

Genome survey and phylogeny of Nematomorpha

**Boris Efeykin^{1,2}, Anastasia Teterina^{1,3}, Kirill Mikhaylov^{2,4}, Olga Popova,
Vladimir Aleshin^{2,4}, Yuri Panchin^{2,4}, Sergei Spiridonov³**

¹ Severtsov Institute of Ecology and Evolution, Moscow 119071, Russian Federation; ² Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow 127994, Russian Federation; ³ University of Oregon, 1585 E 13th Ave, Eugene, OR 9740, USA; ⁴ Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Leninskiye Gory 1-40, Moscow 119991, Russian Federation

e-mail: bocha19@yandex.ru

We study Nematomorpha, or hairworms, – a group of worms parasitize on arthropods. Despite the fact that currently, this group includes about 300 species, the ecology and population structure of these species are practically unknown. We sequenced genomes of four hairy worms: *Gordionus alpestris* from Adygea, *Gordionus wolterstorffii* from the Moscow region, *Gordius* sp. from Primorsky Krai, and *Chordodes* sp. from Sumatra.

We discovered numerous perfect inverted repeats of considerable length (up to 284 bp) embedded within the protein-coding genes in mitochondrial genomes of four Nematomorpha species. Strikingly, both arms of the inverted repeats encode conserved regions of the amino acid sequence. We confirmed enzymatic activity of the respiratory complex I encoded by inverted repeat containing genes. The nucleotide composition of inverted repeats suggests strong selection at the amino acid level in these regions. We conclude that the inverted repeat-containing genes are transcribed and translated into functional proteins. The survey of available mitochondrial genomes reveals that several other organisms possess similar albeit shorter embedded repeats. Mitochondrial genomes of Nematomorpha demonstrate an extraordinary evolutionary compromise where protein function and stringent secondary structure elements within the coding regions are preserved simultaneously.

Additionally, we generated nucleotide sequences of two nuclear (18S and 28S rDNA) and one mitochondrial (COI mtDNA) markers for 62 other individuals of Nematomorpha, which represent 9 genera. These markers were used to reconstruct phylogenetic trees and to identify the species by molecular taxonomy methods (ABGD, GMYC, and PACO). The phylogenetic tree was obtained for the phyletic relationships within Nematomorpha and with other phyla of invertebrates. Demographic history of the population helps to understand the evolution and current status of the species. The individuals assigned to the same species were used in the further population analysis. Whole-genomic data were used to explore population histories via coalescent-HMM in SMC. We reconstructed population histories of *G. alpestris*, *G. wolterstorffii*, *Gordius* sp. and *Chordodes* sp., and showed that their effective population sizes are, approximately, 30000, 50000, 70000, and 110000 individuals; and that their populations have been stable for thousands of generations. We suppose that a lower effective population size may be explained by the greater isolation of these populations. At the moment, we carry out the reconstruction of the evolution-

ary histories of these nematodes by whole-genome analysis to associate these results with the lifestyle and habitat of each parasite.

The study was supported by RFBR (research project 18-34-00984).