ISSN 0006-2979, Biochemistry (Moscow), 2025, Vol. 90, No. 11, pp. 1723-1740 © Pleiades Publishing, Ltd., 2025.
Russian Text © The Author(s), 2025, published in Biokhimiya, 2025, Vol. 90, No. 11, pp. 1843-1861.

Gene Order in Mitochondrial DNA Affects Abundance of their Transcripts (A Case of Marine Nematodes)

Olga V. Nikolaeva[1], Anna S. Ovcharenko[1,2], Polina V. Khorkhordina[1,2], Tatyana S. Miroliubova[1,3], Nataliya S. Sadovskaya[1], Victoria A. Scobeyeva[1,2], Nadya P. Sanamyan[4], Elena G. Panina[5], Kirill V. Mikhailov[1,6], Leonid Yu. Rusin[6], Alexei V. Tchesunov[1,2], and Vladimir V. Aleoshin[1,2,6,a]*

[1]*Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, 119234 Moscow, Russia*

[2]*Faculty of Biology, Lomonosov Moscow State University, 119234 Moscow, Russia*

[3]*Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, 119071 Moscow, Russia*

[4]*Kamchatka Branch of the Pacific Institute of Geography, Far Eastern Branch of the Russian Academy of Sciences, 683000 Petropavlovsk-Kamchatsky, Russia*

[5]*Zoological Institute, Russian Academy of Sciences, 199034 Saint Petersburg, Russia*

[6]*Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, 127051 Moscow, Russia*

[a]*e-mail: Aleshin@genebee.msu.su*

Received July 10, 2025
Revised November 13, 2025
Accepted November 14, 2025

Abstract—Mitochondrial genomes of most animals contain the same set of genes, with all or many protein-coding genes (PCGs) arranged in the same order, forming conserved blocks termed syntenies. Some syntenies have been preserved for hundreds of millions of years and are found in both vertebrates and invertebrates. This evolutionary conservation indicates functional role for PCG arrangement; however, biochemical and/or physiological mechanisms by which the gene order in mtDNA affects viability are unknown. Among animals, there are taxa that have completely lost conserved syntenies in mtDNA. Canonical animal syntenies in mtDNA have not been reported in nematodes, until some were recently discovered in the previously unstudied nematode taxa, including the marine family Thoracostomopsidae (Nematoda, Enoplida). We sequenced the complete mitochondrial genomes of three thoracostomopsid species, determined gene order, and their expression levels from the RNA-seq data for all available family representatives. We found that six species of the Thoracostomopsidae there are three distinct patterns of PCG arrangement, and the relative mRNA levels correlate with the gene order rather than species phylogeny. We hypothesize that the influence of PCG translocations on their expression levels underlies the long-term preservation of mitochondrial syntenies among animals.

DOI: 10.1134/S0006297925602114

Keywords: mtDNA, mitogenome, transcriptome, genome sequencing, RNA-seq, molecular evolution, phylogeny, nematodes, Enoplia, Thoracostomopsidae, *Enoplolaimus*, *Marimermis*, *Thoracostomopsis*, *Trileptium*

INTRODUCTION

Mitochondrial genome (mitogenome) of most multicellular animals is a single circular DNA molecule of about 15 kbp [1], which contains 37 genes encoding seven subunits of respiratory chain complex I, one subunit of complex III, three subunits of complex IV, two subunits of ATP synthase (13 mitochondrial protein-coding genes, mtPCGs), two ribosomal RNAs (rRNAs), and 22 transfer RNAs (tRNAs). Typically, there are almost no gaps between the genes, except for one extended (1-1.5 kbp) non-coding region.

* To whom correspondence should be addressed.

Both strands of animal mitochondrial DNA (mtDNA) can encode genes, but the strands often differ in the gene number and purine/pyrimidine ratio. Based on nucleotide composition, the strands are designated as heavy (H) and light (L). Regulatory elements controlling mtDNA replication and transcription are usually located in the non-coding region. In humans and other mammals, this region contains three strand-specific transcription promoters: two for the L-strand and one for the H-strand [2-4]. In other vertebrates, such as *Xenopus laevis* and chickens, the non-coding region contains only one bidirectional promoter [5]. In lancelets, the extended non-coding region is absent, and regulatory elements are apparently scattered throughout the mtDNA [6]. In invertebrates, other transcription promoter organization patterns are observed: the nematode *Caenorhabditis elegans* has one transcription initiation site on the H-strand and none on the L-strand [7], since all genes are located on the H-strand [8]; in *Drosophila melanogaster*, more than five transcription initiation sites are found throughout the mtDNA [5]. In any case, there are fewer promoters in mtDNA than genes, and transcripts are polycistronic.

During maturation of polycistronic transcripts, each tRNA is released due to hydrolysis at the 5′- and 3′-ends by RNase P and RNase Z, respectively [9], which also leads to the release of monocistronic mRNAs and rRNAs [10, 11]. The origination of mature mRNAs from a few or even a single polycistronic transcript would seem to imply an equimolar ratio of all mRNAs. However, in reality, the mRNA pool for complex I proteins is usually lower than for complex IV, and mRNA levels for the proteins that form complexes in an equimolar ratio may differ substantially [12]. It can be hypothesized that deviation from the expected equimolar ratio is explained by random factors, for example, due to different mRNA stability. However, randomness cannot explain the similar imbalance in different species. A similar ratio of mRNAs between the same mitochondrial genes is characteristic of both vertebrates and invertebrates [13, 14], although small variations are observed even within one species depending on sex, age, and physiological state [12, 15, 16]. In some species, the ratio of transcripts changes dramatically compared to the typical animal pattern; these species also show more or less significant differences in the order of genes in mtDNA [15, 17].

Considering that processing of a long polycistronic precursor into a monocistronic mRNAs occurs before translation, the order of genes initially seems insignificant. However, contrary to expectations, the order of gene arrangement in the mitogenomes of many animal taxa, such as vertebrates, is highly conserved [18]. In other taxa, it is more variable, as in mollusks [19], bryozoans [20], mites [21], tunicates [22], thrips [17, 23], and others. However, both stable and variable mitogenomes usually contain conserved gene blocks (syntenies). Four conserved gene blocks were identified in the mtDNAs of Bilateria, within which composition and arrangement of genes remain constant [24]. The oldest syntenies [24] are preserved in vertebrates and invertebrates, i.e., since the Cambrian period (over 550 million years). Such conservation indicates functional importance of the gene order. It could be assumed that a certain order of genes is coordinated with regulation of their transcription and, ultimately, with the ratio of mitochondrial proteins. However, for example, seven mtDNA-encoded subunits of complex I that are incorporated in the complex in an equimolar ratio, are not grouped into a single conserved block but are split among four different blocks [24]. Three *cox1–cox3* genes encoding complex IV proteins, *cox1–cox3*, are also located in two different blocks. Conservation of the order of mtPCGs in mtDNA appears to be a genuine paradox considering variability of the tRNA gene positions. Transfer RNA genes are located between the mtPCGs and rRNA genes, where they are necessary as markers for processing [10], but most of them are not assigned to conserved blocks and change their location in the mitochondrial genome much more frequently than the mtPCGs and rRNA genes [25]. Thus, selection exerts significantly greater pressure on the location of mtPCGs than on the location of tRNA genes.

The mystery of preservation of the gene order in mtDNA is compounded by the existence of species that have completely lost conserved gene blocks, apparently without visible consequences – functional constraints that have operated for hundreds of millions of years seem to have been relaxed in these species. Paradoxically, individual taxa that “abandoned” the Cambrian legacy in the form of conserved blocks, have acquired specific gene orders that in turn remain preserved fixed it over a long evolutionary times [26-28]. The only known factor to date associated with the changes in the genetic map of animal mtDNA is an increased rate of molecular evolution [1, 24]. Until recently, roundworms (nematodes) could be considered as one of such taxa. More than 200 complete mitogenomes of soil and parasitic nematodes are known; almost all of them are similar to each other in the gene order, but without a single synteny in common with the other animals [8, 29]. Almost all of them belong to the giant, largest in the phylum Nematoda, but isolated terrestrial evolutionary branch – Rhabditia.

Taxonomic representation of the nematode mtDNA has begun to change in recent years due to developments in the genomic sequencing of microscopic organisms and publication of the first mitogenomes of the uncultured free-living marine nematodes [30].

In these mitogenomes, some of the traditional bilaterian gene blocks are found, and for two species of the family Thoracostomopsidae, similar but slightly different genetic maps are shown [30], making these species a promising model for searching for biochemical and physiological consequences of the changes in gene order in mtDNA. The free-living Thoracostomopsidae inhabit silty and sandy bottom sediments in the intertidal, sublittoral, and bathyal zones throughout the World Ocean [31, 32]. Currently, the family includes 3 subfamilies (Thoracostomopsinae, Trileptiinae, and Enoplolaiminae), 16 genera, and 193 species [33]. Based on the structure of the buccal cavity and gut content, the thoracostomopsids are classified as predators: they feed on other nematodes, cyanobacteria, diatoms, rotifers, oligochaetes, turbellarians, and copepods [31]. According to the recent molecular data, parasites of marine invertebrates of the genus *Marimermis* are included in the family Thoracostomopsidae [32, 34, 35].

MATERIALS AND METHODS

Materials. The study included a sample of six nematode species from all three subfamilies of the Thoracostomopsidae family. Specimens of *Enoplolaimus vulgaris* de Man, 1893 (~100 individuals) and *Thoracostomopsis* (*Th.*) *barbata* Ditlevsen, 1918 (~10 individuals) were collected from the littoral and sublittoral zones of the Kandalaksha Bay of the White Sea in August 2023. A specimen of *Marimermis maritima* Rubtzov & Platonova, 1974 was extracted from the coelomic cavity of the sea urchin *Strongylocentrotus polyacanthus* caught near the Matua Island (Kuril Islands, Pacific Ocean) in August 2016. Before the DNA and RNA extraction, the *M. maritima* specimen was stored in 95% ethanol. Data on *Enoplolaimus lenunculus* Wieser, 1959 (SRR24201686), *Trileptium* (*T.*) *ribeirensis* Vilas-Boas, Silva, Alves, Castro & Pinheiro-Junior, 2016 (SRR24201700) [30], and an unidentified Thoracostomopsidae Gen. sp. KK-2019 isolate (SRR8943408) [36] were obtained from the NCBI SRA database.

Sequencing, assembly, and annotation of mitochondrial genomes. Genomic DNA libraries were prepared using an Accel-NGS WGA kit (Swift Biosciences, Inc., USA). RNA extraction from *E. vulgaris* was performed using a RNAqueous-Micro Total RNA Isolation Kit (Thermo Fisher Scientific, USA), from *Th. barbata* using an Arcturus PicoPure Isolation Kit (Thermo Fisher Scientific), and from *M. maritima* using an innuPREP DNA/RNA Mini Kit (Analytik Jena, Germany). Complementary DNA (cDNA) library preparation was performed using an oligo-dT primer and the SMART-Seq v4 Ultra Low Input RNA Kit (Takara Bio, Japan) with 12 (*E. vulgaris* and *M. maritima*) or 15 (*Th. barbata*) PCR cycles. Sequencing of the three thoracostomopsid species was performed using an Illumina NovaSeq 6000 (*E. vulgaris*, *M. maritima*) or Illumina NextSeq (*Th. barbata*) platforms. For *E. vulgaris*, 47.3 million (DNA) and 42.2 million (cDNA) paired-end reads were obtained; for *Th. barbata* we obtained 34.4 million (DNA) and 25.4 million (cDNA) paired-end reads; for *M. maritima* we obtained 29.6 million (DNA) and 36.3 million (cDNA) paired-end reads. Quality of the obtained reads, as well as of the reads from NCBI SRA, was assessed using FastQC [37]. Adapters and incorrect terminal nucleotides were removed using Trimmomatic [38]. Assembly was performed using SPAdes [39] with k-mer lengths of 55, 77, 99, and 123. Target contig search was performed locally using the BLAST package [40]. Mitochondrial genomes were assembled using NOVOPlasty [41]. Mitogenome annotation was performed using MITOS [42]; proteins were predicted according to the invertebrate mitochondrial genetic code. Read mapping to mitogenomes was performed using Bowtie2 [43]. Visualization of mapping and search for post-transcriptional modifications were performed using Tablet [44].

Comparative analysis of mitochondrial genome expression. Read coverage of mitogenomes was assessed using BEDtools [45] based on alignments produced with Bowtie2. After obtaining coverage data for each nucleotide, coverage plots were constructed in MS Excel. To study differential expression, we used total coverage of each gene by reads normalized by length (TPM; transcripts per million) calculated in RSEM [46] for trimmed reads. The *atp8* and *rrnL* genes were not used due to the very low and very high coverage, respectively. Based on the TPM data, principal component analysis was performed using STATISTICA10 (TIBCO Software Inc., USA); Spearman and Kendall rank correlation analysis was performed in the R environment.

Phylogenetic analysis. To determine direction of evolution of the gene order, phylogenetic relationships of the six studied species were examined. For this purpose, nuclear 18S and 28S rRNA genes were obtained from the genomic assemblies, including *E. lenunculus*, *T. ribeirensis*, and the KK-2019 isolate, since only for these markers there is a representative sample of Thoracostomopsidae in the NCBI GenBank database. Alignments were performed using MAFFT [47] with manual correction in BioEdit [48] and subsequent concatenation. For the reconstruction with mitochondrial proteins, we concatenated the alignments of 12 proteins featuring the six Thoracostomopsidae representatives and *Enoplus communis*, which was used as an outgroup. Phylogenetic trees were constructed using IQ-Tree 3.0.1 [49] and MrBayes 3.2.6 [50]. Nucleotide evolution model was selected according

Table 1. Nucleotide composition of thoracostomopsid mitochondrial DNA

Species	GC, %	A, %	T, %	G, %	C, %	AT-skew	GC-skew
Enoplolaimus vulgaris	28	30	42	14	14	−0.17	0
Thoracostomopsis barbata	17	35	48	9	8	−0.15	0.06
Trileptium ribeirensis	20	34	46	10	10	−0.15	0
Enoplolaimus lenunculus	21	33	46	10	11	−0.16	−0.05
Isolate KK-2019*	20	38	42	11	9	−0.05	0.1
Marimermis maritima	35	30	35	17	18	−0.08	−0.03

Note. * Incomplete genome.

to the Bayesian information criterion (BIC) [51] in IQ-Tree: for 18S rRNA, the GTR + F + R3 model; for 28S rRNA, the TVMe + I + G4 model; for mitochondrial proteins, Mtrev + G4. For the Bayesian inference with MrBayes, the following parameters were used: nst = 6, ngammacat = 6, ngen = 3,000,000, burnin = 50% (for nucleotide sequences) and aamodelpr = fixed(mtrev), ngammacat = 4, ngen = 200,000, burnin = 50% (for amino acid sequences). Trees were visualized using MEGA 7 [52].

RESULTS

Structure of thoracostomopsid mitochondrial genomes. The mitochondrial genome of *E. vulgaris* is organized into a single circular DNA molecule of 13,843 bp, containing 37 genes (13 mtPCGs, 2 rRNA genes, and 22 tRNA genes). All genes are located on one strand. Four types of start codons are predicted for mtPCGs: ATA (10 genes), ATG (1 gene), ATC (1 gene), and TTA (1 gene). Eleven mtPCGs terminate with the stop codon TAA, and 2 with TAG.

The mitochondrial genome of *Th. barbata* is organized into a single circular DNA molecule of 14,399 bp, containing the same set of 37 genes commonly present in animals. Thirty-five genes are located on one strand, while *trnY* and *trnD* are located on the other strand. Based on the alignment of translated mtPCGs, four types of start codons are predicted: ATA (6 genes), ATG (3 genes), ATT (3 genes), and GTT (1 gene). Twelve mtPCGs terminate with the stop codon TAA, and 1 gene with TAG.

The mitochondrial genome of *M. maritima* is organized into a single circular DNA molecule of 14,284 bp, containing 37 genes. All genes are located on one strand. Three types of start codons are predicted for mtPCGs: ATA (6 genes), ATG (4 genes), and ATT (3 genes). Eight mtPCGs terminate with the stop codon TAA, and 5 genes with TAG.

The mitochondrial genome of the KK-2019 isolate was assembled into an unclosed contig of 15,589 bp, containing 29 genes, of which 13 are mtPCGs. Eight tRNA genes were not found. All identified genes are located on one strand. Five types of start codons are predicted for mtPCGs: ATA (8 genes), GTA (1 gene), TTA (1 gene), ATT (1 gene), and TGG (1 gene). Start of the *nad5* gene was not found. Twelve mtPCGs terminate with the stop codon TAA, and 1 with TAG.

Nucleotide composition (GC) and AT- and GC-skew in the mtDNA of the six thoracostomopsid species are presented in Table 1.

Nucleotide composition of the *Th. barbata* mtDNA, with GC content of 17%, approaches the lowest value for animals. The GC content of *M. maritima* is 35%. A twofold difference in the GC content among the species of the same family is quite unusual. In all species, thymidine predominates over adenosine in the coding strand of mtDNA, while cytosine and guanine are present in almost equal amounts.

Order of mitochondrial genes. Among the six thoracostomopsid species, three variants of mtPCG and rRNA gene order are observed (Fig. 1). The differences between *E. vulgaris*, *Th. barbata*, and *T. ribeirensis* affect only tRNA genes. The same applies to *E. lenunculatus* and the KK-2019 isolate, representing another variant of mtPCG arrangement. The gene order in *M. maritima* is unique.

In the mitochondrial genomes of thoracostomopsids, conserved gene blocks typical for Bilateria [24] are partially preserved, but they are unusual for nematodes. In four species, the block 2 (*nad4l–nad4–nad5*) is completely preserved, and in *E. lenunculatus* and the KK-2019 isolate, two of the three genes from this block are adjacent. Although the block 1 is fragmented, all six species retain the canonical pair *atp6–cox3* from this block. The block 3 is fragmented in all species, but adjacency of two of its three genes is preserved (*rrnL–nad1* in *E. lenunculatus* and in the KK-2019 isolate, *nd1–rrnS* in the others).

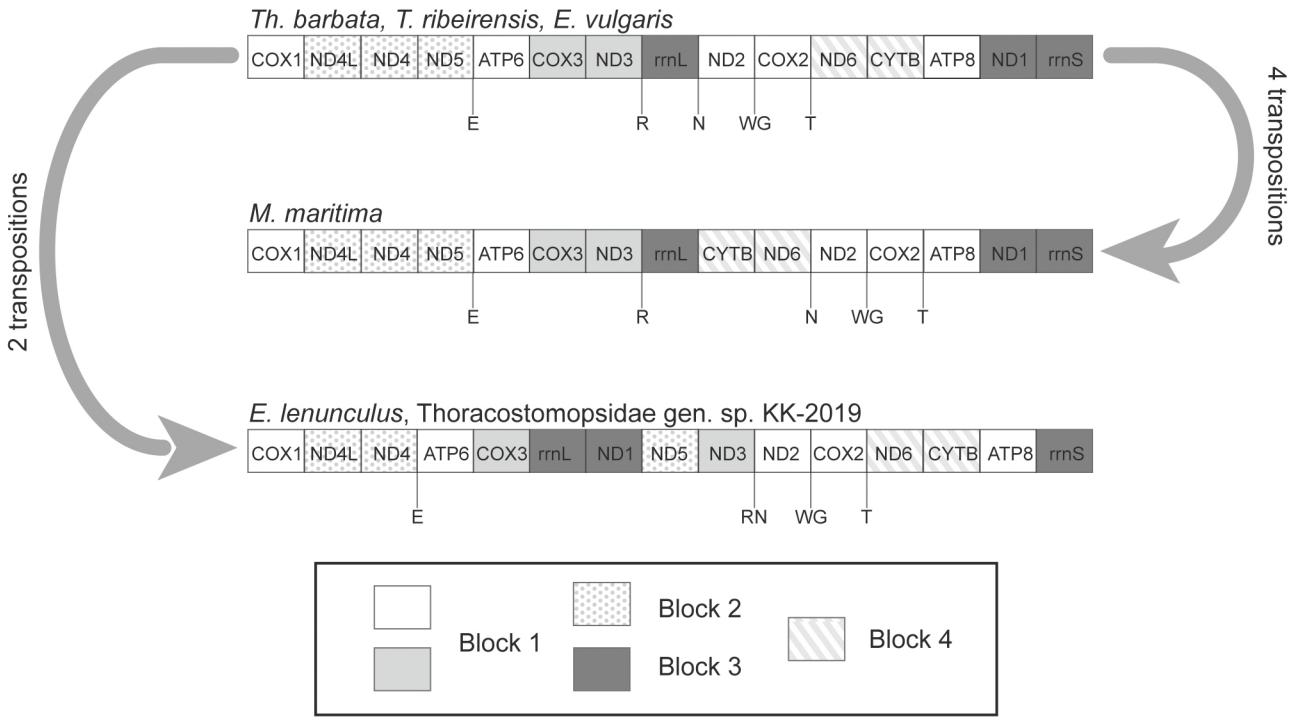

Fig. 1. Evolution of the order of mitochondrial genes in nematodes of the family Thoracostomopsidae. Genes are marked according to the conserved blocks typical for Bilateria [24]. tRNA genes that maintain a neighboring position relative to at least one of the mtPCGs in different species are shown by a single-letter code.

The block 4 (*nad6–cytb*) is preserved in five species; in *M. maritima*, the gene order is changed to *cytb–nad6*.

In all six species, two tRNA genes, *trnW* and *trnG*, are located between *nad3* and *cox2*; another four tRNA genes retain a neighbor at the 5′- or 3′-end (Fig. 1). Position of the remaining tRNA genes is more variable and is specific for each species.

In all six species, the non-coding region is located near the beginning of the *nad1* gene, although its position relative to the neighboring tRNA genes and the *atp8* gene differs. In *E. lenunculus* and the KK-2019 isolate, the non-coding region, together with the *nad1* gene, has moved to a completely different place on the genetic map.

Polarization of gene order evolution. To understand how changes in the order of mtDNA genes affect their expression, it is desirable to know direction of these changes, i.e., which of the three variants observed in modern Thoracostomopsidae is ancestral. It would be imprudent to *a priori* consider the most common variant as ancestral, since it could have arisen in one of the evolutionary lines better represented in the database.

We reconstructed phylogeny of the Thoracostomopsidae family based on the nuclear 18S and 28S rRNA genes, for which the largest amount of representatives are available. Two reconstruction methods, Bayesian and maximum likelihood, led to the similar topologies with the same arrangement of the six species of interest (Fig. 2).

Species with the same genetic map (*E. vulgaris*, *Th. barbata*, and *T. ribeirensis*) do not form a monophyletic group relative to the species with a different map (Fig. 2). One of them, *T. ribeirensis*, together with *Enoploides* species, represents an early branch of Thoracostomopsidae, which is consistent with the results of previous studies on phylogeny of this family [32, 35, 53, 54]. The simplest scenario of evolution (minimum number of gene permutations) is obtained if we assume that the gene order of *T. ribeirensis*, *E. vulgaris*, and *Th. barbata* was inherited from the common ancestor of the family. Let us call this order plesiomorphic. The order of mtPCGs of *M. maritima* and *E. lenunculus* differs from the plesiomorphic one and was acquired by them independently (Fig. 1). Let us call these maps apomorphic. The order of mtPCGs in the KK-2019 isolate and *E. lenunculus* coincide, which is consistent with the relatedness of these species (Fig. 2). Thus, the arrows in Fig. 1 show not only the method of formal transformation of the map, but actual direction of evolution of the gene order.

Mapping of DNA and RNA reads to mitochondrial genomes. Coverage of the mitochondrial genomes by DNA sequencing (DNA-seq) reads fluctuates within a certain range without clear attachment to the gene boundaries (Fig. 3). However, in *Th. barbata* and *M. maritima*, the decrease followed by a sharp

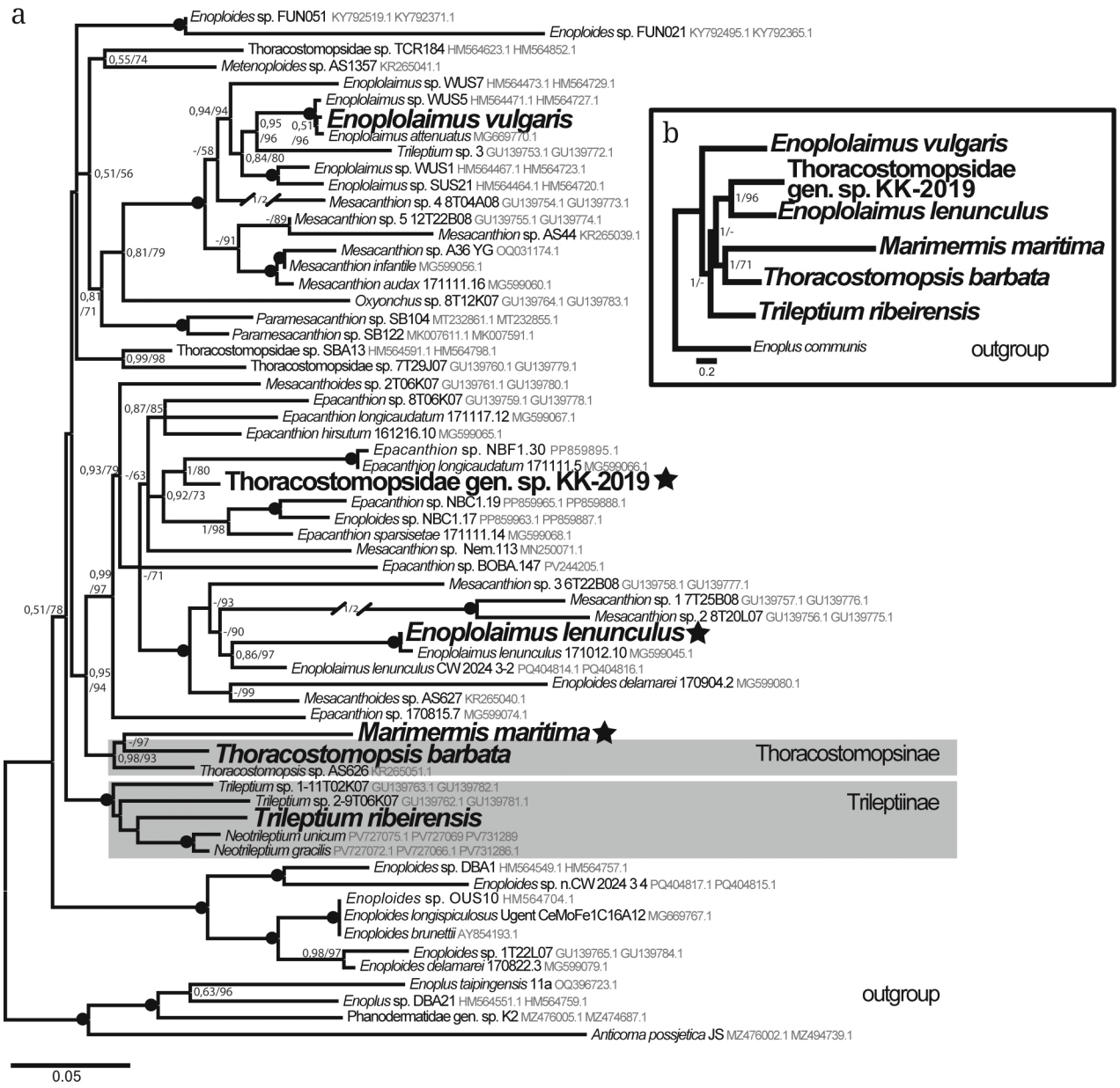

Fig. 2. Phylogenetic tree of the Thoracostomopsidae family based on concatenated 18S and 28S rRNA genes (a) and mitochondrial proteins (b). Numbers at the nodes correspond to Bayesian posterior probability values (first number) and bootstrap support (parametric bootstrap) for the maximum likelihood tree (second number). Values of 1/100 are replaced by ●. Supports below 0.50 or 50 are indicated by a dash; values −/− are omitted. Species for which the mitogenome was sequenced are highlighted in bold; ⋆ species with different order of mtPCGs; gray background highlights representatives of the subfamilies Thoracostomopsinae and Trileptiinae.

increase in the coverage is observed in the non-coding region. For the non-coding region of *Th. barbata*, the increase in coverage could be associated with short tandem repeats in this area. In other species, decrease in the coverage is observed in the non-coding region.

Proportion of the RNA sequencing (RNA-seq) reads mapping to the mitochondrial genome is 0.13% for *Th. barbata*, 22.9% for *M. maritima*, and 32% for *E. vulgaris*. Coverage of the mitochondrial genomes by the RNA-seq reads, unlike DNA-seq, is highly heterogeneous both within one gene and at the level of different genes (Fig. 4). The *rrnL* gene has the highest coverage in all species, while coverage of the *rrnS* gene is significantly lower and comparable with the coverage of mtPCGs. It is expected that rRNA reads should primarily originate from molecules making up mature ribosomes, and, therefore, should be present in the cell in an equimolar ratio. However, "deficiency" of the 12S rRNA has been previously detected in different species, including when cDNA sequencing

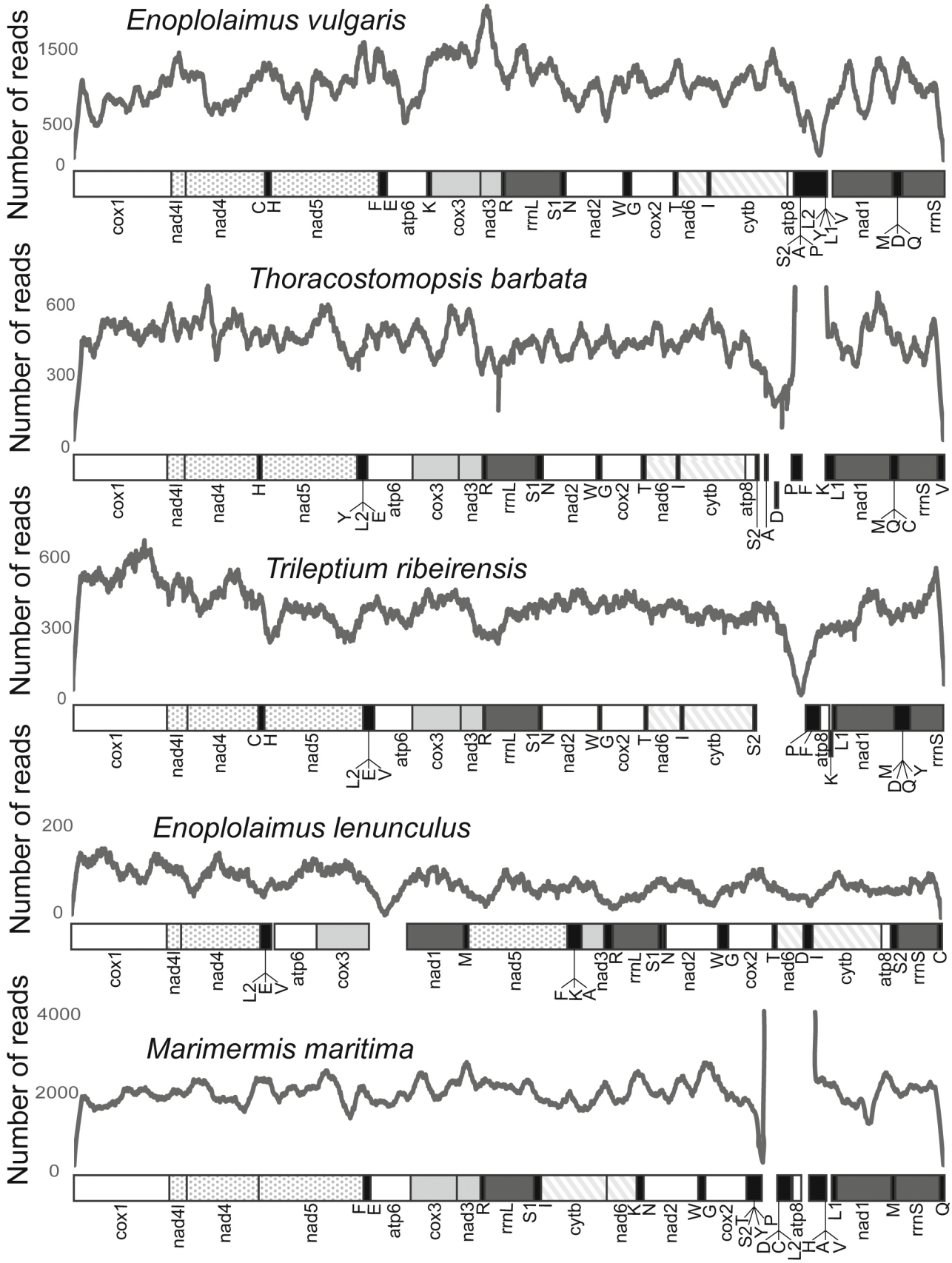

Fig. 3. Coverage of Thoracostomopsidae mitochondrial genomes by the DNA-seq reads. tRNA genes are shown in single-letter code.

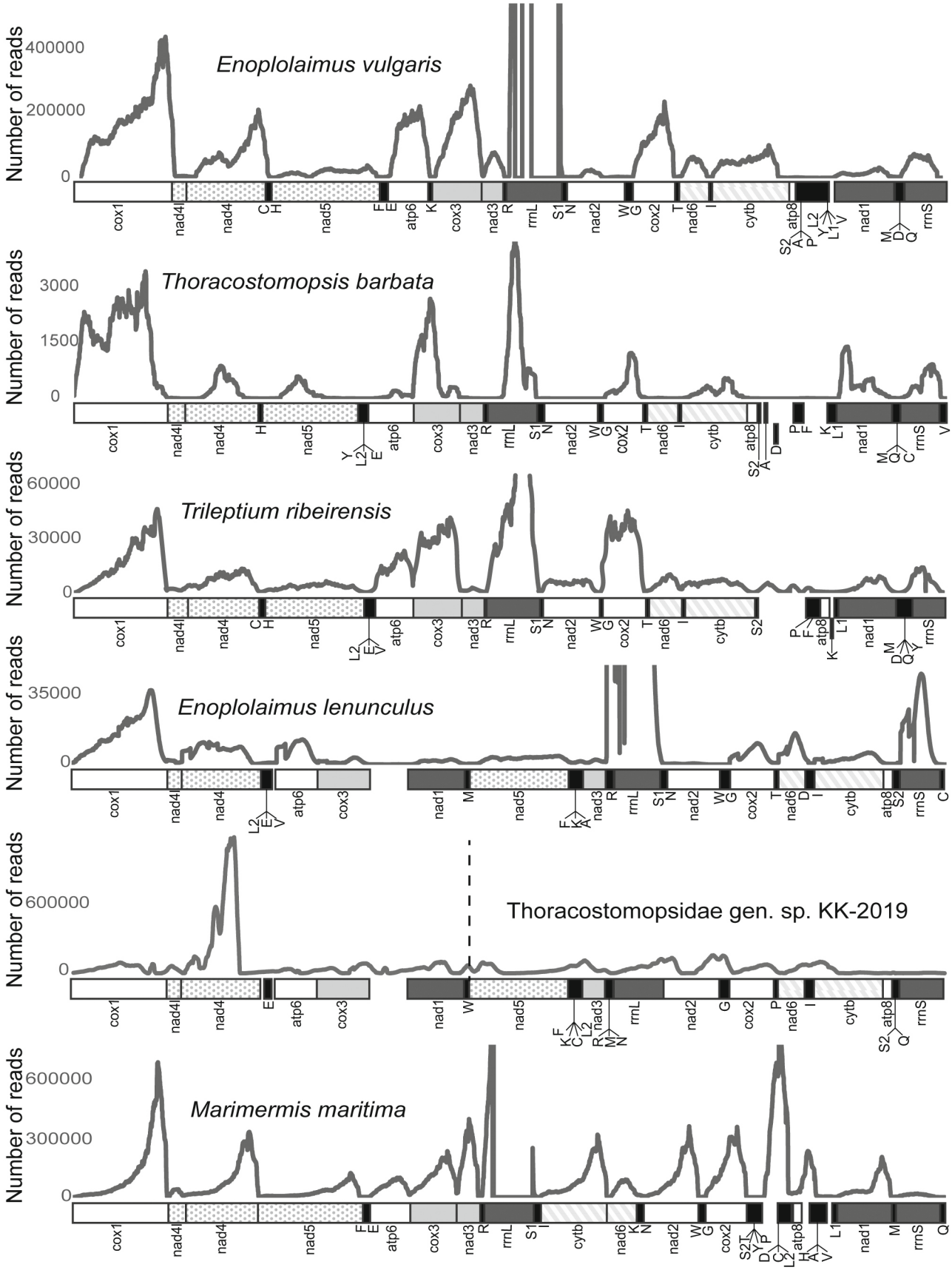

Fig. 4. Coverage of mitochondrial genomes of Thoracostomopsidae by RNA-seq reads. Dashed line indicates a break point in the KK-2019 isolate assembly.

Table 2. Proportions of polyadenylated mitochondrial transcripts (as % of coverage at the modified nucleotide) in the six species of Thoracostomopsidae

Gene	*E. vulgaris*	*Th. barbata*	*T. ribeirensis*	*E. lenunculus*	KK-2019	*M. maritima*
atp6	15.8	5.9	1.7	2.7	–	–
cox1	26.9	33.0	0.5	7.4	–	0.2
cox2	2.7	45.8	–	–	1.8	1.3
cox3	10.7	–	33.5	44.7	–	0.2
cytb	–	–	39.0	19.5	–	0.2
nd1	–	–	49.2	6.7	–	0.1
nd2	–	–	40.1	–	–	1.4
nd3, site 1	–	–	–	4.0	–	–
nd3, site 2	–	–	19.4	4.0	–	–
nd4	28.2	–	23.9	2.4	–	1.2
nd4l	–	–	1.6	0.1	–	0.2
nd5	3.0	–	41.7	6.4	6.8	1.7
nd6	1.2	–	5.4	–	0.5	–
rrnS	–	–	–	–	4.8	–
rrnL	–	27.3	–	–	–	–
trnD	–	–	–	5.7	–	–
trnL2	–	–	–	–	–	0.3
trnR	–	–	–	–	2.5	–
trnS1	–	–	1.0	–	–	–
trnV	–	–	–	–	–	1.1
trnY	–	–	0.3	–	–	–

was performed without enrichment of poly-A transcripts with long-fragment sequencing and using random primers [55, 56]; so, it is unlikely that in the case of thoracostomopsids it is explained by the library preparation artifacts. The lowest coverage is observed in the regions with tRNA genes, which is consistent with the tRNA punctuation model [10]. Among the mtPCGs, the highest level of coverage is observed for *cox1*, *cox2*, and *cox3*, and the lowest for *atp8*. An exception is the KK-2019 isolate with a relatively low level of *cox1* mRNA and high level of *nad4*. Coverage noticeably decreases towards the beginning of each gene. This is explained by more active degradation of mRNAs from the 5′-end, as well as enrichment of libraries with 3′-ends of sequences due to the cDNA preparation method that employs an oligo-dT primer. The RNA of *Th. barbata* and *M. maritima*, extracted from the specimens stored in ethanol for several years, were more degraded.

Mapping of cDNA reads revealed from 4 (*Th. barbata*) to 13 (*T. ribeirensis*) polyadenylation sites, totaling 51 sites (Table 2). mRNA of every protein is polyadenylated, except for *atp8* (modifications of which may not have been detected due to low coverage) in at least two thoracostomopsid species. Thus, the dicistronic mature mRNAs, characteristic of animal mitochondrial transcriptomes, are apparently not present in thoracostomopsids. In *E. lenunculus*, in addition to polyadenylation of the 3′-ends of mRNAs, an additional polyadenylation site was found

Table 3. TPM values for mtPCGs of six Thoracostomopsidae species (rounded to whole numbers)

Gene	*Th. barbata*	*E. lenunculus*	*T. ribeirensis*	*E. vulgaris*	KK-2019	*M. maritima*
cox1	370,422	209,913	120,679	181,041	52,565	68,263
cox2	120,095	50,387	281,626	153,501	48,014	111,296
cox3	251,191	14,364	233,143	178,235	78,702	72,122
cytb	41,575	39,715	41,248	44,589	75,505	52,037
atp6	20,383	132,559	113,650	165,097	186,906	40,601
nad1	111,753	20,458	30,490	12,505	105,815	35,397
nad2	106	2,478	35,902	6,270	28,354	70,957
nad3	3,147	48,833	17,181	100,595	81,445	306,290
nad4	51,011	82,396	43,730	66,985	104,962	53,063
nad4l	4,120	113,462	16,825	15,171	47,107	128,139
nad5	24,456	35,682	20,528	13,527	68,424	14,362
nad6	1,741	249,753	44,998	62,484	122,201	47,473
Sum	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000

in the coding frame of the *nad3* gene. In the individual libraries, polyadenylated rRNAs and tRNAs are present. A total of six polyadenylated tRNAs were found: trnD, trnL2, trnR, trnS1, trnV, and trnY. In *T. ribeirensis*, one modified dicistronic transcript *rrnL–trnS1* was found: some cDNA reads capture poly-A at the 3′-end of *trnS1*, and reverse reads map to the 3′-end of the neighboring *rrnL* gene. The reads produced by the Illumina platform are not long enough to properly detect unprocessed polycistronic precursors of mature RNAs (pre-mRNAs) [55], nevertheless, in another species, the horsehair worm *Parachordodes pustulosus*, similar methods revealed three dicistronic products in the presence of 23 polyadenylation sites [57]. Apparently, maturation of pre-mRNAs differs among distant species.

Polyadenylation sites in the species with the same order of mtPCGs do not coincide. This difference is unlikely to stem from the lack of coverage: 11 polyadenylation sites in *M. maritima* are confirmed by a total of 8 million reads mapped to the mtDNA, while 13 sites in *T. ribeirensis* were found using ~1 million mapped reads. For *Th. barbata*, however, underestimation of polyadenylation sites is likely due to the low read depth of the cDNA library (0.03 million reads mapped to mtDNA).

Proportion of the modified site to the total coverage at of the corresponding nucleotide in the cDNA library is low for genes in *M. maritima*. In other species, this proportion varies from less than 1% to more than 40% within a single library (Table 2). There is no direct correlation between the number of transcripts of a particular gene and the degree of polyadenylation of their ends. We consider these two circumstances as important indirect evidence of suitability of the obtained libraries for transcriptome comparison. In the cDNA libraries constructed using oligo-dT primers, enrichment of polyadenylated transcripts can be expected. In the obtained libraries, the ratio of different cDNAs does not correlate with the degree of their polyadenylation, therefore, this ratio reflects contribution of factors related to the construction methodology and the actual ratio of transcripts *in vivo*.

No other modifications, such as non-templated addition of CCA to the acceptor stem of tRNA, characteristic of animal mitochondria, were detected in thoracostomopsids.

Differential expression of mitochondrial genes. An accepted measure of differential expression assessment is the TPM value. We obtained TPM values for mtPCGs (Table 3) and performed principal component analysis (Fig. 5).

Species with plesiomorphic genetic map – *E. vulgaris*, *Th. barbata*, and *T. ribeirensis* – contribute the most to the first principal component and form a distinct cluster in the projection onto the plane of the first two principal components. The first principal

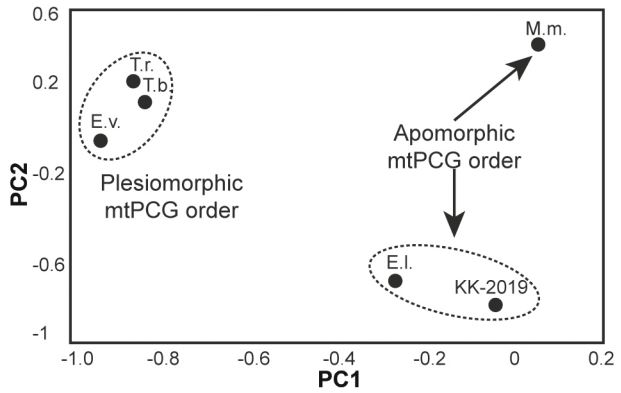

Fig. 5. Principal component analysis of normalized cDNA quantity (TPM) of 12 mtPCGs of six nematode species of Thoracostomopsidae (three variants of the genetic map). Abbreviations: E.v., *Enoplolaimus vulgaris*; T.r., *Trileptium ribeirensis*; T.b., *Thoracostomopsis barbata*; E.l., *Enoplolaimus lenunculus*; KK-2019, Thoracostomopsidae gen. sp., isolate KK-2019; M.m., *Marimermis maritima*.

component explains 39% of total variability in the expression of mitochondrial genes for the species included in the analysis, i.e., we can speak of a similar pattern of mtPCG mRNAs in the species with the plesiomorphic genetic map. Importantly, the species with the plesiomorphic genetic map are not the closest relatives (Fig. 2). The species *E. lenunculus* and the KK-2019 isolate, which have another variant of the genetic map, also form their own cluster on the plane of the first two principal components. The species *M. maritima*, with a unique gene order, has a pattern of expression that differs from all others, including its closest relative *Th. barbata*. Thus, we observe similarity in the patterns of expression of mitochondrial genes in the species with similar genetic maps that transcends their phylogenetic relationships.

As an alternative method of transcriptome comparison, we applied rank correlation analysis, where we ranked mtPCGs according to TPM in each of the transcriptomes, and calculated pairwise (by species) Spearman (r_s) and Kendall (τ) rank correlation coefficients (Table 4).

In both types of analyses, Spearman and Kendall, correlations in the ranks are statistically significant at $p < 0.05$ for one pair of species out of three with the plesiomorphic order of mtPCGs. When calculating significance, we considered it unnecessary to apply a correction for multiple comparisons (Bonferroni or similar) for 15 pairwise comparisons. With a sample size that small, a significant proportion of the results could be false negatives, so we focused not on significance of the correlation, but on the value of its coefficient, expressing the strength of the relationship between the quantities. Between any species with the plesiomorphic gene order, the correlation is higher or at least not lower than in the species with the apomorphic gene order. Given reproducibility of the results obtained by different methods, we cautiously propose the existence of correlation in the expression of mtPCGs in the species with the same gene order, regardless of their degree of relatedness.

DISCUSSION

Distortions in the natural ratio of various RNAs in a generated cDNA library can be attributed to several factors – primarily, selective RNA degradation, selective cDNA synthesis, and context-dependent sequencing effects. The underlying causes are predominantly associated with the inherent molecular properties of RNAs, including secondary structure features, the presence of specific nuclease target motifs or motifs that are difficult for reverse transcriptase or sequencing polymerase to pass through, and the presence and length of the polyadenylated region. The findings

Table 4. Spearman rank correlation coefficients (r_s) and Kendall rank correlation coefficients (τ) for mtPCGs ranked by their TPM values

Species	*E. vulgaris*	*Th. barbata*	*T. ribeirensis*	*E. lenunculus*	KK-2019	*M. maritima*
E. vulgaris	**1**	0.545	0.706*	0.462	0.224	0.294
Th. barbata	0.424	**1**	0.573	−0.028	−0.021	−0.042
T. ribeirensis	0.485*	0.394	**1**	0.217	0.133	−0.007
E. lenunculus	0.394	−0.061	0.121	**1**	0.301	−0.014
KK-2019	0.152	0	0.121	0.212	**1**	−0.497
M. maritima	0.182	−0.03	0.091	0	−0.364	**1**

Note. Above the diagonal with bold values are r_s values, below the diagonal are τ values. Gray cells refer to the species with plesiomorphic arrangement of mtPCGs. * Statistically significant correlation at $p < 0.05$.

of factor and rank analyses, as presented in the paper, lend support to the hypothesis that the relative level of cDNA reads is contingent on the relative position of genes in mtDNA. It is challenging to conceive of a scenario in which selective RNA degradation, selective cDNA synthesis, or context-dependent sequencing effects could result in the observed dependence. This evidence was obtained in a system of fairly closely related species, which renders this and similar systems promising for detailed study of the mitogenome variability and its manifestation through differential expression of RNAs in the mitochondrion. A key issue of fundamental and applied significance is to elucidate the mechanisms of selective regulation of genes that are transcribed in a single round towards a common polycistronic pre-mRNA. The correlation between gene order (i.e., configuration of genes on a genetic map) and their expression levels suggests an indirect relationship, whereby the relocation of one gene affects expression of another, indicating a multifaceted regulatory mechanism, potentially operating at the level of transcription or post-transcriptional modifications (i.e. maturation, or processing) of the polycistron.

Transcriptional regulation. The lifecycle of mitochondrial RNAs unfolds within the giant, microscopically discernible complexes, known as mitochondrial RNA granules [4, 58, 59], where the coordinated processes of transcription, RNA maturation, and mitochondrial ribosome assembly are orchestrated. According to some estimates, up to 300 nucleus-encoded mitochondrial proteins are involved in ensuring the cycle [2], and transcription of the mitogenome is regulated in a complex manner by several transcription factors with chromatin-modulating activity, which have multiple binding sites along the length of mtDNA and change binding activity during ontogeny [5, 60-62]. The binding sites of certain factors and helicases are associated with G-quadruplexes, non-canonical 3D structures that primarily form on the heavy strand of mtDNA and exhibit a non-random, conservative distribution throughout the genome [63]. The involvement of these structures in the regulation of transcription and replication of mtDNA, chromatin remodeling, formation of mutation and deletion breakpoints [64-66] has been demonstrated, indicating their role in gene regulation at a higher-order level of mtDNA organization [67]. The activity of mitochondrial RNA polymerase is modulated by components of this system, impacting the start, speed, and termination of transcription. It stands to reason that alterations in the order of genes bearing conserved regulatory regions could modify the regulatory context of the mitochondrial chromosome, potentially leading to impaired or even arrested transcription of a part of the polycistron, followed by downregulation of downstream genes. Meanwhile, elevated expression levels are anticipated at start of the polycistron, with its first gene expected to exhibit reduced processing dependency.

Post-transcriptional regulation. The maturation of polycistronic pre-mRNA relies on the specific endonuclease-mediated excision of tRNA genes to generate processed fragments with free 3′-ends, which normally undergo polyadenylation by the mitochondrial poly(A) polymerase MTPAP [68]. This highlights the role of polyadenylation and tRNA positioning on the genetic map as related and potentially important factors in the regulation of mitochondrial genes. Polyadenylation is known to contribute to transcript stabilization at several levels: directly by protecting the coding part of the 3′-end from degradation by exonucleases [68] and indirectly via involvement of poly(A)-tails in complex processes of translation regulation (breakdown of aberrant mRNA–mitoribosome complexes [69]; mRNA handoff to the ribosome for translation initiation [70]), routing of processed mRNAs [71], and overall folding and stabilization of the mitochondrial transcriptome by RNA-binding chaperone proteins [72]. A variation in polyadenylation sites during gene rearrangement might exert pleiotropic effects, including gene degradation at 3′-end of the common transcript in a fraction of molecules. Notably, the function of poly(A)-tails could be species-, tissue-, and transcript-specific, depending on their presence or length in different biological contexts, suggesting their role in various, primarily undefined, regulatory mechanisms. For instance, the knockdown of MTPAP in human cell cultures compromises the stability of mRNAs in the *cox1*, *cox2*, *cox3*, and *atp6* genes, while it does not affect *nad3* and even increases the level of *nad1* [73, 74]. Polyadenylation occurs in normal mitochondria for some tRNAs and rRNAs [73]. In fruit flies, the 12S rRNA carries a short poly(A)-tail [68], suggesting its involvement in the regulation of this transcript, although its primary function remains poorly understood and may differ from that in protein-coding genes.

It is important to acknowledge that the expression estimates, as presented in this study, primarily reflect the processed genes with functional polyadenylation sites due to poly(A) enrichment-based library construction. Upstream genes (i.e., those without free 3′-ends) may appear underrepresented, and the more so the greater is the gene's distance from the polycistronic 3′-end, whereas the content of such genes depends on their order along the polycistron and its length.

The orientation of the punctuation tRNA genes plays its own pivotal role in the maturation of the common transcript and, consequently, in achieving functional polyadenylation: the transcription strand must be sense for both tRNA and its flanking genes.

Should the genetic map rearrangement cause altered tRNAs orientation, their primary structures will not fold into hairpins required for recognition and excision by endonucleases [73, 74], the upstream genes will not form free 3′-ends and will not become polyadenylated, which will ultimately disrupt their regulation (as well as affect representation in a poly(A)-library).

It is also of note that the composition of transcripts (revealed in RNA-seq data) does not accurately reflect the composition of their encoded proteins, since translation efficiency also depends on the gene order. A free and polyadenylated 3′-end may be required to complete a functional stop codon from the incomplete T or TA codons in the mitogenome [74] (all stop codons in thoracostomopsids are complete); the ribosome binding efficiency at 5′-untranslated regions of neighboring genes in the polycistronic transcript may require proper sequence overlap with an upstream gene [75, 76], which can change upon gene relocation.

Given the intricacy of the mechanisms that link the mitochondrial gene order and their expression, we believe that chromosomal rearrangements are associated with high lethality risks and require simultaneous compensatory changes in the molecular systems involved. This renders them rare events in which discrete changes are fixed in mitochondrial evolution for over millions of years, i.e., at the level of higher taxa [26-28]. Early mitogenome rearrangements are of particular interest for research into the factors of stability of the mitochondrial system. The first evidence in animal species at a relatively small evolutionary distance is presented in this work.

Taxonomic and morphological conclusions. The obtained cladogram (Fig. 2) allows us to draw some conclusions about the morphological evolution and taxonomy of Thoracostomopsidae.

In the cladogram, two species assigned to the same genus *Enoplolaimus*, *E. vulgaris* and *E. lenunculus*, turned out to be on different branches of the thoracostomopsid tree. The type species of the genus *E. vulgaris* groups with its closest relative *Enoplolaimus attenuatus*, several unidentified *Enoplolaimus* species, and *Mesacanthion*, as well as with *Trileptium* sp. 3 (the latter appears to be a misidentification). *E. lenunculus* (isolates from two studies [30, 35] and unpublished MG599045) and the unidentified thoracostomopsid KK-2019 [36] are combined into a large group with large number of *Epacanthion* species; this group also includes several species of *Mesacanthion*, *Mesacanthoides*, and *Enoploides*. There may be two possible explanations for the grouping of species of different genera together: (1) morphological features of the genus *Enoplolaimus* are not apomorphies, and the genus *Enoplolaimus* itself is not monophyletic; (2) taxonomic identifications of thoracostomopsid species in the GenBank database are partially erroneous. Let us consider both possible options.

(1) According to the latest review of the Thoracostomopsidae family [33], *Enoplolaimus* is characterized by the position of the anterior ten setae at the level of the posterior edge of the head capsule. In *Mesacanthion* and in *Epacanthion*, the anterior setae are located at the level of the anterior edge or the middle of the head capsule. *Enoploides* differs from *Enoplolaimus* and *Mesacanthion* in the shape of mandibles. In *Epacanthion*, the mandibles are of intermediate structure. Relationship between the slight displacement of sensilla or changes in the shape of mandibles, although clear from the point of view of morphological taxonomy, has not been tested for the ability to reflect phylogeny.

(2) The probability of erroneous taxonomic identification always exists, and it could increase significantly if the species is identified from the specimens from an area remote from the type locality. In our case, this is *Enoplolaimus lenunculus*, described from Puget Sound (USA) and then collected for molecular genetic research in China, on the opposite shore of the Pacific Ocean. For the specimens not identified to the species level, and there are many such thoracostomopsids in the GenBank database, accuracy of the genus determination could also be questioned.

Based on the published illustrations (Figs. 6-8 in the article by Meng et al. [35]), arrangement of setae in the Chinese isolate of *E. lenunculus* corresponds to the diagnosis of the genus *Enoplolaimus*. A conflict arises with its distance on the phylogenetic tree from the type species of the genus, *E. vulgaris*. Either *E. lenunculus* should be transferred from the genus *Enoplolaimus* to preserve monophyly of the latter, but then morphological diagnosis of the genus would have to be revised, or the scope of the genus *Enoplolaimus* would have to be expanded to almost all thoracostomopsids, but then a revision of the significance of morphological features would also be required. Regardless of the problems of taxonomy, *E. lenunculus* and the KK-2019 isolate have synapomorphies in the arrangement of genes in mtDNA and belong to one clade (Fig. 2), which did not prevent the KK-2019 isolate from acquiring unique differences in the pattern of mtPCG expression (low level of *cox1* and high level of *nad4*). In the absence of molecular data on the type species of the genus *Trileptium*, we are inclined to consider the genus *Neotrileptium* [32] as a synonym of *Trileptium*. According to the latest review [33], the genus *Metenoploides* is recognized as invalid, but this opinion probably needs to be revised, since the AS1357 isolate, assigned to the genus *Metenoploides* [53], is located on the phylogenetic tree separately from most species of the genus *Enoploides* (Fig. 2a).

According to the phylogenetic reconstruction, division of Thoracostomopsidae into three subfamilies has not been observed. Two of them, Thoracostomopsinae and Trileptiinae, turn out to be branches within Enoplolaiminae. The first two subfamilies are characterized by clear morphological features (fusion of mandibles into a spear and sharp reduction of the stoma, respectively), which could well have arisen as an evolutionary specialization in two phylogenetic branches of Enoplolaiminae. No apomorphies at the morphological level have been proposed for Enoplolaiminae, which is consistent with their paraphyly revealed at the genetic level.

The phylogenetic reconstruction based on 12 mitochondrial proteins (Fig. 2b), as well as on the rRNA genes [32, 34, 35], confirms belonging of the species *M. maritima* to the family Thoracostomopsidae, previously referred to the order Marimermithida.

CONCLUSION

Nematodes of the family Thoracostomopsidae preserve some of the syntenies in the mitochondrial genetic maps commonly seen in other animal phyla, in contrast to the nematodes of the subclass Rhabditia, which have lost these syntenies entirely. Three variants of the genetic map, which involve rearrangements of protein-coding and rRNA genes, were discovered in six species of Thoracostomopsidae. One of the variants is recognized as plesiomorphic, as the species possessing the variant are paraphyletic relative to the species with the alternative gene orders. The two alternative variants arose in evolution independently of each other. Factor analysis and rank analysis based on the normalized gene coverage by cDNA reads (TPM) revealed positive correlations between the expression levels of mitochondrial protein-coding genes in the species with the same gene arrangement, regardless of their relatedness. This indicates the existence of a yet unknown mechanism by which the gene order in mtDNA influences the transcription or mRNA maturation, and ultimately the Darwinian fitness. Phylogenetic analysis confirms the inclusion of marimermithid *M. maritima* in Thoracostomopsidae but refutes the traditional delineation of subfamilies within Thoracostomopsidae.

Abbreviations

mtDNA	mitochondrial DNA
mtPCG	mitochondrial protein-coding genes
pre-mRNA	polycistronic precursor of mature RNAs
TPM	transcripts per million
rRNA	ribosomal RNA
tRNA	transfer RNA
RNA-seq	RNA sequencing
DNA-seq	DNA sequencing
cDNA	complementary DNA

Acknowledgments

The authors are grateful to Yi-Chien Lee for providing mtDNA assemblies of *E. lenunculus* and *T. ribeirensis*; to V. Yu. Shtratnikova, M. A. Kulbachnaya, D. A. Knore, and an anonymous reviewer for constructive criticism of the manuscript; to S. N. Lysenkov for assistance in statistical analysis. Computer calculations were performed using equipment purchased with the funds provided by the state assignment of Lomonosov Moscow State University; animal collection was carried out under state assignment no. 122031100275-4.

Contributions

O. V. Nikolaeva, A. S. Ovcharenko, N. S. Sadovskaya, K. V. Mikhailov, and V. V. Aleoshin – conducted computational experiments; T. S. Miroliubova – conducted laboratory work; N. S. Sadovskaya, V. A. Skobeeva, and A. S. Ovcharenko – conducted statistical analysis; N. P. Sanamyan, E. G. Panina, L. Yu. Rusin, and A. V. Tchesunov – conducted field work; P. V. Khorkhordina, L. Yu. Rusin, A. V. Tchesunov, and V. V. Aleoshin – conducted phylogenetic analysis; all authors – prepared the text and illustrations.

Funding

Field work and material collection was carried out with support of the Zoological Institute of Russian Academy of Sciences (“Taxonomy, biodiversity and ecology of invertebrates of the Russian and adjacent waters of the World Ocean, continental ponds and wetlands”, no. 122031100275-4). Analysis of the structure of mitochondrial genomes and transcriptomes was carried out with financial support of the Russian Science Foundation (grant no. 19-74-20147). Phylogenetic analysis was carried out with financial support of the Russian Science Foundation (grant no. 25-74-20009).

Ethics approval and consent to participate

This work does not contain any studies involving human and animal subjects that are regulated by ethics committee.

Conflict of interest

The authors of this work declare that they have no conflicts of interest.

REFERENCES

1. Lavrov, D. V., and Pett, W. (2016) Animal mitochondrial DNA as we do not know it: mt-genome organization and evolution in nonbilaterian lineages, *Genome*

Biol. Evol., **8**, 2896-2913, https://doi.org/10.1093/gbe/evw195.

2. Pearce, S. F., Rebelo-Guiomar, P., and D'Souza, A. R. (2017) Regulation of mammalian mitochondrial gene expression: recent advances, *Trends Biochem. Sci.*, **42**, 625-639, https://doi.org/10.1016/j.tibs.2017.02.003.
3. Tan, B. G., Mutti, C. D., Shi, Y., Xie, X., Zhu, X., Silva-Pinheiro, P., Menger, K. E., Díaz-Maldonado, H., Wei, W., Nicholls, T. J., Chinnery, P. F., Minczuk, M., Falkenberg, M., and Gustafsson, C. M. (2022) The human mitochondrial genome contains a second light strand promoter, *Mol. Cell*, **82**, 3646-3660.e9, https://doi.org/10.1016/j.molcel.2022.08.011.
4. Falkenberg, M., Larsson, N. G., and Gustafsson, C. M. (2024) Replication and transcription of human mitochondrial DNA, *Annu. Rev. Biochem.*, **93**, 47-77, https://doi.org/10.1146/annurev-biochem-052621-092014.
5. Barshad, G., Marom, S., Cohen, T., and Mishmar, D. (2018) Mitochondrial DNA transcription and its regulation: an evolutionary perspective, *Trends Genet.*, **34**, 682-692, https://doi.org/10.1016/j.tig.2018.05.009.
6. Boore, J. L., Daehler, L. L., and Brown, W. M. (1999) Complete sequence, gene arrangement, and genetic code of mitochondrial DNA of the cephalochordate *Branchiostoma floridae* (Amphioxus), *Mol. Biol. Evol.*, **16**, 410-418, https://doi.org/10.1093/oxfordjournals.molbev.a026122.
7. Blumberg, A., Rice, E. J., Kundaje, A., Danko, C. G., and Mishmar, D. (2017) Initiation of mtDNA transcription is followed by pausing, and diverges across human cell types and during evolution, *Genome Res.*, **27**, 362-373, https://doi.org/10.1101/gr.209924.116.
8. Okimoto, R., Macfarlane, J. L., Clary, D. O., and Wolstenholme, D. R. (1992) The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*, *Genetics*, **130**, 471-498, https://doi.org/10.1093/genetics/130.3.471.
9. Rossmanith, W. (2012) Of P and Z: mitochondrial tRNA processing enzymes, *Biochim. Biophys. Acta*, **1819**, 1017-1026, https://doi.org/10.1016/j.bbagrm.2011.11.003.
10. Ojala, D., Montoya, J., and Attardi, G. (1981) tRNA punctuation model of RNA processing in human mitochondria, *Nature*, **290**, 470-474, https://doi.org/10.1038/290470a0.
11. Montoya, J., Ojala, D., and Attardi, G. (1981) Distinctive features of the 5′-terminal sequences of the human mitochondrial mRNAs, *Nature*, **290**, 465-470, https://doi.org/10.1038/290465a0.
12. Torres, T. T., Dolezal, M., Schlotterer, C., and Ottenwalder, B. (2009) Expression profiling of *Drosophila* mitochondrial genes via deep mRNA sequencing, *Nucleic Acids Res.*, **37**, 7509-7518, https://doi.org/10.1093/nar/gkp856.
13. Nabholz, B., Ellegren, H., and Wolf, J. B. W. (2012) High levels of gene expression explain the strong evolutionary constraint of mitochondrial protein-coding genes, *Mol. Biol. Evol.*, **30**, 272-284, https://doi.org/10.1093/molbev/mss238.
14. Held, J. P., and Patel, M. R. (2020) Functional conservation of mitochondrial RNA levels despite divergent mtDNA organization, *BMC Res. Notes*, **13**, 334, https://doi.org/10.1186/s13104-020-05177-0.
15. Neira-Oviedo, M., Tsyganov-Bodounov, A. G., Lycett, J., Kokoza, V., Raikhel, A. S., and Krzywinski, J. (2011) The RNA-Seq approach to studying the expression of mosquito mitochondrial genes, *Insect Mol. Biol.*, **20**, 141-152, https://doi.org/10.1111/j.1365-2583.2010.01053.x.
16. Wu, X., Zhan, L., Storey, K. B., Zhang, J., and Yu, D. (2025) Differential mitochondrial genome expression of four skink species under high-temperature stress and selection pressure analyses in Scincidae, *Animals (Basel)*, **15**, 999, https://doi.org/10.3390/ani15070999.
17. Liu, Q., Xu, S., He, J., Cai, W., Wang, X., and Song, F. (2024) Full-length transcriptome profiling of the complete mitochondrial genome of *Sericothrips houjii* (Thysanoptera: Thripidae: Sericothripinae) featuring extensive gene rearrangement and duplicated control regions, *Insects*, **15**, 700, https://doi.org/10.3390/insects15090700.
18. Singh, T. R., Shneor, O., and Huchon, D. (2008) Bird mitochondrial gene order: insight from 3 warbler mitochondrial genomes, *Mol. Biol. Evol.*, **25**, 475-477, https://doi.org/10.1093/molbev/msn003.
19. Sun, S., Li, Q., Kong, L., and Yu, H. (2020) Evolution of mitochondrial gene arrangements in Arcidae (Bivalvia: Arcida) and their phylogenetic implications, *Mol. Phylogenet. Evol.*, **150**, 106879, https://doi.org/10.1016/j.ympev.2020.106879.
20. Kutyumov, V. A., Predeus, A. V., Starunov, V. V., Maltseva, A. L., and Ostrovsky, A. N. (2021) Mitochondrial gene order of the freshwater bryozoan *Cristatella mucedo* retains ancestral lophotrochozoan features, *Mitochondrion*, **59**, 96-104, https://doi.org/10.1016/j.mito.2021.02.003.
21. Wang, T., Zhang, S., Pei, T., Yu, Z., and Liu, J. (2019) Tick mitochondrial genomes: structural characteristics and phylogenetic implications, *Parasit. Vectors*, **12**, 451, https://doi.org/10.1186/s13071-019-3705-3.
22. Griggio, F., Voskoboynik, A., Iannelli, F., Justy, F., Tilak, M. K., Turon, X., Pesole, G., Douzery, E. J., Mastrototaro, F., and Gissi, C. (2014) Ascidian mitogenomics: comparison of evolutionary rates in closely related taxa provides evidence of ongoing speciation events, *Genome Biol. Evol.*, **6**, 591-605, https://doi.org/10.1093/gbe/evu041.
23. Liu, Q., Cai, Y. D., Ma, L., Liu, H., Linghu, T., Guo, S., Wei, S., Song, F., Tian, L., Cai, W., and Li, H. (2023) Relaxed purifying selection pressure drives accelerated and dynamic gene rearrangements in thrips

(Insecta: Thysanoptera) mitochondrial genomes, *Int. J. Biol. Macromol.*, **253**, 126742, https://doi.org/10.1016/j.ijbiomac.2023.126742.

24. Bernt, M., Bleidorn, C., Braband, A., Dambach, J., Donath, A., Fritzsch, G., Golombek, A., Hadrys, H., Jühling, F., Meusemann, K., Middendorf, M., Misof, B., Perseke, M., Podsiadlowski, L., von Reumont, B., Schierwater, B., Schlegel, M., Schrödl, M., Simon, S., Stadler, P. F., Stöger, I., and Struck, T. H. (2013) A comprehensive analysis of bilaterian mitochondrial genomes and phylogeny, *Mol. Phylogenet. Evol.*, **69**, 352-364, https://doi.org/10.1016/j.ympev.2013.05.002.
25. Dowton, M., Cameron, S. L., Dowavic, J. I., Austin, A. D., and Whiting, M. F. (2009) Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral, *Mol. Biol. Evol.*, **26**, 1607-1617, https://doi.org/10.1093/molbev/msp072.
26. Popova, O. V., Mikhailov, K. V., Nikitin, M. A., Logacheva, M. D., Penin, A. A., Muntyan, M. S., Kedrova, O. S., Petrov, N. B., Panchin, Y. V., and Aleoshin, V. V. (2016) Mitochondrial genomes of Kinorhyncha: *trnM* duplication and new gene orders within animals, *PLoS One*, **11**, e0165072, https://doi.org/10.1371/journal.pone.0165072.
27. Weigert, A., Golombek, A., Gerth, M., Schwarz, F., Struck, T. H., and Bleidorn, C. (2016) Evolution of mitochondrial gene order in Annelida, *Mol. Phylogenet. Evol.*, **94**, 196-206, https://doi.org/10.1016/j.ympev.2015.08.008.
28. Kang, H., Li, B., Ma, X., and Xu, Y. (2018) Evolutionary progression of mitochondrial gene rearrangements and phylogenetic relationships in Strigidae (Strigiformes), *Gene*, **674**, 8-14, https://doi.org/10.1016/j.gene.2018.06.066.
29. Kern, E. M. A., Kim, T., and Park, J.-K. (2020) The mitochondrial genome in nematode phylogenetics, *Front. Ecol. Evol.*, **8**, 250, https://doi.org/10.3389/fevo.2020.00250.
30. Lee, Y. C., Ke, H. M., Liu, Y. C., Lee, H. H., Wang, M. C., Tseng, Y. C., Kikuchi, T., and Tsai, I. J. (2023) Single-worm long-read sequencing reveals genome diversity in free-living nematodes, *Nucleic Acids Res.*, **51**, 8035-8047, https://doi.org/10.1093/nar/gkad647.
31. Greenslade, P., and Nicholas, W. L. (1991) Some Thoracostomopsidae (Nematoda: Enoplida) from Australia, including descriptions of two new genera and diagnostic keys, *Invertebr. Syst.*, **4**, 1031-1052, https://doi.org/10.1071/IT9901031.
32. Zograf, J. K., Efimova, K. V., and Mordukhovich, V. (2025) Integrative descriptions of two new Thoracostomopsidae species (Nematoda, Enoplida) with the brief discussion on nematode spicules origin, *Zool. Anz.*, **319**, 50-69, https://doi.org/10.1016/j.jcz.2025.08.012.
33. Souza, J. V., and Maria, T. F. (2023) Taxonomic review of Thoracostomopsidae (Nematoda, Enoplida): state of the art, list of valid species and dichotomous keys, *Zootaxa*, **5361**, 463-496, https://doi.org/10.11646/zootaxa.5361.4.2.
34. Tchesunov, A. V., Nikolaeva, O. V., Rusin, L. Y., Sanamyan, N. P., Panina, E. G., Miljutin, D. M., Gorelysheva, D. I., Pegova, A. N., Khromova, M. R., Mardashova, M. V., Mikhailov, K. V., Yushin, V. V., Petrov, N. B., Lyubetsky, V. A., Nikitin, M. A., and Aleoshin, V. V. (2023) Paraphyly of Marimermithida refines primary routes of transition to parasitism in roundworms, *Zool. J. Linn, Soc.*, **197**, 909-923, https://doi.org/10.1093/zoolinnean/zlac070.
35. Meng, Z., Liang, H., and Wang, C. (2025) Phylogenetic analysis within Monhysteridae and Thoracostomopsidae based on rDNA sequences and two new species from the Yellow Sea, China, *Zoosyst. Evol.*, **101**, 1339-1358, https://doi.org/10.3897/zse.101.154881.
36. Smythe, A. B., Holovachov, O., and Kocot, K. M. (2019) Improved phylogenomic sampling of free-living nematodes enhances resolution of higher-level nematode phylogeny, *BMC Evol. Biol.*, **19**, 121, https://doi.org/10.1186/s12862-019-1444-x.
37. Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data, Available online at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc.
38. Bolger, A. M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for illumina sequence data, *Bioinformatics*, **30**, 2114-2120, https://doi.org/10.1093/bioinformatics/btu170.
39. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, *J. Comput. Biol.*, **19**, 455-477, https://doi.org/10.1089/cmb.2012.0021.
40. Altschul, S. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.*, **25**, 3389-3402, https://doi.org/10.1093/nar/25.17.3389.
41. Dierckxsens, N., Mardulyn, P., and Smits, G. (2017) NOVOPlasty: de novo assembly of organelle genomes from whole genome data, *Nucleic Acids Res.*, **45**, e18, https://doi.org/10.1093/nar/gkw955.
42. Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., and Stadler, P. F. (2013) MITOS: improved de novo metazoan mitochondrial genome annotation, *Mol. Phylogenet. Evol.*, **69**, 313-319, https://doi.org/10.1016/j.ympev.2012.08.023.
43. Langmead, B., and Salzberg, S. L. (2012) Fast gapped-read alignment with bowtie 2, *Nat. Methods*, **9**, 357-359, https://doi.org/10.1038/nmeth.1923.

44. Milne, I., Stephen, G., Bayer, M., Cock, P. J., Pritchard, L., Cardle, L., Shaw, P. D., and Marshall, D. (2013) Using Tablet for visual exploration of second-generation sequencing data, *Brief Bioinform.*, **14**, 193-202, https://doi.org/10.1093/bib/bbs012.
45. Quinlan, A. R., and Hall, I. M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features, *Bioinformatics*, **26**, 841-842, https://doi.org/10.1093/bioinformatics/btq033.
46. Li, B., and Dewey, C. N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome, *BMC Bioinformatics*, **12**, 323, https://doi.org/10.1186/1471-2105-12-323.
47. Katoh, K., and Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability, *Mol. Biol. Evol.*, **30**, 772-780, https://doi.org/10.1093/molbev/mst010.
48. Hall, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, *Nucleic Acids Symp. Ser.*, **41**, 95-98.
49. Nguyen, L. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies, *Mol. Biol. Evol.*, **32**, 268-274, https://doi.org/10.1093/molbev/msu300.
50. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., and Huelsenbeck, J. P. (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space, *Syst. Biol.*, **61**, 539-542, https://doi.org/10.1093/sysbio/sys029.
51. Schwarz, G. (1978) Estimating the dimension of a model, *Ann. Statist.*, **6**, 461-464, https://doi.org/10.1214/aos/1176344136.
52. Kumar, S., Stecher, G., and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets, *Mol. Biol. Evol.*, **33**, 1870-1874, https://doi.org/10.1093/molbev/msw054.
53. Smythe, A. B. (2015) Evolution of feeding structures in the marine nematode order Enoplida, *Integr. Comp. Biol.*, **55**, 228-240, https://doi.org/10.1093/icb/icv043.
54. Jeong, R., Tchesunov, A. V., and Lee, W. (2020) Two species of Thoracostomopsidae (Nematoda: Enoplida) from Jeju Island, South Korea, *PeerJ*, **8**, e9037, https://doi.org/10.7717/peerj.9037.
55. Gao, S., Rena, Y., Suna, Y., Wub, Z., Ruan, J., He, B., Zhang, T., Yu, X., Tian, X., and Bu, W. (2016) PacBio full-length transcriptome profiling of insect mitochondrial gene expression, *RNA Biol.*, **13**, 820-825, https://doi.org/10.1080/15476286.2016.1197481.
56. Araujo, N. S., and Arias, M. C. (2019) Mitochondrial genome characterization of *Melipona bicolor*: Insights from the control region and gene expression data, *Gene*, **705**, 55-59, https://doi.org/10.1016/j.gene.2019.04.042.
57. Nikolaeva, O. V., Beregova, A. M., Efeykin, B. D., Miroliubova, T. S., Zhuravlev, A. Y., Ivantsov, A. Y., Mikhailov, K. V., Spiridonov, S. E., and Aleoshin, V. V. (2023) Expression of hairpin-enriched mitochondrial DNA in two hairworm species (Nematomorpha), *Int. J. Mol. Sci.*, **24**, 11411, https://doi.org/10.3390/ijms241411411.
58. Lee, K. W., Okot-Kotber, C., LaComb, J. F., and Bogenhagen, D. F. (2013) Mitochondrial ribosomal RNA (rRNA) methyltransferase family members are positioned to modify nascent rRNA in foci near the mitochondrial DNA nucleoid, *J. Biol. Chem.*, **288**, 31386-31399, https://doi.org/10.1074/jbc.M113.515692.
59. Jedynak-Slyvka, M., Jabczynska, A., and Szczesny, R. J. (2021) Human mitochondrial RNA processing and modifications: overview, *Int. J. Mol. Sci.*, **22**, 7999, https://doi.org/10.3390/ijms22157999.
60. She, H., Yang, Q., Shepherd, K., Smith, Y., Miller, G., Testa, C., and Mao, Z. (2011) Direct regulation of complex I by mitochondrial MEF2D is disrupted in a mouse model of Parkinson disease and in human patients, *J. Clin. Invest.*, **121**, 930-940, https://doi.org/10.1172/JCI43871.
61. Blumberg, A., Sailaja, B. S., Kundaje, A., Levin, L., Dadon, S., Shmorak, S., Shaulian, E., Meshorer, E., and Mishmar, D. (2014) Transcription factors bind negatively-selected sites within human mtDNA genes, *Genome Biol. Evol.*, **6**, 2634-2646, https://doi.org/10.1093/gbe/evu210.
62. Dong, D. W., Pereira, F., Barrett, S. P., Kolesar, J. E., Cao, K., Damas, J., Yatsunyk, L. A., Johnson, F. B., and Kaufman, B. A. (2014) Association of G-quadruplex forming sequences with human mtDNA deletion breakpoints, *BMC Genomics*, **15**, 677, https://doi.org/10.1186/1471-2164-15-677.
63. Butler, T. J., Estep, K. N., Sommers, J. A., Maul, R. W., Moore, A. Z., Bandinelli, S., Cucca, F., Tuke, M. A., Wood, A. R., Bharti, S. K., Bogenhagen, D. F., Yakubovskaya, E., Garcia-Diaz, M., Guilliam, T. A., Byrd, A. K., Raney, K. D., Doherty, A. J., Ferrucci, L., Schlessinger, D., Ding, J., and Brosh, R. M. (2020) Mitochondrial genetic variation is enriched in G-quadruplex regions that stall DNA synthesis *in vitro*, *Hum. Mol. Genet.*, **29**, 1292-1309, https://doi.org/10.1093/HMG/DDAA043.
64. Chatterjee, A., Seyfferth, J., Lucci, J., Gilsbach, R., Preissl, S., Böttinger, L., Mårtensson, C. U., Panhale, A., Stehle, T., Kretz, O., Sahyoun, A. H., Avilov, S., Eimer, S., Hein, L., Pfanner, N., Becker, T., and Akhtar, A. (2016) MOF Acetyl transferase regulates transcription and respiration in mitochondria, *Cell*, **167**, 722-738.e23, https://doi.org/10.1016/j.cell.2016.09.052.
65. Doimo, M., Chaudhari, N., Abrahamsson, S., L'Hôte, V., Nguyen, T. V. H., Berner, A., Ndi, M., Abrahamsson, A., Das, R. N., Aasumets, K., Goffart, S., Pohjoismäki, J. L. O., López, M. D., Chorell, E., and Wanrooij, S. (2023) Enhanced mitochondrial G-quadruplex formation impedes replication fork progression leading

to mtDNA loss in human cells, *Nucleic Acids Res.*, **51**, 7392-7408, https://doi.org/10.1093/nar/gkad535.
66. Falabella, M., Kolesar, J. E., Wallace, C., De Jesus, D., Sun, L., Taguchi, Y. V., Wang, C., Wang, T., Xiang, I. M., Alder, J. K., Maheshan, R., Horne, W., Turek-Herman, J., Pagano, P. J., St Croix, C. M., Sondheimer, N., Yatsunyk, L. A., Johnson, F. B., and Kaufman, B. A. (2019) G-quadruplex dynamics contribute to regulation of mitochondrial gene expression, *Sci. Rep.*, **9**, 5605, https://doi.org/10.1038/s41598-019-41464-y.
67. Mishmar, D., Levin, R., Naeem, M. M., and Sondheimer, N. (2019) Higher order organization of the mtDNA: beyond mitochondrial transcription factor A, *Front. Genet.*, **10**, 1285, https://doi.org/10.3389/fgene.2019.01285.
68. Bratic, A., Clemente, P., Calvo-Garrido, J., Maffezzini, C., Felser, A., Wibom, R., Wedell, A., Freyer, C., and Wredenberg, A. (2016) Mitochondrial polyadenylation is a one-step process required for mRNA integrity and tRNA maturation, *PLoS Genet.*, **12**, e1006028, https://doi.org/10.1371/journal.pgen.1006028.
69. Krüger, A., Remes, C., Shiriaev, D. I., Liu, Y., Spåhr, H., Wibom, R., Atanassov, I., Nguyen, M. D., Cooperman, B. S., and Rorbach, J. (2023) Human mitochondria require mtRF1 for translation termination at non-canonical stop codons, *Nat. Commun.*, **14**, 30, https://doi.org/10.1038/s41467-022-35684-6.
70. Singh, V., Moran, J. C., Itoh, Y., Soto, I. C., Fontanesi, F., Couvillion, M., Huynen, M. A., Churchman, L. S., Barrientos, A., and Amunts, A. (2024) Structural basis of LRPPRC-SLIRP-dependent translation by the mitoribosome, *Nat. Struct. Mol. Biol.*, **31**, 1838-1847, https://doi.org/10.1038/s41594-024-01365-9.
71. Ruzzenente, B., Metodiev, M. D., Wredenberg, A., Bratic, A., Park, C. B., Cámara, Y., Milenkovic, D., Zickermann, V., Wibom, R., Hultenby, K., Erdjument-Bromage, H., Tempst, P., Brandt, U., Stewart, J. B., Gustafsson, C. M., and Larsson, N. G. (2012) LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs, *EMBO J.*, **31**, 443-456, https://doi.org/10.1038/emboj.2011.392.
72. Siira, S. J., Spåhr, H., Shearwood, A. M. J., Ruzzenente, B., Larsson, N. G., Rackham, O., and Filipovska, A. (2017) LRPPRC-mediated folding of the mitochondrial transcriptome, *Nat. Commun.*, **8**, 1532, https://doi.org/10.1038/s41467-017-01221-z.
73. Chang, J. H., and Tong, L. (2012) Mitochondrial poly (A) polymerase and polyadenylation, *Biochim. Biophys. Acta*, **1819**, 992-997, https://doi.org/10.1016/j.bbagrm.2011.10.012.
74. Honarmand, S., and Shoubridge, E. A. (2020) Poly (A) tail length of human mitochondrial mRNAs is tissue-specific and a mutation in LRPPRC results in transcript-specific patterns of deadenylation, *Mol. Genet. Metab. Rep.*, **25**, 100687, https://doi.org/10.1016/j.ymgmr.2020.100687.
75. Temperley, R. J., Wydro, M., Lightowlers, R. N., and Chrzanowska-Lightowlers, Z. M. (2010) Human mitochondrial mRNAs-like members of all families, similar but different, *Biochim. Biophys. Acta*, **1797**, 1081-1085, https://doi.org/10.1016/j.bbabio.2010.02.036.
76. Moran, J. C., Brivanlou, A., Brischigliaro, M., Fontanesi, F., Rouskin, S., and Barrientos, A. (2024) The human mitochondrial mRNA structurome reveals mechanisms of gene expression, *Science*, **385**, eadm9238, https://doi.org/10.1126/science.adm9238.