

The architecture of the large mitochondrial genome of a pelagic amphipod *Macrohectopus branickii*, and mitochondrial genome length variability in animals

Romanova E.V.^{1*}, Bukin Y.S.¹, Mikhailov K.V.^{2,3}, Logacheva M.D.^{2,4}, Aleoshin V.V.^{2,3}, Sherbakov D.Y.^{1,5}

¹Limnological Institute, SB RAS, Irkutsk, Russia

²Belozersky Institute for Physicochemical Biology, Lomonosov Moscow State University, Moscow, Russia

³A.A. Kharkevich Institute for Information Transmission Problems, RAS, Moscow, Russia

⁴Skolkovo Institute of Science and Technology, Moscow, Russia

⁵Novosibirsk State University, Novosibirsk, Russia

* elena_romanova@lin.irk.ru

Key words: long mitochondrial genomes, non-coding regions, animal phyla, amphipods, Lake Baikal

Motivation and Aim: Mitochondrial genomes in animals have various lengths and architecture. Investigation of the different characteristics of mitochondrial genomes may help to understand the mechanisms of mitochondrial genome functioning and maintenance and their transformation in course of evolution. Recent studies of mitochondrial genomes in amphipods (Crustacean) from ancient Lake Baikal showed relatively high variance in length and gene content and order [1]. In this study, we sequenced the mitochondrial genome of a pelagic Baikalian amphipod *Macrohectopus branickii* [2]. We additionally investigated the variability patterns of mitochondrial genome length in amphipods and other animal taxa and assumed common mechanisms of their lengthening.

Methods and Algorithms: The mitochondrial genome of *M. branickii* was assembled from 41 million paired reads generated by Illumina HiSeq 4000 system from a DNA sample of a single specimen. The regions of long repeated sequences were additionally verified by PCR and Sanger sequencing. The statistical analyses of mitochondrial genome length characteristics were made using appropriate sequence data from the RefSeq database (released before 1 January 2020). We assessed the contribution of the coding and non-coding regions to the entire mitochondrial genome length using regression analysis. ANOVA was used to test the dependence of mitochondrial genome lengths on belonging to animal phyla. Additionally, we examined the repeat pattern in the lengthy mitochondrial genomes of different animal phyla.

Results: The complete mitochondrial genome of *M. branickii* has 42,256 bp in total. It has a unique gene order within amphipods, duplications of the four tRNA genes and a part of Cox2 gene, and a long non-coding region that makes up about two-thirds of the genome's size. Most of the non-coding region consists of inverted repeat sequences of different lengths. Statistical analysis showed no dependence of genome length characteristics on the number of sequences in animal phyla (sample size), but there is a significant dependence of genome lengths on the phylum itself.

Conclusion: *M. branickii* was shown to have the longest mitochondrial genome in amphipods so far and one of the longest within animals. We have shown that genome length distributions and the ratio of the longest mitochondrial genomes (upward outliers)

depend on the phylum and are not affected by sampling bias. The diversity in mitochondrial genome length in animal mitochondrial genomes is defined mainly by the lengths of their non-coding regions. A rare pattern of inverted repeats in the mitochondrial genome of *M. branickii* implies a more complex mechanism of the sequence lengthening than mere duplication and loss. The mitochondrial genome of *M. branickii* is a rare intermediate stage of mitochondrial genome evolution bearing signs of multiple duplications, that has not yet been deleted by selective pressure forces. *Acknowledgements:* This work was supported by the Ministry of Science and Higher Education of Russian Federation; project 0279-2021-0010 (121032300196-8).

References

1. Romanova E.V. et al. Hidden cases of tRNA gene duplication and remodeling in mitochondrial genomes of amphipods. *Mol Phylogenet Evol.* 2020;144:106710.
2. Romanova E.V. et al. The mitochondrial genome of a freshwater pelagic amphipod *Macrohectopus branickii* is among the longest in Metazoa. *Genes.* 2021;12(12):2030.