Lake Baikal, which are further assigned as “lineage1” and “lineage2”, and the genus *Gammarus* (Fig. 1). The latter provides two independent ancestors giving rise to the two Baikalian species flocks (Hou and Sket, 2016).

A common ancestor of the Baikalian and *Gammarus* species was defined by the local outgroup *Onisimus nanseni*. The tree topology inferred from the

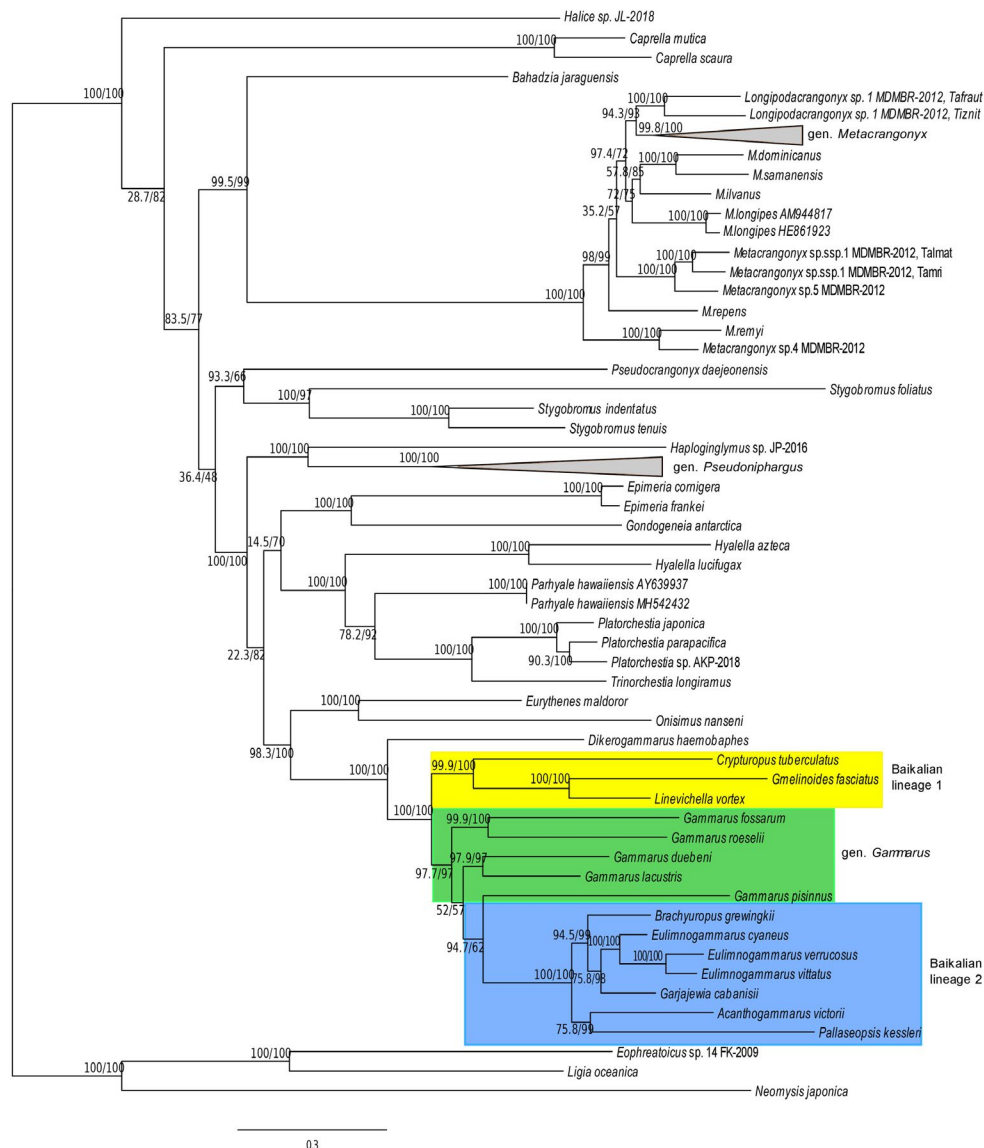


Fig. 1. ML phylogenetic tree of amphipods based on 13 amino acid sequences of mitochondrial protein-coding genes. Numbers at the branches correspond to ultrafast bootstrap support and relative rate test values (%). Clades of two Baikalian amphipod lineages and the representatives of gen. *Gammarus* group are highlighted in colors.

combined amino acid sequences of all mitochondrial protein-coding genes from available amphipod mitogenomes was used as the topological constraint for this study.

Nucleotide sequences of each protein-coding mitochondrial gene in all 94 mitogenomes studied were aligned separately, and the tree inferred from amino acid sequences was used as the topological constraint. IQ-TREE was applied to estimate branch lengths of this tree using a gene-specific model of molecular evolution. This model was determined for each gene using the ModelFinder (Kalyaanamoorthy et al., 2017). The summa of all tree edges connecting terminal leaves with the outgroup was taken to estimate the base substitution rate for each leaf.

Protein-coding genes CO1, CO2, CO3, Atp6, Atp8, Nad1, Nad5, and nad6 were used in the relative rate analysis. The other five standard mitochondrial genes were excluded from the analysis because of incomplete sequences in some crucial amphipod species. The significance of the difference in the substitution rates was tested with the Mann-Whitney non-parametric U-test (Mann and Whitney, 1947) between the groups of edge lengths for both Baikalian lineages and the *Gammarus* group.

3. Results and discussion

ML phylogenetic tree obtained in this study (Fig. 1), has highly supported topology similar to that previously described by Romanova et al. (2020). Minor difference dues to the few recently sequenced amphipod species included in the analysis: (*Dikerogammarus haemobaphes* (MK644228), *Halice* sp. JL-2018 (MH294484), *Gammarus lacustris* (MK354235), and *Gammarus pisinnus* (MK354236). Two latter species allowed us to obtain a better-supported separation between the Baikalian lineages.

For each mitogenome, the edge lengths between the leaf and the outgroup were calculated using ete2 Python library (Huerta-Cepas et al., 2010). The sequences were attributed to three groups as shown in Figure 1 by color boxes. The pairwise differences between the edge lengths of respective genes indicate that all genes evolve at obviously different rates (Fig. 2). There was also some acceleration between the two Baikalian lineages and the *Gammarus* lineage.

To estimate the significance of differences in the substitution rates between the groups of edge lengths for both Baikalian lineages and the *Gammarus* group, the accumulation of all pairwise differences between groups of sequences were calculated (Fig. 3). As the distributions of these differences were far from normal, the Mann - Whitney non-parametric U-test (Mann and Whitney, 1947) was used to find if the distribution median significantly differs from 0 or not (null hypothesis). The analysis showed no statistically significant differences in substitution rates in analyzed gene sequences of Baikalian amphipod lineages, except for the Cox1 and Nad6 genes. In these two cases, the acceleration in lineage 2 was significantly higher than in lineage 1.

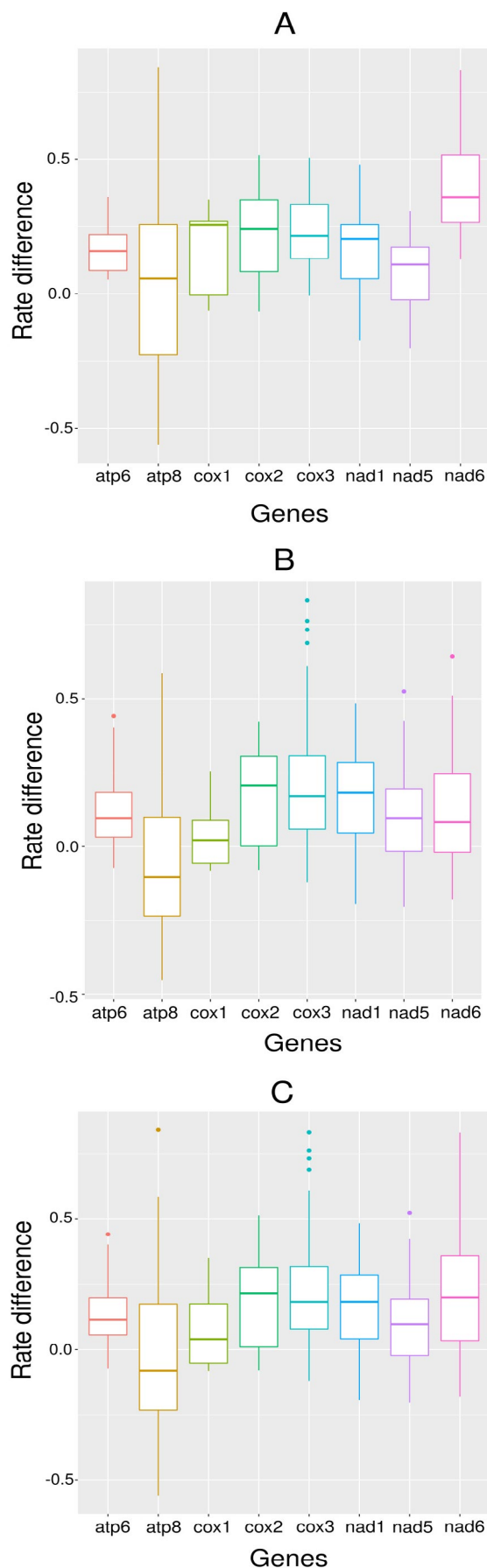


Fig. 2. Box plots of pairwise rate difference values. A – Difference between branch lengths of Baikalian amphipods from lineage 2 and the *Gammarus* group. B - Difference between branch lengths of Baikalian amphipods from lineage 1 and the *Gammarus* group. C – Difference between branch lengths of Baikalian amphipods from both Baikalian lineages and the *Gammarus* group.

The acceleration of substitution rates was detected, however, in both Baikalian lineages compared to the *Gammarus* group, except for Atp8 and Cox1 genes (Fig. 3, Table 1).

The results of the comparison indicate that the rate of molecular evolution in Baikalian amphipods is accelerated in almost all tested genes in parallel. The parallel acceleration of the molecular evolution rate may be explained by either a higher mutation rate in Lake Baikal or by some adaptive constraints specific to the lake. In the first case, one may expect that the difference accumulates faster in faster-evolving genes. In other words, a significant positive correlation between the average amount of difference in the average rates of the group and the rates of molecular evolution would suggest non-adaptive reasons for the acceleration in Lake Baikal. In Figure 4, the gene-specific acceleration in the rate is plotted against the average rates of molecular evolution in each gene. Since the differences were not normally distributed (Fig. 3),

Table 1. The result of testing the hypothesis about the acceleration of the molecular evolution rate in Baikalian amphipod lineages. Bold numbers indicate cases of statistically significant acceleration of substitution rate.

Gene	Pairwise differences between group members (<i>p</i> -values)		
	Lineage2-Lineage1	Lineage2-gen. <i>Gammarus</i>	Lineage1-gen. <i>Gammarus</i>
Atp6	0.1373	1.38E-07	6.00E-05
Atp8	0.1571	0.2383	0.7615
Cox1	0.0018	0.1954	6.70E-03
Cox2	0.055	2.16E-05	1.60E-03
Cox3	0.54	8.63E-08	1.20E-04
Nad1	1	1.07E-05	5.40E-03
Nad5	0.97	1.80E-03	0.04
Nad6	1.90E-06	2.70E-03	6.10E-05

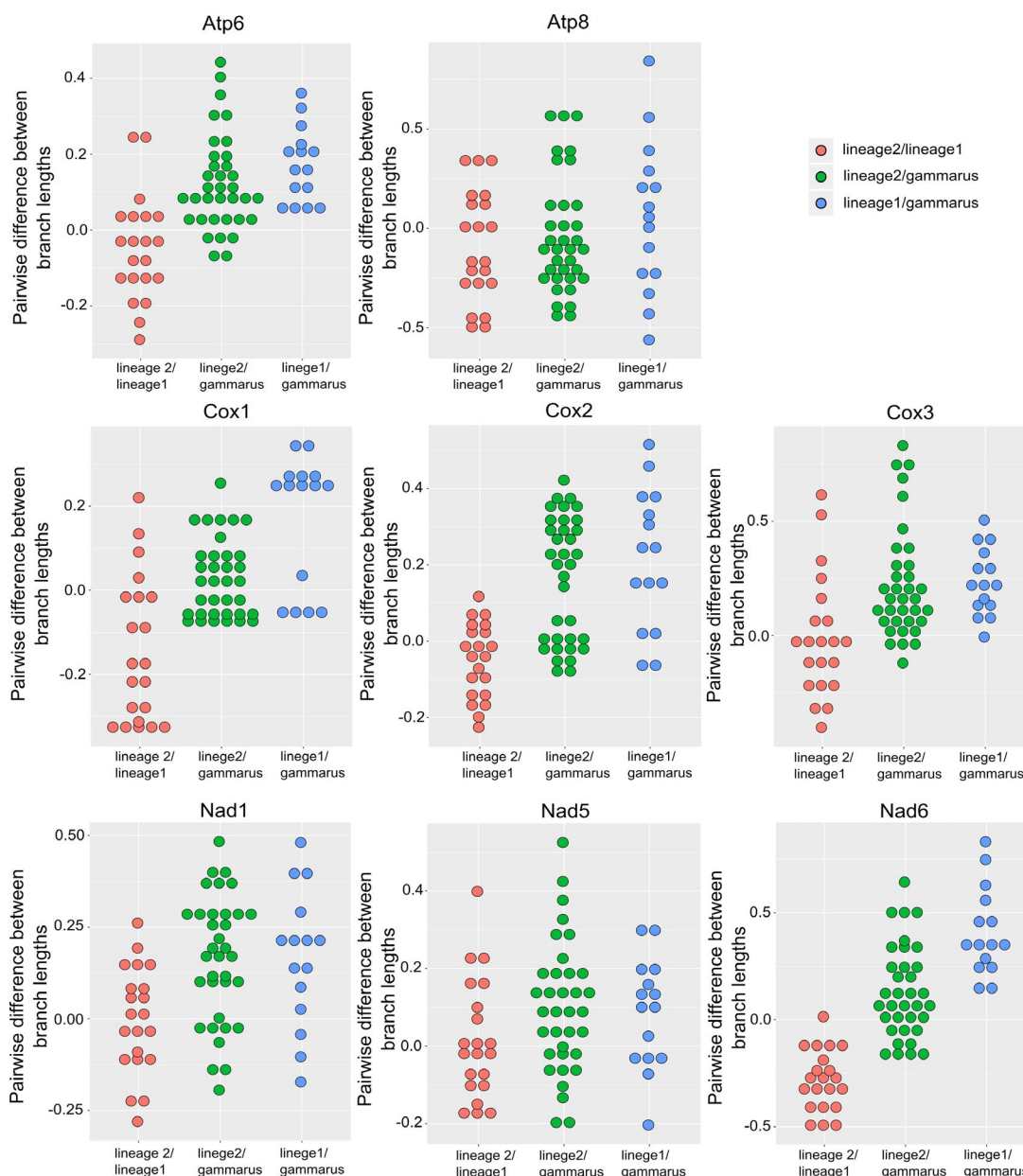


Fig. 3. Pairwise differences between branch lengths in different mitochondrial genes of amphipod groups of interest (Baikalian lineage 1, Baikalian lineage 2, gen. *Gammarus* group).

the Pearson rank correlation was estimated from the data to give a p -value of the hypothesis that it differs from 0 ($p > 0.6$). This favors the second hypothesis explaining the acceleration in molecular evolution by similar adaptive challenges in the parallel Baikalian amphipod lineages.

Our finding corroborates with the results by Naumenko et al. (2017), who detected episodes of positive selection in nuclear genes encoded components of electron transfer chain and ATP synthesis in Baikalian amphipods. It is also worth noticing that all currently available species of gen. *Gammarus* have the same conserved gene order in their mitogenomes, whereas the majority of Baikalian species in both lineages possess different types of gene rearrangements and additional tRNA gene copies (Romanova et al., 2016; 2020). Our analysis supports an opinion about the linkage of high evolutionary rate and gene rearrangements in mitogenomes (Shao et al., 2003; Fourdrilis et al., 2018).

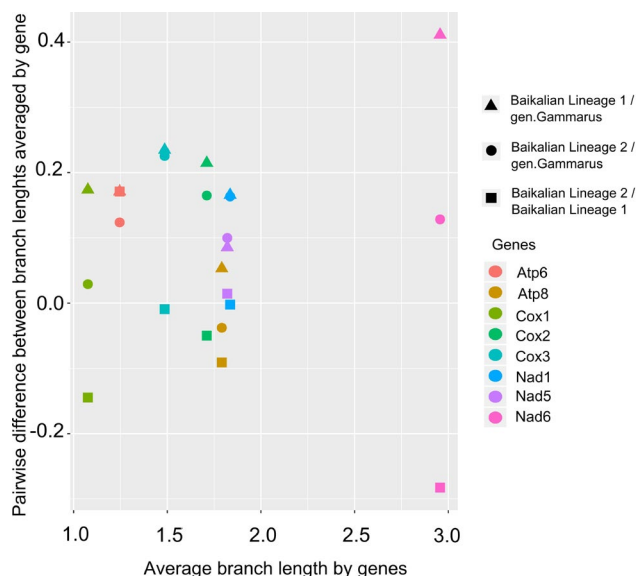
Further studies of higher numbers of mitogenomes in combination with life history data and niche details of the respective species are required to elucidate the acceleration of molecular evolution in Baikalian endemic amphipods. Wider taxonomic sampling is required to test if the acceleration effect is general for Lake Baikal.

Acknowledgements

The study was supported by the governmentally funded project 0345–2019–0004 (AAAA - A16 - 116122110060 - 9).

References

- Bromham L., Cowman P.F., Lanfear R. 2013. Parasitic plants have increased rates of molecular evolution across all three genomes. *BMC Evolutionary Biology* 13. DOI: 10.1186/1471-2148-13-126
- Cock P.A., Antao T., Chang J.T. et al. 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25: 1422-1423. DOI: 10.1093/bioinformatics/btp163
- Cormier A., Wattier R., Teixeira M. et al. 2018. The complete mitochondrial genome of *Gammarus roeselii* (Crustacea, Amphipoda): insights into mitogenome plasticity and evolution. *Hydrobiologia* 825: 197-210. DOI: 10.1007/s10750-018-3578-z
- Ellegren H., Galtier N. 2016. Determinants of genetic diversity. *Nature Reviews Genetics* 17. DOI: 10.1038/nrg.2016.58
- Fazalova V., Nevado B., Peretolchina T. et al. 2010. When environmental changes do not cause geographic separation of fauna: differential responses of Baikalian invertebrates. *BMC Evolutionary Biology* 10: 988-999. DOI: 10.1016/j.ympbev.2018.07.002
- Fourdrilis S., de Frias Martins A.M., Backeljau T. 2018. Relation between mitochondrial DNA hyperdiversity, mutation rate and mitochondrial genome evolution in *Melarhaphé neritoides* (Gastropoda: Littorinidae) and other Caenogastropoda. *Scientific Reports* 8: 1-12. DOI: 10.1038/s41598-018-36428-7
- Gouy M., Guindon S., Gascuel O. 2010. SeaView version 4: a multiplatform graphical user interface for sequence



Naumenko S.A., Logacheva M.D., Popova N.V. et al. 2017. Transcriptome-based phylogeny of endemic Lake Baikal amphipod species flock: fast speciation accompanied by frequent episodes of positive selection. *Molecular Ecology* 26: 536-553. DOI: 10.1111/mec.13927

Nguyen L.T., Schmidt H.A., von Haeseler A. et al. 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32: 268-274. DOI: 10.1093/molbev/msu300

Pinkster S. 1983. The value of morphological characters in taxonomy of *Gammarus*. *Beaufortia* 33: 15-28.

Rabosky D.L. 2019. Phylogenies and diversification rates: variance cannot be ignored. *Systematic Biology* 68: 538-550. DOI: 10.1093/sysbio/syy079

Romanova E.V., Aleoshin V.V., Kamaltynov R.M. et al. 2016. Evolution of mitochondrial genomes in Baikalian amphipods. *BMC Genomics* 17. DOI: 10.1186/s12864-016-3357-z

Romanova E.V., Bukin Y.S., Mikhailov K.V. et al. 2020. Hidden cases of tRNA gene duplication and remodeling in mitochondrial genomes of amphipods. *Molecular Phylogenetics and Evolution* 144. DOI: 10.1016/j.ympev.2019.106710

Shao R., Dowton M., Murrell A. et al. 2003. Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. *Molecular Biology and Evolution* 20: 1612-1619. DOI: 10.1093/molbev/msg176

Sun S., Wu Y., Ge X. et al. 2020. Disentangling the interplay of positive and negative selection forces that shaped mitochondrial genomes of *Gammarus pisinnus* and *Gammarus lacustris*. *Royal Society Open Science* 7. DOI: 10.1098/rsos.190669

Wiens J.J., Scholl J.P. 2019. Diversification rates, clade ages, and macroevolutionary methods. *Proceedings of the National Academy of Sciences* 116: 24400-24400. DOI: 10.1073/pnas.1915908116